HOCHSCHULBÜCHER FÜR MATHEMATIK

HERAUSGEGEBEN VON
H. GRELL, K. MARUHN UND W. RINOW

BAND 7

ENZYKLOPÄDIE DER ELEMENTARMATHEMATIK

REDAKTION:

P. S. ALEXANDROFF
A. I. MARKUSCHEWITSCH
A. J. CHINTSCHIN

BAND I

ARITHMETIK

Mit 18 Abbildungen
Fünfte Auflage

VEB DEUTSCHER VERLAG DER WISSENSCHAFTEN
BERLIN 1970

Академия педагогических наук РСФСР

Энциклопедия элементарной математики

под редакцией

П. С. Александрова, А. И. Маркушевича и А. Я. Хинчина

книга первая Арифметика

Государственное издательство
технико-теоретической литературы
Москва 1951 Ленинград

Übersetzung: Helmut Limberg, Karl-Heinz Rupp und Gerhard Tesch
Endgültige Abfassung des deutschen Textes und wissenschaftliche Redaktion: Dr. Günter Asser

Verantwortlicher Verlagslektor: Ludwig Boll

ES 19 B 1,2
Alle Rechte an dieser Übersetzung beim
VEB Deutscher Verlag der Wissenschaften, Berlin
Printed in the German Democratic Republic
Lizenz-Nr. 206 · 435/170/70
Satz: IV/2/14 VEB Werkdruck Gräfenhainichen · 2349
Druck: VEB Druckerei „Thomas Müntzer" Bad Langensalza
26,70

VORWORT

Die von der Akademie der pädagogischen Wissenschaften der RSFSR vorbereitete „Enzyklopädie der Elementarmathematik" ist als Handbuch für Mathematiklehrer an Oberschulen und für Studenten der physikalisch-mathematischen Fakultäten an pädagogischen Instituten und Lehrerbildungsinstituten gedacht. Sie soll eine systematische Darstellung der wissenschaftlichen Grundlagen der Schulmathematik geben. Hieraus resultieren die Eigenarten dieses Werkes: Es kann nicht als Einführung in den behandelten Stoff dienen, sondern ist für Leser bestimmt, die die Elementarmathematik bereits studiert haben und schon Lehrer der Elementarmathematik sind oder es werden wollen. Die „Enzyklopädie" hält sich in der Regel weder an die Reihenfolge noch an die Methode der Darstellung der Mathematik in der Oberschule, da beide durch die Entwicklung des Auffassungsvermögens der Schüler und die allgemeinbildenden Ziele der Oberschule bedingt sind, d. h. durch Erwägungen, die für den ausgebildeten Fachmann ohne Belang sind. Grundsatz unseres Aufbaus ist es vielmehr, jene Fragen der Mathematik, aus denen die Schulmathematik besteht, in systematischer und möglichst einfacher und zugänglicher Form darzulegen. Daneben sollen auch Fragen angeschnitten werden, die zwar nicht direkt im Unterricht behandelt werden, die aber für das richtige und bewußte Verständnis der Schulmathematik notwendig sind oder Perspektiven auf eine weitere Entwicklung ihres Inhalts und ihrer Methoden eröffnen.

Das ganze Werk ist auf 7 Bände von je 350 bis 450 Seiten berechnet. Obwohl sich diese Bände und ihre Teile einem einheitlichen Plan unterordnen, können sie im wesentlichen unabhängig voneinander benutzt werden. Darüber hinaus können auch die einzelnen Artikel weitgehend unabhängig voneinander gelesen werden. Überdies finden sich in den einzelnen Artikeln Hinweise auf andere Artikel der „Enzyklopädie".[1]) Hier ist der genauere Plan des Gesamtwerkes:

Erster Band. Arithmetik.

Die Entstehung der Bezeichnungssysteme für die Zahlen. Mengen, Gruppen, Ringe und Körper; die theoretischen Grundlagen der Arithmetik. Die Elemente der Zahlentheorie. Kopfrechnen und schriftliches Rechnen; Hilfsmittel für das Rechnen.

Zweiter Band. Algebra.

Vektorräume und lineare Transformationen. Der Ring der Polynome und der Körper der rationalen Funktionen. Numerische und graphische Methoden zum Auflösen von Gleichungen.

[1]) Beim Hinweis auf Artikel desselben Bandes wird neben dem Titel lediglich die entsprechende Seitenzahl angegeben. Beim Hinweis auf andere Bände der Enzyklopädie verwenden wir die Abkürzung „EdEM" und Angabe der Nummer des betreffenden Bandes.

Dritter Band. Analysis.

Funktionen und Grenzwerte; rationale Funktionen, Potenzfunktionen, Exponentialfunktionen und logarithmische Funktionen; die trigonometrischen Funktionen und ihre Umkehrfunktionen. Die Elemente der Differential- und Integralrechnung. Elementare Funktionen einer komplexen Veränderlichen.

Vierter Band. Geometrie, Teil I.

Die Grundbegriffe der Topologie. Grundlagen der Geometrie. Die nichteuklidischen Geometrien. Die Elemente der analytischen und der projektiven Geometrie. Geometrische Transformationen. Die Messung von Flächeninhalten, Längen, Volumina und Oberflächen.

Fünfter Band. Geometrie, Teil II.

Polygone und Polyeder, Kreise und Kugeln. Anwendungen auf Geodäsie und Astronomie. Bemerkenswerte Kurven und Flächen. Konstruktionsaufgaben. Methoden der graphischen Darstellung.

Sechster Band. Verschiedene Fragen.

Kombinatorik. Die Elemente der Wahrscheinlichkeitsrechnung und der mathematischen Statistik. Interessante mathematische Aufgaben. Mathematische Paradoxien und Sophismen. Mathematische Scherze und Spiele.

Siebenter Band. Methodologie und Geschichte der Mathematik.

Die Mathematik und ihre Stellung gegenüber den anderen Wissenschaften, die grundlegenden Etappen ihrer Entwicklung, ihre Methoden und ihre Aufgaben. Abriß der Geschichte der Mathematik. Die Mathematik in der Sowjetunion. Beilage. Terminologisches Wörterbuch.

Der erste Band beginnt mit einem Artikel von I. G. BASCHMAKOWA und A. P. JUSCHKEWITSCH, der kultur-historischen Betrachtungen über die Bezeichnungssysteme für die Zahlen gewidmet ist.

Es folgt ein umfangreicher Artikel von I. W. PROSKURJAKOW, der sich in der Hauptsache mit der theoretischen Begründung der Arithmetik befaßt. In den ersten beiden Kapiteln dieses Artikels werden allgemeine mathematische Begriffe behandelt, deren Bedeutung weit über die Grenzen der Arithmetik hinausgeht und die sowohl im ersten als auch in den weiteren Bänden wiederholt gebraucht werden. Es handelt sich um die Begriffe Menge, Gruppe, Ring und Körper. Eine zentrale Stellung nimmt in diesem Artikel die axiomatische Begründung der natürlichen Zahlen ein, die das theoretische Fundament der gesamten Arithmetik darstellt. Auf der Grundlage der Theorie der natürlichen Zahlen werden sodann in schrittweiser Verallgemeinerung die Theorie der ganzen, der rationalen, der reellen und der komplexen Zahlen entwickelt. Anschließend macht der Verfasser mit einer weiteren Verallgemeinerung des Zahlbegriffs (den hyperkomplexen Zahlen) bekannt. Dieser Artikel gehört zu den schwierigsten und abstraktesten des ganzen vorliegenden Werkes. Die Schwierigkeiten liegen hier allerdings im Wesen der Sache selbst. Der Leser, der an den Fragen nach der logischen Begründung der Arithmetik nicht interessiert ist, kann diesen Artikel übergehen und notfalls in den ersten beiden Kapiteln dieses Artikels nachschlagen.

Der nachfolgende Artikel von A. J. CHINTSCHIN behandelt die elementarsten und wichtigsten Fragen der Zahlentheorie. Hierzu gehören die Fragen der Teilbarkeitstheorie, insbesondere die Theorie der Kettenbrüche und die Fragen der Approximation irrationaler Zahlen durch rationale.

Schließlich ist der Artikel von W. M. BRADIS dem Runden von Zahlen, den Regeln der Näherungsrechnung, der Fehlerrechnung und den Hilfsmitteln des Rechnens, einschließlich dem Rechenschieber, gewidmet.

Als wesentliche Ergänzung zum ersten Band wäre ein Bericht über die historische Entwicklung des Zahlbegriffs anzusehen, der folgende Dinge behandelt: die allmähliche und langwierige Herausbildung des allgemeinen Begriffs der natürlichen Zahl, die Entwicklung des Begriffs der gebrochenen Zahl, die Urform der Theorie der positiven reellen Zahlen, wie sie sich bei den alten Griechen (in den „Elementen" des EUKLID) gebildet hat, und die Entwicklung des Begriffs der negativen und der komplexen Zahl im Zusammenhang mit der Gleichungstheorie und später der analytischen Geometrie und Analysis. Auf alle diese Fragen wird in den einzelnen Artikeln nicht eingegangen. Sie ordnen sich vielmehr in den allgemeinen Abriß der Geschichte der Mathematik ein, der im letzten Bande der „Enzyklopädie" veröffentlicht wird.

Die Redaktion.

INHALTSVERZEICHNIS

I. G. Baschmakowa und A. P. Juschkewitsch

Die Entstehung der Bezeichnungssysteme für die Zahlen

Einleitung 3

§ 1. Das Anfangsstadium der Entwicklung des Zählens 6
§ 2. Nicht-positionelle Bezeichnungssysteme 17
§ 3. Die alphabetischen Bezeichnungssysteme für die Zahlen 21
§ 4. Stellen- oder Positionssysteme 28
§ 5. Die Ausbreitung der positionellen Schreibweise der Zahlen nach Westeuropa und Rußland 39
§ 6. Brüche 46

Zusammenfassung 58

I. W. Proskurjakow

Mengen, Gruppen, Ringe und Körper. Die theoretischen Grundlagen der Arithmetik

Einleitung 63

Kapitel I. Mengen

§ 1. Der Begriff der Menge 65
§ 2. Mengenalgebra 67
§ 3. Funktionen, Abbildungen, Mächtigkeiten 69
§ 4. Endliche und unendliche Mengen 73
§ 5. Geordnete Mengen 79

Kapitel II. Gruppen, Ringe, Körper

§ 6. Gruppen 83
§ 7. Ringe 90
§ 8. Körper 95
§ 9. Der axiomatische Aufbau der Mathematik, Isomorphie 101
§ 10. Angeordnete Ringe und Körper 105

Kapitel III. Die natürlichen Zahlen

§ 11. Das Axiomensystem für die natürlichen Zahlen 112
§ 12. Addition 114
§ 13. Multiplikation 118
§ 14. Ordnung 120
§ 15. Induktive Definitionen. Summe und Produkt aus mehreren Zahlen 123
§ 16. Subtraktion und Division 128
§ 17. Bemerkungen zu dem Axiomensystem der natürlichen Zahlen 130

Kapitel IV. Der Ring der ganzen Zahlen
§ 18. Das Erweiterungsprinzip in der Arithmetik und der Algebra 134
§ 19. Äquivalenzrelationen und Äquivalenzklassen 136
§ 20. Definition des Ringes der ganzen Zahlen 137
§ 21. Eigenschaften der ganzen Zahlen 145

Kapitel V. Der Körper der rationalen Zahlen
§ 22. Definition des Körpers der rationalen Zahlen 148
§ 23. Eigenschaften der rationalen Zahlen 155

Kapitel VI. Der Körper der reellen Zahlen
§ 24. Vollständige und stetige Körper 163
§ 25. Definition des Körpers der reellen Zahlen 176
§ 26. Eigenschaften der reellen Zahlen 186
§ 27. Die axiomatische Charakterisierung der reellen Zahlen 194

Kapitel VII. Der Körper der komplexen Zahlen
§ 28. Definition des Körpers der komplexen Zahlen 198
§ 29. Eigenschaften der komplexen Zahlen 204
§ 30. Hyperkomplexe Zahlen. Quaternionen 211

Literatur 222

A. J. Chintschin

Die Elemente der Zahlentheorie

Kapitel I. Teilbarkeit und Primzahlen
§ 1. Einleitung 227
§ 2. Die eindeutige Zerlegbarkeit einer Zahl in Primfaktoren 228
§ 3. Über Primzahlen 233

Kapitel II. Die Kongruenzmethode
§ 4. Einleitung 241
§ 5. Die Kongruenzen und ihre wichtigsten Eigenschaften 242
§ 6. Klassifikation der Zahlen nach einem gegebenen Modul 246
§ 7. Kongruenzen, die Unbekannte enthalten 251

Kapitel III. Der Euklidische Algorithmus und die Kettenbrüche
§ 8. Der Euklidische Algorithmus 259
§ 9. Die elementare Theorie der Kettenbrüche 265

Kapitel IV. Die Darstellung der Zahlen durch systematische Brüche und durch Kettenbrüche
§ 10. Einleitung 273
§ 11. Die systematischen Brüche 274
§ 12. Die Kettenbrüche 280

Kapitel V. Kettenbrüche und diophantische Approximationen
§ 13. Die Näherungsbrüche als beste Approximation 285
§ 14. Diophantische Approximationen 297

Kapitel VI. Algebraische und transzendente Zahlen
§ 15. Der Satz von Liouville und das erste Auftreten von transzendenten Zahlen 303
§ 16. Die Methode von Cantor 307
§ 17. Die arithmetische Natur der klassischen Konstanten 309

W. M. Bradis

Kopfrechnen und schriftliches Rechnen. Hilfsmittel für das Rechnen

Kapitel I. Allgemeine Bemerkungen über das Rechnen und die Näherungsrechnungen

§ 1. Allgemeine Bemerkungen über das Rechnen in der Schule 317
§ 2. Das Kopfrechnen 319
§ 3. Das schriftliche Rechnen 321
§ 4. Hilfsmittel für das Rechnen 325
§ 5. Näherungswerte 336
§ 6. Verschiedene Methoden zur Abschätzung der Genauigkeit von Näherungswerten 339
§ 7. Die Auswertung von Meßresultaten 342

Kapitel II. Fehlerrechnung

§ 8. Berechnungen mit strenger Fehlerrechnung nach der Methode der Wertschranken 346
§ 9. Berechnungen mit strenger Fehlerrechnung nach der Methode der Fehlerschranken 350
§ 10. Höchstfehler der Resultate der Grundrechenarten an Näherungswerten. Regeln der Ziffernzählung 356
§ 11. Der mittlere quadratische Fehler der Resultate der Grundrechenarten an Näherungszahlen. Das Prinzip von A. N. Krylow 362
§ 12. Verteilung der Fehler in den Rechenresultaten 367
§ 13. Praktische Anwendungen der Regeln über die Ziffernzählung. Aufstellung weiterer Regeln 369

Kapitel III. Verschiedene Fragen

§ 14. Näherungsformeln. Abgekürzte Verfahren für die Grundrechenarten . . . 375
§ 15. Mathematische Tafeln 380
§ 16. Graphische Berechnungen 382
§ 17. Der logarithmische Rechenschieber 384
§ 18. Über den Rechenunterricht in den verschiedenen Schuljahren 389

Literatur 394

Namenverzeichnis 396

Sachverzeichnis 397

I. G. BASCHMAKOWA UND A. P. JUSCHKEWITSCH

DIE ENTSTEHUNG DER BEZEICHNUNGSSYSTEME FÜR DIE ZAHLEN

Einleitung

Der Zweck eines jeden Bezeichnungssystems für die natürlichen Zahlen ist die eindeutige Darstellung jeder natürlichen Zahl mittels einer geringen Anzahl von Grundzeichen. Dies könnte man schon mit Hilfe eines einzigen Grundzeichens „1" („Eins") erreichen, indem man etwa die natürliche Zahl n als Folge aus n Einsen schreibt. Die Addition zweier Zahlen würde sich dann einfach als Aneinanderfügung der zugehörigen Folgen, die Subtraktion als Streichung darstellen. Obwohl diesem Bezeichnungssystem eine sehr einfache Idee zugrunde liegt, erweist es sich für die Praxis als unbrauchbar, da sich in ihm einigermaßen große Zahlen kaum darstellen lassen. Es wurde deshalb auch nur von solchen Völkern benutzt, deren Rechnungen sich lediglich über sehr kleine Bereiche der Zahlenreihe erstreckten.

Die für das praktische Rechnen vollkommenste Schreibweise der Zahlen beruht auf einem Prinzip, welches u. a. auch unserer üblichen dezimalen Zifferndarstellung zugrunde liegt. Bei der dezimalen Zifferndarstellung werden die Zahlen von 1 bis 9 durch die Zahlzeichen (Ziffern) „1", „2", „3", ..., „9" bezeichnet, ferner nimmt man das Zeichen „0" für die Zahl Null. Jede natürliche Zahl läßt sich dann nach dem Prinzip des Stellen- oder Positionswertes allein mit Hilfe dieser zehn Zeichen darstellen.

Dabei benutzt man die Tatsache, daß sich jede natürliche Zahl n eindeutig in der Form

$$n = a_m 10^m + a_{m-1} 10^{m-1} + \cdots + a_1 10 + a_0$$

darstellen läßt, wobei die a_i die Werte 0, 1, 2, ..., 9 annehmen können und $a_m \neq 0$ ist. Dann ergibt sich im (dezimalen) Positionssystem für die Zahl n die Darstellung

$$n = a_m a_{m-1} \ldots a_1 a_0.$$

Jede Ziffer a_i besitzt also im Positionssystem eine Bedeutung, die bestimmt ist 1.) durch ihre Gestalt und 2.) durch ihre Stellung im Schriftbild der Zahl. Wollen wir z. B. die Zahl Viertausend schreiben, so müssen wir an die vierte Stelle von rechts die Ziffer „4" setzen; die übrigen drei Koeffizienten in der dezimalen Entwicklung sind im betrachteten Fall Null, deshalb müssen wir bei der Darstellung unserer Zahl im Positionssystem an die betreffenden Stellen das Zeichen „0" setzen: „4000". Somit kann also z. B. das Zeichen „4", je nach der Stellung, die es im Schriftbild einer Zahl einnimmt, 4 Einer, 4 Zehner, 4 Hunderter usw. bedeuten.

Ungeachtet der scheinbaren Einfachheit dieser Schreibweise ist sie doch das Ergebnis einer langen historischen Entwicklung. An der Schöpfung dieser Schreibweise waren viele Völker, ja man kann sogar sagen, die ganze Menschheit beteiligt. Der bekannte französische Mathematiker und Physiker LAPLACE schrieb: „Die Idee, alle Zahlen durch neun Zeichen auszudrücken, indem man diesen außer ihrer Bedeutung durch ihre Form noch eine Bedeutung durch ihre Stellung beilegt, ist so einfach, daß es namentlich wegen dieser Einfachheit schwer zu begreifen ist, wie bewundernswert sie ist. Wie schwer es war, auf diese Methode zu kommen, sehen wir am Beispiel der großen Genien der griechischen Wissenschaft, ARCHIMEDES und APPOLONIUS, denen diese Idee verborgen blieb."

Ebenso wie im Vorhergehenden die Zahl 10 eine Basis für das dezimale Positionssystem abgab, kann auch jede andere (von 1 verschiedene) natürliche Zahl als

Basis für ein Positionssystem genommen werden. Viele Mathematiker hielten z. B. die Zahl 12 für eine brauchbarere Basis als die Zahl 10, da 12 mehr Teiler als 10 besitzt. Die besonders weite Verbreitung des Dezimalsystems hängt wohl eng mit der Anzahl unserer Finger zusammen, worauf erstmalig ARISTOTELES in seinen „Problemen" aufmerksam machte. Rein sachlich gesehen besitzt es keinerlei nennenswerte Vorzüge, die es vor Positionssystemen mit anderer Basis auszeichnen. Die Auswahl der Basis erweist sich also prinzipiell als willkürlich. Natürlich darf sie nicht zu groß sein, da man sonst zu viele Ziffern benötigt und z. B. das Einmaleins zu umfangreich wird. Andererseits darf sie aber auch nicht zu klein sein.[1])

Einen Beweis dafür, daß nicht in allen Epochen die Bezeichnungsweise der Zahlen mit unserer heutigen übereinstimmte, liefert schon die Betrachtung der Zahlwörter. Bei der sprachlichen Bezeichnung der Zahlen herrscht keineswegs jene Einheitlichkeit wie bei ihrer Ziffernschreibweise. So gibt es im Deutschen[2]) außer den Bezeichnungen für Null und die ersten neun Zahlen: „Null", „Eins", „Zwei", ..., „Neun" spezielle Bezeichnungen für die drei darauffolgenden Zahlen, nämlich „Zehn", „Elf", „Zwölf"[3]) (während in der Schrift z. B. die Zahl Zehn durch „10", d. h. mit Hilfe von Eins und Null bezeichnet wird). Solche speziellen Bezeichnungen existieren auch für eine Reihe von größeren Zahlen: „Hundert", „Tausend", „Million" usw.

Ferner sprechen wir in der deutschen Sprache die Zahlen von 13 bis 19: „Dreizehn", ..., „Neunzehn", d. h., wir sprechen eine gewisse Zahl zwischen 3 und 9 mit dem Zusatz „-zehn". Die Zahlen von 21 bis 99 werden in der deutschen Sprache durchweg praktisch in Übereinstimmung mit ihrer Ziffernschreibweise ausgesprochen[4]): „Einundzwanzig", „Zweiunddreißig" usw.[5]). In der russischen Sprache bilden hier eine Ausnahme die Zahlwörter сорок (Vierzig)[6]) und девяносто (Neunzig)[7]).

In diesem Sinne unterscheiden wir in jedem Bezeichnungssystem zwischen Zahlen, die eine *individuelle* Bezeichnung besitzen, also eine Bezeichnung, die sich nicht

[1]) Die Darstellung der Zahlen mit Hilfe einer beliebigen Basis wurde zuerst von B. PASCAL in seiner Abhandlung „De numeris multiplicibus ex sola characterum numericorum additione agnoscendis" untersucht (1654, veröffentlicht 1665).

[2]) Die nachfolgenden Ausführungen sind sinngemäß auf die deutsche Sprache übertragen worden. Auf die Abweichungen der russischen gegenüber der deutschen Sprache wird, soweit sie im betrachteten Zusammenhang von allgemeinem Interesse sind, in der nächsten Anmerkung bzw. an der betreffenden Stelle eingegangen. *Anm. d. wissenschaftl. Red.*

[3]) Die Zahlwörter „Elf" und „Zwölf" sind aus „ein-lif" (eins über [zehn]) bzw. „zwo-lif" (zwei über [zehn]) entstanden (vgl. M. HEYNE, Deutsches Wörterbuch, Band 1, 2. Aufl., Leipzig 1905, S. 740, und Band 2, 2. Aufl., Leipzig 1906, S. 1464), sind also daher nicht als individuelle Bezeichnungen (siehe unten) anzusehen. In der russischen Sprache hat eine derartige Verschmelzung nicht stattgefunden. Dort bezeichnet man die Zahlen von 11 bis 19 gleichmäßig als один-на-дцать, ..., девят-на-дцать, (eins auf zehn, ..., neun auf zehn), d. h., man spricht dort eine Zahl zwischen 1 und 9 mit dem Zusatz «на десять» (auf zehn). *Anm. d. wissenschaftl. Red.*

[4]) Allerdings, wie auch in anderen germanischen Sprachen, unter „Inversion" der Zehner und Einer. Eine derartige Inversion gibt es z. B. in der englischen, französischen oder russischen Sprache nicht. *Anm. der wissenschaftl. Red.*

[5]) Die Schlußsilbe in „Zwanzig", „Dreißig" usw. hängt mit dem gotischen Wort „tigus" (zehn) zusammen. *Anm. d. wissenschaftl. Red.*

[6]) Die Zahl 40 spielte in Rußland und bei vielen anderen Völkern des Ostens eine besondere Rolle, von der später noch die Rede sein wird.

[7]) Das Wort девяносто ist nicht als individuelles Zahlwort (s. u.) anzusehen, da es vermutlich aus «девять до ста» (neun bis Hundert) entstanden ist.

aus anderen Zahlwörtern zusammensetzt (Knotenzahlen), und Zahlen, die eine *algorithmische* Bezeichnung besitzen, also eine Bezeichnung, die eine Kombination von individuellen Bezeichnungen ist (algorithmische Zahlen). Wir werden sehen, daß sich in dieser Unterscheidung der Bezeichnungen ein wesentlicher Unterschied in ihrer Entstehung bei der geschichtlichen Entwicklung des Zahlbegriffs widerspiegelt.[1])

Analoge Erscheinungen wie in der deutschen und der russischen Sprache zeigen sich natürlich auch in anderen Sprachen. So haben sich z. B. in der französischen Sprache offensichtliche Reste eines nicht-positionellen Zwanziger-Systems erhalten. Es erweist sich nämlich dort die Zahl 20 als eine Zahl mit individueller Bezeichnung, deren Benennung „vingt" nicht aus den Benennungen vorangehender Zahlen gebildet wird. Die Zahl 80 wird „quatre-vingt" (vier-zwanzig), die Zahl 90 wird „quatre-vingt-dix" (vier-zwanzig-zehn) und die Zahl 120 wird gelegentlich „six-vingt" (sechs-zwanzig) gesprochen. In der altfranzösischen Sprache wurde außerdem 140 im Sinne von sieben-zwanzig, die Zahl 160 im Sinne von acht-zwanzig, die Zahl 300 im Sinne von fünfzehn-zwanzig usw. gesprochen. Spuren eines Zwanziger-Systems erhielten sich außer in der französischen Sprache u. a. auch in der englischen und der holländischen Sprache. So bedeutet das englische Wort „score" neben anderen Dingen die Zahl 20 und „three-score" (drei-zwanzig) die Zahl 60. In den skandinavischen Sprachen sind außerdem ausgeprägte Spuren eines Fünfer-Systems zu finden.

Wir können also zusammenfassend feststellen:

1.) Die moderne Ziffernschreibweise entspricht einem strengen Positionssystem, die sprachlichen Bezeichnungssysteme dagegen nicht.

2.) Die Ziffernschreibweise ist streng dezimal, die sprachlichen Bezeichnungssysteme enthalten Spuren eines Zwanziger-Systems bzw. anderer Systeme.

3.) In der modernen Ziffernschreibweise besitzen nur die Zahlen 0, 1, 2, ..., 9 eine individuelle Bezeichnung; in den sprachlichen Bezeichnungssystemen besitzen auch andere Zahlen eine solche Bezeichnung, von denen jede als Basis ihres Stellensystems dient, d. h. als Basis eines gewissen echten Abschnittes der Zahlenreihe (z. B. spricht man in der deutschen Sprache die Zahlen von 100 bis 999 als Kombinationen aus dem Wort „Hundert" und individuellen oder algorithmischen Zahlwörtern für kleinere Zahlen: „Hundert-eins", „Hundertzwölf", „Hunderteinundzwanzig" usw.).

Man kann sagen, daß unsere sprachliche Ausdrucksweise ein früheres Stadium des Rechnens widerspiegelt als unsere moderne Schreibweise. Es ist aber z. B. noch die römische schriftliche Numerierung, die unserem heutigen Positionssystem voranging, ihrer Struktur nach mit der sprachlichen Numerierung der modernen europäischen Völker verwandt.

---

[1]) In der „Arithmetik" von L. Magnizki aus dem Jahre 1703 findet sich eine Einteilung der Zahlen in «персты» (Finger) (Zahlen bis 10), «составы» (zusammengesetzte) (ganze Zehner) und «сочинения» (Verbindungen) (die übrigen Zahlen bis Hundert). Das älteste bekannte Beispiel einer ähnlichen Einteilung stammt aus dem 10. Jahrhundert von Gerbert (digiti [Finger], articuli [Gelenke], compositi [Zusammengesetzte]). Offensichtlich haben wir es hier mit einer ähnlichen Einteilung der Zahlen in solche mit individueller bzw. mit algorithmischer Bezeichnung zu tun. Zweifellos hängen auch die Bezeichnungen „Finger" und „Gelenke" mit der Fingerrechnung zusammen.

Als individuelle Zeichen wurden in der römischen Numerierung benutzt: „I" für die Zahl 1, „V" für die Zahl 5, „X" für die Zahl 10, „L" für die Zahl 50, „C" für die Zahl 100, „D" für die Zahl 500, „M" für die Zahl 1000. Eine Null gab es nicht. Dieses System ist dezimal nicht-positionell und besitzt starke Überreste eines Fünfer-Systems (individuelle Zeichen für 5, 50, 500). Bei der Bildung der algorithmischen Bezeichnungen wird sowohl die Addition als auch die Subtraktion benutzt. Z. B. wird die Zahl 1948 in diesem Bezeichnungssystem folgendermaßen geschrieben: MCMXLVIII.

Etwa in demselben Verhältnis, in dem die römische Schreibweise der Zahlen zur modernen sprachlichen Ausdrucksweise steht, stehen die Zählmethoden vieler sogenannter Urvölker zur römischen Schreibweise. Aus den bisherigen Bemerkungen geht hervor, daß wir zur Aufdeckung der Herkunft der Bezeichnungssysteme (sowohl der modernen Positionssysteme als auch der nicht-positionellen) ethnographisches und sprachwissenschaftliches Material benutzen müssen.

§ 1. Das Anfangsstadium der Entwicklung des Zählens

Der Zahlbegriff ist einer der grundlegenden Begriffe der Mathematik. Er ist auch einer der ältesten Begriffe. Alle Kulturvölker, die eine Schrift besaßen, hatten auch schon einen Begriff von der Zahl und gewisse Systeme zu ihrer Bezeichnung. Über den Zahlbegriff in vorgeschichtlichen Zeiten können wir nur indirekt urteilen. Als Quellen erweisen sich hier erstens die Sprachwissenschaft und zweitens die Ethnographie, die es gestatten, auf Grund des Studiums der Kultur von Völkern, die nach der Klassifikation von ENGELS auf der Stufe der Wildheit und Barbarei stehen, über analoge Perioden des Lebens der Vorfahren der modernen Kulturvölker zu urteilen. Leider war das Sammeln ethnographischen Materials während langer Zeit das ausschließliche Monopol von Missionaren und Kolonisatoren. Als gegen Ende des 19. Jahrhunderts — bedingt durch die Entwicklung der Wissenschaften — auch die Aufmerksamkeit der Gelehrten mit besonderer Stärke auf die vorgeschichtlichen Zeiten des Lebens der Menschheit gerichtet wurde, zeigte es sich, daß es solche sogenannten „Urvölker" fast nicht mehr gab. Die imperialistische Politik der kapitalistischen Staaten hatte zu dieser Zeit zu einer nahezu vollständigen Ausrottung vieler eingeborener Stämme geführt. So war z. B. der australische Stamm der Tasmanier zu Anfang des 20. Jahrhunderts vollständig vernichtet. Genau dasselbe war auch mit dem ehemals zahlenmäßig starken Stamm der Abiponer in Südamerika geschehen.

Bei der Rekonstruktion der ersten Entwicklungsstadien des Zahlbegriffs muß man sich also mit sehr spärlichem Material begnügen. Jedoch ist die Frage nach der Herkunft dieses Begriffes so wichtig, daß auch jenes unvollständige Bild, welches man entwickeln kann, eine große Bedeutung besitzt. Sie ist insbesondere wichtig für die Widerlegung der idealistischen Theorie, daß der Zahlbegriff und sogar die gesamte Folge der natürlichen Zahlen dem Menschen angeboren sind. Bekannt ist in diesem Zusammenhang der folgende Ausspruch von KRONECKER: „Die ganzen

Zahlen hat Gott geschaffen, alles übrige ist das Werk von Menschenhänden." Das Studium der ersten Etappen der Entwicklung des Zahlbegriffs und anderer grundlegender mathematischer Begriffe widerlegt diese und ähnliche bürgerliche Theorien vollständig. Eine objektive Untersuchung zeigt, daß eine enge Verbindung zwischen dem Entstehen dieser Begriffe und der Arbeitspraxis der Horden der Urgesellschaft besteht und daß selbst unsere „Intuition" keine unveränderliche Kategorie ist, daß also die sogenannten „Urbegriffe" keineswegs angeboren sind. Das Studium der ersten Entwicklungsstadien des Zahlbegriffs zeigt uns, daß auch die ganzen Zahlen vom Menschen geschaffen wurden, daß also auch sie das Werk von Menschenhänden sind.

Unter den bürgerlichen Gelehrten des 19. Jahrhunderts (TAYLOR u. a.) war die Meinung verbreitet, der Urmensch habe all sein Wissen über die Welt durch Beobachtung der Naturerscheinungen, Kombination mit früher Gesehenem und logische Schlüsse erhalten. Der Urmensch erschien dabei als ein rein betrachtender Philosoph. In Wirklichkeit begann der Mensch nicht mit dem Theoretisieren, sondern mit der Arbeit, mit dem Kampf um seine Existenz, nicht mit der passiven Wahrnehmung der Natur, sondern aktiv mit ihrer Umgestaltung.

MARX schrieb in den „Randglossen zu Adolph Wagners ‚Lehrbuch der politischen Ökonomie'", daß die Beziehung des Menschen zur Natur von Anfang an keine theoretische, sondern eine praktische, d. h. eine auf die Tätigkeit gegründete, war. „Sie fangen, wie jedes Tier, damit an, zu essen, zu trinken etc., also nicht in einem Verhältnis zu stehen, sondern sich aktiv zu verhalten, sich gewisser Dinge der Außenwelt zu bemächtigen durch die Tat, und so ihr Bedürfnis zu befriedigen." [1])

Über die Arbeit schrieb ENGELS in der „Dialektik der Natur": „Sie ist die erste Grundbedingung alles menschlichen Lebens, und zwar in einem solchen Grade, daß wir in gewissem Sinne sagen müssen: Sie hat den Menschen selbst geschaffen." [2]) Insbesondere wurden im Prozeß der Arbeit solche grundlegenden Begriffe wie der Zahlbegriff, der Begriff der natürlichen Anordnung, der Begriff der Figur geschaffen, wurden einfachste Rechenregeln sowie die Methoden der Längen-, Flächen- und Raummessung herausgearbeitet.

Dabei waren die Begriffe der Zahl und der Figur und ihre grundlegenden Eigenschaften Widerspiegelungen von Eigenschaften und Beziehungen der realen Gegenstände der Umwelt. Hierüber schrieb ENGELS: „Die Begriffe von Zahl und Figur sind nirgends anders hergenommen als aus der wirklichen Welt. Die zehn Finger, an denen die Menschen zählen, also die erste arithmetische Operation vollziehen gelernt haben, sind alles andere, nur nicht eine freie Schöpfung des Verstandes. Zum Zählen gehören nicht nur zählbare Gegenstände, sondern auch schon die Fähigkeit, bei Betrachtung dieser Gegenstände von allen übrigen Eigenschaften abzusehen außer ihrer Zahl — und diese Fähigkeit ist das Ergebnis einer langen geschichtlichen, erfahrungsmäßigen Entwicklung." [3])

Wir werden zeigen, wie der Begriff der Zahl und der Begriff der natürlichen Anordnung auf der ersten Stufe der kulturellen Entwicklung der Menschheit aussahen und wie sie sich — allmählich sich verändernd und vervollkommnend — zu den heutigen Begriffen entwickelten.

---

[1]) K. MARX und F. ENGELS, Werke, Bd. 19, Berlin 1962, S. 362–363.
[2]) F. ENGELS, Dialektik der Natur, Berlin 1952, S. 179.
[3]) F. ENGELS, Anti-Dühring, Berlin 1948, S. 44.

Leider gestatten es die uns zur Verfügung stehenden Materialien nicht, die verschiedenen ersten Etappen der Entwicklung des Rechnens genau mit der von ENGELS in seiner Arbeit „Ursprung der Familie, des Privateigentums und des Staates" gegebenen Periodisierung der vorgeschichtlichen Stufen zu verbinden.

Es gab selbst in unserem Zeitalter noch Völker, deren Sprache nur zwei individuelle Zahlwörter kannte, nämlich für die Zahlen 1 und 2. Bei vielen Stämmen Australiens und Polynesiens war es damit noch in der jüngsten Vergangenheit so. Die weiteren Zahlen wurden durch Kombination dieser Zahlwörter gebildet: 3 = zwei-eins, 4 = zwei-zwei, 5 = zwei-zwei-eins, 6 = zwei-zwei-zwei. So gab es z. B. bei den Bewohnern der westlichen Inseln der Torres-Straße nur die Zahlwörter „urapun" für die Zahl 1 und „okosa" für die Zahl 2. Weiter zählten sie: 3 = okosa-urapun, 4 = okosa-okosa, 5 = okosa-okosa-urapun, 6 = okosa-okosa-okosa. Dieselbe Methode liegt dem ältesten aller Zahlsysteme, dem dyadischen System, zugrunde.[1]) Spuren dieser Rechenmethode finden wir mehrfach in der ägyptischen Multiplikations- und Divisionsmethode sowie in der ägyptischen Bruchrechnung.[2]) Sie äußert sich auch darin, daß in manchen Sprachen — z. B. in der altslawischen — neben der Einzahl und der Mehrzahl noch eine Zweizahl auftritt.[3])

Die Zahlen oberhalb 6 bezeichneten die Inselbewohner der Torres-Straße durch „viel-viel", „Menge" oder „unabzählbar". Bei manchen Stämmen wurden die Bezeichnungen „viel", „unabzählbar" bereits für die Zahlen von 3 ab angewandt.

[1]) Für die mündliche und schriftliche Numerierung ist das dyadische System zu unbequem, da die Bezeichnungen für die Zahlen sehr lang werden (z. B. wird die Zahl 777 in diesem System durch „1 100 001 001" bezeichnet), andererseits besitzt es aber auch bemerkenswerte Vorteile. Die prinzipiellen Vorteile des dyadischen Systems wurden zuerst von LEIBNIZ bemerkt, der auf die besondere Einfachheit der Rechenoperationen in diesem System hinwies (die bekannten Additions- und Multiplikationsschemata reduzieren sich im wesentlichen auf $1 + 1 = 10$, $1 \cdot 1 = 1$; bei der Division ist kein Probieren notwendig). LEIBNIZ empfahl dieses System nicht etwa als Ersatz für das dekadische bei praktischen Rechnungen, hob jedoch hervor, daß „die Berechnung mit Hilfe des Dubletts 0,1 als Entgelt für ihre Längen für die Wissenschaft grundlegend ist, indem sie neue Entdeckungen hervorbringt, die sich später als nützlich erweisen, und zwar sowohl in der Zahlenpraxis als auch besonders in der Geometrie: Der Grund dafür ist die Tatsache, daß bei der Zurückführung der Zahlen auf ihre einfachsten Elemente, 0 und 1, überall eine wunderbare Ordnung zutage tritt". (S. „Explication de l'arithmétique binaire, qui se sert des seuls caractères 0 et 1, avec des remarques sur son utilité", 1703, in „Leibnizens mathematische Schriften", herausgegeben von C. I. GERHARDT, Band VII, Halle 1863, S. 225; siehe auch ebendort den Brief von LEIBNIZ an SCHULENBURG von 1698). In der Tat erweist sich das dyadische System als sehr brauchbar für eine Reihe von theoretischen Untersuchungen. LEIBNIZ konnte jedoch nicht voraussehen, daß das dyadische System auch große Vorteile für die praktische Mathematik in sich birgt. Man hat nämlich in der letzten Zeit elektrische Rechenmaschinen konstruiert, die im dyadischen System arbeiten. Die Leistungsfähigkeit solcher Maschinen kompensiert reichlich den Mehraufwand an Arbeit beim Übergang vom dekadischen System zum dyadischen System am Anfang der Rechnung und beim umgekehrten Übergang zum endgültigen Resultat. Siehe z. B. Л. Д. Кудрявцев, О принципах производства арифметических операций на вычислительных машинах (L. D. KUDRJAWZEW, Über die Prinzipien der Ausführung arithmetischer Operationen auf Rechenmaschinen), Uspechi mat. Nauk 5, Nr. 3 (37), 104—127 (1950). [Vgl. auch Literaturverzeichnis zum letzten Artikel.]

[2]) Vgl. auch die Ausführungen über die Halbierungsmethode der alten russischen Bruchrechnung in § 6.

[3]) In einigen Sprachen gibt es sogar noch eine Dreizahl, die auf ein triadisches System hindeutet. Auf eine Sonderstellung der ersten drei Zahlen weist auch der Umstand hin, daß sich in manchen Sprachen ihre Bezeichnungen — im Unterschied zu den anderen Zahlwörtern — mit dem Geschlecht ändern (deutsch: ein, eine; russisch: один, одна, одно; два, две; lateinisch: tres, tria).

Somit ist auf dieser Kulturstufe die Folge der natürlichen Zahlen endlich[1]) und besteht mitunter nur aus zwei Gliedern. Man darf aber nicht denken, daß die Stämme, die nur Zahlwörter für die Zahlen 1 und 2 besitzen, nicht Gesamtheiten vergleichen können, die aus mehr als zwei oder sechs Dingen bestehen. Der Mensch lernte im üblichen Sinne „zählen", lange bevor Bezeichnungen für die Zahlen entstanden. So berichtet ein Forscher von den Abiponern — von denen bekannt ist, daß sie nur Zahlwörter für die Zahlen 1, 2 und 3 besaßen —, daß sie, wenn sie sich zur Jagd zusammentaten und schon im Sattel saßen, sich umschauten und, wenn etwa einer ihrer zahlreichen Hunde fehlte, diesen riefen. Den Forscher wunderte es besonders, daß es die Abiponer fertig brachten, ohne zählen zu können, sofort festzustellen, daß in der beträchtlichen Meute ein Hund fehlte.

Es war einfach so, daß in jenem Stadium die Anzahl einer Gesamtheit von Dingen eine Eigenschaft war, die mit anderen Eigenschaften dieser Gesamtheit (der Farbe, der Form, der Gestalt usw. ihrer Elemente) untrennbar verknüpft war. Sie charakterisierte also die Gesamtheit erstens bezüglich ihrer Vollständigkeit (ob alle Dinge der betreffenden Gesamtheit vorhanden sind) und zweitens in Ordnungsbeziehung zu anderen Gesamtheiten aus gleichen oder zumindest ähnlichen Dingen (eine Gesamtheit ist größer oder kleiner als eine andere).

Offensichtlich genügte ein solches „Zählen" nur auf jener Stufe der Entwicklung der Menschheit, auf der es — kurz gesagt — nichts zu zählen gab, auf der die Wirtschaft noch auf einem sehr niedrigen Niveau stand und es keine engere Verbindung zwischen den verschiedenen Stämmen gab. Auf der ersten Stufe der Entwicklung erscheint also die Zahl lediglich als spezielle Anzahl-Eigenschaft oder Anzahl-Qualität konkreter Gesamtheiten von Dingen. Eine Ordnungsbeziehung tritt kaum hervor.

Heutzutage findet man keine Völker mehr, deren Zahlbegriff auf dieser ersten Stufe stehengeblieben wäre, die im wesentlichen den ersten beiden Stufen der Wildheit entspricht. Das „Zählen" mit Anzahlen von konkreten Gegenständen blieb nur bei einigen Völkern als Überrest erhalten.

Als Pfeil und Bogen erfunden waren und die Stämme zur systematischen Jagd übergingen, als sich die Menschen in Dörfern niedergelassen hatten und Verbindungen, anfangs zwischen einzelnen Dörfern, später auch zwischen verschiedenen Stämmen entstanden — kurz, beim Übergang zu einer höheren Stufe der Wildheit —, wurde das „Zählen" mit den konkreten Anzahlen unzureichend. Es genügte nicht mehr, die Anzahl von bestimmten, konkreten Gesamtheiten nur nach dem Augenmaß bestimmen zu können, sondern man mußte auch anderen von der Anzahl Mitteilung machen können. Man mußte z. B. benachbarten Stämmen mitteilen können, daß nach einer bestimmten Anzahl von Mondwechseln eine Zusammenkunft zu Verhandlungen oder zu einer gemeinsamen Jagd festgesetzt sei, oder daß alle verbündeten Stämme nach einer bestimmten Frist eine gewisse Anzahl von Kriegern aufstellen sollen. Hierfür benutzten die Aruntier (Australien) und die Polynesier folgende

[1]) Reste davon, daß auch bei den Vorfahren der europäischen Völker die Zahl 7 gleichzeitig als Bezeichnung für eine unbestimmte Anzahl diente, erhielten sich in der russischen Sprache in Form von Sprichwörtern und Redensarten: «Семеро одного не ждут» (Sieben Mann warten nicht auf einen), «Семь раз отмерь, один раз отрежь» (Siebenmal abmessen, einmal abschneiden), «У семи нянек дитя без глаза» (Bei sieben Kindermädchen ist das Kind ohne Aufsicht) usw. In allen diesen Fällen wird das Wort „sieben" offenbar im Sinne von „viel" gebraucht.

Methode: Zur Bezeichnung größerer Zahlen bedienten sich die Eingeborenen verschiedener Körperteile, von denen jeder seine Bezeichnung hatte und der eine wohlbestimmte Stelle in ihrem Bezeichnungssystem einnahm. Eine bestimmte Anzahl so aufgezählter Körperteile bezeichnete dieselbe Anzahl von Kriegern, von Tagen oder von Monaten — je nach dem besonderen Sachverhalt.

Die Zählung begann für gewöhnlich mit dem kleinen Finger der linken Hand, ging über zu den anderen Fingern, dann zum Handgelenk, zum Ellenbogen, zur Schulter usw. bis zum kleinen Finger der rechten Hand und danach — wenn die abzuzählende Gesamtheit noch nicht erschöpft war — in der umgekehrten Reihenfolge zurück. Bei der Regelung von Geschäftsangelegenheiten genügte es dann, daß sich der Eingeborene erinnerte, bis zu welchem Körperteil er bei der Abzählung irgendwelcher Dinge gekommen war, damit er — wenn er die Abzählung, beginnend mit dem kleinen Finger der linken Hand, wiederholte — die betreffende Anzahl rekonstruieren konnte.

Auf diese Weise wurden bei den Insulanern der Torres-Straße die Zahlen bis 33 durch Teile des menschlichen Körpers versinnbildlicht. Bestand eine abzuzählende Gesamtheit aus mehr als 33 Dingen, so bedienten sie sich eines Bündels von Stäbchen. Namentlich die Tatsache, daß sich die Menschen, nach Ausschöpfung der individualisierten Körperteile, mit einem Bündel von Stäbchen (die alle annähernd gleich waren) behalfen, gibt uns den Schlüssel zum Verständnis der Urform einer solchen „lebenden Skala". Es wird nämlich klar, daß sie im Anfang nicht zur Individualisierung der Zahlen benutzt wurde, sondern nur zur Feststellung der Anzahlgleichheit zweier Gesamtheiten, d. h. zur Herstellung einer umkehrbar eindeutigen Beziehung zwischen zwei verschiedenen Gesamtheiten von Dingen.

Überreste eines solchen Zählverfahrens erhielten sich bei vielen Völkern auch noch auf einer höheren Stufe der Entwicklung. So benutzten einige Völker für denselben Zweck eine Leine mit Knoten, andere benutzten Rosenkränze oder Kerbstöcke (hölzerne Stäbe mit Kerben).

So „schrieben" z. B. verschiedene peruanische Stämme die Zahlen mit Hilfe von Knotenleinen (sog. „Quipu", Abb. 1). Die Leinen wurden zu je vieren mit einer fünften Leine verbunden, auf der durch Knoten die Zahl vermerkt wurde, die die Summe der Zahlen auf den ersten vier Leinen war. Die Einer, Zehner und Hunderter einer gegebenen Zahl wurden jeweils durch verschieden geformte Knoten bezeichnet. Bei den Inka (11. bis 16. Jahrhundert n. Z.) wurde mittels solcher Quipu die ganze Buchhaltung „geschrieben". Derartige Knotenleinen dienten nur zum Schreiben der Zahlen.[1]) Die arithmetischen Operationen wurden mit Hilfe von Steinchen oder Maiskörnern ausgeführt.

In diesem Entwicklungsstadium hatte man aber von der Zahl noch keine so allgemeine Auffassung, daß man mit ihr alle untereinander anzahlgleichen Gesamt-

[1]) Herodot berichtet z. B. folgendermaßen von einer Anordnung, die Darius den Ioniern nach dem Übergang über den Fluß Istros während des von ihm gegen die Skythen unternommenen Feldzuges (6. Jahrhundert v. u. Z.) gegeben hat: „Danach band der Kaiser in seinen Leibriemen sechzig Knoten, ließ alle ionischen Fürsten zur Beratung kommen und sagte zu ihnen: Die von mir gefällte Entscheidung bezüglich der Brücke, Ionier, widerrufe ich; nehmt jetzt diesen Riemen und verfahret so: beginnt an jenem Tage, an welchem ich gegen die Skythen ziehe, und bindet jeden Tag einen Knoten auf; sollte ich in der Zwischenzeit nicht zurückgekommen sein, so fahrt, wenn die Anzahl von Tagen verstrichen ist, die durch die Knoten angegeben ist, zurück in die Heimat; bis zu jener Zeit aber beschützt die Brücke und gebt euch alle Mühe bei ihrer Verteidigung und Instandhaltung."

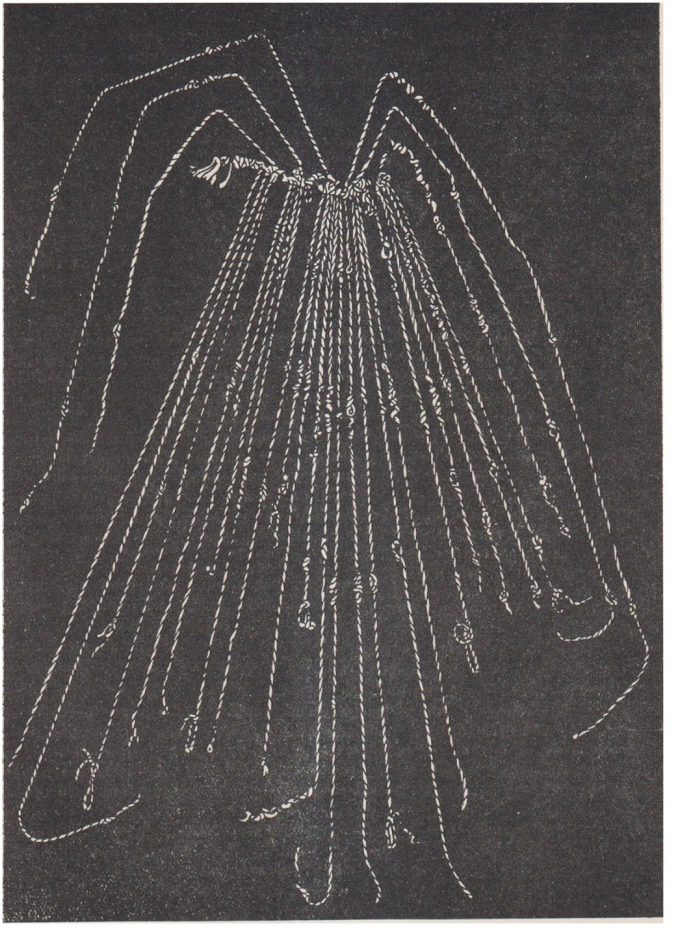


Abb. 1

heiten bezeichnete, sondern man begnügte sich damit, die Anzahlgleichheit zweier Gesamtheiten festzustellen.

Falls eine abzuzählende Gesamtheit eine geringe Anzahl von Gegenständen enthielt (≤ 20), wählte man gewöhnlich aus der Menge der zu dieser Gesamtheit anzahlgleichen Gesamtheiten einen festen Repräsentanten aus und sagte von den übrigen Gesamtheiten dieser Menge, daß sie genau so viele Gegenstände wie die ausgewählte enthalten. Um z. B. auszudrücken, daß in einer Gesamtheit fünf Gegenstände enthalten sind, sagte man, in dieser Gesamtheit seien so viele Gegenstände, wie Finger an einer Hand. Die gemeinsame Eigenschaft aller anzahlgleichen endlichen Mengen — ihre Anzahl — wurde durch die Anzahl einer speziellen Menge, eines Repräsentanten aus ihrer Gesamtheit ausgedrückt.

Es ist interessant, daß bei Stämmen, die auf der erwähnten Stufe der Entwicklung standen, dasselbe Verfahren auch bei der Bildung anderer Begriffe angewandt wurde. Z. B. gab es bei den Tasmaniern keine Worte zur Bezeichnung so allgemeiner Begriffe wie: hart, heiß, kalt, rund usw. Zur Bezeichnung der Härte sagten sie „wie Stein", um auszudrücken, daß ein Gegenstand rund ist, sagten sie „wie der Mond" oder „wie eine Kugel". Genau dasselbe gilt auch für die Bezeichnung der Farben. Auf dieser Stufe gab es auch keine solchen allgemeinen Begriffe wie Baum oder Fisch, sondern nur spezielle Wörter zur Bezeichnung der bekannten Baumarten und Fischsorten.

Diese Phase in der Geschichte der Entstehung der abstrakten Zahlen ist also dadurch charakterisiert, daß die abzuzählenden Mengen von Gegenständen mit Hilfe von Körperteilen — besonders der Finger und Zehen — oder mit Hilfe von Stäbchen, Knotenleinen usw. dargestellt wurden. Trotz ihrer äußersten Primitivität spielte diese Methode der Darstellung eine außerordentliche Rolle in der Entwicklung des Zahlbegriffs. Eine wesentliche Seite dieses Verfahrens ist nämlich, daß alle abzuzählenden Gesamtheiten mit Hilfe eines bestimmten Systems abgezählt werden.

Ein solches Zählverfahren führte in seiner weiteren Entwicklung zur Schaffung des Fünfer-Systems, des Dezimal-Systems und des Zwanziger-Systems. So sagten z. B. die Bewohner von Miraluga (einer Insel der Torres-Straße) „nabiget" für die Zahl 5, „nabiget-nabiget" für die Zahl 10, „nabikoku" für die Zahl 15 und „nabikoku-nabiget" für die Zahl 20. In ihrer Sprache bedeutete „get" die Hand und „koku" der Fuß. Dazu berichtet ein Forscher: „Man darf nicht denken, daß „nabiget" ein Name für die Zahl 5 ist; es drückt vielmehr nur aus, daß es sich gerade um soviel Gegenstände handelt, wie es Finger an der Hand gibt." Nach Mitteilung des bekannten russischen Forschungsreisenden N. N. Miklucho-Maklai zählen die Eingeborenen von Neu-Guinea folgendermaßen: „Ein beliebtes Zählverfahren besteht darin, daß der Papua einen Finger nach dem anderen umbiegt und dabei einen bestimmten Laut ausstößt, z. B. ‚be, be, be, . . .'. Hat er bis fünf gezählt, so sagt er ‚ibon-be' (ibon = Hand). Darauf biegt er die Finger der anderen Hand um und zählt wieder ‚be, be, . . .' bis er zu ‚ibon-ali' (ibon-ali = zwei Hände) gelangt. Dann fährt er fort und zählt solange ‚be, be, . . .' vor sich hin, bis er zu ‚samba-be' und ‚samba-ali' (ein Fuß, zwei Füße) gelangt. Ist es nötig, noch weiter zu zählen, so benutzt der Papua die Finger und Zehen eines Anderen."[1])

[1]) Н. Миклухо-Маклай, Путешествия, (N. Miklucho-Maklai, Reisen), Band 1, Verlag der Akademie der Wissenschaften der UdSSR, 1940, S. 280.

Ähnlich bezeichnen auch viele andere Stämme die Anzahlen. In diesem Entwicklungsstadium des Zahlbegriffs und des Rechnens war das Abzählen mit den Fingern (die sogenannte Instrumentalrechnung) weit verbreitet. Mit der Fingerrechnung hing, wie wir sahen, auch die Einteilung der Zahlen in „Finger" und „Gelenke" zusammen.

Alle bisher behandelten Entwicklungsstadien des Zahlbegriffs können ungefähr der Periode der Wildheit zugeschrieben werden.

Mit der weiteren Entwicklung der Gesellschaft vergrößerte sich auch der Kreis der abzuzählenden Gesamtheiten. Die einfache Feststellung der Anzahlgleichheit und die Handrechnung genügten den neuen Bedürfnissen der Horden nicht mehr; trotzdem erhielt sich die Methode der Fingerrechnung eine lange Zeit hindurch[1]) noch als Überrest. Die eigentliche Entwicklung, die schließlich zur Schaffung der natürlichen Zahlen führte, nahm indes eine andere Richtung. Das entstehende neue System kann als ein Gruppensystem oder als das Rechnen mit Hilfe von abgezählten Gesamtheiten (Stückzahlen) bezeichnet werden.

Ansätze für ein derartiges Zählen finden sich schon auf den früheren Entwicklungsstufen. So bemerkten Forscher bei den Bewohnern der westlichen Inseln der Torres-Straße „deutlich ausgeprägte Anlagen eines Zählens mit Gruppen von zwei Dingen, also mit Paaren". Dasselbe bemerkte ein anderer Forscher bei den Eingeborenen der Halbinsel Prince York. Bei diesen wurden gewisse Gegenstände vorzugsweise in Paaren zusammengefaßt, mit anderen dagegen wurde in Gruppen zu zehn oder hundert Stück gerechnet. Es wurden also bei Rechnungen die Gegenstände — je nach Art — immer in Gruppen einer festen Anzahl vereinigt. Ein Überrest eines solchen Zählverfahrens ist z. B. die Dutzend-Rechnung, die sich für verschiedene Dinge (Hemden, Stühle, Geschirr, Bleistifte, Gummibänder, Schreibfedern usw.) in Europa noch bis in unsere Tage erhalten hat. Hierbei war das Dutzend die Einheit des Rechnens; zwölf Dutzend bildeten ein Gros, zwölf Gros bildeten eine Masse.[2]) Derartige Zählmethoden waren besonders bei den Stämmen verbreitet, die auf den untersten beiden Stufen der Barberei standen. Im Laufe der Zeit erhielt jede derartige Gruppe eine eigene Benennung, die neben der Anzahl stets noch die Art der abgezählten Gegenstände ausdrückte. Die Gruppen dieser Art, mit deren Hilfe gerechnet wurde, wollen wir *Stück-Anzahlen* nennen.

Auf den Fidschi- und den Salomon-Inseln existieren Sammelnamen für zehn willkürlich zusammengestellte Dinge, die aber weder die Anzahl im einzelnen noch die Benennung der zusammengestellten Dinge ausdrücken. Auf den Fidschi-Inseln gibt es außerdem spezielle Bezeichnungen für 100 Kähne, 100 Kokosnüsse, 1000 Kokosnüsse usw. Sind zwei Gruppen von Dingen anzahlgleich, so kommt dies für gewöhnlich auch in den Benennungen der Stück-Anzahlen zum Ausdruck. So bedeutet bei den Eingeborenen Floridas „na-kua" 10 Eier, „na-banara" 10 Körbe mit Lebensmitteln, aber das Wort „na", welches unserem Zahlwort „Zehn" entspräche, wird einzeln nicht gebraucht. In einem der Dialekte der Indianer West-Kanadas bedeutet das Wort „tcha" soviel wie „drei Dinge", „tchane" bedeutet „drei Per-

[1]) Noch im Jahre 1529 erschien in Basel ein — allerdings schon lange vorher geschriebenes — Buch von Beda Venerabilis (der Ehrwürdige) (672—735), in welchem die Rechenmethoden an den Fingern dargelegt wurden, wobei sich diese Rechnungen auf alle Zahlen bis zu einer Million erstreckten. Es ist ferner bekannt, daß die chinesischen und mongolischen Kaufleute die Fingerrechnung noch in einer verhältnismäßig späten Zeit benutzten.

[2]) Einige Sudan-Stämme verwenden auch heute noch ein derartiges Zwölfer-System.

sonen", „tchat" bedeutet „dreimal", „tchatoen" bedeutet „an drei Stellen" usw. Es gibt aber dort kein Wort, welches die abstrakte Zahl 3 bezeichnen würde. Immerhin deutet das Vorhandensein ein und desselben Partikels in den Bezeichnungen verschiedener anzahlgleicher Gesamtheiten darauf hin, daß man in diesem Stadium der Entwicklung schon festzustellen beginnt, daß alle solche Gesamtheiten etwas gemein haben, nämlich die Anzahl. In diesem Stadium der Entwicklung schreibt man aber noch nicht jeder Gesamtheit von Dingen eine Anzahl zu, sondern nur solchen Gesamtheiten, die im Leben der Stämme (z. B. in der Wirtschaft) häufig gebraucht werden. Wenn die Pythagoreer im 6. Jahrhundert v. u. Z. erklärten, daß „alles Zahl ist", so können wir nur sagen, daß im betrachteten Stadium der Entwicklung nicht alles Zahl war.

Die Benennung der Zahlen spiegelt das Wesen der Zahlen in diesem Stadium wider: es waren keine abstrakten Zahlen. Nach und nach wurden die verschiedenen Stück-Anzahlen als neue Einheiten betrachtet, mit denen dann auch gerechnet wurde. Tatsächlich ist ja auch bereits darin, daß eine Gesamtheit von Dingen als eine neue Einheit aufgefaßt wird, die Möglichkeit enthalten, ein Bezeichnungssystem zu schaffen.

Im Laufe der Zeit wurden mittels solcher festen Stück-Anzahlen nicht nur diejenigen Stücke bezeichnet, für welche sie eingeführt waren, sondern auch Gegenstände, die den ursprünglichen in der Form oder dem Gebrauch nach ähnlich waren. Auf diese Weise bildeten sich in den Ursprachen mehrere Reihen von Zahlwörtern. So gibt es in der Sprache der Tschimsian (Britisch-Columbia) sieben verschiedene Reihen von Zahlwörtern, die gebraucht werden zum Zählen (1) von unbestimmten Dingen, (2) von flachen Dingen, (3) von runden Gegenständen und der Zeiteinteilung, (4) von Menschen, (5) von länglichen Dingen (diese Zahlen werden z. B. mittels des unserem Worte „Baum" entsprechenden Wortes gebildet), (6) von Booten und (7) von Maßen. Überreste des Zählens mit Stück-Anzahlen wurden noch in Altgriechenland beobachtet. So erörterte ARISTOTELES in seiner „Metaphysik" die Frage, ob die Einheiten in einer Zahl alle untereinander gleich seien und ob sie in verschiedenen Zahlen gleich oder verschieden sind. Noch DIOPHANT (3. Jahrhundert u. Z.) setzte hinter jede Ziffer, die er schrieb, ein M — den ersten Buchstaben des Wortes $\mu o\nu\alpha\varsigma$ (Einheit) — er sah also noch jede Zahl als eine gewisse Anzahl von untereinander gleichen Einheiten an.

Unter dem Einfluß des Handels begann eine Zahlenreihe alle anderen zu verdrängen. Es war diejenige, die zur Zählung des Geldes diente (das in der ersten Zeit in Muscheln, Vieh o. dgl. bestand). Bei den Joruba (Negerstamm in Zentral-Afrika) haben heute noch die Zahlwörter — wie ein Forscher berichtete — zweierlei Bedeutung, erstens als reine Zahlen und zweitens als Stück-Anzahlen von „Kauri" (Muscheln), die bei den Joruba die Rolle des Geldes spielen. So entstanden im Laufe der Zeit universelle Zahlen, d. h. Zahlen, mit deren Hilfe man beliebige Dinge zählen konnte.

Die Stück-Anzahlen bilden die Urtypen unserer Zahlen, allerdings nur der Zahlen mit individueller Bezeichnung. Hatte man z. B. spezielle Bezeichnungen für Gruppen von zehn oder zwanzig Gegenständen einer gewissen Art oder rechnete man in Dutzenden, so konnten auf die oben beschriebene Weise keine algorithmischen Bezeichnungen für Zahlen wie 17 oder 19 entstehen. Wenn übrigens alle Zahlen nach dem beschriebenen Schema aus Stück-Anzahlen entstanden wären, so hätten sie nebeneinander als untereinander nicht verbundene Begriffe bestanden, und die

anzahlmäßigen Beziehungen zwischen ihnen wären völlig unbekannt geblieben. Wir werden zeigen, daß die algorithmischen Zahlwörter dadurch aus den individuellen Zahlwörtern entstanden sind, daß an den Zahlen mit individuellen Bezeichnungen gewisse Operationen ausgeführt wurden.

Beim Studium der Sprache der Klamathen (Indianerstamm in Nordamerika) und der Sprachen gewisser Stämme Britisch-Columbias zeigte sich, daß diese beim Zählen spezielle Verben — von den Forschern klassifizierende Verben genannt — benutzten, die eine bestimmte Art der Verteilung bezeichneten. Wurden die Rechnungen an einer festen Gruppe von nur zehn Gegenständen durchgeführt, so wurden die Zahlwörter nicht von diesen Verben begleitet. Diese Tatsache erklärten die Forscher als eine Besonderheit des Rechnens der Indianer. Wollten die Indianer aber mehr als zehn Gegenstände zusammenzählen, so bildeten sie auf der Erde zunächst einen Haufen oder eine Reihe aus zehn dieser Gegenstände und begannen dann mit dem elften Gegenstand einen neuen Haufen oder eine neue Reihe. Die klassifizierenden Verben begleiteten auch nicht die Zahlwörter, die ganze Zehner bezeichneten, sie dienten vielmehr nur dazu, die Zahlen zwischen zwei ganzen Zehnern anzuordnen. So mußten die Indianer, um das Vorhandensein von 26 Gegenständen auszudrücken, sagen: „zu zweimal zehn Früchten (oder anderen Gegenständen) füge ich sechs hinzu". Somit entstanden dort die algorithmischen Bezeichnungen tatsächlich als Resultat gewisser Operationen an individuellen Anzahlen. Diese Operationen waren anfangs keineswegs arithmetischer Natur, sondern rein motorisch. Spuren davon erhielten sich in vielen Sprachen. So werden in der russischen Sprache die Zahlen von 11 bis 19 gesprochen als die entsprechende Zahl „auf zehn" («надесять»): два-на-десять (zusammengezogen zu двенадцать, zwei-auf-zehn = zwölf), пять-на-десять (zusammengezogen zu пятнадцать, fünf-auf-zehn = fünfzehn) usw. Hier muß die Partikel «на» (auf) jedenfalls im Sinne von «подожить на» (legen zu) verstanden werden.[1])

Eine gute Illustration des Zählens mittels bestimmter Anordnungen von Dingen bilden die Zahlzeichen der Azteken aus dem 15. und 16. Jahrhundert. Diese bezeichneten durch ⋮ : | · die Zahl 6, durch ⋮ : | · · die Zahl 7 usw. Offensichtlich bestand bei ihnen die grundlegende Gruppe aus fünf Gegenständen. Im Schriftbild trennte ein Strich eine solche Gruppe von den anderen. Der Strich selbst hatte keine ziffernmäßige Bedeutung.

Später ging der unmittelbar motorische Charakter der Operationen mehr und mehr verloren, und ihre arithmetische Bedeutung trat klar in den Vordergrund. So erscheint z. B. in den finnisch-ugrischen Sprachen die Zahl 8 als Differenz aus 10 und 2. Die Zahl 8 wird in diesen Sprachen als „zwei, zehn", die Zahl 80 als „zwei, hundert" und die Zahl 800 als „zwei, zehn, hundert" gesprochen (hier bedeutet „zehn, hundert" die Zahl 1000).

In diesem Entwicklungsstadium ist die Zahlenreihe noch keineswegs homogen. Vielmehr stellen die Zahlen mit individueller Bezeichnung gewissermaßen Inseln dar, von denen aus die Zahlen mit algorithmischer Bezeichnung in verschiedenster Weise gebildet werden. Eine bevorzugte Rolle spielte bei der Bildung der algorithmischen

[1]) Die Partikel «на» hat im Russischen auch die Bedeutung, die das deutsche Wörtchen „mal" bei der Multiplikation hat. Diese Bedeutung ist hier also offenbar nicht gemeint. — *Anm. d. wissenschaftl. Red.*

Bezeichnungen die Addition, jedoch waren in gleicher Weise auch Subtraktion und Multiplikation beteiligt. So besitzen die schon erwähnten Joruba folgendes Bezeichnungssystem:

$$11 = 10 + 1,\ 12 = 10 + 2,\ \ldots,\ 15 = 10 + 5,$$
$$16 = 20 - 4,\ 17 = 20 - 3,\ \ldots,\ 19 = 20 - 1.$$

Die Zahl 20 besitzt eine neue individuelle Bezeichnung, mit deren Hilfe die weiteren Zahlwörter gebildet werden, und zwar werden bei ihrer Bildung sowohl Addition als auch Subtraktion und Multiplikation benutzt. Die Zahl 70 erscheint bei ihnen als $20 \cdot 4 - 10$, die Zahl 190 als $20 \cdot 10 - 10$ usw. Analoge Bezeichnungen finden sich auch in der lateinischen Sprache, in der z. B. die Zahl XVIII als „duo de viginti", d. h. 2 von 20, gelesen wurde.

Die Zahlenreihe war in diesem Entwicklungsstadium aber nicht nur inhomogen, sondern sie war auch endlich, wenngleich sie bei der allmählichen Entwicklung des Zahlbegriffs immer länger wurde. Dabei schoben sich die Worte „viel" und „unabzählbar", die anfangs die Zahlen $\geqq 3$ oder $\geqq 10$ bezeichneten, immer weiter hinaus, indem sie Zahlen $\geqq 100$, $\geqq 1000$ usw. bezeichneten.[1]) Der nächste Schritt bestand dann darin, daß diese Wörter, die ursprünglich eine unbestimmte Anzahl bezeichneten, zu einer Bezeichnung der Zahlen 100 oder 1000 wurden — je nach der Entwicklungsstufe des Bezeichnungssystems des betreffenden Volkes. So bedeutet ein und dasselbe Wort auf der Insel Mengon die Zahl 3, auf den Fidschi-Inseln die Zahl 100 und bei den Maori auf Neuseeland die Zahl 1000. Dies erklärt sich daraus, daß die Maori und die Völker der genannten Inseln — obwohl sie auf verschiedenen Kulturstufen stehen — doch ungefähr dieselbe Sprache sprechen.

Die Reihe der natürlichen Zahlen blieb noch eine lange Zeit hindurch endlich. Aus dem Vorwort zur „Sandrechnung" von ARCHIMEDES geht hervor, daß selbst den Griechen des 3. Jahrhunderts v. u. Z. noch nicht bekannt war, wie man be-

---

[1]) Eines der markantesten Beispiele solcher Grenzzahlen ist die Zahl 40, die in der russischen Sprache zunächst zur Bezeichnung einer unbestimmten großen Anzahl diente. Auf eine solche Rolle dieser Zahl weist sowohl ihre individuelle Benennung in der russischen Sprache als auch ihre als Überrest erhaltene Anwendung bei der Bezeichnung unbestimmter großer Anzahlen von Dingen hin: сорок сороков церквей (vierzig mal vierzig Kirchen), сорок сороков чёрных соболей (vierzig mal vierzig schwarze Zobel). Auf dieselbe Rolle der Zahl 40 weisen auch eine Reihe von religiösen Bräuchen und Aberglauben hin, die mit ihr verbunden sind; z. B. wurde der vierzigste Bär als der letzte im Leben des Jägers angesehen:

„Vierzig Bären nahm er auf den Spieß, beim einundvierzigsten versagte er."

(N. A. NEKRASSOW

In späterer Zeit, als die Zahl 40 längst aufgehört hatte, eine „Grenzzahl" zu sein, spielte sie noch eine große Rolle im russischen Maßsystem: ein Pud enthielt vierzig Pfund, ein „Vierzigerfaß" vierzig Eimer usw.

Die Zahl 40 spielte auch bei vielen Völkern des nahen Ostens die Rolle einer „Grenzzahl". Dies fand z. B. seine Widerspiegelung in dem berühmten armenischen Epos „David Sassunski":

„Msra-Melik stieg in eine Grube.
Da warf man vierzig Büffelfelle auf ihn,
vierzig Mühlsteine warf man auf ihn . . ."

Und das Epos schließt mit einem Gedenken an alle seine Helden:

„Vierzig mal wollen wir uns im Guten an unsere großen Urväter erinnern,
an Sanassar und seinen Bruder Bogdassar wollen wir uns erinnern,
und an Mger den Älteren wollen wir zurückdenken."

liebig große Zahlen ausdrücken kann, z. B. „Zahlen, die größer sind als die Anzahl der Sandkörnchen in einer Kugel, deren Radius gleich der Entfernung der Sonne von einem Fixstern ist". Das wesentliche Ziel des Werkes von ARCHIMEDES bestand gerade in der Begründung eines systematischen Verfahrens zur Konstruktion und zur mündlichen Bezeichnung beliebig großer Zahlen.

§ 2. Nicht-positionelle Bezeichnungssysteme

Als die Schrift entstand, vollzog sich die Entwicklung der Zahlzeichen etwa folgendermaßen: Jede der Zahlen, die durch ein individuelles Zahlwort bezeichnet wurde (sog. Knotenzahl), die also als individueller Begriff existierte, wurde als Basis für ein gewisses Stellensystem beim Zählen genommen. Als Basis des ersten Systems nahm man die kleinste Zahl, die durch ein individuelles Zahlwort bezeichnet wurde. Mit dieser Zahl wurde solange mittels Addition von Einheiten, aber auch mittels Verdoppelung, Verdreifachung usw. gerechnet — kurz, es wurden solange algorithmische Bezeichnungen gebildet, — bis man die zweite Knotenzahl erreichte. Darauf wurde das zweite Stellensystem der Zählung begonnen, dessen Basis diese zweite Knotenzahl war. Die algorithmischen Zeichen dieses zweiten Systems waren Kombinationen aus den ersten beiden Individualzeichen. Mit Hilfe derartiger Zahlzeichen gelangte man bis zur dritten Knotenzahl, die als Basis des dritten Stellensystems diente usw.[1])

Dieses Schema sah mitunter auch etwas anders aus: Die mittels einer Knotenzahl gebildeten algorithmischen Zahlen reihten sich nicht nur nach oben an diese an, sondern auch nach unten, indem zu ihrer Bildung neben der Addition auch die Subtraktion kleinerer Knotenzahlen benutzt wurde.

Eine Schreibweise, die nach dem ersten Schema gebildet wurde, besaßen die Systeme vom Typ der ägyptischen Hieroglyphen (Tafel 1). Individuelle Zeichen waren hier  für die Zahl 1,  für die Zahl 10,  für die Zahl 100 und  für die Zahl 1000[2]), wobei das Symbol für die Zahl 1000 ursprünglich eine unbestimmte Anzahl bezeichnete.

Diese Schreibweise spiegelt die Auffassung wider, daß jedes individuelle Zahlzeichen eine neue Individualität bezeichnet. Aus der Form der Zahlzeichen ist z. B. nicht zu ersehen, daß jede nachfolgende Knotenzahl das Zehnfache der vorangehenden ist. Alle individuellen Zahlzeichen haben absoluten Charakter; be-

[1]) Der Leser, der die CANTORsche Theorie der transfiniten Ordinalzahlen kennt, bemerkt leicht die Ähnlichkeit zwischen dieser Bildungsmethode der Folge der natürlichen Zahlen und dem von CANTOR benutzten Verfahren. In der Tat gibt es auch bei CANTOR zwei Prinzipien zur Bildung von transfiniten Ordinalzahlen: 1. Bildung von Vielfachen und Addition von Einheiten; 2. Einführung neuer Zahlen als Grenzwerte der vorangehenden. Der Unterschied ist nur der, daß in der Folge der natürlichen Zahlen diese neue Zahl schon gegeben ist und immer in endlich vielen Schritten erreicht werden kann, es ist einfach die folgende Knotenzahl.

[2]) Man vermutet, daß die Hieroglyphe eine Maßleine darstellt, die in 100 Teile geteilt war und daß die Hieroglyphe die Blüte einer Lotosblume abbildet.

Zahlzeichen verschiedener Völker

	ägyptische			assyrisch-babylonische	phönizische	syrische	palmyrenische	griech.-herodianische	römische
	Hieroglyphen 1	hieratische 2	demotische 3	4	5	6	7	8	9
1	𓏺	𓂝	𓏺	𒁹	𐤀	𐡀	𐡠	I	I
2	𓎆	𓎆	𓎆	𒁹𒁹	𐤁	𐡁	𐡡	II	II
3	𓎇	𓎇	𓎇	𒁹𒁹𒁹	𐤂	𐡂	𐡢	III	III
4	𓎈	𓎈	𓎈	𒐈	𐤃	𐡃	𐡣	IIII	IV
5	𓎉	𓎉	𓎉	𒐉	𐤄	𐡄	𐡤	Γ	V
6	𓎊	𓎊	𓎊	𒐊	𐤅	𐡅	𐡥	ΓI	VI
7	𓎋	𓎋	𓎋	𒐋	𐤆	𐡆	𐡦	ΓII	VII
8	𓎌	𓎌	𓎌	𒐌	𐤇	𐡇	𐡧	ΓIII	VIII
9	𓎍	𓎍	𓎍	𒐍	𐤈	𐡈	𐡨	ΓIIII	IX
10	𓎆	𓎆	𓎆	𒌋	𐤉	𐡉	𐡩	Δ	X
11	𓎇	𓎇	𓎇	𒌋𒁹	𐤊	𐡊	𐡪	ΔI	XI
15	𓎈	𓎈	𓎈	𒌋𒐉	𐤋	𐡋	𐡫	ΔΓ	XV
20	𓎉	𓎉	𓎉	𒌋𒌋	𐤌	𐡌	𐡬	ΔΔ	XX
30	𓎊	𓎊	𓎊	𒌋𒌋𒌋	𐤍	𐡍	𐡭	ΔΔΔ	XXX
40	𓎋	𓎋	𓎋	𒌋𒌋𒌋𒌋	𐤎	𐡎	𐡮	ΔΔΔΔ	XL
50	𓎌	𓎌	𓎌	𒐼	𐤏	𐡏	𐡯	𐅁	L
60	𓎍	𓎍	𓎍	𒐼𒁹	𐤐	𐡐	𐡰	𐅁Δ	LX
70	𓎎	𓎎	𓎎	𒐼𒁹𒁹	𐤑	𐡑	𐡱	𐅁ΔΔ	LXX
80	𓎏	𓎏	𓎏	𒐼𒁹𒁹𒁹	𐤒	𐡒	𐡲	𐅁ΔΔΔ	LXXX
90	𓎐	𓎐	𓎐	𒐼𒐼	𐤓	𐡓	𐡳	𐅁ΔΔΔΔ	XC
100	𓍢	𓍢	𓍢	𒐼𒐼𒐼	𐤔	𐡔	𐡴	H	C
200	𓍢𓍢	𓍢𓍢	𓍢𓍢	𒐼𒐼𒐼𒐼	𐤕	𐡕	𐡵	HH	CC
400	𓍢𓍢𓍢𓍢	𓍢𓍢𓍢𓍢	𓍢𓍢𓍢𓍢	𒐼𒐼𒐼𒐼𒐼				HHHH	CD
500	𓍢𓍢𓍢𓍢𓍢	𓍢𓍢𓍢𓍢𓍢	𓍢𓍢𓍢𓍢𓍢	𒐼𒐼𒐼𒐼𒐼𒐼				𐅀	D
1000	𓆼	𓆼	𓆼	𒐼𒐼𒐼𒐼𒐼𒐼𒐼				X	M
10 000	𓂭			𒐼𒐼𒐼𒐼𒐼𒐼𒐼𒐼				M	
10^5	𓁨								
10^6	𓁩								
10^7	𓁪								

Tafel 1

zeichnet zehn Einer, aber auf keinen Fall zehn Zehner oder zehn Hunderter. Die algorithmischen Zahlen erhält man im ägyptischen System völlig einheitlich mittels einer einzigen arithmetischen Operation — der Addition. Beispielsweise wird die Zahl 333 in diesem System durch 𓍢𓍢𓎆𓎆𓎆 wiedergegeben.

Das ägyptische System ist noch deshalb von Interesse, weil dort die Zahl Zwei eine besondere Rolle spielte. Sie diente wahrscheinlich ursprünglich als Basis des Bezeichnungssystems. „Drei" war schon der Ausdruck einer unbestimmten Anzahl. Dies ist daraus ersichtlich, daß die Ägypter die Mehrzahl von gewissen Gegenständen oder Begriffen dadurch ausdrückten, daß sie unter die entsprechenden Hieroglyphen drei Striche setzten.

Überreste des dyadischen Systems spiegeln sich auch in der ägyptischen Multiplikationsmethode wider, die in Verdoppelungen und Additionen bestand. Wollten die Ägypter z. B. eine gewisse Zahl n mit 15 multiplizieren, so verfuhren sie (schematisch dargestellt) folgendermaßen:

$$n \cdot 15 = n\,(1 + 2 + 2^2 + 2^3) = n \cdot 1 + n \cdot 2 + n \cdot 2^2 + n \cdot 2^3,$$

d. h., sie stellten den Multiplikator im dyadischen System dar und multiplizierten daraufhin mit jedem Summanden der dyadischen Entwicklung einzeln. Spuren des dyadischen Systems finden sich auch in der ägyptischen Bruchrechnung, von der später noch die Rede sein wird.

Systeme vom Typ der ägyptischen Hieroglyphen sind das phönizische, das palmyrenische, das kretische und das griechisch-herodianische oder attische System (s. Tafel 1).

Die attischen oder herodianischen[1]) Zahlzeichen entstanden, wie schon der Name sagt, in Attika. Die ältesten Schriften in diesem System stammen aus dem 6. Jahrhundert v. u. Z. Das Zeichen für die Einheit ist hier, wie schon im ägyptischen System, ein senkrechter Strich, mit dessen Hilfe die Zahlen bis 4 bezeichnet wurden, Die Zahl 5 wurde durch das Symbol Γ, die Zahl 10 durch Δ, die Zahl 100 durch H, die Zahl 1000 durch X und die Zahl 10000 durch M bezeichnet. Wie jetzt einwandfrei festgestellt wurde (erstmalig machte schon WALLIS im 17. Jahrhundert hierauf aufmerksam), sind diese Zeichen die Anfangsbuchstaben der entsprechenden Zahlwörter. In der Tat heißt fünf auf griechisch *πέντε* (in den attischen Gebieten diente Γ als Bezeichnung des Buchstaben Π, daher wurde dort ΓΕΝΤΕ geschrieben), zehn heißt ΔΕΚΑ, hundert heißt ΗΕΚΑΤΟΝ, tausend heißt ΧΙΛΙΟΙ, und zehntausend heißt ΜΥΡΙΟΙ. Die Zahlen 50, 500, und 5000 schrieb man als Kombinationen aus dem Zeichen für die Zahl 5 und den Zeichen für 10, 100 und 1000: $\Gamma^{\Delta} = 50$, $\Gamma^{H} = 500$, $\Gamma^{X} = 5000$. Alle übrigen Zahlen wurden nach einem rein additiven Prinzip geschrieben. So erschien die Zahl 325 als ΗΗΗΔΔΓ. Diese Schreibweise behauptete sich in Attika bis in das 1. Jahrhundert u. Z., obwohl sie in den anderen griechischen Ländern schon lange vorher durch das bequemere ionische Bezeichnungssystem verdrängt worden war.

Die zweite Art der Darstellung der natürlichen Zahlen finden wir bei den römischen Zahlzeichen. Von ähnlichem Typ (Anwendung der Subtraktion) war, wie wir sahen, auch das System der Joruba. Allerdings befanden sich die Römer auf einer anderen

[1]) HERODIAN war ein griechischer Geschichtsschreiber aus dem 2. bis 3. Jahrhundert u. Z. Aus seinen Werken erhielt die westeuropäische Wissenschaft die erste Kenntnis vom attischen Bezeichnungssystem.

Zahlzeichen verschiedener Völker

	chinesische			Kharosthi-Ziffern	Ziffern der Höhleninschriften von Nasik	Ziffern der Azteken	Ziffern der Maya
	alte 1	kommerzielle 2	wissenschaftliche 3	4	5	6	7
0		○	○				
1	一	𝍠	I	𐨑	—	·	•
2	二	𝍡	II	𐨒	=	··	••
3	三	𝍢	III	𐨓	≡	∴	•••
4	四	𝍣	IIII	×	𑁒	::	••••
5	五	𝍤	IIIII	I×	𑁓	⁘	—
6	六	𝍥	T	II×	𑁔	⁘·	—•
7	七	𝍦	Π		𑁕	⁘··	—••
8	八	𝍧	ΠI	××	𑁖	⁘∴	—•••
9	九	𝍨	ΠII		𑁗	⁘::	—••••
10	十	𝍩	IO	𐨔	∝	◇	=
15	十五	𝍩𝍤	I≡			◇⁘	≡
20	二十	𝍡𝍩	IIO	𐨕	θ	𝙿	
30	三十	𝍢𝍩	IIIO			𝙿◇	
40	四十	𝍣𝍩	IIIIO		×	𝙿𝙿	
50	五十	𝍤𝍩	IIIIIO	𐨖		𝙿𝙿◇	
60	六十	𝍥𝍩	TO	𐨗		𝙿𝙿𝙿	
70	七十	𝍦𝍩	ΠO	𐨘	𑁘	𝙿𝙿𝙿◇	
80	八十	𝍧𝍩	ΠIO			𝙿𝙿𝙿𝙿	
90	九十	𝍨𝍩	ΠIIO			𝙿𝙿𝙿𝙿◇	
100	百	𝍪	IOO	𐨙	𑁙		
200	二百	𝍡𝍪	IIOO	𐨚	𑁚		
400	四百	𝍣𝍪	IIIIOO				
500	五百	𝍤𝍪	IIIIIOO		𑁛		
1000	千	𝍫	IOOO		𑁜		
8000	八千	𝍧𝍫	ΠOOO		𑁝		
10000	万	𝍬	IOOOO				

Tafel 2

Entwicklungsstufe als die Joruba. Das römische Bezeichnungssystem ist sehr alten Ursprungs, wobei bekannt ist, daß früher das Prinzip der Subtraktion in einem noch breiteren Maße angewandt wurde. So wurde z. B. die Zahl 8 durch „IIX" bezeichnet. Selbst die Form der „Ziffern" entlehnten die Römer den früheren Einwohnern Italiens, den Etruskern. Die Zahl 10 wurde von den Etruskern durch + oder × bezeichnet, wobei die Römer die zweite Form entlehnten. Die Zahl 5 schrieben die Etrusker ∨ oder ∧ — dies war die Hälfte des Zeichens für die Zahl 10. Die Zahl 50, die von den Etruskern ↓ geschrieben wurde, wandelte sich anfangs in ⊥, dann in ⊥ und schließlich in L um. Das römische Zeichen „C" für die Zahl 100 entwickelte sich aus dem etruskischen Zeichen ⊕, welches sich anfangs in Ⓒ und schließlich in „C" umwandelte usw.

Es ist interessant zu bemerken, daß die Römer neben der Addition und Subtraktion noch ein originelles Prinzip der „Division" benutzten. So ist das Zeichen für die Zahl 5 die Hälfte des Zeichens für die Zahl 10. Deutlicher tritt dieses Prinzip in dem nichtpositionellen Zwanziger-System der Azteken hervor. Die Zahl 400 wird dort durch  bezeichnet, 300 = , 200 =  und 100 = .

Den Positionssystemen näher stehen schon die Bezeichnungssysteme mit multiplikativer Schreibweise. Hierzu gehört das alte chinesische System, in welchem das multiplikative Prinzip bereits bei den Zehnern beginnt (Tafel 2), sowie das indische System der Kharosthi-Ziffern (Tafel 2), in welchem das multiplikative Prinzip bei den Hundertern anfängt, u. a. Im alt-chinesischen System wurden die Zahlen 20 oder 30, schematisch dargestellt, als 2, 10 bzw. 3, 10 geschrieben usw. Die Zahlen 100, 1000 und 10000 besaßen individuelle Bezeichnungen. Die übrigen zusammengesetzten Zahlen wurden nach einem ähnlichen Schema wie die Vielfachen von 10 bezeichnet. So wurde die Zahl 333 schematisch als 3, 100, 3, 10, 3 geschrieben. In der einheitlichen Bezeichnung der ganzen Zehner, Hunderter usw. kann man schon eine erste Annäherung an ein Positionssystem erblicken. Darüber wird ausführlicher in dem Paragraphen, der den Ursprung der Positionssysteme behandelt, zu sprechen sein. Wir weisen hier nur noch darauf hin, daß auch in unserer sprachlichen Bezeichnungsweise das multiplikative Prinzip eine große Rolle spielt (zwanzig = zwei, — zehn; dreihundert = drei, — hundert usw.).

§ 3. Die alphabetischen Bezeichnungssysteme für die Zahlen

Die vollkommenste Art der nicht-positionellen Systeme stellen — abgesehen von den Systemen mit multiplikativer Schreibweise — die Systeme dar, die zur Bezeichnung der Zahlen irgendein Alphabet benutzten. Als Beispiele hierfür können das ionische (alt-griechische), das slawische (kyrillische und glagolitische), das hebräische, das arabische und auch das georgische und armenische Bezeichnungssystem dienen.

Auch das ägyptisch-hieratische Bezeichnungssystem (s. Tafel 1) ist seiner Struktur nach mit den alphabetischen Bezeichnungssystemen verwandt. Es bestand schon 2000 Jahre v. u. Z. neben dem hieroglyphischen System im alten Ägypten. Angewendet wurde es in Wirtschaftsberichten und anderen offiziellen Dokumenten, aber auch in mathematischen Papyri (die beiden ältesten bekannten mathematischen Pa-

pyri, der Moskauer Papyrus und der Papyrus Rhind, sind im hieratischen System geschrieben), während die Hieroglyphen — wenn man sich so ausdrücken darf — in „Paradefällen", etwa für Aufschriften auf Denkmälern und Obelisken, benutzt wurden. Die hieratische Schrift entstand aus den Hieroglyphen ursprünglich durch Verkürzung und Verschmelzung einzelner Zeichen, wie das bei jeder Schnellschrift der Fall ist, wobei die Zahlen von 1 bis 9, welche in Hieroglyphen durch einfache Wiederholung des Zeichens für die Zahl 1 bezeichnet wurden, besondere individuelle Zeichen erhielten. Es traten hier also erstmalig spezielle Ziffern für die Zahlen der ersten Dekade auf. Solche individuelle Zeichen bildeten sich auch für die ganzen Zehner von 10 bis 90, für die ganzen Hunderter von 100 bis 900, für die Tausender, Zehntausender und Hunderttausend heraus.

Hierin unterscheidet sich also das hieratische Rechensystem prinzipiell vom hieroglyphischen. Auf die Bedeutung dieses neuen Bezeichnungsprinzips, welches man etwa Ziffernprinzip nennen kann, wies 1911 der bedeutende russische Historiker W. W. Bobynin hin.[1]) Unlängst stellte Ch. Boyer in einer Tafel die Entwicklung des Ziffernprinzips (in seiner Terminologie: ciphirisation) und des Positionsprinzips zusammen. „Die Einführung des Ziffernprinzips durch die Ägypter" — schreibt Boyer — „stellt einen entscheidenden Schritt in der Entwicklung des Zählens dar, und dieser Beitrag ist in seiner Tragweite dem Beitrag der Babylonier absolut vergleichbar, die das Positionssystem einführten."[2]) Wie bedeutend die Rolle des Ziffernprinzips ist, wurde schon von Bobynin richtig gezeigt. Hingegen ist die Behauptung Boyers eine starke Übertreibung. Ein Positionssystem hat — wie wir sehen werden — gegenüber einem System vom Typ des hieratischen unbestreitbare Vorteile, selbst dann, wenn im betrachteten Positionssystem die Anzahl der Ziffern klein ist, wie dies bei den Babyloniern der Fall war. Immerhin stellte das Ziffernprinzip einen bedeutenden Fortschritt etwa im Vergleich zum hieroglyphischen Bezeichnungsprinzip dar. Wir werden noch ausführlicher auf seine Vorzüge zu sprechen kommen, wenn wir die Systeme mit alphabetischen Zahlzeichen betrachten. Denselben Typus wie das ägyptisch-hieratische System hatte auch das alte singhalesische Rechensystem. Etwa im 6. Jahrhundert v. u. Z. breitete sich in Ägypten die demotische Schrift aus, die eine Variante der hieratischen Schrift ist. In der griechisch-römischen Epoche war sie in Ägypten allgemein verbreitet.

Ein alphabetisches System für die Zahlen war wahrscheinlich erstmalig in Griechenland in Gebrauch. Die ältesten bekannten Inschriften, die in einem derartigen System geschrieben sind, stammen aus der Mitte des 5. Jahrhunderts v. u. Z. aus Halikarnassos in Kleinasien. In allen alphabetischen Systemen werden die Zahlen von 1 bis 9, die ganzen Zehner und die ganzen Hunderter durch individuelle Zeichen, und zwar durch die aufeinanderfolgenden Buchstaben eines gewissen Alphabetes bezeichnet (Tafel 3). Im griechischen und in den slawischen Bezeichnungssystemen wurde über die Buchstaben, die Zahlen bezeichnen sollten, die also als Ziffern dienten, ein Strich gesetzt, damit man sie von den gewöhnlichen Buchstaben unterscheiden konnte. Alle Zahlen bis 999 wurden mittels des Additionsprinzips als Kombinationen der siebenundzwanzig individuellen Zeichen geschrieben. Da es im gewöhnlichen griechischen Alphabet nur 24 Buchstaben gab, wurden noch drei alte Buchstaben

[1]) В. В. Бобынин, Отзыв о сочинениях Н. М. Бубнова, (W. W. Bobynin, Bemerkungen zu den Werken N. M. Bubnows), St. Petersburg 1911.

[2]) Ch. Boyer, Fundamental Steps in the Development of Numeration, Isis, 1944. Nr. 100, Band 35, S. 158.

als Zahlzeichen verwandt: ϛ (Digamma) für die Zahl 6, ϙ (Koppa) für die Zahl 90 und ϡ (Sampi) für die Zahl 900 (siehe erste Spalte der Tafel 3).

Die Zahl 444 wurde also im griechischen Bezeichnungssystem als $\overline{\upsilon\mu\delta}$ geschrieben. Im römischen System hätte sie die Form CDXLIV und im attischen die Form HHHHΔΔΔΔIIII gehabt.

Schon dieses Beispiel zeigt den unbestreitbaren Vorteil der alphabetischen Bezeichnungssysteme. Daß sie tatsächlich einen neuen, höheren Typ darstellen, beweist auch der Lauf der geschichtlichen Entwicklung. Nachdem in den griechischen Handelskolonien die ionische Bezeichnung aufgekommen war, breitete sie sich auch rasch in Attika aus. Dort verdrängte sie die geweihte Tradition des herodianischen Systems, so daß das ionische System auch bald von denjenigen Gebieten befürwortet wurde, die lange Zeit hindurch in offiziellen Dokumenten nur die Anwendung des herodianischen Systems gestattet hatten.

Wir sehen hierin eine weitere Bestätigung der marxistischen These von der Entwicklung im Kampf der Gegensätze, von der Unüberwindlichkeit des Neuen. Ungeachtet aller Hindernisse, entgegen der Kraft der Tradition überwand das alphabetische System das herodianische und breitete sich in Attika immer weiter aus. Nachdem es unter PTOLEMÄUS in Alexandrien üblich geworden war, verdrängte es das herodianische System gänzlich.

Eine Reihe von Historikern der Mathematik, u. a. M. CANTOR und G. HANKEL, nahmen sogar an, die alphabetische Bezeichnungsweise der Zahlen stelle im Vergleich zur attischen Bezeichnungsweise einen Rückschritt dar. Hieraus zogen sie die Schlußfolgerung, daß die Griechen, denen im allgemeinen ein bedeutender Platz in der Entwicklung der europäischen Kultur zukommt, für die Vervollkommnung des Bezeichnungssystems durchaus nichts geleistet hätten. Diese Ansicht ist indes völlig haltlos. Die Forderungen, die ein bequemes Bezeichnungssystem erfüllen muß, sind nämlich folgende: 1.) Die Zahlen müssen sich schnell und bequem schreiben lassen; 2.) Mit den in diesem System geschriebenen Zahlen muß möglichst mühelos gerechnet werden können; 3.) Das System muß leicht erlernbar sein; 4.) Es muß prinzipiell möglich sein, beliebig große Zahlen in diesem System zu schreiben. Wir sahen, daß das ionische System die erste Forderung erfüllte, speziell ist das Schriftbild der Zahlen im ionischen System wesentlich kürzer als im attischen System. Um festzustellen, ob Rechnungen im ionischen System mühsam sind, eignete sich der französische Mathematik-Historiker P. TANNERY im Jahre 1882 die ionische Bezeichnungsweise an und führte mit ihrer Hilfe die Rechnungen durch, die in der „Kreismessung" des ARCHIMEDES unumgänglich sind. Er gewann dabei die Überzeugung, daß im ionischen System praktische Vorteile verborgen liegen, von welchen man früher kaum etwas vermuten konnte. Er stellte ferner fest, daß in diesem System die Rechenoperationen nicht viel länger sind als unsere nach modernen Rechenschemata ausgeführten. Dieser Meinung P. TANNERYS schließt sich auch T. CHISS an. In der oben erwähnten Arbeit entwickelt BOYER hieraus den Gedanken, daß die leichte Ausführbarkeit der Rechnungen weniger eine Folge des Positionsprinzips als des benutzten Rechenschemas ist. Für nicht zu große Zahlen ist das im gewissen Sinne richtig. Es ist jedoch zu bemerken, daß die modernen Multiplikations- und Divisionsschemata selbst auf dem Positionsprinzip der Zahldarstellung beruhen, d. h., dabei wird in irgendeiner Form dasselbe Positionsprinzip benutzt wie auch in unserer Bezeichnung.

Alphabetische Bezeichnungssysteme für die Zahlen

	ionisch 1	slavisch: kyrillisch 2	slavisch: glagolitisch 3	hebräisch 4	syrisch 5	arabisch 6	georgisch 7	armenisch 8
1	ᾱ	ҁ	ⰰ	א	ܐ	ا	ა	Ա
2	β̄	в	ⰱ	ב	ܒ	ب	ბ	Բ
3	γ̄	г	ⰲ	ג	ܓ	ج	გ	Գ
4	δ̄	д	ⰳ	ד	ܕ	د	დ	Դ
5	ε̄	є	ⰴ	ה	ܗ	ه	ე	Ե
6	ϛ̄	ѕ	ⰵ	ו	ܘ	و	ვ	Զ
7	ζ̄	з	ⰶ	ז	ܙ	ز	ზ	Է
8	η̄	и	ⰷ	ח	ܚ	ح	თ	Ը
9	θ̄	ѳ	ⰸ	ט	ܛ	ط	ი	Թ
10	ῑ	і	ⰹ	י	ܝ	ى	ი	Ժ
20	κ̄	к	ⰺ	כ	ܟ	ك	კ	Ի
30	λ̄	л	ⰻ	ל	ܠ	ل	ლ	Լ
40	μ̄	м	ⰼ	מ	ܡ	م	მ	Խ
50	ν̄	н	ⰽ	נ	ܢ	ن	ნ	Ծ
60	ξ̄	ѯ	ⰾ	ס	ܣ	س	ო	Կ
70	ο̄	о	ⰿ	ע	ܥ	ع	პ	Հ
80	π̄	п	ⱀ	פ	ܦ	ف	ჟ	Ձ
90	ϟ̄	ҁ	ⱁ	צ	ܨ	ص	რ	Ղ
100	ρ̄	р	ⱂ	ק	ܩ	ق	ს	Ճ
200	σ̄	с	ⱃ	ר	ܪ	ر	ტ	Մ
300	τ̄	т	ⱄ	ש	ܫ	ش	უ	Յ
400	ῡ	ѵ	ⱅ	ת	ܬ	ت	ფ	Ն
500	φ̄	ф	ⱆ	ך ק		ث	ქ	Շ
600	χ̄	х	ⱇ	ך ר		خ	ღ	Ո
700	ψ̄	ѱ	ⱈ	ך ש		ذ	ყ	Չ
800	ω̄	ѡ	ⱉ	ך ת		ض	შ	Պ
900	ϡ̄	ц	ⱊ	ץ		ظ	ჩ	Ջ
1000	,ᾱ	,ҁ	ⱋ	א		ع	ც	Ռ
2000	,β̄	,к				غ	ძ	Ս
3000	,γ̄	,г				ف	წ	Վ
4000	,δ̄	,д				ق	ჭ	Տ
5000	,ε̄	,є				ك	ხ	Ր
6000	,ϛ̄	,ѕ				ل	ჯ	Ց
7000	,ζ̄	,з				م	ჰ	Ւ
8000	,η̄	,и				ن	ჱ	Փ
9000	,θ̄	,ѳ				ه	ჲ	Ք
10000	M̄	Ⓐ				و	ჳ	Օ
20000	M̿	Ⓚ				ى	ჴ	Ֆ

Tafel 3

Auch die Tatsache, daß sich die Griechen an Stelle unserer 10 Ziffern 27 Zeichen zu merken hatten, kann nicht als wesentlicher Mangel des alphabetischen Systems gelten, da erstens diese Einprägung nur einmal vorgenommen werden mußte und zweitens für die Zahlen keine wesentlich neuen Zeichen eingeführt wurden. Ihre Bezeichnungen prägten sich zusammen mit dem Alphabet ein.

Gegen die alphabetischen Systeme erhob M. CANTOR noch den folgenden Einwand: Im Dezimal-System folgt aus $2 + 3 = 5$ sofort, daß auch $20 + 30 = 50$ ist. Hingegen ist etwa im ionischen System nicht zu sehen, daß aus $\beta + \gamma = \varepsilon$ sofort $\varkappa + \lambda = \nu$ folgt. Daraus ergibt sich nach seiner Meinung, daß die Griechen sich viel mehr Sonderfälle der Addition und Multiplikation zu merken hatten als wir. Hierzu ist zu bemerken, daß der Prozeß des Zählens sich ja nicht nur auf Zeichen, sondern auch auf Worte erstreckt.[1]) Wenn wir z. B. das Einmaleins lernen, prägen wir uns nicht ein, daß das Zeichen „2", verbunden durch das Zeichen für die Multiplikation mit dem Zeichen „3", das Zeichen „6" ergibt, sondern wir lernen diese Operation in Worten: „zwei mal drei ist sechs", was wir notfalls in Zeichen übersetzen können. Nun war aber bei den Griechen die mündliche Bezeichnung der Zahlen auch nicht schwieriger als die unsrige, deshalb konnten sie ebenso einfach wie wir aus „zwei mal drei gleich sechs" auf „zwanzig mal dreißig gleich sechshundert" schließen.

Für verhältnismäßig kleine Zahlen war also das alphabetische System sowohl hinsichtlich der Schreibweise der Zahlen als auch hinsichtlich des Rechnens einem Positionssystem so gut wie gleichwertig. Hingegen konnte man in einem alphabetischen System hinreichend große Zahlen zunächst nicht niederschreiben. Zu diesem Zweck wurde zu den bisher geschilderten Bezeichnungsprinzipien ein neues Prinzip hinzugefügt.

Die Versuche, in einem alphabetischen System auch Zahlen zu schreiben, die größer als 1000 sind, führten zu Bezeichnungen, die man als Keime eines Positionssystems ansehen kann. So verwandte man in Griechenland zur Bezeichnung der Tausender dieselben Bezeichnungen wie für die Einer, nur daß man zur Unterscheidung links unten einen kleinen Strich heransetzte (s. Tafel 3):

$$\bar{\alpha} = 1, \ \bar{\alpha} = 1000, \ \ldots, \ \bar{\theta} = 9, \ \bar{\theta} = 9000.$$

Mit Hilfe der Buchstaben, die links unten mit einem Strich versehen waren, konnten die Griechen alle Zahlen bis 9999 ausdrücken. Die Zahl 10000 wurde durch $\overset{\alpha}{\mathrm{M}}$ bezeichnet; $\overset{\beta}{\mathrm{M}}$ bedeutete 20000 usw. Hierbei wurde also schon das Multiplikationsprinzip benutzt. An Stelle von $\overset{\beta}{\mathrm{M}}$ schrieb man auch β^{M} oder $\mathrm{M}\beta$. Im ersten Falle setzte man auch häufig für das M einfach einen Punkt hinter den entsprechenden Buchstaben; dann wurde z. B. die Zahl 43458 folgendermaßen geschrieben: $\delta.\overline{\gamma\upsilon\eta}$. Diese letzte Schreibweise, die u. a. von DIOPHANT benutzt wurde, kam einem Positionsprinzip am nächsten.

Die größte Zahl, die man mit Hilfe des ionischen Systems ausdrücken konnte, war $10^6 - 1$. Wenngleich es den Anschein hat, als ob mit der alphabetischen Schreibweise schon der Gedanke aufgekommen ist, den Wert einer Ziffer nicht nur durch ihre Form, sondern auch durch ihre Stellung im Schriftbild einer Zahl auszudrücken, so stellen doch die alphabetischen Systeme nicht den Anfang der Positionssysteme

[1]) M. Я. Выгодский, Арифметика и алгебра в древнем мире, (M. J. WYGODSKI, Arithmetik und Algebra im Altertum), Moskau-Leningrad 1941, Seite 184.

dar. Nur die beiden größten Mathematiker des Altertums, ARCHIMEDES (287—212) und APOLLONIUS (265 ?—170), kamen dem eigentlichen Sinn eines Positionssystems recht nahe. So stellte ARCHIMEDES in seiner „Sandrechnung" eine Zählung in „Oktaven" auf. Alle Zahlen von 1 bis 10^8—1 wurden in der ersten Oktave vereinigt. Die Zahl 10^8 wurde als neue Zähleinheit genommen. Die Zahlen von 10^8 bis 10^{16}—1 bildeten dann die zweite Oktave usw. Dabei wurden die Zahlen der zweiten, dritten und folgenden Oktaven genauso wie die Zahlen der ersten Oktave bezeichnet. Eine ähnliche Gruppierung gab APOLLONIUS in seinem „Schnellzähler", der uns bedauerlicherweise nicht erhalten geblieben ist. Er benutzte an Stelle der Oktaven Quarten (10^4). Alle Zahlen von 1 bis 10^4—1 wurden in der ersten Quarte vereinigt; die Zahlen von 10^4 bis 10^8—1 bildeten die zweite Quarte usw.[1]).

Dennoch gelangte weder ARCHIMEDES noch APOLLONIUS zu einer einheitlichen positionsmäßigen Bezeichnung aller Zahlen mit Hilfe von zehn Zeichen (etwa der ersten zehn Buchstaben des Alphabets) und noch weniger zur Einführung der Null.

Dieser Umstand läßt sich — wie auch M. J. WYGODSKI erwähnt — dadurch erklären, daß das „ionische Bezeichnungssystem innerhalb des Zahlbereiches, in welchem die griechischen Mathematiker zu operieren hatten, den Anforderungen der Praxis vollständig genügte".[2]) Aus diesem Grunde änderten die Griechen ihr Bezeichnungssystem für die ganzen Zahlen auch dann noch nicht ab, als sie schon ihre Brüche mit Hilfe des von den Babyloniern übernommenen Sexagesimalsystems bezeichneten, wobei auch für die Zahl Null ein Zeichen benutzt wurde. Man kann dies auch so erklären, daß das Positionssystem, welches bekanntlich schon lange vor MAXIMUS PLANUDES (13. Jahrhundert) in Byzanz aufgestellt worden war, dort noch nicht genügend Verbreitung gefunden hatte und die alphabetische Bezeichnung allgemein gebräuchlich blieb.

Spuren der alphabetischen Bezeichnung der Zahlen erhielten sich bis in die Gegenwart. Wir numerieren häufig mit Hilfe des Alphabets die „Punkte" eines Vortrages, einer Resolution usw., ähnlich wie damals die vierundzwanzig Lieder der „Iliade" durchnumeriert wurden. Jedoch wird bei uns diese alphabetische Bezeichnung nur für die Ordinalzahlen benutzt. Die Grund- oder Kardinalzahlen werden grundsätzlich nicht durch Buchstaben bezeichnet. Erst recht operieren wir nicht mit Zahlen, die im alphabetischen System geschrieben sind.

Auch im alten Rußland wurden die Zahlen mit Hilfe des Alphabets bezeichnet. Die Bezeichnung mittels der slawischen Alphabete kam etwa im 10. Jahrhundert u. Z. auf. Die Einführung einer solchen Bezeichnung wird bereits dem Begründer des slawischen Alphabets, KYRILLUS (gestorben 869), zugeschrieben. Dieses Bezeichnungssystem der Zahlen war nach dem Muster des ionischen Systems aufgebaut, wie es ehemals bei den Byzantinern in Gebrauch war. Insbesondere dienten nur jene Buchstaben des kyrillischen Alphabets als Zahlzeichen, die einem gewissen Buchstaben des griechischen Alphabets entsprachen. So bezeichnete z. B. der zweite Buchstabe des kyrillischen Alphabets, „Buki" (), keine Zahl (s. Tafel 3). Die Zahl 2 wurde durch den Buchstaben „Wedi" () bezeichnet, der dem Buchstaben β

[1]) Ein ähnliches Prinzip wird auch in dem von APOLLONIUS eingeführten Multiplikationsschema angewendet, das unserem heutigen völlig analog ist. Die Multiplikation zweier Zahlen, die ein Vielfaches von zehn oder hundert sind, wurde dort so ausgeführt, daß man sie auf die Multiplikation ihrer „Wurzeln", d. h. auf die Multiplikation der Zahlen zurückführte, die die Anzahl der Zehner oder Hunderter in ihnen ausdrücken.

[2]) Vgl. Anmerkung 1 auf S. 25: loc. cit. S. 192.

des griechischen Alphabets entsprach, während „Buki" kein Urbild im griechischen Alphabet hatte. Der Buchstabe „Fita" (ϴ), der im kyrillischen Alphabet an vorletzter Stelle stand, bezeichnete die Zahl 9, weil er nämlich dem griechischen Buchstaben θ entsprach, der im ionischen System die Zahl 9 bezeichnete. Bei dem zweiten slawischen Alphabet, das zur Bezeichnung der Zahlen verwendet wurde — dem glagolitischen Alphabet —, finden wir diese Besonderheit nicht.[1]) Hier wurden die Zahlen in streng alphabetischer Reihenfolge bezeichnet. In beiden Systemen wurden im Text die Buchstaben oder Buchstabenverbindungen, die Zahlen bezeichnen sollten, durch ein darüber gesetztes ~ (Abkürzungszeichen in der slawischen Kirchenschrift) kenntlich gemacht. In den westeuropäischen Ländern wurden während dieser Zeit und auch später ausschließlich die römischen Zahlzeichen benutzt, die einem primitiveren Bezeichnungssystem entsprechen.

Zur Bezeichnung der Tausender wurden im kyrillischen System dieselben Buchstaben benutzt wie für die Einer, nur setzte man links unten an sie das Zeichen ҃.

In den slawischen Sprachen bildeten sich zwei Systeme bei der Benennung der höheren Dezimalordnungen heraus: die Benennung als kleine Zahl, die bis 10^6 reichte, und die Benennung als große Zahl oder als große slawische Zahl, die Zahlen bis 10^{48} oder 10^{49} bzw. sogar bis 10^{50} erfaßte („mehr als der menschliche Verstand fassen kann"). Dabei bezeichneten ein und dieselben Wörter in beiden Systemen verschiedene Zahlen. So bedeutete eine „Unmenge" im kleinen System die Zahl 10000, hingegen im großen System eine Million. Eine „Legion" war im ersten System gleich 10 Unmengen, also 100000, im zweiten eine Unmenge Unmengen, also 10^{12}; eine „Leoder" bestand im ersten System aus 10 Legionen, war also eine Million, dagegen im großen System aus einer Legion Legionen (10^{24}). Die weitere Zählung ging über zehn, hundert usw. Leoder bis zu hunderttausend Unmengen Legionen Leoder (10^{47}). Als nächste Einheit diente eine Leoder Leoder (10^{48}), die auch „Krähe" genannt wurde. Die Zahl 10^{50} nannte man mitunter „Spiel".

Aus einer handschriftlichen Grammatik aus dem 17. Jahrhundert sind uns auch die alten Ziffern für die Einheiten höherer Ordnung, die „großen slowenischen Ziffern", bekannt. Waren die Buchstaben des Alphabets, die die Zahlen von 1 bis 9 bezeichneten, mit einem Kreis ○ umgeben, so bezeichneten sie die entsprechende Anzahl von Unmengen, entsprechend diente ein Kreis aus Punkten ◯ zur Kennzeichnung von Legionen und ein Kreis aus Strichen ◯ zur Bezeichnung von Leoder. Das Zeichen ֍ diente zur Bezeichnung einer Krähe und das Zeichen ֎ zur Bezeichnung eines Spiels.[2])

Zweifelsohne sind die beiden Bezeichnungssysteme, die uns aus Handschriften des 17. Jahrhunderts bekannt sind, bedeutend früher entstanden. In den russischen mathematischen Handschriften des 17. Jahrhunderts wurde, nachdem das alphabetische Bezeichnungssystem verdrängt war, bereits unser heutiges verwendet. Mit dem alphabetischen Bezeichnungssystem verschwanden auch die oben angegebenen Namen für die höheren Dezimalordnungen.

[1]) Das glagolitische Alphabet ist eines der slawischen Alphabete. Seine Herkunft ist bis heute noch nicht vollständig geklärt. Man kann annehmen, daß das glagolitische Alphabet ein Vorläufer des kyrillischen ist. Über frühere slawische Zahlbezeichnungen ist nichts Zuverlässiges bekannt.

[2]) В. В. Бобынин, Очерки истории развития физико-математических знаний в России (W. W. Bobynin, Skizzen zur Entwicklungsgeschichte der physikalisch-mathematischen Wissenschaften in Rußland), 1. Aufl., Moskau 1886, Seite 45—47.

Auch in Westeuropa besaßen die Zahlwörter lange Zeit hindurch eine doppelte Bedeutung, die davon abhängig war, in welchem System man die Zahlen betrachtete. Das Wort „Million" z. B. wurde zuerst im 14. Jahrhundert in Italien zur Bezeichnung der „großen Tausend", das ist 1000^2, eingeführt. Ursprünglich diente es wahrscheinlich zur Bezeichnung eines konkreten Maßes von Fässern mit Gold. Im 15. und 16. Jahrhundert breitete sich dieses Wort auch in den anderen europäischen Ländern aus. Der französische Mathematiker NICOLAS CHUQUET führte um 1500 die Worte „Billion", „Trillion", „Quadrillion", ..., „Nonillion" als Benennung für die Potenzen der Million: 1000000^2, 1000000^3, 1000000^4, ..., 1000000^9 ein. In Frankreich wurden daneben die Zahlen ungefähr in der Mitte des 17. Jahrhunderts in Perioden zu je drei Ziffern geteilt. Dabei erhielt das Wort „Billion" die Bedeutung von 10^9 an Stelle von $1000000^2 = 10^{12}$ früher. Die Worte „Trillion", Quadrillion" usw. bedeuteten entsprechend 10^{12} 10^{15}, ... Jedoch bezeichnen in England, Deutschland und anderen nord-, europäischen Ländern auch heute noch diese Worte die Zahlen 10^{12}, 10^{18}, 10^{24}, ...

§ 4. Stellen- oder Positionssysteme

Das älteste uns bekannte Bezeichnungssystem, das nach dem Stellen- oder Positionsprinzip aufgebaut war, ist das Sexagesimalsystem des alten Babylonien, welches annähernd 2000 Jahre v. u. Z. aufkam. Die Babylonier schrieben alle Zahlen mit Hilfe zweier Zeichen, einem „Keil" ▼ mit dem Zahlwert 1 und einem „Winkelhaken" ◀, mit dem Zahlwert 10. Die Zahlen bis 60 wurden als Folgen aus diesen beiden Zeichen geschrieben, und zwar nach demselben additiven Prinzip, auf dem sich z. B. auch das ägyptische Bezeichnungssystem aufbaute. Die Zahl 32 besaß dementsprechend folgende Darstellung: ◀◀▼▼▼.

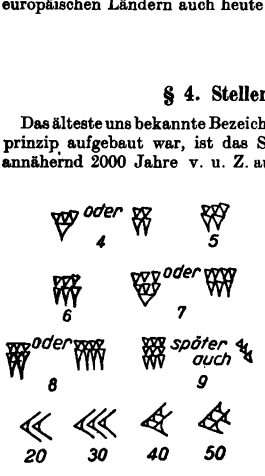


Abb. 2

Geschrieben wurden die Zahlen von links nach rechts, wobei stets zuerst die Zehner und dahinter die Einer gesetzt wurden. Waren in einer Zahl mindestens vier Einer oder Zehner enthalten, so verband man die entsprechenden Zeichen zu einer zusammenhängenden Gruppe (Abb. 2).

Die Zahl 60 wurde wiederum durch das Zeichen ▼ bezeichnet, sie bildete also eine Einheit höherer Ordnung. Für die Zahlen von 60 bis 3600—1, die Vielfache von 60 sind, wurden dieselben Bezeichnungen wie für die Zahlen von 1 bis 59 benutzt, wobei jetzt jedes einzelne Zeichen den sechzigfachen Wert hatte. Also wurde z. B. die Zahl 82 folgendermaßen geschrieben: ▼◀◀▼▼.

Diese Zeichenverbindung konnte aber auch $1\frac{22}{60}$ oder $82 \cdot 60$, allgemein $82 \cdot 60^{\pm k}$, bedeuten. Darüber hinaus bezeichnete sie aber auch die Zahl $60^2 + 22$, allgemein

jede Zahl der Form $60^{\pm k} + 22 \cdot 60^{\pm n}$ usw. Die Positionsschreibweise im Sexagesimalsystem hatte also keinen absoluten Charakter. Der wirkliche Wert einer geschriebenen Zahl mußte aus dem Zusammenhang entnommen werden. Dieser nicht-absolute Charakter der Schreibweise ist durch das Fehlen eines Zeichens für die Null bedingt.

Das Fehlen der Null war in den Anfängen der Entwicklung des Sexagesimalsystems, als man nur mit verhältnismäßig kleinen Zahlen zu operieren hatte, sicher nicht so fühlbar, wie es uns heute anmutet. Man überlegt sich leicht, daß zur eindeutigen Darstellung der Zahlen von 1 bis 3600 im babylonischen Zahlsystem die Null nur 59mal erforderlich wäre. In unserem heutigen Dezimalsystem benötigen wir zum gleichen Zweck die Null 917mal. Rechnet man jedoch mit Zahlen, die größer als 3600 sind, so wächst im babylonischen System das Bedürfnis nach der Null erheblich. Daher ist es nicht verwunderlich, daß sich in späteren Texten, in welchen die Babylonier auf Grund der Anforderungen der Praxis — vor allem der Astronomie — schon mit bedeutend größeren Zahlen operierten, Trennungszeichen  finden. Diese dienten zur Kennzeichnung der Lücken in der sexagesimalen Entwicklung der betrachteten Zahl.[1]) Der absolute Wert der Zahl mußte indes immer noch dem Zusammenhang entnommen werden, da Trennungszeichen niemals am Ende einer geschriebenen Zahl verwendet wurden.

Das Ziffernsystem der Babylonier unterscheidet sich also von unserem heutigen dezimalen Positionssystem (wenn wir einmal vom Unterschied der Basen 60 bzw. 10 absehen) in den beiden folgenden Punkten:

1.) Bei den Babyloniern war das Positionsprinzip nicht vollkommen ausgebaut.

2.) Die Positionsschreibweise der Babylonier hatte wegen des Fehlens der Null keinen absoluten Charakter.

Das Sexagesimalsystem der Babylonier spielte in der Mathematik und in der Astronomie eine große Rolle. Spuren dieses Systems lassen sich bis in unsere Zeit verfolgen. So teilen wir noch heute die Stunde in 60 Minuten, die Minute in 60 Sekunden usw. Ebenso teilen wir — wie ehemals die Babylonier — den Kreis in 360 Grade.

Die Frage nach der Herkunft dieses ersten Positionssystems hat viele Jahre hindurch das Interesse der Wissenschaftler erregt. Es gibt eine ganze Reihe einander widersprechender Hypothesen, die das Aufkommen dieses Systems erklären wollen.

So nahm M. CANTOR anfänglich an, daß die Sumerer (Urbewohner des Euphrattals) das Jahr in 360 Tage und Nächte teilten und daher das Sexagesimalsystem astronomischen Ursprungs sei. Nun ist aber in allen uns erhalten gebliebenen Texten aus dem alten Babylonien nur von Sonnenjahren zu 365 Tagen und Mondjahren zu 354 oder 355 Tagen die Rede. Diese Tatsache wie auch andere Bemerkungen der Kritiker veranlaßten CANTOR, seine Hypothese später zurückzunehmen.

Nach einer Hypothese von G. KEWITSCH lebten im Euphrattal zwei Völker, von denen das eine ein Dezimalsystem, das andere ein System mit der Basis 6 besaß (die Basis 6 leitet KEWITSCH aus der Fingerrechnung ab, wobei er annimmt, daß die geballte Faust die Zahl 6 bezeichnet). Durch Verschmelzung beider Systeme soll dann die „Kompromiß"-Basis 60 entstanden sein. Oder nach E. LÖFFLER: „Durch ihre Veranlagung zur Spekulation beschäftigten sich die sumerischen

[1]) Die ältesten bekannten Texte, in denen das Trennungszeichen verwendet wurde, stammen aus der Perserzeit (6. bis 5. Jahrhundert v. u. Z.).

Priester schon in einer sehr frühen Zeit mit Zahlenspielen. Dabei bemerkten sie, daß unter allen Zahlen unterhalb 100 die Zahl 60 die größte Anzahl von kleinen Teilern, wie 2, 3, 4, 5, 6, besitzt. Diese Entdeckung regte sie dazu an, für *wissenschaftliche Zwecke ein System mit der Basis 60 zu schaffen.*"[1])

Alle diese Hypothesen sind völlig unhistorisch, man könnte ihre Anzahl nach Belieben vermehren[2]). Die zuletzt genannte Hypothese ist sogar antihistorisch, da die Zahlensysteme niemals und nirgends — weder von Gelehrten noch von irgendeiner anderen Klasse und erst recht nicht von Priestern — „für wissenschaftliche Zwecke" geschaffen worden sind. Die Zahlensysteme sind stets das Resultat einer langwierigen historischen Entwicklung gewesen, das Resultat des Schaffens einiger oder aller Völker.

Glaubwürdiger ist eine von O. NEUGEBAUER[3]) im Jahre 1927 dargelegte Hypothese über die Herkunft der Basis 60 und des babylonischen Positionssystems. Ihr liegen u. a. folgende wesentliche Fakten zugrunde:

1.) Im vierten Jahrtausend v. u. Z: hatten die Sumerer ein nicht-positionelles Dezimalsystem mit individuellen Bezeichnungen für die Zahlen 1, 10 und 100. Dabei bedeutete das Wort für die Zahl 100 („*meu*") gleichzeitig eine unbestimmte Anzahl. In dieser Beziehung hatten also die Sumerer ungefähr dasselbe Bezeichnungssystem wie zur gleichen Zeit die Ägypter, nur war es noch äußerst unentwickelt. Daneben hatten die Sumerer Individualbezeichnungen für die Brüche $\frac{1}{2}$, $\frac{1}{3}$ und $\frac{2}{3}$.

2.) In der letzten, dem Auftreten der mathematischen Texte vorangehenden Blütezeit der sumerischen Kultur trat eine neue individuelle Zahl auf: 60. Sie wurde durch dasselbe Zeichen wie die Zahl 1 bezeichnet, aber mit einem deutlich größeren Zahlzeichen als die gewöhnlichen Einheiten. In dieser Entwicklungsperiode wurde das Ziffernsystem sexagesimal, obwohl sich in ihm noch ein individuelles Zeichen für die Zahl 100 findet. Hingegen war das System noch deutlich nicht-positionell. Dieses System besaß eine dem ägyptischen Hieroglyphensystem analoge Struktur, nur waren in Ägypten 1, 10 und 100 Knotenzahlen, während es in Babylonien 1, 10 und 60 waren. Natürlich gab es noch keine Trennungszeichen; sie waren ja auch noch nicht nötig, da die Knotenzahlen absolute Bezeichnungen besaßen.

3.) Im folgenden wird die Zahl 60 durch dasselbe Zeichen bezeichnet wie die Zahl 1 und das einheitliche Bezeichnungsprinzip auch auf die Brüche ausgedehnt. Jedoch erhielten sich noch lange Zeit hindurch die individuellen Bezeichnungen für die Brüche $\frac{1}{2}$, $\frac{1}{3}$, und $\frac{2}{3}$.

Zur Erklärung dieser Fakten betrachtet NEUGEBAUER das babylonische Maßsystem. Insbesondere wendet er seine Aufmerksamkeit den Gewichtsmaßen zu, da diese hier — wie fast überall — die Grundlage für das Geldwesen bildeten.

---

[1]) Е. Леффлер, Цифры и цифровые системы культурных народов в древности и в новое время, (E. LÖFFLER, Ziffern und Ziffernsysteme der Kulturvölker im Altertum und in der Neuzeit), Odessa 1913, Seite 33.(Vgl. auch Archiv Math. Phys. 17 (1911), S. 135. *Anm. d. wissenschaftl. Red.*)

[2]) Siehe die Bemerkung von I. J. TIMTSCHENKO in Ф. Кэджори, История элементарной математики (F. CAJORI, A History of Mathematics, 2. Aufl., New York 1938), Odessa 1918, Seite 313—312

[3]) Siehe O. NEUGEBAUER, Vorlesungen über Geschichte der antiken mathematischen Wissenschaften, Band I: Vorgriechische Mathematik, Berlin 1934, S. 93—110.

NEUGEBAUER kommt zu der Überzeugung, daß bei den beiden Bevölkerungselementen des Stromgebietes von Euphrat und Tigris, den Sumerern und den Akkadern (semitische Völkergruppe, die später das sumerische Bevölkerungselement gänzlich verdrängte), zunächst zwei verschiedene Gewichtssysteme in Gebrauch waren. Die Grundeinheit des einen Systems war der Schekel (Šekel), die des anderen die Mine. Die Reihe der Maßzahlen für diese Einheiten begann mit $\frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1, \ldots, 10$.

Beide Systeme waren ursprünglich Dezimalsysteme. Eine Umrechnungszahl zwischen Schekel und Mine gab es anfangs nicht. Die Maßsysteme existierten unabhängig nebeneinander, und zwar wurde das eine System für kleine Gewichte, das andere für große Gewichte benutzt. Die Entstehung eines zentralisierten babylonischen Staates mit einheitlichem Wirtschaftssystem machte verhältnismäßig frühzeitig eine Normierung des Geld-Gewichts-Systems notwendig. Bei der Aneinanderfügung der Gruppe der kleineren Gewichte und der der großen Gewichte waren naturgemäß Beziehungen aufzustellen, die die Bruchteile $\left(\text{also } \frac{1}{2}, \frac{1}{3} \text{ und } \frac{2}{3}\right)$ der großen Einheit, der Mine, in ganze Vielfache der kleinen Einheit, des Schekels, überführten. Nach der Meinung NEUGEBAUERS wurden die Bruchteile der Mine ganzen Zehnern des Schekels gleichgesetzt, so daß auf diese Weise das Verhältnis zwischen dem Schekel und der Mine auf 1 : 60 festgesetzt wurde (dieses Verhältnis dürfte auch ungefähr dem Verhältnis der ursprünglichen Gewichte entsprochen haben).

Da anfangs kleine und große Einheiten in der Schreibweise entweder durch die Größe der Zahlzeichen oder durch explizite Angabe der Einheit unterschieden wurden, war die Bezeichnung zunächst absolut.

Im Laufe der Zeit breitete sich die Zahlenskala der Gewichte auch in anderen Gebieten aus. Hier hatte es nun keinen Sinn mehr, hinter die Maßzahlen die Bezeichnung eines Gewichtes zu setzen. Von diesem Gesichtspunkt aus erscheint die Positionsbezeichnung als „nichts anderes als der Verzicht auf die ausdrückliche Nennung der Maßeinheiten". Dieser Prozeß verlief im Anfang völlig unbewußt. Es ist sonst nicht zu erklären, warum nicht von vornherein die fehlenden Trennungszeichen eingeführt worden sind. Die Tatsache, daß in alten Texten derartige Trennungszeichen fehlen, beweist vielmehr, „daß man zunächst trotz der formal gleichen Zahlzeichen immer noch bei den einzelnen Stufen an die konkrete Maßbezeichnung dachte".[1])

Denselben Prozeß können wir übrigens auch in fast allen heutigen Sprachen beobachten. So sagen wir statt „2 Mark 20 Pfennige" fast immer „zwei, zwanzig". Die Dimensionen der einzelnen Größen wird hinzugedacht.[2])

Die Trennungszeichen für innere ausfallende Stellen tauchten erst in dem Augenblick auf, als das sexagesimale Positionssystem bewußt zu einem mathematischen System ausgestaltet worden war. Dieses späte Stadium der Entwicklung findet seinen Ausdruck nur in rein mathematischen Texten und erscheint damit als Vollendung des nicht-absoluten sexagesimalen Positionssystems der Babylonier.

---

[1]) Vgl. O. NEUGEBAUER, a. a. O., S. 108.
[2]) Vgl. M. J. WYGODSKI, a. a. O., S. 69.

Nach der dargelegten Hypothese sind also bei der Entstehung des babylonischen Positionssystems die folgenden beiden Hauptetappen wesentlich gewesen:

1.) Die Einführung einer zahlenmäßigen Beziehung zwischen zwei unabhängig nebeneinander bestehenden Maßsystemen[1]).

2.) Der Verzicht auf eine explizite Angabe der Maßeinheiten.

NEUGEBAUER nimmt an, daß dies ganz allgemein die Etappen sind, die bei der geschichtlichen Entwicklung der Positionssysteme durchlaufen wurden: „So erscheint also das sexagesimale Positionssystem der mathematischen Texte als ganz naturgemäßes Endprodukt einer langen Entwicklung, die in prinzipieller Hinsicht gar nicht ausgezeichnet ist vor analogen Prozessen in anderen Kulturen." Allein auf einen günstigen Umstand ist es zurückzuführen, daß ein solches System zuerst in Babylonien geschaffen wurde, nämlich den, „daß die Normierung des Geld-Gewichts-Systems in so eine frühe Entwicklungsphase fällt, daß einerseits die auch im Babylonischen vorliegende, ursprünglich dezimale Struktur die Hundert noch nicht wesentlich überschritten hatte und andererseits die natürlichen Bruchteile auch noch auf die kleine Gruppe $\frac{1}{2}, \frac{1}{3}, \frac{2}{3}$ beschränkt waren"[2]), während kompliziertere Rechenprozesse sich noch nicht herausgebildet hatten.

Es ist interessant, an der dargelegten Hypothese zu verfolgen, wie ihr Verfasser den Entwicklungsprozeß des sexagesimalen Positionssystems mit der Entwicklung der gesellschaftlichen Ökonomie, des Maßsystems, der Geldwirtschaft u. dgl. m. zu verknüpfen sucht. Jedoch kann man deshalb diese Hypothese noch nicht als feststehende Theorie ansehen. M. J. WYGODSKI, der diese Hypothese im großen und ganzen befürwortet, zeigte, daß in ihr durchaus noch „hypothetische Elemente enthalten sind: Die Annahme über die Einführung der Geld-Gewichts-Äquivalente ist bisher durch keinen positiven Beweis bestätigt worden."[3]) Eine Reihe weiterer Einwände gegen diese Hypothese brachte I. N. WESSELOWSKI[4]) in einem Referat vor (z. B. das Vorhandensein einer großen Anzahl nicht-positioneller Niederschriften in der Handelskorrespondenz aus der Zeit nach HAMMURABI, das mit der Annahme über die Entstehung des Positionssystems aus dem Geld-Gewichts-Maßsystem schlecht übereinstimmt). Der Orientalist F. THUREAU-DANGIN glaubte annehmen zu dürfen, daß — entgegen der Meinung NEUGEBAUERS — dem Auftreten des sexagesimalen Systems in der Metrologie das sexagesimale Zählen vorangegangen sei. Jedenfalls fehlt uns bis heute eine allgemeingültige Erklärung für die Herkunft des babylonischen Systems.

Auf alle Fälle war das Auftreten der positionellen Zahlenbezeichnung einer der wichtigsten Marksteine der Kulturgeschichte. Dieses Auftreten kann keinesfalls

---

[1]) Daß im betrachteten Fall diese beiden Maßsysteme ursprünglich bei zwei verschiedenen Völkern (Sumerern und Akkadern) in Gebrauch waren, ist hier nicht wesentlich.

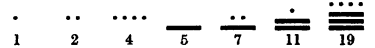
[2]) O. NEUGEBAUER, a. a. O., S. 109.

[3]) M. J. WYGODSKI, a. a. O., S. 69—70.

[4]) I. N. WESSELOWSKI führt die Basis 60 auf das Zählen mit den Fingern und Gelenken der Hand und das Positionsprinzip auf die Verwendung des Abakus zurück, auf die wir weiter unten zu sprechen kommen. Auch I. J. TIMTSCHENKO ließ gelten, daß die „Sexagesimalrechnung von der Fortsetzung eines Zählens an den Fingern der rechten Hand auf die Gelenke der rechten und linken Hand und in umgekehrter Richtung zurück herrührte", ließ jedoch auch als Möglichkeit zu, daß sie sich als Kombination aus einem Vierer- und einem Fünfzehnersystem (wie es z. B. in Bengalen benutzt wurde) entwickelt hat. (Vgl. seine Bemerkung in der russischen Übersetzung des Buches von F. CAJORI.)

ein Zufall gewesen sein. Als Bestätigung hierfür dient die zu verschiedenen Zeiten und unabhängig voneinander verlaufene Entwicklung eines Positionssystems bei wenigstens drei verschiedenen Völkern: 1.) Bei den Babyloniern im Zweistromtal des Euphrat und Tigris — mehr als 2000 Jahre v. u. Z., 2.) zu Beginn unserer Zeitrechnung beim Volke der Maya, den damaligen Bewohnern der Halbinsel Yukatan in Mittelamerika, und 3.) im 8. bis 9. Jahrhundert u. Z. in Indien.

Die Blütezeit der Maya-Kultur fällt in das 6. bis 13. Jahrhundert unserer Zeitrechnung. Zu dieser Zeit benutzten die Maya zwei verschiedene Systeme der Zahlenschrift: 1.) Ein System vom Typ des ägyptischen Bezeichnungssystems, welches im täglichen Leben benutzt wurde, und 2.) ein absolutes Positionssystem, welches hauptsächlich für Kalenderberechnungen verwendet wurde. Charakteristisch für das zweite System ist das Vorhandensein der Null. Als Zeichen für sie diente das Bild eines halbgeschlossenen Auges. Die Basis dieses Positionssystems war die Zahl 20, obwohl auch starke Spuren eines Fünfer-Systems vorhanden waren. Die Zeichen für die ersten neunzehn Zahlen waren Kombinationen aus dem Zeichen für die Zahl 1, einem Punkt (·), und dem Zeichen für die Zahl 5, einem Strich (—). So wurden z. B. geschrieben:



Die Einheit der nächsten Stelle — „Uinalz", d. h. „20 Tage" genannt — war die Zahl 20. Sie wurde durch dasselbe Zeichen wie die Zahl 1 bezeichnet, unter das aber zur Unterscheidung das Zeichen für die Null geschrieben wurde. Die Einheit der nächsten Stelle bildeten nun nicht 20, sondern nur 18 Uinalz, die ein „Tun" (360 Tage) genannt wurde. Dies ist die einzige Abweichung von der Zwanziger-Basis des Systems der Maya. Diese Unregelmäßigkeit erklärt sich daraus, daß das Jahr bei den Maya in 18 Monate zu je 20 Tagen und gesondert 5 Tage eingeteilt wurde. 20 Tun ergaben ein „Katun" oder 7200 Tage, 20 Katun bildeten die Einheit der fünften Stelle, das „Zikel", und schließlich stellten 20 Zikel ein „großes Zikel" oder 2880000 Tage dar. Das System der Maya hatte also die Stelleneinheiten 1, 20, $20 \cdot 18$, $20^2 \cdot 18$, $20^3 \cdot 18$, ...

Bei ihren kalendarischen und chronologischen Berechnungen operierten die Maya bereits mit sehr großen Zahlen. Die größte Zahl, die man in ihren Texten gefunden hat, ist 12489781. Ihr Schriftbild hatte (schematisch) folgendes Aussehen:



12 489 781

$$12489781 = 4\,(18 \cdot 20^4) + 6\,(18 \cdot 20^3) + 14\,(18 \cdot 20^2) + 13\,(18 \cdot 20) + 15 \cdot 20 + 1.$$

Als Vorläufer unseres heutigen Bezeichnungssystems können wir wahrscheinlich das indische System ansehen. Bedauerlicherweise ist uns nur sehr wenig darüber bekannt, wie und wann das indische Positionssystem entstand. Die Untersuchungen hierüber waren schon deshalb sehr mühsam, weil eine große Anzahl gefälschter Schriften auftauchte. So erwiesen sich von 17 alten Textstellen, die die Positionsschreibweise enthielten, nur zwei als echt.

Vor der Entstehung des Positionssystems wurden in verschiedenen Teilen Indiens die sogenannten Ziffern Kharosthi benutzt, die die älteste bekannte Art von indischen Zahlzeichen sind. Man fand sie im Gebiet östlich von Afghanistan und im Nord-Pandschab. Niedergeschrieben wurden sie vermutlich im 3. bis 1. Jahrhundert vor der Zeitwende. Bei den Ziffern Kharosthi handelte es sich um ein nichtpositionelles Dezimalsystem mit Individualzeichen für die Zahlen 1, 4, 10, 20 und 100 (s. Tafel 2). Die Tatsache, daß die Zahlen 4 und 20 Knotenzahlen waren, deutet auf eine Sonderstellung dieser Zahlen hin. Die Zahlen 200, 300 usw. wurden mit Hilfe des Zeichens für die Zahl 100 geschrieben, an das von rechts das Zeichen für die Zahl 2 oder 3 usw. angefügt wurde. Zur Darstellung der ganzen Hunderter wurde also bereits das multiplikative Prinzip benutzt. Man sieht unmittelbar, daß die Ziffern Kharosthi sicher nicht als Urbilder unserer heutigen Ziffern anzusehen sind.

Neben den Ziffern Kharosthi existierte in Indien in ältester Zeit noch ein anderes Bezeichnungssystem für die Zahlen, dessen Zeichen den Buchstaben des sogenannten brahmanischen Alphabets ähnelten (siehe die Ziffern aus den Höhleninschriften von Nasik, Tafel 2). In diesem System gab es spezielle Zeichen für die ersten neun Zahlen sowie für die ganzen Zehner von 10 bis 90 und für die Zahlen 100 und 1000.[1]) Als Bezeichnung für die Zahl 200, 300, 2000 und 3000 diente das Zeichen für die Zahl 100 bzw. 1000 mit zwei oder drei hinzugefügten Strichen. Ähnlich schrieb man die ganzen Hunderter von 400 bis 900 und die ganzen Tausender von 4000 bis 7000 (eine größere Zahl kommt in den uns bekannten Inschriften nicht vor) als Kombinationen aus den Zeichen für 100 bzw. 1000 und den Zeichen für die Zahlen von 4 bis 9. Die Zahl 3451 z. B. wurde (schematisch) als 3 ·„1000“ 4 ·„100“ 5 ·„10“ 1. geschrieben.

Dieses Ziffernsystem war noch bis zum Ende des 19. Jahrhunderts auf Ceylon in Gebrauch, wohin im 3. Jahrhundert die altindische Kultur mit dem Buddhismus gelangt war. Dort hatte sich das alte Ziffernsystem fast unverändert erhalten.

Auch die Herkunft der brahmanischen Ziffern ist keineswegs genau bekannt. Allerdings spricht eine Reihe von Gründen dafür, daß sie aramäischer[2]) Herkunft sind. Vielleicht sind die brahmanischen Ziffern die Urform unserer heutigen Ziffern.

Daneben gab es in Indien noch ein drittes, wortmäßiges Bezeichnungssystem für die Zahlen, auf das wir später zu sprechen kommen.

Die Verwendung eines Zeichens für die Null zum Schreiben der Zahlen im positionellen Dezimalsystem kam in Indien wahrscheinlich um das Jahr 500 unserer Zeitrechnung auf. Es ist möglich, daß ein Zeichen für die Null schon ÄRYABHATA

---

[1]) Es ist anzunehmen, daß diese Ziffern, die auch singhalesische Ziffern genannt werden, aus den Anfangsbuchstaben der entsprechenden Zahlwörter hervorgegangen sind.

[2]) Die Aramäer waren eine arabisch-semitische Völkergruppe, die in der zweiten Hälfte des 2. Jahrtausend v. u. Z. das Territorium des heutigen Syriens und einen Teil Mesopotamiens bewohnte und dort eine Reihe von Staaten bildete.

(476—550) bekannt war.[1]) Die erste genau datierte Handschrift, in der ein Nullzeichen vorkommt, stammt jedoch erst aus dem Jahre 876 (sogenannte Gwalior-Inschrift. *Anm. d. Übers.*). In ihr ist die Zahl 270 in der Form ∋7०[2]) geschrieben worden. Wie kamen aber nun die Inder zu einem absoluten Positionssystem?

N. M. Bubnow führte die Entstehung des Positionssystems auf die Benutzung des Rechenbrettes — des sogenannten Abakus — in Indien und den anderen Ländern des alten Orients zurück. Diese Rechenbretter hatten Längsrillen, deren jede einer bestimmten Dezimalordnung entsprach, in denen sich Jetons (Marken) befanden. Diese waren zunächst ungezeichnet; ihre Anzahl in den einzelnen Rillen gab die Anzahl der Einheiten der entsprechenden Ordnung an.

Nach N. M. Bubnow sind nun unsere Zahlwörter und Zahlzeichen letzten Endes ural-altaischen Ursprungs. Von den ural-altaischen Völkern aus verbreiteten sie sich insbesondere nach Chaldäa (südöstlicher Teil Mesopotamiens). Von hier aus kamen die Zeichen im 3. Jahrhundert v. u. Z. einerseits nach Indien, wo sie ohne Null und ohne Stellenwert als Ziffern benutzt wurden, und andererseits nach Griechenland, wo sie auf den griechischen Abakus mit zunächst unbezeichneten Jetons trafen. Dort sollen dann die Jetons mit den Ziffern von 1 bis 9 versehen und ein zehnter leerer Jeton (Sipos) hinzugefügt worden sein, der das Fehlen von Einheiten einer gewissen Dezimalordnung andeutete. Von Griechenland aus gelangten im Laufe der Zeit die bezeichneten Rechenbretter überall hin. In Indien nun waren die Zeichen des griechischen Abakus als Zahlzeichen — allerdings ohne Stellenwert — bereits lange in Gebrauch. Hier bekamen sie jetzt dadurch einen Stellenwert, daß man die Zeichen der Jetons in der Reihenfolge, in der sie auf dem Abakus lagen, aufs Papier übertrug. Man brauchte dazu ja nur die leeren Stellen verschiedener Dezimalordnung bezeichnen zu können. Im Laufe der Zeit soll man dann dazu übergegangen sein, auch für diese gezeichnete Jetons zu verwenden, und zwar sollen dafür runde Scheiben mit einem Loch in der Mitte, also sozusagen materialisierte Modelle unserer heutigen Null, verwendet worden sein. In der Schrift wurden sie zunächst durch einen dicken Punkt und später durch einen Kreis ○ versinnbildlicht.

Diese Hypothese ist zwar recht scharfsinnig, jedoch wird sie durch die historischen Tatsachen nicht bestätigt. Vielmehr kann man gegen sie die folgenden wesentlichen Einwände ins Feld führen:

1.) Auf allen uns überlieferten antiken Zeichnungen ist der griechische Abakus mit ungezeichneten Jetons dargestellt. Daher liegt kein Grund vor für die Annahme, daß die alten Griechen gezeichnete Jetons verwendet hätten. Im Gegenteil zeigt das Beispiel der russischen Rechenbretter, der chinesischen Suan-Pan und des römischen Abakus, daß gezeichnete Jetons weder zur Fixierung einer Zahl auf einem solchen Rechenbrett noch zur Ausführung der Rechenoperationen überhaupt nötig waren. Es ist nur bekannt, daß im Mittelalter Rechenbretter mit gezeichneten Jetons existierten und weit verbreitet waren.

2.) Es ist höchst unwahrscheinlich, daß sich die aramäischen Ziffern während der ganzen Ausbreitungszeit nach Indien und Griechenland nicht verändert haben sollten.

---

[1]) Ein Zeichen für die Null in Form eines Punktes kommt nachweislich in dem sogenannten Rechenbuch von Bakhšālī vor, dessen genaue Entstehungszeit allerdings nicht bekannt ist. Die Forscher geben hierfür sehr verschiedene Daten an, die vom 2. bis zum 8. und 9. Jahrhundert unserer Zeitrechnung reichen.

[2]) Vgl. D. E. Smith and L. C. Karpinsky, The hindu-arabical numerals, New York 1911. S. 43 bis 44, 52.

3.) In der ältesten Zeit waren Rechenbretter sowohl bei den Ägyptern als auch bei den Griechen, Chinesen und Römern in Gebrauch, ohne daß eines dieser Völker zu irgendeinem Positionssystem übergegangen wäre.[1])

Es ist als glaubwürdiger anzusehen, daß sowohl Abakus als auch Positionssystem ein und desselben Ursprungs und durch denselben historischen Prozeß entstanden sind.

Man kann sich dann die Entstehung der Positionsschreibweise folgendermaßen vorstellen: Das Positionsprinzip ist im Grunde die Vereinigung zweier Prinzipien: 1.) des Multiplikationsprinzips und 2.) des Prinzips, die Benennungen für die Dezimalordnungen beim Schreiben auszulassen. Das Positionsprinzip wird schließlich durch die Einführung der Null vervollständigt. Will man also die Entstehung des Positionsprinzips einigermaßen befriedigend erklären, so muß man vor allem einmal das Auftreten der multiplikativen Bezeichnungsweise zu erklären versuchen, denn diese ist ja gleichermaßen auch die Grundlage für die Darstellung einer Zahl auf dem Abakus.

Wir wollen zunächst auseinandersetzen, worin sich die additive Schreibweise prinzipiell von der multiplikativen unterscheidet. Vom rein algebraischen Standpunkt aus sind z. B. die beiden möglichen Schreibarten $\mathrm{III_X}$ und $\mathrm{X^{III}}$ der Zahl 30 nur ein Ausdruck des distributiven Gesetzes. In der Tat ist:

$$\mathrm{XXX = X + X + X = (I + I + I)\,X = III_X \text{ oder } X^{III}}.$$

Darüber hinaus bringt aber die multiplikative Schreibweise $\mathrm{III_X}$ zum Ausdruck, daß beim Zählen die Zahl Zehn als neue Einheit angesehen wird. Dasselbe gilt entsprechend für das Rechnen mit Hundertern, Tausendern usw. Somit umreißt die multiplikative Schreibweise am klarsten den Tatbestand, daß beim Zählen eine bestimmte Anzahl von Einheiten erster Ordnung zu einer Einheit der folgenden Ordnung zusammengefaßt wird. Entsprechend gewinnt man die Einheit der dritten Ordnung, indem man eine bestimmte Anzahl von Einheiten zweiter Ordnung zusammenfaßt usw. Dadurch wird erreicht, daß ein und dasselbe Zahlzeichen eine bestimmte Quantität von Einheiten verschiedenster Ordnung bezeichnen kann. Hinter dem Zeichen ist dann zu vermerken, von welcher Ordnung die bezeichneten Einheiten sein sollen. Durch diese Bezeichnungsweise wird also unterstrichen, daß die Objekte des Zählens ganz beliebiger Art sein können (z. B. können es Gegenstände, aber auch Mengen einer bestimmten Anzahl von Gegenständen, wie Zehner, Hunderter usw., sein). Das ist aber andererseits gerade die Wurzel des abstrakten Zahlbegriffs: Die Zahl ist die allgemeinste Eigenschaft, die allen den verschiedenen konkreten Mengen gleicher Anzahl gemeinsam ist.

Von der gleichen Art war übrigens — wie wir bereits erwähnten — das Rechnen mit den Stück-Anzahlen. Auch die afrikanischen Neger, die mit Hilfe von Steinchen oder Nüssen zählen, legen diese in Häufchen von je fünf Stück zusammen. Fünf dieser Häufchen vereinigen sie zu einem neuen Haufen usw. Offensichtlich werden hier also zunächst die Steinchen, dann die Häufchen, dann die Haufen usw. gezählt. Bei diesem Zählverfahren muß ganz besonders hervorgehoben werden, daß mit

[1]) Vgl. die Anmerkungen von I. J. Timtschenko in dem bereits zitierten Buche von F. Cajori auf S. 318—320. Dabei muß übrigens hervorgehoben werden, daß I. J. Timtschenko der Meinung war, daß „die Idee des Stellenwertes der Zeichen wahrscheinlich immer von der Verwendung des Abakus ausging, selbst wenn dieser mit ungezeichneten Jetons ausgestattet war" (ebenda).

den Haufen genau so verfahren wird wie mit den einzelnen Steinchen. Nach demselben Verfahren zählt auch der bereits genannte Stamm der Joruba in Zentralafrika. Die Objekte des Zählens sind hier allerdings goldfarbene Muscheln (Kauri), die in Haufen zu je 20 Stück vereinigt werden. Dabei ist noch interessant, daß bei den Jeruba das Wort „zählen" soviel wie „anhäufen", „packen" bedeutet. Ganz ähnlich war es übrigens auch im alten Griechenland. Hier wurde die Arithmetik, als die Wissenschaft von der Zahl, genau getrennt von der Logistik, als der Kunst des Rechnens. Das griechische Wort λογος hat seine Wurzel aber in dem Wort λεγ, was soviel wie „sammeln" bedeutete.

Als eine weitere Illustration zur Entwicklung des Zählens, die zum multiplikativen Prinzip führte, kann das folgende Beispiel von N. N. Miklucho-Maklai dienen. Er berichtet über ein Zählverfahren bei den Eingeborenen Neu-Guineas: Um die Anzahl der Zettelchen zu bestimmen, die der Anzahl der Tage bis zur Rückkehr der Korvette „Witjas" entsprach, verfuhren die Papuas folgendermaßen: „Der erste legte ein Zettelchen nach dem anderen auf sein Knie und sagte dabei jedesmal 'kare, kare' (eins). Ein anderer wiederholte das Wort 'kare' und krümmte dabei jedesmal einen Finger, zuerst an der einen Hand und danach an der anderen. Nachdem er so bis zehn gezählt hatte, senkte er beide Fäuste auf die Knie und sagte 'zwei Hände', wobei ein dritter Papua einen Finger der einen Hand krümmte. Mit den nächsten zehn Zetteln wurde genau so verfahren, wobei der dritte Papua den zweiten Finger krümmte. Analog machten sie es auch mit dem dritten Zehner."[1])

Ein ganz ähnliches Verfahren wenden südafrikanische Stämme beim Zählen ihrer Viehherden an (dieses Beispiel wurde von Zeuthen angeführt). Dazu müssen drei Mann vorhanden sein. Der erste hebt für jeden an ihm vorüberziehenden Kopf der Herde einen Finger seiner Hände und zählt so laufend bis zehn. Der zweite zählt dann die Zehner des ersten und der dritte die Zehner des zweiten, also die Hunderter. Ähnliche Zählmethoden findet man auch in vielen anderen Ländern. Die genannten Beispiele werfen ein Licht auf die Herkunft sowohl des Abakus als auch des Positionssystems. Ersetzt man nämlich die Finger des ersten, zweiten und dritten Zählenden etwa durch Steinchen, die in verschiedene Rillen gelegt werden, oder durch Glasperlen, die auf drei Drähte aufgereiht sind, so erhält man die einfachste Art des Abakus, und zwar in ihrer Urform. Wenn wir andererseits die Finger der Zählenden durch die Zeichen I, X, C bezeichnen, so erhalten wir — wenn wir die Anzahl der gezählten Finger aufs Papier übertragen — die multiplikative Schreibweise. Beispielsweise würde sich die Zahl 323 (schematisch) folgendermaßen schreiben: 3C2X3I.

Wenn das Stadium des Zählens mit Stück-Anzahlen völlig allgemein ist, so muß es auch in Indien durchlaufen worden sein. Jedenfalls gab es in Indien — wie wir sahen — schon von altersher multiplikative Zahlensysteme. So waren das System Kharosthi und das brahmanische System auf dem multiplikativen Prinzip aufgebaut. Dabei wurden in Indien große Zahlen nach derselben Methode geschrieben, die in Babylonien bis zu dem Zeitpunkt benutzt wurde, da man die Benennungen der Ordnungen fortzulassen begann.

Man muß übrigens in Indien dem Problem der Darstellung beliebig großer Zahlen ein erhebliches Interesse (und dies nicht nur unter den Gelehrten) entgegengebracht haben. So wird im Līlavātī, einem berühmten Werk der buddhistischen Literatur,

[1]) N. N. Miklucho-Maklai, a. a. O., S. 58.

der Wettkampf zwischen den Freiern der schönen HOPA (der Landesherrin) beschrieben. Gegenstand des Wettkampfes waren das Schrifttum, die Arithmetik, das Ringen und die Kunst des Pfeilwerfens. Fast die Hälfte der Schilderungen ist den Aufgaben in der Arithmetik gewidmet. So mußten die Wettkämpfer z. B. einen Weg finden, um Zahlen auszudrücken, die größer als 100 Koti (1 Koti = 10^7) waren. Der Sieger, SARWATASSIDDA, erfand eine Zahlentabelle, die durch eine geometrische Reihe mit dem Quotienten 100 dargestellt wurde und die als letztes Glied die Zahl $10^{7+9\cdot46}$ — eine 1 mit 421 Nullen! — hatte. Mit deren Hilfe berechnete er die Anzahl der „primären Atome" der Längeneinheit, wobei er eine Tabelle zur Bezeichnung der betreffenden Zahlen aufstellte.

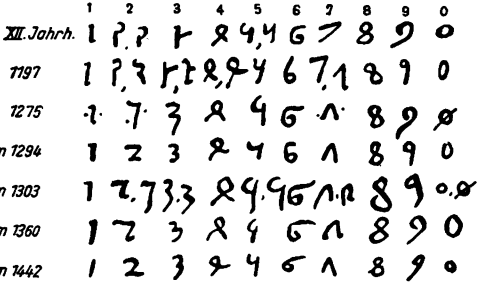
In Indien lag nun ein besonders günstiger Umstand für das Entstehen eines Positionssystems vor. Wir erwähnten schon das dritte, das wortmäßige Bezeichnungssystem für die Zahlen, welches vor allem in den astronomischen und den mathematischen Schriften angewendet wurde. Dieses System dürfte vor dem Beginn des 6. Jahrhunderts unserer Zeitrechnung entstanden sein. Die Zahl Eins wurde in diesem System durch Wörter wie „Mond", „Welt", „Brahma" usw. bezeichnet, also durch Bezeichnungen für Gegenstände, die in der Einzahl vorhanden sind. Die Zahl Zwei wurde durch Wörter wie „Zwillinge", „Augen", „Hände" usw., die Zahl Fünf durch Wörter wie „Sinne", „Speere" (die fünf Speere des Khamadewa, des Liebesgottes der indischen Mythologie) usw. ausgedrückt u. dgl. mehr. Die zusammengesetzten Zahlen wurden in diesem System nach dem Positionsprinzip gebildet. So wurde die Zahl 867 als „giri — raṣa — vasu", d. h. Gebirge (7) — Gerüche (6) — Götter (8) geschrieben. Beim Schreiben begann man stets mit der niedrigsten Ordnung und endete mit der höchsten. Außerdem gab es im Sanskrit (das bei den Indern eine ähnliche Rolle spielte, wie bei uns etwa das mittelalterliche Latein) spezielle Bezeichnungen für die Dezimalordnungen bis zu 10^{16}. So wurde beispielsweise die Zahl 86 789 325 178 sanskritisch folgendermaßen gelesen: 8 kharva, 6 padma, 7 vyrbuda, 8 koti, 9 prayuta, 3 laksha, 2 ayuta, 5 sahasta, 1 gata, 7 daçan, 8. Eine solche Bezeichnungsweise hob bereits die Gleichberechtigung der verschiedenen Ordnungen hervor. Es brauchte also nur noch ein Schriftsystem konstruiert zu werden, das auch der bereits bestehenden wortmäßigen Bezeichnungsweise entsprach. Ein solches System entstand dann auch bei der weiteren Entwicklung der multiplikativen Zahlenschrift

Der Prozeß, der zum Auslassen der Ordnungsbezeichnungen beim Schreiben führte, verlief wahrscheinlich genau so wie in Babylonien. Zur Vollendung des Positionssystems ermangelte es nur noch eines letzten Schrittes — der Einführung der Null. Wegen der kleinen Basis, die die Zahl Zehn nun einmal ist, und wegen des Operierens mit relativ großen Zahlen war — besonders nachdem die Bezeichnungen für die Ordnungen weggelassen wurden — eine solche Einführung unumgänglich. Dabei ist es völlig gleichgültig, ob das Zeichen für die Null ursprünglich ein Abbild des leeren Jetons des Abakus war oder ob es aus einem Punkt entstanden ist, den man an die Stelle der fehlenden Dezimalordnungen gesetzt hatte. Wie dem auch sei, die Einführung der Null war eine völlig unvermeidliche Etappe im gesetzmäßigen Entwicklungsprozeß, der zur Schöpfung des modernen Positionssystems führte.

§ 5. Die Ausbreitung der positionellen Schreibweise der Zahlen nach Westeuropa und Rußland

Wir fragen uns jetzt, auf welchen Wegen das indische Positionssystem nach Europa gelangt ist und wie es hier seinen Einzug hielt.

Das Positionsprinzip breitete sich von Indien zunächst in die benachbarten Länder aus. Dabei übernahmen einige Völker von den Indern nur das Prinzip, behielten aber die Form ihrer Ziffern bei (China), andere übernahmen darüber hinaus von den Indern auch die Form der Ziffern (Tibet, die Mongolei, die Völker des Nahen und des Mittleren Ostens). Die frühesten Handschriften in arabischer Sprache, in denen das Positionsprinzip benutzt wird, stammen aus den Jahren 874 und 878. Da in den verschiedenen Gebieten Indiens Ziffern von recht unterschiedlicher Form in Gebrauch waren, hatten auch die Ziffern, die in der Folgezeit in den östlichen Kalifaten und den maurischen Staaten auf dem Territorium des heutigen Spaniens benutzt wurden, eine recht unterschiedliche Gestalt. Die ostarabischen Ziffern breiteten sich später auf den ganzen islamischen Osten aus, wo sie noch heute — allerdings in etwas veränderter Form — gebraucht werden. Die Ziffern aber, die in den maurischen Staaten verwendet wurden, die sogenannten „Ġubār"-Ziffern (Tafel 4, 1. Zeile), sind die direkten Vorfahren unserer heutigen Ziffern.



Tafel 4

Es erhebt sich nun die Frage nach der Herkunft dieser Ġubār-Ziffern. Wenn sie tatsächlich indischen Ursprungs sind, so können sie nur über die vorderasiatischen Länder nach Spanien gelangt sein. Wie erklärt es sich aber, daß gewisse Abarten der indischen Ziffern in die maurischen Staaten in Spanien gelangten und dabei die vorderasiatischen Länder „übersprangen"?

Nach einer Hypothese von WOEPCKE, die eine Zeitlang weite Anerkennung fand, läßt sich diese Tatsache folgendermaßen erklären:

Bereits im 2. Jahrhundert v. u. Z. bestanden Handelsbeziehungen zwischen Indien und Alexandrien. Dadurch gelangten die indischen Ziffern (ohne Null und Positionsprinzip) nach Alexandrien und von dort nach Rom, Westafrika und Spanien. Als Bestätigung dieses Teils seiner Hypothese wertete WOEPCKE, daß in der „Geo-

metrie" des BOETIUS (480?—524) Ziffern benutzt wurden, deren Form stark an die Gubar-Ziffern erinnert. So berichtet BOETIUS in seinem Buch von einem Abakus, dessen Herstellung er den Pythagoreern zuschreibt und dessen Marken keine Steinchen, sondern bezeichnete Jetons — sogenannte Apices — waren, auf denen Ziffern dargestellt waren. Die Ziffern selbst, die in ihrer Form den Ġubār-Ziffern ähnelten, wurden später auch Apices-Ziffern genannt. Später entstanden in Indien die Null und das Positionsprinzip. Diese wurden von den Völkern des Nahen und Mittleren Ostens zusammen mit den inzwischen in Indien gebräuchlich gewordenen neuen Ziffern entlehnt. In den maurischen Staaten hingegen benutzte man zur Bezeichnung der Zahlen jene Ziffern weiter, die schon längst in Spanien gebräuchlich waren, und übernahm von den östlichen Glaubensbrüdern nur das Zeichen für die Null und das Positionsprinzip.

Von Spanien aus gelangten die Ziffern und das Positionsprinzip in das übrige Europa. Nach Ansicht des Begründers dieser Hypothese erinnert das Wort „Ġubār", was soviel wie „Sand" oder „Staub" bedeutet, an die indische Herkunft der Ziffern (in Indien wurden die Ziffern nämlich in den Sand oder Staub geschrieben). Natürlich waren die Ziffern in Indien selbst mannigfachen Veränderungen unterworfen. Daraus erklärt sich der Unterschied zwischen den Ġubār-Ziffern und den heute in Indien gebräuchlichen „Dewangari"-Ziffern.

Die Hypothese von WOEPCKE hat nun aber zwei schwache Stellen: 1.) Das Vorhandensein von indischen Ziffern in Alexandrien im 2. Jahrhundert v. u. Z. und auch später wird durch keine geschichtlichen Dokumente bewiesen. 2.) Die „Geometrie" des BOETIUS, auf die sich die Hypothese von WOEPCKE maßgeblich stützt, ist — wie jetzt festgestellt wurde — nicht authentisch, sondern eine Fälschung aus dem 11. Jahrhundert. N. M. BUBNOW unterzog in seinem Buche «Арифметическая самостоятельность европейской культуры» (Die arithmetische Selbständigkeit der europäischen Kultur, Kiew 1908, deutsche Ausgabe: Berlin 1914) die Hypothese WOEPCKES einer eingehenden Kritik. Allerdings ist seine eigene Theorie nicht stichhaltiger als die kritisierte. Der bekannte russische Historiker W. W. BOBYNIN unterzog die Hypothese BUBNOWS ihrerseits einer gründlichen Kritik in seinem Buche «Отзыв о сочинениях Н. М. Бубнова и т. д.» (Bemerkungen zu den Werken N. M. BUBNOWS usw., St. Petersburg 1911). W. W. BOBYNIN schrieb in bezug auf die Theorie BUBNOWS: „Die Geschichte unserer Ziffern ist bisher nichts anderes als eine mit willkürlichen Annahmen vermischte Folge von Mutmaßungen, welche — nachdem sie durch Suggestion genügend vorbereitet sind — manchmal den Eindruck aufkommen lassen, daß sie bereits bewiesene Tatsachen seien."

In neuerer Zeit fanden unter der Leitung von S. P. TOLSTOW Ausgrabungen des alten Chowaresm statt. Chowaresm war ein bedeutender Kulturstaat des Altertums auf dem Territorium der heutigen Sowjetunion. Es ist zu hoffen, daß diese Ausgrabungen Licht bringen werden in die Dunkelheit um die Entstehung unserer Ziffern.

Jedenfalls besitzen wir bis heute keine historisch begründete Theorie, welche die Frage nach der Entstehung unserer Ziffern hinreichend befriedigend erklären könnte.

Es ist jedoch unbestreitbar, daß die europäischen Völker ihr Ziffernsystem von den islamischen Staaten im heutigen Spanien bezogen haben. Im Laufe des 10. Jahrhunderts begann ein immer größer werdender Einfluß der Kultur der maurischen Staaten auf Europa sichtbar zu werden. Insbesondere begannen zu dieser Zeit die Ġubār-Ziffern nach Europa einzudringen, wo bis dahin fast ausschließlich die

römischen Ziffern in Gebrauch waren. Da damals in Europa die Kunst des Schreibens nur spärlich verbreitet war und außerdem das schriftliche Rechnen mit den römischen Ziffern sehr unbequem ist (der Leser versuche z. B. zwei vierstellige Zahlen zu multiplizieren, die in römischen Ziffern niedergeschrieben sind), wurde zum Rechnen meist das Rechenbrett — der Abakus — verwendet.

Die Einführung des Rechenbrettes in Europa ist mit dem Namen GERBERT (des späteren Papstes SYLVESTER II.) untrennbar verbunden. Er wird als einer der hervorragendsten Mathematiker des 10. Jahrhunderts (gestorben 1003) angesehen. GERBERT schrieb zwei Werke: „Regeln für das Rechnen auf dem Abakus" und „Büchlein über die Division von Zahlen". In diesen Büchern beschreibt GERBERT die zu seiner Zeit gebräuchlichen Rechenmethoden. Wir wissen, daß in seinem Auftrage ein Abakus in Form einer ledernen Rechentafel hergestellt wurde, die 27 vertikale Spalten besaß. Zu diesem Abakus gehörten Marken aus Horn (Apices), die mit den ersten neun Ziffern versehen waren. Nach anderen Angaben waren es sogar 30 Spalten, von denen 3 für Brüche bestimmt waren. Die übrigen Spalten sollen in Gruppen von je dreien eingeteilt gewesen sein. Die einzelnen Spalten waren mit Buchstaben versehen, M (monas) oder S (singularis), D (decem, 10), C (centum, 100). Dieselben Buchstaben — jedoch mit einem Querstrich versehen — bezeichneten auch den tausendfachen Wert. So bezeichnete beispielsweise C die Spalte für 100000. Die Nachfolger GERBERTS nannte man Abacisten. Im Laufe der nächsten Jahrhunderte wurde das Rechenbrett wesentlich verändert, so wurden die numerierten Jetons durch unbezeichnete ersetzt, an die Stelle des vertikalen Verlaufs der Rillen traten horizontale Rillen. In dieser Form breitete sich der Abakus in Deutschland, Frankreich und England aus.

Obwohl sich die erste Niederschrift in arabisch-indischen Ziffern — allerdings noch ohne Verwendung der Null — bereits in einer Handschrift des spanischen Klosters Albeldo (dem sogenannten codex vigileanus) aus dem Jahre 976 und — jetzt bereits mit einem Zeichen für die Null — in einer Handschrift aus St. Gallen (Universitätsbibliothek Zürich) aus dem 10. Jahrhundert findet, beginnen sich die arithmetischen Verfahren der Völker des Mittleren und Nahen Osten erst zu Beginn des 12. Jahrhunderts in Europa zu festigen. Hierfür besaßen die Übersetzungen der arithmetischen Arbeiten des berühmten chowaresmischen Gelehrten MUḤAMMAD IBN MŪSĀ ALḪWĀRAZMĪ (d. h. aus Chowaresm; um 840 gestorben) eine erhebliche Bedeutung. In diesen Arbeiten wird bereits ein positionelles Bezeichnungssystem für die Zahlen angegeben.[1]) Die Methode der Bezeichnung selbst wurde als „Algorithmus" (eine Verstümmelung des Namens ALḪWĀRAZMĪ) bezeichnet. Eine große Rolle spielten auch das Sammelwerk „Buch der Algorithmen" (Liber alghoarismi) des Spaniers JOHANNES VON SEVILLA aus dem 12. Jahrhundert sowie die mehr populären Werke über Arithmetik von ALEXANDER DE VILLA DEI und JOHANNES DE SACROBOSCO aus dem 13. Jahrhundert. Es bildete sich geradezu eine ganze Schule von Arithmetikern heran, welche das neue Bezeichnungsverfahren der Zahlen und das Rechnen in diesem vertraten. Man nannte sie — im Gegensatz zu den Abacisten — „Algorithmiker", da sie ausdrücklich beim Rechnen auf das Rechenbrett verzichteten. Die Algorithmiker lehrten neben den vier Grundrechenarten der Arithmetik bereits das

[1]) Eine lateinische Übersetzung der Werke von ALḪWĀRAZMĪ, die in der Mitte des 12. Jahrhunderts angefertigt wurde, erschien unter dem Titel: Trattati d'arithmetica, ed. BONCOMPAGNI, Heft I, Roma 1857. (Als Heft II erschien 1858 das im Text genannte Rechenbuch des JOHANNES V. SEVILLA. — *Anm. d. wissenschaftl. Red.*)

Ausziehen der Wurzel. Sie benutzten für die Bruchrechnung — zu einer Zeit, als die Abacisten noch mit römischen Zwölfer-Brüchen arbeiteten — die Sexagesimal-Brüche.

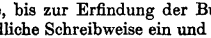
Die neue Bezeichnungsweise setzte sich keineswegs sofort durch. Sie stieß vielmehr auf den heftigen Widerstand sowohl der damaligen scholastischen Wissenschaft als auch der herrschenden Regierungen. So wurde den Florentiner Kaufleuten im Jahre 1299 die Benutzung der indischen Ziffern in der Buchführung offiziell verboten. Gestattet war ihnen ausschließlich die Verwendung der römischen Ziffern oder die wortmäßige Bezeichnung der Zahlen. Noch bis in das 18. Jahrhundert hinein durften in amtlichen Schriftstücken nur römische Ziffern benutzt werden.

Der große Wert der positionellen Schreibweise, der ihren Anhängern wohl klar vor Augen stand, wurde von der großen Masse nicht sofort erkannt. Das Rechnen auf dem Rechenbrett blieb für viele ein unantastbares Privileg. Neben der Neigung zur schablonenhaften Arbeit, dem hohen Preis des Papiers (seine Produktion in Europa begann erst im 12. Jahrhundert) und des Mangels an Schreibwerkzeugen (der Bleistift wurde erst im 16. Jahrhundert erfunden) sind hieran vor allem die überaus langsame Entwicklung der vier Grundrechenarten — insbesondere der Multiplikation und der Division — sowie, bis zur Erfindung der Buchdruckerkunst im 15. Jahrhundert, die so unterschiedliche Schreibweise ein und derselben Ziffer durch die verschiedenen Schreiber schuld.[1])

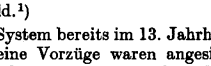
Trotzdem war das neue System bereits im 13. Jahrhundert unter den italienischen Kaufleuten verbreitet. Seine Vorzüge waren angesichts der erheblichen Vereinfachung der arithmetischen Operationen so gewaltig, daß es, ungeachtet aller Schwierigkeiten, nach und nach die alten Rechenverfahren und die alte Bezeichnungsweise verdrängte. Ein interessantes Beispiel für den Sieg der positionellen Bezeichnungsweise ist, daß zu einer gewissen Zeit die römischen Ziffern mit Nullzeichen in positioneller Schreibweise verwendet wurden. So schrieb N. Ocreat (er lebte um die Wende vom 11. zum 12. Jahrhundert) III.III für die Zahl 33, I.O.VIII.IX für die Zahl 1089 usw.[2])

Leonardo von Pisa (Fibonacci), ein überzeugter Anhänger der neuen Schreibweise, schrieb in seinem „Liber abaci" (1202)[3]): „Mit Hilfe der neuen indischen Ziffern 1, 2, . . . , 9 und dem Zeichen für die Null, welche von den Arabern zephirum genannt wird, kann man jede beliebige Zahl schreiben." Hier ist „zephirum" eine Latinisierung des arabischen Wortes „as-ṣifr", das seinerseits eine wörtliche Über-

[1]) In verschiedenen Handschriften finden sich z. B. folgende Ziffern für die Zahl Eins:



und folgende Ziffern für die Zahl Zwei:



[2]) D. Smith and L. Karpinsky, The hindu-arabical numerals, S. 119—120.

[3]) Das Wort „abacus" diente in jener Zeit zur Bezeichnung der Arithmetik überhaupt. Somit ist das Werk des Leonardo eher ein „Buch der Arithmetik" als eine Abhandlung über den Abakus.

setzung des indischen Wortes „sunya" ist, was soviel wie „leer" bedeutet und das auch in Indien zur Bezeichnung der Null diente. Aus dem Worte „zephirum" entwickelte sich später das französische Wort „zéro" und das italienische Wort „zero" (Null). Von dem arabischen Wort „as-ṣifr" leiten sich andererseits auch das Wort „Ziffer" wie auch das französische Wort „chiffre", das englische Wort „cipher" und das russische Wort цифра ab.[1]) Bis in die Mitte des 17. Jahrhunderts wurden diese Wörter ausschließlich zur Bezeichnung der Null verwendet. So benutzte der russische Gelehrte MAGNIZKI in seiner „Arithmetik" zur Bezeichnung der Zahl Null stets das Wort цифра. Das lateinische Wort „nullus" (nichts) für die Null kam erst im 16. Jahrhundert in Gebrauch.[2])

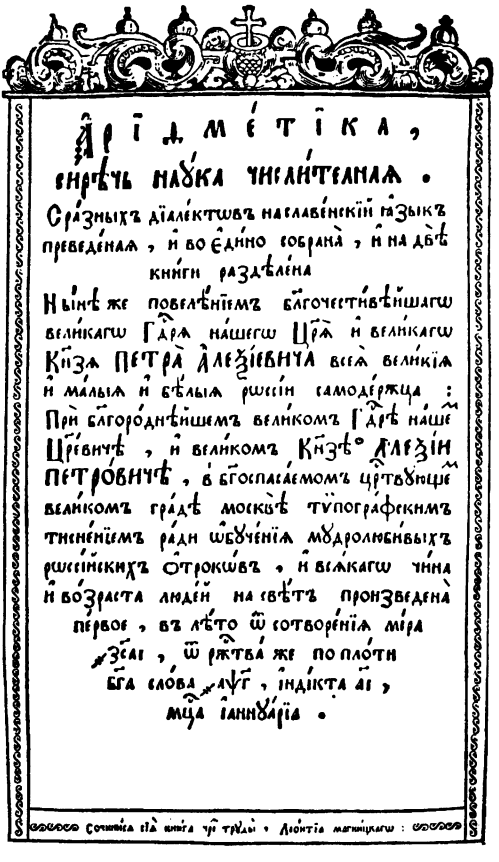
In Deutschland, Frankreich und England wurden die neuen Ziffern vor der zweiten Hälfte, ja bis gegen Ende des 15. Jahrhunderts kaum benutzt. Die ersten Münzen mit indischen Ziffern kamen 1424 in der Schweiz[3]), 1484 in Österreich, 1485 in Frankreich, 1489 in Deutschland und 1551 in England in Umlauf. Die ältesten Inschriften in indischen Ziffern auf Grabsteinen stammen aus den Jahren 1371 aus Baden (Pforzheim — *Anm. d. wissenschaftl. Red.*) und 1388 aus Ulm. Im Jahre 1488 wurde das Buch „De arte numerandi" („Über die Kunst des Rechnens") — auch unter dem Namen „Algorismus" bekannt — gedruckt. In diesem Buch, das man dem bereits erwähnten JOHANNES DE SACROBOSCO (JOHANN HALIFAX) zuschreibt, wurden ohne Beispiele und Beweise die Gesetze der „indischen" Arithmetik dargelegt. Jedoch hatte erst im 16. bis 17. Jahrhundert die neue Bezeichnungsweise der Zahlen die alte so gut wie vollständig verdrängt. So druckte man noch im zweiten Drittel des 16. Jahrhunderts die Kalenderzahlen gewöhnlich in römischen Ziffern.

Die erste in Rußland gefundene mathematische Handschrift stammt aus dem Anfang des 12. Jahrhunderts. Es ist die «Кирика Диакона и Доместика Антониева монастыря учение, им-же ведати человеку числа всех лет» („Lehre des Kyrik, Diakon und Domestik am Antoniter Kloster, wodurch der Mensch die Zahlen aller Jahre lernen kann"). In diesem Leitfaden wurden noch alle Zahlen im alphabetischen System dargestellt. Ebenso steht es mit den aus dem 14. und 15. Jahrhundert stammenden Abschriften des berühmten juristischen Schrifttumsdenkmals, des «Правда Русская» („Das russische Recht"). Das neue Bezeichnungssystem für die Zahlen breitete sich in Rußland nur wenig später als im westlichen Europa aus. In allen, also auch den mathematischen Handschriften des 17. Jahrhunderts fand schon das dezimale Positionssystem Verwendung. Hierüber schrieb W. W. BOBYNIN: „Spuren eines ehemaligen Gebrauchs des alten griechisch-slawischen Systems findet man nur in den ältesten von ihnen, und auch hier nur in einer so schwachen Form, wie etwa der Erklärung der arabischen Ziffern durch die entsprechenden slawischen, oder man begegnet hin und wieder einer Bezeichnung der betrachteten Zahlen durch slawische Ziffern allein oder zusammen mit arabischen. Die Handschriften aus der

[1]) Es ist immerhin bezeichnend, daß im Eifer der Auseinandersetzungen zwischen den Algorithmikern und den Abacisten die Worte „Algorithmus" und „Ziffer" als Synonyme für irgendwelche anderen Dinge dienten.

[2]) In lateinischen Übersetzungen von Werken aus der arabischen Sprache wurden im 17. Jahrhundert Ausdrücke wie „nulla figura", „nullus circulus" (keine Figur, kein Kreis) zur Bezeichnung der Null benutzt.

[3]) In Sizilien, das eng mit den arabischen Staaten verbunden war, wurden hierfür die indischen Ziffern mit Sicherheit bereits 1138 gebraucht.



Titelseite der „Arithmetik“ von Magnizki

zweiten Hälfte des 17. Jahrhunderts weisen nicht einmal mehr diese unbedeutenden Spuren auf."[1]) Im Alltagsleben hingegen bürgerte sich die neue Bezeichnungsweise nicht so schnell ein.

In gedruckten Werken in slawisch-russischer Sprache finden sich die indischen Ziffern erstmalig bei der Numerierung der Seiten in zwei Büchern geistlichen Inhalts, die 1611 in Venedig herausgegeben wurden.[2]) Das erste in einer russischen Druckerei hergestellte Buch, dessen Seitenzahlen indische Ziffern waren, ist ein Psalmenbuch aus dem Jahre 1638. Es wurde in dem kleinen Marktflecken Jewju (Евю) gedruckt. Im Jahre 1647 wurde in Moskau das Buch «Учение и хитрость ратного строения пехотных людей» („Über die Verwendung des Fußvolkes im Kriege") herausgegeben. Darin waren die Ziffern in den Zeichnungen und dem dazugehörigen erläuternden Text bereits indische Ziffern. Allerdings wurden noch lange danach sowohl die „russischen Ziffern" («числа русские») als auch die „Ziffern" («цифирные»), das waren also die indischen Ziffern, benutzt. So wurde noch im Jahre 1702 das Journal über die Belagerung von Noteburg (Нотебург), welches in insgesamt 2000 Exemplaren erschien, zu je 1000 Exemplaren mit arabisch-indischen bzw. slawischen Ziffern gedruckt.

In dem berühmten Leitfaden «Арифметика, сиречь наука числительная. С разных диалектов на славянский язык переведенная, и во едино собрана и на две книги разделена. В лето от сотворения мира 7211, от рождества Бога Слова 1703. Сочинися сия книга через труды Леонтиа Магницкого» („Die Arithmetik oder die Lehre von den Zahlen. Aus verschiedenen Dialekten ins Slawische übersetzt und in eins zusammengefaßt und auf zwei Bücher verteilt. Geschaffen durch die Arbeit des LEONTI MAGNIZKI im Jahre 7211 seit Erschaffung der Welt und 1703 seit der Geburt Christi"), in welchem z. B. noch der große russische Gelehrte M. W. LOMONOSSOW studierte, wurden als Seitenzahlen slawische Ziffern verwendet, während im Text ausschließlich mit indischen Ziffern gerechnet wurde. Bezüglich der Bezeichnungsweise der Zahlen („Numerierung") sagt das Buch folgendes: „Was ist eine Numerierung: Eine Numerierung ist ein Zählen, das jede Zahl mit einem Wort bezeichnet, das aus den folgenden zehn Bezeichnungen oder Bildern erhalten und ausgedrückt werden kann: 1, 2, 3, 4, 5, 6, 7, 8, 9, 0." Wir erwähnen, daß damals das Zählen als fünfte Grundrechenart der Arithmetik galt.

Die ältesten russischen Münzen mit indischen Ziffern sind Goldmünzen im Nennwert von $^1/_4$ Dukaten. Sie tragen die Jahreszahl 1654. Allerdings wurden sie nicht für den eigentlichen Geldverkehr geprägt, sondern hauptsächlich als Geschenke, Belohnungen usw. verwendet. Wenig später kamen in Rußland westeuropäische Taler (sogenannte «ефимки») in Umlauf, die eine Prägung mit der Jahreszahl 1655 trugen. Unter Peter I. verschwanden die slawischen Ziffern fast vollständig von den Münzen. Letztmalig tauchten sie auf Kupfermünzen aus dem Jahre 1718 auf.[3])

In der Zeit nach Peter I. kamen die slawischen Ziffern schnell außer Gebrauch.

---

[1]) W. W. BOBYNIN, Skizzen zur Entwicklungsgeschichte der physikalisch-mathematischen Wissenschaften in Rußland, I. Aufl., S. 43.

[2]) Das erste Druckwerk überhaupt, in dem die Seitenzahlen indische Ziffern waren, stammt aus dem Jahre 1471. Es war ein Werk PETRARCAS, das „Liber de remediis utriusque fortunae", Coloniae 1471 (nach J. TROPFKE, Geschichte der Elementar-Mathematik, Band 1, II. Aufl., Berlin und Leipzig 1921, S. 26.) — *Anm. d. wissenschaftl. Red.*

[3]) Nach I. G. SPASSKI.

§ 6. Brüche

In der modernen Mathematik pflegt man die Brüche als geordnete Paare (m, n) aus ganzen Zahlen einzuführen, für die dann — in bekannter Weise — eine reflexive, symmetrische und transitive Gleichheitsrelation sowie die vier Spezies der Arithmetik definiert werden.[1]) Die ganzen Zahlen erscheinen bei dieser Einführung nachträglich wieder als spezielle Brüche, wie auch die Operationen mit den ganzen Zahlen Spezialfälle der Operationen für die Brüche werden. Nach der Erweiterung des Bereiches der ganzen Zahlen zum Bereich der Brüche (oder der Zahlenpaare) besitzt dann jede Gleichung der Form $ax = b$ — sofern nur $a \neq 0$ ist — für beliebiges ganzzahliges a und b eine eindeutige Lösung.

Allerdings ist das nicht die geschichtliche Entstehungsweise der Brüche. Es ist also keineswegs so, daß die Brüche als Ergebnis der Division ganzer Zahlen entstanden sind oder gar, daß sie geschaffen wurden, damit die Umkehrung der Multiplikation — die Division — immer ausführbar ist. Wäre dies nämlich der Fall gewesen, so müßten sie von Anfang an in rein logischer Weise benutzt worden sein, was sich sowohl in ihrer Behandlung als auch in ihrer Bezeichnungsweise widerspiegeln würde. Historisch betrachtet ist das aber nicht der Tatbestand. Dies zeigt schon ein Blick auf die Tabelle der ägyptischen Brüche (Abb. 3).

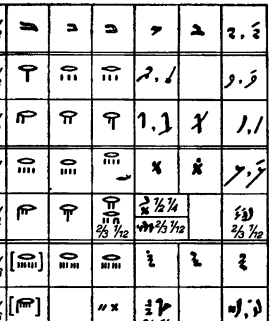


Abb 3.

Zunächst sehen wir, daß die Ägypter überhaupt nur für Brüche der Form $\frac{1}{n}$ und gewisse der Brüche $\frac{n-1}{n}$ Bezeichnungen besaßen. Für $n \geq 5$ wurden alle Brüche der Form $\frac{1}{n}$ auf völlig gleichartige Weise bezeichnet, nämlich durch das Zeichen ⊂, welches soviel wie „Teil" bedeutete. Unter dieses Zeichen wurde dann das Zeichen für n gesetzt. Dementsprechend wurde z. B. die Zahl $\frac{1}{10}$ durch 𓎆 bezeichnet. Bei den Ägyptern waren also alle diese Brüche — im Sinne unserer früheren Terminologie — *algorithmische* Zahlen. Zur Bezeichnung der Zahl $\frac{1}{2}$ benutzten die Ägypter — an Stelle des zu erwartenden Zeichens 𓏴𓏴 — das neue Zeichen 𓋴, während 𓎆 statt dessen die Zahl $\frac{2}{3}$ bezeichnete. Das Zeichen 𓏴, welches rechtmäßig die Zahl 1 zu bezeichnen gehabt hätte, diente tatsächlich zur Bezeichnung von $\frac{1}{3}$. In den gleichen Schriften wurde die Zahl $\frac{3}{4}$

[1]) In dem nachfolgenden Artikel von I. W. Proskurjakow werden die Brüche nach diesem Muster behandelt werden.

durch das Zeichen 𒐜 bezeichnet. In späteren Schriften wurde auch 𒐜 zur Bezeichnung von $\frac{1}{3}$ benutzt, während $\frac{3}{4}$ — dem eigentlichen Prinzip folgend — als $\frac{1}{2}+\frac{1}{4}$ dargestellt wurde. Die abweichenden Bezeichnungen für $\frac{1}{2}$ und $\frac{2}{3}$ dagegen blieben erhalten.

Ganz ähnlich wurden auch bei den Babyloniern (Abb. 4), Griechen und Römern für eine kleine Gruppe von Brüchen, in Abweichung von der allgemeinen Regel, individuelle Bezeichnungen verwendet.

So wurde die Zahl $\frac{1}{2}$ von den Griechen durch das Zeichen $<$ bezeichnet, während alle anderen Brüche der Form $\frac{1}{n}$ ($n \geq 3$) durch den oben rechts mit einem Strich versehenen Buchstaben für die Zahl n bezeichnet wurden (so bezeichnete γ' die Zahl $\frac{1}{3}$, während das Zeichen β', das an sich die Zahl $\frac{1}{2}$ zu bezeichnen gehabt hätte, die Zahl $\frac{2}{3}$ bezeichnete). Die Verhältnisse lagen also hier genauso wie in Ägypten. Ganz analog hierzu wurden in der akkadischen Sprache der Bruch $\frac{2}{3}$ durch „šittâ qâtâ“, d. h. „zwei Hände“, und $\frac{3}{4}$ durch „šalâstâ qâtâ“, d. h. „drei Hände“, bezeichnet. Bei den Römern wurden dieselben Brüche durch die Worte „bes“ (entstanden aus „binae partes“, „zwei Teile“) und „tres partes“ („drei Teile“) ausgedrückt.

Bruch	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
Alte, sumerische Schreibweise	𒐜	𒐜	𒐜
Spätere, akkadische Schreibweise	𒐜	𒐜	𒐜

Abb. 4

Wir kennen keine Sprache, in der das Wort für die Zahl $\frac{1}{2}$ aus dem Wort für die Zahl 2 entstanden wäre. So heißt $\frac{1}{2}$ auf lateinisch „semis“ (2 dagegen „duo“), auf deutsch „einhalb“ (2 dagegen „zwei“), auf russisch «половина» (2 dagegen «два») usw.

Brüche, die eine derartige individuelle Bezeichnung besitzen, wollen wir Individualbrüche (Knotenzahlen) nennen. Auch hier spiegelt — wie wir noch sehen werden — die Unterscheidung zwischen Individualbrüchen und algorithmischen Brüchen die Verschiedenheit ihrer Entstehung wider. Während nämlich die Individualbrüche unmittelbar aus den Bedürfnissen der Praxis entstanden, und zwar als selbständige Zahlgebilde, nicht aber aus den ganzen Zahlen abgeleitet wurden, waren die algorithmischen Brüche das Resultat einer nachfolgenden rechnerischen Behandlung.

Der Umstand, daß die Brüche nicht als Resultat der Division entstanden sind, wird aber nicht nur durch die bereits erwähnten Unterschiede in der Bezeichnungsweise, sondern auch durch gewisse aus der Geschichte bekannte Beispiele der Division ganzer Zahlen belegt. So wird in einer arabischen Handschrift aus dem 12. Jahrhundert u. Z. die folgende Aufgabe behandelt: „Man teile 100 Pfund zu gleichen

Teilen unter 11 Personen". Der Autor erhält bei der Lösung zunächst einen Rest von einem Pfund. Um dieses aufzuteilen, benutzt er nun keineswegs Brüche, sondern schlägt vor, das restliche Pfund in Eier umzutauschen, und zwar — wie er feststellt — gegen 91 Stück. Nachdem jeder der 11 Personen seine 8 Eier bekommen hat, schlägt der Autor vor, die restlichen 3 Eier dem zu geben, der die Mühe des Teilens gehabt hat, oder sie gegen Salz einzutauschen. Auf die gleiche Art verfährt auch ODO VON CLUNY (gestorben 942 oder 943). Um 1001 Pfund unter 100 Personen aufzuteilen, zerkleinert er das restliche Pfund solange in Unzen, Drachmen usw., bis die Anzahl der Teile größer als 100 ist. Da auch hiernach die Division nicht aufgeht, schlägt er vor, daß der verbleibende geringfügige Rest überhaupt weggelassen wird. Die Division führte also hier nicht zu den Brüchen, sondern zur Einführung immer kleinerer benannter Einheiten, während der dann verbleibende unbedeutende Rest einfach weggelassen wurde.

Um die Frage nach der Herkunft der Brüche zu beantworten, müssen wir unser Augenmerk noch auf einen anderen Prozeß richten, der schon in ältester Zeit zusammen mit dem Zählen auftauchte, nämlich auf das Messen (irgendwelcher Größen). In der geschichtlichen Entwicklung entstanden die Brüche durch Meßprozesse. Die Grundlage jeglichen Messens bilden irgendwelche Größensysteme, seien es Längen, Volumina, Gewichte o. dgl. Die Wahl der als Grundlage des jeweiligen Maßsystems dienenden Maßeinheit hing maßgeblich von den konkreten historischen Bedingungen ab.

Die Entwicklung der Maße ging in ungefähr den gleichen Etappen vor sich wie die Entwicklung des Zählens.

Im ersten Entwicklungsstadium der menschlichen Gesellschaft wurden Messungen ausschließlich nach dem Augenmaß vorgenommen. Das Maß wurde also ursprünglich als eine qualitative Eigenschaft des gemessenen Körpers aufgefaßt. Als bei der weiteren Entwicklung der Gesellschaft das Messen nach Augenmaß nicht mehr ausreichte, entstanden gewisse natürliche Maße, welche hauptsächlich von Teilen des menschlichen Körpers abgeleitet wurden: der Länge des Fußes, der Breite der Hand, der Entfernung vom Ellbogen bis zum ausgestreckten Mittelfinger usw.

Von dem Vorhandensein derartiger Maße zeugen Maßbezeichnungen, die sich noch bis heute in Europa erhalten haben: „Fuß" (die Länge desselben), „Zoll" (Breite des Mittelfingers an der Wurzel), „Yard" (Länge des Unterarms), „Palm" (Breite der Handfläche) usw. Zur gleichen Kategorie von Maßen gehört auch das „Machobaja Sashen" (Entfernung zwischen den Mittelfingern der seitwärts ausgestreckten Hände) und das „Kosaja Sashen" (Entfernung vom weit ausgestreckten großen Zeh des linken Fußes bis zum Mittelfinger der nach oben ausgestreckten rechten Hand), welche in Rußland noch lange Zeit in Gebrauch waren. Alle diese Maße dienten ursprünglich zum Vergleich der auszumessenden Größen mit dem entsprechenden Körperteil, aber auch zur Feststellung, welches Vielfache der Maßeinheit in der auszumessenden Größe enthalten ist. Hierzu wurde die Maßeinheit E so oft angetragen, bis ein Vielfaches $nE = E + \cdots + E$ schließlich ungefähr gleich der auszumessenden Größe A $\left(\text{d. h. bis } |A - nE| < \frac{1}{2} E\right)$ war. Dadurch wurde eine Verbindung zwischen dem Messen und dem Zählen hergestellt. Das Bedürfnis nach möglichst genauen Messungen führte dazu, daß die ursprünglichen Maße in zwei, drei oder vier Teile unterteilt wurden. Die als Resultat dieser Teilung

entstandenen kleineren Maßeinheiten bekamen dann individuelle Namen und wurden ihrerseits zu Messungen benutzt.

Somit entstanden die ersten konkreten Brüche als bestimmte Teile gewisser, ganz bestimmter Maße. Erst sehr viel später dienten die Namen dieser konkreten Brüche als Bezeichnungen für die entsprechenden Teile anderer Größen und erst danach zur Bezeichnung der abstrakten Brüche.

„Man darf nicht das ‚четверть' (Tschetwert) und das ‚четь', (Tschet) als Einheit des Flächen- und Hohlmaßes mit dem ‚четверть' oder ‚четь' (Viertel) als Bruch verwechseln" — schreibt L. W. TSCHEREPNIN[1]). So bezeichnete das Wort «полчетверти» (ein halbes Viertel) eine lange Zeit hindurch den Bruch $\frac{1}{8}$, während die Hälfte des Landmaßes Tschetwert durch das Wort «осьмина» (Osmina) bezeichnet wurde, das nur zur Bezeichnung eines ganz bestimmten Land- oder Hohlmaßes diente. Es war also unmöglich, etwa von «осьмина книги» als von „einem Achtel des Buches" oder von «осьмина пути» als „dem achten Teil des Weges" zu sprechen. Erst sehr viel später diente das Wort «осьмина» auch zur Bezeichnung des Bruches $\frac{1}{8}$.

Ganz ähnlich bezeichnete die Unze des römischen Maßsystems ursprünglich nur den zwölften Teil der Geld-Gewichts-Einheit As. Im Laufe der Zeit ging man dann dazu über, auch den zwölften Teil einer beliebigen Größe mit dem Wort „Unze" zu belegen, d. h., man gebrauchte es zur Bezeichnung des abstrakten Bruches $\frac{1}{12}$, so daß es auch möglich wurde, von „fünf Unzen des Weges" und von „sieben Unzen des Buches" zu sprechen.

So waren die Individualbrüche im Anfang nichts anderes als bestimmte Bruchteile bestimmter Maße. Abstrakte Brüche gab es damals noch nicht.

Diese Behauptung wird vollständig bestätigt durch die Betrachtung der Zeichen, die anfangs zur Bezeichnung der Individualbrüche benutzt wurden. Die babylonischen Schriftzeichen für $\frac{1}{2}$, $\frac{1}{3}$ und $\frac{2}{3}$ waren Darstellungen von Gefäßen, also konkreten Hohlmaßen. Die ägyptische Flächeneinheit war das „Setat", ein Quadrat der Seitenlänge von 1 Chet (1 Chet = 100 Ellen). Das Viertel eines Setat wurde „Zerbrochenes" genannt und durch $\times$ bezeichnet. Später war der Begriff „Zerbrochenes" die Bezeichnung für den Bruch schlechthin, während das Zeichen $\times$ im hieratischen System den abstrakten Bruch $\frac{1}{4}$ bezeichnete. Es besteht Grund zur Annahme, daß die Hälfte des Setat ursprünglich durch das Zeichen $<$ (oder $\subset$) dargestellt wurde. Als Einheit der Volumenmessung diente das „Hekat" („Scheffel", ungefähr $4\frac{1}{2}$ Liter). Da diese Einheit sehr klein war, benutzte man gewöhnlich 100 Hekat als Maßeinheit. Die Hälfte und das Viertel dieser Einheit wurden gleichfalls mit Hilfe der Zeichen $<$ und $\times$ bezeichnet, die man unter das Zeichen für 100 Hekat setzte.[2])

Schriftbilder vom Typus $\times$ H, wobei der Buchstabe „H" etwa das Zeichen für das Hekat vertreten möge, sind also völlig analog dem römischen Ausdruck „fünf

---

[1]) Л. В. Черепнин, Русская метрология (L. W. TSCHEREPNIN, Die russische Metrologie), Moskau 1944, Seite 53.

[2]) И. Н. Веселовский, Египетская наука в Греции (I. N. WESSELOWSKI, Die ägyptische Wissenschaft in Griechenland), Veröffentlichungen des Instituts für Geschichte der Naturkunde, Band II, 1948, Seite 437—440.

Unzen des Weges" gebildet. Die Verwendung des Namens (Zeichens) eines bestimmten Teiles eines konkreten Maßes für den gleichen Teil eines anderen Maßes stellt einen wichtigen Schritt auf dem Wege zur Schöpfung der abstrakten Brüche dar.

Eine andere Maßeinheit der Ägypter war das „Chen" $\left(=\frac{1}{10}\text{ Hekat}\right)$. Der 32. Teil des Chen wurde allgemein „Teil" genannt und durch ⊂ bezeichnet. Späterhin wurde — wie wir bereits sahen — dieses Zeichen allgemein zur Bezeichnung der Brüche verwendet.

Die Bezeichnungen für die Individualbrüche waren also anfangs nichts anderes als Bezeichnungen für konkrete Maßeinheiten. Später bezeichneten sie — zunächst in Verbindung mit anderen Maßeinheiten geschrieben — konkrete Teile dieser anderen Maßeinheiten. So bildete sich nur ganz allmählich der (abstrakte) Bruch als das allen gleichartigen Teilen verschiedener Maße Gemeinsame heraus. Es wurde keineswegs sofort erkannt, daß die Darstellung des Bruches als m-faches des n-ten Teiles einer gewissen Größe x, als $\overline{x} = \frac{m}{n}x$, eine rein arithmetische Eigenschaft der Brüche, also unabhängig von den speziellen Eigenschaften des Größenbereiches ist, dem x angehört. Dieser Prozeß verlief viel langsamer als die Entstehung der abstrakten ganzen Zahlen. Es genügt zu bemerken, daß selbst die Römer noch ausschließlich konkrete Brüche benutzten.

Es besteht Grund zur Annahme, daß es ursprünglich nur dyadische Brüche gab. „Der erste Bruch, mit dem die Menschheit bekannt wurde, war — wie unmittelbar einzusehen ist — das Halbe in seinen konkreten Formen, nämlich in Form der Hälfte irgendwelcher realer Gegenstände."[1]) Alle unmittelbar danach erscheinenden Brüche waren Dualbrüche. Diese Etappe der Entwicklung fand ihren Niederschlag in der altägyptischen Metrologie. So unterteilten die Ägypter ihre Flächeneinheit, das Setat, in zwei, vier, acht, sechzehn und zweiunddreißig Teile. Die gleiche Unterteilung besaß auch die Volumeneinheit, das Hekat. Für alle diese Teile gab es individuelle Namen. „Da alle diese Teile konkrete Dinge waren, wurde mit ihnen genau so wie mit den ganzen Gegenständen gerechnet."[2]) Es ist ganz natürlich, daß als Zähler der Brüche zunächst nur die Zahl Eins möglich war. Später wurden zu den genannten Brüchen noch $\frac{1}{3}$ und seine dyadischen Unterteilungen hinzugenommen. Die gleiche Einteilung findet man auch in der späteren ägyptischen Divisionsmethode (Herausbildung der $\frac{1}{2}$- und $\frac{2}{3}$-Reihe). Ähnlich war auch das altrussische Maßsystem aufgebaut. Einheit des Landmaßes war das „Socha" («соха»), das in „Halbsocha", „Drittelsocha", „Viertelsocha", „Halbdrittelsocha", „Halb-Halbdrittelsocha" und „Halb-Halb-Halbdrittelsocha" («пол-пол-полтреть сохи») unterteilt wurde.

Ein anderes Landmaß, das schon genannte Tschetwert, wurde zunächst gedrittelt und dann sukzessive halbiert. Die kleinste Einheit war das „Halb-Halb-Halb-Halb-Halb-Halb-Halb-Halb-Halb-Drittel" («пол-пол-пол-пол-пол-пол-пол-пол-пол-пол-третних»).

Das System der altrussischen Brüche baute sich nach demselben Prinzip wie das Maßsystem auf. Die grundlegenden Brüche waren:

---

[1]) W. W. Bobynin, Bemerkungen zu den Werken N. M. Bubnows, S. 114.
[2]) W. W. Bobynin, loc. cit., S. 115.

Bruch	Russische Bezeichnung	Deutsche Übersetzung
$\frac{1}{2}$	пол	ein Halbes
$\frac{1}{3}$	треть	ein Drittel
$\frac{1}{4}$	четь oder четверть	ein Viertel
$\frac{1}{6}$	полтрети	ein halbes Drittel
$\frac{1}{8}$	полчети oder полчетверти	ein halbes Viertel
$\frac{1}{12}$	пол-полтрети	ein halbes Halbdrittel
$\frac{1}{16}$	пол-полчети	ein halbes Halbviertel
$\frac{1}{24}$	пол-пол-полтрети oder малые трети	ein halbes Halb-Halb-Drittel oder ein kleines Drittel
$\frac{1}{32}$	пол-пол-полчети oder малые чети	ein halbes Halb-Halb-Viertel oder ein kleines Viertel

Die übrigen Brüche wurden in den altrussischen Quellen meistens als Summen oder Differenzen aus diesen Brüchen dargestellt. So wurde $\frac{11}{24}$ (schematisch) als $\frac{1}{3}+\frac{1}{12}+\frac{1}{24}$ und $\frac{29}{96}$ als $\frac{1}{3}-\frac{1}{32}$ geschrieben.

Natürlich wurden alle diese Beziehungen im Original in Worten ausgedrückt. Um auszudrücken, daß an einer gewissen Anzahl von Einheiten eine halbe Einheit fehlte, setzte man vor das betreffende Wort das Wort «пол» (Halb). In diesem Sinne bezeichnete «полтретьи» (Halb-Drei) die Zahl $2\frac{1}{2}$, «полчетверти» die Zahl $3\frac{1}{2}$, «полпяты» die Zahl $4\frac{1}{2}$, «полшесты» die Zahl $5\frac{1}{2}$ usw. Als Überrest dieser Bezeichnungsweise hat sich in Rußland das Wort «полтора» (entstanden aus «полвтора» Halb-Zwei) für $1\frac{1}{2}$ erhalten. Eine ähnliche Bezeichnungsweise gibt es in Rußland (wie auch bei uns in Deutschland — *Anm. d. wissenschaftl. Red.*) in der Zeitrechnung (Halb-Fünf-Uhr, Halb-Sechs-Uhr usw.).[1])

Nach einer Hypothese von I. N. Wesselowski kamen die Ägypter im Zusammenhang mit der Zeitrechnung von den dyadischen Brüchen zu den allgemeinen Brüchen der Form $\frac{1}{n}$. Jedoch wurde auch hier nur der „Zähler" Eins zugelassen. Z. B. konnte im ägyptischen System die Zahl $\frac{15}{17}$ nicht durch ein einheitliches Symbol ausgedrückt werden. Der Begriff $\frac{15}{17}$ als einheitliche Zahl, als einheitlicher Bruch, existierte bei den Ägyptern nicht. Alle Brüche der Form $\frac{m}{n}$ mit $m>1$ stellten sie nach einem

[1]) Eine ähnliche Bildung hat sich auch in der dänischen Sprache erhalten, wo die Zahl 50 durch das Wort „halvtresindstyve" bezeichnet wird, was wörtlich „Halb-Drei mal Zwanzig" bedeutet.

allgemeinen Prinzip als Summen der Form $\sum_i \frac{1}{n_i}$ dar, wobei $n_i \neq n_j$ für $i \neq j$. Da man die Zahl m in ihrer dyadischen Darstellung

$$m = 2^{n_1} + 2^{n_2} + \cdots + 2^{n_k}, \quad n_1 > n_2 > \cdots > n_k \geqq 0$$

geschrieben denken kann, genügt es zur Darstellung von $\frac{m}{n}$ als $\sum \frac{1}{n_i}$, daß man alle Brüche der Form $\frac{2}{n}$ in dieser Gestalt schreiben kann. Tatsächlich finden sich in ägyptischen Papyri auch Tabellen für eine derartige Darstellung von $\frac{2}{n}$ ($n = 3, 5, \ldots, 101$). Es gibt eine ganze Reihe von verschiedenen Hypothesen über die Entstehung dieser Tabellen, auf die wir hier nicht näher eingehen wollen.[1]) Es sei nur erwähnt, daß schon bei der Anfertigung dieser Tabellen, die vor allem praktischen Zwecken dienten, die ägyptischen Rechner zwangsläufig auf zahlentheoretische Probleme geführt wurden.

Auch in Griechenland wurden — genau wie in Ägypten — hauptsächlich Brüche der Form $\frac{1}{n}$ benutzt. Offenbar war anfänglich der Bereich der griechischen Brüche auch allein auf diese beschränkt. Zur Bezeichnung des Bruches $\frac{1}{n}$ verwendeten die Griechen — wie bereits gesagt — den mit einem Strich versehenen Buchstaben für n. So schrieb man $\frac{1}{23}$ als $\kappa\gamma'$ ($\kappa = 20$, $\gamma = 3$).

Heron von Alexandrien (1. bis 2. Jahrhundert v. u. Z.) gebrauchte daneben auch Brüche der Form $\frac{m}{n}$. Zu ihrer Bezeichnung schrieb er erst das Zeichen für m mit einem Strich und wiederholte dann zweimal das Zeichen für n mit zwei Strichen. So schrieb er den Bruch $\frac{2}{5}$ als $\beta'\varepsilon''\varepsilon''$ und $\frac{23}{33}$ als $\kappa\gamma'\lambda\gamma''\lambda\gamma''$ ($\kappa = 20$, $\lambda = 30$, $\gamma = 3$).

Diophant (3. Jahrhundert u. Z.) bezeichnete — wie auch wir heute — die Brüche mit Hilfe eines Bruchstriches, nur schrieb er den Nenner über den Bruchstrich und den Zähler unter denselben. Er schrieb also den Bruch $\frac{21}{25}$ als $\frac{\kappa\varepsilon}{\kappa\alpha}$ ($\kappa\varepsilon = 25, \kappa\alpha = 21$). Der Bruch $\frac{1\,270\,568}{10\,817}$ hätte bei ihm die Form $\frac{\alpha,\omega\iota\zeta}{\varrho\kappa\zeta\cdot\varphi\xi\eta}$ gehabt.

Daneben verwendete Diophant noch eine andere Schreibweise, indem er zunächst den Zähler des Bruches, dann den Nenner und dazwischen das Wort *μοριον* (Teilchen) schrieb. Also:

$$\frac{3\,069\,000}{331\,776} = \overline{\tau\varsigma}\cdot{}^{,}\overline{\theta}\mu o\varrho\cdot\overline{\lambda\gamma}\cdot{}^{,}\overline{\alpha\varphi o\varsigma}.$$

Genau wie in Ägypten stellte man auch in Griechenland häufig die Brüche als Summen von Brüchen mit Zähler Eins dar, z. B. $\frac{9}{20} = \frac{1}{4} + \frac{1}{5} = \delta'\varepsilon'$ (zur Bezeichnung der Addition wurden gewöhnlich die Summanden hintereinander geschrieben).

---

[1]) Vgl. hierzu die bereits erwähnte Arbeit von I. N. Wesselowski und C. A. Яновская, К теории египетских дробей. (S. A. Janowskaja, Zur Theorie der ägyptischen Brüche), Veröffentlichungen des Instituts für Geschichte der Naturkunde, Band I, 1947. (Siehe auch O. Neugebauer, Die Grundlagen der ägyptischen Bruchrechnung, Berlin 1926, und K. Vogel, Die Grundlagen der ägyptischen Mathematik in ihrem Zusammenhang mit der $2/n$-Tabelle des Papyrus Rhind, München 1929. — *Anm. d. wissenschaftl. Red.*)

Für ihre astronomischen Berechnungen benutzten die Griechen die babylonischen Sexagesimalbrüche, auf die wir weiter unten noch zu sprechen kommen.

Wie schon gesagt wurde, besaßen die Römer — im Gegensatz zu den Griechen — nur konkrete Brüche, nämlich die Teile ihrer Geldeinheit As. Diese war in 12 Unzen unterteilt. Später wurde die Unze auch zur Unterteilung anderer Maßgrößen benutzt. In Rom waren jedenfalls nur benannte Brüche bekannt, in dieser Hinsicht blieben die Römer sogar hinter den Ägyptern, also um 1500 Jahre zurück. Zur Bezeichnung der Unze diente ein einfacher Strich „—". Die Hälfte wurde durch den Buchstaben „S" (den Anfangsbuchstaben von „semis") bezeichnet. Die übrigen Zwölfer-Brüche wurden als Kombinationen aus diesen beiden Zeichen gebildet, beispielsweise

$$\frac{7}{12}=\frac{6}{12}+\frac{1}{12}=\frac{1}{2}+\frac{1}{12}=\text{S}-.$$

$\frac{1}{24}$ As wurde „semunzia" genannt, $\frac{1}{36}$ As hieß „duella", $\frac{1}{48}$ As „sicilius", $\frac{1}{72}$ As „sextula".

Eine andere Reihe von Unterteilungen der Grundeinheit war: $1 = \text{As}$, $\frac{1}{12} = \text{unzia}$, $\frac{1}{24} = \text{semunzia}$, $\frac{1}{288} = \text{scripulus}$, $\frac{1}{576} = \text{simplium}$, d. h., jeder dieser Brüche entsteht aus dem vorhergehenden durch abwechselnde Multiplikation des Nenners mit 2 bzw. 12. Die Gewichtseinheit As mit ihrer Unterteilung in Unzen war noch lange Zeit hindurch in den Apotheken gebräuchlich. Auch fanden die Zwölfer-Brüche der Römer bei der mittelalterlichen Abakusrechnung Verwendung. Die Geschichte der römischen Brüche liefert ein Beispiel der „direkten Verwendung eines metrologischen Systems, für welches die Rechenregeln und -methoden der abstrakten Brüche zur Ausführung von Rechenoperationen ausgearbeitet wurden"[1]). Dergleichen beobachtet man auch anderenorts.

Die Inder dehnten die von ihnen erfundene Positionsschreibweise nicht auf die Brüche aus. Vielmehr bezeichneten sie die Stammbrüche genau wie wir, indem sie den Zähler über den Nenner schrieben, nur setzten sie dazwischen keinen Bruchstrich. So wurde der Bruch $\frac{1}{3}$ in der Handschrift von Bakhsālī durch $\frac{1}{3}$ bezeichnet. Zur Bezeichnung gemischter Zahlen schrieben sie die ganze Zahl noch über den Zähler. In diesem Sinne schrieben sie z. B. die Zahl $1\frac{1}{3}$ als $\frac{1}{1}$ $\frac{}{3}$. Diese Schreibweise findet sich erstmalig bei dem tadshikischen Gelehrten Alnasawi (um 1030 u. Z.). Bei ihm ist diese Schreibart so konsequent ausgebildet, daß er sogar in dem Fall, daß eine ganze Zahl fehlte, noch eine Null über den Zähler schrieb; so schrieb er den Bruch $\frac{1}{11}$ als $\frac{0}{1}$ $\frac{}{11}$. Der Bruchstrich findet sich erstmalig bei Alḥaṣṣar (Abū Zaqarijjā Alḥaṣṣār, 12. Jahrhundert u. Z.) erwähnt. Regelmäßig wurde der Bruchstrich jedoch erst bei dem bereits genannten Leonardo von Pisa (um 1250) verwendet. Trotzdem dauerte es noch bis in das 16. Jahrhundert, ehe sich der Bruchstrich allgemein eingebürgert hatte. Im Mittelalter findet sich eine reiche

[1])W. W. Bobynin, Bemerkungen zu den Werken N. M. Bubnows, S. 119.

Anzahl von Methoden zur Bezeichnung der Brüche. Häufig schrieb man Zähler und Nenner noch in römischen Ziffern, wobei oft recht eigenartige Multiplikationsprinzipien angewendet wurden. So findet sich in einem deutschen Rechenbuch aus dem Jahre 1514 der Bruch $\frac{200}{460}$ als $\frac{||^{C}}{|||^{C}\,LX}$ dargestellt. In einer Handschrift aus der Mitte des 14. Jahrhunderts wird $3\,\bar{\bar{5}}$ statt $\frac{3}{5}$ und $4\,\bar{\bar{7}}$ statt $\frac{4}{7}$ geschrieben. Die Zahl $\frac{1}{2}$ wurde oft durch $\div$ bezeichnet; dann bezeichnete z. B. $4 \div$ die Zahl $4\frac{1}{2}$. Bei der Aussprache der Brüche setzte man im Mittelalter stets noch das Wort „Teil" hinzu,[1]) man sprach also von „Zwei Fünf-Teile" für $\frac{2}{5}$.

In Rußland benutzte man während des 16. und 17. Jahrhunderts zur Aussprache der Brüche, deren Nenner zwischen 5 und 10 lag, die Endung «ина», sagte also beispielsweise «седьмина» für $\frac{1}{7}$ und «десятина» für $\frac{1}{10}$. War der Nenner größer als 10, so wurde das Wort «жеребей» hinzugesetzt, man sagte also beispielsweise «пять тринадцатых жеребьев» für $\frac{5}{13}$. In alten russischen Handschriften wurden die Brüche selbst als „Teile" bezeichnet; später nannte man sie auch — entsprechend der lateinischen Bezeichnung „numeri fracti" — „gebrochene Zahlen". Diese Terminologie benutzte auch MAGNIZKI in seiner „Arithmetik".

Eine einheitliche, algorithmische Bezeichnung für beliebige Brüche findet sich zuerst in Babylonien, wo auch die Brüche mit Hilfe des sexagesimalen Positionssystems bezeichnet wurden. Dazu wurde die Einheit in bekannter Weise fortlaufend in 60-tel, 3600-tel usw. unterteilt; in der Schrift wurde dann nur die Anzahl der jeweiligen Teile vermerkt. Die individuellen Bezeichnungen für die genannte kleine Gruppe von Brüchen $\left(\frac{1}{2}, \frac{1}{3}, \frac{2}{3}\text{ usw.}\right)$ wurden im Laufe der Zeit aus den mathematischen Texten fast vollständig verdrängt. Selbst für den Bruch $\frac{1}{2}$ wurde schließlich die algorithmische Darstellung ⟨⟨⟨ (0,30 — da sexagesimales System! *Anm. d. wissenschaftl. Red.*) verwendet. Die sexagesimalen Brüche besaßen den unbestreitbaren Vorzug, daß mit ihnen genau wie mit den ganzen Zahlen gerechnet werden konnte. Deshalb fanden sie späterhin auch außerhalb der Grenzen Babyloniens Verwendung. Wahrscheinlich sind sie nicht später als in der Mitte des 2. Jahrhunderts v. u. Z. — in der Hauptsache durch astronomische Werke — nach Alexandrien gelangt. Da das alphabetische System der Griechen sehr schlecht zur Darstellung auch nur einigermaßen großer Zahlen und zum Rechnen mit ihnen geeignet war, mußten sich die griechischen Astronomen zur Berechnung ihrer Tabellen entweder eines anderen Bezeichnungssystems für die ganzen Zahlen bedienen und dann den Radius des Einheitskreises hinreichend groß wählen (dann konnten sie die Sehnen des Kreises mit genügender Genauigkeit als Vielfache der Maßeinheit des Radius ansehen) oder sie mußten eine andere Darstellungsweise für die Brüche finden. Die griechischen Astronomen entschieden sich für den zweiten Weg. Sie gebrauchten für die ganzen Zahlen nach wie vor die alphabetische Bezeichnungsweise und daneben für die Brüche das sexagesimale System der Babylonier, in welchem sie nur die Form der

[1]) Ein Überrest dieser Bezeichnungsweise ist die Endung „-tel" bei der deutschen Aussprache der Brüche (Drit*tel* usw.). — *Anm. d. wissenschaftl. Red.*

Ziffern änderten. Die Zahlen von 1 bis 59 bezeichneten sie nicht nach dem additiven Prinzip der Babylonier durch ▼ und ◀, sondern mit Hilfe der Buchstaben ihres Alphabets.

Der bekannte griechische Astronom CLAUDIUS PTOLEMAEUS (Mitte des 2. Jahrhunderts u. Z.) teilte den Umfang des Kreises in 360 Teile. Diese Teile bezeichnete er häufig als τμήματα, was soviel wie „Segment" bedeutete, oder er nannte sie einfach μοῖραι, abgekürzt durch μ°. Später bezeichnete man sie einfach durch einen Kreis, wie er heute noch zur Bezeichnung des Grades Verwendung findet. Das Wort „Grad" selbst ist nach Meinung von G. H. F. NESSELMANN (1842) arabischen Ursprungs.

Jeden der erhaltenen 360 Teile (Grade) unterteilte PTOLEMAEUS in 60 Teile, die er λεπτα, d. h. wörtlich „Kleinigkeit", oder „die ersten Sechziger" nannte. Die folgende sexagesimale Unterteilung nannte er „die zweiten Sechziger", die nächste „die dritten Sechziger". Bei der Übersetzung dieser Bezeichnungen ins Lateinische wurde daraus „minuta prima", „minuta secunda", „minuta tertia". Das Wort „minuta" leitet sich von „minuere" (vermindern) ab und ist damit die wörtliche Übersetzung des griechischen λεπτα. Hieraus entstanden unsere heutigen Bezeichnungen „Minute" und „Sekunde". Für gewöhnlich verwendete PTOLEMAEUS nicht die voll ausgeschriebenen Bezeichnungen, sondern Abkürzungen. So schrieb er 37°4′55″ als $\mu^{\circ}\lambda\overline{\zeta}\delta'\nu\varepsilon''$. Manchmal ließ er auch das Zeichen μ° aus und setzte über die Buchstabenverbindung, die die Anzahl der Grade angab, einen Querstrich. Beim Schreiben der sexagesimalen Brüche verwendeten die Griechen das Zeichen *o* zur Bezeichnung nicht vorkommender Sexagesimalstellen. Dieses Zeichen ähnelt in seiner Form unserer heutigen Null. So wurde der Winkel 12°0′23″ als $\iota\overline{\beta}o'\kappa\overline{\delta}''$ geschrieben. Da die ganzen Zahlen weiterhin im üblichen alphabetischen System geschrieben wurden, diente das Zeichen *o* außerdem zur Bezeichnung der Zahl 70. Trotzdem waren bei der Benutzung dieses Zeichens zum Schreiben der Sexagesimalbrüche keine Verwechslungen zu befürchten, da hierbei die Zahl 70 nicht auftreten konnte (weil nämlich an den einzelnen Stellen nur Zahlen bis 59 vorkommen), der Buchstabe *o* (Omikron) konnte also getrost diese neue Bedeutung erhalten. Man nimmt an, daß dieses Nullzeichen als Abkürzung aus οὐδέν (nichts) entstanden ist.

Wir weisen darauf hin, daß bei Kennzeichnung der Ordnungen die Verwendung eines solchen Zeichens überflüssig ist. Dagegen war ein solches Zeichen im babylonischen System, in dem beim Schreiben die Ordnungen nicht unterschieden wurden, sehr wesentlich.

Aber nicht nur die Griechen, sondern auch die Völker des Nahen und Mittleren Ostens und danach auch Westeuropas übernahmen die sexagesimalen Brüche der Babylonier. Allerdings fanden sie dort ausschließlich in wissenschaftlichen Werken Verwendung. Im Mittelalter wurden in Europa zur Darstellung der ganzen Zahlen das Dezimalsystem und zur Darstellung der Brüche das sexagesimale System und verschiedene Darstellungsarten durch Stammbrüche verwendet.[1])

[1]) Die Bruchrechnung galt lange Zeit hindurch als der schwierigste und komplizierteste Teil der Arithmetik. In Deutschland spricht man heute noch vom „in die Brüche geraten", wenn es sich um eine aussichtslose Sache handelt. Auch in den Schulbüchern findet sich die Behandlung der Brüche meistens am Ende der Darstellungen, damit der Schüler vorher ohne allzugroße Anstrengungen die Gesetze der Arithmetik erlernt. Erfahrungsgemäß gelangt bis zu diesem Abschnitt die Mehrzahl der Schüler ohne Mühe.

Erste Andeutungen der Dezimalbrüche finden sich schon bei den Indern. Zur Berechnung der Quadratwurzel aus einer Nicht-Quadratzahl hängten sie an den Radikanden so viele Paare von Nullen wie nötig waren, um genügend viele Stellen der Wurzel zu erhalten. Allerdings dehnten die Inder ihre positionelle Bezeichnungsweise nicht auf diese „Dezimalbrüche" aus, sondern schrieben sie stets mit einem Nenner. Ähnliche Methoden finden sich auch bei den Mathematikern des Mittleren Ostens, z. B. bei ALNASAWĪ.

Ein ähnliches Verfahren zum Ausziehen der Quadratwurzel verwendete in Europa zuerst JOHANNES VON SEVILLA in seinem schon erwähnten „Liber alghoarismi de practica arismetrica" (12. Jahrhundert). In der Mitte des 15. Jahrhunderts benutzten manche (REGIOMONTANUS u. a.) zur Berechnung der trigonometrischen Funktionen einen Kreis vom Radius 10^6 oder 10^7 und bekamen damit de facto die trigonometrischen Größen in Form von Dezimalbrüchen.

Im 15. bis 16. Jahrhundert begegnet man schon häufiger Brüchen mit dezimalen Nennern. So findet man bei einem französischen Mathematiker[1]) des 16. Jahrhunderts schon eine ganze, weit entwickelte Theorie der Brüche mit Dezimalnennern. Auch G. CARDANO (16. Jahrhundert) benutzte bereits derartige Brüche[2]). Der deutsche Schreiber GRAMMATEUS (1523) schlug vor, diese Brüche beim Vergleichen von gemeinen Brüchen zu benutzen. Um festzustellen, welcher der Brüche $\frac{5}{8}$ und $\frac{2}{3}$ größer ist, hängte er an die Zähler Nullen an (spaltete sie also in Dezimalstellen auf) und teilte sie dann durch die Nenner: $\frac{500}{8} = 62\frac{1}{2}$ und $\frac{200}{3} = 66\frac{2}{3}$, also $\frac{2}{3} > \frac{5}{8}$.

Der Franzose ORONTIUS FINAEUS (um 1550) berechnete die Quadratwurzel aus 10, indem er an den Radikanden 10 sechs Nullen hängte, also faktisch mit Dezimalbrüchen rechnete. Indes formte er anschließend den erhaltenen Wert in den entsprechenden Sexagesimalbruch um. Dieses Beispiel zeigt deutlich, daß die Dezimalbrüche, trotz ihres effektiven Auftretens, bis in das letzte Viertel des 16. Jahrhunderts hinein nicht systematisch verwendet wurden.

Den ersten folgerichtigen Gebrauch von Dezimalbrüchen machte der flämische Kaufmann und Ingenieur SIMON STEVIN (1548 bis 1620). Im Jahre 1584 veröffentlichte er eine Prozententabelle in flämischer Sprache, die bald darauf auch in französischer Sprache erschien; im folgenden Jahre publizierte er seine Arbeit „La disme enseignant facilement expédier par nombres entiers sans rompus tous comptes se rencontrans aux affaires des hommes" („Das Dezimalsystem, nach welchem man alle im täglichen Leben vorkommenden Rechnungen in ganzen Zahlen ohne Brüche leicht ausführen kann"). In dieser insgesamt sieben Seiten umfassenden Abhandlung werden u. a. die Dezimalbrüche eingeführt und die Grundrechenarten der Arithmetik an ihnen erklärt. STEVIN erkannte klar den großen Nutzen der Dezimalbrüche, insbesondere bei Verwendung eines dezimalen Maßsystems, dessen erster energischer Verfechter er außerdem war. Sein Wunsch, daß das Dezimalsystem

[1]) Vgl. hierzu GANDZ, The invention of the decimal fractions and application of the exponential calculus by Immanuel Bonfils of Tarascon (Isis, XXV (1), 1936).

[2]) Der Mathematiker und Astronom GIJĀT EDDĪN ALKĀŠĪ aus Samarkand gab bereits um das Jahr 1420 ein mehr als fünfzehnstelliges Verhältnis zwischen der Länge einer Sehne eines Kreises und seinem Radius in Form eines Dezimalbruches an.

von möglichst weiten Kreisen angenommen werden möge, kommt z. B. darin zum Ausdruck, daß er „La disme" nicht in Latein, sondern in der flämischen und französischen Umgangssprache schrieb.

Die Bezeichnung, die STEVIN vorschlug, weicht allerdings noch beträchtlich von der heutigen ab. Statt eines Komma benutzte er die Zahl Null, die er zur Unterscheidung in einen Kreis einschloß. Ferner schrieb er hinter jede Dezimalstelle die Stellenzahl, gleichfalls in einen Kreis eingeschlossen. So schrieb er z. B. die Zahl 35,912 als 35 ⓪ 9 ① 1 ② 2 ③. In dieser Schreibweise kommt also dasselbe Prinzip zur Anwendung, wie in der von STEVIN vorgeschlagenen Bezeichnung der Potenzen von Unbekannten.

Im Jahre 1608 wurde „La disme" von RICHARD NORTON ins Englische übersetzt. 1619 erschien in England ein Werk über Dezimal-Arithmetik von HENRY LIGHT. Trotzdem verdrängten die Dezimalbrüche keineswegs sofort die anderen Bezeichnungsarten. Sie mußten vielmehr — genau wie auch das positionelle Dezimalsystem — einen hartnäckigen Kampf gegen alte Traditionen führen und konnten erst allmählich den Platz erobern, den sie heute einnehmen.

Auf dem Kontinent wurden nach STEVIN die Dezimalbrüche systematisch von J. BÜRGI (1552 bis 1632) in einem handgeschriebenen Werk ungefähr aus dem Jahre 1592 verwendet. BÜRGI benutzte als Trennzeichen zwischen dem ganzen und dem gebrochenen Anteil eine Null, die er über die Einerstelle setzte.

Im Jahre 1603 erschien in Frankfurt a. M. die „Logistica decimalis" („Dezimal-Arithmetik") von J. H. BEYER, in welcher er die vier Grundrechenarten der Arithmetik für die Dezimalbrüche behandelte. BEYER nannte diese Brüche seine eigene Erfindung. Seine Bezeichnungsweise erinnert an die damals vorherrschende Bezeichnungsart der Sexagesimalbrüche. So schrieb er die Zahl 123,459872 als

0 I IIIIIV V VI 0 III VI VI

12 3-4-5-9-8-7-2 oder kürzer als 123-459-872. Die Zeichenverbindung 54 bezeichnete bei ihm die Zahl 0,000054. BEYER bemerkte, daß sich die von ihm benutzten Dezimalbrüche grundlegend von den „gemeinen" Brüchen unterscheiden, bei denen der Zähler über dem Nenner geschrieben wird.[1])

Während des 16. und 17. Jahrhunderts finden sich nebeneinander die verschiedensten Bezeichnungsweisen für die Dezimalbrüche.

Das Komma als Trennzeichen wurde zuerst von BÜRGI und KEPLER (1571—1630) benutzt, die es neben Klammern gebrauchten.

Mit den Dezimalbrüchen befaßte sich auch J. NEPER (NAPIER) (1550—1617), der ihre Theorie in seiner „Rabdologia" niederlegte. Dort werden die Brüche in der Regel nach demselben Prinzip bezeichnet, das wir heute zur Bezeichnung von Graden, Minuten, Sekunden usw. benutzen. Der Bruch 28,675 z. B. wurde von ihm als 28°6′7″5‴ geschrieben und „28 Ganze, 6 Primen, 7 Sekunden, 5 Terzen" gelesen. In dem gleichen Werk benutzte NEPER als Trennzeichen auch einen Punkt, wie er noch heute in England und Amerika verwendet wird.[2]) Jedoch findet man noch im 17. Jahrhundert neben den Dezimalbrüchen auch Sexagesimalbrüche im Gebrauch, welche erst im 18. Jahrhundert gänzlich verdrängt wurden.

[1]) Man nimmt an, daß die deutschen Wörter „Dezimalbrüche", „gemeine Brüche" und „Dezimalrechnung" gleichfalls auf BEYER zurückgehen. (Vgl. J. TROPFKE, a. a. O., Band 1. 2. Aufl., S. 144.) — *Anm. d. wissenschaftl. Red.*

[2]) Ein Punkt als Trennzeichen bei Dezimalbrüchen wurde erstmalig von CHR. CLAVIUS (1593) verwendet.

Die erste russische Darstellung der Dezimalbrüche findet sich in der „Arithmetik" von L. MAGNIZKI. MAGNIZKI unterschied zwischen logistischer oder astronomischer Arithmetik und Dezimalarithmetik. In der ersten behandelte er die Lehre von den Sexagesimalbrüchen, die er in der Form $\overset{0}{51}\ \overset{I}{25}\ \overset{II}{42}\ \overset{III}{51}\ \overset{IV}{25}$ schrieb. Die Dezimal-Arithmetik verwendete er ausschließlich in der Geometrie. Dabei ging er auch auf dezimale Längen- und Flächenmaße ein. Er führte als Längeneinheit die „Rute" („deutscher Sashen") ein mit den Unterteilungen 1 Rute = 10 Fuß, 1 Fuß = 10 Zoll oder Palm, 1 Zoll = 10 Gran oder Sjoren, 1 Gran = 10 Skrupel oder Bruchteile.

Im 18. Jahrhundert breiteten sich die Dezimalbrüche immer weiter aus. Endgültig festigten sie sich nach der Einführung der dezimalen Maß- und Gewichtssysteme.[1]) Heute werden auch im Alltagsleben — besonders in der Statistik — Prozente (das sind ja nichts anderes als Dezimalbrüche) häufiger als gemeine Brüche verwendet.

Zusammenfassung

Mit der Schöpfung der Dezimalbrüche erlangte das positionelle Dezimalsystem seine Vollendung, dieses System, das uns so unentbehrlich zur vollständigen Bezeichnung der Zahlen wurde und das heute allgemein in Wissenschaft und Alltagsleben herrscht. Daneben erhielten sich nur ganz unbedeutende Überreste anderer Bezeichnungssysteme der Zahlen, teils im Sprachgebrauch, teils in gewissen Rechnungen (die sexagesimalen Unterteilungen der Stunde und des Grades, die Verwendung der Stammbrüche $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}$ usw.), manchmal auch bei der Bezeichnung von Ordinalzahlen (mit Hilfe der Buchstaben eines Alphabets).

Mit der Einführung der Dezimalbrüche dehnte sich unsere heutige Bezeichnungsweise der Zahlen, die anfänglich nur zur Darstellung der ganzen Zahlen geschaffen war, auf alle reellen Zahlen[2]) aus. Dazu brauchten keine neuen Prinzipien mehr hinzugenommen zu werden; die Bezeichnung aller Zahlen ist völlig gleichartig.

Das positionelle Dezimalsystem erfüllt, wie leicht einzusehen ist, alle Forderungen, die wir an ein bequem zu verwendendes Bezeichnungssystem für die Zahlen stellen müssen (siehe S. 23). Es ist gleichermaßen gut geeignet zur Darstellung sehr großer wie sehr kleiner Zahlen, die die Menschheit seit dem Beginn der Renaissance — besonders aber in unserem Jahrhundert — bei der Erforschung der entferntesten Himmelskörper einerseits und der inneratomaren Mikrowelt andererseits benötigt. Diese Forderungen erfüllt in einem solchen Maße keines der anderen Bezeichnungssysteme, die wir kennengelernt haben. Deshalb mußten im Laufe der fortschreitenden Entwicklung alle diese Systeme dem positionellen Dezimalsystem weichen, entgegen jeglicher vielhundertjähriger Tradition und sonstiger Hindernisse.

---

[1]) Ein einheitliches Maßsystem wurde zuerst in Frankreich nach der bürgerlichen Revolution von 1789 gesetzlich festgelegt. In der UdSSR wurde das metrische Maßsystem durch einen Beschluß des Rates der Volkskommissare vom 14. Januar 1918 eingeführt. Bis 1926/27 hatte es die alten Maßsysteme vollständig verdrängt. England, die USA und einige andere Länder haben bis zum heutigen Tage noch nicht offiziell das metrische System angenommen. (Bei uns in Deutschland wurde das metrische System im Jahre 1872 gesetzmäßig verankert. *Anm. d. wissenschaftl. Red.*)

[2]) Über die reellen Zahlen siehe Kap. VI des nachfolgenden Artikels von I. W. PROSKURJAKOW

Wir sahen, daß das Positionssystem das Ergebnis einer langen historischen Entwicklung ist, daß an seiner Schöpfung, die sich über mehrere Jahrtausende hinzog, zahlreiche Völker aus aller Welt teilgenommen haben. Die Entstehungsgeschichte des Positionssystems bietet nicht nur im Rahmen unserer Einzeldisziplin — der Mathematik — Interesse, sondern besitzt auch allgemeinere Bedeutung.

1.) Vor allen Dingen widerlegt die wissenschaftliche und objektive Geschichte der Zahlenbezeichnung die idealistische Lehre vom apriorischen Charakter der Zahlen und die Auffassung, sie seien dem menschlichen Bewußtsein angeboren. Die Geschichte der Zahlenbezeichnung zeigt, daß der Begriff der natürlichen Zahl (wie auch der des Bruches) und die Eigenschaften und Gesetze der Zahlen das Resultat von Abstraktionen bestimmter, ganz konkreter, quantitativer Eigenschaften und Verhältnisse der Gegenstände der realen Welt sind.

2.) Weiter zeigt die Geschichte unserer Zahlzeichen, daß die Entwicklung der Bezeichnungssysteme stets vom Vielfältigen und Reichhaltigen zum Einfachen und Einheitlichen verlief. Je niedriger das wirtschaftliche und kulturelle Niveau der Gesellschaft war, desto zahlreicher waren die in ihr gebräuchlichen Bezeichnungssysteme.

Die erste allgemeine Form von Bezeichnungssystemen waren Systeme vom Typ der ägyptischen Hieroglyphen, die auf einem additiven (mitunter auch einem subtraktiven) Prinzip beruhten. Die Phase der hieroglyphischen Bezeichnungsweise der Zahlen, der ein noch sehr primitives mündliches Zählen entspricht, wurde anscheinend in mehr oder minder ausgeprägter Form von allen Völkern durchlaufen. Obwohl das Prinzip dieser Bezeichnungsweise in den verschiedenen Ländern ein und dasselbe war, zeigten sich bei der Wahl der Knotenzahlen, von denen ja jede als Grundlage eines eigenen, besonderen Stellensystems diente, und in den Formen ihrer Bezeichnungen größte Unterschiede.

Die Bezeichnungssysteme vom Typ der Hieroglyphen wurden von den alphabetischen verdrängt, die ihrerseits einen wesentlichen Schritt auf dem Wege zum heutigen universellen Bezeichnungssystem bedeuteten. Sie waren gut geeignet zum Operieren mit nicht allzu großen Zahlen, im Einklang mit den ökonomischen Verhältnissen der antiken Sklavenhaltergesellschaft und des frühen Feudalismus.

Alle alphabetischen Systeme waren Dezimal-Systeme und bauten sich nach einem gemeinsamen Ziffernprinzip auf. Diese beiden Tatbestände waren ein wesentlicher Schritt auf dem Wege zu einer einheitlichen Bezeichnungsweise. Daß aber hierbei jedes Volk sein eigenes Alphabet verwendete, hemmte die Entwicklung einer einheitlichen Bezeichnungsweise.

Schließlich brachte das folgende Entwicklungsstadium das erste einheitliche, in der ganzen Welt verwendete Bezeichnungssystem, unser heutiges dezimales Positionssystem. Einheitlich ist hierbei nicht nur das Prinzip seines Aufbaues, sondern auch die Form der Ziffern.

3.) Damit liefert die Geschichte der Systeme zur Bezeichnung der Zahlen eine ergänzende Illustration zu der marxistischen These von der Entwicklung im Kampf der Gegensätze, im Kampf des Neuen und Progressiven mit dem Alten und Konservativen, von der unüberwindlichen und alles bezwingenden Kraft der fortschrittlichen Elemente der menschlichen Gesellschaft und der menschlichen Kultur. Jedes Bezeichnungssystem fixiert nämlich ein schon vorhandenes mündliches Zählen, das sich weiter entwickelt und mit dem Fortschritt der Gesellschaft immer mehr vervollkommnet. Hierbei bleibt das schriftliche Bezeichnungssystem häufig hinter den

faktisch existierenden Zählmethoden zurück. Unvermeidlich erscheint dann ein neues, vollendeteres Bezeichnungssystem. Wir sahen, mit welcher Mühe sich jedes neue System stets seinen Weg bahnen mußte. So mußte das alphabetische System die jahrhundertealte Tradition der attischen Staaten überwinden. Dem dezimalen Positionssystem standen die reaktionären Kräfte des europäischen Mittelalters entgegen.

4.) Schließlich zeigt die Geschichte der Entstehung des Positionssystems von neuem die Einheitlichkeit der gesellschaftlichen Entwicklung. Als Zugang zum Positionsprinzip dienten — wie wir sahen — die multiplikativen Rechensysteme, die schließlich zum Positionssystem und zur Einführung der Null im alten Zweistromland, bei den Stämmen der Maya, in der spätantiken Epoche und schließlich in Indien führten. Die Bildung des heutigen Positionssystems war also somit keine zufällige Erscheinung, sondern die gesetzmäßige Vollendung eines unumgänglichen historischen Prozesses.

Die Geschichte der Bezeichnungssysteme für die Zahlen zeigt, daß auch unser Bezeichnungssystem geschaffen wurde im Verlaufe der Geschichte der Gesellschaft und der Geschichte der Basen während vieler Jahrhunderte, durch die Anstrengungen Hunderter von Generationen, daß es geschaffen wurde zum Nutzen der ganzen Gesellschaft, aller ihrer Klassen, als ein System, das einheitlich für die ganze Gesellschaft und allen ihren Gliedern gemeinsam ist.[1])

[1]) Weitere Literatur über die Entstehung der Bezeichnungssysteme findet man z. B. in D. J. Struik, Abriß der Geschichte der Mathematik, 2. Aufl., Berlin 1963 (Übersetzung aus dem Amerikanischen), F. A. Willers, Zahlzeichen und Rechnen im Wandel der Zeit, Berlin-Leipzig 1949, und H. Wussing, Mathematik in der Antike, 2. Aufl., Leipzig 1965. — *Anm. d. wissenschaftl. Red.*

I. W. PROSKURJAKOW

MENGEN, GRUPPEN, RINGE UND KÖRPER
DIE THEORETISCHEN GRUNDLAGEN DER ARITHMETIK

EINLEITUNG

Der Begriff der Zahl, der auf sehr frühen Entwicklungsstufen der menschlichen Gesellschaft aus dem Bedürfnis zu zählen heraus entstand, ist eine der grundlegenden Errungenschaften der menschlichen Kultur. Die Zahl ist zu einem dauernden und unentbehrlichen Hilfsmittel unserer gesamten praktischen Tätigkeit geworden. Die Möglichkeit, die Zahlen für das Studium und die Veränderung der uns umgebenden materiellen Welt zu benutzen, beruht darauf, daß der Mensch die Zahlen selbst aus dieser Welt gewonnen hat. Alle Eigenschaften der Zahlen sind nur ein von einer Reihe spezieller konkreter Merkmale befreiter, abstrakter Ausdruck realer Beziehungen der materiellen Welt. So ist die Zahl Fünf nur eine Widerspiegelung der realen Eigenschaft, die den fünf Fingern unserer Hand, einer Blüte mit fünf Blütenblättern und allen anderen Gesamtheiten aus fünf materiellen Gegenständen gemein ist, unabhängig von Form, Größe, Farbe und anderen konkreten Eigenschaften der Gegenstände.

ENGELS schrieb darüber: „... keineswegs aber befaßt sich in der reinen Mathematik der Verstand bloß mit seinen eigenen Schöpfungen und Imaginationen. Die Begriffe von Zahl und Figur sind nirgends anders hergenommen als aus der wirklichen Welt. Die zehn Finger, an denen die Menschen zählen, also die ersten arithmetischen Operationen vollziehen gelernt haben, sind alles andere, nur nicht freie Schöpfungen des Verstandes. Zum Zählen gehören nicht nur zählbare Gegenstände, sondern auch schon die Fähigkeit, bei Betrachtung dieser Gegenstände von allen übrigen Eigenschaften abzusehen, außer von ihrer Zahl — und diese Fähigkeit ist das Ergebnis einer langen geschichtlichen, erfahrungsmäßigen Entwicklung."[1])

Das ist das Wesen der Mathematik vom Standpunkt des dialektischen Materialismus aus.

Entgegengesetzte Meinungen werden von bürgerlich-idealistischer Seite geäußert. Nach diesen Meinungen ist die Mathematik ein Produkt der freien Schöpfung des menschlichen Geistes, und ihre grundlegenden Begriffe sind unserem Verstande a priori eigen, d. h. vor jeder Erfahrung dem Menschen schon bei seiner Geburt gegeben.

Die Unsinnigkeit derartiger Ansichten über die Mathematik wird millionenfach und auf Schritt und Tritt durch unsere gesamte praktische Tätigkeit bewiesen, da in den Anwendungen die Mathematik stets die Ergebnisse zeitigt, die wir von ihr erwarten. Letzteres ist aber nur deshalb der Fall, weil die mathematischen Wahrheiten eine Widerspiegelung objektiver Gesetzmäßigkeiten der Natur sind.

Die Zahl ist das wichtigste Mittel, mit dessen Hilfe die Mathematik die Gesetzmäßigkeiten der realen Welt studiert. Jedoch ist der moderne Zahlbegriff das Resultat eines komplizierten und langwierigen historischen Entwicklungsprozesses.

---

[1]) FRIEDRICH ENGELS, Anti-Dühring, Berlin 1948, S. 44.

Nach den natürlichen Zahlen entstanden die gebrochenen Zahlen, später die irrationalen Zahlen und schließlich die negativen und die komplexen Zahlen. Der vorliegende Artikel wird nur ganz kursorisch die historische Entwicklung des Zahlbegriffs andeuten, da seine eigentliche Aufgabe darin besteht, das logische Wesen des modernen Zahlbegriffs aufzudecken. Der Leser findet hier auch keine Zusammenstellung von Eigenschaften der Zahlen, die für ihn neu sind. Nicht die Vermittlung neuer Erkenntnisse, sondern die Begründung derjenigen Eigenschaften der Zahlen, die jedem von der Schulbank her bekannt sind, ist das Hauptziel dieses Artikels.

Gerade der Beweis der einfachsten Eigenschaften der Zahlen, wie z. B. des kommutativen oder des assoziativen Gesetzes der Addition, verlangt eine genaue Definition der Zahl und stößt daher auf bemerkenswerte Schwierigkeiten. Nichtsdestoweniger glauben wir, daß der Lehrer, der täglich wiederholt von diesen Eigenschaften der Zahlen Gebrauch macht, eine Vorstellung davon besitzen muß, wie sie bewiesen werden. Dies ist auch für fähige, mathematisch interessierte Schüler der oberen Klassen für das allgemeine mathematische Wissen nützlich. Aus demselben Grunde kann der Artikel auch den Studenten der pädagogischen Institute empfohlen werden. Da der Aufbau der reellen Zahlen an die Vorlesungen über Differential- und Integralrechnung und der Aufbau der komplexen Zahlen an die Vorlesungen über höhere Algebra heranführt, können die entsprechenden Kapitel dieses Artikels auch den Studenten der physikalisch-mathematischen Fakultäten der Universitäten empfohlen werden.

Außer der Begründung der elementaren Eigenschaften der Zahlen hat der Artikel zum Ziel, den Leser an die grundlegenden Ideen und die Begriffe der modernen Mathematik heranzuführen. Hierzu gehören in erster Linie der Begriff der Isomorphie sowie die Begriffe der Menge, der Gruppe, des Ringes und des Körpers. Die konsequente Anwendung der angegebenen allgemeinen Begriffe erlaubt es, vielmalige und anstrengende Wiederholungen ein und derselben Überlegung beim Beweis analoger Eigenschaften von Zahlen verschiedener Art zu vermeiden, und gestattet dem Leser, die Eigenschaften verschiedener Zahlenbereiche vom allgemeinen Standpunkt aus zu erfassen. Selbstverständlich bereitet einem Leser, der mit diesen Begriffen nicht vertraut ist, eine solche Darlegung zusätzliche Schwierigkeiten, da er diese neue Auffassung von der Zahl wirklich erlernen muß. Die Kenntnis dieser Ideen und Begriffe der modernen Mathematik stellt den wesentlichen Teil jenes Neuen dar, das der Leser im vorliegenden Artikel kennenlernt. Die Ausführungen über die Begründung des Zahlbegriffs vom Standpunkt der Theorie der Ringe und Körper aus dürften, so scheint es uns, auch für den Fachmann von Interesse sein.

Im ersten Kapitel werden die wichtigsten Grundbegriffe der Mengenlehre zusammengestellt. Das zweite Kapitel behandelt den Begriff der Gruppe, den Begriff des Ringes und den des Körpers, wobei in allgemeiner Form die Eigenschaften algebraischer Operationen untersucht werden, die später beim Studium der verschiedenen Zahlenbereiche benutzt werden. In den folgenden Kapiteln werden nacheinander die natürlichen, die ganzen, die rationalen, die reellen und die komplexen Zahlen eingeführt. Im letzten Paragraphen werden schließlich noch die Quaternionen betrachtet und die Frage nach der Möglichkeit einer nochmaligen Erweiterung des Zahlenbereiches erörtert.

Da wir eine streng logische Begründung der Eigenschaften der Zahlen beabsichtigen, weisen wir bei der Benutzung schon bewiesener Eigenschaften in Klammern auf die entsprechenden Sätze aus den vorhergehenden Kapiteln hin. Daher muß

ein Leser, der die Richtigkeit der Begründung einer gegebenen Eigenschaft nachprüfen will, entweder den ganzen vorangegangenen Teil des Artikels lesen oder zumindest aber jene Teile, die in diesen Hinweisen angegeben sind. Jedoch kann ein Leser, der sich speziell für die Begründung der Eigenschaften von Zahlen einer gegebenen Art interessiert und die Eigenschaften der vorhergehenden Zahlenbereiche als bekannt ansieht, nach den ersten beiden Kapiteln und § 19 aus Kapitel IV, in welchem Begriffe eingeführt werden, die für das Verständnis alles weiteren notwendig sind, sofort das ihn interessierende Kapitel lesen. Bei einer solchen Lektüre braucht man seine Aufmerksamkeit nicht auf die Hinweise in Klammern zu richten, da die Eigenschaften, von denen die Rede ist, jedem von der Schule her bekannt sind. So kann man, wenn man etwa die Eigenschaften der rationalen Zahlen als bekannt annimmt, nach den ersten beiden Kapiteln und § 19 sofort das Kapitel VI über die reellen Zahlen lesen; wenn man die Eigenschaften der reellen Zahlen als bekannt annimmt, kann man das Kapitel VII über die komplexen Zahlen lesen.

Kapitel I

MENGEN

§ 1. Der Begriff der Menge

In den verschiedenen Gebieten der Mathematik werden häufig die Objekte nicht als einzelne Individuen, sondern in ihrer Gesamtheit untersucht. Manchmal faßt man Objekte, die eine gewisse Eigenschaft gemein haben, zu einer Gesamtheit zusammen und studiert sie in dieser Gesamtheit.

So ist die Gesamtheit aller natürlichen Zahlen in der umfassenderen Gesamtheit der ganzen Zahlen enthalten. Wenn man diesen Zahlenbereich erweitert, kommt man zur Gesamtheit der rationalen, zur Gesamtheit der reellen und schließlich zur Gesamtheit der komplexen Zahlen. In der Algebra betrachtet man z. B. die Gesamtheit der Polynome und die Gesamtheit der rationalen Funktionen. Wenn man in der Geometrie die Eigenschaften eines Dreiecks studiert, so abstrahiert man häufig von seiner Lage in der Ebene oder auch von seiner Größe und erhält so Sätze, die für alle untereinander kongruenten bzw. alle untereinander ähnlichen Dreiecke gültig sind. In anderen Fällen betrachtet man Gesamtheiten von Punkten, die irgendeine gemeinsame Eigenschaft besitzen (geometrische Örter) usw.

Wir beschränken uns hier darauf, die Anfangsgründe der Mengenlehre darzustellen, und verweisen den interessierten Leser, der sie näher kennenlernen will, auf die Bücher von P. S. ALEXANDROFF [1] und N. N. LUSIN [2].

Eine Menge ist eine Gesamtheit von Objekten, die als ein Ganzes betrachtet wird. Diese Worte darf man nicht als Definition des Begriffs Menge auffassen, denn weshalb soll das Wort „Gesamtheit" besser sein als das Wort „Menge"? Der Mengenbegriff wird vielmehr als ein Grundbegriff angesehen, der nicht auf andere Begriffe zurückgeführt wird. Die Objekte, aus denen eine gegebene Menge besteht, heißen ihre *Elemente*. Die grundlegende Beziehung zwischen einem Element a und einer dieses Element enthaltenden Menge A wird durch $a \in A$ (in Worten: „a ist Element

der Menge A" oder „a gehört zu A" oder „A enthält a [als Element]") zum Ausdruck gebracht. Wenn a nicht Element von A ist, so schreibt man $a \bar{\in} A$ (in Worten: „a ist nicht in A als Element enthalten" usw.). Eine Menge kann durch Aufzählung aller ihrer Elemente gegeben sein. Zur Bezeichnung derartiger Mengen verwenden wir geschweifte Klammern. So bezeichnet $\{a, b, c\}$ die Menge aus den Elementen a, b, c. Eine entsprechende Schreibweise wird auch im Falle unendlicher Mengen angewendet, wobei die nicht aufgeschriebenen Elemente durch Punkte ersetzt werden. So wird die Menge der natürlichen Zahlen durch $\{1, 2, 3, \ldots\}$, die Menge der geraden Zahlen durch $\{2, 4, 6, \ldots\}$ angedeutet, wobei die Punkte im ersten Falle die fehlenden natürlichen Zahlen und im zweiten Falle nur die fehlenden geraden Zahlen bezeichnen.

Mengen A und B heißen *gleich*, wenn sie aus denselben Elementen bestehen, wenn also jedes Element von A auch der Menge B angehört und umgekehrt auch jedes Element von B zu A gehört. In diesem Fall schreibt man $A = B$. Eine Menge ist also eindeutig durch ihre Elemente bestimmt und hängt nicht von deren Reihenfolge ab. Z. B. kann also die Menge aus den Elementen a, b, c auf folgende sechs Arten geschrieben werden:

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

Aus formalen Gründen führt man noch die sogenannte *leere Menge* ein, das ist eine „Menge", die kein einziges Element enthält. Wir werden sie durch das Zeichen „0" bezeichnen (Verwechslungen mit der Zahl Null sind nicht zu befürchten, da die Bedeutung des Zeichens in jedem Falle klar ist).

Ist jedes Element einer Menge A in der Menge B enthalten, so nennen wir A eine *Teilmenge* oder *Untermenge* von B und B eine *Obermenge* von A. Man schreibt dafür $A \subseteq B$ bzw. $B \supseteq A$ (in Worten: „A ist enthalten in B", „B enthält A"). Ist $A \subseteq B$ und $B \subseteq A$, so ist $A = B$. Die leere Menge ist per definitionem Teilmenge jeder beliebigen Menge.

Wenn jedes Element der Menge A in B enthalten ist, aber B wenigstens ein Element enthält, das nicht zu A gehört, wenn also $A \subseteq B$ und $A \neq B$ ist, so nennen wir A eine *echte Teilmenge* von B bzw. B eine *echte Obermenge* von A. In diesem Fall schreibt man $A \subset B$ bzw. $B \supset A$. Offensichtlich ist $A \neq 0$ gleichbedeutend mit $A \supset 0$; beides besagt nämlich, daß die Menge A nicht leer ist.

Wir bemerken noch, daß man das Element a sorgfältig von der Menge $\{a\}$ unterscheiden muß, die a als einziges Element enthält. Diese Unterscheidung wird nicht nur deshalb gefordert, weil Element und Menge verschiedene Rollen spielen (die Beziehung $a \in A$ ist nicht symmetrisch), sondern vor allem aus der Notwendigkeit heraus, Widersprüche zu vermeiden. Betrachten wir nämlich einerseits die Menge $A = \{a, b\}$ und andererseits die Menge $\{A\}$, die als einziges Element die Menge A enthält, so kämen wir — würden wir die Menge A nicht von der Menge $\{A\}$ unterscheiden — zu dem paradoxen Ergebnis, daß A einerseits die beiden Elemente a, b, andererseits genau das eine Element A enthielte. Deswegen wenden wir auch nicht die Schreibweise $a \subset A$ an, sondern behalten die Schreibweise $a \in A$ bei.

Beispiele von Mengen. Man kann beliebig viele Beispiele für Mengen anführen. So kann man von der Menge aller Buchstaben eines vorliegenden Buches sprechen, in der ein und derselbe Buchstabe auf verschiedenen Seiten des Buches oder in verschiedenen Zeilen einer Seite als zwei verschiedene Elemente angesehen werden, oder von der Menge aller Menschen auf der Erde, wobei man allerdings voraussetzen

muß, daß im betrachteten Moment weder ein Mensch stirbt noch ein Mensch geboren wird, oder von der Menge der Wassermoleküle in einem Glase usw.

Das waren alles Beispiele für endliche Mengen. Beispiele für unendliche Mengen sind, neben der bereits erwähnten Menge der natürlichen Zahlen und der Menge der geraden Zahlen, die Menge der rationalen Zahlen, die Menge der reellen Zahlen usw.

Es seien nun a und b zwei reelle Zahlen mit $a < b$. Die Menge aller reellen Zahlen x mit $a \leq x \leq b$ nennen wir das *abgeschlossene Intervall* mit den Enden a, b und bezeichnen sie durch $[a, b]$. Die Menge (a, b) aller x mit $a < x < b$ nennen wir das *offene Intervall* mit den Enden a, b. Weiter bezeichnen wir als *halboffene Intervalle* die Menge $[a, b)$ aller x mit $a \leq x < b$ und die Menge $(a, b]$ aller x mit $a < x \leq b$. Wir führen noch zwei Symbole ein: $+\infty$ (plus Unendlich) und $-\infty$ (minus Unendlich). Sie bezeichnen keine Zahlen, sondern dienen nur zur bequemeren Schreibweise. Nichtsdestoweniger vereinbaren wir, daß wir zu ihrer leichteren Behandlung davon sprechen, daß $+\infty$ größer als jede reelle Zahl und daß $-\infty$ kleiner als jede reelle Zahl ist. Diese Zeichen ermöglichen es uns, bequeme Bezeichnungen für unbeschränkte halboffene und offene Intervalle einzuführen, die ganz den obigen für beschränkte Intervalle entsprechen. Es bedeuten nämlich: $[a, +\infty)$ die Menge aller reellen Zahlen x mit $a \leq x$, $(-\infty, b]$ die Menge aller x mit $x \leq b$, $(a, +\infty)$ die Menge aller x mit $a < x$, $(-\infty, b)$ die Menge aller x mit $x < b$ und $(-\infty, +\infty)$ die Menge aller reellen Zahlen.

§ 2. Mengenalgebra

Unter der *Vereinigung* zweier Mengen A und B verstehen wir die Menge der Elemente, die zu wenigstens einer der Mengen A, B gehören (d. h. entweder zu A oder zu B oder zu beiden). Man bezeichnet sie durch $A \cup B$, gelesen: „Vereinigung aus (von) A und B". Der *Durchschnitt* der Mengen A und B ist die Menge der Elemente, die sowohl zu A als auch zu B gehören. Man bezeichnet ihn durch $A \cap B$, gelesen: „Durchschnitt von A und B". Die *Differenz* aus den Mengen A und B ist die Menge der Elemente, die zwar zu A, aber nicht zu B gehören. Man bezeichnet sie durch $A \setminus B$, gelesen: „Differenz aus A und B".[1])

Beispiel 1. Es sei A das Intervall $[1, 3]$ und B das Intervall $[2, 4]$; dann ist die Vereinigung $A \cup B$ das Intervall $[1, 4]$, der Durchschnitt $A \cap B$ das Intervall $[2, 3]$, die Differenz $B \setminus A$ das halboffene Intervall $(3, 4]$.

Beispiel 2. Es sei A die Menge aller Rechtecke, B die Menge aller Rhomben in einer gewissen Ebene. Dann ist $A \cap B$ die Menge aller Quadrate, $A \setminus B$ die Menge aller Rechtecke mit ungleichen Seiten und $B \setminus A$ die Menge aller Rhomben mit ungleichen Winkeln.

Die Bildung der Vereinigung bzw. des Durchschnitts von Mengen hat viele Eigenschaften mit der Addition und Multiplikation von Zahlen gemein; so gelten für sie das kommutative, das assoziative und beide distributiven Gesetze.

Die Definition der Vereinigung und des Durchschnitts überträgt sich wörtlich von zwei Mengen auf eine beliebige endliche oder unendliche Menge von Mengen.

[1]) Viele Autoren verwenden die Bezeichnungen $A + B$, AB, $A - B$, was aber in der Algebra wegen der möglichen Verwechslung mit den algebraischen Operationen unzweckmäßig ist.

Um zu einer besseren Ausdrucksweise zu kommen, nennen wir Mengen, deren Elemente ihrerseits Mengen sind, *Mengensysteme* oder kurz *Systeme*. Unter der *Vereinigung eines Mengensystems* (der Mengen dieses Systems) verstehen wir die Menge der Elemente, die wenigstens einer der Mengen des gegebenen Systems angehören. Entsprechend verstehen wir unter dem *Durchschnitt eines Mengensystems* die Menge der Elemente, die in jeder der Mengen des betreffenden Systems enthalten sind.

Hierfür haben sich folgende Bezeichnungen eingebürgert: Ist das vorgelegte System endlich — bestehend etwa aus den Mengen $A_1, A_2, \ldots, A_n$ —, so bezeichnet man die Vereinigung S und den Durchschnitt D durch:

$$S = A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i,$$

$$D = A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i.$$

Im Falle einer unendlichen Folge $A_1, A_2, \ldots, A_n, \ldots$, d. h. eines Systems, dessen Mengen mittels der natürlichen Zahlen durchnumeriert sind, schreibt man:

$$S = A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots = \bigcup_{i=1}^{\infty} A_i,$$

$$D = A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots = \bigcap_{i=1}^{\infty} A_i.$$

Beispiel 3. Es sei A_n die Menge der Punkte einer Ebene, die innerhalb eines Kreises vom Radius 2^n um den Nullpunkt der Ebene liegen, wobei n alle ganzen Zahlen zwischen $-\infty$ und $+\infty$ annehmen kann. Dann ist die Vereinigung $\bigcup_{n=1}^{+\infty} A_n$ die Menge aller Punkte der Ebene, während der Durchschnitt $\bigcap_{n=1}^{+\infty} A_{-n}$ genau einen Punkt, nämlich den Nullpunkt der Ebene enthält.

Schließlich schreibt man im Fall eines beliebigen Systems $\{A_m\}_{m \in M}$ von Mengen A_m, bei dem die Indizes m einer gewissen Menge M angehören:

$$S = \bigcup_{m \in M} A_m, \qquad D = \bigcap_{m \in M} A_m.$$

Beispiel 4. Es sei X die Menge aller positiven reellen Zahlen x und A_x die Menge aller Punkte des Kreises mit dem Radius x um den Nullpunkt der Ebene. Dann ist wieder die Vereinigung $\bigcup_{x \in X} A_x$ die Menge aller Punkte der Ebene, während der Durchschnitt $\bigcap_{x \in X} A_x$ wieder aus nur einem Element, nämlich dem Nullpunkt der Ebene besteht.

§ 3. Funktionen, Abbildungen, Mächtigkeiten

Eine ebenso bedeutende Rolle wie der Mengenbegriff spielt in der Mathematik der Funktionsbegriff. Was ist nun eine Funktion? Man hört oft, eine Funktion sei „eine veränderliche Größe, die von einer anderen veränderlichen Größe (Argument) abhängt". In den Anwendungen auf konkrete Funktionen — z. B. $y = \sin x$ —, wie sie in der Schule behandelt werden, genügt diese „Definition" vollständig und kann im Unterricht deshalb verwendet werden. Unsere Aufgabe besteht jedoch in einer genaueren Erläuterung dieses Begriffes und in der Darlegung seiner modernen Definition. Wenn man zunächst einmal die Funktion

$$y = \sin^2 x + \cos^2 x$$

betrachtet, so stellt man fest, daß ihr Wert schon nicht mehr von dem Wert von x abhängt. Ferner ist es üblich, unter „Größen" solche Objekte zu verstehen, die man untereinander (der Größe nach) vergleichen kann, zwischen denen also eine Beziehung „größer" bzw. „kleiner" besteht. Indessen betrachtet man in der Mathematik auch Funktionen, für deren Argumente und Werte solche Beziehungen nicht erklärt sind, wie z. B. Funktionen, die auf der Menge der komplexen Zahlen oder auf einer ganz beliebigen Menge definiert sind. Eine genauere Analyse des Funktionsbegriffs zeigt, daß an der Funktion keineswegs die Veränderung des Wertes mit der Veränderung des Argumentes wesentlich ist, sondern die Tatsache der Zuordnung selber, auf Grund derer jedem Wert des Argumentes eindeutig ein ihm entsprechender Funktionswert zugeordnet ist. So kann man die Funktion

$$y = \sin^2 x + \cos^2 x$$

einfach dadurch definieren, daß man verlangt, daß jeder reellen Zahl x die Zahl 1 zugeordnet wird. Eine Zuordnung ist hierbei eine Vorschrift, nach welcher jedem Element x einer gewissen Menge X eindeutig ein gewisses Objekt entspricht (das dem gegebenen Element „zugeordnet" ist). Diese Worte sollen nur den Begriff der Zuordnung erläutern, sind aber nicht als deren Definition anzusehen. Vielmehr nehmen wir hier auch den Begriff der Zuordnung, wie schon den Begriff der Menge, als einen Grundbegriff an, der keiner Definition unterliegt. Dann lautet die vollständige Definition der Funktion folgendermaßen:

Definition 1. *Eine auf einer gewissen Menge X erklärte (oder definierte) Funktion ist eine Vorschrift, die jedem Element x der Menge X ein gewisses Objekt $f(x)$ zuordnet.*

Die Menge X heißt der *Definitionsbereich der Funktion*, und die Menge Y der Objekte, die den Elementen der Menge X zugeordnet sind, heißt der *Wertevorrat der Funktion*.

Beispiel 1. Es sei $y = \sin x$. Definitionsbereich dieser Funktion ist die Menge aller reellen Zahlen, Wertevorrat das abgeschlossene Intervall $[-1, +1]$.

Beispiel 2. Es sei $y = \operatorname{tg} x$. Definitionsbereich dieser Funktion ist die Menge aller reellen Zahlen, die verschieden sind von den Zahlen der Form $n\pi + \frac{\pi}{2}$, wobei n beliebig ganz ist (für diese Zahlen ist die Funktion bekanntlich nicht erklärt), Wertevorrat ist die Menge aller reellen Zahlen.

Beispiel 3. Die DIRICHLETsche Funktion:

$$f(x) = \begin{cases} 0, & \text{falls } x \text{ rational ist,} \\ 1, & \text{falls } x \text{ irrational ist.} \end{cases}$$

Definitionsbereich ist hier die Menge aller reellen Zahlen, Wertevorrat ist die zweielementige Menge $\{0, 1\}$.

Es ist bemerkenswert, daß der geniale russische Mathematiker N. I. LOBATSCHEWSKI schon vor mehr als 100 Jahren eine ähnliche Funktion angegeben hat. Gegenüber der damaligen Auffassung der Funktion als eines analytischen Ausdrucks (d. h. einer geschlossenen Formel) hob er die Bedeutung der Zuordnung bei der Definition des Funktionsbegriffes hervor.

„Dieser allgemeine Begriff" — schrieb LOBATSCHEWSKI über den Funktionsbegriff — „verlangt, daß eine Zahl, die für jeden Wert von x gegeben ist und die sich zusammen mit x ändert, eine Funktion von x genannt wird. Der Wert der Funktion kann durch einen analytischen Ausdruck oder durch irgendeine andere Bedingung gegeben sein, nach der aus der Gesamtheit der Zahlen eine ausgewählt werden kann; schließlich kann sie in Form irgendeiner Abhängigkeit bestehen, die an sich unbekannt bleibt."[1])

Dem Funktionsbegriff nahe verwandt ist der Begriff der Abbildung.

Definition 2. *Gegeben seien Mengen X und Y. Eine Vorschrift, durch welche jedem Element $x \in X$ (eindeutig) ein Element $y \in Y$ zugeordnet wird, heißt eine Abbildung der Menge X in die Menge Y; ist speziell jedes Element $y \in Y$ wenigstens einem $x \in X$ zugeordnet, so heißt diese Vorschrift eine Abbildung von X auf Y.*[2])

Das einem Element x bei einer Abbildung zugeordnete Element y heißt *das Bild des Elementes x*, und x heißt *ein Urbild des Elementes y*. Man schreibt dafür $x \to y$ oder $y = f(x)$. Die Menge A aller Elemente $x \in X$, die ein und dasselbe Bild $y \in Y$ besitzen, heißt *das volle Urbild des Elementes y*.

Beispiel 4. Es sei D die Menge der reellen Zahlen. Die Vorschrift $x \to |x|$ ist eine Abbildung von D in sich und eine Abbildung von D auf die Menge der nichtnegativen reellen Zahlen. Urbild der Zahl 0 ist allein die Zahl 0, während jede reelle Zahl $y > 0$ zwei Urbilder, nämlich $+y$ und $-y$ besitzt.

Beispiel 5. Es sei jedem Punkt eines Quadrates in der Ebene sein Projektionspunkt auf die Grundlinie zugeordnet. Wir erhalten dadurch eine Abbildung des Quadrates auf eine Strecke. Das volle Urbild jedes Punktes der Grundlinie besteht aus der Gesamtheit der Punkte des Quadrates, die auf der in jenem Punkte errichteten Senkrechten zur Grundlinie liegen.

Die Beispiele 4 und 5 zeigen, daß bei einer Abbildung der Menge X in die Menge Y einerseits gewisse Elemente aus Y kein Urbild zu besitzen brauchen, andererseits aber auch gewisse Elemente mehrere (unter Umständen sogar unendliche viele) Urbilder haben können. Trifft weder das eine noch das andere zu, so nennen wir die Abbildung *eineindeutig* (oder auch *umkehrbar eindeutig*). Wir erhalten also folgende Definition:

Definition 3. *Eine eineindeutige Beziehung zwischen den Mengen X und Y (oder eine eineindeutige Abbildung von X auf Y) ist eine Vorschrift (oder Abbildung), die die folgenden drei Eigenschaften besitzt*: 1.) *Jedem Element der Menge X ist ein*

[1]) Н. И. Лобачевский, Об исчезании тригонометрических строк (N. I. LOBATSCHEWSKI, Über die Konvergenz trigonometrischer Reihen), Wissenschaftliche Schriften der Universität Kasan, II, 1834.

[2]) Man spricht heute häufig schon dann von einer Abbildung, wenn jedem Element der Menge X **gewisse** Elemente der Menge Y zugeordnet sind, und nennt Abbildungen der oben angegebenen Art *eindeutige Abbildungen*. *Anm. d. wissenschaftl. Red.*

und nur ein Element der Menge Y zugeordnet; 2.) *Je zwei voneinander verschiedenen Elementen der Menge X sind stets zwei voneinander verschiedene Elemente der Menge Y zugeordnet*; 3.) *Jedes Element der Menge Y ist einem Element der Menge X zugeordnet.*

Wir bemerken, daß die ersten beiden Bedingungen eine eineindeutige Abbildung von X auf eine gewisse Teilmenge von Y charakterisieren. In diesem Falle spricht man auch von einer *eineindeutigen Abbildung von X in Y.*

Ist $y = f(x)$ eine eineindeutige Abbildung von X auf Y, so kann man jedem $y \in Y$ das eindeutig bestimmte Element $x \in X$ zuordnen, dessen Bild y bei der Abbildung f ist. Diese Zuordnung heißt die *zu f entgegengesetzte* oder *inverse Abbildung* und wird mit f^{-1} bezeichnet. Als Übungsaufgabe beweise man, daß f^{-1} eine eineindeutige Abbildung von Y auf X ist und daß die zu f^{-1} inverse Abbildung die ursprüngliche Abbildung f ist.

Definition 4. *Mengen X und Y heißen gleichmächtig (oder äquivalent), wenn man zwischen ihren Elementen eine eineindeutige Zuordnung herstellen kann. Wir drücken die Gleichmächtigkeit der Mengen X und Y durch $X \sim Y$ aus.*

Von gleichmächtigen Mengen sagt man auch, daß sie dieselbe Mächtigkeit besitzen. Die leere Menge ist sich selbst und nur sich selbst gleichmächtig.

Bemerkung. Wir definierten eben den Begriff der Gleichmächtigkeit, nicht aber den Begriff der Mächtigkeit. Wir können sagen, daß die Mächtigkeit einer Menge X diejenige Eigenschaft ist, die allen zu X gleichmächtigen Mengen und nur diesen gemein ist. Übrigens reicht der Begriff der Gleichmächtigkeit stets aus.

Die Beziehung der Gleichmächtigkeit besitzt die folgenden drei grundlegenden Eigenschaften:

1. Reflexivität: $X \sim X$;
2. Symmetrie: *Wenn $X \sim Y$ ist, so ist $Y \sim X$*;
3. Transitivität: *Wenn $X \sim Y$ und $Y \sim Z$ ist, so ist $X \sim Z$.*

Zum Beweis von 1. ordnen wir jedes Element $x \in X$ sich selbst zu (identische Abbildung), was bereits eine eineindeutige Abbildung von X auf sich ergibt. Der Beweis der beiden anderen Eigenschaften sei dem Leser überlassen.

Die Mächtigkeit einer Menge charakterisiert sozusagen die „Anzahl" ihrer Elemente. Hierbei kann unter Umständen eintreten, daß „ein Teil gleich dem Ganzen" ist, d. h., eine Menge kann die gleiche Mächtigkeit wie eine ihrer echten Teilmengen besitzen.

Beispiel 6. Die Funktion $y = 10\,x$ ergibt, wenn x die reellen Zahlen von 0 bis 1 durchläuft, eine eineindeutige Abbildung des Intervalls $[0, 1]$ auf das zehnmal so lange Intervall $[0, 10]$. Also ist im Sinne der Mächtigkeit die „Anzahl" der Punkte beider Intervalle gleich.

Beispiel 7. Je zwei Intervalle $[a, b]$ und $[c, d]$ und je zwei Intervalle (a, b) und (c, d) sind gleichmächtig.

Zum Beweis braucht man nur die Funktion

$$y = c + \frac{d-c}{b-a}(x-a)$$

zu betrachten. Zunächst entspricht bei dieser Abbildung jeder reellen Zahl x eine eindeutig bestimmte Zahl y, wobei — wie man leicht ausrechnet — die Zuordnungen $a \to c$ und $b \to d$ bestehen. Es sei nun

$$x_1 \to y_1, \quad x_2 \to y_2 \quad \text{und} \quad x_1 < x_2.$$

Da entsprechend der Definition des abgeschlossenen bzw. offenen Intervalls $a < b$ und $c < d$ gilt, ist $\frac{d-c}{b-a} > 0$. Mithin ist unter den angegebenen Voraussetzungen $y_1 < y_2$. Ist also $a \leqq x \leqq b$ (oder $a < x < b$), so ist auch $c \leqq y \leqq d$ (bzw. $c < y < d$). Also entsprechen den Zahlen des Intervalls $[a, b]$ die Zahlen des Intervalls $[c, d]$, und verschiedene Zahlen des Intervalls $[a, b]$ gehen in verschiedene Zahlen des Intervalls $[c, d]$ über (entsprechend für offene Intervalle). Schließlich besitzt die inverse Abbildung

$$x = a + \frac{b-a}{d-c}(y - c)$$

dieselben Eigenschaften, woraus folgt, daß es auch zu jedem y aus $[c, d]$ ein (und auch nur ein) Urbild x aus $[a, b]$ gibt. Damit ist bewiesen, daß $[a, b] \sim [c, d]$ (und entsprechend $(a, b) \sim (c, d)$) gilt.

Beispiel 8. Die Funktion $y = \operatorname{tg} x$ liefert die Äquivalenz des Intervalls

$$\left(-\frac{\pi}{2}, +\frac{\pi}{2}\right)$$

mit der Menge aller reellen Zahlen.

Beispiel 9. Indem wir die jeweils untereinanderstehenden Zahlen einander zuordnen:

$$\begin{array}{rrrrrrr} 1, & 2, & 3, & \ldots, & n & , & \ldots, \\ 2, & 4, & 6, & \ldots, & 2n & , & \ldots, \\ 1, & 3, & 5, & \ldots, & 2n-1, & & \ldots, \\ 10, & 100, & 1000, & \ldots, & 10^n & , & \ldots, \\ 2, & 3, & 5, & \ldots, & p_n & , & \ldots \end{array}$$ (p_n n-te Primzahl),

schließen wir, daß die Menge aller natürlichen Zahlen, die Menge aller geraden Zahlen, die Menge aller ungeraden Zahlen, die Menge aller Potenzen von 10 und die Menge aller Primzahlen dieselbe Mächtigkeit besitzen, obwohl die erste von ihnen eine echte Obermenge der übrigen ist.

Beispiel 10. Die Menge der natürlichen Zahlen ist gleichmächtig der Menge der rationalen Zahlen.

Dazu denken wir uns jede von 0 verschiedene rationale Zahl in der Form $\frac{p}{q}$ mit teilerfremden p und q und positivem q (d. h., wir schlagen das Vorzeichen zum Zähler) geschrieben. Von den möglichen Darstellungen der Zahl 0, nämlich $0 = \frac{0}{1} = \frac{0}{2} = \cdots$ wählen wir eine, etwa $\frac{0}{1}$ aus. Damit ist jede rationale Zahl eindeutig in der Form $\frac{p}{q}$ dargestellt (insbesondere erhält man für $q = 1$ die ganzen Zahlen). Als *Höhe* der Zahl $\frac{p}{q}$ bezeichnen wir die natürliche Zahl $|p| + q$, wobei $|p|$ den absoluten Betrag von p bezeichnet. Dann können wir die rationalen Zahlen in einer Folge anordnen, indem wir nämlich die Zahlen einmal nach wachsender Höhe und die Zahlen gleicher Höhe nach wachsendem Zähler aufschreiben. Auf diese Weise erhalten wir die Folge:

$$0, -1, +1, -2, -\frac{1}{2}, +\frac{1}{2}, +2, -3, -\frac{1}{3}, +\frac{1}{3}, +3,$$
$$-4, -\frac{3}{2}, -\frac{2}{3}, -\frac{1}{4}, +\frac{1}{4}, +\frac{2}{3}, +\frac{3}{2}, +4, \ldots$$

Da es zu vorgegebener Höhe n nur endlich viele Zahlen gibt (nämlich höchstens $2(n-1)$, da der Zähler schlimmstenfalls von $-(n-1)$ bis $+(n-1)$ ausschließlich 0 läuft), stehen nur endlich viele Zahlen einer gegebenen Höhe in dieser Folge. Daher lassen sich die rationalen Zahlen in der Tat auf die angegebene Weise mit Hilfe der natürlichen Zahlen durchnumerieren, womit die behauptete Gleichmächtigkeit bewiesen ist.

§ 4. Endliche und unendliche Mengen

Alle im vorangehenden Paragraphen erwähnten Mengen, die einer echten Teilmenge gleichmächtig sind, waren unendlich. Wir werden sogleich sehen, daß das kein Zufall ist (Satz 1). Zunächst müssen wir aber eine einwandfreie Definition der Begriffe „endliche Menge" und „unendliche Menge" geben. Dazu müssen wir allerdings eine wesentliche Eigenschaft der natürlichen Zahlen benutzen, deren systematische Begründung erst in Kapitel III gegeben wird. Der Leser muß sich davon überzeugen, daß sich in unseren Überlegungen kein Zirkelschluß befindet. Dazu genügt es nachzuprüfen, daß bei der Begründung der Eigenschaften der natürlichen Zahlen in Kapitel III keine Ergebnisse aus den ersten beiden Kapiteln benutzt werden, die hier mit Hilfe der natürlichen Zahlen erhalten wurden.

Definition 1. *Unter einem Abschnitt der Folge der natürlichen Zahlen verstehen wir die Menge der natürlichen Zahlen, die kleiner oder gleich einer gewissen natürlichen Zahl n sind. Speziell bezeichnen wir den durch die Zahl n bestimmten Abschnitt durch* $|1, n|$.

Definition 2. *Eine Menge heißt endlich, wenn sie leer ist oder wenn sie einem Abschnitt der Folge der natürlichen Zahlen gleichmächtig ist. Eine Menge heißt unendlich, wenn sie nicht endlich ist.*

Mit anderen Worten: Eine endliche Menge (die nicht leer ist) ist eine Menge, deren Elemente man aufzählen kann, die man also durchnumerieren kann: $a_1, a_2 \ldots, a_n$, wobei alle Elemente mit Hilfe der Zahlen von 1 bis n numeriert sind, bei der Numerierung alle Zahlen von 1 bis n benutzt werden und verschiedene Elemente verschiedene Nummern haben. Eine unendliche Menge dagegen ist eine Menge, deren Elemente man nicht so aufzählen kann.

Auf Grund der Symmetrie und der Transitivität der Gleichmächtigkeit folgt unmittelbar, daß jede einer endlichen (bzw. unendlichen) Menge gleichmächtige Menge ebenfalls endlich (bzw. unendlich) ist.

Satz 1. (Hauptsatz über endliche Mengen) *Eine endliche Menge ist keiner ihrer echten Teilmengen und keiner ihrer echten Obermengen gleichmächtig.*

Beweis: Jede der beiden Behauptungen dieses Satzes folgt unmittelbar aus der anderen. Ist nämlich $A \sim B$ und $A \supset B$, so folgt nach dem Gesagten aus der Endlichkeit einer dieser Mengen die Endlichkeit der anderen. Es genügt also zu zeigen, daß keine endliche Menge einer ihrer echten Teilmengen gleichmächtig ist. Dieser Satz gilt für die leere Menge $A = 0$, da diese überhaupt keine echte Teilmenge besitzt. Es sei also $A \neq 0$. Dann ist nach Definition der endlichen Menge die Menge A (wenigstens) einem Abschnitt $|1, n|$ der Folge der natürlichen Zahlen gleichmächtig.

Wir beweisen durch vollständige Induktion über n[1]), daß sich die Menge A auf keine ihrer echten Teilmengen B eineindeutig abbilden läßt. Für $n = 1$ ist dies klar: Weil nämlich wegen $A \sim |1, 1|$ die Menge A dann nur ein einziges Element enthält, ist $B = 0$ die einzige echte Teilmenge von A, und diese ist der Menge A nicht gleichmächtig.

Wir nehmen nun an, daß der zu beweisende Satz bereits für die Zahl n gilt, und zeigen, daß er dann auch für $n+1$ Gültigkeit hat. Es sei also $A \sim |1, n+1|$ und f eine eineindeutige Abbildung von A auf eine echte Teilmenge B von A. Wir können dann die Menge A mit Hilfe der Zahlen von 1 bis $n+1$ durchnumerieren und erhalten:

$$A = \{a_1, a_2, \ldots, a_{n+1}\}.$$

Der Fall $B = 0$ führt unmittelbar auf einen Widerspruch. Es muß also $B \neq 0$ sein. Ohne Beschränkung der Allgemeinheit können wir annehmen, daß $a_{n+1} \in B$ ist. Andernfalls wählen wir nämlich ein Element $b \in B$ aus und konstruieren uns eine neue Menge B_1, die dadurch aus B entsteht, daß wir b durch a_{n+1} ersetzen; entsprechend konstruieren wir eine neue Abbildung f_1, die für alle Elemente aus A mit f übereinstimmt, ausgenommen das Element $a \in A$ mit $f(a) = b$, für das wir $f_1(a) = a_{n+1}$ setzen. Offenbar ist das so konstruierte f_1 eine eineindeutige Abbildung der Menge A auf die echte Teilmenge B_1, die a_{n+1} enthält. Ferner können wir annehmen, es sei $f(a_{n+1}) = a_{n+1}$. Ist nämlich $f(a_i) = a_{n+1}$ und $f(a_{n+1}) = a_j$, so konstruieren wir die neue Abbildung f_1, die für alle Elemente aus A, mit Ausnahme von a_i und a_{n+1}, mit f zusammenfällt, während wir $f_1(a_i) = a_j$ und $f_1(a_{n+1}) = a_{n+1}$ setzen. Es sei also $a_{n+1} \in B$ und $f(a_{n+1}) = a_{n+1}$, ferner sei $A' = A \setminus \{a_{n+1}\}$ und $B' = B \setminus \{a_{n+1}\}$. Da B eine echte Teilmenge von A ist, gibt es ein $a' \in A \setminus B$. Wegen $a_{n+1} \in B$ ist $a' \neq a_{n+1}$. Mithin ist $a' \in A' \setminus B'$. Also ist auch B' eine echte Teilmenge von A'. Da nun $f(a_{n+1}) = a_{n+1}$ ist, vermittelt f eine eineindeutige Abbildung von A' auf B'. Dann ergibt aber $A' = \{a_1, a_2, \ldots, a_n\} \sim |1, n|$ einen Widerspruch zur Induktionsannahme. Also kann es keine Abbildung f mit den angegebenen Eigenschaften geben, womit unsere Behauptung und damit der ganze Satz bewiesen ist.

Aus Satz 1 folgt unmittelbar

Satz 2. *Jede nichtleere endliche Menge ist einem und nur einem Abschnitt der Folge der natürlichen Zahlen gleichmächtig.*

Beweis: Nach Definition 2 ist jede nichtleere endliche Menge A wenigstens einem Abschnitt der Folge der natürlichen Zahlen gleichmächtig. Wäre sie nun zwei verschiedenen Abschnitten gleichmächtig, wäre also $A \sim |1, m|$, $A \sim |1, n|$, $m \neq n$, so wäre auf Grund der Eigenschaften der Gleichmächtigkeit $|1, m| \sim |1, n|$, im Widerspruch zu Satz 1, da von zwei verschiedenen Abschnitten der Folge der natürlichen Zahlen stets einer von beiden eine echte Teilmenge des anderen ist.

Definition 3. *Die für jede nichtleere endliche Menge A eindeutig bestimmte natürliche Zaul n mit $A \sim |1, n|$ nennen wir die Anzahl der Elemente der Menge A. Als Anzahl der Elemente der leeren Menge sieht man die Zahl 0 an.*

Auf Grund der Eigenschaften der Gleichmächtigkeit sind zwei endliche Mengen dann und nur dann gleichmächtig, wenn sie die gleiche Anzahl von Elementen be-

[1]) Wir bemerken, daß wir die Induktion nicht über „die Anzahl der Elemente von A" führen dürfen, da der Begriff „Anzahl der Elemente" erst unter Anwendung von Satz 1 eingeführt wird.

sitzen. Daher kann man die Mächtigkeit einer endlichen Menge auch mit Hilfe der Anzahl ihrer Elemente definieren.

Satz 3. *Jede Teilmenge einer endlichen Menge ist endlich, jede Obermenge einer unendlichen Menge ist unendlich.*

Beweis. Jede der beiden Behauptungen dieses Satzes folgt aus der anderen. Wenn z. B. die erste der Behauptungen richtig ist, so ist es auch die zweite: Ist A unendlich und $A \subseteq B$, so ist auch B unendlich, anderenfalls wäre nämlich auf Grund der ersten Behauptung des Satzes A endlich. Es genügt deshalb, die erste Behauptung zu beweisen. Es sei also A endlich und $B \subseteq A$. Ist $A = 0$, so ist auch $B = 0$, und die Behauptung gilt. Es sei also ferner $A \supset 0$. Dann gibt es eine natürliche Zahl n derart, daß $A \sim |1, n|$. Wir führen den weiteren Beweis durch vollständige Induktion über n. Für $n = 1$ gilt der Satz, da dann A genau ein Element enthält, also $B = 0$ oder $B = A$ ist. Die Behauptung sei für ein gewisses n richtig. Wir zeigen, daß sie dann auch für $n + 1$ gilt. Es sei also f eine eineindeutige Abbildung von A auf den Abschnitt $|1, n + 1|$. Ist $B = A$, so sind wir fertig. Wir können also zusätzlich voraussetzen, daß $B \subset A$. Dann gibt es ein Element $a \in A \setminus B$. Wir können annehmen, daß $f(a) = n + 1$ ist. Gilt nämlich $f(a') = n + 1$ mit $a' \in A$, $a' \neq a$, und ist $f(a) = i$, so gehen wir zu einer neuen Abbildung f_1 über, indem wir $f_1(a) = n + 1$, $f_1(a') = i$ und $f_1(x) = f(x)$ für alle übrigen Elemente x aus A setzen. Es sei also $f(a) = n + 1$. Wir betrachten dann die Menge $A' = A \setminus \{a\}$. Augenscheinlich vermittelt f eine eineindeutige Abbildung der Menge A' auf den Abschnitt $|1, n|$. Außerdem ist $B \subseteq A'$. Folglich ist nach Induktionsannahme die Menge B endlich. Damit ist Satz 3 vollständig bewiesen.

Auf Grund von Satz 3 hat es einen Sinn, von der Anzahl der Elemente einer beliebigen Teilmenge einer gegebenen endlichen Menge zu sprechen. Wir behaupten nun:

Satz 4. *Die Anzahl der Elemente einer endlichen Menge A ist stets größer als die Anzahl der Elemente jeder ihrer echten Teilmengen B.*

Beweis. Es sei m die Anzahl der Elemente von A und n die Anzahl der Elemente von B. Angenommen, es wäre $n \geq m$. Wegen $A \supset B$ ist $A \neq 0$, also $m > 0$ und $A \sim |1, m|$. Auf Grund von $n \geq m > 0$ gilt dann:

$$B \sim |1, n|. \tag{1}$$

Durch die eineindeutige Abbildung von A auf den Abschnitt $|1, m|$ wird die Menge B ebenfalls eineindeutig auf eine gewisse echte Teilmenge B' des Abschnittes $|1, m|$ abgebildet, so daß also

$$B \sim B' \tag{2}$$

gilt. Aus $B' \subset |1, m|$ und $m \leq n$ folgt

$$B' \subset |1, n|. \tag{3}$$

Aus (1) und (2) ergibt sich, daß $B' \sim |1, n|$ ist, was zusammen mit (3) einen Widerspruch zu Satz 1 liefert, da sich der Abschnitt $|1, n|$ einer echten Teilmenge B' gleichmächtig erweisen würde.

Bisher haben wir noch nicht die Existenz auch nur einer unendlichen Menge bewiesen. Jedoch folgt aus Satz 1 unmittelbar

Satz 5. *Die Menge N der natürlichen Zahlen und damit jede Menge, die eine zu N gleichmächtige Teilmenge enthält, ist unendlich.*

Beweis. Die Menge N ist unendlich, da die Abbildung $f(n) = n + 1$ eine eineindeutige Abbildung der Menge $N = \{1, 2, 3, \ldots\}$ auf ihre echte Teilmenge $N_1 = \{2, 3, 4, \ldots\}$ ist. Damit ist aber auch jede Menge N', die zu N gleichmächtig ist, unendlich und folglich — nach Satz 3 — auch jede Menge, die eine zu N gleichmächtige Teilmenge N' enthält.

Beispiele. Die Menge der reellen Zahlen und die Menge der komplexen Zahlen enthalten die Menge N der natürlichen Zahlen und sind daher unendlich. Auch das Intervall $[0, 1]$ ist unendlich, da es die zu N gleichmächtige Menge N' der Zahlen $\frac{1}{n}$ $(n = 1, 2, 3, \ldots)$ enthält.

Definition 4. *Eine Menge heißt abzählbar (unendlich), wenn sie gleichmächtig der Menge der natürlichen Zahlen ist.*

Eine abzählbare Menge ist — mit anderen Worten — eine Menge, deren Elemente man mit Hilfe aller natürlichen Zahlen so durchnumerieren kann, daß alle Zahlen dabei benutzt werden und verschiedene Elemente stets verschiedene Nummern erhalten. Eine abzählbare Menge kann man also immer in der Form

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}$$

schreiben. Die Beispiele am Ende des vorangehenden Paragraphen zeigen, daß sowohl die Menge der geraden als auch die Menge der ungeraden Zahlen und sogar die Menge der rationalen Zahlen abzählbar ist.

Definition 5. *Eine Menge heißt überabzählbar oder nichtabzählbar, wenn sie weder endlich noch abzählbar ist.*

Das folgende Beispiel zeigt, daß es tatsächlich überabzählbare Mengen gibt.[1])

Die Menge der reellen Zahlen ist nichtabzählbar. Zunächst bemerken wir, daß auf Grund der Beispiele 2 und 7 des vorangehenden Paragraphen folgt, daß die Menge der reellen Zahlen gleichmächtig dem Intervall $(0, 1)$ ist. Es genügt also zu zeigen, daß das Intervall $(0, 1)$ überabzählbar ist.

Wir benutzen hierbei die bekannte Tatsache, daß sich jede Zahl des Intervalls $(0, 1)$ in Form eines endlichen oder unendlichen Dezimalbruches

$$0, a_1 a_2 a_3 \cdots$$

darstellen läßt, wobei mindestens eine der Ziffern a_i verschieden von Null ist (da die Zahl $0 = 0{,}000\cdots$ nicht dem Intervall $(0, 1)$ angehört). Dazu müssen wir bemerken, daß für jede Zahl, die sich in Form eines endlichen Dezimalbruches schreiben läßt, noch eine andere Schreibweise möglich ist, bei der von einer gewissen Stelle ab alle Ziffern a_i gleich 9 sind. Zum Beispiel ist

$$0{,}53000\cdots = 0{,}52999\cdots$$

Für alle übrigen Zahlen (das sind also alle irrationalen Zahlen und die rationalen Zahlen, die sich als periodischer Bruch mit einer von 9 verschiedenen Periode darstellen lassen) ist diese Schreibweise völlig eindeutig.[2]) Wir wählen nun von den möglichen Schreibweisen für die Zahlen der ersten Art eine aus, etwa die in Form

[1]) Es gibt sogar unendlich viele verschiedene Mächtigkeiten von unendlichen Mengen, auf die wir hier aber nicht näher eingehen wollen. Vgl. dazu die schon erwähnten Bücher [1] oder [2] (oder auch irgendein anderes einschlägiges Buch über Mengenlehre — *Anm. d. wissenschaftl. Red.*).

[2]) Vgl. A. J. Chintschin, Elemente der Zahlentheorie. EdEM Band 1, S. 274.

eines endlichen Dezimalbruches. Dann lassen sich alle Zahlen des Intervalls (0, 1) in eindeutiger Weise in der Form

$$0, a_1 a_2 a_3 \ldots,$$

schreiben, wobei nicht alle a_i gleich Null sind und von keiner Stelle ab nur die Ziffer 9 steht. Umgekehrt liefert auch jeder derartige Dezimalbruch eine Zahl des Intervalls (0, 1).

Man sieht unmittelbar ein, daß das Intervall (0, 1) eine unendliche Menge ist, da in ihm die zur Menge der natürlichen Zahlen gleichmächtige Menge

$$N' = \left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right\}$$

enthalten ist (vgl. Satz 5).

Wir zeigen nun, daß das Intervall (0, 1) überdies nicht abzählbar ist.

Angenommen nämlich, es wäre abzählbar, so könnte man die Elemente des Intervalls folgendermaßen durchnumerieren:

$$(0, 1) = \{c_1, c_2, c_3, \ldots\},$$

wobei sich jede der Zahlen c_n in angegebener Weise als Dezimalbruch darstellen läßt:

$$\begin{aligned} c_1 &= 0, a_{11} a_{12} a_{13}, \ldots, \\ c_2 &= 0, a_{21} a_{22} a_{23} \ldots, \\ c_3 &= 0, a_{31} a_{32} a_{33} \ldots, \\ &\ldots\ldots\ldots\ldots \\ c_n &= 0, a_{n1} a_{n2} a_{n3} \ldots, \\ &\ldots\ldots\ldots\ldots \end{aligned} \tag{4}$$

Wir konstruieren nun eine Zahl

$$c = 0, b_1 b_2 b_3 \ldots$$

auf folgende Weise: Zunächst nehmen wir eine Ziffer b_1, die verschieden ist von a_{11}, 0 und 9; dann nehmen wir eine Ziffer b_2, die von a_{22}, 0 und 9 verschieden ist; dann wählen wir b_3 verschieden von a_{33}, 0 und 9; allgemein wählen wir b_n verschieden von a_{nn}, 0 und 9. Das Vorhandensein von zehn Ziffern gewährt, daß stets für die Auswahl genügend viele Möglichkeiten vorhanden sind (und zwar stehen bei unserer Vorschrift jedesmal noch sieben Ziffern zur Auswahl). Der so konstruierte Bruch besitzt dann gleichfalls die oben angegebenen Eigenschaften, und zwar sogar in dem verschärften Maße, daß er keine der Ziffern 0 und 9 enthält, d. h., die Zahl c gehört sicher dem Intervall (0, 1) an. Die Darstellung dieser Zahl c ist aber verschieden von allen Darstellungen in (4). In der Tat unterscheidet sich diese Darstellung von c von der von c_1, da $b_1 \neq a_{11}$, von der von c_2, da $b_2 \neq a_{22}$ usw. Nun war aber die Darstellung der Zahlen des Intervalls durch Brüche des angegebenen Typus eindeutig, also gilt:

$$c \neq c_1,\ c \neq c_2,\ c \neq c_3, \ldots, c \neq c_n, \ldots$$

Wir haben damit gezeigt, daß die Zahl c nicht unter den Zahlen (4) vorkommt, während wir angenommen hatten, daß in (4) alle Zahlen des Intervalls durchnumeriert sind. Der erhaltene Widerspruch beweist die Behauptung.

Unter allen unendlichen Mengen erweisen sich die abzählbaren Mengen als die kleinsten im folgenden Sinne:

Satz 6. *Jede unendliche Menge besitzt eine abzählbare Teilmenge.*

Beweis. Es sei M eine beliebige unendliche Menge. Dann ist $M \neq 0$. Wir wählen nun ein beliebiges Element aus M aus und nennen es a_1. Angenommen, wir hätten aus M schon n paarweise verschiedene Elemente $a_1, a_2, \ldots, a_n$ ausgewählt. Da M unendlich ist, ist

$$M \setminus \{a_1, a_2, \ldots, a_n\} \neq 0,$$

und man kann ein weiteres Element

$$a_{n+1} \in M \setminus \{a_1, a_2, \ldots, a_n\}$$

auswählen. Dieses ist von allen bereits ausgewählten Elementen verschieden. Auf Grund des Prinzips der vollständigen Induktion ist damit gezeigt, daß es zu jeder natürlichen Zahl n eine Teilmenge $A_n = \{a_1, a_2, \ldots, a_n\}$ von M gibt, die genau n Elemente enthält, wobei die Menge A_{n+1} aus der Menge A_n durch Hinzunahme eines weiteren Elementes a_{n+1} aus M hervorgeht. Dann ist offensichtlich

$$A = \bigcup_{n=1}^{\infty} A_n = \{a_1, a_2, \ldots, a_n, \ldots\}$$

eine abzählbare Teilmenge von M.

Hieraus kann man jetzt leicht folgern, daß die für endliche Mengen bewiesene Eigenschaft, keine gleichmächtige echte Teilmenge zu besitzen (Satz 1), auch nur für endliche Mengen gilt. Wir behaupten also:

Satz 7. *Jede unendliche Menge M ist einer ihrer echten Teilmengen gleichmächtig.*

Beweis. Auf Grund von Satz 6 enthält M eine abzählbare Teilmenge

$$A = \{a_1, a_2, \ldots, a_n, \ldots\}.$$

Es sei dann $M \setminus A = B$. Wir betrachten die folgende Abbildung f der Menge M in sich:

$$f(a_n) = a_{n+1}\ (n = 1, 2, \ldots), \quad f(b) = b\ (b \in B).$$

Offensichtlich ist f eine eineindeutige Abbildung der Menge M auf ihre echte Teilmenge $M \setminus \{a_1\}$, womit die Behauptung unseres Satzes bewiesen ist.

Wir geben jetzt die folgende neue Definition der endlichen und der unendlichen Menge:

Definition 2'. *Eine Menge heißt endlich, wenn sie keiner ihrer echten Teilmengen gleichmächtig ist. Eine Menge heißt unendlich, wenn sie nicht endlich ist.*

Aus den Sätzen 1 und 7 folgt die Gleichwertigkeit dieser Definition mit der früheren Definition 2. In der Tat ist eine Menge, die nach Definition 2 endlich ist, auf Grund von Satz 1 auch nach Definition 2' endlich. Umgekehrt: Ist eine Menge nach Definition 2' endlich, so muß sie auch nach Definition 2 endlich sein, denn wäre sie nach Definition 2 unendlich, so wäre sie auf Grund von Satz 7 auch nach Definition 2' unendlich, was ja nicht der Fall sein sollte. Also sind beide Endlichkeits-Definitionen gleichwertig. Hieraus folgt sofort (durch gegenteilige Annahme) die Gleichwertigkeit auch der Definitionen der unendlichen Menge.

Wir bemerken, daß die Definition 2' den (allerdings nur formalen) Vorzug gegenüber der Definition 2 besitzt, daß sie allein mit den Begriffsbildungen der allgemeinen Mengenlehre formuliert ist, während die Definition 2 bekannte Eigenschaften der natürlichen Zahlen voraussetzt.

§ 5. Geordnete Mengen

Bisher haben wir nur die Eigenschaften der Mengen studiert, die sich auf die Grundbeziehung gründen, die zwischen den Mengen und ihren Elementen besteht. Wir haben indes noch keine Beziehungen betrachtet, die zwischen den Elementen ein und derselben Menge erklärt sind; alle Elemente einer Menge waren bis jetzt völlig gleichberechtigt. In der Mathematik kommen jedoch solche — wie man sagt — „reine" Mengen nur selten vor. Meistens betrachtet man Mengen, zwischen deren Elementen ganz bestimmte Beziehungen, ganz bestimmte Abhängigkeiten bestehen. So können zwei Geraden einer Ebene, wie man sie in der Geometrie studiert, sich schneiden oder parallel sein. Zwischen den Punkten einer Geraden ist eine Beziehung erklärt, die man durch die Worte „einer der drei Punkte liegt zwischen den beiden anderen" ausdrückt. In der Arithmetik bestehen zwischen Zahlen die Beziehungen $a + b = c$ und $ab = c$ usw. Eine der wichtigsten Beziehungen zwischen Zahlen ist ihre Anordnung. Die Zahlen irgendeiner (reellen) Gesamtheit sind auf natürliche Weise in einer bestimmten Reihenfolge, nämlich ihrer Größe nach, angeordnet. So ist für die Menge der natürlichen Zahlen die Anordnung

$$1, 2, 3, \ldots$$

die natürliche Ordnung.

In diesem Paragraphen wollen wir den Begriff der Ordnung in seiner allgemeinsten Form, und zwar sogleich für beliebige Mengen betrachten.

Definition 1. *Eine Menge M heißt geordnet, wenn zwischen ihren Elementen eine Beziehung $a < b$*[1]) *(in Worten: „a geht b voran") erklärt ist, die die folgenden Eigenschaften besitzt:* 1.) *Zwischen Elementen a und b besteht eine und nur eine der drei Beziehungen $a = b$, $a < b$, $b < a$;* 2.) *Für irgend drei Elemente a, b, c gilt: wenn $a < b$ und $b < c$ ist, so ist $a < c$. Auch die leere Menge gilt als geordnet.*

Bemerkung. Das Zeichen „=" verwenden wir stets zur Bezeichnung der Identität, des Zusammenfallens der Elemente. Die Schreibweise $a = b$ bedeutet also einfach, daß die Buchstaben a und b ein und dasselbe Element der Menge M bezeichnen. Daher ist die Bedingung 1.) gleichwertig der Bedingung, daß zwischen zwei voneinander verschiedenen Elementen eine und nur eine der Beziehungen $a < b$ oder $b < a$ besteht.[2])

Geht das Element a dem Element b voran, so sagt man auch, daß b auf a folgt, und schreibt $b > a$. Die Beziehung $a > b$ besitzt, wie man leicht nachprüft, die zu 1.) und 2.) analogen Eigenschaften. Man kann auch sie als Grundbeziehung einführen und mit ihrer Hilfe die Beziehung $a < b$ definieren (siehe unten, § 9).

Wenn man in einer geordneten Menge die Rollen der Beziehungen $<$ und $>$ vertauscht, d. h. also, wenn man die Elemente a, die einem Element b vorangehen, als dessen Nachfolger ansieht und umgekehrt, so erhält man eine neue geordnete Menge

[1]) Man verwechsle diese Schreibweise nicht mit der genauso bezeichneten Beziehung „kleiner" zwischen Zahlen.

[2]) Der Leser sehe sich hierauf noch einmal die Formulierung der Bedingung 1.) an. Es darf hier keineswegs heißen „Zwischen *zwei* Elementen a und b . . .", da zwei Elemente stets voneinander verschieden sind (sonst wären es nicht zwei!). Die Bedingung 2.) ist in der angegebenen Form gleichfalls korrekt. — *Anm. d. wissenschaftl. Red.*

M', deren Ordnung man als die zur Ordnung von M *entgegengesetzte* bezeichnet. So ist zur oben betrachteten Ordnung der natürlichen Zahlen die Ordnung

$$\ldots, 3, 2, 1$$

entgegengesetzt.

Geordnete Mengen, die aus denselben Elementen bestehen, aber verschieden geordnet sind, werden als verschieden angesehen. Daher ist es bei der Angabe einer geordneten Menge durch ihre Elemente stets notwendig, auch die jeweils betrachtete Ordnung zu vermerken. Wir wollen vereinbaren, daß in Zukunft die Schreibweise von links nach rechts stets die entsprechende Anordnung der Elemente angibt, und behalten die Bezeichnungen durch geschweifte Klammern bei. So kann man die Menge der natürlichen Zahlen einmal in ihrer natürlichen Anordnung betrachten, zum anderen kann man sie in der entgegengesetzten Ordnung betrachten, man kann aber auch die ungeraden Zahlen in ihrer natürlichen Anordnung den geraden Zahlen in natürlicher Anordnung voranstellen oder umgekehrt die geraden Zahlen den ungeraden voranstellen, man kann schließlich hierbei die entgegengesetzten Ordnungen betrachten usw. Wir kommen auf diese Weise zu den folgenden geordneten Mengen:

$$\{1, 2, 3, \ldots\}, \tag{1}$$

$$\{\ldots, 3, 2, 1\}, \tag{2}$$

$$\{1, 3, 5, \ldots, 2, 4, 6, \ldots\}, \tag{3}$$

$$\{1, 3, 5, \ldots, 6, 4, 2\}, \tag{4}$$

$$\{\ldots, 5, 3, 1, 2, 4, 6, \ldots\}, \tag{5}$$

$$\{\ldots, 5, 3, 1, \ldots, 6, 4, 2\}. \tag{6}$$

Ein Element, das keinen Vorgänger besitzt, heißt ein *erstes Element*. Ein Element, das keinen Nachfolger besitzt, heißt ein *letztes Element*. Zwei (verschiedene) Elemente a und b heißen *benachbart*, wenn es kein Element c gibt mit $a < c < b$ oder $b < c < a$. Sind a und b benachbart und ist $a < b$, so sagt man, daß a ein *unmittelbarer Vorgänger* von b ist und daß b ein *unmittelbarer Nachfolger* von a ist (unmittelbar vorangeht bzw. nachfolgt). Die geordnete Menge (1) besitzt ein erstes Element, aber kein letztes, die Menge (2) dagegen ein letztes Element, aber kein erstes. Die Menge (4) besitzt sowohl ein erstes als auch ein letztes, hingegen die Menge (5) weder ein erstes noch ein letztes Element. Die Menge (3) besitzt zwei Elemente ohne unmittelbaren Vorgänger, die Menge (6) zwei Elemente ohne unmittelbaren Nachfolger. In allen diesen Mengen besitzt jedes Element ein benachbartes. Hingegen besitzt etwa in der nach der Größe ihrer Elemente geordneten Menge der rationalen Zahlen kein Element ein benachbartes Element, da zwischen zwei rationalen Zahlen a und b stets noch eine weitere, z. B. $\frac{a+b}{2}$ liegt.

Für „$a = b$ oder $a < b$“ schreibt man: $a \leq b$, entsprechend für „$a = b$ oder $a > b$“: $a \geq b$. Aus Definition 1 erhalten wir unmittelbar:

Satz 1: *Wenn $a \leq b$ und $b \leq a$ ist, so ist $a = b$.*

Satz 2. *Wenn $a \leq b$ und $b \leq c$ ist, so ist $a \leq c$. Wenn $a \geq b$ und $b \geq c$ ist, so ist $a \geq c$. Besteht in einer der Voraussetzungen eigentliche Ungleichheit, so auch in der jeweiligen Behauptung.*

Definition 2. *Geordnete Mengen A und B heißen ähnlich, wenn man zwischen ihnen eine eineindeutige Beziehung herstellen kann, die die Ordnung erhält, d. h. eine Zuordnung, für die aus $a_1 \to b_1$, $a_2 \to b_2$ und $a_1 < a_2$ folgt, daß $b_1 < b_2$ ist.*

Aus Definition 2 folgt unmittelbar, daß alle Mengen, die nur ein einziges Element enthalten, untereinander ähnlich sind und daß die leere Menge nur sich selbst ähnlich ist. Man sagt von ähnlichen Mengen auch, sie hätten *den gleichen Ordnungstypus*. Die Ähnlichkeit von Mengen bezeichnen wir durch $A \approx B$.

Es soll dem Leser überlassen bleiben, die folgenden Eigenschaften der Ähnlichkeit von geordneten Mengen zu beweisen:

I. Reflexivität: $A \approx A$.

II. Symmetrie: *Wenn* $A \approx B$ *ist, so ist* $B \approx A$.

III. Transitivität: *Wenn* $A \approx B$ *und* $B \approx C$ *ist, so ist* $A \approx C$.

Ein Vergleich der Definition der Ähnlichkeit mit der Definition der Gleichmächtigkeit (§ 3, Definition 4) zeigt, daß die erste die zweite einschließt, daß also gilt:

Satz 3. *Ähnliche Mengen sind gleichmächtig; aus* $A \approx B$ *folgt* $A \sim B$.

Die Umkehrung dieses Satzes ist keineswegs richtig. So sind die Mengen (1) und (2) gleichmächtig (sie sind ja als ungeordnete Mengen sogar gleich), aber nicht ähnlich, denn die Menge (1) besitzt ein erstes Element, aber die Menge (2) nicht, während bei einer ähnlichen Abbildung das erste Element der einen Menge in das erste Element der anderen Menge übergehen müßte. Jedoch gilt für endliche Mengen auch die Umkehrung des Satzes 3, also:

Satz 4. *Wenn zwei endliche geordnete Mengen gleichmächtig sind, so sind sie auch ähnlich.*

Dieser Satz ist auf Grund der Eigenschaften I. bis III. der Ähnlichkeit eine unmittelbare Folge aus dem folgenden Satz 7. Als eine gewisse Umkehrung des Satzes 3 für beliebige Mengen kann der folgende Satz gelten:

Satz 5. *Ist die Menge A einer geordneten Menge B gleichmächtig, so gibt es eine Ordnung für A — also eine Beziehung zwischen den Elementen von A, die den Bedingungen* 1.) *und* 2.) *aus Definition* 1 *genügt — derart, daß die so geordnete Menge A der Menge B ähnlich ist.*[1])

Beweis. Es seien a_1 und a_2 beliebige Elemente der Menge A und b_1 und b_2 die ihnen vermöge der als existent vorausgesetzten eineindeutigen Abbildung von A auf B zugeordneten Elemente aus B. Wir setzen nun fest, daß $a_1 < a_2$ genau dann gelten soll, wenn $b_1 < b_2$ (im Sinne der vorausgesetzten Ordnung von B — *Anm. d. wissenschaftl. Red.*) ist. Man prüft leicht nach, daß die so definierte Ordnungsbeziehung den Bedingungen 1.) und 2.) genügt und daß die (so geordnete — *Anm. d. wissenschaftl. Red.*) Menge A der geordneten Menge B ähnlich ist.

Satz 6. *Jede geordnete endliche (nichtleere) Menge A besitzt ein erstes und ein letztes Element.*

Beweis. Angenommen, die Menge A besäße kein letztes Element. Dann wählen wir ein beliebiges Element $a_1 \in A$ aus. Da es nicht letztes Element sein kann, gibt es ein $a_2 \in A$ mit $a_1 < a_2$; da auch a_2 nicht letztes Element sein kann, gibt es ein

[1]) Man kann sogar zeigen, daß jede beliebige Menge — wie man sagt — wohlgeordnet werden kann (siehe [1], S. 99). Dieser Beweis würde aber den Rahmen unseres Artikels überschreiten.

$a_2 \in A$ mit $a_2 < a_3$ usw. Allgemein kann man zu jedem a_n ein $a_{n+1} \in A$ finden mit $a_n < a_{n+1}$. Man kann also induktiv für jedes n ein a_n finden. Es sei nun

$$N' = \{a_1, a_2, a_3, \ldots\}$$

die Menge der so erhaltenen Elemente. Auf Grund von Eigenschaft 2.) ergibt sich: Wenn $i < k$ ist, so ist $a_i < a_k$, woraus auf Grund von Eigenschaft 1.) folgt, daß $a_i \neq a_k$ ist. Das besagt aber, daß die Menge N' der Menge der natürlichen Zahlen gleichmächtig ist. Dann wäre aber die Menge A unendlich (§ 4, Satz 5), was der Voraussetzung des Satzes widerspricht. Ganz analog beweist man die Existenz eines ersten Elementes.

Satz 7. *Jede endliche Menge kann geordnet werden. Alle endlichen Mengen mit der gleichen Anzahl $n > 0$ von Elementen sind dem Abschnitt $|1, n|$ der Folge der natürlichen Zahlen und damit untereinander ähnlich.*

Beweis. Die leere Menge ist per definitionem geordnet. Sei also $A \neq 0$ eine beliebige endliche Menge, also etwa $A \sim |1, n|$. Der Abschnitt $|1, n|$ ist offenbar eine (durch die übliche $<$-Beziehung) geordnete Menge. Dann kann die Menge A nach Satz 5 geordnet werden. Es sei jetzt A eine beliebige endliche geordnete Menge mit einer Anzahl $n > 0$ von Elementen. Nach Satz 6 besitzt die Menge A ein erstes Element a_1. Ist $n > 1$, so ist die Menge

$$A_1 = A \setminus \{a_1\} \neq 0$$

und enthält damit — wiederum nach Satz 6 — gleichfalls ein erstes Element a_2, wobei $a_1 < a_2$. Angenommen nun, wir hätten schon ein Element a_i gefunden, so ist — sofern $i < n$ ist —

$$A_i = A \setminus \{a_1, a_2, \ldots, a_i\} \neq 0,$$

und A_i enthält nach Satz 6 ein erstes Element a_{i+1}, wobei $a_i < a_{i+1}$. Auf diese Weise können wir für jedes $i \leq n$ ein Element a_i bestimmen, und es ist

$$A_n = \{a_1, a_2, \ldots, a_n\} \sim |1, n| \sim A.$$

Da die Menge A voraussetzungsgemäß endlich ist, ist sie keiner echten Teilmenge gleichmächtig (§ 4, Satz 1), es muß also gelten:

$$A = A_n = \{a_1, a_2, \ldots, a_n\}.$$

Wenn $i < k$ ist, so ist offensichtlich $a_i < a_k$, d. h., die geordnete Menge A ist dem Abschnitt $|1, n|$ sogar ähnlich.

Dieser Satz hat zur Folge, daß alle $n!$ Permutationen einer Menge von n Elementen den gleichen Ordnungstypus haben.

Kapitel II

GRUPPEN, RINGE, KÖRPER

§ 6. Gruppen

In der Arithmetik und in der Algebra haben wir es mit Objekten verschiedenster Art zu tun: mit ganzen, rationalen, reellen und komplexen Zahlen, mit Polynomen, rationalen Funktionen usw. Dabei werden in erster Linie die Eigenschaften der vier Grundoperationen, der Addition, Subtraktion, Multiplikation und Division studiert.

Es zeigt sich nun, daß diese Operationen für die verschiedenen Objekte doch weitgehend dieselben Eigenschaften besitzen. Daher ist die Betrachtung abstrakter Gebilde mit den uns interessierenden Eigenschaften — wie sie der modernen Algebra eigen ist — durchaus natürlich und zweckmäßig.

Hinzu kommt noch, daß man in einer solchen abstrakten Form die Bedeutung und die gegenseitige Abhängigkeit der betreffenden Eigenschaften leichter klären kann, da der jeweilige Sachverhalt in konkreten Zahlenbereichen, im Bereich der Polynome usw. durch spezielle weitere Eigenschaften kompliziert wird, die die zu untersuchenden Eigenschaften verschleiern.

In den späteren Kapiteln wollen wir die verschiedenen Zahlenbereiche behandeln. Um die Bedeutung ihrer unterschiedlichen Eigenschaften klarer herausarbeiten zu können und um eine Wiederholung ein und derselben Überlegung in jedem dieser Bereiche zu vermeiden, schicken wir in diesem Kapitel die grundlegenden Begriffe der abstrakten Algebra voraus. Wir verweisen den Leser, der diese Fragen eingehender studieren will, auf den Artikel von L. J. OKUNJEW[1]) und die Bücher L. J. OKUNJEW [3] und B. L. VAN DER WAERDEN [4].

Vom Standpunkt der Mengenlehre aus ist jede der vier Grundoperationen nichts anderes als eine gewisse Beziehung zwischen irgend drei Elementen der betrachteten Menge (siehe Anfang § 5). Diese Beziehungen unterscheiden sich jedoch in einem wesentlichen Punkte von anderen (z. B. von der in § 5 untersuchten Ordnungsbeziehung), nämlich darin, daß in allen vier Fällen zu je zwei Elementen ein eindeutig bestimmtes drittes (das Resultat der jeweiligen Operation) gehört, das mit den beiden anderen ein Tripel bildet, das in der betrachteten Beziehung steht. Derartige Beziehungen haben einen besonderen Namen erhalten:

Definition 1. *Eine Vorschrift, die jedem geordneten Paar a, b von Elementen einer Menge M ein eindeutig bestimmtes Element derselben Menge zuordnet, heißt eine in M definierte algebraische Operation.*

Unter Verwendung des Funktionsbegriffs (§ 3, Definition 1) kann man auch kürzer sagen, daß eine in M definierte algebraische Operation eine Funktion auf der Menge der geordneten Paare von Elementen aus M mit Werten aus M ist.

Beispiele für algebraische Operationen sind: Die Addition $a + b = c$, die Subtraktion $a - b = c$, die Multiplikation $ab = c$, die Division $\frac{a}{b} = c$ im Bereiche der reellen Zahlen, wobei zu beachten ist, daß bei der Division der Fall $b = 0$ aus-

[1]) L. J. OKUNJEW, Der Ring der Polynome und der Körper der rationalen Funktionen, EdEM Band 2.

geschlossen werden muß, da hierfür die Division nicht definiert ist. Weitere Beispiele sind die Addition, Subtraktion, Multiplikation, Division für komplexe Zahlen, die Vektoraddition nach der Parallelogrammregel, die Addition, Subtraktion und Multiplikation von Polynomen usw.

Unter Umständen können zwei oder mehrere Operationen dadurch verknüpft sein, daß in ihnen die Rolle von gegebenen und gesuchten Elementen vertauscht ist: Wenn $a + b = c$ ist, so ist $c - a = b$; wenn $ab = c$ ist, so ist $\frac{c}{b} = a$ usw. Dieser Zusammenhang zwischen Operationen führt uns zum Begriff der Umkehroperation. Die genaue Definition hierfür lautet:

Vorgegeben sei eine Operation, die jedem geordneten Paar a, b von Elementen aus M ein gewisses Element c aus M zuordnet. Die beiden Operationen, die man aus der gegebenen dadurch erhält, daß man die Rolle von a und c bzw. von b und c (also eines der gegebenen Elemente mit dem gesuchten) vertauscht, heißen die *Umkehroperationen* der gegebenen Operation.

Somit besteht die erste Umkehroperation in der Zuordnung von a zu dem geordneten Paar b, c und die zweite in der Zuordnung von b zu dem geordneten Paar a, c. Bekanntlich brauchen die Umkehroperationen weder immer ausführbar zu sein, noch brauchen sie eindeutig zu sein. So sind im Bereich der natürlichen Zahlen die Addition und Multiplikation erklärt, während die Umkehroperationen Subtraktion und Division nicht unbeschränkt ausführbar sind.

Eine Operation heißt *kommutativ*, wenn sie den Paaren a, b und b, a dasselbe Resultat zuordnet. Wir werden später sehen, daß für kommutative Operationen die eine Umkehroperation dann und nur dann existiert, wenn die andere existiert, und daß dann beide zusammenfallen. Für nichtkommutative Operationen ist das keineswegs der Fall.

So ist im Bereich der positiven reellen Zahlen die Operation $f(a, b) = a^b$ nicht kommutativ, da im allgemeinen $a^b \neq b^a$ ist. Die erste Umkehroperation $f_1(c, b) = \sqrt[b]{c}$ existiert, während die zweite Umkehroperation $f_2(c, a) = {}_a\log c$ nicht definiert ist für $a = 1$ und $c \neq 1$ und für solche a und c, für die ${}_a\log c \leq 0$ (wir haben ja unsere Operation nur auf der Menge der positiven reellen Zahlen betrachtet). In den Fällen, in denen auch die zweite Operation existiert, fällt sie nicht mit der ersten zusammen.

In ein und derselben Menge können mehrere algebraische Operationen gegeben sein. Da wir aber die allgemeinen Eigenschaften der Addition und der Multiplikation von Zahlen studieren wollen, betrachten wir anfangs Mengen mit nur einer algebraischen Operation. So gelangen wir zu dem ersten fundamentalen Begriff der modernen Algebra, zum Begriff der Gruppe.

Definition 2. *Eine nichtleere Menge G heißt eine Gruppe, wenn in ihr eine algebraische Operation — Multiplikation genannt — erklärt ist, die jedem geordneten Paar a, b von Elementen aus G ein eindeutig bestimmtes Element ab aus G — das Produkt aus a und b — zuordnet und die die folgenden Eigenschaften besitzt:*

I. (Assoziatives Gesetz) $a(bc) = (ab)c$[1]);

II. (Gesetz der beiderseitigen Umkehrbarkeit) *Für beliebiges a und b aus G sind die Gleichungen $ax = b$ und $ya = b$ lösbar, d. h., zu jedem a, b aus G gibt es in G ein c und ein d derart, daß $ac = b$ und $da = b$ ist.*

[1]) Das Zeichen „=" bezeichnet nach wie vor das Zusammenfallen der Elemente.

Ist die Gruppenoperation kommutativ, also $a\,b = b\,a$ für beliebiges a, b aus G, so heißt die Gruppe *kommutativ*[1]).

Wir führen einige Beispiele für Gruppen an:

Beispiel 1. Die Mengen der ganzen, der rationalen, der reellen und der komplexen Zahlen bilden Gruppen bezüglich der Addition von Zahlen als Gruppenoperation „Multiplikation". Keine dieser Mengen erweist sich als Gruppe bezüglich der Multiplikation von Zahlen, da die Gleichung $0x = 1$ in keinem dieser Bereiche eine Lösung besitzt.

Beispiel 2. Die Mengen der rationalen, reellen und komplexen Zahlen, die verschieden sind von der Zahl 0, bilden Gruppen bezüglich der Multiplikation von Zahlen.

Beispiel 3. Die Menge G aus den beiden Elementen e und a bildet bezüglich der durch die Gleichungen $ee = aa = e$ und $ea = ae = a$ festgelegten Operation eine Gruppe.

Beispiel 4. Es sei G die Menge aller eineindeutigen Abbildungen der Menge M auf sich (§ 3, Definition 3). Das einem Element $a \in M$ vermöge der Abbildung $s \in G$ zugeordnete Element möge mit as bezeichnet werden. Das Produkt st von Abbildungen s und t aus G soll die Abbildung sein, die man erhält, wenn man die gegebenen Abbildungen hintereinander ausführt (zuerst s, dann t), d. h., es ist

$$a(st) = (as)\,t$$

für jedes $a \in M$[2]). Bezüglich dieser Multiplikation erweist sich die Menge G als Gruppe. In der Tat gilt das assoziative Gesetz I., denn sind r, s und t irgendwelche Abbildungen aus G, so gilt für jedes a aus M:

$$a\,[r\,(st)] = (ar)\,(st) = [(ar)\,s]\,t$$

und ebenso:

$$a\,[(rs)\,t] = [a\,(rs)]\,t = [(ar)\,s]\,t,$$

d. h., es ist

$$a\,[r\,(st)] = a\,[(rs)\,t]$$

für jedes a aus M. Also ist $r(st) = (rs)\,t$ (beide Abbildungen ergeben sich, indem man die Abbildungen r, s, t hintereinander ausführt).

Wir beweisen jetzt, daß in G auch das Gesetz der beiderseitigen Umkehrbarkeit II. gilt. Seien dazu s und t beliebige Abbildungen aus G. Zu der eineindeutigen Abbildung s existiert die inverse Abbildung s^{-1}, die gleichfalls eineindeutig ist (§ 3). Namentlich gilt: Ist $as = b$, so ist $bs^{-1} = a$. Offensichtlich ist also $ss^{-1} = s^{-1}s = e$, wobei e die identische Abbildung von M auf sich ist, also $ex = xe = x$ für beliebige Abbildungen x aus G gilt. Nehmen wir nun einmal an, daß es in G eine Abbildung u mit $su = t$ gibt. Wenn wir diese Gleichung von links mit s^{-1} multiplizieren, so erhalten wir:

$$s^{-1}(su) = s^{-1}t.$$

Auf Grund des assoziativen Gesetzes folgt $s^{-1}(su) = (s^{-1}s)u = eu = u$, d. h., es muß — wenn es überhaupt eine Lösung u gibt — $u = s^{-1}\,t$ sein. Also kann die

[1]) Kommutative Gruppen nennt man auch *abelsch*.

[2]) Man kann das Produkt st auch so definieren, daß zuerst die Abbildung t und dann die Abbildung s ausgeführt wird. Dann schreibt man das einem Element $a \in M$ vermöge der Abbildung s zugeordnete Element besser als sa.

Gleichung $sx = t$ höchstens die Lösung $s^{-1}t$ haben. Diese Abbildung genügt aber auch in der Tat der Gleichung $sx = t$, da $s(s^{-1}t) = (ss^{-1})t = et = t$ ist. Analog beweist man, daß die Gleichung $ys = t$ die eindeutig bestimmte Lösung $y = ts^{-1}$ besitzt.

Also ist G tatsächlich eine Gruppe. Sie heißt *die Transformationsgruppe der Menge M*. Ist M eine endliche Menge, so spricht man auch von der *Permutationsgruppe von M*.

Für den Fall, daß M mehr als zwei Elemente enthält, ist die Permutationsgruppe von M nicht kommutativ. So enthält die Permutationsgruppe der drei Zahlen 1, 2, 3 sechs Elemente. Bezeichnen wir jede Permutation durch ein zweizeiliges Schema, in dem wir unter jede Zahl die ihr vermöge der Permutation zugeordnete Zahl schreiben, so erhalten wir die Elemente dieser Gruppe in der Form

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}, \begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}, \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}, \begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}, \begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}, \begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}.$$

Dann ist

$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix} \text{ und } \begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\cdot\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$

d. h., der Wert eines Produktes hängt im allgemeinen von der Reihenfolge der Faktoren ab.

Die Permutationsgruppen spielen eine große Rolle in der Algebra. Mit ihrer Hilfe gelang es dem französischen Mathematiker Evariste Galois (1811—1832), die Frage nach der Auflösbarkeit einer algebraischen Gleichung durch Radikale endgültig zu klären (und im allgemeinen Falle negativ zu entscheiden — *Anm. d. wissenschaftl. Red.*).

Folgerungen aus dem assoziativen und dem kommutativen Gesetz.

Auf Grund des assoziativen Gesetzes I. kann man von dem Produkt aus Elementen a, b, c einer Gruppe G in dem Sinne sprechen, daß man darunter eines der beiden (untereinander gleichen) Produkte $a(bc)$ oder $(ab)c$ versteht und dafür ohne Klammern abc schreibt. Man kann jedoch auch ohne Benutzung des assoziativen Gesetzes das allgemeine Produkt

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

induktiv für n beliebige Elemente $a_1, a_2, \ldots, a_n$ einer Gruppe G definieren (die Rechtfertigung der induktiven Definitionen wird in Kapitel III gegeben), und zwar:

Definition 3. $\prod_{i=1}^{1} a_i = a_1$ *für jedes Element* a_1 *aus* G; $\prod_{i=1}^{n+1} a_i = \left(\prod_{i=1}^{n} a_i\right) a_{n+1}$.

Entsprechend dieser Definition ist:

$$a_1 a_2 a_3 = (a_1 a_2) a_3,$$

$$a_1 a_2 a_3 a_4 = [(a_1 a_2) a_3] a_4,$$

$$a_1 a_2 a_3 a_4 a_5 = \{[(a_1 a_2) a_3] a_4\} a_5 \text{ usw.}$$

Das assoziative Gesetz erlaubt es dann, ein Produkt aus zwei (mehrgliedrigen) Produkten als ein Produkt über alle vorkommenden Elemente darzustellen, d. h.:

$$(a_1 a_2 \cdots a_m)\,(a_{m+1}\, a_{m+1} \cdots a_n) = a_1 a_2 \cdots a_n$$

oder kürzer:

$$\prod_{i=1}^{m} a_i \prod_{k=1}^{n} a_{m+k} = \prod_{i=1}^{m+n} a_i. \tag{1}$$

Wir beweisen die Gleichung (1) bei festgehaltenem m durch vollständige Induktion über n. Für $n = 1$ folgt die Gleichung (1) unmittelbar aus Definition 3. Aus der Annahme, daß (1) bereits für die Zahl n gilt, ergibt sich auf Grund der Definition 3 und des assoziativen Gesetzes:

$$\prod_{i=1}^{m} a_i \prod_{k=1}^{n+1} a_{m+k} = \prod_{i=1}^{m} a_i \left(\prod_{k=1}^{n} a_{m+k} \cdot a_{m+n+1} \right) = \left(\prod_{i=1}^{m} a_i \prod_{k=1}^{n} a_{m+k} \right) a_{m+n+1}$$
$$= \prod_{i=1}^{m+n} a_i \cdot a_{m+n+1} = \prod_{i=1}^{m+n+1} a_i,$$

womit (1) auch für die Zahl $n+1$ bewiesen ist.

Man kann ganz ähnlich auch das Produkt einer beliebigen endlichen Anzahl von Gruppenelementen mit einer anderen Verteilung der Klammern definieren und entsprechend seine Unabhängigkeit von der Verteilung der Klammern beweisen [5].

In einer kommutativen Gruppe G hängt das Produkt von n Elementen nicht von der Reihenfolge der Faktoren ab, genauer: Ist $f(i)$ eine beliebige eineindeutige Abbildung der Menge $1, 2, \ldots, n$ auf sich, so ist

$$\prod_{i=1}^{n} a_i = \prod_{i=1}^{n} a_{f(i)}. \tag{2}$$

Wir wollen hier den Beweisgedanken nur kurz skizzieren und seine genaue Durchführung dem Leser überlassen. Erstens beweist man unter Benutzung des kommutativen Gesetzes, daß sich ein Produkt aus n Elementen nicht ändert, wenn man zwei benachbarte Elemente miteinander vertauscht. Dabei ist die oben eingeführte, in der Schreibweise unberücksichtigt gelassene Klammerung genau zu beachten. Zweitens zeigt man, daß sich jede Vertauschung von irgend zwei Elementen auf eine gewisse Anzahl von Vertauschungen benachbarter Elemente zurückführen läßt. Drittens beweist man, daß jede Permutation der Faktoren auf eine gewisse Anzahl von Vertauschungen von je zwei Elementen zurückgeführt werden kann.

Folgerungen aus dem Gesetz der beiderseitigen Umkehrbarkeit. Wir bemerken zunächst, daß in der Formulierung der Bedingung II. nicht die Forderung nach der Existenz der Umkehroperationen der Gruppenmultiplikation enthalten ist, da II. nur die Existenz, aber nicht die Eindeutigkeit der Elemente c und d verlangt. Zum Beweis der Eindeutigkeit dieser Elemente führen wir den Begriff des Einselementes und den Begriff des inversen Elementes ein:

Definition 4. *Ein Element e einer Gruppe G heißt ein Einselement von G, wenn $ea = ae = a$ für alle a aus G gilt. Ist e ein Einselement von G, so nennt man ein Element a^{-1} aus G ein zum Gruppenelement a inverses Element (bezüglich e), wenn $aa^{-1} = a^{-1}a = e$.*

Satz 1. *In jeder Gruppe G gibt es ein und nur ein Einselement e. Zu jedem Gruppenelement a gibt es ein und nur ein inverses Element a^{-1}. Die nach dem Gesetz der beiderseitigen Umkehrbarkeit* II. *existierenden Lösungen der Gleichungen $ax = b$ und $ya = b$ sind für jedes a und b aus G eindeutig bestimmt.*

Beweis. e sei eine Lösung der Gleichung $yb = b$ für ein gewisses Element b aus G, also $eb = b$. Ferner sei c eine Lösung der Gleichung $bx = a$ für ein beliebiges a aus G, also $bc = a$. Dann ist

$$ea = e(bc) = (eb)c = bc = a.$$

Also gilt $ea = a$ für alle a aus G. Entsprechend zeigt man, daß es in G ein Element e' gibt, für das $ae' = a$ für alle a aus G gilt. Dann ist $e = ee' = e'$, d. h., e ist ein Einselement der Gruppe G. Sind nun e_1 und e_2 irgendwelche Einselemente von G, so ist $e_1 = e_1 e_2 = e_2$, womit die Eindeutigkeit des Einselementes e bewiesen ist.

Auf Grund des Gesetzes der beiderseitigen Umkehrbarkeit gibt es zu jedem a aus G Elemente b und c, für die $ba = e$ und $ac = e$ gilt. Dann ist

$$b = be = b(ac) = (ba)c = ec = c,$$

also $b = c$. Mithin besitzt das Element $b = a^{-1}$ die Eigenschaft $aa^{-1} = a^{-1}a = e$, d. h., es gibt ein zu a inverses Element. Sind nun b und c irgendwelche zu a inverse Elemente, so zeigt man genau wie oben, daß $b = bac = c$ ist, womit auch die Eindeutigkeit der inversen Elemente bewiesen ist.

Schließlich seien c_1 und c_2 beliebige Lösungen der Gleichung $ax = b$, es sei also $ac_1 = b$ und $ac_2 = b$, d. h. $ac_1 = ac_2$. Indem man diese letzte Gleichung von links mit a^{-1} multipliziert, erhält man $c_1 = c_2$. Genau so beweist man die Eindeutigkeit der Lösung der Gleichung $ya = b$.

Damit ist Satz 1 in allen Teilen bewiesen.

Wir vermerken noch, daß man aus der Existenz eines Einselementes und aus der Existenz von inversen Elementen unter Benutzung des assoziativen Gesetzes das Gesetz von der beiderseitigen Umkehrbarkeit in G beweisen kann. In der Tat ist dann $a^{-1}b$ eine Lösung der Gleichung $ax = b$ und ba^{-1} eine Lösung der Gleichung $ya = b$. Man kann also eine Gruppe auch als eine Menge definieren, in der eine assoziative Operation gegeben ist, bezüglich der ein Einselement und inverse Elemente vorhanden sind.

In Beispiel 1, den Gruppen von Zahlen bezüglich der Addition, ist Einselement in allen Fällen die Zahl 0 und das zu einer Zahl a inverse Element die Zahl $-a$. In Beispiel 2, den Gruppen von Zahlen bezüglich der Multiplikation, ist Einselement in allen Fällen die Zahl 1 und das zu einer Zahl a inverse Element die Zahl $\frac{1}{a}$. In Beispiel 3 ist Einselement das Element e und jedes der Elemente e und a zu sich selbst invers. In Beispiel 4 ist Einselement die identische Abbildung der Menge M auf sich und das zu einer Abbildung s inverse Element die inverse Abbildung s^{-1}.

Unter der *n-ten Potenz eines Gruppenelementes* a verstehen wir das Produkt aus den n gleichen Faktoren a. Sie wird durch a^n bezeichnet.

In dieser Form gilt die Definition zunächst nur für jede natürliche Zahl n. Im Falle $n = 0$ soll $a^0 = e$ sein, wobei e das Einselement der Gruppe G ist. Ferner soll für negative ganze $n = -m$ die Potenz $a^n = a^{-m}$ als $(a^{-1})^m$ oder auch als $(a^m)^{-1}$ erklärt sein. Beide Definitionen sind gleichwertig, da

$$a^m (a^{-1})^m = \underbrace{(aa\cdots a)}_{m\text{-mal}} \underbrace{(a^{-1}a^{-1}a^{-1}\cdots a^{-1})}_{m\text{-mal}} = e$$

ist, was $(a^{-1})^m = (a^m)^{-1}$ zur Folge hat.

Die Eigenschaft (1) des allgemeinen Produktes führt im Falle, daß alle Faktoren des Produktes gleich sind, auf die bekannte Potenzregel

$$a^m a^n = a^{m+n}. \tag{3}$$

Durch vollständige Induktion über n beweist man leicht:

$$(a^m)^n = a^{mn}. \tag{4}$$

In kommutativen Gruppen gilt außerdem auch noch die dritte Potenzregel

$$(ab)^n = a^n b^n, \tag{5}$$

die sich unmittelbar als Spezialfall der Vertauschungsregel (2) erweist. Wir haben bisher nur angegeben, wie man die Gleichungen (3), (4) und (5) für natürliche Zahlen m und n beweist. Jedoch gelten diese Gleichungen auch für beliebige ganze Zahlen m und n, wie man durch Nachprüfung aller Fälle $m \gtreqless 0$, $n \gtreqless 0$ bestätigt.

Aus der eindeutigen Lösbarkeit der Gleichungen $ax = b$ und $ya = b$ folgt die Existenz der beiden Umkehroperationen der Gruppenmultiplikation von G. Ist G eine kommutative Gruppe, so fallen die beiden Umkehroperationen zusammen. Ist nämlich c die Lösung der Gleichung $ax = b$, so ist $ac = b$. Dann ist aber auch $ca = b$, also c auch die (eindeutig bestimmte) Lösung der Gleichung $ya = b$.

Definition 5. *Die Umkehroperation der Multiplikation einer kommutativen Gruppe G heißt Division. Das den Elementen a und b vermöge der Division zugeordnete Element, d. h. die Lösung der Gleichungen $ax = b$ und $ya = b$, wird der Quotient der Elemente b und a genannt und durch $b : a$ oder $\frac{b}{a}$ bezeichnet.*

Die additive Schreibweise. Häufig bezeichnet man die Gruppenoperation auch durch $a + b$ und nennt sie dann *Addition*. In diesem Fall spricht man von der additiven Schreibweise der Gruppe. Meistens setzt man dann die Gruppe als kommutativ voraus.[1]) Bei der Verwendung der additiven Schreibweise spricht man an Stelle des Einselementes vom *Nullelement* 0 und an Stelle des inversen Elementes a^{-1} vom *entgegengesetzten Element* $-a$. Ferner spricht man an Stelle der Potenz a^n vom *Vielfachen* na (man darf dabei aber nicht na als Produkt von n und a auffassen, da die ganze Zahl n nicht Element von G zu sein braucht). Hierbei ist na nichts anderes als eine Abkürzung für

$$na = \underbrace{a + a + a + \cdots + a}_{n\text{-mal}}.$$

In einer additiv geschriebenen Gruppe bezeichnet man außerdem die allgemeine Summe durch

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

Schließlich verändert sich entsprechend auch die Form der Gleichungen (1) bis (5), wobei speziell die Gleichungen (3) bis (5) die Gestalt

$$(m + n)a = ma + na, \tag{6}$$

$$m(na) = (mn)a, \tag{7}$$

$$n(a + b) = na + nb \tag{8}$$

[1]) Die Gruppe wird dann oft als *Modul* bezeichnet. — *Anm. d. wissenschaftl. Red.*

annehmen. Die Umkehroperation einer additiv geschriebenen kommutativen Gruppe heißt *Subtraktion*, wobei das den Elementen a und b zugeordnete Element, also die Lösung der Gleichungen $a + x = b$ und $y + a = b$, die *Differenz* der Elemente b und a genannt und durch $b - a$ bezeichnet wird.

Untergruppen. Definition 6. *Eine Teilmenge H einer Gruppe G heißt eine Untergruppe von G, wenn sie bezüglich der Gruppenoperation von G selbst eine Gruppe ist.*

Um von einer gegebenen Teilmenge H einer Gruppe G festzustellen, ob sie Untergruppe von G ist, kann man das folgende Kriterium benutzen:

Satz 2. *Eine nicht-leere Teilmenge H einer Gruppe G ist dann und nur dann Untergruppe von G, wenn* 1.) *das Produkt irgendwelcher Elemente a, b aus H wieder zu H gehört und wenn* 2.) *mit einem Element a auch das zu a inverse Element a^{-1} zu H gehört.*

Beweis. Die Notwendigkeit dieser Bedingungen ist unmittelbar klar. Es sei also H eine beliebige nicht-leere Teilmenge von G, die die Bedingungen 1.) und 2.) erfüllt. Dann gibt es ein Element a in H, und es gehört wegen Bedingung 2.) auch a^{-1} und wegen Bedingung 1.) auch $aa^{-1} = e$ zu H. Somit besitzt H ein Einselement und außerdem zu jedem beliebigen Element a ein inverses. Da sich das assoziative Gesetz von G automatisch auf H überträgt, ist dann (nach einer früheren Bemerkung — *Anm. d. wissenschaftl. Red.*) H eine Gruppe und also eine Untergruppe von G.

Wir beschränken uns hier auf die angegebenen grundlegenden Eigenschaften der Gruppen und verweisen den interessierten Leser auf die Spezialliteratur ([6] und [7]).

§ 7. Ringe

Im vorangehenden Paragraphen haben wir die Eigenschaften einer einzelnen algebraischen Operation betrachtet. In Zahlenbereichen, mit denen wir uns später ausschließlich beschäftigen werden, hat man aber stets zwei Operationen — die Addition und die Multiplikation — zur Verfügung, die zudem durch das distributive Gesetz miteinander verknüpft sind. In diesem und dem folgenden Paragraphen wollen wir deshalb in ganz allgemeiner Form Mengen betrachten, in denen zwei Operationen mit gewissen Eigenschaften erklärt sind. Dabei beschränken wir uns auf den — für Zahlen allein interessierenden — Fall, daß die betrachteten Operationen kommutativ sind.

Definition 1. *Eine nicht-leere Menge R heißt ein Ring, wenn in ihr zwei algebraische Operationen erklärt sind, eine Addition — die jedem geordneten Paar a, b von Elementen aus R ein eindeutig bestimmtes Element $a + b$, die Summe aus a und b, zuordnet — und eine Multiplikation — die jedem geordneten Paar a, b von Elementen aus R ein eindeutig bestimmtes Element ab, das Produkt aus a und b, zuordnet —, die die folgenden Bedingungen erfüllen*:

I. (Kommutatives Gesetz der Addition) $a + b = b + a$:

II. (Assoziatives Gesetz der Addition) $a + (b + c) = (a + b) + c$;

III. (Gesetz der Umkehrbarkeit der Addition) *Für jedes a und b aus R besitzt die Gleichung $a + x = b$ (wenigstens) eine Lösung, d. h., zu jedem a und b aus R gibt es ein $c \in R$ mit $a + c = b$;*

IV. (Kommutatives Gesetz der Multiplikation)[1]) $ab = ba$;

V. (Assoziatives Gesetz der Multiplikation) $a(bc) = (ab)c$;

VI. (Distributives Gesetz der Multiplikation bezüglich der Addition) $(a + b)c = ac + bc$.

Beispiele für Ringe. Bezüglich der üblichen Addition und Multiplikation erweisen sich als Ringe:

1. Die Menge aller ganzen Zahlen.
2. Die Menge aller rationalen Zahlen.
3. Die Menge aller reellen Zahlen.
4. Die Menge aller komplexen Zahlen.
5. Die Menge, die nur aus der Zahl 0 besteht.
6. Die Menge aller geraden Zahlen und allgemein die Menge aller ganzen Zahlen, die Vielfache einer gewissen Zahl n sind.
7. Die Menge aller komplexen Zahlen der Form $a + bi$ mit ganzzahligem a und b (der sogenannte Ring der ganzen komplexen Zahlen).
8. Die Menge aller reellen Zahlen der Form $a + b\sqrt{2}$ mit ganzzahligem a und b.

Die Menge aller natürlichen Zahlen und auch die Menge aller positiven rationalen Zahlen bilden keinen Ring bezüglich der üblichen Addition und Multiplikation, da dort die Bedingung III. nicht erfüllt ist.

9. Eine große Rolle spielt in der Algebra der Ring der Polynome in einer oder auch mehreren Unbekannten mit Koeffizienten aus einem gewissen Ringe R.

Dabei werden als Addition und Multiplikation die üblichen — von der Schule her bekannten — Rechenoperationen für Polynome verwendet. Diese Rechenoperationen sind bereits über einem Ring ausführbar, da sie zurückgeführt werden auf die Addition und die Multiplikation der Koeffizienten der Polynome, die ihrerseits dem Ringe R angehören, in dem diese Operationen ausführbar sind.

10. Die Menge aller geordneten Paare (a, b) von ganzen Zahlen bildet einen Ring, sofern man die Operationen durch die Gleichungen

$$(a, b) + (c, d) = (a + c, b + d), \quad (a, b)\,(c, d) = (ac, bd)$$

definiert.

Der Leser möge für alle diese Beispiele nachprüfen, daß die Bedingungen I. bis VI. erfüllt sind.

Zunächst besitzen die Addition und die Multiplikation in einem Ring sicher alle die Eigenschaften, die wir im vorangehenden Paragraphen ganz allgemein aus dem assoziativen und dem kommutativen Gesetz gefolgert haben. Insbesondere kann man in einem Ring die Summe und das Produkt einer beliebigen endlichen Anzahl von Elementen definieren (vgl. § 6, Definition 3), für die dann die Rechenregeln gelten, die der Regel (1) aus § 6 entsprechen und bei denen zudem der Wert unabhängig von der Reihenfolge der gegebenen Elemente ist [§ 6, (2)].

Die Bedingungen I. bis III. besagen, daß jeder Ring eine kommutative Gruppe bezüglich der Addition bildet. Mithin gibt es in jedem Ring ein Element 0, das

[1]) In der Literatur wird die Bezeichnung „Ring" auch für Mengen mit nichtkommutativer und sogar nichtassoziativer Multiplikation verwendet. Auch die Formulierungen der anderen Bedingungen variieren von Fall zu Fall. Am Ende dieses Artikels werden wir beim Studium gewisser Verallgemeinerungen des Zahlbegriffs Ringe mit nichtkommutativer Multiplikation verwenden.

Nullelement oder kurz die Null des Ringes, für das

$$a+0=0+a=a$$

ist für jedes a. Ferner gibt es zu jedem a ein *entgegengesetztes Element* $-a$ mit

$$a+(-a)=(-a)+a=0.$$

Als Spezialfall der endlichen Summe bzw. des endlichen Produkts erhalten wir, wenn nämlich alle Summanden bzw. Faktoren gleich sind, das n-fache na und die n-te Potenz a^n eines Elementes a. Dabei ist zu beachten, daß man in einem Ring allgemein die Potenz a^n nur für natürliche Zahlen n definieren kann, da die Ausdehnung dieser Definition auf Zahlen $n \leq 0$ (vgl. § 6) maßgeblich die Existenz eines Einselementes und des inversen Elementes a^{-1} benutzte, die in einem Ring nicht vorhanden zu sein brauchen. Dementsprechend gelten in einem Ring die Potenzregeln (3) bis (5) aus § 6 nur für natürliche Exponenten. Dagegen bleiben auch in einem Ring (als einer Gruppe bezüglich der Addition) der Begriff des n-fachen na eines Elementes a und die Regeln (6) bis (8) aus § 6 für beliebige ganze Zahlen sinnvoll.

Aus den Gesetzen I. bis III. für die Addition folgt (wie in jeder kommutativen Gruppe), daß in jedem Ring eine Subtraktion als Umkehroperation der Addition existiert. Hingegen besitzt in einem Ring die Multiplikation im allgemeinen keine Umkehroperation, wie die Beispiele des Ringes der ganzen Zahlen und des Ringes der Polynome zeigen.

Folgerungen aus dem distributiven Gesetz. Bisher haben wir die beiden Operationen eines Ringes unabhängig voneinander auf ihre Eigenschaften hin untersucht. Wir gehen jetzt dazu über, ihren Zusammenhang zu studieren. Dieser Zusammenhang wird offensichtlich durch das distributive Gesetz VI. hergestellt.

Zunächst ergibt sich aus VI. und IV. die folgende zweite Form des distributiven Gesetzes:

$$a(b+c)=ab+ac.$$

Beide Formen des distributiven Gesetzes gelten auch für die Differenz, also:

$$(a-b)c=ac-bc, \quad a(b-c)=ab-ac. \tag{1}$$

Zum Beweis der ersten dieser beiden Gleichungen haben wir nur zu zeigen, daß das Element $(a-b)c$ der charakteristischen Eigenschaft der Differenz aus den Elementen ac und bc genügt. In der Tat ist

$$bc+(a-b)c=[b+(a-b)]c=ac.$$

Ganz analog beweist man die zweite Gleichung.

Wir zeigen jetzt, daß die Null eines Ringes bezüglich der Multiplikation eine wesentliche Eigenschaft mit der Zahl Null gemein hat:

Satz 1. *Ist ein Faktor eines Produktes gleich Null, so ist das ganze Produkt gleich Null, d. h., es ist*

$$a \cdot 0=0, \quad 0 \cdot a=0 \tag{2}$$

für jedes a.

Es genügt, die erste dieser Gleichungen zu beweisen, die zweite ergibt sich durch Anwendung von IV. aus der ersten: Nach Definition der Null und der Differenz

ist $0 = b - b$ für jedes Element b des Ringes. Dann ist aber

$$a \cdot 0 = a(b - b) = ab - ab = 0.$$

Hingegen ist die Umkehrung von Satz 1, die ja für Zahlen auch gilt, in beliebigen Ringen nicht erfüllt; mit anderen Worten: das Produkt zweier Elemente eines Ringes kann gleich Null sein, ohne daß einer der Faktoren gleich Null ist. So ist in dem oben als Beispiel 10 angeführten Ring aller Paare (a, b) von ganzen Zahlen offensichtlich das Paar $(0, 0)$ Nullelement. Sind also a und b irgendwelche von Null verschiedene ganze Zahlen, so sind die Paare $(a, 0)$ und $(0, b)$ beide von der Null des Ringes verschieden, hingegen ist $(a, 0)\,(0, b) = (0, 0)$.

Definition 2. *Elemente $a \neq 0$ und $b \neq 0$ eines Ringes, für die $ab = 0$ gilt, heißen Nullteiler. Ein Ring ohne Nullteiler wird Integritätsbereich genannt.*

Satz 2. *Ist $a \neq 0$ und a kein Nullteiler, so folgt aus $ab = ac$, daß $b = c$ ist.*

Beweis. Aus $ab = ac$ folgt, daß $ab - ac = 0$, also daß $a(b - c) = 0$. Da $a \neq 0$ und a kein Nullteiler ist, so muß $b - c = 0$, also $b = c$ sein.

Im folgenden werden wir es ausschließlich mit Ringen ohne Nullteiler zu tun haben. In ihnen folgt also aus $ab = ac$ stets $b = c$, sofern nur $a \neq 0$ ist.

Für die Multiplikation gelten ferner die üblichen Vorzeichenregeln[1]):

$$a(-b) = -ab, \quad (-a)\,b = -ab, \quad (-a)\,(-b) = ab. \tag{3}$$

Zum Beweis der ersten Gleichung benutzt man, daß

$$ab + a(-b) = a[b + (-b)] = a \cdot 0 = 0$$

ist, woraus sich unmittelbar $a(-b) = -ab$ ergibt. Die zweite Gleichung erhält man daraus gemäß

$$(-a)\,b = b\,(-a) = -ba = -ab.$$

Schließlich erhält man die dritte Gleichung aus den ersten beiden vermöge

$$(-a)\,(-b) = -(-a)\,b = -(-ab) = ab.$$

Durch vollständige Induktion kann man das distributive Gesetz auf Summen aus einer beliebigen endlichen Anzahl von Summanden und auch auf Produkte aus zwei Summen ausdehen. Man kommt dann zu folgenden Gleichungen:

$$\left(\sum_{i=1}^{n} a_i\right) b = \sum_{i=1}^{n} a_i b, \quad a \sum_{i=1}^{n} b_i = \sum_{i=1}^{n} a b_i, \quad \left(\sum_{i=1}^{n} a_i\right)\left(\sum_{k=1}^{m} b_k\right) = \sum_{i=1}^{n}\left(\sum_{k=1}^{m} a_i b_k\right). \tag{4}$$

Hieraus und aus den anderen Eigenschaften der Vielfachenbildung [§ 6, (7)] folgt, sofern man in den Summen alle Summanden gleich, also

$$a_i = a \quad (i = 1, 2, \ldots, n), \qquad b_k = b \quad (k = 1, 2, \ldots, m)$$

setzt:

$$\left.\begin{aligned} (na)\,b &= a\,(nb) = n\,(ab), \\ (na)\,(mb) &= n\,[m\,(ab)] = (nm)\,(ab). \end{aligned}\right\} \tag{5}$$

[1]) Wir bemerken, daß es hier wenig Sinn hat, die Termini „positives" und „negatives" Element von den Zahlen her zu übernehmen. Wir werden vielmehr diese Begriffe für beliebige Ringe erst in § 10 einführen. Solange die Elemente a und $-a$ noch völlig gleichberechtigt sind, ist jedes von ihnen dem anderen entgegengesetzt. Wenn wir also $-a$ mit b bezeichnen, so müssen wir notgedrungen a mit $-b$ bezeichnen.

In Kapitel IV benötigen wir noch die folgenden Eigenschaften der Differenz zwischen den Elementen eines Ringes:

Satz 3 (Eigenschaften der Differenz) *In jedem beliebigen Ring besitzt die Differenz von Elementen die folgenden Eigenschaften:*

a) *Es ist* $a - b = c - d$ *dann und nur dann, wenn* $a + d = b + c$ *ist;*

b) $(a - b) + (c - d) = (a + c) - (b + d)$;

c) $(a - b) - (c - d) = (a + d) - (b + c)$;

d) $(a - b)(c - d) = (ac + bd) - (ad + bc)$.

Beweis. Addiert man zur Gleichung $a - b = c - d$ auf beiden Seiten das Element $b + d$, so erhält man die Gleichung $a + d = b + c$. Umgekehrt erhält man aus dieser zweiten Gleichung die erste, wenn man auf beiden Seiten das Element $(-b) + (-d)$ hinzufügt. Damit ist a) bewiesen. Ähnlich beweist man die Gleichungen b), c) und d).

Unterringe. Definition 3. *Eine Teilmenge M eines Ringes R heißt ein Unterring, wenn sie bezüglich der Addition und der Multiplikation des Ringes R selbst ein Ring ist.*

In diesem Sinne ist der Ring aller geraden Zahlen ein Unterring des Ringes aller ganzen Zahlen und dieser seinerseits ein Unterring des Ringes aller rationalen Zahlen. Um festzustellen, ob eine gegebene Teilmenge eines gewissen Ringes ein Unterring desselben ist, ist es keineswegs nötig, die Gültigkeit aller Ringeigenschaften in der gegebenen Teilmenge nachzuprüfen. Vielmehr überträgt sich der größte Teil von ihnen automatisch von einem Ring auf jede seiner Teilmengen. Daher ist es zweckmäßiger, das folgende Kriterium zu benutzen:

Satz 4. *Damit eine nicht-leere Teilmenge M eines Ringes R ein Unterring von R ist, ist notwendig und hinreichend, daß die Summe, die Differenz und das Produkt von Elementen aus M wieder zu M gehören.*

Beweis. Zum Beweis der Notwendigkeit dieser Bedingungen sei M ein beliebiger Unterring von R. Die Addition in M fällt dann mit der Addition in R zusammen. Da die Umkehroperation der Addition in einem Ring eindeutig bestimmt ist, fällt auch die Subtraktion in M mit der Subtraktion in R zusammen. Dann müssen aber die Summe, die Differenz und das Produkt (im Sinne der Operationen von R) von Elementen aus M wieder zu M gehören, anderenfalls wäre eine von diesen Operationen für gewisse Elemente aus M nicht in M ausführbar, was der Definition des Ringes (siehe Definition 1) bzw. der aus ihr fließenden Ausführbarkeit der Subtraktion widersprechen würde.

Zum Beweis der Umkehrung sei M eine Menge, die den Bedingungen unseres Satzes genügt. Da die Summe und das Produkt (bezüglich R) von Elementen aus M wieder zu M gehören, kann man sie auch als Resultat einer gewissen Addition bzw. Multiplikation in M ansehen. Dadurch wird in der Menge M geradezu eine Addition und eine Multiplikation definiert. Die Eigenschaften I., II., IV., V. und VI. übertragen sich offenbar automatisch von R auf jede Teilmenge, gelten also sicher in M. Seien nun a und b beliebige Elemente aus M. Nach Voraussetzung gehört dann auch $b - a = c$ zu M. Aus den Eigenschaften der Differenz in R folgt

$$a + (b - a) = b \quad \text{oder} \quad a + c = b.$$

Mithin ist also auch die Bedingung III. in M erfüllt und M ein Unterring des Ringes R.

§ 8. Körper

Die im vorangehenden Paragraphen angeführten Beispiele für Ringe zeigen, daß die verschiedenen Ringe bezüglich der Umkehroperation der Multiplikation völlig verschiedene Eigenschaften besitzen (im Gegensatz zur Addition). So ist im Ring aller ganzen Zahlen die Division nur in Ausnahmefällen möglich, wobei jedoch alle Elemente des Ringes durch $+1$ und -1 teilbar sind. Im Ring aller rationalen Zahlen dagegen ist die Division (von der Division durch 0 abgesehen) stets möglich. Die Frage nach der Umkehroperation der Multiplikation führt uns auf den wichtigsten Spezialfall des Ringes, den Körper.

Definition 1. *Ein Ring K heißt ein Körper, wenn er die folgenden Eigenschaften besitzt:*

VII. (Umkehrbarkeit der Multiplikation) *Für jedes a und b aus K besitzt, sofern nur $a \neq 0$ ist, die Gleichung $ax = b$ (wenigstens) eine Lösung, d. h., zu jedem $a \neq 0$ und jedem b gibt es ein $q \in K$ mit $aq = b$.*

VIII. *K enthält wenigstens ein von Null verschiedenes Element.*

Beispiele von Körpern. Von den Beispielen 1 bis 10 für Ringe aus dem vorangehenden Paragraphen erweisen sich nur die in den Beispielen 2, 3 und 4 angeführten Ringe, also der Ring aller rationalen, der Ring aller reellen und der Ring aller komplexen Zahlen als Körper. In Beispiel 5 ist die Bedingung VIII. nicht erfüllt, weil kein Element $a \neq 0$ vorhanden ist. In den übrigen Beispielen ist die Bedingung VII. verletzt. Wir führen hier noch folgende Beispiele von Körpern an:

1. Die Menge aller komplexen Zahlen $a + bi$ mit rationalem a und b (der sogenannte Körper der rationalen komplexen Zahlen, vgl. § 7, Beispiel 7).

2. Die Menge aller reellen Zahlen der Form $a + b\sqrt{2}$ mit rationalem a und b (vgl. § 7, Beispiel 8).

3. Die Menge aller rationalen Funktionen mit reellen Koeffizienten in einer oder auch mehreren Veränderlichen.

4. Die Menge aus den beiden Elementen 0 und 1 bezüglich der folgendermaßen definierten Operationen:

$$0+0=1+1=0, \quad 0+1=1+0=1,$$
$$0\cdot 0=0\cdot 1=1\cdot 0=0, \quad 1\cdot 1=1.$$

Wir überlassen es wiederum dem Leser, die Bedingungen I. bis VIII. in allen Fällen nachzuprüfen.

Da die Körper spezielle Ringe sind, gelten alle in § 7 für Ringe abgeleiteten Sätze auch für Körper. Außerdem ergeben sich aus der Bedingung VII. Sätze über die Multiplikation, die wörtlich den in § 7 aus der Bedingung III. für die Addition abgeleiteten Sätzen entsprechen.

Wie jeder Ring, bildet auch jeder Körper eine kommutative Gruppe bezüglich der Addition. Darüber hinaus bilden die von Null verschiedenen Elemente eines Körpers auch eine Gruppe bezüglich der Multiplikation. In der Tat: Ist $a \neq 0$ und $b \neq 0$, so besitzt die Gleichung $ax = b$ eine Lösung q, die gleichfalls von Null verschieden ist, da $a0 = 0 \neq b$ ist (§ 7, Satz 1); damit folgt die Behauptung unmittelbar aus den Bedingungen IV., V. (§ 7, Definition 1) und VII. Die Gruppe aller Elemente eines Körpers bezüglich der Addition heißt die *additive Gruppe*, die Gruppe der von Null verschiedene Elemente eines Körpers bezüglich der Multiplikation die *multi-*

plikative Gruppe des Körpers. Ein Körper ist vollständig festgelegt durch Vorgabe dieser beiden Gruppen sowie der Produktè der Null mit allen anderen Elementen und der Forderung des distributiven Gesetzes für beliebige Elemente, einschließlich der Null. Hieraus folgt bereits, daß das Produkt eines beliebigen Elementes mit der Null gleich Null sein muß (§ 7, Satz 1).

Allein aus den Eigenschaften der multiplikativen Gruppe folgt bereits, daß es in jedem Körper ein Einselement gibt (§ 6, Satz 1), also ein Element e mit $ae = ea = a$ für alle a aus K. Für die Elemente $a \neq 0$ folgt dies aus den Eigenschaften des Einselementes einer Gruppe und für $a = 0$ aus den Eigenschaften der Null bezüglich der Multiplikation. Ferner gibt es zu jedem $a \neq 0$ ein inverses Element a^{-1} mit $aa^{-1} = a^{-1}a = e$.

Hierbei sind sowohl das Einselement e als auch das inverse Element a^{-1} eines vorgegebenen Elementes a eindeutig bestimmt. Wenn es nämlich in einem Ring überhaupt ein Einselement gibt, so auch nur eines, was sich sofort aus $e_1 = e_1e_2 = e_2$ für Einselemente e_1, e_2 ergibt. Entsprechend: Wenn es zu einem Element a eines Ringes mit Einselement ein inverses Element gibt, so auch nur eines; denn sind b und c inverse Elemente zu a, so ist $b = bac = c$.

Jedoch braucht es in einem Ring mit Einselement keine inversen Elemente zu geben, wie das Beispiel des Ringes der ganzen Zahlen zeigt. Ebenso gibt es auch Ringe ohne Einselement, wie den Ring aller geraden Zahlen oder den Ring aller ganzen Zahlen, die Vielfache einer Zahl $n > 1$ sind.

Wenn es in einem Ring R ein Einselement $e \neq 0$ und zu jedem $a \neq 0$ ein inverses Element a^{-1} gibt, so bildet die Menge der von der Null des Ringes verschiedenen Elemente des Ringes eine Gruppe bezüglich der Multiplikation (§ 6), und der Ring R ist ein Körper.

Da die multiplikative Gruppe eines Körpers kommutativ ist, besitzt die Multiplikation eine eindeutig bestimmte Umkehroperation, die Division. Dabei ist der Quotient $\frac{b}{a}$ für beliebiges $a \neq 0$ und beliebiges b eindeutig bestimmt. Für $b \neq 0$ folgt dies aus den allgemeinen Eigenschaften der multiplikativen Gruppe des Körpers (§ 6), und für $b = 0$ ist $\frac{0}{a} = 0$, da $a0 = 0$ ist. Die Forderung $a \neq 0$, die in die Bedingung VII. eingeht, verletzt in einem gewissen Sinne die Symmetrie in den Eigenschaften der Addition und der Multiplikation. Es ist jedoch unmöglich, die genannte Symmetrie etwa dadurch herzustellen, daß man die Forderung $a \neq 0$ fallen läßt. In der Tat besitzt die Gleichung $ax = b$ für $a = 0$ und $b \neq 0$ in keinem Körper und auch in keinem Ring mit wenigstens einem von Null verschiedenen Element eine Lösung. Wäre nämlich q eine Lösung dieser Gleichung, so wäre $aq = 0q = 0 = b$, was der Voraussetzung $b \neq 0$ widerspricht. Daher ist die Division durch Null nicht ausführbar, wenn der Dividend von Null verschieden ist. Hingegen kann der Quotient $\frac{0}{0}$ jedes beliebige Element des betrachteten Ringes sein, da für jedes q gilt: $0q = 0$.

Satz 1. *In einem Körper gibt es keine Nullteiler* (§ 7, Definition 2), *d. h., wenn* $ab = 0$ *ist, so ist* $a = 0$ *oder* $b = 0$.

Beweis. Aus $ab = 0$ und $a \neq 0$ folgt, wenn man beide Seiten der Gleichung mit a^{-1} multipliziert, $1 \cdot b = a^{-1}0$, also $b = 0$.

Mithin ist jeder Körper ein Ring ohne Nullteiler. Die Umkehrung dieser Behauptung gilt indes im allgemeinen nicht: Es gibt Ringe ohne Nullteiler, die keine Körper sind (z. B. der Ring der ganzen Zahlen). Sie gilt jedoch ganz allgemein für endliche Ringe, genauer:

Satz 2. *Jeder endliche Ring ohne Nullteiler, der mindestens zwei Elemente enthält, ist ein Körper.*

Beweis. Offenbar genügt es, die Bedingung VII. zu verifizieren. Sei also $a \neq 0$. Dann ordnen wir jedem Element x des Ringes das Element $y = ax$ zu. Ist $x_1 \neq x_2$, so ist auch $y_1 \neq y_2$, da andernfalls $ax_1 = ax_2$ und mithin $x_1 = x_2$ wäre (§ 7, Satz 2). Also ist die Zuordnung $x \to y$ eine eineindeutige Abbildung des ganzen Ringes R auf eine gewisse Teilmenge M, also $R \sim M$. Da nach Satz 1 aus § 4 eine endliche Menge R keiner ihrer echten Teilmengen gleichmächtig ist, muß also $R = M$ sein, d. h., zu jedem Element $b \in R$ gibt es in R ein Element q mit $q \to b$, also $aq = b$, womit VII. bereits bewiesen ist.

Da die von Null verschiedenen Elemente eines Körpers bezüglich der Multiplikation eine kommutative Gruppe bilden, ist in einem Körper für jedes Element $a \neq 0$ die Potenz a^n für beliebige ganze Exponenten n definiert, wobei die üblichen Potenzgesetze gelten [§ 6, (3) — (5)].

Für die Quotienten aus Elementen eines beliebigen Körpers gelten dieselben Rechenregeln, wie in der üblichen Bruchrechnung. In Kapitel V. benötigen wir speziell die folgenden Eigenschaften des Quotienten:

Satz 3. (Eigenschaften des Quotienten) a) *Wenn* $b \neq 0$ *und* $d \neq 0$ *ist, so gilt* $\frac{a}{b} = \frac{c}{d}$ *dann und nur dann, wenn* $ad = bc$ *ist.*

b) *Wenn* $b \neq 0$, $d \neq 0$ *ist, so ist* $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$;

c) *Wenn* $b \neq 0$, $d \neq 0$ *ist, so ist* $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$;

d) *Wenn* $b \neq 0$, $c \neq 0$, $d \neq 0$ *ist, so ist* $\frac{a}{b} : \frac{c}{d} = \frac{ad}{bc}$.

Beweis. Wenn man beide Seiten der Gleichung $\frac{a}{b} = \frac{c}{d}$ mit bd multipliziert, so erhält man $ad = bc$. Ist umgekehrt eine Gleichung $ad = bc$ mit $b \neq 0$ und $d \neq 0$ vorgegeben, so erhält man, wenn man $\frac{a}{b} = x$ und $\frac{c}{d} = y$ setzt: $bdx = ad$, $bdy = bc$, also $bdx = bdy$. Indem man beide Seiten dieser letzten Gleichung zunächst mit b^{-1} und dann mit d^{-1} multipliziert, erhält man $x = y$, also $\frac{a}{b} = \frac{c}{d}$.

Damit ist die Behauptung a) bereits bewiesen. Die Behauptungen b) und c) beweist man ähnlich wie den zweiten Teil der Behauptung a). Zum Beweis der Behauptung d) braucht man sich offenbar nur davon zu überzeugen, daß

$$\frac{a}{b} = \frac{c}{d} \cdot \frac{ad}{bc}$$

ist, was aber bereits aus c) und a) folgt. Damit ist dann Satz 3 in allen Teilen bewiesen.

Die Charakteristik eines Körpers. In manchen Körpern gibt es Elemente $a \neq 0$, für die $na = 0$ für gewisse ganze Zahlen $n \neq 0$ gilt. So ist im Körper aus den beiden Elementen 0 und e (siehe Beispiel 4 am Anfang dieses Paragraphen) $2e = e + e = 0$. Wesentlich ist nun der folgende

Satz 4. *In jedem Körper K tritt genau einer der beiden nachstehenden Fälle ein*:

a) *Für jedes Element $a \neq 0$ und jede ganze Zahl $n \neq 0$ ist das Vielfache na ebenfalls von Null verschieden.*

b) *Es gibt eine eindeutig bestimmte Primzahl*[1] *p derart, daß $pa = 0$ ist für jedes Element a.*

Beweis. Wir zeigen, daß, sofern der Fall a) nicht eintritt, sofern es also ein Element $a \neq 0$ und eine ganze Zahl $n \neq 0$ mit $na = 0$ gibt, der Fall b) eintreten muß. Wegen $a \neq 0$ gibt es zu jedem $b \in K$ ein q mit $aq = b$. Dann ist auf Grund von Gleichung (5) aus § 7

$$nb = n(aq) = (na)q = 0 \cdot q = 0.$$

Es genügt also zu zeigen, daß der Fall b) für irgendein Element $a \neq 0$, also etwa für das Einselement e, eintritt.

Wie wir gesehen haben, ist $ne = 0$, also auch $(-n)e = -ne = 0$. Eine der Zahlen $n, -n$ ist aber positiv, es gibt also eine natürliche Zahl k mit $ke = 0$. Sei nun p die kleinste derartige natürliche Zahl.[2])

Wir zeigen, daß p notwendig eine Primzahl ist. Zunächst ist sicher $p \neq 1$, da $1 \cdot e = e \neq 0$ ist, während $pe = 0$ sein solte. Angenommen nun, es wäre p durch eine Zahl q teilbar, also $p = qr$, wobei $1 < q < p$ und damit auch $1 < r < p$ ist. Dann wäre auf Grund von Gleichung (5) aus § 7

$$pe = (qr)(ee) = (qe)(re) = 0.$$

Da es nun aber in einem Körper keine Nullteiler gibt (Satz 1), wäre dann $qe = 0$ oder $re = 0$ im Widerspruch dazu, daß p die kleinste natürliche Zahl sein sollte, für die $pe = 0$ ist.

Sei schließlich k eine beliebige natürliche Zahl mit $ke = 0$. Wir können dann k durch p mit Rest dividieren, also k in der Form $k = pq + r$ darstellen, wobei der Rest r der Bedingung $0 \leq r < p$ genügt. Dann erhalten wir aus Gleichung (6) aus § 6 und Gleichung (5) aus § 7:

$$ke = (pq + r)e = (pq)e + re = q(pe) + re = 0 + re = re = 0.$$

Dann muß aber $r = 0$ sein, da $r > 0$ der Wahl von p widerspräche. Also ist $k = pq$, d. h. k durch p teilbar. Es kann also, wenn k von p verschieden ist, k keine Primzahl sein. Damit ist gezeigt, daß p die einzige Primzahl mit $pe = 0$ ist.

Dieser letzte Satz führt uns auf die folgende

Definition 2. *Unter der Charakteristik des Körpers K verstehen wir die Zahl* 0, *falls für jedes Element $a \neq 0$ des Körpers K und jede ganze Zahl $n \neq 0$ auch $na \neq 0$ ist, und die Primzahl p, falls $pa = 0$ für alle Elemente a aus K gilt.*

Da für die Zahl 1 und jede ganze Zahl n die Gleichung $1 \cdot n = n$ gilt, haben alle Zahlkörper die Charakteristik 0.

Beispiel eines Körpers der Charakteristik $p > 0$. Es sei n eine beliebige natürliche Zahl größer als 1. Dann können wir die ganzen Zahlen dadurch in Klassen einteilen, daß wir alle Zahlen, die bei Division durch n denselben Rest lassen, in

[1]) Eine Primzahl ist eine von 1 verschiedene natürliche Zahl, die nur durch 1 und sich selbst teilbar ist.

[2]) Daß jede nichtleere Menge von natürlichen Zahlen eine kleinste Zahl enthält, wird in Kapitel III bewiesen.

einer Klasse zusammenfassen. Bezeichnen wir mit (r) die Klasse der Zahlen, die bei Division durch n den Rest r ergeben, so kommen wir auf die n paarweise verschiedenen Klassen (0), (1), (2), ..., $(n-1)$. Offenbar gehören Zahlen a und b dann und nur dann derselben Klasse an, wenn ihre Differenz $a-b$ durch n teilbar ist.[1]) Ferner verstehen wir unter C_n die Menge der so definierten Klassen. Wir führen nun in C_n eine Addition und eine Multiplikation ein. Sind (r) und (s) derartige Klassen aus C_n, wobei etwa die Klasse (r) die Zahl a und die Klasse (s) die Zahl b enthalten möge, so verstehen wir unter der Summe $(r)+(s)$ der gegebenen Klassen diejenige Klasse, die die Zahl $a+b$ enthält und unter dem Produkt $(r)\cdot(s)$ die Klasse, die die Zahl ab enthält. Die Summe und das Produkt von Klassen sind eindeutig bestimmt, d. h. unabhängig von der Wahl der Repräsentanten a und b aus den betreffenden Klassen. In der Tat: Sind a und a' Zahlen aus der Klasse (r) und b und b' Zahlen aus der Klasse (s), so sind die Zahlen $a-a'$ und $b-b'$ durch n teilbar; dann sind aber auch

$$(a+b)-(a'+b')=(a-a')+(b-b')$$

und

$$ab-a'b'=(ab-a'b)+(a'b-a'b')=(a-a')\,b+a'(b-b')$$

durch n teilbar, d. h., auch die Zahlen $a+b$ und $a'+b'$ sowie ab und $a'b'$ gehören jeweils derselben Klasse an.

Die Ringeigenschaften I. bis VI. aus § 7, Definition 1, sind automatisch erfüllt, da sie für die ganzen Zahlen gelten und die Operationen zwischen den Klassen mit Hilfe der entsprechenden Operationen zwischen den Repräsentanten definiert sind. Also ist C_n unter allen Umständen ein Ring. Er heißt der Restklassenring nach dem Modul n (auch: Restklassenring modulo n). Die Null des Ringes C_n ist offensichtlich die Klasse (0) der durch n teilbaren Zahlen.

Ist die Zahl $n=kl$ zusammengesetzt, so enthält der Ring C_n Nullteiler, da zwar $(k)\neq 0$, $(l)\neq 0$, aber $(k)\cdot(l)=(0)$ ist. Ist hingegen $n=p$ eine Primzahl, so enthält der Ring C_n keine Nullteiler; denn ist $(r)\cdot(s)=0$, so ist rs durch p teilbar, also r oder s durch p teilbar und mithin $(r)=0$ oder $(s)=0$.

Weiter enthält der Ring C_p genau p Elemente, ist also endlich. Auf Grund von Satz 2 ist er dann aber ein Körper. Da für jedes r die Klasse $p(r)$ die Zahl pr enthält, die durch p teilbar ist, ist $p(r)=0$ für jede Klasse (r) des Körpers C_p.

Damit ist bewiesen, daß C_p ein Körper der Charakteristik p ist.

Unterkörper, Primkörper. Definition 3. *Eine Teilmenge M eines Körpers K heißt ein Unterkörper (oder Teilkörper) von K, wenn sie bezüglich der Addition und der Multiplikation des Körpers K einen Körper bildet. Dann heißt K auch ein Oberkörper oder eine Erweiterung des Körpers M.*

In diesem Sinne ist der Körper der rationalen Zahlen ein Unterkörper des Körpers der reellen Zahlen und dieser seinerseits ein Unterkörper des Körpers der komplexen Zahlen.

Satz 5. *Damit eine Teilmenge M eines Körpers K, die wenigstens zwei Elemente enthält, ein Unterkörper von K ist, ist notwendig und hinreichend, daß die Summe, das Produkt, die Differenz und der Quotient (soweit er in K erklärt ist) von Elementen aus M wieder zu M gehört.*

[1]) Wir haben es hier also mit nichts anderem als mit den Restklassen nach dem Modul n zu tun (vgl. den Artikel von A. J. Chintschin in diesem Buche).

Der Beweis hierfür verläuft dem Beweis des entsprechenden Satzes für Ringe (§ 7, Satz 4) völlig analog. Wir verzichten deshalb auf seine Wiedergabe.

Jeder Unterkörper M eines Körpers K enthält die Null als Differenz $a - a$ für $a \in M$ und das Einselement als Quotient $\frac{a}{a}$ für $a \in M$ mit $a \neq 0$.

Satz 6.[1]) *Der Durchschnitt* (im mengentheoretischen Sinn, § 2) *eines beliebigen Systems von Unterkörpern des Körpers K ist wieder ein Unterkörper von K.*

Beweis. Es sei $\{M_s\}$ ein beliebiges System von Unterkörpern, wobei die Indizes s eine gewisse Menge S durchlaufen, und $D = \bigcap_{s \in S} M_s$ der Durchschnitt der Körper M_s des gegebenen Systems. Da 0 und 1 in jedem der Unterkörper M_s enthalten sind, sind sie auch in D enthalten, d. h., D enthält wenigstens zwei Elemente. Sind nun a und b irgendwelche Elemente aus D, so sind sie in jedem M_s enthalten. Dann sind auf Grund von Satz 5 aber auch $a + b$, $a - b$, ab und, falls $b \neq 0$ ist, auch $\frac{a}{b}$ in jedem M_s und damit in D enthalten. Dann ist aber auf Grund von Satz 5 auch D ein Unterkörper von K.

Definition 4. *Ein Körper heißt Primkörper, wenn er keinen Unterkörper außer sich selbst enthält.*

Als Beispiele für Primkörper können der Körper der rationalen Zahlen und die Restklassenkörper nach einer Primzahl dienen:

Ein beliebiger Unterkörper M des Körpers K der rationalen Zahlen enthält nämlich sicher die Zahl 1, also auch alle ihre Vielfachen $n \cdot 1 = n$, d. h. alle ganzen Zahlen und damit auch alle Quotienten mit von Null verschiedenem Nenner, d. h. alle rationalen Zahlen. Also ist $K = M$ und K Primkörper. Entsprechend enthält jeder Unterkörper M des Restklassenkörpers C_p die Restklasse (1), die das Einselement von C_p ist, und damit auch jede Klasse (r) als r-faches der Klasse (1). Also ist auch hier $C_p = M$ und C_p Primkörper.

Man kann beweisen, daß in einem gewissen, näher zu präzisierenden Sinne die genannten Körper auch alle Primkörper erschöpfen.

Satz 7. *Jeder Körper enthält genau einen Primkörper.*

Beweis. Zunächst enthält jeder Körper K mindestens einen Unterkörper (z. B. K selbst). Es sei nun D der Durchschnitt aller Unterkörper des Körpers K. Nach Satz 6 ist D Unterkörper von K und auf Grund seiner Definition in jedem Unterkörper von K enthalten. Wäre nun M ein beliebiger, von D verschiedener Unterkörper von D, dann wäre auf Grund von Definition 3 offenbar M auch Unterkörper von K, also D in M enthalten, was unmöglich ist. Also ist D ein Primkörper von K. Ist nun D' ein beliebiger Primkörper von K, so ist der Durchschnitt $D'' = D \cap D'$ ebenfalls ein Unterkörper des Körpers K, wobei $D'' \subseteq D$ und $D'' \subseteq D'$ ist. Auf Grund von Definition 3 ist dann aber D'' Unterkörper sowohl von D als auch von D', also ist — da D und D' Primkörper sind — $D = D'' = D'$, womit die Eindeutigkeit des Primkörpers bewiesen ist.

---

[1]) Der entsprechende Satz gilt auch für Ringe, d. h., auch der Durchschnitt eines beliebigen Systems von Unterringen eines Ringes R ist ein Unterring von R. Der Beweis hierfür verläuft dem oben für Körper geführten völlig analog und sei deshalb dem Leser überlassen.

§ 9. Der axiomatische Aufbau der Mathematik. Isomorphie

In den einzelnen mathematischen Theorien untersucht man Mengen von Objekten, zwischen denen gewisse Beziehungen erklärt sind, die diese oder jene Eigenschaften besitzen. Der Inhalt einer solchen Theorie besteht dann in der Definition von Beziehungen (oder Begriffen) durch andere und im Beweis von Eigenschaften dieser Beziehungen (oder Begriffe) auf Grund anderer Eigenschaften. So kann man in der Theorie der geordneten Mengen jede der Beziehungen „größer" und „kleiner" durch die andere definieren oder mit ihrer Hilfe den Begriff „erstes Element" einführen (§ 5), so kann man in der Theorie der Ringe durch die Beziehung $a + b = c$ die Beziehung $a - b = c$ definieren und mit ihrer Hilfe den Begriff „Null" einführen usw.

Es ist unmittelbar klar, daß man nicht alle Begriffe und Beziehungen definieren kann, wie man auch nicht alle ihre Eigenschaften allein mit Hilfe der Logik beweisen kann: Jede Definition führt den zu definierenden Begriff auf andere Begriffe zurück, jeder Beweis einer gegebenen Eigenschaft benutzt irgendwelche andere Eigenschaften. Deshalb muß man in jeder mathematischen Theorie gewisse Beziehungen (oder Begriffe) ohne Definition an die Spitze stellen; das sind die sogenannten *Grundbeziehungen* oder *Grundbegriffe* (der betr. Theorie). Genau so muß man gewisse Eigenschaften dieser Grundbeziehungen ohne Beweis voraussetzen; das sind die sogenannten *Grundeigenschaften* oder *Axiome* (genau genommen sind die Axiome die Sätze, in denen die Grundeigenschaften formuliert werden — *Anm. d. wissenschaftl. Red.*). Die Aufstellung der Grundbegriffe und der Axiome bildet das Fundament jeder mathematischen Theorie. Aus ihnen wird alles weitere auf rein logischem Wege gefolgert.

Eine Eigentümlichkeit des modernen Aufbaus der mathematischen Wissenschaften, die ihnen ihren abstrakten Charakter verleiht, ist das Studium der uns interessierenden Begriffe und Eigenschaften in Anwendung auf beliebige Mengen, auf denen sie definiert sind. Dabei ist es für die betrachtete Theorie völlig gleichgültig, welche konkrete Bedeutung die Elemente der jeweils betrachteten Mengen haben und was für konkrete Eigenschaften sie (abgesehen von den in der betreffenden mathematischen Theorie untersuchten) besitzen. So haben wir in den drei vorangehenden Paragraphen Gruppen, Ringe und Körper als Mengen von Elementen mit gewissen Operationen (Addition und Multiplikation) definiert, die einige grundlegende Eigenschaften besitzen. Entsprechend steht es auch z. B. mit dem axiomatischen Aufbau der Geometrie (vgl. [8] und [9]), wobei die Punkte, Geraden und Ebenen irgendwelche Objekte sind, deren Natur für den formalen Aufbau der Geometrie völlig gleichgültig ist; es sind lediglich zwischen ihnen einige grundlegende Beziehungen („ein Punkt liegt auf einer Geraden" usw.) erklärt, die gewissen grundlegenden Bedingungen (den Axiomen der Geometrie) genügen.

Es könnte nun scheinen, als ob es viele Theorien der Ringe und Körper und viele Geometrien gebe, je nachdem, welche konkrete Menge man als Grundlage der jeweiligen Theorie nimmt. Der Ausweg aus dieser scheinbaren Schwierigkeit ergibt sich aber bereits aus dem oben Gesagten und besteht in einer genauen Festlegung des Inhalts einer gegebenen mathematischen Theorie. In einer vorgelegten Theorie studiert man — wie bereits gesagt — nicht alle Eigenschaften der Elemente einer Menge, sondern nur diejenigen, die auf den Grundbeziehungen der vorgelegten Theorie beruhen, soweit sie aus den Axiomen für die Grundbeziehungen folgen. Alle übrigen Eigenschaften (die an sich natürlich von größter Wichtigkeit sein können) gehören einfach nicht zum Gegenstand der Untersuchungen der vor-

liegenden Theorie, von ihnen wird im betrachteten Zusammenhang abstrahiert. Es wird also vom Standpunkt der betreffenden Theorie aus nicht zwischen solchen Mengen unterschieden, für deren Elemente gewisse Grundbeziehungen (die jeweils von den konkreten Eigenschaften der Elemente der betreffenden Menge abhängen können) erklärt sind, die mit den Beziehungen der betreffenden Theorie die Grundeigenschaften gemein haben.

Da jedoch im konkreten Spezialfall die Grundbeziehungen in einer Menge ausgehend von den konkreten Eigenschaften ihrer Elemente definiert werden, behandelt somit eine jede Theorie, auch wenn sie die Eigenschaften der Grundbeziehungen in ganz abstrakter Form studiert, gewisse konkrete Eigenschaften einer ganzen Klasse von konkreten Mengen. Diese dialektische Einheit von Abstraktem und Konkretem ist jeder Wissenschaft eigen, aber wohl in der Mathematik am deutlichsten ausgeprägt. Schließlich studiert ja die Mathematik nicht alle Eigenschaften der materiellen Körper, sondern nur diejenigen, die eine anzahlmäßige Beurteilung oder eine räumliche Beschreibung gestatten. Namentlich sind die für die ganze Mathematik grundlegenden Begriffe der Zahl und der Figur solche abstrakten Ausdrücke für derartige Eigenschaften materieller Körper. Somit bleibt auch für die moderne Mathematik, ungeachtet des abstrakten Charakters ihres Aufbaus, die von ENGELS für die Mathematik gegebene Definition gültig: „Die reine Mathematik hat zum Gegenstand die Raumformen und Quantitätsverhältnisse der wirklichen Welt, also einen sehr realen Stoff. Daß dieser Stoff in einer höchst abstrakten Form erscheint, kann seinen Ursprung aus der Außenwelt nur oberflächlich verdecken."[1] Die Tatsache, daß Mengen, zwischen deren Elementen Beziehungen mit gleichartigen Eigenschaften erklärt sind, im Rahmen einer vorgelegten mathematischen Theorie nicht unterschieden werden, erfährt ihre Präzisierung in dem nachfolgenden allgemeinen Isomorphiebegriff:

Definition 1. *Mengen M und M', in denen gewisse Beziehungen erklärt sind, die in einem System S von Beziehungen zusammengefaßt sein mögen, heißen isomorph bezüglich des gegebenen Systems von Beziehungen (kurz: isomorph; in Zeichen: $M \cong M'$) wenn es eine eineindeutige Zuordnung zwischen ihren Elementen gibt, bei der die Beziehungen aus S erhalten bleiben, also eine eineindeutige Abbildung, bei der Elemente aus M, die in einer gewissen Beziehung aus S stehen, in Elemente aus M' übergehen, die in derselben Beziehung stehen, und umgekehrt.*

Man kann dann sagen, daß die in einer axiomatisch aufgebauten Theorie behandelten Mengen nur bis auf Isomorphie bezüglich des Systems der Grundbeziehungen der betreffenden Theorie festgelegt sind.

Der Isomorphiebegriff besitzt offenbar die drei folgenden grundlegenden Eigenschaften:

1. $M \cong M$,
2. *Wenn* $M \cong M'$ *ist, so ist* $M' \cong M$,
3. *Wenn* $M \cong M'$ *und* $M' \cong M''$ *ist, so ist* $M \cong M''$.

Beispielsweise führt die Definition 1 im Falle, daß keine Beziehungen vorgegeben sind (also das System S die leere Menge ist), auf die Gleichmächtigkeit der Mengen M und M' (§ 3). Ist als einzige Beziehung die Beziehung „a geht b voran" vorgegeben,

[1] F. ENGELS, Anti-Dühring, Berlin 1952, S. 44/45.

so führt Definition 1 für den Fall, daß diese Beziehung die entsprechenden Axiome erfüllt, auf die Ähnlichkeit (§ 5).

Die Tatsache, daß der Begriff der Isomorphie die Gleichwertigkeit zweier Mengen bezüglich der betrachteten Eigenschaften aufdeckt, kann in Form des folgenden allgemeinen Satzes formuliert werden:

Sind die Mengen M und M' bezüglich eines gewissen Systems S von Beziehungen isomorph, so überträgt sich jede Eigenschaft der Menge M, die mit Hilfe von Beziehungen aus S (oder auch mit Hilfe von Beziehungen, die durch Beziehungen aus S definiert sind) formuliert ist, auf die Menge M', und umgekehrt.

Wir wollen diesen allgemeinen Satz an Hand eines konkreten Beispiels untersuchen:

Es sei in den Mengen M und M' eine Beziehung „größer" definiert, bezüglich der die Mengen M und M' isomorph sind. Dann ist, wenn M geordnet ist, wenn also in M die Bedingungen 1). und 2.) aus § 5 erfüllt sind, auch die Menge M' geordnet.

Beweis der Bedingung 1.). Es seien a' und b' Elemente aus M' und a und b die entsprechenden Elemente aus M. Da die Bedingung 1.) nach Voraussetzung in M erfüllt ist, gilt in M genau eine der Beziehungen $a = b$, $a > b$ oder $b > a$. Die Abbildung von M auf M' erhält nun die Beziehung „größer", d. h., es gilt mindestens eine der Beziehungen $a' = b'$, $a' > b'$ oder $b' > a'$. Würde nun in M' mehr als eine dieser Beziehungen gelten, so würde, da auch die Abbildung von M' auf M die Beziehung „größer" erhält, zwischen a und b mehr als eine der Beziehungen gelten, was der Bedingung 1.) für M widerspricht.

Beweis der Bedingung 2.). Ist $a' > b'$ und $b' > c'$, so ist auch $a > b$ und $b > c$. In M ist dann $a > c$, also ist auch $a' > c'$.

Wir wollen jetzt die Isomorphie von Gruppen, Ringen und Körpern betrachten. Da hier die Beziehungen $a + b = c$ und $ab = c$ der zusätzlichen Bedingung genügen, daß es zu jedem a und b genau ein c mit $a + b = c$ bzw. $ab = c$ gibt (diese beiden Bedingungen sind — genau genommen — zwei ergänzende Axiome), kann hier die Definition der Isomorphie gegenüber der Definition 1 etwas vereinfacht werden. Es braucht hier — kurz gesagt — nur verlangt zu werden, daß die Grundbeziehungen beim Übergang *von M zu M'* erhalten bleiben. Wir wollen uns dabei auf den Fall der Ringe und Körper beschränken, den wir bei den späteren Untersuchungen der Zahlbereiche nur benötigen (der Fall der Gruppen unterscheidet sich von dem behandelten nur darin, daß an Stelle der beiden Operationen eine einzige tritt). Wir definieren also:

Definition 2. *Ein Ring (oder Körper) R heißt isomorph dem Ring (bzw. Körper) R' (in Zeichen: $R \cong R'$), wenn es eine eineindeutige Abbildung von R auf R' gibt, bei der der Summe und dem Produkt von Elementen aus R die Summe und das Produkt der entsprechenden Elemente aus R' zugeordnet sind,*

Wir wollen uns überlegen, daß diese Definition tatsächlich ein Spezialfall der Definition 1 ist. Dazu brauchen wir uns nur davon zu überzeugen, daß die inverse Abbildung von R' auf R ebenfalls Summe und Produkt erhält. In R' gelte also $a' + b' = c'$, wobei den Elementen a', b', c' vermöge der inversen Abbildung die Elemente a, b, c aus R zugeordnet seien. Es ist zu zeigen, daß $a + b = c$ ist. Angenommen, es wäre $a + b = d \neq c$. Dann würde auf Grund von Definition 2 gelten: $a' + b' = d' \neq c'$, was der Eindeutigkeit der Addition in R' widersprechen würde.

Wir sehen, daß wir bei diesem Beweis in keiner Weise von den Axiomen I. bis VI. des Ringes Gebrauch gemacht haben. Aus diesem Grunde übertragen sich Definition 2 und die anschließenden Überlegungen wörtlich auf beliebige Mengen, in denen je zwei (oder auch mehr — *Anm. d. wissenschaftl. Red.*) algebraische Operationen, eine Addition und eine Multiplikation, vorgegeben sind.

Satz 1. *Es seien R und R' beliebige Mengen, in denen zwei Operationen (Addition und Multiplikation) definiert sind, bezüglich derer R und R' isomorph* (im Sinne von Definition 2) *sind. R ist dann und nur dann ein Ring (oder Körper), wenn R' Ring (bzw. Körper) ist.*

Beweis. Es genügt offenbar zu zeigen, daß R' den Axiomen I. bis VI. bzw. I. bis VIII. (§ 7, Definition 1 und § 8, Definition 1) genügt, falls dies für R gilt. In allen Fällen (abgesehen von Axiom VIII., wofür die Behauptung trivial ist) verläuft der Beweis völlig gleichartig. Wir beweisen als Beispiel das Axiom III. für R'. Es seien also a' und b' beliebige Elemente aus R' und a und b ihre Urbilder in R. Da nach Voraussetzung in R das Axiom III. gilt, gibt es ein Element $c \in R$ mit $a + c = b$. Ist nun $c \to c'$, dann ist auf Grund der Isomorphie auch $a' + c' = b'$, d. h., c' ist Lösung der Gleichung $a' + x' = b'$. Also erfüllt auch R' die Bedingung III. Dem Leser sei empfohlen, die Gültigkeit der anderen Axiome in R' zu beweisen.

Zusammen mit den Grundeigenschaften bleiben also bei der Isomorphie auch alle anderen Eigenschaften erhalten, die Folgerungen aus den Grundeigenschaften sind. So entspricht bei einem Isomorphismus der Ringe R und R' die Null von R' der Null von R; falls R ein Einselement enthält, so enthält auch R' ein Einselement, und die Einselemente entsprechen sich wechselseitig. In der Tat: Aus $a + 0 = a$ in R folgt $a' + 0' = a'$ in R' und aus $a \cdot 1 = a$ in R folgt $a' \cdot 1' = a'$ in R', und zwar für jedes a' aus R'.

Eine große Bedeutung für die Konstruktion der Zahlkörper besitzt der folgende, fast offensichtliche

Satz 2. *Es sei R ein Unterring des Ringes S und R' ein zu R isomorpher Ring, der mit S kein Element gemein hat. Dann gibt es zu jeder isomorphen Abbildung f des Ringes R auf den Ring R' einen Ring S', der R' als Unterring enthält und der zu S isomorph ist. Dabei gibt es eine isomorphe Abbildung g des Ringes S auf S', die auf R mit der gegebenen Abbildung f zusammenfällt, d. h., für die $g(a) = f(a)$ für alle a aus R gilt. Ist S ein Körper, so ist auch S' ein Körper. Ist R ein Unterkörper von S, so ist auch R' ein Unterkörper von S'.*

Beweis. Es sei S' die Menge, die aus S entsteht, wenn man in ihr die Elemente von R durch die Elemente von R' ersetzt, also $S' = (S \setminus R) \cup R'$. Wir konstruieren nun die folgende Abbildung g der Menge S auf die Menge S': Ist $a \in S \setminus R$, so sei $g(a) = a$; ist $a \in R$, so sei $g(a) = f(a)$, wobei $f(a)$ dasjenige Element aus R' ist, das dem Element a vermöge des Isomorphismus f zugeordnet ist.

Da f eine eineindeutige Abbildung von R auf R' und g bereits eine eineindeutige Abbildung von $S \setminus R$ auf sich ist und da schließlich S und R' voraussetzungsgemäß kein gemeinsames Element besitzen (es würde hier schon genügen, daß $S \setminus R$ und R' kein Element gemein haben), ist g eine eineindeutige Abbildung von S auf S'.

Die Addition und die Multiplikation definieren wir in S' mittels der entsprechenden Operationen aus S auf dem Wege ihrer Übertragung von S auf S' vermöge der Abbildung g, d. h., wir setzen

$$g(a) + g(b) = g(a + b), \quad g(a)g(b) = g(ab) \tag{1}$$

für beliebige Elemente a und b aus S. Wegen der Eineindeutigkeit der Abbildung g gibt es zu jedem Element a' aus S' ein eindeutig bestimmtes Element a aus S, für das $g(a) = a'$ ist; mithin erfassen wir mit $g(a)$ und $g(b)$ alle Elemente aus S', und die Gleichungen (1) definieren algebraische Operationen in S'.

Gleichzeitig zeigen die Gleichungen (1), daß S' und S bezüglich der Addition und der Multiplikation isomorph sind und daß damit auf Grund des vorangehenden Satzes S' ein Ring ist. Ist S ein Körper, so ist aus dem gleichen Grunde auch S' ein Körper.

Wir zeigen nun, daß die Operationen in S' für die Elemente aus R' mit den dort vorgegebenen Operationen zusammenfallen. Da f eine isomorphe Abbildung von R auf R' ist, ist

$$f(a)+f(b)=f(a+b), \qquad f(a)f(b)=f(ab) \tag{2}$$

für alle Elemente a und b aus R. Wenn nun in (1) die Elemente $g(a)$ und $g(b)$ dem Ring R' angehören, so gehören a, b und damit $a+b$ und ab dem Ring R an. Dann fallen aber nach Konstruktion von g die Gleichungen (1) mit den Gleichungen (2) zusammen, in denen die Addition und die Multiplikation auf der linken Seite die Operationen aus dem Ring R' sind. Damit ist bereits bewiesen, daß in R' die Operationen zusammenfallen. Damit ist weiter gezeigt, daß R' ein Unterring von S' ist und daß, falls R ein Unterkörper von S ist, auch R' ein Körper und damit ein Unterkörper von S' ist. Damit ist der behauptete Satz in allen Teilen bewiesen.

§ 10. Angeordnete Ringe und Körper

Bisher haben wir Mengen ohne jede Beziehungen zwischen ihren Elementen (§ 1 bis § 4), Mengen mit einer Ordnungsbeziehung (§ 5) und Mengen mit einer oder zwei algebraischen Operationen (§ 6 bis § 9) betrachtet. Die wichtigste Rolle in der Mathematik spielen indes Zahlenmengen, in denen gleichzeitig eine Ordnungsbeziehung und algebraische Operationen erklärt sind. Wir betrachten jetzt geordnete Ringe und Körper, in denen außerdem noch eine Verbindung zwischen der Ordnungsbeziehung und den Operationen besteht.

Mit der Ordnungsbeziehung in einem Ring (§ 7, Definition 1 und 3) sind die Begriffe „positiv", „negativ" und „absoluter Betrag" für die Elemente eng verbunden.

Das Vorhandensein der Operationen gestattet es, die Einführung der Ordnung in einem Ring etwas zu vereinfachen. Es zeigt sich nämlich, daß es genügt, die Ordnung der Elemente bezüglich der Null zu kennen. Ferner muß man, um die bekannten Eigenschaften der Zahlen zu erhalten, den Operationen und der Ordnung eine weitere Forderung auferlegen, die eine Verbindung zwischen den Operationen und der Ordnung herstellt. Wir definieren dazu:

Definition 1. *Ein Ring (oder Körper) heißt angeordnet, wenn für seine Elemente eine Eigenschaft — positiv zu sein — definiert ist, die den folgenden Forderungen genügt:*

IX. *Für jedes Element $a \in R$ gilt genau eine der drei Beziehungen: $a = 0$, a ist positiv, $-a$ ist positiv.*

X. *Wenn a und b positiv sind, so sind auch $a+b$ und ab positiv.*

a heißt negativ, wenn $-a$ positiv ist.

Satz 1. *Definiert man in einem angeordneten Ring R eine Ordnungsbeziehung durch die Festsetzung, daß $a>b$ dann und nur dann gelten soll, wenn $a-b$ positiv ist, so ist R eine geordnete Menge* (§ 5), *bei der die Null kleiner als alle positiven und größer als alle negativen Elemente ist.*

Beweis. Es seien a und b beliebige Elemente aus R. Ist $a-b=0$, so ist $a=b$. Ist $a-b$ positiv, so ist $a>b$. Ist $-(a-b)=b-a$ positiv, so ist $b>a$. Auf Grund von Eigenschaft IX. folgt dann, daß stets genau einer dieser drei Fälle eintritt [§ 5, Bedingung (1)]. Ferner, wenn $a>b$ und $b>c$ ist, so sind $a-b$ und $b-c$ positiv. Auf Grund von Eigenschaft X. ist dann auch $(a-b)+(b-c)=a-c$ positiv und damit $a>c$ [§ 5, Bedingung (2)]. Also ist in der Tat R eine geordnete Menge.

Ist a positiv, so folgt aus $a=a-0$, daß $a>0$ ist; ist a negativ, so folgt aus $-a=0-a$, daß $0>a$, also $a<0$ ist.

Dieser Satz zeigt, daß die Bedingungen IX. und X. hinreichen, um in R eine Ordnung einzuführen, wobei die Bedingung X. gerade die von den Zahlen her bekannte Verknüpfung zwischen der Ordnung und den Ringoperationen herstellt.

Satz 2. (Monotoniegesetze für die Addition und die Multiplikation) *Für beliebige Elemente a, b, c eines angeordneten Ringes R folgt aus* a) $a>b$ *bzw.* $a=b$ *bzw.* $a<b$ *jeweils* b) $a+c>b+c$ *bzw.* $a+c=b+c$ *bzw.* $a+c<b+c$ *und, falls* $c>0$ *ist,* c) $ac>bc$ *bzw.* $ac=bc$ *bzw.* $ac<bc$ *und entsprechend, falls* $c<0$ *ist,* d) $ac<bc$ *bzw.* $ac=bc$ *bzw.* $ac>bc$.

Beweis. Ist $a>b$, so ist

$$(a+c)-(b+c)=a-b>0,$$

d. h. also $a+c>b+c$. Ist $a=b$, so ist wegen der Eindeutigkeit der Addition $a+c=b+c$. Ist $a<b$, so ist $b>a$, also auf Grund des ersten Falles

$$b+c>a+c, \quad a+c<b+c.$$

Damit ist b) vollständig bewiesen.

Ist $a>b$ und $c>0$, so ist $a-b>0$, also auf Grund der Bedingung X. auch

$$(a-b)\,c=ac-bc>0, \quad ac>bc.$$

Ist hingegen $c<0$, so ist $-c>0$ und auf Grund der Vorzeichenregeln für die Multiplikation [§ 7, Gleichung (3)]

$$bc-ac=(b-a)\,c=[-(b-a)]\,(-c)=(a-b)\,(-c)>0,$$

$$bc>ac, \quad ac<bc.$$

Damit sind die ersten beiden Fälle von c) und d) bewiesen. Die übrigen Fälle erledigen sich wie unter b).

Es gelten nun auch die Umkehrungen von Satz 2, also:

Satz 3. *Aus $a+c>b+c$ bzw. $a+c=b+c$ bzw. $a+c<b+c$ folgt $a>b$ bzw. $a=b$ bzw. $a<b$. Aus $ac>bc$ bzw. $ac=bc$ bzw. $ac<bc$ folgt, falls $c>0$ ist, $a>b$ bzw. $a=b$ bzw. $a<b$ und entsprechend, falls $c<0$ ist, $a<b$ bzw. $a=b$ bzw. $a>b$.*

Beweis. In Satz 2 besitzen die Voraussetzungen a) die Eigenschaft, daß stets eine von ihnen (und auch nur eine von ihnen, was hier aber unwichtig ist) zutrifft, und die Folgerungen [in jedem der Fälle b), c) und d) für sich] die Eigenschaft, daß sie sich paarweise ausschließen. Für jedes Theorem dieser Art gelten ganz allgemein auch die Umkehrungen[1]), wie man jeweils durch einen indirekten Beweis zeigt. Als Beispiel zeigen wir, daß aus $ac = bc$ im Falle $c > 0$ auch $a = b$ folgt. Wir nehmen dazu an, es wäre $a \neq b$. Dann tritt eine der Voraussetzungen $a > b$ oder $a < b$ des Satzes 2 in Funktion. Im Fall, daß $a > b$ ist, besagt aber Satz 2, daß dann $ac > bc$, und im Fall, daß $a < b$ ist, daß $ac < bc$ ist, was jeweils einen Widerspruch zu $ac = bc$ ergibt, weil dadurch die Ungleichungen $ac > bc$ und $ac < bc$ ausgeschlossen sind.

Korollar 1. *In einem angeordneten Ring folgt aus*

a) $$a - b \gtreqless c - d,$$

daß jeweils

b) $$a + d \gtreqless b + c$$

ist, und umgekehrt.

In der Tat: Fügt man in a) auf beiden Seiten die Summe $b + d$ hinzu, so erhält man b). Die Umkehrung gilt, da in a) alle Fälle erschöpft werden und die Aussagen unter b) sich paarweise ausschließen.

Korollar 2. *In einem angeordneten Körper folgt, falls $bd > 0$ ist, aus*

a) $$\frac{a}{b} \gtreqless \frac{c}{d},$$

daß jeweils

b) $$ad \gtreqless bc$$

ist, und umgekehrt.

Der Beweis hierfür verläuft analog dem vorhergehenden.

Aus Satz 2 folgen ferner die von den Zahlen her bekannten Rechenregeln für das Rechnen mit Ungleichungen:

Satz 4. *Wenn $a > b$ und $c > d$ ist, so ist $a + c > b + d$ und, wenn die Elemente a, b, c, d positiv sind, auch $ac > bd$. Sind die Elemente a, b, c, d alle negativ, so ist unter den angegebenen Voraussetzungen $ac < bd$. Schließlich kann man überall $>$ mit $<$ vertauschen.*

Beweis. Wenn $a > b$ ist, so ist auf Grund von Satz 2 auch $a + c > b + c$, und wenn zudem $c > d$ ist, so ist $b + c > b + d$ und damit $a + c > b + d$. Entsprechend zeigt man, daß für positive a, b, c, d auch $ac > bd$ gilt. Sind hingegen a, b, c, d negativ, so folgt aus $a > b$, daß $ac < bc$ ist, und aus $c > d$, daß $bc < bd$ und mithin $ac < bd$ ist.

Eine unmittelbare Folgerung aus Satz 3 ist

Satz 5. *In einem angeordneten Ring gibt es keine Nullteiler* (§ 7, Definition 2).

Beweis. Wenn $ab = 0$ ist, so ist $ab = a0$, also muß nach Satz 3, falls $a \neq 0$, also $a > 0$ oder $a < 0$ ist, $b = 0$ sein.

---

[1]) Dies ist die Aussage des sog. HAUBERschen Theorems der formalen Logik. — *Anm. d. wiss. Red.*

Satz 6. *Die Charakteristik* (§ 8, Definition 2) *eines angeordneten Körpers K ist Null.*

Beweis. Es sei $a \in K$ und $a \neq 0$. Ist $a > 0$, so ist auf Grund von Bedingung X. für jedes natürliche n auch $na > 0$ und, da $(-n)a = -na$ ist, auch allgemein $na \neq 0$ für beliebiges ganzes $n \neq 0$.

Satz 7. *In einem angeordneten Ring ist stets die Summe der Quadrate einer endlichen Anzahl von Elementen (also insbesondere das Quadrat jedes Elementes) größer oder gleich Null. Eine Summe von Quadraten ist dann und nur dann gleich Null, wenn alle gegebenen Elemente gleich Null sind.*

Beweis. Für ein einzelnes Element gilt der Satz. Ist nämlich $a_1 = 0$, so ist auch $a_1^2 = 0$; ist hingegen $a_1 \neq 0$, also $a_1 > 0$ oder $-a_1 > 0$, so ist

$$a_1^2 = a_1 a_1 = (-a_1)(-a_1) > 0.$$

Für $n = 1$ Element gilt also der Satz. Angenommen nun, er sei bereits für Summen von Quadraten aus n Gliedern bewiesen. Dann ist aber auch

$$\sum_{i=1}^{n+1} a_i^2 = \sum_{i=1}^{n} a_i^2 + a_{n+1}^2 \geq 0,$$

als Summe aus nicht-negativen Summanden (Bedingung X.). Ist einer der beiden Summanden > 0, so ist auch die Summe > 0. Das heißt, daß die Summe nur dann gleich Null ist, wenn beide Summanden gleich Null sind, wenn also

$$\sum_{i=1}^{n} a_i^2 = 0 \quad \text{und} \quad a_{n+1}^2 = 0$$

ist. Dann ist aber auf Grund des bereits Bewiesenen $a_{n+1} = 0$ und nach Induktionsvoraussetzung $a_1 = a_2 = \cdots = a_n = 0$.

Definition 2. *Unter dem absoluten Betrag eines Elementes a eines angeordneten Ringes (oder Körpers) verstehen wir das nicht-negative unter den Elementen a und $-a$. Der absolute Betrag des Elementes a wird mit $|a|$ bezeichnet.*

Auf Grund dieser Definition ist $|0| = 0$ und $|a| > 0$, falls $a \neq 0$ ist.

Satz 8 (Dreiecksungleichung) *Der absolute Betrag einer Summe aus endlich vielen Elementen eines angeordneten Ringes ist stets höchstens gleich der Summe der absoluten Beträge der Summanden. Gleichheit tritt dann und nur dann ein, wenn alle Summanden nicht-negativ oder wenn alle Summanden nicht-positiv sind. Der absolute Betrag eines Produktes aus endlich vielen Elementen ist gleich dem Produkt der absoluten Beträge der Faktoren.*

Beweis. Wir beschränken uns hier auf die Summe und das Produkt aus zwei Elementen, da die Durchführung des Induktionsschlusses keinerlei Schwierigkeiten bereitet. Wir zeigen also, daß

$$|a + b| \leq |a| + |b| \tag{1}$$

ist und daß hier Gleichheit dann und nur dann eintritt, wenn $a \geq 0$ und $b \geq 0$ oder wenn $a \leq 0$ und $b \leq 0$ ist. Wir zeigen außerdem, daß stets

$$|ab| = |a| \cdot |b| \tag{2}$$

ist.

Wenn $a \geqq 0$ und $b \geqq 0$ ist, so ist auch $a+b \geqq 0$ und mithin

$$|a+b| = a+b = |a|+|b|.$$

Ist $a \leqq 0$ und $b \leqq 0$, so ist $-a \geqq 0$ und $-b \geqq 0$ und damit

$$-(a+b) = (-a)+(-b) \geqq 0,$$

woraus folgt, daß

$$|a+b| = -(a+b) = (-a)+(-b) = |a|+|b|$$

ist. Also gilt in diesen beiden Fällen (1) mit dem Gleichheitszeichen. Da (1) in a und b völlig symmetrisch ist, genügt es, von den restlichen beiden Fällen $a > 0$, $b < 0$ und $a < 0$, $b > 0$ einen, etwa den ersten, zu betrachten. Auf Grund von Satz 2 erhalten wir, wenn wir zu der Ungleichung $b < -b$ das Element a addieren:

$$a+b < a+(-b) = |a|+|b|.$$

Entsprechend erhalten wir, wenn wir zu der Ungleichung $-a < a$ das Element $-b$ hinzufügen:

$$-(a+b) = (-a)+(-b) < a+(-b) = |a|+|b|.$$

Nun ist aber $|a+b|$ gleich einem der Elemente $a+b$ oder $-(a+b)$, also jedenfalls

$$|a+b| < |a|+|b|.$$

Mithin gilt in den restlichen Fällen (1) mit dem $<$-Zeichen.

Die Gleichung (2) ist sicher dann erfüllt, wenn wenigstens eines der Elemente a oder b gleich Null ist. Es genügt also, die folgenden drei Fälle zu untersuchen:

1.) $a > 0$, $b > 0$. Dann ist auf Grund von Bedingung X. $ab > 0$ und mithin

$$|ab| = ab = |a| \cdot |b|.$$

2.) $a < 0$, $b < 0$, also $-a > 0$, $-b > 0$ und damit $(-a)(-b) > 0$. Dann ist auf Grund der Vorzeichenregel (3) aus § 7

$$|ab| = |(-a)(-b)| = (-a)(-b) = |a| \cdot |b|.$$

3.) $a > 0$, $b < 0$. Dann ist $-b > 0$, $a(-b) > 0$ und mithin

$$|ab| = |-ab| = |a(-b)| = a(-b) = |a| \cdot |b|.$$

Aus der Ungleichung (1) folgt unmittelbar

$$\big||a|-|b|\big| \leqq |a \pm b| \leqq |a|+|b| \tag{3}$$

für beliebige Elemente a, b eines angeordneten Ringes R. In der Tat: Da $a+b = a-(-b)$ und $|b| = |-b|$ ist, genügt es, (3) für den Fall der Differenz $a-b$ zu beweisen. Aus $a = (a-b)+b$ und $b = (b-a)+a$ folgt auf Grund von (1), daß

$$|a| \leqq |a-b|+|b| \quad \text{und} \quad |b| \leqq |b-a|+|a| = |a-b|+|a|$$

ist, woraus sich

$$|a|-|b| \leqq |a-b| \quad \text{und} \quad |b|-|a| \leqq |a-b|$$

und damit

$$\big||a|-|b|\big| \leqq |a-b| = |a+(-b)| \leqq |a|+|b|$$

ergibt.

Bemerkung. Entsprechend gelten auch die aus der elementaren Algebra bekannten Regeln für den Vergleich und das Operieren mit sogenannten „relativen Zahlen" durch den Vergleich und das Operieren mit ihren absoluten Beträgen in beliebigen angeordneten Ringen R.

Jedes positive Element des Ringes R ist größer als jedes negative Element, wie sich unmittelbar aus dem Vergleich mit der Null ergibt. Von zwei positiven Elementen ist dasjenige größer, das den größeren absoluten Betrag besitzt, da der absolute Betrag eines positiven Elementes gleich dem Element selbst ist. Von zwei negativen Elementen ist dasjenige größer, das den kleineren absoluten Betrag besitzt. In der Tat: Sind a und b negativ, so ist $a - b = (-b) - (-a) = |b| - |a|$ und daher $a > b$ dann und nur dann, wenn $|a| < |b|$ ist.

Wenn wir in Symmetrie zur Bezeichnung des entgegengesetzten Elementes das entgegengesetzte Element von $-a$, also a selbst, mit $+a$ bezeichnen, so können wir jedes Element mit Hilfe seines absoluten Betrages in der Form $a = \pm |a|$ ausdrücken. Dabei steht das Zeichen $+$ für positive und das Zeichen $-$ für negative Elemente. In diesem Sinne können wir von dem Vorzeichen eines gegebenen Elementes sprechen. Dabei gelten die folgenden Rechengesetze:

Will man zwei Elemente gleichen Vorzeichens addieren, so muß man ihre absoluten Beträge addieren und dem Resultat das Vorzeichen der Summanden geben. In der Tat: Ist $a > 0$ und $b > 0$, so ist dies klar; ist dagegen $a < 0$ und $b < 0$, so ist $a + b = (-|a|) + (-|b|) = -(|a| + |b|)$.

Will man zwei Elemente verschiedenen Vorzeichens addieren, so muß man von dem größeren der absoluten Beträge den kleineren subtrahieren (bei Gleichheit der absoluten Beträge ist die Summe gleich Null!) und dem Resultat das Vorzeichen des Summanden mit dem größeren absoluten Betrag geben. Ist nämlich etwa $a > 0$ und $b < 0$, so ist, falls $|a| > |b|$ ist,

$$a + b = a - (-b) = +(|a| - |b|)$$

und, falls $|a| < |b|$ ist,

$$a + b = -(-b - a) = -(|b| - |a|).$$

Will man von einem Element ein anderes subtrahieren, so muß man zum ersten Element das entgegengesetzte Element des zweiten addieren. Dies gilt, wie wir sahen, in ganz beliebigen Ringen.

Will man zwei Elemente miteinander multiplizieren (durcheinander dividieren), so muß man den absoluten Betrag des ersten Elementes mit dem absoluten Betrag des zweiten Elementes multiplizieren (durch ihn dividieren) und als Vorzeichen das Zeichen $+$ wählen, wenn beide Elemente gleiches Vorzeichen haben; dagegen nimmt man als Vorzeichen das Zeichen $-$, wenn die Vorzeichen voneinander verschieden sind. Für die Multiplikation folgt dies aus den Vorzeichenregeln, die in beliebigen Ringen gelten [§ 7, Gleichung (3)], da

$$ab = (\pm |a|) \cdot (\pm |b|)$$

ist. Für die Division (falls sie ausführbar ist) ergibt es sich folgendermaßen: Wenn $\frac{a}{b} = c$ ist, so ist $a = bc$, also $|a| = |b|\,|c|$, woraus sich $\frac{|a|}{|b|} = |c|$ ergibt. Bei der Multiplikation mit einem positiven Element bleibt das Vorzeichen erhalten, während es sich bei der Multiplikation mit einem negativen Element ändert. Daher folgt

aus $a = bc$, daß bei gleichen Vorzeichen von a und b der Quotient c positiv und bei verschiedenen Vorzeichen negativ ist.

Wir haben damit gesehen, daß die üblichen Rechenregeln des Zahlenrechnens mit Ungleichungen und absoluten Beträgen nicht nur für Zahlen, sondern auch für die Elemente eines ganz beliebigen angeordneten Ringes gelten. Sie erweisen sich als unmittelbare Folgerungen aus den Axiomen I. bis VI., IX. und X.

Es gibt aber noch eine wesentliche Eigenschaft der Zahlen, die nicht allgemein auf angeordnete Ringe übertragen werden kann. Wir meinen die Gültigkeit des sogenannten Archimedischen Axioms, nach welchem man mit Hilfe jeder positiven Zahl (so klein sie auch immer sein mag), wenn man sie nur hinreichend oft zu sich selbst addiert, jede vorgegebene Zahl (wie groß sie auch immer ist) übertreffen kann. Ringe, die eine analoge Eigenschaft besitzen, bedürfen einer besonderen Definition:

Definition 3. *Ein Ring (Körper) heißt archimedisch angeordnet, wenn er die folgende Eigenschaft besitzt:*

XI. (Archimedisches Axiom) *Zu jedem Element a und jedem Element $b > 0$ des Ringes gibt es eine natürliche Zahl n derart, daß $nb > a$ ist.*

Im Fall eines Körpers läßt sich diese Bedingung etwas vereinfachen; es genügt dort nämlich, die Bedingung jeweils für das Einselement e des Körpers zu fordern, d. h., in einem Körper ist die Bedingung XI. gleichwertig mit

XI′. *Zu jedem Element a des Körpers gibt es eine natürliche Zahl n mit $ne > a$.*

In der Tat: Wenn $b > 0$ ist, so gibt es eine natürliche Zahl n, für die $ne > \frac{a}{b}$ ist. Multipliziert man diese Ungleichung mit $b > 0$, so ergibt sich $nb > a$.

Beispiel 1. Der Ring aller ganzen Zahlen, der Körper aller rationalen Zahlen und der Körper aller reellen Zahlen sind archimedisch angeordnet (die Beweise hierfür finden sich in den entsprechenden Kapiteln).

Beispiel 2. Es sei R der Ring aller Polynome

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

mit rationalen Koeffizienten (bezüglich der üblichen Addition und Multiplikation von Polynomen). Wir nennen ein Polynom $f(x)$ positiv, wenn sein höchster (nicht verschwindender) Koeffizient a_n positiv ist. Man sieht leicht ein, daß dann die Bedingungen IX. und X. der Definition 1 erfüllt sind, d. h., daß R ein angeordneter Ring ist. Es ist dann $1 > 0$, aber $n \cdot 1 = n < x$ für jede natürliche (und auch jede rationale) Zahl n, da $x - n > 0$ ist. Also ist R ein nicht-archimedisch angeordneter Ring. Die Brüche der Gestalt $\frac{f(x)}{g(x)}$, wobei $f(x)$ und $g(x)$ Polynome aus dem Ring R sind, bilden einen Körper K. Der Leser mag beweisen, daß der Körper K angeordnet ist, wenn man einen Bruch $\frac{f(x)}{g(x)}$ als positiv bezeichnet, falls $f(x)$ und $g(x)$ bei der oben angegebenen Anordnung von R gleiches Vorzeichen besitzen. Da wieder $n \cdot 1 < x$ für jede natürliche Zahl n gilt, ist K ein nicht-archimedisch angeordneter Körper.

Kapitel III

DIE NATÜRLICHEN ZAHLEN

§ 11. Das Axiomensystem für die natürlichen Zahlen

Der axiomatische Aufbau einer vorgelegten mathematischen Theorie (siehe § 9) beginnt mit einer Aufzählung der (nicht definierten) Grundbeziehungen und einer Zusammenstellung der (ohne Beweis verwendeten) Grundeigenschaften oder Axiome für die betrachteten Beziehungen. Beim axiomatischen Aufbau der Theorie der natürlichen Zahlen wird eine Grundbeziehung eingeführt, für welche vier Axiome gefordert werden, und zwar:

Definition 1. *Als natürliche Zahlen bezeichnen wir die Elemente jeder nichtleeren Menge N, für deren Elemente eine Beziehung „b folgt (unmittelbar) auf a" (oder „b ist unmittelbarer Nachfolger von a") erklärt ist (jede auf a folgende Zahl werde mit a' bezeichnet), die den folgenden Axiomen genügt:*

I. *Es gibt eine Zahl* 1, *die nicht unmittelbarer Nachfolger einer anderen ist, d. h., für die* $a' \neq 1$ *für jede Zahl a gilt.*[1])

II. *Zu jeder Zahl a gibt es einen eindeutig bestimmten Nachfolger a', d. h., wenn* $a = b$ *ist, so ist* $a' = b'$.

III. *Jede Zahl ist Nachfolger höchstens einer Zahl, d. h., wenn* $a' = b'$ *ist, so ist* $a = b$.

IV. (Axiom der vollständigen Induktion) *Jede Menge M von natürlichen Zahlen, die*

A.) *die Zahl* 1 *enthält,*

B.) *mit einer Zahl a auch ihren Nachfolger a' enthält,*

enthält alle natürlichen Zahlen, ist also gleich der Menge N.

Das hier angegebene Axiomensystem für die natürlichen Zahlen stellt eine nur unwesentliche Abänderung des im Jahre 1891 von dem italienischen Mathematiker und Logiker G. PEANO vorgeschlagenen Axiomensystems dar.

Unsere Definition könnte deshalb als schlecht erscheinen, weil nach ihr die Elemente jeder Menge N, die die angegebenen Axiome erfüllt, als natürliche Zahlen bezeichnet werden.

Tatsächlich gibt es auch verschiedene Mengen, die der Definition 1 genügen. Alle diese Mengen sind indes bezüglich der Grundbeziehung „b folgt auf a" untereinander isomorph (vgl. § 9, Definition 1), besitzen also in bezug auf diese Beziehung völlig gleichartige Eigenschaften, sofern diese aus den Axiomen I. bis IV. folgen. Den Beweis hierfür wie auch alle anderen Fragen, die das Axiomensystem selbst betreffen, verschieben wir an das Ende dieses Kapitels (§ 17) und ziehen zunächst eine Reihe von Folgerungen aus dem Axiomensystem.

Zuerst wollen wir die Bedeutung des Axioms der vollständigen Induktion näher erläutern. Üblicherweise führt man die Beweise durch vollständige Induktion folgendermaßen. Ist ein gewisser Satz, in dessen Formulierung eine natürliche Zahl n

---

[1]) Das Zeichen „=" bezeichnet nach wie vor das Zusammenfallen der Elemente, das Zeichen „≠" ihre Verschiedenheit.

eingeht (wie etwa der binomische Satz), für jede Zahl n zu beweisen, so beweist man diesen Satz zunächst für $n = 1$ und zeigt dann, daß dieser Satz für die Zahl $n + 1$ gilt, sofern er für die Zahl n gilt. Danach sieht man den betreffenden Satz als für jede Zahl n bewiesen an. Für gewöhnlich begründet man das so: Der Satz gilt für die Zahl 1, also auch für die Zahl 2; da er für die Zahl 2 gilt, gilt er dann aber auch für die Zahl 3 und damit auch für die Zahl 4 usw. Was bedeutet nun aber hier das „usw."? Können wir etwa, wenn wir so fortschreiten, nach und nach alle natürlichen Zahlen ausschöpfen? Offenbar doch nicht, denn es gibt ja unendlich viele natürlichen Zahlen. Das Axiom IV. der vollständigen Induktion ist nun gerade das formale Hilfsmittel, um derartige Sätze sofort für die ganze unendliche Gesamtheit der natürlichen Zahlen zu beweisen. Es gilt nämlich folgender Satz:

Satz 1 (Satz über die Rechtfertigung induktiver Beweise) *Wenn ein gewisses Theorem T, in dessen Formulierung eine natürliche Zahl n eingeht, für die Zahl 1 bewiesen ist und wenn außerdem bewiesen ist, daß das Theorem für die Zahl n'*[1]) *gilt unter der Voraussetzung, daß es für die Zahl n gilt, so gilt das Theorem T für alle natürlichen Zahlen n.*

Beweis. Es sei M die Menge aller natürlichen Zahlen, für die das betrachtete Theorem T gilt. Dann ist A.) die Zahl 1 in M enthalten, da das Theorem für die Zahl 1 bewiesen ist. B.) Es möge die Zahl n in M enthalten sein, also das Theorem für die Zahl n gelten. Dann ist aber bewiesen, daß das Theorem auch für die Zahl n' gilt, d. h., dann gehört auch die Zahl n' zu M. Also erfüllt die Menge M die Bedingungen A.) und B.) des Axioms IV. Auf Grund von Axiom IV. enthält mithin die Menge M alle natürlichen Zahlen, d. h., (so war ja die Menge M gerade definiert) das Theorem T gilt für jede natürliche Zahl n. Damit ist aber Satz 1 bereits bewiesen.

Definition 2. *Wir sagen: „a geht b (unmittelbar) voran"* (oder *„a ist unmittelbarer Vorgänger von b"*), *wenn b unmittelbarer Nachfolger von a ist.*

Axiom I. besagt dann gerade, daß die Zahl 1 keinen unmittelbaren Vorgänger besitzt. Sie ist aber auch die einzige Zahl mit dieser Eigenschaft.

Satz 2. *Zu jeder Zahl $a \neq 1$ gibt es genau eine Zahl, die ihr unmittelbar vorangeht.*

Beweis. Es sei M die Menge, die die Zahl 1 und alle die natürlichen Zahlen enthält, die einen unmittelbaren Vorgänger besitzen.

A.) Die Zahl 1 gehört zu M. B.) Wenn a zu M gehört, so gehört auch a' zu M, da die Zahl a ein unmittelbarer Vorgänger von a' ist (die Voraussetzung, daß a zu M gehört, wird hier offenbar nicht benötigt). Auf Grund von Axiom IV. enthält dann M alle natürlichen Zahlen. Also besitzt jede Zahl $a \neq 1$ wenigstens einen unmittelbaren Vorgänger. Die Eindeutigkeit der vorangehenden Zahl ergibt sich unmittelbar aus Axiom III., nach welchem jede Zahl höchstens einen Vorgänger besitzt.

Satz 3. *Sind die Nachfolger von a und b voneinander verschieden, so sind auch a und b voneinander verschieden, d. h. wenn $a' \neq b'$ ist, so ist auch $a \neq b$.*

Beweis. Wäre nämlich $a = b$, so wäre auf Grund von Axiom II. $a' = b'$.

Satz 4. *Sind zwei Zahlen voneinander verschieden, so sind auch ihre Nachfolger voneinander verschieden, d. h., wenn $a \neq b$ ist, so ist auch $a' \neq b'$.*

Beweis. Wäre nämlich $a' = b'$, so wäre auf Grund von Axiom III. auch $a = b$.

[1]) Um zu zeigen, daß $n' = n + 1$ ist, muß zuvor die Addition für natürliche Zahlen eingeführt werden.

Satz 5. *Jede Zahl ist von ihrem Nachfolger verschieden, d. h., es ist $a \neq a'$ für jedes a.*

Beweis. Es sei M die Menge der Zahlen, für die dieser Satz gilt.

A.) Auf Grund von Axiom I. ist $1 \neq 1'$, also 1 Element von M. B.) Wenn a zu M gehört, so ist $a \neq a'$, also nach Satz 4 auch $a' \neq (a')'$, d. h. a' Element von M. Nach Axiom IV. enthält dann M alle natürlichen Zahlen, d. h., es ist $a \neq a'$ für jede natürliche Zahl a.

§ 12. Addition

Definition. *Unter einer Addition im Bereich der natürlichen Zahlen verstehen wir eine Vorschrift (Abbildung), die jedem geordneten Paar von natürlichen Zahlen a, b eine und nur eine natürliche Zahl $a + b$ so zuordnet, daß*

1.) $a + 1 = a'$ *für jede natürliche Zahl a,*

2.) $a + b' = (a + b)'$ *für beliebiges a und b.*

Die Zahlen a und b heißen Summanden, die Zahl $a + b$ ihre Summe.[1])

Es erhebt sich sofort die Frage, ob es überhaupt eine derartige Vorschrift gibt, und wenn dies der Fall ist, ob es auch nur eine solche Vorschrift gibt. Die obenstehende Definition ist ein Beispiel für eine sogenannte induktive Definition. Ist nämlich a eine beliebige natürliche Zahl, so definiert die Bedingung 1.) die Zahl $a + 1$ und die Bedingung 2.) die Zahl $a + b'$ unter der Voraussetzung, daß bereits die Zahl $a + b$ definiert ist. Somit kann man — so scheint es wenigstens — auf Grund von Axiom IV. die Zahl $a + b$ bei festem a für beliebiges b als definiert ansehen, und da a ganz beliebig war, so ist dann $a + b$ für beliebiges a und b definiert. Dies dachten auch die Begründer des axiomatischen Aufbaus der natürlichen Zahlen, PEANO und seine Schüler. Ähnliche Begründungen findet man auch heute noch in den meisten Lehrbüchern der Mathematik. Allein, diese Überlegungen enthalten einen Fehler. Wenn wir nämlich das Axiom der vollständigen Induktion benutzen, müssen wir jedesmal genau die Menge angeben, für welche die Gültigkeit der Bedingungen A.) und B.) zu beweisen ist. So bestand in Satz 1 aus § 11 die Menge M aus allen den natürlichen Zahlen, für die das vorgelegte Theorem T gilt. Wir konnten dann zeigen, daß unter den angegebenen Voraussetzungen über T die Bedingungen A.) und B.) erfüllt waren, womit dann das Theorem T für alle natürlichen Zahlen n bewiesen war. Auch der Einwand wurde damit hinfällig, daß wir beim Beweis des Theorems für $n + 1$ angenommen hätten, daß das Theorem bereits für n bewiesen sei, obgleich wir es dafür noch gar nicht bewiesen haben. Denn wir benutzen ja nicht, daß das Theorem für n gilt, sondern beweisen nur, entsprechend der bedingenden Form der Eigenschaft B.), die folgende bedingende Aussage: „Wenn das Theorem T für n gilt, so gilt es auch für $n + 1$."

Wir wollen jetzt aufzuklären versuchen, auf welche Menge M wir das Axiom IV. im Fall der Definition der Addition anwenden müßten. Kann man sagen, daß

[1]) Die Addition ist mithin ein Spezialfall des allgemeinen Begriffs der algebraischen Operation (§ 6, Definition 1) oder des noch allgemeineren Begriffs der Funktion (§ 3, Definition 1).

für ein festes a die Menge M aus den Zahlen b besteht, für die $a+b$ definiert ist? Offenbar doch nicht, denn wir wollen ja gerade erst zeigen, daß die Zahl $a+b$ durch die Eigenschaften 1.) und 2.) eindeutig bestimmt ist (wir dürfen — kurz gesagt — bei der Definition von M nichts benutzen, von dem wir noch gar nicht wissen, ob es so etwas überhaupt gibt — *Anm. d. wissenschaftl. Red.*). Hierin besteht gerade der Unterschied zwischen den induktiven Definitionen und den induktiven Beweisen, bei denen die Menge M der Zahlen, für die das Theorem T gilt, eine fest umrissene Bedeutung hat, unabhängig davon, ob das Theorem bewiesen ist oder nicht. Die Aussage: „Für gegebenes a gibt es zu jedem b eine Zahl $a+b$ mit den Eigenschaften 1.) und 2.)" kann doch offenbar nur bedeuten: „Für gegebenes a gibt es eine Vorschrift, die jeder Zahl b eine Zahl $a+b$ zuordnet und für die die Bedingungen 1.) und 2.) erfüllt sind." Diese Aussage betrifft nun aber nicht eine, sondern sofort alle Zahlen b und kann deshalb nicht durch vollständige Induktion über b mit einfachem Hinweis auf die Bedingungen 1.) und 2.) bewiesen werden. Dafür bezieht sich diese Aussage aber auf eine vorgelegte Zahl a, und man kann daher versuchen, sie durch vollständige Induktion über a zu beweisen (was wir auch weiter unten tun werden). Wir behaupten ja keineswegs, daß die Aussage, nach welcher die Bedingungen 1.) und 2.) eine eindeutig bestimmte Zahl $a+b$ definieren, falsch ist, sondern nur, daß der Beweis durch vollständige Induktion über b fehlerhaft ist. Man kann sogar ganz allgemein zeigen, daß jede induktive Definition, die sich auf die Ordnung der natürlichen Zahlen stützt, einwandfrei ist (siehe § 15). Da wir aber den Begriff der Ordnung erst unter Benutzung der Addition einführen (siehe § 14), müssen wir die Frage nach der Existenz der Addition auf einem anderen Wege lösen.

Satz 1. *Im Bereich der natürlichen Zahlen gibt es genau eine Addition, d. h., es gibt im Bereich der natürlichen Zahlen eine und nur eine Vorschrift, die beliebigen Zahlen a und b eine eindeutig bestimmte Zahl $a+b$ zuordnet und die folgenden Bedingungen erfüllt:*

1.) $a+1=a'$ *für jedes* a,

2.) $a+b'=(a+b)'$ *für jedes* a *und* b.

Mit anderen Worten: Die Addition ist stets eindeutig ausführbar.

Beweis. a) Zunächst beweisen wir, daß es zu vorgegebenem a höchstens eine Vorschrift gibt, die jeder Zahl b eine Zahl x_b zuordnet und für die $x_1=a'$ und $x_{b'}=(x_b)'$ für beliebiges b gilt. Es sei y_b eine beliebige Vorschrift mit denselben Eigenschaften, also eine Vorschrift, für die $y_1=a'$ und $y_{b'}=(y_b)'$ für beliebiges b ist. Ferner sei M die Menge aller Zahlen b, für die $x_b=y_b$ ist.

A.) $x_1=a'=y_1$, d. h., 1 gehört zu M.

B.) Wenn b zu M gehört, so ist $x_b=y_b$, also auf Grund von Axiom II. auch $(x_b)'=(y_b)'$ und mithin $x_{b'}=(x_b)'=(y_b)'=y_{b'}$, also auch b' Element von M.

Dann enthält nach Axiom IV. die Menge M alle natürlichen Zahlen, d. h., es ist $x_b=y_b$ für alle b. Damit ist die Einzigkeit der Addition für gegebenes a bewiesen. Da a ganz beliebig gewählt war, gilt sie dann für beliebiges a und b.

b) Wir zeigen jetzt, daß es für vorgelegtes a eine (und wegen a) auch nur eine) Vorschrift gibt, die jeder Zahl b eine eindeutig bestimmte Zahl $a+b$ zuordnet und die die Bedingungen $a+1=a'$, $a+b'=(a+b)'$ für beliebiges b (und vorgegebenes a) erfüllt. Es sei M die Menge derjenigen natürlichen Zahlen a, für die es eine solche Vorschrift gibt (die dann wegen a) eindeutig bestimmt ist).

A.) Für $a = 1$ setzen wir $a + b = b'$ für beliebiges b. Diese Vorschrift besitzt offenbar die verlangten Eigenschaften, denn es ist

$$a + 1 = 1' = a', \quad a + b' = (b')' = (a + b)'.$$

Also gehört 1 zu M.

B.) Wenn a zu M gehört, so gibt es eine Vorschrift, die jeder Zahl b eine Zahl $a + b$ zuordnet und für die $a + 1 = a'$, $a + b' = (a + b)'$ ist. Dann setzen wir $a' + b = (a + b)'$. Diese Vorschrift besitzt dann die entsprechenden Eigenschaften bezüglich a'; in der Tat ist

$$a' + 1 = (a + 1)' = (a')', \quad a' + b' = (a + b')' = [(a + b)']' = (a' + b)'.$$

Auf Grund von Axiom IV. enthält dann M alle natürlichen Zahlen, d. h., zu jedem a gibt es eine Vorschrift, die jeder Zahl b eine Zahl $a + b$ zuordnet und für die

$$a + 1 = a', \; a + b' = (a + b)'$$

für das gegebene a und beliebiges b gilt. Da die Zahl a ganz beliebig gewählt ist, ist damit die eindeutige Existenz einer Vorschrift[1]) bewiesen, die jedem geordneten Paar a, b von natürlichen Zahlen eine Zahl $a + b$ zuordnet und die die Bedingungen 1.) und 2.) erfüllt, was zu beweisen war.

Satz 2. (Assoziatives Gesetz der Addition)

$$(a + b) + c = a + (b + c).$$

Beweis. Es bestehe für gegebenes a und b die Menge M aus allen denjenigen Zahlen c, für die diese Gleichung gilt.

A.) $\quad (a + b) + 1 = (a + b)' = a + b' = a + (b + 1);$

M enthält also die Zahl 1.

B.) Wenn c zu M gehört, so ist $(a + b) + c = a + (b + c)$. Dann ist aber auch

$$(a + b) + c' = [(a + b) + c]' = [a + (b + c)]' = a + (b + c)' = a + (b + c'),$$

d. h. c' Element von M.

Auf Grund von Axiom IV. gilt dann $(a + b) + c = a + (b + c)$ für beliebiges a, b und c.

Satz 3. (Kommutatives Gesetz der Addition)

$$a + b = b + a.$$

Beweis. a) Wir zeigen zunächst durch vollständige Induktion über a, daß $a + 1 = 1 + a$ ist. Es sei M die Menge aller Zahlen a, für die dies gilt. A.) Offensichtlich gehört 1 zu M. B.) Wenn a zu M gehört, so ist $a + 1 = 1 + a$. Dann ist aber

$$a' + 1 = (a + 1) + 1 = (1 + a) + 1 = (1 + a)' = 1 + a',$$

also auch a' Element von M. Mithin ist auf Grund von Axiom IV. $a + 1 = 1 + a$ für alle a.

---

[1]) Man braucht dazu nur jedem geordneten Paar a, b von natürlichen Zahlen die Zahl $a + b$ zuzuordnen, die der Zahl b vermöge der (eindeutig bestimmten) zu a gehörigen Vorschrift zugeordnet ist. — *Anm. d. wissenschaftl. Red.*

b) Wir zeigen nun durch vollständige Induktion über b, daß $a + b = b + a$ ist. Dazu sei M die Menge derjenigen natürlichen Zahlen b, für die dies für ein vorgegebenes a gilt. A.) Auf Grund des unter a) bewiesenen gehört 1 zu M. B.) Wenn b zu M gehört, so ist $a + b = b + a$, also gilt nach Satz 2:

$$a + b' = (a + b)' = (b + a)' = b + a' = b + (a + 1) = b + (1 + a) = \\ = (b + 1) + a = b' + a,$$

d. h., dann gehört auch b' zu M. Axiom IV. liefert unmittelbar die Behauptung.

Satz 4. $a + b \neq b$.

Beweis. Dieser Satz gilt für $b = 1$, da auf Grund von Axiom I. $a + 1 = a' \neq 1$ ist. Wenn $a + b \neq b$ ist, so ist auf Grund von Satz 4 aus § 11 auch

$$a + b' = (a + b)' \neq b'.$$

Satz 5. *Für natürliche Zahlen a und b gilt genau einer der folgenden drei Fälle:* 1.) $a = b$; 2.) *Es gibt eine natürliche Zahl k derart, daß $a = b + k$ ist;* 3.) *Es gibt eine natürliche Zahl l derart, daß $b = a + l$ ist.*

Beweis. Aus Satz 4 folgt zunächst unmittelbar, daß nicht mehr als einer der Fälle 1.), 2.) und 3.) eintreten kann. Für 1.) und 2.) bzw. 3.) ist dies offensichtlich. Würden etwa 2.) und 3). gleichzeitig gelten, so wäre

$$a = b + k = (a + l) + k = a + (l + k),$$

was wiederum dem Satz 4 widerspricht.

Wir zeigen nun, daß auch stets einer dieser Fälle eintritt. Sei M für gegebenes a die Menge der b, für die einer der Fälle 1.), 2.) oder 3.) gilt.

A.) Ist $a = 1$, so tritt für $b = 1$ der Fall 1.) ein. Ist $a \neq 1$, so gibt es nach Satz 2 aus § 11 ein k derart, daß $a = k' = k + 1 = 1 + k$ ist, d. h., es gilt dann für $b = 1$ der Fall 2.). Also gehört 1 zu M.

B.) Wenn b zu M gehört, so gilt: Entweder ist $a = b$ und damit $b' = b + 1 = a + 1$, d. h., es gilt für b' der Fall 3.). Oder es ist $a = b + k$; dann ist, falls $k = 1$ ist, $a = b + 1 = b'$, d. h., dann tritt für b' der Fall 1.) ein; ist hingegen $k \neq 1$, so ist $k = m'$ und mithin

$$a = b + m' = b + (m + 1) = b + (1 + m) = (b + 1) + m = b' + m,$$

d. h., dann tritt für b' der Fall 2.) ein. Oder es ist $b = a + l$; dann ist

$$b' = (a + l)' = a + l',$$

d. h., dann tritt auch für b' der Fall 3.) ein. In allen Fällen gehört also b' zu M. Damit ist der behauptete Satz bewiesen.

Unter Benutzung dieses letzten Satzes könnten wir bereits jetzt eine Ordnung für natürliche Zahlen definieren und ihre grundlegenden Eigenschaften beweisen (siehe § 14), jedoch wollen wir hier zunächst noch die Multiplikation für natürliche Zahlen einführen, um später sofort den Begriff der Ordnung in seinen Beziehungen zu den beiden Grundoperationen studieren zu können.

Aufgaben. Man beweise unter Benutzung der Definitionen

$$2 = 1', \quad 3 = 2', \quad 4 = 3', \quad 5 = 4', \quad 6 = 5'$$

und der induktiven Definition der Summe, daß

$$1 + 1 = 2, \quad 1 + 2 = 3, \quad 2 + 2 = 4, \quad 2 + 3 = 5, \quad 2 + 4 = 3 + 3 = 6$$

ist.

§ 13. Multiplikation

Definition. *Unter einer Multiplikation im Bereich der natürlichen Zahlen verstehen wir eine Vorschrift, die jedem geordneten Paar von natürlichen Zahlen a, b eine und nur eine natürliche Zahl ab* (oder $a \cdot b$ oder $a \times b$) *zuordnet und die die folgenden Bedingungen erfüllt*:

1.) $a \cdot 1 = a$ *für jede natürliche Zahl* a;

2.) $ab' = ab + a$ *für beliebiges* a *und* b.

Die Zahl a heißt der Multiplikand, die Zahl b der Multiplikator, beide zusammen auch Faktoren; die Zahl ab heißt ihr Produkt.

Auf den ersten Blick mag es seltsam erscheinen, daß hier diese komplizierte induktive Definition gegeben wird, anstatt bei der üblichen Schul-Definition des Produktes ab als Summe von b Summanden, die alle gleich dem Multiplikanden a sind, zu bleiben. Was soll hier aber „b Summanden" bedeuten, wobei offenbar b die Rolle einer Kardinalzahl spielt? Der Begriff „Anzahl der Summanden" hat für uns eine ganz fest umrissene Bedeutung, er bezeichnet nämlich die Mächtigkeit der Menge der Summanden (siehe § 3, Definition 4). Es ist zwar richtig, daß wir für endliche Mengen (mit denen wir es ja bei der Definition der Multiplikation zu tun haben) eine andere Definition für „Anzahl der Elemente" gegeben haben (siehe § 4, Definition 3) und daß wir gezeigt haben, daß sie mit dem Begriff „Anzahl der Elemente" als Mächtigkeit einer gewissen endlichen Menge übereinstimmt. Aber wir haben dabei wesentlich den Begriff des Abschnittes $|1, n|$ der Folge der natürlichen Zahlen als Menge der natürlichen Zahlen, die nicht größer als n sind, benutzt. Dieser Begriff setzt also bereits eine bestimmte Ordnung der Menge der natürlichen Zahlen voraus. Es ist nun weiter richtig, daß wir die Ordnung der natürlichen Zahlen vor der Multiplikation einführen könnten und daß wir dann mittels Definition 3 aus § 4 eine Vorschrift angeben könnten, die es gestattet, die natürlichen Zahlen mit den Mächtigkeiten endlicher Mengen zu identifizieren. Auf diese Weise würden die natürlichen Zahlen einen anzahlmäßigen Charakter erhalten. Indes bedarf die Arithmetik der natürlichen Zahlen eines solchen nicht. Man kann vielmehr die natürlichen Zahlen allein auf Grund der Definition 1 aus § 11 (durch sukzessives Weiterzählen von der Zahl 1 aus — *Anm. d. wissenschaftl. Red.*) aufbauen, ohne den Begriff der Mächtigkeit zu benutzen. Die auf diesem Wege konstruierten Zahlen heißen *Ordinalzahlen* (Ordnungszahlen) im Unterschied zu den Mächtigkeiten, den sogenannten *Kardinalzahlen*.

Damit die Theorie der natürlichen Zahlen kein leerer Formalismus bleibt, sondern zu dem grundlegenden Mittel der praktischen Tätigkeit des Menschen wird, das sie tatsächlich ist, ist es unumgänglich, eine Beziehung zwischen den Mächtigkeiten der endlichen Mengen und den unabhängig von ihnen konstruierten natürlichen Ordinalzahlen herzustellen und diesen damit einen anzahlmäßigen Sinn beizulegen. Darin besteht gerade die Bedeutung der Definition 3 und des Satzes 2 aus § 4.

Bezüglich der induktiven Definition der Multiplikation gelten alle die Bemerkungen, die wir im vorangehenden Paragraphen anläßlich der Definition der Addition gemacht haben. Insbesondere folgt aus ihr noch nicht unmittelbar die Existenz einer Vorschrift mit den angegebenen Eigenschaften. Aus diesem Grunde ist auch hier das dem Satz 1 aus § 12 entsprechende Theorem von prinzipieller Bedeutung.

Satz 1. ***Es gibt genau eine Multiplikation im Bereich der natürlichen Zahlen, d. h., die Multiplikation ist stets eindeutig ausführbar.***

Beweis. a) Wir zeigen zunächst wieder, daß es zu vorgegebenem a höchstens eine Vorschrift gibt, die jeder Zahl b eine Zahl x_b zuordnet und die die Bedingungen $x_1 = a$, $x_{b'} = x_b + a$ erfüllt. Ist nämlich y_b eine beliebige Vorschrift, die diese Bedingungen erfüllt, so verstehen wir unter M die Menge der b, für die $x_b = y_b$ ist.

A.) Wegen $x_1 = a = y_1$ gehört 1 zu M.

B.) Wenn b zu M gehört, so ist $x_{b'} = x_b + a = y_b + a = y_{b'}$, d. h., dann gehört auch b' zu M.

Auf Grund von Axiom IV. ist dann $x_b = y_b$ für jedes b. Damit ist die Einzigkeit der Multiplikation für gegebenes a bewiesen und, da a ganz beliebig gewählt war, auch für jedes a und b.

b) Wir zeigen nun, daß es zu jedem a eine (und wegen a) auch nur eine) Vorschrift gibt, die jeder Zahl b eine Zahl ab zuordnet und die Bedingungen $a \cdot 1 = a$, $ab' = ab + a$ für beliebiges b (und das vorgegebene a) erfüllt. Dazu sei M die Menge aller Zahlen a, für die es eine solche (und wegen a) dann auch nur diese) Vorschrift gibt.

A.) Für $a = 1$ setzen wir $ab = b$ für alle b. Diese Vorschrift besitzt offenbar die verlangten Eigenschaften, denn es ist

$$a \cdot 1 = 1 = a, \quad ab' = b' = b + 1 = ab + a$$

und mithin 1 Element von M.

B.) Wenn a zu M gehört, so gibt es zu jedem b eine Zahl ab, wobei $a \cdot 1 = a$ und $ab' = ab + a$ ist. Für die Zahl a' wählen wir dann die folgende Vorschrift: Jeder Zahl b wird die Zahl $a'b = ab + b$ zugeordnet. Diese Vorschrift besitzt dann offenbar auch die verlangten Eigenschaften, denn es ist

$$a' \cdot 1 = a \cdot 1 + 1 = a + 1 = a',$$

$$a' \cdot b' = ab' + b' = (ab + a) + b' = ab + (a + b') = ab + (a + b)' = ab + (b + a)'$$
$$= ab + (b + a') = (ab + b) + a' = a' \cdot b + a'.$$

Also gehört dann auch a' zu M.

Damit ist zu jedem a eine Vorschrift konstruiert, die für das vorgegebene a und jedes b, also für jedes a und b die verlangten Bedingungen erfüllt, was zu beweisen war.

Satz 2. (Rechtsseitiges distributives Gesetz)

$$(a + b)\, c = ac + bc.$$

Beweis. Wir beweisen diesen Satz für gegebenes a und b durch vollständige Induktion über c.

A.) $(a + b) \cdot 1 = a + b = a \cdot 1 + b \cdot 1$, d. h., der Satz gilt für $c = 1$.

B.) Wenn der Satz für c gilt, so ist $(a + b)\, c = ac + bc$. Auf Grund des assoziativen und des kommutativen Gesetzes der Addition gilt dann

$$(a + b)\, c' = (a + b)\, c + (a + b) = (ac + bc) + (a + b)$$
$$= (ac + a) + (bc + b) = ac' + bc',$$

d. h., unser Satz gilt auch für c'.

Auf Grund von Axiom IV. ist damit unser Satz bewiesen.

Satz 3. (Kommutatives Gesetz der Multiplikation)

$$ab = ba.$$

Beweis. a) Zunächst beweisen wir den Satz für $a = 1$ durch vollständige Induktion über b, also $1 \cdot b = b \cdot 1$. Es sei M die Menge aller b mit dieser Eigenschaft.

A.) 1 gehört zu M.

B.) Wenn $1 \cdot b = b \cdot 1$ ist, so ist

$$1 \cdot b' = 1 \cdot b + 1 = b \cdot 1 + 1 = b + 1 = b' = b' \cdot 1,$$

also auch b' Element von M.

b) Jetzt beweisen wir durch vollständige Induktion über a, daß $ab = ba$ für alle b gilt. Dazu sei M die Menge aller natürlichen Zahlen a mit $ab = ba$.

A.) Wegen a) ist 1 Element von M.

B.) Wenn a zu M gehört, so ist $ab = ba$, also auf Grund von Satz 2:

$$a' \cdot b = (a + 1)\, b = ab + 1 \cdot b = ba + b \cdot 1 = ba + b = ba',$$

d. h., auch a' ist Element von M.

Satz 4. (Linksseitiges distributives Gesetz)

$$c(a + b) = ca + cb.$$

Der Beweis hierfür ergibt sich unmittelbar aus Satz 2 und Satz 3.

Satz 5. (Assoziatives Gesetz der Multiplikation)

$$(ab)c = a(bc).$$

Beweis. Es sei M für vorgegebenes a und b die Menge der c, für die diese Gleichung gilt.

A.) $(ab) \cdot 1 = ab = a(b \cdot 1)$, d. h., 1 gehört zu M.

B.) Wenn c zu M gehört, so ist $(ab)c = a(bc)$. Durch Anwendung von Satz 4 erhalten wir dann

$$(ab)c' = (ab)c + ab = a(bc) + ab = a(bc + b) = a(bc'),$$

d. h., dann ist auch c' Element von M.

Damit ist Satz 5 bewiesen.

Aufgaben. Man zeige unter Benutzung der früheren Definitionen $2 = 1'$, $3 = 2'$, $4 = 3'$, ..., daß $2 \cdot 2 = 4$, $3 \cdot 2 = 6$ ist.

§ 14. Ordnung

Bei der Definition der natürlichen Zahlen (§ 11, Definition 1) gingen wir von der Grundbeziehung „b folgt auf a" aus. Schon die Wahl des Wortes „folgt" läßt auf einen Zusammenhang zwischen dieser Grundbeziehung und dem Begriff der Ordnung schließen, wie wir ihn in § 5 für beliebige Mengen eingeführt haben. Zwar zeigen die Axiome II. und III., daß sich die Beziehung „folgt auf" für Zahlen von der genauso

benannten Beziehung der Ordnung unterscheidet, da sie lediglich jedes Element mit höchstens zwei „benachbarten" Elementen verbindet, denn auf Grund von Axiom II. folgt auf jede Zahl nur eine andere, und auf Grund von Axiom III. ist jede Zahl Nachfolger höchstens einer anderen. Man kann indes im Bereich der natürlichen Zahlen eine Ordnung einführen, die für a und a' mit der Grundbeziehung „folgt auf" zusammenfällt. Für diese neue Beziehung verwendet man üblicherweise das Wort „größer".

Definition. *Wenn es zu gegebenen Zahlen a und b eine Zahl k gibt derart, daß $a = b + k$ ist, so nennen wir a größer als b bzw. b kleiner als a und schreiben $a > b$ bzw. $b < a$. Wenn $a > b$ oder $a = b$ ist, so schreiben wir $a \geq b$. Wenn $a < b$ oder $a = b$ ist, so schreiben wir $a \leq b$.*

Satz 1. a) *Für Zahlen a, b gilt eine und nur eine der drei Beziehungen: $a = b$, $a > b$ oder $b > a$.* b) *Wenn $a > b$ und $b > c$ ist, so ist $a > c$.*

Mit anderen Worten: Die oben definierte Beziehung „größer" ist eine Ordnung der Menge N der natürlichen Zahlen im Sinne von Definition 1 aus § 5 (daß wir in § 5 diese Grundbeziehung durch das Zeichen „$<$" ausgedrückt haben, ist ohne Bedeutung).

Beweis. Die Behauptung a) ist nur eine andere Schreibweise für Satz 5 aus § 12. Der Beweis für die Behauptung b) ergibt sich folgendermaßen: Ist $a > b$ und $b > c$, so ist $a = b + k$ und $b = c + l$ und mithin

$$a = b + k = (c + l) + k = c + (l + k),$$

also

$$a > c.$$

Die Beziehung „größer" fällt für unmittelbar aufeinander folgende Zahlen mit der Beziehung „folgt auf" zusammen, da $a' = a + 1$, also $a' > a$ ist. Was die Verknüpfung der Ordnung mit der Addition und der Multiplikation anbetrifft, so bleiben auch für die natürlichen Zahlen die meisten der in § 10 für angeordnete Ringe bewiesenen Sätze gültig. Da aber die natürlichen Zahlen keinen Ring bilden, müssen wir hier diese Sätze (die wir dort aus den Ringeigenschaften abgeleitet haben) neu beweisen.

Satz 2. (Monotoniegesetze für die Addition und die Multiplikation) *Wenn* a) $a \gtreqless b$ *ist, so ist entsprechend* b) $a + c \gtreqless b + c$ *und* c) $ac \gtreqless bc$.

Beweis. 1.) Wenn $a > b$ ist, so ist $a = b + k$ und mithin

$$a + c = (b + k) + c = c + (b + k) = (c + b) + k = (b + c) + k,$$

also $a + c > b + c$ und ferner $ac = (b + k)\, c = bc + kc > bc$.

2.) Wenn $a = b$ ist, dann ist wegen der Eindeutigkeit der Addition und der Multiplikation $a + c = b + c$ und $ac = bc$.

3.) Wenn $a < b$ ist, so ist $b > a$ und auf Grund des unter 1.) Bewiesenen $b + c > a + c$ und $bc > ac$, also $a + c < b + c$ und $ac < bc$.

Ferner gelten auch die Umkehrungen der Behauptungen von Satz 2.

Satz 3. *Wenn $a + c \gtreqless b + c$ oder $ac \gtreqless bc$ ist, so ist $a \gtreqless b$.*

Beweis. Da in Satz 2 die Voraussetzungen alle Möglichkeiten ausschöpfen und die Behauptungen sich paarweise ausschließen, gelten auch die Umkehrungen (siehe den Beweis von Satz 3 aus § 10).

Aus Satz 2 erhält man durch wörtliche Wiederholung des Beweises von Satz 4 aus § 10 die bekannten Rechenregeln für das Rechnen mit Ungleichungen.

Satz 4. *Wenn* $a \gtreqless b$ *und* $c \gtreqless d$ *ist, so ist entsprechend* $a + c \gtreqless b + d$ *und* $ac \gtreqless bd$.

Satz 5. *Die Zahl* 1 *ist die kleinste natürliche Zahl, d. h., es ist* $a \geq 1$ *für jedes* a.

Beweis. Wenn $a \neq 1$ ist, so ist (§ 11, Satz 2) $a = b' = b + 1 > 1$.

Satz 6. *In der Menge der natürlichen Zahlen gilt das Archimedische Axiom* (§ 10, Definition 3), *d. h., zu jedem* a *und* b *gibt es ein* c *derart, daß* $bc > a$ *ist.*

Beweis. Dazu genügt es, ein $c > a$ zu nehmen, da dann aus $b \geq 1$ (Satz 5) auf Grund von Satz 2 und Satz 4 die Ungleichung $bc > a \cdot 1 = a$ folgt.

Satz 7. *Bezüglich der oben definierten Ordnung der natürlichen Zahlen sind die Zahlen* a *und* $a + 1$ *benachbart* (§ 5), *d. h., es gibt keine Zahl* b *mit* $a + 1 > b > a$. *Insbesondere folgt aus* $b > a$, *daß* $b \geq a + 1$ *ist, und aus* $b < a + 1$, *daß* $b \leq a$ *ist.*

Beweis. Wenn $b > a$ ist, so ist $b = a + k$. Auf Grund von Satz 5 ist $k \geq 1$, also ist auf Grund von Satz 2 auch $a + k \geq a + 1$, d. h. $b \geq a + 1$. Wegen Satz 1 kann dann aber der Fall $a + 1 > b$ nicht eintreten, was zu beweisen war.

Sehr oft wird der folgende Satz benötigt:

Satz 8. *In jeder nichtleeren Menge* A *von natürlichen Zahlen gibt es eine kleinste, also eine Zahl, die kleiner als alle anderen Zahlen der betrachteten Menge ist.*

Beweis. Es sei M die Menge aller natürlichen Zahlen a mit $a \leq b$ für alle Zahlen b der Menge A. Auf Grund von Satz 5 gehört 1 zu M. Es gehören aber sicher nicht alle Zahlen zu M, denn wenn b zu A gehört, so gehört die Zahl $a = b + 1 > b$ nicht zu M. Also gibt es eine Zahl a derart, daß a zu M gehört, aber $a + 1$ nicht zu M gehört (denn anderenfalls würde M nach Axiom IV. alle natürlichen Zahlen enthalten). Da a zu M gehört, ist $a \leq b$ für alle b aus A. Ferner gehört a auch zu A, da anderenfalls $a < b$ für alle b aus A gelten würde und auf Grund von Satz 7 dann $a + 1 \leq b$ für alle b aus A wäre, d. h. $a + 1$ zu M gehören würde, was der Wahl von a widerspricht.

Auf diesen Satz gründet sich eine zweite Form der induktiven Beweise:

Satz 9. (Vgl. § 11, Satz 1.) *Wenn ein gewisses Theorem* T *für die Zahl* 1 *gilt und wenn* T *unter der Voraussetzung, daß* T *für alle Zahlen kleiner als* n *gilt, auch für die Zahl* n *gilt, so gilt* T *für alle natürlichen Zahlen.*

Beweis. Würde T nicht für alle natürlichen Zahlen gelten, so wäre die Menge M der Zahlen n, für die es nicht gilt, nicht leer. Nach Satz 8 enthielte dann diese Menge eine kleinste Zahl n. Da n zu M gehört, gilt für n das Theorem T nicht, also ist $n > 1$. Nun sollte aber n die kleinste Zahl aus M sein, also gilt das Theorem T für alle Zahlen, die kleiner als n sind, also nach Voraussetzung auch für n, was jedoch nicht der Fall sein sollte.

Nach Einführung der Ordnung für die natürlichen Zahlen können wir die ursprüngliche Fassung der induktiven Beweise, d. h. Satz 1 aus § 11 folgendermaßen modifizieren:

Satz 10. *Wenn ein gewisses Theorem* T *für die Zahl* k *gilt und wenn es unter der Voraussetzung, daß es für eine beliebige Zahl* $n \geq k$ *gilt, auch für* $n + 1$ *gilt, so gilt dieses Theorem für jede natürliche Zahl* $n \geq k$.

Beweis. Würde das Theorem T nicht für alle Zahlen $n \geq k$ gelten, so wäre die Menge A der Zahlen $n \geq k$, für die T nicht gilt, nicht leer. Dann gäbe es nach Satz 8 eine kleinste Zahl $l \geq k$, für die das Theorem T nicht gilt. Dann muß aber sogar

$l > k$ sein. Auf Grund von Satz 5 ist dann $l \neq 1$ und besitzt mithin einen unmittelbaren Vorgänger n (§ 11, Satz 2), d. h. eine Zahl n, für die $n' = n + 1 = l$ ist. Hierbei ist $n \geq k$, denn wäre $n < k$, so wäre auf Grund von Satz 7 die Zahl $l = n + 1 \leq k$. Wegen $l = n + 1$ ist $n < l$, also gehört n nicht zu A, d. h., für n gilt das Theorem T. Dann müßte nach Voraussetzung aber auch für $n + 1 = l$ das Theorem T gelten. Dieser Widerspruch beweist unseren Satz.

Eine analoge Modifizierung gestattet auch die zweite Fassung der induktiven Beweise (d. h. Satz 9), und zwar:

Satz 11. *Wenn ein gewisses Theorem T über natürliche Zahlen für die Zahl k gilt und wenn es unter der Voraussetzung, daß es für alle Zahlen a mit $k \leq a < n$ gilt, auch für die Zahl n gilt, so gilt dieses Theorem für alle natürlichen Zahlen n mit $n \geq k$.*

Der Beweis hierfür verläuft analog dem Beweis für Satz 10 und sei daher dem Leser überlassen.

In Ergänzung zum Satz 8 gilt:

Satz 12. *Jede nichtleere, nach oben beschränkte Menge A von natürlichen Zahlen enthält eine größte Zahl* (dabei wird unter einer nach oben beschränkten Menge eine Menge verstanden, deren sämtliche Elemente kleiner sind als eine gewisse natürliche Zahl k).

Beweis. Es sei B die Menge der natürlichen Zahlen, die nicht kleiner sind als die Zahlen der Menge A. Da A nach oben beschränkt ist, ist B nicht leer. Auf Grund von Satz 8 enthält B eine kleinste Zahl b. Nach Definition von B ist $b \geq a$ für jedes a aus A. Wir zeigen nun, daß die Zahl b zu A gehört und folglich größte Zahl in A ist. Wäre nämlich b nicht Element von A, so wäre $b > a$ für alle a aus A. Auf Grund von Satz 7 wäre dann $b - 1 \geq a$ für alle a aus A, also würde dann noch $b - 1$ zu B gehören, was zusammen mit $b - 1 < b$ einen Widerspruch zur Wahl von b ergibt.

§ 15. Induktive Definitionen. Summe und Produkt aus mehreren Zahlen

Zwei wichtige Spezialfälle der induktiven Definition haben wir bereits anläßlich der Definition der Addition und der Multiplikation kennengelernt. In beiden Fällen ging es darum, für ein vorgelegtes a eine gewisse Funktion $f(b)$ von b zu konstruieren (deren Werte natürliche Zahlen sind), welche folgende Eigenschaften besitzt: 1.) Für $b = 1$ nimmt die Funktion einen vorgeschriebenen Wert an (im Falle der Addition war $f(1) = a'$, im Falle der Multiplikation war $f(1) = a$). 2.) Der Wert der Funktion für eine von 1 verschiedene Zahl ist durch eine Rekursionsgleichung eindeutig durch den Wert der Funktion für die vorangehende Zahl festgelegt (im Falle der Addition war $f(b') = [f(b)]'$, im Falle der Multiplikation war $f(b') = f(b) + a$).

Anläßlich der Definition der Addition (§ 12) haben wir bereits gesagt, daß aus einer solchen Definition noch nicht unmittelbar (als einfache Folgerung aus dem Axiom IV. der vollständigen Induktion) die Existenz und Einzigkeit einer Funktion mit den Eigenschaften 1.) und 2.) folgt. Indes konnten wir im Falle der Addition und

der Multiplikation auf anderem Wege sowohl die Existenz als auch die Einzigkeit der entsprechenden Funktionen beweisen. Nachdem wir für die natürlichen Zahlen die Ordnung eingeführt haben, können wir ganz allgemein die induktiven Definitionen rechtfertigen, und zwar in einer noch etwas allgemeineren Form als im Falle der Addition und der Multiplikation.

Definition 1. *Unter einer induktiven Definition (oder Konstruktion) einer Funktion $f(a)$ auf der Menge der natürlichen Zahlen verstehen wir ihre Definition durch die folgenden beiden Angaben:*

1.) *Der Wert der Funktion für das Argument 1 ist vorgegeben, etwa $f(1) = x_1$.*

2.) *Es ist ein System S von Rekursionsgleichungen vorgegeben, mit dessen Hilfe sich der Wert $f(a)$ der Funktion f für eine natürliche Zahl $a > 1$ eindeutig aus den Werten $f(b)$ für die natürlichen Zahlen $b < a$ bestimmen läßt.*

Wir bemerken, daß die Werte einer induktiv definierten Funktion $f(a)$ keineswegs natürliche Zahlen zu sein brauchen. Sie können vielmehr Elemente eines beliebigen Ringes oder noch allgemeiner Elemente einer beliebigen Menge A sein, zwischen deren Elementen gewisse Beziehungen erklärt sind, die in die Rekursionsgleichungen des Systems S eingehen.

Daß eine solche induktive Definition tatsächlich eine (und zwar eindeutig bestimmte) Funktion $f(a)$ definiert, zeigt der folgende

Satz 1. (Satz über die Rechtfertigung der induktiven Definitionen)[1] *Zu jedem vorgegebenen System S von Rekursionsgleichungen gibt es eine und nur eine Funktion $f(a)$, die auf der Menge der natürlichen Zahlen erklärt ist und die in Definition 1 angegebenen Bedingungen erfüllt.*

Wir beweisen dazu zunächst das folgende

Lemma. *Vorgegeben seien* a) *eine natürliche Zahl n,* b) *ein Element x_1 aus einer gewissen Menge A,* c) *falls $n > 1$ ist, ein System S von Rekursionsgleichungen, durch welche für jede natürliche Zahl a mit $1 < a \leq n$ der Menge der Elemente x_b mit $b < a$ aus A eindeutig ein Element x_a der Menge A zugeordnet ist.*[2] *Dann gibt es eine und nur eine Funktion $f_n(a)$, die auf dem Abschnitt*[3] $|1, n|$ *definiert ist, deren Werte der Menge A angehören und die folgende Eigenschaften besitzt:* 1.) $f_n(1) = x_1$, 2_n.) *Wenn $n > 1$ und $1 < a \leq n$ ist, so ist der Wert $f_n(a)$ mit den Werten $f_n(b)$ für $b < a$ durch die Rekursionsgleichungen des Systems S verknüpft.*

Beweis des Lemmas. Es sei M die Menge aller natürlichen Zahlen n, für die das Lemma gilt.

A.) Für $n = 1$ entfallen die Bedingung c) und die Eigenschaft 2_n.). In diesem Fall ist offensichtlich $f_1(1) = x_1$ die einzige Funktion auf dem Abschnitt $|1, 1|$, die die Eigenschaft 1.) besitzt. Also gehört 1 zu M.

B.) Wenn n zu M gehört, so gilt das Lemma für n. Die Voraussetzungen a), b) und c) seien nun für die Zahl $n + 1$ erfüllt. Dann sind eo ipso diese Voraussetzungen [und zwar bezüglich desselben Systems S wie in Voraussetzung c) für $n + 1$ und des-

[1] Dieser Satz geht zurück auf R. Dedekind, „Was sind und was sollen die Zahlen?", Braunschweig 1888, 1893, 1911, 1930, Berlin, 1965. Auf dieses Büchlein sei der interessierte Leser besonders hingewiesen, ferner auf das Buch von Rózsa Péter, „Rekursive Funktionen", 2. Aufl., Berlin 1957, das wegen seiner elementaren Art besonders dem noch nicht so weit fortgeschrittenen Leser zu empfehlen ist. — *Anm. d. wissenschaftl. Red.*

[2] Unter Umständen brauchen die rekursiven Beziehungen für $a > n$ gar nicht erklärt zu sein.

[3] Unter dem Abschnitt $|1, n|$ der Folge der natürlichen Zahlen verstehen wir (entsprechend § 4, Definition 1) die Menge der natürlichen Zahlen $a \leq n$.

selben x_1 wie in Voraussetzung b) für $n+1$] auch für die Zahl n erfüllt. Mithin gibt es eine und nur eine auf dem Abschnitt $|1, n|$ definierte Funktion $f_n(a)$, die die Eigenschaften 1.) und 2_n.) besitzt. Wir konstruieren nun eine Funktion $f_{n+1}(a)$ auf folgende Weise: Für jedes $a \leq n$ setzen wir $f_{n+1}(a) = f_n(a)$. Den Wert $f_{n+1}(n+1)$ definieren wir durch die Werte $f_{n+1}(a)$ mit $a < n+1$ mittels der Rekursionsgleichungen des gegebenen Systems S. Dies ist möglich, da die Voraussetzung c) für die Zahl $n+1$ erfüllt sein soll. Damit ist die Funktion $f_{n+1}(a)$ auf dem Abschnitt $|1, n+1|$ erklärt, auf dem sie nach Konstruktion auch die Eigenschaften 1.) und 2_{n+1}.) besitzt. Ist nun $g(a)$ eine beliebige Funktion, die auf dem Abschnitt $|1, n+1|$ definiert ist und die Eigenschaften 1.) und 2_{n+1}.) besitzt, so ist sie erst recht auf dem Abschnitt $|1, n|$ erklärt und besitzt dort die Eigenschaften 1.) und 2_n.). Wegen der Einzigkeit der Funktion $f_n(a)$ (für n sollte ja das Lemma bereits gelten), ist $g(a) = f_n(a)$ für alle $a \leq n$. Ferner sollte $g(a)$ die Eigenschaft 2_{n+1}.) besitzen. Also ist der Wert $g(n+1)$ eindeutig durch die Werte $g(a)$ mit $a < n+1$ festgelegt. Nun ist aber für $a < n+1$, also für $a \leq n$ bereits

$$g(a) = f_n(a) = f_{n+1}(a).$$

Dann ist aber auch $g(n+1) = f_{n+1}(n+1)$. Also stimmt auf dem ganzen Abschnitt $|1, n+1|$ die Funktion $g(a)$ mit der Funktion $f_{n+1}(a)$ überein, womit die Einzigkeit der Funktion $f_{n+1}(a)$ bewiesen ist. Damit ist das Lemma für die Zahl $n+1$ bewiesen, d. h., $n+1$ gehört der Menge M an. Auf Grund von Axiom IV. enthält dann M alle natürlichen Zahlen, d. h., das Lemma gilt für jede natürliche Zahl n.

Beweis von Satz 1. Die Bedingungen 1.) in Definition 1 und im Lemma fallen zusammen. Die Bedingung 2.) aus Definition 1 folgt daraus, daß die Bedingung c) des Lemmas für jedes $n > 1$ gilt. Entsprechend dem Lemma gibt es nämlich zu jedem n eine und nur eine auf dem Abschnitt $|1, n|$ definierte Funktion $f_n(a)$, die die Eigenschaften 1.) und 2_n.) besitzt. Ist nun $m < n$, so ist die Funktion $f_n(a)$ auch auf dem Abschnitt $|1, m|$ als Teilmenge des Abschnittes $|1, n|$ erklärt und besitzt dort die Eigenschaften 1.) und 2_n.) und folglich auch die Eigenschaft 2_m.). Auf Grund der Einzigkeit der Funktion $f_m(a)$ ist dann $f_n(a) = f_m(a)$ für alle $a \leq m$. Also besitzen alle Funktionen $f_n(a)$, die an der Stelle a definiert sind (für die also $n \geq a$ ist), dort ein und denselben Wert. Den gemeinsamen Wert $f_n(a)$ aller f_n mit $n \geq a$ nehmen wir nun als Wert der gesuchten Funktion f an der Stelle a.[1]) Der Wert $f(1)$ ist dann also $f_n(1)$ für beliebiges n, und da alle Funktionen $f_n(a)$ die Eigenschaft 1.) besitzen, so besitzt auch $f(a)$ die Eigenschaft 1.). Ist $a > 1$ und $n \geq a$, so ist $f(a) = f_n(a)$, also genügt $f(a)$ auch den Rekursionsgleichungen und besitzt ebenfalls die Eigenschaft 2.). Ist schließlich $g(a)$ eine beliebige, auf der Menge der natürlichen Zahlen definierte Funktion, die die Eigenschaften 1.) und 2.) aus Definition 1 besitzt, so ist sie auf jedem Abschnitt $|1, n|$ definiert und besitzt dort die Eigenschaften 1.) und 2_n.). Wegen der Einzigkeit einer solchen Funktion ist dann aber $g(a) = f_n(a) = f(a)$ für alle $n \geq a$, also $g(a) = f(a)$ für jedes a. Damit ist gezeigt, daß es höchstens eine und daher genau eine Funktion $f(a)$ mit den verlangten Eigenschaften gibt.

Auf das oben bewiesene Lemma gründet sich die Einführung der allgemeinen Summe und des allgemeinen Produktes aus endlich vielen natürlichen Zahlen.

---

[1]) Als geschlossene Formel könnten wir etwa $f(n) = f_n(n)$ setzen. — *Anm. d. wissenschaftl. Red*

Definition 2. *Vorgegeben sei eine natürliche Zahl n und eine Folge $a_1, a_2, \ldots, a_n$*[1] *von natürlichen Zahlen.*[2]
Die Summe dieser Zahlen, in Zeichen:

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

wird definiert durch die Bedingungen

$$\sum_{i=1}^{1} a_i = a_1, \tag{1}$$

$$\sum_{i=1}^{k+1} a_i = \sum_{i=1}^{k} a_i + a_{k+1} \tag{2}$$

für jedes $k < n$.

Das Produkt dieser Zahlen, in Zeichen:

$$a_1 a_2 \cdots a_n = \prod_{i=1}^{n} a_i,$$

wird definiert durch die Bedingungen

$$\prod_{i=1}^{1} a_i = a_1, \tag{3}$$

$$\prod_{i=1}^{k+1} a_i = \left(\prod_{i=1}^{k} a_i\right) a_{k+1} \tag{4}$$

für jedes $k < n$.

Die Bedingungen (1) und (3) legen den Wert der betrachteten Funktionen von k für $k = 1$ fest. Die Bedingungen (2) und (4) spielen die Rolle der Rekursionsgleichungen in Voraussetzung c) des obigen Lemmas. Auf Grund des Lemmas gibt es eindeutig bestimmte Funktionen $\sum_{i=1}^{k} a_i$ und $\prod_{i=1}^{k} a_i$, die auf dem Abschnitt $|1, n|$ erklärt sind und die die Eigenschaften (1) und (2) bzw. (3) und (4) besitzen. Daher hat die Definition 2 eine ganz präzise Bedeutung.

Bemerkung. Bisher haben wir beim Aufbau der Arithmetik der natürlichen Zahlen (beginnend mit § 11) keinerlei Sätze aus den ersten beiden Kapiteln benutzt. Da wir jetzt andererseits alle die Begriffe und Fakten aus der Theorie der natürlichen Zahlen (namentlich den Begriff des Abschnittes, des induktiven Beweises und der induktiven Definition) behandelt haben, die wir in den ersten beiden Kapiteln benutzten, können wir uns beim weiteren Aufbau der Theorie der natürlichen Zahlen auf die Ergebnisse aus den ersten beiden Kapiteln stützen, ohne Zirkelschlüsse befürchten zu müssen. Insbesondere gelten für die allgemeine Summe

[1] Genau genommen ist dies eine auf dem Abschnitt $|1, n|$ definierte Funktion $f(b) = a_b$.
[2] Diese Definition und alle Ergebnisse dieses Paragraphen übertragen sich wortwörtlich auf beliebige Ringe und allgemein auf beliebige Mengen, in denen eine kommutative und assoziative Addition und Multiplikation erklärt sind.

und das allgemeine Produkt die grundlegenden Eigenschaften, die wir früher [§ 6, (1) und (2)] (allein aus dem assoziativen und dem kommutativen Gesetz — *Anm. d. wissenschaftl. Red.)* abgeleitet haben:

$$\sum_{i=1}^{m} a_i + \sum_{i=1}^{n} a_{m+i} = \sum_{i=1}^{m+n} a_i; \qquad \prod_{i=1}^{m} a_i \cdot \prod_{i=1}^{n} a_{m+i} = \prod_{i=1}^{m+n} a_i, \tag{5}$$

$$\sum_{i=1}^{n} (a_i + b_i) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i; \qquad \prod_{i=1}^{n} (a_i b_i) = \prod_{i=1}^{n} a_i \cdot \prod_{i=1}^{n} b_i. \tag{6}$$

Falls in einer Summe bzw. in einem Produkt alle Summanden bzw. Faktoren untereinander gleich sind, erhalten wir die Definition des Vielfachen bzw. der entsprechenden Potenz der natürlichen Zahl a. Für sie gelten dann die bekannten Rechenregeln [§ 6, (3) bis (8)].

Als Definition des Vielfachen und der Potenz einer Zahl a gelten also die Gleichungen

$$an = \sum_{i=1}^{n} a. \tag{7}$$

$$a^n = \prod_{i=1}^{n} a. \tag{8}$$

Nun ist aber die Bezeichnung an in (7) von früher her zunächst mit einer anderen Bedeutung belastet; so haben wir nämlich das Produkt aus den natürlichen Zahlen a und n bezeichnet. Wir werden indes zeigen, daß beide Erklärungen der Schreibweise an auf dasselbe hinauslaufen. Wenn wir dies bewiesen haben, so kommen wir — indem wir der Zahl n (im Sinne der Mächtigkeit von Mengen) eine anzahlmäßige Bedeutung beilegen — zur üblichen Schul-Definition des Produktes an als Summe von n Summanden, die alle gleich a sind. Analog kommt man zur Definition der Potenz a^n als Produkt aus n Faktoren, die alle gleich a sind. Wir beweisen also:

Satz 2. *Für beliebige Zahlen a und n gilt die Gleichung*

$$an = \sum_{i=1}^{n} a, \tag{9}$$

wobei die linke Seite von (9) das Produkt aus den Zahlen a und n (im Sinne der Definition aus § 13) *bedeutet. Insbesondere ist*

$$n = 1 \cdot n = \sum_{i=1}^{n} 1,$$

d. h., jede natürliche Zahl n ist Summe aus n Einsen.

Beweis. Für $n = 1$ kommen wir entsprechend der Bedingung (1) der Definition des Produktes aus § 13 und der Bedingung (1) der allgemeinen Summe auf

$$a \cdot 1 = a = \sum_{i=1}^{1} a.$$

Ist $an = \sum_{i=1}^{n} a$, so ist auf Grund von Bedingung (2) der Produktdefinition aus § 13

und der Bedingung (2) für die allgemeine Summe

$$a(n+1) = an + a = \sum_{i=1}^{n} a + a = \sum_{i=1}^{n+1} a.$$

Damit ist nach Axiom IV. die Behauptung bewiesen.

§ 16. Subtraktion und Division

Nachdem wir die grundlegenden Fragen der Arithmetik der natürlichen Zahlen behandelt haben, die im Zusammenhang mit ihrem axiomatischen Aufbau gewisse Schwierigkeiten enthalten, wollen wir noch kurz bei den Eigenschaften der Umkehroperationen verweilen.

Definition 1. *Als Subtraktion im Bereich der natürlichen Zahlen bezeichnen wir die Umkehroperation der Addition, also die Vorschrift, die den Zahlen a und b die Zahl $a - b$ (genannt Differenz aus a und b) zuordnet, für die*

$$(a - b) + b = a \tag{1}$$

ist.

Auf Grund der Definition der Ordnung und von Satz 3 aus § 14 erhalten wir hieraus unmittelbar:

Satz 1. *Die Subtraktion $a - b$ ist dann und nur dann ausführbar, wenn $a > b$ ist. Falls die Differenz existiert, ist sie eindeutig bestimmt.*

Aus (1) erhalten wir ferner

$$a - b < a. \tag{2}$$

Hierbei wie auch im folgenden ist (wenn nicht ausdrücklich anders vermerkt) stillschweigend vorausgesetzt, daß alle vorkommenden Differenzen existieren.

Weiter erhalten wir

$$(a - b)c = ac - bc,$$

da

$$(a - b)c + bc = ac. \tag{3}$$

Aus (1) und (3) folgt, daß

a) $a - b = c - d$ dann und nur dann gilt, wenn $a + d = b + c$ ist; (4)

b) $$(a - b) + (c - d) = (a + c) - (b + d);$$

c) $$(a - b) - (c - d) = (a + d) - (b + c);$$

d) $$(a - b)(c - d) = (ac + bd) - (ad + bc).$$

Satz 2. *Wenn* a) $b \gtrless c$ *ist, so ist entsprechend* b) $a - b \lessgtr a - c$, *und umgekehrt.*

Beweis. Wir zeigen zunächst, daß a) aus b) folgt. Dazu addieren wir auf beiden Seiten von b) die Zahl $b + c$ und erhalten (§ 14, Satz 2) $a + c \lessgtr a + b$ und daraus (§ 14, Satz 3) $c \lessgtr b$, $b \gtrless c$. Entsprechend folgt auch b) aus a).

Definition 2. *Als Division bezeichnen wir die Umkehroperation der Multiplikation, die den Zahlen a und b die Zahl* $\frac{a}{b}$ *oder* $a : b$ *(genannt Quotient aus a und b) zuordnet, für die*

$$\frac{a}{b} \cdot b = a \tag{5}$$

ist.

Wegen $1 \leqq b$ ist stets

$$a \leqq ab, \tag{6}$$

wobei das Gleichheitszeichen nur im Fall $b = 1$ auftritt.

Hieraus und aus (5) folgt

$$\frac{a}{b} \leqq a, \tag{7}$$

wobei auch hier Gleichheit nur im Fall $b = 1$ eintritt.

Wie im Fall der Subtraktion sei auch hier vorausgesetzt, daß alle hingeschriebenen Quotienten tatsächlich existieren.

Satz 3. *Wenn* a) $b \gtreqless c$ *ist, so ist entsprechend* b) $\frac{a}{b} \lesseqgtr \frac{a}{c}$, *und umgekehrt.*

Der Beweis hierfür verläuft analog dem Beweis von Satz 2.

Weiterhin gelten für den Quotienten die folgenden Rechenregeln:

a) Es ist $\frac{a}{b} = \frac{c}{d}$ genau dann, wenn $ad = bc$ ist; (8)

b) $$\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd};$$

c) $$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd};$$

d) $$\frac{a}{b} : \frac{c}{d} = \frac{ad}{bc}.$$

Sie werden unter Benutzung des Satzes 3 aus § 14 wie die entsprechenden Eigenschaften in einem Körper bewiesen (§ 7, Satz 8). Ferner zeigt man, daß in den Gleichungen b), c) und d) aus der Existenz der Quotienten auf der linken Seite die Existenz des Quotienten auf der rechten Seite folgt.

Schließlich ergibt sich aus (6) und Satz 3 aus § 14:

Satz 4. *Damit der Quotient* $\frac{a}{b}$ *existiert, ist notwendig* (aber — wie wir gleich sehen werden — nicht hinreichend), *daß* $a \geqq b$ *ist. Wenn der Quotient existiert, so ist er eindeutig bestimmt.*

Daß aus $a \geqq b$ noch nicht die Existenz des Quotienten $\frac{a}{b}$ folgt, zeigen bereits sehr einfache Beispiele. So ergibt sich mit Hilfe der Zahlen

$$2 = 1', \quad 3 = 2', \quad 4 = 3',$$

daß es keine Zahl a gibt, für die $2a = 3$ ist. Wegen (6) müßte nämlich $a < 3$ sein, d. h., es müßte $a = 1$ oder $a = 2$ sein. Hingegen ist $2 \cdot 1 = 2$ und $2 \cdot 2 = 4$.

Dieser Sachverhalt bedingt einen tiefergehenden Unterschied zwischen den Eigenschaften der Subtraktion und denen der Division. Er führt auf Eigenschaften der Zahlen, die in der sogenannten *Teilbarkeitstheorie* behandelt werden.[1])

§ 17. Bemerkungen zu dem Axiomensystem der natürlichen Zahlen

Ausgehend von den Axiomen I. bis IV. (§ 11) haben wir im Vorangehenden die Arithmetik der natürlichen Zahlen aufgebaut. Wir wollen jetzt noch einmal auf die Fragen der axiomatischen Begründung dieser Theorie zurückkommen.

Zur Beurteilung des Axiomensystems einer axiomatisch aufgebauten mathematischen Theorie sind drei grundlegende Fragen zu lösen (die allerdings von ungleicher Schwierigkeit und Bedeutung sind), und zwar die Frage nach der Widerspruchsfreiheit, die Frage nach der Vollständigkeit und die Frage nach der Unabhängigkeit des vorgelegten Axiomensystems.

Widerspruchsfreiheit. Um ein vorgelegtes Axiomensystem überhaupt als brauchbar annehmen zu können, muß man sich vor allem davon überzeugen, daß die auf seiner Grundlage aufgebaute Theorie keine Widersprüche enthält, daß man also mit Hilfe dieser Axiome nicht zwei sich einander ausschließende Sätze beweisen kann. Wie kann man nun aber überhaupt die Widerspruchsfreiheit (in diesem Sinne) der Axiome eines vorgelegten Systems beweisen? Wir wollen diese Frage zunächst an Hand des Beispiels der ebenen Geometrie untersuchen. Bei ihrem axiomatischen Aufbau werden gewisse Grundbeziehungen zwischen Punkten und Geraden („Ein Punkt liegt auf einer Geraden", „Ein Punkt einer Geraden liegt zwischen zwei anderen Punkten dieser Geraden" usw.) formal (abstrakt) untersucht. Diese Begriffsbildungen sind untereinander durch die Axiome eines gewissen Systems verbunden. Eine andere axiomatische Theorie behandelt den Körper der reellen Zahlen. Im Rahmen der analytischen Geometrie werden nun bekanntlich den Punkten der Ebene Paare von reellen Zahlen (die Koordinaten der Punkte) und den Geraden gewisse Gleichungen (die Gleichungen der Geraden) zugeordnet. Dabei entsprechen den Grundbeziehungen zwischen Punkten und Geraden ganz bestimmte Zahlenbeziehungen der Zahlenpaare und der Gleichungen, und zwar so, daß den Axiomen der Geometrie Sätze der Theorie der reellen Zahlen entsprechen, die auf Grund der Axiome der reellen Zahlen beweisbar sind. Damit ordnet sich eine axiomatisch aufgebaute Theorie (die Geometrie der Ebene) als Teil in eine andere (die Theorie der reellen Zahlen) ein. Würde nun die Geometrie im oben angegebenen Sinne einen Widerspruch enthalten, so würde sich der entsprechende Widerspruch auch in der Theorie der reellen Zahlen finden (d. h., man könnte dann auch auf Grund der Axiome für die reellen Zahlen zwei einander ausschließende Sätze beweisen). Wenn also das Axiomensystem für die reellen Zahlen widerspruchsfrei ist, so gilt dies auch für das Axiomensystem der Geometrie. Damit ist in einem gewissen Sinne die Widerspruchsfreiheit der Axiome der Geometrie bewiesen.

Die Darstellung einer axiomatisch begründeten Theorie im Rahmen einer anderen Theorie, wie wir sie eben am Beispiel der Geometrie der Ebene und der Arithmetik der reellen Zahlen kennengelernt haben, spielt in der Mathematik eine große Rolle,

[1]) Über die Eigenschaften der Teilbarkeit vgl. den Artikel von A. J. Chintschin in diesem Bande.

und zwar nicht nur bei der Reduktion der Widerspruchsfreiheit einer gewissen Theorie auf die Widerspruchsfreiheit einer anderen Theorie. Wir geben daher die folgende

Definition 1. *Eine Menge, zwischen deren Elementen gewisse Beziehungen erklärt sind derart, daß alle Axiome einer axiomatisch begründeten Theorie erfüllt sind, heißt ein Modell (eine Interpretation) dieser Theorie.*

Die Konstruktion eines Modells einer gegebenen axiomatisch begründeten Theorie löst — streng genommen — nicht die Frage ihrer Widerspruchsfreiheit, sondern reduziert diese nur auf die Frage der Widerspruchsfreiheit jener Theorie, mittels derer das Modell konstruiert ist.

Die Widerspruchsfreiheit der Theorie der natürlichen Zahlen kann nicht mit formal-logischen Mitteln (zumindest nicht mit Mitteln, die weniger anspruchsvoll als die Mittel der Theorie der natürlichen Zahlen sind — *Anm. d. wissenschaftl. Red.*) bewiesen werden. Jedoch hat die vielhundertjährige menschliche Praxis gezeigt, daß die Theorie der natürlichen Zahlen keine Widersprüche enthält und daß sie mit den tatsächlichen Gegebenheiten der realen Welt übereinstimmt.

Vollständigkeit. Als zweites erhebt sich die Frage, inwieweit ein Axiomensystem eine vorgelegte Theorie gut beschreibt. Kann man vielleicht sogar mit Hilfe des gegebenen Axiomensystems jeden Satz, der mittels der Grundbegriffe der vorgelegten Theorie formuliert ist, beweisen oder widerlegen? Der österreichische Mathematiker K. GÖDEL hat im Jahre 1931 bewiesen, daß für eine Reihe von Theorien, darunter auch für die axiomatisch begründete Theorie der natürlichen Zahlen, eine vollständige Beschreibung im angegebenen Sinne (grundsätzlich) nicht erreicht werden kann, d. h., es gibt stets Sätze, die mit gegebenen Mitteln nicht entschieden werden können. Wir werden daher einen anderen Begriff von Vollständigkeit einführen, und zwar werden wir ein System dann als vollständig bezeichnen, wenn es eine Menge völlig beschreibt, d. h. bis auf Isomorphie eindeutig festlegt.

Definition 2. *Ein Axiomensystem heißt vollständig*[1]*, wenn je zwei Modelle dieses Axiomensystems isomorph* (§ 9, Definition 1) *sind.*

Ein Beispiel für ein in diesem Sinne unvollständiges Axiomensystem ist das System aus den Bedingungen I. bis VI. aus § 7, mit deren Hilfe wir den Begriff des Ringes festgelegt haben. Tatsächlich gibt es Ringe (z. B. endliche und unendliche), die untereinander nicht isomorph sind. Es ist daher nicht verwunderlich, daß man sich in der Theorie der Ringe hauptsächlich für die Beschreibung aller Typen von Ringen interessiert.

Wir wollen jetzt zeigen, daß das System der Axiome I. bis IV. für die natürlichen Zahlen im angegebenen Sinne vollständig ist. Seien dazu N_1 und N_2 zwei beliebige Modelle für dieses Axiomensystem. Die Zahlen dieser Modelle wollen wir gleichfalls durch die Indizes 1 bzw. 2 unterscheiden. Wir konstruieren dann auf dem Bereich N_1 induktiv (§ 15, Definition 1) eine Funktion $f(x_1)$ mit Werten aus N_2, für die 1.) $f(1_1) = 1_2$ und 2.) $f(a_1') = [f(a_1)]'$ ist.

Auf Grund von Satz 1 aus § 15 gibt es genau eine Funktion mit den angegebenen Eigenschaften. Wir zeigen, daß die Vorschrift $f(a_1) = a_2$ ein Isomorphismus zwischen N_1 und N_2 ist. Wenn $a_1 \neq 1_1$ ist, so ist $a_1 = b_1'$ und nithin

$$f(a_1) = f(b_1') = [f(b_1)]' \neq 1_2.$$

[1]) Zum Unterschied gegenüber dem vorher erwähnten Begriff von Vollständigkeit spricht man in der deutschen Literatur hier meistens von *Kategorizität*. — *Anm. d. wissenschaftl. Red.*

Also besitzt 1_2 bei dieser Abbildung genau ein Urbild in N_1, nämlich 1_1. Wenn nun a_2 bereits das eindeutig bestimmte Urbild a_1 besitzt, dann ist

$$f(a_1') = [f(a_1)]' = a_2',$$

also besitzt dann a_2' wenigstens ein Urbild in N_1. Ist aber b_1 ein beliebiges Urbild von a_2', so ist nach 1.) $b_1 \neq 1_1$, also $b_1 = c_1'$ und

$$a_2' = f(b_1) = f(c_1') = [f(c_1)]';$$

auf Grund von Axiom III. ist dann $a_2 = f(c_1)$ und, da a_1 das einzige Urbild von a_2 sein sollte, $c_1 = a_1$, also nach Axiom II. $b_1 = c_1' = a_1'$. Folglich ist dann a_1' das einzige Urbild von a_2'. Auf Grund des Axioms IV. der vollständigen Induktion besitzt mithin jedes Element aus N_2 bei der betrachteten Abbildung genau ein Urbild in N_1. Die Zuordnung $f(a_1) = a_2$ ist also eine eineindeutige Abbildung von N_1 auf N_2. Auf Grund von 2.) erhält die Abbildung $f(a_1) = a_2$ von N_1 auf N_2 die Grundbeziehung „folgt auf". Es bleibt also nur noch zu zeigen, daß dies auch die inverse Abbildung $f^{-1}(a_2) = a_1$ tut. Wegen $f(a_1') = [f(a_1)]' = a_2'$ ist aber in der Tat $f^{-1}(a_2') = a_1'$, d. h., auch die inverse Abbildung erhält die Beziehung „folgt auf".

Das Axiomensystem I. bis IV. ist also tatsächlich im angegebenen Sinne vollständig. Über die Bedeutung dieser Tatsache haben wir bereits in § 11 gesprochen. Dank der Vollständigkeit des Axiomensystems I. bis IV. sind alle Modelle für die natürlichen Zahlen gleichwertig (man kann also, unbeschadet der Gültigkeit der Sätze, römische oder arabische Ziffern benutzen, die Zahlen im dezimalen oder im dyadischen System darstellen usw.).

Unabhängigkeit. Weniger bedeutungsvoll und von mehr praktischem als prinzipiellem Wert ist die Frage nach der Unabhängigkeit eines Axiomensystems. Bei der Auswahl eines gewissen Axiomensystems für eine gegebene Theorie möchte man nämlich mit einer möglichst kleinen Anzahl von Sätzen auskommen, die man als Axiome nimmt. Erweist sich z. B. eines der Axiome als aus den übrigen Axiomen beweisbar, so braucht man es nicht mehr unter den Axiomen aufzuführen.

Definition 3. *Ein Axiomensystem heißt unabhängig, wenn keines seiner Axiome aus den übrigen beweisbar ist.*

Den Beweis für die Unabhängigkeit eines gegebenen Axiomensystems kann man folgendermaßen führen: Man gibt zu jedem Axiom ein Modell an, in dem alle übrigen Axiome gelten, in dem aber das ausgewählte Axiom nicht gilt. Wäre dieses Axiom nämlich eine Folgerung aus den übrigen, so gäbe es offenbar kein derartiges Modell.

Nach dieser Methode wollen wir jetzt die Unabhängigkeit der Axiome I. bis IV. für die natürlichen Zahlen beweisen. Dazu bemerken wir zunächst bezüglich des Beweises für die Unabhängigkeit des Axioms I., daß, falls das Axiom I. nicht gilt, das Axiom IV. in unserer Fassung inhaltsleer wird, da es, wenn es keine natürliche Zahl 1 gibt, auch keine Menge M von natürlichen Zahlen geben kann, die die Zahl 1 enthält. Deshalb modifizieren wir zum Beweis der Unabhängigkeit des Axioms I. von den übrigen Axiomen die Formulierung des Axioms IV. etwas und ersetzen es durch das folgende Axiom

IV'. *Jede nichtleere Menge M von natürlichen Zahlen, die die Bedingungen*

A'.) *Wenn es eine Zahl 1 gibt, die auf keine andere Zahl folgt, so gehört sie zu M;*

B.) *Wenn eine Zahl a zu M gehört, so gehört auch ihr unmittelbarer Nachfolger a' zu M;*

erfüllt, enthält alle natürlichen Zahlen.

Offensichtlich sind das Axiomensystem I. bis III. und IV. und das Axiomensystem I. bis III. und IV'. äquivalent, d. h., aus dem ersten Axiomensystem folgen die Axiome des zweiten Systems, und umgekehrt (dazu genügt es zu zeigen, daß aus den Axiomen I. bis III. und IV. das Axiom IV'. folgt und daß aus den Axiomen I. bis III. und IV'. das Axiom IV. folgt). Ist von zwei untereinander äquivalenten Axiomensystemen das eine widerspruchsfrei bzw. vollständig, so auch das andere. Mithin ist also das Axiomensystem I. bis III. und IV'. widerspruchsfrei und vollständig. Wir zeigen jetzt abschließend auch seine Unabhängigkeit.

1. Unabhängigkeit des Axioms I. Es sei N die Menge aus den drei Elementen a, b, c mit folgender Nachfolgerbeziehung:

$$a' = b, \quad b' = c, \quad c' = a.$$[1])

Da jedes der Elemente Nachfolger eines anderen ist, ist Axiom I. nicht erfüllt. Hingegen gelten in diesem Modell die Axiome II., III., und IV'. Ist nämlich $M \neq 0$ eine beliebige Teilmenge von N, also z. B. $b \in M$, so ist nach B.) auch $b' = a \in M$ und $c' = a \in M$, also $M = N$.

2. Unabhängigkeit des Axioms II. Es sei N die Menge aus den beiden Elementen a und b, $a' = b$ und a erstes Element. Dann sind alle Axiome außer II. erfüllt, da b keinen Nachfolger besitzt.

3. Unabhängigkeit des Axioms III. Es sei N die Menge aus den vier Elementen a, b, c, d mit der Nachfolgerbeziehung

$$a' = b, \quad b' = c, \quad c' = d, \quad d' = b.$$

Axiom III. gilt nicht, da b sowohl Nachfolger von a als auch von d ist, d. h., aus $a' = d'$ folgt keineswegs $a = d$. Die übrigen Axiome gelten indes, und zwar spielt a die Rolle des ersten Elementes.

4. Unabhängigkeit des Axioms IV' (oder auch IV). Es sei N die Menge aller natürlichen Zahlen $1, 2, 3, \ldots, n, \ldots$ und aller Zahlen der Form $n + \frac{1}{2}$ mit ganzem n. Die Nachfolgerbeziehung sei für die natürlichen Zahlen die übliche und für die anderen Elemente von N die Beziehung

$$\left(n + \frac{1}{2}\right)' = n + \frac{3}{2}.$$

Axiom IV'. ist nicht erfüllt. In der Tat: Die Rolle des ersten Elementes spielt die Zahl 1 selbst (sie allein ist nicht Nachfolger einer anderen Zahl). Die Menge M aller natürlichen Zahlen erfüllt dann offensichtlich die Bedingungen A'.) und B.) (oder auch A.) und B.) aus Axiom IV.), aber sie enthält nicht alle Elemente der Menge N.

Das Axiomensystem I. bis III. und IV. für die natürlichen Zahlen ist also auch unabhängig.

---

[1]) Man könnte entsprechend jede endliche Menge mit einer Elementanzahl ≥ 2 und zyklischer Anordnung der Elemente nehmen.

Kapitel IV

DER RING DER GANZEN ZAHLEN

§ 18. Das Erweiterungsprinzip in der Arithmetik und der Algebra

Der Zahlbegriff ist im Laufe einer langen geschichtlichen Entwicklung entstanden. Die natürlichen Zahlen wurden von den Menschen bereits auf einer frühen Entwicklungsstufe als Mittel zum Rechnen benutzt. Die altgriechischen Mathematiker kannten bereits die natürlichen Zahlen wie auch die Brüche, jedoch waren ihnen die negativen Zahlen noch unbekannt. Positive und negative Zahlen (in Form von „Guthaben" und „Schulden") wurden zuerst von den Indern verwendet (ĀRYABHATA [geb. 476], BRAHMAGUPTA [588 ?—660], BHĀSKARA [geb. 1114]).

Die heutige Bezeichnung der positiven und negativen Zahlen durch die Zeichen „+" und „—" findet sich erstmalig gegen Ende des 14. Jahrhunderts in dem Rechenbuch des JOHANNES WIDMANN. Trotzdem erkannten viele Mathematiker noch des 16. Jahrhunderts die negativen Zahlen nicht an. So beschränkte sich der französische Mathematiker F. VIETA (1540—1603) bei der Aufstellung der Beziehungen zwischen den Wurzeln und den Koeffizienten einer Gleichung auf den Fall positiver Wurzeln. Vollständiges Bürgerrecht erhielten die negativen Zahlen in der Mathematik erst im 17. Jahrhundert.

Die gebrochenen Zahlen traten also in der Mathematik eher in Erscheinung als die negativen Zahlen. Dies erklärt sich daraus, daß die gebrochenen Zahlen im Zusammenhang mit Messungen entstanden sind.

Abweichend von ihrer geschichtlichen Entstehung, behandeln wir hier die ganzen Zahlen vor den Brüchen, da sie vom logischen Standpunkt aus einfacher beschaffen sind als diese.

Die natürlichen Zahlen stellen das Fundament dar, auf welchem alle anderen Zahlenbereiche aufgebaut werden können. Wir werden im folgenden der Reihe nach die ganzen, die rationalen, die reellen und schließlich die komplexen Zahlen einführen. Jeder der genannten Zahlenbereiche wird sich dabei als eine Erweiterung des vorangehenden erweisen. Hierbei sind wir bestrebt, die Erweiterung jeweils so vorzunehmen, daß gewisse allgemeine Eigenschaften erfüllt sind. Die Menge B soll nämlich nur dann als Erweiterung der Menge A angesehen werden, wenn folgendes gilt:

1. A ist Teilmenge von B.

2. Die uns jeweils interessierenden Operationen oder allgemein die Beziehungen zwischen den Elementen von A sind auch für die Elemente von B erklärt und fallen für die Elemente von A (jetzt als Elemente von B aufgefaßt) mit den ursprünglichen, vor der Erweiterung betrachteten Operationen in A zusammen.

3. In B ist eine weitere Operation ausführbar, die in A entweder gar nicht oder zumindest nicht unbeschränkt ausführbar ist.

Diese bisherigen Forderungen präzisieren nur das Ziel, das man sich stets bei der Erweiterung eines Bereiches steckt. Einige Beispiele mögen dies näher erläutern: Im Bereich der natürlichen Zahlen ist — wie wir bereits gesehen haben — die Subtraktion nicht unbeschränkt ausführbar, während sie im Bereich der ganzen Zahlen immer ausführbar ist. Im Bereich der ganzen Zahlen wiederum ist die Division

nicht unbeschränkt ausführbar, während sie im Bereich der rationalen Zahlen (vom Fall der Division durch Null abgesehen, die prinzipiell unmöglich ist) stets ausführbar ist. Im Bereich der rationalen Zahlen dagegen sind gewisse Grenzübergänge nicht ausführbar, die dann im Bereich der reellen Zahlen ausführbar sind. Im Bereich der reellen Zahlen schließlich existiert nicht immer die Wurzel, während im Bereich der komplexen Zahlen diese Operation stets ausführbar ist.

Zu den angegebenen Forderungen tritt ferner noch die nachfolgende Vollständigkeitsbedingung hinzu:

4. Die Erweiterung B ist unter allen Erweiterungen von A, die die Bedingungen 1. bis 3. erfüllen, minimal und bis auf Isomorphie eindeutig durch A bestimmt.

In diesem Sinne erweitern wir die Menge der natürlichen Zahlen zunächst nur zur Menge der ganzen Zahlen und nicht gleich zur Menge der reellen oder der komplexen Zahlen.

Die Menge der ganzen Zahlen setzt sich zusammen aus der Menge der positiven (oder natürlichen) Zahlen, der Menge der negativen Zahlen und der Zahl 0. Der Begriff der negativen (gleichgültig jetzt, ob ganzen, rationalen oder reellen) Zahl hängt zusammen mit der Messung von Größen in zwei entgegengesetzten Richtungen, wie z. B. der Länge von Strecken, die auf einer Geraden von einem festen Punkt aus nach links oder nach rechts abgetragen sind, oder den Werten auf der Temperaturskala, die oberhalb oder unterhalb des Nullpunktes liegen usw. Man verabredet in solchen Fällen, daß man die Größen einer dieser Bedeutungen oder Richtungen mittels der gewöhnlichen Zahlen mißt, die dann positive Zahlen genannt werden, und daß man die Größen in der anderen Richtung mittels der gleichen Zahlen mißt, die man jedoch zur Unterscheidung von den Zahlen der ersten Art mit einem Zeichen „—" versieht und negative Zahlen nennt. Schließlich führt man rein formal eine Zahl 0 ein, die die positiven von den negativen Zahlen trennt. Ohne hier ausführlicher auf die Einführung von solchen „relativen Zahlen" einzugehen, vermerken wir nur, daß ein derartiger Aufbau wohl am natürlichsten wäre, da er der Entstehung der negativen Zahlen Rechnung trägt, und daß er sich auch streng formal durchführen läßt. So könnte man zur Konstruktion der ganzen „relativen Zahlen" eine eineindeutige Zuordnung zwischen den natürlichen Zahlen $a, b, c, \ldots$ und gewissen neuen Objekten $\bar{a}, \bar{b}, \bar{c}, \ldots$ herstellen, ein neues Objekt 0 einführen und danach die Summe, das Produkt und die „größer"-Beziehung nach den bekannten Schulregeln erklären, für die man dann (durch Betrachtung aller Fälle) die üblichen Rechen- und Ordnungsgesetze beweist.

Hier wollen wir, um in den ganzen nachfolgenden Ausführungen eine einheitliche Linie zu wahren, eine andere Konstruktionsmethode benutzen. Wenn nämlich nach der Erweiterung in B eine Operation unbeschränkt ausführbar sein soll, die in A nur beschränkt ausführbar ist, so können wir in B formal nach den gleichen Gesetzen operieren wie in A in den Fällen, in denen dort die betreffende Operation erklärt ist. Eine solche formale Übertragung der alten Gesetze auf den neuen Bereich führt nun andererseits auch zu einer Konstruktion der gesuchten Erweiterung. So ist die Differenz $a - b$ zweier natürlicher Zahlen stets eindeutig bestimmt durch das Zahlenpaar a, b. Diese Zahlenpaare werden wir nun geradezu als Ausgangspunkt für die Konstruktion der ganzen Zahlen verwenden. Der gleiche Gedanke liegt dann auch der Konstruktion sowohl der rationalen wie auch der komplexen Zahlen und der rationalen Funktionen zugrunde. Man bezeichnet diese Konstruktionsmethode gemeinhin als *Theorie der Paare*. Wir bemerken noch, daß in

allen angeführten Fällen die Konstruktion nicht sofort zur gesuchten Erweiterung B des Bereiches A führt. Vielmehr erhält man zunächst nur einen zu B isomorphen Bereich B', der eine zu A isomorphe Teilmenge A' enthält. Die gesuchte Erweiterung B erhält man dann aus B', indem man A' durch die isomorphe Menge A ersetzt. Bevor wir indes die Konstruktion der ganzen Zahlen nach diesem Muster vornehmen, müssen wir noch einige Bemerkungen über die grundlegenden Eigenschaften der Gleichheit machen.

§ 19. Äquivalenzrelationen und Äquivalenzklassen

Unter der Gleichung $a = b$ zwischen Elementen einer gewissen Menge verstehen wir hier stets diejenige Beziehung zwischen den Elementen der Menge, die im Zusammenfallen der Elemente, also in ihrer Identität, besteht.[1])

Hieraus ergeben sich allein auf Grund der Logik die folgenden fundamentalen Eigenschaften der Gleichheit: a) Es ist stets $a = a$ (Reflexives Gesetz oder Satz der Identität); b) Wenn $a = b$ ist, so ist auch $b = a$ (Symmetrie); c) Wenn $a = b$ und $b = c$ ist, so ist $a = c$ (Transitives Gesetz).

Diese Eigenschaften kommen aber — wie wir bereits wissen — auch anderen Beziehungen zu, nämlich z. B. der Gleichmächtigkeit $A \sim B$ (§ 3), der Ähnlichkeit von Mengen $A \approx B$ (§ 5) und der Isomorphie $A \cong B$ (§ 9).

Für alle derartige Beziehungen gilt der folgende fundamentale

Satz. *Jede in einer Menge M erklärte Äquivalenzrelation $a \sim b$ (in Worten: a ist äquivalent b), d. h. eine Relation, die die Bedingungen*

1.) $a \sim a$;

2.) *wenn $a \sim b$ ist, so ist auch $b \sim a$;*

3.) *wenn $a \sim b$ und $b \sim c$ ist, so ist auch $a \sim c$;*

erfüllt, definiert eine eindeutig bestimmte Einteilung von M in paarweise elementefremde, nicht leere Teilmengen (Äquivalenzklassen) derart, daß alle Elemente ein und derselben Klasse untereinander äquivalent sind und jeweils Elemente aus verschiedenen Teilmengen nicht äquivalent sind.

Umkehrung. *Zu jeder Einteilung einer Menge M in zueinander fremde, nicht leere Teilmengen läßt sich eine Äquivalenzrelation definieren, bezüglich der die vorgegebene Einteilung von M die Einteilung in Äquivalenzklassen ist.*

Beweis. a) Vorgegeben sei eine beliebige Äquivalenzrelation. Für jedes Element $a \in M$ verstehen wir unter M_a die Menge aller Elemente x mit $x \sim a$. Auf Grund der Bedingung 1.) ist $a \in M_a$, d. h., jedes Element von M gehört mindestens einer dieser Teilmengen an. Sei nun $b \in M_a$ und $c \in M_a$. Dann ist $b \sim a$ und $c \sim a$, also wegen 2.) auch $a \sim c$ und mithin wegen 3.) $b \sim c$. Es sind also je zwei Elemente aus M_a untereinander äquivalent. Wenn $a \sim b$ ist, so ist $M_a = M_b$. In der Tat: Ist $c \in M_a$, so ist $c \sim a$, also wegen $a \sim b$ auf Grund von 3.) auch $c \sim b$, d. h., es ist $c \in M_b$; ist andererseits $c \in M_b$, so ist $c \sim b$, außerdem ist wegen $a \sim b$ auf

[1]) Viele Autoren sehen den Begriff der Gleichheit als einen Begriff an, der keiner Definition, sondern einer axiomatischen Beschreibung unterliegt.

Grund von 2.) $b \sim a$, also wegen 3.) $c \sim a$, d. h., es ist $c \in M_a$. Hieraus erhalten wir: Wenn $b \in M_a$ ist, so ist $M_a = M_b$, was besagt, daß alle Elemente der Menge M_a gleichermaßen zur Festlegung dieser Menge dienen können. Wenn also M_a und M_b ein gemeinsames Element c besitzen, so ist einerseits $M_c = M_a$, andererseits $M_c = M_b$, also $M_a = M_b$. Es können also zwei voneinander verschiedene dieser Mengen kein gemeinsames Element besitzen, also auch keine untereinander äquivalenten Elemente enthalten. Die Elemente verschiedener Mengen sind untereinander nicht äquivalent.

b) Vorgegeben sei eine Einteilung von M in paarweise elementefremde Teilmengen. Eine Äquivalenzrelation in M werde wie folgt definiert: Es sei $a \sim b$ dann und nur dann, wenn a und b derselben Teilmenge der gegebenen Einteilung angehören. Es ist unmittelbar klar, daß die Äquivalenzklassen nach dieser Äquivalenzrelation gleich den Teilmengen der gegebenen Einteilung sind.

Der eben bewiesene Satz wird im folgenden wiederholte Anwendung finden. Er gestattet es, die hier durchgeführten Überlegungen in jedem konkreten Spezialfall fortzulassen.

§ 20. Definition des Ringes der ganzen Zahlen

Im Bereich der natürlichen Zahlen ist die Umkehroperation der Addition, die Subtraktion, nicht unbeschränkt ausführbar (§ 16, Satz 1). Wir stellen uns jetzt die Aufgabe, die Menge N der natürlichen Zahlen so zu einer Menge C zu erweitern, daß in C die Addition und die Multiplikation dieselben Eigenschaften wie im Bereich der natürlichen Zahlen besitzen, daß aber auch die Subtraktion immer ausführbar ist. Das bedeutet, daß C ein Ring (§ 7, Definition 1) sein muß. Unter allen derartigen Erweiterungen wählen wir die kleinste im Sinne der nachfolgenden Definition aus:

Definition 1. *Unter einem Ring der ganzen Zahlen verstehen wir einen minimalen Ring C, der die Menge N aller natürlichen Zahlen enthält, d. h. eine Menge mit den Eigenschaften:* 1.) *C enthält N;* 2.) *C ist ein Ring;* 3.) *Die Addition und die Multiplikation in C stimmen für die in C enthaltenen natürlichen Zahlen mit der Addition und Multiplikation für natürliche Zahlen überein*; 4.) *Der Ring C enthält keinen von sich selbst verschiedenen Unterring, der auch noch die Menge N enthält.*

Die Elemente von C heißen ganze Zahlen.

Auf Grund dieser Definition ist es keineswegs klar, ob es überhaupt einen solchen Ring C gibt, und wenn es einen derartigen Ring gibt, ob er eindeutig bestimmt ist. Wir zeigen zunächst unter Voraussetzung der Existenz eines derartigen Ringes, daß er bis auf Isomorphie eindeutig bestimmt ist.

Satz 1. *Ein die Menge der natürlichen Zahlen enthaltender*[1]) *Ring C ist dann und nur dann ein Ring der ganzen Zahlen (d. h. minimal), wenn jedes seiner Elemente eine Differenz aus natürlichen Zahlen ist.*

[1]) Hier wie auch im folgenden soll die Redewendung „Der Ring C enthält die Menge der natürlichen Zahlen" bzw. „Ein gewisser Ring ist in einem anderen enthalten" stets so verstanden werden, daß die Operationen in den Teilmengen mit den entsprechenden Operationen in den Obermengen übereinstimmen.

Beweis. A.) Wenn der Ring C die Menge N der natürlichen Zahlen enthält und wenn jedes Element von C eine Differenz aus natürlichen Zahlen ist, so ist C minimal, weil jeder Unterring von C, der gleichfalls die Menge N enthält, auch alle Differenzen aus natürlichen Zahlen enthält (§ 7, Satz 4) und mithin gleich C ist.

B.) Sei umgekehrt C ein minimaler Ring, der die Menge der natürlichen Zahlen enthält. In jedem Ring besitzt die Differenz aus Ringelementen die folgenden Eigenschaften (§ 7, Satz 3):

$$\left.\begin{array}{ll} \text{a)} & \text{Es ist } a-b=c-d \text{ dann und nur dann, wenn } a+d=b+c \text{ ist:} \\ \text{b)} & (a-b)+(c-d)=(a+c)-(b+d): \\ \text{c)} & (a-b)-(c-d)=(a+d)-(b+c): \\ \text{d)} & (a-b)(c-d)=(ac+bd)-(ad+bc). \end{array}\right\} \quad (1)$$

Wenn wir unter R die Menge aller Elemente aus C verstehen, die Differenzen aus natürlichen Zahlen sind, so gehört wegen (1) die Summe, die Differenz und das Produkt von Elementen aus R wieder zu R, es ist also R ein Unterring von C. Ferner läßt sich jede natürliche Zahl a als Differenz aus zwei natürlichen Zahlen, etwa in der Form $a=(a+b)-b$ darstellen, wobei b eine beliebige natürliche Zahl ist. Da voraussetzungsgemäß die Operationen in N und C übereinstimmen, enthält der Ring R die Menge N, und folglich ist $R=C$ auf Grund der Minimalität von C. Das besagt aber gerade, daß jede ganze Zahl Differenz aus natürlichen Zahlen ist.

Satz 2. *Alle minimalen Ringe, die die Menge der natürlichen Zahlen enthalten, sind untereinander isomorph, d. h., es gibt bis auf Isomorphie nur einen einzigen Ring der ganzen Zahlen.*

Beweis. Dazu seien C_1 und C_2 beliebige derartige Ringe. Auf Grund des vorangehenden Satzes ist jedes Element von C_1 wie auch jedes Element von C_2 gleich einer Differenz aus natürlichen Zahlen. Wir erklären nun eine Abbildung f zwischen den Ringen C_1 und C_2: Wenn $c_1 \in C_1$ und $c_1=a-b$ in C_1 ist, wobei a und b natürliche Zahlen sind, und wenn $a-b=c_2$ in C_2[1]) ist, so setzen wir $f(c_1)=c_2$. Dabei ist c_2 unabhängig von der speziellen Wahl der Zahlen a und b. Wenn nämlich außerdem $c_1=c-d$ ist, so ist $a-b=c-d$ und mithin wegen (1) $a+d=b+c$ in C_1 und damit auch in C_2 (da $a+d$ und $b+c$ natürliche Zahlen sind — *Anm. d. wissenschaftl. Red.*); folglich ist auch $a-b=c-d$ in C_2. Entsprechend ist, wenn $c_1 \neq d_1$ ist, wegen (1) auch $f(c_1) \neq f(d_1)$. Da jedes $c_1 \in C_1$ eine Differenz aus natürlichen Zahlen ist und da dasselbe auch in C_2 der Fall ist, ist f eine eineindeutige Abbildung von C_1 auf C_2. Auf Grund von b) und d) ist

$$f(c_1+d_1)=f(c_1)+f(d_1) \text{ und } f(c_1 d_1)=f(c_1)f(d_1)$$

für beliebige c_1 und d_1 aus C_1, d. h., die Abbildung f ist sogar eine isomorphe Abbildung zwischen den Ringen C_1 und C_2 (§ 9, Definition 2). Wir beweisen etwa die erste dieser beiden Gleichungen. Dazu sei $c_1=a-b$, $d_1=c-d$ in C_1. Dann ist in C_2

$$f(c_1)=a-b,\ f(d_1)=c-d$$

und mithin

$$f(c_1)+f(d_1)=(a-b)+(c-d)=(a+c)-(b+d).$$

[1]) Aus $c_1=a-b$ und $c_2=a-b$ kann man keineswegs auf $c_1=c_2$ schließen, da die Differenz in C_1 und C_2 auf ganz verschiedene Art erklärt sein kann.

Andererseits ist in C_1

$$c_1 + d_1 = (a + c) - (b + d),$$

d. h., die Elemente $c_1 + d_1 \in C_1$ und $f(c_1) + f(d_1) \in C_2$ sind beide gleich der Differenz der natürlichen Zahlen $a + c$ und $b + d$. Dann ist aber auf Grund der Definition der Abbildung f in der Tat

$$f(c_1 + d_1) = f(c_1) + f(d_1).$$

Entsprechend beweist man die zweite Gleichung. Damit ist dann der behauptete Satz in allen Teilen bewiesen.

Bemerkung. Die isomorphe Abbildung f besitzt zusätzlich noch die Eigenschaft, daß sie auf der Menge N gleich der identischen Abbildung ist, d. h., daß bei der Abbildung von C_1 auf C_2 jede natürliche Zahl in sich übergeführt wird. In der Tat: Wenn $c_1 = a - b$ in C_1 und $c_2 = a - b$ in C_2 ist, so sind, falls $a > b$ ist, die Elemente c_1 und c_2 gleich ein und derselben natürlichen Zahl, es ist also $c_2 = f(c_1) = a - b = c_1$.

Satz 3. *Jeder Ring R, der die Menge N der natürlichen Zahlen enthält, enthält auch einen Ring der ganzen Zahlen.*

Beweis. Der Durchschnitt aller die Menge N umfassenden Unterringe des Ringes R ist ein Unterring von R (§ 8, Satz 6), der gleichfalls die Menge N enthält. Dieser Ring ist minimal, da er (nach Definition des Durchschnitts) in jedem Unterring von R enthalten ist, der die Menge N enthält. Er ist also auf Grund von Definition 1 ein Ring der ganzen Zahlen.

Wir haben bisher noch nicht gezeigt, daß es überhaupt einen Ring der ganzen Zahlen gibt, d. h., wir haben noch keinen Ring (noch kein Modell) konstruiert, der die verlangten Bedingungen erfüllt. Diesem Ziel wollen wir uns jetzt zuwenden. Satz 1 legt uns bereits die Konstruktion eines der untereinander isomorphen Ringe der ganzen Zahlen nahe. Wenn nämlich C ein Ring der ganzen Zahlen ist, so sind alle Elemente von C Differenzen aus natürlichen Zahlen. Man könnte nun als Elemente des gesuchten Ringes die Symbole „$a - b$" für die Differenzen selbst nehmen. Das ist aber unzweckmäßig, weil auf Grund von (1) unter gewissen Bedingungen zwei verschiedene Symbole dieselbe ganze Zahl bezeichnen (es ist $a - b = c - d$ dann und nur dann, wenn $a + d = b + c$ ist), was unseren Vereinbarungen über die Gleichheit widersprechen würde (wir wollten unter Gleichheit stets die Identität, das genaue Übereinstimmen verstehen), und weil ferner die Bezeichnung „$a - b$" für die Subtraktion in dem gesuchten Ring vorbehalten bleiben soll.

Statt dessen nehmen wir als Ausgangselemente für die Konstruktion die geordneten Paare (a, b) aus natürlichen Zahlen. Es sei also M die Menge aller derartigen Paare. In der Menge dieser Paare erklären wir eine Äquivalenzrelation derart, daß äquivalente Paare und nur solche eine und dieselbe Differenz aus natürlichen Zahlen des gesuchten Ringes festlegen, indem wir gemäß (1) verlangen, daß

$$(a, b) \sim (c, d) \tag{2}$$

dann und nur dann gelten soll, wenn $a + d = b + c$ ist.

Ferner erklären wir in der Menge der Zahlenpaare eine Addition und eine Multiplikation entsprechend der Addition und Multiplikation von Differenzen aus natür-

lichen Zahlen in dem gesuchten Ringe. Wir setzen also gemäß b) und d) fest:

$$(a, b) + (c, d) = (a + c, b + d), \tag{3}$$

$$(a, b)\,(c, d) = (ac + bd, ad + bc). \tag{4}$$

Satz 4. *Die Addition und die Multiplikation im Bereich der Paare sind kommutativ, assoziativ und durch das distributive Gesetz miteinander verknüpft.*

Beweis. Die behaupteten Eigenschaften folgen unmittelbar aus den entsprechenden Eigenschaften für die natürlichen Zahlen und können unschwer nachgeprüft werden. Als Beispiel beweisen wir hier das assoziative Gesetz für die Multiplikation:

$$\begin{aligned}[(a, b)\,(c, d)]\,(e, f) &= (ac + bd, ad + bc)\,(e, f)\\ &= (ace + bde + adf + bcf, acf + bdf + ade + bce);\\ (a, b)\,[(c, d)\,(e, f)] &= (a, b)\,(ce + df, cf + de)\\ &= (ace + adf + bcf + bde, acf + ade + bce + bdf).\end{aligned}$$

Da die nach den Umformungen erhaltenen Paare übereinstimmen, ist in der Tat

$$[(a, b)\,(c, d)]\,(e, f) = (a, b)\,[(c, d)\,(e, f)].$$

Die oben unter (2) erklärte Äquivalenzrelation erfüllt die Bedingungen 1. bis 3. des Satzes aus § 19. In der Tat ist

1.) $(a, b) \sim (a, b)$, da $a + b = b + a$ ist.

2.) Wenn $(a, b) \sim (c, d)$ ist, so ist auch $(c, d) \sim (a, b)$, da mit $a + d = b + c$ auch $c + b = d + a$ ist.

3.) Wenn $(a, b) \sim (c, d)$ und $(c, d) \sim (e, f)$ ist, so ist auch $(a, b) \sim (e, f)$, da aus $a + d = b + c$ und $c + f = d + e$ sofort $a + d + c + f = b + c + d + e$ und hieraus $a + f = b + e$ (§ 14, Satz 3) folgt.

Die betrachtete Äquivalenzrelation liefert also eine Einteilung der Menge M aller geordneten Paare von natürlichen Zahlen in Klassen äquivalenter Paare. Diese Klassen bezeichnen wir im folgenden durch kleine griechische Buchstaben α, β, γ, δ, ...

Definition 2. *Es sei C_0 die Menge aller Klassen äquivalenter Paare der Menge M. Unter der Summe (bzw. dem Produkt) der Klassen α und β aus C_0 werde die Klasse $\alpha + \beta$ (bzw. $\alpha\beta$) verstanden, die die Summen (bzw. die Produkte) aus je einem Paar der Klasse α und einem Paar der Klasse β enthält.*

Hierbei müssen wir uns, wie in jedem Fall, in dem man Operationen zwischen Klassen mittels Operationen zwischen ihren Repräsentanten erklärt, davon überzeugen, daß das Resultat der Operationen unabhängig ist von der speziellen Wahl der Repräsentanten. Offenbar haben wir hierzu den folgenden Satz zu beweisen:

Satz 5. *Wenn $(a_1, b_1) \sim (a_2, b_2)$ und $(c_1, d_1) \sim (c_2, d_2)$ ist, so ist*

$$(a_1, b_1) + (c_1, d_1) \sim (a_2, b_2) + (c_2, d_2),$$
$$(a_1, b_1)\,(c_1, d_1) \sim (a_2, b_2)\,(c_2, d_2).$$

Beweis. Zunächst zeigen wir, daß aus $(a_1, b_1) \sim (a_2, b_2)$ folgt, daß für jedes Paar (c, d) die Beziehungen $(a_1, b_1) + (c, d) \sim (a_2, b_2) + (c, d)$ und $(a_1, b_1)\,(c, d) \sim (a_2, b_2)\,(c, d)$ gelten.

Wenn nämlich $a_1 + b_2 = a_2 + b_1$ ist, so ist

$$(a_1 + c) + (b_2 + d) = (a_2 + c) + (b_1 + d),$$

d. h.

$$(a_1, b_1) + (c, d) \sim (a_2, b_2) + (c, d).$$

Multipliziert man die Gleichung $a_1 + b_2 = a_2 + b_1$ mit c bzw. nach Vertauschung der Seiten mit d, so erhält man

$$a_1 c + b_2 c = b_1 c + a_2 c, \quad b_1 d + a_2 d = a_1 d + b_2 d.$$

Addiert man diese beiden Gleichungen, so erhält man

$$(a_1 c + b_1 d) + (a_2 d + b_2 c) = (b_1 c + a_1 d) + (a_2 c + b_2 d)$$

und damit

$$(a_1, b_1)(c, d) \sim (a_2, b_2)(c, d).$$

Zweimalige Anwendung der eben bewiesenen Beziehungen ergibt unter Benutzung des bereits bewiesenen kommutativen Gesetzes der Addition bzw. Multiplikation

$$(a_1, b_1) + (c_1, d_1) \sim (a_2, b_2) + (c_1, d_1) \sim (a_2, b_2) + (c_2, d_2),$$
$$(a_1, b_1)(c_1, d_1) \sim (a_2, b_2)(c_1, d_1) \sim (a_2, b_2)(c_2, d_2).$$

Damit ist gezeigt, daß durch Definition 2 in der Tat im Bereich C_0 der Klassen äquivalenter Paare von natürlichen Zahlen in eindeutiger Weise eine Addition und eine Multiplikation erklärt sind.

Satz 6. *Die Menge C_0 bildet bezüglich der angegebenen Operationen einen Ring.*

Beweis. Wir haben zu zeigen, daß in C_0 die Axiome I. bis VI. (§ 7, Definition 1) erfüllt sind. Da die Operationen zwischen den Klassen aus C_0 mit Hilfe der entsprechenden Operationen zwischen Repräsentanten dieser Klassen erklärt sind, ist die Gültigkeit der Axiome I., II., IV., V. und VI. eine unmittelbare Folge aus Satz 4. Es bleibt also zu zeigen, daß in C_0 auch das Axiom III. erfüllt ist. Dazu betrachten wir zunächst einmal zwei beliebige Paare (a, b) und (c, d). Damit für ein gewisses Paar (x, y) die Gleichung $(a, b) + (x, y) = (c, d)$ erfüllt ist, muß $a + x = c$ und $b + y = d$ sein, d. h., es muß $a < c$ und $b < d$ sein. Wenn also auch nur eine der Ungleichungen $a \geqq c$ oder $b \geqq d$ erfüllt ist, so kann es kein derartiges Paar (x, y) geben. Somit ist die Subtraktion im Bereich der Paare nicht unbeschränkt ausführbar, die Paare selbst bilden also keinen Ring. Trotzdem bildet die Menge C_0 einen Ring. Es seien uns also Klassen α und β aus C_0 vorgegeben, wobei die Klasse α etwa das Paar (a, b) und die Klasse β das Paar (c, d) enthalten möge. Gesucht ist eine Klasse γ derart, daß $\alpha + \gamma = \beta$ ist. Wenn (x, y) ein Paar aus der gesuchten Klasse γ ist, so ist es durchaus nicht nötig, daß die Gleichung $(a, b) + (x, y) = (c, d)$ erfüllt ist; vielmehr genügt es, daß die Äquivalenz $(a, b) + (x, y) \sim (c, d)$ besteht. Setzen wir zunächst einmal die Existenz eines Paares (x, y) voraus, das die angegebene Bedingung erfüllt. Wegen $(a + x, b + y) \sim (c, d)$ muß dann $a + x + d = b + y + c$ oder $(a + d) + x = (b + c) + y$ sein. Wegen (2) ist das aber gleichbedeutend mit $(x, y) \sim (b + c, a + d)$.

Auf Grund von Satz 5 genügt es, die Existenz auch nur eines Paares (x, y) nachzuweisen, das der Bedingung $(a, b) + (x, y) \sim (c, d)$ genügt. Das Paar $(b + c, a + d)$ leistet nun aber in der Tat bereits das Verlangte, da

$$(a, b) + (b + c, a + d) = (a + b + c, b + a + d) \sim (c, d)$$

ist.

Damit ist gezeigt, daß es wirklich eine Klasse γ gibt, für die $\alpha + \gamma = \beta$ ist.

Aus der Existenz einer Klasse γ mit $\alpha + \gamma = \beta$ folgt bereits auf Grund eines ganz allgemeinen Satzes (§ 6, Satz 1), daß sie auch eindeutig bestimmt ist.

Wir wollen uns jetzt der Frage nach der Null und den entgegengesetzten Elementen im Ring C_0 zuwenden.

Seiner Definition gemäß ist das Nullelement von C_0 diejenige Klasse 0, für die $\alpha + 0 = \alpha$ ist für jede beliebige Klasse α. Wenn die Klasse α das Paar (a, b) und die Nullklasse 0 das Paar (x, y) enthält, so muß also $(a, b) + (x, y) \sim (a, b)$ sein. Entsprechend dem Beweis des vorangehenden Satzes erhält man, wenn man dort (c, d) durch (a, b) ersetzt:

$$(x, y) \sim (b + a, a + b) = (a + b, a + b) = (k, k).$$

Andererseits gilt aber auch wegen (2) für jede natürliche Zahl k:

$$(a, b) + (k, k) \sim (a, b).$$

Die Nullklasse von C_0 besteht also aus den und nur den Paaren aus gleichen natürlichen Zahlen.

Die zu einer Klasse α entgegengesetzte Klasse $-\alpha$ ist diejenige Klasse, für die $\alpha + (-\alpha) = 0$ ist. Enthält die Klasse α das Paar (a, b) und die Klasse $-\alpha$ das Paar (x, y), so ist also $(a, b) + (x, y) = (k, k)$ für eine gewisse natürliche Zahl k. Hier kann man an Stelle des Zeichens „$\sim$“ für die Äquivalenz tatsächlich das Gleichheitszeichen „$=$“ schreiben, da wegen (2) alle zu einem Paar (k, k) äquivalenten Paare gleichfalls aus gleichen Zahlen bestehen. Es muß also $a + x = b + y$ sein, d. h., es muß $(x, y) \sim (b, a)$ sein. Das Paar (b, a) leistet andererseits auch das Verlangte, weil

$$(a, b) + (b, a) = (a + b, b + a)$$

tatsächlich der Klasse 0 angehört.

Wir wollen das Paar (b, a) *entgegengesetzt* dem Paar (a, b) nennen. Dem Übergang von einem Paar (a, b) zu einem zu diesem äquivalenten Paar entspricht bei dem zu (a, b) entgegengesetzten Paar der Übergang zu einem dem entgegengesetzten Paar äquivalenten Paar. Ferner ist jedes Paar aus der Klasse $-\alpha$ einem eindeutig bestimmten Paar aus der Klasse α entgegengesetzt. Die zu einer Klasse α entgegengesetzte Klasse $-\alpha$ besteht also aus genau den Paaren, die zu einem Paar aus der Klasse α entgegengesetzt sind.

Der von uns konstruierte Ring C_0 ist isomorph jedem Ring der ganzen Zahlen. Wenn es nur darauf ankommt, einen Ring der ganzen Zahlen bis auf Isomorphie zu kennen, so kann man C_0 selbst als Ring der ganzen Zahlen ansehen. Wir wollen hier jedoch erreichen, daß bei der Erweiterung eines vorgegebenen Zahlenbereiches zu einem neuen der neue Bereich stets eindeutig durch den gegebenen Bereich bestimmt ist, d. h., wir wollen aus allen Modellen für die Erweiterung ein ganz bestimmtes auswählen. Unter diesen verschärften Bedingungen genügt der Ring C_0 noch nicht der Definition 1, da in C_0 die natürlichen Zahlen nicht enthalten sind. Die Elemente von C_0 sind vielmehr Klassen von äquivalenten Paaren aus natürlichen Zahlen.

Da die natürlichen Zahlen selbst noch nicht als Elemente in C_0 enthalten sind, hat man, um aus C_0 einen Ring der ganzen Zahlen (Definition 1) zu erhalten, in C_0 die Menge N der natürlichen Zahlen „einzubetten“.

Dazu legen wir zunächst in C_0 eine der Menge der natürlichen Zahlen isomorphe Teilmenge fest. Eine von der Nullklasse des Ringes C_0 verschiedene Klasse α besteht nach dem Gesagten aus Paaren (a, b) von natürlichen Zahlen, für die $a \neq b$ ist. Wir nennen α eine *Klasse erster Art* (oder positive Klasse), wenn $a > b$ ist, und eine *Klasse zweiter Art* (oder negative Klasse), wenn $a < b$ ist. Diese Definition ist unabhängig von der speziellen Wahl des Repräsentanten (a, b) aus der Klasse α. Wenn nämlich $(a, b) \sim (c, d)$ ist, so ist $a + d = b + c$; wenn also $a > b$ ist, so ist auch $c > d$ (§ 16, Satz 2), während aus $a < b$ folgt, daß auch $c < d$ ist. Seien nun N_1 bzw. N_2 die Mengen der Klassen erster bzw. zweiter Art. Wir zeigen, daß die Menge N_1, also die Menge der Klassen erster Art bezüglich der Addition und der Multiplikation der Menge N der natürlichen Zahlen isomorph ist. Dazu betrachten wir die folgende eineindeutige Abbildung zwischen den Mengen N_1 und N: Wenn die Klasse α aus N_1 das Paar (a, b) enthält, so ist nach Definition der Klassen erster Art $a > b$. Es gibt also eine natürliche Zahl k derart, daß $a = b + k$ ist (§ 14). Wir setzen dann fest, daß $f(\alpha) = k$ ist. Diese Festsetzung ist unabhängig von der speziellen Wahl des Repräsentanten (a, b) aus der Klasse α; ist nämlich $(a, b) \sim (c, d)$, also $a + d = b + c$, so ist wegen $a = b + k$ dann $b + k + d = b + c$, also in der Tat $c = d + k$. Bei dieser Abbildung entsprechen weiter verschiedenen Klassen auch verschiedene natürliche Zahlen; wenn nämlich die Klasse α das Paar (a, b) und die Klasse β das Paar (c, d) enthält und wenn $f(\alpha) = f(\beta) = k$ ist, so ist $a = b + k$, $c = d + k$, also

$$a + d + k = b + k + c,\ a + d = b + c,\ (a, b) \sim (c, d),\ \alpha = \beta.$$

Jede natürliche Zahl k ist überdies Bild einer eindeutig bestimmten Klasse α aus N_1, nämlich der Klasse aller Paare der Form $(a + k, a)$. Die Abbildung f ist also eine umkehrbar eindeutige Abbildung von N_1 auf N (§ 3, Definition 3).

Wir zeigen jetzt, daß die so erklärte Abbildung f ein Isomorphismus bezüglich der Addition und Multiplikation von N_1 auf N ist, d. h., daß die Gleichungen

$$f(\alpha) + f(\beta) = f(\alpha + \beta),\ f(\alpha) f(\beta) = f(\alpha\beta) \tag{5}$$

gelten.

Wenn die Klasse α das Paar $(a + k, a)$ und die Klasse β das Paar $(b + l, b)$ enthält, so enthält die Klasse $\alpha + \beta$ das Paar $(a + b + k + l, a + b)$, und es ist

$$f(\alpha + \beta) = k + l = f(\alpha) + f(\beta).$$

Entsprechend enthält die Klasse $\alpha\beta$ das Paar

$$(a + k, a) \cdot (b + l, b) = (ab + kb + al + kl + ab, ab + kb + ab + al) = (c + kl, c),$$

wobei $c = 2ab + al + bk$ ist. Mithin ist auch

$$f(\alpha\beta) = kl = f(\alpha) f(\beta).$$

Nach diesen Vorbereitungen sind wir in der Lage, endgültig einen Ring C der ganzen Zahlen zu konstruieren. Die hierfür notwendigen Überlegungen verlaufen analog den entsprechenden Überlegungen für beliebige Ringe (§ 9, Satz 2). Es sei also C die Menge, die man aus dem Ring C_0 erhält, indem man die Klassen erster Art durch die ihnen vermöge der Abbildung f entsprechenden natürlichen Zahlen ersetzt. Wenn wir darüber hinaus $f(\alpha) = \alpha$ für die Klassen zweiter Art und die Nullklasse setzen, so erhalten wir eine umkehrbar eindeutige Abbildung von C_0 auf C.

Die Addition und die Multiplikation in C erklären wir durch die Gleichungen

$$f(\alpha)+f(\beta)=f(\alpha+\beta),\ f(\alpha)f(\beta)=f(\alpha\beta). \tag{5'}$$

Hierbei sollen nach wie vor α und β beliebige Elemente des Ringes C_0 sein. Dann sind $f(\alpha)$ und $f(\beta)$ gewisse Elemente der Menge C. Da die Summe $\alpha+\beta$ und das Produkt $\alpha\beta$ für die Elemente von C_0 eindeutig festgelegt sind, wird durch die Gleichungen (5') tatsächlich für die Elemente von C eine Addition und eine Multiplikation definiert.

Durch die angegebenen Festsetzungen wird also die Menge C zu einem Bereich mit zwei Verknüpfungen. Die Gleichungen (5') zeigen darüber hinaus, daß die Menge C bezüglich der so erklärten Addition und Multiplikation isomorph dem Ringe C_0, also selbst ein Ring (§ 9, Satz 1) ist.

Satz 7. *Der eben konstruierte Ring C ist ein Ring der ganzen Zahlen.*

Beweis. Es ist zu zeigen, daß C die in Definition 1 am Anfang dieses Paragraphen festgelegten Bedingungen erfüllt. Es wurde bereits gezeigt, daß 1.) C die Menge N der natürlichen Zahlen enthält und daß 2.) C ein Ring ist.

Sind $k=f(\alpha)$ und $l=f(\beta)$ natürliche Zahlen, so sind α und β Klassen erster Art. Dann stimmen die Gleichungen (5'), die im Ring C die Summe $k+l$ und das Produkt $k\cdot l$ festlegen, mit den entsprechenden Gleichungen (5) überein, in denen die Summe bzw. das Produkt auf der linken Seite die Summe und das Produkt von natürlichen Zahlen im Sinne der §§ 12, 13 sind. Damit ist gezeigt, daß 3.) die Addition und die Multiplikation für natürliche Zahlen im Ringe C mit den üblichen Operationen für natürliche Zahlen zusammenfallen.

Wir wollen schließlich zeigen, daß jedes Element aus C gleich einer Differenz aus natürlichen Zahlen ist. Dazu beachten wir, daß sich jedes Element aus C in der Form $f(\alpha)$ darstellen läßt, wobei α eine Klasse aus C_0 und f die oben definierte Abbildung ist. Die Klasse α möge dann das Paar (k, l) von natürlichen Zahlen enthalten und es möge $k=f(\beta)$ und $l=f(\gamma)$ sein. Nach Definition von f besteht dann β aus allen Paaren der Form $(b+k, b)$ und γ aus allen Paaren der Form $(c+l, c)$; folglich enthält die Klasse $\alpha+\gamma$ alle Paare $(k, l)+(c+l, c)=(k+c+l, l+c)$, die andererseits gerade in β liegen. Es ist also $\alpha+\gamma=\beta$. Gemäß der Definition der Addition in C, d. h. gemäß (5')[1] ist somit

$$f(\alpha)+f(\gamma)=f(\beta),\ \text{d. h. } f(\alpha)=f(\beta)-f(\gamma)=k-l.^{2})$$

Da nun aber jeder Unterring von C, der die Menge N der natürlichen Zahlen enthält, auch alle Differenzen aus natürlichen Zahlen enthält, also mit C zusammenfällt, gilt:

4.) Der Ring C enthält keinen Unterring, der N enthält und der von C verschieden ist.

Wir haben also mit dem Ring C tatsächlich einen der untereinander isomorphen Ringe der ganzen Zahlen konstruiert. Seine Elemente (d. h. ganze Zahlen) sind 1.) die natürlichen Zahlen, 2.) die Zahl 0 in Form der Klasse aller Paare aus gleichen natürlichen Zahlen und 3.) alle Klassen zweiter Art, also alle Klassen äquivalenter Paare (a, b) von natürlichen Zahlen mit $a<b$. Damit ist die Frage nach der Existenz eines Ringes der ganzen Zahlen beantwortet.

[1]) Wir weisen darauf hin, daß man hier nicht etwa die Gleichungen (5) verwenden darf, weil α nicht notwendig von erster Art zu sein braucht.

[2]) Für die in C enthaltenen Klassen zweiter Art und die Nullklasse 0 bedeutet der bewiesene Sachverhalt, daß die Klasse, die das Paar (k, l) enthält, gleich der Differenz $k-l$ ist.

So gut dem Leser auch immer der Ring der ganzen Zahlen bekannt sein mag, er wird sich in dem eben konstruierten Ring C doch nur schwer zurechtfinden. Wir wollen daher im nachfolgenden Paragraphen eine Reihe von einfachen Eigenschaften dieses Ringes studieren und dabei sehen, daß er sich in keiner Weise von der bekannten Gesamtheit der ganzen Zahlen unterscheidet.

§ 21. Eigenschaften der ganzen Zahlen

Bemerkung 1. Für die ganzen Zahlen als Elemente eines Ringes gelten zunächst alle Verknüpfungsregeln, die wir in § 7 für beliebige Ringe bewiesen haben. So ist das Produkt der Null mit einer beliebigen Zahl stets Null [§ 7, (2)], es gelten die üblichen Gesetze für die Multiplikation [§ 7, (3)] usw.

Satz 1. *Die natürlichen Zahlen* 1, 2, 3, ..., *die Zahl* 0 *und die Zahlen* —1, —2, —3, ..., *also die den natürlichen Zahlen entgegengesetzten Zahlen, erschöpfen bereits alle ganzen Zahlen, d. h., für jedes Element* $a \in C$ *gilt einer und nur einer der drei folgenden Fälle:* a *ist eine natürliche Zahl,* $a = 0$ *oder* $-a$ *ist eine natürliche Zahl.*

Beweis. Es sei $a = f(\alpha)$, wobei α eine Klasse des Ringes C_0 ist.[1]) Wir haben oben gezeigt, daß α von erster Art, gleich der Nullklasse oder von zweiter Art ist. Diese Fälle schließen sich auch gegenseitig aus; ist nämlich (k, l) ein Repräsentant der Klasse α, so schließen sich (§ 14, Satz 1) die Beziehungen $k > l$, $k = l$ und $k < l$ gegenseitig aus. Ist α speziell von zweiter Art, so ist $k < l$. Dann enthält die zu α entgegengesetzte Klasse $-\alpha$ das Paar (l, k), wobei dann $l > k$ ist, d. h., dann ist $-\alpha$ von erster Art. Ferner ist bei dem Isomorphismus f das Bild des entgegengesetzten Elementes gleich dem entgegengesetzten Element des Bildes, d. h.,

$$f(-\alpha) = -f(\alpha) = -a.$$

Wenn α von erster Art ist, so ist $a = f(\alpha)$ nach Definition von f eine natürliche Zahl; wenn $\alpha = 0$ ist, so ist $a = \alpha = 0$; wenn α von zweiter Art ist, so ist $-\alpha$ von erster Art und $-a = -f(\alpha) = f(-\alpha)$ eine natürliche Zahl.

Satz 2. *Der Ring der ganzen Zahlen ist ein Integritätsbereich mit Einselement, und zwar spielt die natürliche Zahl* 1 *die Rolle des Einselementes.*

Beweis. Wir wollen im folgenden, soweit es notwendig erscheint, statt a auch $+a$ schreiben. Es soll jetzt gezeigt werden, daß das Produkt ab zweier ganzer Zahlen nur dann Null ist, wenn mindestens einer der Faktoren Null ist. Sei also $a \neq 0$ und $b \neq 0$. Auf Grund des vorangehenden Satzes ist dann $a = \pm c$ und $b = \pm d$, wobei c und d natürliche Zahlen sind. Also ist $ab = \pm cd$, wobei das Zeichen „+" bei gleichen Vorzeichen von a und b und das Zeichen „—" bei ungleichem Vorzeichen von a und b steht. Es ist $cd \neq 0$, weil ein Produkt aus natürlichen Zahlen wieder eine natürliche Zahl ist. Dann ist aber auch $ab \neq 0$.

[1]) Wir verwenden im folgenden zur Bezeichnung auch der von natürlichen Zahlen verschiedenen ganzen Zahlen sowohl griechische als auch lateinische Buchstaben.

Ferner ist $a \cdot 1 = a$ für jede ganze Zahl a. Falls a eine natürliche Zahl ist, gilt dies auf Grund der Definition der Multiplikation für natürliche Zahlen (§ 13). Falls $a = 0$ ist, gilt dies auf Grund von $a \cdot 1 = 0 \cdot 1 = 0 = a$. Falls $a = -b$ ist, wobei b eine natürliche Zahl ist, gilt dies auf Grund von $a \cdot 1 = (-b) \cdot 1 = -(b \cdot 1) = -b = a$. Damit ist der behauptete Satz bewiesen.

Wir behandeln als nächstes den Begriff der positiven und der negativen Zahl sowie den Vergleich ganzer Zahlen ihrer Größe nach.

Satz 3. *Der Ring C der ganzen Zahlen kann auf eine und nur eine Weise angeordnet werden* (§ 10, Definition 1). *Dabei sind die natürlichen Zahlen sämtlich positiv, während die ihnen entgegengesetzten Zahlen* $-1, -2, -3, \ldots$ *negativ sind.*

Beweis. Wir erhalten eine Anordnung des Ringes C, wenn wir alle natürlichen Zahlen und nur diese als positiv ansehen. Nach Satz 1 ist dann nämlich entweder a positiv, $a = 0$ oder $-a$ positiv, d. h., dann ist das Axiom IX. (§ 10) in der Tat erfüllt. Da die Summe und das Produkt natürlicher Zahlen wieder natürliche Zahlen sind, so gilt außerdem auch das Axiom X. Da die natürlichen Zahlen sämtlich positiv sind, sind die ihnen entgegengesetzten Zahlen sämtlich negativ.

Die angegebene Anordnung ist aber auch die einzig mögliche: Dazu sei eine beliebige Anordnung des Ringes C vorgegeben. Auf Grund von Axiom IX. ist dann eine der Zahlen $+1$ oder -1 positiv. Dann ist nach Axiom X. die Zahl $1 = 1 \cdot 1 = (-1) \cdot (-1)$ als Produkt positiver Zahlen selbst positiv. Damit ist dann aber auch jede natürliche Zahl n als Summe aus n Einsen (§ 15, Satz 2) positiv, so daß die entgegengesetzten Zahlen $-n$ wegen Axiom IX. nicht positiv sein können. Nun bilden aber auf Grund von Satz 1 bereits die Zahlen 0 und $\pm n$, wo n alle natürlichen Zahlen durchläuft, den ganzen Ring C. Es sind also bei der angenommenen Anordnung von C die natürlichen Zahlen und nur diese positiv, d. h., die betrachtete Anordnung von C stimmt mit der am Anfang dieses Beweises eingeführten Anordnung von C überein.

Bemerkung 2. Als Elemente eines angeordneten Ringes besitzen die ganzen Zahlen alle Eigenschaften, die wir in § 10 für die Elemente eines beliebigen angeordneten Ringes abgeleitet haben. So erhält man eine Ordnung in der Menge der ganzen Zahlen, bezüglich der die Zahl 0 kleiner als alle positiven Zahlen und größer als alle negativen Zahlen ist, wenn man festsetzt, daß dann und nur dann $a > b$ sein soll, wenn $a - b$ positiv ist (§ 10, Satz 1). Für diese Ordnung gelten dann die Monotoniegesetze und die üblichen Gesetze für das Rechnen mit Ungleichungen (§ 10, Satz 2 bis Satz 4). Ferner kann man den absoluten Betrag $|a|$ einer ganzen Zahl a als die positive der beiden Zahlen $\pm a$ einführen (siehe § 10, Definition 2), der dann die üblichen Eigenschaften besitzt und für den die bekannten Rechen- und Vergleichsregeln zwischen Zahlen und ihren absoluten Beträgen gelten (§ 10, Satz 8 und die nachfolgende Bemerkung).

Satz 4. *Die Ordnung der natürlichen Zahlen im Ring der ganzen Zahlen stimmt mit der üblichen Ordnung der natürlichen Zahlen* (§ 14) *überein.*

Beweis. Sind a und b ganze Zahlen mit $a > b$, so ist $a - b = k$ eine positive, also eine natürliche Zahl und $a = b + k$. Wenn also a und b speziell natürliche Zahlen sind, so ist dies gleichbedeutend damit, daß $a > b$ ist im Sinne der Definition aus § 14.

Da die Menge aller ganzen Zahlen keine kleinste Zahl enthält, gilt für sie der Satz 8 aus § 14 nicht. Um für die ganzen Zahlen eine Behauptung ähnlicher Art aufstellen zu können, hat man weitergehende Voraussetzungen zu machen.

Definition. *Eine Menge A von ganzen Zahlen heißt nach oben beschränkt (bzw. nach unten beschränkt, bzw. beschränkt), wenn es eine ganze Zahl k gibt derart, daß $k > x$ (bzw. $k < x$, bzw. wenn es zwei ganze Zahlen k und l gibt derart, daß $k < x < l$) für alle Zahlen x aus A gilt. Auch die leere Menge soll als beschränkt gelten.*

Satz 5. *Jede nichtleere, nach oben (bzw. nach unten, bzw. schlechthin) beschränkte Menge A von ganzen Zahlen enthält eine größte (bzw. eine kleinste, bzw. eine größte und eine kleinste) Zahl.*

Beweis. Sei A nach oben beschränkt. Wenn die Menge A eine natürliche Zahl enthält, so ist die Menge der in A enthaltenen natürlichen Zahlen nicht leer und besitzt mithin eine größte Zahl a (§ 14, Satz 2). Offenbar ist in diesem Fall a auch größte Zahl von ganz A. Enthält A keine natürliche Zahl, aber die Zahl 0, so ist 0 auch größte Zahl von A. Enthält schließlich A nur negative Zahlen, so ist die Menge B der zu den Zahlen aus A entgegengesetzten Zahlen eine Menge von natürlichen Zahlen, die dann eine kleinste Zahl b enthält (§ 14, Satz 8), also eine Zahl b, für die $b \leq y$ für alle y aus B gilt. Multiplizieren wir alle diese Ungleichungen mit -1, so erhalten wir $-b \geq -y$ oder, wenn wir $a = -b$ und $x = -y$ setzen, $a \geq x$ für alle x aus A. Ist hingegen A nach unten beschränkt, so ist gemäß der oben stehenden Definition B nach oben beschränkt und enthält auf Grund des Bewiesenen eine größte Zahl b. Dann ist aber $a = -b$ kleinste Zahl von A. Ist schließlich A beschränkt, so ist A sowohl nach oben als auch nach unten beschränkt und enthält damit eine größte und eine kleinste Zahl.

Auf diesem Satz beruhen die verschiedenen Formen der einseitigen und zweiseitigen Induktion im Bereich der ganzen Zahlen. So gilt beispielsweise:

Satz 6. *Es sei T eine Aussage über ganze Zahlen, die für eine gewisse ganze Zahl a gilt.*

a) *Wenn aus der Voraussetzung, daß die Aussage T für eine beliebige Zahl $x \geq a$ richtig ist, folgt, daß T auch für die Zahl $x + 1$ gilt, so gilt die Aussage T für alle ganzen Zahlen $b \geq a$.*

b) *Wenn aus der Voraussetzung, daß die Aussage T für eine beliebige Zahl $x \leq a$ gilt, folgt, daß T auch für die Zahl $x - 1$ gilt, so gilt die Aussage T für alle ganzen Zahlen $b \leq a$.*

c) *Wenn aus der Voraussetzung, daß für beliebiges x_1 und x_2 mit $x_1 \leq a \leq x_2$ die Aussage T für alle Zahlen x mit $x_1 < x < x_2$ richtig ist, folgt, daß die Aussage T auch für die Zahlen x_1 und x_2 gilt, so gilt die Aussage T für alle ganzen Zahlen b.*

Beweis. Alle drei Behauptungen lassen sich nach ein und derselben Methode beweisen. Als Beispiel führen wir den Beweis für den Fall c) vor. Würde unter den angegebenen Voraussetzungen die Aussage T nicht für alle ganzen Zahlen gelten, so gäbe es eine ganze Zahl b, für die die Aussage T falsch ist. Dann ist sicher $b \neq a$. Es möge etwa $b > a$ sein (im Falle $b < a$ schließt man entsprechend); ferner sei A die Menge der ganzen Zahlen $x > a$, für die die Aussage T nicht gilt. Die Menge A ist dann durch die Zahl a nach unten beschränkt und nicht leer, weil sie nach Annahme die Zahl b enthält. Auf Grund des vorangehenden Satzes besitzt dann die Menge A eine kleinste Zahl x_2. Setzt man etwa $a - 1$ gleich x_1, so gilt die Aussage T für alle ganzen Zahlen x mit $x_1 < x < x_2$, wobei außerdem $x_1 < a < x_2$ ist. Nach Voraussetzung gilt dann aber T auch für x_1 und x_2. Die Zahl x_2 sollte aber zu A gehören, d. h., für x_2 sollte die Aussage T nicht gelten. Damit haben wir einen Widerspruch zu unserer Annahme erhalten, womit c) bewiesen ist.

Satz 7. *Der Ring der ganzen Zahlen ist archimedisch angeordnet* (§ 10, Definition 3).

Beweis. Es seien a und b beliebige ganze Zahlen und $b > 0$. Ist $a \leq 0$, so ist $1 \cdot b = b > a$. Ist $a > 0$, so sind a und b natürliche Zahlen, für die wir das Archimedische Axiom bereits bewiesen haben (§ 14, Satz 6), d. h., in diesem Falle gibt es eine natürliche Zahl n derart, daß $n \cdot b > a$ ist.

Auf die Teilbarkeitseigenschaften der ganzen Zahlen wollen wir hier nicht näher eingehen. Sie finden sich in dem nachfolgenden Artikel von A. J. CHINTSCHIN.

Kapitel V

DER KÖRPER DER RATIONALEN ZAHLEN

§ 22. Definition des Körpers der rationalen Zahlen

In diesem Kapitel wollen wir den Körper der rationalen Zahlen aufbauen. Die gebrochenen Zahlen entstanden lange vor den negativen Zahlen bereits im frühen Altertum. Ihre Entstehung war bedingt durch das Messen von Größen, die sich nicht als ganzzahlige Vielfache der Maßeinheit ausdrücken ließen. Wir bemerken, daß sich die Reihenfolge unserer Untersuchungen von der Reihenfolge, in der der hier behandelte Stoff in der Schule dargeboten wird, unterscheidet, da wir zunächst die ganzen Zahlen eingeführt haben und erst jetzt die rationalen Zahlen definieren, während man in der Schule üblicherweise die negativen ganzen Zahlen erst nach den (positiven) gebrochenen Zahlen behandelt. Wir haben diesen Weg gewählt, um möglichst schnell zu einem Zahlenbereich (den ganzen Zahlen) zu gelangen, der einen Ring bildet, um im weiteren die in Kapitel II dargelegte allgemeine Theorie anwenden zu können. Wir weisen jedoch darauf hin, daß man ohne wesentliche Veränderung der Überlegungen auch die Plätze der in § 20 konstruierten „relativen Zahlen" und der rationalen Zahlen aus dem vorliegenden Paragraphen vertauschen kann, womit man zu einer Darstellung kommt, die mit der in der Schule üblichen Reihenfolge übereinstimmt.

Die Erweiterung der Menge der ganzen Zahlen zur Menge der rationalen Zahlen verläuft nach den in § 18 für beliebige Erweiterungen dargelegten allgemeinen Prinzipien, wobei die hierfür nötigen Überlegungen den Überlegungen bei der Erweiterung der Menge der natürlichen Zahlen zur Menge der ganzen Zahlen völlig analog sind. Der Unterschied besteht lediglich darin, daß dort von den Eigenschaften der Addition die Rede war, während wir uns hier mit den Eigenschaften der Multiplikation zu beschäftigen haben.

In der Menge der ganzen Zahlen ist die Umkehroperation der Multiplikation, die Division, nicht unbeschränkt ausführbar, selbst dann nicht, wenn man voraussetzt, daß der Teiler von Null verschieden ist. Wir stellen uns jetzt die Aufgabe, den Ring C der ganzen Zahlen zu einer Menge Γ zu erweitern, in der eine Addition und eine Multiplikation erklärt sind, die dieselben Eigenschaften wie die Addition und die Multiplikation im Bereich der ganzen Zahlen besitzen, in der aber außerdem die Division durch alle vom Nullelement des Ringes C verschiedenen Elemente

möglich ist. Dies besagt nichts anderes, als daß Γ ein Körper (§ 8, Definition 1) ist. Unter allen derartigen Erweiterungen wählen wir zudem die kleinste im Sinne der nachfolgenden Definition aus.

Definition 1. *Unter einem Körper der rationalen Zahlen verstehen wir einen minimalen Körper Γ, der den Ring C der ganzen Zahlen enthält, also eine Menge Γ mit folgenden Eigenschaften*: 1.) *Γ enthält die Menge C*; 2.) *Γ ist ein Körper*; 3.) *Die Addition und die Multiplikation in Γ stimmen für die in Γ enthaltenen ganzen Zahlen mit der Addition und Multiplikation für ganze Zahlen überein*; 4.) *Der Körper Γ enthält keinen von sich selbst verschiedenen Unterkörper, der ebenfalls noch die Menge C enthält. Die Elemente des Körpers Γ heißen rationale Zahlen.*

Auf Grund dieser Definition ist es keineswegs klar, ob es überhaupt einen derartigen Körper Γ gibt, und wenn es einen solchen Körper gibt, ob er eindeutig bestimmt ist. Wir zeigen zunächst wieder unter Voraussetzung der Existenz eines solchen Körpers, daß er bis auf Isomorphie eindeutig bestimmt ist.

Satz 1. (Vgl. § 20, Satz 1.) *Ein den Ring C der ganzen Zahlen enthaltender*[1]) *Körper Γ ist dann und nur dann ein Körper der rationalen Zahlen (d. h. minimal), wenn jedes seiner Elemente ein Quotient aus ganzen Zahlen ist.*

Beweis. A.) Wenn der Körper Γ den Ring C der ganzen Zahlen enthält und jedes seiner Elemente ein Quotient aus ganzen Zahlen ist, so ist Γ minimal, weil jeder in Γ enthaltene Unterkörper, der gleichfalls den Ring C enthält, auch alle Quotienten aus ganzen Zahlen enthält (§ 8, Satz 5) und mithin gleich Γ ist.

B.) Sei umgekehrt Γ ein minimaler Körper, der den Ring der ganzen Zahlen enthält. In jedem Körper besitzt der Quotient (falls die Nenner von Null verschieden sind) die folgenden Eigenschaften (§ 8, Satz 3):

$$\left.\begin{array}{ll} \text{a)} & \text{Ist } b \neq 0 \text{ und } d \neq 0, \text{ so ist } \frac{a}{b} = \frac{c}{d} \text{ genau dann, wenn } ad = bc \text{ ist;} \\ \text{b)} & \text{wenn } b \neq 0 \text{ und } d \neq 0 \text{ ist, so ist } \frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}; \\ \text{c)} & \text{wenn } b \neq 0 \text{ und } d \neq 0 \text{ ist, so ist } \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}; \\ \text{d)} & \text{wenn } b \neq 0,\ c \neq 0 \text{ und } d \neq 0 \text{ ist, so ist } \frac{a}{b} : \frac{c}{d} = \frac{ad}{bc}. \end{array}\right\} \quad (1)$$

Wenn wir unter M die Menge aller Elemente aus Γ verstehen, die gleich einem Quotienten aus ganzen Zahlen sind, so gehören wegen (1) die Summe, die Differenz, das Produkt und der Quotient (falls der Nenner von Null verschieden ist) von Elementen aus M wieder zu M, es ist also M ein Unterkörper des Körpers Γ (§ 8, Satz 5). Ferner läßt sich jede ganze Zahl a als Quotient aus zwei ganzen Zahlen, etwa in der Form $a = \frac{ab}{b}$ darstellen, wobei b eine beliebige von Null verschiedene ganze Zahl sein kann. Da aber die Operationen in C und Γ übereinstimmen, enthält folglich M den Ring der ganzen Zahlen, und es ist $M = \Gamma$ auf Grund der Minimalität von Γ. Das besagt aber gerade, daß jede rationale Zahl Quotient aus ganzen Zahlen ist.

[1]) Hier wie im folgenden soll das „Enthaltensein" stets so verstanden werden, daß die Operationen in den Teilmengen mit den entsprechenden Operationen in den Obermengen übereinstimmen.

Satz 2. (Vgl. § 20, Satz 2.) *Alle minimalen Körper, die den Ring C der ganzen Zahlen enthalten, sind untereinander isomorph, d. h., es gibt bis auf Isomorphie nur einen einzigen Körper der rationalen Zahlen.*

Beweis. Dazu seien Γ_1 und Γ_2 beliebige derartige Körper. Auf Grund des vorangehenden Satzes ist jedes Element von Γ_1 wie auch jedes Element von Γ_2 gleich einem Quotienten aus ganzen Zahlen. Wie folgt erklären wir eine Abbildung f zwischen den Körpern Γ_1 und Γ_2: Wenn $c_1 \in \Gamma_1$ und $c_1 = \frac{a}{b}$ in Γ_1 ist, wobei a und b ganze Zahlen sind, und wenn $c_2 = \frac{a}{b}$ in Γ_2 ist, so soll $f(c_1) = c_2$ sein. Wegen der völligen Analogie des weiteren Beweises zum Beweis von Satz 2 aus § 20 beschränken wir uns auf den Hinweis, daß aus der obigen Eigenschaft a) die Eineindeutigkeit der Abbildung f folgt. Aus der Eigenschaft b) ergibt sich

$$f(c_1 + d_1) = f(c_1) + f(d_1),$$

und schließlich aus der Eigenschaft c)

$$f(c_1 d_1) = f(c_1) f(d_1),$$

beides für beliebige Elemente c_1 und d_1 aus Γ_1. Damit ist dann aber bereits die Isomorphie der Körper Γ_1 und Γ_2 bewiesen.

Satz 3. (Vgl. § 20, Satz 3.) *Jeder Körper K, der den Ring C der ganzen Zahlen enthält, enthält auch einen Körper Γ der rationalen Zahlen.*

Beweis. Der Durchschnitt aller C enthaltenden Unterkörper von K ist ein gleichfalls den Ring C enthaltender Unterkörper von K, der zudem minimal ist, weil er in jedem in K enthaltenen Unterkörper enthalten ist. Er ist also auf Grund von Definition 1 ein Körper der rationalen Zahlen.

Wir wenden uns jetzt dem Beweis der Existenz eines Körpers der rationalen Zahlen zu. Wie im Falle des Ringes der ganzen Zahlen geben wir auch hier ein Beispiel (ein Modell) für einen der Definition 1 genügenden Körper an.

Die Konstruktion eines der untereinander isomorphen Körper der rationalen Zahlen wird wiederum durch Satz 1 nahegelegt, demzufolge die Elemente eines Körpers Γ der rationalen Zahlen Quotienten aus ganzen Zahlen sein müssen. Überdies werden die Verknüpfungsoperationen für diese Quotienten durch die Formeln (1) festgelegt.

Als Ausgangselemente für die Konstruktion des Körpers der rationalen Zahlen verwenden wir wie in § 20 geordnete Paare (a, b), jetzt aber aus ganzen Zahlen, wobei wir voraussetzen, daß die zweite Zahl b von Null verschieden ist. Es sei also M die Menge aller dieser Paare. Zwischen den Paaren der Menge M definieren wir eine Äquivalenzrelation, eine Addition und eine Multiplikation, entsprechend der Gleichheit, der Addition und der Multiplikation der aus den Zahlen der Paare bestehenden Quotienten in dem zu konstruierenden Körper. Wir setzen also gemäß (1) fest, daß

$$(a, b) \sim (c, d) \tag{2}$$

dann und nur dann gelten soll, wenn $ad = bc$ ist, und daß

$$(a, b) + (c, d) = (ad + bc, bd), \tag{3}$$

$$(a, b)\,(c, d) = (ac, bd) \tag{4}$$

sein soll. Die Paare auf der rechten Seite von (3) und (4) liegen tatsächlich in M, weil für beliebige ganze Zahlen b und d aus $b \neq 0$ und $d \neq 0$ folgt, daß auch $bd \neq 0$ ist (§ 21, Satz 2).

Satz 4. *Die Addition und die Multiplikation der Paare sind kommutativ und assoziativ, während das distributive Gesetz in Form der Äquivalenz*

$$[(a, b) + (c, d)]\,(e, f) \sim (a, b)\,(e, f) + (c, d)\,(e, f) \tag{5}$$

gilt.

Beweis. Alle diese Behauptungen beweist man unmittelbar unter Benutzung der Eigenschaften der ganzen Zahlen als Elemente eines Ringes (§ 20, Definition 1) durch Rückgang auf die Definitionen. Als Beispiel führen wir den Beweis für die Äquivalenz (5) vor: Durch Umformung der rechten bzw. der linken Seite von (5) erhält man

$$[(a, b) + (c, d)]\,(e, f) = (ad + bc, bd)\,(e, f) = (ade + bce, bdf),$$

$$(a, b)\,(e, f) + (c, d)\,(e, f) = (ae, bf) + (ce, df) = (aedf + bfce, bfdf),$$

woraus sich auf Grund von Definition (2) unmittelbar die behauptete Äquivalenz ergibt.

Die Äquivalenzrelation (2) für die Paare besitzt die grundlegenden Eigenschaften einer solchen, nämlich:

1.) $(a, b) \sim (a, b)$, weil $ab = ba$ ist.

2.) Wenn $(a, b) \sim (c, d)$ ist, so ist $(c, d) \sim (a, b)$, weil aus $ad = bc$ folgt, daß $cb = da$ ist.

3.) Wenn $(a, b) \sim (c, d)$ und $(c, d) \sim (e, f)$ ist, so ist $(a, b) \sim (e, f)$; multipliziert man nämlich die Gleichung $ad = bc$ mit f und die Gleichung $cf = de$ mit b, so erhält man $adf = bcf = bde$, d. h. $adf = bde$, woraus sich (wegen $d \neq 0$) $af = be$ ergibt.

Die Äquivalenzrelation (2) definiert in M eine Einteilung in Klassen äquivalenter Elemente, die wir im folgenden durch kleine griechische Buchstaben $\alpha, \beta, \gamma, \delta, \ldots$ bezeichnen wollen.

Definition 2. *Es sei Γ_0 die Menge aller Klassen äquivalenter Paare der Menge M. Unter der Summe (bzw. dem Produkt) der Klassen α und β aus Γ_0 verstehen wir die Klasse $\alpha + \beta$ (bzw. $\alpha\beta$), die die Summen (bzw. die Produkte) aus je einem Paar der Klasse α und einem Paar der Klasse β enthält.*

Wie im vorangehenden Kapitel, so folgt auch hier die Unabhängigkeit dieser Definition von der speziellen Wahl der Repräsentanten aus den Klassen α und β aus dem folgenden

Satz 5. *Wenn $(a_1, b_1) \sim (a_2, b_2)$ und $(c_1, d_1) \sim (c_2, d_2)$ ist, so ist*

$$(a_1, b_1) + (c_1, d_1) \sim (a_2, b_2) + (c_2, d_2)$$

$$(a_1, b_1)\,(c_1, d_1) \sim (a_2, b_2)\,(c_2, d_2).$$

Beweis. Wie früher (§ 20, Satz 5) hat man nur zu zeigen, daß für ein beliebiges Paar (c, d) die Äquivalenzen

$$(a_1, b_1) + (c, d) \sim (a_2, b_2) + (c, d)$$

und

$$(a_1, b_1)\,(c, d) \sim (a_2, b_2)\,(c, d)$$

gelten.

Zunächst ist auf Grund der Definition (2) der Äquivalenz $(a_1, b_1) \sim (a_2, b_2)$

$$a_1 b_2 = a_2 b_1.$$

Multipliziert man diese Gleichung mit d, so ergibt sich

$$a_1 b_2 d = a_2 b_1 d.$$

Addiert man zu dieser letzten Gleichung $b_1 c b_2$, so erhält man

$$a_1 b_2 d + b_1 c b_2 = a_2 b_1 d + b_1 c b_2.$$

Schließlich multiplizieren wir diese letzte Gleichung mit d und klammern die gemeinsamen Faktoren der Summanden aus. Dann ergibt sich

$$(a_1 d + b_1 c)\, b_2 d = (a_2 d + b_2 c)\, b_1 d$$

und hieraus

$$(a_1 d + b_1 c, b_1 d) \sim (a_2 d + b_2 c, b_2 d).$$

Multipliziert man andererseits die Gleichung $a_1 b_2 = a_2 b_1$ mit cd, so erhält man

$$(a_1 c)\,(b_2 d) = (a_2 c)\,(b_1 d)$$

und damit

$$(a_1 c, b_1 d) \sim (a_2 c, b_2 d).$$

Somit wird also durch Definition 2 in der Menge Γ_0 der Klassen äquivalenter Paare von ganzen Zahlen tatsächlich eine eindeutig bestimmte Addition und Multiplikation erklärt.

Satz 6. *Die Menge Γ_0 bildet bezüglich der in Definition 2 erklärten Operationen der Addition und Multiplikation einen Körper.*

Beweis. Wir haben zu zeigen, daß in Γ_0 die Axiome I. bis VI. (§ 7, Definition 1), VII. und VIII. (§ 8, Definition 1) erfüllt sind. Da die Verknüpfungen zwischen den Klassen aus Γ_0 repräsentantenweise erklärt sind, folgt die Gültigkeit der Axiome I., II., IV., V. und VI. unmittelbar aus Satz 4. Offenbar enthält Γ_0 auch mehr als ein Element, d. h., das Axiom VIII. ist erfüllt. Ferner gilt in Γ_0 das Axiom III., denn wenn die Klasse α das Paar (a, b) und die Klasse β das Paar (c, d) enthält, so genügt wegen

$$(a, b) + (bc - ad, bd) = (abd + b^2 c - abd, b^2 d) \sim (c, d)$$

die Klasse γ, die das Paar $(bc - ad, bd)$ enthält, der Gleichung $\alpha + \gamma = \beta$. Damit ist bereits bewiesen, daß Γ_0 ein Ring ist.

Wir wollen zunächst die Frage klären, welche Klassen das Nullelement und die entgegengesetzten Elemente dieses Ringes sind. Zunächst stellt man unmittelbar fest, daß alle Paare der Form $(0, b)$ untereinander äquivalent sind. Andererseits muß aber auch jedes Paar (x, y), das einem Paar der Form $(0, b)$ äquivalent ist, diese Form besitzen, da aus $xb = y \cdot 0$ und $b \neq 0$ die Bedingung $x = 0$ folgt. Somit bildet die Menge der Paare der Form $(0, b)$ eine Klasse, die offenbar gerade Nullelement von Γ_0 ist. Weiterhin ist klar, daß zu der das Paar (a, b) enthaltenden Klasse α die Klasse entgegengesetzt ist, die das Paar $(-a, b)$ enthält. Sie möge mit $-\alpha$ bezeichnet werden.

Wir wollen jetzt die Gültigkeit des Axioms VII. in Γ_0 nachweisen. Dazu seien uns zwei Klassen α und β vorgegeben, wobei zudem α vom Nullelement von Γ_0

verschieden sei. Wenn die Klasse α das Paar (a, b) und die Klasse β das Paar (c, d) enthält, so ist (neben b und d) $a \neq 0$. Mithin gehört das Paar (bc, ad) zu der bei der Definition von Γ_0 betrachteten Menge M, und zwar möge es der Klasse γ angehören. Wegen

$$(a, b)\,(bc, ad) = (abc, abd) \sim (c, d)$$

ist dann aber $\alpha\gamma = \beta$, womit auch die Gültigkeit von Axiom VII. nachgewiesen ist. Es ist also in der Tat Γ_0 ein Körper.

Wir wollen noch die Frage behandeln, welche Klasse aus Γ_0 Einselement ist und was die inversen Elemente in Γ_0 sind. Wenn $\alpha\varepsilon = \alpha$ ist, wo α verschieden von Null sei und das Paar (a, b) enthalte, also $a \neq 0$ ist, so ist $(a, b)\,(x, y) \sim (a, b)$, also $abx = aby$, d. h. $x = y$, wenn (x, y) eine beliebiges Paar aus ε ist. Andererseits erfüllen alle Paare (x, x) mit $x \neq 0$ die Beziehung

$$(a, b)\,(x, x) \sim (a, b).$$

Die Menge aller dieser Paare bildet also eine Klasse, die in Γ_0 zudem die Rolle des Einselementes spielt.

Zu der das Paar (a, b) mit $a \neq 0$ enthaltenden Klasse α ist die Klasse invers, die das Paar (b, a) enthält, da $(a, b)\,(b, a) = (ab, ab)$ Element der Einsklasse von Γ_0 ist.

Der soeben konstruierte Körper Γ_0 ist zwar zu jedem Körper der rationalen Zahlen isomorph, aber noch kein Körper der rationalen Zahlen selbst, da die ganzen Zahlen nicht unter seinen Elementen vorkommen. Unser Ziel wird es daher sein, den Ring C der ganzen Zahlen in den Körper Γ_0 einzubetten. Dazu müssen wir zunächst eine dem Ring C der ganzen Zahlen isomorphe Teilmenge von Γ_0 finden. Hierzu betrachten wir die Klassen α, die ein Paar (b, c) enthalten, in dem b durch c teilbar, also etwa $b = ac$ ist. Offenbar sind je zwei Paare (ac_1, c_1) und (ac_2, c_2) äquivalent. Umgekehrt hat aber auch jedes zu einem Paar (ac, c) äquivalente Paar die Form (ac_1, c_1), da aus $(b_1, c_1) \sim (ac, c)$ zunächst $b_1 c = c_1 ac$ und hieraus $b_1 = ac_1$ folgt. Die betrachtete Klasse α besteht also aus genau den Paaren der Form (ac, c), wobei a eine gegebene ganze Zahl und c eine beliebige ganze Zahl $\neq 0$ ist.

Es sei nun C' die Menge aller Klassen von Paaren (b, c) mit durch c teilbarem b. Jeder Klasse α aus C' ordnen wir dann die ganze Zahl a zu, für die das Paar (ac, c) in der Klasse α liegt. Da für beliebiges c_1 und c_2 (ungleich Null) die Beziehung $(ac_1, c_1) \sim (ac_2, c_2)$ gilt, wird durch diese Festsetzung eine eindeutige Abbildung $a = f(\alpha)$ von der Menge C' in die Menge C der ganzen Zahlen definiert. Dabei entsprechen überdies zwei verschiedenen Klassen auch verschiedene ganze Zahlen und ist jede ganze Zahl a Bild einer gewissen Klasse, nämlich der Klasse, die das Paar (ac, c) enthält. Die Abbildung f ist also eine umkehrbar eindeutige Abbildung von C' auf C. Darüber hinaus stellt aber f sogar einen Isomorphismus zwischen C' und C bezüglich der oben definierten Klassenverknüpfungen einerseits und der Addition und Multiplikation von ganzen Zahlen andererseits her. Dazu haben wir zu zeigen, daß

$$f(\alpha) + f(\beta) = f(\alpha + \beta), \quad f(\alpha) \cdot f(\beta) = f(\alpha\beta) \tag{6}$$

ist. In der Tat: Wenn die Klasse α das Paar (ac, c) und die Klasse β das Paar (bc, c) enthält, so enthält die Klasse $\alpha + \beta$ das Paar

$$(ac, c) + (bc, c) = ((a + b)c^2, c^2)$$

und die Klasse $\alpha\beta$ das Paar

$$(ac, c)\,(bc, c) = (abc^2, c^2),$$

woraus sich

$$f(\alpha+\beta) = a+b = f(\alpha)+f(\beta)$$

und

$$f(\alpha\beta) = ab = f(\alpha)\cdot f(\beta)$$

ergeben.

Nach diesen Vorbereitungen sind wir in der Lage, nun endgültig den Körper Γ der rationalen Zahlen zu definieren. Es sei nämlich Γ die Menge, die man aus Γ_0 dadurch erhält, daß man jede Klasse aus C' durch die ihr vermöge der Abbildung f entsprechende ganze Zahl ersetzt. Zur Definition der Operationen in Γ setzen wir die Abbildung f auf die ganze Menge Γ_0 fort, indem wir $f(\alpha) = \alpha$ für alle die Klassen α aus Γ_0 setzen, die nicht in C' liegen. Dadurch wird f zu einer umkehrbar eindeutigen Abbildung von Γ_0 auf Γ. Dann können wir die Operationen in Γ einfach durch die Gleichungen

$$f(\alpha)+f(\beta) = f(\alpha+\beta), \quad f(\alpha)\cdot f(\beta) = f(\alpha\beta) \tag{7}$$

festlegen. Hierbei sind α und β beliebige Elemente aus Γ_0 und folglich $f(\alpha)$ und $f(\beta)$ beliebige Elemente aus Γ. Mithin werden durch die Gleichungen (7) tatsächlich Verknüpfungen in Γ erklärt.

Satz 7. *Die Menge Γ bildet bezüglich der durch die Gleichungen* (7) *in Γ erklärten Addition und Multiplikation einen Körper der rationalen Zahlen.*

Beweis. Dazu ist zu zeigen, daß Γ die Bedingungen 1.) bis 4.) aus Definition 1 erfüllt.

1.) Nach Konstruktion enthält Γ den Ring C der ganzen Zahlen.

2.) Γ ist ein Körper. Die Gleichungen (7), durch die in Γ die Addition und die Multiplikation erklärt wurden, zeigen nämlich, daß die Menge Γ bezüglich dieser Verknüpfungen dem Körper Γ_0 isomorph und daher selbst ein Körper ist (§ 9, Satz 1).

3.) Für die in Γ enthaltenen ganzen Zahlen stimmen die Addition und Multiplikation in Γ mit der entsprechenden Operation für ganze Zahlen überein. In der Tat: Bei der Abbildung f sind die ganzen Zahlen die Bilder der Elemente der Teilmenge C' von Γ_0; sind aber α und β Klassen aus C', so stimmen für sie die Gleichungen (7) mit den Gleichungen (6) überein, in denen die Addition bzw. Multiplikation auf der linken Seite die in § 20 definierten Verknüpfungen zwischen ganzen Zahlen sind.

4.) Der Körper Γ der rationalen Zahlen enthält keinen von sich selbst verschiedenen Unterkörper, der den Ring C der ganzen Zahlen enthält. Hierfür beweisen wir, daß jedes Element von Γ ein Quotient aus ganzen Zahlen ist. Dazu beachten wir, daß sich jedes Element aus Γ in der Form $f(\alpha)$ darstellen läßt, wo α eine gewisse Klasse aus Γ_0 ist. Diese Klasse α möge etwa das Paar (k, l) ganzer Zahlen enthalten, wobei $l \neq 0$ ist. Ferner sei $k = f(\beta)$ und $l = f(\gamma)$. Auf Grund der Definition der Abbildung f besteht dann β aus allen Paaren der Form (kc, c) und γ aus allen Paaren der Form (lc, c). Folglich enthält die Klasse $\alpha\gamma$ das Paar

$$(k, l)\,(lc, c) = (klc, lc) \sim (kc, c),$$

d. h., es ist $\alpha\gamma = \beta$. Auf Grund der Definition der Multiplikation in Γ [zweite der Gleichungen (7)] erhält man daraus $f(\alpha) \cdot f(\gamma) = f(\beta)$, es ist also in der Tat

$$f(\alpha) = \frac{f(\beta)}{f(\gamma)} = \frac{k}{l}.$$

Wenn also ein beliebiger Unterkörper des Körpers Γ alle ganzen Zahlen und damit auch alle Quotienten aus ganzen Zahlen enthält, so enthält er nach dem Bewiesenen den Körper Γ und ist mithin gleich Γ.

Damit ist der behauptete Satz in allen Teilen bewiesen.

Im Verlauf der vorangehenden Überlegungen haben wir einen der untereinander isomorphen Körper der rationalen Zahlen konstruiert. Seine Elemente sind einmal die ganzen Zahlen und zum anderen die Klassen äquivalenter Paare (a, b) von ganzen Zahlen, in denen $b \neq 0$ und a nicht durch b teilbar ist. Hiermit haben wir die Frage nach der Existenz eines Körpers der rationalen Zahlen, d. h. eines Körpers, der die Definition 1 erfüllt, beantwortet. Es bleibt noch übrig, für die rationalen Zahlen die üblichen Bezeichnungen durch Brüche einzuführen und zu zeigen, daß die so konstruierten Zahlen die gewöhnlichen, allgemein bekannten Eigenschaften besitzen.

§ 23. Eigenschaften der rationalen Zahlen

Wir wollen zunächst für die als Elemente des im vorangehenden Paragraphen konstruierten Körpers Γ definierten rationalen Zahlen die übliche Bezeichnung durch Brüche einführen. Wir haben gesehen, daß jede rationale Zahl a Bild einer gewissen Klasse α des Körpers Γ_0, also $a = f(\alpha)$ ist. Nun ist jede Klasse eindeutig durch jedes beliebige in ihr enthaltene Paar (k, l) von ganzen Zahlen bestimmt, wobei definitionsgemäß $l \neq 0$ ist. Somit legt jedes derartige Paar (k, l) in eindeutiger Weise eine rationale Zahl a fest. In diesem Sinne wollen wir im folgenden zur Bezeichnung der Zahl a die Schreibweise $\frac{k}{l}$ verwenden. Die Zeichen $\frac{k}{l}$, wobei k und l ganze Zahlen sind und $l \neq 0$ ist, sollen *Brüche* genannt werden.[1])

Wir haben andererseits das Zeichen $\frac{k}{l}$ zur Bezeichnung des Quotienten aus k und l im Körper Γ benutzt. Dies kann jedoch nicht zu Verwechslungen führen. Wenn nämlich $a = f(\alpha)$ ist und die Klasse α das Paar (k, l) enthält, so ist (wie wir am Ende des vorangehenden Paragraphen gezeigt haben) $a = \frac{k}{l}$, wobei jetzt $\frac{k}{l}$ der Quotient aus k und l im Körper Γ ist. Nach dem Gesagten bezeichnen alle Brüche, die man aus den Paaren einer Klasse α erhält, dieselbe rationale Zahl $a = f(\alpha)$.

---

[1]) Im Gegensatz zur üblichen Auffassung der Brüche als Zahlen einer besonderen Kategorie werden hier die Brüche nicht als Zahlen, sondern als Zeichen zur Bezeichnung der Zahlen angesehen. In der Tat können verschiedene Brüche ein und dieselbe Zahl bezeichnen. So ist $\frac{2}{3} = \frac{4}{6} = \frac{6}{9} \cdots$ (Hierbei bezeichnet das Gleichheitszeichen jedoch nun nicht die Identität der Brüche — die Brüche sind Zeichen! —, sondern die Identität der durch sie bezeichneten rationalen Zahlen. — *Anm. d. wissenschaftl. Red.*)

Somit ist auf Grund der Definition der Äquivalenz (2) von Paaren

$$\frac{a}{b} = \frac{c}{d} \tag{1}$$

dann und nur dann, wenn $ad = bc$ ist. Als fundamentale Eigenschaft der Brüche folgt hieraus, daß für jedes $c \neq 0$

$$\frac{a}{b} = \frac{ac}{bc} \tag{2}$$

ist. Auf dieser Eigenschaft beruht bekanntlich das Kürzen von Brüchen und die Möglichkeit, verschiedene Brüche auf einen gemeinsamen Nenner zu bringen.

Wir bemerken, daß $a = \frac{k}{l}$ eine ganze Zahl ist, wenn k durch l teilbar ist. Als einfachste Darstellung einer ganzen Zahl a in Form eines Bruches erhalten wir dementsprechend die Darstellung $a = \frac{a}{1}$. Neben der Bezeichnung der ganzen Zahlen durch Brüche soll jedoch auch die frühere Bezeichnungsweise beibehalten werden. So ist

$$\frac{6}{3} = \frac{4}{2} = \frac{2}{1} = 2, \quad \frac{-15}{3} = \frac{-5}{1} = -5.$$

Da jeder Bruch $\frac{k}{l}$ eine wohlbestimmte rationale Zahl, nämlich den Quotienten der Zahlen k und l im Körper Γ, bezeichnet, gelten für die Addition, Subtraktion, Multiplikation und Division der durch Brüche bezeichneten Zahlen die Regeln (1) b), c), d) aus § 22, also die bekannten Verknüpfungsregeln für Brüche.

Eine rationale Zahl, die keine ganze Zahl ist, wollen wir eine *gebrochene Zahl* nennen (wir unterscheiden also in der Terminologie wohl zwischen einem „Bruch" und einer „gebrochenen Zahl"). Die Gesamtheit der rationalen Zahlen setzt sich also aus den ganzen und den gebrochenen Zahlen zusammen.

Bemerkung 1. Für die rationalen Zahlen als Elemente des Körpers Γ gelten alle die in den Paragraphen 7 und 8 für die Elemente eines beliebigen Ringes oder Körpers bewiesenen Sätze. So also u. a. bei der Multiplikation die bekannten Vorzeichenregeln [§ 7, (3)]. Ferner gibt es in Γ ein Einselement, welches übrigens gleich der Zahl 1 ist, da diese vermöge des Isomorphismus f der Einsklasse von Γ_0 zugeordnet ist [weil nämlich diese Klasse aus allen Paaren $(c, c) = (c \cdot 1, c)$ mit $c \neq 0$ besteht]. Überdies gibt es zu jeder Zahl $\frac{a}{b} \neq 0$ eine inverse, die gleich der Zahl $\frac{b}{a}$ ist. Schließlich gibt es in Γ keine Nullteiler (§ 8, Satz 1) usw.

Wir wenden uns als nächstes der Frage nach der Anordnung des Körpers der rationalen Zahlen zu.

Satz 1. *Der Körper Γ der rationalen Zahlen kann auf eine und nur eine Weise angeordnet werden. Bei dieser Anordnung ist eine Zahl $a = \frac{k}{l}$ dann und nur dann positiv, wenn die ganze Zahl kl positiv ist. Für die in Γ enthaltenen ganzen Zahlen stimmt diese Anordnung mit der früheren* (§ 21, Satz 3) *Anordnung der ganzen Zahlen überein.*

Beweis. Eine rationale Zahl $a = \frac{k}{l}$ mit $k \neq 0$ werde als positiv angesehen, wenn k und l gleiches Vorzeichen besitzen, wenn also beide positiv oder beide negativ

sind. Es soll also $a = \frac{k}{l}$ dann und nur dann positiv (in Zeichen: $a > 0$) sein, wenn $kl > 0$ ist im Sinne der Anordnung der ganzen Zahlen. Diese Festsetzung ist unabhängig von der speziellen Wahl des zur Bezeichnung von a benutzten Bruches. Wenn nämlich $a = \frac{k_1}{l_1} = \frac{k_2}{l_2}$ und $k_1 l_1 > 0$ ist, so ergibt sich nach Multiplikation dieser letzten Ungleichung mit der positiven ganzen Zahl l_2^2

$$k_1 l_1 l_2^2 = (k_1 l_2)\,(l_1 l_2) = (k_2 l_1)\,(l_1 l_2) = k_2 l_2 l_1^2 > 0$$

und hieraus auf Grund von $l_1^2 > 0$ die Ungleichung $k_2 l_2 > 0$ (§ 10, Satz 3).

Diese Definition der positiven rationalen Zahlen erfüllt, wie wir sofort sehen werden, die Axiome IX. und X. aus § 10.

Dazu sei $a = \frac{k}{l}$. Da für die ganzen Zahlen das Axiom IX. erfüllt ist, gilt einer und nur einer der drei nachfolgenden Fälle:

$$kl > 0, \quad kl = 0, \quad -kl > 0.$$

Wenn $kl > 0$ ist, so ist $a > 0$; wenn $kl = 0$ ist, so ist $k = 0$ und mithin $a = 0$; wenn $-kl > 0$ ist, so ist $-a = \frac{-k}{l} > 0$. Also gilt das Axiom IX. auch für die rationalen Zahlen.

Wenn

$$a_1 = \frac{k_1}{l_1} > 0 \quad \text{und} \quad a_2 = \frac{k_2}{l_2} > 0$$

ist, so ist auch

$$a_1 + a_2 = \frac{k_1 l_2 + k_2 l_1}{l_1 l_2} > 0,$$

weil

$$(k_1 l_2 + k_2 l_1)\, l_1 l_2 = (k_1 l_1)\, l_2^2 + (k_2 l_2)\, l_1^2 > 0$$

ist. Ebenso ist auch

$$a_1 a_2 = \frac{k_1 k_2}{l_1 l_2} > 0,$$

weil

$$(k_1 k_2)\,(l_1 l_2) = (k_1 l_1)\,(k_2 l_2) > 0$$

ist. Es gilt also für die rationalen Zahlen auch das Axiom X. Damit ist gezeigt, daß vermöge der obigen Festsetzung der Körper Γ ein angeordneter Körper ist.

Man sieht unmittelbar ein, daß die Gültigkeit der Axiome IX. und X. sich von einem beliebigen Ring oder Körper auf jeden seiner Unterringe überträgt. Daher erzeugt die obige Anordnung des Körpers der rationalen Zahlen eine ganz bestimmte Anordnung des in ihm enthaltenen Ringes C der ganzen Zahlen. Nun läßt sich aber der Ring der ganzen Zahlen nur auf eine einzige Weise anordnen (§ 21, Satz 3), d. h., bezüglich jeder Anordnung des Körpers der rationalen Zahlen (also auch speziell bezüglich der oben betrachteten) stimmt die Anordnung der im Körper der rationalen Zahlen enthaltenen ganzen Zahlen mit der früheren Anordnung des Ringes C (§ 21) überein.

Wir wollen nun zeigen, daß die oben erklärte Anordnung des Körpers der rationalen Zahlen auch die einzige ist. Dazu sei uns eine beliebige Anordnung des Körpers der rationalen Zahlen vorgegeben. Diese muß nach dem Gesagten für die in I' enthaltenen ganzen Zahlen mit der Anordnung der ganzen Zahlen aus § 21 übereinstimmen. Wenn nun $\frac{k}{l} > 0$ ist im Sinne der vorgegebenen Anordnung, so ergibt sich nach Multiplikation dieser Ungleichung mit $l^2 > 0$, daß $kl > 0$ im Sinne der vorgegebenen Anordnung, also auch im Sinne der Anordnung der ganzen Zahlen sein muß. Ist umgekehrt $kl > 0$, so muß auch $\frac{k}{l} > 0$ sein, weil anderenfalls $-\frac{k}{l} \geq 0$ wäre, woraus sich nach Multiplikation mit $l^2 > 0$ die Ungleichung $-kl \geq 0$ ergäbe, was der Voraussetzung $kl > 0$ widerspricht. Es ist also eine rationale Zahl $a = \frac{k}{l}$ im Sinne der vorgegebenen Anordnung dann und nur dann positiv, wenn die ganze Zahl kl positiv ist. Die vorgegebene Anordnung stimmt also — wie behauptet — mit der oben betrachteten Anordnung der rationalen Zahlen überein.

Bemerkung 2. Die rationalen Zahlen besitzen also alle Eigenschaften der Elemente eines angeordneten Körpers, wie wir sie in § 10 dargelegt haben. So erhält man eine Ordnung des Körpers der rationalen Zahlen, wenn man festsetzt, daß dann und nur dann $a > b$ sein soll, wenn $a - b$ positiv ist. Bezüglich dieser Ordnung ist die Zahl 0 kleiner als alle positiven und größer als alle negativen rationalen Zahlen (§ 10, Satz 1). Für diese Ordnung gelten ferner die Monotoniegesetze und die bekannten Rechenregeln für das Rechnen mit Ungleichungen (§ 10, Satz 2 bis Satz 4). Der Körper der rationalen Zahlen ist als angeordneter Körper von der Charakteristik 0 (§ 10, Satz 6). Schließlich kann man auch für die rationalen Zahlen den absoluten Betrag einer Zahl a als die nichtnegative der Zahlen $\pm a$ definieren und für diesen die üblichen Vergleichsregeln (Dreiecksungleichung usw.) sowie die Regeln für die Rückführung des Rechnens mit Zahlen auf das Rechnen mit ihren Beträgen beweisen (§ 10, Satz 8 und die nachfolgende Bemerkung).

Es sei nun K ein beliebiger Körper der Charakteristik Null (§ 8, Definition 2) und e das Einselement von K. Wir definieren dann das Produkt ax eines Elementes x des Körpers K mit einer rationalen Zahl a auf folgende Weise: Wenn $a = \frac{k}{l}$ ist, wobei k und l ganze Zahlen sind und $l \neq 0$ ist, so ist auch $le \neq 0$, und wir können

$$ae = \frac{k}{l} e = \frac{ke}{le}, \quad ax = (ae)\, x$$

setzen. Falls a eine ganze Zahl ist, stimmt diese Definition mit der entsprechenden Definition aus § 7 überein. Wenn nämlich $a = \frac{k}{l}$ ist, so ist $al = k$ und wegen (5) aus § 7

$$(ae)\,(le) = (al)\, e = ke,$$

also $ae = \frac{ke}{le}$; mithin ist $(ae)\, x = a(ex) = ax$, d. h., für ganzzahliges a stimmt diese neue Definition in der Tat mit der Definition aus § 7 überein.

Die Elemente der Form ae heißen *ganze* bzw. *rationale Elemente* des Körpers K, je nachdem, ob a eine ganze oder eine rationale Zahl ist.

Satz 2. *Jeder Körper Q der Charakteristik Null enthält einen und nur einen dem Körper Γ der rationalen Zahlen isomorphen Unterkörper Π. Dieser Unterkörper Π besteht aus allen rationalen Elementen ae des Körpers Q, und es gibt nur eine isomorphe Abbildung von Π auf Γ, nämlich die Abbildung, die dem Element ae die rationale Zahl a zuordnet. Demzufolge besitzt speziell der Körper der rationalen Zahlen keine von sich selbst verschiedenen Unterkörper, d. h., Γ ist Primkörper* (§ 8, Definition 4), *und es gibt nur eine isomorphe Abbildung von Γ auf sich, nämlich die identische Abbildung. Der Körper Q ist ferner isomorph einem Körper P, der den Körper Γ als Unterkörper enthält, wobei jede isomorphe Abbildung von Q auf P die Elemente von Π in der oben beschriebenen Weise auf die Elemente von Γ abbildet. Ist der Körper Q außerdem angeordnet, so kann auch der Körper P angeordnet werden und zwar so, daß ein Isomorphismus von P auf Q die Anordnung erhält.*

Beweis. Zunächst gilt für beliebige ganze Zahlen m und n [§ 6, (6) und § 7, (5)]:

a) $$me + ne = (m+n)e, \quad (me)(ne) = (mn)e.$$

Da die Charakteristik von Q gleich Null sein sollte, ist $ne \neq 0$ für jede ganze Zahl $n \neq 0$. Ist also $m \neq n$ so ist $m - n \neq 0$ und mithin $me - ne = (m-n)e \neq 0$. Daraus folgt, daß die Zuordnung $n \leftrightarrow ne$ eine umkehrbar eindeutige Abbildung zwischen dem Ring C der ganzen Zahlen und der Menge S der ganzen Elemente von Q, und wegen a) sogar ein Isomorphismus ist.

Ebenso folgen aus a) und den Regeln für die Addition und die Multiplikation von Quotienten in einem Körper [§ 8, Satz 3b) und c)] für beliebige rationale Zahlen $a = \frac{k}{l}$ und $b = \frac{m}{n}$ die Gleichungen

b) $$ae + be = (a+b)e, \quad (ae)(be) = (ab)e,$$

weil

$$ae + be = \frac{ke}{le} + \frac{me}{ne} = \frac{(ke)(ne) + (le)(me)}{(le)(ne)} = \frac{(kn+lm)e}{(ln)e} = (a+b)e,$$

$$(ae)(be) = \frac{ke}{le} \cdot \frac{me}{ne} = \frac{(ke)(me)}{(le)(ne)} = \frac{(km)e}{(ln)e} = (ab)e$$

ist.

Wenn $a = \frac{k}{l} \neq 0$ ist, so ist $k \neq 0$ und mithin $ae = \frac{ke}{le} \neq 0$. Hieraus folgt genau wie oben, daß mit $a \neq b$ auch $ae \neq be$ ist. Daher ist die Zuordnung $a \leftrightarrow ae$ eine umkehrbar eindeutige Abbildung von dem Körper Γ auf die Menge Π der rationalen Elemente von Q, die wegen b) sogar ein Isomorphismus ist. Da Γ ein Körper ist, ist dann auch Π ein Körper (§ 9, Satz 1).

Sei nun Γ auf irgendeine Weise isomorph auf einen gewissen Unterkörper Π' des Körpers Q abgebildet. Der Zahl 1 entspricht dann notwendig das Einselement e des Körpers Q. Auf Grund der Isomorphie-Eigenschaften ist dann aber auch

$$n = 1 + \cdots + 1 \leftrightarrow e + \cdots + e = ne$$

und

$$-n \leftrightarrow -ne = (-n)e, \quad 0 \leftrightarrow 0 = 0 \cdot e$$

(in der letzten Zuordnungsvorschrift steht links die Zahl 0 und rechts das Nullelement von Q). Es ist also $n \leftrightarrow ne$ für jede ganze Zahl n. Dann ist aber auch

$a = \frac{m}{n} \leftrightarrow \frac{me}{ne} = ae$ für jede rationale Zahl $a = \frac{m}{n}$. Folglich ist Π' gleich Π und jeder Isomorphismus zwischen Γ und Π gleich dem Isomorphismus $a \leftrightarrow ae$.

Da der Körper Q den zu Γ isomorphen Körper Π als Unterkörper enthält, ist Q isomorph dem Körper P, den man erhält, wenn man in Q die Elemente von Π durch die ihnen entsprechenden Zahlen aus Γ ersetzt (§ 9, Satz 2). Hierbei muß jeder Isomorphismus von P auf Q für die Elemente aus Γ mit dem oben angegebenen Isomorphismus übereinstimmen, weil Γ nur dem Unterkörper Π von Q isomorph ist und es zudem nur eine isomorphe Abbildung von Γ auf Π gibt.

Ist schließlich der Körper Q angeordnet und $y = f(x)$ eine beliebige isomorphe Abbildung von P auf Q, so erhält man — wie man leicht einsieht — eine Anordnung des Körpers P, wenn man diejenigen Elemente x aus P als positiv auszeichnet, deren Bilder $y = f(x)$ in Q positiv sind. Es ist klar, daß sich diese Anordnung bei dem Isomorphismus f überträgt.

Damit ist der behauptete Satz in allen Teilen bewiesen.

Dieser Satz zeigt, daß unter allen Körpern der Charakteristik Null der Körper der rationalen Zahlen im angegebenen Sinne minimal ist. Wenn wir diese Körper nur bis auf Isomorphie betrachten, so können wir sagen, daß der Körper der rationalen Zahlen Unterkörper jedes Körpers der Charakteristik 0 ist.

Satz 3. *Der Körper Γ der rationalen Zahlen ist (bezüglich seiner einzig möglichen Anordnung) archimedisch angeordnet.*

Beweis. Zum Nachweis der Gültigkeit des Archimedischen Axioms in Γ brauchen wir — wie in jedem angeordneten Körper — nur zu zeigen, daß es zu jeder rationalen Zahl c eine natürliche Zahl n gibt, die größer als c ist. Dann gibt es nämlich zu jedem a und b mit $b > 0$ eine natürliche Zahl n mit $n > \frac{a}{b}$, woraus sich nach Multiplikation mit b sofort $nb > a$ ergibt.

Sei also a eine beliebige rationale Zahl. Ist $a \leq 0$, so ist $n > a$ für jede natürliche Zahl n. Ist hingegen $a > 0$, so läßt sich a als Bruch der Form $a = \frac{k}{l}$ darstellen, wobei k und l natürliche Zahlen sind. Wenn nämlich $a > 0$ ist, so ist auf Grund von Satz 1 das Produkt $kl > 0$, d. h., k und l besitzen gleiches Vorzeichen, wobei nach (2) das Vorzeichen von k und l noch gleichermaßen verändert werden kann. Wenn aber $l \geq 1$ ist, so ist wegen $a > 0$ die natürliche Zahl $k = al \geq a$, woraus sich für $n = k + 1$ die Ungleichung $n > a$ ergibt, was zu beweisen war.

Eine Teilbarkeitstheorie ist im Körper der rationalen Zahlen, wie auch in jedem anderen Körper, ohne Interesse und führt nur zu dem Schluß, daß jede Zahl durch jede andere von Null verschiedene Zahl teilbar ist.

Für die Anwendungen der Mathematik in der Technik und den anderen Wissenschaften benötigt man durchaus nicht alle rationalen Zahlen, sondern kann sich beispielsweise auf diejenigen Zahlen beschränken, die sich in der Form endlicher Dezimalbrüche darstellen lassen. Bei allen Messungen und bei Berechnungen angewandter Natur braucht man jeweils das Resultat nur bis zu einem gewissen Genauigkeitsgrad zu kennen. Durch die Zahlen der genannten Art kann man aber in der Tat bereits jede beliebige Genauigkeit erreichen. Zur Präzisierung dieser Behauptung geben wir zunächst die folgende

Definition. *Vorgegeben sei eine natürliche Zahl n. Alle Zahlen der Form mn^k, wobei m und k beliebige ganze Zahlen sind, nennen wir n-adisch rational oder kurz n-rational.*

Im Falle $n = 2, 3, 10$ sprechen wir speziell von dyadisch-, triadisch- bzw. dekadisch-rationalen Zahlen.

Für $k = 0$ ergibt sich, daß jede ganze Zahl n-rational ist, wie auch immer die natürliche Zahl n gewählt ist.

Im folgenden werden wir zeigen, daß man bei allen Näherungsrechnungen die rationalen Zahlen durch die n-rationalen Zahlen ersetzen kann. Wir beweisen die nachfolgenden beiden Behauptungen nicht speziell für den Körper Γ der rationalen Zahlen, sondern gleich in einer etwas allgemeineren Form, in der wir sie im nachfolgenden Kapitel benötigen werden.

Satz 4. *Vorgegeben sei ein archimedisch angeordneter Körper K, der den Körper Γ der rationalen Zahlen enthält. Ferner sei a ein Element aus K und n eine beliebige natürliche Zahl größer als 1. Dann gibt es zu jeder ganzen Zahl k eine ganze Zahl m derart, daß*

$$m n^k \leq a < (m+1)\, n^k$$

ist.

Beweis. Aus $n > 1 > 0$ folgt, daß $n^k > 0$ ist. Weil voraussetzungsgemäß K archimedisch angeordnet ist, gibt es natürliche Zahlen l_1 und l_2 derart, daß $l_1 n^k > a$ und $l_2 n^k > -a$, also $(-l_2) n^k < a$ ist. Folglich enthält die Menge A der ganzen Zahlen l, für die $l n^k \leq a$ ist, die Zahl $-l_2$, ist also nicht leer. Ferner ist diese Menge A nach oben beschränkt, weil man aus $l n^k \leq a < l_1 n^k$ sofort auf $l < l_1$ schließen kann. Daher enthält die Menge A eine größte Zahl m (§ 21, Satz 5). Dann gehört also m zu A, während die Zahl $m + 1 > m$ nicht mehr zu A gehört. Nach Definition von A ist dies gleichbedeutend mit

$$m n^k \leq a < (m+1) n^k,$$

was zu beweisen war.

Satz 5. *Vorgegeben sei ein archimedisch angeordneter Körper K, der den Körper Γ der rationalen Zahlen enthält. Ferner sei n eine beliebige natürliche Zahl größer als 1. Dann gibt es zu jedem positiven Element a des Körpers K eine natürliche Zahl k derart, daß $\frac{1}{n^k} < a$ ist.*

Beweis. Wir zeigen zunächst, daß für jede natürliche Zahl $n > 1$ und jede ganze Zahl k die Ungleichung

$$n^k > k \tag{3}$$

gilt.

Da stets $n^k > 0$ ist, gilt diese Ungleichung für alle ganzen Zahlen $k \leq 0$. Für die restlichen ganzen Zahlen, also für die natürlichen Zahlen beweisen wir sie bei gegebenem n durch vollständige Induktion über k. Wegen $n^1 = n > 1$ gilt sie für $k = 1$. Wenn diese Ungleichung für die natürliche Zahl k gilt, wenn also $n^k > k$ ist, so ist

$$n^{k+1} = n \cdot n^k > n k \geq 2k = k + k \geq k + 1,$$

d. h., unsere Ungleichung gilt auch für $k + 1$.

Wenn nun $a > 0$ ist, so gibt es auf Grund des Archimedischen Axioms eine natürliche Zahl k mit $1 < ka$. Wegen (3) ist dann aber auch $1 < n^k a$. Multipliziert man diese Ungleichung mit $n^{-k} > 0$, so erhält man $n^{-k} < a$, was zu beweisen war.

Wir weisen darauf hin, daß wegen Satz 2 die letzten beiden Sätze auch in beliebigen archimedisch angeordneten Körpern K gelten, sofern man in ihrer Formulierung die rationalen Zahlen durch die rationalen Elemente von K (also die Zahlen r durch die Elemente re, wobei e das Einselement von K ist) ersetzt.

Aus den Sätzen 4 und 5 folgt, daß bei allen Näherungsrechnungen an Stelle der rationalen Zahlen n-rationale Zahlen mit einem beliebigen $n > 1$ genommen werden können. Insbesondere lassen sich hierzu die durch endliche Dezimalbrüche dargestellten Zahlen ($n = 10$) verwenden, was auch in der Praxis häufig geschieht. Wir wollen dies hier noch etwas näher ausführen: Man sagt, das Resultat einer Rechnung sei durch rationale Zahlen mit einer Genauigkeit c festgelegt, wobei c eine rationale Zahl > 0 ist, wenn man zwei rationale Zahlen a und b mit $b > a$ (die untere bzw. obere Fehlergrenze der Rechnung) gefunden hat derart, daß $b - a < c$ ist und das Ergebnis der Rechnung zwischen a und b liegt. Auf Grund von Satz 5 gibt es dann eine ganze Zahl k mit

$$n^k < \frac{c-(b-a)}{2}.$$

Ferner kann man nach Satz 4 zwei ganze Zahlen l und m finden derart, daß

$$a_1 = ln^k \leq a < (l+1)n^k \text{ und } (m-1)n^k \leq b < mn^k = b_1$$

ist. Da nach Konstruktion von a_1 und b_1 das Intervall (a_1, b_1) das Intervall (a, b) umfaßt, liegt das Ergebnis der Rechnung zwischen a_1 und b_1. Darüber hinaus ist aber auch

$$\begin{aligned} b_1 - a_1 &= (b_1 - b) + (b - a) + (a - a_1) \\ &\leq n^k + (b-a) + n^k < (b-a) + 2\,\frac{c-(b-a)}{2} = c, \end{aligned}$$

d. h., die n-rationalen Zahlen a_1 und b_1 approximieren das Ergebnis der Rechnung mit demselben Genauigkeitsgrad c wie die rationalen Zahlen a und b. Durch entsprechende Überlegungen kann man übrigens auch die Zahl c noch durch eine n-rationale Zahl ersetzen.

Für die genaue Darstellung des Ergebnisses einer Rechnung sind jedoch im allgemeinen nicht nur die n-rationalen Zahlen, sondern sogar alle rationalen Zahlen unzureichend.

Gesucht sei z. B. die Länge einer Strecke MN bezüglich der Strecke AB als Einheitsstrecke. Die gesuchte Länge stellt sich dann als das Verhältnis der Strecken MN und AB dar. Sind die Strecken AB und MN überhaupt vergleichbar (kommensurabel), so gibt es ein Grundmaß CD, welches p mal in MN und q mal in AB enthalten ist. Dann ist das Verhältnis $MN : AB = \frac{p}{q}$ eine rationale Zahl. Ist umgekehrt das Verhältnis $MN : AB = \frac{p}{q}$ rational, so erhält man ein gemeinsames Maß der Strecken MN und AB, wenn man die Strecke AB in q gleichlange Teile teilt (von denen jeder p mal in MN enthalten ist), dann sind also MN und AB vergleichbar. Aus der Geometrie ist jedoch bekannt, daß es inkommensurable Strecken gibt. So ist z. B. die Diagonale eines Quadrates stets inkommensurabel mit den Seiten des Quadrates. Wählt man nämlich eine Seite des Quadrates als Maßeinheit der Streckenmessung, so läßt sich ihr Verhältnis zur Diagonalen durch keine rationale Zahl darstellen.

Die rationalen Zahlen reichen auch nicht dazu aus, um aus jeder positiven rationalen Zahl oder auch nur aus jeder natürlichen Zahl die Wurzel zu ziehen. In der Tat: Wenn p eine Primzahl und n eine natürliche Zahl größer als 1 ist, so kann $\sqrt[n]{p}$ keine rationale Zahl sein. Anderenfalls wäre nämlich (sofern man für gerades n den positiven Wert der Wurzel wählt) $\sqrt[n]{p} = \frac{q}{r}$, wobei q und r natürliche Zahlen sind. Dann wäre aber $p = \frac{q^n}{r^n}$ und mithin

$$pr^n = q^n. \tag{4}$$

Wenn nun p in der Primzahlzerlegung von q genau a mal und in der Primzahlzerlegung von r genau b mal vorkommt, so würde die Zahl p auf der linken Seite von (4) $na+1$ mal und auf der rechten Seite von (4) nb mal als Faktor auftreten. Nun ist aber $na+1 \neq nb$, da zwar nb, aber nicht $na+1$ durch n teilbar ist; es würde also der Faktor p in der Primzahlzerlegung der rechten und der linken Seite von (4) in verschiedener Anzahl auftreten, was dem Satz über die eindeutige Zerlegbarkeit einer natürlichen Zahl in Primfaktoren[1]) widerspricht.

Im folgenden Kapitel werden wir uns mit der Erweiterung des Körpers der rationalen Zahlen zum Körper der reellen Zahlen beschäftigen, in dem unter anderem der genaue Vergleich von Strecken und das Ausziehen der Wurzel aus einer positiven Zahl stets möglich sein werden.

Kapitel VI

DER KÖRPER DER REELLEN ZAHLEN

§ 24. Vollständige und stetige Körper

Die Existenz inkommensurabler Strecken war bereits im alten Griechenland bekannt. Das Bestreben, genaue Zahlenwerte für ihren Vergleich zu erhalten, hätte notwendig zum Begriff der irrationalen Zahl geführt. Dieser Schritt zu ihrer Begründung wurde indes von den Gelehrten des Altertums nicht getan. Vielmehr gaben sie in ihrem Streben nach einer strengen Begründung der mathematischen Aussagen diesen eine geometrische Form. Als Beispiel dieser eigenartigen geometrischen Algebra können die „Elemente" EUKLIDS gelten.

Im Mittelalter benutzten auch die Inder irrationale Ausdrücke, aber auch sie blieben in der Frage nach ihrer Begründung an der Oberfläche. Erst mit der Entwicklung der Infinitesimalrechnung im 17. und 18. Jahrhundert wurden die reellen Zahlen zum Gegenstand eingehenderer Untersuchungen. Dabei operierte man zunächst mit ihnen in rein anschaulicher Weise, indem man sie auf die Punkte einer Geraden abbildete.

---

[1]) Vgl. den nachfolgenden Artikel von A. J. CHINTSCHIN.

In der zweiten Hälfte des 19. Jahrhunderts reifte das Bedürfnis nach einem exakten Aufbau der Theorie der reellen Zahlen soweit heran, daß ihre Begründung von einer Reihe von Mathematikern (DEDEKIND, CANTOR, WEIERSTRASS) in Angriff genommen wurde. Ihre Begründungsarten, die in der äußeren Form vollkommen verschieden sind, sind aber insofern gleichwertig, als sie zu einander isomorphen Zahlenbereichen führen. Wir werden im folgenden den Aufbau CANTORS vorführen, der eng an den Begriff des Grenzwertes anknüpft, den wir oben schon betrachtet haben. In der Literatur findet man auch häufig den Aufbau der reellen Zahlen nach DEDEKIND behandelt, mit dem sich der Leser gut an Hand des Buches von DEDEKIND [10] bekannt machen kann. Eine vortreffliche Darstellung der DEDEKINDschen Theorie, die reich an methodologischen Hinweisen ist, findet sich in dem Buch [11] von A. J. CHINTSCHIN (der deutsche Leser sei auf das Buch „Irrationalzahlen" von O. PERRON, 4. Aufl., Berlin 1960, und auf die „Grundlagen der Analysis", von E. LANDAU, Leipzig 1930, hingewiesen. — *Anm. d. wissenschaftl. Red.*).

Wir haben bereits am Ende des § 23 gezeigt, daß sich die Verhältnisse von Strecken und die Wurzeln aus positiven Zahlen nicht immer durch rationale Zahlen ausdrücken lassen. Unser Ziel wird es im folgenden sein, den Körper Γ der rationalen Zahlen zum Körper D der reellen Zahlen zu erweitern, in dem diese (und eine umfangreiche Klasse weiterer) Aufgaben stets lösbar sein werden.

Um zu einer möglichst naheliegenden Definition des Körpers der reellen Zahlen zu gelangen, wollen wir zunächst an Hand der beiden detaillierten Aufgaben herausarbeiten, was für Eigenschaften die Zahlen haben müssen, die diese Aufgaben lösen.

Gesucht sei also das Verhältnis der Strecken AB und MN. Dazu messen wir auf der Strecke MN vom Punkte M aus die Strecke $MM_1 = AB$ ab, sodann von M_1 aus in derselben Richtung die Strecke $M_1M_2 = AB$ usw. Auf Grund der geometrischen Fassung des Axioms von ARCHIMEDES gibt es eine natürliche Zahl n derart, daß wir, wenn wir n mal die Strecke AB auf diese Weise abgetragen haben, eine Strecke $n \cdot AB > MN$ erhalten. Jedenfalls ist also die Menge der ganzen Zahlen k, für die $k \cdot AB \leq MN$ ist, nach oben beschränkt. Sie ist auch nicht leer, da die Zahl 0 sicher in ihr enthalten ist. Also enthält diese Menge eine größte Zahl a_0 (§ 21, Satz 5) d. h., wenn wir $a_0 + 1 = b_0$ setzen, so ist

$$a_0 \cdot AB \leq MN < b_0 \cdot AB.$$

Daher liegt das gesuchte Verhältnis $MN : AB$ sicher zwischen a_0 und b_0. Nun teilen wir die Strecke AB in 10 gleiche Teile und wiederholen mit einem von ihnen, etwa A_1B_1, die vorangehenden Überlegungen. Auf diese Weise kommen wir zu ganzen Zahlen a_1' und $b_1' = a_1' + 1$, für die

$$a_1' \cdot A_1B_1 \leq MN < b_1' \cdot A_1B_1$$

oder, wenn wir

$$a_1 = \frac{a_1'}{10}, \quad b_1 = \frac{b_1'}{10}$$

setzen,

$$a_1 \cdot AB \leq MN < b_1 \cdot AB, \quad b_1 - a_1 = \frac{1}{10}$$

gilt. Aus

$$10a_0 \cdot A_1B_1 = a_0 \cdot AB \leq MN < b_0 \cdot AB = 10b_0 \cdot A_1B_1$$

erhalten wir auf Grund der Definition von a_1':

$$10a_0 \leq a_1' < 10b_0,$$

woraus sich zunächst

$$b_1' = a_1' + 1 \leq 10b_0$$

und dann

$$a_0 \leq \frac{a_1'}{10} = a_1 \text{ und } b_0 \geq \frac{b_1'}{10} = b_1$$

ergibt. Wenn wir dieses Verfahren fortlaufend wiederholen, so erhalten wir zwei Zahlenfolgen a_n und b_n, die die Bedingungen

$$\left.\begin{array}{ll} \text{a)} & a_0 \leq a_1 \leq a_2 \leq \ldots, \\ \text{b)} & b_0 \geq b_1 \geq b_2 \geq \ldots, \\ \text{c)} & b_n - a_n = \frac{1}{10^n}, \quad n = 0, 1, 2, \ldots \end{array}\right\} \tag{1}$$

erfüllen. Dabei liegt das gesuchte Verhältnis der Strecken MN und AB stets zwischen a_n und b_n, und zwar nähern sich die Zahlen dieser Folgen für wachsendes n immer mehr dem Wert des gesuchten Verhältnisses. Wenn nämlich eine positive rationale Zahl ε vorgegeben ist, so kann man eine natürliche Zahl n_0 finden derart, daß sich die Zahlen a_n und b_n für alle $n > n_0$ voneinander (und damit auch von dem gesuchten Verhältnis) um weniger als ε unterscheiden. In der Tat gibt es ein n_0, für welches $\frac{1}{10^{n_0}} < \varepsilon$ (§ 23, Satz 5) und mithin $b_n - a_n = \frac{1}{10^n} < \varepsilon$ für alle $n > n_0$ gilt.

Als nächstes soll $\sqrt[k]{a}$ bestimmt werden, wobei a eine positive rationale Zahl und k eine natürliche Zahl größer als 1 sei. Dabei soll zunächst nur vom positiven Wert der Wurzel die Rede sein. Vorgegeben sei eine ganze Zahl $n \geq 0$. Dann ist $10^{-n} > 0$, und es gibt nach dem Archimedischen Axiom eine Zahl m, so daß $m \cdot 10^{-n} > a + 1$ ist. Nun gilt für jede rationale Zahl $b > 1$ und jede natürliche Zahl $k > 1$ die Ungleichung $b^{k-1} > 1$ (§ 10, Satz 4), also $b^k > b$. Daher ist

$$(m \cdot 10^{-n})^k > m \cdot 10^{-n} > a + 1 > a.$$

Es sei nun A die Menge aller ganzen Zahlen l, für die $(l \cdot 10^{-n})^k \leq a$ ist. Die Menge dieser Zahlen ist nach oben beschränkt und nicht leer, da sie die Zahl 0 enthält. Also enthält sie eine größte Zahl a_n'. Setzen wir

$$b_n' = a_n' + 1, \quad a_n = a_n' \cdot 10^{-n}, \quad b_n = b_n' \cdot 10^{-n},$$

so ist

$$a_n^k \leq a < b_n^k.$$

Mithin liegt die gesuchte Wurzel $\sqrt[k]{a}$ zwischen a_n und b_n. Ferner ist $b_n - a_n = 10^{-n}$. Da die Zahlen der Form $m \cdot 10^{-n}$ gleich den Zahlen der Form

$$m' \cdot 10^{-(n+1)}$$

sind, ist

$$a_n = a_n' \cdot 10^{-n} = 10 \cdot a_n' \cdot 10^{-(n+1)} \leq a_{n+1}' \cdot 10^{-(n+1)} = a_{n+1}.$$

Da andererseits

$$a < (b'_n \cdot 10^{-n})^k = (10 \cdot b_n \cdot 10^{-(n+1)})^k$$

ist, ist auch

$$a'_{n+1} < 10 \cdot b'_n$$

und daher

$$b'_{n+1} = a'_{n+1} + 1 \leq 10 \cdot b'_n,$$

$$b_{n+1} = b'_{n+1} \cdot 10^{-(n+1)} \leq 10 \cdot b'_n \cdot 10^{-(n+1)} = b_n.$$

Wir erhalten also wiederum Folgen a_n und b_n, die den Bedingungen (1) genügen. Berücksichtigen wir ferner, daß die gesuchte Wurzel für jedes n zwischen a_n und b_n liegt, so können wir über die Annäherung der Zahlen dieser Folgen an den Wert der Wurzel dasselbe aussagen wie im Fall des Streckenverhältnisses.

Das eigentliche Problem liegt jedoch darin, daß es im allgemeinen keine rationale Zahl gibt, gegen die die Zahlen a_n und b_n in der beschriebenen Weise streben. Vielmehr müssen wir, wenn es für beliebige Folgen a_n und b_n von rationalen Zahlen, die den Bedingungen (1) genügen, eine derartige Zahl geben soll, neue (nichtrationale) Zahlen einführen. Zu diesem Zweck müssen wir zunächst den Begriff der Zahlenfolge näher präzisieren und seine wichtigsten Eigenschaften feststellen.

Definition 1. *Unter einer Folge von Elementen einer gegebenen nichtleeren Menge M verstehen wir eine Funktion* (§ 3, Definition 1) $f(n) = a_n$, *die auf der Menge N der natürlichen Zahlen erklärt ist und deren Werte der Menge M angehören. Mit anderen Worten: Eine Folge ist eine Vorschrift, die jeder natürlichen Zahl n eindeutig ein gewisses Element a_n der Menge M zuordnet.*

Im folgenden bezeichnen wir Folgen durch $a_1, a_2, a_3, \ldots$ oder kürzer durch $\{a_n\}$. Das Element a_n heißt *das n-te Glied* der Folge $\{a_n\}$.

Wir bemerken, daß die Glieder einer Folge nicht notwendig verschiedene Elemente der Menge M sein müssen.

Zunächst einige Beispiele für Folgen:

1.) Die Folge der natürlichen Zahlen $1, 2, 3, \ldots = \{n\}$.

2.) $1, \frac{1}{2}, \frac{1}{3}, \ldots = \left\{\frac{1}{n}\right\}$.

3.) $1, 0, 1, 0, \ldots = \{a_n\}$, wobei a_n der Rest bei der Division von n durch 2 ist.

4.) $+1, -2, +3, -4, \ldots = \{n \cdot (-1)^{n+1}\}$.

5.) $3, \frac{3}{2}, \frac{7}{3}, \frac{7}{4}, \frac{11}{5}, \frac{11}{6}, \ldots = \left\{2 + \frac{(-1)^{n+1}}{n}\right\}$.

6.) $2, 3, 5, 7, \ldots = \{p_n\}$, wobei p_n die n-te Primzahl ist. Hier können wir keine geschlossene Formel für das n-te Glied der Folge angeben. Trotzdem ist diese Folge einwandfrei definiert. Man braucht hierzu nur das Prinzip der induktiven Definitionen zu benutzen (§ 15, Definition 1). Setzt man nämlich $f(1) = 2$ und verlangt man, daß $f(n)$ die kleinste Primzahl sein soll, die größer als $f(n-1)$ ist, so besagt Satz 1 aus § 15, daß es genau eine Funktion auf der Menge der natürlichen Zahlen gibt, die diese Forderungen erfüllt. Dieses Beispiel zeigt darüber hinaus, daß auch eine Funktion nicht unbedingt durch eine geschlossene Formel gegeben zu sein braucht, auf Grund derer der Wert durch das Argument bestimmt ist.

Die nachfolgenden Begriffe haben in beliebigen Mengen keinen Sinn, sie beziehen sich nur auf geordnete Mengen bzw. angeordnete Ringe. Wir beschränken uns bei unseren Ausführungen auf angeordnete Körper, die den Körper der rationalen Zahlen enthalten.

In diesem Sinne verstehen wir in diesem Paragraphen unter K einen angeordneten Körper, der den Körper Γ der rationalen Zahlen als Unterkörper enthält. Alles in diesem Paragraphen über einen derartigen Körper K Ausgesagte bleibt (auf Grund der in § 23, Satz 2 bewiesenen Isomorphie) auch für beliebige angeordnete Körper Q richtig, sofern man überall die rationalen Zahlen r durch die ihnen entsprechenden Elemente re ersetzt, wobei e das Einselement des Körpers Q ist.

Definition 2. *Eine Folge $\{a_n\}$ aus Elementen eines Körpers K heißt nach oben* (entsprechend: *nach unten*) *beschränkt, wenn es in K ein Element a gibt derart, daß $a_n < a$* (entsprechend: $a_n > a$) *für alle n. Sie heißt beschränkt, wenn sie sowohl nach oben als auch nach unten beschränkt ist, wenn es also in K ein Element $a > 0$ gibt derart, daß $|a_n| < a$ für alle n.*

Die unter Beispiel 4.) angeführte Folge ist weder nach oben noch nach unten beschränkt. Die in den Beispielen 2.), 3.) und 5.) angegebenen Folgen sind beschränkt.

Als nächstes führen wir einen der wichtigsten Grundbegriffe der gesamten Mathematik ein.

Definition 3. *Ein Element a des Körpers K heißt Grenzwert der Folge $\{a_n\}$ von Elementen aus K, wenn es zu jedem positiven ε aus K eine natürliche Zahl n_0 gibt (die durchaus von der Wahl von ε abhängen kann) derart, daß $|a_n - a| < \varepsilon$ für alle $n > n_0$. Man schreibt dafür $a = \lim\limits_{n \to \infty} a_n$ (in Worten: „a ist Limes von a_n für n gegen Unendlich“) oder kurz $a = \lim a_n$ („a ist Limes von a_n“). Eine Folge a_n, die den Grenzwert a besitzt, heißt gegen a konvergent oder kurz konvergent. Eine Folge, die (in K) keinen Grenzwert besitzt, heißt divergent.*

Von den oben betrachteten Folgen sind nur zwei konvergent, nämlich die Folge 2.) mit dem Grenzwert 0 und die Folge 5.) mit dem Grenzwert 2. In der Tat: Für die Folge 2.) ist

$$|a_n - 0| = |a_n| = a_n = \frac{1}{n},$$

und für die Folge 5.) ist

$$|a_n - 2| = \left|\frac{(-1)^{n+1}}{n}\right| = \frac{1}{n}.$$

Da es im Körper der rationalen Zahlen auf Grund des Archimedischen Axioms (§ 23, Satz 3) zu jedem rationalen $\varepsilon > 0$ eine natürliche Zahl $n_0 > \frac{1}{\varepsilon}$ gibt, ist dann $\frac{1}{n} < \frac{1}{n_0} < \varepsilon$ für alle $n > n_0$.

Die in Beispiel 3.) angegebene Folge ist divergent. Zwar gibt es zu jedem $\varepsilon > 0$ und jedem n_0 ein $n' > n_0$, derart, daß $|a_{n'} - 0| = 0 < \varepsilon$, und ein $n'' > n_0$ derart, daß $|a_{n''} - 1| = 0 < \varepsilon$ ist, aber sofern nur $\varepsilon \leq 1$ ist, gibt es kein n_0 derart, daß eine dieser Ungleichungen für alle $n > n_0$ erfüllt wäre. In der Tat: Wäre z. B. $|a_n - 0| = |a_n| < \varepsilon \leq 1$, so wäre $a_n = 0$ und mithin $a_{n+1} = 1$ und damit $|a_{n+1} - 0| = 1 \geq \varepsilon$.

Der Begriff des Grenzwertes hat einige Ähnlichkeit mit dem Begriff der algebraischen Operation (§ 6, Definition 1). Dort haben wir jedem geordneten Paar von

Elementen einer gewissen Menge ein Element derselben Menge zugeordnet, während hier gewissen Systemen von Elementen, die nach dem Typus der natürlichen Zahlen $\{1, 2, 3, \ldots\}$ geordnet sind, ein gewisses Element entspricht. In diesem Sinne spricht man häufig von „der Operation des Grenzüberganges". Natürlich ist sie keine algebraische Operation im Sinne von Definition 1 aus § 6 mehr.

Es drängt sich nun sofort die Frage nach der Ausführbarkeit und der Eindeutigkeit dieser Operation des Grenzüberganges auf. Wir sahen schon am Beispiel der Folge 3.), daß nicht jede Folge einen Grenzwert besitzt. Wir wollen deshalb zunächst die Frage nach der Eindeutigkeit behandeln. Hier gilt:

Satz 1. *Wenn eine Folge von Elementen des Körpers K überhaupt einen Grenzwert besitzt, so auch nur einen.*

Beweis. Es sei $\lim a_n = a$ und $b \neq a$. Wir zeigen, daß b nicht Grenzwert der Folge $\{a_n\}$ sein kann. Anschaulich würde man so argumentieren, daß die Elemente a_n, da sie gegen a streben, für hinreichend große Indizes von b abweichen. Formal geht dieser Beweis folgendermaßen vor sich: Da $a \neq b$ ist, ist $|a - b| > 0$ und damit $\left|\frac{a-b}{2}\right| > 0$. Wäre nun auch $\lim a_n = b$, so gäbe es einerseits eine natürliche Zahl n_1 mit $|a_n - a| < \left|\frac{a-b}{2}\right|$ für alle $n > n_1$ und andererseits eine natürliche Zahl n_2 mit $|a_n - b| < \left|\frac{a-b}{2}\right|$ für alle $n > n_2$. Ist nun n_0 die größere der Zahlen n_1 und n_2, so wäre für alle $n > n_0$:

$$|a - b| = |(a - a_n) + (a_n - b)| \leq |a - a_n| + |a_n - b|$$
$$< \left|\frac{a-b}{2}\right| + \left|\frac{a-b}{2}\right| = |a - b|,$$

also $|a - b| < |a - b|$, was nicht der Fall ist.

Die Frage nach der Existenz des Grenzwertes wollen wir vorläufig noch zurückstellen und zunächst einige Eigenschaften der Operation des Grenzüberganges unter der Voraussetzung ihrer Ausführbarkeit ableiten.

Satz 2. a) *Wenn eine der Folgen $\{a_n\}$ oder $\{b_n\}$ von Elementen des Körpers K konvergiert und wenn $\lim (a_n - b_n) = 0$ ist, so konvergiert auch die andere, und es ist $\lim a_n = \lim b_n$.*

Umgekehrt: Wenn Folgen $\{a_n\}$ und $\{b_n\}$ gegen denselben Grenzwert konvergieren, so ist $\lim (a_n - b_n) = 0$.

Ferner: Wenn die Folgen $\{a_n\}$ und $\{b_n\}$ von Elementen aus K konvergieren, so ist

b) $\lim (a_n \pm b_n) = \lim a_n \pm \lim b_n$;

c) $\lim (a_n \cdot b_n) = \lim a_n \cdot \lim b_n$;

d) $\lim \frac{a_n}{b_n} = \frac{\lim a_n}{\lim b_n}$, *sofern $b_n \neq 0$ für alle n und $\lim b_n \neq 0$ ist.*

Die Konvergenz der Folgen auf der linken Seite der Gleichungen b), c) *und* d) *braucht nicht vorausgesetzt zu werden, sondern folgt bereits aus der Konvergenz der Folgen $\{a_n\}$ und $\{b_n\}$.*

e) *Wenn $\lim a_n > \lim b_n$ ist, so gibt es ein Element $\varepsilon > 0$ in K und eine natürliche Zahl n_0 derart, daß $a_n - b_n > \varepsilon$ für alle $n > n_0$. Wenn es eine natürliche Zahl n_0 gibt derart, daß $a_n \geq b_n$ für alle $n > n_0$ ist, so ist auch $\lim a_n \geq \lim b_n$.*

Beweis. a) Es möge etwa die Folge $\{a_n\}$ konvergieren und $\lim a_n = a$ sein. Dann gibt es zu jedem $\varepsilon > 0$ aus K natürliche Zahlen n_1 und n_2 derart, daß $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n > n_1$ und $|a_n - b_n| < \frac{\varepsilon}{2}$ für alle $n > n_2$ ist. Ist nun n_0 die größere der Zahlen n_1 und n_2, so ist

$$|b_n - a| = |(b_n - a_n) + (a_n - a)| \leq |b_n - a_n| + |a_n - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für alle $n > n_0$ und mithin

$$\lim b_n = a = \lim a_n.$$

Die zweite Behauptung von a) folgt unmittelbar aus b).

Es seien jetzt Folgen $\{a_n\}$ und $\{b_n\}$ vorgegeben mit $\lim a_n = a$ und $\lim b_n = b$.

b) Zu jedem $\varepsilon > 0$ gibt es natürliche Zahlen n_1 und n_2 derart, daß $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n > n_1$ und $|b_n - b| < \frac{\varepsilon}{2}$ für alle $n > n_2$ ist. Ist nun n_0 die größere der Zahlen n_1 und n_2, so gilt für jedes $n > n_0$:

$|(a_n \pm b_n) - (a \pm b)| = |(a_n - a) \pm (b_n - b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$,

also ist

$$\lim (a_n \pm b_n) = a \pm b = \lim a_n \pm \lim b_n.$$

c) Wir zeigen zunächst, daß die als konvergent vorausgesetzte Folge $\{a_n\}$ beschränkt ist (vgl. Definition 2). Da $\lim a_n = a$ ist, gibt es eine natürliche Zahl p derart, daß $|a_n - a| < 1$ für alle $n > p$ ist. Dann ist aber für alle $n > p$:

$$|a_n| = |(a_n - a) + a| \leq |a_n - a| + |a| < 1 + |a|.$$

Ferner gibt es unter den endlich vielen Elementen $|a_1|, |a_2|, \ldots, |a_p|, 1 + |a|$ des Körpers K ein größtes Element a' (§ 5, Satz 6). Setzen wir nun $c = a' + 1$, so ist $c \geq 1 > 0$ und $|a_n| < c$ für alle n.

Zum Beweis der Behauptung c) greifen wir ein beliebiges Element $d > |b|$, etwa $d = |b| + 1$ heraus. Offenbar ist $d > 0$. Da $\lim a_n = a$ und $\lim b_n = b$ ist, gibt es zu jedem $\varepsilon > 0$ aus K natürliche Zahlen n_1 und n_2 derart, daß $|a_n - a| < \frac{\varepsilon}{2d}$ für alle $n > n_1$ und $|b_n - b| < \frac{\varepsilon}{2c}$ für alle $n > n_2$ ist. Nehmen wir nun als n_0 die größere der Zahlen n_1 und n_2, so ist für jedes $n > n_0$:

$$\begin{aligned} |a_n b_n - ab| &= |(a_n b_n - a_n b) + (a_n b - ab)| \\ &\leq |a_n b_n - a_n b| + |a_n b - ab| = |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| \\ &< c \cdot \frac{\varepsilon}{2c} + \frac{\varepsilon}{2d} \cdot d = \varepsilon, \end{aligned}$$

also ist

$$\lim (a_n b_n) = ab = \lim a_n \cdot \lim b_n.$$

d) Wir zeigen zunächst, daß es unter der Voraussetzung $\lim b_n = b \neq 0$ eine natürliche Zahl n_1 gibt derart, daß $|b_n| > \frac{|b|}{2}$ für alle $n > n_1$. Dazu beachten wir, daß es eine natürliche Zahl p gibt mit $|b_n - b| < \frac{|b|}{2}$ für alle $n > p$. Wäre nun die

zu beweisende Behauptung falsch, so gäbe es eine Zahl $q > p$ derart, daß $|b_q| \leqq \frac{|b|}{2}$ ist. Dann wäre aber

$$|b| = |(b - b_q) + b_q| \leqq |b - b_q| + |b_q| < \frac{|b|}{2} + \frac{|b|}{2} = |b|,$$

also $|b| < |b|$, was nicht der Fall ist.

Ferner ist die Folge $\{a_n\}$, da sie konvergent ist, beschränkt, d. h., es gibt in K ein Element $c > 0$ mit $|a_n| < c$ für jedes n.

Schließlich gibt es wegen $\lim a_n = a$ und $\lim b_n = b$ zu jedem $\varepsilon > 0$ aus K natürliche Zahlen n_2 und n_3 mit $|a_n - a| < \frac{\varepsilon \cdot |b|}{2}$ für alle $n > n_2$ und $|b_n - b| < \frac{\varepsilon \cdot b^2}{4c}$ für alle $n > n_3$ (da aus $b \neq 0$ jedenfalls $b^2 = |b|^2 > 0$ folgt).

Ist nun n_0 die größte unter den Zahlen n_1, n_2 und n_3, so gilt für jedes $n > n_0$:

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| = \left|\frac{a_n b - b_n a}{b_n b}\right| = \frac{|(a_n b - a_n b_n) + (a_n b_n - b_n a)|}{|b_n b|}$$

$$\leqq \frac{|a_n b - a_n b_n|}{|b_n b|} + \frac{|a_n b_n - b_n a|}{|b_n b|} = \frac{|a_n|\,|b - b_n|}{|b|\,|b_n|} + \frac{|a_n - a|}{|b|}$$

$$< \frac{c\,\frac{\varepsilon b^2}{4c}}{\frac{|b|}{2}\,|b|} + \frac{\frac{\varepsilon |b|}{2}}{|b|} = \varepsilon,$$

also ist

$$\lim \frac{a_n}{b_n} = \frac{a}{b} = \frac{\lim a_n}{\lim b_n}.$$

e) Es sei $a > b$. Dann ist $\varepsilon = \frac{a - b}{3} > 0$. Mithin gibt es natürliche Zahlen n_1 und n_2 derart, daß $|a_n - a| < \varepsilon$ für jedes $n > n_1$ und $|b_n - b| < \varepsilon$ für alle $n > n_2$ ist. Nehmen wir als n_0 die größere der Zahlen n_1 und n_2, so wäre, wenn $a_n - b_n \leqq \varepsilon$ für ein gewisses $n > n_0$ gelten würde:

$$a - b = (a - a_n) + (a_n - b_n) + (b_n - b) < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon = a - b,$$

was nicht der Fall ist. Also ist in der Tat $a_n - b_n > \varepsilon$ für alle $n > n_0$.

Ist umgekehrt $a_n - b_n \geqq 0$ für alle n, die größer als ein gewisses n_0 sind, so gäbe es, wenn $a < b$ wäre, nach dem bereits Bewiesenen ein $\varepsilon > 0$ und eine natürliche Zahl n_1 derart, daß $b_n - a_n > \varepsilon > 0$ für alle $n > n_1$ ist. Wählen wir nun eine Zahl n, die größer als n_0 und größer als n_1 ist, so erhalten wir einerseits $a_n \geqq b_n$ und zum andern $b_n > a_n$, was nicht möglich ist. Folglich muß $a \geqq b$ sein, womit der behauptete Satz in allen Teilen bewiesen ist.

Wenn eine Folge überhaupt einen Grenzwert besitzt, so müssen sich mit wachsenden Indizes die Glieder dieser Folge, indem sie dem Grenzwert beliebig nahe kommen, auch untereinander beliebig nähern. Folgen dieser Art wollen wir einen besonderen Namen geben:

Definition 4. *Eine Folge $\{a_n\}$ von Elementen eines Körpers K heißt eine Fundamentalfolge (oder* CAUCHY*-Folge) in K, wenn es zu jedem $\varepsilon > 0$ aus K eine natürliche Zahl n_0 (die noch von ε abhängen kann) gibt derart, daß $|a_p - a_q| < \varepsilon$ für alle p und q, die größer als n_0 sind.*

Satz 3. *Jede konvergente Folge von Elementen des Körpers K ist eine Fundamentalfolge.*

Beweis. Wenn $\lim a_n = a$ ist, so gibt es zu jedem $\varepsilon > 0$ aus K eine natürliche Zahl n_0 mit $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n > n_0$. Wenn nun $p > n_0$ und $q > n_0$ ist, so gilt auf Grund der Dreiecksungleichung [§ 10, (3)]:

$$|a_p - a_q| = |(a_p - a) - (a_q - a)| \leq |a_p - a| + |a_q - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

also ist $\{a_n\}$ in der Tat eine Fundamentalfolge.

Dieser Satz gibt uns also eine notwendige Bedingung für die Konvergenz einer Folge: Damit eine Folge konvergiert, ist notwendig, daß sie eine Fundamentalfolge ist. Diese Bedingung ist aber für beliebige Körper K keineswegs hinreichend. So gibt es im Körper der rationalen Zahlen — wie wir gleich sehen werden — Fundamentalfolgen, die in diesem Körper keinen Grenzwert besitzen.

Wir kehren dazu noch einmal zu den Aufgaben zurück, die wir uns am Anfang dieses Kapitels stellten. In jedem der beiden Fälle haben wir zwei Folgen von rationalen (sogar dekadisch rationalen) Zahlen $\{a_n\}$ und $\{b_n\}$ konstruiert, die die Bedingungen (1) erfüllen. Man sieht unmittelbar, daß jede dieser Folgen eine Fundamentalfolge ist. Da es nämlich zu jedem rationalen $\varepsilon > 0$ eine natürliche Zahl n_0 gibt mit $\frac{1}{10^{n_0}} < \varepsilon$ (§ 23, Satz 5), gilt für jedes p und q mit z. B. $p \geq q > n_0$:

$$|a_p - a_q| = a_p - a_q < b_p - a_q \leq b_{n_0} - a_{n_0} = \frac{1}{10^{n_0}} < \varepsilon$$

und entsprechend $|b_p - b_q| < \varepsilon$.

Wenn nun eine der Aufgaben eine rationale Zahl c als Lösung besitzt, so muß c Grenzwert sowohl der Folge $\{a_n\}$ als auch der Folge $\{b_n\}$ sein. Dann ist nämlich im Falle der Strecken einerseits $c \cdot AB = MN < b_n \cdot AB$, also $c < b_n$, und andererseits $a_n \cdot AB \leq MN = c \cdot AB$, also $a_n \leq c$. Entsprechend folgt im Falle der Wurzel aus $c^k = a$, daß $a_n \leq c < b_n$ ist, da $a_n > c$ die Ungleichung $a_n^k > c^k = a$ und $b_n \leq c$ die Ungleichung $b_n^k \leq c^k = a$ zur Folge hätte, was jedoch der Konstruktion der Zahlen a_n und b_n widerspricht. Aus $a_n \leq c < b_n$ folgt, daß es zu jedem $\varepsilon > 0$ ein n_0 derart gibt, daß $\frac{1}{10^{n_0}} < \varepsilon$ und damit für jedes $n > n_0$

$$|a_n - c| = c - a_n < b_n - a_n \leq b_{n_0} - a_{n_0} = \frac{1}{10^{n_0}} < \varepsilon$$

und entsprechend $|b_n - c| < \varepsilon$ gilt, daß also tatsächlich

$$\lim a_n = \lim b_n = c$$

ist. Die gesuchte Lösung kann also, falls es überhaupt eine gibt, in der angegebenen Weise durch einen Grenzübergang erhalten werden.

Umgekehrt: Wenn z. B. die Folge $\{a_n\}$ einen rationalen Grenzwert c besitzt, so ist auch $\lim b_n = c$ und c eine Lösung der betreffenden Aufgabe. In der Tat: Wenn $\lim a_n = c$ ist, so ist $a_n \leq c \leq b_n$ für jedes n; denn wäre für ein gewisses n_1 etwa $a_{n_1} > c$, so wäre für jedes $n > n_1$

$$a_n > a_{n_1} > c, \qquad |a_n - c| = a_n - c \geq a_{n_1} - c,$$

oder wäre für ein gewisses n_2 etwa $b_{n_2} < c$, so wäre für jedes $n > n_2$

$$a_n < b_n \leq b_{n_2} < c, \qquad |a_n - c| = c - a_n > c - b_{n_2},$$

was beides der Definition des Grenzwertes widerspricht. Aus $a_n \leq c \leq b_n$ folgt aber, wie wir bereits oben sahen, $\lim a_n = \lim b_n = c$.

Wir werden im Fall der Wurzel später auf Grund eines allgemeineren Satzes zeigen, daß die so bestimmte Zahl c im Bereich der reellen Zahlen eine Lösung der gestellten Aufgabe ist. Hier beweisen wir zunächst nur, daß $c^k = a$ für rationales $a > 0$ und natürliches $k > 1$ ist, falls die am Anfang dieses Paragraphen konstruierten Folgen $\{a_n\}$ und $\{b_n\}$ den rationalen Grenzwert c besitzen. Angenommen nämlich, es wäre $c^k < a$, so wäre auf Grund von Satz 2 c) wegen $\lim b_n = c$ auch $\lim b_n^k = c^k$ und mithin gäbe es eine natürliche Zahl n_0 mit $|b_n^k - c^k| < a - c^k$ für alle $n > n_0$. Wegen $b_n \geqq c \geqq a_n \geqq 0$ ist aber $b_n^k \geqq c^k$, also wäre sogar

$$|b_n^k - c^k| = b_n^k - c^k < a - c^k,$$

d. h. $b_n^k < a$, im Widerspruch zur Konstruktion von b_n. Entsprechend führt man den Fall $c^k > a$ zum Widerspruch. Damit ist gezeigt, daß $c^k = a$, also $c = \sqrt[k]{a}$ ist.

Wenn es also für ein vorgegebenes rationales $a > 0$ keine rationale Zahl c gibt mit $c^k = a$ (siehe das Ende von § 23), so besitzen die für dieses a und k konstruierten Folgen $\{a_n\}$ und $\{b_n\}$ im Körper der rationalen Zahlen keinen Grenzwert, obwohl sie Fundamentalfolgen sind.

Für das Streckenverhältnis ist zu zeigen, daß dann, wenn die für die Strecken AB und MN konstruierten Folgen $\{a_n\}$ und $\{b_n\}$ von rationalen Zahlen gegen eine rationale Zahl c konvergieren, die Zahl c gleich dem Verhältnis dieser Strecken, also $c \cdot AB = MN$ ist. Angenommen, dies wäre nicht der Fall und es wäre z. B. $c \cdot AB < MN$, also $c \cdot AB = MN_1$, wobei MN_1 ein echter Teil der Strecke MN ist. So klein die Strecke N_1N auch immer wäre, stets gäbe es auf Grund des Archimedischen Axioms eine natürliche Zahl k derart, daß $k \cdot N_1N > AB$ ist. Nun ist $10^k > k$ [§ 23, (3)], also wäre auch $10^k \cdot N_1N > AB$. Hieraus folgte, daß $\frac{AB}{10^k} < N_1N$ sein müßte. Die Zahl a_k war nun so gewählt, daß

$$a_k \cdot AB \leqq MN < b_k \cdot AB$$

ist, wobei

$$b_k - a_k = \frac{1}{10^k}$$

gilt. Also wäre wegen $a_k \leqq c$

$$b_k \cdot AB = a_k \cdot AB + (b_k - a_k) \cdot AB \leqq c \cdot AB + \frac{AB}{10^k} < MN_1 + N_1N = MN,$$

im Widerspruch zu $b_k \cdot AB > MN$. Entsprechend erhalten wir einen Widerspruch, wenn wir annehmen, es sei $c \cdot AB > MN$. Also ist $c \cdot AB = MN$.

Wenn also die Strecken AB und MN inkommensurabel sind, so kann ihr Verhältnis nicht durch eine rationale Zahl ausgedrückt werden, und mithin besitzen die für sie konstruierten Folgen $\{a_n\}$ und $\{b_n\}$ von rationalen Zahlen keinen rationalen Grenzwert, obwohl sie Fundamentalfolgen sind.

Wir haben an Hand dieser Beispiele gezeigt, daß es im Körper der rationalen Zahlen Fundamentalfolgen gibt, die keinen Grenzwert besitzen.

Definition 5. *Ein angeordneter Körper K heißt vollständig, wenn in ihm die folgende Bedingung erfüllt ist:*

XII. (Vollständigkeitsaxiom) *Jede Fundamentalfolge von Elementen aus K ist konvergent, d. h., sie besitzt in K einen Grenzwert.*

Aus dem Gesagten folgt:

Satz 4. *Der Körper Γ der rationalen Zahlen ist nicht vollständig.*

Wir haben oben zwei Beweise für diesen Satz gegeben, indem wir für die inkommensurablen Strecken und für die rationalen Zahlen, die nicht k-te Potenz einer rationalen Zahl sind, Fundamentalfolgen konstruiert haben, die nicht konvergieren. Im Falle des Streckenverhältnisses stützten wir uns dabei auf Sätze aus der Geometrie, die wir hier nicht systematisch begründet haben. Der andere Beweis dagegen fußte auf von uns bewiesenen Eigenschaften der rationalen Zahlen und ist daher lückenlos.

Bemerkung. Die in diesem Paragraphen eingeführten Begriffe des Grenzwertes, der Fundamentalfolge und des damit zusammenhängenden Begriffs des vollständigen Körpers unterscheiden sich in einem wesentlichen Punkte von den früher betrachteten Begriffen der algebraischen Operation, der Anordnung und der archimedischen Anordnung: Vorgegeben sei ein Körper K und ein Unterkörper K' von K. Wenn für die Elemente a, b, c des Unterkörpers K' die Beziehung $a + b = c$ gilt, so gilt sie vermöge der Definition des Unterkörpers (§ 8, Definition 3) auch in dem Oberkörper K. Umgekehrt: Wenn in K die Beziehung $a + b = c$ gilt und wenn die Elemente a, b, c dem Unterkörper K' angehören, so gilt auch in K' die Beziehung $a + b = c$. Dasselbe gilt für die Beziehung $ab = c$. Ist zudem der Körper K angeordnet, so überträgt sich die Anordnung auch auf K', indem wir nämlich ein Element a aus K' als positiv ansehen, wenn es im Oberkörper K positiv ist. Man zeigt, daß bezüglich dieser Anordnung in K' die Axiome IX. und X. (§ 10, Definition 1) der Anordnung erfüllt sind, daß also in der Tat K' ein angeordneter Körper ist. Auch die Eigenschaft einer Anordnung von K', archimedisch zu sein, hängt nicht davon ab, ob wir K' als selbständigen Körper oder als Unterkörper von K ansehen. In der Tat gilt für die Elemente e und a aus K' die Beziehung $ne > a$ dann und nur dann in K', wenn sie in K (bezüglich derselben Ordnung) erfüllt ist. In diesem Sinne sind die in Kapitel II eingeführten Begriffsbildungen *absolut*; sie hängen nicht von dem umgebenden Körper ab. Anders steht es mit den im vorliegenden Paragraphen eingeführten Begriffsbildungen; sie sind durchaus davon abhängig, in welchem Körper man die jeweiligen Elemente betrachtet. In diesem Sinne sind diese Begriffe *relativ*. So besagt die Relation $\lim a_n = a$, daß es zu jedem Element $\varepsilon > 0$ des Körpers K eine natürliche Zahl n_0 derart gibt, daß $|a_n - a| < \varepsilon$ für alle $n > n_0$ ist. Entsprechend kommt in der Definition der Fundamentalfolge ein Element $\varepsilon > 0$ des Körpers K vor. Die Frage, welche Elemente dabei zur Konkurrenz kommen, hängt hier maßgeblich von der Wahl des Körpers K ab. Es liegt kein Grund zu der Annahme vor, daß für die Elemente einer Folge $\{a_n\}$ und ein Element a aus einem Unterkörper K' des Körpers K die Relation $\lim a_n = a$ und die Eigenschaft der Folge $\{a_n\}$, Fundamentalfolge zu sein, in K und in K' zusammenfallen werden. Klar ist nur, daß aus dem Bestehen einer dieser Bedingungen in K das Bestehen derselben Bedingung in K' folgt. Wenn eine dieser Bedingungen für alle $\varepsilon > 0$ aus K und gewisse Elemente aus K' erfüllt ist, so gilt sie insbesondere für jedes $\varepsilon > 0$ aus K'. Man muß sich nur vor dem umgekehrten Schluß hüten. Wir wollen dies noch durch ein Beispiel unterstreichen.

Sei K der Körper der rationalen Funktionen (d. h. der algebraischen Brüche) $\frac{f(x)}{g(x)}$, wobei $f(x)$ und $g(x)$ Polynome mit rationalen Koeffizienten sind. Eine Funktion $\frac{f(x)}{g(x)}$ heißt positiv, wenn die höchsten Koeffizienten von $f(x)$ und $g(x)$ gleiches

Vorzeichen haben. Auf diese Weise erhalten wir eine Anordnung des Körpers K, die allerdings nicht archimedisch ist, da für jede natürliche Zahl n die Ungleichung $x - n = \frac{x-n}{1} > 0$ besteht, woraus sich $n \cdot 1 < x$ ergibt. Die Funktion x ist auch größer als jede rationale Zahl. Wenn also $a > 0$ und rational ist, so ist auch $a^{-1} > 0$ rational und mithin $a^{-1} < x$. Wenn wir diese letzte Ungleichung mit $\frac{a}{x} > 0$ multiplizieren, so ergibt sich, daß $\frac{1}{x} < a$ ist. Es ist also $\frac{1}{x}$ kleiner als jede positive rationale Zahl. Der Körper K enthält ferner den Körper Γ der rationalen Zahlen als Unterkörper. In Γ konvergiert die Folge $\left\{\frac{1}{n}\right\}$, $n = 1, 2, 3, \ldots$, gegen die Zahl 0, ist also eine Fundamentalfolge. In K dagegen ist $\left|\frac{1}{n}\right| = \frac{1}{n} > \frac{1}{x}$ für jedes n. Daher kann in K die Folge $\left\{\frac{1}{n}\right\}$ nicht den Grenzwert 0 besitzen. Sie kann dort aber auch keinen anderen Grenzwert besitzen, da sie in K keine Fundamentalfolge ist. In der Tat: Für $p \neq q$ ist die Zahl

$$\left|\frac{1}{p} - \frac{1}{q}\right| > 0$$

und rational. Dann ist aber stets $\left|\frac{1}{p} - \frac{1}{q}\right| > \frac{1}{x}$. Man zeigt übrigens leicht, daß in K eine Folge $\{a_n\}$ von rationalen Zahlen dann und nur dann eine Fundamentalfolge ist, wenn sie stationär ist, d. h., wenn es eine rationale Zahl a und eine natürliche Zahl n_0 derart gibt, daß $a_n = a$ für alle $n > n_0$ ist. Dann ist offensichtlich auch $\lim a_n = a$. Wenn wir diese Operation des Grenzüberganges vom Körper K auf den Unterkörper Γ übertragen, so erhalten wir einen vollständigen Körper, während er im Sinne der Definition 5 unvollständig ist.

In einem Falle sind indes auch die in diesem Paragraphen eingeführten Begriffe absolut. Es gilt nämlich:

Satz 5. *Damit die Begriffe des Grenzwertes und der Fundamentalfolge in einem Körper K mit denselben Begriffen in jedem seiner Unterkörper K' übereinstimmen, ist notwendig und hinreichend, daß die Anordnung des Körpers K archimedisch ist.*[1])

Beweis. Wenn der Körper K nicht archimedisch angeordnet ist, so gibt es in K ein Element c derart, daß $n < c$ für jede natürliche Zahl n ist. Da nun aber der Körper Γ der rationalen Zahlen archimedisch angeordnet ist, ist sogar $a < c$ für jede rationale Zahl a. Dann ist $\frac{1}{c} < \frac{1}{a}$ für jede rationale Zahl $a > 0$, wie wir unmittelbar aus $a < c$ durch Multiplikation mit $\frac{1}{ac} > 0$ erhalten. Es ist also $0 < \frac{1}{c} < b$, wobei $b = \frac{1}{a}$ eine beliebige positive rationale Zahl ist. Im Körper der rationalen Zahlen konvergiert die Folge $\left\{\frac{1}{n}\right\}$, $n = 1, 2, 3, \ldots$, gegen 0, ist also dort Fundamentalfolge.

[1]) Aus dem Beweis des Satzes ergibt sich sogar, daß die archimedische Anordnung dafür notwendig ist, daß die Begriffe des Grenzwertes und der Fundamentalfolge im Körper K mit denselben Begriffen in dem in ihm enthaltenen Körper Γ der rationalen Zahlen zusammenfallen. Mit anderen Worten: Wenn jede konvergente Fundamentalfolge aus Γ auch in K konvergiert, so ist der Körper K archimedisch angeordnet. Hiervon werden wir am Anfang des nächsten Paragraphen Gebrauch machen.

Sie ist aber keine Fundamentalfolge im Körper K, besitzt also dort insbesondere auch keinen Grenzwert. In der Tat: Setzen wir $\varepsilon = \frac{1}{c} > 0$, so ist für alle $p \neq q$ die Zahl $\left|\frac{1}{p} - \frac{1}{q}\right| > \varepsilon$. Es gibt also keine Zahl n_0 mit $\left|\frac{1}{p} - \frac{1}{q}\right| < \varepsilon$ für alle p und q, die größer als n_0 sind. Damit ist die Notwendigkeit der angegebenen Bedingung bewiesen.

Sei nun K ein beliebiger archimedisch angeordneter Körper. Wir zeigen, daß die Eigenschaft einer Folge $\{a_n\}$, konvergent oder Fundamentalfolge zu sein, unabhängig davon ist, in welchem Unterkörper K' die Elemente a_n und (für den Fall der Konvergenz) der Grenzwert $a = \lim a_n$ enthalten sind. Aus dem Bestehen dieser Eigenschaften in K folgt unmittelbar ihr Bestehen in K'. Es sei nun umgekehrt $\lim a_n = a$ in K'. Wir zeigen, daß dasselbe dann auch in K der Fall ist. Dazu sei $\varepsilon > 0$ ein beliebiges Element aus K. Da K voraussetzungsgemäß archimedisch angeordnet ist, gibt es eine natürliche Zahl $n > \frac{1}{\varepsilon}$, woraus $0 < \frac{1}{n} = \varepsilon' < \varepsilon$ folgt. Die Zahl ε' liegt dann in jedem Unterkörper des Körpers K, also folglich auch in K'. Da in K' die Beziehung $\lim a_n = a$ gelten sollte, gibt es eine natürliche Zahl n_0 mit $|a_n - a| < \varepsilon' < \varepsilon$ für jedes $n > n_0$. Mithin ist auch in K die Eigenschaft $\lim a_n = a$ erfüllt, was zu beweisen war.

Definition 6. *Ein Körper heißt stetig, wenn er vollständig und archimedisch angeordnet ist.*

In einem stetigen Körper sind die am Anfang dieses Paragraphen gestellten Aufgaben, die Bestimmung des Verhältnisses zweier Strecken und der Wurzel aus einem positiven Element, stets lösbar. Auf die Aufgabe der Bestimmung der Wurzel gehen wir in § 26 näher ein. Wir wollen hier abschließend nur noch etwas über die Bestimmung des Verhältnisses zweier Strecken sagen. Wenn es uns gelingt, den Körper Γ der rationalen Zahlen zu einem stetigen Körper K zu erweitern, so sind auf Grund des zuletzt bewiesenen Satzes die oben für die Strecken AB und MN konstruierten Folgen $\{a_n\}$ und $\{b_n\}$ nicht nur in Γ, sondern auch in K Fundamentalfolgen. Da aber zudem der Körper K vollständig ist, besitzen sie dann in K einen gemeinsamen Grenzwert c [Satz 2 a)]. Dieses Element c kann man dann auf Grund seiner Konstruktion als das Verhältnis der gegebenen Strecken ansehen, d. h., man kann annehmen, daß $MN : AB = c$ oder $MN = c \cdot AB$ ist. Diese neue Definition des Streckenverhältnisses stimmt, wie wir gleichfalls oben gezeigt haben, im Falle kommensurabler Strecken mit der früheren Definition (siehe Ende des § 23) überein. Während aber die frühere Definition nur für kommensurable Strecken geeignet war, existiert das gemäß der neuen Definition bestimmte Element des Körpers K für beliebige Strecken, unabhängig davon, ob sie kommensurabel sind oder nicht. In diesem Sinne sprechen wir davon, daß in einem stetigen Körper das Verhältnis zweier Strecken stets bestimmbar ist. Die Behandlung dieser Aufgabe sollte nur die Wichtigkeit des Begriffs des stetigen Körpers illustrieren. Wir können hier nicht daran denken, auf diese geometrische Aufgabe ausführlicher einzugehen.

Wir wollen noch ohne Beweis vermerken, daß das oben definierte Streckenverhältnis alle notwendigen Eigenschaften besitzt. Für beliebige Strecken AB und CD und beliebige Elemente $c > 0$ und $d > 0$ eines stetigen Körpers K gilt nämlich:

a) Wenn $c < d$ ist, so ist $c \cdot AB < d \cdot AB$;

b) $(c + d)\, AB = c \cdot AB + d \cdot AB$;

c) $c(AB + CD) = c \cdot AB + c \cdot CD$.

Ferner gibt es zu jeder Strecke AB und zu jedem Element $c > 0$ aus K eine Strecke MN mit $MN : AB = c$.

Die Aufgabe, den Umfang des Kreises zu bestimmen, läßt sich auf die Bestimmung der Länge einer Strecke zurückführen. Dazu konstruieren wir zu einem vorgegebenen Kreis zwei Folgen regelmäßiger Vielecke (nämlich der umbeschriebenen und der einbeschriebenen), indem wir die Seitenanzahl laufend verdoppeln. Wenn wir in K die Länge einer Strecke (in Form des Verhältnisses dieser Strecke zu einer vorgegebenen Einheitsstrecke — *Anm. d. wissenschaftl. Red.*) bestimmen können, so können wir auch den Umfang a_n des einbeschriebenen und den Umfang b_n des umbeschriebenen regelmäßigen 2^{n+1}-Ecks bestimmen. Auf Grund einfacher elementargeometrischer Überlegungen kann man zeigen, daß $a_1 < a_2 < \cdots$, daß $b_1 > b_2 > \cdots$ und daß ferner $a_n < b_n$ und $\lim (b_n - a_n) = 0$ ist. Hieraus folgt unmittelbar, daß die Folgen $\{a_n\}$ und $\{b_n\}$ von Elementen des Körpers K Fundamentalfolgen sind. Auf Grund der Vollständigkeit von K besitzen sie in K einen gemeinsamen Grenzwert c, der definitionsgemäß den Umfang des Kreises festlegt. Analog zeigt man, daß man in K auch die Länge eines Kreisbogens bestimmen kann. Umgekehrt kann man beweisen, daß es zu jedem Element zwischen 0 und dem Umfang des Kreises, also zu jedem Element c' aus K mit $0 < c' < c$ einen Kreisbogen der Peripherie des gegebenen Kreises gibt, der die Länge c' besitzt. In diesem Sinne ist in einem stetigen Körper auch die Aufgabe der Rektifikation eines Kreisbogens lösbar.

Im folgenden Paragraphen werden wir sehen, daß speziell der Körper der reellen Zahlen ein stetiger Körper ist.

§ 25. Definition des Körpers der reellen Zahlen

Im Körper Γ der rationalen Zahlen ist — wie wir gesehen haben — die Operation des Grenzüberganges nicht unbeschränkt ausführbar, d. h., der Körper Γ ist nicht vollständig (§ 24, Satz 4). Wenn wir den allgemeinen Prinzipien der Erweiterung eines Zahlenbereiches, die wir in § 18 aufgestellt haben, folgen wollen, so müssen wir den Körper Γ zu einem Körper D erweitern, in dem eine Anordnung definiert ist und in dem jede Fundamentalfolge einen Grenzwert besitzt. Ferner sollen die Fundamentalfolgen aus Γ, für die in Γ bereits die Operation des Grenzüberganges ausführbar ist, in D denselben Grenzwert besitzen wie in Γ. Damit müssen die Fundamentalfolgen aus Γ auch Fundamentalfolgen in D sein. Mithin muß D ein vollständiger und archimedisch angeordneter Körper, also ein stetiger Körper sein (§ 24, Satz 5). Genau wie im Falle der ganzen (§ 20) und der rationalen (§ 22) Zahlen werden wir auch hier den kleinsten Körper mit dieser Eigenschaft suchen. Dabei wird sich jedoch herausstellen, daß diese Minimalbedingung von selbst erfüllt ist, da die Forderung der Stetigkeit den Körper D bereits eindeutig bis auf Isomorphie festlegt. Deshalb erübrigt es sich, die Minimalbedingung in die Definition von D mit aufzunehmen. Wir definieren also:

Definition 1. *Unter einem Körper der reellen Zahlen verstehen wir einen stetigen Körper D, der den Körper Γ der rationalen Zahlen als Unterkörper enthält. Die Elemente von D heißen reelle Zahlen.*

Der Beweis für die Existenz und die Eindeutigkeit eines Körpers D mit diesen Eigenschaften verläuft analog wie im Fall des Ringes der ganzen Zahlen (§ 20) und des Körpers der rationalen Zahlen (§ 22). Wir beginnen mit dem Beweis der Eindeutigkeit.

Satz 1. *Ein angeordneter Körper K, der den Körper Γ der rationalen Zahlen enthält*[1]*, ist dann und nur dann archimedisch angeordnet, wenn jedes Element des Körpers K Grenzwert einer Folge von rationalen Zahlen ist.*

Beweis. a) Sei das Element a des Körpers K Grenzwert der Folge $\{a_n\}$ von rationalen Zahlen. Dann gibt es eine natürliche Zahl k derart, daß $|a_k - a| < 1$ ist, woraus sich

$$a \leq |a| = |(a - a_k) + a_k| \leq |a - a_k| + |a_k| < 1 + |a_k|$$

ergibt. Da $1 + |a_k|$ eine rationale Zahl ist und da der Körper der rationalen Zahlen archimedisch angeordnet ist, gibt es eine natürliche Zahl n mit $1 + |a_k| < n$. Dann ist aber auch $a < n$ und mithin K archimedisch angeordnet (§ 10, XI'.).

b) Sei nun umgekehrt der Körper K archimedisch angeordnet. Dann gibt es zu jedem Element a aus K und zu jeder natürlichen Zahl n natürliche Zahlen m_1 und m_2 mit

$$m_1 \cdot \frac{1}{n} > a \text{ und } m_2 \cdot \frac{1}{n} > -a,$$

also $(-m_2) \cdot \frac{1}{n} < a$. Folglich ist die Menge A der ganzen Zahlen l mit $l \cdot \frac{1}{n} \leq a$ durch die Zahl m_1 nach oben beschränkt und nicht leer, da sie die Zahl $-m_2$ enthält. Daher gibt es in A eine größte Zahl m (§ 21, Satz 5). Offensichtlich ist dann $\frac{m}{n} \leq a < \frac{m+1}{n}$. Indem wir von dieser Ungleichung die Zahl $\frac{m}{n}$ subtrahieren, erhalten wir $0 \leq a - \frac{m}{n} \leq \frac{1}{n}$. Wir setzen nun $\frac{m}{n} = a_n$ und zeigen, daß $\lim a_n = a$ ist. Zunächst gibt es zu jedem $\varepsilon > 0$ aus K eine natürliche Zahl n_0 mit $n_0 > \frac{1}{\varepsilon}$. Dann gilt aber für alle $n > n_0$:

$$|a_n - a| = a - a_n < \frac{1}{n} < \frac{1}{n_0} < \varepsilon,$$

d. h., es ist tatsächlich in K die Relation $\lim a_n = a$ erfüllt.

Satz 2. *Alle Körper der reellen Zahlen sind untereinander isomorph, d. h., der Begriff des Körpers der reellen Zahlen ist durch Definition* 1 *eindeutig bis auf Isomorphie festgelegt. Darüber hinaus gibt es zwischen zwei Körpern D_1 und D_2 der reellen Zahlen nur eine isomorphe Abbildung, die die Anordnung erhält. Bei diesem speziellen Isomorphismus bleiben die rationalen Zahlen elementweise fest. Insbesondere gibt es nur eine isomorphe Abbildung eines Körpers der reellen Zahlen auf sich, die die Anordnung erhält, nämlich die identische Abbildung.* (Auf Grund von Satz 2 aus § 23 gilt dieser Satz auch für beliebige stetige Körper, sofern man die rationalen Zahlen durch die rationalen Elemente ersetzt.)[2]

---

[1] Die Bedingung $K \supseteq \Gamma$ kann man hier noch dahingehend abschwächen, daß man verlangt, der Körper K enthalte die rationalen Elemente eines anderen Körpers.

[2] Wir werden in § 26 zeigen, daß die Bedingung, daß die Anordnung erhalten bleibt, entbehrlich ist, da der Körper der reellen Zahlen überhaupt nur eine Anordnung gestattet.

Beweis. Wir konstruieren auf folgende Weise eine Abbildung des Körpers D_1 in den Körpern D_2: Es sei d_1 ein beliebiges Element aus dem Körper D_1. Da D_1 archimedisch angeordnet ist, gibt es nach Satz 1 eine Folge $\{a_n\}$ von rationalen Zahlen a_n mit $d_1 = \lim a_n$. Dann ist $\{a_n\}$ eine Fundamentalfolge in D_1 und damit auch in seinem Unterkörper Γ. Weil $\Gamma \subset D_2$ und D_2 archimedisch angeordnet ist, ist die Folge $\{a_n\}$ als Fundamentalfolge in Γ auch Fundamentalfolge in D_2 (§ 24, Satz 5). Da D_2 vollständig ist, gibt es in D_2 ein Element d_2 mit $\lim a_n = d_2$. Wir setzen dann $f(d_1) = d_2$.

Als erstes zeigen wir, daß die Definition von d_2 unabhängig ist von der Wahl der Folge $\{a_n\}$. Wenn nämlich $\{b_n\}$ eine beliebige Folge von rationalen b_n ist, für die $\lim b_n = d_1$ ist, so ist $\lim a_n = \lim b_n$, also $\lim (a_n - b_n) = 0$ [§ 24, Satz 2a)] in D_1 und damit auch in Γ. Genau wie oben schließen wir, daß dann $\lim (a_n - b_n) = 0$ und mithin $\lim a_n = \lim b_n = d_2$ in D_2 ist. Ist speziell d_1 eine rationale Zahl, so ist unter anderem $\lim a_n = d_1$ für die Folge $\{a_n\}$ mit $a_n = d_1$ für alle n. Also ist $f(d_1) = d_1$, d. h., die Abbildung f läßt die rationalen Zahlen elementweise fest.

Wenn $c_1 \neq d_1$ und $c_1 = \lim a_n$, $d_1 = \lim b_n$ ist, so ist $\lim (a_n - b_n) \neq 0$, also in D_2 $\lim a_n \neq \lim b_n$ und mithin $f(c_1) \neq f(d_1)$. Also ist f eine eineindeutige Abbildung von D_1 in D_2. Sie ist jedoch im allgemeinen noch abhängig von der Definition des Grenzwertes in D_1 bzw. in D_2, also insbesondere von den Anordnungen dieser Körper.

Wir zeigen als nächstes, daß f eine isomorphe Abbildung von D_1 in D_2 ist. Dazu ist zu zeigen, daß für beliebige Elemente c_1 und d_1 aus D_1 die Beziehungen

$$f(c_1 + d_1) = f(c_1) + f(d_1), \quad f(c_1 d_1) = f(c_1) f(d_1)$$

gelten. Dies folgt aber unmittelbar aus Satz 2* b) bzw. c) in § 24. Wenn nämlich $c_1 = \lim a_n$ und $d_1 = \lim b_n$ ist, so erhalten wir, indem wir die Definition der Abbildung f benutzen:

$$\begin{aligned} f(c_1 + d_1) &= f(\lim a_n + \lim b_n) = f[(\lim (a_n + b_n)] \\ &= \lim f(a_n + b_n) = \lim [f(a_n) + f(b_n)] \\ &= \lim f(a_n) + \lim f(b_n) = f(\lim a_n) + f(\lim b_n) \\ &= f(c_1) + f(d_1) \end{aligned}$$

und entsprechend die zweite Gleichung.

Wir zeigen weiter, daß die so konstruierte Abbildung die Anordnung erhält. Dazu sei $c_1 < d_1$, $c_1 = \lim a_n$, $d_1 = \lim b_n$ im Körper D_1. Dann gibt es ein n_0 derart, daß $a_n < b_n$ für alle $n > n_0$ [§ 24, Satz 2 e)] und daher $\lim a_n \leq \lim b_n$ in D_2, d. h. also $f(c_1) \leq f(d_1)$ ist. Da außerdem $c_1 \neq d_1$ ist, ist $f(c_1) \neq f(d_1)$ und damit $f(c_1) < f(d_1)$.

Als nächstes zeigen wir, daß f die einzige isomorphe Abbildung von D_1 in D_2 ist, die die Anordnung erhält. Ist nämlich g eine beliebige derartige Abbildung, so bildet g den in D_1 enthaltenen Körper der rationalen Zahlen auf die rationalen Elemente des Körpers D_2 ab, wobei die rationale Zahl r aus D_1 in das Element re übergeht, wenn e Einselement von D_2 ist (§ 23, Satz 2). Nun enthält D_2 voraussetzungsgemäß den Körper Γ, also ist $e = 1$ und $re = r \cdot 1 = r$, d. h., es ist $g(r) = r$ für jede rationale Zahl r. Wäre nun die Abbildung g von der Abbildung f verschieden, so gäbe es ein Element d_1 in D_1, für das $a_2 = f(d_1) \neq g(d_1) = b_2$ ist. Wir konstruieren eine rationale Zahl c, die zwischen a_2 und b_2 liegt. Dazu sei ohne Beschränkung der Allgemeinheit $a_2 < b_2$. Indem wir genau wie unter b) im Beweis des Satzes 1 vorgehen, finden wir zunächst eine natürliche Zahl n mit $\frac{1}{n} < b_2 - a_2$ und danach eine

ganze Zahl m, für die $\frac{m}{n} \leq a_2 < \frac{m+1}{n}$ ist. Setzen wir schließlich $c = \frac{m+1}{n}$, so ist

$$a_2 < c = \frac{m}{n} + \frac{1}{n} < a_2 + (b_2 - a_2) = b_2.$$

Da $c = f(c)$ ist und f nach dem oben Bewiesenen die Anordnung erhält, folgt aus $f(d_1) = a_2 < c$, daß $d_1 < c$ ist. Da andererseits auch $g(c) = c$ ist und g gleichfalls die Anordnung erhält, ist dann

$$g(d_1) = b_2 < g(c) = c,$$

was der Konstruktion der Zahl c widerspricht.

Bisher haben wir an noch keiner Stelle von der Vollständigkeit des Körpers D_1 Gebrauch gemacht. Deshalb gilt das bisher Bewiesene auch in jedem beliebigen archimedisch angeordneten Körper. Es bleibt lediglich noch zu beweisen, daß die oben konstruierte Abbildung f eine isomorphe Abbildung des Körpers D_1 auf den Körper D_2 ist. Hierbei werden wir wesentlich die Vollständigkeit von D_1 benutzen. Wir müssen zu jedem Element d_2 aus D_2 ein Element d_1 aus D_1 finden, für das $f(d_1) = d_2$ ist. Da D_2 archimedisch angeordnet ist, gibt es auf Grund von Satz 1 in D_2 eine Folge $\{a_n\}$ von rationalen a_n, für die $d_2 = \lim a_n$ ist. Die Fundamentalfolge $\{a_n\}$ aus D_2 ist wegen $\Gamma \subset D_2$ Fundamentalfolge in Γ und damit auch in D_1. Wegen der Vollständigkeit von D_1 gibt es in D_1 ein Element d_1 mit $d_1 = \lim a_n$. Nach Definition von f ist aber dann $f(d_1) = d_2$. Damit ist der behauptete Satz in allen Teilen bewiesen.

Satz 3. *Jeder archimedisch angeordnete Körper K ist isomorph einem Unterkörper eines Körpers D der reellen Zahlen. Es gibt nur eine isomorphe Abbildung von K in D, die die Anordnung erhält. Insbesondere kann der Körper K nur auf eine Weise, nämlich identisch, isomorph unter Erhaltung der Ordnung auf sich abgebildet werden.*[1])

Beweis. Zunächst kann der Körper K isomorph unter Erhaltung der Ordnung auf einen angeordneten Körper Q abgebildet werden, der den Körper der rationalen Zahlen enthält. Da K archimedisch angeordnet ist, ist dann auch Q archimedisch angeordnet. Für Q ergibt sich der Satz aber nebenbei aus dem Beweis von Satz 2, wenn wir dort D_1 durch Q ersetzen und beachten, daß wir die Vollständigkeit von D_1 nur im letzten Teil des Beweises benutzt haben. Dieser letzte Teil fällt aber in dem jetzt betrachteten Satz weg. Auf Grund der Isomorphie von K und Q ist Satz 3 dann auch für den Körper K richtig.

Wenn es also überhaupt einen Körper D der reellen Zahlen gibt, so gibt es (von isomorphen Abbildungen abgesehen) auch nur einen derartigen Körper.

Als nächstes wenden wir uns diesem Existenzbeweis zu. Wie auch im Falle der ganzen und der rationalen Zahlen genügt es, einen Körper (ein Modell) zu konstruieren, der der Definition 1 genügt.

Es gibt mehrere Konstruktionsmethoden für einen derartigen Körper. Wir bringen hier eine Methode, die auf G. CANTOR zurückgeht.

---

[1]) Im Unterschied zu Satz 2 darf hier die Bedingung, daß die Abbildung die Ordnung erhält, nicht weggelassen werden. In der Tat: Ist K der Körper aller Zahlen der Form $a + b\sqrt{2}$ mit rationalem a und b, so ist die Abbildung $a + b\sqrt{2} \leftrightarrow a - b\sqrt{2}$ ein Isomorphismus bezüglich der Addition und der Multiplikation, der nicht die Identität ist. Dies ist möglich, da die Abbildung nicht die in K als Unterkörper des Körpers der reellen Zahlen gegebene Anordnung erhält, weil z. B. $1 + \sqrt{2} > 0$, aber $1 - \sqrt{2} < 0$ ist.

Die Konstruktion eines der untereinander isomorphen Körper der reellen Zahlen wird durch Satz 1 nahegelegt. Wenn nämlich D ein derartiger Körper ist, so muß jedes Element aus D Grenzwert einer Fundamentalfolge von rationalen Zahlen sein. Andererseits muß auch jede derartige Fundamentalfolge wegen der Stetigkeit von D in D einen Grenzwert besitzen.

Als Ausgangselemente zum Aufbau des Körpers D der reellen Zahlen wählen wir die Fundamentalfolgen $a_1, a_2, a_3, \ldots = \{a_n\}$ aus rationalen Zahlen, d. h. die Folgen, für die es zu jedem rationalen $\varepsilon > 0$ eine natürliche Zahl n_0 derart gibt, daß $|a_p - a_q| < \varepsilon$ ist für alle p und q, die größer als n_0 sind (§ 24, Definition 4). Es sei M die Menge aller dieser Folgen. Für die Elemente aus M definieren wir eine Äquivalenzrelation, eine Addition und eine Multiplikation, die uns dann die Gleichheit, die Addition und die Multiplikation derjenigen Elemente des Körpers D liefern werden, die gleich den Grenzwerten dieser Folgen sind [§ 24, Satz 2 a), b), c)]. Es soll nämlich

$$\{a_n\} \sim \{b_n\} \tag{1}$$

dann und nur dann gelten, wenn

$$\lim (a_n - b_n) = 0$$

ist. Ferner soll

$$\{a_n\} + \{b_n\} = \{a_n + b_n\} \tag{2}$$

und

$$\{a_n\} \cdot \{b_n\} = \{a_n b_n\} \tag{3}$$

sein.

Zunächst wollen wir uns davon überzeugen, daß die Gleichungen (2) und (3) tatsächlich Operationen in der Menge M definieren, d. h., daß die Folgen auf der rechten Seite wiederum Fundamentalfolgen sind. Für den Fall der Addition wählen wir eine beliebige rationale Zahl $\varepsilon > 0$. Da die Folgen $\{a_n\}$ und $\{b_n\}$ Fundamentalfolgen sind, gibt es natürliche Zahlen n_1 und n_2 derart, daß $|a_p - a_q| < \frac{\varepsilon}{2}$ für alle $p, q > n_1$ und $|b_p - b_q| < \frac{\varepsilon}{2}$ für alle $p, q > n_2$. Ist n_0 die größere der Zahlen n_1 und n_2, so gilt für alle $p, q > n_0$

$$|(a_p + b_p) - (a_q + b_q)| \leq |a_p - a_q| + |b_p - b_q| < \varepsilon,$$

d. h., die Folge $\{a_n + b_n\}$ ist eine Fundamentalfolge.

Im Fall der Multiplikation zeigen wir zunächst, daß jede Fundamentalfolge $\{c_n\}$ beschränkt (§ 24, Definition 2) ist. In der Tat: Es gibt eine natürliche Zahl n_0 derart, daß $|c_p - c_q| < 1$ ist für alle $p, q > n_0$. Dann ist aber für jedes $n > n_0$

$$|c_n| = |(c_n - c_{n_0+1}) + c_{n_0+1}| \leq |c_n - c_{n_0+1}| + |c_{n_0+1}| < 1 + |c_{n_0+1}|.$$

Indem wir die rationale Zahl c so bestimmen, daß c größer ist als alle Zahlen $|c_1|, |c_2|, \ldots, |c_{n_0}|, |c_{n_0+1}| + 1$ (also z. B. die Summe aller dieser Zahlen plus 1 nehmen), finden wir, daß $|c_n| < c$ ist für alle n.

Es gibt also rationale Zahlen a und b derart, daß $|a_n| < a$ und $|b_n| < b$ ist für alle n. Ferner gibt es zu jeder rationalen Zahl $\varepsilon > 0$ natürliche Zahlen n_1 und n_2 mit $|a_p - a_q| < \frac{\varepsilon}{2b}$ für alle $p, q > n_1$ und $|b_p - b_q| < \frac{\varepsilon}{2a}$ für alle $p, q > n_2$. Setzen wir

wieder n_0 gleich der größeren der Zahlen n_1 und n_2, so gilt für alle $p, q > n_0$

$$|a_p b_p - a_q b_q| = |(a_p b_p - a_p b_q) + (a_p b_q - a_q b_q)|$$
$$\leq |a_p| \cdot |b_p - b_q| + |b_q| \cdot |a_p - a_q| < a \cdot \frac{\varepsilon}{2a} + b \cdot \frac{\varepsilon}{2b} = \varepsilon,$$

d. h., auch $\{a_n b_n\}$ ist eine Fundamentalfolge.

Wir nennen eine Fundamentalfolge *positiv*, wenn es eine rationale Zahl $\varepsilon > 0$ und eine natürliche Zahl n_0 gibt mit $a_n > \varepsilon$ für alle $n > n_0$.

Die Äquivalenzrelation (1) zwischen den Folgen besitzt offensichtlich die grundlegenden Eigenschaften der Gleichheit (§ 19), nämlich

1. $\{a_n\} \sim \{a_n\}$, da $\lim (a_n - a_n) = 0$ ist.

2. Wenn $\{a_n\} \sim \{b_n\}$ ist, so ist auch $\{b_n\} \sim \{a_n\}$; denn wenn $\lim (a_n - b_n) = 0$ ist, so ist auch $\lim (b_n - a_n) = 0$, weil $|a_n - b_n| = |b_n - a_n|$ ist.

3. Ist $\{a_n\} \sim \{b_n\}$ und $\{b_n\} \sim \{c_n\}$, so ist $\{a_n\} \sim \{c_n\}$; wenn nämlich $\lim (a_n - b_n) = 0$ und $\lim (b_n - c_n) = 0$ gilt, so gilt auch

$$\lim (a_n - c_n) = \lim [(a_n - b_n) + (b_n - c_n)]$$
$$= \lim (a_n - b_n) + \lim (b_n - c_n) = 0 + 0 = 0$$

[§ 24, Satz 2 b)].

Auf Grund des Hauptsatzes aus § 19 bestimmt also die Relation (1) eine Einteilung der Menge M in Klassen äquivalenter Folgen. Derartige Klassen werden wir im folgenden durch kleine griechische Buchstaben $\alpha, \beta, \gamma, \delta, \ldots$ bezeichnen.

Definition 2. *Es sei D_0 die Menge aller Klassen äquivalenter Folgen aus M. Unter der Summe (dem Produkt) der Klassen α und β verstehen wir diejenige Klasse $\alpha + \beta$ (bzw. $\alpha\beta$), die die Summe (bzw. das Produkt) aus einer Folge der Klasse α und einer Folge der Klasse β enthält. Eine Klasse heißt positiv, wenn eine Folge aus dieser Klasse positiv ist.*

Zunächst zeigen wir, daß die Summe, das Produkt und die Eigenschaft, positiv zu sein, unabhängig von der Wahl der Repräsentanten aus den jeweiligen Klassen sind. Es sei also etwa

$$\{a_n\} \sim \{b_n\} \text{ und } \{c_n\} \sim \{d_n\},$$

d. h.

$$\lim (a_n - b_n) = 0 \text{ und } \lim (c_n - d_n) = 0.$$

Dann ist auch

$$\lim [(a_n + c_n) - (b_n + d_n)] = \lim (a_n - b_n) + \lim (c_n - d_n) = 0,$$

d. h. $\{a_n + c_n\} \sim \{b_n + d_n\}$.

Da die Folge $\{c_n\}$ eine Fundamentalfolge ist, ist sie beschränkt, d. h., es gibt eine rationale Zahl $c > 0$ mit $|c_n| < c$ für alle n. Ist nun eine beliebige rationale Zahl $\varepsilon > 0$ vorgegeben, so gibt es eine natürliche Zahl n_0 mit $|a_n - b_n| < \frac{\varepsilon}{c}$ für alle $n > n_0$. Dann ist aber für alle $n > n_0$

$$|a_n c_n - b_n c_n| = |a_n - b_n| \cdot |c_n| < \frac{\varepsilon}{c} \cdot c = \varepsilon.$$

Folglich ist $\lim (a_n c_n - b_n c_n) = 0$, d. h.

$$\{a_n c_n\} \sim \{b_n c_n\}.$$

Indem wir das kommutative Gesetz und noch einmal das soeben Bewiesene benutzen, erhalten wir daraus

$$\{a_n c_n\} \sim \{b_n c_n\} = \{c_n b_n\} \sim \{d_n b_n\} = \{b_n d_n\}.$$

Wenn schließlich die Folge $\{a_n\}$ positiv und ferner $\{a_n\} \sim \{b_n\}$ ist, so gibt es ein rationales $\varepsilon > 0$ und eine natürliche Zahl n_1 mit $a_n > \varepsilon$ für alle $n > n_1$. Ferner gibt es zu dem gegebenen ε eine natürliche Zahl n_2 mit $|a_n - b_n| < \frac{\varepsilon}{2}$ für alle $n > n_2$. Ist wieder n_0 die größere der Zahlen n_1 und n_2, so ergibt sich, daß für alle $n > n_0$

$$b_n = a_n - (a_n - b_n) \geq a_n - |a_n - b_n| > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}$$

ist, d. h., auch die Folge $\{b_n\}$ ist positiv.

Somit führen die in Definition 2 zunächst repräsentantenweise erklärten Operationen der Addition und Multiplikation und die mit Hilfe von Folgen definierte Anordnung zu Beziehungen in der Menge D_0 der Klassen von äquivalenten Folgen.

Satz 4. *Die Menge D_0 bildet bezüglich der in Definition 2 erklärten algebraischen Operationen der Addition und der Multiplikation und bezüglich der dort angegebenen Anordnung einen stetigen Körper* (§ 24, Definition 6).

Beweis. Es ist zu zeigen, daß in D_0 die Bedingungen I. bis XII. (§ 7, Definition 1; § 8, Definition 1; § 10, Definitionen 1 und 3; § 24, Definition 5) erfüllt sind. Da die Operationen (2) und (3) zwischen den Folgen durch die entsprechenden Operationen zwischen ihren Gliedern erklärt sind, folgen aus den Ringeigenschaften I. bis VI. für die rationalen Zahlen die entsprechenden Eigenschaften für die Menge M und damit auch für die Menge D_0. Somit sind jedenfalls M und D_0 Ringe.

Wir wollen uns nun zunächst überlegen, was die Null und das Entgegengesetzte eines Elementes aus D_0 sind. Offensichtlich ist das Nullelement von D_0 die Klasse, die die Fundamentalfolge $\{0\} = 0, 0, 0, \ldots$ enthält; wir wollen sie im folgenden durch (0) bezeichnen. Diese Klasse besteht aus allen den Folgen $\{a_n\}$, die der Folge $\{0\}$ äquivalent sind, also aus allen Folgen, für die $\lim a_n = 0$ ist. Derartige Folgen nennt man *Nullfolgen*. Also ist einerseits jede Folge aus der Klasse (0) Nullfolge. Andererseits ist aber auch jede Nullfolge als konvergente Fundamentalfolge mit der Folge $\{0\}$ äquivalent und gehört mithin zur Klasse (0).

Wenn die Klasse α die Folge $\{a_n\}$ enthält, so enthält die zu α entgegengesetzte Klasse $-\alpha$ sicher die zur Folge $\{a_n\}$ entgegengesetzte Folge $\{-a_n\}$. Darüber hinaus enthält sie dann auch alle zur Folge $\{-a_n\}$ äquivalenten Folgen. Wegen $a_n - b_n = -[(-a_n) - (-b_n)]$ ist nun $\{a_n\} \sim \{b_n\}$ genau dann, wenn $\{-a_n\} \sim \{-b_n\}$ ist, d. h., die Klasse $-\alpha$ besteht aus genau den Folgen, die zu einer Folge der Klasse α entgegengesetzt sind.

Die Bedingung VII. folgt nun nicht, wie oben die Bedingungen I. bis VI., aus der entsprechenden Bedingung für die rationalen Zahlen. Wenn nämlich nicht alle Glieder der Folge $\{a_n\}$ aus M gleich Null sind, so ist die Folge $\{a_n\}$ von der Folge $\{0\}$ verschieden, die wir oben als Nullelement des Ringes M erkannt haben. Jedoch besitzt die Gleichung $\{a_n\}\{x_n\} = \{b_n\}$ sicher schon dann keine Lösung, wenn nur $a_1 = 0$ und $b_1 \neq 0$ ist. Folglich kann der Ring M kein Körper sein. Trotzdem können wir zeigen, daß D_0 ein Körper ist. Seien dazu $\alpha \neq (0)$ und β beliebige Klassen aus D_0. Ferner sei $\{a_n\}$ ein beliebiger Repräsentant aus α und $\{b_n\}$ ein beliebiger Repräsentant aus β. Dann gibt es eine rationale Zahl $a > 0$ und eine natürliche Zahl n_1 mit $|a_n| > a$

für alle $n > n_1$. Andernfalls gäbe es nämlich zu jeder rationalen Zahl $\varepsilon > 0$ eine natürliche Zahl p derart, daß $|a_n - a_q| < \frac{\varepsilon}{2}$ ist für alle $n, q > p$, und wir könnten ein $q > p$ finden, für das $|a_q| < \frac{\varepsilon}{2}$ ist. Dann wäre aber für alle $n > p$

$$|a_n| = |(a_n - a_q) + a_q| \leq |a_n - a_q| + |a_q| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

und daher $\lim a_n = 0$, im Widerspruch dazu, daß $\{a_n\}$ der Klasse $\alpha \neq (0)$ angehören sollte.

Ohne Beschränkung der Allgemeinheit können wir sogar annehmen, daß $a_n \neq 0$ für alle n ist. In der Tat können ja wegen $|a_n| > a > 0$ für alle $n > n_1$ nur endlich viele Glieder a_n (mit $n \leq n_1$) gleich Null sein. Wenn wir diese endlich vielen Glieder durch beliebige von Null verschiedene rationale Zahlen ersetzen, so erhalten wir eine offensichtlich zu $\{a_n\}$ äquivalente Folge, die also zur Klasse α gehört und deren sämtliche Glieder von Null verschieden sind.

Wir zeigen nun, daß die Folge $\{c_n\} = \left\{\frac{b_n}{a_n}\right\}$ eine Fundamentalfolge ist. Zunächst ist die Folge $\{b_n\}$ als Fundamentalfolge beschränkt, d. h., es gibt eine rationale Zahl b mit $|b_n| < b$ für alle n. Ist nun eine beliebige rationale Zahl $\varepsilon > 0$ vorgegeben, so gibt es, da $\{a_n\}$ und $\{b_n\}$ Fundamentalfolgen sind, natürliche Zahlen n_2 und n_3 mit $|a_p - a_q| < \frac{a^2 \varepsilon}{2b}$ für alle $p, q > n_2$ und $|b_p - b_q| < \frac{a\varepsilon}{2}$ für alle $p, q > n_3$. Ist n_0 die größte unter den Zahlen n_1, n_2, n_3, so gilt für jedes $p, q > n_0$

$$\left|\frac{b_p}{a_p} - \frac{b_q}{a_q}\right| = \left|\frac{b_p a_q - a_p b_q}{a_p a_q}\right| = \left|\frac{b_p a_q - a_p b_p}{a_p a_q} + \frac{a_p b_p - a_p b_q}{a_p a_q}\right|$$

$$\leq \frac{|b_p| \cdot |a_q - a_p|}{|a_p| \cdot |a_q|} + \frac{|b_p - b_q|}{|a_q|} < \frac{b \cdot \frac{a^2 \varepsilon}{2b}}{a^2} + \frac{\frac{a\varepsilon}{2}}{a} = \varepsilon,$$

also ist in der Tat $\{c_n\} = \left\{\frac{b_n}{a_n}\right\}$ eine Fundamentalfolge.

Ist γ die $\{c_n\}$ enthaltende Klasse, so ist wegen $\{a_n\}\{c_n\} = \{b_n\}$ auch $\alpha\gamma = \beta$. Damit ist die Eigenschaft VII. bewiesen.

Die Eigenschaft VIII. ist trivialerweise erfüllt, weil D_0 offensichtlich mehr als zwei Elemente enthält.

Wir beweisen als nächstes, daß in D_0 auch die Bedingung IX. erfüllt ist. Dazu ist zu zeigen, daß für jede Klasse α einer und nur einer der folgenden drei Fälle eintritt: α ist positiv, $-\alpha$ ist positiv, $\alpha = (0)$. Wir nehmen an, es sei weder α noch $-\alpha$ positiv und $\{a_n\}$ eine beliebige Folge aus der Klasse α. Ferner sei ε eine beliebige rationale Zahl > 0. Da $\{a_n\}$ eine Fundamentalfolge ist, gibt es eine natürliche Zahl n_0 derart, daß $|a_p - a_q| < \frac{\varepsilon}{2}$ für alle $p, q > n_0$. Da α nicht positiv ist, gibt es ein $r > n_0$ mit $a_r \leq \frac{\varepsilon}{2}$. Da $-\alpha$ nicht positiv ist, gibt es ein $s > n_0$ mit $-a_s \leq \frac{\varepsilon}{2}$. Somit gilt für jedes $n > n_0$ einerseits

$$a_n = a_r + (a_n - a_r) \leq a_r + |a_n - a_r| < \varepsilon$$

und andererseits

$$-a_n = (a_s - a_n) - a_s \leq |a_s - a_n| + (-a_s) < \varepsilon.$$

Daher ist $|a_n| < \varepsilon$ für alle $n > n_0$, also $\lim a_n = 0$ und damit $\alpha = (0)$.

Es tritt also stets mindestens einer der oben genannten Fälle ein. Wenn nun die Klasse α positiv ist, so gibt es eine rationale Zahl $a > 0$ und ein n_0 derart, daß $a_n > a$ und mithin $-a_n < -a$ ist für alle $n > n_0$. Damit ist sowohl ausgeschlossen, daß $\lim a_n = 0$, also $\alpha = (0)$ ist, als auch, daß $-\alpha$ positiv ist. Entsprechend zeigt man, daß die beiden anderen Fälle ausgeschlossen sind, wenn $-\alpha$ positiv ist. Damit ist dann gezeigt, daß sich alle drei Fälle gegenseitig ausschließen, d. h. die Bedingung IX. erfüllt ist.

Die Bedingung X. ist erfüllt, da die Summe und das Produkt positiver Folgen offensichtlich positiv sind.

Damit ist gezeigt, daß D_0 ein angeordneter Körper ist. Durch die Festsetzung, daß $\alpha > \beta$ dann und nur dann gelten soll, wenn $\alpha - \beta$ eine positive Klasse ist, führen wir in D_0 eine Ordnung ein, bezüglich der die positiven Elemente und nur diese größer als Null sind (§ 10, Satz 1). Man sieht weiter unmittelbar, daß Einselement des Körpers D_0 die Klasse ist, die die Folge $\{1\} = 1, 1, 1, \ldots$ und alle mit ihr äquivalenten Folgen, d. h. alle Folgen $\{a_n\}$ mit $\lim a_n = 1$ enthält; diese Klasse wollen wir mit (1) bezeichnen.

Wir zeigen als nächstes, daß in D_0 das Archimedische Axiom XI. gilt. Dazu sei α die Klasse, die die Folge $\{a_n\}$ enthält. Oben haben wir gezeigt, daß jede Fundamentalfolge beschränkt ist. Daher gibt es eine rationale Zahl a mit $|a_n| < a$ und folglich $a - a_n > 0$ für alle n. Da nun im Körper der rationalen Zahlen das Archimedische Axiom gilt, gibt es eine natürliche Zahl $k > a + 1$. Dann ist aber $k - a_n > 1$ für alle n und folglich die Klasse $k \cdot (1) - \alpha$ positiv, d. h. $k \cdot (1) > \alpha$. Also ist im Körper D_0 auch die Eigenschaft XI. erfüllt.

Schließlich zeigen wir noch, daß in D_0 auch das Vollständigkeitsaxiom XII. (§ 24, Definition 5) erfüllt ist. Dazu eine Vorbemerkung: Wenn die Klasse α eine Folge $\{a_n\}$ enthält, für die $a_n \geq 0$ ist für alle n, die größer als ein gewisses n_0 sind, so ist $\alpha \geq (0)$, da die Annahme, es wäre $\alpha < (0)$, offensichtlich zu einem Widerspruch führt. Daher ist $\alpha \geq \beta$, wenn α eine Folge $\{a_n\}$ und β eine Folge $\{b_n\}$ enthält, für die $a_n \geq b_n$ ist für alle n, die größer als ein gewisses n_0 sind. Entsprechend wie wir oben die Klasse, die die Folge $\{0\}$ bzw. $\{1\}$ enthält, mit (0) bzw. (1) bezeichnet haben, bezeichnen wir für ein vorgegebenes rationales a mit (a) die Klasse, die die Folge $\{a\} = a, a, a, \ldots$ enthält. Folgen, deren sämtliche Glieder gleich sind, nennen wir *stationär*. Offensichtlich ist die Zuordnung $a \leftrightarrow (a)$ eine isomorphe Abbildung des Körpers Γ auf die Menge Γ' aller der Klassen, die eine stationäre Folge enthalten. Insbesondere ist dann auch Γ' ein Körper (§ 9, Satz 1).

Im Körper D_0 sind — wie in jedem archimedisch angeordneten Körper — die Begriffe des Grenzwertes und der Fundamentalfolge erklärt und stimmen mit den entsprechenden Begriffen in jedem seiner Unterkörper überein (§ 24, Definitionen 3 und 4, Bemerkung und Satz 5).

Wir zeigen, daß unter der Voraussetzung, daß die Klasse α die Folge $\{a_n\}$ enthält, in D_0 $\lim (a_n) = \alpha$ ist. Dazu sei $\varepsilon > (0)$ ein beliebiges Element des Körpers D_0, das etwa die Folge $\{e_n\}$ enthalten möge. Dann gibt es eine rationale Zahl $e > 0$ und eine natürliche Zahl m derart, daß $e_n > e$ ist für alle $n > m$, d. h., es ist $\varepsilon \geq (e)$. Wir wählen nun eine rationale Zahl ε' mit $e > \varepsilon' > 0$ $\left(\text{z. B. } \varepsilon' = \frac{e}{2}\right)$. Dann ist $(\varepsilon') < (e) \leq \varepsilon$. Da die Folge $\{a_n\}$ eine Fundamentalfolge ist, gibt es eine natürliche Zahl n_0 mit $|a_p - a_q| < \varepsilon'$ für alle $p, q > n_0$. Mithin ist für alle $n > n_0$ sowohl $a_p - a_n < \varepsilon'$ als auch $a_n - a_q < \varepsilon'$ für beliebige $p, q > n_0$. Indem wir von den Folgen zu den sie

enthaltenden Klassen übergehen, erhalten wir gemäß unserer Vorbemerkung

$$\alpha - (a_n) \leq (\varepsilon') \text{ und } (a_n) - \alpha \leq (\varepsilon'),$$

d. h., es ist $|(a_n) - \alpha| \leq (\varepsilon') < \varepsilon$ für alle $n > n_0$, also $\lim (a_n) = \alpha$.

Wenn wir jetzt noch zeigen, daß in D_0 jede Fundamentalfolge von Elementen (a_n) des Unterkörpers Γ' einen Grenzwert besitzt, so erhalten wir mühelos die Vollständigkeit von D_0. Dazu sei $\{\alpha_n\}$ eine beliebige Fundamentalfolge von Elementen des Körpers D_0. Nach dem bereits Bewiesenen ist jede der Klassen α_n gleich dem Grenzwert einer Folge von Klassen des Unterkörpers Γ', d. h., es gibt zu jeder natürlichen Zahl n $\left[\text{wegen } \left(\frac{1}{n}\right) > (0)\right]$ ein Element (a_n) aus Γ' derart, daß $|\alpha_n - (a_n)| < \left(\frac{1}{n}\right)$ ist. Wir zeigen nun, daß die Folge $\{(a_n)\}$ der so erhaltenen (a_n) in D_0 eine Fundamentalfolge bildet. Sei dazu $\varepsilon > (0)$ ein beliebiges Element aus D_0. Auf Grund der bereits bewiesenen Gültigkeit des Archimedischen Axioms gibt es dann eine rationale Zahl $e > 0$ mit $(e) < \varepsilon$. Aus demselben Grunde gibt es eine natürliche Zahl n_1 mit $n_1 > \frac{3}{e}$ bzw. $\frac{1}{n_1} < \frac{e}{3}$. Ferner gibt es, da $\{\alpha_n\}$ eine Fundamentalfolge ist, eine natürliche Zahl n_2 derart, daß $|\alpha_p - \alpha_q| < \left(\frac{e}{3}\right)$ ist für alle $p, q > n_2$. Ist n_0 die größere der Zahlen n_1 und n_2, so ist für alle $p, q > n_0$

$$\begin{aligned} |(a_p) - (a_q)| &\leq |(a_p) - \alpha_p| + |\alpha_p - \alpha_q| + |\alpha_q - (a_q)| \\ &< \left(\frac{1}{p}\right) + \left(\frac{e}{3}\right) + \left(\frac{1}{q}\right) < \left(\frac{e}{3}\right) + \left(\frac{e}{3}\right) + \left(\frac{e}{3}\right) = (e) < \varepsilon. \end{aligned}$$

Auf Grund der Isomorphie der Körper Γ und Γ' (bei der offenbar die Anordnung erhalten bleibt) ist dann auch die entsprechende Folge $\{a_n\}$ der rationalen Zahlen a_n eine Fundamentalfolge. Es sei α die Klasse aus D_0, die die Folge $\{a_n\}$ enthält. Wie wir oben gezeigt haben, ist dann $\lim (a_n) = \alpha$. Andererseits ist aber $\lim [(a_n) - \alpha_n] = 0$, denn zu jedem $\varepsilon > (0)$ aus D_0 können wir eine rationale Zahl $e > 0$ und eine natürliche Zahl n_0 bestimmen mit $(e) < \varepsilon$ und $\frac{1}{n_0} < e$; dann ist aber in der Tat $|(a_n) - \alpha_n| < \left(\frac{1}{n}\right) < (e) < \varepsilon$ für alle $n > n_0$. Mithin konvergiert auch die Folge α_n, und es ist

$$\lim \alpha_n = \lim (a_n) = \alpha.$$

Damit ist die Bedingung XII. und mithin Satz 4 bewiesen.

Der Körper D_0 stimmt bis auf Isomorphie mit jedem Körper der reellen Zahlen überein. Jedoch enthält er nicht den Körper Γ der rationalen Zahlen, von dem wir bei seiner Konstruktion ausgingen. Die Elemente des Körpers D_0 sind Klassen äquivalenter Fundamentalfolgen von rationalen Zahlen, aber unter keinen Umständen rationale Zahlen selbst.

Indes haben wir oben gezeigt, daß der Körper D_0 den zum Körper Γ der rationalen Zahlen isomorphen Körper Γ' der stationären Folgen enthält. Daher gibt es auch einen Körper D, der dem Körper D_0 (bezüglich der Addition und der Multiplikation) isomorph ist und der den Körper Γ der rationalen Zahlen als Unterkörper enthält (§ 9, Satz 2). Übertragen wir mittels der isomorphen Abbildung f des Körpers D auf den Körper D_0 die Anordnung von D_0 auf den Körper D vermöge der Festsetzung, daß ein Element d aus D dann und nur dann positiv sein soll, wenn das ihm

entsprechende Element $f(d) = d_0$ des Körpers D_0 positiv ist, so wird D zu einem angeordneten Körper, wobei der Isomorphismus f die Anordnung erhält. Die so erhaltene Anordnung von D erzeugt nun ihrerseits eine Anordnung des Unterkörpers Γ von D, die aber mit der früher für die rationalen Zahlen definierten Anordnung zusammenfällt, da der Körper Γ nur eine Anordnung zuläßt (§ 23, Satz 1). Bei der isomorphen Abbildung von D auf D_0 wird der Körper Γ auf einen gewissen Unterkörper Γ'' von D_0 abgebildet. Da aber andererseits Γ isomorph zu Γ'' ist und Γ nur eine isomorphe Abbildung in D_0 zuläßt (§ 23, Satz 2), ist $\Gamma'' = \Gamma'$, und der rationalen Zahl a aus Γ entspricht bei der isomorphen Abbildung f die Klasse (a) aus Γ'. Da bei der isomorphen Abbildung von D auf D_0 die Anordnung erhalten bleibt, gelten in D alle Eigenschaften der Anordnung gleichermaßen wie in D_0, insbesondere das Archimedische Axiom, die Konvergenz der Fundamentalfolgen, die Äquivalenz von Folgen und die Vollständigkeit. Insbesondere überträgt sich also die Stetigkeit des Körpers D_0 auf den Körper D.

Damit ist die Konstruktion eines Körpers der reellen Zahlen abgeschlossen. Seine Elemente, d. h. reelle Zahlen, sind erstens die rationalen Zahlen und zweitens die Klassen von Fundamentalfolgen aus rationalen Zahlen, die keinen rationalen Grenzwert besitzen.

Aus den Eigenschaften des Körpers D_0 folgt, daß in D jede Fundamentalfolge $\{a_n\}$ aus rationalen Zahlen einen Grenzwert besitzt. Dieser Grenzwert ist entweder eine rationale Zahl (wenn nämlich die Folge $\{a_n\}$ bereits im Körper Γ der rationalen Zahlen konvergiert — *Anm. d. wissenschaftl. Red.*) oder sie ist diejenige Klasse, der die gegebene Folge $\{a_n\}$ angehört.

§ 26. Eigenschaften der reellen Zahlen

Im Körper D der reellen Zahlen sind zunächst alle Eigenschaften erfüllt, die wir in Kapitel II für beliebige angeordnete Körper bewiesen haben. So gibt es z. B. in D keine Nullteiler (§ 7, Definition 2 und Satz 2; § 8, Satz 1). In D gibt es positive und negative Zahlen (§ 10, Definition 1) und eine Ordnung, bezüglich der die positiven Zahlen größer als Null und die negativen Zahlen kleiner als Null sind (§ 10, Satz 1). Für diese Ordnung gelten ferner die Monotoniegesetze und die üblichen Rechenregeln für das Rechnen mit Ungleichungen (§ 10, Satz 2 bis 4). Außerdem ist in D das Quadrat jeder von Null verschiedenen Zahl positiv (§ 10, Satz 7). Schließlich ist für die Zahlen aus D ein absoluter Betrag definiert (§ 10, Definition 2), der die üblichen Eigenschaften besitzt, für den also insbesondere die Regeln über die Rückführung des Rechnens mit Zahlen auf das Rechnen mit ihren Beträgen gelten (§ 10, Satz 8 und die nachfolgende Bemerkung).

Die nichtrationalen reellen Zahlen nennen wir *irrationale Zahlen*.

Wir stellen uns jetzt die Aufgabe, aus beliebigen reellen Zahlen die Wurzel zu ziehen. Die Lösung dieser Aufgabe erhalten wir als Spezialfall einer anderen Aufgabe, der Frage nach den Argumenten, für die eine vorgegebene stetige Funktion einen gegebenen Wert annimmt. Der Begriff der stetigen Funktion spielt neben dem Begriff des Grenzwertes in den mathematischen Untersuchungen eine überragende Rolle.

Den allgemeinen Funktionsbegriff haben wir schon früher (§ 3, Definition 1) kennengelernt. Hier werden wir es ausschließlich mit Funktionen zu tun haben, deren Argumente und Werte dem Körper der reellen Zahlen entnommen sind.

Definition 1. *Unter einer auf der Menge X von reellen Zahlen definierten reellen Funktion (oder Funktion einer reellen Veränderlichen) $y = f(x)$* (oder kurz *f*) *verstehen wir eine Vorschrift, durch die jeder Zahl x aus X eindeutig eine reelle Zahl $y = f(x)$ zugeordnet wird. Die Zahl x heißt ein Argument von f und die Zahl y der Wert von f für das Argument x* (oder *im Punkte x*).

In diesem Paragraphen verstehen wir unter einer Funktion — ohne daß wir dies immer ausdrücklich betonen — stets eine reelle Funktion im angegebenen Sinne.

Definition 2. *Eine auf einer Menge X von reellen Zahlen erklärte Funktion $y = f(x)$ heißt stetig im Punkte x_0 der Menge X, wenn es zu jeder reellen Zahl $\varepsilon > 0$ eine reelle Zahl $\delta > 0$ gibt derart, daß die Ungleichung $|f(x) - f(x_0)| < \varepsilon$ für alle x der Menge X erfüllt ist, die der Bedingung $|x - x_0| < \delta$ genügen. Die Funktion $y = f(x)$ heißt stetig auf der ganzen Menge X, wenn sie in jedem Punkte x_0 aus X (also für alle reellen Zahlen x_0 aus X) stetig ist.*

Der Begriff der stetigen Funktion ist mit dem Begriff des Grenzwertes durch den folgenden Satz verknüpft:

Satz 1. *Eine auf einer Menge X erklärte Funktion f ist dann und nur dann stetig im Punkte x_0 aus X, wenn für jede Zahlenfolge $\{x_n\}$ aus X mit $\lim\limits_{n\to\infty} x_n = x_0$ auch $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$ ist. Die Funktion f ist dann und nur dann stetig auf ganz X, wenn für jede Folge $\{x_n\}$ aus X und jeden Punkt x_0 aus X mit $\lim\limits_{n\to\infty} x_n = x_0$ auch $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$ ist.*

Beweis. Es genügt offenbar, den Teil des Satzes zu beweisen, der sich auf die Stetigkeit der Funktion in einem Punkte bezieht.

a) Es sei die Funktion f stetig im Punkte x_0 und $\{x_n\}$ eine beliebige Folge aus X mit $\lim x_n = x_0$. Ferner sei eine beliebige reelle Zahl $\varepsilon > 0$ vorgegeben. Auf Grund der Definition der Stetigkeit gibt es dazu ein $\delta > 0$ derart, daß $|f(x) - f(x_0)| < \varepsilon$ ist für alle x aus X mit $|x - x_0| < \delta$. Nach Definition des Grenzwertes (§ 24, Definition 3) gibt es zu diesem δ eine natürliche Zahl n_0 derart, daß $|x_n - x_0| < \delta$ ist für alle $n > n_0$. Wir können also zu jedem vorgegebenen $\varepsilon > 0$ eine natürliche Zahl n_0 finden mit $|f(x_n) - f(x_0)| < \varepsilon$ für alle $n > n_0$. Dann ist aber auf Grund der Definition des Grenzwertes

$$\lim_{n\to\infty} f(x_n) = f(x_0).$$

b) Es sei jetzt umgekehrt $\lim\limits_{n\to\infty} f(x_n) = f(x_0)$ für jede gegen x_0 konvergierende Folge $\{x_n\}$ aus X. Angenommen, die Funktion f wäre nicht stetig im Punkte x_0. Dann gäbe es eine reelle Zahl $\varepsilon > 0$ derart, daß es zu jeder reellen Zahl $\delta > 0$ eine Zahl x aus X gibt, für die zwar $|x - x_0| < \delta$, aber nicht $|f(x) - f(x_0)| < \varepsilon$, also $|f(x) - f(x_0)| \geq \varepsilon$ ist. Mithin gäbe es zu jeder natürlichen Zahl n eine Zahl x_n aus X, für die

$$|x_n - x_0| < \frac{1}{n} \tag{1}$$

und

$$|f(x_n) - f(x_0)| \geq \varepsilon \tag{2}$$

ist

Da im Körper der reellen Zahlen das Archimedische Axiom gilt (§ 25, Definition 1), gibt es zu jeder reellen Zahl $\varepsilon_0 > 0$ eine natürliche Zahl n_0 mit $n_0 > \frac{1}{\varepsilon_0}$. Wegen (1) wäre also $|x_n - x_0| < \frac{1}{n} < \frac{1}{n_0} < \varepsilon_0$ für alle $n > n_0$, d. h. $\lim x_n = x_0$. Daraus folgte voraussetzungsgemäß $\lim f(x_n) = f(x_0)$, im Widerspruch zu (2). Also ist in der Tat f stetig im Punkte x_0.

Unter der Summe, der Differenz, dem Produkt, dem Quotienten zweier auf einer Menge X erklärter Funktionen f_1 und f_2 verstehen wir diejenige Funktion f, die jeder Zahl x aus X die Summe beziehungsweise die Differenz, das Produkt oder den Quotienten der Werte der Funktionen f_1 und f_2 im Punkte x zuordnet, d. h., für die $f(x)$ gleich

$$f_1(x) + f_2(x), \quad f_1(x) - f_2(x), \quad f_1(x) \cdot f_2(x), \quad \frac{f_1(x)}{f_2(x)}$$

für jedes x aus X ist (dabei müssen wir im Falle des Quotienten voraussetzen, daß $f_2(x) \neq 0$ ist für alle x aus X).

Aus Satz 1 und den bekannten Eigenschaften des Grenzwertes [§ 24, 2 b), c), d)] folgt unmittelbar

Satz 2. *Die Summe, die Differenz und das Produkt von auf einer Menge X erklärten stetigen Funktionen f_1 und f_2 sind gleichfalls auf der Menge X erklärt und stetig. Der Quotient von auf einer Menge X stetigen Funktionen f_1 und f_2 ist in allen den Punkten x von X stetig, in denen $f_2(x) \neq 0$ ist.*

Wir wollen zunächst einige Beispiele für stetige Funktionen betrachten:

Beispiel 1. Die Funktionen $f(x) = x^k$ mit ganzem $k \geq 0$ sind auf der Menge aller reellen Zahlen definiert und daselbst stetig. Für $k = 0$ ist $f(x) = 1$ für alle x und mithin stetig, weil wegen $|f(x) - f(x_0)| = 0$ jede konstante Funktion stetig ist. Weiterhin ist klar, daß auch die Funktion $f(x) = x$ stetig ist. Hieraus ergibt sich durch vollständige Induktion über k auf Grund von Satz 2 die Stetigkeit von x^k für jede ganze Zahl $k \geq 0$.

Beispiel 2. Aus Beispiel 1 und Satz 2 ergibt sich durch vollständige Induktion über die Anzahl der Glieder leicht die Stetigkeit aller Polynome

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

mit reellen Koeffizienten $a_0, a_1, \ldots, a_n$. Hieraus ergibt sich dann, wiederum auf Grund von Satz 2, daß alle Funktionen $\frac{f(x)}{g(x)}$, wobei $f(x)$ und $g(x)$ Polynome mit reellen Koeffizienten sind, auf der Menge X aller x stetig sind, für die $g(x) \neq 0$ ist. Die Polynome der hier betrachteten Art nennt man *ganze rationale Funktionen* und Quotienten aus solchen *gebrochene rationale Funktionen.*

Beispiel 3. Die Funktionen $\sin x$ und $\cos x$ sind auf der Menge aller reellen Zahlen erklärt und daselbst stetig. Die Funktionen $\operatorname{tg} x$ und $\operatorname{ctg} x$ sind an allen den Stellen stetig, an denen sie erklärt sind, an denen also $\cos x \neq 0$ bzw. $\sin x \neq 0$ ist. Um diese Behauptungen beweisen zu können, müssen wir zunächst eine genaue Definition der genannten Funktionen geben.

Jeder Winkel α bestimmt im Sinne der Geometrie einen gewissen Bogen des Einheitskreises, also des Kreises vom Radius 1. Da nun der Körper der reellen Zahlen stetig ist, gibt es eine reelle Zahl x, die gleich der Länge des betrachteten Bogens ist. Diese Zahl wird das *Bogenmaß des Winkels* α genannt. Andererseits gehört zu

jeder reellen Zahl x ein gewisser Bogen der Länge x und zu diesem ein bestimmter Zentriwinkel α, dessen zugehöriges Bogenmaß gerade die Zahl x ist. Wenn wir in bekannter Weise auch Winkel zulassen, die größer als 360° oder negativ sind, so erhalten wir hierdurch eine eineindeutige Zuordnung zwischen der Menge aller reellen Zahlen und den Winkeln, wobei jeder Zahl x der Zentriwinkel α mit dem Bogenmaß x entspricht. Daher versteht man häufig unter einem Winkel nicht die geometrische Figur, sondern die sein Bogenmaß ausdrückende reelle Zahl. Dann ist $\sin x$ eine Funktion, die jeder reellen Zahl x eine ganz bestimmte reelle Zahl, nämlich (Abb. 1) das Verhältnis der Strecke MP zum Radius MO unter den aus der Trigonometrie bekannten Vereinbarungen über das Vorzeichen zuordnet. Entsprechend werden die anderen trigonometrischen Funktionen erklärt. Wir weisen nochmals darauf hin, daß die prinzipielle Schwierigkeit bei der Definition der trigonometrischen Funktionen in der Messung der Kreisbögen besteht, die aber im Körper der reellen Zahlen als stetigem Körper möglich ist.[1])

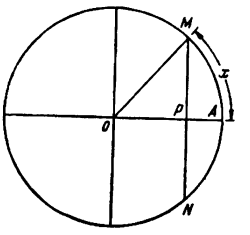


Abb. 1

Die Zuordnung zwischen den Winkeln einerseits und den Bogenmaßen andererseits ist so beschaffen, daß dabei der Summe $\alpha+\beta$ zweier Winkel die Summe $x+y$ ihrer Bogenmaße und dem Vielfachen $a\alpha$ eines Winkels α das Produkt ax des Bogenmaßes x von α und der Zahl a entspricht. Hieraus folgt, daß die bekannten trigonometrischen Formeln (wie Additionstheoreme usw.) für die trigonometrischen Funktionen als Funktionen des Winkels auch für die trigonometrischen Funktionen als Funktionen des Bogenmaßes gelten.

Zum Beweis der Stetigkeit der Funktion $\sin x$ zeigen wir zunächst, daß stets $|\sin x| \leq |x|$ ist. Da $\sin(-x) = -\sin x$ ist, kann man sich dabei auf die Zahlen $x \geq 0$ beschränken; da stets $|\sin x| \leq 1$ ist, genügt es schließlich, die Zahlen x mit $0 \leq x \leq 1$ zu betrachten. Alle zu diesen gehörigen Winkel liegen im ersten Quadranten des Einheitskreises. Dann ist aber offenbar die Länge der Strecke MP gleich der halben Länge der Strecke MN, die zum Bogen $MAN = 2x$ gehört (Abb. 1). Nun sind alle Polygonzüge, die man in den Winkelraum MAN einbeschreiben kann, mindestens so lang wie die Sehne MN. Dann ist aber auch die Länge $2x$ des Bogens MAN als Grenzwert der Längen der einbeschriebenen Polygonzüge nicht kleiner als die Sehne MN. Es ist also $MN \leq 2x, \frac{MN}{2} \leq x$, also $\sin x \leq x$. Nun sollte aber $0 \leq x \leq 1$, also $x > 0$ und $\sin x > 0$ sein. Folglich ist sogar $|\sin x| \leq |x|$, was zu beweisen war.

Es sei uns jetzt eine beliebige reelle Zahl $\varepsilon > 0$ vorgegeben. Dann setzen wir $\delta = \varepsilon$. Auf Grund der Formel

$$\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\cdot\sin\frac{\alpha-\beta}{2}$$

folgt somit unter Benutzung von $|\cos\alpha| \leq 1$, daß für alle x und x_0 mit $|x - x_0| < \delta$

[1]) In den Vorlesungen über Differential- und Integralrechnung wird meistens eine andere Definition dieser Funktionen (etwa durch unendliche Reihen) gegeben, die nicht mit der Winkelmessung oder sonstigen geometrischen Fragen zusammenhängt.

die Ungleichung

$$|\sin x - \sin x_0| = |2 \cos \frac{x + x_0}{2} \cdot \sin \frac{x - x_0}{2}| \leq |x - x_0| < \delta = \varepsilon$$

gilt, womit die Stetigkeit der Funktion sin x bewiesen ist.

Die Stetigkeit der Funktion cos x kann man entsprechend oder einfacher unter Benutzung der Beziehung $\cos x = \sin\left(\frac{\pi}{2} - x\right)$ beweisen. Aus der Stetigkeit der Funktionen sin x und cos x folgt nach Satz 2 die Stetigkeit der Funktionen tg x und ctg x an allen Stellen, an denen sie definiert sind.

Aus den angegebenen Beispielen erkennt man bereits, wie groß die Klasse der stetigen Funktionen ist. Für alle diese Funktionen beweisen wir nun die folgende weittragende Behauptung:

Satz 3. (Satz vom Zwischenwert) *Es sei $f(x)$ eine auf dem abgeschlossenen Intervall $[a, b]$ definierte und stetige Funktion ($[a, b]$ ist die Menge aller reellen Zahlen x mit $a \leq x \leq b$,* siehe Ende von § 1) *und $f(a) = \alpha$ und $f(b) = \beta$. Dann gibt es zu jeder reellen Zahl γ des Intervalls $[\alpha, \beta]$ (falls $\alpha \leq \beta$ ist) bzw. des Intervalls $[\beta, \alpha]$ (falls $\beta \leq \alpha$ ist) eine Zahl c des Intervalls $[a, b]$, für die $f(c) = \gamma$ ist. Mit anderen Worten: Jede auf einem abgeschlossenen Intervall definierte und stetige Funktion nimmt alle Werte an, die zwischen den Werten der Funktion an den Endpunkten des Intervalls liegen.*

Beweis. Wenn $\alpha = \beta$ ist, so kommt nur der Fall $\alpha = \gamma = \beta$ in Frage und $c = a$ oder $c = b$ leistet das Verlangte. Es sei also $\alpha < \beta$ (im Falle $\beta < \alpha$ verläuft der Beweis entsprechend). Wenn $\gamma = \beta$ ist, so kann man $c = b$ setzen und ist fertig. Wir können daher annehmen, daß $\alpha \leq \gamma < \beta$ ist. Wir schließen jetzt weiter nach der sogenannten Halbierungsmethode. Dazu konstruieren wir zwei Folgen $\{a_n\}$ und $\{b_n\}$ von reellen Zahlen, deren Glieder sämtlich im Intervall $[a, b]$ liegen und für die

$$f(a_n) \leq \gamma < f(b_n), \tag{3}$$

$$a_n \leq a_{n+1}, \quad b_n \geq b_{n+1}, \tag{4}$$

$$b_n - a_n = \frac{b - a}{2^n} \tag{5}$$

für alle natürlichen Zahlen n gilt.

Dazu setzen wir $a_0 = a$ und $b_0 = b$. Wenn wir bereits Zahlen a_n und b_n im Intervall $[a, b]$ gefunden haben, so liegt auch $\frac{a_n + b_n}{2}$ im Intervall $[a, b]$, und es ist daher f an dieser Stelle erklärt. Wenn nun

$$f\left(\frac{a_n + b_n}{2}\right) \leq \gamma$$

ist, so setzen wir

$$a_{n+1} = \frac{a_n + b_n}{2} \text{ und } b_{n+1} = b_n.$$

Ist hingegen

$$f\left(\frac{a_n + b_n}{2}\right) > \gamma,$$

so soll

$$a_{n+1} = a_n \text{ und } b_{n+1} = \frac{a_n + b_n}{2}$$

ein. Durch diese Bedingungen sind die Folgen $\{a_n\}$ und $\{b_n\}$ eindeutig festgelegt (§ 15, Satz 1). Sie erfüllen außerdem — wie wir gleich sehen werden — die Be-

dingungen (3), (4) und (5). Die Gültigkeit von (4) folgt unmittelbar aus der Definition von a_{n+1} und b_{n+1}. Die Beziehungen (3) und (5) beweisen wir durch vollständige Induktion über n. Da $\alpha \leq \gamma < \beta$ ist, gelten (3) und (5) für $n = 0$. Wir nehmen nun an, die Beziehungen (3) und (5) seien bereits für n bewiesen, es sei also gezeigt, daß $f(a_n) \leq \gamma < f(b_n)$ und $b_n - a_n = \frac{b-a}{2^n}$ gilt. Nach Definition von a_{n+1} und b_{n+1} ist aber dann offenbar auch

$$f(a_{n+1}) \leq \gamma < f(b_{n+1}), \qquad b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2} = \frac{b-a}{2^{n+1}},$$

womit (3) und (5) für $n+1$ und damit für jedes n bewiesen sind.

Aus (4) folgt, daß $a_p \leq a_q$ ist für alle p und q mit $p < q$. Hieraus und aus (5) folgt, daß $\{a_n\}$ eine Fundamentalfolge ist. In der Tat: Da der Körper der reellen Zahlen archimedisch angeordnet ist, gibt es zu jedem $\varepsilon > 0$ eine natürliche Zahl n_0 derart, daß $\frac{1}{2^{n_0}} < \frac{\varepsilon}{b-a}$ (§ 23, Satz 5) und mithin $\frac{b-a}{2^{n_0}} < \varepsilon$ ist (wegen $\alpha \neq \beta$ ist nämlich auch $a \neq b$ und damit $b - a > 0$). Wenn also $p \leq q$ ist, so ist

$$|a_p - a_q| = a_q - a_p < b_p - a_p = \frac{b-a}{2^p} < \varepsilon$$

für alle $p, q > n_0$.

Auf Grund der Vollständigkeit des Körpers der reellen Zahlen besitzt in ihm die Folge $\{a_n\}$ einen Grenzwert c. Aus (5) folgt leicht (durch abermalige Anwendung von Satz 5 aus § 23), daß $\lim (a_n - b_n) = 0$ und mithin [§ 24, Satz 2 a)] auch $\{b_n\}$ konvergent und $\lim a_n = \lim b_n = c$ ist. Hieraus folgt gemäß Satz 1 wegen der Stetigkeit der Funktion $f(x)$ auf dem Intervall $[a, b]$, daß auch $\lim f(a_n) = \lim f(b_n) = f(c)$ ist. Andererseits folgt aber aus (3), daß $\lim f(a_n) \leq \gamma \leq \lim f(b_n)$ [§ 24, Satz 2 c)], also $f(c) \leq \gamma \leq f(c)$, d. h. $f(c) = \gamma$ ist, was zu beweisen war.

Als Beispiel für die vielfältigen Anwendungsmöglichkeiten dieses Satzes folgern wir aus ihm die Existenz der Wurzel aus einer positiven reellen Zahl und eines Winkels zu vorgegebenem Wert des Sinus, von der wir im nächsten Kapitel Gebrauch machen werden.

Satz 4. *Zu jeder reellen Zahl $a > 0$ und jeder natürlichen Zahl n gibt es eine und nur eine reelle Zahl $b > 0$ derart, daß $b^n = a$ ist. Mit anderen Worten: Für positives a besitzt $\sqrt[n]{a}$ einen und nur einen positiven Wert b. Ist n eine gerade Zahl, so besitzt $\sqrt[n]{a}$ noch genau einen negativen Wert, nämlich $-b$. Wenn $a = 0$ ist, so gibt es genau einen Wert der Wurzel, nämlich $\sqrt[n]{a} = 0$. Wenn $a < 0$ ist, so gibt es für ungerades n einen und nur einen Wurzelwert, der dann negativ ist, und für gerades n keinen reellen Wert für $\sqrt[n]{a}$.*

Beweis. Für jede natürliche Zahl n ist die Funktion $f(x) = x^n$ auf der Menge aller reellen Zahlen erklärt und daselbst stetig. Insbesondere ist sie also auf jedem abgeschlossenen Intervall definiert und stetig. Sei zunächst $a > 0$. Dann setzen wir $c = a + 1$. Aus $c > 1 > 0$ folgt $c^{n-1} \geq 1$ (wobei das Gleichheitszeichen nur im Falle $n = 1$ steht) und mithin $c^n \geq c > a$. Wir wenden nun den Satz vom Zwischenwert auf die Funktion x^n im Intervall $[0, c]$ an. Da $0^n < a < c^n$ ist, gibt es auf Grund dieses Satzes eine Zahl b aus dem Intervall $[0, c]$, für die $b^n = a$, also $\sqrt[n]{a} = b$ ist.

Insbesondere ist also offensichtlich $b > 0$. Wenn nun $b' > 0$ und $b' \neq b$ ist, so ist im Falle $b' < b$ auch $b'^n < b^n$ und im Falle $b' > b$ auch $b'^n > b^n$ (§ 10, Satz 4), also sicher $b'^n \neq b^n$, d. h. $b'^n \neq a$. Damit ist gezeigt, daß der gemäß der vorangehenden Überlegungen gefundene positive reelle Wert von $\sqrt[n]{a}$ der einzige derartige Wert ist.

Für den Fall, daß n eine gerade Zahl ist, ist außerdem

$$(-b)^n = [(-b)^2]^{\frac{n}{2}} = (b^2)^{\frac{n}{2}} = b^n = a,$$

also $-b$ ein weiterer Wert von $\sqrt[n]{a}$, der dann negativ ist. Wenn nun $b' < 0$ und $b' \neq -b$ ist, so ist im Falle $b' < -b$ umgekehrt $b'^2 > b^2 > 0$ und mithin

$$b'^n = (b'^2)^{\frac{n}{2}} > (b^2)^{\frac{n}{2}} = b^n = a,$$

während im Fall $-b < b' < 0$ entsprechend $b'^n < a$ ist. Somit gibt es im betrachteten Fall auch genau einen negativen Wert für die Wurzel, nämlich $\sqrt[n]{a} = -b$.

Wenn n eine ungerade Zahl ist, so besitzt $\sqrt[n]{a}$ keinen negativen Wert, weil aus $b' < 0$ folgt, daß $b'^n < 0 < a$ ist.

Wenn $a = 0$ ist, so ist einerseits $0^n = 0 = a$, während andererseits 0 auch der einzige Wert von $\sqrt[n]{0}$ ist, weil der Körper der reellen Zahlen nullteilerfrei ist und mithin aus $b^n = 0$ unmittelbar $b = 0$ folgt.

Wenn $a < 0$ und n ungerade ist, so gibt es auf Grund des bereits Bewiesenen eine und nur eine positive reelle Zahl b, für die $b^n = -a$ gilt. Dann ist

$$(-b)^n = (-1)^n b^n = a.$$

Wenn jedoch $b' \neq b$ ist, so schließt man genau wie oben, daß auch $b'^n \neq (-b)^n = a$ ist. Im zuletzt betrachteten Fall besitzt also $\sqrt[n]{a}$ den einzigen Wert $-b$.

Wenn schließlich $a < 0$ und n gerade ist, so kann $\sqrt[n]{a}$ keinen reellen Wert besitzen. Da nämlich der Körper der reellen Zahlen ein angeordneter Körper (§ 10, Definition 1) ist, so gilt für jede reelle Zahl b die Ungleichung $b^2 \geq 0$ (§ 10, Satz 7), woraus sich unmittelbar $b^n = (b^2)^{\frac{n}{2}} \geq 0$, also $b^n \neq a$ ergibt.

Wir wollen uns nun der Frage nach der Existenz eines Winkels mit vorgegebenem Wert des Sinus zuwenden. Hier gilt

Satz 5. *Zu jeder reellen Zahl a des Intervalls* $[0, 1]$ *gibt es genau eine Zahl b aus dem Intervall* $\left[0, \frac{\pi}{2}\right]$, *für die* $a = \sin b$ *ist.*

Beweis. Da die Funktion $f(x) = \sin x$ auf der Menge aller reellen Zahlen erklärt und stetig ist, ist sie insbesondere auch im Intervall $\left[0, \frac{\pi}{2}\right]$ erklärt und stetig. Wegen $\sin 0 \leq a \leq \sin \frac{\pi}{2}$ gibt es auf Grund des Satzes vom Zwischenwert eine Zahl b des Intervalls $\left[0, \frac{\pi}{2}\right]$, für die $\sin b = a$ ist. Zum Beweis der Einzigkeit von b benutzen wir die aus der Trigonometrie bekannte Eigenschaft, daß die Funktion $\sin x$ für von

0 bis $\frac{\pi}{2}$ fortschreitendes x ebenfalls wächst. Diese Tatsache kann man etwa folgendermaßen beweisen: Wenn $0 \leq x_1 < x_2 \leq \frac{\pi}{2}$ ist, so entspricht dem größeren Winkel auch der größere Bogen. Außerdem ist dann $0 \leq 2x_1 < 2x_2 \leq \pi$. Für derartige Winkel gehört zu dem größeren Bogen auch die größere Sehne. Die Hälfte der zum Bogen der Länge $2x$ gehörigen Sehne ist nun aber gerade die zum Winkel x gehörige Gegenkathete. Damit ist aber einleuchtend, daß $\sin x_1 < \sin x_2$ ist. Wenn also $b' \neq b$ eine andere Zahl des Intervalles $\left[0, \frac{\pi}{2}\right]$ ist, so ist im Falle $b < b'$ auch $\sin b < \sin b'$ und im Falle $b' < b$ auch $\sin b' < \sin b$, also sicher $\sin b' \neq \sin b$ und damit $\sin b' \neq a$.

Zum Schluß wollen wir noch einige Eigenschaften des Körpers der reellen Zahlen als eines stetigen angeordneten Körpers beweisen.

Satz 6. *Der Körper der reellen Zahlen kann (unter Beibehaltung der obigen Addition und Multiplikation) nur auf eine Weise angeordnet werden und läßt nur eine isomorphe Abbildung (bezüglich der Addition und Multiplikation) auf sich zu, nämlich die identische Abbildung.*

Beweis. Es sei D der in der üblichen Weise (§ 25, Definition 2) angeordnete Körper der reellen Zahlen und D' ein Körper, der aus denselben Elementen wie D besteht, in dem die Addition und die Multiplikation mit der Addition und der Multiplikation in D übereinstimmt, der aber ganz beliebig angeordnet ist. Da die Addition in D mit der Addition in D' übereinstimmt, ist das Nullelement von D auch Nullelement von D'. Ist weiter $a > 0$ (also positiv) in D, so gibt es auf Grund von Satz 4 eine Zahl b derart, daß $a = b^2$ in D und — da die Multiplikation in D mit der Multiplikation in D' übereinstimmt — auch in D' ist. Nun ist D' ein angeordneter Körper, also ist a als Quadrat von b größer oder gleich Null (§ 10, Satz 7) und wegen $a \neq 0$ sogar positiv in D'. Ist hingegen $a < 0$ in D, so ist $-a > 0$ in D und damit auch in D', also ebenfalls $a < 0$ in D'. Hieraus folgt, daß mit $a > 0$ in D' auch $a > 0$ in D gilt, weil $a \leq 0$ in D auch $a \leq 0$ in D' zur Folge hätte. Somit ist a in D' dann und nur dann positiv, wenn a in D positiv ist, d. h., die Anordnung von D' stimmt mit der Anordnung von D überein. D läßt also tatsächlich nur eine Anordnung zu.

Es sei nun $x' = f(x)$ eine (bezüglich der Addition und Multiplikation) isomorphe Abbildung des Körpers D der reellen Zahlen auf sich oder auch auf einen seiner Unterkörper D_0. Wenn $a > 0$ ist, so ist $a = b^2$ mit $b \neq 0$. Auf Grund der Isomorpheieigenschaften ist dann aber

$$a' = f(a) = f(b^2) = [f(b)]^2 > 0,$$

d. h., der Isomorphismus f erhält auch die Anordnung.

Wir zeigen nun zunächst, daß zwischen zwei reellen Zahlen a und b stets eine rationale Zahl c liegt. In der Tat: Wenn etwa $a < b$ ist, so ist $b - a > 0$. Mithin gibt es nach dem Archimedischen Axiom eine natürliche Zahl n mit $n > \frac{1}{b-a}$, also $\frac{1}{n} < b - a$. Weiter gibt es aus demselben Grunde natürliche Zahlen m_1 und m_2 derart, daß

$$m_1 \cdot \frac{1}{n} > a \text{ und } m_2 \cdot \frac{1}{n} > -a,$$

also $(-m_2)\,\frac{1}{n} < a$ ist. Daher ist die Menge A der ganzen Zahlen k mit $k \cdot \frac{1}{n} > a$ nicht leer (weil m_1 zu A gehört) und durch die Zahl $-m_2$ nach unten beschränkt. Folglich enthält die Menge A eine kleinste Zahl m (§ 21, Satz 5), für die also $\frac{m-1}{n} \leqq a < \frac{m}{n}$ ist. Daraus folgt, daß

$$a < \frac{m}{n} = \frac{m-1}{n} + \frac{1}{n} < a + (b-a) = b$$

ist, also die rationale Zahl $\frac{m}{n}$ zwischen a und b liegt.

Bei einer isomorphen Abbildung f des Körpers der reellen Zahlen auf sich wird der Körper der rationalen Zahlen identisch auf sich abgebildet (§ 23, Satz 2). Wäre nun f nicht die identische Abbildung von D auf sich, so gäbe es eine reelle Zahl a, für die $f(a) = b \neq a$ ist. Es möge z. B. $a < b$ sein. Nach dem Bewiesenen gäbe es also eine rationale Zahl c mit $a < c < b$, woraus sich

$$a - c < 0 < b - c \text{ und } c - a > 0 > c - b$$

ergäbe. Nun ist aber

$$f(c-a) = f(c) - f(a) = c - b,$$

d. h., die Zahl $c - a > 0$ würde in die Zahl $c - b < 0$ übergehen, was nach dem bereits Bewiesenen nicht der Fall sein kann.

Das Rechnen mit reellen Zahlen als Klassen äquivalenter Fundamentalfolgen von rationalen Zahlen ist praktisch nicht sehr geeignet, da die Darstellung der reellen Zahlen durch Fundamentalfolgen ziemlich viel Platz erfordert. In der Praxis verwendet man daher zur Bezeichnung der reellen Zahlen durchweg die bequemeren Dezimalbrüche.[1])

§ 27. Die axiomatische Charakterisierung der reellen Zahlen

Im Verlauf der vorangehenden Überlegungen haben wir zunächst die natürlichen Zahlen als Elemente einer beliebigen Menge eingeführt, zwischen deren Elementen eine Beziehung „folgt auf" erklärt ist, die den Peanoschen Axiomen genügt (§ 11, Definition 1). Einen derartigen Aufbau einer mathematischen Theorie nennt man *axiomatisch*. Im weiteren Verlauf definierten wir dann unter Benutzung der natürlichen Zahlen die ganzen Zahlen, die rationalen Zahlen und die reellen Zahlen. In allen drei Fällen ergab sich der neue Zahlenbereich aus dem vorhergehenden, indem man weitergehende Forderungen stellte, die den neuen Bereich bis auf Isomorphie eindeutig festlegten. In jedem der Fälle überzeugten wir uns durch Angabe eines Modelles (eines konkreten Beispiels) von der Existenz eines Bereiches mit den verlangten Eigenschaften. Wegen der Isomorphie aller die jeweiligen Bedingungen erfüllenden Mengen konnten wir in allen Fällen das Modell selbst als entsprechenden Zahlenbereich ansehen. Ein derartiger Aufbau eines Zahlenbereiches wird *konstruktiv* oder auch *genetisch* genannt. Es erhebt sich nun die Frage, ob man jeden der genannten Zahlenbereiche auch axiomatisch festlegen kann.

[1]) Vgl. hierzu den Artikel von A. J. Chintschin, Kap. IV.

Bei der Erweiterung der Zahlenbereiche haben wir jeweils gewisse Forderungen (die Ausführbarkeit der Subtraktion, die Ausführbarkeit der Division, die Stetigkeit) aufgestellt, wobei die jeweilige Erweiterung außerdem minimal sein sollte. Im Fall der reellen Zahlen nun haben wir gesehen, daß sich diese Minimalbedingung erübrigte, daß sie von selbst erfüllt war. Das bedeutet, daß die Menge der reellen Zahlen bereits durch gewisse Eigenschaften ihrer Elemente eindeutig bis auf Isomorphie festgelegt ist. Man kann also gerade diese Eigenschaften der reellen Zahlen zu ihrer axiomatischen Charakterisierung benutzen. Die reellen Zahlen werden somit durch die Eigenschaften eines angeordneten stetigen Körpers axiomatisch festgelegt. Die in diesem Begriff zusammengefaßten Eigenschaften liefern eine Definition des Körpers der reellen Zahlen.

Definition. *Unter einem Körper der reellen Zahlen verstehen wir eine nichtleere Menge D, zwischen deren Elementen eine Addition (welche Elementen a und b aus D ein eindeutig bestimmtes Element $a+b$ aus D, die Summe aus a und b, zuordnet), eine Multiplikation (welche Elementen a und b aus D das Produkt ab zuordnet) und eine Eigenschaft „a ist positiv" erklärt sind, die den folgenden Bedingungen genügen:*

I. (Kommutatives Gesetz der Addition) $a+b=b+a$.

II. (Assoziatives Gesetz der Addition) $a+(b+c)=(a+b)+c$.

III. (Gesetz der Umkehrbarkeit der Addition) *Zu beliebigen Elementen a und b aus D gibt es ein Element c aus D derart, daß $a+c=b$ ist.*

IV. (Kommutatives Gesetz der Multiplikation) $ab=ba$.

V. (Assoziatives Gesetz der Multiplikation) $a(bc)=(ab)c$.

VI. (Distributives Gesetz der Multiplikation bezüglich der Addition) $a(b+c)=ab+ac$.

Die bisher genannten Bedingungen besagen, daß D ein Ring ist. Damit ist insbesondere die Multiplikation der Elemente von D mit natürlichen Zahlen (Vielfachenbildung) erklärt. Ferner gibt es ein eindeutig bestimmtes Element 0 mit $a+0=0+a=a$ für jedes a aus D. Weiter gibt es zu jedem Element a aus D ein eindeutig bestimmtes entgegengesetztes Element $-a$, für das $a+(-a)=(-a)+a=0$ ist. Schließlich gibt es zu gegebenen Elementen a und b aus D ein eindeutig bestimmtes Element $b-a$ (die Differenz aus b und a), für das $a+(b-a)=(b-a)+a=b$ ist.

Weiter verlangen wir:

VII. (Gesetz der Umkehrbarkeit der Multiplikation) *Zu Elementen a und b aus D mit $a\neq 0$ gibt es ein Element q aus D derart, daß $aq=b$ ist.*

VIII (Mächtigkeitsbedingung) *Die Menge D enthält wenigstens zwei Elemente.*

Die Bedingungen I. bis VIII. besagen, daß D ein Körper (§ 8, Definition 1) ist. Damit ist insbesondere der Begriff des Unterkörpers (§ 8, Definition 3) von D festgelegt.

Als nächstes verlangen wir:

IX. *Jedes Element a aus D besitzt eine und nur eine der folgenden drei Eigenschaften: a ist positiv, $a=0$, $-a$ ist positiv.*

X. *Die Summe und das Produkt aus positiven Elementen sind positiv.*

Die Bedingungen I. bis X. besagen, daß D ein angeordneter Körper ist. Speziell können wir also in D dadurch eine Ordnung einführen, daß wir $a > b$ dann und nur dann setzen, wenn $a - b$ positiv ist (§ 10, Satz 1).

Als nächstes verlangen wir:

XI. (Archimedisches Axiom) *Zu jedem a und zu jedem $b > 0$ aus D gibt es eine natürliche Zahl n, derart daß $nb > a$ ist.*

Auf Grund der Bedingungen I. bis XI. ist D ein archimedisch angeordneter Körper. Damit sind in D der Begriff des Grenzwertes einer Folge und der Begriff der Fundamentalfolge erklärt, und zwar unabhängig davon, ob man die Elemente einer Folge als Elemente eines etwaigen Unterkörpers von D oder als Elemente von D auffaßt (§ 24, Satz 5).

Schließlich verlangen wir:

XII. (Axiom der Vollständigkeit) *Jede Fundamentalfolge aus Elementen der Menge D besitzt in D einen Grenzwert.*

Die Axiome I. bis XII. besagen, daß D ein angeordneter stetiger Körper (§ 24, Definition 6) ist.

Die zuletzt gegebene Definition setzt den Begriff der natürlichen Zahl voraus, da sonst das Archimedische Axiom keinen Sinn hätte. Weiter unten werden wir ein Axiomensystem angeben, in das der Begriff der natürlichen Zahl nicht eingeht.

Es erhebt sich nun sofort die Frage nach der Widerspruchsfreiheit, der Vollständigkeit und der Unabhängigkeit des angegebenen Axiomensystems.

Zum Nachweis der Widerspruchsfreiheit der Axiome I. bis XII. haben wir lediglich ein Modell für diese anzugeben (§ 17, Definition 1). Der in § 25 konstruierte Körper D_0 (§ 25, Definition 2 und Satz 4) ist aber gerade ein derartiges Modell. Nun basierte die Konstruktion von D_0 auf dem Körper der rationalen Zahlen. Wenn wir jedoch für diesen seine konstruktive Definition, d. h. den Körper Γ_0 (§ 22, Definition 2) benutzen, in dem der Ring der ganzen Zahlen durch den konstruktiv definierten Ring C_0 (§ 20, Definition 2) ersetzt ist, so können wir die Konstruktion des Körpers D_0 letztlich auf die natürlichen Zahlen zurückführen. Damit ist die Widerspruchsfreiheit des obigen Axiomensystems (im modelltheoretischen Sinne) auf die Widerspruchsfreiheit des Axiomensystems für die natürlichen Zahlen zurückgeführt.

Zum Beweis der Vollständigkeit (Kategorizität) zeigen wir, daß je zwei Modelle dieses Axiomensystems isomorph sind (§ 17, Definition 3). Im wesentlichen haben wir das bereits bewiesen. Wenn nämlich K_1 und K_2 zwei Interpretationen des Systems der Axiome I. bis XII. (also zwei stetige Körper) sind, so gibt es ein Modell Γ des Körpers der rationalen Zahlen und zwei Körper D_1 und D_2 derart, daß D_1 und D_2 den Körper Γ als Unterkörper enthalten und (bezüglich der Addition, Multiplikation und Ordnung) dem Körper K_1 bzw. K_2 isomorph sind (§ 23, Satz 2). Infolge dieses Isomorphismus sind D_1 und D_2 stetige Körper und als Oberkörper von Γ untereinander isomorph (§ 25, Satz 2). Dann sind aber auf Grund der Transitivität der Isomorphiebeziehung auch K_1 und K_2 (bzgl. der Addition, Multiplikation und Ordnung) einander isomorph. Damit ist die Vollständigkeit des Axiomensystems I.—XII. bewiesen.

Mit der Widerspruchsfreiheit und Vollständigkeit ist gleichzeitig bewiesen, daß das Axiomensystem in der Tat den Körper der reellen Zahlen charakterisiert und damit als Fundament für die Theorie der reellen Zahlen dienen kann. Der weitere Aufbau vollzieht sich dann nach dem Muster des vorangehenden Paragraphen.

Die Frage nach der Unabhängigkeit des obigen Axiomensystems ist nicht von der gleichen prinzipiellen Bedeutung wie die Frage der Widerspruchsfreiheit und der Vollständigkeit und möge daher unberücksichtigt bleiben. Es sei nur darauf hingewiesen, daß die Axiome XI. und XII. von den übrigen Axiomen unabhängig sind. Wir hatten die Stetigkeit eines angeordneten Körpers mit Hilfe des Archimedischen Axioms und des Axioms der Vollständigkeit (§ 24, Definition 6) erklärt. Es gibt daneben noch eine ganze Reihe von anderen Möglichkeiten, den Begriff des stetigen Körpers zu erklären. Wir wollen hier noch zwei von ihnen erwähnen. Dazu ist es notwendig, eine Reihe von neuen Begriffsbildungen einzuführen.

Unter einem *Schnitt in einer geordneten Menge* (also insbesondere in einem angeordneten Körper) P verstehen wir ein Paar von nichtleeren Teilmengen X, Y von P, die kein gemeinsames Element besitzen und deren Vereinigung (§ 2) die Menge P ist, d. h., für die

$$X \cap Y = 0, \quad X \cup Y = P$$

gilt und für die zudem $x < y$ für beliebiges $x \in X$ und beliebiges $y \in Y$ ist.

Wenn a größtes Element in X ist und Y kein kleinstes Element besitzt oder wenn a kleinstes Element von Y ist und X kein größtes Element besitzt, so nennen wir a eine *Schnittzahl des gegebenen Schnittes.*

Ein Element b einer geordneten Menge P heißt ein *Häufungspunkt der Menge A*, wenn es zu jedem b_1 und b_2 mit $b_1 < b < b_2$ unendlich viele Elemente a aus A gibt, für die $b_1 < a < b_2$ ist.

Man sieht unmittelbar ein, daß in einem angeordneten Körper K diese letzte Definition mit der nachfolgenden äquivalent ist:

Ein Element b ist dann und nur dann Häufungspunkt der Menge A, wenn es zu jedem Element $\varepsilon > 0$ aus K unendlich viele Elemente a aus A gibt, für welche $|a - b| < \varepsilon$ ist.

Eine Teilmenge A einer geordneten Menge (also insbesondere eines angeordneten Körpers) P heißt *beschränkt*, wenn es in P ein Element b_1 und ein Element b_2 gibt derart, daß $b_1 < a < b_2$ ist für alle $a \in A$.

In einem angeordneten Körper K sind dann die drei folgenden Aussagen gleichwertig:

a) *In K gelten die Axiome XI. und XII.*

b) (DEDEKIND) *Jeder Schnitt in K besitzt eine Schnittzahl.*

c) (WEIERSTRASS) *Jede unendliche beschränkte Teilmenge von K besitzt mindestens einen Häufungspunkt.*

Man kann also den Körper der reellen Zahlen durch die Bedingungen I. bis X. und eine der Bedingungen a), b) oder c) axiomatisch charakterisieren. Der Äquivalenzbeweis für die Aussagen a), b) und c) findet sich z. B. in dem Buch von I. W. PROSKURJAKOW [5].

Den Körper der rationalen Zahlen kann man axiomatisch als Primkörper der Charakteristik 0 festlegen. In der Tat: Jeder derartige Körper ist gleich seinem aus allen rationalen Elementen bestehenden Unterkörper und mithin dem Körper Γ der rationalen Zahlen isomorph (§ 23, Satz 2).

Schließlich kann man den Ring der ganzen Zahlen axiomatisch als Ring R mit Einselement e charakterisieren, der keinen von sich selbst verschiedenen Unterring mit Einselement enthält und in dem $ne \neq 0$ ist für jede natürliche Zahl n. Man

zeigt nämlich leicht, daß die Menge der Elemente ne isomorph ist der Menge N der natürlichen Zahlen bezüglich der Addition und der Multiplikation und folglich R einen dem Ring C der ganzen Zahlen isomorphen Unterring R_0 enthält (§ 20, Satz 3). Da aber R_0 ein Einselement enthält, so stimmt R_0 mit R überein, und es ist R dem Ring der ganzen Zahlen isomorph.

Kapitel VII

DER KÖRPER DER KOMPLEXEN ZAHLEN

§ 28. Definition des Körpers der komplexen Zahlen

Bereits im Altertum stieß man bei der Lösung von quadratischen Gleichungen auf Gleichungen, deren Wurzeln — wie wir heute sagen würden — komplexe Zahlen sind. Derartige Aufgaben sah man damals als unlösbar an. Die in der ersten Hälfte des 16. Jahrhunderts von italienischen Mathematikern gefundenen Formeln zur Auflösung kubischer Gleichungen durch Radikale führten dann auf die Darstellung reeller Wurzeln von Gleichungen mit reellen Koeffizienten durch Quadratwurzeln aus negativen Zahlen. Dadurch wurden die Mathematiker jener Zeit dazu angeregt, sich mit diesen neuen Zahlen, die sie „imaginär", „unmöglich", „eingebildet" usw. nannten, zu befassen. Dabei verwendeten sie für das Rechnen mit diesen Zahlen dieselben Regeln, wie für das Rechnen mit reellen Zahlen. Die Bedeutung jener Zahlen blieb ihnen jedoch unklar, was schon ihre Benennung verrät. So nennt CARDANO die neuen Zahlen „falsche, wahrhaft sophistische" Zahlen. Eine erste formale Begründung des Rechnens mit reellen und komplexen Zahlen findet sich in der „Algebra" des italienischen Mathematikers RAFAELLO BOMBELLI (1572). Die anschaulich-geometrische Darstellung dieser Zahlen (durch Punkte oder Vektoren der Ebene) wurde indes erst am Anfang des 19. Jahrhunderts gefunden.[1])

Danach gingen die Untersuchungen über komplexe Zahlen sehr schnell voran, sodaß heute die Theorie der Funktionen einer komplexen Veränderlichen einen fundamentalen Bestandteil der Mathematik bildet. Diese Theorie findet ihre mannigfachen Anwendungen in den verschiedensten Wissenschaften, so u. a. in der Aerodynamik. Dabei sind die komplexen Zahlen ebenso gut begründet, wie etwa die rationalen oder die reellen Zahlen.

Wir haben gesehen, daß sich im Körper der reellen Zahlen nicht immer die Wurzel aus einer reellen Zahl bestimmen läßt, weil nämlich eine Wurzel geraden Grades aus einer negativen Zahl keinen reellen Wert besitzt, d. h., für reelles $a < 0$ und gerades natürliches n gibt es keine reelle Zahl b, für die $b^n = a$ ist (§ 26, Satz 4). Gemäß unseren in § 18 entwickelten allgemeinen Prinzipien der Erweiterung eines gegebenen

[1]) Die erste geometrische Begründung für das Rechnen mit komplexen Zahlen gaben 1799 der dänische Landmesser CASPAR WESSEL und unabhängig von ihm im Jahre 1806 der französische Mathematiker J.-R. ARGAND. Allgemeine Anerkennung fand diese Auslegung aber erst nach einer weiteren, 1831 durch C. F. GAUSS gegebenen Begründung.

Zahlenbereichs werden wir jetzt den Körper D der reellen Zahlen zum Körper K der komplexen Zahlen erweitern. Dabei wird es sich darum handeln, die Operation des Wurzelziehens in K zu einer unbeschränkt ausführbaren Operation zu machen. Als wesentlich neues Resultat bei dieser Erweiterung ergibt sich, daß auch in den Fällen, in denen diese Operation bereits in D ausführbar ist, in K noch weitere Lösungen auftreten, und zwar wird in K die $\sqrt[n]{a}$ für jedes $a \neq 0$ und beliebiges natürliches n genau n Werte besitzen.[1])

Wir werden sehen, daß es genügt, D zu einem Körper K zu erweitern, in dem $\sqrt{-1}$ mindestens einen Wert besitzt, d. h. zu einem Körper, in dem es ein Element i gibt, für das $i^2 = -1$ ist. Unter allen derartigen Körpern werden wir zudem den kleinsten im Sinne der nachfolgenden Definition auswählen:

Definition 1. *Unter einem Körper der komplexen Zahlen verstehen wir einen minimalen Körper K, der den Körper D der reellen Zahlen enthält und in dem es ein Element i gibt, für das $i^2 = -1$ ist. Ein Körper der komplexen Zahlen ist also eine Menge K, die die folgenden Bedingungen erfüllt:*

1.) *K ist ein Körper, der den Körper D der reellen Zahlen als Unterkörper enthält und in dem es ein Element i mit $i^2 = -1$ gibt.*

2.) *Der Körper K enthält keinen von sich selbst verschiedenen Unterkörper, der ebenfalls die Bedingung 1 erfüllt.*

Die Elemente von K heißen komplexe Zahlen.

Wir beweisen zunächst wieder die Eindeutigkeit (wie immer, bis auf Isomorphie) des so definierten Körpers K.

Satz 1. *Ein Körper K, der den Körper D der reellen Zahlen enthält*[2]*) und in dem es ein Element i mit $i^2 = -1$ gibt, ist dann und nur dann minimal, wenn sich jedes Element x aus K in der Form*

$$x = a + bi \tag{1}$$

darstellen läßt, wobei a, b reelle Zahlen sind. Diese Darstellung ist zudem eindeutig, d. h., zu jedem x aus K gibt es auch nur ein geordnetes Paar a, b von reellen Zahlen, das der Gleichung (1) genügt.

Beweis. a) Es sei jedes Element x aus K in der Form (1) mit reellen a und b darstellbar und P ein beliebiger Unterkörper von K, der den Körper D der reellen Zahlen enthält und in dem es ein Element j mit $j^2 = -1$ gibt. Da dann $i^2 = j^2 = -1$ ist, so ist $(i+j)(i-j) = i^2 + ij - ij - j^2 = 0$. Weil K als Körper keine Nullteiler besitzt (§ 8, Satz 1), muß also $i + j = 0$ oder $i - j = 0$, also $j = \pm i$ sein. Jedes x aus K ist somit wegen $x = a + bi = a \pm bj$ auch Element von P und daher ist $P = K$. Damit ist die Minimaleigenschaft des Körpers K bewiesen.

---

[1]) Die Werte von $\sqrt[n]{a}$ sind offenbar gerade die Wurzeln der Gleichung $x^n - a = 0$. Gleichungen dieser Art werden üblicherweise „reine" Gleichungen genannt. Somit sind also im Körper der komplexen Zahlen alle reinen Gleichungen auflösbar. Darüber hinaus gilt die stärkere Aussage, daß in K alle algebraischen Gleichungen, d. h. alle Gleichungen der Form $f(x) = 0$, wobei $f(x)$ irgendein Polynom vom Grad $n \geq 1$ mit beliebigen komplexen Koeffizienten ist, auflösbar sind. Der Beweis dieses Satzes findet sich in EdEM Band 2, L. J. Okunjew, Der Ring der Polynome und der Körper der rationalen Funktionen, Kap. I, § 6.

[2]) Wie immer soll die Redeweise, ein Ring sei in einem anderen enthalten, so verstanden werden, daß die Operationen in beiden Ringen übereinstimmen.

b) Sei nun umgekehrt der Körper K minimal. Wir zeigen, daß dann jedes Element x aus K in der Form (1) darstellbar ist. Dazu sei M die Menge der Elemente aus K, die sich in der Form (1) darstellen lassen. Wie wir sogleich zeigen werden, gilt für diese Elemente

$$\left.\begin{array}{ll} \text{a)} & a+bi=c+di \text{ dann und nur dann, wenn } a=c \text{ und } b=d \text{ ist;} \\ \text{b)} & (a+bi)\pm(c+di)=(a\pm c)+(b\pm d)\,i; \\ \text{c)} & (a+bi)\,(c+di)=(ac-bd)+(ad+bc)\,i; \\ \text{d)} & \dfrac{a+bi}{c+di}=\dfrac{ac+bd}{c^2+d^2}+\dfrac{bc-ad}{c^2+d^2}\,i, \text{ falls } c+di\neq 0 \text{ ist.} \end{array}\right\} \quad (2)$$

In der Tat: Wenn $a \neq c$ und $b = d$ ist, so ist auf Grund der eindeutigen Ausführbarkeit der Addition und der Multiplikation im Körper K natürlich $a + bi = c + di$. Wenn umgekehrt $a + bi = c + di$ ist, so ist, wenn $b = d$ ist, $bi = di$ und mithin auch $a = c$; wäre aber $b \neq d$, so wäre $i = \frac{a-c}{d-b}$ ein Element des Körpers D der reellen Zahlen, was wegen $i^2 = -1 < 0$ nicht der Fall sein kann, weil das Quadrat einer reellen Zahl niemals negativ ist (§ 10, Satz 7). Wenn also $a + bi = c + di$ ist, so ist $b = d$ und $a = c$, womit die Behauptung a) bewiesen ist.

Da wegen der üblichen Eigenschaften der Null offenbar $0 + 0i = 0$ ist, so folgt aus a) insbesondere, daß dann und nur dann $a + bi = 0$ ist, wenn $a = 0$ und $b = 0$ ist.

Die Rechenregeln b) und c) ergeben sich unmittelbar aus den Eigenschaften der Addition und der Multiplikation im Körper K.

Wenn $c + di \neq 0$ ist, so ist $c \neq 0$ oder $d \neq 0$, also auch $c - di \neq 0$ und $c^2 + d^2 > 0$. Multipliziert man den Zähler und den Nenner auf der linken Seite von d) mit $c - di \neq 0$, so ändert sich der Quotient nicht. Danach kann er aber leicht auf die rechts stehende Form gebracht werden.

Eine unmittelbare Folge aus a) ist die Eindeutigkeit der Darstellung (1).

Aus b), c) und d) folgt nun sofort, daß die Summe und die Differenz sowie das Produkt und der Quotient (mit von Null verschiedenem Nenner) von Elementen der obigen Menge M wieder zu M gehören, also M ein Unterkörper von K ist (§ 8, Satz 5). Da außerdem wegen $a = a + 0i$ alle reellen Zahlen und wegen $i = 0 + 1i$ auch i zu M gehören, so ist $K = M$ auf Grund der Minimalität von K, also jedes Element von K in der Form (1) darstellbar.

Satz 2. *Alle Körper der komplexen Zahlen sind untereinander isomorph, d. h., der Begriff des Körpers der komplexen Zahlen ist durch Definition* 1 *eindeutig bis auf Isomorphie festgelegt.*

Beweis. Es seien K_1 und K_2 beliebige Körper der komplexen Zahlen, wobei K_1 das Element i_1 und K_2 das Element i_2 enthalten möge, für das $i_1^2 = -1$ bzw. $i_2^2 = -1$ ist. Auf Grund des vorangehenden Satzes lassen sich dann alle Elemente aus K_1 in der Form $a + bi_1$ und alle Elemente aus K_2 in der Form $a + bi_2$ darstellen, wobei a und b beliebige reelle Zahlen sind und die jeweilige Darstellung zudem eindeutig ist. Hieraus folgt bereits unmittelbar, daß die Zuordnung $f(a + bi_1) = a + bi_2$ eine eineindeutige Abbildung von K_1 auf K_2 ist. Nun zeigen die Gleichungen (2) b) und c), daß die Addition und die Multiplikation in K_1 und K_2 in gleicher Weise auf die Addition und die Multiplikation von reellen Zahlen zurückgeführt werden. Hieraus folgt leicht, daß die Abbildung f sogar ein Isomorphismus ist. Dazu haben wir zu

zeigen, daß

$$f(x_1+y_1)=f(x_1)+f(y_1), \quad f(x_1y_1)=f(x_1)f(y_1)$$

für beliebiges x_1 und beliebiges x_2 aus K_1 gilt. Da der Beweis beider Gleichungen völlig analog verläuft, wollen wir nur die erste beweisen. Es sei dazu

$$x_1=a+bi_1, \; y_1=c+di_1.$$

Dann ist

$$f(x_1)=a+bi_2, \; f(y_1)=c+di_2$$

und mithin

$$\begin{aligned} f(x_1+y_1) &= f[(a+bi_1)+(c+di_1)]=f[(a+c)+(b+d)\,i_1] \\ &= (a+c)+(b+d)i_2=(a+bi_2)+(c+di_2)=f(x_1)+f(y_1), \end{aligned}$$

was zu beweisen war.

Bemerkung. Bei dem angegebenen Isomorphismus wird offensichtlich jede reelle Zahl in sich und das Element i_1 aus K_1 in das Element i_2 aus K_2 übergeführt.

Satz 3. *Jeder Körper P, der den Körper D der reellen Zahlen enthält und in dem es ein Element i mit $i^2=-1$ gibt, enthält einen Körper der komplexen Zahlen.*

Beweis. Dazu sei K die Menge aller der Elemente aus P, die sich in der Form $a+bi$ mit reellem a und b darstellen lassen. Entsprechend dem Beweis von Satz 1, Teil b) überzeugt man sich davon, daß K ein Unterkörper von P ist, der zudem den Körper der reellen Zahlen und das Element i enthält. Da sich außerdem jedes Element von K in der Form $a+bi$ darstellen läßt, ist auf Grund von Satz 1 der Körper K minimal im Sinne der Definition 1 und mithin ein Körper der komplexen Zahlen. Damit ist der behauptete Satz bewiesen.

Wir wollen als nächstes die Existenz eines Körpers der komplexen Zahlen nachweisen. Dazu genügt es, wie schon im Falle der ganzen, der rationalen und der reellen Zahlen, ein Modell (ein Beispiel) eines der Definition 1 genügenden Körpers anzugeben. Man könnte hier versucht sein, als Elemente dieses Körpers die Symbole $a+bi$ zu nehmen, wobei a, b reelle Zahlen sind und i ein Symbol ist, das der Bedingung $i^2=-1$ genügt. Dann müßte jedoch gezeigt werden, daß in diesen Körper die reellen Zahlen so eingebettet werden können, daß das Symbol $a+bi$ gleich der Summe aus a und dem Produkt von b mit i in dem genannten Körper ist. Wegen der Unbestimmtheit des Symbols i würde es sich zudem um eine rein formale Konstruktion handeln. Wir wollen deshalb einen anderen Weg einschlagen, der zwar dem angedeuteten sehr nahe kommt, sich aber darin von diesem unterscheidet, daß alle verwendeten Symbole eine ganz konkrete Bedeutung haben.

Der obige Satz 1 legt bereits die Konstruktion eines der untereinander isomorphen Körper der komplexen Zahlen nahe. Da jedes Element des gesuchten Körpers sich in der Form $a+bi$ darstellen lassen muß, wird es eindeutig durch das Paar (a, b) von reellen Zahlen festgelegt, wobei außerdem verschiedenen Paaren auch verschiedene Elemente entsprechen. Somit werden wir es hier nicht nötig haben, für die Paare eine Äquivalenzrelation zu erklären und zu den Klassen aus äquivalenten Paaren überzugehen, wie wir es im Falle der ganzen und der rationalen Zahlen getan haben.

Definition 2. *Es sei K_0 die Menge aller geordneten Paare (a, b) von reellen Zahlen. In der Menge K_0 erklären wir eine Addition und eine Multiplikation durch die*

Gleichungen

$$(a, b) + (c, d) = (a + c, b + d), \tag{3}$$

$$(a, b)\,(c, d) = (ac - bd, ad + bc). \tag{4}$$

Die Verknüpfungen in K_0 sind offensichtlich entsprechend den Verknüpfungen in dem gesuchten Körper erklärt, die durch die Gleichungen (2) b) und c) festgelegt wurden.

Satz 4. *Die Menge K_0 bildet bezüglich der Verknüpfungen* (3) *und* (4) *einen Körper.*

Beweis. Dazu haben wir zu zeigen, daß in K_0 die Bedingungen I. bis VIII. (§ 7, Definition 1 und § 8, Definition 1) erfüllt sind.

Da die Addition der Paare gliedweise erklärt ist, folgen die Bedingungen I. bis III. für die Paare unmittelbar aus den entsprechenden Eigenschaften der reellen Zahlen.

Die Gültigkeit der Bedingungen IV. bis VI. verifiziert man durch einfaches Ausrechnen. Als Beispiel beweisen wir das distributive Gesetz der Multiplikation bezüglich der Addition (Bedingung VI.):

$$\begin{aligned}[(a, b) + (c, d)]\,(e, f) &= (a + c, b + d)\,(e, f)\\ &= [(a + c)\,e - (b + d)f, (a + c)f + (b + d)\,e]\\ &= (ae + ce - bf - df, af + cf + be + de);\end{aligned}$$

$$\begin{aligned}(a, b)\,(e, f) + (c, d)\,(e, f) &= (ae - bf, af + be) + (ce - df, cf + de)\\ &= (ae - bf + ce - df, af + be + cf + de).\end{aligned}$$

Da die jeweils zuletzt erhaltenen Paare übereinstimmen, gilt in der Tat VI.

Damit ist bereits gezeigt, daß K_0 ein Ring ist. Man sieht leicht ein, daß das Paar $(0, 0)$ in K_0 die Rolle des Nullelementes spielt und daß die entgegengesetzten Paare und die Differenz zweier Paare durch die Formeln

$$-(a, b) = (-a, -b), \quad (a, b) - (c, d) = (a - c, b - d)$$

festgelegt werden.

Wir fragen jetzt nach der Umkehrbarkeit der Multiplikation (Bedingung VII.) Dazu seien uns zwei beliebige Paare (a, b) und (c, d) vorgegeben, wobei $(a, b) \neq (0, 0)$ ist. Diese letzte Forderung besagt nichts anderes, als daß $a \neq 0$ oder $b \neq 0$ ist.[1]) Da a und b reelle Zahlen sind, so ist unter dieser Bedingung $a^2 + b^2 > 0$ (§ 10, Satz 7). Gesucht ist ein Paar (x, y), für das

$$(a, b)\,(x, y) = (c, d)$$

ist. Setzen wir zunächst einmal voraus, daß es ein derartiges Paar (x, y) gibt, so muß

$$(ax - by, ay + bx) = (c, d)$$

sein, woraus sich $ax - by = c$ und $bx + ay = d$ ergibt. Wenn wir dieses Gleichungssystem nach x und y auflösen, so erhalten wir

$$x = \frac{ac + bd}{a^2 + b^2}, \qquad y = \frac{ad - bc}{a^2 + b^2}.$$

Damit ist gezeigt, daß es unter der Voraussetzung der Existenz eines der Gleichung (5)

[1]) Wie für die Elemente einer beliebigen Menge, setzen wir auch hier fest, daß das Gleichheitszeichen die Identität und das Ungleichheitszeichen die Verschiedenheit bezeichnet. Es ist also im Fall der Paare $(x, y) = (z, t)$ dann und nur dann, wenn $x = z$ und $y = t$ ist.

genügenden Paares (x, y) auch nur ein derartiges Paar gibt, nämlich das Paar aus den soeben angegebenen Zahlen x und y. Man prüft nun leicht nach, daß dieses Paar auch tatsächlich die Gleichung (5) befriedigt, da

$$(a, b)\left(\frac{ac+bd}{a^2+b^2}, \quad \frac{ad-bc}{a^2+b^2}\right)$$
$$= \left(\frac{a^2c+abd-bad+b^2c}{a^2+b^2}, \quad \frac{a^2d-abc+bac+b^2d}{a^2+b^2}\right) = (c, d)$$

ist.

Damit ist für K_0 auch die Eigenschaft VII. nachgewiesen.

Da es schließlich in K_0 sicher mehr als ein Element gibt, so ist auch die Bedingung VIII. erfüllt und K_0 in der Tat ein Körper.

Einselement von K_0 ist das Paar $(1, 0)$, weil

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

ist.

Der Körper K_0 ist bis auf Isomorphie ein Körper der komplexen Zahlen. Er selbst erfüllt allerdings noch nicht die Bedingungen der Definition 1, da er nicht den Körper der reellen Zahlen enthält. Um dies zu erreichen, werden wir wiederum den Körper D in den Körper K_0 einbetten. Dazu sei D' die Menge aller Paare des Körpers K_0, die die Form $(a, 0)$ haben. Aus den Gleichungen (3) und (4), die die Addition und die Multiplikation in K_0 definieren, liest man unmittelbar ab, daß die Zuordnung $a \leftrightarrow (a, 0)$ eine isomorphe Abbildung von D auf D' ist. Daher ist D' ein Körper (§ 9, Satz 1), und folglich gibt es einen Körper K, der den Körper D als Unterkörper enthält und der so auf K_0 isomorph abgebildet werden kann, daß jeder reellen Zahl a aus D das Paar $(a, 0)$ aus D' entspricht (§ 9, Satz 2).

Satz 5. *Der so definierte Körper K ist ein Körper der komplexen Zahlen.*

Beweis. Nach Konstruktion umfaßt der Körper K den Körper D. Ferner enthält K das Paar $(0, 1)$. Dieses Paar wollen wir mit i bezeichnen; es ist also $i = (0, 1)$. In K_0 ist nun

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

Da bei der oben genannten isomorphen Abbildung von K_0 auf K dem Paar $(-1, 0)$ aus K_0 die Zahl -1 aus K entspricht, ist in K also $i^2 = -1$. Somit besitzt K die Eigenschaft 1.) aus Definition 1.

Es bleibt also lediglich zu zeigen, daß K auch die Minimalbedingung 2.) erfüllt. Nach Satz 1 genügt es hierzu zu beweisen, daß sich jedes Element x aus K in der Form $x = a + bi$ mit reellem a und b darstellen läßt. Dazu möge dem Element x aus K bei der isomorphen Abbildung von K auf K_0 das Paar (a, b) aus K_0 entsprechen. Offenbar ist dann in K_0

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

Auf Grund der Zuordnung $a \leftrightarrow (a, 0)$ ist dann aber in K in der Tat $x = a + bi$, was zu beweisen war.

§ 29. Eigenschaften der komplexen Zahlen[1])

Der Körper der komplexen Zahlen besitzt als Körper alle die in den §§ 7 und 8 für beliebige Ringe und Körper bewiesenen Eigenschaften.

Da der Körper der komplexen Zahlen den Körper der reellen Zahlen umfaßt, ist seine Charakteristik gleich Null.

Der Körper der komplexen Zahlen ist sicher kein angeordneter Körper, da in einem solchen $a^2 > 0$ für jedes a ist (§ 10, Satz 7), während im Körper der komplexen Zahlen $i^2 = -1$ ist.

Die geometrische Darstellung der komplexen Zahlen. In einer Ebene wählen wir zwei beliebige, aufeinander senkrecht stehende Geraden, eine horizontale Gerade Ox und eine vertikale Gerade Oy aus, die sich im Punkte O schneiden mögen (Abb. 2). Ferner wählen wir eine beliebige Strecke MN als Einheit der Längenmessung. Dann lassen sich alle komplexen Zahlen als Punkte der Ebene Oxy darstellen. Dazu trägt man, falls $z = a + bi$ eine gegebene komplexe Zahl ist, auf der Geraden Ox vom Punkte O aus die Strecke OA der Länge $|a|$ ab, und zwar nach rechts, wenn $a > 0$ ist, und nach links, wenn $a < 0$ ist. Entsprechend wird auf der Geraden Oy die Strecke OB der Länge $|b|$ abgetragen, und zwar nach oben, wenn $b > 0$ ist, und nach unten, wenn $b < 0$ ist. Durch den Punkt A legt man sodann eine zur Geraden Oy parallele Gerade und durch B eine zu Ox parallele Gerade. Der Schnittpunkt Z dieser beiden Geraden dient als Bild der komplexen Zahl $z = a + bi$. Man überzeugt sich leicht davon, daß jeder Punkt der gewählten Ebene eine gewisse komplexe Zahl darstellt und daß die so erklärte Zuordnung zwischen den komplexen Zahlen und den Punkten der Ebene Oxy umkehrbar eindeutig ist. Es wird also — kurz gesagt — die Zahl $z = a + bi$ durch den Punkt $Z(a, b)$ mit den rechtwinkligkartesischen Koordinaten a und b dargestellt.

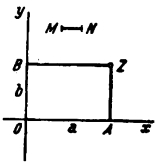


Abb. 2

Bei dieser Zuordnung entsprechen offenbar den reellen Zahlen und nur diesen die Punkte der Geraden Ox, während die Zahlen der Form bi, die sogenannten *rein imaginären Zahlen*, und nur diese durch die Punkte der Geraden Oy dargestellt werden. Daher heißt die Gerade Ox die *reelle* und die Gerade Oy die *imaginäre Achse*. Die nach rechts weisende Richtung der Geraden Ox und die nach oben weisende Richtung von Oy werden als *positiv* bezeichnet, während die nach links weisende Richtung von Ox und die nach unten weisende Richtung von Oy *negativ* heißen. Der Punkt O heißt der *Koordinatenursprung*. Die Geraden Ox und Oy werden *Koordinatenachsen* genannt.

Im folgenden werden wir bei Beweisen für irgendwelche Eigenschaften der komplexen Zahlen nicht unmittelbar mit ihrer geometrischen Darstellung operieren, sondern diese nur zur Veranschaulichung der jeweiligen Eigenschaft benutzen.

---

[1]) Wir werden hier nur die elementaren Eigenschaften der komplexen Zahlen behandeln. Der Leser, der sich mit weiteren Eigenschaften dieser Zahlen (z. B. der Teilbarkeitstheorie im Ring der sogenannten ganzen komplexen Zahlen) eingehender beschäftigen möchte, sei auf das Buch von R. O. Kusmin und D. K. Faddejew [13] verwiesen.

Die trigonometrische Darstellung der komplexen Zahlen. Definition. *Unter einer trigonometrischen Darstellung einer komplexen Zahl z verstehen wir ihre Schreibweise in der Form*

$$z = r(\cos\alpha + i\sin\alpha),$$

wobei r und α reelle Zahlen sind und $r \geq 0$ ist. Die Zahl r heißt der absolute Betrag (oder kurz: *Betrag*) *von z und die Zahl α das Argument oder der Arcus von z.*

Satz 1. *Jede komplexe Zahl z läßt sich in trigonometrischer Darstellung schreiben. Dabei ist der Betrag von z eindeutig durch z bestimmt und dann und nur dann gleich Null, wenn $z = 0$ ist. Für $z = 0$ kann Argument von z jede beliebige reelle Zahl sein, während für $z \neq 0$ das Argument von z eindeutig bis auf ganze Vielfache von 2π bestimmt ist.*

Beweis. Wenn $z = 0$ ist, so ist $0 \cdot (\cos\alpha + i\sin\alpha)$ für beliebiges α eine trigonometrische Darstellung von z. Wenn umgekehrt $r(\cos\alpha + i\sin\alpha) = 0$ ist, so ist $r = 0$, da wegen $\sin^2\alpha + \cos^2\alpha = 1$ stets $\cos\alpha + i\sin\alpha \neq 0$ ist. Damit sind die die Zahl $z = 0$ betreffenden Behauptungen unseres Satzes bewiesen.

Sei also jetzt $z = a + bi \neq 0$. Dann sind nicht a und b beide Null, und es ist $a^2 + b^2 > 0$. Mithin besitzt $\sqrt{a^2 + b^2}$ im Körper der reellen Zahlen zwei Werte, und zwar einen positiven und einen negativen Wert (§ 26, Satz 4). Den positiven Wert dieser Wurzel wollen wir mit r bezeichnen. Da dann $a^2 \leq r^2$ und $b^2 \leq r^2$ ist, so ist $\left|\frac{a}{r}\right| \leq 1$ und $\left|\frac{b}{r}\right| \leq 1$. Mithin gibt es eine Zahl α_0 derart, daß $0 \leq \alpha_0 \leq \frac{\pi}{2}$ und $\sin\alpha_0 = \left|\frac{b}{r}\right|$ ist (§ 26, Satz 5). Da außerdem $\left(\frac{a}{r}\right)^2 + \left(\frac{b}{r}\right)^2 = 1$ ist, so ist $\frac{a}{r} = \pm\cos\alpha_0$ und $\frac{b}{r} = \pm\sin\alpha_0$.

Wenn nun $\frac{a}{r} = \cos\alpha_0$ und $\frac{b}{r} = \sin\alpha_0$ ist, so setzen wir $\alpha = \alpha_0$; wenn $\frac{a}{r} = -\cos\alpha_0$ und $\frac{b}{r} = \sin\alpha_0$ ist, so setzen wir $\alpha = \pi - \alpha_0$; wenn $\frac{a}{r} = \cos\alpha_0$ und $\frac{b}{r} = -\sin\alpha_0$ ist, so setzen wir $\alpha = -\alpha_0$; wenn schließlich $\frac{a}{r} = -\cos\alpha_0$ und $\frac{b}{r} = -\sin\alpha_0$ ist, so setzen wir $\alpha = \pi + \alpha_0$. In allen Fällen erhalten wir eine reelle Zahl α derart, daß

$$\frac{a}{r} = \cos\alpha, \quad \frac{b}{r} = \sin\alpha$$

und mithin

$$z = a + bi = r\left(\frac{a}{r} + i\,\frac{b}{r}\right) = r(\cos\alpha + i\sin\alpha)$$

ist. Also besitzt z jedenfalls eine trigonometrische Darstellung. Offenbar erhält man aus einer trigonometrischen Darstellung einer Zahl z weitere Darstellungen derselben Zahl z, wenn man zu α eine Zahl der Form $2k\pi$ addiert, wobei k eine beliebige ganze Zahl ist.

Wir beweisen als nächstes die Eindeutigkeit des Betrages. Dazu sei

$$a + bi = r(\cos\alpha + i\sin\alpha).$$

Dann ist offensichtlich

$$a = r\cos\alpha, \quad b = r\sin\alpha. \tag{1}$$

Quadriert man diese beiden Gleichungen und addiert sie anschließend, so ergibt

sich $a^2 + b^2 = r^2$, also $r = \sqrt{a^2 + b^2}$. Gemäß der Voraussetzung $r \geqq 0$ kommt von den beiden Werten der Wurzel nur der positive Wert in Frage, womit die Eindeutigkeit von r bewiesen ist.

Sind schließlich zwei trigonometrische Darstellungen einer und derselben komplexen Zahl z gegeben, so ist

$$r(\cos \alpha_1 + i \sin \alpha_1) = r(\cos \alpha_2 + i \sin \alpha_2),$$

woraus im Falle $z \neq 0$ zunächst $r \neq 0$ und dann

$$\cos \alpha_1 = \cos \alpha_2, \quad \sin \alpha_1 = \sin \alpha_2$$

folgt, was seinerseits — wie man aus der Trigonometrie weiß — die Gleichung $\alpha_1 = \alpha_2 + 2k\pi$ mit einer gewissen ganzen Zahl k zur Folge hat.

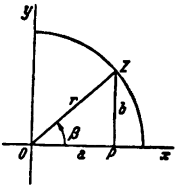


Abb. 3

Wir wollen jetzt die Frage nach der geometrischen Bedeutung des Betrages und des Argumentes untersuchen. Dazu möge der Zahl $z = r(\cos \alpha + i \sin \alpha)$ der Punkt Z der Ebene Oxy entsprechen (Abb. 3). Wir verbinden dann den Punkt Z durch eine Gerade mit dem Koordinatenursprung O und fällen vom Punkte Z auf die reelle Achse Ox das Lot ZP. Wenn $z = a + bi$ ist, so ist die Länge der Strecke OP gleich $|a|$ und die Länge der Strecke ZP gleich $|b|$. Daher ist

$$OZ^2 = OP^2 + ZP^2 = a^2 + b^2 = r^2$$

und mithin $r = OZ$. Der Betrag von z ist also, geometrisch betrachtet, nichts anderes als der Abstand des Punktes Z vom Koordinatenursprung. Wenn außerdem β das Bogenmaß des Winkels zwischen der Geraden OZ und der positiven Richtung der reellen Achse ist, so sieht man — wenn man noch den Kreis vom Radius r um den Punkt O als Mittelpunkt zeichnet —, daß die Zahlen a und b ihrem absoluten Betrage nach mit der Länge der Ankathete bzw. der Gegenkathete des Winkels β im Dreieck OPZ übereinstimmen. Somit muß wegen (1)

$$\cos \beta = \frac{a}{r} = \cos \alpha, \quad \sin \beta = \frac{b}{r} = \sin \alpha$$

sein, woraus sich $\alpha = \beta + 2k\pi$ ergibt. Damit ist gezeigt, daß das Argument von z bis auf ganze Vielfache von 2π gleich dem von der Geraden OZ und der positiven Richtung der reellen Achse gebildeten Winkel ist. Aus dem Gesagten folgt unmittelbar, daß der Betrag und das Argument einer komplexen Zahl z nichts anderes sind als die Polarkoordinaten des ihr entsprechenden Punktes Z in dem System, dessen Pol der Koordinatenursprung O ist und dessen Polarachse in die positive Richtung der reellen Achse Ox weist.

Besonders übersichtlich wird in der trigonometrischen Darstellung der komplexen Zahlen ihre Multiplikation. Hier gilt:

Satz 2. *Bei der Multiplikation einer beliebigen endlichen Anzahl von komplexen Zahlen multiplizieren sich ihre Beträge, während sich ihre Argumente addieren.*

Beweis. Wir beschränken uns hier auf ein Produkt aus zwei Faktoren. Der allgemeine Fall folgt hieraus unmittelbar durch vollständige Induktion.

Wir zeigen also, daß

$$[r_1(\cos\alpha_1 + i\sin\alpha_1)]\cdot[r_2(\cos\alpha_2 + i\sin\alpha_2)] = (r_1r_2)[\cos(\alpha_1+\alpha_2) + i\sin(\alpha_1+\alpha_2)] \quad (2)$$

ist. Dies folgt aber unmittelbar daraus, daß

$$\begin{aligned}(\cos\alpha_1 + i\sin\alpha_1)(\cos\alpha_2 + i\sin\alpha_2) &= \\ &= (\cos\alpha_1\cos\alpha_2 - \sin\alpha_1\sin\alpha_2) + i(\cos\alpha_1\sin\alpha_2 + \sin\alpha_1\cos\alpha_2) \\ &= \cos(\alpha_1+\alpha_2) + i\sin(\alpha_1+\alpha_2)\end{aligned}$$

ist. Da nun mit $r_1 \geq 0$ und $r_2 \geq 0$ auch $r_1r_2 \geq 0$ ist, so ist in der Tat r_1r_2 der Betrag und $\alpha_1+\alpha_2$ das Argument des Produktes der gegebenen Zahlen. Damit ist aber bereits der behauptete Satz für den Fall zweier Faktoren bewiesen.

Aus diesem Satz folgt unmittelbar

Satz 3. *Bei der Division komplexer Zahlen dividieren sich ihre Beträge, während sich ihre Argumente subtrahieren, d. h., es ist*

$$\frac{r_1(\cos\alpha_1 + i\sin\alpha_1)}{r_2(\cos\alpha_2 + i\sin\alpha_2)} = \frac{r_1}{r_2}[\cos(\alpha_1-\alpha_2) + i\sin(\alpha_1-\alpha_2)]. \quad (3)$$

Beweis. Wie jede komplexe Zahl, so läßt sich auch der Quotient zweier gegebener komplexer Zahlen in der trigonometrischen Form darstellen. Sei diese in unserem Fall $r_0(\cos\alpha_0 + i\sin\alpha_0)$. Auf Grund der Definition des Quotienten ist dann aber

$$\begin{aligned}r_1(\cos\alpha_1 + i\sin\alpha_1) &= r_2(\cos\alpha_2 + i\sin\alpha_2)\cdot r_0(\cos\alpha_0 + i\sin\alpha_0) \\ &= r_2r_0[\cos(\alpha_2+\alpha_0) + i\sin(\alpha_2+\alpha_0)],\end{aligned}$$

woraus sich nach Einbeziehung evtl. Vielfacher von 2π in α_0 unmittelbar $r_1 = r_2r_0$ und $\alpha_1 = \alpha_2 + \alpha_0$, also $r_0 = \frac{r_1}{r_2}$, $\alpha_0 = \alpha_1 - \alpha_2$ ergibt, womit der behauptete Satz bewiesen ist.

Für den Fall, daß in einem Produkt alle Faktoren übereinstimmen, erhält man aus Satz 2 die sogenannte MOIVREsche Formel

$$[r(\cos\alpha + i\sin\alpha)]^n = r^n(\cos n\alpha + i\sin n\alpha). \quad (4)$$

Mit Hilfe dieser Formel läßt sich jetzt leicht die Frage nach den Werten der Wurzel aus einer komplexen Zahl beantworten. Hierüber gibt Auskunft der folgende

Satz 4. *Es sei z eine beliebige komplexe Zahl und n eine natürliche Zahl. Dann besitzt $\sqrt[n]{z}$ im Körper der komplexen Zahlen für $z = 0$ den einzigen Wert $\sqrt[n]{z} = 0$ und für $z \neq 0$ genau n verschiedene Werte, und zwar sind dies, wenn*

$$z = r(\cos\alpha + i\sin\alpha)$$

ist, die Werte

$$z_k = \sqrt[n]{r}\left(\cos\frac{\alpha+2k\pi}{n} + i\sin\frac{\alpha+2k\pi}{n}\right) \quad (k = 0, 1, 2, \ldots, n-1). \quad (5)$$

Beweis. Es ist einerseits $0^n = 0$ und andererseits — da der Körper der komplexen Zahlen als Körper nullteilerfrei ist — mit $x^n = 0$ auch $x = 0$. Also ist in der Tat die Zahl 0 der einzige Wert von $\sqrt[n]{z}$ im Falle $z = 0$.

Es sei jetzt

$$z = r(\cos\alpha + i\sin\alpha) \neq 0.$$

Dann ist $r \neq 0$ und α bis auf ganze Vielfache von 2π eindeutig bestimmt. Wir

setzen zunächst voraus, daß x ein Wert von $\sqrt[n]{z}$ im Körper der komplexen Zahlen ist, daß also $x^n = z$ ist. Auf Grund von Satz 1 können wir dann x in seiner trigonometrischen Darstellung in der Form

$$x = r'(\cos\alpha' + i\sin\alpha'), \quad r' > 0$$

ansetzen. Nach der MOIVREschen Formel (4) ist dann

$$r'^n(\cos n\alpha' + i\sin n\alpha') = r(\cos\alpha + i\sin\alpha),$$

woraus sich

$$r'^n = r, \quad n\alpha' = \alpha + 2k\pi, \quad r' = \sqrt[n]{r}, \quad \alpha' = \frac{\alpha + 2k\pi}{n}$$

ergibt. Dabei kann man sich auf ganze Zahlen k mit $0 \leq k \leq n-1$ beschränken. Wenn man nämlich k durch n mit Rest dividiert, so ist $k = nq + k_1$, wobei q und k_1 ganze Zahlen sind und $0 \leq k_1 \leq n-1$ ist; dann ist aber

$$\alpha' = \frac{\alpha + 2k\pi}{n} = \frac{\alpha + 2k_1\pi}{n} + 2q\pi;$$

und es kann — da das Argument von x nur bis auf ganze Vielfache von 2π festgelegt ist — das Argument von x gleich $\frac{\alpha + 2k_1\pi}{n}$ gesetzt werden.

Wir finden also, daß

$$x = \sqrt[n]{r}\left(\cos\frac{\alpha + 2k\pi}{n} + i\sin\frac{\alpha + 2k\pi}{n}\right) \quad (0 \leq k \leq n-1)$$

sein muß. Wenn $\sqrt[n]{z}$ überhaupt einen Wert im Körper der komplexen Zahlen besitzt, so stimmt er mit einer der Zahlen z_k aus (5) überein.

Andererseits sieht man auch leicht ein, daß alle Zahlen z_k aus (5) tatsächlich Werte von $\sqrt[n]{z}$ sind, und dies sogar für beliebiges ganzes k. In der Tat ist

$$z_k^n = \left(\sqrt[n]{r}\right)^n\left(\cos\frac{\alpha + 2k\pi}{n} + i\sin\frac{\alpha + 2k\pi}{n}\right)^n = r(\cos\alpha + i\sin\alpha) = z.$$

Wir müssen jedoch noch zeigen, daß alle z_k mit $k = 0, 1, 2, \ldots, n-1$ untereinander verschieden sind. Wenn aber $k \neq l$ und $z_k = z_l$ ist, so folgt aus Satz 1 wegen $r \neq 0$, daß

$$\frac{\alpha + 2k\pi}{n} = \frac{\alpha + 2l\pi}{n} + 2m\pi$$

mit einer gewissen ganzen Zahl m ist, und hieraus, daß $k - l = mn$ ist. Wäre nun $0 \leq k < n$ und $0 \leq l < n$, so wäre $|k - l| < n$ und mithin $|mn| < n$, also $|m| < 1$. Da m eine ganze Zahl ist, müßte also $m = 0$ und mithin $k = l$ sein, was im Widerspruch zur Voraussetzung $k \neq l$ steht.

Damit ist Satz 4 in allen Teilen bewiesen.

Aus (5) kann man noch unmittelbar die geometrische Bedeutung der Werte von $\sqrt[n]{z}$ für $z \neq 0$ ablesen. Da die Beträge aller z_k gleich $\sqrt[n]{r}$ sind, liegen die ihnen entsprechenden Punkte auf einem Kreise vom Radius $\sqrt[n]{r}$ um den Koordinatenursprung als Mittelpunkt. Da sich außerdem die Argumente zweier benachbarter Zahlen

z_k und z_{k+1} um $\frac{2\pi}{n}$ unterscheiden, liegen die zugehörigen Punkte auf den Ecken eines dem eben genannten Kreise einbeschriebenen regelmäßigen n-Ecks, wobei ferner eine der Ecken der der Zahl z_0 entsprechende Punkt ist, die das Argument $\frac{\alpha}{n}$ besitzt. Durch diese letzte Angabe ist aber bereits die Lage der anderen Ecken des n-Ecks eindeutig festgelegt.

Nach dieser geometrischen Interpretation der komplexen Werte von $\sqrt[n]{z}$ erscheinen jetzt die früheren Aussagen über die Wurzeln einer reellen Zahl (§ 26, Satz 4) in einem ganz neuen Licht. Gesucht seien also die reellen Werte von $\sqrt[n]{z}$ für reelles $z \neq 0$. Diese Werte werden durch diejenigen Ecken des oben genannten regelmäßigen n-Ecks repräsentiert, die auf der reellen Achse liegen. Daher ist zunächst klar, daß es höchstens zwei derartige Werte geben kann, und wenn es zwei Werte gibt, daß sie gleichen absoluten Betrag, aber verschiedenes Vorzeichen besitzen. Wenn nun $z > 0$ ist, so ist das Argument α von z gleich Null, und die z_0 darstellende Ecke des regelmäßigen n-Ecks liegt auf der positiven reellen Halbachse. Ist dabei n eine gerade natürliche Zahl, so liegt offensichtlich auch die gegenüberliegende Ecke auf der reellen Achse, und wir erhalten zwei reelle Werte von $\sqrt[n]{z}$. Wenn jedoch n ungerade ist, so kann keine weitere Ecke auf der reellen Achse liegen, und daher gibt es in diesem Fall nur einen reellen Wert der Wurzel. Wenn dagegen $z < 0$ ist, so ist das Argument α von z gleich π. Nun ist ein z_k dann und nur dann reell, wenn sein Argument ein ganzes Vielfaches von π ist. Für ungerades n ist das Argument $\frac{\pi + 2k\pi}{n} = \pi\frac{2k+1}{n}$ von z_k genau dann ein ganzes Vielfaches von π, wenn $k = \frac{n-1}{2}$ ist, und man erhält dann einen reellen Wert der Wurzel mit dem Argument π, also einen negativen Wert. Dagegen kann für gerades n das Argument $\pi\frac{2k+1}{n}$ kein ganzzahliges Vielfaches von π sein, d. h., in diesem Falle gibt es keinen reellen Wert von $\sqrt[n]{z}$.

Eigenschaften des Betrages einer komplexen Zahl. Auch den Betrag einer komplexen Zahl z bezeichnet man gewöhnlich mit $|z|$. Für reelles z sind dabei keine Verwechslungen mit der früheren Bezeichnung des absoluten Betrages einer reellen Zahl zu befürchten. Wenn nämlich $z = a + bi$ eine reelle Zahl ist, so ist $b = 0$, und der Betrag (im Sinne des Betrages einer komplexen Zahl) von z ist

$$|z| = \sqrt{a^2 + b^2} = \sqrt{a^2} = |a|,$$

d. h., der Betrag der reellen Zahl z ist gleich dem absoluten Betrag von z (im Sinne des absoluten Betrages einer reellen Zahl).

Die komplexen Zahlen

$$z = a + bi \text{ und } \bar{z} = a - bi$$

heißen zueinander *konjugiert*. Offenbar besitzen konjugiert komplexe Zahlen denselben absoluten Betrag. Ferner ist das Produkt zweier konjugiert komplexer Zahlen gleich dem Quadrat ihres absoluten Betrages:

$$z\bar{z} = (a + bi)(a - bi) = a^2 + b^2 = r^2.$$

Hieraus folgt, daß

$$|z| = \sqrt{z\bar{z}} \tag{6}$$

ist.

Der absolute Betrag der komplexen Zahlen besitzt ferner analoge Eigenschaften wie der absolute Betrag der Elemente eines angeordneten Körpers (§ 10, Satz 8). So ist

$$|xy| = |x| \cdot |y|, \tag{7}$$

$$|x + y| \leq |x| + |y| \tag{8}$$

für beliebige komplexe Zahlen x und y.

In der Tat: Die Gleichung (7) ist offensichtlich in Satz 2 mitbewiesen. Will man jedoch nicht auf die trigonometrische Darstellung der komplexen Zahlen zurückgehen, sondern als Definition des absoluten Betrages die Gleichung (6) benutzen, so kann man zum Beweis von (7) auch folgendermaßen schließen:

$$|xy| = \sqrt{(xy)\,\overline{(xy)}} = \sqrt{xy\,\bar{x}\,\bar{y}} = \sqrt{x\bar{x}}\,\sqrt{y\bar{y}} = |x| \cdot |y|.$$

Zum Beweis der Ungleichung (8) zeigen wir zunächst, daß stets

$$|1 + z| \leq 1 + |z| \tag{9}$$

ist. Dazu sei $z = a + bi$. Dann ist

$$|z| = \sqrt{a^2 + b^2} \geq \sqrt{a^2} = |a|,$$

$$|1 + z|^2 = (1 + z)\,(1 + \bar{z}) = 1 + (z + \bar{z}) + z\bar{z} = 1 + 2a + |z|^2$$

$$\leq 1 + 2 \cdot |z| + |z|^2 = (1 + |z|)^2,$$

woraus sich unmittelbar $|1 + z| \leq 1 + |z|$, also (9) ergibt.

Nun zum Beweis von (8)! Falls $x = 0$ ist, so gilt (8) trivialerweise. Es sei also $x \neq 0$.

Dann ist

$$|x + y| = |x\,(1 + x^{-1}y)\,| = |x| \cdot |1 + x^{-1}y| \leq |x|(1 + |x^{-1}y|) =$$

$$= |x| + |x|\,|x^{-1}y| = |x| + |x x^{-1} y| = |x| + |y|,$$

was zu beweisen war.

Bei der Definition des Grenzwertes einer Folge, der Fundamentalfolge und der Vollständigkeit eines Körpers (§ 24, Definition 3 bis 5) haben wir lediglich den Begriff des absoluten Betrages benutzt, während zum Nachweis ihrer wichtigsten Eigenschaften nur die in Satz 8 aus § 10 bewiesenen Eigenschaften des absoluten Betrages verwendet wurden, also daß $|a| > 0$ ist für $a \neq 0$ und daß

$$|ab| = |a|\,|b|, \quad |a + b| \leq |a| + |b|$$

ist. Die Formeln (7) und (8) zeigen nun, daß auch der absolute Betrag im Bereich der komplexen Zahlen diese Eigenschaften besitzt. Deshalb kann man genauso im Körper der komplexen Zahlen die obengenannten Begriffe einführen und für sie die früheren Eigenschaften beweisen. Ebenfalls übertragen sich auch die fundamentalen Begriffe der Differential- und Integralrechnung vom Körper der reellen Zahlen auf den Körper der komplexen Zahlen. Alles dies wird in der umfangreichen und schönen Theorie der Funktionen einer komplexen Veränderlichen (Funktionentheorie) behandelt.[1])

[1]) Siehe EdEM Band 3, W. L. Gontscharow, Elementare Funktionen im Bereich der komplexen Zahlen.

§ 30. Hyperkomplexe Zahlen. Quaternionen

In diesem Paragraphen werden wir sowohl den Begriff des Vektorraumes und seine wichtigsten Eigenschaften als auch den Begriff des Polynoms mit komplexen oder reellen Koeffizienten benötigen. Wir werden hier alle erforderlichen Eigenschaften präzise formulieren, bezüglich ihrer Beweise sei der Leser jedoch auf den zweiten Band der „Enzyklopädie" verwiesen.[1])

Wir haben gesehen, daß sich jede komplexe Zahl eindeutig in der Form

$$a + bi = a \cdot 1 + bi$$

darstellen läßt (§ 28, Satz 1), d. h., jede komplexe Zahl läßt sich aus 1 und i linear mit reellen Koeffizienten a und b kombinieren. Nachdem die komplexen Zahlen in der Mathematik weite Anerkennung gefunden hatten, entstand natürlich die Frage, ob man nicht Zahlen angeben könnte, die noch allgemeiner als die komplexen Zahlen sind und die sich sämtlich durch n dieser Zahlen mit beliebigen reellen Koeffizienten linear ausdrücken lassen. In der Mitte des 19. Jahrhunderts konstruierte der englische Mathematiker HAMILTON derartige Zahlen für $n = 4$, die er Quaternionen nannte. Jedoch mußte hierbei auf die Gültigkeit des kommutativen Gesetzes der Multiplikation verzichtet werden. Später wurde gezeigt, daß dies keineswegs ein Zufall ist: Der Körper der reellen Zahlen (für $n = 1$) und der Körper der komplexen Zahlen (für $n = 2$) sind die einzigen kommutativen Körper der genannten Art.

Mit Rücksicht auf die Quaternionen und noch allgemeinerer Systeme, die in der modernen Algebra eine wichtige Rolle spielen, wollen wir in diesem Paragraphen unter einem Ring eine allgemeinere Bildung als in § 7 verstehen. Wir wollen nämlich auf die Gültigkeit des kommutativen Gesetzes der Multiplikation (§ 7, Definition 1, IV.) verzichten. Dann muß man allerdings an Stelle des bisherigen distributiven Gesetzes (§ 7, VI.) die beiden Forderungen

VI'. $$(a + b)\,c = ac + bc, \qquad c(a + b) = ca + cb$$

stellen. Entsprechend soll auch der Begriff des Körpers verallgemeinert werden. Dabei haben wir an die Stelle des Gesetzes der Umkehrbarkeit der Multiplikation (§ 8, Definition 1, VII.) die schärfere Forderung

VII'. *Für jedes $a \neq 0$ und für jedes b besitzt jede der Gleichungen $ax = b$ und $ya = b$ eine Lösung*

zu setzen. Im Unterschied zu den Ringen hat sich für derartige Körper eine besondere Terminologie eingebürgert: Man nennt eine Menge K, in der eine Addition und eine Multiplikation erklärt sind, die die Bedingungen I. bis III. und V. aus § 7 sowie die Bedingungen VI'. und VII'. erfüllen, und die mehr als ein Element enthält, einen *Schiefkörper*.

Offensichtlich bilden die von Null verschiedenen Elemente eines Schiefkörpers bezüglich der Multiplikation eine (im allgemeinen nichtkommutative) Gruppe. Daher gibt es auch im Schiefkörper — genau wie in einem kommutativen Körper — zu jedem von Null verschiedenen Element ein eindeutig bestimmtes inverses.

Definition 1. *Eine Menge R heißt ein n-dimensionaler Vektorraum über einem gegebenen Körper K, wenn zwischen den Elementen von R ein Addition erklärt ist, bezüglich der R eine kommutative Gruppe* (§ 6, Definition 2) *bildet, und zwischen*

[1]) Siehe EdEM Band 2, A. I. USKOW, Vektorräume und lineare Transformationen.

den Elementen von R und den Elementen von K eine Multiplikation erklärt ist, die den folgenden Bedingungen genügt:

1.) *Das Produkt ax eines beliebigen Elementes a aus K und eines beliebigen Vektors x aus R gehört wieder zu R.*

2.) $a(x+y) = ax + ay$ *für jedes a aus K und jedes x und y aus R.*

3.) $(a+b)\,x = ax + bx$ *für jedes a und b aus K und jedes x aus R.*

4.) $(ab)\,x = a(bx)$ *für jedes a und b aus K und jedes x aus R.*

5.) *In R gibt es n Elemente $e_1, \ldots, e_n$ (eine Basis von R) derart, daß sich jedes x aus R eindeutig in der Form*

$$x = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$$

mit Elementen $a_1, a_2, \ldots, a_n$ aus K darstellen läßt. Die Elemente $a_1, a_2, \ldots, a_n$ heißen die Komponenten des Vektors x.

Aus dieser Definition folgt unmittelbar, daß der Addition zweier Vektoren die Addition ihrer Komponenten und der Multiplikation eines Vektors mit einem Element aus K die Multiplikation seiner sämtlichen Komponenten mit diesem Element entspricht. Aus diesem Grunde kann ein n-dimensionaler Vektorraum über einem Körper K auch als Gesamtheit der n-Tupel $(a_1, a_2, \ldots, a_n)$ von Elementen aus K bezüglich der so erklärten Addition und Multiplikation mit Elementen aus K definiert werden.

Definition 2. *Ein n-dimensionaler Vektorraum R über dem Körper K heißt eine Algebra (oder ein hyperkomplexes System) vom Rang n über K, wenn zwischen den Elementen aus R neben der Addition eine Multiplikation erklärt ist und R bezüglich der Addition und dieser Multiplikation einen (nicht notwendig kommutativen) Ring bildet, wobei die Multiplikation in R mit der Multiplikation der Elemente von R mit Elementen aus K durch die folgende Bedingung verknüpft sein soll:*

6.) $$(ax)\,y = x(ay) = a(xy)$$

für jedes a aus K und jedes x und y aus R. Wenn hierbei R sogar ein Schiefkörper ist, so spricht man von einer Divisionsalgebra.

Aus 6.) folgt, daß

$$(ax)\,(by) = (ab)\,(xy) \tag{1}$$

für beliebige Elemente a und b aus K und beliebige Elemente x und y aus R gilt.

Hieraus ergibt sich auf Grund der distributiven Gesetze VI'., daß das Produkt beliebiger Elemente aus R bereits durch die Produkte der Basiselemente festgelegt ist. Wenn nämlich

$$x = \sum_{i=1}^{n} a_i e_i, \quad y = \sum_{j=1}^{n} b_j e_j$$

ist, so ist

$$xy = \left(\sum_{i=1}^{n} a_i e_i\right)\left(\sum_{j=1}^{n} b_j e_j\right) = \sum_{i,j=1}^{n} (a_i b_j)\,(e_i e_j). \tag{2}$$

Dabei können die Produkte $e_i e_j$ als Elemente aus R auf Grund von 5.) linear durch die Basiselemente in der Form

$$e_i e_j = \sum_{k=1}^{n} c_{ijk} e_k \tag{3}$$

ausgedrückt werden, wobei die c_{ijk} gewisse Elemente aus K sind, die ihrerseits die Produkte $e_i e_j$ eindeutig bestimmen.

Durch (2) und (3) wird andererseits in R eine Multiplikation erklärt, die jedenfalls die distributiven Gesetze VI'. und die Bedingung 6.) erfüllt. Damit diese Multiplikation das assoziative Gesetz für alle Elemente aus R erfüllt, genügt es offensichtlich, daß diese Multiplikation für die Basiselemente assoziativ ist. Dies führt auf die Bedingungen

$$(e_i e_j)\, e_k = e_i (e_j e_k) \quad (i, j, k = 1, 2, \ldots, n). \tag{4}$$

Ersetzt man in (4) die Produkte der Basiselemente durch ihre Ausdrücke (3), so erhält man Bedingungen für die c_{ijk}, die notwendig und hinreichend für die Gültigkeit des assoziativen Gesetzes für die durch (2) und (3) in R definierte Multiplikation sind. Somit ist eine Algebra vom Range n über einem Körper K vollständig festgelegt durch Vorgabe des Körpers K, des Ranges n und von n^3 Elementen c_{ijk} $(i, j, k = 1, 2, \ldots, n)$ aus K, die den Bedingungen (3) und (4) genügen, wobei $e_1, e_2, \ldots, e_n$ eine gegebene Basis des Raumes R ist. Die n^3 Elemente c_{ijk} aus K heißen die *Strukturkonstanten* der Algebra R.

Wir geben zunächst einige einfache Beispiele für den Begriff der Algebra an.

Beispiel 1. Der Körper D der reellen Zahlen bildet bezüglich der Addition und Multiplikation der reellen Zahlen einen Vektorraum über D mit dem Basiselement 1. Faßt man daneben das Produkt zweier Zahlen x und y auch als Produkt der Vektoren x und y auf, so wird D zu einer Algebra vom Range 1 über D. Mit dem Basiselement 1 erhält man $c_{111} = 1$ als einzige Strukturkonstante. Wählt man dagegen eine Zahl $a \neq 0$ als Basiselement, so folgt aus $a^2 = a \cdot a$, daß $c'_{111} = a$ die neue Strukturkonstante ist. Offensichtlich ist D eine Divisionsalgebra, die zudem kommutativ ist.

Beispiel 2. Der Körper K der komplexen Zahlen bildet bezüglich der Addition von komplexen Zahlen und der Multiplikation einer komplexen Zahl mit einer reellen Zahl einen zweidimensionalen Vektorraum über dem Körper D der reellen Zahlen mit den Basiselementen 1 und i, da sich jede komplexe Zahl eindeutig in der Form $a \cdot 1 + bi$ mit reellem a und b darstellen läßt. Man erhält eine Algebra vom Range 2 über D, wenn als Produkt zweier Vektoren x und y das übliche Produkt $x \cdot y$ von komplexen Zahlen genommen wird. Aus den Multiplikationsregeln für die Basiselemente $(1 \cdot 1 = 1,\ i \cdot i = -1,\ 1 \cdot i = i \cdot 1 = i)$ findet man als Werte für die $2^3 = 8$ Strukturkonstanten zu der betrachteten Basis:

$$c_{111} = 1, \quad c_{112} = 0, \quad c_{121} = 0, \quad c_{122} = 1,$$
$$c_{211} = 0, \quad c_{212} = 1, \quad c_{221} = -1, \quad c_{222} = 0.$$

Offensichtlich ist K eine kommutative Divisionsalgebra.

Beispiel 3. Der Schiefkörper der Quaternionen. Es gibt noch eine weitere Divisionsalgebra über dem Körper D der reellen Zahlen, und zwar ist sie vom Range 4. Es ist dies die Quaternionenalgebra Q.

Wir nehmen der Einfachheit halber an, daß Q den Körper der reellen Zahlen enthält. Dann können wir als erstes Basiselement die Zahl 1 nehmen, während wir die übrigen Basiselemente durch i, j, k bezeichnen wollen. Gemäß dieser Verabredung stellt sich jede Quaternion in der Form

$$q = a + bi + cj + dk \tag{5}$$

dar, wobei a, b, c, d beliebige reelle Zahlen sind. Zur vollständigen Beschreibung der Algebra genügt es, die Multiplikationstabelle für die Basiselemente anzugeben. Wir setzen

$$\left.\begin{array}{lll} i^2 = j^2 = k^2 = -1, & & \\ ij = k, \quad jk = i, & & ki = j, \\ ji = -k, kj = -i, & & ik = -j \end{array}\right\} \tag{6}$$

und verlangen, daß die Zahl 1 die üblichen Multiplikationseigenschaften haben soll, daß also $1 \cdot 1 = 1$, $1 \cdot i = i \cdot 1 = i$ usw. ist.

Wegen

$$ij = -ji, \quad jk = -kj, \quad ki = -ik$$

ist die Algebra Q nicht kommutativ.

Es bleibt zu zeigen, daß die Multiplikation für die Basiselemente assoziativ ist, daß also (4) gilt. Da die Festsetzung (6) symmetrisch in i, j, k ist, braucht man von jedem der Fälle, daß in dem Produkt $(e_i e_j) e_k$ alle Elemente untereinander gleich sind, daß zwei von ihnen gleich sind und daß alle drei voneinander verschieden sind, jeweils nur einen zu betrachten, wobei der Fall, daß eines der Elemente gleich 1 ist, auch noch außer Betracht bleiben kann. Es genügt also nachzuprüfen, daß

$$(ii)i = i(ii), \quad (ii)j = i(ij), \quad (ij)i = i(ji),$$
$$(ji)i = j(ii), \quad (ij)k = i(jk)$$

ist. Die Durchführung dieser Rechnungen möge dem Leser überlassen bleiben.

Wir wollen nun zeigen, daß die Algebra Q ein Schiefkörper ist. Aus der Tatsache, daß die Zahl 1 bei der Multiplikation mit den Basiselementen die üblichen Eigenschaften besitzt, folgt unter Benutzung von 2.) und 6.) auf Grund der Darstellung (5) der Quaternionen q, daß $1 \cdot q = q \cdot 1 = q$ ist für jede Quaternion q, d. h., die Zahl 1 ist Einselement des Ringes Q.

Die Quaternion

$$\bar{q} = a - bi - cj - dk$$

heißt *konjugiert* zur Quaternion

$$q = a + bi + cj + dk.$$

Aus der Multiplikationstabelle (6) folgt unter Benutzung der distributiven Gesetze und der Beziehung (1), daß

$$q\bar{q} = \bar{q}q = a^2 + b^2 + c^2 + d^2$$

ist.

Die Zahl

$$N(q) = a^2 + b^2 + c^2 + d^2$$

heißt die *Norm* der Quaternion $q = a + bi + cj + dk$. Offenbar ist $N(q) = N(\bar{q})$ und $N(q) \geq 0$, wobei $N(q) > 0$ ist für alle $q \neq 0$. Da nun für jede Quaternion $q \neq 0$

$$q \cdot \frac{\bar{q}}{N(q)} = \frac{\bar{q}}{N(q)} \cdot q = 1$$

ist, besitzt jede Quaternion $q \neq 0$ ein inverses Element

$$q^{-1} = \frac{\bar{q}}{N(q)}.$$

Hieraus folgt bereits (siehe § 6), daß die Menge der von Null verschiedenen Quaternionen bezüglich der Multiplikation eine Gruppe bildet. Dann ist aber der Ring Q ein Schiefkörper und mithin eine Divisionsalgebra über dem Körper D der reellen Zahlen.

Bezüglich weiterer Eigenschaften der Quaternionen, besonders bezüglich ihrer geometrischen Darstellung sei der Leser auf das Buch von E. CÉSARO [14] verwiesen.

Wir haben bisher drei Divisionsalgebren über dem Körper der reellen Zahlen kennen gelernt, und zwar jeweils eine vom Range 1, 2 und 4. Es gilt nun der bemerkenswerte Satz, daß es keine weiteren Algebren dieser Art gibt, genauer, daß jede Divisionsalgebra über dem Körper der reellen Zahlen einer dieser drei Algebren isomorph ist.

Zum Beweis dieses Satzes machen wir zunächst einige allgemeine Bemerkungen über Algebren über einem beliebigen Körper. Sollten einem Leser die abstrakten Ausführungen für beliebige Körper Mühe bereiten, so möge er sich am Körper der reellen Zahlen orientieren.

Bemerkung 1. *Jedes Element x einer Algebra R über einem Körper K ist Wurzel eines Polynoms mit Koeffizienten aus K, die nicht alle gleich Null sind.*

In der Tat: Wenn n der Rang von R über K ist, so sind je $n+1$ Elemente aus R linear abhängig. Insbesondere sind also die Elemente $x, x^2, \ldots, x^{n+1}$ linear abhängig, d. h.; es gibt in K Elemente $a_1, a_2, \ldots, a_{n+1}$, die nicht sämtlich gleich Null sind und für die $a_1 x + a_2 x^2 + \cdots + a_{n+1} x^{n+1} = 0$ ist. Das besagt aber nichts anderes, als daß x Wurzel des Polynoms $a_1 z + a_2 z^2 + \cdots + a_{n+1} z^{n+1}$ mit Koeffizienten aus K ist.

Bemerkung 2. *Es seien $f(z), g(z), h(z)$ Polynome in z mit Koeffizienten aus einem gewissen Körper K.*[1]) *Dann bleiben die Gleichungen $f(z) + g(z) = h(z)$ und $f(z) \cdot g(z) = h(z)$ richtig, wenn man die Unbestimmte z durch ein beliebiges Element x aus einer Algebra R über dem Körper K ersetzt.*

In der Tat folgt aus 3.) unmittelbar, daß $f(x) + g(x) = h(x)$ ist. Da $x^m \cdot x^n = x^{m+n}$ ist [§ 6, (3)], ergibt sich entsprechend auf Grund der Gültigkeit der distributiven Gesetze VI'. aus (1) und 3.), daß $f(x) \cdot g(x) = h(x)$ ist.

Bemerkung 3. *Wenn eine Algebra R über einem Körper K ein Einselement e besitzt (speziell also, wenn R eine Divisionsalgebra ist), so ist R isomorph einer Algebra R', die den Körper K enthält.*

Aus (1) und 3.) folgt nämlich, daß

$$a_1 e + a_2 e = (a_1 + a_2) e,$$

$$(a_1 e)(a_2 e) = (a_1 a_2)(ee) = (a_1 a_2) e$$

ist. Mithin ist die Menge K' der Elemente der Algebra R, die die Form ae haben, isomorph dem Körper K (§ 9, Definition 2). Also gibt es nach Satz 2 aus § 9 (für dessen Gültigkeit die Kommutativität der Multiplikation unwesentlich ist) einen Ring R', der den Körper K enthält und dem Ring R isomorph ist. Wir definieren dann das Produkt $a\,x'$ eines Elementes x' aus R' mit einem Element a aus K als das-

[1]) Die Definition eines Polynoms mit Koeffizienten aus einem beliebigen Körper K wie auch die Definition der Addition und Multiplikation derartiger Polynome entsprechen völlig den üblichen Definitionen für Polynome, deren Koeffizienten Zahlen sind. Der einzige Unterschied besteht darin, daß die Koeffizienten eben keine Zahlen, sondern Elemente aus K sind.

jenige Element aus R', das dem Produkt ax aus R bei obigem Isomorphismus entspricht, wenn x das x' entsprechende Element aus R ist. Man sieht leicht ein, daß damit R' zu einer Algebra über K wird, in der zudem für die in R' enthaltenen Körperelemente x' das zuletzt definierte Produkt ax' mit dem Produkt in K übereinstimmt. Daher ist das Einselement von K zugleich Einselement von R'.

Bemerkung 4. *Jede Divisionsalgebra ist nullteilerfrei* (§ 7, Definition 2).

Dies beweist man genau wie in einem kommutativen Körper: Wenn nämlich $xy = 0$ und $x \neq 0$ ist, so ergibt sich nach Multiplikation dieser Gleichung mit x^{-1}, daß $y = 0$ ist.

Satz 1. *Jede kommutative Divisionsalgebra R über dem Körper D der reellen Zahlen ist entweder dem Körper D der reellen Zahlen oder dem Körper K der komplexen Zahlen isomorph und damit vom Rang* 1 *oder* 2 *über D. Ist umgekehrt R eine Divisionsalgebra vom Rang* 1 *oder* 2 *über dem Körper D der reellen Zahlen, so ist R dem Körper der reellen oder dem Körper der komplexen Zahlen isomorph und damit kommutativ.*

Beweis. a) Es sei R eine Divisionsalgebra über dem Körper D der reellen Zahlen, die den Körper D enthält, die aber von D verschieden ist. Wir wollen zeigen, daß es zu jedem Element x aus R, das nicht zu D gehört, reelle Zahlen a und b gibt derart, daß $a \neq 0$ ist und das Element $i = ax + b$ die Eigenschaft $i^2 = -1$ besitzt.

Für das Folgende sei betont, daß wir hierbei nicht von der Kommutativität der Algebra R Gebrauch machen.

Auf Grund von Bemerkung 1 ist x Wurzel eines Polynoms $f(z)$ mit reellen Koeffizienten, die nicht sämtlich Null sind. Es ist nun bekannt, daß sich jedes derartige Polynom als Produkt aus Polynomen ersten und zweiten Grades mit reellen Koeffizienten darstellen läßt.[1]) Es sei also

$$f(z) = f_1(z) f_2(z) \cdots f_k(z)$$

eine solche Zerlegung. Nach Bemerkung 2 ist dann

$$f(x) = f_1(x) f_2(x) \cdots f_k(x).$$

Da nun $f(x) = 0$ ist und nach Bemerkung 4 die Divisionsalgebra R keine Nullteiler besitzt, muß es ein $l \leq k$ geben, für das $f_l(x) = 0$ ist. Wäre nun x Wurzel eines Polynoms ersten Grades, also etwa des Polynoms $z - c$, so wäre $x - c = 0$ und mithin $x = c$, d. h., es wäre x Element von D, was unserer Voraussetzung widerspricht. Folglich muß x Wurzel eines Polynoms zweiten Grades sein, d. h., es muß reelle Zahlen p und q geben derart, daß $x^2 + px + q = 0$ und $\frac{p^2}{4} - q < 0$ ist, letzteres, weil anderenfalls x Wurzel eines Polynoms ersten Grades wäre, was nicht der Fall sein kann. Setzt man nun

$$\frac{p^2}{4} - q = -t^2$$

mit reellem t, so ist

$$\left(x + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q = -t^2,$$

---

[1]) Siehe EdEM Band 2, L. J. Okunjew, Der Ring der Polynome und der Körper der rationalen Funktionen, Kap. I, § 6.

woraus sich nach Division durch t^2 die Gleichung

$$\left(\frac{x+\frac{p}{2}}{t}\right)^2 = -1$$

ergibt. Setzen wir schließlich noch $a = \frac{1}{t}$ und $b = \frac{p}{2t}$, so erhalten wir in $i = ax + b$ ein Element aus R, das der Bedingung $i^2 = -1$ genügt, was zu beweisen war.

b) Es sei R eine kommutative Divisionsalgebra über dem Körper D der reellen Zahlen, die den Körper D enthält. Wenn $R \neq D$ ist, so gibt es in R ein Element i, für das $i^2 = -1$ ist. Daher gehört i nicht zu D, und die Elemente 1 und i sind linear unabhängig. Es sei ferner R_0 die Menge aller Elemente x aus R, die sich in der Form $x = a + bi$ mit reellem a und b darstellen lassen. Offenbar ist R_0 eine Algebra vom Rang 2 über D und dem Körper K der komplexen Zahlen isomorph. Wir wollen nun zeigen, daß $R = R_0$ ist.

Wenn wir in $x = a + bi$ für b die Zahl Null nehmen, so wird $x = a$, d. h., R_0 enthält den Körper der reellen Zahlen. Es sei nun x ein Element der Algebra R, das nicht zu D gehört. Nach dem unter a) Bewiesenen gibt es reelle Zahlen c und d derart, daß $c \neq 0$ ist und das Element $j = cx + d$ die Eigenschaft $j^2 = -1$ besitzt. Da nach Voraussetzung in R das kommutative Gesetz der Multiplikation gilt, ist $ij = ji$ und mithin

$$(i + j)(i - j) = i^2 - ij + ji - j^2 = 0.$$

Da nun gemäß Bemerkung 4 die Algebra R keine Nullteiler besitzt, so muß $i - j = 0$ oder $i + j = 0$, also $j = \pm i$ sein. Aus dem Gesagten ergibt sich unmittelbar, daß dann das Element

$$x = -\frac{d}{c} + \frac{1}{c} j = -\frac{d}{c} \pm \frac{1}{c} i$$

zu R_0 gehört und daher $R = R_0$ ist.

Damit ist gezeigt, daß R entweder gleich dem Körper D der reellen Zahlen ist oder dem Körper K der komplexen Zahlen isomorph ist.

Wenn schließlich R eine beliebige kommutative Divisionsalgebra über dem Körper D der reellen Zahlen ist, so gibt es nach Bemerkung 3 eine zu R isomorphe Algebra R', die den Körper der reellen Zahlen enthält (und auf Grund der Isomorphie ebenfalls kommutativ und Divisionsalgebra ist). Daraus folgt aber unmittelbar, daß dann R entweder dem Körper D der reellen Zahlen oder dem Körper K der komplexen Zahlen isomorph ist.

c) Es sei R eine beliebige Divisionsalgebra vom Rang 1 über dem Körper D der reellen Zahlen, die den Körper D enthält. Wählt man als Basiselement die Zahl 1, so ergibt sich, daß jedes Element x aus R in der Form $x = a \cdot 1 = a$ mit reellem a dargestellt werden kann und mithin $R = D$ ist. Auf Grund der Bemerkung 3 ist dann aber jede Divisionsalgebra, die vom Rang 1 über D ist, isomorph dem Körper D der reellen Zahlen.

d) Es sei R eine beliebige Divisionsalgebra vom Rang 2 über dem Körper D der reellen Zahlen, die den Körper D enthält. Dann ist sicher $R \neq D$. Nach dem unter a) Bewiesenen gibt es in R ein Element i mit $i^2 = -1$. Dann sind die Elemente 1 und i jedenfalls linear unabhängig. Andernfalls gäbe es nämlich von Null verschiedene reelle Zahlen a_1 und a_2 derart, daß $a_1 \cdot 1 + a_2 i = 0$ ist (da R als Divisionsalgebra

nullteilerfrei ist, folgt aus $a_1 = 0$, daß $a_2 = 0$ ist, und umgekehrt). Dann wäre aber $i = -\frac{a_1}{a_2}$, d. h. i Element von D, was wegen $i^2 = -1$ nicht der Fall sein kann (§ 10, Satz 2). Da nun in einem n-dimensionalen Vektorraum je n linear unabhängige Vektoren eine Basis bilden[1]) und R eine Algebra vom Rang 2 sein sollte, bilden die Elemente 1 und i aus R eine Basis von R, d. h., jedes Element x aus R ist eindeutig in der Form $x = a + bi$ mit reellem a und b darstellbar. Sind schließlich $x = a + bi$ und $y = c + di$ beliebige Elemente aus R, so ergibt sich auf Grund von 3.) bzw. auf Grund der distributiven Gesetze VI'., 3.) und 6.), daß

$$x + y = (a + c) + (b + d)\,i,$$
$$xy = (ac - bd) + (ad + bc)\,i$$

ist. Damit ist gezeigt, daß R dem Körper K der komplexen Zahlen isomorph ist. Nach Bemerkung 3 gilt dies genauso für beliebige Divisionsalgebren vom Rang 2 über dem Körper D der reellen Zahlen, da jede derartige Algebra isomorph einer Algebra ist, die den Körper D enthält.

Damit ist Satz 1 in allen Teilen bewiesen.

Wenn man von der Gültigkeit des kommutativen Gesetzes der Multiplikation absieht, so gibt es noch genau eine weitere Divisionsalgebra über dem Körper der reellen Zahlen, nämlich die Quaternionenalgebra. Es gilt also

Satz 2. (Satz von FROBENIUS) *Jede Divisionsalgebra R über dem Körper der reellen Zahlen ist isomorph dem Körper D der reellen Zahlen, dem Körper K der komplexen Zahlen oder dem Schiefkörper Q der Quaternionen und mithin vom Rang* 1, 2 *oder* 4 *über D.*

Beweis. Es sei R eine beliebige Divisionsalgebra vom Rang n über dem Körper D der reellen Zahlen, die den Körper D enthält. Falls $n = 1$ ist, so ist auf Grund von Satz 1 die Algebra R dem Körper D der reellen Zahlen isomorph. Wenn $n = 2$ ist, so ist R dem Körper K der komplexen Zahlen isomorph. Es möge also der Rang n von R größer als 2 sein. Dann ist sicher $R \neq D$. Nach dem unter a) Bewiesenen gibt es daher in R ein Element i, für das $i^2 = -1$ ist. Genau wie unter d) zeigt man dann, daß 1 und i linear unabhängig sind. Da nun der Rang größer als 2 ist, gibt es in R ein Element x, das sich nicht in der Form $a + bi$ mit reellem a und b darstellen läßt. Zu diesem gibt es, wie unter a) bewiesen wurde, reelle Zahlen a' und b' derart, daß $a' \neq 0$ ist und das Element $x_1 = a'x + b'$ die Eigenschaft $x_1^2 = -1$ besitzt. Hierbei ist auch das Element x_1 nicht in der Form $a + bi$ mit reellem a und b darstellbar, weil es sonst auch für

$$x = \frac{1}{a'}\,x_1 - \frac{b'}{a'}$$

eine derartige Darstellung gäbe. Mithin sind sogar die Elemente $1, i, x_1$ linear unabhängig: Wenn nämlich

$$a_1 \cdot 1 + a_2 i + a_3 x_1 = 0$$

mit gewissen reellen Zahlen a_1, a_2, a_3 gilt, so muß $a_3 = 0$ sein (weil sich sonst x_1 linear aus 1 und i kombinieren ließe); dann muß aber wegen der bereits bewiesenen linearen Unabhängigkeit von 1 und i auch $a_1 = a_2 = 0$ sein. Genau wie unter a)

[1]) Siehe EdEM Band 2, A. I. USKOW, Vektorräume und lineare Transformationen.

überzeugt man sich davon, daß die Elemente $i + x_1$ und $i - x_1$ Wurzeln quadratischer Gleichungen mit reellen Koeffizienten sind, woraus sich

$$(i + x_1)^2 = p(i + x_1) + q,$$
$$(i - x_1)^2 = r(i - x_1) + s$$

ergibt. Daher ist

$$\left.\begin{aligned} -2 + ix_1 + x_1 i &= p(i + x_1) + q, \\ -2 - ix_1 - x_1 i &= r(i - x_1) + s \end{aligned}\right\} \tag{7}$$

mit gewissen reellen Zahlen p, q, r, s.

Addiert man diese beiden Gleichungen, so erhält man

$$-4 = (p + r)i + (p - r)x_1 + (q + s),$$

woraus sich wegen der linearen Unabhängigkeit von $1, i, x_1$

$$p + r = 0, \quad p - r = 0, \quad \text{d. h.} \quad p = r = 0$$

ergibt. Hieraus und aus (7) folgt, daß

$$ix_1 + x_1 i = 2t \tag{8}$$

ist, wobei $t = \frac{1}{2}(q + 2)$ eine reelle Zahl ist. Wir setzen dann $x_2 = x_1 + ti$. Auch die Elemente $1, i, x_2$ sind linear unabhängig, weil anderenfalls die Elemente $1, i, x_1$ linear abhängig wären. Aus (8) folgt ferner unmittelbar, daß

$$x_2^2 = -1 + t(ix_1 + x_1 i) - t^2 = t^2 - 1$$

ist. Dabei ist $t^2 - 1$ eine negative reelle Zahl. Wäre nämlich $t^2 - 1 > 0$, so gäbe es eine reelle Zahl u mit $t^2 - 1 = u^2$; dann wäre aber

$$(x_2 + u)(x_2 - u) = x_2^2 - u^2 = 0$$

und mithin $x_2 = \pm u$ eine reelle Zahl, was der linearen Unabhängigkeit von $1, i, x_2$ widerspricht.

Wir können also $x_2^2 = -c^2$ mit reellem c setzen. Außerdem sei $j = \frac{1}{c}x_2$. Dann ist $j^2 = -1$, und die Elemente $1, i, j$ sind linear unabhängig, weil ja $1, i, x_2$ linear unabhängig waren. Weiter gilt wegen (8)

$$ij + ji = i \cdot \frac{1}{c}(x_1 + ti) + \frac{1}{c}(x_1 + ti)\, i$$
$$= \frac{1}{c}(ix_1 + ti^2 + x_1 i + ti^2) = \frac{1}{c}(2t - t - t) = 0,$$

d. h., es ist

$$ij = -ji. \tag{9}$$

Schließlich setzen wir $ij = k$ und zeigen, daß sich k nicht linear durch $1, i, j$ ausdrücken läßt. Wäre nämlich $k = a + bi + cj$ mit gewissen reellen Zahlen a, b, c, so müßte, wie sich nach linksseitiger Multiplikation dieser Gleichung mit i ergäbe,

$$ik = i(ij) = -j = ai - b + ck = ai - b + c(a + bi + cj)$$

und mithin

$$(ac - b) + (a + bc)i + (c^2 + 1)j = 0$$

sein. Da aber $1, i, j$ linear unabhängig sind, müßte $c^2 + 1 = 0$ also $c^2 = -1$ sein, im Widerspruch dazu, daß c reell ist. Hieraus folgt aber unmittelbar, daß die Elemente $1, i, j, k$ linear unabhängig sind und daß der Rang von R über D nicht kleiner als 4 ist.

Wir wollen nun zunächst zeigen, daß für die Elemente i, j, k die Multiplikationstabelle (6) erfüllt ist.

Wir wissen bereits, daß $i^2 = j^2 = -1$ und $ij = k$ ist. Wegen (9) ist ferner

$$k^2 = (ij)(ij) = i(ji)j = i(-ij)j = -1,$$

$$jk = j(ij) = j(-ji) = i, \quad ki = (ij)i = (-ji)i = j,$$

$$ji = -ij = -k, \quad kj = (ij)j = -i, \quad ik = i(ij) = -j.$$

Somit gelten also alle Beziehungen aus (6). Am Schluß der Bemerkung 3 haben wir ferner darauf hingewiesen, daß die Zahl 1 Einselement der Algebra R ist. Daher bildet die Menge Q der Elemente x aus R, die sich in der Form $x = a + bi + cj + dk$ mit reellem a, b, c, d darstellen lassen, den Schiefkörper der Quaternionen.

Wir wollen schließlich zeigen, daß $R = Q$ ist. Dazu nehmen wir an, es wäre $R \neq Q$. Dann gibt es in R ein Element y, das nicht zu Q gehört. Nach dem unter a) Bewiesenen gibt es zu y reelle Zahlen a und b derart, daß $a \neq 0$ ist und das Element $l = ay + b$ die Eigenschaft $l^2 = -1$ besitzt. Dieses Element l muß ebenfalls außerhalb Q liegen, weil sonst $y = \frac{1}{a} l - \frac{b}{a}$ zu Q gehören würde. Ähnlich wie unter (8) findet man, daß es reelle Zahlen a, b, c gibt, für die

$$il + li = a, \quad jl + lj = b, \quad kl + lk = c$$

ist. Hieraus folgt, daß

$$lk = l(ij) = (li)j = (a - il)j = aj - i(lj)$$
$$= aj - i(b - jl) = aj - bi + kl = aj - bi + c - lk,$$

also

$$2lk = c - bi + aj$$

ist. Multipliziert man diese Gleichung von rechts mit k, so ergibt sich

$$-2l = ai + bj + ck,$$

d. h., l ist Element von Q, was aber nach dem Gesagten nicht der Fall sein dürfte. Also ist unsere Annahme falsch, und es ist $R = Q$.

Mithin ist also $R = D$, $R = K$ oder $R = Q$.

Wenn R eine beliebige Divisionsalgebra über dem Körper D der reellen Zahlen ist, so gibt es eine zu R isomorphe Algebra R' (die offensichtlich auch Divisionsalgebra ist), die den Körper D der reellen Zahlen enthält und nach dem Bewiesenen gleich D, K oder Q ist. Dann ist aber R isomorph dem Körper D der reellen Zahlen, dem Körper K der komplexen Zahlen oder dem Schiefkörper Q der Quaternionen, was zu beweisen war.

Ersetzt man in den Beispielen 1 bis 3 den Körper D der reellen Zahlen durch den Körper Γ der rationalen Zahlen, so erhält man drei Divisionsalgebren über dem Körper der rationalen Zahlen, nämlich den Körper Γ selbst, den Körper der komplexen Zahlen der Form $a + bi$ mit rationalem a und b (den sogenannten GAUSSschen Zahlkörper) und den Schiefkörper der rationalen Quaternionen der Form $a + bi + cj + dk$ mit rationalem a, b, c und d.

Ersetzt man in den Beispielen 1 bis 3 den Körper D der reellen Zahlen durch den Körper K der komplexen Zahlen, so erhält man in Beispiel 1 den Körper K selbst, während man in den Beispielen 2 und 3 Algebren erhält, die bereits keine Divisionsalgebren mehr sind. In der Tat: Auf Grund der Bemerkung 1 ist jedes Element x aus einer Algebra über K, die den Körper K enthält, Wurzel eines Polynoms $f(z)$ mit komplexen Koeffizienten. Nun weiß man, daß sich jedes Polynom mit komplexen Koeffizienten als Produkt aus Polynomen ersten Grades darstellen läßt, deren Koeffizienten gleichfalls komplexe Zahlen sind.[1]) Wenn nun R eine Divisionsalgebra über K ist, so findet man wie unter a), daß x Wurzel eines Polynoms ersten Grades mit komplexen Koeffizienten, also eine komplexe Zahl sein muß. Also ist dann $R = K$. Somit gilt:

Wenn R eine Algebra über dem Körper der komplexen Zahlen vom Rang $n > 1$ ist, so ist R keine Divisionsalgebra.

[1]) Siehe EdEM Band 2, L. J. Okunjew, Der Ring der Polynome und der Körper der rationalen Funktionen, Kap. I, § 8.

LITERATUR

[1] Александров, П. С., Введение в общую теорию множеств и функций (ALEXANDROFF, P. S., Einführung in die allgemeine Mengenlehre und die Theorie der reellen Funktionen), Moskau-Leningrad 1948.
Deutsche Übersetzung: ALEXANDROFF, P. S., Einführung in die Mengenlehre und die Theorie der reellen Funktionen, 4. Aufl., Berlin 1967.

[2] Лузин, Н. Н., Теория функций действительного переменного (LUSIN, N. N., Theorie der Funktionen einer reellen Veränderlichen), Moskau 1948.

[3] Окунев, Л. Я., Основы современной алгебры (OKUNJEW, L. J., Grundlagen der modernen Algebra), Moskau 1941.

[4] VAN DER WAERDEN, B. L., Algebra, Band I, 7. Aufl., Berlin-Göttingen-Heidelberg 1966.

[5] Проскуряков, И. В., Числа и многочлены (PROSKURJAKOW, I. W., Zahlen und Polynome), Moskau 1949.

[6] Шмидт, О. Ю., Абстрактная теория групп (SCHMIDT, O. J., Abstrakte Gruppentheorie), Moskau-Leningrad 1933. Nachdruck in Шмидт, О. Ю., Избранные труды (SCHMIDT, O. J., Ausgewählte Werke), Moskau 1959.

[7] Курош, А. Г., Теория групп (KUROSCH, A. G., Gruppentheorie), Moskau-Leningrad 1944 (2. Aufl., Moskau-Leningrad 1953).
Deutsche Übersetzung nach der 1. Auflage: KUROSCH, A. G., Gruppentheorie, Nachdruck, Berlin 1955.

[8] Ефимов, Н. В., Высшая геометрия (JEFIMOW, N. W., Höhere Geometrie), 2. Aufl., Moskau-Leningrad 1949.
Deutsche Übersetzung: EFIMOW, N. W., Höhere Geometrie, Berlin 1960.

[9] Костин, В. И., Основания геометрии (KOSTIN, W. I., Grundlagen der Geometrie), 2. Aufl., Moskau-Leningrad 1948.

[10] DEDEKIND, R., Stetigkeit und irrationale Zahlen, Braunschweig 1964 und Berlin 1967.

[11] Хинчин, А. Я., Восемь лекций по математическому анализу (CHINTSCHIN, A. J., Acht Vorlesungen über Analysis), Moskau-Leningrad 1943.

[12] Курош, А. Г., Курс высшей алгебры (KUROSCH, A. G., Kurs der höheren Algebra), 2. Aufl., Moskau-Leningrad 1950.

[13] Кузьмин, Р. О., и Д. К. Фаддеев, Алгебра и арифметика комплексных чисел (KUSMIN, R. O., und D. K. FADDEJEW, Algebra und Arithmetik der komplexen Zahlen), Moskau 1939.

[14] CESÀRO, E., Elementares Lehrbuch der algebraischen Analysis und der Infinitesimalrechnung, Leipzig 1904.

Der deutsche Leser sei noch verwiesen auf:

ASSER, G., Einführung in die höhere Mathematik, Fernstudienbriefe der Pädagogischen Hochschule Potsdam, 2. Aufl., Berlin 1965.

BACHMANN, F., Aufbau des Zahlensystems, Enzyklopädie der mathematischen Wissenschaften, Band I, 1, 3; 2. Aufl., Leipzig und Berlin 1939.

BALDUS, R., Nichteuklidische Geometrie, 3. Aufl., Berlin 1953.

BECK, H., Einführung in die Axiomatik der Algebra, Berlin und Leipzig 1926.

BIEBERBACH, L., Einführung in die analytische Geometrie, 4. Aufl., Bielefeld 1950.

–, Projektive Geometrie, Leipzig 1931.

CARATHÉODORY, C., Vorlesungen über reelle Funktionen, 2. Aufl., Leipzig und Berlin 1927.

FICHTENHOLZ, G. M., Differential- und Integralrechnung I, 3. Aufl., Berlin 1968 (Übersetzung aus dem Russischen).

FRAENKEL, A. A., Einleitung in die Mengenlehre, 3. Aufl., Berlin 1928.

GÖRKE, L., Mengen, Relationen, Funktionen, 2. Aufl., Berlin 1967.

HAHN, H., Reelle Funktionen I, Leipzig 1935.

HASSE, H., Höhere Algebra I, II, 5. bzw. 4. Aufl., Berlin 1963 bzw. 1962.

HASSE, M., Grundbegriffe der Mengenlehre und Logik, 3. Aufl., Leipzig 1967.

HAUPT, D., Mengenlehre – leicht verständlich. Eine Einführung in die Grundzüge der Mengenlehre, 2. Aufl., Leipzig 1967.

HAUPT, O., Einführung in die Algebra I, II, 3. bzw. 2. Aufl., Leipzig 1956 bzw. 1954.

HAUSDORFF, F., Grundzüge der Mengenlehre, Leipzig 1914.

–, Mengenlehre, 3. Aufl., Berlin und Leipzig 1935.

HEFFTER, L., Grundlagen und analytischer Aufbau der Geometrie, 2. Aufl., Leipzig 1950.

Hessenberg, G., Grundlagen der Geometrie, Berlin und Leipzig 1930.
Hilbert, D., Grundlagen der Geometrie, 7. Aufl., Leipzig und Berlin 1930, 9. Aufl., Stuttgart 1961.
Kamke, E., Mengenlehre, 4. Aufl., Berlin 1962.
Klaua, D., Allgemeine Mengenlehre, Berlin 1964.
Kochendörffer, R., Einführung in die Algebra, 3. Aufl., Berlin 1966.
Knopp, K., Elemente der Funktionentheorie, 7. Aufl., Berlin 1966.
–, Theorie und Anwendung der unendlichen Reihen, 5. Aufl., Berlin-Göttingen-Heidelberg 1965.
Landau, E., Grundlagen der Analysis, Leipzig 1930.
Lenz, H., Grundlagen der Elementarmathematik, 2. Aufl., Berlin 1966.
Liebmann, H., Nichteuklidische Geometrie, 3. Aufl., Berlin und Leipzig 1923.
Loewy, A., Lehrbuch der Algebra, 1. Teil, Leipzig 1915.
Lugowski, H., und H. J. Weinert, Grundzüge der Algebra, Teil I (3. Aufl.), II, III (2. Aufl.), Leipzig 1966, 1965 bzw. 1967.
von Mangoldt, H., und K. Knopp, Einführung in die höhere Mathematik I, II, III, 12. Aufl. (Nachdruck), Leipzig 1965, 1964 bzw. 1965.
Norden, A. P., Elementare Einführung in die Lobatschewskische Geometrie, Berlin 1958 (Übersetzung aus dem Russischen).
Perron, O., Algebra I, II, 3. Aufl., Berlin 1951.
Pickert, G., Analytische Geometrie, 6. Aufl., Leipzig 1967.
Prüfer, H., Projektive Geometrie, 2. Aufl., Leipzig 1953.
Rédei, L., Algebra; Erster Teil, Leipzig 1959.
Schröder, K. (Hrsg.), Mathematik für die Praxis I, 3. Aufl., Berlin 1966.
Speiser, A., Die Theorie der Gruppen von endlicher Ordnung, 4. Aufl., Basel 1956.
Vogel, A., Klassische Grundlagen der Analysis, Leipzig 1952.
Weber, H., und J. Wellstein, Enzyklopädie der Elementar-Mathematik I, II, III, 5., 3. bzw. 3. Aufl., Leipzig und Berlin 1934, 1915 bzw. 1924.
Zassenhaus, H., Lehrbuch der Gruppentheorie I, Leipzig und Berlin 1937 (Nachdr. Hamburg 1948).

A. J. CHINTSCHIN

DIE ELEMENTE DER ZAHLENTHEORIE

Kapitel I

TEILBARKEIT UND PRIMZAHLEN

§ 1. Einleitung

Vom frühen Altertum bis in unsere Tage ist ein großer Teil der aktuellen Probleme der Theorie der ganzen Zahlen direkt oder indirekt mit dem Begriff der Teilbarkeit verknüpft. Der Grund hierfür ist leicht einzusehen, ist doch von den vier Grundrechenarten nur die Division im Bereich der ganzen Zahlen nicht unbeschränkt ausführbar. Daher kann man vernünftigerweise nur bezüglich der Division die Frage stellen, unter welchen Bedingungen sie ein Ergebnis besitzt. Schon die elementarsten Begriffe der Teilbarkeitstheorie, wie der Begriff des größten gemeinsamen Teilers, des kleinsten gemeinsamen Vielfachen, der teilerfremden Zahlen, der absoluten Primzahlen u. dgl. m. haben bezüglich der anderen Grundrechenarten kein Analogon, da diese Rechenarten unbeschränkt ausführbar sind.

Die historische Entwicklung der Arithmetik zeigt, daß die Teilbarkeitstheorie der Zahlen, die anfänglich von den einfachsten und natürlichsten, mit dem Begriff der Teilbarkeit verbundenen Aufgaben ausging, sehr schnell und unvermeidlich zur Bildung komplizierterer Begriffe und zur Aufstellung neuer schwieriger Aufgaben kam. Es ist bekannt, daß es hier Aufgaben gibt, die in ihrer Natürlichkeit bereits dem Schüler einleuchten und von ihm begriffen werden und deren Lösung teilweise dem jahrhundertelangen Bemühen der größten Gelehrten versagt blieb. Hieraus erklärt sich, daß das Studium der ganzen Zahlen den Mathematikern aller Zeiten ein unerschöpfliches Forschungsfeld bot und daß die Untersuchung ihrer Probleme die hervorragendsten Geister anzog.

Bei diesen Forschungen nahm die russische und später die sowjetische wissenschaftliche Schule einen führenden Platz ein, den die sowjetische Schule auch heute noch innehat. Seit den Zeiten EULERs bis in unsere Tage liefert die russische Akademie der Wissenschaften glänzende Beispiele für neue tiefe Problemstellungen und wirksame Lösungsmethoden. Es genügt hier, die bedeutenden Schöpfungen der hervorragenden TSCHEBYSCHEWschen Schule der arithmetischen Forschung zu erwähnen, die in den letzten hundert Jahren in der Welt einmalig war und deren Tradition in den Händen der glänzenden Schule der sowjetischen Mathematiker es auch heute noch zu den weitest reichenden Erfolgen bringt.[1])

Wir wollen im folgenden in einem kurzen Abriß die Entwicklung einiger Stücke der Teilbarkeitslehre von der ältesten Zeit bis in unsere Tage verfolgen und dabei dort, wo es nötig ist, unsere Aufmerksamkeit den methodischen und pädagogischen Fragen zuwenden.

[1]) Die wichtigsten Etappen in der Entwicklung dieser Schule sind eingehend behandelt in dem Buch Б. Н. Делоне, Петербургская школа теории чисел, (B. N. DELAUNAY, Die Petersburger zahlentheoretische Schule), Verlag der Akad. d. Wiss. d. UdSSR, Moskau-Leningrad 1947.

§ 2. Die eindeutige Zerlegbarkeit einer Zahl in Primfaktoren

In den mannigfachen Zweigen der Teilbarkeitstheorie spielt der Satz von der eindeutigen Zerlegbarkeit der Zahlen in Primfaktoren eine hervorragende Rolle:

Fundamentalsatz. *Jede von 1 verschiedene natürliche Zahl läßt sich als Produkt von Primzahlen darstellen. Diese Darstellung ist, von der Reihenfolge der Faktoren abgesehen, eindeutig.*

Die letzte Behauptung besagt dabei folgendes: Wenn

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s$$

ist, wobei die p_i und q_j Primzahlen sind, so ist $s = r$, und die Zahlen q_j unterscheiden sich höchstens in der Reihenfolge ihrer Anordnung von den Zahlen p_i.

Bemerkung. Eine Zahl $p > 1$ heißt *Primzahl* (oder *absolut prim*), wenn sie keine anderen natürlichen Teiler als p und 1 besitzt. Alle anderen Zahlen, die größer als 1 sind, heißen *zusammengesetzt*. Die Zahl 1 nimmt insofern eine Sonderstellung ein, als sie weder zu den Primzahlen noch zu den zusammengesetzten Zahlen gerechnet wird. Bedauerlicherweise wurde noch bis vor kurzem in fast allen Schullehrbüchern die Zahl 1 zu den Primzahlen gezählt, und es gibt auch heute noch unter den Methodikern einflußreiche Anhänger dieser Tradition, obgleich ihre grobe Fehlerhaftigkeit wiederholt nachgewiesen wurde. Die Frage, ob die Zahl 1 eine Primzahl ist oder nicht, ist keineswegs — wie es zunächst scheinen könnte — eine Frage der Terminologie oder der Konvention. Wenn wir die Zahl 1 unter die Primzahlen aufnehmen, werden fast alle Sätze über Primzahlen einfach falsch. Wir wollen dies am Beispiel des oben formulierten Fundamentalsatzes zeigen: Wenn die Zahl 1 als Primzahl genommen wird, so besitzt z. B. die Zahl 5 die folgenden unendlich vielen Primzahlzerlegungen:

$$5 = 1 \cdot 5 = 1 \cdot 1 \cdot 5 = 1 \cdot 1 \cdot 1 \cdot 5 = \cdots.$$

Diese sind tatsächlich alle voneinander verschieden, einfach deshalb, weil die Anzahl der Faktoren in ihnen verschieden ist.

Wir wollen jetzt einen Beweis für den Fundamentalsatz geben, der in vielen Beziehungen von methodischem Interesse ist.

Die Möglichkeit einer derartigen Zerlegung ist sehr leicht zu beweisen. Dazu sei n eine beliebige natürliche Zahl > 1. Unter ihren Teilern gibt es dann Zahlen, die größer als 1 sind (z. B. n selbst). Es sei p_1 der kleinste von ihnen. Offensichtlich ist p_1 eine Primzahl; andernfalls besäße nämlich p_1 einen Teiler a mit $1 < a < p_1$, der aber als Teiler von p_1 auch Teiler von n wäre, was der Definition von p_1 widerspricht. Also ist $n = p_1 n_1$, wobei p_1 eine Primzahl ist. Ist $n_1 > 1$, so verfahren wir mit n_1 genauso wie eben mit n und stellen n_1 in der Form $n_1 = p_2 n_2$ dar, wobei p_2 eine Primzahl ist. Dann ist $n = p_1 p_2 n_2$. Wenn auch noch $n_2 > 1$ ist, so setzen wir den Prozeß in der angegebenen Weise fort. Da hierbei $n > n_1 > n_2 > \cdots$ ist, muß der Prozeß nach endlich vielen Schritten abbrechen, d. h., es muß nach endlich vielen Schritten ein gewisses $n_k = 1$ sein. Dann ist aber

$$n = p_1 p_2 \cdots p_k,$$

wobei $p_1, p_2, \ldots, p_k$ Primzahlen sind. Damit ist die Möglichkeit der Zerlegung einer vorgegebenen natürlichen Zahl $n > 1$ in Primfaktoren bereits bewiesen.

Wir müssen uns jetzt noch von der Eindeutigkeit einer derartigen Zerlegung überzeugen, was sich als eine bedeutend schwierigere Aufgabe erweist. Historisch

ist interessant, daß erst verhältnismäßig spät bemerkt wurde, daß diese Eindeutigkeitsaussage keineswegs trivial ist (und damit eines effektiven Beweises bedarf), nachdem man sie schon lange Zeit hindurch als eine offensichtliche Tatsache benutzt hatte. Anscheinend hat GAUSS als erster nachdrücklich darauf hingewiesen, daß die Unmöglichkeit zweier wesentlich verschiedener Zerlegungen ein und derselben Zahl in Primfaktoren nicht trivial ist, sondern eines strengen Beweises bedarf. Selbst z. B. ein so hervorragender Gelehrter wie LEGENDRE, der kurz vor GAUSS wirkte, hat das nicht bemerkt. Die weitere Entwicklung der Zahlentheorie zeigte, in welchem Maße GAUSS nicht nur vom formal-logischen, sondern auch vom inhaltlichen Standpunkt aus recht hatte. Im Verlaufe des 19. Jahrhunderts gingen nämlich die Mathematiker dazu über, die Gesetze der Teilbarkeit auch in Gebieten zu untersuchen, die komplizierter sind als der Bereich der natürlichen Zahlen, wie z. B. im Bereich der sogenannten ganzen algebraischen Zahlen. Die Gesetze, die hier gelten, erinnern zwar noch in vielem an die Gesetzmäßigkeiten im Bereich der natürlichen Zahlen, es zeigen sich jedoch auch gewisse Unterschiede. So hat man auch dort Primzahlen und kann auch jede Zahl in Primfaktoren zerlegen, aber diese Zerlegung ist dort im allgemeinen *nicht mehr eindeutig*. Dieser Umstand schuf in der Arithmetik der algebraischen Zahlen neue eigenartige Schwierigkeiten, die man von der gewöhnlichen Arithmetik der natürlichen Zahlen her nicht kannte und die erst in neuerer Zeit erfolgreich bewältigt wurden.

Den Eindeutigkeitsbeweis für die Zerlegung der natürlichen Zahlen in Primfaktoren stützt man gewöhnlich auf den folgenden überaus wichtigen Satz, der auch in vielen anderen Fällen in der Zahlentheorie nützliche Dienste leistet:

Satz 1. *Wenn die natürlichen Zahlen a und b teilerfremd sind, so gibt es ganze Zahlen x und y derart, daß*

$$ax - by = 1$$

ist.

Man beweist diesen Satz meistens unter Benutzung des Euklidischen Algorithmus oder der Theorie der Kettenbrüche. Wir kommen darauf in Kapitel III zurück. Hier wollen wir einen anderen, methodisch sehr lehrreichen Beweis geben, der auf GAUSS zurückgeht und der keinen Gebrauch von irgendwelchen Algorithmen macht.

Es sei d die kleinste positive Zahl, die sich bei beliebiger Wahl der Zahlen x und y in der Form

$$d = ax - by \tag{1}$$

darstellen läßt. Wir wollen zeigen, daß $d = 1$ ist. Hierzu genügt es, da a und b voraussetzungsgemäß teilerfremd sind, d. h. keinen anderen gemeinsamen positiven Teiler als 1 besitzen, zu zeigen, daß sowohl a als auch b durch d teilbar ist. Da in (1) die Zahlen a und b völlig gleichberechtigt sind, reicht es hin, für eine der beiden Zahlen — etwa für a — zu zeigen, daß sie durch d teilbar ist.

Dazu möge a durch d mit dem Quotienten m und dem Rest r teilbar, also

$$a = md + r \quad (0 \leq r < d)$$

sein. Hieraus folgt, daß

$$r = a - md = a - m(ax - by) = a(1 - mx) - b(-my) = ax' - by'$$

ist, wobei

$$x' = 1 - mx, \quad y' = -my$$

gesetzt wurde.

Also kann r in der Form $ax' - by'$ mit ganzem x' und y' dargestellt werden. Da $r < d$ und definitionsgemäß d die kleinste positive Zahl ist, die sich in der Form $ax - by$ darstellen läßt, muß $r = 0$, also $a = md$ sein, d. h., a ist tatsächlich durch d teilbar, was zu beweisen war.

Wir bemerken, daß wir gleich etwas mehr bewiesen haben, nämlich einen Satz, der sich auf beliebige ganze Zahlen a und b bezieht (die nicht notwendig teilerfremd zu sein brauchen):

Die kleinste positive Zahl d, die sich in der Form $ax - by$ mit ganzen a und b darstellen läßt, ist der größte gemeinsame Teiler der Zahlen a und b.

Daß d ein gemeinsamer Teiler von a und b ist, haben wir gerade oben bewiesen. Andererseits zeigt aber die Relation (1), daß auch jeder gemeinsame Teiler von a und b ein Teiler von d ist, d. h., d ist tatsächlich der größte gemeinsame Teiler.

Wir weisen noch darauf hin, daß die hier behandelten Fragen im Zusammenhang mit dem in Kapitel III betrachteten Euklidischen Algorithmus ein ausgezeichnetes, nicht schwieriges und zugleich interessantes Material für mathematische Zirkel an Oberschulen darstellen können.

Den Satz 1 benutzen wir jetzt zum Beweis des folgenden sehr wichtigen Satzes der Teilbarkeitstheorie (der schon EUKLID bekannt war):

Satz 2. *Wenn die Zahlen a und b teilerfremd sind und wenn b ein Teiler von ac ist, so ist c durch b teilbar.*

In der Tat: Auf Grund von Satz 1 gibt es ganze Zahlen x und y derart, daß

$$ax - by = 1$$

ist, woraus sich

$$acx - bcy = c$$

ergibt. Da nach Voraussetzung ac durch b teilbar ist, gibt es eine ganze Zahl k mit $ac = bk$. Dann ist

$$c = acx - bcy = bkx - bcy = b(kx - cy),$$

woraus zu ersehen ist, daß b ein Teiler von c ist.

Wenn p eine Primzahl und a eine beliebige natürliche Zahl ist, so tritt offensichtlich stets einer der folgenden beiden Fälle ein: Es ist p ein Teiler von a, oder a und p sind teilerfremd. Wenn nämlich a und p nicht teilerfremd sind, so besitzen sie einen gemeinsamen Teiler $d > 1$; da aber p eine Primzahl ist, die nur durch 1 und p selbst teilbar ist, muß $d = p$ sein, und a wird von p geteilt.

Auf Grund dieser einfachen Bemerkung erhalten wir aus Satz 2:

Folgerung: *Wenn das Produkt ab durch eine Primzahl p teilbar ist, so ist wenigstens einer der Faktoren a oder b durch p teilbar.*

Wenn nämlich z. B. a nicht durch p teilbar ist, so ist nach der eben gemachten Bemerkung p zu a teilerfremd. Dann folgt aber auf Grund von Satz 2 aus der Teilbarkeit von ab durch p tatsächlich die Teilbarkeit von b durch p, was zu beweisen war.

Dieses Teilbarkeitsgesetz, das wir soeben für ein Produkt aus zwei Faktoren bewiesen haben, läßt sich leicht durch vollständige Induktion auf eine beliebige Anzahl von Faktoren ausdehnen. Wenn z. B. das Produkt abc durch eine Primzahl p teilbar ist und a nicht von p geteilt wird, so muß nach dem Bewiesenen das Produkt bc durch p teilbar sein. Dann ist aber — wie wir wissen — b oder c durch p teilbar. Auf diese Weise haben wir in endlich vielen Schritten erschlossen, daß

aus der Teilbarkeit des Produktes abc durch p die Teilbarkeit wenigstens eines der Faktoren a, b, c durch p folgt. Offensichtlich kann man in gleicher Weise von drei zu vier Faktoren, von vier zu fünf Faktoren und allgemein zu einer beliebigen endlichen Anzahl von Faktoren übergehen.[1]) Wir können damit die folgende allgemeine Aussage als bewiesen ansehen, die im wesentlichen das Ziel der vorangehenden Überlegungen ist:

Satz 3. *Wenn ein Produkt aus endlich vielen Faktoren durch eine Primzahl p teilbar ist, so ist wenigstens einer der Faktoren durch p teilbar.*

Aus Satz 3 folgt unmittelbar die Eindeutigkeit der Zerlegung einer beliebigen, von 1 verschiedenen natürlichen Zahl in Primfaktoren. Wenn nämlich

$$n = p_1 p_2 \cdots p_r = q_1 q_2 \cdots q_s \tag{2}$$

ist, wobei alle p_i und q_j Primzahlen sind, so müssen wir zeigen, daß sich die Zahlen $q_1, q_2, \ldots, q_s$ höchstens in ihrer Reihenfolge von den Zahlen $p_1, p_2, \ldots, p_r$ unterscheiden. Wenn wir zudem voraussetzen, daß die Zahlen p_i und q_j im Sinne der obigen Entwicklung angeordnet sind (daß also $p_1 \leq p_2 \leq \cdots \leq p_r$ und $q_1 \leq q_2 \leq \cdots \leq q_s$ ist), so müssen wir zeigen, daß $r = s$ und $p_i = q_i$ ist für alle i von 1 bis r. In dieser Weise werden wir im folgenden verfahren.

Wir zeigen zunächst, daß $p_1 = q_1$ ist. Angenommen, es wäre $q_1 > p_1$. Wegen (2) ist das Produkt $q_1 q_2 \cdots q_s$ durch p_1 teilbar. Daher ist nach Satz 3 wenigstens eine der Zahlen q_j durch p_1 teilbar. Da alle q_j Primzahlen sind, muß jede von ihnen, die durch p_1 teilbar ist, gleich p_1 sein. Das ist aber nach unserer Annahme nicht möglich, da gemäß unserem Ansatz

$$p_1 < q_1 \leq q_2 \leq \cdots \leq q_s$$

ist. Ebenso kann nicht $p_1 > q_1$ sein. Also ist $p_1 = q_1$. Hieraus und aus (2) folgt, daß

$$p_2 p_3 \cdots p_r = q_2 q_3 \cdots q_s$$

ist. Mit Hilfe derselben Überlegungen können wir offensichtlich zeigen, daß $p_2 = q_2$ und damit

$$p_3 p_4 \cdots p_r = q_3 q_4 \cdots q_s$$

ist. Hieraus folgt dann $p_3 = q_3$ usw. Dieses Verfahren können wir so lange fortsetzen, wie auf beiden Seiten noch Primfaktoren stehen. Es bricht ab, wenn auf einer Seite die Primfaktoren erschöpft sind. Es ist nun unmittelbar klar, daß dies auf beiden Seiten gleichzeitig geschehen muß, daß also $r = s$ sein muß. Wäre nämlich z. B. $r < s$, so würden wir auf Grund des Gesagten schließen können, daß $p_1 = q_1, p_2 = q_2, \ldots, p_r = q_r$ ist, und nach Kürzen dieser gleichen Faktoren, daß

$$1 = q_{r+1} q_{r+2} \cdots q_s$$

wäre. Das ist aber offensichtlich nicht der Fall. Also ist in der Tat

$$r = s$$

und

$$p_1 = q_1, \quad p_2 = q_2, \cdots, p_r = q_r,$$

d. h., beide Zerlegungen stimmen überein. Damit ist der Fundamentalsatz der Teilbarkeitstheorie vollständig bewiesen.

[1]) Bezüglich der strengen Fassung des Satzes von der vollständigen Induktion vgl. den vorangehenden Artikel von I. W. Proskurjakow, § 11, Satz 1. — *Anm. d. wissenschaftl. Red.*

Wir sehen, daß wir bei seinem Beweis wesentlich von Satz 2 Gebrauch gemacht haben. Alle bekannten Beweise des Fundamentalsatzes basieren in irgendeiner Form auf ihm. Sie unterscheiden sich lediglich in den Wegen, die zu Satz 2 führen. Der oben angegebene Beweis von Satz 2 stützte sich wesentlich auf Satz 1. Methodisch ist dieser Weg von größtem Interesse, da er nicht von den Eigenschaften des größten gemeinsamen Teilers und des kleinsten gemeinsamen Vielfachen zweier Zahlen Gebrauch macht, ja es lag überhaupt kein Grund dazu vor, diese Begriffe einzuführen. Darüber hinaus gestattet bekanntlich gerade der Fundamentalsatz über die eindeutige Primzahlzerlegung eine möglichst einfache und durchsichtige Einführung des Begriffs des größten gemeinsamen Teilers und des kleinsten gemeinsamen Vielfachen.

Immerhin ist es methodologisch von großem Interesse, daß der entscheidende Satz 2 auch auf einem ganz anderen Wege bewiesen werden kann, der nicht den Satz 1 benutzt, sondern auf den elementaren Eigenschaften des kleinsten gemeinsamen Vielfachen beruht. Wir wollen daher auch diesen Beweis vorführen.

Dazu muß man vor allen Dingen die Struktur der Gesamtheit aller gemeinsamen Vielfachen zweier gegebener Zahlen a und b untersuchen, d. h. der Gesamtheit aller Zahlen, die sowohl durch a als auch durch b teilbar sind. Dazu sei m die kleinste positive Zahl, die durch a und b teilbar ist (also das kleinste gemeinsame Vielfache der Zahlen a und b) und m' ein beliebiges gemeinsames Vielfaches von a und b. Ferner sei q der Quotient und r der Rest bei der Division von m' durch m, also

$$m' = qm + r \quad (0 \leq r < m)$$

und mithin

$$r = m' - qm.$$

Da nun m und m' beide durch a und b teilbar sind, so ist auch r durch a und b teilbar. Weil aber $r < m$ und m das kleinste positive gemeinsame Vielfache der Zahlen a und b ist, muß $r = 0$ und daher $m' = qm$ sein, d. h., jedes gemeinsame Vielfache der Zahlen a und b ist ein Vielfaches der Zahl m. Andererseits ist aber auch jede Zahl der Form qm ein gemeinsames Vielfaches der Zahlen a und b, d. h., *die Gesamtheit aller Vielfachen der Zahlen a und b ist genau gleich der Gesamtheit aller Zahlen, die Vielfaches einer gewissen Zahl m sind* (die dann das kleinste gemeinsame Vielfache der Zahlen a und b ist).

Wir zeigen nun, daß *das kleinste gemeinsame Vielfache zweier teilerfremder Zahlen gleich ihrem Produkt ist.*

Dazu seien a und b zwei beliebige teilerfremde Zahlen und m ihr kleinstes gemeinsames Vielfaches. Da das Produkt ab ein gemeinsames Vielfaches von a und b ist, ist nach dem vorher Gesagten

$$ab = qm,$$

wobei q eine gewisse ganze Zahl ist. Nun zeigen aber die sich hieraus ergebenden Relationen

$$\frac{a}{q} = \frac{m}{b}, \quad \frac{b}{q} = \frac{m}{a},$$

daß q ein gemeinsamer Teiler der teilerfremden Zahlen a und b ist (m war nach Voraussetzung das kleinste gemeinsame Vielfache von a und b, also sind $\frac{m}{b}$ und $\frac{m}{a}$

ganze Zahlen — *Anm. d. wissenschaftl. Red.*). Folglich ist $q = 1$ und $m = ab$, was zu beweisen war.

Hieraus ergibt sich nun ganz leicht der obige Satz 2. Wenn nämlich ac ein gemeinsames Vielfaches der teilerfremden Zahlen a und b ist, so wird ac von dem kleinsten gemeinsamen Vielfachen dieser Zahlen, also — wie wir eben festgestellt haben — von dem Produkt ab geteilt. Es ist aber die Teilbarkeit von ac durch ab gleichbedeutend mit der Teilbarkeit von c durch b, womit Satz 2 bewiesen ist.

§ 3. Über Primzahlen

Die Primzahlen spielen in der Teilbarkeitstheorie die Rolle der primären, unzerlegbaren Elemente, aus denen sich mittels der Multiplikation alle natürlichen Zahlen gewinnen lassen, wie dies der oben aufgestellte Fundamentalsatz zeigt. Die Menge der Primzahlen kann also mit einigem Recht als multiplikative Basis (d. h. als eine Basis, mit deren Hilfe sich alle Zahlen multiplikativ darstellen lassen) der Menge der natürlichen Zahlen angesprochen werden. Diese grundlegende Eigenschaft der Menge der Primzahlen hat zu allen Zeiten das Interesse der Forscher erregt. Was ist das nun für eine Menge, wieviel Elemente enthält sie, wie ordnen sich ihre Elemente in die Reihe aller natürlichen Zahlen ein und welche Gesetzmäßigkeiten herrschen bezüglich der Verteilung der Primzahlen gegenüber den zusammengesetzten Zahlen? Diese Fragen bewegten die Gelehrten aller Epochen von der Antike bis in unsere Tage. Auch heute stehen sie in beträchtlichem Maße im Mittelpunkt der arithmetischen Forschung, namentlich weil ihre Lösung sich als außerordentlich schwierig erwiesen hat.

Zuerst erhebt sich selbstverständlich die Frage, ob es endlich oder unendlich viele Primzahlen gibt. Hierbei ist hervorzuheben, daß der von uns bewiesene Fundamentalsatz darüber — zumindest unmittelbar — nichts aussagt. Er ist weder eine Bestätigung für die Endlichkeit noch für die Unendlichkeit der Menge der Primzahlen. Die genannte Aufgabe war ein erstes natürliches Problem der Theorie der Primzahlen, das bereits von den Mathematikern des Altertums gelöst wurde. Wir führen hier die einfache und scharfsinnige Lösung vor, die sich bei EUKLID findet (womit nicht gesagt sein soll, daß sie von EUKLID stammt — *Anm. d. wissenschaftl. Red.*) und die die Unendlichkeit der Menge der Primzahlen beweist. Obwohl die moderne Ausdrucksweise für das Unendliche, das in der Mathematik eine so zentrale Stellung einnimmt, EUKLID natürlich fremd war, hat er das Wesentliche dieses Satzes folgendermaßen treffend formuliert: *Die Menge der Primzahlen ist größer als jede gegebene Menge derselben* (in EUKLIDS „Elementen", Buch IX, Satz 20).

Dazu seien $p_1, p_2, \ldots, p_k$ gewisse, endlich viele Primzahlen. Wir werden zeigen, daß es eine Primzahl p gibt, die unter den vorgegebenen Zahlen nicht vorkommt. Zunächst betrachten wir die Zahl $P + 1$, wobei $P = p_1 p_2 \cdots p_k$ ist, und bezeichnen mit p den kleinsten von 1 verschiedenen Teiler dieser Zahl. Offensichtlich ist p eine Primzahl, die außerdem mit keiner der Zahlen $p_1, p_2, \ldots, p_k$ zusammenfallen kann, weil p ein Teiler von $P + 1$ ist, mithin bei Division durch jede der Zahlen $p_1, p_2, \ldots p_k$ den Rest 1 ergibt und folglich durch keine der vorgegebenen Zahlen teilbar sein kann. Damit haben wir in p eine Primzahl gefunden, die noch nicht

in der vorgegebenen (endlichen) Menge von Primzahlen enthalten ist, wie die Behauptung EUKLIDs verlangt.

Bezüglich der Anordnung der Primzahlen in der Reihe der natürlichen Zahlen soll hier nur eine Tatsache erwähnt werden, die sich genau so einfach wie der Satz von EUKLID beweisen läßt: *In der Reihe der natürlichen Zahlen gibt es beliebig lange Intervalle, die keine Primzahl enthalten, also aus lauter zusammengesetzten Zahlen bestehen.*

In der Tat: Wenn $n > 1$ eine beliebige natürliche Zahl ist, so enthält die Folge

$$n! + 2, \quad n! + 3, \quad n! + 4, \quad \ldots, \quad n! + n$$

(die offensichtlich ein Intervall in der Menge der natürlichen Zahlen von der Länge $n - 1$ bildet) keine Primzahl, weil $n! + 2$ durch 2, $n! + 3$ durch 3 usw. und schließlich $n! + n$ durch n teilbar ist, wobei in allen Fällen der Teiler kleiner ist als die durch ihn teilbare Zahl.

Die Frage nach der Verteilung der Primzahlen in der Menge der natürlichen Zahlen war in der Zeit von EUKLID bis EULER so gut wie unbeantwortet geblieben. Man hatte lediglich mit dem Ziel, irgendeine Gesetzmäßigkeit in der Reihenfolge der Primzahlen zu finden, Tabellen dieser Zahlen aufgestellt, die mit der Zahl 2 begannen und bis zu sehr großen Zahlen reichten (heute bis ungefähr 10 Millionen). Das Studium dieser Tabellen zeigte, daß beim Fortschreiten in der Folge der natürlichen Zahlen die Primzahlen im Mittel immer seltener und seltener werden. Das gilt aber nur im Mittel. In Wirklichkeit geht der anzahlmäßige Rückgang der Primzahlen außerordentlich unregelmäßig vor sich. Nach bedeutender Auflockerung zeigen sich immer wieder Verdichtungen, wobei bis heute noch keine Gesetzmäßigkeit in der Reihenfolge dieser Verdichtungen und Verdünnungen festgestellt worden ist.

Das alles gibt dem Problem der Verteilung der Primzahlen seine historisch bekannte, treibende Kraft.

Das bedeutendste Resultat EULERS auf diesem Gebiet ist die theoretische Begründung dieser allmählichen Verminderung der Anzahl der Primzahlen in den fernsten Teilen der Folge der natürlichen Zahlen, die sich — wie schon gesagt — beim Studium der Primzahltabellen gezeigt hatte. Im folgenden wollen wir mit $\pi(n)$ die Anzahl der Primzahlen bezeichnen, die die vorgegebene Zahl n nicht übertreffen. In diesem Sinne ist also $\pi(10) = 4$ und $\pi(23) = 9$. Dann können wir das Verhältnis $\frac{\pi(n)}{n}$ (das offensichtlich immer zwischen Null und Eins liegt) als einen Bruch ansehen, der uns die „mittlere Dichte" der Primzahlen im Abschnitt der natürlichen Zahlen von 1 bis n festlegt: Je kleiner dieser Bruch ist, um so kleiner ist die Anzahl der Primzahlen im Abschnitt $(1, n)$ im Verhältnis zur Anzahl der natürlichen Zahlen dieses Abschnittes.

Satz von EULER. *Wenn die Zahl n unbeschränkt wächst, so geht $\frac{\pi(n)}{n}$ gegen 0.*

Das besagt nichts anderes, als daß in allen hinreichend großen Abschnitten der Folge der natürlichen Zahlen die überwiegende Mehrheit der Zahlen zusammengesetzt ist und nur ein geringfügiger Anteil zur Menge der Primzahlen gehört.

Als wesentliches Hilfsmittel zum Beweis des Satzes von EULER schicken wir zunächst den folgenden Satz voran:

Lemma. *Es sei $p_1, p_2, p_3, \ldots$ die Menge der Primzahlen in ihrer natürlichen Reihenfolge* (also $p_1 = 2$, $p_2 = 3$, $p_3 = 5$ usw.). *Dann geht*

$$\Pi_n = \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_n}\right)$$

gegen 0, falls n über alle Grenzen wächst.

Zum Beweis bemerken wir zunächst, daß für beliebiges k

$$\frac{1}{1 - \frac{1}{p_k}} = 1 + \frac{1}{p_k} + \frac{1}{p_k^2} + \cdots = \sum_{s=0}^{\infty} \frac{1}{p_k^s}$$

ist. Daher ist

$$\frac{1}{\Pi_n} = \left(\sum_{s=0}^{\infty} \frac{1}{p_1^s}\right) \left(\sum_{s=0}^{\infty} \frac{1}{p_2^s}\right) \cdots \left(\sum_{s=0}^{\infty} \frac{1}{p_n^s}\right).$$

Hier steht auf der rechten Seite ein Produkt aus n absolut konvergenten Reihen, das man bekanntlich[1]) genau wie ein Produkt aus endlichen Summen gliedweise ausmultiplizieren kann. Dann erhält man eine Reihe, deren allgemeines Glied die Form

$$\frac{1}{p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}}$$

hat, wobei $\alpha_1, \alpha_2, \ldots, \alpha_n$ beliebige Zahlen aus der Folge 0, 1, 2, ... sind. Wir erhalten also

$$\frac{1}{\Pi_n} = \sum_{\alpha_1, \alpha_2, \ldots, \alpha_n} \frac{1}{p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}},$$

wobei die Summation über alle Kombinationen der Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_n$ (in beliebiger Reihenfolge) zu erstrecken ist. In der Form

$$p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_n^{\alpha_n}$$

können nun offenbar bei beliebiger Wahl der α_i genau die natürlichen Zahlen dargestellt werden, die keine anderen Primteiler als $p_1, p_2, \ldots, p_n$ besitzen. Insbesondere kommen unter diesen Zahlen alle natürlichen Zahlen vor, die nicht größer als p_n sind. Unter den Summanden der rechten Seite finden sich also bestimmt die Brüche $\frac{1}{m}$ mit $1 \leq m \leq p_n$. Daher ist

$$\frac{1}{\Pi_n} > \sum_{m=1}^{p_n} \frac{1}{m}.$$

Nun ist bekanntlich[2]) die Reihe $\sum_{m=1}^{\infty} \frac{1}{m}$ (die sogenannte „harmonische Reihe") divergent. Wir werden also zu jedem noch so kleinen positiven ε ein n und damit ein p_n finden können, für das

$$\sum_{m=1}^{p_n} \frac{1}{m} > \frac{1}{\varepsilon}$$

[1]) Siehe EdEM Band 3, Differential- und Integralrechnung.
[2]) Vgl. die vorangehende Fußnote.

ist. Hieraus und aus der vorangehenden Ungleichung folgt:

$$\frac{1}{\Pi_n} > \frac{1}{\varepsilon}$$

und mithin:

$$\Pi_n < \varepsilon.$$

Da aus der Gültigkeit dieser Ungleichung für ein gewisses n auch die Gültigkeit für alle größeren natürlichen Zahlen folgt, ist in Anbetracht der beliebigen Wahl von ε der behauptete Hilfssatz bewiesen.

Wir gehen jetzt zum Beweis des Satzes von EULER über. Dazu bezeichnen wir mit P_n das Produkt $p_1 p_2 \cdots p_n$ der ersten n Primzahlen.

Für den Beweis ist es zunächst notwendig, die Anzahl Q_n der Zahlen aus der Folge

$$1, 2, \ldots, P_n \tag{3}$$

zu bestimmen, die durch keine der Primzahlen $p_1, p_2, \ldots, p_n$ teilbar sind. Es zeigt sich, daß diese Anzahl

$$Q_n = P_n \Pi_n = P_n \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_n}\right)$$

ist. Eine ausführliche Begründung dieser Formel wird in Kapitel II gegeben werden (siehe Seite 251).

Es mögen nun r und s beliebige natürliche Zahlen sein. Notwendig und hinreichend dafür, daß die Zahl $sP_n + r$ durch eine der Zahlen $p_1, p_2, \ldots, p_n$ teilbar wird, ist dann offenbar, daß r durch diese Zahl teilbar ist. Daher sind in der Folge

$$sP_n + 1, \quad sP_n + 2, \ldots, sP_n + P_n = (s+1)P_n$$

für beliebiges $s \geq 0$ genau so viele Zahlen durch keine der Primzahlen $p_1, p_2, \ldots, p_n$ teilbar wie in der Folge (3), also Q_n.

Zu einer vorgegebenen, beliebig kleinen positiven Zahl ε bestimmen wir zunächst eine Zahl n so, daß

$$\Pi_n = \left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_n}\right) < \varepsilon$$

wird, was nach dem vorangehenden Lemma möglich ist. Wenn dann q eine beliebige natürliche Zahl ist, so wollen wir q durch P_n mit Rest dividieren, wobei sich der Quotient s und der Rest r ergeben möge, also

$$q = sP_n + r \quad (0 \leq r < P_n)$$

ist. Unser Ziel ist es, die Anzahl $\pi(q)$ der Primzahlen im Abschnitt $(1, q)$ zu bestimmen. Unter diesen können zunächst die n Primzahlen $p_1, p_2, \ldots, p_n$ vorkommen. Alle übrigen Primzahlen des Abschnittes $(1, q)$ sind dann durch keine der Primzahlen $p_1, p_2, \ldots, p_n$ teilbar. Von diesen gibt es aber — wie wir bereits wissen — in jedem der Intervalle $(1, P_n)$, $(P_n + 1, 2P_n)$, $(2P_n + 1, 3P_n), \ldots,$ $((s-1)P_n + 1, sP_n)$ genau $Q_n = P_n \Pi_n$ Stück, also im ganzen Abschnitt $(1, sP_n)$ genau $sQ_n = sP_n \Pi_n$. Daraus folgt, daß es im Abschnitt $(1, q) = (1, sP_n + r)$ nicht mehr als

$$sQ_n + r < sP_n \Pi_n + P_n$$

Primzahlen dieser Art geben kann. Auf diese Weise finden wir, daß

$$\pi(q) < n + s P_n \Pi_n + P_n$$

und mithin

$$\frac{\pi(q)}{q} < \frac{n + s P_n \Pi_n + P_n}{s P_n + r} \leq \frac{n + P_n + s P_n \Pi_n}{s P_n} = \frac{n + P_n}{s P_n} + \Pi_n$$

ist.

Wenn wir nun die Zahl q (und damit die Zahl s) hinreichend groß wählen (wobei aber n fest bleibt!), so kann der erste Summand auf der rechten Seite kleiner als ε gemacht werden. Da andererseits die Zahl n so gewählt war, daß $\Pi_n < \varepsilon$ ist, so ist

$$\frac{\pi(q)}{q} < 2\varepsilon$$

für alle hinreichend großen q, d. h., es ist

$$\lim_{q \to \infty} \frac{\pi(q)}{q} = 0,$$

wie der Satz von EULER behauptet.

Dieses bemerkenswerte Resultat sagt aus, daß in der Folge der natürlichen Zahlen die Primzahlen im bekannten Sinne sehr „spärlich" vertreten sind, spärlicher z. B. als die Glieder einer beliebigen arithmetischen Folge (mit beliebig großer Differenz). Jedoch gilt das — wie schon betont — nur im Mittel. Wir haben keinerlei Gründe, uns die Sache so vorzustellen, daß irgendwo, sehr weit entfernt in der Folge der natürlichen Zahlen die Primzahlen ganz „einsam" stehen, ohne daß einer ihrer nähergelegenen Nachbarn — weder nach der einen noch nach der anderen Seite — eine Primzahl ist. Im Gegenteil zeigt das Studium der Primzahltabellen, daß von Zeit zu Zeit in der Reihe der natürlichen Zahlen Primzahlen auftreten, die sehr nahe beieinander liegen, ja daß sogar immer wieder „Primzahlzwillinge" auftreten, die sich nur um zwei Einheiten unterscheiden, wie (5, 7), (41, 43), (101, 103) usw. Die Frage, ob es unendlich viele derartige Zwillinge gibt, ist bis zum heutigen Tage unbeantwortet. Wir haben jedoch keinerlei Gründe zu der Annahme, daß die Existenz von unendlich vielen solchen Zahlen unmöglich ist.

Wir wollen hier noch die folgende interessante Bemerkung machen. Das gleiche Lemma, das uns oben zum EULERschen Satz führte, der etwas über die relative „Spärlichkeit" der Verteilung der Primzahlen in der Folge der natürlichen Zahlen aussagt, gibt uns auch die Möglichkeit eines Beweises (und zwar auf viel einfacherem Wege) dafür, daß die Primzahlen in der Folge der natürlichen Zahlen auch hinreichend dicht liegen: Aus der Theorie der unendlichen Produkte ist bekannt[1]), daß aus der Tatsache, daß für unbeschränkt wachsendes n die Folge der Produkte

$$\left(1 - \frac{1}{p_1}\right) \cdot \left(1 - \frac{1}{p_2}\right) \cdots \left(1 - \frac{1}{p_n}\right)$$

gegen Null strebt, die Divergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{p_n}$ folgt, daß also die Folge der

[1]) Siehe A. Я. Хинчин, Восемь лекций по математическому анализу, (A. J. CHINTSCHIN, Acht Vorlesungen über mathematische Analysis), Gostechisdat, 1948. Der deutsche Leser sei verwiesen auf MANGOLDT-KNOPP, Einführung in die höhere Mathematik, Band 2, 9. Auflage, Leipzig 1949, S. 295. — *Anm. d. wissenschaftl. Red.*

Summen

$$\sum_{i=1}^{n} \frac{1}{p_i}$$

mit wachsendem n über alle Grenzen wächst. In dieser Beziehung verhält sich somit die Menge der Primzahlen genau so wie die Menge aller natürlichen Zahlen (Divergenz der harmonischen Reihe $1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}+\cdots$), anders als z. B. die Menge der Quadrate $1^2, 2^2, 3^2, \ldots$ (die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert). Das zeigt, daß die Primzahlen in gewissem Sinne „dichter" liegen als die Quadratzahlen.

Alle diese elementar beweisbaren Tatsachen, die aus dem Anfang des 19. Jahrhunderts stammen, geben jedoch nur ein höchst unklares Bild von der Verteilung der Primzahlen in der Folge der natürlichen Zahlen. Die Mathematiker dieser Epoche dachten jedoch schon früh an eine viel bedeutendere Errungenschaft, nämlich für die Funktion $\pi(n)$ (die Anzahl der Primzahlen, die nicht größer als n sind) eine gute Näherungsformel in Form eines möglichst einfachen analytischen Ausdrucks in Abhängigkeit von n anzugeben, um die Wachstumsverhältnisse der Funktion $\pi(n)$ ohne Schwierigkeiten erforschen zu können. So ergab sich der sehr verlockende Gedanke, zu der Funktion $\pi(n)$ einen „asymptotisch gleichen" analytischen Ausdruck anzugeben, d. h. eine analytische Funktion $\varphi(n)$ zu finden, für die

$$\frac{\pi(n)}{\varphi(n)} \to 1 \quad (n \to \infty)$$

gilt (die — m. a. W. — von gleicher Größenordnung unendlich wird wie die Funktion $\pi(n)$). Indes verfügte die Wissenschaft jener Epoche noch nicht über die zur Lösung dieser bedeutenden Aufgabe unentbehrlichen Hilfsmittel. Das Studium der Primzahltabellen zeigte, daß es unter den elementaren Funktionen eine sehr einfache gibt, nämlich $\frac{n}{\ln n}$, die für große n (in den Grenzen der Tabellen) eine gute Annäherung an die Funktion $\pi(n)$ ergibt. Viele der bedeutendsten Gelehrten jener Epoche (LEGENDRE, GAUSS) versuchten beharrlich, diese zunächst empirisch gefundenen Zusammenhänge exakt zu begründen. Diese Aufgabe erwies sich jedoch als maßlos schwierig, und bis zur Mitte des 19. Jahrhunderts wurde nicht ein Resultat in dieser Richtung erzielt; dann aber trieben die großartigen Untersuchungen des russischen Mathematikers P. L. TSCHEBYSCHEW die Entwicklung voran.

Das Ziel, das — wie gesagt — jedem der Forscher vorschwebte, war der Beweis des sogenannten Primzahlsatzes, der Relation

$$\frac{\pi(n)}{\frac{n}{\ln n}} = \frac{\pi(n)\ln n}{n} \to 1 \quad (n \to \infty). \tag{4}$$

Auf dem Wege zu diesem Ziel, das in dieser Zeit noch nicht erreicht werden konnte, bewies TSCHEBYSCHEW als erster einige entscheidende Tatsachen, von denen wir nur die beiden folgenden erwähnen:

1. Wenn der Grenzwert

$$\lim_{n \to \infty} \frac{\pi(n)\ln n}{n}$$

existiert, so ist er gleich Eins.

2. Für alle hinreichend großen n ist

$$\ln 2 < \frac{\pi(n) \ln n}{n} < 2 \ln 2.$$

Die Tragweite dieser hervorragenden Leistungen, die schon sehr nahe an das endgültige Ergebnis (4) heranführen, wird noch besonders durch zwei Umstände unterstrichen: Erstens, daß vor TSCHEBYSCHEW absolut nichts in der genannten Richtung zu beweisen geglückt war, daß er also alle Ideen und Methoden seiner Beweise selbst schaffen mußte, da er keine Vorgänger hatte. Zweitens, daß TSCHEBYSCHEW seine so bedeutenden Resultate durch völlig elementare arithmetische Überlegungen, ohne Hilfsmittel der höheren Mathematik in Anspruch zu nehmen, erhielt.

Unmittelbar nach den Arbeiten TSCHEBYSCHEWs erschienen die Untersuchungen des deutschen Mathematikers B. RIEMANN, die vollkommen neue, komplizierte analytische Hilfsmittel zur Behandlung von Problemen der Verteilung der Primzahlen enthielten. Dabei ist jedoch zu bemerken, daß RIEMANN selbst seine Methoden nicht als Resultat arithmetischer Forschung fand. So kam es, daß erst später, gegen Ende des 19. Jahrhunderts, die große Tragweite der RIEMANNschen Methoden in Verbindung mit der sich zu dieser Zeit stürmisch entwickelnden Theorie der Funktionen einer komplexen Veränderlichen erkannt wurde. Mit Hilfe dieser Methoden gelang es schließlich im Jahre 1894 dem französischen Mathematiker HADAMARD, das so lange verfolgte Ziel zu erreichen und die Relation (4) zu beweisen, auf Grund derer die Funktion $\frac{n}{\ln n}$ tatsächlich ein asymptotischer Ausdruck für die arithmetische Funktion $\pi(n)$ ist.

Die weiteren Untersuchungen in dieser Richtung beschäftigen sich bis zum heutigen Tage mit einer Präzisierung dieses Ergebnisses, d. h. mit der Frage nach einer genaueren Abschätzung der Differenz

$$\frac{\pi(n) \ln n}{n} - 1,$$

von der man auf Grund des HADAMARDschen Satzes nur weiß, daß sie für $n \to \infty$ beliebig klein wird. Sehr bedeutende Resultate in diesem Zusammenhang lieferte in den letzten Jahren die von I. M. WINOGRADOW — einem der größten Schöpfer der modernen Arithmetik — geleitete sowjetische zahlentheoretische Schule.

Ein anderer Zweig in der Entwicklung der Theorie der Primzahlen, der gleichfalls von dem Satz von EUKLID über die Existenz unendlich vieler Primzahlen ausging, behandelt die Frage nach der Anzahl der Primzahlen in gewissen Teilmengen der Menge der natürlichen Zahlen, d. h. unter den natürlichen Zahlen einer gewissen vorgegebenen Form. Ein klassisches Ergebnis in dieser Richtung ist der Satz von DIRICHLET über die Existenz von unendlich vielen Primzahlen in jeder arithmetischen Progression, deren erstes Glied zur Differenz teilerfremd ist. Es ist bis heute noch nicht gelungen, dieses Resultat (für das wir — an dieser Stelle möge darauf hingewiesen werden — noch keinen rein arithmetischen Beweis kennen)[1]) wesentlich zu verbessern. Der Satz von DIRICHLET besagt, daß es für teilerfremdes a und b unter den Zahlen der Form $ax + b$ (wobei x eine beliebige ganze

[1]) Inzwischen wurde ein von analytischen Hilfsmitteln freier Beweis für den DIRICHLETschen Primzahlsatz gegeben von H. ZASSENHAUS: Über die Existenz von Primzahlen im arithmetischen Progressionen, Comment. Math. Helvetici 22 (1949), 232—259. — *Anm. d. wissenschaftl. Red.*

Zahl ist) unendlich viele Primzahlen gibt. Eine naheliegende Verallgemeinerung dieses Resultates wäre die Untersuchung der quadratischen Ausdrücke, also der Ausdrücke der Form $a x^2 + b x + c$ auf dieselbe Frage hin. Wir kennen selbst für die wesentlichsten Spezialfälle dieser Aufgabe bis heute keine Methoden zu ihrer Behandlung. Sogar die Frage, ob es unter den Zahlen der Form $x^2 + 1$, also in der Folge 2, 5, 10, 17, 26, 37, ... unendlich viele Primzahlen gibt, ist noch ungelöst.

Ein weiterer, sehr intensiv bearbeiteter Zweig der Theorie der Primzahlen beschäftigt sich mit den Fragen, die im Umkreis des berühmten GOLDBACHschen Problems liegen. Man hatte bereits sehr früh bemerkt, daß sich alle geraden Zahlen, von der Zahl 4 angefangen, als Summe aus zwei Primzahlen darstellen lassen ($4 = 2 + 2$, $6 = 3 + 3$, $8 = 3 + 5$, $10 = 3 + 7$, $12 = 5 + 7$ usw.) und daß folglich alle ungeraden Zahlen als Summen aus drei Primzahlen dargestellt werden können. Das GOLDBACHsche Problem besteht nun in der Lösung der Frage, ob dies tatsächlich für alle geraden (bzw. ungeraden) Zahlen gilt.

Noch vor 25 Jahren hatte es den Anschein, als ob die Mathematik keinerlei Methoden zur Lösung dieser überaus schwierigen Aufgabe kenne. Nach einigen unfruchtbaren Versuchen, die sich über mehrere Jahrhunderte erstreckten, wurde im Jahre 1930 durch den jungen sowjetischen Mathematiker L. G. SCHNIRELMANN ein bemerkenswerter Fortschritt in Richtung der Lösung des GOLDBACHschen Problems erreicht. Er bewies die Existenz einer Konstanten k mit der Eigenschaft, daß sich jede von 1 verschiedene natürliche Zahl als Summe aus höchstens k Primzahlen darstellen läßt. Bis zu der Arbeit SCHNIRELMANNs war es nicht gelungen, an dieses Resultat, geschweige denn an das GOLDBACHsche Problem nur heranzukommen. Das ist umso bemerkenswerter, als die ganzen Untersuchungen SCHNIRELMANNs mit so elementaren arithmetischen Methoden durchgeführt wurden, daß sie in derselben Form schon gut hundert Jahre früher, zur Zeit TSCHEBYSCHEWs, hätten durchgeführt werden können.

Die Konstante k, die sich auf Grund der Untersuchungen SCHNIRELMANNs unmittelbar abschätzen ließ, erwies sich zunächst als sehr groß. Daher beschäftigten sich sofort viele Wissenschaftler mit dem Problem, diese Konstante nach den gleichen elementaren Methoden besser abzuschätzen und weiter herabzusetzen. Im Laufe einiger Jahre wurde sie bis auf 69 herabgedrückt.

Jedoch konnte bereits im Jahre 1936 I. M. WINOGRADOW mit von ihm selbst geschaffenen analytischen Hilfsmitteln die GOLDBACHsche Vermutung für hinreichend große ungerade ganze Zahlen vollständig beweisen; er zeigte also, daß alle hinreichend großen ungeraden Zahlen als Summe dreier Primzahlen dargestellt werden können. Daraus folgt unmittelbar, daß sich alle hinreichend großen geraden Zahlen als Summe aus vier Primzahlen darstellen lassen. Auf diese Weise wurde die Konstante k von SCHNIRELMANN bis auf 4 herabgesetzt. Wenn man die historische Berühmtheit des GOLDBACHschen Problems und die zu seiner Lösung gemachten Anstrengungen berücksichtigt, muß man dieses Resultat I. M. WINOGRADOWs als einen der bedeutendsten Erfolge sowjetischer Mathematiker anerkennen.

Kapitel II

DIE KONGRUENZMETHODE

§ 4. Einleitung

Die besondere Schwierigkeit, die die Probleme der Zahlentheorie zu allen Zeiten auszeichnete, veranlaßte die Forscher, nach immer neuen Methoden in diesem Zweig der Mathematik zu suchen. So kommt es, daß wir heute in der Zahlentheorie eine solche Vielfalt an Methoden besitzen wie wohl in keiner anderen mathematischen Disziplin. Für alle diese Methoden ist charakteristisch, daß ihre Anwendungsmöglichkeiten verhältnismäßig beschränkt sind. Jede derartige Methode kann in der Regel nur zur Lösung eines mehr oder weniger engen Kreises von untereinander verwandten Aufgaben verwendet werden. Sobald die Grenzen eines solchen Kreises überschritten werden, muß nach neuen, zuweilen ganz anders gearteten Methoden gesucht werden.

Man kann die verschiedenen Verfahren zunächst ganz grob nach ihrer gegenständlichen Natur einteilen. Da haben wir in erster Linie die elementaren arithmetischen Methoden (Sieb des ERATOSTHENES, Euklidischer Algorithmus usw.). Daneben besitzen wir eine ganze Reihe von analytischen Methoden größter Leistungsfähigkeit. Schließlich treten heute immer mehr auch geometrische Methoden in den Vordergrund, wie sie von MINKOWSKI in die arithmetische Forschung eingeführt wurden. Diese Methoden können natürlich in anderen Beziehungen sehr unterschiedlich sein. So ist ein Teil von ihnen Ausdruck gewisser anschaulich-inhaltlicher Ideen (eine solche Methode ist z. B. die „Geometrie der Zahlen" von MINKOWSKI), während ein anderer als Grundlage für gewisse formale Verfahren dient. Dabei begegnet man selbstverständlich auch Methoden von gemischtem Typus.

Unter den formalen elementar-arithmetischen Methoden hat eine besondere Bedeutung die von GAUSS [Disquisitiones Arithmeticae, siehe Literaturverzeichnis Nr. 13, *die Wiss. Red.*] geschaffene sogenannte Kongruenzmethode erlangt. Diese Methode ist in erster Linie ein formaler Apparat, dem zwar eine sehr einfache Idee zugrunde liegt, der aber bedeutenden technischen Wert besitzt. Wenn man ihn beherrscht, kann man in vielen Fällen relativ einfach Resultate erhalten, die auf anderem Wege nur schwer und umständlich zu erreichen sind. Daher erweist sich in Verbindung mit den jeweiligen Einzelheiten die Theorie der Kongruenzen als ein formales Instrument, dessen Kenntnis eine merkliche Hilfe bei der Lösung fast aller Probleme der Zahlentheorie darstellt, und ohne dessen reichhaltige Methoden wir viele Aufgaben kaum bearbeiten würden. Man kann daher ohne Übertreibung die Theorie der Kongruenzen (beinahe als einzige) als eine universelle Methode der Zahlentheorie ansehen, wenn man nämlich darunter versteht, daß die Anwendung der Verfahren dieser Theorie wesentliches leisten kann und in fast allen Gebieten der Arithmetik eine erhebliche formale Hilfe gewährt. In allen Teilen der Zahlentheorie würde eine Arbeit ohne Benutzung der Elemente dieser Theorie in gewissem Sinne als „primitiv" und rückständig empfunden werden. Dabei darf man jedoch den Wert der Theorie der Kongruenzen auch nicht überschätzen. Selbst die vollkommenste Beherrschung dieser Methode kann dem Forscher nur als technisches Rüstzeug dienen, ohne daß sie ihm irgendeine leitende Idee zu geben vermag.

In diesem Kapitel wollen wir einen kurzen, nur auf das wesentlichste beschränkten Abriß der Grundbegriffe der Theorie der Kongruenzen und ihrer hauptsächlichen Anwendungen geben. Eine sowohl ihrer Vollständigkeit als auch ihrer Tiefe nach unübertroffene Darstellung der Theorie der Kongruenzen findet der Leser in der klassisch gewordenen „Theorie der Kongruenzen" von P. L. TSCHEBYSCHEW.[1])

§ 5. Die Kongruenzen und ihre wichtigsten Eigenschaften

Der Begriff der Kongruenz hat seine Grundlage in der einfachen Bemerkung, daß Zahlen a und b, die bei Division durch eine natürliche Zahl m denselben Rest r lassen, hinsichtlich der Fragen der Teilbarkeit in Beziehung zur Zahl m eine Reihe von gleichen Eigenschaften besitzen. Die wichtigste Eigenschaft besteht darin, daß alle gemeinsamen Teiler von a und m auch gemeinsame Teiler von b und m sind und umgekehrt. Insbesondere ist der größte gemeinsame Teiler (a, m) der Zahlen a und m gleich dem größten gemeinsamen Teiler (b, m) der Zahlen b und m.

In der Tat: Wenn wir voraussetzen, daß die Zahlen a und b bei Division durch m denselben Rest r lassen, so ist

$$a = km + r, \quad b = lm + r$$

und mithin

$$a - b = (k - l)\, m$$

durch m teilbar. Hieraus folgt, daß

$$a = qm + b, \quad b = a - qm,$$

also jeder Teiler von m und b auch Teiler von a und jeder Teiler von m und a auch Teiler von b ist.

Diese gemeinsame Eigenschaft der Zahlen a und b, die sich aus ihrer „Restgleichheit" bei der Division durch m ergibt, erweist sich als so fundamental, daß es sich lohnt, den Begriff der Restgleichheit formal festzulegen, indem man ihr einen besonderen Namen und ein besonderes Zeichen gibt. Es ist üblich geworden, Zahlen a und b, die bei der Division durch m denselben Rest ergeben, *kongruent modulo m* zu nennen und dies durch

$$a \equiv b \pmod{m}$$

zu bezeichnen.

Ein Beispiel: $3 \equiv -17 \pmod 5$.

Die Kongruenz (also die Restgleichheit) zweier Zahlen nach einem gegebenen Modul m macht sie — wie wir gesehen haben — in ihren Beziehungen zu der Zahl m in einem ganz bestimmten Sinne gleichartig. Die Kongruenzrelation ist also eine gewisse Ähnlichkeit zwischen den Zahlen, und die leitende Idee der Theorie der Kongruenzen besteht gerade in der Auswahl dieser Relation und der Verwertung ihrer wichtigsten Eigenschaften. Man muß sich jedoch fest einprägen, daß der Begriff der Kongruenz stets mit einem bestimmten Modul verbunden ist, daß also jene

[1]) П. Л. Чебышев, Полное собрание сочинений, т. I, Теория чисел, (P. L. TSCHEBYSCHEW, Gesammelte Werke, Band I, Zahlentheorie) Verlag d. Akad. Wiss. d. UdSSR, 1944.

Verwandtschaft oder Ähnlichkeit von Zahlen, von der hier die Rede ist, nicht für die Zahlen selbst, sondern nur in ihren Beziehungen zur Zahl m charakteristisch ist. Zahlen, die kongruent nach einem Modul m sind, werden in der Regel nach einem anderen Modul m' nichts miteinander gemein haben.

Bei der Definition der Kongruenz kann die Forderung der Restgleichheit gleichwertig durch eine andere Forderung ersetzt werden, nämlich die Forderung, daß die Differenz der gegebenen Zahlen durch den Modul m teilbar ist, eine Forderung, die wohl zur Kontrolle in konkreten Fällen besser geeignet ist als die ursprüngliche. So ist es im oben angeführten Beispiel nicht notwendig, den Rest von 3 und -17 bei der Division durch 5 zu bestimmen, sondern man braucht sich nur davon zu überzeugen, daß die Differenz dieser Zahlen, also $3 - (-17) = 20$, durch 5 teilbar ist.

Die nachfolgenden fundamentalen Sätze zeigen, daß man mit den Kongruenzen in weitem Rahmen genau so wie mit gewöhnlichen Gleichungen operieren kann.

Satz 1. *Kongruenzen nach ein und demselben Modul kann man seitenweise addieren, subtrahieren und multiplizieren.*

Wenn also

$$\left.\begin{aligned} a &\equiv b \pmod m, \\ a' &\equiv b' \pmod m \end{aligned}\right\} \tag{1}$$

ist, so ist

$$\begin{aligned} a \pm a' &\equiv b \pm b' \pmod m, \\ aa' &\equiv bb' \pmod m. \end{aligned}$$

Wegen (1) ist nämlich

$$\begin{aligned} a - b &= mq, \\ a' - b' &= mq', \end{aligned}$$

also

$$(a \pm a') - (b \pm b') = (a - b) \pm (a' - b') = m\,(q \pm q')$$

und mithin

$$a \pm a' \equiv b \pm b' \pmod m;$$

entsprechend ist

$$aa' - bb' = a\,(a' - b') + b'\,(a - b) = (aq' + b'q)\,m$$

und mithin

$$aa' \equiv bb' \pmod m,$$

was zu beweisen war.

Bemerkung. Insbesondere kann man also zu beiden Seiten einer Kongruenz ein und dieselbe Zahl addieren und beide Seiten einer Kongruenz mit ein und derselben Zahl multiplizieren.

Wir haben den Satz 1 oben nur für den Fall zweier Kongruenzen bewiesen. Natürlich überträgt er sich von zwei unmittelbar auf drei, vier usw., allgemein von n auf $n + 1$ Kongruenzen, womit er nach dem Prinzip der vollständigen Induktion für jede beliebige endliche Anzahl von Kongruenzen bewiesen ist.

Folgerung. *Wenn*

$$a \equiv b \pmod{m} \tag{2}$$

ist, so ist

$$a^k \equiv b^k \pmod{m}$$

für jede natürliche Zahl k und für $k = 0$.

Zum Beweis genügt es, die Kongruenz (2) k-mal seitenweise mit sich selbst zu multiplizieren.

Wenn wir die bisher erhaltenen Resultate miteinander kombinieren, so erhalten wir offensichtlich den folgenden wichtigen

Satz 2. *Wenn $P(x)$ ein beliebiges Polynom mit ganzzahligen Koeffizienten ist, so folgt aus*

$$x \equiv y \pmod{m},$$

daß auch

$$P(x) \equiv P(y) \pmod{m}$$

ist.

Dieser Satz ist für die Arithmetik in der Schule von großem Interesse, bildet er doch die theoretische Grundlage für die Herleitung der wichtigsten Teilbarkeitskriterien. Wenn nämlich die Zahl n im Dezimalsystem von links nach rechts durch die Ziffern $a, b, c, \ldots, k, l$ dargestellt wird, so ist

$$n = a \cdot 10^s + b \cdot 10^{s-1} + c \cdot 10^{s-2} + \cdots + k \cdot 10 + l.$$

Da nun

$$10 \equiv 1 \pmod{3} \quad \text{und} \quad \pmod{9}$$

ist, so ergibt sich aus Satz 2, daß

$$n \equiv a + b + c + \cdots + k + l \pmod{3} \quad \text{und} \quad \pmod{9},$$

also jede Zahl ihrer Quersumme kongruent modulo 3 und modulo 9 ist. Hieraus folgt speziell, daß der größte gemeinsame Teiler einer Zahl n und der Zahl 3 (oder 9) gleich dem größten gemeinsamen Teiler der Quersumme dieser Zahl und 3 (oder 9) ist. Insbesondere ist eine Zahl n dann und nur dann durch 3 (oder 9) teilbar, wenn ihre Quersumme durch 3 (oder 9) teilbar ist. Durch eine ähnliche Überlegung erhält man aus der Kongruenz

$$10 \equiv -1 \pmod{11}$$

gemäß Satz 2:

$$n \equiv (-1)^s a + (-1)^{s-1} b + \cdots - k + l \pmod{11}$$

und hieraus unmittelbar das bekannte Kriterium für die Teilbarkeit einer Zahl durch 11.

Satz 1 sagt zunächst nur etwas aus über die unbeschränkte Ausführbarkeit der Operationen Addition, Subtraktion und Multiplikation von Kongruenzen; er sagt jedoch nichts über die vierte Grundrechenart, die Division. Insbesondere wissen wir nicht, ob die Division beider Seiten einer Kongruenz durch ein und dieselbe Zahl (unter der Voraussetzung, daß sie ohne Rest durchführbar ist) erlaubt ist. Es ist kein Zufall, daß wir die Behandlung dieser Frage zurückgestellt haben. Wir

stoßen hier zum erstenmal auf eine Situation, in der sich die Kongruenzen anders als die Gleichungen verhalten. Daher wollen wir uns über diese Frage eingehend Klarheit verschaffen.

Einfache Beispiele zeigen bereits, daß die Division von Kongruenzen, auch wenn sie ohne Rest möglich ist, nicht immer durchführbar ist. So sind in der Kongruenz

$$45 \equiv 27 \pmod 6$$

beide Seiten durch die Zahl 9 teilbar. Wenn wir jedoch diese Kongruenz seitenweise durch 9 teilen, so erhalten wir das falsche Ergebnis

$$5 \equiv 3 \pmod 6.$$

Wir wollen jetzt sogleich den allgemeinen Fall betrachten. Dazu sei

$$a \equiv b \pmod m, \tag{2}$$

wobei die Zahlen a und b durch ein und dieselbe Zahl d teilbar seien, also

$$a = da', \quad b = db'$$

ist. Es erhebt sich dann die Frage, unter welchen Bedingungen die Kongruenz (2) durch d gekürzt werden kann, d. h. unter welchen Bedingungen

$$a' \equiv b' \pmod m$$

ist. Zunächst besagt offenbar die Kongruenz (2), daß die Differenz $a - b = d(a' - b')$ durch m teilbar ist. Unter welchen Bedingungen folgt nun hieraus, daß $a' - b'$ durch m teilbar ist? Hierauf gibt der Satz 1 aus Kapitel I eine Antwort. Er sagt uns, daß das sicher dann der Fall ist, wenn d und m teilerfremd sind. Wir erhalten also als wichtige Regel: *Man kann beide Seiten einer Kongruenz durch einen gemeinsamen Teiler kürzen, wenn dieser Teiler zum Modul m teilerfremd ist.* Umgekehrt können wir sagen, daß die Division beider Seiten einer Kongruenz durch eine zum Modul m nicht teilerfremde Zahl d im allgemeinen zu einem falschen Ergebnis führt. So zeigt uns das obige Beispiel, daß die Division zu einem falschen Resultat führte, weil der Teiler 9 nicht zum Modul 6 teilerfremd ist.

Dieser eben festgestellte Unterschied zwischen dem Verhalten der Kongruenzen und dem der Gleichungen hat seine Ursache in dem überaus wichtigen Umstand, daß sich die Kongruenzen einem der Hauptprinzipien der Teilbarkeitstheorie nicht unterordnen, nämlich dem Prinzip, *daß das Produkt zweier Zahlen nur dann gleich Null ist, wenn mindestens einer der Faktoren gleich Null ist.* In der Theorie der Kongruenzen würde das entsprechende Prinzip offensichtlich lauten: Wenn das Produkt zweier Zahlen kongruent Null modulo m ist, so ist mindestens einer der Faktoren kongruent Null modulo m. Nun besagt aber „kongruent Null modulo m" nichts anderes als die Teilbarkeit durch m, so daß das genannte Prinzip gleichwertig ist mit: Wenn das Produkt zweier Zahlen durch m teilbar ist, so ist wenigstens einer der Faktoren durch m teilbar. Aber das ist im allgemeinen falsch. So ist $4 \cdot 15 = 60$ durch 6 teilbar, während weder 4 noch 15 durch 6 teilbar ist. Eine ähnliche unerlaubte Verwendung dieses Prinzips führte uns — wie man leicht einsieht — auch im vorigen Beispiel zu dem falschen Resultat.

Jedoch zeigt uns Satz 3 aus Kapitel I, daß es einen Fall gibt, in dem dieses Prinzip gilt. Wenn nämlich der Modul eine Primzahl p ist, so folgt in der Tat aus der Teilbarkeit eines Produktes zweier Zahlen durch p unbedingt die Teilbarkeit mindestens eines der Faktoren durch p. Diese bemerkenswerte Tatsache hat zur Folge, daß

die Kongruenzen nach einem Primzahlmodul den Gleichungen bedeutend weitergehend analog sind als die Kongruenzen nach einem zusammengesetzten Modul.

Insbesondere kann man hier in gewissem Sinne sagen, daß sich die Frage nach der Möglichkeit der Division beider Seiten einer Kongruenz durch eine gewisse Zahl im Fall eines Primzahlmoduls genauso beantworten läßt wie die entsprechende Frage für Gleichungen. In der Tat: Wir haben bereits oben gesehen, daß man beide Seiten einer Kongruenz immer dann durch eine Zahl d teilen kann, wenn d zum Modul m teilerfremd ist. Wenn nun m eine Primzahl ist, so ist das „zu m teilerfremd sein" gleichbedeutend mit dem „nicht durch m teilbar sein", oder — was wiederum damit gleichwertig ist — mit dem „nicht kongruent Null modulo m sein". Es ist also im Falle des Primzahlmoduls lediglich die Division durch Zahlen verboten, die kongruent Null nach dem gegebenen Modul sind. Bei der Analogie zwischen den Kongruenzen und den Gleichungen entspricht aber diesen Zahlen die Zahl Null, durch die auch in der gewöhnlichen Teilbarkeitstheorie nicht dividiert werden kann. Wir sehen damit, daß sich in diesem Fall — wie auch in vielen anderen Fällen — die Kongruenzen nach einem Primzahlmodul genauso wie die Gleichungen verhalten. Was jedoch die Kongruenzen nach einem zusammengesetzten Modul betrifft, so beobachten wir hier wesentlich andere Gesetzmäßigkeiten als in der Theorie der Gleichungen.

§ 6. Klassifikation der Zahlen nach einem gegebenen Modul

Wir haben Zahlen nach einem gegebenen Modul m kongruent genannt, wenn sie bei Division durch m denselben Rest lassen. Hieraus ergibt sich unmittelbar, daß aus der gemeinsamen Kongruenz modulo m zweier Zahlen a und b zu einer Zahl c folgt, daß auch $a \equiv b \pmod{m}$ ist. Dieser Umstand gestattet es uns, alle Zahlen, die einer gegebenen Zahl a (und damit auch untereinander) kongruent modulo m sind, in einer Klasse zusammenzufassen. Auf diese Weise wird die Menge aller ganzen Zahlen in Restklassen modulo m eingeteilt. Alle Zahlen einer und derselben Klasse sind dabei kongruent modulo m, während Zahlen aus verschiedenen Klassen stets untereinander inkongruent modulo m sind.[1])

[1]) Der Leser, der den vorangehenden Artikel von I. W. Proskurjakow (im folgenden als P. zitiert) studiert hat, sei auf folgendes hingewiesen: Man überzeugt sich leicht davon, daß die oben festgestellte *Drittengleichheit* [wenn $a \equiv c \pmod{m}$ und $b \equiv c \pmod{m}$ ist, so ist $a \equiv b \pmod{m}$] in Verbindung mit dem für die Restgleichheit ebenfalls geltenden reflexiven Gesetz [$a \equiv a \pmod{m}$] gleichwertig ist der Gültigkeit des Gesetzes der Symmetrie [wenn $a \equiv b \pmod{m}$ ist, so ist $b \equiv a \pmod{m}$] und des transitiven Gesetzes [wenn $a \equiv b \pmod{m}$ und $b \equiv c \pmod{m}$ ist, so ist $a \equiv c \pmod{m}$]. Die Kongruenzrelation ist also eine Äquivalenzrelation im Bereich der ganzen Zahlen im Sinne von P. § 19. Insbesondere gilt also für die Kongruenzrelation der Hauptsatz der Theorie der Äquivalenzrelationen (P. § 19, Satz). Die Restklassen modulo m sind dann nichts anderes als die Äquivalenzklassen der Kongruenzrelation. Führt man schließlich zwischen den Restklassen modulo m durch Rechnen mit Repräsentanten eine Addition und eine Multiplikation ein (vgl. P. § 8, Beispiel eines Körpers der Charakteristik $p > 0$), so erhält man einen Ring aus genau m Elementen, den *Restklassenring modulo m*. Dieser Ring ist (wie bereits in P. § 8 gezeigt wurde) dann und nur dann ein Körper, wenn m eine Primzahl ist. Dies ist der tiefere Grund für die am Ende des vorangehenden Paragraphen festgestellten Abweichungen zwischen den Kongruenzen nach einem Primzahlmodul und den Kongruenzen nach einem zusammengesetzten Modul. Andererseits erklärt sich aus der Ring- bzw. Körpereigenschaft die weitgehende Analogie zwischen den Gleichungen und den Kongruenzen (vgl. die Ausführungen über Ringe und Körper in P. § 7 und § 8). — *Anm. d. wissenschaftl. Red.*

Wie die Anzahl und die gegenseitigen Beziehungen der Klassen, läßt sich auch ihre innere Struktur leicht übersehen. Da alle Zahlen ein und derselben Klasse bei der Division durch m denselben Rest lassen und als Reste bei der Division durch den Modul m nur die Zahlen

$$0, 1, 2, \ldots, m-1$$

in Frage kommen, muß m die Anzahl der Restklassen modulo m sein. Dabei ist jede Klasse durch den jeweiligen Rest r ($0 \leq r \leq m-1$) eindeutig festgelegt, und die zugehörige Klasse besteht aus allen Zahlen der Form $mx + r$, wobei x eine beliebige ganze Zahl ist. Offensichtlich bilden diese Zahlen eine nach beiden Seiten unendliche arithmetische Folge der Differenz m. Auf diese Weise erscheint die Zerlegung der Menge der ganzen Zahlen in Restklassen modulo m als nichts anderes als eine Zerlegung in m arithmetische Folgen der Differenz m.

Wenn wir aus jeder der Klassen willkürlich eine Zahl auswählen, so erhalten wir eine Menge von m Zahlen, die dadurch ausgezeichnet ist, daß keine zwei Zahlen dieser Menge einander kongruent modulo m sind und daß andererseits jede ganze Zahl einer Zahl der betrachteten Mengen kongruent modulo m ist. Eine derartige Menge von ganzen Zahlen nennt man ein *vollständiges Restsystem modulo m*. Offenbar gibt es unendlich viele vollständige Restsysteme nach einem gegebenen Modul. So können als vollständiges Restsystem modulo 3 jedes der Tripel (0, 1, 2), (10, 11, 12), (—4, 6, —5) und noch unendlich viele weitere dienen.

In vielen Fragen der Teilbarkeitstheorie verhalten sich nun — wie wir bereits in einigen Fällen festgestellt haben — die Zahlen aus ein und derselben Restklasse völlig gleichartig und können nach Belieben durch einander ersetzt werden. Für derartige Fragen ist es daher gänzlich gleichgültig, welches der unzählig vielen vollständigen Restsysteme man für die jeweiligen Überlegungen auswählt. Dieser Umstand, der eine gewisse Willkür schafft, kann nun in vielen Fällen zu einer wesentlichen Vereinfachung der Rechnungen ausgenutzt werden (z. B — wie im Falle der Regeln über die Teilbarkeit einer Zahl durch 3, 9 oder 11 — können unter Umständen große Zahlen durch bedeutend kleinere ersetzt werden).

Für die theoretischen Untersuchungen über den Begriff des vollständigen Restsystems spielt eine erhebliche Rolle der folgende

Satz 3. *Wenn die Zahlen a und m zueinander teilerfremd sind und die Zahl x in dem Ausdruck $ax + b$ ein vollständiges Restsystem modulo m durchläuft, so durchlaufen auch die Werte dieses Ausdrucks ein vollständiges Restsystem modulo m.*

Da die Anzahl der Werte des Ausdrucks $ax + b$ gleich m ist, genügt es, sich davon zu überzeugen, daß sie ein Restsystem modulo m bilden, d. h., daß die Werte paarweise verschiedenen Restklassen modulo m angehören: Würde für zwei Zahlen x_1 und x_2 aus verschiedenen Restklassen modulo m

$$ax_1 + b \equiv ax_2 + b \pmod{m}$$

gelten, so wäre auch

$$ax_1 \equiv ax_2 \pmod{m};$$

da nun aber nach Voraussetzung a und m teilerfremd sind, so kann man diese letzte Kongruenz durch a kürzen und erhält

$$x_1 \equiv x_2 \pmod{m},$$

was nicht der Fall sein sollte. Damit ist bereits der behauptete Satz 3 bewiesen. Wir werden bald wichtigen Anwendungen dieses Satzes begegnen.

Wir haben bereits gesehen, daß alle Zahlen einer Restklasse mit dem Modul dieselben gemeinsamen Teiler, also insbesondere denselben größten gemeinsamen Teiler besitzen. Wenn also eine Zahl aus einer Restklasse zum Modul teilerfremd ist, so sind auch alle anderen Zahlen aus dieser Restklasse zum Modul teilerfremd. Es hat daher einen Sinn, von der Teilerfremdheit einer Restklasse zum Modul m zu sprechen. Unter einem *reduzierten* (im Gegensatz zum vollständigen) Restsystem nach einem gegebenen Modul m verstehen wir dann eine Menge von ganzen Zahlen, die aus jeder zu m teilerfremden Restklasse genau einen Vertreter enthält. Eine einfache Methode, um ein reduziertes Restsystem zu erhalten, besteht offensichtlich darin, daß man aus der Folge der Zahlen

$$1, 2, \ldots, m,$$

die jedenfalls ein vollständiges Restsystem modulo m darstellt, diejenigen Zahlen auswählt, die zu m teilerfremd sind. Hieraus folgt, daß die Anzahl der zu m teilerfremden Restklassen (und damit die Anzahl der Elemente jedes reduzierten Restsystems) gleich der Anzahl der natürlichen Zahlen kleiner oder gleich m ist, die zu m teilerfremd sind. Diese Anzahl, die offenbar nur von m abhängt, wird üblicherweise mit $\varphi(m)$ bezeichnet. Die so erhaltene zahlentheoretische Funktion (die sogenannte *Eulersche Funktion — Anm. d. wissenschaftl. Red.*) stellt eine der wichtigsten Funktionen der Zahlentheorie überhaupt dar. Wir werden später zeigen, wie man den Wert dieser Funktion für ein gegebenes Argument m einfach aus der Primzahlzerlegung von m bestimmen kann.

Für jedes reduzierte Restsystem gilt der folgende, dem Satz 3 für vollständige Restsysteme analoge

Satz 4. *Wenn die Zahlen a und m teilerfremd sind und die Zahl x in dem Ausdruck ax ein reduziertes Restsystem modulo m durchläuft, so durchlaufen auch die Werte dieses Ausdrucks ein reduziertes Restsystem modulo m.*

Man muß dabei jedoch beachten, daß die in Satz 3 noch auftretende Zahl b in Satz 4 gleich Null ist. Das zeigt, daß die in Satz 3 formulierte Aussage über vollständige Restsysteme doch weitergehend ist als die in Satz 4 aufgestellte Behauptung über reduzierte Restsysteme.

Zum Beweis von Satz 4 genügt es zu bemerken, daß mit x auch alle $\varphi(m)$ Produkte ax zu m teilerfremd sind und sämtlich — wie bereits in Satz 3 gezeigt wurde — verschiedenen Restklassen modulo m angehören.

Als eine einfache, aber interessante und überaus wichtige Folgerung aus Satz 4 erhalten wir einen Satz, der auf Euler zurückgeht. Dazu seien a und m zwei teilerfremde Zahlen und

$$r_1, r_2, \ldots, r_s \tag{3}$$

ein reduziertes Restsystem modulo m (hier ist zur Abkürzung $\varphi(m) = s$ gesetzt). Dann bilden nach Satz 4 auch die Zahlen

$$ar_1, ar_2, \ldots, ar_s \tag{4}$$

ein reduziertes Restsystem modulo m, d. h., jede der Zahlen (4) ist einer der Zahlen (3) kongruent modulo m, also etwa

$$\left.\begin{aligned} ar_1 &\equiv r_{i_1} \\ ar_2 &\equiv r_{i_2} \\ &\cdots\cdots\cdots \\ ar_s &\equiv r_{i_s} \end{aligned}\right\} \pmod{m},$$

wobei die Zahlen $i_1, i_2, \ldots, i_s$ eine Permutation der Zahlen $1, 2, \ldots, s$ bilden. Wenn wir diese Kongruenzen seitenweise miteinander multiplizieren, so erhalten wir

$$a^{\varphi(m)} r_1 r_2 \cdots r_s \equiv r_{i_1} r_{i_2} \cdots r_{i_s} = r_1 r_2 \cdots r_s \pmod{m}.$$

Da nun jedes r_i zu m teilerfremd ist, ist auch ihr Produkt zu m teilerfremd, und wir können diese letzte Kongruenz durch das Produkt $r_1 r_2 \cdots r_s$ kürzen. Auf diese Weise kommen wir zu dem erwähnten Satz von EULER, der besagt, daß *für teilerfremde Zahlen a und m stets*

$$a^{\varphi(m)} \equiv 1 \pmod{m}$$

ist.

Beispiele: $\varphi(10) = 4$; $3^4 = 81 \equiv 1 \pmod{10}$; $7^4 = 2401 \equiv 1 \pmod{10}$.

Im Spezialfall, daß der Modul eine Primzahl p ist, ist der Satz bereits vor EULER von FERMAT ausgesprochen worden. Wenn nämlich p eine Primzahl ist, so sind alle Zahlen der Folge $1, 2, \ldots, p$ mit Ausnahme von p selbst zu p teilerfremd, so daß $\varphi(p) = p - 1$ ist, und wir erhalten den

Satz von FERMAT. *Wenn p eine Primzahl ist und a nicht von p geteilt wird, so ist*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Bemerkung. Diesen letzten Satz bezeichnet man häufig als „kleinen FERMATschen Satz", im Unterschied zum sogenannten „großen FERMATschen Satz", nach dem die Gleichung $x^n + y^n = z^n$ für ganzzahliges $n > 2$ keine ganzzahlige Lösung x, y, z besitzt (diese Behauptung, für die FERMAT nach eigenen Aussagen einen Beweis besaß, konnte bis heute noch nicht bewiesen werden). Zu dieser Bezeichnung ist zu sagen, daß — wenn es nach der Wichtigkeit dieser Sätze im Rahmen der Gesamtentwicklung der Mathematik ginge — wahrscheinlich die umgekehrte Terminologie angebrachter wäre. Denn wenn wirklich der „große Fermat" einmal bewiesen sein sollte, so dürfte diese Tatsache — soweit hier eine Voraussage möglich ist — der Wissenschaft kaum eine Stütze für weitere Errungenschaften bieten. Er wird wahrscheinlich immer mehr oder weniger isoliert dastehen. Demgegenüber stellt der eben bewiesene „kleine FERMATsche Satz" schon seit langem ein wichtiges Hilfsmittel der Forschung dar, und dies nicht nur in der Theorie der ganzen Zahlen, sondern auch in wesentlich umfangreicheren Gebieten der Arithmetik und der Algebra.

Wir gehen jetzt zur Berechnung der Funktion $\varphi(m)$ über, die die Anzahl der zu m teilerfremden natürlichen Zahlen, die nicht größer als m sind, angibt.

Dazu zeigen wir zunächst, daß *für beliebige teilerfremde Zahlen m und n*

$$\varphi(mn) = \varphi(m)\,\varphi(n)$$

ist.

Zur Berechnung von $\varphi(mn)$ ordnen wir die Zahlen von 1 bis mn in folgender übersichtlichen Tabelle an

1	2	$\cdots$	k	$\cdots$ m
$m+1$	$m+2$	$\cdots$	$m+k$	$\cdots$ $2m$
$2m+1$	$2m+2$	$\cdots$	$2m+k$	$\cdots$ $3m$
$\cdots\cdots$	$\cdots\cdots$		$\cdots\cdots$	$\cdots$
$(n-1)m+1$	$(n-1)m+2$	$\cdots$	$(n-1)m+k$	$\cdots$ nm

und versuchen, die Anzahl der zu mn teilerfremden Zahlen in dieser Tabelle zu bestimmen. Damit nun eine Zahl zu mn teilerfremd ist, muß sie sowohl zu m als auch zu n teilerfremd sein. Andererseits sind aber auch die Zahlen, die zu m und zu n teilerfremd sind, zu mn teilerfremd, da nach Voraussetzung m und n teilerfremd sind. Wir können also unsere Berechnung so durchführen, daß wir zunächst in der obigen Tabelle die Zahlen auswählen, die zu m teilerfremd sind, und dann unter diesen diejenigen Zahlen bestimmen, die zu n teilerfremd sind. So wollen wir auch im folgenden vorgehen.

In unserer Tabelle gehören nun offensichtlich alle Zahlen, die in einer Spalte stehen, ein und derselben Restklasse modulo m an. Das bedeutet, daß sie entweder alle zu m teilerfremd sind oder alle einen mit m gemeinsamen Teiler haben, der verschieden von 1 ist. Wir können daher von „den Spalten, die zu m teilerfremd sind" reden. Die Anzahl dieser Spalten ist nun sehr einfach zu bestimmen: Da nämlich in der ersten Zeile unserer Tabelle die Zahlen $1, 2, \ldots, m$ stehen, ist die Anzahl der zu m teilerfremden Spalten gleich der Anzahl der zu m teilerfremden Zahlen in der Folge $1, 2, \ldots, m$, also gerade gleich $\varphi(m)$.

Wir wählen dann eine beliebige dieser $\varphi(m)$ zu m teilerfremden Spalten aus. Dies möge etwa die Spalte aus den Zahlen

$$k, m + k, 2m + k, \ldots, (n - 1)\, m + k \tag{5}$$

sein. Gefragt ist nach der Anzahl der zu n teilerfremden Zahlen in dieser Spalte. Da nun die Zahlen dieser Spalte sämtlich die Form $mx + k$ besitzen, wobei x die Zahlen $0, 1, 2, \ldots, n - 1$ durchläuft, bilden nach Satz 3 die Zahlen dieser Spalte ein vollständiges Restsystem modulo n. Weil aber jedes vollständige Restsystem modulo n genau $\varphi(n)$ zu n teilerfremde Zahlen enthält, enthält die betrachtete Spalte und damit jede zu m teilerfremde Spalte $\varphi(n)$ Zahlen, die zu n teilerfremd sind.

Wir fassen zusammen: Unsere Tabelle enthält $\varphi(m)$ Spalten, die zu m teilerfremd sind, und jede derartige Spalte enthält $\varphi(n)$ Zahlen, die zu n teilerfremd sind. Also enthält die Tabelle genau $\varphi(m) \cdot \varphi(n)$ Zahlen, die sowohl zu m als auch zu n teilerfremd sind, und mithin genau $\varphi(m) \cdot \varphi(n)$ Zahlen, die zu mn teilerfremd sind. Es ist also in der Tat

$$\varphi(mn) = \varphi(m)\, \varphi(n),$$

was zu beweisen war.

Nun bereitet es aber keine wesentlichen Schwierigkeiten mehr, einen geschlossenen Ausdruck für die Funktion $\varphi(m)$ anzugeben. Dazu denken wir uns die Zahl m in ihrer Primzahlzerlegung hingeschrieben:

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r},$$

wobei $p_1, p_2, \ldots, p_r$ paarweise verschiedene Primzahlen und $\alpha_1, \alpha_2, \ldots, \alpha_r$ gewisse natürliche Zahlen sind. Nach der bereits bewiesenen Eigenschaft der Funktion $\varphi(m)$ ist also

$$\varphi(m) = \varphi(p_1^{\alpha_1})\, \varphi(p_2^{\alpha_2}) \cdots \varphi(p_r^{\alpha_r}). \tag{6}$$

Dabei ist $\varphi(p_i^{\alpha_i})$ $(1 \leq i \leq r)$ die Anzahl der natürlichen Zahlen, die nicht größer als $p_i^{\alpha_i}$ und zu $p_i^{\alpha_i}$ teilerfremd, d. h. nicht durch p_i teilbar sind. Zur Berechnung dieser Anzahl bemerken wir, daß in der Folge der Zahlen

$$1, 2, \ldots, p_i^{\alpha_i}$$

nur die Zahlen

$$p_i,\ 2p_i,\ 3p_i,\ \ldots,\ p_i^{\alpha_i-1}p_i = p_i^{\alpha_i},$$

also $p_i^{\alpha_i-1}$ Stück, durch p_i teilbar sind, während die restlichen

$$p_i^{\alpha_i} - p_i^{\alpha_i-1}$$

Zahlen dieser Folge nicht von p_i geteilt werden, also zu $p_i^{\alpha_i}$ teilerfremd sind. Damit haben wir gezeigt, daß

$$\varphi(p_i^{\alpha_i}) = p_i^{\alpha_i} - p_i^{\alpha_i-1} = p_i^{\alpha_i}\left(1 - \frac{1}{p_i}\right)$$

ist, woraus sich gemäß (6) die Gleichung

$$\begin{aligned}\varphi(m) &= p_1^{\alpha_1}\left(1-\frac{1}{p_1}\right)\ p_2^{\alpha_2}\left(1-\frac{1}{p_2}\right)\cdots p_r^{\alpha_r}\left(1-\frac{1}{p_r}\right)\\ &= m\left(1-\frac{1}{p_1}\right)\cdot\left(1-\frac{1}{p_2}\right)\cdots\left(1-\frac{1}{p_r}\right)\end{aligned}$$

ergibt.

Damit ist die Aufgabe der Berechnung von $\varphi(m)$ in befriedigender Weise gelöst. Es ist an der gefundenen Formel interessant, daß der von uns erhaltene Ausdruck von $\varphi(m)$ nicht explizit von den Exponenten α_i abhängt, zu denen die Primfaktoren p_i in m vorkommen. Es ist also zur Berechnung von $\varphi(m)$ nicht notwendig, jene Exponenten zu kennen, sondern man braucht nur zu wissen, welche Primzahlen in die Primzahlzerlegung von m eingehen.

Beispiel.

$$120 = 2^3\cdot 3\cdot 5;$$

$$\varphi(120) = 120\left(1-\frac{1}{2}\right)\cdot\left(1-\frac{1}{3}\right)\cdot\left(1-\frac{1}{5}\right) = 120\cdot\frac{1}{2}\cdot\frac{2}{3}\cdot\frac{4}{5} = 32.$$

Wir bemerken noch folgendes: Anläßlich des Beweises des Satzes von EULER in § 3 benötigten wir die Anzahl Q_n derjenigen natürlichen Zahlen von 1 bis $P_n = p_1p_2\cdots p_n$, die durch keine der Primzahlen $p_1, p_2, \ldots, p_n$ teilbar sind. Es ist offensichtlich, daß diese Anzahl keine andere als die Zahl $\varphi(P_n)$ ist. Dann ist aber die Formel für Q_n, die wir auf Seite 236 ohne Beweis angegeben haben, ein Spezialfall unserer allgemeinen Formel für $\varphi(m)$.

§ 7. Kongruenzen, die Unbekannte enthalten

Man unterteilt die Gleichungen gewöhnlich in Identitäten und Bestimmungsgleichungen. Eine ganz entsprechende Unterteilung können wir auch bei den Kongruenzen vornehmen. Wir unterscheiden demgemäß zwischen identischen Kongruenzen (das sind Kongruenzen, die entweder keine Variable enthalten oder bei beliebigen Werten für die in ihnen enthaltenen Variablen erfüllt werden) und Kongruenzen, die Unbekannte enthalten (das sind Kongruenzen, die nur bei ganz spezieller Wahl der Werte für die in ihnen auftretenden Variablen befriedigt werden).

Als Beispiele für identische Kongruenzen mögen die Kongruenzen

$$103 \equiv 1 \pmod{17}, \quad (a+b)^2 \equiv a^2 \pmod{b}$$

dienen, während ein Beispiel für eine Kongruenz in einer Unbekannten die Kongruenz

$$x^2 + 1 \equiv 0 \pmod{10}$$

ist.

Wir werden im folgenden ausschließlich Kongruenzen in einer Unbekannten betrachten. Eine solche Kongruenz nennen wir *algebraisch vom Grade n*, wenn sie die Form

$$P(x) \equiv 0 \pmod{m}$$

hat, wobei $P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n$ ein Polynom n-ten Grades mit ganzzahligen Koeffizienten $a_0, a_1, \ldots, a_n$ und $a_0 \not\equiv 0 \pmod{m}$ (also a_0 nicht durch den Modul m teilbar) ist. Der Begriff der algebraischen Kongruenz n-ten Grades ist also völlig analog dem Begriff der algebraischen Gleichung n-ten Grades, bei dem entsprechend verlangt wird, daß der Koeffizient von x^n verschieden von Null ist.

Auf Grund von Satz 2 ergibt sich unmittelbar, daß mit einer Zahl x_0 auch jede Zahl x, die zu x_0 kongruent modulo m ist, eine gegebene algebraische Kongruenz modulo m erfüllt. Für die algebraischen Kongruenzen ist also charakteristisch, daß ihre Wurzeln ganze Restklassen nach dem gegebenen Modul bilden. Daher ist es üblich, als *Lösung* einer algebraischen Kongruenz modulo m nicht eine einzelne Zahl, sondern die ganze Restklasse (modulo m) zu bezeichnen, der die Zahl angehört, die jene Kongruenz befriedigt. Dementsprechend versteht man unter der Anzahl der Lösungen einer algebraischen Kongruenz modulo m nicht die Anzahl der Zahlen, die diese Kongruenz erfüllen (hier gibt es entweder gar keine oder gleich unendlich viele), sondern die Anzahl der Restklassen modulo m, die eine die Kongruenz erfüllende Zahl enthalten.

Wir wollen zunächst den wichtigsten Spezialfall der algebraischen Kongruenz, den der linearen Kongruenz (d. h. der Kongruenz ersten Grades) in einer Unbekannten, eingehend betrachten, die wir in der allgemeinen Form

$$ax \equiv b \pmod{m} \tag{7}$$

ansetzen können.

Wenn die Zahl a zum Modul m teilerfremd ist, so durchläuft nach Satz 3 mit x auch das Produkt ax ein vollständiges Restsystem modulo m, so daß eines und nur eines dieser Produkte kongruent der gegebenen Zahl b ist. Unsere Kongruenz hat also in diesem Fall, analog den Gleichungen ersten Grades in einer Unbekannten, genau eine Lösung.

Eine Möglichkeit zum effektiven Auffinden der Lösung gibt uns der EULERsche Satz. Da nämlich

$$a^{\varphi(m)} \equiv 1 \pmod{m}$$

ist, so ist

$$b a^{\varphi(m)} \equiv b \pmod{m},$$

und wir sehen, daß die Zahl $b a^{\varphi(m)-1}$ die Kongruenz (7) löst. Da die Eindeutigkeit

dieser Lösung bereits feststeht, wird die vollständige Lösung durch

$$x \equiv b a^{\varphi(m)-1} \pmod{m} \tag{8}$$

gegeben.

Offenbar ist der betrachtete Spezialfall die Regel, wenn m eine Primzahl ist. Im Fall des Primzahlmoduls fällt nämlich die bei der Definition des Grades einer Kongruenz gestellte Forderung, daß der Modul kein Teiler des höchsten Koeffizienten ist, mit der Forderung, daß der Modul zum höchsten Koeffizienten teilerfremd ist, zusammen. Daher *besitzt jede Kongruenz ersten Grades nach einem Primzahlmodul m stets genau eine Lösung*, die auf Grund des FERMATschen Satzes durch

$$x \equiv b a^{m-2} \pmod{m} \tag{9}$$

gegeben wird. Es muß allerdings bei dieser Gelegenheit darauf hingewiesen werden, daß in den meisten Fällen die praktische Bestimmung der Lösung mit Hilfe der Formeln (8) oder (9) nicht den kürzesten Weg darstellt. Dieser führt vielmehr über den Euklidischen Algorithmus (vgl. Kapitel III). Die vorangehenden Überlegungen zeigen jedoch, daß sich im Falle eines Primzahlmoduls auch die Lösung linearer Kongruenzen nach ähnlichen Gesetzen wie die Lösung linearer Gleichungen vollzieht.

Als Lösung der Kongruenz (7) kann jede Zahl x dienen, für die die Differenz $ax - b$ durch m teilbar ist, also die Form my hat, wobei y eine gewisse ganze Zahl ist. Somit ist die Aufgabe, eine Lösung der Kongruenz (7) zu finden, gleichwertig mit der Aufgabe, die Gleichung

$$ax - b = my$$

oder, was dasselbe ist, die Gleichung

$$ax - my = b \tag{10}$$

in ganzen Zahlen x und y zu lösen. (10) ist offensichtlich die allgemeine Form der Gleichung ersten Grades in zwei Unbekannten. Wir können also feststellen, daß alle Aussagen der Theorie der linearen Kongruenzen in einer Unbekannten im angegebenen Sinne auch als Aussagen der Theorie der „diophantischen Gleichungen" (das ist die Lehre von den ganzzahligen Lösungen algebraischer Gleichungen) ersten Grades in zwei Unbekannten gedeutet werden können. Insbesondere kann das von uns erhaltene Hauptresultat auf folgende Weise als Resultat über diophantische Gleichungen formuliert werden:

Wenn die Zahlen a und m teilerfremd sind, so besitzt die Gleichung (10) *stets ganzzahlige Lösungen. Ist (x_0, y_0) eine beliebige ganzzahlige Lösung von* (10), *so erhält man alle ganzzahligen Lösungen von* (10) *in der Form*

$$x = x_0 + mk, \quad y = y_0 + ak,$$

wobei k eine beliebige ganze Zahl ist.

Im Spezialfall, daß $b = 1$ ist, haben wir die Aufgabe, eine Lösung der Gleichung (10) (für teilerfremdes a und m) zu finden, bereits in Kapitel I behandelt. Dort haben wir (Satz 1) die Existenz der Lösung mit Hilfe des GAUSSschen Verfahrens erschlossen. Hier haben wir einen zweiten Beweis für dieselbe Tatsache gegeben.[1])

[1]) Dem Umstand, daß dieser neue Beweis einen Hinweis auf Satz 2 aus Kapitel I enthält, entspringt kein Zirkelschluß, da wir in Kapitel I gezeigt haben (S. 232), daß Satz 2 auch unabhängig von Satz 1 bewiesen werden kann.

Dieser neue Beweis ist aber in einem erheblichen Maße befriedigender als der erste, da er uns neben der Existenz einer Lösung auch einen Überblick über sämtliche ganzzahligen Lösungen der Gleichung (10) gibt. Der hier angegebene Weg zur Auffindung der Lösungen ist aber — wie bereits erwähnt — nicht der kürzeste. Wir werden deshalb im nächsten Kapitel nochmals auf diese Aufgabe zurückkommen und dort einen einfachen Weg zur praktischen Bewältigung der gestellten Aufgabe zeigen.

Wir kommen jetzt auf den Fall zurück, daß der größte gemeinsame Teiler d der Zahlen a und m größer als 1, also

$$(a, m) = d > 1$$

ist. Dann ist $a = da'$, $m = dm'$, wobei $(a', m') = 1$ ist, also a' und m' teilerfremd sind. Wenn in diesem Fall d kein Teiler von b ist, so besitzt die Kongruenz (7) keine Lösung. Ist nämlich x eine Lösung der Kongruenz (7), so ist

$$ax - b = my, \quad b = ax - my = d(a'x - m'y)$$

und d muß ein Teiler von b sein. Falls aber d ein Teiler von b ist, so ist $b = db'$ und die Teilbarkeit von $ax - b = d(a'x - b')$ durch $m = dm'$ mit der Teilbarkeit von $a'x - b'$ durch m' gleichbedeutend, d. h., die Kongruenz (7) ist mit der Kongruenz

$$a'x \equiv b' \pmod{m'}$$

gleichwertig. Da aber in dieser letzten Kongruenz $(a', m') = 1$ ist, besitzt sie — wie wir bereits wissen — als Lösung genau eine Restklasse modulo m'. Die Zahlen, die die Kongruenz (7) lösen, bilden also im betrachteten Fall eine Restklasse modulo $m' = \frac{m}{d}$. Jedoch sieht man unmittelbar, daß jede Restklasse modulo m' in d Restklassen modulo m zerfällt. In der Tat: Wenn man die Zahlen einer Restklasse modulo m' in der Form $x_0 + m'z$ darstellt, wobei z eine beliebige ganze Zahl ist, so gehören offensichtlich die Zahlen

$$x_0, x_0 + m', x_0 + 2m', \ldots, x_0 + (d-1)m'$$

aus dieser Restklasse modulo m' sämtlich verschiedenen Restklassen modulo m an, während für die weiteren Zahlen die Beziehungen

$$x_0 + dm' = x_0 + m \equiv x_0 \pmod{m},$$

$$x_0 + (d+1)m' = x_0 + m + m' \equiv x_0 + m' \pmod{m}$$

usw. gelten. Wir stellen also fest, daß im zuletzt betrachteten Fall die Kongruenz (7) im Sinne der üblichen Terminologie genau d Lösungen besitzt.

Damit ist die Frage nach der Lösungsmannigfaltigkeit einer Kongruenz ersten Grades in einer Unbekannten in allen Einzelheiten geklärt. Das Ergebnis können wir übersichtlich in dem folgenden Satz zusammenfassen:

Satz 5. *Wenn $(a, m) = d$ ist, so besitzt die Kongruenz (7) entweder genau d Lösungen, wenn nämlich b durch d teilbar ist, oder gar keine Lösung, wenn b nicht durch d teilbar ist.*

In diesem Satz ist der vorher betrachtete Fall $d = 1$ enthalten, ohne irgendwelche Nebenbedingungen zu erfordern.

Wir können dieses Resultat auch gleichwertig in der Sprache der diophantischen Gleichungen ersten Grades in zwei Unbekannten ausdrücken:

Wenn $(a, m) = d$ ist, so besitzt die Gleichung

$$a x - m y = b$$

für den Fall, daß b durch d teilbar ist, unendlich viele ganzzahlige Lösungen. Ist (x_0, y_0) eine beliebige Lösung dieser Gleichung, so erhält man alle Lösungen in der Form

$$x = x_0 + k\frac{m}{d}, \quad y = y_0 + k\frac{a}{d},$$

wobei k eine beliebige ganze Zahl ist. Wenn d jedoch kein Teiler von b ist, so besitzt die Gleichung keine ganzzahlige Lösung.

Wir wollen uns jetzt den algebraischen Kongruenzen höheren Grades zuwenden. Dabei beschränken wir uns auf den Fall eines Primzahlmoduls p, da nur hier eine entfernte Analogie zu den Verhältnissen der Gleichungstheorie vorhanden ist. Wir werden uns also mit Kongruenzen der Form

$$P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n \equiv 0 \pmod{p} \tag{11}$$

beschäftigen, wobei p eine Primzahl und a_0 nicht durch p teilbar ist.

Zunächst beweisen wir für derartige Kongruenzen einen Satz, der dem sogenannten „Satz von BÉZOUT" für algebraische Gleichungen analog ist:

Satz 6. *Wenn $x \equiv a \pmod{p}$ eine Lösung der Kongruenz (11) ist, so gibt es ein Polynom $Q(x)$ vom Grad $n-1$ mit ganzzahligen Koeffizienten derart, daß identisch*

$$P(x) \equiv (x - a)\, Q(x) \pmod{p} \tag{12}$$

ist.

Der Beweis dieses Satzes verläuft völlig analog dem Beweis des entsprechenden Satzes für algebraische Gleichungen. Wenn man nämlich das Polynom $P(x)$ im üblichen algebraischen Sinne durch das Binom $x - a$ dividiert, so erhält man als Quotienten ein Polynom $Q(x)$ vom Grad $n-1$ mit ganzzahligen Koeffizienten und als Rest eine ganze Zahl r, so daß also identisch

$$P(x) = (x - a)\, Q(x) + r$$

ist. Setzt man hier $x = a$, so erhält man

$$r = P(a) \equiv 0 \pmod{p}$$

und mithin

$$P(x) \equiv (x - a)\, Q(x) \pmod{p},$$

was zu beweisen war.

Wenn nun die Kongruenz (11) neben der Lösung $x \equiv a \pmod{p}$ noch eine weitere, von dieser verschiedene Lösung $x \equiv b \pmod{p}$ besitzt, so erhalten wir, wenn wir in der Kongruenz (12) $x = b$ setzen,

$$(b - a) Q(b) \equiv 0 \pmod{p}.$$

Da aber b eine von a verschiedene Lösung der Kongruenz (11) sein sollte, ist p kein Teiler von $b - a$ und mithin

$$Q(b) \equiv 0 \pmod{p},$$

d. h., $x \equiv b \pmod{p}$ ist eine Lösung der Kongruenz $Q(x) \equiv 0 \pmod{p}$. Folglich ist nach dem bereits bewiesenen Satz 6

$$Q(x) \equiv (x - b)\, R(x) \pmod{p}, \tag{13}$$

wobei $R(x)$ ein Polynom vom Grad $n - 2$ mit ganzen Koeffizienten ist. Aus (12) und (13) folgt dann aber, daß identisch

$$P(x) \equiv (x - a)\,(x - b)\, R(x) \pmod{p}$$

ist.

Wenn wir dieses Verfahren hinreichend oft fortsetzen, so erhalten wir folgenden allgemeinen Satz:

Wenn die Kongruenz (11) *die* $k \leq n$ *paarweise verschiedenen Lösungen* $x \equiv x_i \pmod{p}$ $(1 \leq i \leq k)$ *besitzt, so ist identisch*

$$P(x) \equiv (x - x_1)\,(x - x_2) \cdots (x - x_k)\, L(x) \pmod{p},$$

wobei $L(x)$ *ein Polynom vom Grad* $n - k$ *mit ganzen Koeffizienten ist.* Wir weisen bei dieser Gelegenheit darauf hin, daß die Polynome $Q(x)$, $R(x)$, $L(x)$ sämtlich den gleichen höchsten Koeffizienten a_0 wie das Polynom $P(x)$ besitzen, weil jedes dieser Polynome aus dem Polynom $P(x)$ durch sukzessive Division durch Binome der Form $x - a$ hervorgeht. Der obige Satz führt unverzüglich zu dem folgenden wichtigen Resultat:

Satz 7. *Eine Kongruenz vom Grad n nach einem Primzahlmodul kann höchstens n Lösungen besitzen.*

In der Tat: Würde die Kongruenz (11) die $n+1$ paarweise verschiedenen Lösungen $x \equiv x_i \pmod{p}$ $(1 \leq i \leq n+1)$ besitzen, so wäre zunächst gemäß der eben durchgeführten Überlegung für den Fall $k = n$ identisch

$$P(x) \equiv a_0\,(x - x_1)\,(x - x_2) \cdots (x - x_n) \pmod{p}.$$

Setzen wir hier $x = x_{n+1}$ und beachten wir, daß $P(x_{n+1}) \equiv 0 \pmod{p}$ sein sollte, so ergibt sich, daß

$$a_0\,(x_{n+1} - x_1)\,(x_{n+1} - x_2) \cdots (x_{n+1} - x_n) \equiv 0 \pmod{p}$$

sein müßte, was aber der Annahme, daß die x_i paarweise verschiedene Lösungen der Kongruenz (11) sind, widerspricht, da in diesem Fall kein Faktor der linken Seite, also auch nicht die linke Seite selbst, durch p teilbar sein kann.

Dieser Satz, der weitgehende Analogien zwischen den algebraischen Kongruenzen nach einem Primzahlmodul und den algebraischen Gleichungen zeigt, gilt nicht mehr für zusammengesetzte Moduln. Hier haben wir bereits gezeigt, daß eine Kongruenz ersten Grades nach einem zusammengesetzten Modul mehr als eine Lösung besitzen kann.

Ein interessantes Beispiel für den Fall, daß von einem Typus von algebraischen Kongruenzen tatsächlich die Höchstzahl der möglichen Lösungen erreicht wird, liefert uns der Satz von FERMAT. Nach ihm erfüllen nämlich für jede beliebige Primzahl p alle Zahlen, die nicht durch p teilbar sind, die Kongruenz

$$x^{p-1} - 1 \equiv 0 \pmod{p}. \tag{14}$$

Diese Zahlen bilden aber, modulo p gerechnet, gerade $p - 1$ Restklassen, so daß also die Kongruenz (14) für den Fall, daß p eine Primzahl ist, tatsächlich $p - 1$

Lösungen besitzt. Hieraus folgt — wie wir gleichfalls bereits wissen —, daß identisch

$$x^{p-1} - 1 \equiv (x-1)(x-2)\cdots(x-p+1) \pmod{p}$$

ist. Wenn wir hier speziell $x = 0$ setzen, so erhalten wir:

$$-1 \equiv (-1)^{p-1}(p-1)! \pmod{p}.$$

Falls nun $p > 2$ ist, so ist $(-1)^{p-1} = 1$ und mithin

$$(p-1)! + 1 \equiv 0 \pmod{p}, \tag{15}$$

während im Falle $p = 2$ die Kongruenz (15) auf

$$2 \equiv 0 \pmod{2}$$

führt, was gleichfalls gilt.. Wir sehen also, daß die Kongruenz (15) für jede Primzahl p erfüllt ist. Das ist die Aussage des sogenannten WILSONschen Satzes, der darüber hinaus sogar behauptet, daß die Kongruenz (15) charakteristisch für die Primzahlen ist. D s ist deshalb der Fall, weil die Kongruenz (15) für keine zusammengesetzte Zahl gilt; wenn nämlich p eine zusammengesetzte Zahl ist, so kann — wie man leicht verifiziert — p kein Teiler von $(p-1)! + 1$ sein.[1])

Es ist bis heute nicht gelungen, das WILSONsche Kriterium für irgendwelche theoretische Schlußfolgerungen auszunutzen, wenngleich es selbst natürlich von erheblichem Interesse ist.

Der Satz von FERMAT führt im betrachteten Problemkreis noch zu einer Reihe weiterer sehr allgemeiner und wichtiger Folgerungen. Wenn wir nämlich die Kongruenz

$$x^{p-1} \equiv 1 \pmod{p}$$

(die gemäß dem FERMATschen Satz für alle Zahlen x gilt, die nicht durch die Primzahl p teilbar sind) mit x multiplizieren, so erhalten wir die Kongruenz

$$x^p \equiv x \pmod{p},$$

die jetzt aber für alle ganzen Zahlen x gilt, da sie für die Zahlen, die durch p teilbar sind, trivialerweise erfüllt ist. Daher ist

$$x^p \equiv x \pmod{p}$$

eine Kongruenz, die für alle ganzen Zahlen gilt. Mit anderen Worten:
Für jede ganze Zahl x ist $x^p - x$ durch p teilbar.

Hieraus folgt, daß *man sich bei der Untersuchung von algebraischen Kongruenzen nach einem Primzahlmodul p auf Kongruenzen beschränken kann, deren Grad nicht größer als $p-1$ ist.* In der Tat: Wenn $P(x)$ ein Polynom eines Grades $\geq p$ mit ganzen Koeffizienten ist, so ergibt sich nach Division von $P(x)$ durch $x^p - x$ als Quotient ein Polynom $Q(x)$ und als Rest ein Polynom $R(x)$ mit gleichfalls ganzen Koeffizienten. Wir erhalten also

$$P(x) = (x^p - x)\,Q(x) + R(x),$$

wobei der Grad von $R(x)$ höchstens $p-1$ ist. Da nun für alle ganzen Zahlen

[1]) In der Tat: Wenn p einen Teiler d mit $1 < d < p$ besitzt, so ist offensichtlich $(p-1)!$ durch d teilbar. Dann kann aber $(p-1)! + 1$ nicht durch d und mithin auch nicht durch p teilbar sein.

$x^p - x \equiv 0 \pmod p$ gilt, ist jede Lösung der Kongruenz

$$P(x) \equiv 0 \pmod p$$

gleichzeitig Lösung von

$$R(x) \equiv 0 \pmod p$$

und umgekehrt. Auf diese Weise haben wir also tatsächlich die Frage nach den Lösungen einer algebraischen Kongruenz eines Grades $\geqq p$ gleichwertig auf die Frage nach den Lösungen einer Kongruenz eines Grades $< p$ zurückgeführt, wobei die neue Kongruenz zudem in einfacher Weise aus der gegebenen Kongruenz gewonnen werden kann.

Es sei also jetzt

$$P(x) \equiv 0 \pmod p \tag{16}$$

eine algebraische Kongruenz eines Grades $n < p$. Als letzte Frage wollen wir hier untersuchen, wie man feststellen kann, ob die Anzahl der Lösungen dieser Kongruenz gleich n oder kleiner als n ist (größer als n kann sie — wie wir bereits wissen — nicht sein). Dabei können wir ohne Beschränkung der Allgemeinheit annehmen, daß der höchste Koeffizient des Polynoms $P(x)$ gleich 1 ist. In der Tat: Dieser Koeffizient a_0 ist nach Voraussetzung nicht durch p teilbar. Daher garantieren uns unsere Überlegungen über lineare Kongruenzen die Existenz einer Zahl a, für die $a_0 a \equiv 1 \pmod p$ ist. Wenn wir dann die Kongruenz (16) mit a multiplizieren, so erhalten wir die zu ihr bezüglich der Lösungsmannigfaltigkeit gleichwertige Kongruenz

$$a P(x) \equiv 0 \pmod p,$$

deren höchster Koeffizient $a_0 a$ ist, der kongruent 1 modulo p ist und daher in der Kongruenz durch die Zahl 1 ersetzt werden kann.

Wir setzen also im Folgenden voraus, daß in der Kongruenz (16) der höchste Koeffizient von $P(x)$ gleich 1 ist. Dann erhalten wir bei der Division von $x^p - x$ durch $P(x)$ als Quotienten und als Rest Polynome mit ganzzahligen Koeffizienten, die w.r mit $M(x)$ und $N(x)$ bezeichnen wollen, so daß also

$$x^p - x = P(x) M(x) + N(x)$$

ist, wobei offensichtlich $M(x)$ ein Polynom höchstens vom Grade $p - n$ und $N(x)$ ein Polynom höchstens vom Grade $n - 1$ ist.

Wir nehmen nun zunächst an, daß die Kongruenz (16) genau n Lösungen besitzt. Da aber $x^p - x \equiv 0 \pmod p$ ist, erfüllt jede dieser n Lösungen auch die Kongruenz

$$N(x) = x^p - x - P(x) M(x) \equiv 0 \pmod p.$$

Dann muß aber jeder Koeffizient des Polynoms $N(x)$ durch p teilbar sein, da anderenfalls $N(x) \equiv 0 \pmod p$ eine algebraische Kongruenz eines Grades $< n$ wäre, die genau n Lösungen besitzt — was nicht möglich ist. Wir haben also festgestellt, daß alle Koeffizienten des Polynoms $N(x)$ durch p teilbar sein müssen.

Es sei nun umgekehrt bekannt, daß alle Koeffizienten des Polynoms $N(x)$ durch p teilbar sind, daß also identisch $N(x) \equiv 0 \pmod p$ ist. Dann ist auch identisch

$$P(x) M(x) \equiv 0 \pmod p, \tag{17}$$

d. h., alle p Restklassen modulo p erfüllen die Kongruenz. Nun ist aber offen-

sichtlich jede Lösung der Kongruenz (17) Lösung wenigstens einer der Kongruenzen

$$P(x) \equiv 0 \pmod{p}, \quad M(x)^t \equiv 0 \pmod{p}$$

und mithin die Summe der Anzahlen der Lösungen dieser beiden Kongruenzen mindestens gleich p. Da aber die zweite höchstens $p - n$ Lösungen besitzt, kann die Anzahl der Lösungen der Kongruenz (16) nicht kleiner als n sein und muß mithin genau n sein.

Wir kommen also auf diese Weise zu folgendem Kriterium:

Satz 8. *Damit die Kongruenz* (16) *vom Grad* $n < p$ *mit höchstem Koeffizienten* 1 *genau* n *Lösungen besitzt, ist notwendig und hinreichend, daß alle Koeffizienten des Polynoms, das sich als Rest bei der Division von* $x^p - x$ *durch* $P(x)$ *ergibt, durch* p *teilbar sind.*

Kapitel III

DER EUKLIDISCHE ALGORITHMUS UND DIE KETTENBRÜCHE

§ 8. Der Euklidische Algorithmus

Die elementare Arithmetik lehrt im wesentlichen zwei Methoden zur Bestimmung des größten gemeinsamen Teilers zweier gegebener Zahlen. Die eine Methode geht über die Primzahlzerlegung der gegebenen Zahlen, wobei der größte gemeinsame Teiler in bekannter Weise aus den Primfaktoren der gegebenen Zahlen bzw. eines Teils derselben kombiniert wird. Die andere Methode besteht in einer Reihe sukzessiver Divisionen der gegebenen Zahlen: Zunächst wird die eine der gegebenen Zahlen durch die andere mit Rest dividiert; sodann dividiert man die zweite der gegebenen Zahlen durch den Rest der ersten Division; anschließend dividiert man den Rest der ersten Division durch den Rest der zweiten Division usw. Da bei diesem Verfahren jeder Rest kleiner ist als der vorhergehende und da alle Reste nichtnegativ sind, muß man nach endlich vielen Schritten zu einem Rest kommen, der gleich Null ist. Wir werden zeigen, daß der letzte positive Rest in diesem Prozeß der größte gemeinsame Teiler der gegebenen Zahlen ist. Diese Methode zur Bestimmung des größten gemeinsamen Teilers, die man gewöhnlich als „Euklidischen Algorithmus" bezeichnet, ist offensichtlich außerordentlich einfach. Für ihre Anwendung ist es nicht notwendig, die Primzahlzerlegung der gegebenen Zahlen zu kennen. Darin besteht ihr wesentlicher Vorteil gegenüber der zuerst genannten Methode. Ungeachtet seiner Einfachheit und des ihm anhaftenden elementaren Zuges, besitzt der Euklidische Algorithmus schon auf Grund seines Wesens und seiner Geschichte einen erheblichen methodologischen Wert. Er kann als Grundlage für die gesamte Teilbarkeitstheorie, einschließlich der auf ihr begründeten Theorie der diophantischen Gleichungen ersten Grades in zwei Unbekannten (vgl. S. 253) dienen. Ferner ist er — wie bekannt — ein unentbehrliches Hilfsmittel zur Bestimmung des gemeinsamen Maßes zweier Größen und damit ein Grundpfeiler jeglicher Theorie des

Messens. Schließlich dient er als Ausgangspunkt für die mächtigste aller Methoden der Arithmetik der irrationalen Zahlen, der Theorie der Kettenbrüche, und besitzt damit einen weiteren unmittelbar praktischen Wert.

Wir wollen deshalb hier den Euklidischen Algorithmus und seine wichtigsten arithmetischen Anwendungen eingehend und mit aller Sorgfalt betrachten.

Vorgegeben seien zwei Zahlen a und b, von denen die zweite positiv sei. Wir wollen dann a durch b mit Rest dividieren, wobei der Quotient und der Rest der Division mit a_1 bzw. r_1 bezeichnet werde, so daß also

$$a = ba_1 + r_1 \quad (0 \leq r_1 < b) \tag{1}$$

ist. Wenn nun $r_1 > 0$ ist, so dividieren wir b durch r_1, wobei wir mit a_2 den Quotienten und mit r_2 den Rest dieser Division bezeichnen, so daß also

$$b = r_1 a_2 + r_2 \quad (0 \leq r_2 < r_1) \tag{2}$$

ist. Wenn nun auch r_2 noch nicht Null ist, so dividieren wir r_1 durch r_2 und erhalten entsprechend

$$r_1 = r_2 a_3 + r_3 \quad (0 \leq r_3 < r_2)$$

usw. Da nun $b > r_1 > r_2 > r_3 \cdots \geq 0$ ist, muß der so begonnene Prozeß nach endlich vielen Schritten abbrechen, d. h., wir müssen früher oder später zu einem Rest gelangen, der gleich Null ist. Es möge etwa erstmalig $r_{n+1} = 0$ sein, so daß also

$$r_{n-1} = r_n a_{n+1}$$

ist. Dann ist r_n der größte gemeinsame Teiler der Zahlen a und b. Zum Beweis dieser Tatsache zeigen wir zunächst, daß a und b durch r_n teilbar sind: Gemäß der zuletzt erhaltenen Gleichung ist sicher r_{n-1} durch r_n teilbar. Dann zeigt die vorangehende (bisher nicht hingeschriebene) Gleichung

$$r_{n-2} = r_{n-1} a_n + r_n, \tag{3}$$

daß auch r_{n-2} durch r_n teilbar ist. Indem wir die von uns konstruierte Gleichungskette (in umgekehrter Reihenfolge) weiter verfolgen, finden wir schrittweise, daß auch $r_{n-3}, r_{n-4}, \ldots, r_1$ und schließlich auch a und b durch r_n teilbar sind. Es bleibt uns noch zu zeigen, daß jeder Teiler d der Zahlen a und b auch Teiler von r_n ist (dann ist offensichtlich gezeigt, daß r_n der größte gemeinsame Teiler von a und b ist). Dazu verfolgen wir wiederum die von uns konstruierte Gleichungskette, dieses Mal aber von oben nach unten. Da die Gleichung (1) auch in der Form

$$r_1 = a - ba_1$$

geschrieben werden kann, erweist sich jeder gemeinsame Teiler d von a und b auch als Teiler von r_1. Dann zeigt aber die Gleichung (2) entsprechend, daß d ein Teiler von r_2 ist usw. Schließlich finden wir am Ende der Rechnung auf Grund der Gleichung (3), daß die Teilbarkeit von r_{n-2} und r_{n-1} durch d die Teilbarkeit von r_n durch d zur Folge hat.

Das hier geschilderte Verfahren zur Bestimmung des größten gemeinsamen Teilers zweier Zahlen mittels des Euklidischen Algorithmus erweist sich in den Anwendungen als kürzer und damit vorteilhafter als die andere Methode. Interessant ist es noch, zu bemerken, daß die eben durchgeführten Überlegungen zeigen, daß jeder gemeinsame Teiler zweier Zahlen auch Teiler des größten gemeinsamen Teilers dieser Zahlen ist.

Wir wollen uns nun davon überzeugen, daß der Euklidische Algorithmus als Grundlage der gesamten Teilbarkeitstheorie dienen kann. Die Gleichung (1) zeigt uns, daß die Zahl r_1 als „Linearkombination" der Zahlen a und b dargestellt werden kann, daß also die Zahl r_1 in der Form $ax + by$ mit ganzem x und y (nämlich $x = 1$ und $y = -a_1$) geschrieben werden kann. Dann folgt aber aus Gleichung (2), daß auch $r_2 = b - r_1 a_2$ als Linearkombination der Zahlen a und b erhalten werden kann, nämlich

$$r_2 = b - (ax + by)\, a_2 = a\,(-a_2 x) + b\,(1 - a_2 y).$$

Indem wir die Gleichungskette des Euklidischen Algorithmus weiter verfolgen, finden wir, daß auch die Zahlen $r_3, r_4, \ldots$ sich in der Form $ax + by$ ausdrücken lassen. Schließlich erlaubt es uns Gleichung (3), aus der Darstellbarkeit von r_{n-2} und r_{n-1} als Linearkombination von a und b auf die Darstellbarkeit von r_n in dieser Form zu schließen. Auf diese Weise haben wir einen neuen Beweis für die bereits aus Kapitel I bekannte Tatsache, daß der größte gemeinsame Teiler zweier Zahlen stets als Linearkombination dieser Zahlen darstellbar ist, erhalten. Wenn speziell die Zahlen a und b teilerfremd sind, so kommen wir erneut auf Satz 1 aus Kapitel I (S. 229).

Wir haben damit einen dritten Beweis dieses Satzes gegeben, der zudem ein Verfahren zur Bestimmung der in ihm auftretenden Zahlen x und y enthält. Im Rahmen der Theorie der Kettenbrüche werden wir noch eine bequemere Anordnung der Operationen zum Auffinden dieser Zahlen kennenlernen.

Aus Satz 1 folgt leicht, wie wir in Kapitel I gezeigt haben, der dortige Satz 2, der als Grundlage für den Fundamentalsatz über die eindeutige Zerlegbarkeit der Zahlen in Primfaktoren und damit für die Theorie der Teilbarkeit diente. Damit ist Satz 1 auch das Fundament für die Theorie der diophantischen Gleichungen ersten Grades in zwei Unbekannten.

Die Bedeutung des Euklidischen Algorithmus reicht jedoch weit über die Grenzen der Arithmetik der natürlichen Zahlen hinaus. Ganz abgesehen davon, daß der Euklidische Algorithmus als Grundlage für die Teilbarkeitstheorie in den verschiedensten algebraischen Bereichen dient, bildet er in der Algebra das Fundament für die Teilbarkeitstheorie der Polynome in einer Veränderlichen. Da diese Frage sich unmittelbar an die Schulalgebra anschließt, wollen wir sie hier ausführlich behandeln. Dabei können wir uns in den meisten Punkten unserer Darstellung recht kurz fassen, indem wir auf die fast vollständige Analogie zu den vorangehenden Überlegungen hinweisen.

Die Objekte unserer Rechenoperationen sind jetzt nicht Zahlen, sondern Polynome der Gestalt

$$P(x) = a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n,$$

wobei die Koeffizienten $a_0, a_1, \ldots, a_n$ rationale Zahlen sind. Wir nennen $P(x)$ ein Polynom vom Grad n, wenn $a_0 \neq 0$ ist.

Es seien nun $A(x)$ und $B(x)$ beliebige derartige Polynome, wobei das Polynom $B(x)$ keine Konstante (also kein Polynom, dessen sämtliche Koeffizienten außer dem absoluten Glied gleich Null sind) sei. Dividiert man nun nach dem bekannten elementaren Verfahren $A(x)$ durch $B(x)$, so erhält man zwei Polynome $Q(x)$ und $R(x)$ derart, daß identisch in x

$$A(x) = Q(x)\, B(x) + R(x) \tag{4}$$

ist, wobei der Grad des Polynoms $R(x)$ (des Restes) kleiner als der Grad des Polynoms $B(x)$ (des Teilers) ist. Die Polynome $Q(x)$ (der Quotient) und $R(x)$ (der Rest) sind dabei eindeutig durch die Gleichung (4) festgelegt. In der Tat: Wenn neben der Relation (4) noch eine weitere derartige Gleichung, etwa

$$A(x) = Q'(x)\,B(x) + R'(x)$$

gilt, wobei der Grad von $R'(x)$ gleichfalls kleiner als der Grad von $B(x)$ ist, so ist identisch

$$R(x) + Q(x)\,B(x) = R'(x) + Q'(x)\,B(x)$$

und mithin

$$[Q(x) - Q'(x)]\,B(x) = R'(x) - R(x).$$

Da nun der Grad des Polynoms auf der rechten Seite dieser letzten Gleichung kleiner als der Grad von $B(x)$ ist, ist das Bestehen dieser Gleichung nur möglich, wenn $Q(x)$ und $Q'(x)$ und mithin auch $R(x)$ und $R'(x)$ übereinstimmen.

Da die Koeffizienten der Polynome $Q(x)$ und $R(x)$ aus den Koeffizienten von $A(x)$ und $B(x)$ mittels der rationalen Operationen gewonnen werden, sind sie gleichfalls rationale Zahlen.

Im folgenden wollen wir unter der Gleichheit zweier Polynome die Gleichheit der jeweiligen Koeffizienten der gleichen Potenzen von x verstehen. Insbesondere bedeutet also $P(x) = 0$, daß sämtliche Koeffizienten von $P(x)$ gleich Null sind.

Wenn in der Gleichung (4) das Polynom $R(x) = 0$ ist, wenn also $A(x)$ in der Form $B(x)\,Q(x)$ darstellbar ist, so sagt man, das Polynom $A(x)$ sei durch das Polynom $B(x)$ (ohne Rest) *teilbar*. Mithin ist also die Teilbarkeit im Bereich der Polynome analog wie im Bereich der ganzen Zahlen definiert.

Wenn r eine von Null verschiedene rationale Zahl und $P(x)$ ein Polynom mit rationalen Koeffizienten ist, so ist auch das Polynom $\frac{P(x)}{r}$ ein Polynom mit rationalen Koeffizienten. Dann zeigt aber die Beziehung

$$A(x) = r \cdot \frac{A(x)}{r},$$

daß jedes Polynom $A(x)$ durch jede (von Null verschiedene) rationale Zahl und durch jedes Polynom der Form $\frac{A(x)}{r}$ teilbar ist. Die rationalen Zahlen spielen also in der Teilbarkeitstheorie der Polynome dieselbe Rolle wie die Zahlen ± 1 in der Teilbarkeitstheorie der ganzen Zahlen. Es sind dementsprechend zwei Polynome *teilerfremd* zu nennen, wenn sie außer rationalen Zahlen keine gemeinsamen Teiler besitzen. Ferner nennen wir ein Polynom $P(x)$ *absolut prim* (oder *irreduzibel*), wenn es keine Teiler außer den rationalen Zahlen und den Polynomen der Form $\frac{P(x)}{r}$ besitzt. Dabei sollen jedoch die rationalen Zahlen nicht zu den irreduziblen Polynomen gerechnet werden, da es — wie auch in der Teilbarkeitstheorie der ganzen Zahlen — bequemer ist, die „Eins" nicht als Primelement anzusehen.

Die Relation (4), die in ihrer äußeren Struktur der Beziehung zwischen dem Dividenden, dem Divisor, dem Quotienten und dem Rest in der Teilbarkeitstheorie der ganzen Zahlen völlig analog ist, kann auch hier — zunächst rein formal — als

Ausgangspunkt für einen Euklidischen Algorithmus dienen, der gleichfalls die Teilbarkeitstheorie der Polynome bereits im Keime enthält. Im Fall der ganzen Zahlen war hierbei das entscheidende Moment, daß bei jeder Division der Rest stets kleiner als der Divisor war. Hierauf stützte sich die Endlichkeit des Algorithmus. Im Falle der Polynome leistet der Grad des Restes die entsprechenden Dienste, da der Grad des Restes stets kleiner als der Grad des Divisors ist. Da nun der Grad bei der fortschreitenden Division immer kleiner wird, so muß er nach endlich vielen Schritten zu Null werden, womit auch hier die Endlichkeit des Prozesses gesichert ist.

Die formale Seite des Algorithmus ist dabei die gleiche wie im Fall der ganzen Zahlen, so daß wir sie nur in großen Zügen ins Gedächtnis zurückzurufen brauchen: Es wird zunächst $A(x)$ durch $B(x)$, dann $B(x)$ durch den ersten Rest, dann der erste Rest durch den zweiten Rest usw. dividiert. Solange der Rest keine rationale Zahl ist (also einen positiven Grad besitzt), ist der Grad des nachfolgenden Restes kleiner als der Grad des betrachteten Restes. Wenn wir jedoch zu einem Rest des Grades Null (d. h. zu einer rationalen Zahl) gekommen sind, so ist der nachfolgende Rest die Zahl Null selbst. Wir können also feststellen, daß ein Rest früher oder später die Zahl Null sein wird. Wenn wir unter $R_n(x)$ den letzten von Null verschiedenen Rest (der insbesondere eine rationale Zahl sein kann) verstehen, so erhalten wir — wie im Falle der ganzen Zahlen —

$$R_{n-2}(x) = Q_n(x)\, R_{n-1}(x) + R_n(x),$$

$$R_{n-1}(x) = Q_{n+1}(x)\, R_n(x).$$

Ausgehend von der letzten Gleichung, überzeugt man sich davon, daß gemäß der vorangehenden Gleichung $R_{n-2}(x)$ und schließlich, nach endlich vielen Schritten, daß auch die Ausgangspolynome $A(x)$ und $B(x)$ durch $R_n(x)$ teilbar sind. Es ist also $R_n(x)$ ein gemeinsamer Teiler der gegebenen Polynome. Wenn man entsprechend die Gleichungskette des Euklidischen Algorithmus in ihrer ursprünglichen Reihenfolge verfolgt, so findet man wie im Fall der ganzen Zahlen, daß jeder gemeinsame Teiler $D(x)$ der Polynome $A(x)$ und $B(x)$ auch Teiler des Polynoms $R_n(x)$ ist. Mithin ist $R_n(x)$ ein gemeinsamer Teiler der Polynome $A(x)$ und $B(x)$, der von jedem anderen gemeinsamen Teiler dieser Polynome geteilt wird. In diesem Sinne wollen wir $R_n(x)$ als *größten* gemeinsamen Teiler von $A(x)$ und $B(x)$ bezeichnen; diese Definition ist insofern ganz natürlich, als wir Polynome nicht ihrer Größe nach vergleichen können und somit die gewöhnliche Definition des größten gemeinsamen Teilers nicht aus der Theorie der ganzen Zahlen in das Gebiet der Polynome übernommen werden kann (während andererseits die hier gegebene Festsetzung auch zur Charakterisierung des größten gemeinsamen Teilers zweier ganzer Zahlen benutzt werden kann — *Anm. d. wissenschaftl. Red.*). Übrigens besitzt auch hier der größte gemeinsame Teiler eine Reihe von Maximaleigenschaften, auf die wir jetzt zu sprechen kommen wollen. Interessant ist dabei, daß wir auf diesem Wege die Frage nach der Eindeutigkeit des größten gemeinsamen Teilers zweier Polynome sehr einfach werden lösen können. Während diese Frage im Bereich der ganzen Zahlen trivial entschieden werden kann, erfordert sie jetzt natürlich eine gesonderte Betrachtung.

Zunächst ist unmittelbar klar, daß mit einem größten gemeinsamen Teiler $D(x)$ der Polynome $A(x)$ und $B(x)$ auch jedes Polynom der Form $rD(x)$, wobei r eine beliebige von Null verschiedene rationale Zahl ist, ein größter gemeinsamer Teiler

dieser Polynome ist. Wir sehen also, daß zwei Polynome stets unendlich viele größte gemeinsame Teiler besitzen, nämlich sicher die rationalen Vielfachen des gemäß dem Euklidischen Algorithmus gefundenen größten gemeinsamen Teilers (wir werden weiter unten sehen, daß dies auch alle größten gemeinsamen Teiler der gegebenen Polynome sind — *Anm. d. wissenschaftl. Red.*). Das ist aber keineswegs verwunderlich, da wir schon erkannt haben, daß in der Teilbarkeitstheorie der Polynome sämtliche rationalen Zahlen die Rolle der Zahlen ± 1 spielen. Der Unterschied zwischen den bisher gefundenen verschiedenen größten gemeinsamen Teilern zweier Polynome ist insofern trivial, als sie bezüglich ihrer Teilbarkeitseigenschaften nur unwesentlich voneinander verschieden sind. Man sieht nun leicht ein, daß *jeder gemeinsame Teiler $D'(x)$ der Polynome $A(x)$ und $B(x)$, der sich von ihrem größten gemeinsamen Teiler $D(x)$ wesentlich* (d. h. nicht nur bis auf eine multiplikative Konstante) *unterscheidet, einen Grad besitzt, der kleiner als der Grad von $D(x)$ ist.* In der Tat: Wenn $D(x)$ eine rationale Zahl ist, so ist auch $D'(x)$ als Teiler von $D(x)$ eine rationale Zahl und mithin nur unwesentlich von $D(x)$ verschieden. Wir können also annehmen, daß der Grad von $D(x)$ positiv ist. Dann besitzt, da die Polynome $D(x)$ und $D'(x)$ wesentlich verschieden sein sollen, der Quotient der Division von $D(x)$ durch $D'(x)$ einen positiven Grad und mithin $D'(x)$ einen kleineren Grad als $D(x)$, wie behauptet. Daraus folgt nun aber unmittelbar, daß *je zwei größte gemeinsame Teiler zweier Polynome sich nur durch einen konstanten Faktor unterscheiden*, der größte gemeinsame Teiler zweier Polynome also bis auf einen konstanten Faktor eindeutig bestimmt ist.

Damit kann auch hier — genau wie in der Theorie der ganzen Zahlen — der Euklidische Algorithmus, mit dessen Hilfe man den (im wesentlichen eindeutigen) größten gemeinsamen Teiler zweier Polynome finden kann, als Grundlage für die gesamte Teilbarkeitstheorie dienen. Zunächst folgt nämlich aus dem Euklidischen Algorithmus in völliger Analogie zu dem obigen Beweis im Bereich der ganzen Zahlen, daß der größte gemeinsame Teiler zweier Polynome stets als Linearkombination dieser Polynome dargestellt werden kann. Die Koeffizienten dieser Kombination sind jetzt natürlich gewisse Polynome mit rationalen Koeffizienten. Daraus folgt insbesondere als Analogon zu Satz 1 aus Kapitel I, daß *es zu teilerfremden Polynomen $A(x)$ und $B(x)$ stets Polynome $X(x)$ und $Y(x)$ gibt, für die*

$$A(x)X(x) + B(x)Y(x) = 1$$

ist.

Hieraus ergibt sich nun vor allem, daß sich auch die Frage nach den Lösungen einer diophantischen Gleichung ersten Grades in zwei Unbekannten (die Bezeichnung *diophantisch* soll hier — wie üblich — besagen, daß es sich um die ganzen, also in unserem Falle um die Polynomlösungen der Gleichung handelt und nicht etwa um die rationalen Funktionen, die die Gleichung erfüllen — *Anm. d. wissenschaftl. Red.*) im Bereich der Polynome entsprechend wie im Falle der ganzen Zahlen beantwortet. Wenn wir im Bereich der Polynome den Weg verfolgen, den wir in Kapitel I für die ganzen Zahlen vorgezeichnet haben, so erhalten wir leicht einen Satz über Polynome, der dem Satz 2 aus Kapitel I entspricht. Daraus ergibt sich dann auch hier — genau wie dort — die (bis auf die Reihenfolge der Faktoren und jetzt einen konstanten Faktor) eindeutige Zerlegbarkeit eines Polynoms in Primfaktoren (irreduzible Polynome), die als Fundament für die weitere Teilbarkeitstheorie dient.

Wir erwähnen noch, daß sich auch der Satz EUKLIDS über die Unendlichkeit der Menge der Primzahlen mitsamt seines Beweises wörtlich auf den Bereich der Polynome übertragen läßt. Jedoch kann man hier die Existenz unendlich vieler irreduzibler Polynome auch einfacher erschließen, indem man feststellt, daß alle Polynome ersten Grades irreduzibel sind. Der Beweis hierfür sei dem Leser überlassen.

§ 9. Die elementare Theorie der Kettenbrüche

Wir kehren nun zu dem Bereich der ganzen Zahlen zurück und schreiben uns noch einmal sorgfältig die Gleichungskette auf, die zur Bestimmung des größten gemeinsamen Teilers führte:

$$\begin{aligned} a &= ba_1 + r_1, \\ b &= r_1a_2 + r_2, \\ r_1 &= r_2a_3 + r_3, \\ &\cdots\cdots\cdots \\ r_{n-2} &= r_{n-1}a_n + r_n, \\ r_{n-1} &= r_na_{n+1}, \end{aligned}$$

wobei

$$b > r_1 > r_2 > \cdots > r_n > 0 \tag{5}$$

ist.

Diese Gleichungskette können wir offensichtlich gleichwertig in das folgende Gleichungssystem umschreiben:

$$\begin{aligned} \frac{a}{b} &= a_1 + \frac{r_1}{b}, \\ \frac{b}{r_1} &= a_2 + \frac{r_2}{r_1}, \\ \frac{r_1}{r_2} &= a_3 + \frac{r_3}{r_2}, \\ &\cdots\cdots\cdots \\ \frac{r_{n-2}}{r_{n-1}} &= a_n + \frac{r_n}{r_{n-1}}, \\ \frac{r_{n-1}}{r_n} &= a_{n+1}. \end{aligned}$$

In dieser Form läßt sich aber jede unserer Gleichungen als eine einfache arithmetische Operation deuten. Es wird nämlich einfach aus den unechten Brüchen auf der linken Seite jeweils der ganze Teil ausgesondert, d. h., die unechten Brüche auf der linken Seite werden als Summen aus einer ganzen Zahl und einem echten Bruch dargestellt. In der Tat: Gemäß (5) stellen die Brüche auf der linken Seite unechte Brüche dar, während gleichfalls wegen (5) die zweiten Summanden auf der rechten Seite echte Brüche sind. Wir weisen noch darauf hin, daß (angefangen von der

zweiten Gleichung) die linke Seite jeder Gleichung gleich dem reziproken zweiten Summanden der vorangehenden Gleichung ist.

Wir stellen uns jetzt die Aufgabe, das Verhältnis $\frac{a}{b}$ der Zahlen a und b allein mit Hilfe der Zahlen $a_1, a_2, \ldots, a_n$ auszudrücken. Das läßt sich nun sehr leicht erreichen. Auf Grund der ersten Gleichung ist nämlich zunächst

$$\frac{a}{b} = a_1 + \frac{1}{\frac{b}{r_1}}.$$

Setzen wir hier für $\frac{b}{r_1}$ den entsprechenden Ausdruck aus der zweiten Gleichung ein, so finden wir, daß

$$\frac{a}{b} = a_1 + \frac{1}{a_2 + \frac{r_2}{r_1}} = a_1 + \frac{1}{a_2 + \frac{1}{\frac{r_1}{r_2}}}$$

ist. Jetzt können wir entsprechend für $\frac{r_1}{r_2}$ den Ausdruck aus der dritten Gleichung einsetzen und erhalten

$$\frac{a}{b} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{r_3}{r_2}}} = a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{\frac{r_2}{r_3}}}}$$

usw. So finden wir schließlich:

$$\frac{a}{b} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots + \cfrac{1}{\frac{r_{n-1}}{r_n}}}} = a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots + \cfrac{1}{a_{n+1}}}} \tag{6}$$

Ein Ausdruck, wie er auf der rechten Seite der Gleichung (6) steht, heißt ein *Kettenbruch*. Hierbei ist a_1 eine ganze Zahl (die auch negativ oder Null sein kann), während $a_2, a_3, \ldots, a_{n+1}$ natürliche Zahlen sind.

Damit erhält der Euklidische Algorithmus für uns einen ganz neuen und sehr bedeutenden Wert, gibt er uns doch die Möglichkeit, jeden beliebigen einfachen Bruch in Form eines Kettenbruches darzustellen, wobei er gleichzeitig eine Anweisung zur faktischen Berechnung der Glieder liefert.

In der elementaren Arithmetik werden eine ganze Reihe von Möglichkeiten für die Darstellung der rationalen Zahlen gelehrt. Da ist zunächst die einfachste Darstellungsart in Form der gewöhnlichen Brüche, ferner die Darstellung in Form der periodischen unendlichen und der endlichen Dezimalbrüche (oder allgemeiner in Form beliebiger systematischer Brüche, d. h. als Brüche bezüglich einer beliebigen Grundzahl) und schließlich in der Form von Prozentzahlen. Neben diesen verschiedenen Darstellungsformen, von denen jede ihre Vorzüge hat, haben wir jetzt als neue Darstellungsform die Kettenbrüche kennengelernt, die sowohl für theo-

retische Überlegungen als auch in den praktischen Anwendungen eine wichtige Rolle spielen. Daraus erklärt es sich, daß die Lehre von den Kettenbrüchen einen erheblichen Umfang angenommen hat, der sich auch heute noch ständig vergrößert.

Als Abkürzung für den auf der rechten Seite der Gleichung (6) stehenden Kettenbruch schreibt man gewöhnlich

$$[a_1;\ a_2,\ a_3,\ \ldots,\ a_{n+1}],$$

wobei das Semikolon hinter a_1 die Zahl a_1 als „ganzen Teil" der durch den Kettenbruch dargestellten Zahl hervorheben soll. Die Zahlen $a_1, a_2, \ldots, a_{n+1}$ heißen die *Glieder* oder auch die *Teilnenner* oder schließlich auch die *unvollständigen Quotienten* des gegebenen Kettenbruches. Der letzte Name erinnert noch an ihre Entstehung aus dem Euklidischen Algorithmus.

Zunächst erhebt sich naturgemäß die Frage nach der Eindeutigkeit der Darstellung einer gegebenen rationalen Zahl durch einen Kettenbruch, also die Frage, ob zwei verschiedene Kettenbrüche ein und dieselbe Zahl darstellen können, d. h., ob zwei verschiedene Kettenbrüche untereinander gleich sein können. Sehr einfache Beispiele wie

$$\frac{1}{2} = \frac{1}{1+\frac{1}{1}}$$

oder kurz

$$[0;\ 2] = [0;\ 1,\ 1]$$

zeigen bereits, daß das durchaus der Fall sein kann. Allgemein: Wenn $a_{n+1} > 1$ ist, so ist

$$a_{n+1} = (a_{n+1} - 1) + \frac{1}{1}$$

und mithin

$$[a_1; a_2, \ldots, a_{n+1}] = [a_1; a_2, \ldots, a_{n+1} - 1,\ 1].$$

Diese zweifache Möglichkeit der Darstellung einer rationalen Zahl als Kettenbruch ist indes von ebenso trivialem Charakter wie etwa die Möglichkeit der Darstellung

$$1 = 0{,}999\ldots$$

in der Theorie der Dezimalbrüche. Entsprechend wie wir dort die Doppeldeutigkeit aufheben können durch ein Verbot von unendlichen Dezimalbrüchen, die mit einer unendlichen Folge von Neunen abschließen, wollen wir hier verabreden, daß Kettenbrüche, in denen das letzte Glied a_{n+1} gleich 1 ist, außer Betracht bleiben sollen. Das ist sinnvoll, da der Bruch

$$[a_1; a_2, \ldots, a_{n-1}, a_n,\ 1]$$

gleich dem Bruch

$$[a_1; a_2, \ldots, a_{n-1}, a_n + 1]$$

ist, dessen letztes Glied größer als 1 ist.

Wir zeigen nun, daß nach dieser Festsetzung die Eindeutigkeit der Darstellung einer gegebenen rationalen Zahl als Kettenbruch gewährleistet ist. Dazu sei

$$[a_0; a_1, a_2, \ldots, a_r] = [b_0; b_1, b_2, \ldots, b_s],$$

wobei $a_r > 1$ und $b_s > 1$ ist. Zu zeigen ist, daß dann $r = s$ und $a_0 = b_0, a_1 = b_1, \ldots, a_r = b_r$ ist.

Aus

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_r}}} = b_0 + \cfrac{1}{b_1 + \cfrac{1}{\ddots + \cfrac{1}{b_s}}}$$

folgt zunächst, daß die ganzen Teile in beiden Brüchen gleich sein müssen. Denn der zweite Summand des linker Hand stehenden Kettenbruches ist ein echter Bruch, da er dann und nur dann gleich 1 wäre, wenn $r = 1$ und $a_1 = 1$ ist, was wir indes durch unsere Forderung $a_r > 1$ ausgeschlossen haben. Mithin ist der ganze Teil des linker Hand stehenden Bruches die Zahl a_0. Entsprechend schließt man, daß der ganze Teil des rechter Hand stehenden Kettenbruches die Zahl b_0 ist und daß mithin $a_0 = b_0$ ist. Wenn das aber der Fall ist, so ist

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_r}}} = b_1 + \cfrac{1}{b_2 + \cfrac{1}{\ddots + \cfrac{1}{b_s}}},$$

woraus wir auf dieselbe Weise, auf die wir eben auf $a_0 = b_0$ geschlossen haben, auf $a_1 = b_1$ schließen können. Wenn wir diesen Prozeß hinreichend oft fortsetzen, so erhalten wir offensichtlich nach und nach alles, was zu beweisen ist. Damit können wir die Eindeutigkeit der Darstellung einer rationalen Zahl in Form eines Kettenbruches als bewiesen ansehen.

Der wichtigste Grundbegriff der gesamten Theorie der Kettenbrüche ist der Begriff des *Näherungsbruches*. Dazu sei uns ein beliebiger Kettenbruch

$$\frac{a}{b} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}} = [a_0; a_1, a_2, \ldots a_n]$$

vorgegeben. Wir betrachten dann die Folge der Kettenbrüche $[a_0]$, $[a_0; a_1]$, $[a_0; a_1, a_2]$ $\ldots$, $[a_0; a_1, a_2, \ldots, a_{n-1}]$, $[a_0; a_1, a_2, \ldots, a_n]$. Der letzte Bruch in dieser Folge ist offensichtlich der gegebene Kettenbruch, während man die vorhergehenden durch „Abbrechen" nach einem gewissen Teilnenner a_k erhält. Man kann nun jeden dieser Kettenbrüche ohne weiteres berechnen, d. h. in der Form eines gewöhnlichen Bruches angeben. Dabei findet man

$$[a_0] = a_0 = \frac{a_0}{1},$$

$$[a_0; a_1] = a_0 + \frac{1}{a_1} = \frac{a_0 a_1 + 1}{a_1},$$

$$[a_0; a_1, a_2] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}} = \frac{(a_0 a_1 + 1) a_2 + a_0}{a_1 a_2 + 1}$$

usw. Die auf diese Weise erhaltene Folge von gewöhnlichen Brüchen nennt man die Folge der *Näherungsbrüche* des gegebenen Kettenbruches (oder auch der durch

ihn dargestellten Zahl $\frac{a}{b}$). Wir wollen diese Brüche in der angegebenen Reihenfolge durch $\frac{p_0}{q_0}, \frac{p_1}{q_1}, \frac{p_2}{q_2}, \ldots, \frac{p_n}{q_n}$ bezeichnen, so daß also insbesondere

$$\left.\begin{aligned} &p_0 = a_0,\ p_1 = a_0 a_1 + 1,\ p_2 = (a_0 a_1 + 1)\, a_2 + a_0, \\ &q_0 = 1,\ q_1 = a_1,\qquad\quad q_2 = a_1 a_2 + 1 \end{aligned}\right\} \tag{7}$$

und schließlich

$$\frac{p_n}{q_n} = \frac{a}{b}$$

ist. Aus den Relationen (7) ergibt sich unmittelbar, daß

$$p_2 = p_1 a_2 + p_0,$$
$$q_2 = q_1 a_2 + q_0$$

ist. Es ist nun von außerordentlicher Wichtigkeit, daß dieses Bildungsgesetz ganz allgemein für die Zähler und die Nenner der Näherungsbrüche gilt, daß also für jedes k $(2 \leq k \leq n)$

$$\left.\begin{aligned} p_k &= p_{k-1} a_k + p_{k-2}, \\ q_k &= q_{k-1} a_k + q_{k-2} \end{aligned}\right\} \tag{8}$$

ist.

Für $k = 2$ haben wir diese Relationen bereits bewiesen. Es bleibt also zu zeigen, daß aus der Gültigkeit von (8) für ein gewisses $k < n$ die Gültigkeit für $k + 1$ folgt. Nun ist aber für jedes $k < n$

$$\frac{p_k}{q_k} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_k}}},$$

$$\frac{p_{k+1}}{q_{k+1}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_k + \cfrac{1}{a_{k+1}}}}},$$

d.h., $\frac{p_{k+1}}{q_{k+1}}$ geht dadurch aus $\frac{p_k}{q_k}$ hervor, daß man a_k durch $a_k + \frac{1}{a_{k+1}}$ ersetzt. Da aber gemäß (8)

$$\frac{p_k}{q_k} = \frac{p_{k-1} a_k + p_{k-2}}{q_{k-1} a_k + q_{k-2}} \tag{9}$$

ist, wobei die hier eingehenden Zahlen $p_{k-1}, p_{k-2}, q_{k-1}, q_{k-2}$ offensichtlich nur von $a_0, a_1, \ldots, a_{k-1}$, aber nicht von a_k abhängen, so ist

$$\frac{p_{k+1}}{q_{k+1}} = \frac{p_{k-1}\left(a_k + \frac{1}{a_{k+1}}\right) + p_{k-2}}{q_{k-1}\left(a_k + \frac{1}{a_{k+1}}\right) + q_{k-2}} = \frac{(p_{k-1} a_k + p_{k-2})\, a_{k+1} + p_{k-1}}{(q_{k-1} a_k + q_{k-2})\, a_{k+1} + q_{k-1}} = \frac{p_k a_{k+1} + p_{k-1}}{q_k a_{k+1} + q_{k-1}},$$

d. h., p_{k+1} und q_{k+1} genügen den Formeln, die man aus (8) erhält, wenn man dort k durch $k+1$ ersetzt.

Bemerkung. Wir werden uns bald davon überzeugen, daß die Brüche (9) sämtlich reduziert (also unkürzbar) sind, wenn man die Zahlen p_k und q_k gemäß den Rekursionsformeln (8) bestimmt. Bis dahin wollen wir unter dem Zähler und dem Nenner eines Näherungsbruches stets die Zahlen p_k und q_k verstehen, die sich rekursiv aus den Formeln (8) bestimmen, wobei wir zunächst die Frage nach der Reduziertheit der Brüche (9) ganz außer acht lassen.

Den Bruch $\frac{p_k}{q_k}$ nennen wir den *Näherungsbruch k-ter Ordnung*. Offensichtlich ist dieser Bruch eine Funktion der Zahlen $a_0, a_1, \ldots, a_k$. Die Gleichung (9) zeigt überdies, daß wegen der bereits erwähnten Unabhängigkeit der $p_{k-1}, q_{k-1}, p_{k-2}, q_{k-2}$ von a_k der Bruch $\frac{p_k}{q_k}$ bei festem $a_0, a_1, \ldots, a_{k-1}$ (also nach Wahl fester Zahlenwerte für diese Variablen) eine gebrochene lineare Funktion von a_k ist. Das ist im wesentlichen die Aussage der Rekursionsformeln (8), auf denen sich die ganze Theorie der Kettenbrüche aufbaut.

Wir wollen jetzt die wichtigsten Eigenschaften der Näherungsbrüche zusammenstellen. Dazu führen wir die folgende Bezeichnung ein:

$$q_k p_{k+1} - p_k q_{k+1} = \Delta_k \quad (k = 0, 1, \ldots, n-1).$$

Auf Grund der Rekursionsgleichungen (8) ist dann

$$\Delta_k = q_k(p_k a_{k+1} + p_{k-1}) - p_k(q_k a_{k+1} + q_{k-1}) = -(q_{k-1} p_k - p_{k-1} q_k) = -\Delta_{k-1},$$

d. h., sämtliche Δ_k besitzen denselben absoluten Betrag, während ihr Vorzeichen laufend wechselt. Indem wir noch feststellen, daß gemäß (7)

$$\Delta_0 = q_0 p_1 - p_0 q_1 = 1$$

ist, kommen wir zu der folgenden wichtigen Aussage:

Satz 1. $\Delta_k = (-1)^k \quad (0 \leq k < n)$.

Hieraus folgt nun unmittelbar, daß sämtliche Näherungsbrüche reduziert sind. In der Tat: Wären etwa die Zahlen p_k und q_k durch ein und dieselbe Zahl $d > 1$ teilbar, so wäre auch Δ_k durch d teilbar, was nach Satz 1 unmöglich ist.

Wenn speziell unser Ausgangsbruch $\frac{a}{b}$ reduziert ist, so ist $a = p_n$ und $b = q_n$ und mithin

$$\Delta_{n-1} = q_{n-1} a - p_{n-1} b = (-1)^{n-1}.$$

Damit haben wir (nun zum vierten Mal) den Satz bewiesen, daß es zu teilerfremden ganzen Zahlen a und b stets ganze Zahlen x und y gibt derart, daß

$$ax + by = 1$$

ist. Der hier erhaltene Beweis liefert uns zudem eine Methode zur Bestimmung der Zahlen x und y, die in den meisten Fällen der praktischen Anwendung bei weitem am kürzesten ist. Wir sehen, daß man dazu lediglich $\frac{a}{b}$ in Form eines Kettenbruches darstellen muß und dann, wenn $a = p_n$ und $b = q_n$ ist, $x = (-1)^{n-1} q_{n-1}$ und $y = (-1)^n p_{n-1}$ setzen kann.

Beispiel. Es sei $a = 52$, $b = 23$. Dann ist

$$\frac{52}{23} = 2 + \frac{6}{23} = 2 + \frac{1}{3 + \frac{5}{6}} = 2 + \frac{1}{3 + \frac{1}{1 + \frac{1}{5}}} = \frac{p_3}{q_3},$$

$$\frac{p_2}{q_2} = 2 + \frac{1}{3 + \frac{1}{1}} = \frac{9}{4};$$

$$x = 4, \quad y = -9, \quad ax + by = 1.$$

Satz 1 gibt uns ferner einen vollständigen Überblick über die Anordnung der Näherungsbrüche. Wegen $\Delta_k = (-1)^k$ ist nämlich

$$\frac{\Delta_k}{q_k q_{k+1}} = \frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k} = \frac{(-1)^k}{q_k q_{k+1}}.$$

d. h., für gerades k ist $\frac{p_{k+1}}{q_{k+1}} > \frac{p_k}{q_k}$, während für ungerades k umgekehrt $\frac{p_{k+1}}{q_{k+1}} < \frac{p_k}{q_k}$ ist. Wenn wir also von den Näherungsbrüchen $\frac{p_0}{q_0}, \frac{p_1}{q_1}$ ausgehen, so können wir zunächst feststellen, daß $\frac{p_1}{q_1} > \frac{p_0}{q_0}$ ist (Abb. 1). Ferner ist $\frac{p_2}{q_2} < \frac{p_1}{q_1}$; da aber der Abstand $\left|\frac{p_1}{q_1} - \frac{p_2}{q_2}\right| = \frac{1}{q_1 q_2}$ von $\frac{p_1}{q_1}$ zu $\frac{p_2}{q_2}$ kleiner als der Abstand $\left|\frac{p_1}{q_1} - \frac{p_0}{q_0}\right| = \frac{1}{q_0 q_1}$ von $\frac{p_1}{q_1}$ zu $\frac{p_0}{q_0}$ ist, so liegt $\frac{p_2}{q_2}$ zwischen $\frac{p_0}{q_0}$ und $\frac{p_1}{q_1}$, wie in Abb. 1 dargestellt ist.

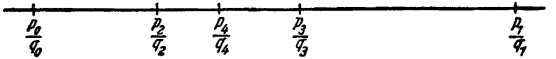


Abb. 1

Wenn wir weiter die Lage des Bruches $\frac{p_3}{q_3}$ ermitteln wollen, so beachten wir zunächst, daß $\frac{p_2}{q_2} > \frac{p_3}{q_3}$ sein muß; da aber andererseits

$$\left|\frac{p_3}{q_3} - \frac{p_2}{q_2}\right| = \frac{1}{q_2 q_3} < \frac{1}{q_1 q_2} = \left|\frac{p_1}{q_1} - \frac{p_2}{q_2}\right|$$

ist, so muß $\frac{p_3}{q_3}$ näher an $\frac{p_2}{q_2}$ als an $\frac{p_1}{q_1}$ liegen, d. h., $\frac{p_3}{q_3}$ liegt zwischen $\frac{p_2}{q_2}$ und $\frac{p_1}{q_1}$, wie auch in Abb. 1 dargestellt ist. In entsprechender Weise können wir diese Überlegung bis zu $\frac{p_n}{q_n} = \frac{a}{b}$ fortsetzen. Als Schlußfolgerung aus dieser Überlegung erhalten wir offensichtlich:

Satz 2. *Alle Näherungsbrüche gerader Ordnung von $\frac{a}{b}$ sind kleiner als $\frac{a}{b}$, und ihre Größe nimmt mit wachsender Ordnung zu. Alle Näherungsbrüche ungerader Ordnung sind größer als $\frac{a}{b}$, und ihre Größe nimmt mit wachsender Ordnung ab.*

Hieraus folgt insbesondere, daß die Zahl $\frac{a}{b}$ zwischen je zwei aufeinanderfolgenden Näherungsbrüchen $\frac{p_k}{q_k}$ und $\frac{p_{k+1}}{q_{k+1}}$ liegt. Folglich ist der Abstand zwischen $\frac{a}{b}$ und jedem dieser Brüche kleiner als der Abstand zwischen diesen beiden Brüchen, der gleich $\frac{1}{q_k q_{k+1}}$ ist. Auf diese Weise kommen wir zu der folgenden Aussage, die in vielen Fragen der Approximation von Zahlen eine erhebliche Rolle spielt:

Satz 3.

$$\left|\frac{a}{b}-\frac{p_k}{q_k}\right|<\frac{1}{q_k q_{k+1}}.$$

Die Theorie der Kettenbrüche entwickelte sich historisch aus dem Bedürfnis heraus, einen Bruch mit großem Zähler und Nenner möglichst gut durch einen Bruch mit bedeutend kleinerem Zähler und Nenner anzunähern. Der Begründer der Theorie der Kettenbrüche, CHRISTIAN HUYGENS, hatte sich nämlich die Aufgabe gestellt, ein Zahnradmodell des Sonnensystems zu bauen. Dabei begegnete er einer eigentümlichen Schwierigkeit: Um das Verhältnis der Umlaufszeiten zweier ineinander greifender Zahnräder gleich dem Verhältnis der Umlaufszeiten (um die Sonne) der durch sie dargestellten Planeten zu machen, müssen die Anzahlen der Zähne dieser Räder in dem gleichen Verhältnis stehen. Das Verhältnis wird jedoch von so großen Zahlen gebildet, daß es technisch unmöglich ist, Räder mit derartigen „astronomischen" Zähnezahlen herzustellen. Deshalb war es notwendig, sich auf ein Näherungsmodell zu beschränken, wobei die Anzahlen der Zähne einerseits im Rahmen des technisch Möglichen liegen mußten, andererseits aber das Verhältnis dieser Zahlen doch recht nahe dem Verhältnis der sehr großen Zahlen liegen sollte. Hierzu gibt nun die Theorie der Kettenbrüche die notwendigen Anweisungen. Es seien nämlich a und b die gegebenen großen Zahlen, deren Verhältnis durch das Verhältnis der kleineren Zahlen c und d approximiert werden soll, wobei aus technischen oder anderen Gründen die Zahl d nicht größer als z. B. 100 sein darf. Dann stellen wir $\frac{a}{b}$ in Form eines Kettenbruches dar und berechnen die aufeinanderfolgenden Näherungsbrüche. Es möge nun $q_k \leq 100$, aber $q_{k+1} > 100$ sein. Dann setzen wir $c = p_k$ und $d = q_k$. Satz 3 gestattet es uns, den Fehler abzuschätzen, den wir machen, wenn wir das Verhältnis $\frac{a}{b}$ durch das Verhältnis $\frac{c}{d}$ ersetzen.

Beispiel. $a = 1355$, $b = 946$. Dann ist

$$\frac{a}{b} = [1; 2, 3, 5, 8, 3],$$

$$\frac{p_3}{q_3}=\frac{53}{37}, \qquad \frac{p_4}{q_4}=\frac{434}{303},$$

$$\left|\frac{a}{b}-\frac{p_3}{q_3}\right|=\left|\frac{1355}{946}-\frac{53}{37}\right|<\frac{1}{q_3 q_4}=\frac{1}{11211}<0{,}0001.$$

Hätten wir zur Lösung dieser Aufgabe an Stelle von Kettenbrüchen Dezimalbrüche benutzt, so hätten wir für die gleiche Genauigkeit einen Bruch mit wenigstens vier Stellen hinter dem Komma benutzen müssen, also einen Bruch mit dem Nenner 10000. Das hätte aber in unserem Falle gar keinen Sinn gehabt, da der anzunähernde Bruch nur den Nenner 946 besitzt. Wir sehen also, daß der Apparat der Kettenbrüche für derartige Aufgaben nicht nur den Dezimalbrüchen gegenüber den Vorteil der bequemeren Handhabung bietet, sondern daß er auch in vielen Fällen (wie

in dem angeführten Beispiel) noch dann anwendbar ist, wenn die Dezimalbrüche zwangsläufig versagen. Grund dafür ist einfach, daß bei den Kettenbrüchen als Nenner der Näherungsbrüche beliebige Zahlen auftreten können, während bei den endlichen Dezimalbrüchen die Nenner auf Potenzen von 10 beschränkt sind, daß also bei den Kettenbrüchen eine erheblich größere Freiheit der Auswahl besteht. Dabei darf man jedoch nicht aus den Augen verlieren, daß in anderen Beziehungen die Dezimalbrüche (und die anderen systematischen Brüche) erhebliche Vorteile gegenüber den Kettenbrüchen bieten. So ist allgemein bekannt, wie leicht sich die arithmetischen Operationen an den in Dezimalbruchdarstellung geschriebenen Zahlen ausführen lassen, während es z. B. bis heute noch nicht gelungen ist, die allgemeine Regel für die Addition zweier Kettenbrüche zu finden. Der richtige Standpunkt besteht hier — wie überall — nicht darin, daß man über die relativen Vorteile zweier leistungsfähigen Methoden rechtet, sondern darin, daß man beide beherrscht und jede im richtigen Augenblick und am richtigen Orte anzuwenden vermag.

Kapitel IV

DIE DARSTELLUNG DER ZAHLEN DURCH SYSTEMATISCHE BRÜCHE UND DURCH KETTENBRÜCHE

§ 10. Einleitung

Beide Prinzipien, sowohl das Prinzip der systematischen Brüche (also speziell der Dezimalbrüche) als auch das Prinzip der Kettenbrüche, erfüllen nur dann vollständig ihren Zweck, wenn man mit ihrer Hilfe die Gesamtheit aller reellen Zahlen (d. h. auch die irrationalen Zahlen) erfaßt. Während wir nämlich für die Darstellung der rationalen Zahlen in den gewöhnlichen Brüchen einen formalen Apparat besitzen, der den weitaus meisten Anforderungen gerecht wird und der nur in Ausnahmefällen die Benutzung anderer formaler Hilfsmittel erfordert, besitzen wir für die irrationalen Zahlen keine entsprechende Darstellungsweise. Für sie wird erst durch die systematischen Brüche und durch die Kettenbrüche das formale Rüstzeug zu ihrer Erforschung geliefert.

Wir haben bereits betont, daß der wesentliche Vorteil der systematischen Brüche darin liegt, daß sich an ihnen die wichtigsten arithmetischen Operationen sehr einfach ausführen lassen. Wenn wir uns hingegen die Aufgabe stellen, die arithmetische Struktur einer gegebenen Zahl, sei sie rational oder irrational, zu untersuchen, so wäre es äußerst unzweckmäßig, sie in Form eines systematischen Bruches (etwa eines Dezimalbruches) anzusetzen, da die Reihenfolge der Ziffern in einem solchen Bruch nicht die absolute Natur dieser Zahl charakterisiert, sondern nur ihre Wechselbeziehungen zu dem ausgewählten Zahlensystem. Hierzu genügt es, daran zu erinnern, daß unter Umständen ein und dieselbe Zahl in einem Zahlensystem durch einen endlichen und in einem anderen Zahlensystem durch einen unendlichen

systematischen Bruch dargestellt wird. Hingegen haben wir in den Kettenbrüchen eine Darstellungsweise der Zahlen, die von keinem speziellen Zahlensystem abhängt, sondern ausschließlich durch ihre eigentliche arithmetische Natur bedingt ist und sich mithin sehr gut zur Erforschung ihrer arithmetischen Struktur eignet.

In der Oberschule lernen zwar die Schüler die Grundrechenarten an den Dezimalbrüchen kennen, jedoch wird ihnen fast nichts über die theoretischen Grundlagen dieses Apparates gesagt. Das ist dadurch bedingt, daß die Dezimalbrüche schon verhältnismäßig früh behandelt werden, während die strenge Begründung der Theorie der Dezimalbrüche eine Gewöhnung an formales Denken erfordert, die zu dieser Zeit das Verständnis der Schüler weit übersteigt. Daher wird jene Theorie in den Mathematiklehrbüchern der Oberschulen entweder gar nicht oder nur an Beispielen und ohne tiefere Überlegungen über ihre Anwendungen gebracht. Jedoch kann ein Lehrer selbst einen solchen Unterricht über Dezimalbrüche nicht sicher führen, wenn ihm nicht die theoretische Grundlage der Lehre von den Dezimalbrüchen in allen ihren Teilen klar ist.

Das gilt in gleichem Maße auch von den zur Zeit im Unterrichtsplan der Oberschulen gar nicht vorgesehenen Kettenbrüchen; gehören doch die Kettenbrüche, die eine der lebensfähigsten und wirksamsten arithmetischen Schöpfungen der klassischen Epoche darstellen und die heute ein wichtiges Hilfsmittel mathematischer Forschung sind, schon lange zum Besitz jedes gebildeten Mathematikers.

§ 11. Die systematischen Brüche

Als Grundzahl des betrachteten Zahlensystems wollen wir eine beliebige ganze Zahl $k > 1$ nehmen. In diesem Sinne werden wir von der Darstellung der Zahlen als *k-adische* Brüche sprechen. In den Anwendungen wählt man gewöhnlich $k = 10$ (Dezimalbrüche). Für den Aufbau der Theorie ist es jedoch gänzlich belanglos, welche Zahl man als Grundzahl des Zahlensystems nimmt, so daß wir keinen Anlaß haben, diese Zahl näher zu fixieren. Unser Ziel ist es, jede reelle Zahl α in diesem System darzustellen. Zur Vereinfachung der Rechnungen und der Schreibweise wollen wir annehmen, daß $0 \leq \alpha < 1$ ist. Offenbar bedeutet das keine wesentliche Einschränkung der Allgemeinheit unserer Überlegungen, da die unter Umständen notwendige Ergänzung durch eine ganze Zahl keine zusätzlichen Schwierigkeiten bereitet.

Definition 1. *Die Zahlen* $0, 1, 2, \ldots, k-1$ *nennen wir Ziffern.*

Definition 2. *Vorgegeben sei eine unendliche Folge von Ziffern*

$$a_1, a_2, \ldots, a_n, \ldots \quad (0 \leq a_n < k).$$

Dann heißt die unendliche Reihe

$$\frac{a_1}{k} + \frac{a_2}{k^2} + \cdots + \frac{a_n}{k^n} + \cdots \tag{1}$$

ein k-adischer Bruch (oder kurz Bruch), wenn nicht von einer gewissen Stelle ab alle a_n *gleich* $k-1$ *sind.*

Einen derartigen Bruch schreibt man gewöhnlich in der Form

$$0, a_1 a_2 \ldots a_n \ldots.$$

Der Grund, der uns zum Ausschluß derjenigen Brüche veranlaßt, die auf eine unendliche Folge von Ziffern $k-1$ enden, ist aus der elementaren Theorie der Dezimalbrüche hinlänglich bekannt. Wir können nämlich eine Zahl, die durch einen solchen Bruch dargestellt wird, stets auch mit Hilfe eines anderen Bruches darstellen, der diese Eigenschaft nicht besitzt.[1]) Wir werden also die eindeutige Darstellbarkeit der Zahlen durch Brüche nur dann erreichen können, wenn wir eine dieser beiden Darstellungsarten der genannten Zahlen ausschließen.

Definition 3. *Der Bruch* (1) *heißt endlich, wenn von einer gewissen Stelle ab alle Ziffern* a_n *gleich Null sind. Ein Bruch, der nicht endlich ist, heißt unendlich.*

Im Fall eines endlichen Bruches schreibt man üblicherweise den Rest aus lauter Nullen nicht mit, d. h., man schreibt für $0, a_1 a_2 \dots a_n 000 \dots$ kurz $0, a_1 \dots a_n$.

Da für jedes $n \geq 1$

$$0 \leq \frac{a_n}{k^n} \leq \frac{k-1}{k^n} < \frac{1}{k^{n-1}}$$

ist, konvergiert die Reihe $\sum_{i=1}^{\infty} \frac{a_i}{k^i}$, d. h., wenn wir $s_0 = 0$ und für jedes $n \geq 1$

$$s_n = \sum_{i=1}^{n} \frac{a_i}{k^i}$$

setzen, so existiert jedenfalls der Grenzwert

$$\lim_{n \to \infty} s_n = \alpha.$$

Definition 4. *Wenn* $\sum_{i=1}^{\infty} \frac{a_i}{k^i} = \lim_{n \to \infty} s_n = \alpha$ *ist, so sagt man, daß der Bruch* (1) *die Zahl* α *darstellt oder kurz, daß er gleich* α *ist.*

In diesem Sinne stellt jeder k-adische Bruch eine gewisse reelle Zahl dar.

Satz 1. *Wenn der Bruch* (1) *die Zahl* α *darstellt, so ist für jedes* $n \geq 0$

$$0 \leq \alpha - s_n < \frac{1}{k^n}.$$

Bemerkung. Setzt man in Satz 1 speziell $n = 0$, so sieht man, daß stets $0 \leq \alpha < 1$ ist. Die von uns betrachteten k-adischen Brüche stellen also nur reelle Zahlen aus diesem Intervall dar.

Beweis. Da stets

$$\alpha - s_n = \sum_{i=n+1}^{\infty} \frac{a_i}{k^i}$$

ist, ist offensichtlich $\alpha - s_n \geq 0$; da sich andererseits unter den Ziffern $a_{n+1}, a_{n+2}, \dots$ stets solche befinden, die kleiner als $k-1$ sind, ist

$$\alpha - s_n < \sum_{i=n+1}^{\infty} \frac{k-1}{k^i} = \frac{k-1}{k^{n+1}} \left\{ 1 + \frac{1}{k} + \frac{1}{k^2} + \cdots \right\} = \frac{1}{k^n},$$

womit Satz 1 bewiesen ist.

[1]) So ist im Fall der Dezimalbrüche $0{,}1999 \dots = 0{,}2000 \dots$.

Satz 2. *Damit der Bruch* (1) *die Zahl* α *darstellt, ist notwendig und hinreichend, daß für alle* $n \geq 1$

$$a_n = [k^n \alpha] - k\,[k^{n-1}\alpha] \tag{2}$$

ist.

Bemerkung. Wie üblich, verwenden wir das Zeichen $[x]$ zur Bezeichnung der größten ganzen Zahl, die nicht größer als x ist. In diesem Sinne ist also $[\pi] = 3$, $[-\pi] = -4$ usw.

Beweis. 1.) Es seien die Zahl α und der Bruch (1) so gewählt, daß für jedes $n \geq 1$ die Relation (2) erfüllt ist. Dann ist

$$s_n = \sum_{i=1}^{n} \frac{a_i}{k^i} = \sum_{i=1}^{n} \frac{[k^i\alpha] - k\,[k^{i-1}\alpha]}{k^i} = \sum_{i=1}^{n} \left\{ \frac{[k^i\alpha]}{k^i} - \frac{[k^{i-1}\alpha]}{k^{i-1}} \right\} = \frac{[k^n\alpha]}{k^n} - [\alpha] = \frac{[k^n\alpha]}{k^n}$$

und mithin

$$\alpha - s_n = \frac{k^n\alpha - [k^n\alpha]}{k^n}.$$

Da nun aber für jedes x offensichtlich $0 \leq x - [x] < 1$ ist, ist

$$0 \leq \alpha - s_n < \frac{1}{k^n}, \quad (n = 1, 2, \ldots)$$

und es konvergiert s_n gegen α, falls n gegen Unendlich geht, d. h., der Bruch (1) stellt die Zahl α dar.

2.) Es möge nun umgekehrt der Bruch (1) die Zahl α darstellen. Dann ist auf Grund von Satz 1 für jedes $n \geq 0$

$$0 \leq \alpha - s_n < \frac{1}{k^n}$$

und daher

$$k^n s_n \leq k^n \alpha < k^n s_n + 1.$$

Weil aber $k^n s_n$ eine ganze Zahl ist, ergibt diese letzte Ungleichung, daß

$$k^n s_n = [k^n \alpha]$$

ist. Nun ist aber für jedes $n \geq 1$

$$s_n - s_{n-1} = \frac{a_n}{k^n}, \quad a_n = k^n (s_n - s_{n-1}),$$

so daß also in der Tat

$$a_n = k^n s_n - k(k^{n-1} s_{n-1}) = [k^n \alpha] - k\,[k^{n-1}\alpha]$$

ist, was noch zu beweisen war.

Satz 3. *Jede reelle Zahl* α *des Intervalls* $0 \leq \alpha < 1$ *läßt sich durch einen und nur einen k-adischen Bruch darstellen.*

Beweis. Die Eindeutigkeit der Darstellung ist eine unmittelbare Folge aus Satz 2, nach welchem die Ziffern a_n, die den darstellenden Bruch bilden, gemäß der Formel (2) eindeutig durch die Zahl α festgelegt sind. Was die Möglichkeit der Darstellung betrifft, so genügt es hierzu — wiederum wegen Satz 2 —, zu zeigen, daß die Zahlen a_n, die nach Formel (2) bestimmt sind, für jedes α aus dem

angegebenen Intervall als Ziffern eines k-adischen Bruches dienen können. Es ist also zu zeigen, daß für alle $n \geq 1$

$$0 \leq a_n < k$$

ist und daß von keiner Stelle ab alle a_n gleich $k-1$ sind.

Dazu setzen wir für ein gegebenes $n \geq 1$

$$[k^{n-1}\alpha] = g, \quad k^{n-1}\alpha = g + \beta \quad (0 \leq \beta < 1),$$

so daß also

$$[k^n\alpha] = [kg + k\beta] = kg + [k\beta]^1),$$
$$k\,[k^{n-1}\alpha] = kg$$

und mithin

$$a_n = [k^n\alpha] - k\,[k^{n-1}\alpha] = [k\beta]$$

ist. Da aber $0 \leq \beta < 1$ ist, gilt

$$0 \leq a_n < k.$$

Wegen $0 \leq \beta < 1$ gibt es ferner eine natürliche Zahl q derart, daß

$$1 - \frac{1}{k^{q-1}} \leq \beta < 1 - \frac{1}{k^q}$$

und daher

$$k^{q-1}\beta \geq k^{q-1} - 1, \quad k^q\beta < k^q - 1$$

ist. Hieraus folgt aber unmittelbar (wir erinnern daran, daß $k^{n-1}\alpha = g + \beta$ war), daß

$$\begin{aligned} a_{n+q-1} &= [k^{n+q-1}\alpha] - k\,[k^{n+q-2}\alpha] = [k^q(g+\beta)] - k\,[k^{q-1}(g+\beta)] \\ &= k^q g + [k^q\beta] - k^q g - k\,[k^{q-1}\beta] < k^q - 1 - k(k^{q-1} - 1) = k - 1 \end{aligned}$$

ist. Da hier n ganz beliebig ist, sehen wir, daß es im betrachteten Bruch hinter jeder Ziffer a_i stets noch eine Ziffer gibt, die kleiner als $k-1$ ist, womit wir Satz 3 bewiesen haben.

Damit ist gezeigt, daß die systematischen Brüche zu beliebiger Grundzahl als ein formales Mittel zur Darstellung der reellen Zahlen dienen können, da sie die beiden Hauptforderungen, nämlich die nach der Existenz und die nach der Eindeutigkeit der Darstellung der reellen Zahlen, befriedigen. Wir wollen jetzt zu einer Reihe von Sätzen kommen, die eine Verbindung zwischen der arithmetischen Natur einer gegebenen Zahl und dem sie darstellenden systematischen Bruch herstellen.

Satz 4. *Jeder periodische k-adische Bruch stellt eine rationale Zahl dar.*

Beweis. Wir wollen den Bruch (1) *periodisch* nennen, wenn es Zahlen $r \geq 0$ und $s \geq 1$ gibt derart, daß $a_n = a_{n+s}$ ist für alle $n > r$. Dann kann man eine Zahl α, die durch einen derartigen periodischen Bruch dargestellt wird, auch in der Form

$$\alpha = \sum_{i=1}^{\infty} \frac{a_i}{k^i} = \sum_{i=1}^{r} \frac{a_i}{k^i} + \sum_{i=1}^{s} \frac{a_{r+i}}{k^{r+i}} \left\{1 + \frac{1}{k^s} + \frac{1}{k^{2s}} + \cdots\right\} = \sum_{i=1}^{r} \frac{a_i}{k^i} + \sum_{i=1}^{s} \frac{a_{r+i}}{k^{r+i}} \cdot \frac{1}{1 - \frac{1}{k^s}}$$

schreiben, woraus unmittelbar ersichtlich ist, daß α rational ist, was zu beweisen war.

[1]) Dies folgt daraus, daß für jede ganze Zahl a mit $x = a + y$ auch $[x] = a + [y]$ ist.

Bemerkung. Der Bruch (1) heißt *reinperiodisch*, wenn die Zahl $r = 0$ ist, und *gemischtperiodisch*, wenn $r > 0$ ist. Im Falle eines reinperiodischen Bruches, also im Falle $r = 0$, führt die letzte Gleichung für α auf

$$\alpha = \frac{a_1 k^{s-1} + a_2 k^{s-2} + \cdots + a_s}{k^s - 1}.$$

Da nun die Zahl $k^s - 1$ jedenfalls zu k teilerfremd ist, erhalten wir die

Ergänzung. *Jeder reinperiodische Bruch stellt eine rationale Zahl $\alpha = \frac{a}{b}$ dar, in der b zu k teilerfremd ist.*

Satz 5. *Wenn $\alpha = \frac{a}{b}$ eine rationale Zahl ist, in der b zu k teilerfremd ist, so wird α durch einen reinperiodischen Bruch dargestellt.*

Beweis. Auf Grund des Eulerschen Satzes (S. 249) ist

$$k^{\varphi(b)} \equiv 1 \pmod{b}.$$

Wenn wir also $\varphi(b) = h$ setzen, so ist

$$k^h - 1 = bq,$$

wobei q eine gewisse ganze Zahl ist. Daher ist

$$\alpha = \frac{a}{b} = \frac{aq}{bq} = \frac{aq}{k^h - 1} = \frac{aq}{k^h} \cdot \frac{1}{1 - \frac{1}{k^h}} = \frac{aq}{k^h}\left\{1 + \frac{1}{k^h} + \frac{1}{k^{2h}} + \cdots\right\}$$

und mithin

$$k^h \alpha = aq\left\{1 + \frac{1}{k^h} + \frac{1}{k^{2h}} + \cdots\right\} = aq + \frac{aq}{k^h}\left\{1 + \frac{1}{k^h} + \frac{1}{k^{2h}} + \cdots\right\} = aq + \alpha,$$

woraus sich für beliebiges $n \geq 1$

$$\begin{aligned} a_{n+h} &= [k^{n+h}\alpha] - k\,[k^{n+h-1}\alpha] = [k^n(aq + \alpha)] - k\,[k^{n-1}(aq + \alpha)] \\ &= k^n aq + [k^n \alpha] - k^n aq - k\,[k^{n-1}\alpha] = [k^n \alpha] - k\,[k^{n-1}\alpha] = a_n \end{aligned}$$

ergibt, was zu beweisen war.

Bemerkung. Man kann jedoch aus dem Beweis von Satz 5 nicht entnehmen, daß die Periode des Bruches, der die Zahl α darstellt, notwendig gleich $h = \varphi(b)$ ist. Vielmehr ist es durchaus möglich, daß die Periode dieses Bruches kleiner als h ist, so daß also h aufeinanderfolgende Ziffern des Bruches nicht eine, sondern mehrere Perioden enthalten. Die Frage, wie sich zu gegebenem a und b die Länge der Periode bestimmt, ist von außerordentlichem Interesse, wir können hier jedoch nicht daran denken, sie eingehend zu behandeln. Dasselbe gilt für die entsprechende Frage im Zusammenhang mit dem nachfolgenden Satz 6.

Satz 6. *Wenn die Zahl b nicht zur Zahl k teilerfremd ist, so wird die Zahl $\alpha = \frac{a}{b}$ (wobei a und b ganz und teilerfremd seien) durch einen gemischtperiodischen Bruch dargestellt.*

Beweis. Es sei $\frac{a'}{b'}$ die reduzierte (unkürzbare) Form der Zahl $k^r \alpha = k^r \frac{a}{b}$. Wenn wir hier die Zahl r hinreichend groß wählen, so können wir erreichen, daß sich dabei alle in b vorhandenen gemeinsamen Primfaktoren der Zahlen b und k wegkürzen, so daß wir also b' und k als teilerfremd annehmen können. Dann sei $[k^r \alpha] = q$ und

$$k^r \alpha = \frac{a'}{b'} = q + \frac{a''}{b'} \qquad \left(0 \leq \frac{a''}{b'} < 1\right). \tag{3}$$

Nach Satz 5 ist der die Zahl $\frac{a''}{b'}$ darstellende k-adische Bruch $0, b_1 b_2 \ldots b_n \ldots$ reinperiodisch, also etwa $b_{n+h} = b_n$ $(n = 1, 2, \ldots)$. Bezeichnen wir mit $0, a_1 a_2 \ldots a_n \ldots$ den die Zahl α darstellenden Bruch, so erhalten wir gemäß (2) und (3) für jedes $n \geq 1$

$$a_{r+n} = [k^{r+n}\alpha] - k[k^{r+n-1}\alpha] = \left[k^n\left(q + \frac{a''}{b'}\right)\right] - k\left[k^{n-1}\left(q + \frac{a''}{b'}\right)\right]$$

$$= \left[k^n \frac{a''}{b'}\right] - k\left[k^{n-1}\frac{a''}{b'}\right] = b_n,$$

und mithin

$$a_{r+n+h} = a_{r+n} \quad (n = 1, 2, \ldots).$$

Damit ist gezeigt, daß der die Zahl $\alpha = \frac{a}{b}$ darstellende Bruch $0, a_1 a_2 \cdots a_n \cdots$ periodisch ist. Er ist aber sicher nicht reinperiodisch, da sonst — gemäß der Ergänzung zu Satz 4 — die Zahl b zu k teilerfremd sein müßte. Damit ist Satz 6 bewiesen.

Bemerkung. Wenn alle Primfaktoren, die in der Zahl b enthalten sind, auch in der Zahl k enthalten sind, so ist im Beweis vom Satz 6 die Zahl b' gleich Eins. Wegen $0 \leq \frac{a''}{b'} < 1$ muß dann aber $a'' = 0$ sein, so daß auch alle b_n gleich Null sind. Dann ist aber gemäß (4) auch

$$a_{r+n} = 0 \quad (n = 1, 2, \ldots),$$

d. h., die Zahl $\alpha = \frac{a}{b}$ wird durch einen endlichen Bruch dargestellt. Hiervon gilt nun offensichtlich auch die Umkehrung: Wird nämlich die Zahl α durch einen endlichen Bruch dargestellt, so ist $\alpha = \frac{g}{k^n}$, wobei g und n gewisse ganze Zahlen sind und $n \geq 1$ ist. Wenn also $\frac{a}{b}$ die reduzierte Form von α ist, dann besitzt b keinen Primteiler, der nicht auch Primteiler von k ist. Wir erhalten also als

Ergänzung. *Damit die Zahl α durch einen endlichen k-adischen Bruch dargestellt wird, ist notwendig und hinreichend, daß α rational ist und daß in der reduzierten Form $\alpha = \frac{a}{b}$ die Zahl b keine anderen Primfaktoren enthält als die Zahl k.*

Die eben bewiesenen Sätze zeigen, daß sich die irrationalen Zahlen und nur diese als nichtperiodische k-adische Brüche darstellen. Das ist aber auch die einzige Beziehung, die zwischen der arithmetischen Natur einer reellen Zahl und dem sie darstellenden systematischen Bruch unabhängig von der gewählten Grundzahl besteht. Das dabei gefundene Gesetz besagt, daß unabhängig von der Grundzahl des gewählten Zahlensystems die rationalen Zahlen eine periodische und die irrationalen Zahlen eine nichtperiodische Darstellung besitzen. Alle anderen Besonderheiten der Darstellung einer rationalen Zahl, die wir betrachtet haben, hingen bereits von arithmetischen Beziehungen der rationalen Zahl zur gewählten Grundzahl ab.

§ 12. Die Kettenbrüche

In Kapitel III haben wir gesehen, daß sich jede rationale Zahl eindeutig in Form eines Kettenbruches $[a_0; a_1, \ldots, a_n]$ darstellen läßt und daß umgekehrt jeder Kettenbruch eine gewisse rationale Zahl darstellt. Wenn wir also mit den Kettenbrüchen auch die irrationalen Zahlen erfassen wollen, so müssen wir uns zwangsläufig nach einer Erweiterung des Begriffs des Kettenbruches umsehen. Es liegt nun recht nahe, auch hier unendliche Kettenbrüche, d. h. Ausdrücke der Form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \ddots}}}} = [a_0; a_1, a_2, \ldots, a_n, \ldots], \tag{5}$$

wobei a_0 eine ganze Zahl ist und $a_1, a_2, \ldots, a_n, \ldots$ natürliche Zahlen sind, in den Kreis der Betrachtungen aufzunehmen.

Es ist klar, daß ein Ausdruck dieser Art so lange keine bestimmte Bedeutung besitzt, als wir keine Verabredung in dieser Frage treffen.[1])

Bevor wir zu einer zweckmäßigen Vereinbarung über den Sinn des Ausdrucks (5) kommen, bemerken wir, daß wir zunächst rein formal, ohne daß wir dem Symbol (5) vorher irgendeinen realen Inhalt beilegen, die Folge der „Näherungsbrüche"

$$\frac{p_0}{q_0} = \frac{a_0}{1}, \quad \frac{p_1}{q_1} = a_0 + \frac{1}{a_1},$$

$$\frac{p_2}{q_2} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2}}, \ldots, \quad \frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}, \ldots$$

bilden können. Alle diese Näherungsbrüche haben einen ganz bestimmten realen Sinn. Sie unterscheiden sich durch nichts von den in Kapitel III betrachteten Näherungsbrüchen, da sich auch dort $\frac{p_n}{q_n}$ als Funktion der Glieder $a_0, a_1, \ldots, a_n$ erwies. Ihre Form hing in keiner Weise davon ab, ob a_n letztes Glied ist oder ob auf a_n noch ein weiteres Glied a_{n+1} folgt. Der ganze Unterschied besteht darin, daß im Falle eines endlichen Kettenbruches die Folge der Näherungsbrüche endlich ist und das letzte Glied der Folge gleich dem „Wert" des gegebenen Kettenbruches ist, während im Falle des unendlichen Kettenbruches die Folge der Näherungsbrüche eine unendliche Folge von rationalen Zahlen ist und wir über seinen „Wert" bis jetzt noch nicht verfügt haben.

Insbesondere besitzen auch die hier betrachteten Näherungsbrüche die folgenden in Kapitel III festgestellten Eigenschaften:

---

[1]) Wir erinnern daran, daß wir den systematischen Brüchen, die wir zunächst rein formal als Folgen von Ziffern $a_1, a_2, \ldots, a_n, \ldots$ eingeführt hatten, dadurch einen Sinn beilegten, daß wir als Wert dieses Bruches die reelle Zahl $\sum_{n=1}^{\infty} \frac{a_n}{k^n}$ vereinbarten.

1.) Für jedes $n \geq 2$ ist

$$p_n = p_{n-1}a_n + p_{n-2},$$

$$q_n = q_{n-1}a_n + q_{n-2}.$$

2.) Für jedes $n \geq 0$ ist der Bruch $\frac{p_n}{q_n}$ reduziert und

$$\Delta_n = q_n p_{n+1} - p_n q_{n+1} = (-1)^n.$$

3.) Mit wachsender Ordnung nehmen die Näherungsbrüche gerader Ordnung zu und die Näherungsbrüche ungerader Ordnung ab. Dabei ist außerdem jeder Näherungsbruch gerader Ordnung kleiner als jeder Näherungsbruch ungerader Ordnung.

Diese Eigenschaften der Näherungsbrüche geben uns die Möglichkeit, den Wert eines unendlichen Kettenbruches zweckmäßig zu definieren und eine brauchbare Theorie der unendlichen Kettenbrüche aufzustellen.

Zunächst können wir auf Grund von Eigenschaft 3.) feststellen, daß die Folge der Näherungsbrüche gerader Ordnung eine monoton wachsende Folge

$$\frac{p_0}{q_0} < \frac{p_2}{q_2} < \cdots < \frac{p_{2k}}{q_{2k}} < \cdots$$

rationaler Zahlen bildet, die durch die Zahl $\frac{p_1}{q_1}$ nach oben beschränkt ist. Daher existiert der Grenzwert

$$\lim_{k \to \infty} \frac{p_{2k}}{q_{2k}} = \alpha.$$

Entsprechend stellen wir fest, daß die Folge der Näherungsbrüche ungerader Ordnung monoton fallend ist, daß also

$$\frac{p_1}{q_1} > \frac{p_3}{q_3} > \cdots > \frac{p_{2k+1}}{q_{2k+1}} > \cdots$$

ist, und daß diese Folge durch $\frac{p_2}{q_2}$ nach unten beschränkt ist, daß also auch

$$\lim_{k \to \infty} \frac{p_{2k+1}}{q_{2k+1}} = \beta$$

existiert. Dann ist für jedes $k \geq 0$

$$\frac{p_{2k}}{q_{2k}} < \alpha \leq \beta < \frac{p_{2k+1}}{q_{2k+1}}$$

und daher nach Eigenschaft 2)

$$0 \leq \beta - \alpha < \frac{p_{2k+1}}{q_{2k+1}} - \frac{p_{2k}}{q_{2k}} = \frac{1}{q_{2k}q_{2k+1}}.$$

Nach Eigenschaft 1.) ist nun aber $q_n > q_{n-1}$ $(n \geq 1)$, so daß q_n für $n \to \infty$ unbeschränkt wächst. Die rechte Seite der letzten Gleichung kann also für hinreichend großes k beliebig klein gemacht werden. Da $\beta - \alpha$ aber nicht von k abhängt, muß somit $\alpha = \beta$ sein.

Wir haben damit gezeigt, daß für jeden unendlichen Kettenbruch der Grenzwert

$$\lim_{n \to \infty} \frac{p_n}{q_n} = \alpha$$

existiert, den wir als Wert des Kettenbruches ansehen wollen. In diesem Sinne wollen wir auch sagen, daß der Kettenbruch (5) die Zahl α „darstellt".

Offensichtlich ist in diesem Fall für jedes $k \geq 0$

$$\frac{p_{2k}}{q_{2k}} < \alpha < \frac{p_{2k+1}}{q_{2k+1}}.$$

Hieraus folgt, daß für jedes $n \geq 0$

$$\left|\alpha - \frac{p_n}{q_n}\right| < \left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_n}{q_n}\right| = \frac{1}{q_n q_{n+1}}$$

ist.

Da $\frac{p_n}{q_n} = [a_0; a_1, a_2, \ldots, a_n]$ $(n \geq 0)$ ist, können wir den Kettenbruch (5) auch in der folgenden übersichtlichen Form schreiben, die klar den Zusammenhang zwischen dem Kettenbruch und seinen Näherungsbrüchen zum Ausdruck bringt:

$$[a_0; a_1, a_2, \ldots, a_n, \ldots] = \lim_{n \to \infty} [a_0; a_1, a_2, \ldots, a_n].$$

Wir fassen die bisherigen Ergebnisse zusammen:

Satz 7. *Jeder unendliche Kettenbruch* (5), *in dem a_0 eine beliebige ganze Zahl ist und $a_1, a_2, \ldots, a_n, \ldots$ natürliche Zahlen sind, stellt eine eindeutig bestimmte reelle Zahl dar, die durch die Grenzbeziehung*

$$\lim_{n \to \infty} \frac{p_n}{q_n} = \lim_{n \to \infty} [a_0; a_1, a_2, \ldots, a_n]$$

festgelegt ist. Dabei ist für jedes $k \geq 0$

$$\frac{p_{2k}}{q_{2k}} < \alpha < \frac{p_{2k+1}}{q_{2k+1}}$$

und für jedes $n \geq 0$

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}}.$$

Im vorangehenden setzten wir den Kettenbruch (5) als gegeben voraus und bestimmten die durch ihn dargestellte reelle Zahl. Wir wenden uns jetzt der umgekehrten Aufgabe zu. Vorgegeben sei uns also eine beliebige reelle Zahl α. Gefragt ist nach der Anzahl der Kettenbrüche, die diese Zahl darstellen (also insbesondere danach, ob es überhaupt einen derartigen Kettenbruch gibt) und wie sie bestimmt werden können.

Dazu setzen wir zunächst voraus, daß wir die Zahl α bereits in Form eines unendlichen Kettenbruches (5) dargestellt haben, und versuchen die Teilnenner a_n dieses Kettenbruches durch α auszudrücken. Da wir bereits wissen (Satz 7), daß jeder unendliche Kettenbruch eine gewisse reelle Zahl darstellt, können wir insbesondere die Zahlen

$$\alpha_1 = [a_1; a_2, a_3, \ldots, a_n, \ldots], \quad \alpha_2 = [a_2; a_3, a_4, \ldots, a_n, \ldots]$$

und allgemein

$$\alpha_k = [a_k; a_{k+1}, a_{k+2}, \ldots, a_{k+s}, \ldots] \quad (k \geq 1)$$

bilden. Es ist nun sehr leicht, eine Beziehung zwischen den Zahlen α_k und α_{k+1} herzustellen. Wenn wir nämlich mit $\frac{u_r}{v_r}$ den r-ten Näherungsbruch des Kettenbruches,

der die Zahl α_k darstellt, und entsprechend mit $\frac{u'_r}{v'_r}$ die Näherungsbrüche des Kettenbruches von α_{k+1} bezeichnen, so ist

$$\frac{u_r}{v_r} = [a_k; a_{k+1}, a_{k+2}, \ldots, a_{k+r}] = a_k + \frac{1}{[a_{k+1}; a_{k+2}, \cdots, a_{k+r}]} = a_k + \frac{1}{\frac{u'_{r-1}}{v'_{r-1}}}.$$

Nun ist nach Definition von α_k und α_{k+1}

$$\lim_{r\to\infty} \frac{u_r}{v_r} = \alpha_k, \quad \lim_{r\to\infty} \frac{u'_{r-1}}{v'_{r-1}} = \alpha_{k+1};$$

wenn wir also (bei festem k) in der vorangehenden Gleichung r unbeschränkt wachsen lassen, so wird ($\alpha_0 = \alpha$ gesetzt)

$$\alpha_k = a_k + \frac{1}{\alpha_{k+1}} \quad (k \geq 0). \tag{6}$$

Wenn speziell $k \geq 1$ ist, so ist $a_k \geq 1$, d. h., es ist $\alpha_k > 1$ für $k = 1, 2, \ldots$. Dann folgt aber aus (6) $\left(\text{wegen } \frac{1}{\alpha_{k+1}} < 1\right)$:

$$a_k < \alpha_k < a_k + 1 \quad (k \geq 0),$$

also speziell, daß

$$a_k = [\alpha_k] \quad (k \geq 0) \tag{7}$$

ist.

Wenn also die Zahl α in irgendeiner Weise in Form eines unendlichen Kettenbruches dargestellt ist, so berechnen sich die Teilnenner dieses Bruches einfach nach dem folgenden rekursiven Verfahren: 1.) $a_0 = [\alpha]$; 2.) Wenn bereits die Zahlen α_i und a_i für $i \leq k$ bestimmt sind, so berechnet sich α_{k+1} gemäß Relation (6) und danach a_{k+1} nach (7) als $[\alpha_{k+1}]$.

Wir gelangen auf diese Weise zu folgendem wichtigen Schluß: Wenn eine reelle Zahl α überhaupt in Form eines unendlichen Kettenbruches dargestellt werden kann, so sind die Teilnenner dieses Kettenbruches eindeutig durch die Zahl α bestimmt. Daraus folgt, daß die Darstellung einer reellen Zahl in Form eines unendlichen Kettenbruches, wenn eine solche existiert, auch eindeutig ist. Mit anderen Worten: *Es gibt keine zwei verschiedenen unendlichen Kettenbrüche, die ein und dieselbe reelle Zahl darstellen.*

Wir weisen darauf hin, daß im Falle einer rationalen Zahl $\alpha = \frac{a}{b}$ unser Prozeß der sukzessiven Konstruktion der Zahlen $\alpha_1, \alpha_2, \ldots$ nach Formel (6) sich durch nichts von dem Prozeß unterscheidet, mit dessen Hilfe wir in Kapitel III (S. 265) die Zahl $\frac{a}{b}$ in einen Kettenbruch entwickelt haben. Die Zahlen $\alpha_0, \alpha_1, \alpha_2, \ldots$ sind dort nichts anderes als die Zahlen $\frac{a}{b}, \frac{b}{r_1}, \frac{r_1}{r_2}, \ldots$. Nun haben wir dort aber festgestellt, daß für ein rationales $\alpha = \frac{a}{b}$ das Verfahren nach endlich vielen Schritten abbricht, d. h., daß eine der Zahlen $\alpha_k = \frac{r_{k-1}}{r_k}$ ganz und mithin $a_k = \alpha_k$ ist, also die Zahl α_{k+1} nicht mehr nach Formel (6) bestimmt werden kann. Andererseits haben wir jetzt gesehen, daß im Falle eines unendlichen Kettenbruches das durch

Formel (6) charakterisierte Verfahren niemals enden kann, sondern sich unbeschränkt fortsetzt. Aus dieser Gegenüberstellung können wir offensichtlich entnehmen, daß keine rationale Zahl durch einen unendlichen Kettenbruch dargestellt werden kann, daß also *alle Zahlen, die durch einen unendlichen Kettenbruch dargestellt werden, irrational sind.*

Zum Abschluß des hier angeschnittenen Problemkreises bleibt uns also lediglich noch zu zeigen, daß sich wirklich jede irrationale Zahl durch einen unendlichen Kettenbruch darstellen läßt (die Eindeutigkeit einer solchen Darstellung haben wir bereits bewiesen). Es sei also α eine beliebige irrationale Zahl. Dann bestimmen wir, ausgehend von α nach dem geschilderten Verfahren, also mit Hilfe der Relationen (6) und (7), die Zahlenfolgen α_k und a_k. Dieser Prozeß kann sicher nicht abbrechen, d. h., es kann sich kein α_k als ganze Zahl erweisen, da sonst α eine rationale Zahl wäre. Es ist also

$$0 < \frac{1}{\alpha_{k+1}} = \alpha_k - a_k = \alpha_k - [\alpha_k] < 1,$$

d. h. $\alpha_{k+1} > 1$ und mithin $a_{k+1} \geq 1$ für jedes $k \geq 0$.

Oben haben wir ferner gesehen, daß

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{\alpha_{n+1}}}}},$$

$$\frac{p_{n+1}}{q_{n+1}} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n + \cfrac{1}{a_{n+1}}}}}$$

ist. Hieraus folgt, daß wir die Zahl α erhalten, wenn wir in $\frac{p_{n+1}}{q_{n+1}}$ als Funktion von $a_0, a_1, \ldots, a_n, a_{n+1}$ das Argument a_{n+1} durch α_{n+1} (mit gleichem $a_0, a_1, \ldots, a_n$) ersetzen. Da aber

$$\frac{p_{n+1}}{q_{n+1}} = \frac{p_n a_{n+1} + p_{n-1}}{q_n a_{n+1} + q_{n-1}}$$

ist und $p_{n-1}, p_n, q_{n-1}, q_n$ nur von $a_0, a_1, \ldots, a_n$ abhängen, ist

$$\alpha = \frac{p_n \alpha_{n+1} + p_{n-1}}{q_n \alpha_{n+1} + q_{n-1}} \tag{8}$$

und daher

$$\left|\alpha - \frac{p_n}{q_n}\right| = \left|\frac{p_n \alpha_{n+1} + p_{n-1}}{q_n \alpha_{n+1} + q_{n-1}} - \frac{p_n}{q_n}\right| = \frac{|q_n p_{n-1} - p_n q_{n-1}|}{(q_n \alpha_{n+1} + q_{n-1}) q_n} = \frac{1}{(q_n \alpha_{n+1} + q_{n-1}) q_n} < \frac{1}{q_n^2},$$

woraus sich unmittelbar

$$\lim_{n \to \infty} \frac{p_n}{q_n} = \alpha$$

ergibt. Damit ist gezeigt, daß der konstruierte Kettenbruch tatsächlich die irrationale Zahl α darstellt.

Die zuletzt erhaltenen Ergebnisse können wir in folgendem Satz zusammenfassen:

Satz 8. *Jede reelle Zahl α läßt sich auf genau eine Weise in Form eines Kettenbruches darstellen. Dieser Kettenbruch ist endlich, wenn α rational ist, und unendlich, wenn α irrational ist.*

Wir weisen darauf hin, daß der von uns angegebene Beweis streng konstruktiv ist, d. h., daß er nicht nur die Existenz des eine gegebene Zahl darstellenden Kettenbruches zeigt, sondern auch ein Verfahren enthält, das in Form eines einfachen Algorithmus alle Teilnenner zu berechnen gestattet.

Wir sehen, daß die in den Sätzen 7 und 8 niedergelegten Grundgesetze über die Darstellung reeller Zahlen durch Kettenbrüche bemerkenswert einfach sind: Zu jeder reellen Zahl gibt es einen eindeutig bestimmten Bruch und zu jedem Bruch eine eindeutig bestimmte reelle Zahl; dabei entsprechen den rationalen Zahlen die endlichen und den irrationalen Zahlen die unendlichen Kettenbrüche. Wir haben bereits hervorgehoben, daß diese Einfachheit und Harmonie durch die Eigenschaften des formalen Apparates bedingt sind, vor allem durch seinen absoluten Charakter, der an kein spezielles Zahlensystem gebunden ist.

Unabhängig von der wünschenswerten Aufnahme der elementaren Theorie der Kettenbrüche in den Unterrichtsplan der Oberschulen gibt die Lehre von den Kettenbrüchen einen vortrefflichen Stoff für die Arbeit in mathematischen Zirkeln oder anderer außerplanmäßiger Form ab. Hierfür ist besonders der Inhalt des nachfolgenden Kapitels V zu empfehlen. Die Erfahrung hat gezeigt, daß die Theorie der Kettenbrüche und ihre Anwendungen auf die approximative Darstellung der reellen Zahlen stets das Interesse und die Aufmerksamkeit der Schüler zu fesseln vermögen. Das tritt immer dann deutlich in Erscheinung, wenn sich das Resultat in Form einfacher und effektiver Sätze formulieren läßt, die zudem noch leicht und elegant beweisbar sind. Dies ist aber bei fast allen Problemen der Fall, die mit der Darstellung der Zahlen als Kettenbrüche zusammenhängen.

Kapitel V

KETTENBRÜCHE UND DIOPHANTISCHE APPROXIMATIONEN

§ 13. Die Näherungsbrüche als beste Approximation

Will man in der Praxis eine gegebene reelle Zahl mit Hilfe eines rationalen Bruches angenähert ausdrücken, so bedient man sich meistens der Dezimalbrüche. Dazu wird die gegebene Zahl (sei sie nun rational oder irrational) in einen Dezimalbruch entwickelt. Ist dieser Dezimalbruch endlich und außerdem nicht allzu lang, so erübrigt sich eine Näherung, da der erhaltene Dezimalbruch dann eine bequeme und überdies absolut genaue Darstellung der gegebenen Zahl liefert. Wenn hingegen der erhaltene Dezimalbruch unendlich oder, obwohl endlich, sehr lang ist und sich damit für praktische Rechnungen nicht eignet, wird er in der Praxis nach den bekannten Regeln „gerundet", d. h., man beschränkt sich in den Rechnungen auf eine hinreichend große Anzahl von Dezimalstellen und läßt die übrigen unberücksichtigt. Eine solche „Rundung" ist nichts anderes als eine Annäherung der gegebenen Zahl durch eine rationale Zahl mit relativ kleinem Zähler und Nenner, wobei man einen gewissen Fehler in Kauf nimmt.

Welcher Fehler wird aber dabei zugelassen? Unser Ziel besteht doch darin, bei einer solchen angenäherten Darstellung eine möglichst große Genauigkeit (d. h. einen möglichst kleinen Fehler) bei kleinstmöglichem Zähler und Nenner des annähernden Bruches zu erreichen. Daher ist es ganz natürlich, daß man gewöhnlich die Größe des Fehlers mit der Größenordnung des Nenners eines solchen Bruches vergleicht (man könnte ebensogut den Zähler nehmen, denn offensichtlich haben die Zähler von annähernden Brüchen mit Nennern gleicher Größenordnung ebenfalls gleiche Größenordnung; den Nenner wählt man deshalb, weil er stets als positiv vorausgesetzt ist, während der Zähler auch negativ sein kann). Wenn wir uns nun beim Runden eines Dezimalbruches auf n Stellen hinter dem Komma beschränken, so bedeutet das, daß die gegebene Zahl durch einen Bruch mit dem Nenner $q = 10^n$ angenähert wird. Bekanntlich übersteigt der Fehler dabei nicht die halbe Einheit der zuletzt berücksichtigten Ordnung, d. h., er ist nicht größer als

$$\frac{1}{2 \cdot 10^n} = \frac{1}{2q}.$$

Dieses Ergebnis wollen wir jetzt mit dem Resultat vergleichen, das uns die Kettenbrüche liefern. Wenn hier $\frac{p}{q}$ einer der Näherungsbrüche der gegebenen Zahl α ist, so ist — wie wir bereits wissen (siehe Kapitel IV, Satz 7) —

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{qq'}, \tag{1}$$

wobei q' der Nenner des auf $\frac{p}{q}$ folgenden Näherungsbruches ist. Wegen $q' > q$ ist also erst recht

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

Wir sehen also, daß bei der Annäherung der Zahl α durch den Näherungsbruch $\frac{p}{q}$ der Fehler nicht größer als $\frac{1}{q^2}$ ist. Da nun bei beliebigem Nenner q die Zahl $\frac{1}{q^2}$ erheblich kleiner ist als $\frac{1}{2q}$, finden wir, daß die Approximation durch Näherungsbrüche bei gleicher Größenordnung des Nenners einen Fehler liefert, der im allgemeinen erheblich kleiner als der Fehler bei Annäherung durch einen Dezimalbruch (oder einen anderen systematischen Bruch) ist.

Das folgende Beispiel möge dies näher erläutern: Gesucht sei ein Näherungswert der Zahl π in Form eines rationalen Bruches, dessen Nenner die Zahl 100 nicht übersteigt. Bekanntlich liefert uns die Dezimalbruchentwicklung die Lösung in Form des Bruches

$$3{,}14 = \frac{314}{100}$$

mit einem Fehler von etwa 0,0016 (einem Fehler, der sicher größer als 0,0015 ist — *Anm. d. wissenschaftl. Red.*). Wenn wir andererseits π in Form eines Kettenbruches darstellen, so finden wir

$$\pi = [3; 7, 15, \ldots]$$

und mithin

$$\frac{p_0}{q_0} = \frac{3}{1}, \quad \frac{p_1}{q_1} = \frac{22}{7}, \quad \frac{p_2}{q_2} = \frac{333}{106}.$$

Da der Nenner des Näherungsbruches nicht größer als 100 sein soll, müssen wir uns mit dem Bruch $\frac{22}{7}$ begnügen, für den aber immerhin nach (1)

$$\left|\pi - \frac{22}{7}\right| < \frac{1}{7 \cdot 106} < 0{,}0014$$

gilt. Wir stellen also fest, daß die Annäherung durch den Näherungsbruch, trotz des erheblich kleineren Nenners (7 statt 100), wesentlich besser ist als die Annäherung durch den Dezimalbruch. Dieses Beispiel zeigt auch ganz deutlich, wodurch das bedingt ist: Wenn wir eine gegebene Zahl durch einen Dezimalbruch annähern, so kommen nur Nenner der Form 10^n in Frage, d. h., wir können bei der Auswahl des Nenners in keiner Weise der arithmetischen Natur der anzunähernden Zahl Rechnung tragen. Hätten wir an Stelle der Zahl π in unserem Beispiel eine andere Zahl genommen, so wäre bei der Annäherung durch einen Dezimalbruch der Nenner derselbe, nämlich 100 geblieben, während wir, ausgehend von ihrer Kettenbruchdarstellung, bei ihrer Approximation durch Näherungsbrüche wahrscheinlich einen ganz anderen Nenner erhalten hätten, der in irgendeiner Weise die arithmetische Natur der gegebenen Zahl berücksichtigt. Die Nenner der Näherungsbrüche verschiedener Zahlen sind im allgemeinen verschieden, und diese größere Freiheit in der Wahl des Nenners hat eben die erheblich besseren Resultate zur Folge. Wir müssen also für die Bequemlichkeit des Rechnens in systematischen Brüchen auf der anderen Seite ziemlich teuer bezahlen.

Wenn auch der Vergleich zwischen der Qualität der Approximation durch systematische Brüche und der durch Kettenbrüche entschieden zugunsten der Kettenbrüche ausfällt, so scheint diese Einschätzung doch nur relativ zu sein, und es bleibt die Frage nach der Möglichkeit noch kräftigerer Methoden davon gänzlich unberührt. Daher wollen wir jetzt die absolute Güte der Approximation durch Näherungsbrüche unabhängig von anderen speziellen Approximationsmethoden untersuchen. Zunächst überzeugen wir uns davon, daß die Näherungsbrüche einer gegebenen Zahl für sie die beste Annäherung im folgenden präzisen Sinne darstellen:

Satz 1. *Wenn $\frac{p_n}{q_n}$ ein Näherungsbruch der Zahl α ist, so ist jeder Bruch mit einem Nenner $q < q_n$ weiter von α entfernt als $\frac{p_n}{q_n}$.*

Jeder Näherungsbruch liefert uns also eine Approximation, die besser ist als die Annäherung durch jeden Bruch mit kleinerem Nenner. Eine weitere Verbesserung kann demnach nur um den Preis einer Vergrößerung des Nenners (und damit auch des Zählers) des annähernden Bruches erreicht werden.

Zum Beweis dieses Satzes überzeugen wir uns zunächst davon, daß jeder Näherungsbruch $\frac{p_n}{q_n}$ näher an α liegt als der vorangehende Näherungsbruch $\frac{p_{n-1}}{q_{n-1}}$. In der Tat: Nach Satz 7 aus Kapitel IV (S. 282) liegt der Bruch $\frac{p_{n+1}}{q_{n+1}}$ zwischen $\frac{p_{n-1}}{q_{n-1}}$ und α. Daher ist

$$\left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| > \left|\frac{p_{n+1}}{q_{n+1}} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{|q_{n-1}(p_n a_{n+1} + p_{n-1}) - p_{n-1}(q_n a_{n+1} + q_{n-1})|}{q_{n-1} q_{n+1}}$$

$$= \frac{a_{n+1}}{q_{n-1} q_{n+1}} > \frac{1}{q_n q_{n+1}} > \left|\alpha - \frac{p_n}{q_n}\right|, \qquad (2)$$

was zu beweisen war.

Wir zeigen jetzt, daß für jedes q mit $0 < q < q_n$

$$\left|\alpha - \frac{p}{q}\right| > \left|\alpha - \frac{p_n}{q_n}\right| \tag{3}$$

ist. Für den Fall, daß $\frac{p}{q} = \frac{p_{n-1}}{q_{n-1}}$ ist, fällt die Ungleichung (3) mit der bereits bewiesenen Ungleichung (2) zusammen. Wir können also annehmen, daß $\frac{p}{q} \neq \frac{p_{n-1}}{q_{n-1}}$ ist.

Dann ist

$$\left|\frac{p}{q} - \frac{p_{n-1}}{q_{n-1}}\right| = \frac{|pq_{n-1} - qp_{n-1}|}{qq_{n-1}} \geq \frac{1}{qq_{n-1}} > \frac{1}{q_n q_{n-1}} > \left|\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}}\right|,$$

d. h., $\frac{p}{q}$ ist von $\frac{p_{n-1}}{q_{n-1}}$ weiter entfernt als $\frac{p_n}{q_n}$, kann also nicht im Intervall $\left(\frac{p_{n-1}}{q_{n-1}}, \frac{p_n}{q_n}\right)$ liegen, in dem aber — wie wir wissen — die Zahl α liegt. $\frac{p}{q}$ liegt also entweder auf derselben Seite von α wie $\frac{p_n}{q_n}$ und weiter von α entfernt als $\frac{p_n}{q_n}$, so daß also trivialerweise

$$\left|\alpha - \frac{p}{q}\right| > \left|\alpha - \frac{p_n}{q_n}\right|$$

ist, oder $\frac{p}{q}$ liegt auf derselben Seite von α wie $\frac{p_{n-1}}{q_{n-1}}$ und weiter entfernt von α als $\frac{p_{n-1}}{q_{n-1}}$, so daß wegen (2)

$$\left|\alpha - \frac{p}{q}\right| > \left|\alpha - \frac{p_{n-1}}{q_{n-1}}\right| > \left|\alpha - \frac{p_n}{q_n}\right|$$

ist, was zu beweisen war.

Wir haben bisher die Güte der Approximation einer Zahl α durch einen Bruch $\frac{p}{q}$ mittels der Größe der Differenz $\left|\alpha - \frac{p}{q}\right|$ gemessen. Das ist ohne Zweifel die natürlichste Art. Jedoch können wir zu diesem Zweck auch die Größe $|q\alpha - p|$ benutzen, die sich von der vorhergehenden nur durch den Faktor q unterscheidet. In diesem Sinne können wir z. B. das Näherungsgesetz

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n q_{n+1}}$$

für den Näherungsbruch $\frac{p_n}{q_n}$ der Zahl α gleichwertig durch die Ungleichung

$$|q_n \alpha - p_n| < \frac{1}{q_{n+1}}$$

ersetzen. So trivial diese Umformung auch immer erscheinen mag, die geschichtliche Entwicklung der behandelten Probleme zeigt doch, daß die systematische Verwendung der Größe $|q\alpha - p|$ zur Abschätzung der Approximation der Zahl α durch den Bruch $\frac{p}{q}$ erhebliche Vorteile gegenüber der Benutzung der Größe $\left|\alpha - \frac{p}{q}\right|$ bietet. Einerseits ergeben sich nämlich auf diesem Wege häufig Näherungsgesetze einer viel geschlosseneren Form (wir werden sofort ein Beispiel hierfür kennenlernen), und andererseits — und das ist besonders wichtig — eröffnet dieser Weg

außerordentlich reiche weitere Möglichkeiten. Er führt ganz natürlich zur Bildung neuer Begriffe und Aufgabenstellungen, die das betrachtete Wissenschaftsgebiet in bedeutendem Maße bereichern. Wir werden auch hierfür im folgenden eine Reihe von Beispielen kennenlernen.

Zunächst untersuchen wir von diesem neuen Standpunkt aus die Frage, die im vorangehenden zu Satz 1 führte. Dazu sei $\frac{p_n}{q_n}$ ein Näherungsbruch der Zahl α und $0 < q < q_n$. Kann man auch hier behaupten, daß für jedes p

$$|q\alpha - p| > |q_n\alpha - p_n| \tag{4}$$

ist? Zur Beantwortung dieser Frage betrachten wir das Gleichungssystem

$$\left.\begin{aligned} p_n x + p_{n+1} y &= p \\ q_n x + q_{n+1} y &= q \end{aligned}\right\} \tag{5}$$

in den Unbekannten x und y. Wegen $p_n q_{n+1} - q_n p_{n+1} = \pm 1$ besitzt dieses System genau ein Lösungspaar (x, y) von ganzen Zahlen x und y. Wir erweitern unsere Aufgabenstellung nun gleich dahingehend, daß wir zulassen, daß q eine beliebige natürliche Zahl ist, die kleiner als q_{n+1}, aber verschieden von q_n ist. Ohne daß wir die Allgemeinheit unserer Frage im geringsten einschränken, können wir offenbar voraussetzen, daß der Bruch $\frac{p}{q}$ reduziert ist (also p und q teilerfremd sind). Dann ergibt sich unmittelbar, daß die als Lösung des Systems (5) bestimmten ganzen Zahlen x und y verschiedene Vorzeichen haben müssen, also $xy < 0$ ist. In der Tat: Wenn $xy > 0$ oder $x = 0$ wäre, so wäre auf Grund der zweiten Gleichung von (5) $q \geqq q_{n+1}$, was nicht der Fall sein sollte. Wäre $y = 0$, so wäre (da dann wegen der Reduziertheit des Bruches $\frac{p}{q}$ jedenfalls $x = 1$ sein müßte) $q = q_n$, was gleichfalls ausgeschlossen ist. Es ist also wirklich $xy < 0$.

Nun folgt aus (5), daß

$$q\alpha - p = x(q_n\alpha - p_n) + y(q_{n+1}\alpha - p_{n+1})$$

ist. Da nun nach Satz 7 aus Kapitel IV die Zahlen $q_n\alpha - p_n$ und $q_{n+1}\alpha - p_{n+1}$ verschiedene Vorzeichen besitzen und auch $xy < 0$ ist, haben beide Summanden auf der rechten Seite dieser Gleichung gleiches Vorzeichen, so daß also sogar

$$|q\alpha - p| = |x| \cdot |q_n\alpha - p_n| + |y| \cdot |q_{n+1}\alpha - p_{n+1}|$$

ist. Hieraus folgt, daß die Ungleichung (4) nur in einem Fall nicht gilt, wenn nämlich $q_{n+1}\alpha - p_{n+1} = 0$ und $|x| = 1$ ist, in dem dann

$$|q\alpha - p| = |q_n\alpha - p_n|$$

wird. Wegen (5) tritt dieser Fall jedoch nur dann ein, wenn

$$x = -1, \quad y = 1, \quad q = q_{n+1} - q_n, \quad p = p_{n+1} - p_n$$

ist.

Wir finden also, daß für alle q mit $0 < q < q_{n+1}$

$$|q\alpha - p| \geqq |q_n\alpha - p_n|$$

gilt.[1]) Hierbei tritt Gleichheit genau dann ein, wenn $p = p_n$, $q = q_n$ oder wenn $\alpha = \frac{p_{n+1}}{q_{n+1}}$, $q = q_{n+1} - q_n = (a_{n+1} - 1)\, q_n + q_{n-1}$ ist. In diesem letzten Fall ist a_{n+1} letzter Teilnenner des endlichen Kettenbruches

$$[a_0; a_1, a_2, \ldots, a_{n+1}] = \alpha$$

und mithin nach Definition des Kettenbruches (Kapitel III, S. 267) größer als Eins, so daß also $q \geq q_n + q_{n-1} > q_n$ ist.

Zusammenfassend können wir feststellen, daß

$$|q\alpha - p| > |q_n\alpha - p_n| \quad (0 < q < q_n), \tag{6}$$

$$|q\alpha - p| \geq |q_n\alpha - p_n| \quad (0 < q < q_{n+1}) \tag{7}$$

gilt.

Im folgenden wollen wir einen echten Bruch $\frac{l}{k}$ eine *beste Näherung zweiter Art der Zahl* α nennen, wenn für jedes ganze k' mit $0 < k' < k$ und jedes ganze l'

$$|k'\alpha - l'| > |k\alpha - l|$$

gilt. Die Ungleichung (6) zeigt dann, daß jeder Näherungsbruch der Zahl α eine beste Näherung zweiter Art ist. Die Bezeichnung „zweiter Art" soll darauf hinweisen, daß es sich bei der Abschätzung der Güte der Approximation um die Abschätzung nach der Differenz $|k\alpha - l|$ handelt im Gegensatz zu den Näherungen „erster Art", von denen Satz 1 handelt und die nach der Differenz $\left|\alpha - \frac{l}{k}\right|$ abgeschätzt werden.

Nach dem Gesagten stellen die Näherungsbrüche „beste Näherungen" sowohl erster als auch zweiter Art dar. Wesentlich ist jedoch, daß bezüglich der Näherung zweiter Art die Näherungsbrüche auch die einzigen derartigen Brüche sind: *Jede beste Näherung zweiter Art ist ein Näherungsbruch.*

In der Tat: Wenn der Bruch $\frac{p}{q}$ kein Näherungsbruch der Zahl α ist, so gibt es eine natürliche Zahl n derart, daß $q_n < q < q_{n+1}$ ist.[2]) Gemäß (7) ist dann aber

$$|q_n\alpha - p_n| \leq |q\alpha - p| \quad (q_n < q)$$

und daher $\frac{p}{q}$ keine beste Näherung zweiter Art der Zahl α. Zusammenfassend können wir folgenden Satz aufstellen:

---

[1]) Der im vorangehenden ausgeschlossene Fall $q = q_n$ ordnet sich hier offenbar trivial ein, da stets $\left|\alpha - \frac{p}{q_n}\right| \geqq \left|\alpha - \frac{p_n}{q_n}\right|$ ist, wobei Gleichheit dann und nur dann eintritt, wenn $p = p_n$ ist. — *Anm. d. wissenschaftl. Red.*

[2]) Der Fall, daß $q = q_n$, aber $p \neq p_n$ ist, kann außer Betracht bleiben, da man leicht einsieht, daß ein solcher Bruch $\frac{p}{q}$ keine beste Näherung zweiter Art sein kann, weil z. B. für $q = q_n > 1$ aus $p \neq p_n$ folgt, daß

$$|q\alpha - p| = |(q_n\alpha - p_n) + (p_n - p)| \geq 1 - |q_n\alpha - p_n| > 1 - \frac{1}{q_{n+1}} > \frac{1}{q_n} > |q_{n-1}\alpha - p_{n-1}|$$

ist.

Anm. d. wissenschaftl. Red.: Auch der Fall, daß alle q_n kleiner als q sind, erledigt sich entsprechend, da dann der α darstellende Kettenbruch endlich, also $\alpha = \frac{p_k}{q_k}$ und mithin $0 = |q_k\alpha - p_k| \leq |q\alpha - p|$ ist.

Satz 2. *Die Näherungsbrüche einer gegebenen Zahl α und nur diese sind beste Näherungen zweiter Art der Zahl α.*

Man kann zeigen, daß Satz 1 keine entsprechende Verschärfung zuläßt: Neben den Näherungsbrüchen einer Zahl können durchaus noch andere Brüche als beste Näherungen erster Art auftreten. Der hier betrachtete Fall liefert uns ein Beispiel für ein Problem, bei dem die Verwendung der Differenzen vom Typ $|q\alpha - p|$ zu einfacheren und gesetzmäßigeren Resultaten führt als die Benutzung der Differenzen vom Typ $\left|\alpha - \frac{p}{q}\right|$.

Durch Satz 2 werden wir nun aber an eine ganz neue Fragestellung herangeführt: Vorgegeben sei eine reelle Zahl α und ein gewisser rationaler Bruch $\frac{p}{q}$. Wie kann man feststellen, ob $\frac{p}{q}$ ein Näherungsbruch der Zahl α ist? Satz 2 gibt hierauf eine sehr einfache Antwort: Man prüfe nach, ob der Bruch $\frac{p}{q}$ eine beste Näherung zweiter Art der Zahl α ist. Diese Antwort kann uns jedoch nur wenig befriedigen, verlangt sie doch von uns, daß wir den gegebenen Bruch mit allen Brüchen eines kleineren Nenners vergleichen. Wir suchen hingegen nach einem Kriterium, das uns die aufgeworfene Frage unter alleiniger Benutzung der Zahl α und des Bruches $\frac{p}{q}$ beantwortet.

Ein vortreffliches Beispiel einer einfachen Bedingung dieser Art gibt uns der bereits auf LEGENDRE zurückgehende

Satz 3. *Wenn $q > 0$ und*

$$|q\alpha - p| \leq \frac{1}{2q}$$

ist, so ist $\frac{p}{q}$ ein Näherungsbruch der Zahl α.

Unter Benutzung des Satzes 2 können wir stattdessen auch behaupten: *Wenn $q > 0$ und*

$$|q\alpha - p| \leq \frac{1}{2q}$$

ist, so ist $\frac{p}{q}$ eine beste Näherung zweiter Art der Zahl α.

Beweis. Es sei $\frac{p'}{q'}$ ein beliebiger rationaler Bruch mit $q' > 0$ und

$$|q'\alpha - p'| \leq |q\alpha - p| \leq \frac{1}{2q}. \tag{a}$$

Wir haben zu zeigen, daß $q' \geq q$ ist.

Zunächst folgt aus Ungleichung (a), daß

$$\left|\alpha - \frac{p'}{q'}\right| \leq \frac{1}{2qq'}, \quad \left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2q^2}$$

ist. Da wir ohne Einschränkung voraussetzen können, daß $q \neq q'$ (und mithin $\frac{p}{q} \neq \frac{p'}{q'}$, $|p'q - q'p| \geq 1$) ist, gilt

$$\frac{1}{qq'} \leq \left|\frac{p'}{q'} - \frac{p}{q}\right| \leq \left|\frac{p'}{q'} - \alpha\right| + \left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2qq'} + \frac{1}{2q^2}$$

und damit

$$\frac{1}{q'} \leqq \frac{1}{2q'} + \frac{1}{2q}, \quad \frac{1}{2q'} \leqq \frac{1}{2q},$$

also $q' \geqq q$, was zu beweisen war.

Satz 3 gibt uns eine einfache Bedingung dafür, daß ein Bruch $\frac{p}{q}$ ein Näherungsbruch der Zahl α ist. Diese Bedingung ist jedoch nicht charakteristisch für die Näherungsbrüche: Eine Zahl α besitzt im allgemeinen unendlich viele Näherungsbrüche, die diese Bedingung nicht erfüllen. Man kann sich sogar davon überzeugen, daß es kein allgemeines Kriterium der Form

$$|q\alpha - p| < \frac{a}{q}$$

mit konstantem a gibt.

Dennoch kann man für die Näherungsbrüche der Zahl α eine hinreichend einfache charakteristische (d. h. notwendige und hinreichende) Bedingung angeben. Dazu stellen wir den gegebenen Bruch $\frac{p}{q}$ in Form eines Kettenbruches

$$\frac{p}{q} = [a_0; a_1, a_2, \ldots, a_n]$$

dar und bezeichnen mit $\frac{p'}{q'}$ den letzten Näherungsbruch dieser Darstellung, setzen also

$$\frac{p'}{q'} = [a_0; a_1, a_2, \ldots, a_{n-1}].$$

Dann gilt der folgende, gleichfalls auf LEGENDRE zurückgehende

Satz 4. *Der Bruch $\frac{p}{q}$ ist dann und nur dann ein Näherungsbruch der Zahl α, wenn*

$$|q\alpha - p| < \frac{1}{q+q'} \tag{8}$$

ist.

Bemerkung. Da stets $q + q' < 2q$ ist, ergibt sich übrigens Satz 3 als unmittelbare Folgerung aus Satz 4.

Beweis. 1.) Es sei $\frac{p}{q}$ ein Näherungsbruch der Zahl α und $\frac{p'}{q'}$ der vorangehende Näherungsbruch. Wir haben dann bereits des öfteren festgestellt, daß

$$|q\alpha - p| < \frac{1}{q''}$$

gilt, wobei q'' der Nenner des auf $\frac{p}{q}$ folgenden Näherungsbruches ist.[1]) Gemäß dem Bildungsgesetz der Näherungsbrüche kann dann q'' in der Form $aq + q'$ dargestellt werden, wobei $a \geqq 1$ ist, so daß also speziell $q'' \geqq q + q'$ und damit

$$|q\alpha - p| < \frac{1}{q+q'}$$

ist. Hiermit ist gezeigt, daß für die Näherungsbrüche die Bedingung (8) notwendig ist.

[1]) Hierbei müssen wir allerdings voraussetzen, daß $q\alpha \neq p$ ist. Das ist aber keine Einschränkung der Allgemeinheit, da im Falle $q\alpha = p$ die Ungleichung (8) trivial erfüllt ist.

2.) Es möge nun umgekehrt die Bedingung (8) erfüllt sein. Dann können wir stets (und zwar auf genau eine Weise) eine reelle Zahl β bestimmen, für die

$$\alpha = \frac{p\beta + p'}{q\beta + q'} \tag{9}$$

gilt (dazu braucht man nur die Gleichung (9) nach β aufzulösen). Es sei

$$\beta = [a_{n+1}; a_{n+2}, \ldots]$$

die Kettenbruchdarstellung der Zahl β. Wir zeigen zunächst, daß $a_{n+1} \geq 1$ ist. Dazu genügt es offensichtlich, zu zeigen, daß $\beta \geq 1$ ist: Gemäß Ungleichung (8) ist

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{p\beta + p'}{q\beta + q'} - \frac{p}{q}\right| = \frac{1}{q(q\beta + q')} < \frac{1}{q(q + q')}$$

und mithin

$$q\beta + q' > q + q',$$

also sogar $\beta > 1$.

Nun folgt aus

$$\frac{p}{q} = [a_0; a_1, \ldots, a_n]$$

[vgl. Kapitel IV, Gleichung (8), S. 284], daß

$$\frac{p\beta + p'}{q\beta + q'} = [a_0; a_1, \ldots, a_n, \beta]^{1})$$

und daher

$$\alpha = [a_0; a_1, \ldots, a_n, \beta]$$

ist. Wenn also

$$\frac{p'_s}{q'_s} = [a_{n+1}; a_{n+2}, \ldots, a_{n+s}]$$

die Näherungsbrüche der Zahl β sind, so ist

$$\beta = \lim_{s \to \infty} \frac{p'_s}{q'_s}$$

und daher

$$\alpha = \lim_{s \to \infty} \left[a_0; a_1, \ldots, a_n, \frac{p'_s}{q'_s}\right] = \lim_{s \to \infty} [a_0; a_1, \ldots, a_n, a_{n+1}, \ldots, a_{n+s}].$$

Das besagt aber nichts anderes, als daß α durch den Kettenbruch

$$[a_0; a_1, \ldots, a_n, a_{n+1}, \ldots]$$

dargestellt wird, der zudem dann der einzige Bruch ist, der die Zahl α darstellt. Insbesondere ist also

$$\frac{p}{q} = [a_0; a_1, \ldots, a_n]$$

ein Näherungsbruch der Zahl α, womit gezeigt wurde, daß die LEGENDREsche Bedingung auch hinreichend ist.

---

[1]) Die Zahl β ist hier im allgemeinen nicht ganz. Das darf den Leser aber nicht stutzig machen, da die von uns gegebene Deutung des Symbols $[a_0, a_1, a_2, \ldots, a_n, \beta]$ für beliebiges $\beta \neq 0$ einen Sinn gibt.

Wir haben schon darauf hingewiesen, daß die in Satz 3 angegebene Bedingung nicht notwendig ist (oder, was dasselbe ist, daß Satz 3 keine Umkehrung zuläßt). Aber wenn auch nicht jeder Näherungsbruch einer Zahl die dort angegebene Eigenschaft zu besitzen braucht, so besitzt sie doch stets wenigstens einer von zwei aufeinanderfolgenden Näherungsbrüchen. Das ist die Aussage des folgenden, von VAHLEN aufgestellten Satzes:

Satz 5. *Wenn* $\frac{p}{q}$ *und* $\frac{p'}{q'}$ *zwei aufeinanderfolgende Näherungsbrüche der Zahl* α *sind, so gilt wenigstens eine der Ungleichungen*

$$\left|\alpha-\frac{p}{q}\right|<\frac{1}{2q^2},\quad \left|\alpha-\frac{p'}{q'}\right|<\frac{1}{2q'^2}.$$

Beweis. Da die Brüche $\frac{p}{q}$ und $\frac{p'}{q'}$ auf verschiedenen Seiten der Zahl α liegen, würde aus

$$\left|\alpha-\frac{p}{q}\right|\geq\frac{1}{2q^2},\quad \left|\alpha-\frac{p'}{q'}\right|\geq\frac{1}{2q'^2}$$

folgen, daß

$$\left|\frac{p'}{q'}-\frac{p}{q}\right|=\frac{1}{qq'}\geq\frac{1}{2}\left(\frac{1}{q^2}+\frac{1}{q'^2}\right)$$

und mithin

$$\frac{1}{q^2}+\frac{1}{q'^2}-\frac{2}{qq'}=\left(\frac{1}{q}-\frac{1}{q'}\right)^2\leq 0$$

gilt, was offensichtlich falsch ist. Damit ist Satz 5 bereits bewiesen.

Dadurch ist gezeigt, daß jede irrationale Zahl α unendlich viele Näherungsbrüche $\frac{p}{q}$ besitzt, für die

$$\left|\alpha-\frac{p}{q}\right|<\frac{1}{2q^2}$$

ist. Hier erhebt sich nun sofort die Frage nach der Möglichkeit einer weiteren Verschärfung dieses Resultats. Kann man etwa die Konstante $\frac{1}{2}$ auf der rechten Seite dieser Ungleichung durch eine kleinere Konstante c ersetzen, für die noch dasselbe gilt, d. h., gibt es eine kleinere Konstante c derart, daß es zu jeder irrationalen Zahl unendlich viele Näherungsbrüche $\frac{p}{q}$ mit

$$\left|\alpha-\frac{p}{q}\right|<\frac{c}{q^2}$$

gibt? Ferner: Wenn es solche Zahlen $c<\frac{1}{2}$ gibt, wie sieht die kleinste derartige Zahl aus? Die Lösung dieses Problems bedeutet die Bestimmung des im bekannten Sinne besten universellen (d. h. für jede irrationale Zahl α geltenden) Näherungsgesetzes für die Approximation reeller Zahlen durch rationale Brüche.

HURWITZ konnte zeigen, daß der kleinstmögliche Wert der Konstanten c die Zahl $\frac{1}{\sqrt{5}}$ ist. Kurz darauf bewies BOREL, daß von drei aufeinanderfolgenden Näherungsbrüchen einer Zahl α wenigstens einer die genannte Näherung ergibt. Es gilt also

Satz 6. *Wenn* $\frac{p_n}{q_n}, \frac{p_{n+1}}{q_{n+1}}, \frac{p_{n+2}}{q_{n+2}}$ *drei aufeinanderfolgende Näherungsbrüche der Zahl* α *sind, so gilt wenigstens eine der Ungleichungen*

$$\left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{\sqrt{5}\, q_n^2}, \quad \left|\alpha - \frac{p_{n+1}}{q_{n+1}}\right| < \frac{1}{\sqrt{5}\, q_{n+1}^2}, \quad \left|\alpha - \frac{p_{n+2}}{q_{n+2}}\right| < \frac{1}{\sqrt{5}\, q_{n+2}^2}.$$

Zu jeder Zahl $c < \frac{1}{\sqrt{5}}$ *gibt es eine irrationale Zahl* α *derart, daß die Ungleichung*

$$\left|\alpha - \frac{p}{q}\right| < \frac{c}{q^2}$$

nur für endlich viele rationale Brüche $\frac{p}{q}$ *erfüllt ist.*

Beweis. Zum Beweis der ersten Behauptung des Satzes 6 benötigen wir das folgende

Lemma. *Von den beiden Quotienten* $\frac{q_{n+1}}{q_n}$ und $\frac{q_{n+2}}{q_{n+1}}$ *ist mindestens einer größer als die Zahl* $\gamma = \frac{\sqrt{5}+1}{2}$.

In der Tat: Aus $\frac{q_{n+1}}{q_n} < \gamma$, $\frac{q_{n+2}}{q_{n+1}} < \gamma$[1]) würde wegen $1 + \frac{1}{\gamma} = \gamma$ folgen, daß

$$\gamma > \frac{q_{n+2}}{q_{n+1}} = a_{n+2} + \frac{q_n}{q_{n+1}} > 1 + \frac{1}{\gamma} = \gamma$$

ist, was nicht gilt. Damit ist das Lemma bewiesen.

Für den nun folgenden Beweis setzen wir für jedes $i \geq 1$

$$\psi_i = \alpha_{i+1} + \frac{q_{i-1}}{q_i},$$

wobei

$$\alpha_{i+1} = [a_{i+1}; a_{i+2}, a_{i+3}, \ldots]$$

dieselbe Bedeutung wie in Kapitel IV (S. 283) hat. Dort (S. 284) haben wir gezeigt, daß

$$\left|\alpha - \frac{p_i}{q_i}\right| = \frac{1}{q_i(q_i\alpha_{i+1} + q_{i-1})} = \frac{1}{\psi_i q_i^2}$$

ist. Wir haben also unsere Behauptung bewiesen, wenn wir gezeigt haben, daß wenigstens eine der Zahlen $\psi_n, \psi_{n+1}, \psi_{n+2}$ größer als $\sqrt{5}$ ist.

Da nun $\alpha_{i+1} = a_{i+1} + \frac{1}{\alpha_{i+2}}$ und $\frac{q_{i+1}}{q_i} = a_{i+1} + \frac{q_{i-1}}{q_i}$ ist, so gilt

$$\psi_i = \alpha_{i+1} + \frac{q_{i-1}}{q_i} = \frac{1}{\alpha_{i+2}} + \frac{q_{i+1}}{q_i}. \tag{10}$$

Dabei übertrifft nach dem oben bereits bewiesenen Lemma wenigstens einer der Quotienten $\frac{q_{n+1}}{q_n}$ oder $\frac{q_{n+2}}{q_{n+1}}$ den Wert γ. Zunächst möge $\frac{q_{n+1}}{q_n} > \gamma$ sein. Da die Funktion $u + \frac{1}{u}$ für $u > 1$ monoton wächst, ist für den Fall, daß $\alpha_{n+2} \geq \frac{q_{n+1}}{q_n}$ ist,

$$\psi_{n+1} = \alpha_{n+2} + \frac{q_n}{q_{n+1}} \geq \alpha_{n+2} + \frac{1}{\alpha_{n+2}} > \gamma + \frac{1}{\gamma} = \sqrt{5}$$

[1]) Wegen der Irrationalität der Zahl γ ist hier das Gleichheitszeichen selbstverständlich unmöglich.

und für den Fall, daß $\alpha_{n+2} < \frac{q_{n+1}}{q_n}$ ist, nach (10)

$$\psi_n = \frac{1}{\alpha_{n+2}} + \frac{q_{n+1}}{q_n} > \frac{1}{\frac{q_{n+1}}{q_n}} + \frac{q_{n+1}}{q_n} > \gamma + \frac{1}{\gamma} = \sqrt{5}.$$

Wenn hingegen $\frac{q_{n+1}}{q_n} < \gamma$ ist, so muß $\frac{q_{n+2}}{q_{n+1}} > \gamma$ sein, und dieselbe Kette von Überlegungen mit jeweils um 1 größerem Index zeigt, daß $\psi_{n+2} > \sqrt{5}$ oder $\psi_{n+1} > \sqrt{5}$ ist. Damit ist die erste Behauptung aus Satz 6 bewiesen.

Zum Beweis der zweiten Behauptung betrachten wir die Zahl $\alpha = \gamma$. Wegen $\gamma = 1 + \frac{1}{\gamma}$ ist $\alpha = [1; 1, 1, 1, \ldots]$, d. h. $a_n = 1$ und $\alpha_n = \alpha$ für jedes n. Hieraus folgt, daß

$$\frac{p_0}{q_0} = \frac{1}{1}, \quad \frac{p_1}{q_1} = \frac{2}{1}, \quad \frac{p_2}{q_2} = \frac{3}{2},$$

also $q_1 = p_0$, $q_2 = p_1$ ist. Auf Grund der Formeln $p_{n+1} = p_n + p_{n-1}$, $q_{n+1} = q_n + q_{n-1}$ folgt hieraus durch vollständige Induktion, daß für jedes n

$$q_{n+1} = p_n$$

und mithin

$$\frac{q_{n+1}}{q_n} = \frac{p_n}{q_n} \to \alpha = \gamma \quad (n \to \infty)$$

gilt. Daher ist für alle hinreichend großen n, wie klein $\varepsilon > 0$ auch immer gewählt sein mag,

$$\frac{q_{n+1}}{q_n} < \gamma + \varepsilon,$$

also auf Grund von (10)

$$\psi_n < \frac{1}{\gamma} + \gamma + \varepsilon = \sqrt{5} + \varepsilon.$$

Hieraus folgt aber unmittelbar $\left(\text{unter Benutzung von } \left|\gamma - \frac{p_n}{q_n}\right| = \frac{1}{\psi_n q_n^2}, \text{ vgl. S. 284} \right.$ — *Anm. d. wissenschaftl. Red.*$\left.\right)$, daß für alle hinreichend großen n

$$\left|\gamma - \frac{p_n}{q_n}\right| > \frac{1}{(\sqrt{5} + \varepsilon)\, q_n^2}$$

ist. Wenn also $c < \frac{1}{\sqrt{5}}$ ist, so kann es unter den Näherungsbrüchen der Zahl γ nur endlich viele geben, für die

$$\left|\gamma - \frac{p_n}{q_n}\right| < \frac{c}{q_n^2}$$

ist, während von den Brüchen, die keine Näherungsbrüche von γ sind, nach Satz 3 keiner diese Ungleichung erfüllen kann. Damit ist auch die zweite Behauptung aus Satz 6 bewiesen.

§ 14. Diophantische Approximationen

Wir sahen bereits, daß der auf den ersten Blick so trivial anmutende Übergang von der Differenz $\left|\alpha - \frac{p}{q}\right|$ zur Differenz $|q\alpha - p|$ (als Maß für die Genauigkeit einer Approximation) uns einerseits eine gesetzmäßigere Lösung einer früher gestellten Aufgabe ermöglichte und zum anderen uns in natürlicher Weise an ganz neue Probleme heranführte. Damit ist aber der eigentliche Wert dieses Überganges noch gar nicht voll erkannt. Die prinzipiell wichtigen und historisch bedeutungsvollen Probleme der hierdurch angebahnten Entwicklung liegen weit außerhalb der Grenzen nicht nur der Theorie der Kettenbrüche, sondern auch des Problemkreises der Annäherung reeller Zahlen durch rationale Brüche. Daher müssen wir auch hier auf diese überaus weittragenden Schlußfolgerungen zu sprechen kommen.

Anläßlich der Suche nach ganzen Zahlen p und q, für die die Differenz $q\alpha - p$ ihrem Betrage nach beliebig klein wird (diese Frage haben wir doch im wesentlichen in der ersten Hälfte des vorliegenden Kapitels behandelt), beschäftigten wir uns — so kann man offenbar unser Bemühen auch kennzeichnen — mit der angenäherten Lösung der Gleichung

$$x\alpha - y = 0 \tag{11}$$

in ganzen Zahlen x und y bei gegebener reeller Zahl α. Dabei setzten wir voraus, daß x positiv ist, und schlossen die triviale genaue Lösung $x = y = 0$ aus. Wenn die Zahl α rational ist, so besitzt die Gleichung (11) stets unendlich viele nichttriviale (d. h. von der im folgenden ein für allemal ausgeschlossenen trivialen Lösung $x = y = 0$ verschiedene) genaue Lösungen. Wenn dagegen α irrational ist, so kann die Gleichung (11) außer der trivialen Lösung keine genauen Lösungen besitzen. Daher erhebt sich hier ganz naturgemäß die Frage nach der angenäherten Lösung dieser Gleichung und den dabei geltenden Gesetzmäßigkeiten. Mit diesen Fragen haben wir uns im vorangehenden beschäftigt. Den in diesem Zusammenhang gewonnenen Satz 6 können wir von unserem neuen Standpunkt aus folgendermaßen formulieren: *Für irrationales α besitzt die Gleichung* (11) *unendlich viele angenäherte Lösungen $x > 0$, y mit*

$$|x\alpha - y| < \frac{1}{\sqrt{5}\,x},$$

und die Zahl $\frac{1}{\sqrt{5}}$ ist die kleinste positive Konstante, die in diesem Zusammenhang auftreten kann.

Lejeune Dirichlet gebührt das Verdienst, eine Methode von bewunderungswerter Einfachheit und Tragweite geschaffen zu haben, mit deren Hilfe man die Aufgabe der näherungsweisen Lösung der Gleichung (11) (und — wie wir bald sehen werden — auch vieler anderer ähnlicher Aufgaben) ohne Benutzung der Kettenbrüche oder anderer spezieller Algorithmen bewältigen kann.

Es sei verlangt,

$$|x\alpha - y| < \frac{1}{t} \quad (x > 0) \tag{12}$$

zu machen, wobei t eine gegebene (große) natürliche Zahl ist. Wie groß müssen wir dazu die Zahl x (und damit auch y) wählen? Um hierauf eine Antwort zu finden, lassen wir x der Reihe nach die Zahlen $0, 1, 2, \ldots, t$ durchlaufen und bestimmen

zu jedem dieser Werte den gebrochenen Anteil

$$x\alpha - [x\alpha]$$

des Produktes $x\alpha$. Auf diese Weise erhalten wir $t+1$ Zahlen $x\alpha - [x\alpha]$, die offensichtlich sämtlich die Ungleichung

$$0 \leq x\alpha - [x\alpha] < 1 \quad (0 \leq x \leq t)$$

erfüllen. Wenn wir nun das Intervall $[0, 1]$ in t gleiche Teile (der Länge $\frac{1}{t}$) unterteilen, wobei wir zu jedem Teilintervall den linken, aber nicht den rechten Endpunkt hinzurechnen, gehört offensichtlich jede der Zahlen $x\alpha - [x\alpha]$ genau einem der Teilintervalle an. Da wir aber $t+1$ Zahlen und nur t Intervalle haben, müssen in wenigstens einem Teilintervall zwei Zahlen $x_1\alpha - [x_1\alpha]$ und $x_2\alpha - [x_2\alpha]$ enthalten sein. Dann ist aber der Betrag der Differenz dieser beiden Zahlen kleiner als die Länge des sie enthaltenden Teilintervalls, also kleiner als $\frac{1}{t}$. Nehmen wir an, es sei $x_1 < x_2$, und setzen wir $x_2 - x_1 = x$, $[x_2\alpha] - [x_1\alpha] = y$, so finden wir, daß

$$|x\alpha - y| < \frac{1}{t}$$

ist, wobei offensichtlich $0 < x \leq t$ sein muß. Wir gelangen auf diese Weise zu der folgenden, von DIRICHLET stammenden Aussage:

Satz 7. *Es sei α eine beliebige reelle und t eine beliebige natürliche Zahl. Dann gibt es stets ganze Zahlen x und y derart, daß*

$$|x\alpha - y| < \frac{1}{t} \tag{13}$$

gilt, wobei $0 < x \leq t$ ist.

Man kann also die Ungleichung (12) immer erfüllen, wobei man x nicht größer als die gegebene Zahl t zu wählen braucht. Aus der Ungleichung (13) ergibt sich offensichtlich die Ungleichung

$$|x\alpha - y| < \frac{1}{x},$$

deren Bestehen für geeignet gewählte Lösungen uns bereits aus der Theorie der Kettenbrüche bekannt ist. Wir haben sie jetzt mit Hilfe der DIRICHLETschen Methode ohne irgendeinen Algorithmus bewiesen. Andererseits läßt sich Satz 7 auch leicht mit Hilfe der Kettenbrüche beweisen: Dazu seien $\frac{p_n}{q_n}$ die Näherungsbrüche der Zahl α. Wir wählen dann die Zahl n so (was offenbar stets möglich ist), daß $q_n \leq t < q_{n+1}$ ist. Dann ist — wie bekannt —

$$|q_n\alpha - p_n| < \frac{1}{q_{n+1}} < \frac{1}{t} \quad (q_n \leq t),$$

was den Satz 7 beweist. Der Hauptvorteil der Methode DIRICHLETs besteht nun keineswegs nur in ihrer Einfachheit (obwohl dies schon ein hinreichendes Moment für ihre Bedeutung wäre). Ihre Tragweite erweist sich erst dann in vollem Maße, wenn wir von der elementaren Gleichung (11) zu komplizierteren und allgemeineren Aufgaben übergehen.

DIRICHLET bemerkte, daß man sich dieselbe Aufgabe, die wir eben für die Gleichung (11) gelöst haben, auch allgemein für Gleichungen der Form

$$x_1\alpha_1 + x_2\alpha_2 + \ldots + x_n\alpha_n - y = 0$$

stellen kann, wobei $\alpha_1, \alpha_2, \ldots, \alpha_n$ gegebene reelle Zahlen und $x_1, x_2, \ldots, x_n, y$ Veränderliche für ganze Zahlen seien. Dabei soll natürlich die triviale Lösung $x_1 = x_2 = \cdots = x_n = y = 0$ wiederum ausgeschlossen sein. Kann man auch hier durch entsprechende (nichttriviale) Wahl von ganzen Zahlen x_i und y die Differenz

$$\sum_{i=1}^{n} x_i\alpha_i - y$$

ihrem absoluten Betrage nach beliebig klein machen, und, wenn ja, wie muß man dabei die Zahlen x_i (und folglich auch y) wählen? Es wäre hoffnungslos, diese Aufgabe mit Hilfe der Kettenbrüche lösen zu wollen: Die Geschichte des hier betrachteten Gebietes der Mathematik zeigt, daß es keinen allgemeinen Algorithmus für eine derartige gleichzeitige Approximation mehrerer Irrationalitäten $\alpha_1, \alpha_2, \ldots, \alpha_n$ gibt (und wahrscheinlich auch nicht geben kann), wie wir ihn für den Fall einer einzelnen Irrationalität α in Form der Kettenbrüche besitzen. Die Methode DIRICHLETS läßt sich dagegen — wie wir sogleich zeigen werden — sehr leicht der neuen Aufgabe anpassen und führt ebenso einfach zu ihrer Lösung wie im Fall der vorhergehenden Grundaufgabe.

Dazu lassen wir die Variablen $x_1, x_2, \ldots, x_n$ unabhängig voneinander die Zahlen $0, 1, \ldots, t$ durchlaufen (wobei t eine gegebene natürliche Zahl ist). Wir erhalten auf diese Weise offensichtlich $(t+1)^n$ Wertsysteme $(x_1, x_2, \ldots, x_n)$ und folglich ebenso viele Summenwerte $\sum_{i=1}^{n} x_i\alpha_i$ und ebenso viele gebrochene Anteile

$$\sum_{i=1}^{n} x_i\alpha_i - \left[\sum_{i=1}^{n} x_i\alpha_i\right]. \tag{14}$$

(Von diesen Anteilen können unter Umständen einige gleich sein — dieselbe Bemerkung gilt übrigens auch für den oben betrachteten Grundfall —, jedoch ändert dies nichts am folgenden Beweis.)

Wir teilen nun das Intervall $[0, 1]$ in t^n gleiche Teile $\left(\text{der Länge } \frac{1}{t^n}\right)$, wobei jeweils wieder der linke, aber nicht der rechte Endpunkt zum Teilintervall hinzugezählt werden soll. Dann liegt wieder jeder gebrochene Anteil in genau einem Teilintervall, und da die Anzahl $(t+1)^n$ dieser Anteile größer ist als die Anzahl t^n der Teilintervalle, enthält mindestens eines der Teilintervalle zwei der gebrochenen Anteile (14). Wenn wir schließlich die Werte der Veränderlichen dieser beiden gebrochenen Anteile mit $x_1', x_2', \ldots, x_n'$ bzw. $x_1'', x_2'', \ldots, x_n''$ bezeichnen und

$$x_i'' - x_i' = x_i \quad (1 \leq i \leq n), \qquad \left[\sum_{i=1}^{n} x_i''\alpha_i\right] - \left[\sum_{i=1}^{n} x_i'\alpha_i\right] = y$$

setzen, so erhalten wir wie früher:

$$\left|\sum_{i=1}^{n} x_i\alpha_i - y\right| < \frac{1}{t^n},$$

wobei $|x_i| \leq t$ und nicht alle x_i gleich Null sind $\left(\text{was wir bequem in } \sum_{i=1}^{n} x_i^2 > 0 \text{ zu-}\right.$

sammenfassen können). Wir kommen auf diese Weise zu folgendem Satz, der ebenfalls von DIRICHLET aufgestellt wurde:

Satz 8. *Wenn $\alpha_1, \alpha_2, \ldots, \alpha_n$ beliebige reelle Zahlen sind und t eine beliebige natürliche Zahl ist, so gibt es ganze Zahlen $x_1, x_2, \ldots, x_n, y$ derart, daß*

$$\left|\sum_{i=1}^{n} x_i\alpha_i - y\right| < \frac{1}{t^n}, \quad |x_i| \leq t \quad (1 \leq i \leq n), \quad \sum_{i=1}^{n} x_i^2 > 0$$

ist.

Wenn wir die größte der Zahlen $|x_1|, |x_2|, \ldots, |x_n|$ mit x bezeichnen, so erhalten wir aus Satz 8 offensichtlich die

Folgerung. *Unter den Voraussetzungen des Satzes 8 gibt es unendlich viele*[1]) *Wertsysteme $x_1, x_2, \ldots, x_n, y$ mit*

$$\left|\sum_{i=1}^{n} x_i\alpha_i - y\right| < \frac{1}{x^n}.$$

Diese Ungleichung, die die aus der Theorie der Kettenbrüche bekannte Ungleichung

$$|q\alpha - p| < \frac{1}{q}$$

verallgemeinert, läßt sich nicht mit bekannten algorithmischen Methoden beweisen und stellt daher ein bemerkenswertes Anwendungsbeispiel der DIRICHLETschen Methode dar.

Nachdem dieser erste Schritt einer Verallgemeinerung erst einmal getan war, eröffnete sich bald ein unerschöpfliches Forschungsfeld. Von der allgemeinen linearen homogenen Gleichung, mit der auch wir uns eben beschäftigt haben, ging man zunächst einerseits zu Systemen von solchen Gleichungen und andererseits zu inhomogenen Gleichungen über. Schließlich beschäftigte man sich nicht nur mit linearen Gleichungen, sondern auch mit algebraischen Gleichungen höheren Grades

---

[1]) Dies kann man nach dem Muster des Beweises von EUKLID für die Unendlichkeit der Menge der Primzahlen folgendermaßen erschließen: Es seien $x_{1\mu}, x_{2\mu}, \ldots, x_{n\mu}, y_\mu$ $(\mu = 1, 2, \ldots, m)$ beliebige derartige Wertsysteme, unter denen ohne Beschränkung der Allgemeinheit $\left|\sum_{i=1}^{n} x_{im}\alpha_i - y_m\right|$ minimal sei. Wenn $\left|\sum_{i=1}^{n} x_{im}\alpha_i - y_m\right| = 0$ ist, so sind wir fertig, da dann alle Wertsysteme $k x_{1m}, k x_{2m}, \ldots, k x_{nm}, k y_m$, wobei k eine beliebige ganze und von Null verschiedene Zahl ist, das Verlangte leisten. Ist hingegen $\left|\sum_{i=1}^{n} x_{im}\alpha_i - y_m\right| \neq 0$, so gibt es eine natürliche Zahl t derart, daß $\frac{1}{t^n} \leq \left|\sum_{i=1}^{n} x_{im}\alpha_i - y_m\right|$ ist. Zu dieser Zahl t gibt es nach Satz 8 dann ein Wertsystem $x_{1\,m+1}, x_{2\,m+1}, \ldots, x_{n\,m+1}, y_{m+1}$ derart, daß $\left|\sum_{i=1}^{n} x_{i\,m+1}\alpha_i - y_{m+1}\right| < \frac{1}{t^n}$, $|x_{i\,m+1}| \leq t$ $(1 \leq i \leq n)$ und $\sum_{i=1}^{n} x_{i\,m+1}^2 > 0$ ist. Dann ist aber — wegen der Wahl von $\left|\sum_{i=1}^{n} x_{im}\alpha_i - y_m\right|$ — sicherlich das Wertsystem $x_{1\,m+1}, x_{2\,m+1}, \ldots, x_{n\,m+1}, y_{m+1}$ von allen Wertsystemen $x_{1\mu}, x_{2\mu}, \ldots, x_{n\mu}, y_\mu$ $(\mu = 1, 2, \ldots, m)$ verschieden und daher $\left|\sum_{i=1}^{n} x_{i\,m+1}\alpha_i - y_{m+1}\right| < \frac{1}{x^n}$, sofern x die größte der Zahlen $|x_{i\,m+1}|$ ist. — *Anm. d. wissenschaftl. Red.*

und danach mit gewissen transzendenten Gleichungen. Damit wurde die Lehre von der näherungsweisen Lösung von Gleichungen durch ganze Zahlen geschaffen, die nach einem Vorschlag von MINKOWSKI heute üblicherweise *Theorie der diophantischen Approximationen* genannt wird. Dieser Name ist sehr treffend, weil man schon von altersher unter *diophantischen Gleichungen* Gleichungen in Unbestimmten für ganze Zahlen versteht. Die Begründer dieser Theorie waren DIRICHLET, TSCHEBYSCHEW, KRONECKER und MINKOWSKI. Insbesondere wurde durch MINKOWSKI eine tiefe und sehr zugkräftige Methode für dieses Gebiet geschaffen, die in der systematischen Anwendung geometrischer Gesetze auf Probleme der Zahlentheorie (der sogenannten „Geometrie der Zahlen") besteht.

Die inhomogene lineare Grundaufgabe dieses Gebietes besteht offensichtlich darin, die angenäherten Lösungen der Gleichung

$$x\alpha - y - \beta = 0 \tag{15}$$

zu finden, wobei α und β gegebene reelle Zahlen und x und y Unbestimmte für ganze Zahlen sind. Diese Aufgabe wurde erstmalig von dem großen russischen Mathematiker TSCHEBYSCHEW gestellt und auch eingehend untersucht. Wenn wir voraussetzen, daß $0 < \beta < 1$ ist (wodurch offenbar die Allgemeinheit der Aufgabenstellung nicht eingeschränkt wird), so ist $y = [x\alpha]$, und die von TSCHEBYSCHEW gestellte Aufgabe besteht in der Bestimmung solcher ganzzahligen Werte von x, für die der gebrochene Anteil des Produktes $x\alpha$ nahe bei der vorgegebenen Zahl β liegt, analog wie wir früher die Werte x suchten, für die dieser Anteil möglichst nahe bei Null liegt. Schon daß es möglich sein soll, für ein beliebiges $\varepsilon > 0$ passende ganze Zahlen x und y mit

$$|x\alpha - y - \beta| < \varepsilon \tag{16}$$

zu finden (d. h. die Möglichkeit, die Gleichung (15) durch ganze Zahlen beliebig genau zu lösen), ist hier zunächst keineswegs klar. Wenn z. B. die Zahl $\alpha = \frac{a}{b}$ rational ist, so läßt sich auch $x\alpha - y$ für beliebiges ganzes x und y als rationaler Bruch mit Nenner b darstellen; falls also β von dem nächsten Bruch mit dem Nenner b den positiven Abstand δ hat, läßt sich die Ungleichung

$$|x\alpha - y - \beta| < \delta$$

nicht durch ganze Zahlen x und y erfüllen.

Um so bemerkenswerter ist es, daß *sich für irrationales α die Ungleichung* (16) *für beliebiges β und beliebiges $\varepsilon > 0$ mit passend gewählten ganzen Zahlen x und y realisieren läßt.* Am einfachsten läßt sich das durch folgende Überlegung beweisen: Aus der Theorie der Kettenbrüche wissen wir, daß es zu jeder irrationalen Zahl α ganze Zahlen p und q gibt derart, daß

$$0 < q\alpha - p < \varepsilon$$

ist. Setzen wir zur Abkürzung $q\alpha - p = \lambda$ und betrachten wir die Zahlenfolge

$$0, \lambda, 2\lambda, \ldots, k\lambda, \ldots,$$

so finden wir eine Zahl $k \geq 0$, für die

$$k\lambda \leq \beta < (k+1)\lambda$$

und mithin

$$|k\lambda - \beta| = |kq\alpha - kp - \beta| < \lambda < \varepsilon$$

ist. Wenn wir also $x = kq$ und $y = kp$ setzen, so ist die Ungleichung (16) erfüllt

Bis jetzt haben wir uns nur mit der Frage nach der Möglichkeit einer angenäherten Lösung der Gleichung (15) mit beliebiger Genauigkeit befaßt, ohne dabei nach der Größe der dazu erforderlichen Zahlen x und y zu fragen. Man kann jedoch auch in dieser Richtung noch fortschreiten und Gesetze aufstellen, die analog den Sätzen für die homogene Aufgabe sind, die wir im vorangehenden erhalten haben. Wir beschränken uns hier auf den Beweis eines Satzes von TSCHEBYSCHEW, der ein wichtiges Resultat seiner Untersuchungen darstellt.

Satz 9. *Wenn α eine irrationale Zahl und β eine beliebige reelle Zahl ist, so gibt es unendlich viele ganze Zahlen x und y, für die*

$$|x\alpha - y - \beta| < \frac{3}{|x|}$$

ist.

Beweis. Es sei $\frac{p}{q}$ ein Näherungsbruch der Zahl α, so daß also

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}$$

ist. Wir bezeichnen dann mit r die dem Produkt $q\beta$ am nächsten gelegene ganze Zahl, für die mithin

$$|q\beta - r| \leq \frac{1}{2}, \quad \left|\beta - \frac{r}{q}\right| \leq \frac{1}{2q} \tag{17}$$

gilt. Aus Kapitel II wissen wir, daß wir die Kongruenz

$$px \equiv r \pmod{q}$$

durch eine ganze Zahl x erfüllen können, die in einem beliebig vorgegebenen Intervall der Länge q liegt. In diesem Sinne sei

$$px \equiv r \pmod{q}, \quad \frac{q}{2} \leq x < \frac{3q}{2},$$

so daß also

$$q > \frac{2}{3}x \tag{18}$$

und außerdem

$$px - r = qy \tag{19}$$

ist. Dann folgt aus (17), (18) und (19):

$$\begin{aligned} |x\alpha - y - \beta| &= \left|x\left(\alpha - \frac{p}{q}\right) - \left(\beta - \frac{r}{q}\right) + \frac{px - qy - r}{q}\right| \\ &= \left|x\left(\alpha - \frac{p}{q}\right) - \left(\beta - \frac{r}{q}\right)\right| < \frac{x}{q^2} + \frac{1}{2q} < \frac{3}{x}, \end{aligned}$$

wobei gemäß $x \geq \frac{1}{2}q$ die Zahlen x und y noch beliebig groß gewählt werden können. Damit ist Satz 9 bewiesen.

Die Konstante 3 auf der rechten Seite der Ungleichung kann noch bedeutend herabgesetzt werden. MINKOWSKI hat für die nicht durch ganze Zahlen genau lösbaren Gleichungen (15) gezeigt, daß der beste Wert der Konstanten $\frac{1}{4}$ ist.

Kapitel VI

ALGEBRAISCHE UND TRANSZENDENTE ZAHLEN

§ 15. Der Satz von Liouville und das erste Auftreten von transzendenten Zahlen

Jede rationale Zahl $\frac{a}{b}$ ist offensichtlich Wurzel einer Gleichung ersten Grades mit ganzen Koeffizienten, nämlich der Gleichung

$$bx - a = 0,$$

und umgekehrt ist jede Wurzel einer solchen Gleichung eine gewisse rationale Zahl. Auf diese Weise können wir die Menge der rationalen Zahlen als Menge der Wurzeln aller Gleichungen ersten Grades mit ganzen Koeffizienten charakterisieren. Von diesem Standpunkt aus werden wir natürlich geneigt sein, als einfachste Irrationalzahlen diejenigen anzunehmen, welche quadratischen Gleichungen

$$ax^2 + bx + c = 0 \tag{1}$$

mit ganzen Koeffizienten a, b, c genügen. Tatsächlich sind ja auch die ersten Irrationalitäten, mit denen man schon auf der Schule bekannt wird, Quadratwurzeln aus ganzen oder, allgemeiner, rationalen Zahlen. Zwar ist jede derartige Zahl

$$\alpha = \sqrt{\frac{p}{q}} \tag{2}$$

Wurzel der quadratischen Gleichung

$$qx^2 - p = 0,$$

die lediglich einen Spezialfall der Gleichung (1) darstellt, jedoch kann man bekanntlich jede Wurzel einer beliebigen Gleichung (1) mit ganzzahligen Koeffizienten a, b, c rational durch Irrationalitäten vom Typus (2) ausdrücken. Weiter lernt man in der Schule Wurzeln dritten, vierten Grades usw. aus rationalen Zahlen kennen. Derartige Irrationalitäten sind entsprechend Wurzeln von Gleichungen dritten, vierten Grades usw. mit ganzen Koeffizienten (jedoch kann man hier im allgemeinen nicht mehr die Wurzeln einer beliebigen Gleichung rational durch derartige Irrationalitäten ausdrücken — *Anm. d. wissenschaftl. Red.*).

Wenn wir in den vorangehenden Überlegungen noch einen Schritt weitergehen, so kommen wir zu denjenigen reellen Zahlen, die Wurzel einer beliebigen algebraischen Gleichung

$$a_0 x^n + a_1 x^{n-1} + \cdots + a_{n-1} x + a_n = 0 \tag{3}$$

mit ganzzahligen Koeffizienten $a_0, a_1, \ldots, a_n$ sind. Derartige Zahlen nennen wir *algebraisch*, und zwar heißt genauer eine reelle Zahl α *algebraisch vom Grad* (oder *von der Ordnung*) n, wenn α Wurzel einer Gleichung (3) vom Grad n, aber keiner Gleichung desselben Typus von einem kleineren Grad ist. In diesem Sinne sind die rationalen Zahlen algebraisch vom Grad (oder von der Ordnung) Eins, während z. B. $\sqrt{2}$ algebraisch vom Grad Zwei ist usw.

Die erste und Hauptfrage nach der Einführung dieses Begriffes lautet offensichtlich: Gibt es reelle Zahlen, die nicht algebraisch sind, die also keiner Gleichung (3) mit ganzen Koeffizienten genügen? Die erste Antwort auf diese Frage in Form eines Beispiels für eine nichtalgebraische Zahl wurde in der Mitte des 19. Jahrhunderts von LIOUVILLE gegeben. Dazu schlug er folgenden Weg ein: Zunächst zeigte er, daß bei der Annäherung algebraischer Zahlen durch rationale Brüche gewisse ganz spezifische Gesetzmäßigkeiten auftreten, und gab dann einfache Beispiele für Zahlen an, die diesen Gesetzmäßigkeiten nicht genügen und folglich nicht algebraisch sein können. Die nichtalgebraischen reellen Zahlen heißen *transzendent* (d. h. die Grenzen (der Algebra) überschreitend). Die transzendenten Zahlen LIOUVILLES, deren Konstruktion wir gleich vorführen, bildeten historisch die ersten Beispiele für derartige Zahlen.

Es sei α eine algebraische Zahl vom Grade $n > 1$, die die Gleichung (3) (und keine Gleichung niedrigeren Grades) erfüllen möge. Wenn wir die linke Seite von (3) zur Abkürzung mit $f(x)$ bezeichnen, so gilt also $f(\alpha) = 0$. Würde nun das Polynom $f(x)$ neben α eine rationale Nullstelle $\frac{a}{b}$ besitzen, so wäre $f(x)$ nach dem Satz von BÉZOUT durch $x - \frac{a}{b}$ teilbar, und wir hätten

$$f(x) = \left(x - \frac{a}{b}\right) f_1(x),$$

wobei $f_1(x)$ ein Polynom mit rationalen Koeffizienten vom Grad $n-1$ ist. Da $f(\alpha) = 0$ und $\alpha - \frac{a}{b} \neq 0$ ist, müßte $f_1(\alpha) = 0$ sein. Wenn dann g der Hauptnenner der Koeffizienten des Polynoms $f_1(x)$ ist, so ist $g f_1(x)$ ein Polynom vom Grad $n-1$ mit ganzen Koeffizienten, für das gleichfalls $g f_1(\alpha) = 0$ sein müßte, was nicht der Fall ist, da α nach Voraussetzung eine algebraische Zahl der Ordnung n sein sollte. Wir haben damit gefunden, daß das Polynom $f(x)$ keine rationale Nullstelle besitzen kann.

Auf Grund des Satzes von BÉZOUT ist wegen $f(\alpha) = 0$ identisch

$$f(x) = (x - \alpha)\varphi(x),$$

wobei $\varphi(x)$ ein Polynom vom Grad $n-1$ ist. Wenn wir nun $x = \frac{p}{q}$ setzen, so erhalten wir

$$f\left(\frac{p}{q}\right) = \left(\frac{p}{q} - \alpha\right)\varphi\left(\frac{p}{q}\right). \tag{4}$$

Wir wählen jetzt den rationalen Bruch $\frac{p}{q}$ aus dem Intervall $(\alpha - 1, \alpha + 1)$ und bezeichnen den größten Wert der Funktion $|\varphi(x)|$ in diesem Intervall mit β, so daß also insbesondere

$$\left|\varphi\left(\frac{p}{q}\right)\right| \leq \beta \tag{5}$$

ist. Da nach dem Bewiesenen

$$\left|f\left(\frac{p}{q}\right)\right| = \left|\frac{a_0 p^n + a_1 p^{n-1} q + \cdots + a_n q^n}{q^n}\right| > 0$$

ist, muß $\left|f\left(\frac{p}{q}\right)\right| \geq \frac{1}{q^n}$ sein, und die Relation (4) ergibt in Verbindung mit der Un-

gleichung (5)

$$\left|\alpha-\frac{p}{q}\right| \geq \frac{1}{\beta q^n}.$$

Dies gilt für jeden beliebigen rationalen Bruch $\frac{p}{q}$ aus dem Intervall $(\alpha-1, \alpha+1)$. Liegt jedoch $\frac{p}{q}$ außerhalb dieses Intervalls, so ist

$$\left|\alpha-\frac{p}{q}\right| \geq 1 \geq \frac{1}{q^n}.$$

Wenn also λ die kleinere der Zahlen 1 und $\frac{1}{\beta}$ bezeichnet, so gilt für jeden rationalen Bruch $\frac{p}{q}$

$$\left|\alpha-\frac{p}{q}\right| \geq \frac{\lambda}{q^n}. \tag{6}$$

Damit haben wir den folgenden Satz bewiesen, der die Grundlage für die LIOUVILLEsche Konstruktionsmethode bildet:

Satz von LIOUVILLE: *Zu jeder algebraischen Zahl vom Grade n gibt es eine positive Zahl λ derart, daß für alle rationalen Brüche $\frac{p}{q}$ die Ungleichung* (6) *gilt.*

Wir wollen zunächst die Aussage dieses Satzes noch etwas genauer beleuchten: Als universelles (d. h. für jede irrationale Zahl α geltendes) Approximationsgesetz haben wir früher festgestellt, daß es unendlich viele rationale Brüche $\frac{p}{q}$ gibt, für die

$$\left|\alpha-\frac{p}{q}\right| < \frac{1}{q^2}$$

ist. Damit ist nicht gesagt, daß nicht für spezielle Irrationalzahlen α ein viel besseres Approximationsgesetz gelten kann. Wir werden bald sehen, daß diese Relation keine Grenze darstellt, daß es vielmehr Irrationalitäten gibt, die durch rationale Brüche besser approximiert werden können, als irgendein vorgegebenes Approximationsgesetz festlegt. Der Satz von LIOUVILLE sagt nun gerade, daß die algebraischen Irrationalitäten nicht beliebig genau durch rationale Brüche approximiert werden können; die Genauigkeit wird durch ein ganz bestimmtes Gesetz beschränkt (und zwar hängt die Genauigkeit vom Grad der betrachteten algebraischen Zahl ab). So gibt es z. B. für jede quadratische Irrationalität α ($n=2$) eine Zahl $\lambda>0$ derart, daß für jedes p und q

$$\left|\alpha-\frac{p}{q}\right| \geq \frac{\lambda}{q^2}$$

ist, d. h., die quadratischen Irrationalitäten (z. B. $\sqrt{2}$, $\sqrt{5}$ usw.) können auf keinen Fall besser als durch das „universelle" Gesetz approximiert werden.

Wenn es uns gelingt, eine reelle Zahl zu konstruieren, die für jedes n eine bessere Annäherung gestattet, als nach dem Satz von LIOUVILLE möglich ist, so ist diese Zahl notwendig transzendent. Es zeigt sich nun, daß eine solche Zahlenart (die sogenannten „transzendenten Zahlen von LIOUVILLE") sehr leicht konstruiert werden kann. Dazu kann jedes Verfahren benutzt werden, daß zur näherungsweisen Darstellung reeller Zahlen durch rationale Brüche dient (jedoch kann man auch die Konstruktion ohne ein solches Verfahren durchführen).

Wir wollen hier zunächst die Methode der Kettenbrüche verwenden: Dazu setzen wir $a_0 = 0$ und bestimmen die weiteren Glieder eines Kettenbruches rekursiv durch folgende Festsetzung: wenn bereits die Teilnenner $a_0, a_1, \ldots, a_n$ (und damit der Näherungsbruch $\frac{p_n}{q_n}$) bestimmt sind, so soll $a_{n+1} = q_n^n$ sein. Der auf diese Weise konstruierte Kettenbruch $[a_0; a_1, a_2, \ldots, a_n, \ldots]$ stellt — wie wir wissen — eine gewisse irrationale Zahl α dar. Wir zeigen, daß diese Zahl α transzendent ist.

In der Tat: Bekanntlich ist für jedes k

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k q_{k+1}} = \frac{1}{q_k (q_k a_{k+1} + q_{k-1})} < \frac{1}{q_k^2 a_{k+1}},$$

woraus gemäß $a_{k+1} = q_k^k$

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^{2+k}}$$

folgt. Wenn nun $\lambda > 0$ und die natürliche Zahl n beliebig vorgegeben sind, so ist, wenn die Zahl $k \geq n$ so groß gewählt wird, daß $\frac{1}{q_k^k} < \lambda$ ist,

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^{k+2}} < \frac{\lambda}{q_k^k} \leq \frac{\lambda}{q_k^n};$$

daher kann auf Grund des Satzes von Liouville α keine algebraische Zahl vom Grad n und — da n ganz willkürlich ist — überhaupt keine algebraische Zahl sein.

Wir führen jetzt dieselbe Konstruktion noch einmal ohne die Verwendung eines Algorithmus durch. Dazu beginnen wir mit einem ganz beliebigen reduzierten Bruch $\frac{p_1}{q_1}$ und umgeben ihn mit dem Intervall

$$\Delta_1 = \left(\frac{p_1}{q_1} - \frac{1}{q_1}, \frac{p_1}{q_1} + \frac{1}{q_1}\right).$$

Im Intervall Δ_1 wählen wir einen reduzierten Bruch $\frac{p_2}{q_2}$ mit $q_2 > q_1$ und umgeben ihn mit dem Intervall

$$\Delta_2 = \left(\frac{p_2}{q_2} - \frac{1}{q_2^2}, \frac{p_2}{q_2} + \frac{1}{q_2^2}\right).$$

Wenn wir bereits das Intervall Δ_{n-1} konstruiert haben, so wählen wir in diesem einen reduzierten Bruch $\frac{p_n}{q_n}$ derart, daß $q_n > q_{n-1}$ ist, und umgeben ihn mit dem Intervall

$$\Delta_n = \left(\frac{p_n}{q_n} - \frac{1}{q_n^n}, \frac{p_n}{q_n} + \frac{1}{q_n^n}\right).$$

Diesen Prozeß können wir unbeschränkt fortsetzen. Wir zeigen nun, daß jedes Intervall Δ_n vollständig in dem vorhergehenden Intervall Δ_{n-1} $(n \geq 2)$ enthalten ist. In der Tat: Wenn z. B. der rechte Endpunkt von Δ_n rechts von dem rechten Endpunkt von Δ_{n-1} läge (oder mit ihm zusammenfiele), so müßte

$$\frac{p_n}{q_n} < \frac{p_{n-1}}{q_{n-1}} + \frac{1}{q_{n-1}^{n-1}} \leq \frac{p_n}{q_n} + \frac{1}{q_n^n}$$

und daher

$$0 < \frac{p_{n-1}}{q_{n-1}} + \frac{1}{q_{n-1}^{n-1}} - \frac{p_n}{q_n} \leq \frac{1}{q_n^n}$$

sein; da der Mittelteil dieser Ungleichung ein positiver Bruch mit dem Nenner $q_{n-1}^{n-1} q_n$ ist, kann er nicht kleiner als $\frac{1}{q_{n-1}^{n-1} q_n}$ sein, so daß also

$$\frac{1}{q_{n-1}^{n-1} q_n} \leq \frac{1}{q_n^n}$$

sein müßte, woraus sich $q_n \leq q_{n-1}$ ergäbe, was dem Gang unserer Konstruktion widerspricht.

Es ist also Δ_n vollständig in Δ_{n-1} enthalten, so daß — da die Länge der Intervalle Δ_n für $n \to \infty$ gegen 0 strebt — es eine eindeutig bestimmte reelle Zahl α gibt, die in allen Intervallen Δ_n $(n = 1, 2, 3, \ldots)$ enthalten ist.

Es sei jetzt wieder ein $\lambda > 0$ und eine natürliche Zahl n beliebig vorgegeben. Wenn wir dann $k > n$ so groß wählen, daß $\frac{1}{q_k} < \lambda$ wird, dann ist, da α in Δ_k liegt,

$$\left| \alpha - \frac{p_k}{q_k} \right| < \frac{1}{q_k^k} = \frac{1}{q_k} \cdot \frac{1}{q_k^{k-1}} < \frac{\lambda}{q_k^n}.$$

Da $\lambda > 0$ und n ganz beliebig sind, ist mithin α nach dem Satz von LIOUVILLE transzendent.

§ 16. Die Methode von CANTOR

Etwa 20 bis 25 Jahre nach den Arbeiten LIOUVILLES erschienen die Untersuchungen G. CANTORS, in denen er den Grundstein für einen neuen wichtigen Zweig der Mathematik, die Mengenlehre, legte. Eine der ersten Früchte dieser neuen Theorie war ein neuer und außerordentlich origineller Beweis für die Existenz transzendenter Zahlen. Er enthält Überlegungen, die bis dahin noch niemals in der Wissenschaft angewendet worden waren, und wurde in der Folge zum Vorbild einer ganzen Reihe weiterer fruchtbringender Konstruktionen.

Zuerst richtete CANTOR seine Aufmerksamkeit auf die Tatsache, daß man die algebraischen Zahlen abzählen kann. Was soll das bedeuten? Es bedeutet, daß man jeder algebraischen Zahl eine sie registrierende Nummer geben kann. Selbstverständlich bildet die Menge aller algebraischen Zahlen eine unendliche Menge, da zu ihr alle rationalen Zahlen und damit insbesondere alle natürlichen Zahlen gehören. Andererseits steht uns auch eine unbeschränkte Menge von Nummern, nämlich die Menge der natürlichen Zahlen, zur Verfügung. Es zeigt sich nun, daß man mit Hilfe dieser unendlichen Folge von Nummern alle algebraischen Zahlen durchnumerieren kann, ähnlich wie man mit Hilfe der ersten zehn natürlichen Zahlen die Finger der Hände durchnumerieren kann. Aber ist das nicht trivial? Es gibt unendlich viele natürliche Zahlen, können wir nicht mit ihrer Hilfe, da sie uns einen unbeschränkten Vorrat an Nummern bereitstellen, überhaupt jede Menge von Gegenständen durchnumerieren? CANTOR zeigte (und das ist vielleicht

das wesentlichste seiner ersten Resultate), daß dies nicht der Fall ist. Die Menge C aller reellen Zahlen kann nicht mit Hilfe der unendlichen Menge der natürlichen Zahlen durchnumeriert werden. Aus dem Gesagten ergibt sich aber unmittelbar die Existenz von transzendenten Zahlen, denn anderenfalls fiele die Menge C mit der Menge A aller algebraischen Zahlen zusammen, und die Menge C wäre abzählbar.

Wir wollen uns nun zunächst überlegen, wie man die Menge der algebraischen Zahlen durchnumerieren kann. Dazu beachten wir, daß jede algebraische Zahl Wurzel einer algebraischen Gleichung der Form (3) ist, wobei $a_0, a_1, \ldots, a_n$ ganze Zahlen sind und $a_0 \neq 0$ ist. Als „Höhe" dieser Gleichung bezeichnen wir die positive ganze Zahl

$$h = n + |a_0| + |a_1| + \cdots + |a_n|.$$

Wenn nun die Höhe h einer Gleichung vorgegeben ist, so sind durch sie der Grad und die absoluten Beträge der Koeffizienten beschränkt. Es kann also jeweils nur eine endliche Anzahl von Gleichungen der Form (3) zu vorgegebener Höhe h geben. In diesem Sinne wird die Gesamtheit der Gleichungen der Höhe 3 durch die Gleichungen

$$x^2 = 0, \quad -x^2 = 0, \quad x+1 = 0, \quad x-1 = 0,$$
$$-x+1 = 0, \quad -x-1 = 0, \quad 2x = 0, \quad -2x = 0$$

erschöpft. Dieser Umstand gestattet es uns zunächst, alle Gleichungen (3) durchzunumerieren: Der kleinste Wert der Höhe einer Gleichung ist offensichtlich $h = 2$. Zu dieser Höhe gibt es aber nur zwei Gleichungen, nämlich $x = 0$ und $-x = 0$. Diese Gleichungen versehen wir mit den Nummern 1 bzw. 2. Danach gehen wir zu den oben genannten acht Gleichungen der Höhe $h = 3$ über und geben ihnen der Reihe nach die Nummern 3, 4, 5, 6, 7, 8, 9, 10. Anschließend gehen wir zu den Gleichungen der Höhe $h = 4$ über und numerieren sie schrittweise durch die Zahlen 11, 12, ... usw. Da es zu jeder Höhe h nur endlich viele Gleichungen gibt, erhält bei unbeschränkter Fortsetzung dieses Prozesses jede Gleichung vom Typus (3) eine frühere oder spätere, sie registrierende Nummer. Auf diese Weise kann man also tatsächlich alle Gleichungen (3) durchnumerieren.

Wir kommen nun zur Numerierung der (reellen) algebraischen Zahlen. Diese Numerierung gelingt uns, weil nach dem Fundamentalsatz der Algebra jede Gleichung der Form (3) endlich viele (und zwar höchstens n) Wurzeln besitzt. Wir betrachten zunächst die Gleichung mit der Nummer 1. Sie möge n_1 paarweise verschiedene reelle Wurzeln besitzen, die wir mit $\alpha_1, \alpha_2, \ldots, \alpha_{n_1}$ bezeichnen wollen. Dann kommen wir zur Gleichung der Nummer 2. Diese möge n_2 reelle Wurzeln besitzen, die voneinander und von den reellen Wurzeln der Gleichung Nummer 1 verschieden sind und die wir mit $\alpha_{n_1+1}, \alpha_{n_1+2}, \ldots, \alpha_{n_1+n_2}$ bezeichnen wollen. Anschließend bezeichnen wir mit

$$\alpha_{n_1+n_2+1}, \quad \alpha_{n_1+n_2+2}, \quad \ldots, \quad \alpha_{n_1+n_2+n_3}$$

die reellen Wurzeln der Gleichung mit der Nummer 3, die voneinander und von den Wurzeln der Gleichungen mit den Nummern 1 und 2 verschieden sind usw. Wenn wir diesen Prozeß unbeschränkt fortsetzen, so erhalten wir eine Folge

$$\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots \tag{7}$$

von reellen Zahlen, in der jede reelle Zahl, die irgendeiner Gleichung vom Typus (3) genügt, einmal und nur einmal vorkommt. Diese Menge von reellen Zahlen ist

also nichts anderes als die Menge aller algebraischen Zahlen, die wir damit wirklich durchnumeriert haben.

Nachdem die Numerierung der Menge der algebraischen Zahlen gelungen ist, bereitet die Konstruktion von transzendenten Zahlen keinerlei Schwierigkeit mehr. Dazu habe der Kettenbruch, der die Zahl α_k darstellt, die Form

$$\alpha_k = [a_0^{(k)};\ a_1^{(k)},\ a_2^{(k)},\ \ldots,\ a_n^{(k)},\ \ldots],$$

wobei wir (nur für dieses Mal) verabreden wollen, daß in den Fällen, in denen α_k eine rationale Zahl ist, in denen unser Kettenbruch also ein letztes Glied $a_n^{(k)}$ besitzt, $a_{n+1}^{(k)} = a_{n+2}^{(k)} = \cdots = 0$ gesetzt wird (damit nämlich jede reelle Zahl in Form eines unendlichen Kettenbruches erscheint). Wenn wir dann für jedes $k > 0$

$$b_k = a_k^{(k)} + 1$$

setzen, so besitzt der Kettenbruch

$$[0; b_1, b_2, \ldots, b_n, \ldots],$$

mit b_1 begonnen, sämtlich positive Glieder, stellt also eine gewisse reelle Zahl β dar. Wir zeigen nun, daß β transzendent ist. In der Tat: Da die Folge (7) nach Konstruktion sämtliche algebraischen Zahlen enthält, so müßte die Zahl β, falls sie algebraisch wäre, gleich einer der Zahlen α_k aus der Folge (7) sein; dann wäre aber, wegen der Eindeutigkeit der Darstellung der Zahlen als Kettenbrüche, $b_n = a_n^{(k)}$ für jedes $n \geq 0$, also insbesondere $b_k = a_k^{(k)}$, was der Definition der Zahl β widerspricht. Damit ist gezeigt, daß die Zahl β wirklich transzendent ist.

Es ist klar, daß wir bei dieser Konstruktion einer transzendenten Zahl die Definition der Zahlen b_k noch beliebig variieren können. Das zeigt, daß man mittels des CANTORschen Verfahrens beliebig viele transzendente Zahlen erhält.

§ 17. Die arithmetische Natur der klassischen Konstanten

Im vorangehenden war es uns stets ohne erhebliche Schwierigkeiten gelungen, Zahlen mit einer vorgegebenen arithmetischen Struktur zu konstruieren. Wir können im Rahmen des Möglichen beliebig viele Beispiele von Zahlen angeben, die nachweisbar irrational oder transzendent sind und die durch rationale Brüche sehr gut oder auch nicht allzu gut approximierbar sind usw. Im Gegensatz dazu ist es unvergleichlich schwieriger, die arithmetische Natur einer Zahl zu bestimmen, die in der mathematischen Praxis unter keineswegs arithmetischen Gesichtspunkten auftritt, die also sozusagen von außen in die Arithmetik hineinkommt. Ist z. B. die Zahl π, die in der Geometrie als das Verhältnis von Umfang und Durchmesser des Kreises definiert wird, rational oder irrational? Hierüber sagt ihre Definition zunächst gar nichts aus. Oder ist die Zahl e, die in der Analysis als

$$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

oder

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \cdots \tag{8}$$

definiert wird, algebraisch oder transzendent, und inwieweit ist sie durch rationale Brüche approximierbar? Auch ihre Definition läßt hierüber nichts erkennen. Die Zahlen e und π spielen jedoch in der Mathematik eine so fundamentale Rolle, daß man über ihre arithmetische Natur möglichst genauen Aufschluß erhalten möchte. Daraus erklärt es sich, daß die Frage nach der arithmetischen Struktur dieser und auch anderer „klassischer Konstanten" seit langem die Gemüter der Wissenschaftler bewegt. Die Schwierigkeit dieser Aufgabe spornte jedoch — wie es stets der Fall zu sein pflegt — die Forscher zu weiteren Bemühungen an. Wenngleich auch heute noch viele dieser Aufgaben ungelöst sind, so kann sich doch die Forschung mit vollem Recht bedeutender Erfolge in dieser Richtung rühmen, Erfolge, die den Anstrengungen der größten Geister vornehmlich des letzten Jahrhunderts zu verdanken sind.

Verhältnismäßig früh wurde erkannt, daß die Zahlen e und π irrational sind. Der Beweis der Irrationalität von e kann sogar so leicht erbracht werden, daß wir ihn hier vorführen können. Als Ausgangspunkt wählen wir die Definition (8) von e als Summe einer unendlichen Reihe.

Wäre die Zahl e rational, also $e = \frac{a}{b}$, so wäre die Zahl $b!e$ ganz; aus (8) ergibt sich jedoch

$$b!e = b! + \frac{b!}{1!} + \frac{b!}{2!} + \cdots + \frac{b!}{b!} + \\ + \frac{b!}{(b+1)!} + \frac{b!}{(b+2)!} + \cdots,$$

wobei alle Summanden der ersten Zeile auf der rechten Seite ganze Zahlen sind; stände nun auf der linken Seite eine ganze Zahl, so wäre auch die Summe in der zweiten Zeile, also die Größe

$$\alpha = \frac{1}{b+1} + \frac{1}{(b+1)(b+2)} + \frac{1}{(b+1)(b+2)(b+3)} + \cdots$$

eine ganze Zahl, im Widerspruch dazu, daß diese Summe positiv und andererseits

$$\alpha < \frac{1}{b+1} + \frac{1}{(b+1)^2} + \frac{1}{(b+1)^3} + \cdots = \frac{1}{b} \leq 1,$$

also $0 < \alpha < 1$ ist. Damit ist die Irrationalität von e bewiesen.

Der Beweis für die Irrationalität von π ist bedeutend verwickelter, da wir für π keine so einfache und arithmetisch übersehbare Darstellung kennen, wie wir sie für die Zahl e in Form der Reihe (8) besitzen. Er wurde erst in dem Augenblick gefunden, als es gelang, die Zahl $\frac{\pi}{4}$ in Form eines Kettenbruches darzustellen. Da sich dieser Kettenbruch als unendlich erweist, ist die Zahl $\frac{\pi}{4}$ und mithin auch die Zahl π irrational.

Noch unverhältnismäßig schwieriger war die Frage zu lösen, ob die Zahlen e und π algebraisch oder transzendent sind. Zuerst bewies im Jahre 1873 HERMITE mittels komplizierter analytischer Methoden, die auf der Betrachtung gewisser bestimmter Integrale mit unendlichen Grenzen beruhen, die Transzendenz der Zahl e. In Anbetracht der fundamentalen Rolle, die die Zahl e in der Differential- und Integralrechnung spielt, darf die Benutzung von Integralen zur Lösung der Frage nach der arithmetischen Natur dieser Zahl nicht besonders verwunderlich

erscheinen. Neun Jahre später bewies LINDEMANN die Transzendenz der Zahl π, indem er das Verfahren von HERMITE weiter ausbaute. Dieser Zugang zur Zahl π durch das HERMITEsche Verfahren mag zunächst recht seltsam anmuten, da nach ihrer Definition die Zahlen e und π nichts miteinander zu tun haben; sie sind aus ganz verschiedenen Gebieten in die Arithmetik gekommen. Bemerkenswert ist nun, daß die Zahlen e und π in der Analysis eine ganz enge Verknüpfung durch die Formel $e^{\pi i} = -1$ erfahren. Auf diese Weise kann man jede dieser Zahlen auf die andere zurückführen, ein Umstand, den sich LINDEMANN in seinem Beweis zunutze machte.

Nachdem damit die Transzendenz der wichtigsten beiden klassischen Konstanten bewiesen war, ereignete sich auf diesem Gebiet lange Zeit hindurch nichts Neues. Wie kümmerlich es noch heute in diesem Gebiet mit allgemeinen Methoden steht, kann man daraus ersehen, daß wir gegenwärtig noch nichts über die arithmetische Natur solcher Zahlen wie $e + \pi$, $e - \pi$ oder $e\pi$ wissen (es ist sogar unbekannt, ob sie rational oder irrational sind). 1907 gab HILBERT einen wesentlich einfacheren Beweis für die Transzendenz der Zahlen e und π als vorher HERMITE und LINDEMANN. HURWITZ gelang es dann, einen Beweis zu finden, der im wesentlichen ohne Integrale auskommt (jedoch dafür ausgiebig die Mittel der Differentialrechnung benutzt). Wenn auch alle diese Untersuchungen außerordentlich tiefliegend und zudem wesentlich eleganter als die vorangehenden waren, so waren sie doch methodisch keineswegs ergiebiger als diese; sie vermochten keine neuen Resultate aufzudecken.

Auf dem internationalen Mathematiker-Kongreß von 1900 lenkte HILBERT in seinem berühmten Vortrag über aktuelle mathematische Probleme die Aufmerksamkeit darauf, daß wir von den von außen in die Arithmetik eingedrungenen Zahlen e und π bislang genug wissen, während über die arithmetische Natur gewisser Zahlen sozusagen innerarithmetischer Herkunft bis dahin nichts bekannt war. Als Beispiel verwies er auf die Zahl $2^{\sqrt{2}}$ und allgemein auf Zahlen der Form α^β, wobei α und β algebraische Zahlen, α verschieden von Null und Eins, und β irrational ist (eine derartige Zahl ist z. B. e^π, da aus $e^{\pi i} = -1 = i^2$ folgt, daß $e^\pi = i^{-2i}$ ist). Diese Aufgabe von HILBERT wurde sehr berühmt, jedoch waren im Laufe von 30 Jahren in dieser Richtung keine Erfolge zu verzeichnen. Es zeigte sich zu diesem Problemkreis kein Zugang. Dann öffnete in den dreißiger Jahren der damals noch sehr junge sowjetische Mathematiker A. O. GELFOND das Tor mit einem weittragenden allgemeinen Verfahren, das sich auf die Theorie der Funktionen einer komplexen Veränderlichen gründete. Nachdem es anfangs nur die Lösung für einige spezielle Klassen der HILBERTschen Aufgabe zuließ, gelang 1936 der Beweis für alle von HILBERT genannten Zahlen. GELFOND zeigte, daß alle diese Zahlen ausnahmslos transzendent sind. Die bedeutenden Verfahren, die der sowjetische Gelehrte begründet hatte, waren von ihm schrittweise so weit vervollkommnet worden, daß dieser Beweis gelang. Heute geben sie die Möglichkeit, die Transzendenz noch vieler weiterer ganzer Klassen von Zahlen nachzuweisen.

LITERATUR

[1] Полное собрание сочинений П. Л. Чебышева, т. I, Теория чисел (Vollständige gesammelte Werke von P. L. TSCHEBYSCHEW, Band I, Zahlentheorie), Moskau-Leningrad 1944.
In diesem Bande sind alle klassischen Werke P. L. TSCHEBYSCHEWS über Zahlentheorie enthalten, und zwar seine „Theorie der Kongruenzen", die in dem vortrefflich geschriebenen Kurs der Zahlentheorie klar und verständlich dargestellt ist, seine Untersuchungen über das Verteilungsgesetz der Primzahlen in der Folge der natürlichen Zahlen und vieles andere.

[2] Виноградов, И. М., Основы теории чисел (WINOGRADOW, I. M., Grundzüge der Zahlentheorie), 5. überarbeitete Aufl., Moskau-Leningrad 1949, in 6. verbesserter Aufl. 1952 erschienen. Deutsche Übersetzung nach der 6. Auflage: WINOGRADOW, I. M., Elemente der Zahlentheorie, Berlin [und München] 1955.
Kurzes, aber klar und einfach geschriebenes Lehrbuch der Zahlentheorie, das eine große Zahl origineller Aufgaben und Fragen mit Lösungen enthält.

[3] Арнольд, И. В., Теория чисел (ARNOLD, I. W., Zahlentheorie), Moskau 1939.
Ein gutes Lehrbuch der Zahlentheorie, das außer einem elementaren Kurs viele Zusammenfassungen aus anderen Teilen der modernen Arithmetik, teilweise ohne Beweis, enthält. Am Schluß des Buches finden sich zahlreiche Übungen zur Zahlentheorie.

[4] Диксон, Л. Е., Введение в теорию чисел (DICKSON, L. E., Einführung in die Zahlentheorie), überarbeitete Übersetzung der 1. englischen Auflage, Tbilissi 1941.
Originalausgabe: DICKSON, L. E., Introduction to the theory of numbers, Cambridge 1931. Deutsche Übersetzung: DICKSON, L. E., Einführung in die Zahlentheorie, Leipzig und Berlin 1931.
Das Buch enthält eine wertvolle Aufgabensammlung, besonders über diophantische Gleichungen höheren Grades. Die russische Ausgabe ist mit ausführlichen Lösungen versehen.

[5] Делоне, Б. Н., Петербургская школа теории чисел (DELAUNAY, B. N., Die Petersburger zahlentheoretische Schule), Moskau-Leningrad 1947.
Das Buch macht in chronologischer Reihenfolge mit den Werken der größten russischen Gelehrten, die auf dem Gebiet der Zahlentheorie gearbeitet haben, bekannt: P. L. TSCHEBYSCHEW, A. N. KORKIN, E. I. SOLOTARJOW, A. A. MARKOFF, G. F. WORONOI und I. M. WINOGRADOW.

[6] Шнирельман, Л. Г., Простые числа (SCHNIRELMAN, L. G., Primzahlen), Moskau-Leningrad 1940.
Eine kleine (60 Seiten oktav), aber alles Wichtige enthaltende Broschüre, die eine Einführung in die Zahlentheorie gibt. Es sind enthalten: Die Grundzüge der Theorie der Kongruenzen, die Arithmetik der komplexen Zahlen, einige Ergebnisse TSCHEBYSCHEWS mit vereinfachten Beweisen u. a. Mit Ausnahme des letzten Paragraphen ist die ganze Broschüre für Schüler der zehnten Klasse verständlich.

[7] Хинчин, А. Я., Цепные дроби (CHINTSCHIN, A. J., Kettenbrüche), 2. Aufl., Moskau-Leningrad 1950.
Deutsche Übersetzung: KHINTCHINE, A. J., Kettenbrüche, Leipzig 1956.
Eine systematische Darstellung der Theorie der Kettenbrüche. Für die Lektüre des letzten Kapitels über die metrische Theorie der Kettenbrüche wird beim Leser die Kenntnis der Integralrechnung und der Maßtheorie vorausgesetzt.

[8] Хинчин, А. Я., Три жемчужины теории чисел (CHINTSCHIN, A. J., Drei Perlen der Zahlentheorie), Moskau-Leningrad 1948.
Deutsche Übersetzung: CHINTSCHIN, A. J., Drei Perlen der Zahlentheorie, Berlin 1950.
Das Büchlein ist dem Beweis dreier berühmter Probleme der Zahlentheorie gewidmet (einem Satz über die arithmetische Progression, einem Satz über die Dichte einer Summe von Zahlenfolgen und dem WARINGschen Problem). Alle diese Beweise sind erst im Laufe der vierziger Jahre gefunden worden. Sie sind sämtlich elementar, erfordern aber vom Leser große Aufmerksamkeit und Konzentrationsfähigkeit.

[9] Кузьмин, Р. О., и Д. К. Фаддеев, Арифметика и алгебра комплексных чисел (KUSMIN, R. O., und D. K. FADDEJEW, Arithmetik und Algebra der komplexen Zahlen), Leningrad 1939.
Der erste Teil dieses Buches behandelt die Algebra der komplexen Zahlen und die Eigenschaften der elementaren Funktionen einer komplexen Veränderlichen sowie Begriffe, die mit den Quaternionen zusammenhängen. Der zweite, kürzere Teil ist der Teilbarkeitstheorie im Bereich der ganzrationalen, der ganzen komplexen Zahlen und der Zahlen der Form $a + b\varrho$, wobei a und b ganze Zahlen sind und ϱ eine dritte Einheitswurzel ist, und schließlich weiteren Verallgemeinerungen der Teilbarkeitstheorie gewidmet. Fast der gesamte Inhalt dieses zweiten Teiles ist in erheblich kürzerer Form bereits in dem oben erwähnten Büchlein von SCHNIRELMAN enthalten.
[10] Хинчин, А. Я., Великая теорема Ферма (CHINTSCHIN, A. J., Der große Fermatsche Satz), Moskau-Leningrad 1932. Die Darlegungen im Haupttext der Broschüre erfordern vom Leser keine größeren Kenntnisse aus der Zahlentheorie. Lediglich die Ergänzungen, die sich mit den Untersuchungen KUMMERS in der Theorie der algebraischen Zahlen befassen, verlangen einige Grundkenntnisse aus dieser Theorie.

Der deutsche Leser sei noch verwiesen auf:

BACHMANN, P., Das Fermatproblem in seiner bisherigen Entwicklung, Berlin und Leipzig 1919.
DEDEKIND, R., Was sind und was sollen die Zahlen? [Braunschweig 1964 und] Berlin 1967.
—, Stetigkeit und irrationale Zahlen, [Braunschweig 1964 und] Berlin 1967.
—, Über die Theorie der ganzen algebraischen Zahlen, [Braunschweig und] Berlin 1964.
DIRICHLET-DEDEKIND, Vorlesungen über Zahlentheorie, 4. Aufl., Braunschweig 1894.
DYNKIN, E. B., und W. A. USPENSKI, Mathematische Unterhaltungen, II. Aufgaben aus der Zahlentheorie, 4. Aufl., Berlin 1968 (Übersetzung aus dem Russischen).
GAUSS, C. F., Disquisitiones Arithmeticae, Leipzig 1801 (Untersuchungen über höhere Arithmetik, deutsch von Maser, Berlin 1889).
GELFOND, A. O., Die Auflösung von Gleichungen in ganzen Zahlen, 4. Aufl., Berlin 1968 (Übersetzung aus dem Russischen).
HASSE, H., Vorlesungen über Zahlentheorie, 2. Aufl., Berlin-Göttingen-Heidelberg 1954.
HASSE, H., Zahlentheorie, 2. Aufl., Berlin 1963.
HOLZER, L., Zahlentheorie I, II, Leipzig 1958 bzw. 1959.
JUNG, H. W., Zahlentheorie, 2. Aufl., Leipzig 1952.
KOKSMA, J. F., Diophantische Approximationen, Berlin 1936.
LANDAU, E., Vorlesungen über Zahlentheorie, Erster Band (Aus der elementaren Zahlentheorie), Dritter Band (Aus der algebraischen Zahlentheorie), Leipzig 1927.
—, Diophantische Gleichungen mit endlich vielen Lösungen. Neu herausgegeben von A. WALFISZ, Berlin 1959.
LIETZMANN, W., Der pythagoreische Lehrsatz, mit einem Ausblick auf das Fermatsche Problem, 8. Aufl., Leipzig 1966.
LOEWY, A., Lehrbuch der Algebra, Leipzig 1915.
MINKOWSKI, H., Geometrie der Zahlen, Leipzig 1896.
PERRON, O., Irrationalzahlen, 4. Aufl., Berlin 1960.
—, Die Lehre von den Kettenbrüchen, 2. Aufl., Leipzig 1929.
PRACHAR, K., Primzahlverteilung, Berlin-Göttingen-Heidelberg 1957.
SCHOLZ, A., Einführung in die Zahlentheorie, 4. Aufl., Berlin 1966.
SKOLEM, TH., Diophantische Gleichungen, Berlin 1938.
SPECHT, W., Elementare Beweise der Primzahlsätze, Berlin 1956.

W. M. BRADIS

KOPFRECHNEN UND SCHRIFTLICHES RECHNEN
HILFSMITTEL FÜR DAS RECHNEN

Kapitel I

ALLGEMEINE BEMERKUNGEN ÜBER DAS RECHNEN UND DIE NÄHERUNGSRECHNUNGEN

§ 1. Allgemeine Bemerkungen über das Rechnen in der Schule[1])

Der Mathematikunterricht in den unteren Schulklassen ist fast ausschließlich dem Rechnen gewidmet. Der gegenwärtig geltende Unterrichtsplan sieht vor, daß ein Schüler nach Abschluß der ersten fünf Klassen im Rechnen nur noch mit den logarithmischen Rechenmethoden vertraut gemacht zu werden braucht, wofür genügend Zeit in der neunten Klasse vorgesehen ist. Auf weitere Fragen, die mit der Theorie und Praxis des Rechnens zusammenhängen, wird in der Schule nicht eingegangen. Da in den höheren Klassen in dieser Richtung nichts mehr getan wird, verlieren die Schüler häufig im Laufe der Zeit auch die wenigen Rechenkenntnisse und -fertigkeiten, die sie in den unteren Klassen erworben haben. Daher ist es nicht verwunderlich, daß Universitäten, und insbesondere technische Hochschulen hinsichtlich der Rechenkultur über die ungenügende Vorbildung der Jugend, die die Oberschulen durchlaufen hat, vielfach Klage führen.

Beim Studium der Mathematik darf man ihre praktischen Anwendungen nicht außer acht lassen, da jede Lösung einer praktischen Aufgabe, die von mathematischen Überlegungen Gebrauch macht, im Endeffekt eine Rechnung mit Zahlen ist. Daher muß es erste und oberste Aufgabe des allgemeinbildenden mathematischen Unterrichtes sein, das richtige, schnelle und ohne unnützen Zeit- und Kräfteaufwand durchzuführende Rechnen zu lehren; denn ein gewisses Minimum an guten und beständigen Fertigkeiten im Rechnen hat jeder zivilisierte Mensch nötig. Dieses Minimum erweitert sich wesentlich für alle technischen und physikalisch-mathematischen Fachrichtungen und in etwas geringerem Maße auch für alle anderen naturwissenschaftlichen Disziplinen. Überdies darf man aber auch den erzieherischen Wert einer richtig angesetzten Rechenarbeit nicht außer acht lassen: Durch sie

---

[1]) In der Originalausgabe handelt es sich hier naturgemäß um eine Kritik des Rechenunterrichtes der sowjetischen Schule. Sie gilt aber in gleichem Maße auch für den Rechenunterricht an den deutschen Schulen. Der erste Absatz dieses Paragraphen der Originalausgabe, der den Gebrauch des russischen Wortes «счёт» (Zählen, Rechnen) für den vorliegenden Artikel regelt, wurde in der deutschen Übersetzung ausgelassen.— *Anm. d. wissenschaftl. Red.*

wird besonders gut das Verantwortungsgefühl entwickelt; indem man — wie es sich gebührt — jeden Schritt rationalisiert, werden die Fähigkeiten in der Selbstkontrolle, nämlich das Vermögen, seine Fehler zu entdecken und zu korrigieren, sowie die Fähigkeit, konzentriert zu arbeiten, entfaltet.

Auf dem Gebiet der Rechentechnik entfernt sich unsere Oberschule vom Leben. Zu den primitiven arithmetischen Fertigkeiten fügt sie lediglich die Fähigkeit hinzu, sich der logarithmischen Rechenmethoden zu bedienen. Inzwischen hat aber die Benutzung der Logarithmen schon längst die universelle Bedeutung verloren, die ihr im 18. und 19. Jahrhundert zukam. An die Stelle der Logarithmentafeln traten der logarithmische Rechenschieber, der im letzten halben Jahrhundert sich überall dort eingebürgert hat, wo eine Genauigkeit der Rechnungen von drei bis vier Stellen ausreicht, ferner die verschiedenen Rechenmaschinen, mit deren Hilfe man Resultate mit größerer Genauigkeit erreichen kann, sowie Nomogramme, die Rechnungen nach bestimmten Formeln erstaunlich beschleunigen.

Es gibt ferner eine Reihe von mathematischen Tafeln, die bereits dem Schüler der Grundschule vollkommen zugänglich sind und die die alltägliche Rechenarbeit wesentlich erleichtern, die aber tatsächlich in der Schule sehr selten verwendet werden (Tafeln der Quadrate, Kuben, Quadrat- und Kubikwurzeln, der Reziproken, der Kreisbögen und Kreisflächen u. a. m.).

Aber noch in einem anderen wesentlichen Punkt löst sich das schulmäßige Rechnen vom Leben. In der Schule werden ausschließlich Operationen an vorgegebenen, genau fixierten Zahlen gelehrt, während man es in der Praxis in der überwiegenden Mehrzahl von Fällen mit Zahlen zu tun hat, die nur angenähert genau den Wert einer uns unbekannten realen Größe darstellen. So kann man die Anzahl der Gegenstände einer Menge nur dann genau auszählen, wenn sie relativ klein ist, während das Auszählen einer mehr oder weniger umfangreichen Menge von Gegenständen häufig ernsthafte, oft unüberwindliche Schwierigkeiten bietet und man sich zufrieden geben muß, wenn man einen angenäherten Wert für die Anzahl der Gegenstände einer solchen Menge angeben kann; und was die Messungen betrifft, so liefern diese stets nur Näherungswerte für die gemessenen Größen. In diesem Zusammenhang tauchen nun unvermeidlich die folgenden Fragen auf: Wie kann man in einem vorliegenden Fall die Genauigkeit einer gegebenen Näherungszahl (d. h. einer Zahl, die näherungsweise den genauen Wert einer betrachteten Größe ausdrückt) abschätzen? Wie schätzt man die Genauigkeit des Resultates einer Rechnung mit Näherungszahlen ab? Von welcher Genauigkeit müssen vorgegebene Näherungszahlen sein, damit das Resultat einer Rechnung mit ihnen eine gewisse verlangte Genauigkeit besitzt? Wie führt man am zweckmäßigsten die Grundrechenarten an Näherungszahlen aus?

Auf alle diese Fragen gibt die Schulmathematik keine Antwort. In diesen Punkten scheint eine grundlegende Revision der Unterrichtspläne, die den reinen Rechenunterricht betreffen, unumgänglich zu sein, und es bleibt zu hoffen, daß eine solche Revision nicht mehr lange auf sich warten läßt. Aber auch im Rahmen der gegenwärtig geltenden Unterrichtspläne kann der Lehrer noch viel für die Verbesserung der Rechenkultur seiner Schüler tun. Grundbedingung hierfür ist allerdings, daß er selbst die entsprechenden Kenntnisse und Fertigkeiten besitzt.

Der vorliegende Artikel soll einen kurzen Überblick über das dafür wesentliche Material geben. Die folgenden drei Paragraphen behandeln das genaue Rechnen,

alle weiteren die Näherungsrechnung. Der Umfang des Artikels zwingt uns dazu, uns stets nur auf die prinzipielle Seite der gestreiften Fragen und auf ein Minimum von erläuternden Beispielen zu beschränken. Der Leser, der sich mit den Einzelheiten vertraut machen möchte, sei auf die am Schluß dieses Artikels angegebene Literatur verwiesen. Hinweise im Text auf diese Literatur erfolgen in eckigen Klammern.

§ 2. Das Kopfrechnen

Die heute üblichen Verfahren zur Durchführung der arithmetischen Grundrechenarten an mehrstelligen, im Dezimalsystem dargestellten natürlichen Zahlen laufen letzten Endes auf die Verwendung von auswendig gelernten Additions- und Multiplikationstabellen der einstelligen Zahlen hinaus. Sie setzen das Aufschreiben der vorgegebenen Zahlen, der erhaltenen Zwischenergebnisse und des Endresultates voraus. Eine Rechnung heißt *Kopfrechnung*, wenn sie ohne jegliches Schreiben durchgeführt wird. Die Fertigkeit in dieser Art des Rechnens ist von größtem Wert, und zwar sowohl in rein praktischer Hinsicht, da es im täglichen Leben unvergleichlich häufiger benötigt wird als das schriftliche Rechnen, als auch im Hinblick auf die Ausbildung jener Fähigkeiten, die durch den Mathematikunterricht im allgemeinen gepflegt werden sollen: die Auffassungsgabe, die Aufmerksamkeit, die Initiative usw. Während die schriftlichen Berechnungen nach streng definierten Regeln ausgeführt werden und sich im beträchtlichen Maße als mechanische Arbeit erweisen, läßt das Kopfrechnen einen großen Spielraum für die Findigkeit und die Beobachtungsgabe und stellt ernsthafte Anforderungen an die Aufmerksamkeit sowie an die Fähigkeit der Selbstkontrolle. Eine Mittelstellung zwischen Kopfrechnen und schriftlichem Rechnen nimmt das sogenannte „halbschriftliche" Rechnen ein, bei dem nur die vorgegebenen Zahlen und das Endresultat aufgeschrieben werden. Da sich jedoch keine scharfe Grenze zwischen Kopfrechnen und halbschriftlichem Rechnen ziehen läßt, wollen wir im folgenden zwischen beiden nicht unterscheiden.

Üblicherweise eignen sich die Schüler in den ersten vier Klassen der Grundschule die einfachsten Verfahren des Kopfrechnens mit natürlichen Zahlen bis 100 und mit Brüchen der Nenner 2, 4, 8, 3, 6, 10 an. Später werden diese Fertigkeiten jedoch kaum weiter entwickelt, ja nicht einmal verwertet, so daß sie häufig wieder schwinden. Manchmal sehen wir in der Oberschule Schüler nach Bleistift und Papier greifen, wenn sie die Summe zweier zweistelliger natürlicher Zahlen finden sollen. Bezüglich des Kopfrechnens erscheint es zweckmäßig, die folgenden drei Forderungen in den Vordergrund zu stellen:

a) Es ist notwendig, die Schüler aller Klassen der Oberschule dazu anzuhalten, alle Berechnungen, die sich ohne ungewöhnliche Anstrengung bewältigen lassen, im Kopf durchzuführen. Hierzu gehören in erster Linie die vier Grundrechenarten an Zahlen bis etwa 100. Die größte Schwierigkeit bereitet hier sicherlich die Multiplikation zweistelliger Zahlen, die sich aber ebenfalls bei einiger Übung mit der bekannten Regel der „Multiplikation über kreuz" bewältigen läßt. Wenn man z. B. 53 mit 74 multiplizieren will, so bildet man zunächst das Produkt der Einer ($3 \cdot 4 = 12$); die Ziffer 2 schreibt man hin und behält einen Zehner im Sinn; weiter

bestimmt man das Produkt aus Zehner und Einer ($5 \cdot 4 = 20$), Einer und Zehner ($3 \cdot 7 = 21$), bildet die Summe $1 + 20 + 21 = 42$ (Zehner) und schreibt die Ziffer 2 links neben die bereits hingeschriebene Ziffer 2, während man die 4 (Hunderter) im Sinn behält; schließlich bildet man das Produkt der Zehner ($5 \cdot 7 = 35$) und dann die Summe $4 + 35 = 39$, die man wiederum links von der früher erhaltenen Zahl 22 niederschreibt, so daß man als Endresultat die Zahl 3922 erhält.

Wie überall muß jedoch auch hier ein vernünftiges Maß eingehalten werden.

b) Man sei bemüht, in jeder Art und Weise verschiedenartige, besondere Kopfrechenverfahren zu pflegen, die von den individuellen Eigenschaften der Zahlen, aber auch den kommutativen, assoziativen und distributiven Gesetzen sowie von gesetzmäßigen Änderungen des Ergebnisses bei Abänderung der gegebenen Zahlen Gebrauch machen. Derartige Verfahren sind erheblich ökonomischer (nicht nur hinsichtlich der verwendeten Kreide, sondern auch bezüglich der aufgewendeten Kopfarbeit) als die allgemeinen Verfahren. Hier einige Beispiele (die ausführliche Schreibweise dient nur zur Erläuterung des Verfahrens!):

$$32704 \cdot 25 = 3270400 : 4 = 817600,$$

weil nämlich

$$a \cdot 25 = (a \cdot 100) : 4$$

ist.

$$47 + 48 + 49 + 50 + 51 + 52 + 53 + 54$$
$$= (47 + 54) + (48 + 53) + (49 + 52) + (50 + 51) = 101 \cdot 4 = 404.$$

$$735 + 99 = 735 + 100 - 1 = 835 - 1 = 834.$$

$$15\tfrac{1}{2} + 32\tfrac{7}{16} + 1\tfrac{1}{3} + 2\tfrac{1}{6} = (15 + 32 + 1 + 2) + \left(\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{6}\right) + \tfrac{7}{16}$$
$$= 50 + 1 + \tfrac{7}{16} = 51\tfrac{7}{16}.$$

$$14\tfrac{3}{4} \cdot 8 = \left(14 + \tfrac{3}{4}\right) \cdot 8 = 14 \cdot 8 + \tfrac{3}{4} \cdot 8 = 112 + 6 = 118.$$

$$135\tfrac{3}{7} : 15 = \left(135 + \tfrac{3}{7}\right) : 15 = 135 : 15 + \left(\tfrac{3}{7} : 3\right) : 5 = 9 + \tfrac{1}{7} : 5 = 9\tfrac{1}{35}.$$

c) Man ergänze laufend den Vorrat an besonderen Verfahren, indem man — in dem Maße, in dem sie bekannt werden — verschiedene Formeln der Algebra praktisch auswertet, aber auch indem man neue und praktisch wertvolle Formeln entwickelt:

$$54 \cdot 46 = (50 + 4) \cdot (50 - 4) = 50^2 - 4^2 = 2484.$$
$$97^2 = (100 - 3)^2 = 10000 - 600 + 9 = 9409.$$

Wenn $b + c = 10$ ist, so ist

$$(10a + b)(10a + c) = 100a(a + 1) + bc$$

und dementsprechend

$$9993 \cdot 9997 = 100 \cdot 999 \cdot 1000 + 3 \cdot 7 = 99900021,$$
$$82 \cdot 88 = 100 \cdot 8 \cdot 9 + 16 = 7216.$$

Wegen

$$(10+a)(10+b) = 100 + 10(a+b) + ab$$

ist z. B.

$$13 \cdot 16 = 100 + 10 \cdot 9 + 18 = 208,$$

oder anders

$$13 \cdot 16 = 16 \cdot 10 + 16 \cdot 3 = 160 + 48 = 208.$$

$$\sqrt{74^2 - 70^2} = \sqrt{144 \cdot 4} = 12 \cdot 2 = 24.$$

Zum Abschluß dieses Paragraphen erwähnen wir noch eine Regel, an die sich mit großem Nutzen für die Rechenfertigkeit bereits eine ganze Reihe von Lehrern halten: Die Schüler aller Klassen haben bei jeder Zahlenrechnung mit einer grob genäherten Abschätzung des gesuchten Ergebnisses zu beginnen, wobei sie alle gegebenen Zahlen auf ein bis zwei Stellen runden und sämtliche Grundrechenarten im Kopf ausführen (Überschlagsrechnung). Wenn also z. B.

$$x = \sqrt{0{,}0045 \cdot 7{,}5132 : (2{,}0719 \cdot 0{,}864)}$$

berechnet werden soll, so findet man zunächst im Kopf

$$x \approx \sqrt{0{,}004 \cdot 8 : (2 \cdot 0{,}9)} = \sqrt{0.016 : 0{,}9} = \sqrt{0{,}16 : 9} = 0{,}4 : 3 \approx 0{,}13$$

und hinterher schriftlich den genaueren Wert (wenn man sich etwa auf vier Stellen beschränkt, $x \approx 0{,}1374$). Derartige Abschätzungen verhüten wesentlich grobe Rechenfehler und werden von Angehörigen ingenieur-technischer Berufe sehr geschätzt.

Wer sich eingehender mit den verschiedenen Kopfrechenverfahren beschäftigen möchte, sei auf die Bücher [1], [2a], [3] und die Arbeit [4] verwiesen.

§ 3. Das schriftliche Rechnen

Hinsichtlich des schriftlichen Rechnens, sowohl mit ganzen als auch mit gebrochenen Zahlen, steht es in der Schule wesentlich günstiger als hinsichtlich des Kopfrechnens. Hier stellen die gegenwärtig in der Schule gelehrten Algorithmen für die Grundrechenarten an mehrstelligen Zahlen (Schemata der Durchführung und des Aufschreibens) die besten unter vielen, zu verschiedenen Zeiten vorgeschlagenen Varianten dar, die auch in der Mehrzahl der Fälle bereits in der Grundschule hinreichend sicher erlernt werden. Der Mathematikunterricht an den Grund- und Oberschulen bietet genügend Anwendungsmöglichkeiten für sie und garantiert damit die Erhaltung der technischen Fertigkeiten. Es sollen jedoch auch hier einige Wünsche methodischer Art hinsichtlich der schriftlichen Durchführung der Grundrechenarten an mehrstelligen Zahlen ausdrücklich erwähnt werden:

a) Während man die Schüler bei der Wiederholung der Arithmetik oder in den Anfängen der Algebra mit den kommutativen, assoziativen und distributiven Gesetzen der Addition und Subtraktion bekannt macht, erläutere man, wie diese Eigenschaften bei den üblichen Algorithmen der Addition und Subtraktion mehrstelliger Zahlen verwendet werden; ist es doch außerordentlich wichtig, daß die Schüler nicht nur den Mechanismus der Algorithmen einwandfrei beherrschen

sondern auch ihre theoretische Basis verstehen. Diese Frage läßt sich naturgemäß bei der erstmaligen Einführung der Grundrechenarten nicht vollständig klären. Besonders lehrreich ist die Behandlung dieser Verhältnisse bei der Multiplikation und Division.

b) Wenngleich man immer und überall eine sorgfältige Niederschrift verlangen wird, muß man dieser bei der Durchführung der Grundrechenarten an mehrstelligen Zahlen besondere Aufmerksamkeit zuwenden. Es darf insbesondere auch keine nachlässige Niederschrift der Berechnungen im Konzept geduldet werden, da diese eine beständige Fehlerquelle darstellt, indem die Fehler bei der Übertragung ins Reine mechanisch wiederholt werden.

c) Wenn man von den Schülern verlangt, daß sie keine Rechenfehler begehen, so muß man sie auch mit vernünftigen Methoden zur Kontrolle ihrer Berechnungen vertraut machen. Keine Rechnung darf als erledigt angesehen werden, ehe nicht das Resultat durch irgendein Kontrollverfahren überprüft ist. Die Addition kontrolliert man durch die Addition, allerdings bei anderer Reihenfolge der Summanden; die Subtraktion kontrolliert man durch die Addition; die Multiplikation kontrolliert man durch die Multiplikation (mit umgestellten Faktoren); die Division kontrolliert man durch die Multiplikation des Divisors mit dem Quotienten und — falls vorhanden — Addition des Restes; das Radizieren kontrolliert man durch das Potenzieren.

Sehr nützlich ist auch die Kontrolle des Ergebnisses mit der Neuner- oder Elferprobe. Dabei werden die vorgegebenen Zahlen durch ihre Reste bei der Division durch 9 bzw. 11 ersetzt; nimmt man dann an den Resten die in Frage stehenden Grundrechenarten vor, so muß man im Fall der Fehlerlosigkeit der kontrollierten Rechnung zu Zahlen kommen, die bei der Division durch 9 bzw. 11 denselben Rest lassen wie das erhaltene Resultat. Die Zahlen 9 und 11 nimmt man deshalb, weil sich der Rest der Division einer gegebenen Zahl durch 9 bzw. 11 besonders einfach bestimmen läßt: Der Rest bei der Division einer beliebigen, im Dezimalsystem angegebenen Zahl durch 9 stimmt nämlich überein mit dem Rest bei der Division ihrer Quersumme durch 9. So ist der Rest bei der Division der Zahl 4138097 durch 9 gleich dem Rest bei der Division der Zahl $4+1+3+8+0+9+7=32$ durch 9 und dieser wiederum gleich dem Rest bei der Division von $3+2=5$ durch 9. Entsprechend erhält man den Rest bei der Division einer gegebenen, im Dezimalsystem geschriebenen Zahl durch 11 mittels der „alternierenden" Quersumme[1]). Diese Summe wird aus den mit wechselnden Vorzeichen genommenen Ziffern der gegebenen Zahl gebildet, wobei man mit der Ziffer der Einer beginnt (damit das Auftreten von negativen Zahlen vermieden wird, kann man dabei jeweils 11 hinzufügen, wenn von einer Zahl eine größere abgezogen werden muß). Zum Beispiel ist der Rest bei der Division der Zahl 4138097 durch 11 gleich dem Rest bei der Division der Zahl $7-9+0-8+3-1+4+11=7$ durch 11. Bezüglich weiterer Einzelheiten über die Neunerprobe vgl. das Buch [5].

Grobe Rechenfehler können bereits dadurch vermieden werden, daß man oben die erwähnte, leicht im Kopf durchzuführende „Überschlagsrechnung" vornimmt, deren Durchführung vor jeder vollständigen Berechnung wir empfohlen haben.

[1]) Vgl. S. 244, A. J. Chintschin, Elemente der Zahlentheorie.

d) Wir wollen noch einige verbreitete Fehler erwähnen, die bei schriftlichen Rechnungen mit ganzen und gebrochenen Zahlen begangen werden:

Häufig werden bei der Division zwischen den Ziffern des Quotienten Nullen vergessen. Zum Beispiel findet man bei der Division von 3708 durch 18 den falschen Quotienten 26 an Stelle von 206. Es ist interessant, daß ähnliche Fehler bei der Durchführung der Rechnung im Kopf nicht angetroffen werden, wie in dem genannten einfachen Beispiel: $3708 : 18 = (3600 + 108) : 18 = 200 + 6 = 206$.

Mitunter verzerrt man bei der Division den Rest durch das Streichen von Nullen am Ende des Dividenden und des Divisors. Zum Beispiel ersetzt man bei der Division von 650 durch 110 diese Zahlen durch 65 und 11, erhält den richtigen Quotienten 5 und den falschen Rest 10 (an Stelle von 100).

Häufig werden ohne jede Veranlassung bei der Addition und der Subtraktion vorgegebene gemischte Zahlen in unechte Brüche umgewandelt. Zwar wird dadurch das Resultat nicht beeinflußt, jedoch werden die Rechnungen verwickelter.

Manchmal beschränkt man sich fälschlich bei der Multiplikation von gemischten Zahlen nur auf die Multiplikation der ganzen mit der ganzen Zahl und des Bruches mit dem Bruch. Auf diese Weise erhält man bei der Multiplikation von $5\frac{2}{3}$ mit $2\frac{3}{4}$ wegen $5 \cdot 2 = 10$ und $\frac{2}{3} \cdot \frac{3}{4} = \frac{1}{2}$ das falsche Ergebnis $10\frac{1}{2}$ an Stelle von $\frac{17 \cdot 11}{3 \cdot 4} = 15\frac{7}{12}$.

Es steht außer Frage, daß eine gründliche Prüfung dieser und ähnlicher Rechenfehler für den Schüler äußerst lehrreich ist.

e) Bei jeder komplizierten schriftlichen Rechnung ist eine zweckmäßige Niederschrift des Rechnungsganges von großer Bedeutung. Eine unordentliche, verstreute Niederschrift ist auch im Konzept unter allen Umständen unzulässig. Jede Eintragung muß so klar sein, daß sich nicht nur der Rechner selbst, sondern auch jeder andere unmittelbar zurechtfindet. Die Niederschrift muß so vollständig wie notwendig und so knapp wie möglich sein, sie muß — kurz gesagt — alles Notwendige und nichts Überflüssiges enthalten. Sehr nützlich ist die Aufstellung eines vorläufigen Schemas, d. h. einer Einteilung auf einem Blatt Papier, die Auskunft über den ganzen Verlauf der Rechnung gibt, wobei jedes Zwischenergebnis einen wohlbestimmten Platz einnimmt; alle Nebenrechnungen, die nicht im Kopf unmittelbar verfolgt werden können, werden an einer besonderen Stelle ausgeführt. Wünschenswert sind Hinweise auf die bei den Berechnungen verwendeten Hilfsmittel (Tafeln, Geräte, Zeichnungen). Den Abschluß einer jeden Berechnung hat eine Probe zu bilden; aber auch Überlegungen über die Genauigkeit des erhaltenen Resultates sind unumgänglich, da in der überwiegenden Mehrzahl der Fälle — auch bei genau vorgegebenen Zahlen — das Endergebnis nur angenähert zu sein pflegt (hiermit werden wir uns im folgenden noch ausführlich beschäftigen). Beim Übertragen der ausgeführten Rechnung ins Reine empfiehlt es sich, jedes Zwischenergebnis erneut zu prüfen, indem man zur Kontrolle jede Rechnung wiederholt und das neue Ergebnis mit dem Konzept vergleicht.

Hier ein Beispiel für die Niederschrift der Lösung einer Aufgabe (9. Klasse):

Aufgabe: Man bestimme die Länge x der Seite des Quadrates, das dem Dreieck mit den Seiten $a = 89$ m, $b = 321$ m, $c = 395$ m flächengleich ist.

Graphische Lösung (Maßstab: 1 Einheit $= 20$ m). Nachdem wir aus den drei gegebenen Seiten das Dreieck ABC konstruiert haben (Abb. 1), zeichnen wir $CD \perp AB$ und tragen auf der Verlängerung von AB die Strecke $BE = 0{,}5\ CD$ ab; sodann konstruieren wir über der Strecke AE den Halbkreis mit dem Durchmesser AE; die Halbsehne GB, die senkrecht auf AE steht, stellt die gesuchte Länge dar. Ihre Länge beträgt in der Zeichnung 4,8 Einheiten, so daß $x \approx 96$ m ist. Die letzte Stelle ist bereits unsicher, da eine Ungenauigkeit von GB um 0,1 Einheiten, d. h. eine Ungenauigkeit des angenäherten Wertes von x um 2 m durchaus möglich ist.

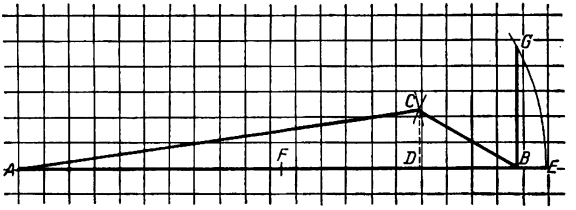


Abb. 1

Rechnerische Lösung. Wir verwenden die Formeln

$$s^2 = p(p-a)(p-b)(p-c);\quad 2p = a+b+c;\quad x = \sqrt{s}.$$

Berechnung mittels vierstelliger Logarithmentafel:

a	89
b	321
c	395
$2p$	805
p	402,5
$p-a$	313,5
$p-b$	81,5
$p-c$	7,5
$(p-a)+(p-b)+(p-c) =$	402,5
$= 3p-2p = p$	(Probe)

$\lg p$	2,6047
$\lg (p-a)$	2,4962
$\lg (p-b)$	1,9112
$\lg (p-c)$	0,8751
$\lg s^2 = \lg x^4$	7,8872
$\lg x$	1,9718
x	93,72

Ergebnis: $x \approx 93{,}72$.

Sind die vorgegebenen Zahlenwerte genau, so kann man in der Lösung für die ersten drei Stellen garantieren, jedoch ist die vierte Stelle wegen der Verwendung vierstelliger Logarithmen nicht völlig sicher. Es ergibt sich eine befriedigende Übereinstimmung zwischen der graphischen und der rechnerischen Lösung.

Eine Kontrollrechnung mit siebenstelligen Logarithmen ergibt: $x \approx 93{,}71424$.

Dem Verlangen nach einer rationellen Niederschrift jeder Berechnung und jeder Lösung von Aufgaben wird in der Schule eine viel zu geringe Beachtung geschenkt. Es ist unbedingt notwendig, dem Schüler gut durchdachte Muster von Lösungsgängen vorzuführen, aber keineswegs zu fordern, daß er diese blind nachahmt, sondern ihn zu weiteren Verbesserungen an ihnen anzuhalten.

§ 4. Hilfsmittel für das Rechnen

Gegenwärtig haben bereits eine ganze Anzahl von Geräten und Maschinen eine weite Verbreitung gefunden, die zahlreiche mathematische Operationen automatisch oder halbautomatisch ausführen. Das beginnt bei der Addition (und Subtraktion) mehrstelliger Zahlen, die mit vollem Erfolg auf allen bekannten Rechenbrettern ausgeführt werden kann, und geht bis zur Durchführung der sehr komplizierten Berechnungen der harmonischen Analyse und der Integration partieller Differentialgleichungen mittels der großen Rechenautomaten, von denen jeder eine Vielzahl qualifizierter Rechner ersetzt. Darf man angesichts dieser Tatsache in der Schule die äußerst nützliche Mechanisierung des Rechnens völlig ignorieren? Sogar im Rahmen der Grundschule ist es möglich, drei noch näher zu betrachtende einfache Hilfsmittel des mechanischen Rechnens zu benutzen (Rechenbrett, Nepersche Streifen, Sprossenradmaschine). Mit dem Gebrauch des logarithmischen Rechenschiebers, der heute zum unentbehrlichen Werkzeug jedes Ingenieurs und Technikers gehört, müssen die Schüler der 9. und 10. Klasse eingehend vertraut gemacht werden. Wünschenswert ist, daß der Lehrer selbst ihn beständig, sowohl bei seiner Hausarbeit als auch vor den Augen der Schüler aller Klassen, benutzt. Auf dieses Gerät werden wir in § 17 ausführlich zu sprechen kommen.

Das Rechenbrett ist ein schönes und bequemes Hilfsmittel für die Addition und die Subtraktion mehrstelliger Zahlen. Solange nur mit diesen beiden Grundrechenarten gerechnet wird, konkurriert es mit Erfolg — was die Schnelligkeit der Rechnung betrifft — mit fast allen anderen gebräuchlichen Rechenmaschinen. Das kommt daher, daß man bei der Addition und Subtraktion auf Rechenmaschinen jede der gegebenen Zahlen einzeln einzustellen hat, um sie danach durch Drehen der Kurbel zu addieren oder zu subtrahieren, während auf dem Rechenbrett nur die erste Zahl eine besondere Einstellung erfordert. Allerdings ist das Rechenbrett nicht vollautomatisiert; insbesondere erfordern die Übertragung der Zehner, d. h. der Ersatz von jeweils zehn Steinen, die sich auf einer Speiche angehäuft haben, durch einen Stein der nächsten Speiche und der entsprechende Übergang bei der Subtraktion die Aufmerksamkeit des Rechners; sie stellen daher eine Hauptfehlerquelle dar. Jedoch verkürzt der Gebrauch des Rechenbrettes die für die Durchführung mehrerer Additionen und Subtraktionen aufzuwendende Zeit derart, daß es wegen seines einfachen Aufbaus und seiner einfachen Bedienungsweise weite Verbreitung gefunden hat. In der Sowjetunion müssen laut Lehrplan alle Absolventen der Grundschule die Fertigkeit besitzen, an Rechenbrettern zu arbeiten. In der Oberschule wird das Rechenbrett im allgemeinen nicht mehr benutzt, jedoch wäre es wünschenswert, daß man die Fertigkeit im Rechnen auf dem Rechenbrett auch in der Oberschule pflegt, zumal der beständige Gebrauch des Rechenbretts eine bemerkenswerte Einsparung an Zeit mit sich bringt, die man zur Lösung von Aufgaben, besonders in der 5. Klasse, verwenden könnte. Auch die Multiplikation in Form aufeinanderfolgender Additionen und die Division in Form aufeinanderfolgender Subtraktionen lassen sich mit Erfolg auf dem Rechenbrett ausführen (dabei beachte man in besonderen Fällen die jeweiligen Vereinfachungen: so kann man z. B. das Produkt aus 365 und 17 auf dem Rechenbrett einfach dadurch berechnen, daß man $3650 + 3650 - 365 - 365 - 365$ bildet). Anfangs sollte man sich jedoch auf die Addition und Subtraktion beschränken.

In dem gleichen unvollständigen, aber doch bedeutungsvollen Maße, in dem das Rechenbrett die Addition und die Subtraktion mechanisiert, kann man die Multiplikation und die Division durch ein anderes, allerdings weniger verbreitetes, einfaches Gerät, die NEPERschen Streifen, mechanisieren. Das ist ein Satz Streifen, wie er in Abb. 2 dargestellt ist, wobei von jedem Streifen einige Exemplare vorhanden sein müssen. Die Streifen tragen oben die Ziffern von 0 bis 9 und darunter die Produkte dieser Zahlen mit allen einstelligen Zahlen. Dabei wird in jedem Produkt die Zehnerziffer links etwas oberhalb der Einerziffer hingeschrieben und durch einen schrägen Strich abgetrennt. Will man nun z. B. die Zahl 37214 mit einer beliebigen anderen natürlichen Zahl multiplizieren, so legt man die Streifen, die oben die Ziffern 3, 7, 2, 1, 4 tragen, der Reihe nach nebeneinander, wie Abb. 3 zeigt. Dann

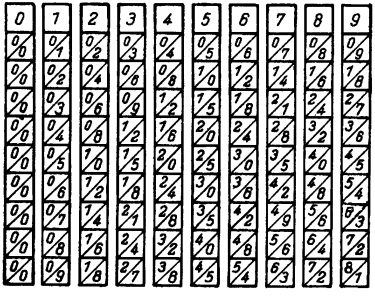


Abb. 2

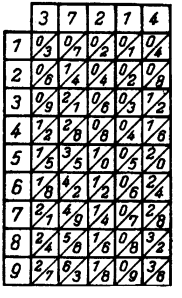


Abb. 3

kann man das Produkt der vorgegebenen Zahl mit jeder der Zahlen 1, 2, . . ., 9 unmittelbar ablesen, wobei die Übertragung der Zehner jeweils im Kopf vorgenommen wird. Zunächst müssen daher die Produkte von rechts nach links gelesen werden. So erhält man z. B. das Produkt 37214 mal 7 in der Form: 8 Einer, $7 + 2 = 9$ Zehner, $4 + 0 = 4$ Hunderter, $9 + 1 = 10$ Tausender, $1 + 1 + 4 = 6$ Zehntausender, 2 Hunderttausender. Bei einiger Übung wird man aber das Produkt unmittelbar auch von links nach rechts ablesen können. Auf diese Weise erhält man mittels der NEPERschen Streifen zunächst das fertige Produkt jeder beliebigen Zahl mit allen einstelligen Zahlen. Will man mehrstellige Zahlen mit mehrstelligen Zahlen multiplizieren, so legt man für den Multiplikanden die Streifen hin, multipliziert diesen in angegebener Weise mit den jeweiligen Stellen des Multiplikators und addiert in passender Weise. Indem man die NEPERschen Streifen mit dem Rechenbrett kombiniert, kann man die Multiplikation so weit mechanisieren, daß das Ergebnis fertig erhalten wird.

Unvergleichlich weitergehend als bei Rechenbrett und NEPERschen Streifen ist die Mechanisierung des Rechnens bei den sogenannten Sprossenradmaschinen. Wir geben im folgenden eine Beschreibung der in der Sowjetunion hergestellten Maschine „Felix", so genannt zu Ehren von F. E. DZIERZYNSKI. Diese Maschine ist in Abb. 4 dargestellt. Sie stellt eine Weiterentwicklung einer Maschine dar, die gegen

Ende des 19. Jahrhunderts von dem Petersburger Ingenieur W. T. ODHNER gebaut wurde. Die Sprossenradmaschine hat (auch bei uns — *Anm. d. wissenschaftl. Red.*) eine weite Verbreitung gefunden und ist gegenwärtig in fast allen Büros zu finden. Die Aneignung des Bauprinzips und der Arbeitsweise der Sprossenradmaschine erfordert im ganzen etwa 15 bis 20 Minuten, und es wäre wünschenswert, daß alle Absolventen der Oberschule mit ihrer Bedienung vertraut gemacht werden. Während der Gebrauch der NEPERschen Streifen die Durchführung der Multiplikation und Division etwa in der halben Zeit ermöglicht, braucht man bei Benutzung der Sprossenradmaschine nur ungefähr den 10. bis 12. Teil der Zeit (im Vergleich zur Durchführung der Grundrechenarten auf dem Papier). Eine Berechnung, die ohne Verwendung der Sprossenradmaschine eine ganze Stunde Arbeit beansprucht, kann also mit ihrer Hilfe in 5 bis 6 Minuten bewältigt werden, und zwar mit einer unvergleichlich geringeren Gefahr eines Fehlers.

Im folgenden geben wir die aus dem Buch [2a] des Autors entnommene Beschreibung des Bauprinzips und der Arbeitsweise der Maschine „Felix" wieder.

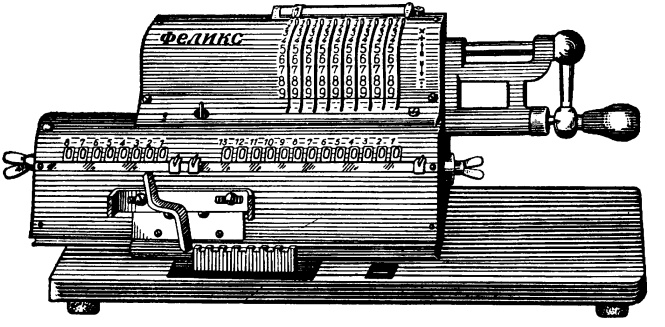


Abb. 4

Auf der Vorderseite der Maschine befindet sich das *Einstellwerk*. In Abb. 4 sind neun Handgriffe sichtbar, die dem Einstellmechanismus angehören und die in neun Einschnitten gleiten. Auf der linken Seite eines jeden Einschnittes sind von oben nach unten die Ziffern von 0 bis 9 angeordnet. Durch Schieben der Handgriffe kann man im Einstellwerk jede neunstellige ganze oder gebrochene Dezimalzahl einstellen. Beim Einstellen von Brüchen bedient man sich eines metallischen „Komma", das oben zwischen zwei benachbarte Einschnitte geschoben werden kann. Rechts neben dem Einstellwerk befindet sich eine Kurbel mit Handgriff. Diese läßt sich nur drehen, wenn man zuvor den Handgriff etwas nach rechts herausgezogen hat. Nach jeder für die jeweilige Rechnung erforderlichen Anzahl von Umdrehungen ist der Handgriff der Kurbel unbedingt wieder in die „Normallage" zu bringen, in der er auf der Zeichnung dargestellt ist. Unterhalb des Einstellwerkes befindet sich ein *Schlitten*, der mit zwei Reihen von Schaulöchern ausgestattet

ist: Rechts in der Abbildung sieht man 13 Schaulöcher, die wir *Resultatwerk* nennen wollen, da in ihnen die Ergebnisse der Addition, Subtraktion und Multiplikation erscheinen; links sieht man 8 Schaulöcher, die wir *Umdrehungszählwerk* nennen wollen, da in ihnen die Triebwellenumdrehungen gezählt werden (statt Umdrehungszählwerk spricht man mitunter auch von *Quotientenwerk*, da in ihm die Ergebnisse bei der Division abgelesen werden — *Anm. d. wissenschaftl. Red.*). Rechts und links am Schlitten befindet sich je eine *Löschvorrichtung*. Durch Umlegen der Hebel werden die im Resultatwerk bzw. im Umdrehungszählwerk stehenden Ziffern durch Nullen ersetzt (gelöscht). Unterhalb der Schaulöcher gleiten auf einer Leiste metallene Kommata. Unter dieser Leiste befindet sich eine Vorrichtung (der Transporter), mit deren Hilfe man den Schlitten jeweils um eine oder auch um mehrere Intervalle zwischen benachbarten Einschnitten des Einstellwerkes verschieben kann.

Vor jeder Umdrehung der Triebwelle muß man sich davon überzeugen, ob beide Hebel der Löschvorrichtung in horizontaler Richtung stehen (wenn man diese Lage erreicht hat, rastet die Löschvorrichtung hörbar ein) und ob die Leiste des Transporters in einen der Zwischenräume zwischen den Zähnen des darunterliegenden Zahnkammes eingerastet ist (auch das geschieht hörbar). Wenn man diese Anweisung nicht beachtet, so läßt sich das *Triebwerk* nicht drehen. Versucht man dennoch mit Gewalt die Kurbel zu drehen, so tritt unverzüglich Maschinenschaden ein. Links vom Einstellwerk ist auf der Höhe der unteren Einschnittenden ein Knopf sichtbar. Er hat die Aufgabe, die Einstellung der Handgriffe des Einstellwerkes auf die Nullage zu beschleunigen: Wenn man diesen Knopf nach links drückt und gleichzeitig die Kurbel vorsichtig auf sich zu dreht, so werden nach Rückkehr der Kurbel in Normallage alle Handgriffe auf Null stehen. Man muß nur daran denken, daß nach etwa ein Drittel Umdrehung der Kurbel, wenn alle Handgriffe „ausgerichtet" sind, der Knopf losgelassen werden muß. Rechts vom Einstellwerk sind auf dem Mantel der Maschine zwei entgegengesetzt gerichtete Pfeile sichtbar, die mit den Zeichen für die Grundrechenarten versehen sind (+ und × nach oben, — und : nach unten). Diese Pfeile zeigen die Richtung an, in der das Triebwerk bei Durchführung der jeweiligen Operation zu drehen ist. Diese Richtungen werden wir *positiv* (Kurbel aus der Normallage auf den Rechner zu) und *negativ* (vom Rechner weg) nennen. Damit haben wir alle Teile der Maschine genannt, mit denen man beim Rechnen unmittelbar zu tun hat. Wir können hier auf ihre innere Konstruktion nicht näher eingehen, wollen jedoch darauf hinweisen, daß ihr Hauptbestandteil das in Abb. 5 dargestellte *Sprossenrad* (ODHNERsches Rad) ist. Dieses Rad besitzt eine veränderliche Anzahl nach außen tretender Zähne, und zwar treten jeweils so viele Zähne nach außen, wie die Ziffer angibt, auf die der Handgriff im Einschnitt eingestellt ist. In Abb. 5 treten aus dem Rad sechs Zähne hervor, während die anderen eingezogen sind. Wenn wir den Handgriff im Einschnitt bewegen, so verändern wir die Anzahl der Zähne auf dem Zahnrad, das mit diesem Handgriff versehen ist. In jeder Sprossenradmaschine befinden sich so viele Sprossenräder wie Handgriffe.

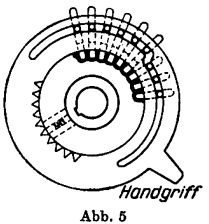


Abb. 5

Wenn wir im ersten (äußersten rechten) Einschnitt den Handgriff auf die Ziffer 3 stellen und eine Umdrehung des Triebwerkes in positiver Richtung vollführen, so

drehen wir damit ein Rad um drei Zähne, auf dessen Reifen Ziffern angebracht sind, die durch das erste (äußerste rechte) Schauloch zu sehen sind. Wenn vorher in diesem Schauloch die Ziffer 0 zu sehen war, so erscheint jetzt die Ziffer 3. Eine zweite Umdrehung des Triebwerkes in gleicher Richtung dreht dieses Rad um drei weitere Zähne und statt der Ziffer 3 erscheint im Schauloch die Ziffer 6: Wir haben die Addition $3 + 3$, oder was dasselbe ist, die Multiplikation $3 \cdot 2$ durchgeführt. Eine weitere Umdrehung des Triebwerkes liefert bereits $6 + 3 = 9$ oder $3 \cdot 3 = 9$. Bei der vierten Umdrehung gehen am ersten Schauloch die Ziffern 9, 0, 1, 2 vorüber (das Rad hat eine vollständige Umdrehung gemacht und beginnt mit der zweiten), und im zweiten Schauloch des Resultatwerkes erscheint die Ziffer 1. Hierbei hat sich die Tätigkeit des Mechanismus der Zehnerübertragung eingeschaltet, der den empfindlichsten Teil jeder Rechenmaschine darstellt. Als Resultat unserer Rechnung erhalten wir $9 + 3 = 12$ oder $3 \cdot 4 = 12$.

Jetzt ist es nicht mehr schwierig, den Verlauf der Grundrechenarten auf der betrachteten Maschine zu verstehen. Um zwei Zahlen zu addieren, muß man (1) im Resultatwerk Nullen einstellen (indem man die rechte Löschvorrichtung bedient), (2) im Einstellwerk den ersten Summanden einstellen (üblicherweise beginnt man dabei mit der letzten Ziffer am ersten, d. h. äußersten rechten Handgriff, was jedoch nicht unbedingt notwendig ist), (3) diesen Summanden in das Resultatwerk einführen (durch eine Drehung des Triebwerkes in positiver Richtung), (4) den zweiten Summanden in das Einstellwerk einführen, (5) durch eine weitere Umdrehung des Triebwerkes in positiver Richtung den zweiten Summanden in das Resultatwerk überführen; im Schauloch des Resultatwerkes erscheint dann die gesuchte Summe. Entsprechend wird in fünf Schritten die Subtraktion vorgenommen, wobei man lediglich nach der Einstellung des Subtrahenden (Schritt (4)) in Schritt (5) die Kurbel in negativer Richtung zu drehen hat. Es ist klar, daß man zu einer einmal erhaltenen Zahl noch nach Belieben weitere Zahlen addieren (oder von ihr subtrahieren) kann.

Die Multiplikation mit einer einstelligen Zahl wird als eine wiederholte Addition ausgeführt. Wenn man also z. B. eine Zahl mit 9 multiplizieren will, so dreht man das Triebwerk 9mal in positiver Richtung. Will man eine gegebene (im Einstell- und Resultatwerk eingestellte) Zahl mit einer zweistelligen Zahl, etwa mit 39, multiplizieren, so macht man von der Möglichkeit Gebrauch, den Schlitten bezüglich des oberen Teiles der Maschine, der den Einstellmechanismus enthält, zu verschieben: nachdem man den Schlitten mit Hilfe des Transporters um ein Intervall nach rechts verschoben hat, dreht man die Kurbel dreimal in positiver Richtung und erhält im Resultatwerk das Produkt der vorgegebenen Zahl mit 30. Dann braucht man nur noch den Schlitten in Normallage zu bringen — das erste Schauloch des Resultatwerkes befindet sich dabei wiederum unter dem ersten Einschnitt — und noch 9 Umdrehungen der Kurbel in positiver Richtung zu vollführen. Im Resultatwerk erscheint dann das Produkt der gegebenen Zahl mit 39.

Die Multiplikation einer gegebenen Zahl mit 39 erfordert in der eben beschriebenen Weise $3 + 9 = 12$ Umdrehungen der Kurbel; diese Zahl verringert sich auf 5, wenn wir bei nach rechts verschobenem Schlitten vier Umdrehungen ausführen (also mit 40 multiplizieren) und darauf — nachdem der Schlitten in Normalstellung gebracht wurde — eine Umdrehung in negativer Richtung ausführen. Auf Grund dieses Verfahrens, das in der Praxis meistens angewendet wird, braucht man die Kurbel bei jeder Schlittenstellung nicht mehr als 5mal in einer Richtung zu drehen.

So wird die Multiplikation mit 9, 8, 7, 6 durch die Multiplikation mit 10 und nachfolgende Subtraktion des 1-, 2-, 3-, 4fachen Multiplikanden ersetzt.

Die Anzahl der Umdrehungen des Triebwerkes wird in den Schaulöchern des Umdrehungszählwerkes (links) registriert. Bei der Durchführung einer Multiplikation wird vorher die linke Löschvorrichtung des Umdrehungszählwerkes betätigt, d. h., alle Ziffern der linken Schaulöcher werden auf Null gestellt. Dann kann man das Multiplikationsverfahren folgendermaßen beschreiben: Nachdem man im Einstellwerk den Multiplikanden eingestellt hat, kombiniert man die Bewegung des Schlittens und die Umdrehungen des Triebwerkes so, daß in den Schaulöchern des Umdrehungszählwerkes der Multiplikator auftaucht; dann zeigt das Resultatwerk das Produkt an. Dabei muß man jedoch beachten, daß beim Drehen der Kurbel in negativer Richtung im Umdrehungszählwerk rote Ziffern erscheinen, so daß man unter Umständen den Multiplikator in einer etwas seltsamen Form einstellt. So taucht bei der Multiplikation mit $39 = 40 - 1$ im Umdrehungszählwerk die Zahl $4\bar{1}$ auf, wobei wir hier durch $\bar{1}$ die rote Ziffer 1 bezeichnen. Bei der Multiplikation mit der Zahl 8376 stellen wir im Umdrehungszählwerk die Zahl $1\bar{2}4\bar{2}4$ ein, wobei wiederum $\bar{2}$ und $\bar{4}$ die roten Ziffern bezeichnen. Bei Benutzung der roten Ziffern haben wir im zuletzt betrachteten Fall $1 + 2 + 4 + 2 + 4 = 13$ Umdrehungen zu vollführen, während es bei ausschließlicher Benutzung von schwarzen Ziffern $8 + 3 + 7 + 6 = 24$ Umdrehungen wären.

Genau wie man mittels der Sprossenradmaschine die Multiplikation als wiederholte Addition ausführt, kann man mit ihr die Division als wiederholte Subtraktion ausführen. So heißt 17 durch 3 teilen, erkennen, wie oft man von 17 die Zahl 3 abziehen kann (bis man einen Rest erhält, der kleiner als der Divisor ist). Dementsprechend stellt man bei der Division den Dividenden im Resultatwerk und den Divisor im Einstellwerk ein und subtrahiert, bis im Resultatwerk eine Zahl steht, die kleiner als der Divisor ist. Im Umdrehungszählwerk kann man dann den Quotienten als Anzahl der vollführten Umdrehungen ablesen. Bei der Division mehrstelliger Zahlen kann man, genau wie bei der Multiplikation, eine Verringerung der notwendigen Umdrehungen durch Betätigung des Schlittens erreichen. Will man z. B. 243 558 durch 913 dividieren, so stellt man den Dividenden in die äußersten linken Schaulöcher des Resultatwerkes ein; alle anderen Schaulöcher des Resultatwerkes und alle Schaulöcher des Umdrehungszählwerkes müssen Nullen zeigen. Indem wir mittels des metallenen Kommas (entsprechend der Stellenzahl des Divisors) die ersten drei Stellen des Dividenden absondern, erhalten wir die Zahl 243; da diese Zahl kleiner als der Divisor ist, nehmen wir noch die vierte Stelle hinzu, d. h., wir sondern 2435 ab. Nachdem wir den Schlitten so weit wie möglich nach rechts verschoben haben, stellen wir den Divisor 913 so im Einstellwerk ein, daß er von 2435 abgezogen werden kann (also 9 unter 4). Dann nehmen wir die Subtraktion so lange vor, wie es möglich ist, d. h., bis wir als Rest eine Zahl erhalten die kleiner als der Divisor ist. Im Beispiel finden wir nach zwei Umdrehungen (in negativer Richtung) als Rest die Zahl 609. Dann verschieben wir den Schlitten um eine Stelle nach links und das Komma um eine Stelle nach rechts und wiederholen die Operation. Nach sechs Umdrehungen erhalten wir im Beispiel den Rest 617. Darauf rücken wir den Schlitten nochmals um eine Stelle nach links und das Komma um eine Stelle nach rechts und wiederholen die fortlaufende Subtraktion, wobei wir im Beispiel nach sechs Umdrehungen den Rest 700 erhalten. Damit ist im Bei-

spiel die Division in ganzen Zahlen beendet. Als Quotient ergibt sich im Umdrehungszählwerk 266 und als Rest im Resultatwerk 700. Setzt man die Division in angegebener Weise weiter fort, so erhält man Zehntel, Hundertstel usw. des Quotienten. Dabei braucht man bei der Division nicht auf die hintereinander erscheinenden Reste zu achten. Man braucht die Kurbel nur bis zum Ertönen eines Glockenzeichens zu drehen, das die erste überschüssige Umdrehung anzeigt, und dann diese überschüssige Umdrehung durch eine Umdrehung in positiver Richtung auszugleichen.

Schließlich wollen wir noch auf das Ausziehen der Quadratwurzel mittels der Sprossenradmaschine zu sprechen kommen.

Das am häufigsten dabei angewendete Verfahren beruht auf der leicht nachprüfbaren Identität

$$1 + 3 + 5 + 7 + \cdots + (2n - 3) + (2n - 1) = n^2,$$

nach der die Summe der ersten n ungeraden Zahlen gleich dem Quadrat von n ist. Um die Quadratwurzel einer gegebenen natürlichen Zahl zu berechnen, muß man also von ihr der Reihe nach die Zahlen 1, 3, 5, 7, ... abziehen, bis man als Rest eine Zahl erhält, die kleiner als der folgende Subtrahend ist. Die Anzahl der ausgeführten Subtraktionen ist dann gerade die gesuchte Quadratwurzel (genauer die Quadratwurzel aus der größten in der gegebenen Zahl enthaltenen Quadratzahl). Eine passende Verschiebung des Schlittens gestattet auch hier häufig, die Anzahl der Subtraktionen wesentlich zu verringern. Die Einzelheiten des Verfahrens wollen wir wieder an einem Beispiel demonstrieren.

Es möge nach $\sqrt{5234096}$ gefragt sein. Wir übertragen den Radikanden, entsprechend wie bei der Division den Dividenden, auf die äußersten linken Stellen des Resultatwerkes (alle übrigen Schaulöcher des Resultatwerkes werden mit Nullen besetzt) und löschen die Ziffern des Umdrehungszählwerkes. Dann verschieben wir den Schlitten möglichst weit nach rechts und trennen die erste Stelle des Radikanden (die Ziffer 5) mittels des Kommas ab. Daraufhin beginnen wir von 5 die ungeraden Zahlen zu subtrahieren, wobei wir nach zwei Subtraktionen $5 - 1 - 3 = 1$ erhalten. Eine weitere Subtraktion ist nicht möglich. Dann verschieben wir den Schlitten um eine Stelle nach links, aber das Komma im Radikanden um zwei Stellen nach rechts. Den letzten Subtrahenden (3) vergrößern wir um 1 und stellen in dem rechts benachbarten Einschnitt die Ziffer 1 ein. Dann fahren wir fort, mit der Zahl 41 beginnend, ungerade Zahlen zu subtrahieren. Nach zwei Umdrehungen hören wir auf, da wir die Zahl $123 - 41 - 43 = 39$ erhalten, die kleiner als die nächste ungerade Zahl 45 ist. Wiederum rücken wir den Schlitten um eine Stelle nach links und das Komma um zwei Stellen nach rechts, vergrößern den letzten Subtrahenden 43 um 1 und stellen in dem rechts anschließenden Einschnitt den Handgriff auf die Ziffer 1. Sodann subtrahieren wir, mit der Zahl 441 beginnend, weiter ungerade Zahlen, wobei wir (nach acht Subtraktionen) bei der Zahl 455 anhalten müssen. Schließlich rücken wir nochmals den Schlitten um eine Stelle nach links und das Komma um zwei Stellen nach rechts (jetzt befindet sich das Komma hinter der letzten Ziffer des Restes 36696) und subtrahieren nacheinander die Zahlen 4561, 4563, ..., 4573. Dann finden wir im Resultatwerk den endgültigen Rest 3727 und im Umdrehungszählwerk die gesuchte Wurzel 2287. Eine Kontrollrechnung ergibt:

$$2287^2 + 3727 = 5234096.$$

Ferner überzeugen wir uns davon, daß

$$2288^2 = (2287 + 1)^2 = 2287^2 + 2 \cdot 2287 + 1 = 2287^2 + 4575$$

größer als der gegebene Radikand ist.

Von den anderen Rechenmaschinen, die — wie die Maschine „Felix" — für die Ausführung der arithmetischen Grundrechenarten bestimmt sind, erwähnen wir noch eine Rechenmaschine mit kontinuierlicher Zehnerübertragung, die von dem berühmten russischen Mathematiker P. L. TSCHEBYSCHEW im Jahre 1882 erfunden und konstruiert wurde. Bei jeder Bewegung eines der Rechenräder der Maschine bewegt sich ein Rechenrad der nächst höheren Kategorie mit einer zehnfach geringeren Geschwindigkeit; in diesem Sinne erfolgt also die Übertragung der Zehner einer Kategorie auf die nächste kontinuierlich; der Übertragungsmechanismus dieser Maschine ist wesentlich einfacher als der Übertragungsmechanismus in der von ODHNER erfundenen Maschine. Eine ausführliche Beschreibung der TSCHEBYSCHEWschen Maschine findet sich in Band IV seiner Gesammelten Werke.[1])

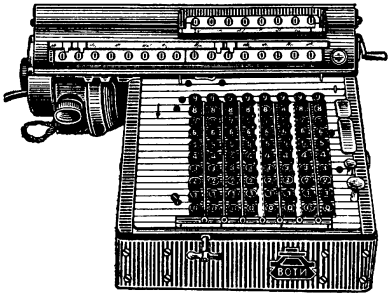


Abb. 6

Daneben gibt es heute eine große Anzahl weiterer Maschinen, mit deren Hilfe man die arithmetischen Grundrechenarten ausführen kann. Sie unterscheiden sich voneinander in dem Grad ihrer Mechanisierung und in verschiedenen Ergänzungsmechanismen, die Antworten drucken, die Einstellung der Werte kontrollieren, Zwischenergebnisse liefern usw. Besondere Verbreitung haben die Maschinen mit Volltastatur gefunden, von denen eine in Abb. 6 gezeigt ist. Das Triebwerk wird bei der abgebildeten Maschine nicht, wie bei der Rechenmaschine „Felix", mit der Hand, sondern durch einen Elektromotor angetrieben. Die kleine Kurbel (auf der Abbildung rechts) ist lediglich zum Löschen des Resultatwerkes bestimmt. Die Maschine addiert und multipliziert nach einem Druck auf die Taste mit dem Pluszeichen und subtrahiert und dividiert nach einem Druck auf die Taste mit dem Minuszeichen. Dank der Arbeit des Motors liefert die Tastenrechenmaschine die Resultate wesentlich schneller als die von Hand betriebenen Rechenmaschinen; daher ist das Rechnen auf den Tastenmaschinen weniger ermüdend.

Für die in den verschiedenen Wissenschaften und Zweigen der Technik, besonders in der Astronomie, der Optik und der Statistik auftretenden Rechenaufgaben sind eine große Zahl von speziellen Rechenmaschinen konstruiert und gebaut worden, die unvergleichlich komplizierter und vollkommener als die von uns beschriebenen Rechenmaschinen sind. Dabei empfiehlt es sich jedesmal, eine besondere Maschine

[1]) П. Л. Чебышев, Полное собрание сочинений, Band IV, Теория механизмов. (P. L. TSCHEBYSCHEW, Gesammelte Werke IV, Theorie der Mechanismen), Verlag d. Akad. Wiss. d. UdSSR, 1948.

zu konstruieren und zu bauen, wenn sich in der Praxis eine solche Aufgabe bei verschiedenen Ausgangszahlen häufig wiederholt, besonders dann, wenn es darauf ankommt, die Lösung möglichst schnell zu ermitteln. Mitunter stellt eine solche Maschine eine Kombination einer Reihe von einfacheren Maschinen dar, von denen jede eine ganz bestimmte Operation ausführt und die Resultate an die nächste Maschine weitergibt, wobei natürlich alles automatisch vor sich geht. Auf diese Weise haben die mannigfachen Aufgaben der Wissenschaft und Technik zu einem neuen Zweig des Präzisionsmaschinenbaus geführt, dem die Konstruktion und Produktion der verschiedensten „rechnunglösenden Mechanismen" obliegt.

„Jedem rechnunglösenden Mechanismus", so schreiben N. A. KOBRINSKI und L. A. LJUSTERNIK[1]), „liegt die Herstellung eines physikalischen Modells einer gewissen mathematischen Beziehung, eines physikalischen Prozesses, der diese Beziehung darstellt, zugrunde. So wird die Addition von Zahlen realisiert durch die Addition von Winkeln oder Strecken, durch die Addition von Strömen, die in einem Knoten eines elektrischen Leitersystems zusammentreffen usw. Die Multiplikation von Zahlen kann man modellmäßig erfassen, indem man die Faktoren als Winkel bei der Drehung eines Hebelarmes bzw. als Länge des Hebelarmes und das Produkt als Weglänge des Endes des Hebelarmes deutet oder indem man die Faktoren als Spannung bzw. Leitvermögen und ihr Produkt gemäß dem OHMschen Gesetz als Stromstärke darstellt usw. . . . Dabei beschreibt i. a. ein und dieselbe mathematische Operation verschiedene physikalische Prozesse, von denen dann umgekehrt jeder als physikalisches Modell der mathematischen Operation dienen kann. Man wird sich naturgemäß vornehmlich mit physikalischen Prozessen befassen, die leicht zu untersuchen und zu messen sind . . . Interessant zu bemerken ist, daß man früher mechanische Modelle zur Berechnung elektrischer Vorgänge gebaut hat, während heute, nach der stürmischen Entwicklung der elektrischen Meßtechnik, umgekehrt elektrische Geräte als Modelle für mechanische Systeme dienen."

In vielen Fällen der Praxis muß man den Inhalt eines von einer beliebigen Kurve begrenzten ebenen Flächenstücks bestimmen. Der Geodät tut dies auf der Landkarte, der Energie-Ingenieur auf dem Indikatordiagramm, das die Abhängigkeit des Druckes im Zylinder einer Dampfmaschine oder eines Verbrennungsmotors von der Kolbenlage anzeigt; in den Lederfabriken errechnet man die Produktion durch Messung des Flächeninhaltes der einzelnen gegerbten Häute. Zur Vereinfachung und Beschleunigung der Lösung dieser Aufgaben sind viele Spezialgeräte gebaut worden, die unter dem Namen „Planimeter" bekannt sind und Rechenmaschinen für kontinuierliche Operationen darstellen. Sie arbeiten automatisch; man kann das gesuchte Resultat ablesen, nachdem der Fahrstift des Gerätes die vorliegende Kontur umschrieben hat oder die zu messende Kontur (z. B. die Kontur einer gegerbten Haut) die Walken des Gerätes passiert hat.

Eine große Zahl von Aufgaben in diskreten (nichtkontinuierlichen) Größen, die letzten Endes Aufgaben in natürlichen Zahlen sind, werden mit Maschinen eines anderen Typs, den sogenannten rechenanalytischen Maschinen gelöst. Ihre Projektierung, Produktion und Anwendung hat in den letzten Jahren einen erheblichen Aufschwung genommen. Es gibt besondere Fabriken, die diese Maschinen herstellen. Gebraucht werden sie in Recheninstituten usw.

[1]) Н. А. Кобринский и Л. А. Люстерник, Математическая техника. (N. A. KOBRINSKI und L. A. LJUSTERNIK, Mathematische Technik), Uspechi mat. nauk, Bd. I, Heft 5/6 (15/16), (Neue Serie) 1946.

Betrachten wir zwei Beispiele für die Anwendung rechenanalytischer Maschinen:

Im Rahmen einer Volkszählung sei ein umfangreiches Material in Form von Fragebogen gesammelt, die von jeder gezählten Person Angaben über Geschlecht, Alter, Nationalität, Bildung, Beruf usw. enthalten. Zur Auswertung soll nun festgestellt werden, wieviel Männer und Frauen, wieviel Menschen in jeder Altersgruppe u. a. m. vorhanden sind. Das Sortieren der Fragebogen von Hand und die Aufschlüsselung in Gruppen erfordert wegen ihrer Vielzahl einen maßlos großen Zeitaufwand und liefert Resultate, die bei weitem nicht frei von zufälligen Fehlern sind. Daher ist es ganz natürlich, daß man Maschinen gebaut hat, die diese Berechnungen automatisch und mit sehr großer Geschwindigkeit ausführen. Dazu werden alle Angaben, die im Verlauf der Zählung für jeden einzelnen Menschen ermittelt wurden, auf besondere *Lochkarten* übertragen, auf denen jedem speziellen Merkmal an einer bestimmten Stelle ein kleines Loch (oder auch eine Gruppe von Löchern) entspricht. Diese Lochkarten laufen anschließend mit großer Geschwindigkeit durch die rechenanalytische Maschine. Die Maschinen arbeiten mit einem Motor und zählen fehlerlos zusammen, wieviel Karten, die an bestimmten Stellen Löcher besitzen, die Maschine passiert haben, d. h., wie vielen der registrierten Menschen ein ganz bestimmtes Merkmal zukommt.

Als zweites Beispiel nehmen wir die sogenannten „Differenzentafeln", die bei der Lösung vieler Aufgaben zu bilden sind. Das sind Zahlentafeln, in denen die auftretenden Zahlen Funktionswerte ganz konkreter Funktionen sind. Nachstehend findet man eine solche Tafel der Differenzen erster, zweiter und dritter Ordnung für die Funktion $f(x) = 10^4\sqrt{x}$; die Werte sind jeweils auf ganze Zahlen gerundet und entsprechen den ganzzahligen Argumentwerten von 10 bis 20.

x	$f(x) = 10^4\sqrt{x}$	$\Delta f(x)$	$\Delta^2 f(x)$	$\Delta^3 f(x)$
10	31623			
		1543		
11	33166		−68	
		1475		8
12	34641		−60	
		1415		6
13	36056		−54	
		1361		6
14	37417		−48	
		1313		5
15	38730		−43	
		1270		4
16	40000		−39	
		1231		3
17	41231		−36	
		1195		4
18	42426		−32	
		1163		1
19	43589		−31	
		1132		
20	44721			

Die Zahlen in der Spalte unter $\Delta f(x)$, die *erste Differenzen* oder *Differenzen erster Ordnung* genannt werden, berechnen sich als Differenzen benachbarter Funktionswerte (von jedem wird der vorangehende subtrahiert); so ist $33166 - 31623 = 1543$; $34641 - 33166 = 1475$ usw. Wenn wir mit der Spalte der ersten Differenzen entsprechend verfahren, so erhalten wir die Spalte der *zweiten Differenzen*, darauf die der *dritten*, *vierten* usw. Mitunter muß man die Differenzentafeln bis zu den Differenzen fünfter, sechster, siebenter und höherer Ordnung aufstellen. Das ist eine Arbeit, die prinzipiell höchst einfach ist, andererseits aber eine ungeheure Mühe bereitet und eine sorgfältige Prüfung jedes Schrittes erfordert.

Vor einigen Jahren wendete I. N. JANSHUL, ein sowjetischer Spezialist für rechnunglösende Mechanismen, eine der rechenanalytischen Maschinen, den sogenannten „Tabulator", zur automatischen Aufstellung von Differenzenschemata an. Für jeden Wert der vorgelegten Funktion wird dabei nach einem besonderen Schlüssel eine Lochkarte hergestellt; die Lochkarten werden dann in die Maschine gegeben, die automatisch die Differenzentafel druckt. Wenn man alle Hilfsoperationen in Betracht zieht, die bei der Arbeit des Automaten benötigt werden, so zeigt sich, daß nach diesem Verfahren stündlich rund 2000 Differenzen geliefert werden, wobei jede Differenz bis zu 7 Stellen haben kann. Das übertrifft den bis dahin geltenden Schnelligkeitsrekord für eine solche Arbeit um das 3- bis 4fache.

Die verschiedenen rechenanalytischen Maschinen (Tabuliermaschinen, Multiplier u. a.) gestatten es, sehr schnell und genau allerlei Tafeln zusammenzustellen. Die gegebenen Größen werden auf Lochkarten übertragen, und die Maschinen liefern automatisch (mit Motorantrieb) die fertigen, gedruckten Tafeln.

Über die Geschichte der rechenanalytischen Maschinen und der Fortschritte, die in der UdSSR hinsichtlich ihrer Projektierung, Produktion und Anwendung erzielt wurden, kann man sich in den Artikeln informieren, die in der Zeitschrift „Успехи математических наук" (Uspechi mat. nauk) in den Jahren 1946 und 1947 erschienen sind. Eine ausführliche Bibliographie über Fragen der Mechanisierung des Rechnens im allgemeinen findet sich in dem Buch „Mathematische Instrumente" von FR. A. WILLERS[1]).

Abschließend wollen wir noch eine Maschine für die Auflösung eines Systems von Gleichungen ersten Grades erwähnen. Prinzipiell erfordert die Auflösung eines solchen Systems nichts weiter als die vier Grundrechenarten über dem Koeffizientenbereich; sie kann durchaus von einem Schüler der siebenten Klasse bewältigt werden. Wenn jedoch viele Gleichungen und Unbekannte vorliegen (was bei den in der Praxis auftretenden Aufgaben häufig der Fall zu sein pflegt, es sind nicht selten Systeme in einigen zehn Unbekannten aufzulösen) und wenn als Koeffizienten mehrstellige Zahlen auftreten, so wird die Lösung des Systems zu einer Angelegenheit, die viele Monate angestrengter Arbeit beansprucht. Im Laufe der Zeit sind eine ganze Reihe von Maschinen gebaut worden, die diese Arbeit wesentlich schneller

[1]) FR. A. WILLERS, Mathematische Instrumente, München 1943. Eine russische Übersetzung dieses Buches ist 1949 in Moskau erschienen. In dieser Übersetzung ist in zahlreichen Anmerkungen auf die Entwicklung in der UdSSR hingewiesen und das Schriftenverzeichnis entsprechend vervollständigt. Eine Neubearbeitung des genannten Buches, in der die Anmerkungen der russischen Übersetzung zum Teil verarbeitet sind, ist 1951 im Akademie-Verlag Berlin unter dem Titel „Mathematische Maschinen und Instrumente" erschienen *(Anm. d. wissenschaftl. Red.)*.

bewältigen. Erwähnt sei hier eine Maschine, die im Jahre 1940 von dem sowjetischen Konstrukteur W. M. PROSCHKO[1]) gebaut wurde. Mit ihrer Hilfe kann man Systeme von 10 Gleichungen ersten Grades in 10 Unbekannten und mit dreistelligen Koeffizienten lösen.

In der UdSSR wird gegenwärtig von einer Reihe von wissenschaftlichen Zentren aus intensiv an der weiteren Vervollkommnung der bekannten Rechenmaschinen und an der Konstruktion neuer Typen gearbeitet. Eine hervorragende Stelle nimmt dabei die Abteilung für Näherungsrechnung des mathematischen Instituts der Akademie der Wissenschaften der UdSSR ein. Eingehend bekannt machen kann man sich mit diesen Arbeiten an Hand eines Aufsatzes von K. A. SEMENDJAJEW, der in dem Band «Математика в СССР за тридцать лет» („Dreißig Jahre Mathematik in der UdSSR") enthalten ist, der vom Staatlichen Verlag für technisch-theoretische Literatur im Jahre 1948 herausgegeben wurde.

§ 5. Näherungswerte

Bei der experimentellen Bestimmung des Wertes einer unbekannten Größe x handelt es sich um ihre Berechnung oder ihre Messung. Dabei erhält man — wie wir schon oben bemerkt haben — nur in sehr einfachen Ausnahmefällen genaue Werte, während man in der Regel gezwungen ist, sich mit angenäherten Werten zufrieden zu geben: Statt der genauen Gleichung $x = a$ erhält man eine Näherungsgleichung $x \approx a$. In der Schulmathematik wird dieser Charakter der meisten Zahlen, mit denen wir es in der Praxis zu tun haben, völlig ignoriert. Man kommt jedoch nicht um die Tatsache herum, daß man schon in der fünften Klasse bei der Behandlung der Brüche auf Quotienten stößt, die auf unendliche Dezimalbrüche führen. Um diese Dezimalbrüche in der Praxis zu verwenden oder auch nur weitere Operationen an ihnen vorzunehmen, muß man sie notgedrungen runden. Damit stößt man aber bereits bei der Behandlung der rationalen Zahlen in der Schulmathematik ganz zwangsläufig auf das Rechnen mit Näherungswerten. Im Verlauf des weiteren Unterrichtes tauchen dann die verschiedensten irrationalen Zahlen, wie Wurzeln, Logarithmen und deren Numeri, die Zahl π, Werte der trigonometrischen Funktionen und der Kreisfunktionen usw. auf. Daher ist das Fehlen eines besonderen Abschnittes über Näherungsrechnung in der Schulmathematik ein ernsthafter Mangel der gegenwärtigen Unterrichtspläne, der sich höchst ungünstig auf die mathematische Allgemeinbildung der schulentlassenen Jugend auswirkt.

Hier zunächst ein typisches Beispiel aus der physikalischen Praxis: Gefragt ist nach dem (möglichst genauen) mittleren spezifischen Gewicht δ des Materials eines Stückes runden Drahtes, dessen Durchmesser mittels einer Mikrometerschraube zu $2r \approx 0{,}48$ mm, dessen Länge mittels eines Millimetermaßes zu $h \approx 264{,}4$ mm und dessen Gewicht mittels einer Laboratoriumswaage zu $p \approx 0{,}423$ g bestimmt wurde. Man hat dann lediglich die Berechnung nach den Formeln $\delta = \frac{p}{v}$, $v = \pi r^2 h$ auszuführen, wobei p in Gramm und h und r in Zentimetern auszudrücken sind.

[1]) В. М. Прошко, Приборы для определения корней системы линейных уравнений (W. M. PROSCHKO, Instrumente zur Bestimmung der Wurzeln eines Systems linearer Gleichungen), Uspechi mat. nauk, Bd. I, Heft 5/6 (15/16), (Neue Serie) 1946.

Bei dieser Berechnung tauchen nun eine Reihe von Schwierigkeiten auf. Erstens: Mit welcher Genauigkeit ist π zu nehmen? Um — wie verlangt — die Größe δ möglichst genau zu bestimmen, wird man den genauesten Wert nehmen, den man üblicherweise in den Lehrbüchern angegeben findet, nämlich 3,14159, wobei jedoch zweifelhaft bleibt, ob man nicht besser noch mehr Stellen verwendet. Dann findet man

$$r \approx 0{,}24 \text{ mm}, \; r^2 \approx 0{,}0576 \text{ mm}^2, \; \pi r^2 \approx 0{,}180955584 \text{ mm}^2,$$

$$v = 47{,}8446564096 \text{ mm}^3 = 0{,}0478446564096 \text{ cm}^3.$$

Es taucht nun eine zweite Schwierigkeit auf: Die erhaltenen Zahlen besitzen sehr viele Dezimalstellen; kann man sie nicht in irgendeiner Weise runden? Aussagen darüber, welche Rundungen im vorliegenden Fall zulässig sind, macht die Schulmathematik nicht, deshalb nehmen wir vorsichtshalber keine Rundungen vor. Dann stoßen wir aber bei der Division von $p \approx 0{,}423$ g durch den gefundenen Wert v auf eine dritte Schwierigkeit: Wie viele Ziffern hat man in dem entstehenden Quotienten zu nehmen, der durch einen unendlichen Dezimalbruch ausgedrückt wird? Die Division kann zwar unbeschränkt fortgesetzt werden, aber es ist klar, daß wir bei den angenähert gemessenen Werten von r, h, p auch nur einen angenäherten Wert für δ erhalten können. Daher wäre es unsinnig, wenn man nicht nach einer gewissen Anzahl von Dezimalstellen abbrechen würde; sonst würde man ja die Tatsache, daß die gegebenen Größen nur angenähert sind, völlig ignorieren. Solange man aber über keine feste Grundlage für die Bestimmung der notwendigen Anzahl von Dezimalstellen verfügt, befindet man sich in einem peinlichen Zustand der Unsicherheit, der keinesfalls den Vorstellungen von einer exakten Wissenschaft, wie die Mathematik es sein soll, entspricht. Wenn wir im Beispiel nach der Berechnung der Hundertstelziffer abbrechen, so stellen wir fest, daß der Rest bei der Division größer als der halbe Quotient ist. Wir werden also geneigt sein, den Quotienten auf Hundertstel nach oben zu runden, und erhalten $\delta \approx 8{,}84$. Jedoch haben wir keinerlei Grund für die Behauptung, daß alle Ziffern dieser Zahl genau sind, noch können wir aussagen, die unberücksicht gelassenen Stellen (Tausendstel und die weiteren) verdienten kein Vertrauen mehr.

Alle bei Rechnungen mit Näherungswerten auftretenden Schwierigkeiten von der in dem Beispiel gezeigten Art lassen sich durch einige sehr einfache Regeln aus dem Wege räumen, die bereits von den Schülern der fünften Klasse durchaus bewältigt werden können und die alle Berechnungen wesentlich vereinfachen. Über sie wird in den §§ 10—13 eingehender zu sprechen sein. Wendet man diese Regeln auf die eben behandelte Aufgabe an, so findet man, daß $\delta \approx 8{,}8$ ist mit der Sicherheit, daß in der Zehntelziffer ein irgendwie bedeutsamer Fehler sehr unwahrscheinlich ist, daß aber die unberücksichtigt gelassenen Stellen (Hundertstel, Tausendstel usw.) kein Vertrauen mehr verdienen. Wenn wir nämlich die etwas langwierige, aber prinzipiell sehr einfache „Methode der Wertschranken" anwenden, die selbst von den Schülern der Grundschule durchaus verstanden werden kann, so stellen wir folgendes fest: Nimmt man nämlich an, es sei $0{,}475 \text{ mm} < 2r < 0{,}485$ mm, $264{,}3 \text{ mm} < h < 264{,}5$ mm, $0{,}422 \text{ g} < p < 0{,}424$ g, so kann man mit absoluter Sicherheit behaupten, daß das gesuchte spezifische Gewicht zwischen 8,63 und 9,05 liegt und darum $\delta \approx 8{,}8$ ($\pm 0{,}25$) sein muß, d. h., der Näherungswert von δ gleich 8,8 ist und von dem gesuchten genauen Wert auf jeden Fall um weniger als 0,25 abweicht (die ausführliche Berechnung wird auf S. 348 nachgeholt).

Im üblichen Arithmetikunterricht lernen die Schüler, Dezimalbrüche auf eine bestimmte Ordnung oder — was dasselbe ist — auf eine bestimmte Dezimalstelle zu runden. Diese Operation wird im Zusammenhang mit der Division behandelt,[1]) besitzt aber auch unabhängig von ihr eine erhebliche Bedeutung. Sie besteht darin, daß man in einer gegebenen ganzen oder gebrochenen Zahl alle Ziffern vernachlässigt, die sich rechts von der Ziffer der betreffenden Ordnung finden. Wenn dabei die erste vernachlässigte Ziffer gleich 0, 1, 2, 3 oder 4 ist, so ist der gerundete Wert kleiner als der vorgegebene Wert, der gegebene Wert wird *nach unten gerundet* oder *abgerundet*; wenn hingegen diese Ziffer gleich 5, 6, 7, 8 oder 9 ist, so erhöht man die letzte berücksichtigte Ziffer um 1, der gegebene Wert wird *nach oben gerundet* oder *aufgerundet*. In beiden Fällen ist der absolute Betrag der Differenz zwischen vorgegebenem Wert und gerundetem Wert (der *Fehler* des Rundens) kleiner als die halbe Einheit der zuletzt berücksichtigten Ordnung (auf die gerundet wurde). Eine Ausnahme bildet nur der Fall, daß genau die Ziffer 5 (und nichts weiter) vernachlässigt wird, in dem die Differenz genau gleich der halben Einheit der zuletzt berücksichtigten Ordnung ist; hier wendet man mitunter die „Regel der geraden Ziffern" an: ist die zuletzt berücksichtigte Ziffer gerade, so wird abgerundet, ist die zuletzt berücksichtigte Ziffer ungerade, so wird aufgerundet. Mittels dieser zusätzlichen Regel erhält man als Näherungszahlen beim Runden der Zahl 345,0715 auf Hunderter, Zehner, Einer, Zehntel, Hundertstel, Tausendstel der Reihe nach 300; 350; 345; 345,1; 345,07; 345,072, während die Zahl 345,0725 bei Abrunden auf Tausendstel ebenfalls 345,072 ergibt.[2])

Damit wir später nicht nochmals auf das Runden zu sprechen kommen müssen, erwähnen wir bereits hier, daß das Runden auf eine bestimmte Ordnung bzw. eine bestimmte Dezimalstelle gleichbedeutend dem Runden mit einer bestimmten Anzahl von „wesentlichen Ziffern" ist. Dabei verstehen wir unter den *wesentlichen Ziffern* einer Näherungszahl alle ihre Ziffern, mit Ausnahme der Nullen links und jener Nullen rechts, die vernachlässigte oder unbekannte Ziffern vertreten. So enthält die Zahl 3,14 zwei Dezimalstellen, aber drei wesentliche Ziffern. Beim Runden der Zahl 345,0715 mit einer, zwei, drei, vier, fünf, sechs wesentlichen Ziffern erhält man der Reihe nach gerade die Zahlen, die oben angegeben sind. Wenn wir die Zahl 7893 mit einer, zwei, drei wesentlichen Ziffern runden, erhalten wir der Reihe nach 8000; 7900; 7890, wobei die Nullen rechts an die Stelle der vernachlässigten Ziffern treten und unwesentliche Ziffern darstellen (besser wäre es, 8 ? ? ?; 79 ? ?; 789 ? zu schreiben). Dagegen besitzt die Zahl 37,0, die zum Beispiel die mittels eines Fieberthermometers (mit einer in Zehntelgrad unterteilten Skala) gemessene Temperatur ausdrückt, drei wesentliche Ziffern, die (geschriebene) Null rechts ist wesentlich (während erst die nicht geschriebenen weiteren Nullen unwesentlich sind — *Anm. d. wissenschaftl. Red.*)

Aus gewissen Gründen sieht man mitunter die Ziffer „1" dann nicht als wesentliche Ziffer an, wenn sie die Ziffer höchster Ordnung einer Näherungszahl ist, d. h. wenn sie als erste Ziffer links steht (wir kommen hierauf auf S. 341 zurück). Legen wir diese Regel zugrunde, so besitzt z. B. die Zahl 12,47 nicht vier, sondern nur drei wesentliche Ziffern.

---

[1]) Siehe А. П. Киселёв, Арифметика. (A. P. Kisseljow, Arithmetik) 1946, § 169 und 170.

[2]) Die Bezeichnungen „Aufrunden" für das Runden *nach oben*, „Abrunden" für das Runden *nach unten* und „Runden" schlechthin entsprechen den Vorschriften des Normblattes (DIN 1333, Rundung und Kürzung von Zahlen). *(Anm. d. wissenschaftl. Red.)*

§ 6. Verschiedene Methoden zur Abschätzung der Genauigkeit von Näherungswerten

Wenn wir in irgendeinem Zusammenhang auf eine Näherungsgleichung der Form $x \approx a$ stoßen, so müssen wir in erster Linie ihren genauen Sinn klären. Was bedeutet im betrachteten Zusammenhang die Redeweise: „x ist angenähert gleich einer gegebenen Zahl“ ?

Eine Näherungsgleichung $x \approx a$ erhält sicher dann einen bestimmten Sinn, wenn sie von der Angabe *einer Schranke des absoluten Fehlers* begleitet wird, d. h. von der Angabe einer positiven Zahl Δa, die zu a addiert eine Zahl $a + \Delta a$ ergibt, die nachweisbar größer als der wahre (uns unbekannte) Wert von x, also eine sogenannte *obere Schranke* von x ($\mathrm{OS}(x)$) ist, und die von a subtrahiert eine Zahl $a - \Delta a$ ergibt, die nachweisbar kleiner als der wahre Wert von x, also eine *untere Schranke* von x ($\mathrm{US}(x)$) ist. In diesem Sinne bedeutet die Näherungsgleichung $x \approx 27{,}4\ (\pm 0{,}1)$ kg, daß $27{,}4\,\mathrm{kg} - 0{,}1\,\mathrm{kg} = 27{,}3\,\mathrm{kg}$ kleiner als x, aber $27{,}4\,\mathrm{kg} + 0{,}1\,\mathrm{kg} = 27{,}5\,\mathrm{kg}$ größer als x ist. Die üblicherweise verwendete Schreibweise $x \approx a\,(\pm\, \Delta a)$ ist somit gleichbedeutend mit der doppelten Ungleichung $a - \Delta a < x < a + \Delta a$. Wenn wir andererseits je eine $\mathrm{US}(x)$ und eine $\mathrm{OS}(x)$ kennen, d. h., wenn wir eine doppelte Ungleichung der Form $q < x < p$ haben, so finden wir leicht als Näherungswert $a = \frac{p+q}{2}$ mit $\Delta a = \frac{p-q}{2}$ als Schranke des absoluten Fehlers. Die Näherungsgleichung $x \approx a\,(\pm \Delta a)$ besagt also folgendes: „x ist angenähert gleich a mit einer Schranke des absoluten Fehlers von Δa“ oder „x ist angenähert gleich a und weicht von a nach oben und nach unten um weniger als Δa ab“.

Manchmal ersetzt man die echten Ungleichungen $a - \Delta a < x < a + \Delta a$ auch durch die schwächeren Ungleichungen $a - \Delta a \leq x \leq a + \Delta a$.

Nach einer Tradition, die sich seit der Zeit von GAUSS fest eingebürgert hat, besitzen alle in mathematischen Tabellen aufgeführten Näherungszahlen Schranken des absoluten Fehlers, die gleich der halben Einheit der zuletzt geschriebenen Ordnung sind. Wenn wir z. B. in einer Logarithmentafel den Wert $\lg 7 \approx 0{,}8451$ finden, so können wir sicher sein, daß der wahre Wert von $\lg 7$ sich von 0,8451 um weniger als $0{,}0001 : 2 = 0{,}00005$ unterscheidet und daß folglich $0{,}84505 < \lg 7 < 0{,}84515$ ist. Wenn wir entsprechend in einer Tafel den Wert $\mathrm{tg}\, 89°59' \approx 3438$ finden, so ist mit Sicherheit $3437{,}5 < \mathrm{tg}\, 89°59' < 3438{,}5$.

Daneben ist noch eine andere Definition der Schranke des absoluten Fehlers möglich und gebräuchlich, die offensichtlich der oben angegebenen äquivalent ist. Wenn wir nämlich $x = a + \xi$ setzen, so ist $\xi = x - a$ (wenn x unbekannt ist, so ist auch ξ unbekannt). Wir nennen ξ den *absoluten Fehler* oder auch kurz *Fehler* der Näherungszahl a. Unter einer Schranke des absoluten Fehlers werden wir dann eine beliebige positive Zahl Δa verstehen, die der Ungleichung $|\xi| < \Delta a$, also $|x - a| < \Delta a$, d. h. der doppelten Ungleichung $-\Delta a < x - a < \Delta a$ genügt. Wenn wir zu jedem Glied dieser Ungleichung die Zahl a addieren, so erhalten wir in Übereinstimmung mit der ersten Definition der Schranke des absoluten Fehlers die Ungleichungen $a - \Delta a < x < a + \Delta a$. Da man offenbar entsprechend von diesen letzten Ungleichungen zu der Ungleichung $|x - a| < \Delta a$ gelangt, sind beide Definitionen der Schranke des absoluten Fehlers in der Tat äquivalent.

Die Angabe einer Schranke des Fehlers gestattet es, die Genauigkeit verschiedener Näherungen ein und desselben unbekannten Wertes zu vergleichen: Je kleiner Δa ist, desto genauer ist der Näherungswert a. Wenn wir einmal ermittelt haben, daß $x \approx 5{,}64\,(\pm\,0{,}01)$ ist, und ein andermal feststellen, daß $x \approx 5{,}63183\,(\pm\,0{,}00002)$ ist, so können wir sagen, daß die zweite Näherung $0{,}01 : 0{,}00002 = 500$mal so genau ist wie die erste. Hingegen genügt für den Vergleich der Genauigkeit von Näherungen für verschiedene Zahlen die Angabe einer Schranke des absoluten Fehlers schon nicht mehr. Wenn z. B. bekannt ist, daß zwei Längenmessungen mit ein und derselben Schranke des absoluten Fehlers von 1 mm durchgeführt wurden, und zwar im ersten Fall ein Draht von etwa 2 mm Durchmesser und im zweiten Fall eine geodätische Basis von etwa 1 km Länge gemessen wurden, so ist unmittelbar klar, daß die erste Messung sehr grob war — die Schranke des absoluten Fehlers stellt volle 50% des Näherungswertes dar —, während die zweite Messung außerordentlich genau ist — die Schranke des absoluten Fehlers beträgt nur 0,0001% des gemessenen Wertes. Auf diese Weise gelangen wir ganz naturgemäß zum Begriff der *Schranke des relativen Fehlers.* Sie wird erklärt als das Verhältnis einer Schranke des absoluten Fehlers zum Näherungswert, d. h. als $\frac{\Delta a}{a}$, (oder auch als Verhältnis einer Schranke des absoluten Fehlers zum wahren Wert, d. h. als $\frac{\Delta a}{x}$, was aber praktisch auf dasselbe hinausläuft) und wird üblicherweise in Prozenten ausgedrückt. Die Angabe einer Schranke des relativen Fehlers ist in der Praxis weit verbreitet. So findet man auf Einzelteilen für Rundfunkempfänger (Widerstände, Kondensatoren, Selbstinduktionsspulen usw.) häufig Aufschriften von der Art: „$200\,\Omega \pm 10\%$", was bedeutet, daß der vorliegende Widerstand von $200\,\Omega$ um nicht mehr als 10% dieser Größe, also um $20\,\Omega$ abweicht und mithin zwischen $180\,\Omega$ und $220\,\Omega$ liegt.

Wir bemerken, daß sich die Schranke des relativen Fehlers einer Näherungszahl beim Übergang zu einer anderen Maßeinheit nicht ändert. Insbesondere ändert sich ihre Größe beim Verschieben des Kommas nicht, da sich hierbei sowohl a als auch Δa um denselben Faktor vergrößern oder verkleinern.

Die Angabe einer Schranke des absoluten bzw. des relativen Fehlers liefert zwei grundlegende Verfahren für die Charakterisierung der Genauigkeit von Näherungszahlen. In der Praxis wendet man diese beiden Verfahren verhältnismäßig selten an. Unvergleichlich häufiger charakterisiert man die Genauigkeit einer Näherungszahl durch einfache Angabe der Anzahl ihrer Ziffern, die zuverlässig sind, und zwar sagt man, alle Ziffern einer betrachteten Näherungszahl seien *zuverlässig*, wenn die halbe Einheit der Ordnung der letzten Ziffer dieser Näherungszahl eine Schranke des absoluten Fehlers dieser Zahl ist. Daher ist eine Näherungszahl, deren sämtliche Ziffern zuverlässig sind, das Resultat eines Rundens des entsprechenden wahren, uns unbekannten Wertes auf eine gewisse Ordnung. Wie gesagt, sind alle in mathematischen Tafeln aufgenommenen Zahlen von dieser Art. Durch Angabe der wesentlichen Ziffern einer solchen Näherungszahl wird von selbst die Genauigkeit dieser Zahl charakterisiert. In diesem Sinne wird man unter einer zweistelligen, dreistelligen, allgemein k-stelligen Näherungszahl eine Zahl verstehen, die zwei, drei, allgemein k wesentliche Ziffern besitzt. Diese Charakterisierung der Genauigkeit einer Näherungszahl besitzt den Vorteil, daß sie keine zusätzlichen Angaben erfordert; die Niederschrift der Zahl spricht für sich selbst.

Die Angabe der Anzahl der zuverlässigen Ziffern und der Kommastellung ist gleichbedeutend mit der Angabe einer Schranke des absoluten Fehlers. So ist 0,0005 eine Schranke des absoluten Fehlers der mit vier zuverlässigen Ziffern und dem Komma hinter der ersten wesentlichen Ziffer angegebenen Quadratwurzeln aus den Zahlen 1 bis 100. Daneben läßt sich auch ein Schluß über eine Schranke des relativen Fehlers ziehen; zwar ist er weniger bestimmt als der Schluß über die Schranke des absoluten Fehlers; trotzdem ist er überaus wesentlich für die Begründung einiger praktischer Regeln. In der Tat: Wenn uns eine k-stellige Näherungszahl a vorgegeben ist, deren sämtliche Ziffern zuverlässig sind, und wenn das Komma hinter ihrer letzten Ziffer steht (wie wir wissen, ändert ein Verschieben des Kommas den relativen Fehler nicht), so genügt diese Zahl a den Ungleichungen $10^{k-1} \leq a < 10^k$, und es ist $\Delta a = 0{,}5$; daher ist jede Schranke des relativen Fehlers dieser Näherungszahl zwischen $0{,}5 \cdot 10^{-k}$ und $0{,}5 \cdot 10^{-(k-1)}$ eingeschlossen. Wenn wir dies in Prozenten ausdrücken, so erhalten wir:

k	2	3	4	5	6
$\frac{\Delta a}{a}$	von 0,5% bis 5%	von 0,05% bis 0,5%	von 0,005% bis 0,05%	von 0,0005% bis 0,005%	von 0,00005% bis 0,0005%

Wenn man die Ziffer 1 als Ziffer höchster Ordnung in einer Näherungszahl als nicht wesentlich ansieht, so verringern sich die Werte der Schranken von $\frac{\Delta a}{a}$ in dieser Tabelle jeweils auf die Hälfte, da unter dieser Voraussetzung die schärferen Ungleichungen $2 \cdot 10^{k-1} \leq a < 2 \cdot 10^k$ gelten.

Mitunter findet man eine etwas andere Definition des Terminus „zuverlässige Ziffer". So lesen wir in dem Artikel von P. S. ALEXANDROFF und A. N. KOLMOGOROFF [6]: „Man sagt, eine Ziffer eines vorgegebenen Näherungswertes für eine Zahl sei zuverlässig, wenn der absolute Fehler die Größe der Einheit der betreffenden Ordnung nicht überschreitet." In der Praxis trifft man jedoch häufig auf Näherungszahlen, bei denen der (absolute) Fehler erheblich größer sein kann; so führt schon die Addition von vier Summanden, von denen jeder einen Fehler von höchstens einem halben Hundertstel aufweist, auf eine Summe, deren Fehler in der Nähe von zwei Hundertstel liegen kann. Wenn nun die Wahrscheinlichkeit großer Fehlerwerte in der letzten Ziffer der Näherungszahl viel geringer ist als die kleiner Fehlerwerte, so bleibt auch die letzte Ziffer erhalten, da sie in einem bestimmten Maße Vertrauen verdient, wenn auch keine Gewähr für ihre Zuverlässigkeit vorhanden ist. Wir werden dies noch an vielen Zahlenbeispielen sehen.

Wir haben bisher drei Verfahren zur Charakterisierung der Genauigkeit einer Näherungszahl kennengelernt: 1.) Die Angabe einer Schranke des absoluten Fehlers, 2.) die Angabe einer Schranke des relativen Fehlers, 3.) die Angabe der Anzahl der Ziffern, die zuverlässig sind. Diese letzte Methode wird im täglichen Leben bei weitem am häufigsten verwendet. Im folgenden werden wir noch zwei weitere Verfahren kennenlernen, die große theoretische Bedeutung besitzen, bisher aber in der Schule überhaupt nicht verwendet werden. Es sind die Charakterisierung der Genauigkeit einer Näherungszahl durch Angabe ihres mittleren quadratischen Fehlers und die Charakterisierung durch Angabe der Wahrscheinlichkeit für das Auftreten von gewissen Werten des Fehlers.

§ 7. Die Auswertung von Meßresultaten

In einfachen Fällen, mit denen man es zwangsläufig bereits in der Schule zu tun bekommt, erhält man bei Messungen angenäherte Resultate, deren Genauigkeit sich leicht durch Angabe einer Schranke des absoluten Fehlers charakterisieren läßt. Wenn wir z. B. mit einem Millimeterlineal die Länge x eines Bleistiftes messen und feststellen, daß sie zwischen 178 mm und 179 mm, und zwar näher an 179 mm liegt, so finden wir, daß $x \approx 179 (\pm 0{,}5)$ mm ist. Streben wir eine noch kleinere Schranke des absoluten Fehlers an, so werden wir sagen, daß $x \approx 178{,}75 (\pm 0{,}25)$ mm ist. Bei jeder Wägung ermittelt man, bei welcher größtmöglichen Belastung der einen Waagschale mit zur Verfügung stehenden Gewichten die andere Waagschale mit der Last ein Übergewicht zeigt und welche minimale Zulage von Gewichten ein Übergewicht der Waagschale mit den Gewichten hervorruft. Wenn wir z. B. ein kleines Gewicht von 0,1 g benutzen und bemerken, daß bei einer Belastung mit 67,6 g die Last überwiegt, bei einer Belastung mit 67,7 g dagegen die Gewichte überwiegen, so erhalten wir damit eine obere und eine untere Schranke des gesuchten Gewichtes. Hieraus erhalten wir dann aber auch unmittelbar einen Näherungswert für das gesuchte Gewicht, der sich zu $(67{,}6 + 67{,}7) : 2 = 67{,}65$ g ergibt, und eine Schranke des absoluten Fehlers, die gleich $(67{,}7 - 67{,}6) : 2 = 0{,}05$ g ist.

Man wird nun stets bestrebt sein, unter gegebenen Verhältnissen eine größtmögliche Genauigkeit der Messung sicherzustellen. Deshalb teilt man das kleinste Teilintervall, das sich auf der Skala des verwendeten Meßgerätes befindet, nach Augenmaß nochmals in 5 oder 10 Teile und erhält damit noch eine näherungsweise weitere Dezimalstelle des gesuchten Wertes. Dabei wird sich allerdings bei Wiederholung der Messung häufig ein Wert ergeben, der von den Resultaten vorangehender Messungen etwas verschieden ist, da sowohl bei der Abschätzung nach Augenmaß als auch in der Einstellgenauigkeit geringfügige Fehler auftreten können (z. B. kann bei Messungen mit einem Millimeterlineal der Nullpunkt des Lineals um einige Millimeterteile vom Anfangspunkt des zu messenden Objektes abweichen) und unter Umständen auch das zu messende Objekt nicht vollständig bestimmt ist (so werden z. B. mitunter bei Längenmessungen die Enden der zu messenden Strecke durch dünne Striche markiert, die immerhin eine Dicke von 0,1 mm haben und deren Mitte man nach Augenmaß bestimmen muß). In komplizierteren Fällen nimmt diese Schwankung im Meßresultat erheblich zu. Das ist z. B. der Fall, wenn man im Gelände eine längere Strecke mißt, in die das verwendete Maß (20 m Band) einige Male hineinpaßt, oder wenn man das mittlere Gewicht von hundert Weizenkörnern bestimmen möchte, die auf gut Glück aus einem bestimmten Haufen Weizen herausgenommen werden.

Wenn wir also im Verlauf wiederholter Messungen eine Reihe von mehr oder weniger gut übereinstimmenden Näherungswerten $a_1, a_2, \ldots, a_n$ für ein und dieselbe unbekannte Größe x gefunden haben, so erhebt sich die Frage nach einer Auswertung der erhaltenen Meßresultate mit dem Ziel, 1.) für x einen möglichst guten Näherungswert a zu finden und 2.) eine Charakterisierung der Genauigkeit der Näherungsgleichung $x \approx a$ zu gewinnen.

Dabei ist zunächst zu untersuchen, ob die bei den Messungen erhaltenen Werte keinen systematischen Fehler aufweisen, der durch eine beständig wirkende Ursache bedingt ist. Wenn man z. B. bei Längenmessungen einen Millimeterstab

benutzt, dessen Teilintervalle etwas kürzer sind als die üblichen, so erhält man laufend Werte, die zu groß sind. Wenn man einen hölzernen Gegenstand mit Messinggewichten wägt, so erhält man Werte, die auf Grund des beständigen Verlustes an Gewicht durch die verdrängte Luft (ARCHIMEDisches Gesetz in Gasen) kleiner als die wahren Werte sind. Fehler dieser Art müssen naturgemäß berücksichtigt und zunächst durch Einführung entsprechender Korrekturen beseitigt werden. Hat man z. B. festgestellt, daß die 100 Teilintervalle des verwendeten Millimeterlineals nicht 100 mm sondern nur 98,5 mm messen, so muß man alle mit den Teilungen dieses Maßstabes gemessenen Resultate mit 0,985 multiplizieren, um diese Resultate in Millimetern zu erhalten.

Nach Beseitigung derartiger systematischer Fehler verbleiben die zufälligen Fehler, die durch die Verschiedenheit der einzelnen Meßresultate bedingt sind. Wenn alle Werte $a_1, a_2, \ldots, a_n$ unter gleichen Bedingungen gewonnen wurden und in gleichem Maße Vertrauen verdienen, nimmt man als wahrscheinlichsten Wert der Größe x das arithmetische Mittel $a = (a_1 + a_2 + \cdots + a_n) : n$ der gemessenen Werte, kürzer $a = \frac{1}{n}\left(\sum_{i=1}^{n} a_i\right)$ geschrieben. Dabei brauchen die Werte a_i durchaus nicht sämtlich verschieden zu sein. Wenn a_1 sich n_1 mal, a_2 sich n_2 mal, allgemein a_i sich n_i mal wiederholt, so erhalten wir offensichtlich aus der Definition des arithmetischen Mittels

$$a = \frac{1}{n} \cdot \left(\sum_{i=1}^{k} a_i n_i\right),$$

wobei $n = n_1 + n_2 + \cdots + n_k$ ist.

Dieses in der Praxis weit verbreitete Verfahren ist theoretisch vollkommen gerechtfertigt.[1]) Daneben finden auch noch einige andere Methoden Verwendung; so ordnet man mitunter die erhaltenen Werte ihrer Größe nach und bildet den Mittelwert zwischen dem größten und dem kleinsten Wert (den „Median").

Nachdem wir angenommen haben, daß $x \approx a$, $a = (a_1 + a_2 + \cdots + a_n) : n$ ist, müssen wir feststellen, von welcher Genauigkeit diese Näherungsgleichung ist. In der Schule kann man eine Abschätzung der Genauigkeit dadurch gewinnen, daß man die Zahl a mit den einzelnen Meßresultaten $a_1, a_2, \ldots, a_n$ vergleicht. Dabei rundet man das Mittel a dadurch ab, daß man alle Ziffern beibehält, die in allen Werten a_i unverändert oder fast unverändert vorkommen, und alle folgenden wegläßt.

Daneben kann die Berechnung der Abweichungen vom Mittel, d. h. der Differenzen $a - a_i$, empfohlen werden. Die Summe dieser Differenzen ist, wie man leicht erkennt, gleich Null (Kontrolle!). Der Mittelwert aus den absoluten Werten dieser Differenzen („der mittlere Fehler") kann dann zur Charakterisierung der Genauigkeit des gefundenen Mittels benutzt werden. Mit einer bestimmten Wahrscheinlichkeit, die wir weiter unten berechnen werden, kann man behaupten, daß der wahre Wert von x von a um weniger als dieser mittlere Fehler verschieden ist.

[1]) Vgl. EdEM Bd. 6, B. W. GNEDENKO, Elemente der Wahrscheinlichkeitsrechnung und Statistik.

Beispiel: Eine zu ermittelnde Größe sei fünfmal gemessen worden. Die Meßresultate sind in der zweiten Spalte der folgenden Tabelle aufgeführt:

i	a_i	$a - a_i$	$(a - a_i)^2$
1	4,781	+ 0,0022	0,00000484
2	4,795	— 0,0118	13924
3	4,769	+ 0,0142	20164
4	4,792	— 0,0088	7744
5	4,779	+ 0,0042	1764
Summe der absoluten Werte	23,916	0,0412	0,00044080
Mittel	4,7832	0,00824	

Welche Ziffern des gefundenen Mittels $a = 4{,}7832$ kann man beibehalten? Die Ziffer der Einer (4) und der Zehntel (7) wiederholen sich in jedem einzelnen Meßresultat und sind daher unbedingt zuverlässig. Die Ziffern der Hundertstel schwanken, allerdings recht unbedeutend (von 6 bis 9), deshalb verdient die dritte Ziffer des Mittels (die Hundertstelziffer 8) auch einiges Vertrauen und wird beibehalten. Dann erhebt sich die Frage nach der Zuverlässigkeit der Tausendstelziffer (3). Sie ist schon sehr unzuverlässig und kann daher wegbleiben; üblicherweise zieht man es jedoch vor, auch die erste anzuzweifelnde Ziffer noch aufzuführen (meistens aus dem Bestreben, evtl. auftretende Rundungsfehler zu vermeiden). Die Ziffern dagegen, die rechts von der ersten anzuzweifelnden Ziffer stehen, im vorliegenden Falle also die Zehntausendstelziffer (2), werden weggelassen. Folglich ist im betrachteten Beispiel $x \approx a = 4{,}783$. Wir haben also eine Näherungszahl mit vier wesentlichen Ziffern (drei Dezimalstellen) erhalten, wobei die ersten drei Ziffern zuverlässig sind und die vierte zweifelhaft ist. Diese Folgerung wird auch durch die Berechnung des mittleren Fehlers bestätigt. Die Abweichungen vom Mittel sind in der vorstehenden Tabelle in der dritten Spalte aufgeführt. Die Summe der positiven Abweichungen ist + 0,0206, die der negativen Abweichungen — 0,0206, ihre Summe ist — wie es sein muß — Null. Mithin ist die Summe der absoluten Werte der Abweichungen $0{,}0206 \cdot 2 = 0{,}0412$ und der mittlere Fehler $0{,}0412 : 5 = 0{,}00824$; nach Abrunden auf eine wesentliche Ziffer erhalten wir also 0,008. Dieses Resultat kann als Bestätigung für die oben gezogenen Schlußfolgerungen über die Zuverlässigkeit der Ziffer der Hundertstel und die Fragwürdigkeit der Ziffer der Tausendstel angesehen werden.

Auf derartige Abschätzungen der Genauigkeit des arithmetischen Mittels muß man sich zwangsläufig in der Schule beschränken. Ein genaueres, in der Fehlerrechnung ganz wesentliches Verfahren für die Behandlung von gleichberechtigten Meßresultaten ist im folgenden enthalten (wir machen hier nur Angaben über die praktische Handhabung dieses Verfahrens und verweisen den, der sich näher mit der Theorie des Verfahrens vertraut machen möchte, auf das Buch [7]; zum Verständnis dieses Buches sind allerdings einige Kenntnisse aus der Wahrscheinlichkeitsrechnung erforderlich, die der Leser dem Buche [8] entnehmen kann[1]).

[1]) Vgl. auch EdEM Bd. 6, B. W. Gnedenko, Elemente der Wahrscheinlichkeitsrechnung und der Statistik.

Nachdem man den Mittelwert $a = \frac{\sum a_i}{n}$ und die Abweichungen vom Mittel $a - a_i$ bestimmt hat, berechnet man die Quadrate dieser Abweichungen und anschließend den „mittleren quadratischen Fehler" s gemäß der Formel $s^2 = \frac{\sum (a-a_i)^2}{n-1}$; diese Zahl s kann zur Charakterisierung der Genauigkeit der ganzen Meßreihe benutzt werden. Weiter findet man nach der Formel $s_a = s : \sqrt{n}$ „den mittleren quadratischen Fehler des arithmetischen Mittels". Die Wahrscheinlichkeit α dafür, daß die Ungleichung $a - \varepsilon < x < a + \varepsilon$ besteht, d. h. die Wahrscheinlichkeit dafür, daß der gesuchte Wert x von dem gefundenen Mittelwert a um weniger als eine gegebene Zahl ε (nach oben oder nach unten) abweicht, hängt dann allein von dem Verhältnis $t = \varepsilon : s_a$ und der Anzahl n der Messungen (oder — was dasselbe ist — von der Zahl $k = n - 1$) ab. Die Zahl α kann durch eine ziemlich komplizierte Formel ausgedrückt werden, und es gibt Tabellen, aus denen man für gegebenes $k = n - 1$ und t den Wert von α ersehen kann und mittels derer man auch umgekehrt für gegebenes k und α den Wert von t und damit von ε bestimmen kann. Wir geben einen Auszug aus einer solchen Tafel, der dem Buch [7] entnommen ist:

$k = n-1$	$t = 2{,}0$	$t = 2{,}5$	$t = 3{,}0$	$t = 3{,}5$	$k = n-1$	$t = 2{,}0$	$t = 2{,}5$	$t = 3{,}0$	$t = 3{,}5$
1	0,7048	0,7578	0,7952	0,8228	11	0,9292	0,9704	0,9880	0,9950
2	8164	8704	9046	9276	12	9314	9720	9890	9956
3	8606	9122	9424	9606	13	9332	9737	9898	9960
4	8838	9332	9600	9752	14	9348	9740	9904	9964
5	8980	9456	9700	9828	15	9360	9754	9910	9968
6	9076	9534	9760	9872	16	9372	9764	9916	9970
7	9144	9590	9800	9900	17	9382	9770	9920	9972
8	9194	9630	9830	9920	18	9392	9776	9924	9974
9	9234	9662	9850	9932	19	9400	9782	9926	9976
10	9266	9686	9866	9942	20	9408	9788	9930	9978
					∞	9545	9876	9973	9995

In dem oben betrachteten Beispiel war

$$n = 5,\ k = 4,\ a = 4{,}7832,\ s^2 = 0{,}00044080 : 4 = 0{,}00011020,\ s = 0{,}0105,$$

$$s_a = 0{,}0105 : \sqrt{5} = 0{,}0021 \cdot \sqrt{5} = 0{,}00469.$$

Wir wollen untersuchen, bei welchem Wert von ε die Wahrscheinlichkeit dafür, daß $a - \varepsilon < x < a + \varepsilon$ ist, gleich 0,96 ist. Die Tabelle zeigt, daß bei $k = 4$ die Wahrscheinlichkeit gleich 0,96 ist, wenn $t = 3{,}0$ und mithin $\varepsilon = t \cdot s_a = 3{,}0 \cdot 0{,}00469 \approx 0{,}0141$ ist. Folglich kann man mit einer Wahrscheinlichkeit 0,96 behaupten, daß der wahre Wert von x von dem gefundenen Mittel $a = 4{,}7832$ um weniger als $\varepsilon = 0{,}0141$ verschieden ist. Mit anderen Worten: Von 100 Chancen sind 96 dafür, daß x von $a = 4{,}7832$ um weniger als $\varepsilon = 0{,}0141$ verschieden ist, und nur 4 dagegen. Aus der Tabelle ist ersichtlich, daß bei konstantem t die Wahrscheinlichkeit α um so größer ist, je größer die Anzahl n der Messungen ist, und daß bei konstanter Anzahl n von Messungen die Wahrscheinlichkeit α mit der Zahl $t = \varepsilon : s_a$, d. h. mit wachsendem ε und fallendem s_a, wächst.

Untersuchen wir noch, wie groß die Wahrscheinlichkeit dafür ist, daß der wahre Wert von x vom Mittel $a = 4{,}7832$ um weniger als die Größe des gefundenen mittleren Fehlers 0,00824 abweicht. Jetzt ist $\varepsilon = 0{,}00824$, $t = 0{,}00824 : 0{,}00469 = 1{,}76$. Die Tabelle zeigt, daß dann α etwas kleiner als 0,8838, also vielleicht $\alpha \approx 0{,}85$ ist.

Zum Abschluß dieses Paragraphen bemerken wir, daß auch für die Schule jenes sehr einfache Verfahren, mit dessen Hilfe man die Genauigkeit einer einzelnen Messung abschätzen kann und von dem am Anfang dieses Paragraphen die Rede war, bereits eine große Bedeutung besitzt. Wenn wir dahin kommen, daß ein Schüler der Oberschule lernt, *bei jeder Messung eine Schranke des absoluten Fehlers anzugeben* oder — was dasselbe ist — eine untere und eine obere Schranke für die gesuchte, unbekannte Zahl zu bestimmen, so wird bereits ein großer Schritt in der Richtung getan sein, daß im Schulunterricht die Mathematik nicht nur rein formal übermittelt wird.

Kapitel II

FEHLERRECHNUNG

§ 8. Berechnungen mit strenger Fehlerrechnung nach der Methode der Wertschranken

Wenn wir irgendeine Berechnung mit angenäherten Daten durchführen, so ist das Resultat notwendigerweise auch nur angenähert. In diesem Resultat treten sowohl die Fehler der ursprünglichen Zahlen als auch etwaige „Berechnungsfehler" auf; letztere werden z. B. unvermeidlich durch das Auf- und Abrunden hervorgerufen, das im Verlauf der Rechnung vorgenommen wird. In diesem Zusammenhang taucht eine Frage von erstrangiger Bedeutung auf: Wie schätzt man die Genauigkeit des Resultates von Rechnungen mit solchen angenäherten Daten ab? Das sowohl hinsichtlich der Strenge als auch der Zugänglichkeit beste Verfahren einer derartigen „Fehlerrechnung" ist ohne Zweifel die *Methode der Wertschranken*. Wenn man eine untere und eine obere Schranke (*US* und *OS*) einer jeden der vorgegebenen Zahlen kennt, so kann man (zumindest in einfacheren Fällen) eine *US* und eine *OS* des Resultates jeder Grundrechenart an ihnen bestimmen und erhält letzten Endes eine *US* und eine *OS* des gesuchten Endergebnisses. Dieses Verfahren verwendete bereits ARCHIMEDES in seiner berühmten „Kreismessung". Er begnügte sich nicht damit, für das Verhältnis vom Umfang zum Durchmesser jedes Kreises den Näherungswert $22:7$ zu bestimmen, sondern zeigte, daß das wahre Verhältnis, das wir heute mit dem Buchstaben π bezeichnen, größer als $3\frac{10}{71}$ und kleiner als $3\frac{1}{7}$ ist, d. h., er stellte fest, daß $US(\pi) = 3\frac{10}{71}$ und $OS(\pi) = 3\frac{1}{7}$ ist.

„ARCHIMEDES berechnet der Reihe nach die Seite des umschriebenen Sechsecks, des Zwölfecks, des Vierundzwanzigecks, des Achtundvierzigecks und des Sechsundneunzigecks, ausgedrückt durch den Durchmesser, und zwar gibt er mit feinem mathematischem Gefühl das (immer nur näherungsweise bestimmbare) Verhältnis des Durchmessers zur Seite des umschriebenen Polygons jedesmal etwas zu klein an, wodurch er für den Umfang des betreffenden Polygons und um so mehr für den Kreisumfang jedesmal eine sichere obere Grenze gewinnt ... Um die untere Grenze für das Verhältnis des Kreisumfanges zum Durchmesser festzustellen, bedient sich ARCHIMEDES der entsprechenden eingeschriebenen Polygone, mit dem Sechseck beginnend und dem Sechsundneunzigeck abschließend. Bei diesen Berech-

nungen aber wählt ARCHIMEDES mit derselben bewußten Sicherheit die auftretenden Quadratwurzelwerte jedesmal so, daß die betreffende Polygonseite etwas zu klein angegeben wird."[1])

Vom theoretischen Standpunkt aus ist die Methode der Wertschranken äußerst einfach. Sie benutzt allein die bereits aus den ersten Klassen der Schule bekannten Sätze über die Änderung der Resultate der Grundrechenarten in Abhängigkeit von Änderungen der verknüpften Zahlen. Wenn wir uns auf die ersten vier Grundrechenarten beschränken, so erhalten wir die folgenden Sätze über Schranken (die Unbekannten x und y seien als positiv vorausgesetzt):

$$US(x+y) = US(x) + US(y) \qquad OS(x+y) = OS(x) + OS(y)$$
$$US(x-y) = US(x) - OS(y) \qquad OS(x-y) = OS(x) - US(y)$$
$$US(x\,y) = US(x) \cdot US(y) \qquad OS(x\,y) = OS(x) \cdot OS(y)$$
$$US\left(\frac{x}{y}\right) = \frac{US(x)}{OS(y)} \qquad OS\left(\frac{x}{y}\right) = \frac{OS(x)}{US(y)}$$

Hinzu kommen noch drei weitere Sätze, die unmittelbar aus der Definition von US und OS folgen: 1.) US darf nur abgerundet, OS nur aufgerundet werden. 2.) Je kleiner die Differenz $OS(x) - US(x)$ ist, um so genauer ist x bestimmt. 3.) Es empfiehlt sich, als Näherungswert von x das arithmetische Mittel der Zahlen $US(x)$ und $OS(x)$ oder eine Zahl, die diesem Mittel sehr nahe liegt, zu nehmen.

In einfachen Fällen bietet die Anwendung der Methode der unteren und oberen Schranke keine Schwierigkeiten, wie das folgende Beispiel vollständig erläutert:

Beispiel 1. Gesucht ist $x = \frac{a+b}{(a-b)\,c}$, wobei $a = 3\frac{4}{7}$, $b = 3\frac{5}{11}$, $c = 28\frac{1}{3}$ ist und die genauen Werte von a, b, c durch Näherungswerte ersetzt sind, die bis auf Hundertstel genau sind.

Lösung.

	US	OS
a	3,57	3,58
b	3,45	3,46
$a+b=m$	7,02	7,04
$a-b$	0,11	0,13
c	28,33	28,34
$(a-b)\,c=n$	3,11	3,69
$x=\frac{m}{n}$	1,90	2,27

$$\begin{array}{rr} 2{,}27 & 2{,}27 \\ +\ 1{,}90 & -\ 1{,}90 \\ \hline 4{,}17 : 2 = 2{,}085 & 0{,}37 : 2 = 0{,}185 \end{array}$$

$x \approx 2{,}085\ (\pm\ 0{,}185)$
$x \approx 2{,}1\ \ (\pm\ 0{,}2)$.

Erläuterung. Nachdem wir die Ungleichungskette $1{,}90 < x < 2{,}27$ gefunden haben, ist es natürlich, als Näherungswert für x das arithmetische Mittel aus den erhaltenen Schranken, nämlich 2,085, zu nehmen. Hierbei kann jedoch der Eindruck entstehen, es sei x auf Tausendstel genau bestimmt. Diese Täuschung wird dadurch vermieden, daß wir angeben, wie groß die größtmögliche Abweichung des wahren Wertes von diesem Mittel ist, d. h., daß wir eine *Schranke des absoluten Fehlers* des Mittels bestimmen. Eine solche Schranke ist, wie wir wissen, die halbe

[1]) Ф. Рудио, О квадратуре круга (F. RUDIO, Über die Quadratur des Kreises), übersetzt aus dem Deutschen unter der Redaktion und mit Anmerkungen von S. N. BERNSTEIN, 3. Aufl., GTTI 1936, S. 31—32. Originalausgabe: F. RUDIO, Das Problem von der Quadratur des Zirkels, Vierti. Natf. Ges. Zürich 35 (1890), 1—61, insbesondere S. 10, 11—12.

Differenz der erhaltenen oberen und unteren Schranke. Dann erhalten wir als Lösung die Näherungsgleichung $x \approx 2{,}085\ (\pm\ 0{,}185)$. Runden wir schließlich den gefundenen Wert noch so auf, daß nur eine anzuzweifelnde Ziffer verbleibt, so erhalten wir $x \approx 2{,}1$, und wegen $2{,}1 - 1{,}9 = 0{,}2$; $2{,}27 - 2{,}1 = 0{,}17$ schließlich $x \approx 2{,}1\ (\pm\ 0{,}2)$.

Da uns im vorliegenden Beispiel die genauen Werte a, b, c bekannt sind, können wir das erhaltene Resultat mit dem genauen Wert von x, der gleich $2\frac{31}{255} = 2{,}1215\cdots$ ist, vergleichen. Wir sehen, daß dieser genaue Wert wirklich innerhalb der von uns angegebenen Schranken 1,90 und 2,27 liegt. Der von uns gefundene Näherungswert 2,1 unterscheidet sich von dem wahren Wert sehr wenig. Auf Grund unserer allgemeinen Überlegungen können wir garantieren, daß er vom wahren Wert um weniger als 0,2 verschieden ist; tatsächlich übersteigt jedoch der Unterschied kaum 0,02. In unserem Fall stellt sich also die Abschätzung des Fehlers als zu grob heraus, eine Tatsache, die man übrigens fast immer beobachtet.

Als zweites Beispiel führen wir die vollständigen Rechnungen an, die für die Lösung der auf S. 337 behandelten Aufgabe erforderlich sind: Zu bestimmen war dort $\delta = p : v$, $v = \pi r^2 h$, wobei $2r \approx 0{,}48\ (\pm\ 0{,}005)$ mm, $h \approx 264{,}4\ (\pm\ 0{,}1)$ mm, $p \approx 0{,}423\ (\pm\ 0{,}001)$ g ist. Dabei müssen die Werte von r und h in Zentimetern ausgedrückt werden.

Die nachfolgenden Werte von πr^2 sind aus einer Tafel für Kreisflächen entnommen.

	US	*OS*
$2r$	0,0475	0,0485
πr^2	0,001772	0,001847
h	26,43	26,45
$v = \pi r^2 h$	0,0468	0,0489
p	0,422	0,424
$\delta = p : v$	8,63	9,05

9,05	9,05
$+\ 8{,}63$	$-\ 8{,}63$
$17{,}68 : 2 = 8{,}84$	$0{,}42 : 2 = 0{,}21$

$\delta \approx 8{,}84\ (\pm\ 0{,}21)$
$\delta \approx 8{,}8\ \ (\pm\ 0{,}25)$.

Schließlich wollen wir ein noch etwas schwierigeres Beispiel betrachten, das Schülern etwa der 9. und 10. Klassen zugänglich ist.

Beispiel 2. Zu berechnen sind auf vier Dezimalstellen genau die Werte von $\sin 3°$ und $\cos 3°$ unter Benutzung der Gleichungen $a_6 = r$ und $a_{10} = 0{,}5\, r\,(\sqrt{5} - 1)$, wobei a_6 bzw. a_{10} die Länge der Seiten des einem Kreis vom Radius r einbeschriebenen regelmäßigen Sechs- bzw. Zehnecks ist.

Lösung.

$\sin 30° = 0{,}5\, a_6 : r = 0{,}5$; $\cos 30° = \sqrt{1 - 0{,}25} = 0{,}5\,\sqrt{3}$;

$\sin 18° = 0{,}5\, a_{10} : r = 0{,}25\,(\sqrt{5} - 1)$; $\cos 18° = 0{,}25\,\sqrt{10 + \sqrt{20}}$;

$\sin 15° = \sqrt{0{,}5\,(1 - \cos 30°)} = 0{,}5\,\sqrt{2 - \sqrt{3}} = 0{,}25\,(\sqrt{6} - \sqrt{2})$;

$\cos 15° = \sqrt{0{,}5\,(1 + \cos 30°)} = 0{,}5\,\sqrt{2 + \sqrt{3}} = 0{,}25\,(\sqrt{6} + \sqrt{2})$;

$\sin 3° = \sin(18° - 15°) = \sin 18° \cos 15° - \cos 18° \sin 15°$;

$\cos 3° = \cos 18° \cos 15° + \sin 18° \sin 15°$.

Berechnung.

$$\sin 15° = 0{,}25\,(\sqrt{6} - \sqrt{2}),\quad \cos 15° = 0{,}25\,(\sqrt{6} + \sqrt{2}).$$

	US	*OS*
$\sqrt{6}$	2,44948	2,44950
$\sqrt{2}$	1,41420	1,41422
$a_1 = \sqrt{6} - \sqrt{2}$	1,03526	1,03530
$b_1 = \sqrt{6} + \sqrt{2}$	3,86368	3,86372
$\sin 15° = a_1 : 4$	0,25881	0,25883
$\cos 15° = b_1 : 4$	0,96592	0,96593

Resultat (auf vier Dezimalstellen genau):
$\sin 15° \approx 0{,}2588$,
$\cos 15° \approx 0{,}9659$.

Berechnung.

$$\sin 18° = 0{,}25\,(\sqrt{5} - 1),\quad \cos 18° = 0{,}25\,\sqrt{10 + \sqrt{20}}.$$

	US	*OS*
$\sqrt{5}$	2,23606	2,23608
$a_2 = \sqrt{5} - 1$	1,23606	1,23608
$\sin 18° = a_2 : 4$	0,30901	0,30902
$\sqrt{20}$	4,47213	4,47215
$b_2 = 10 + \sqrt{20}$	14,47213	14,47215
$c_2 = \sqrt{b_2}$	3,80421	3,80424
$\cos 18° = c_2 : 4$	0,95105	0,95106

Resultat (auf vier Dezimalstellen genau):
$\sin 18° \approx 0{,}3090$,
$\cos 18° \approx 0{,}9511$.

Berechnung.

$$\sin 3° = \sin 18° \cos 15° - \cos 18° \sin 15°,$$
$$\cos 3° = \cos 18° \cos 15° + \sin 18° \sin 15°.$$

	US	*OS*
$a_3 = \sin 18°$	0,30901	0,30902
$b_3 = \cos 15°$	0,96592	0,96593
$c_3 = \cos 18°$	0,95105	0,95106
$d_3 = \sin 15°$	0,25881	0,25883
$a_3 b_3$	0,29847	0,29849
$c_3 d_3$	0,24614	0,24616
$\sin 3° = a_3 b_3 - c_3 d_3$	0,05231	0,05235
$b_3 c_3$	0,91863	0,91865
$a_3 d_3$	0,07997	0,07998
$\cos 3° = b_3 c_3 + a_3 d_3$	0,99860	0,99863

Resultat (auf vier Dezimalstellen genau):
$\sin 3° \approx 0{,}0523$,
$\cos 3° \approx 0{,}9986$.

Zur Probe könnte man die Summe $s = \sin^2 3° + \cos^2 3°$ bilden und sich überzeugen, daß $US(s) < 1$ und $OS(s) > 1$ ist, wie es auch sein muß. Einfacher ist es jedoch, sich Auskunft aus einer Tafel zu holen. In einer vierstelligen Tafel finden wir die gerade von uns erhaltenen Werte, während wir in einer sechsstelligen Tafel (PETERS) die Werte $\sin 3° = 0{,}052336$, $\cos 3° = 0{,}998630$ finden, was unsere Ergebnisse vollständig rechtfertigt. Wir bemerken, daß ein noch genauerer Wert für $\cos 3°$ der Wert 0,99862995 ist.

§ 9. Berechnungen mit strenger Fehlerrechnung nach der Methode der Fehlerschranken

Bei aller Strenge und Zugänglichkeit erweisen sich Berechnungen nach der Methode der Wertschranken doch als sehr umfangreich, da man in jedem Schritt US und OS bestimmen muß, also jede Rechnung wiederholt durchzuführen ist. Daher erhebt sich unmittelbar die folgende Frage: Ist es nicht möglich, unter Umgehung der Berechnungen von $US(x)$ und $OS(x)$ eine Methode der Berechnung mit strenger Fehlerrechnung anzugeben, die es gestattet, einen Näherungswert und eine Schranke des absoluten Fehlers für die gesuchte Zahl in Abhängigkeit von Näherungswerten und Schranken des absoluten Fehlers der vorgegebenen Zahlen zu bestimmen?

Es sei etwa bekannt, daß $x \approx a\ (\pm \Delta a)$ und $y \approx b\ (\pm \Delta b)$ ist. Gefragt ist nach Näherungswerten für $x + y$, $x - y$, xy, $\frac{x}{y}$, x^n, $\sqrt[n]{x}$ (n natürliche Zahl) und nach einer Charakterisierung ihrer Genauigkeit.

Zunächst erhalten wir aus den Ungleichungen

$$a - \Delta a < x < a + \Delta a \tag{1}$$

$$b - \Delta b < y < b + \Delta b \tag{2}$$

durch gliedweise Addition

$$a + b - (\Delta a + \Delta b) < x + y < a + b + (\Delta a + \Delta b). \tag{3}$$

Ferner erhält man, wenn man alle Glieder der Ungleichungen (2) mit (-1) multipliziert,

$$-b - \Delta b < -y < -b + \Delta b$$

und nach Addition zu (1)

$$a - b - (\Delta a + \Delta b) < x - y < a - b + (\Delta a + \Delta b). \tag{4}$$

Aus den Ungleichungen (3) und (4) erhalten wir den folgenden

Satz 1. *Jede Summe aus jeweils einer Schranke des absoluten Fehlers zweier Näherungszahlen ist eine Schranke des absoluten Fehlers der Summe und der Differenz dieser Näherungszahlen.*

Der angegebene Satz läßt sich offensichtlich unmittelbar auch auf Summen beliebig vieler Näherungszahlen verallgemeinern.

Nehmen wir alle Glieder der Ungleichungen (1) und (2) als positiv an und multiplizieren wir sie gliedweise, so erhalten wir

$$a\,b - (a\,\Delta b + b\,\Delta a) + \Delta a\,\Delta b < x\,y < a\,b + (a\,\Delta b + b\,\Delta a) + \Delta a\,\Delta b. \tag{5}$$

Das Produkt $\Delta a \Delta b$ auf der linken Seite kann unbeschadet der Gültigkeit der Ungleichungen wegbleiben. Wenn die Zahlen Δa und Δb bedeutend kleiner als a und b sind, was in der Praxis stets der Fall zu sein pflegt, so lassen wir jedoch auch das Produkt $\Delta a \Delta b$ auf der rechten Seite als eine Zahl „zweiter (Größen-)Ordnung" fort. Dadurch entziehen wir allerdings die Methode der Fehlerschranken jener einwandfreien Strenge, die dem oben entwickelten Verfahren der Wertschranken innewohnt. Wir erhalten somit die neuen Ungleichungen

$$a\,b - (a\,\Delta b + b\,\Delta a) < x\,y < a\,b + (a\,\Delta b + b\,\Delta a), \tag{6}$$

oder nach einer einfachen Umformung

$$-\left(\frac{\Delta a}{a} + \frac{\Delta b}{b}\right) < \frac{x\,y - a\,b}{a\,b} < \frac{\Delta a}{a} + \frac{\Delta b}{b}. \tag{7}$$

Diese letzten Ungleichungen führen uns auf den folgenden

Satz 2. *Jede Summe aus jeweils einer Schranke des relativen Fehlers zweier Näherungszahlen ist eine Schranke des relativen Fehlers ihres Produktes.*

Um zu einem analogen Satz über den Quotienten zu gelangen, wollen wir zunächst eine Schranke des relativen Fehlers der Zahl $1 : b$ in Abhängigkeit von gegebener Schranke des relativen Fehlers $\frac{\Delta b}{b}$ der Zahl b bestimmen.

Wenn wir wiederum alle Glieder der Ungleichungen (2) als positiv annehmen, so können wir sie auch in der Form

$$\frac{1}{b + \Delta b} < \frac{1}{y} < \frac{1}{b - \Delta b}$$

schreiben und erhalten der Reihe nach

$$\frac{1}{b+\Delta b} - \frac{1}{b} < \frac{1}{y} - \frac{1}{b} < \frac{1}{b-\Delta b} - \frac{1}{b},$$

$$-\frac{\Delta b}{b\,(b+\Delta b)} < \frac{1}{y} - \frac{1}{b} < \frac{\Delta b}{b\,(b-\Delta b)},$$

$$-\frac{\Delta b}{b+\Delta b} < \left(\frac{1}{y} - \frac{1}{b}\right) : \frac{1}{b} < \frac{\Delta b}{b-\Delta b}.$$

In der letzten Ungleichungskette unterscheidet sich der Bruch auf der linken Seite von dem größeren Bruch $\frac{\Delta b}{b}$ nur durch eine Zahl zweiter Größenordnung, nämlich $\frac{(\Delta b)^2}{b\,(b+\Delta b)}$, und entsprechend der Bruch auf der rechten Seite von dem kleineren Bruch $\frac{\Delta b}{b}$ um die Zahl $\frac{(\Delta b)^2}{b\,(b-\Delta b)}$ zweiter Größenordnung. Durch einen erneuten geringfügigen Verstoß gegen die Strenge der Überlegungen gelangen wir also zu den Ungleichungen

$$-\frac{\Delta b}{b} < \left(\frac{1}{y} - \frac{1}{b}\right) : \frac{1}{b} < \frac{\Delta b}{b}, \tag{8}$$

gemäß denen jede Schranke des relativen Fehlers von b auch eine Schranke des relativen Fehlers von $1 : b$ ist. Indem wir den Quotienten $\frac{a}{b}$ als Produkt $a \cdot \frac{1}{b}$ auffassen, gelangen wir zu

Satz 3. *Jede Summe aus jeweils einer Schranke des relativen Fehlers zweier Näherungszahlen ist eine Schranke des relativen Fehlers ihres Quotienten.*

Eine einfache Folgerung aus Satz 2 ist die Ungleichungskette

$$\frac{-n\,\Delta a}{a} < \frac{x^n - a^n}{a^n} < \frac{n\,\Delta a}{a},$$

d. h.,

Satz 4. *Jedes Produkt aus einer Schranke des relativen Fehlers der Basis mit dem (natürlichen) Exponenten ist eine Schranke des relativen Fehlers der Potenz.*

Setzen wir schließlich $y = \sqrt[n]{x}$, so ist $x = y^n$, $\frac{\Delta x}{x} = \frac{n\,\Delta y}{y}$, und mithin

$$\frac{\Delta y}{y} = \frac{\Delta x}{x} : n,$$

so daß wir auf folgenden Satz geführt werden:

Satz 5. *Jeder Quotient aus einer Schranke des relativen Fehlers des Radikanden und dem (natürlichen) Wurzelexponenten ist eine Schranke des relativen Fehlers der Wurzel.*

Wir wollen die Anwendung dieser Sätze an einem einfachen Beispiel demonstrieren:

Beispiel 1. Zu berechnen ist unter Angabe einer Schranke des relativen Fehlers

$$t = \sqrt{\frac{2\,h\,d}{g\,(d - d_1)}},$$

wobei $h \approx 25{,}3\ (\pm 0{,}1)$, $d \approx 19{,}32\ (\pm 0{,}01)$,
$d_1 \approx 0{,}998\ (\pm 0{,}0005)$, $g \approx 982\ (\pm 0{,}5)$
sei.

Lösung. Hier ist

$$\frac{\Delta t}{t} = \frac{0{,}5\,\Delta u}{u},\quad u = \frac{2\,h\,d}{g\,(d - d_1)}\ \text{(nach Satz 5)};$$

$$\frac{\Delta u}{u} = \frac{\Delta\,(2\,h\,d)}{2\,h\,d} + \frac{\Delta\,(g\,v)}{g\,v},\quad v = d - d_1\ \text{(nach Satz 3)};$$

$$\frac{\Delta\,(2\,h\,d)}{2\,h\,d} = \frac{\Delta h}{h} + \frac{\Delta d}{d}\ \text{(nach Satz 2 unter Berücksichtigung von } \Delta 2 = 0);$$

$$\frac{\Delta\,(g\,v)}{g\,v} = \frac{\Delta g}{g} + \frac{\Delta v}{v}\ \text{(nach Satz 2)};$$

$$\Delta v = \Delta\,(d - d_1) = \Delta d + \Delta d_1\ \text{(nach Satz 1)}.$$

Also schließlich $\frac{\Delta t}{t} = 0{,}5\left[\frac{\Delta h}{h} + \frac{\Delta d}{d} + \frac{\Delta g}{g} + \frac{\Delta d + \Delta d_1}{d - d_1}\right]$.

Den Näherungswert bestimmen wir mit Hilfe von vierstelligen Logarithmen, die Schranke des relativen Fehlers mit dem Rechenschieber.

d	19,32	lg 2	0,3010
d_1	0,998	lg h	1,4031
$v = d - d_1$	18,322	lg d	1,2860
lg v	1,2630	$-$ lg $(g\,v)$	$\bar{5}$,7449
lg g	2,9921	lg t^2	$\bar{2}$,7350
lg $(g\,v)$	4,2551	lg t	$\bar{1}$,3675
		t	0,2331

$\Delta h = 0{,}1$; $\Delta d = 0{,}01$; $\Delta d_1 = 0{,}0005$; $\Delta g = 0{,}5$;

$$\begin{aligned}\Delta t : t &= 0{,}5 \cdot [0{,}396 + 0{,}052 + 0{,}051 + 0{,}057]\,\% \\ &= 0{,}5 \cdot 0{,}454\% \\ &= 0{,}227\%\end{aligned}$$

$\Delta t = 0{,}227\%$ von 0,2331
$= 0{,}00053$.

Antwort.

$$t \approx 0{,}2331\ (\pm\ 0{,}00053) \approx 0{,}233\ (\pm\ 0{,}00053)$$

oder schließlich

$$t \approx 0{,}233\ (\pm\ 0{,}001).$$

Die vierte wesentliche Ziffer des Ergebnisses, das mit Hilfe von vierstelligen Logarithmen bestimmt wurde, ist unzuverlässig und daher weggelassen worden.

Zur Probe und zum Vergleich der Methoden lösen wir dieselbe Aufgabe nochmals mittels des Verfahrens der Wertschranken und ohne Benutzung von Logarithmen.

	US	*OS*
h	25,2	25,4
$2h$	50,4	50,8
d	19,31	19,33
$2hd = a$	973,224	981,964
d_1	0,9975	0,9985
$d - d_1$	18,3115	18,3325
g	981,5	982,5
$g(d - d_1) = b$	17972	18012
$a : b = t^2$	0,05403	0,05465
t	0,2325	0,2338

$$t \approx 0{,}2331\ (\pm\ 0{,}0007)$$

$$\begin{array}{r} 0{,}2338 \\ +\ 0{,}2325 \\ \hline 0{,}4663 : 2 = 0{,}23315 \end{array}$$

$$\begin{array}{r} 0{,}2338 \\ -\ 0{,}2325 \\ \hline 0{,}0013 : 2 = 0{,}00065 \end{array}$$

Hier sind alle Fehlerquellen berücksichtigt. Das Resultat stimmt in beiden Fällen praktisch überein.

Alle Formeln für die Berechnung einer Schranke des Fehlers, die wir oben durch ganz elementare Überlegungen abgeleitet haben, lassen sich unter Benutzung der Differentialrechnung wesentlich einfacher ableiten.

Dazu sei $f(x, y)$ eine gegebene differenzierbare Funktion der beiden Veränderlichen x und y; x_0 und y_0 seien beliebige Elemente aus ihrem Definitionsbereich. Wir setzen dann $x = x_0 + \alpha$, $y = y_0 + \beta$, $|\alpha| < \Delta x$, $|\beta| < \Delta y$, wobei wir die gegebenen Zahlen Δx und Δy als so klein im Verhältnis zu x_0 und y_0 annehmen, daß ihre Potenzen und Produkte vernachlässigt werden können. Wir stellen uns die Aufgabe, den größten Wert des absoluten Betrages der Differenz $f(x, y) - f(x_0, y_0)$ zu berechnen, unter der Nebenbedingung, daß der Zuwachs der „Inkremente" $\alpha = x - x_0$ und $\beta = y - y_0$ absolut genommen die Werte Δx bzw. Δy nicht überschreitet.

Wie aus der Differentialrechnung[1]) bekannt ist, setzt sich die Differenz $f(x, y) - f(x_0, y_0)$ (der „Zuwachs" der Funktion $f(x, y)$) aus zwei Anteilen zusammen: Erstens aus einem Hauptanteil, den man *vollständiges Differential* der Funktion $f(x, y)$ nennt und der sich nach der Formel

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy$$

berechnet (x und y sind dabei in den Ableitungen an den Stellen x_0 und y_0 zu nehmen), und zweitens aus Gliedern höherer Ordnung. Hierbei sind dx und dy die Differentiale der Argumente x und y oder, was dasselbe ist, die Änderungen der Inkremente $x - x_0$ und $y - y_0$, die wir hier mit den Buchstaben α und β bezeichnet haben (während die Zeichen Δx und Δy zur Bezeichnung der oberen Grenzen dieser Änderungen dienen). Wenn Δx und Δy sehr kleine Zahlen sind, was wir im folgenden voraussetzen wollen, so sind auch die Zahlen α und β sehr klein, und alle Glieder höherer Ordnung können vernachlässigt werden. Wir finden also, daß

$$f(x_0 + \alpha, y_0 + \beta) - f(x_0, y_0) \approx \frac{\partial f}{\partial x}\,\alpha + \frac{\partial f}{\partial y}\,\beta$$

ist.

[1]) Vgl. EdEM Bd. 3, Artikel „Differential- und Integralrechnung".

Wenden wir hierauf den bekannten Satz über den absoluten Betrag einer Summe an („Der Betrag einer Summe ist höchstens gleich der Summe der Beträge der Summanden") und beachten wir, daß der absolute Betrag eines Produktes gleich dem Produkt der absoluten Beträge der Faktoren ist, so erhalten wir, wenn wir noch $|\alpha|$ durch Δx und $|\beta|$ durch Δy abschätzen,

$$|f(x_0+\alpha, y_0+\beta) - f(x_0, y_0)| \leq \left|\frac{\partial f}{\partial x}\right| \cdot |\alpha| + \left|\frac{\partial f}{\partial y}\right| \cdot |\beta|$$
$$< \left|\frac{\partial f}{\partial x}\right| \cdot \Delta x + \left|\frac{\partial f}{\partial y}\right| \cdot \Delta y.$$

Der zahlenmäßige Wert dieses letzten Ausdrucks (den man erhält, wenn man x und y durch x_0 bzw. y_0 ersetzt) ist dem Betrage nach größer als alle unter den angegebenen Voraussetzungen auftretenden Werte der Differenz zwischen dem unbekannten genauen Wert $f(x_0+\alpha, y_0+\beta)$ und dem bekannten Näherungswert $f(x_0, y_0)$, kann also als eine Schranke des absoluten Fehlers der Zahl $f(x_0, y_0)$ als Näherungswert für $f(x, y)$ aufgefaßt werden.

Dieser Schluß läßt sich leicht von einer Funktion zweier Variablen auf eine Funktion mit beliebig vielen Argumenten übertragen, wobei wir die Formel

$$\Delta f(x_0, y_0, z_0, \ldots) \approx \left|\frac{\partial f}{\partial x}\right| \cdot \Delta x + \left|\frac{\partial f}{\partial y}\right| \cdot \Delta y + \left|\frac{\partial f}{\partial z}\right| \cdot \Delta z + \cdots \qquad \text{(A)}$$

erhalten, mit deren Hilfe sich gleichfalls eine Schranke des absoluten Fehlers berechnen läßt. Die einzelnen Glieder auf der rechten Seite geben den Anteil jeder einzelnen Variablen am Gesamtfehler an. Lassen wir in (A) die Funktion $f(x, y)$ der Reihe nach die Funktionen $x+y$, $x-y$, xy, $\frac{x}{y}$, x^n, $\sqrt[n]{x}$ durchlaufen, so kommen wir unmittelbar auf die oben angegebenen Sätze 1 bis 5. Setzen wir z. B. $f(x, y) = \frac{x}{y}$, so wird $\frac{\partial f}{\partial x} = \frac{1}{y}$, $\frac{\partial f}{\partial y} = -\frac{x}{y^2}$, und die Formel (A) liefert (für positives x und y): $\Delta\left(\frac{x}{y}\right) = \frac{1}{y}\Delta x + \frac{x}{y^2}\Delta y$ oder, nach Division durch $\frac{x}{y}$,

$$\Delta\left(\frac{x}{y}\right) : \frac{x}{y} = \frac{\Delta x}{x} + \frac{\Delta y}{y},$$

d. h. den Satz 3.

Hier noch ein Beispiel einer unmittelbaren Anwendung der Formel (A).

Beispiel 2. Zu berechnen ist die Seite t eines Dreiecks, wenn man seine anderen beiden Seiten $x \approx 25{,}0(\pm 0{,}2)$ mm, $y \approx 30{,}0(\pm 0{,}2)$ mm und den Winkel zwischen ihnen

$$z \approx 60°{,}0\ (\pm 0°{,}5)$$

kennt.

Lösung. Zur Berechnung eines Näherungswertes verwenden wir die Formel

$$t^2 = x^2 + y^2 - 2xy\cos z$$

und benutzen eine vierstellige Tafel. Dann erhalten wir $t \approx 27{,}84$ mm. Die Differentiation ergibt:

$$\frac{\partial t}{\partial x} = \frac{x - y\cos z}{t}, \quad \frac{\partial t}{\partial y} = \frac{y - x\cos z}{t}, \quad \frac{\partial t}{\partial z} = \frac{xy\sin z}{t},$$

und wir finden als Schranke des absoluten Fehlers des Näherungswertes t gemäß Formel (A)

$$\Delta t = 0{,}359 \cdot 0{,}2 + 0{,}628 \cdot 0{,}2 + 23{,}3 \cdot 0{,}00873 = 0{,}072 + 0{,}126 + 0{,}203 = 0{,}401.$$

Hierbei ist Δz im Bogenmaß ausgedrückt.

Wir finden also das Resultat $t \approx 27{,}84\ (\pm 0{,}401)$ mm oder nach dem üblichen Runden $t \approx 27{,}8(\pm 0{,}40)$ mm; d. h., man kann garantieren, daß die Länge der gesuchten Dreiecksseite von 27,8 mm um weniger als 0,40 mm abweicht. Dasselbe Ergebnis erhält man bei Anwendung der Methode der Wertschranken.

Bisher hatten wir es mit Aufgaben zu tun, in denen aus bekannten Schranken des Fehlers der gegebenen Zahlen eine Schranke des Fehlers im Resultat zu bestimmen war. Manchmal hat man indes auch die umgekehrte Aufgabe zu lösen,

d. h., man hat festzustellen, welche Genauigkeit die Ausgangszahlen haben müssen, damit das Resultat eine vorgegebene Genauigkeit besitzt. Bei der Lösung derartiger Aufgaben besitzt nun die Methode der Fehlerschranken erhebliche Vorzüge gegenüber dem Verfahren der Wertschranken. Wir wollen hier auf derartige Aufgaben nicht näher eingehen, sondern uns auf die Betrachtung eines Beispiels beschränken. Der Leser, der sich mit Einzelheiten vertraut machen möchte, sei auf die Bücher [9] und [2b] verwiesen.

Mit welcher Genauigkeit muß man das Gewicht p (in Gramm) und das Volumen v (in Kubikzentimeter) eines Stückes Blei ermitteln, um sein spezifisches Gewicht δ nach der Formel $\delta = \frac{p}{v}$ mit einem Fehler von weniger als einem halben Prozent berechnen zu können?

Auf Grund von Satz 2 ist

$$\frac{\Delta\delta}{\delta} = \frac{\Delta p}{p} + \frac{\Delta v}{v},$$

d. h., man muß gemäß unserer Aufgabe die Werte p und v so genau bestimmen, daß die Summe der Schranken ihrer relativen Fehler 0,5% nicht überschreitet. Da man beim Wägen mit Leichtigkeit eine viel größere Genauigkeit erreichen kann als bei der Volumenmessung, wird man den zulässigen Höchstfehler bei der Bestimmung des Gewichtes zu ein Zehntel des Gesamtfehlers ansetzen, also zu 0,05%, und die restlichen 0,45% dem Fehler bei der Bestimmung des Volumens zuschreiben. Wenn also das Gewicht des betreffenden Stückes Blei sich grob genähert als 400 g herausstellt und sein Volumen nahe an 40 cm³ liegt, so ist das Gewicht mit einem Fehler zu bestimmen, der 0,05% von 400 g, also 0,2 g nicht überschreitet, das Volumen dagegen mit einem Fehler, der 0,18 cm³ nicht überschreitet. Die verlangte Genauigkeit bei der Bestimmung des spezifischen Gewichts läßt sich also garantiert dann erreichen, wenn eine Waage zur Verfügung steht, die bei einer Belastung mit 400 g auf 0,2 g anschlägt, und wenn ein Gerät zur Messung des Volumens vorhanden ist, das mit einer Genauigkeit bis zu 0,1 cm³ abzulesen gestattet.

Wir haben bisher zwei Verfahren kennengelernt, mit deren Hilfe man ganz bestimmte Schlüsse über die Genauigkeit der Rechenresultate ziehen kann, wenn man die Genauigkeit der gegebenen Zahlen kennt, d. h., wir haben zwei Methoden der „genauen Fehlerrechnung" kennengelernt. Welches der beiden Verfahren, die Methode der Wertschranken oder die Methode der Fehlerschranken, verdient nun den Vorzug?

Augenscheinliche Vorzüge der Methode der Wertschranken sind: 1.) Die außerordentliche Einfachheit, mit der man ihre ganze Theorie auf ein Grundprinzip zurückführen kann, dessen praktische Anwendung selbst dem nur wenig ausgebildeten Rechner keinerlei Schwierigkeiten bereitet; 2.) ihre uneingeschränkte Anwendbarkeit auf alle numerischen Rechnungen, von den einfachsten bis zu den kompliziertesten; 3.) die absolute Strenge, mit der man bei Anwendung dieser Methode unbedingt zuverlässige Resultate erhalten kann, dank der Möglichkeit, sowohl die Fehler, die von der Ungenauigkeit der Ausgangszahlen herrühren, als auch die Rundungsfehler zu berücksichtigen; 4.) die Kontrollmöglichkeit für die Richtigkeit der Rechnungen, die aus dem Vergleich der Resultate der beiden parallelen Reihen von Operationen fließt.

Demgegenüber übertrifft die Methode der Fehlerschranken das Verfahren der Wertschranken in der Hinsicht, daß sie 1.) erlaubt, im voraus den Fehler zu berechnen, der von der Ungenauigkeit der vorgegebenen Zahlen herrührt, und dadurch eine mehr oder weniger zuverlässige Aussage über die Genauigkeit macht, mit der die Rechnung durchzuführen ist; daß sie 2.) festzustellen erlaubt, welcher Anteil des Gesamtfehlers im Resultat von den Fehlern der einzelnen vorgegebenen Näherungszahlen hervorgerufen wird.

Auf den ersten Blick scheint es, daß ein wesentlicher Nachteil der Methode der Wertschranken in der Notwendigkeit liegt, alle Berechnungen doppelt durchführen zu müssen. Wenn wir jedoch an ein und derselben Aufgabe die Fehlerrechnung nach der Methode der Wertschranken mit der Fehlerrechnung nach der Methode der Fehlerschranken vergleichen, so finden wir, daß die Quantität der Rechnungen in beiden Fällen ungefähr die gleiche ist. Das liegt daran, daß die Berechnung einer Schranke des Fehlers ebenfalls einige Mühe erfordert. Zwar kann diese Berechnung häufig dadurch wesentlich vereinfacht werden, daß man gröbere Näherungen benutzt, jedoch ergeben sich dadurch entweder sehr unsichere Resultate oder die Schranken des Fehlers wachsen unnütz an. Allerdings bemerkt man, daß bei Berechnungen nach einer Formel, die nur Grundrechenarten zweiter und dritter Stufe (Multiplikation, Division und Potenzieren, Radizieren) enthält, die Berechnungen nach der Methode der Fehlerschranken im allgemeinen entschieden schneller durchgeführt werden können als nach der Methode der Wertschranken. Anders steht es, wenn in die Formel neben den Grundrechenarten zweiter und dritter Stufe auch die Grundrechenarten erster Stufe (Addition, Subtraktion) eingehen.

In Fällen, in denen es nicht auf eine absolute Sicherheit, sondern nur auf eine mehr oder weniger große Wahrscheinlichkeit des Ergebnisses ankommt, wie dies z. B. bei der Auswertung von Versuchen und Beobachtungen der Fall ist, benutzt man meistens die Methode der Fehlerschranken. Bei Berechnungen dagegen, bei denen eine absolute Sicherheit unumgänglich (und im Wesen der Sache möglich) ist, wie z. B. bei der Berechnung mathematischer Tafeln, ist es besser, die Methode der Wertschranken zu benutzen. In didaktischer Hinsicht hat die Methode der Wertschranken unbestreitbare Vorteile gegenüber dem Verfahren der Fehlerschranken, so daß man, zumal für eine erste Einführung in die Methoden der genauen Fehlerrechnung, unbedingt das Verfahren der Wertschranken empfehlen muß.

§ 10. Höchstfehler der Resultate der Grundrechenarten an Näherungswerten. Regeln der Ziffernzählung

Wenn bekannt ist, wie viele zuverlässige Ziffern jede vorgegebene Näherungszahl besitzt, und wenn außerdem diese Ziffern selbst bekannt sind, so können wir in jedem Einzelfall auf Grund der oben betrachteten Verfahren feststellen, wieviel Ziffern des Resultates Vertrauen verdienen, und wir können es dann entsprechend runden. In diesem Zusammenhang erhebt sich naturgemäß die folgende Frage: Kann man aus der Kenntnis der Anzahl der zuverlässigen Ziffern jeder gegebenen Zahl ohne Kenntnis dieser Zahlen selbst irgendwelche Schlüsse über die Genauigkeit

des Resultates ziehen? Es zeigt sich, daß derartige Schlüsse durchaus möglich sind und daß sie in der Praxis des Rechnens mit Erfolg angewendet werden können. Meistens spricht man in diesem Zusammenhang von „Regeln der Ziffernzählung" bzw. von „Regeln für das Runden der Resultate der Grundrechenarten an Näherungszahlen". Um zu diesen Regeln zu gelangen, müssen wir feststellen, welchen maximalen Fehler die Resultate der Grundrechenarten annehmen können, wenn die Anzahl von zuverlässigen Ziffern in den Komponenten gegeben ist. Wir werden also unter dem „Höchstfehler" des Resultates einer jeden Grundrechenart die obere Grenze seines absoluten Fehlers bei beliebigen Komponenten mit vorgegebener Anzahl von zuverlässigen Ziffern verstehen und uns mit der Berechnung dieses Höchstfehlers in den verschiedenen Fällen beschäftigen. Den Höchstfehler werden wir stets mit dem Buchstaben ε bezeichnen.

Sehr einfach läßt sich der Höchstfehler einer algebraischen Summe berechnen. Sind die Komponenten (Summanden und Subtrahenden) mit einer gegebenen Anzahl von Dezimalstellen vorgegeben, so ist $\varepsilon = 0{,}5 \cdot n \cdot 10^{-k}$, wenn die Komponente mit der kleinsten Anzahl von Dezimalstellen k Dezimalstellen (k nicht-negativ, ganz) besitzt und im ganzen n Komponenten vorhanden sind. Der wirkliche absolute Fehler der Summe ist dann und nur dann gleich diesem Höchstfehler, wenn alle Komponenten k Dezimalstellen besitzen, jede den maximalen Fehler von einer halben Einheit der Ordnung der letzten Ziffer annimmt und alle Fehler dasselbe Vorzeichen besitzen. In diesem Fall kann der Höchstfehler im allgemeinen nicht herabgesetzt werden.

Diese Überlegungen geben uns bereits eine hinreichende Begründung für die folgende praktische Regel:

Regel 1 der Ziffernzählung. *Bei der Addition und Subtraktion von Näherungszahlen sind im Resultat nur so viele Dezimalstellen beizubehalten, wie in einer vorgegebenen Näherungszahl mit kleinster Anzahl von Dezimalstellen vorhanden sind.*

Wir erinnern bei dieser Gelegenheit daran, daß man unter den Dezimalstellen einer Zahl die Ziffern der Zahl versteht, die rechts vom Komma stehen. Dabei denken wir uns alle vorgegebenen Näherungszahlen so vorgegeben, daß nur diejenigen Ziffern aufgeführt sind, die Vertrauen verdienen. Hierbei empfiehlt es sich, ganze Zahlen, in denen rechts eine gewisse Anzahl von Nullen unbekannte Ziffern ersetzt, in Form eines Produktes mit einer gewissen Zehnerpotenz mit ganzzahligem Exponenten zu schreiben (z. B. ist es besser, die Näherungszahl 347000, falls sie Ergebnis des Rundens einer gewissen genauen Zahl auf die Ordnung der Tausender ist, in der Form $347 \cdot 10^3$ oder $3{,}47 \cdot 10^5$ o. dgl. zu schreiben).

Die Zulässigkeit des Rundens gemäß der angegebenen Regel wird augenscheinlich, wenn man irgendein konkretes Beispiel betrachtet. Dabei empfiehlt es sich, die unbekannten Ziffern der vorgegebenen Näherungszahlen durch ein besonderes Zeichen, etwa durch ein Fragezeichen, anzudeuten. Es sei z. B. nach der Summe der unten angegebenen Näherungszahlen gefragt, von denen die erste der auf drei Dezimalstellen gerundete Wert einer uns unbekannten Zahl ist, die zweite entsprechend der auf eine Dezimalstelle und die dritte der auf zwei Dezimalstellen gerundete Wert uns unbekannter Zahlen ist. Würden wir im vorliegenden Falle die Addition so ausführen, als wenn es sich um genaue Komponenten handelte, so erhielten wir die Zahl 87,943, in der naturgemäß die Hundertstel- und Tausendstel-

ziffern keinerlei Vertrauen verdienen und wegzulassen sind, wie gerade die Regel 1 verlangt.

$$\begin{array}{r} 0{,}423\,?\,. \\ +\,72{,}8\,?\,?\,?\,.\,. \\ \hline 14{,}72\,?\,?\,.\,. \\ \hline 87{,}943\,?\,.\,. \\ \hline 87{,}9 \end{array}$$

Im vorliegenden Beispiel kann der wahre absolute Fehler der Summe die halbe Einheit der letzten aufgeführten Ordnung nur wenig übertreffen. Man kann jedoch leicht Beispiele konstruieren, in denen er einige Einheiten dieser Ordnung ausmacht. Inwieweit verdient also die letzte Ziffer Vertrauen? Auf diese Frage werden wir in den §§ 11 und 12 zu sprechen kommen.

Wir gehen nun zur Multiplikation über und formulieren folgenden Satz über den Höchstfehler:

Satz 1. *Das Produkt zweier Näherungszahlen mit je k zuverlässigen Ziffern besitzt einen Höchstfehler von 5,5 Einheiten der Ordnung seiner k-ten wesentlichen Ziffer; dieser Wert des Höchstfehlers vermindert sich auf 5,05 Einheiten, wenn eine der Näherungszahlen k zuverlässige Ziffern und die andere $(k+1)$ zuverlässige Ziffern besitzt; er vermindert sich auf 5 Einheiten, wenn einer der Faktoren k zuverlässige Ziffern besitzt und der andere genau ist.*

Hier zunächst ein Beispiel, in dem der absolute Fehler des Produktes in der Nähe des im Satz angegebenen Höchstfehlers liegt:

$$x = 100{,}499,\quad y = 9{,}99499,\quad x\,y = 1004{,}486\cdots,\quad a = 100,\quad b = 9{,}99,\quad a\,b = 999.$$

Hier weicht das Produkt der Näherungszahlen a, b von dem Produkt der genauen Zahlen x, y um $5{,}486\cdots$ Einheiten der Ordnung der dritten wesentlichen Ziffer des Produktes ab.

Wir wollen den Beweis von Satz 1 für den Fall $k = 3$ ausführen (man sieht leicht, daß der angegebene Beweis auch für beliebiges k Gültigkeit hat und dort nur etwas größere Schreibarbeit erfordert). Zunächst ist unmittelbar klar, daß die Stellung des Kommas in den gegebenen Faktoren gleichgültig ist; um etwas Bestimmtes vor Augen zu haben, wollen wir voraussetzen, daß das Komma im ersten Faktor hinter der dritten wesentlichen Ziffer und im zweiten Faktor hinter der ersten wesentlichen Ziffer steht. Die genauen Werte der angenäherten Faktoren a, b seien x und y, so daß $x = a + \alpha$ und $y = b + \beta$ ist, wobei α und β die wirklichen absoluten Fehler der Näherungszahlen a und b sind. Gemäß unserer Voraussetzung über die Kommastellung ist dann

$$100 \le a \le 999;\ 1{,}00 \le b \le 9{,}99;\ |\alpha| \le 0{,}5;\ |\beta| \le 0{,}005.$$

Wir wollen nun eine Abschätzung der Differenz $xy - ab = a\beta + b\alpha + \alpha\beta$ in Einheiten der Ordnung der dritten wesentlichen Ziffer des Produktes ab geben. Offensichtlich erhalten wir:

$$|xy - ab| \le 0{,}005\,a + 0{,}5\,b + 0{,}0025$$

oder

$$|xy - ab| \le 0{,}005\,(a + 100\,b) + 0{,}0025.$$

Wir treffen nun eine Fallunterscheidung, je nachdem ob (1) das Produkt ab drei oder (2) das Produkt ab vier Ziffern vor dem Komma besitzt. Die Ungleichungen $100 \cdot 1{,}00 \le ab \le 999 \cdot 9{,}99$ zeigen, daß genau diese beiden Fälle möglich sind.

Im ersten Fall ist $ab \le 999{,}99$; $100\,b \le \frac{p^2}{a}$, wobei $p^2 = 99999$ ist, und mithin

$$a + 100\,b \le a + \frac{p^2}{a}.$$

Durch Differentiation der Funktion $a + \frac{p^2}{a}$ nach a finden wir, daß sie bei stetigem Wachsen der Variablen a von 100 bis 999 zunächst monoton fällt (und zwar vom Werte 1099,99 bei $a = 100$ bis zum Wert $2p = 632{,}4\ldots$ bei $a = p = 316{,}2\ldots$) und danach monoton wächst (vom Werte $2p$ bei $a = p$ bis zum Werte $1099 + \frac{99}{999}$ bei $a = 999$). Wenn wir beachten, daß die Summe $a + 100\,b$ nur ganze Werte annimmt, so kommen wir zu dem Schluß, daß der größtmögliche Wert dieser Summe gleich 1099 und daher

$$|xy - ab| \leq 0{,}005 \cdot 1099 + 0{,}0025 = 5{,}4975 < 5{,}5$$

ist, womit die erste Behauptung von Satz 2 für den Fall bewiesen ist, daß das Produkt drei wesentliche Ziffern vor dem Komma besitzt.

Im zweiten Fall, wenn also das Produkt nicht drei, sondern vier Stellen vor dem Komma besitzt, läßt sich der größtmögliche Fehler des Produktes weit einfacher berechnen. Dann ist nämlich

$$|xy - ab| \leq 0{,}005 \cdot 999 + 0{,}5 \cdot 9{,}99 + 0{,}0025 = 9{,}9925 < 10,$$

wobei im Fall von vier wesentlichen Ziffern vor dem Komma die Zehnerziffer die dritte wesentliche Ziffer ist, so daß wir gezeigt haben, daß die Zahl $|xy - ab|$ kleiner ist als eine Einheit der Ordnung der dritten wesentlichen Ziffer des Produktes. Damit ist die erste Behauptung von Satz 2 auch für den zweiten Fall bewiesen.

Wenn einer der Näherungsfaktoren $k = 3$ und der andere $k + 1 = 4$ zuverlässige Ziffern besitzt, so erhalten wir durch entsprechende Überlegungen wie oben:

$$100 \leq a \leq 999;\ 1{,}000 \leq b \leq 9{,}999;\ |\alpha| \leq 0{,}5;\ |\beta| \leq 0{,}0005;$$
$$|xy - ab| \leq 0{,}0005\,a + 0{,}5\,b + 0{,}00025 = 0{,}0005(a + 1000\,b) + 0{,}00025.$$

Wenn dann das Produkt ab drei Ziffern vor dem Komma besitzt, so ist

$$ab \leq 999{,}999;\ 1000\,b \leq \frac{p^2}{a},$$

wobei $p^2 = 999999$ ist, und mithin

$$a + 1000\,b \leq a + \frac{p^2}{a}.$$

Diese letzte Summe ist, falls a von 100 bis 999 wächst, monoton fallend, da ihr Minimum bei $a = p = 999{,}99\ldots$ liegt. Folglich ist ihr größtmöglicher Wert gleich $100 + 9999{,}99 = 10099{,}99$, und da die Summe $a + 1000\,b$ nur ganze Werte annimmt, kann sie die Zahl 10099 nicht übertreffen. Daraus ergibt sich, daß

$$|xy - ab| \leq 0{,}0005 \cdot 10099 + 0{,}00025 = 5{,}04975 < 5{,}05$$

ist. Wenn dagegen das Produkt ab vier Ziffern vor dem Komma besitzt, so ist

$$|xy - ab| \leq 0{,}4995 + 4{,}9995 + 0{,}00025 = 5{,}49925 < 5{,}5$$

oder 0,55 Einheiten der Ordnung der dritten wesentlichen Ziffer des Produktes. Damit ist auch die zweite Behauptung unseres Satzes bewiesen.

Zum Beweis der dritten Behauptung schließen wir folgendermaßen:

$$x = a + \alpha;\ 100 \leq a \leq 999;\ |\alpha| \leq 0{,}5;\ y = b;\ 1 \leq b < 10;$$
$$|xy - ab| \leq 0{,}5\,b \leq 5.$$

Wenn das Produkt drei Stellen vor dem Komma besitzt, so ist sein Fehler nicht größer als 5 Einheiten der Ordnung seiner dritten wesentlichen Ziffer; wenn es dagegen vier Stellen vor dem Komma besitzt, also seine dritte wesentliche Ziffer die Zehnerziffer ist, so ist sein Fehler höchstens gleich 0,5 Einheiten der Ordnung der dritten wesentlichen Ziffer.

Damit ist Satz 1 vollständig bewiesen.

Aus dem oben formulierten Satz können wir bereits eine Reihe von praktischen Schlußfolgerungen ziehen. Wenn eine vorgegebene Näherungszahl k zuverlässige wesentliche Ziffern besitzt und eine andere vorgegebene Zahl nicht weniger genau

ist, d. h. entweder eine Näherungszahl mit mindestens k zuverlässigen wesentlichen Ziffern oder genau ist, so hat es keinen Sinn, im Produkt mehr als k wesentliche Ziffern beizubehalten, ja schon die k-te wesentliche Ziffer ist zweifelhaft. In diesem Zusammenhang taucht die Frage auf, ob es sich überhaupt lohnt, diese k-te wesentliche Ziffer beizubehalten (diese Frage wird in den §§ 11 und 12 positiv beantwortet). Ferner kann man im Fall zweier Faktoren mit verschiedener Anzahl von wesentlichen Ziffern, ohne Nachteil für die Genauigkeit des Resultates, vor der Rechnung den genaueren Faktor derart runden, daß er nur eine wesentliche Ziffer mehr als der weniger genaue Faktor besitzt: Wenn der weniger genaue Faktor k zuverlässige wesentliche Ziffern besitzt, so vergrößert sich der Höchstfehler des Resultates dabei nur wenig, nämlich von 5 auf 5,05 Einheiten der Ordnung der k-ten wesentlichen Ziffer des Produktes. Allerdings lohnt es sich in jedem Falle, die $(k+1)$-te wesentliche Ziffer des genaueren Faktors beizubehalten, da sich sonst der Höchstfehler des Produktes unnütz vergrößert, nämlich von 5,05 auf 5,5 Einheiten der Ordnung der k-ten wesentlichen Ziffer.

Wenn wir in analoger Weise den Quotienten untersuchen, erhalten wir:

Satz 2. *Der Quotient zweier Näherungszahlen mit je k zuverlässigen wesentlichen Ziffern besitzt einen Höchstfehler von 10 Einheiten seiner k-ten wesentlichen Ziffer. Dieser Höchstfehler verringert sich auf 5,5 Einheiten, wenn eine der gegebenen Näherungszahlen k und die andere $(k+1)$ zuverlässige wesentliche Ziffern besitzt. Er verringert sich auf 5,22 Einheiten, wenn der Dividend genau ist und der Divisor k zuverlässige wesentliche Ziffern besitzt, und auf 5 Einheiten, wenn der Divisor genau ist und der Dividend k zuverlässige wesentliche Ziffern besitzt.*

Auf den Beweis dieses Satzes, der analog dem Beweis des vorangehenden Satzes verläuft, wollen wir hier verzichten. Die angegebenen Abschätzungen des Höchstfehlers von Produkt und Quotient von Näherungszahlen führen uns auf folgende praktische Regel:

Regel 2 der Ziffernzählung: *Bei der Multiplikation und Division von Näherungszahlen sind im Resultat so viele wesentliche Ziffern beizubehalten, wie in der vorgegebenen Näherungszahl mit der kleinsten Anzahl von wesentlichen Ziffern vorhanden sind.*

Der Zweifel an der Zweckmäßigkeit der Festsetzung, daß auch die k-te wesentliche Ziffer beim Produkt noch beizubehalten ist, wobei der Fehler ja bis 5,5 Einheiten der betreffenden Ordnung betragen kann, und noch mehr beim Quotienten, bei den er 10 Einheiten erreichen kann, wird durch Überlegungen beseitigt, auf die wir in den §§ 11 und 12 eingehen werden.

Die Brauchbarkeit der Regeln 1 und 2 macht man sich am besten an Hand von konkreten Beispielen klar, wobei man zweckmäßigerweise die unbekannten Ziffern durch Fragezeichen ersetzt. Wir führen vier derartige Beispiele an, die ohne jede Erklärung verständlich sind:

9,56?		3,143?		936	?:218? = 4,29		65	4,7?:2,6? = 250	
×2,18?		0,85?		872	?		52	?	(besser $2{,}5\cdot 10^2$)
?????		??????		64	??		13	47	
7	648?	1	5715?	43	6?		13	0?	
9	56?	25	144?	20	4??			47	
191	2?	26	7155??	19	62?				
208	408??	2,7			78?				
20,8									

Der Übersichtlichkeit halber haben wir die zuverlässigen Ziffern durch vertikale punktierte Linien abgetrennt. Näheres über Beispiele dieser Art kann man in dem Buch [2a] nachlesen.

Wenn wir zu den Grundrechenarten dritter Stufe (Potenzieren und Radizieren) übergehen und uns auf den Fall beschränken, daß als Exponent der Potenz und der Wurzel die Zahlen 2 und 3 auftreten, so beweisen wir mühelos die beiden folgenden Sätze:

Satz 3. *Das Quadrat bzw. der Kubus einer Näherungszahl mit k zuverlässigen wesentlichen Ziffern besitzt einen Höchstfehler von* 3,5 *bzw.* 7,2 *Einheiten der Ordnung seiner k-ten wesentlichen Ziffer.*

Satz 4. *Die Quadrat- bzw. die Kubikwurzel aus einer Näherungszahl mit k zuverlässigen wesentlichen Ziffern besitzt einen Höchstfehler von* 0,81 *bzw.* 0,79 *Einheiten der Ordnung ihrer k-ten wesentlichen Ziffer.*

Diese Sätze bilden die Grundlage für die folgenden Regeln der Ziffernzählung:

Regel 3 der Ziffernzählung. *Beim Potenzieren mit* 2 *und* 3 *sind im Resultat so viele Ziffern als wesentlich beizubehalten, wie die Basis zuverlässige wesentliche Ziffern besitzt.*

Regel 4 der Ziffernzählung. *Bei der Berechnung der Quadrat- und Kubikwurzel aus einer Näherungszahl sind im Resultat so viele Ziffern beizubehalten, wie im Radikanden zuverlässige wesentliche Ziffern vorhanden sind.*

Auf Grund der oben angegebenen Werte des Höchstfehlers und der Überlegungen aus den §§ 11 und 12 können wir zu diesen Regeln noch folgendes bemerken: Die letzte Ziffer eines in der angegebenen Weise gerundeten Quadrates und besonders eines so gerundeten Kubus ist weniger zuverlässig als die letzte Ziffer der Basis; hingegen ist die letzte Ziffer der in angegebener Weise gerundeten Quadrat- und besonders der so abgerundeten Kubikwurzel zuverlässiger als die letzte Ziffer des Radikanden.

Auf die praktische Anwendung der Regeln der Ziffernzählung kommen wir erst in § 13 zu sprechen. Wir wollen jedoch bereits hier darauf hinweisen, daß sich beim Runden des Resultates nach den Regeln der Ziffernzählung die oben angegebenen Höchstfehler jeweils um maximal 0,5 Einheiten vergrößern. In der Tabelle auf S. 366 sind sowohl die oben angegebenen Höchstfehler als auch die so vergrößerten Höchstfehler aufgenommen.

Interessant ist noch die Analogie, die zwischen den Regeln der Ziffernzählung und den Sätzen aus § 9 über die Schranken des absoluten und relativen Fehlers besteht: Bei der Addition und Subtraktion von Näherungszahlen addierten sich die durch die Anzahl der zuverlässigen Dezimalstellen festgelegten Schranken ihres absoluten Fehlers. Bei der Multiplikation und Division dagegen waren die durch die Anzahl der zuverlässigen, geltenden Ziffern bestimmten Schranken des relativen Fehlers der gegebenen Näherungszahlen zu addieren. Die Multiplikation der Schranke des relativen Fehlers mit dem Exponenten beim Potenzieren und die Division durch den Wurzelexponenten im Fall des Radizierens machen die oben bemerkte Verminderung der Genauigkeit im ersten Fall und ihre Vergrößerung im zweiten Fall begreiflich.

§ 11. Der mittlere quadratische Fehler der Resultate der Grundrechenarten an Näherungszahlen. Das Prinzip von A. N. Krylow

In den Fällen, in denen wir die Möglichkeit haben, außer einer Schranke des Fehlers, d. h. seines größtmöglichen Wertes, auch den wahren Wert des Resultates festzustellen, bemerken wir im allgemeinen, daß der wahre Fehler bedeutend kleiner ist als der theoretische Höchstfehler. Diese Erscheinung wird um so markanter, je mehr Näherungszahlen in die Berechnung eingehen. Wenn wir z. B. die Summe der vierstelligen Logarithmen der zwanzig Zahlen von 11 bis 30 einschließlich bestimmen, so ist 0,00005 eine Schranke des absoluten Fehlers jedes Logarithmus, also $0{,}00005 \times 20 = 0{,}001$ eine Schranke des absoluten Fehlers der Summe. Da die Summe der genannten Logarithmen 25,8638 ist, können wir also dafür garantieren, daß der wahre Wert dieser Summe größer als 25,8628 und kleiner als 25,8648 ist. Wiederholen wir dieselbe Rechnung mit achtstelligen Logarithmen, so erhalten wir die Summe 25,86389705, d. h., der wahre Fehler der ersten Summe macht nicht einmal ein Zehntausendstel aus, stellt also ungefähr ein Zehntel des oben berechneten theoretischen Wertes dar.

Diese Differenz zwischen dem wahren Fehler und dem Höchstfehler erklärt sich daraus, daß wir beim Bestimmen des Höchstfehlers die ungünstigsten Umstände vorausgesetzt haben. So haben wir im eben betrachteten Beispiel als Schranke des Fehlers jedes Summanden die halbe Einheit der Ordnung seiner letzten Ziffer angesetzt. Indes können die wahren Fehler dieser Summanden alle möglichen Werte zwischen $-0{,}5$ und $+0{,}5$ Einheiten dieser Ordnung annehmen, und sie nehmen auch in der Tat verschiedene dieser Werte an. Dabei kommen positive Abweichungen ungefähr in derselben Häufigkeit vor wie negative Abweichungen und heben sich damit in mehr oder weniger bedeutendem Maße auf. Der Anhäufung von Fehlern läuft also eine gewisse gegenseitige Kompensation parallel, so daß die Wahrscheinlichkeit für das Auftreten eines großen, nahe dem Höchstfehler gelegenen Fehlers im Resultat äußerst gering ist. Zwar kann man bei künstlicher Wahl der Summanden Fehler konstruieren, die dem Höchstfehler beliebig nahe kommen, jedoch ist das Auftreten von solchen künstlichen Auswahlen in der Praxis sehr unwahrscheinlich.

Mit den Methoden der Wahrscheinlichkeitsrechnung kann man übrigens sehr genau berechnen, wie oft ein gegebener Wert des Fehlers in der Summe auftritt. Die Resultate dieser theoretischen Untersuchungen wurden durch direkte Experimente bestätigt. So hat man z. B. folgenden Versuch durchgeführt: Man hat 440 Summen aus je zwanzig Logarithmen berechnet, und zwar einmal mit fünf Dezimalstellen und zum anderen mit sieben Dezimalstellen. Sodann hat man die Differenzen der jeweiligen Summen von fünf- und siebenstelligen Logarithmen berechnet, d. h. einen Näherungswert für die Fehler der Summen aus fünfstelligen Logarithmen bestimmt:

Der Fehler der Summe liegt zwischen	Nach der Theorie	Anzahl der Fälle in der Wirklichkeit
0 und 100	56,14%	65%
100 und 200	31,72%	28%
200 und 300	10,13%	6%
300 und 400	1,82%	1%
400 und 500	0,18%	0%
500 und 1000	0,01%	0%

Die vorangehende Tabelle zeigt ein genügend gutes Übereinstimmen der Resultate des Versuches mit den theoretischen Werten. Die Fehler sind dabei in Zehnmillionstel (d. h. Einheiten der Ordnung der letzten Ziffer der siebenstelligen Logarithmen) angegeben. Die theoretische Schranke des absoluten Fehlers jeder Summe aus zwanzig Summanden mit je fünf zuverlässigen Dezimalstellen ist gleich $0{,}5 \cdot 10^{-5} \cdot 20$, d. h. 1000 Zehnmillionstel. Jedoch wird diese Schranke in keinem der 440 Fälle erreicht. Bei praktischen Berechnungen muß man mit der angegebenen Wahrscheinlichkeit, mit einem großen, d. h. dem Höchstfehler relativ nahen Fehler rechnen.

Da die genaue Fehlerrechnung, wie wir in den §§ 8 und 9 gesehen haben, manche zusätzliche Arbeit erfordert, wendet man sie in der Praxis nur gelegentlich an. Gewöhnlich begnügt man sich damit, die Berechnungen auf eine bestimmte Anzahl von Ziffern (wesentliche Ziffern oder Dezimalstellen) vorzunehmen, wobei man im Resultat eine oder auch zwei unzuverlässige Ziffern beibehält (vgl. den Schluß des Artikels [6]).

In der Praxis läßt man üblicherweise zu, daß die auftretenden Zahlen Fehler besitzen, die die Einheit der Ordnung ihrer letzten Ziffer nicht überschreiten. So finden wir in dem Buch [10] von A. N. Krylow:

„Das Resultat jeder Berechnung und Messung wird durch eine gewisse Zahl ausgedrückt. Dabei werden verabredungsgemäß alle Zahlen so niedergeschrieben, daß man aus der Form der Niederschrift bereits den Grad der Genauigkeit ablesen kann, und zwar benutzt man in der Regel eine Schreibweise, bei der alle wesentlichen Ziffern mit Ausnahme der letzten zuverlässig sind und die letzte Ziffer nur um höchstens eine Einheit der betreffenden Ordnung angezweifelt werden kann."

Wenn man diese letzte Forderung buchstäblich nimmt, so ist sie nur sehr schwer zu erfüllen. Um sie zu befolgen, müßte man entweder eine fortlaufende genaue Fehlerrechnung durchführen oder bei jedem Schritt die Resultate stark runden. So besitzen z. B. die vierstelligen Logarithmen, die man als Summen von drei vierstelligen Logarithmen erhält, einen Höchstfehler von $1\frac{1}{2}$ Einheiten der Ordnung der letzten Stelle, und man müßte sie beim starren Festhalten an der genannten Regel auf drei Dezimalstellen runden. Fügen wir jedoch in die obige Regel den Zusatz „im Mittel" ein, so erhalten wir ein Prinzip von entscheidender Bedeutung, mit dessen Hilfe man eine ganze Reihe von praktischen Regeln für das Rechnen mit Näherungszahlen vernünftig begründen kann. Dieses „Grundprinzip für gewöhnliche Rechnungen", d. h. für Rechnungen ohne genaue Fehlerrechnung, formulieren wir in seiner endgültigen Gestalt folgendermaßen:

Prinzip von A. N. Krylow. *Jede Näherungszahl ist so zu schreiben, daß alle ihre wesentlichen Ziffern mit evtl. Ausnahme der letzten zuverlässig sind und die letzte Ziffer im Mittel um höchstens eine Einheit der betreffenden Ordnung zweifelhaft ist.*

Diesen Zusatz „im Mittel" werden wir dabei so verstehen, daß hier nicht von einer Schranke des Fehlers, sondern vom mittleren quadratischen Fehler, d. h. von der Quadratwurzel aus dem Mittelwert der Quadrate der möglichen Fehler die Rede ist.

Um eine klare Vorstellung von dem mittleren quadratischen Fehler zu erhalten, wollen wir als Beispiel den mittleren quadratischen Fehler beim Runden durch Vernachlässigung einer Ziffer berechnen. Dabei sehen wir alle möglichen Werte der fortgelassenen Ziffer als gleich wahrscheinlich an, d. h., wir setzen voraus, daß (bei einer

großen Zahl von derartigen Vernachlässigungen) alle Werte gleich oft auftreten. Es sind also folgende Werte für den Fehler beim Runden (in Einheiten der Ordnung der zuletzt berücksichtigten Ziffer) gleich wahrscheinlich: $-0{,}5$; $-0{,}4$; $-0{,}3$; $-0{,}2$; $-0{,}1$; 0; 0,1; 0,2; 0,3; 0,4; 0,5. Insgesamt sind dies 11 Fehlerwerte. Wir bilden dann die Summe der Quadrate dieser Fehlerwerte, teilen sie durch 11 und ziehen aus dem Quotienten die Quadratwurzel. Dies ergibt dann gerade den gesuchten mittleren quadratischen Fehler beim Runden, also

$$\sqrt{\frac{2}{11}(0{,}25 + 0{,}16 + 0{,}09 + 0{,}04 + 0{,}01)} = \sqrt{\frac{2 \cdot 0{,}55}{11}} = \sqrt{0{,}1} \approx 0{,}316.$$

Besteht das Runden nicht im Fortlassen einer, sondern zweier Ziffern, so treten an die Stelle der 11 Fehlerwerte die 101 Fehlerwerte von $-0{,}50$ bis $+0{,}50$ (Einheiten der Ordnung der zuletzt berücksichtigten Ziffer) und der mittlere quadratische Fehler ergibt sich zu 0,292. Bei dieser Berechnung kann man die an sich notwendige Addition der Quadrate durch Verwendung der Formel

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

vermeiden.

Wenn schließlich das Abrunden im Fortlassen einer unendlichen Anzahl von Ziffern besteht, so erhält man als mittleren quadratischen Fehler durch Übergang zum Limes (oder unter Benutzung der Integralrechnung) den Wert $\sqrt{3} : 6 = 0{,}289$.

Für die Anwendung des Prinzips von A. N. Krylow in seiner oben angegebenen Formulierung auf die Grundrechenarten an Näherungszahlen müssen wir die mittleren quadratischen Fehler der Resultate der einzelnen Grundrechenarten bestimmen. Wir wollen hier nur den mittleren quadratischen Fehler einer Summe aus n angenäherten Summanden bestimmen, deren jeder das Ergebnis des Rundens einer gegebenen genauen Zahl auf k Dezimalstellen ist.

Dazu sei uns zunächst ein Näherungswert a vorgegeben, der k zuverlässige Dezimalstellen besitzt und den wir als Ergebnis des Rundens einer Zahl x_i mit $m > k$ Dezimalstellen ansehen wollen. Dann sind insgesamt $10^{m-k} + 1 = p$ derartige Zahlen x_i möglich, von denen jede gleichwahrscheinlich ist. Ist z. B. $a = 74{,}62$; $k = 2$; $m = 5$, so ist $p = 1001$; $x_1 = 74{,}61500$; $x_2 = 74{,}61501$; $x_3 = 74{,}61502$; ...; $x_{1000} = 74{,}62499$; $x_{1001} = 74{,}62500$. Der wahre absolute Fehler $x_i - a$ des Wertes a kann also gleichfalls p Werte annehmen, nämlich die p gleichwahrscheinlichen Werte von $x_1 - a = -0{,}5 \cdot 10^{-k}$ bis $x_p - a = +0{,}5 \cdot 10^{-k}$, wobei

$$\sum_{i=1}^{p}(x_i - a) = 0, \qquad \frac{\sum_{i=1}^{p}(x_i - a)^2}{p} = \sigma_a^2$$

ist.

Hier bezeichnet σ_a den mittleren quadratischen Fehler der Zahl a, der im Falle $k = 2$, $m = 5$ augenscheinlich

$$\sigma_a = \sqrt{\frac{2 \cdot \frac{1}{6} \cdot 500 \cdot 501 \cdot 1001}{1001}} \cdot 10^{-5} \approx 0{,}289 \cdot 10^{-3}$$

ist.

Für $m \to \infty$ erhält man entsprechend:

$$\sigma_a \to \frac{1}{6}\sqrt{3} \cdot 10^{-k}.$$

Daneben betrachten wir jetzt eine zweite Näherungszahl b, die — vollkommen unabhängig von der ersten — ebenfalls k zuverlässige Dezimalstellen besitze und das Ergebnis des Rundens der genauen Zahl y_j mit ebenfalls $m > k$ Dezimalstellen sei. Dann sind auch für y_j im ganzen $10^{m-k} + 1 = p$ Werte möglich, die wiederum gleiche Wahrscheinlichkeit besitzen. Wir erhalten also ebenso wie im ersten Fall

$$\sum_{j=1}^{p} (y_j - b) = 0, \qquad \frac{\sum_{j=1}^{p} (y_j - b)^2}{p} = \sigma_b^2$$

und damit $\sigma_b = \sigma_a$.

Schließlich stellt die Summe $c = a + b$ einen gewissen Näherungswert einer genauen Summe $z_{ij} = x_i + y_j$ dar. Dabei kann z_{ij} ein beliebiger von p^2 möglichen und gleichwahrscheinlichen Werten sein (ein beliebiger der p möglichen Werte x_i wird mit einem beliebigen der p möglichen Werte von y_j kombiniert).

Wir halten nun in der Gleichung

$$(z_{ij} - c)^2 = [(x_i - a) + (y_j - b)]^2 =$$
$$= (x_i - a)^2 + (y_j - b)^2 + 2(x_i - a)(y_j - b)$$

den Index j fest und summieren die p Gleichungen über i von 1 bis p. Dann erhalten wir

$$\sum_{i=1}^{p} (z_{ij} - c)^2 = \sum_{i=1}^{p} (x_i - a)^2 + p(y_j - b)^2 + 2(y_j - b) \sum_{i=1}^{p} (x_i - a);$$

hier ist der dritte Summand auf der rechten Seite gleich Null. Danach summieren wir alle Gleichungen nach dem Index j über die p Werte von 1 bis p und erhalten

$$\sum_{j=1}^{p} \sum_{i=1}^{p} (z_{ij} - c)^2 = p \sum_{i=1}^{p} (x_i - a)^2 + p \sum_{j=1}^{p} (y_j - b)^2.$$

Nach gliedweiser Division durch p^2 ergibt sich hieraus

$$\sigma_c^2 = \sigma_a^2 + \sigma_b^2,$$

wonach unter den angegebenen Voraussetzungen *das Quadrat des mittleren quadratischen Fehlers einer Summe gleich der Summe der Quadrate der mittleren quadratischen Fehler der Summanden ist.*

Dieses Ergebnis (das — wie man leicht sieht — auch unter allgemeineren Voraussetzungen erhalten bleibt) läßt sich unmittelbar auch auf Summen aus einer beliebigen Anzahl von Summanden verallgemeinern. Wenn uns eine Summe aus n Summanden vorgegeben ist, die die obigen Bedingungen erfüllen, so erhalten wir

$$\sigma_s^2 = \sigma_a^2 + \sigma_b^2 + \cdots + \sigma_n^2, \quad \sigma_a = \sigma_b = \cdots = \sigma_n, \quad \sigma_s^2 = n\sigma_a^2,$$
$$\sigma_s = \sigma_a \sqrt{n}.$$

d. h., *der mittlere quadratische Fehler einer Summe aus gleich-genauen Summanden ist proportional der Quadratwurzel aus der Anzahl der Summanden.* Hieraus folgt, daß der mittlere quadratische Fehler einer Summe aus n Näherungszahlen, deren jede das Ergebnis des Rundens einer gegebenen genauen Zahl auf k Dezimal-

stellen ist, gleich $\frac{1}{6} \cdot \sqrt{3} \cdot \sqrt{n} \cdot 10^{-k}$ oder angenähert $0{,}289\ \sqrt{n} \cdot 10^{-k}$ ist. Auf Grund des Prinzips von A. N. KRYLOW muß also im betrachteten Fall die Ungleichung $\frac{1}{6} \sqrt{3} \cdot \sqrt{n} \leq 1$ gelten, die unmittelbar auf $n \leq 12$ führt. Wenn nicht mehr als 12 angenäherte Summanden vorgegeben sind, von denen jeder durch Runden auf ein und dieselbe Dezimalordnung erhalten wurde, so kann man in der Summe alle Dezimalstellen beibehalten. In der Praxis wird die Anzahl 12 natürlich häufig überschritten. Das oben beschriebene Experiment mit der Addition von $n = 20$ Logarithmen zeigt, daß auch dort noch alle Stellen der Summe erhalten bleiben.

Interessant ist eine Gegenüberstellung der Werte des Höchstfehlers ε und der Werte des mittleren quadratischen Fehlers σ einer Summe aus n Summanden. Wir geben nachfolgend eine Tabelle für n von 2 bis 12 an, in der ε und σ in Einheiten der Ordnung der k-ten Dezimalstelle ausgedrückt sind:

n	2	3	4	5	6	7	8	9	10	11	12
ε	1	1,5	2	2,5	3	3,5	4	4,5	5	5,5	6
σ	0,409	0,501	0,578	0,647	0,708	0,765	0,818	0,867	0,915	0,960	1,000

Die angegebene Berechnung des Wertes von σ für eine Summe aus n Summanden vereinfacht sich wesentlich, wenn man einfache Sätze aus der Wahrscheinlichkeitsrechnung und die Integralrechnung benutzt. Dann bereitet auch die Berechnung des mittleren quadratischen Fehlers für die Resultate der anderen Grundrechenarten keinerlei Schwierigkeiten. Wir wollen uns hier auf die Angabe einer Tabelle beschränken, in der die Werte des Höchstfehlers und des mittleren quadratischen Fehlers gegenübergestellt sind:

Grundrechenart	ε	σ	ε_1
Algebraische Summe aus n Summanden	$0{,}5\,n$	$0{,}289\ \sqrt{n}$	$0{,}5\,n$
Produkt zweier k-stelliger Näherungszahlen	5,5	0,626	6
Produkt einer k-stelligen Näherungszahl mit einer genauen Zahl	5	0,442	5,5
Produkt einer k-stelligen Näherungszahl mit einer $(k+1)$-stelligen Näherungszahl	5,05	0,445	5,55
Quotient der Division einer k-stelligen Näherungszahl durch eine k-stellige Näherungszahl	10	0,576	10,5
Quotient der Division einer k-stelligen Näherungszahl durch eine genaue Zahl	5	0,151	5,5
Quotient der Division einer k-stelligen Näherungszahl durch eine $(k+1)$-stellige Näherungszahl	5,5	0,391	6
Quotient der Division einer genauen Zahl durch eine k-stellige Näherungszahl	5,22	0,425	5,72
Quotient der Division einer $(k+1)$-stelligen Näherungszahl durch eine k-stellige Näherungszahl	5,5	0,427	6
Quadrat einer k-stelligen Näherungszahl	3,5	0,705	4
Kubus einer k-stelligen Näherungszahl	7,13	1,059	7,63
Quadratwurzel aus einer k-stelligen Näherungszahl	0,81	0,221	1,31
Kubikwurzel aus einer k-stelligen Näherungszahl	0,79	0,185	1,29

Alle Zahlen, die in den Spalten ε, σ, ε_1 auftreten, sind in Einheiten der Ordnung der k-ten wesentlichen Ziffer ausgedrückt. Eine Ausnahme bilden nur die Zahlen in

der ersten Zeile, die in Einheiten der Ordnung der k-ten Dezimalstelle ausgedrückt sind. Die Gegenüberstellung von Höchstfehler ε und mittlerem quadratischem Fehler σ bestätigt die Zweckmäßigkeit der in § 10 aufgestellten Regeln der Ziffernzählung, da sie zeigt, daß alle dort beibehaltenen Ziffern dem Prinzip von A. N. KRYLOW genügen. Die Zahlen in der Spalte ε_1 gehen aus den Zahlen der Spalte ε durch Addition von 0,5 hervor und geben an, welchen Höchstfehler man begeht, wenn man das Resultat auf k wesentliche Ziffern rundet.

§ 12. Verteilung der Fehler in den Rechenresultaten

Die in der Tabelle in § 11 gegebene Gegenüberstellung von mittlerem quadratischem Fehler und Höchstfehler zeigt, daß *große Fehler* (d. h. Fehler, die dem Höchstfehler nahe kommen) *eine sehr kleine Wahrscheinlichkeit besitzen*: Wenn z. B. das Produkt zweier k-stelliger Näherungszahlen einen Höchstfehler von 5,5 und der mittlere quadratische Fehler dagegen nur 0,626 Einheiten der Ordnung der k-ten wesentlichen Ziffer beträgt, so muß in der überwiegenden Mehrzahl der Fälle der wirkliche Fehler sehr gering sein und nur in seltenen Fällen sich dem Höchstfehler nähern. Hierbei möchte man sich naturgemäß noch ein genaueres Bild von der Verteilung der Fehler machen, d. h., man möchte wissen, wie häufig Fehler vorkommen, die in einem vorgegebenen Intervall liegen.

Die Lösung dieser Aufgabe erfordert erhebliche Mittel aus der Wahrscheinlichkeitsrechnung, und wir müssen uns darauf beschränken, auf das Buch [10] von A. N. KRYLOW hinzuweisen, in dem dieses Problem für die Summe gelöst ist (S. 195 bis 204). Durch Anwendung der dort angegebenen Formeln auf eine Summe aus zwanzig Summanden, die sämtlich aus genauen Zahlen durch Runden auf die gleiche Dezimalstelle erhalten werden, gewinnt man gerade die oben (S. 362) angegebene Fehlerverteilung: Einen Fehler von höchstens einer Einheit der Ordnung der letzten Dezimalstelle begeht man in 56,14% aller Fälle; einen Fehler von einer bis zwei Einheiten begeht man in 31,72% aller Fälle, einen Fehler von zwei bis drei Einheiten in 10,13% aller Fälle, einen Fehler von drei bis vier Einheiten in 1,82% aller Fälle, einen Fehler von vier bis fünf Einheiten in 0,18% aller Fälle; einen Fehler von fünf bis zehn Einheiten begeht man schließlich nur in 0,01% aller Fälle, d. h. im Mittel nur einmal unter 10000 derartigen Additionen; ein Fehler, der zehn Einheiten übertrifft, ist schon nicht mehr möglich. Entsprechend findet man bei der Multiplikation zweier k-stelliger Zahlen, die durch Runden genauer Zahlen auf k wesentliche Ziffern gewonnen werden, im Resultat einen Fehler von höchstens einer Einheit der Ordnung der k-ten wesentlichen Ziffer des Produktes in 91,51% aller Fälle, einen Fehler von ein bis zwei Einheiten in 5,87% aller Fälle, einen Fehler von zwei bis drei Einheiten in 2,09% aller Fälle, einen Fehler von drei bis vier Einheiten in 0,47% aller Fälle und einen Fehler von vier bis fünfeinhalb Einheiten in 0,06% aller Fälle; ein Fehler von 5,5 Einheiten ist bereits Höchstfehler.

Die Richtigkeit dieser letzten Berechnung läßt sich gleichfalls leicht am Experiment bestätigen. Die Übereinstimmung zwischen Experiment und Theorie ist um so

größer, je mehr Produkte in Betracht gezogen werden. Bei kollektiver Arbeit von 20 bis 30 Personen beansprucht man für die Durchführung eines solchen Experimentes etwa 15 bis 20 Minuten. Jeder Teilnehmer multipliziert zunächst eine gewisse Anzahl von Paaren beliebiger, etwa dreistelliger Zahlen. Dann rundet er die Ausgangszahlen auf z. B. zwei wesentliche Ziffern und multipliziert die so gerundeten Zahlen von neuem. Wenn man schließlich das Komma in sämtlichen Faktoren so setzt, daß die Produkte zwei Ziffern links vom Komma besitzen, so findet man als Fehler die Differenzen aus den so abgeänderten genauen Produkten und den so abgeänderten Produkten der zweistelligen Näherungszahlen. Wenn wir z. B. von den Zahlen 492 und 927 ausgehen, so erhalten wir als Produkt 456084; als zweistellige Näherungszahlen finden wir 49 und 93 mit dem Produkt 4557; dann setzen wir in jedem Faktor das Komma hinter die erste wesentliche Ziffer und erhalten $4{,}92 \cdot 9{,}27 = 45{,}6084$ und $4{,}9 \cdot 9{,}3 = 45{,}57$; der Fehler zwischen dem genauen Produkt und dem Produkt der Näherungszahlen beträgt dann 0,0384 Einheiten der Ordnung der zweiten wesentlichen Ziffer des Produktes. In Fortführung des Experimentes hat man dann nur noch die Anzahl aller Fälle zusammenzuzählen, die Anzahlen der Fälle zu bestimmen, in denen die Differenzen zwischen 0 und 1, 1 und 2, 2 und 3, 3 und 4, 4 und 5,5 Einheiten liegen, und danach die entsprechenden Prozente zu bestimmen.

Hier das Ergebnis eines derartigen Experimentes, das an 200 Paaren von fünfstelligen Zahlen durchgeführt wurde, die im Laufe des Experiments auf drei wesentliche Ziffern abgerundet wurden. Die Fehler der Produkte lagen zwischen 0 und 1 (Einheiten der Ordnung der dritten wesentlichen Ziffer des Produktes) in 186 Fällen, d. h. 93% aller Fälle (nach der Theorie müßten es 91,51% sein); der Fehler lag zwischen 1 und 2 Einheiten in 10 Fällen, d. h. 5% aller Fälle (nach der Theorie 5,87%); der Fehler lag zwischen 2 und 3 Einheiten in 3 Fällen, d. h. 1,5% aller Fälle (nach der Theorie 2,09%); der Fehler lag zwischen 3 und 4 Einheiten in nur einem Fall, d. h. 0,5% aller Fälle (nach der Theorie 0,47%); ein Fehler von über 4 Einheiten trat im Experiment nicht auf (die Theorie liefert für das Intervall von 4 bis 5,5 den Wert 0,06%).

Wir finden also eine befriedigende Übereinstimmung zwischen den theoretischen Untersuchungen über die Fehlerverteilung bei der Summe und beim Produkt mit den experimentellen Daten. Ähnlich liegen die Dinge auch beim Quotienten, beim Quadrat, beim Kubus, bei der Quadratwurzel und bei der Kubikwurzel. Auf diese Weise erhalten die Regeln der Ziffernzählung eine neue Bestätigung. Wenn man sie befolgt, kann man zwar nicht mehr für die Zuverlässigkeit der letzten Ziffer des Resultats garantieren, jedoch ist in der Mehrzahl der Fälle der Fehler in der letzten Ziffer nur unbedeutend, und es wäre unklug, sie völlig wegzulassen, wie es auch unklug wäre, mehr Ziffern beizubehalten, als die Regeln empfehlen.

Es versteht sich natürlich von selbst, daß in Fällen, in denen eine vollständig sichere Berechnung des Resultates erforderlich ist, die Regeln über die Ziffernzählung nicht angewendet werden dürfen. Hier ist eine genaue Fehlerrechnung nach der Methode der Wertschranken oder der Methode der Fehlerschranken unerläßlich. Bei gewöhnlichen Berechnungen hingegen, bei denen üblicherweise keine genaue Fehlerrechnung durchgeführt wird, machen die Regeln der Ziffernzählung zuverlässige Angaben über das zweckmäßigste Runden aller erhaltenen Resultate.

§ 13. Praktische Anwendungen der Regeln über die Ziffernzählung. Aufstellung weiterer Regeln

Die Regeln 1 bis 4, die wir in § 10 aufgestellt haben, geben an, wie man die Resultate der Grundrechenarten an Näherungszahlen zu runden hat. Dabei werden die im Resultat vorhandenen Fehler manchmal vergrößert, manchmal verkleinert.

So ist z. B. für die gegebenen Zahlen $x = 33{,}1$ und $y = 2{,}52$ das wahre Produkt $xy = 83{,}412$. Wenn wir beide Zahlen auf zwei wesentliche Ziffern abrunden, so erhalten wir $a = 33$ und $b = 2{,}5$ mit dem Produkt 82,5, das von dem genauen Produkt um die Differenz $83{,}412 - 82{,}5 = 0{,}912$ abweicht. Ersetzen wir nun dieses Produkt durch das gemäß Regel 2 auf zwei wesentliche Ziffern gerundete Produkt der Näherungszahlen, also durch 82, so vergrößert sich diese Differenz auf $83{,}412 - 82 = 1{,}412$. Im vorliegenden Fall wird also der wirkliche Fehler des Resultates durch das Runden gemäß Regel 2 vergrößert. Nehmen wir dagegen z. B. das Produkt $1{,}41 \cdot 1{,}73 = 2{,}4393$, wobei die Faktoren als Näherungszahlen mit drei wesentlichen Ziffern für die genauen Werte $\sqrt{2} = 1{,}41421\ldots$ und $\sqrt{3} = 1{,}73205\ldots$ aufgefaßt seien, und runden wir dieses Produkt gemäß Regel 2, so zeigt es sich, daß das aufgerundete Produkt 2,44 von dem genauen Produkt $\sqrt{2} \cdot \sqrt{3} = \sqrt{6} = 2{,}43948\ldots$ etwas weniger abweicht als das nicht aufgerundete Produkt. Eine genauere Untersuchung zeigt allerdings, daß im allgemeinen ein Runden die Genauigkeit verschlechtert, so daß man, wenn es nur um die Genauigkeit ginge, besser daran täte, die angenäherten Resultate nicht zu runden. In der Praxis des Rechnens spielt aber neben der Genauigkeit auch die Einfachheit des Resultates eine bedeutende Rolle: Der Verzicht auf ein Runden zieht die Notwendigkeit nach sich, laufend mit Zahlen großer Ziffernanzahl, ja Zahlen mit unendlich vielen Ziffern umgehen zu müssen. Dadurch würden alle Rechnungen äußerst schwierig, wenn nicht gänzlich undurchführbar werden. Man kann zeigen, daß der ungünstige Einfluß des Rundens fast unmerklich wird, wenn man folgende Regel anwendet:

Regel 5 der Ziffernzählung. *In allen Zwischenergebnissen (d. h. in allen Ergebnissen, die als Grundwerte für weitere Rechenoperationen im Gang der Gesamtrechnung dienen) behalte man jeweils eine Ziffer mehr bei, als die Regeln* 1 *bis* 4 *empfehlen.*

Bei der Anwendung dieser Regel empfiehlt es sich, die auftretenden überschüssigen Ziffern besonders zu kennzeichnen, z. B. sie etwas kleiner zu schreiben. Im Endergebnis wird die überschüssige Ziffer dann natürlich weggelassen. Das Beibehalten von mehr als einer Zusatzziffer ist nur bei besonders komplizierten Berechnungen gerechtfertigt. In der Mehrzahl der Fälle ist dies ohne Einfluß auf die Genauigkeit des Endergebnisses und bedeutet nur eine unnütze Vergrößerung des Arbeitsumfanges.

In eine analoge Lage geraten wir, wenn die Ausgangszahlen eine verschiedene Anzahl von Ziffern (Dezimalstellen bei der Addition und Subtraktion, wesentliche Ziffern bei den anderen Grundrechenarten) besitzen.

Regel 6 der Ziffernzählung. *Es empfiehlt sich, die genaueren Ausgangszahlen derart zu runden, daß sie nur eine überschüssige Ziffer gegenüber der ungenauesten Ausgangszahl besitzen.*

Durch Vergleich der Werte des mittleren quadratischen Fehlers des Resultates (siehe die Tafel auf S. 366) zeigt man, daß es für die Genauigkeit des Resultates ohne Belang ist, wenn man mehrere Zusatzziffern beibehält. Der hinzukommende Arbeitsaufwand läßt es sogar nachteilig erscheinen. Ein Verzicht auf die Zusatzziffer vermindert hingegen die Genauigkeit.

Wir wollen jetzt unter Benutzung der Regeln der Ziffernzählung die Aufgabe lösen, die wir auf S. 352 mit genauer Fehlerrechnung gelöst haben. Gesucht ist $t = \sqrt{a : b}$, $a = 2hd$, $b = g(d - d_1)$, wobei $h \approx 25{,}3$, $d \approx 19{,}32$, $d_1 \approx 0{,}998$, $g \approx 982$.

Die vollständige Lösung lautet:

$2h$	50,6	19,32	19,32	18,322	977,6 : 17990	
$a = 2hd$	977,6	50,6	− 0,998	982	9776 : 179900 = 0,054302 ···	
$d - d_1$	18,322	11 592	18,322	36 644	8995	≈ 0,05430
$b = g(d - d_1)$	17990	966 0	18,322	146576	7810	$\sqrt{0{,}05430} \approx 0{,}2330$
$a : b$	0,05430	977,592		164898	7196	(nach Tafel)
$t = \sqrt{a : b}$	0,233	977,6		17992,204	6140	
				17990	5097	
					4300	

Antwort: $t \approx 0{,}233$

Diese Antwort stimmt vollständig mit der Antwort überein, die wir oben mit dem Verfahren der Wertschranken erhalten haben (siehe S. 352, wo wir gefunden haben: $t \approx 0{,}2331\ (\pm\ 0{,}0007)$).

Als zweites Beispiel wollen wir die Aufgabe lösen, die wir auf S. 348 nach dem Verfahren der Wertschranken gelöst haben: Zu bestimmen ist $\delta = \frac{p}{v}$, $v = \pi r^2 h$, wobei $2r \approx 0{,}048$ cm, $h = 26{,}44$ cm, $p \approx 0{,}423$ g ist. Auf S. 337 haben wir bereits die Schwierigkeiten erwähnt, die unvermeidlich auftreten, wenn man die besonderen Gegebenheiten ignoriert, die bei der Durchführung der Grundrechenarten an Näherungszahlen zu beachten sind.

Im vorgegebenen Fall sind $2r$, h, p mit zwei, vier, bzw. drei wesentlichen Ziffern gegeben. Mit wieviel Ziffern soll man π nehmen? Da unter den durchzuführenden Operationen die Multiplikation und die Division auftreten, wenden wir die Regel 6 an und beschränken uns auf den Wert $\pi = 3{,}14$. Die vollständige Berechnung lautet:

$2r$	0,048	0,000576 · 3,14	0,001809 · 26,44	0,423 : 0,04783
r	0,024	2304	7236	42300 : 4783 = 8,84 ···
r^2	0,000576	576	7236	38264
π	3,14	1728	10854	40360
πr^2	0,00180	0,00180864	3618	38264
h	26,44	0,00180	0,04782996	20960
$v = \pi r^2 h$	0,04783		0,04783	19132
p	0,423			1828
$\delta = \frac{p}{v}$	8,8			

Antwort: $\delta \approx 8{,}8$.

Auch hier stimmt die Antwort vollständig mit der Antwort des Verfahrens der Wertschranken überein. Man muß sich jedoch vor Augen halten, daß bei der Anwendung der Methode der Wertschranken (im betrachteten Fall, wie auch ganz allgemein) die Antwort völlig bestimmt ausfällt, indem sie ganz präzise Angaben über den größtmöglichen Fehler macht. Dagegen gewährleisten die Regeln der Ziffernzählung, auch wenn sie auf dasselbe Ergebnis hinführen, nur mit einer großen Wahrscheinlichkeit, daß der Fehler in der letzten beibehaltenen Ziffer des Resultats ziemlich klein ist. Wenn eine solche unvollständige Bestimmung des Resultats sich als unzulänglich erweist, so ist die genaue Fehlerrechnung unumgänglich. Aber selbst in solchen Fällen ist die Verwendung der Regeln der Ziffernzählung noch nützlich, da man sich mit ihrer Hilfe relativ schnell eine erste Orientierung über die Anzahl der zu erwartenden zuverlässigen Ziffern des Resultats verschaffen und dementsprechend feststellen kann, mit wieviel Ziffern die ganze Berechnung zu führen ist.

Wir wollen noch ein Beispiel für die Zweckmäßigkeit der Regel über das Runden der Zwischenergebnisse (Regel 5) betrachten:

Bei der Auflösung der Gleichung $f(x) = x^4 + 2x^3 - 5x^2 + 3x - 7 = 0$ sei eine ihrer Wurzeln, nämlich $x_1 \approx -3{,}71$, bis auf Hundertstel genau bestimmt worden. Zur Kontrolle des gefundenen Wertes ist $f(x_1)$ zu bestimmen.

x	$-3{,}71$	$-3{,}71$	$-3{,}71$
x^2	$13{,}8$	$13{,}76$	$13{,}764$
x^3	$-51{,}1$	$-51{,}06$	$-51{,}064$
$2x^3$	-102	$102{,}1$	$-102{,}13$
$-5x^2$	$-69{,}0$	$-68{,}80$	$-68{,}820$
$3x$	$-11{,}1$	$-11{,}18$	$-11{,}130$
-7	$-7{,}0$	$-7{,}00$	$-7{,}000$
$2x^3 - 5x^2 + 3x - 7$	-189	$-189{,}00$	$-189{,}08$
x^4	190	$189{,}2$	$189{,}45$
$f(x)$	$+1$	$+0{,}2$	$+0{,}37$
Bei Vernachlässigung der Zusatzziffern:	$+1$	0	0

Wir führen die dazu notwendigen Berechnungen (unter Verwendung der Regeln der Ziffernzählung) in drei Varianten durch: Zunächst ohne Zusatzziffer, dann mit einer Zusatzziffer, schließlich mit zwei Zusatzziffern.

Wir sehen, daß die Verwendung einer Zusatzziffer das Resultat etwas verändert hat, während die Verwendung von zwei Zusatzziffern keine weitere Veränderung bewirkt.

Es bleibt noch ein Fall der Verwendung von Zusatzziffern zu besprechen, die Berechnungen auf eine vorgegebene Genauigkeit. Wenn man die Ausgangszahlen mit mehr oder minder willkürlicher Ziffernanzahl nehmen kann, die Genauigkeit des Resultats aber im voraus verlangt ist, so können wir niemals für die Genauigkeit der letzten Ziffer des Resultates garantieren, sofern wir die Ausgangszahlen nur mit einer solchen Anzahl von Ziffern nehmen, daß das Resultat gemäß den Regeln 1 bis 4 die erforderliche Anzahl von Ziffern erhält, sofern wir also die Ausgangszahlen sozusagen „knapp" wählen. Die Regeln der Ziffernzählung garantieren uns nur, daß ein bedeutender Fehler im Resultat viel unwahrscheinlicher ist als ein kleiner. Diese Unsicherheit in der letzten Ziffer verschwindet weitgehend, wenn wir in allen Ausgangszahlen noch eine Zusatzziffer hinzunehmen. Eine Hinzunahme von zwei Zusatzziffern verspricht im allgemeinen Fall keinen Gewinn an Genauigkeit, sie vergrößert nur die Rechenarbeit (nur im Fall von sehr komplizierten Rechnungen empfiehlt sich die Verwendung von jeweils zwei Zusatzziffern).

Regel 7 der Ziffernzählung. *Wenn die Genauigkeit der Ausgangszahlen willkürlich ist, so muß man, um ein Resultat mit k zuverlässigen Ziffern zu erhalten, die Ausgangszahlen so wählen, daß sie gemäß den Regeln 1 bis 4 auf ein Resultat mit $(k+1)$ Ziffern führen.*

Hier ein Beispiel für eine Berechnung mit vorgegebener Genauigkeit des Resultates: Unter Verwendung der Reihe

$$\lg(1+x) = M\left(x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5 - \cdots\right),$$

wobei $M = 0{,}43429448\ldots$, $-1 < x \leq +1$ ist, ist der vierstellige Logarithmus der Zahl 7 zu bestimmen.

Es ist nicht möglich, sogleich $x = 6$ zu nehmen, da die angegebene Reihe nur für solche x konvergiert, die (absolut) kleiner als 1 sind, und daher auch nur für die

Berechnung der entsprechenden Logarithmen verwendet werden kann. Deshalb werden wir zunächst lg 0,7 bestimmen. Dazu setzen wir $x = -0{,}3$ und führen alle Berechnungen mit einer zusätzlichen Dezimalstelle, also mit fünf $(4 + 1 = 5)$ Dezimalstellen. Wir ziehen dazu alle Glieder der Reihe in den Kreis der Betrachtungen, die beim Runden auf fünf Dezimalstellen nicht verschwinden:

x	$-0{,}30000$	x	$-0{,}30000$
x^2	$0{,}09000$	$-x^2:2$	$-0{,}45000$
x^3	$-0{,}02700$	$x^3:3$	$-0{,}00900$
$x^4 = (x^2)^2$	$0{,}00810$	$-x^4:4$	$-0{,}00202$
$x^5 = x^3 \cdot x^3$	$-0{,}00243$	$x^5:5$	$-0{,}00049$
$x^6 = (x^3)^2$	$0{,}00073$	$-x^6:6$	$-0{,}00012$
$x^7 = x^3 \cdot x^4$	$-0{,}00022$	$x^7:7$	$-0{,}00003$
$x^8 = (x^4)^2$	$0{,}00007$	$-x^8:8$	$-0{,}00001$
$x^9 = x^4 \cdot x^5$	$-0{,}00002$	$x^9:9$	$-0{,}00000$
$x^{10} = (x^5)^2$	$0{,}00001$	S	$-0{,}35667$
M	$0{,}43429$	MS	$-0{,}15490$

Indem wir die Zusatzziffer im Resultat vernachlässigen, erhalten wir dann lg $0{,}7 = -0{,}1549$ und damit

$$\lg 7 = \lg(0{,}7 \cdot 10) = -0{,}1549 + 1 = 0{,}8451.$$

Denselben Wert für lg 7 finden wir in der vierstelligen Logarithmentafel. Wenn wir dieselbe Berechnung mit genauer Fehlerrechnung durchführen wollten, so müßten wir auch noch das Restglied der Reihe berücksichtigen.

Einer besonderen Erwähnung bedarf noch der sogenannte *„Genauigkeitsverlust bei der Subtraktion"*. Bei der Subtraktion zweier Näherungszahlen, die sich nur wenig voneinander unterscheiden und die eine gleiche Anzahl von Dezimalstellen besitzen, ergeben sich in der Differenz weniger wesentliche Ziffern, als vorher in jeder Komponente vorhanden waren. Wenn wir also eine solche Differenz mit einer bestimmten Anzahl von wesentlichen Ziffern berechnen wollen, müssen wir die Komponenten mit einer Stellenzahl kennen, die bedeutend größer ist.

Gesucht sei z. B. der Wert $x = \dfrac{\operatorname{tg}\alpha - \sin\alpha}{(\operatorname{arc}\alpha)^3}$ für $\alpha = 5°$ mit drei wesentlichen Ziffern (mit arc α wird dabei das Bogenmaß von α bezeichnet).

Um den Quotienten auf drei wesentliche Ziffern zu erhalten, muß man gemäß Regel 7 der Ziffernzählung den Dividenden und den Divisor mit vier wesentlichen Ziffern nehmen. Um nun die Differenz $\operatorname{tg} 5° - \sin 5° = 0{,}0875 - 0{,}0872$ mit vier wesentlichen Ziffern zu erhalten, müssen wir nicht, wie wir es eben ansetzten, tg 5° und sin 5° nur mit vier Dezimalstellen, sondern mit sieben Dezimalstellen nehmen. Dagegen genügt es, den Wert von arc 5° mit fünf Dezimalstellen zu kennen. Die weitere Berechnung ist aus dem folgenden Schema unmißverständlich zu ersehen:

arc 5°	0,08727
tg 5°	0,0874887
sin 5°	0,0871557
$a = \operatorname{tg} 5° - \sin 5°$	0,0003330
$b = (\operatorname{arc} 5°)^3$	0,0006646
$x = \dfrac{a}{b}$	0,501

Zu demselben Resultat gelangt man im vorliegenden Fall viel leichter, wenn man den Zähler der gegebenen Funktion auf eine Form bringt, die zum Logarithmieren geeignet ist, und für die Rechnung vierstellige Logarithmen benutzt. Auf diese Weise findet man:

$$x = \frac{\sin 5^\circ (1 - \cos 5^\circ)}{\cos 5^\circ \cdot (\text{arc } 5^\circ)^3} = \frac{2 \operatorname{tg} 5^\circ \sin^2 2^\circ 30'}{(\text{arc } 5^\circ)^3} = \frac{2 \operatorname{tg} 5^\circ \sin^2 2^\circ 30'}{(\pi : 36)^3} \text{ und } x \approx 0{,}5012,$$

oder nach Abrunden auf drei Dezimalstellen, $x \approx 0{,}501$.

Wir wollen noch auf eine Regel hinweisen, mit deren Hilfe man die Genauigkeit einer logarithmischen Berechnung beurteilen kann. Wenn man mittels Logarithmen den Wert eines eingliedrigen[1]) Ausdrucks berechnet, dessen Komponenten genau sind, so wird man einen um so kleineren Fehler begehen, je mehr Stellen die Mantissen der verwendeten Tafel besitzen. Daß hierbei im allgemeinen Fall ein Fehler auftreten wird, erklärt sich daraus, daß alle Logarithmen, ausgenommen die Logarithmen der Zahlen 1, 10, 100 usw., Näherungszahlen sind. Das Experiment wie auch die theoretische Untersuchung zeigen, daß bei Benutzung einer k-stelligen Logarithmentafel der im Resultat auftretende Rechenfehler eine Unzuverlässigkeit der k-ten wesentlichen Ziffer des Resultates bewirkt. Wenn wir z. B. den Wert von $x = 70 : 19$ unter Benutzung von 3-, 4-, 5-, 7-, 12-stelligen Logarithmentafeln berechnen[2]), so finden wir:

k	3	4	5	7	12
lg 70	1,845	1,8451	1,84510	1,8450980	1,845098040014
lg 19	1,279	1,2788	1,27875	1,2787536	1,278753600953
lg x	0,566	0,5663	0,56635	0,5663444	0,566344439061
x	3,68	3,684	3,6842	3,684210	3,68421052632

im Vergleich zum wahren Wert

$$x = 70 : 19 = 3{,}684210526315789473 \cdots,$$

einem periodischen Bruch mit einer Periode von 18 Ziffern. Wir sehen also, daß alle Ziffern der erhaltenen angenäherten Resultate genau sind. Jedoch kann man leicht beliebig viele Rechenbeispiele mit k-stelligen Logarithmen anführen, in denen das gewonnene Resultat von wahren Wert um ein bis zwei Einheiten der Ordnung der k-ten Stelle abweicht. So findet man bei der Berechnung des Quotienten $17 : 6 = 2{,}833 \cdots$ mittels vierstelliger Logarithmen den Wert 2,832, und bei der Berechnung der Potenz $3{,}9^3 = 59{,}319$ mittels fünfstelliger Logarithmen den Wert 59,317.

Will man also mit Hilfe von Logarithmen sicher ein Resultat mit k zuverlässigen, wesentlichen Ziffern bestimmen, so muß man eine mindestens $(k + 1)$-stellige Logarithmentafel benutzen. Diese Überlegungen führen uns auf Regel 8 der Ziffernzählung, die wir unten angeben. Wenn man diese Regel befolgt, so hat man eine

[1]) Hierunter soll verstanden werden, daß in dem betrachteten Ausdruck keine Addition oder Subtraktion vorkommt. — *Anm. d. wissenschaftl. Red.*

[2]) Die zwölfstelligen Logarithmen sind entnommen aus: М. Ф. Субботин, Многозначные таблицы логарифмов. (M. F. Subbotin, Mehrstellige Logarithmentafeln), Verlag d. Akad. d. Wiss. d. UdSSR, 1940.

Gewähr dafür, daß man nicht durch die Verwendung von Logarithmen die Genauigkeit des Resultates vermindert, die auf Grund der Genauigkeit der gegebenen Komponenten erreichbar ist. Desgleichen gibt uns diese Regel die Möglichkeit an die Hand, durch Verwendung einer Logarithmentafel von geeigneter Stellenzahl, nämlich einer Stellenzahl, die der Genauigkeit der Komponenten angepaßt ist, unnötige Rechenarbeit zu vermeiden.

Hier noch eine Zusammenfassung der acht Regeln der Ziffernzählung, die wir oben betrachtet haben:

1. *Bei der Addition und Subtraktion von Näherungszahlen sind im Resultat nur so viele Dezimalstellen beizubehalten, wie in der Ausgangszahl mit kleinster Anzahl von Dezimalstellen vorhanden sind.*

Bemerkung. „Dezimalstellen" einer Zahl sind diejenigen ihrer Ziffern, die rechts vom Komma stehen.

2. *Bei der Multiplikation und Division von Näherungszahlen sind im Resultat nur so viele wesentliche Ziffern beizubehalten, wie in der Ausgangszahl mit kleinster Anzahl von wesentlichen Ziffern vorhanden sind.*

Bemerkung. „Wesentliche Ziffern" einer Zahl sind alle ihre Ziffern mit Ausnahme der Nullen, die links von ihrer ersten, von Null verschiedenen Ziffer stehen.

3. *In der zweiten und dritten Potenz einer Näherungszahl sind nur so viele wesentliche Ziffern beizubehalten, wie in der vorgegebenen Zahl (der Basis) vorhanden sind.*

Bemerkung. Die letzte Ziffer des Quadrates und besonders des Kubus ist weniger zuverlässig als die letzte Ziffer der Basis.

4. *In der Quadrat- und Kubikwurzel aus einer Näherungszahl sind so viele wesentliche Ziffern beizubehalten, wie in der vorgegebenen Zahl (dem Radikanden) vorhanden sind.*

Bemerkung. Die letzte Ziffer der Quadrat- und besonders der Kubikwurzel ist zuverlässiger als die letzte Ziffer des Radikanden.

5. *Bei der Berechnung von Zwischenergebnissen empfiehlt es sich, jeweils eine Ziffer mehr beizubehalten, als die vorangehenden Regeln angeben.*

Bemerkung. Im Endergebnis ist diese Zusatzziffer wegzulassen. Es empfiehlt sich, die Zusatzziffern jeweils besonders zu kennzeichnen.

6. *Wenn einige Ausgangszahlen mehr Dezimalstellen (im Fall der Grundrechenarten erster Stufe) bzw. mehr wesentliche Ziffern (im Fall der Grundrechenarten zweiter und dritter Stufe) als die anderen besitzen, so sind sie vorher so zu runden, daß sie gegenüber den anderen nur eine überschüssige Ziffer besitzen.*

7. *Wenn die Genauigkeit der Ausgangszahlen willkürlich ist, aber verlangt ist, daß das Resultat k Ziffern besitzt, so sind die Ausgangszahlen so zu wählen, daß man gemäß der Regeln 1 bis 4 ein Resultat mit $(k + 1)$ Ziffern erhält.*

8. *Bei der Berechnung des Wertes eines eingliedrigen Ausdrucks mittels Logarithmen hat man eine Logarithmentafel zu benutzen, deren Mantissen eine um 1 größere Stellenanzahl besitzen als die Ausgangszahl mit kleinster Anzahl von wesentlichen Ziffern wesentliche Ziffern besitzt. Im Endergebnis ist die letzte wesentliche Ziffer wegzulassen.*

Bemerkung. Bei den Regeln der Ziffernzählung sind alle Nullen unberücksichtigt zu lassen, die am Ende der Näherungszahlen stehen und unbekannte Ziffern ersetzen.

Kapitel III

VERSCHIEDENE FRAGEN

§ 14. Näherungsformeln. Abgekürzte Verfahren für die Grundrechenarten

Bei praktischen Rechnungen finden eine Reihe von Näherungsformeln häufige Verwendung. Ihre Begründung ist selbst den Schülern der höheren Klassen der Oberschule bereits zugänglich. So prüft man leicht, daß identisch

$$\frac{1}{1+x} = 1 - x + \frac{x^2}{1+x}$$

ist. Hieraus folgt, daß man in den Fällen, in denen die Zahl x so klein ist, daß man bei der üblichen Rechengenauigkeit ihr Quadrat vernachlässigen kann, den Quotienten $1 : (1 + x)$ durch die Differenz $1 - x$ ersetzen darf.

Wenn wir beachten, daß die Gleichung

$$\frac{x^2}{1+x} = 0{,}005$$

(genähert) die Wurzeln 0,0733 und $-0{,}0683$ besitzt, so können wir feststellen, daß diese Ersetzung einen maximalen Fehler von der halben Ordnung der zweiten Dezimalstelle (also einem halben Hundertstel) bedingt, wenn x im Intervall von $-0{,}0683$ bis 0,0733 liegt, wenn also der absolute Betrag von x die Zahl 0,068 nicht überschreitet. Analog stellt man fest, daß der Fehler der Näherungsgleichung $1 : (1 + x) \approx 1 - x$ den Wert $0{,}5 \cdot 10^{-3}$ nicht übertrifft, wenn $|x| < 0{,}022$ ist, und daß er $0{,}5 \cdot 10^{-4}$ nicht übertrifft, wenn $|x| < 0{,}0074$ ist. Mittels dieser Formel können wir z. B. leicht den Quotienten $1 : 0{,}997$ berechnen, indem wir $x = -0{,}003$ setzen; dann erhalten wir $1 - x = 1 + 0{,}003$ und können mit Sicherheit behaupten, daß $1 : 0{,}997 \approx 1{,}0030$ ist (die unmittelbare Division ergibt $1{,}0030090\cdots$). Die Näherungsformel $1 : (1 + x) \approx 1 - x$ erleichtert also, wie wir sehen, in gewissen Fällen die Division sehr erheblich; sie gestattet es aber auch — und das ist noch wichtiger — die Änderung des Quotienten bei geringfügiger Änderung des Divisors zu berechnen: Wenn die Zahl Δb sehr klein im Vergleich zu b ist, so ist

$$\frac{a}{b+\Delta b} = \frac{a}{b\left(1+\frac{\Delta b}{b}\right)} = \frac{a}{b} \cdot \frac{1}{1+\frac{\Delta b}{b}} \approx \frac{a}{b}\left(1-\frac{\Delta b}{b}\right),$$

$$\frac{a}{b+\Delta b} - \frac{a}{b} \approx -\frac{a}{b} \cdot \frac{\Delta b}{b}.$$

Ebenso erhalten wir, wenn wir das Quadrat x^2 vernachlässigen, die Näherungsformel

$$(1+x)^2 \approx 1 + 2x.$$

Wenn wir in dieser Näherungsformel $2x$ durch y ersetzen und auf beiden Seiten die Wurzel ziehen, so erhalten wir die neue Näherungsformel

$$\sqrt{1+y} \approx 1 + 0{,}5\, y.$$

Im folgenden geben wir eine Liste der wichtigsten Näherungsformeln an, die besonders häufig beim praktischen Rechnen auftreten. Dabei geben wir gleichzeitig die maximalen Werte des Betrages des Argumentes x an, für die der Fehler bei Anwendung der betreffenden Formel den Wert $0{,}5 \cdot 10^{-k}$ für $k = 2, 3, 4$ nicht übersteigt. Man berechnet sie am einfachsten, indem man für eine Folge von wachsenden Werten des Argumentes den Wert der linken und der rechten Seite der Formeln bestimmt (und zwar unter Benutzung von geeigneten Tafeln).

	Formel	$k=2$	$k=3$	$k=4$
I	$(1+x)(1+y) \approx 1+x+y, \quad \lvert x\rvert \geq \lvert y\rvert$	0,07	0,022	0,007
II	$(1+x)(1+y)(1+z) \approx 1+x+y+z, \quad \lvert x\rvert \geq \lvert y\rvert \geq \lvert z\rvert$	0,04	0,012	0,004
III	$(1+x)^2 \approx 1+2x$	0,07	0,022	0,007
IV	$(1+x)^3 \approx 1+3x$	0,04	0,012	0,004
V	$\frac{1}{1+x} \approx 1-x$	0,07	0,022	0,007
VI	$\frac{1}{1+x} \approx 1-x+x^2$	0,18	0,081	0,037
VII	$\sqrt{1+x} \approx 1+\frac{1}{2}x$	0,21	0,064	0,020
VIII	$\sqrt{1+x} \approx 1+\frac{1}{2}x-\frac{1}{8}x^2$	0,46	0,20	0,09
IX	$\sqrt[3]{1+x} \approx 1+\frac{1}{3}x$	0,22	0,068	0,021
X	$\sqrt[3]{1+x} \approx 1+\frac{1}{3}x-\frac{1}{9}x^2$	0,47	0,21	0,09
XI	$\sin x \approx x$	17°	8°,2	3°,8
XII	$\sin x \approx x-\frac{1}{6}x^3$	51°	32°	20°
XIII	$\cos x \approx 1$	5°,7	1°,8	0°,5
XIV	$\cos x \approx 1-\frac{1}{2}x^2$	33°	18°	10°
XV	$\operatorname{tg} x \approx x$	14°	6°,4	3°,0
XVI	$\operatorname{tg} x \approx x+\frac{1}{3}x^3$	29°	18°	11°
XVII	$\lg(1+x) \approx 0{,}4343\,x$	0,15	0,048	0,015
XVIII	$\ln(1+x) \approx x$	0,10	0,031	0,010
XIX	$e^x \approx 1+x$	0,09	0,031	0,010
XX	$10^x \approx 1+2{,}303\,x$	0,04	0,014	0,004
XXI	$\lg\frac{1+x}{1-x} \approx 0{,}8686\,x$	0,25	0,119	0,055
XXII	$\ln\frac{1+x}{1-x} \approx 2x$	0,19	0,090	0,042

Hierbei bezeichnet lg den BRIGGschen Logarithmus und ln den natürlichen Logarithmus. In den trigonometrischen Formeln ist x im Bogenmaß ausgedrückt.

In sehr vielen Spezialfällen erleichtern die Näherungsformeln die Berechnungen ganz erheblich. Demselben Zweck dienen die sogenannten „abgekürzten Verfahren" für die Multiplikation, Division und das Ausziehen der Quadratwurzel. Sie sind fast immer anwendbar, besonders aber dann, wenn es sich bei den Ausgangszahlen um Näherungszahlen handelt. Hinsichtlich der Frage nach der Bedeutung dieser Verfahren für die Schule gehen die Meinungen sehr auseinander. So schenkt Prof. W. L. GONTSCHAROW in seinem Buch [5] diesen Verfahren große Beachtung.

Wenn man sich jedoch vergegenwärtigt, daß diese Verfahren, auf die schon seit sehr langer Zeit hingewiesen worden ist, keinerlei Verbreitung in der Schule gefunden haben, obgleich sie in fast jedem Buch für Lehrer behandelt werden, das der Rechentechnik gewidmet ist, so scheint die zweite Meinung zutreffender, nach der die genannten Verfahren im Vergleich zu den Rechengeräten und den Tafelwerken eine nur zweitrangige Bedeutung besitzen. Wie dem auch sei, für den Lehrer wird es stets nützlich sein, sie zu kennen.

Es sei z. B. nach den ersten vier wesentlichen Ziffern des Produktes $29{,}97 \cdot 2{,}738$ gefragt. Nachfolgend vollziehen wir diese Multiplikation einmal nach der üblichen Methode (links) und zum anderen nach dem „abgekürzten Verfahren" (rechts).

$29{,}97 \cdot 2{,}738$
23976
8991
20979
5994
82,05786
82,06

$29{,}97 \cdot 2{,}738$
8372
5994
2098
90
24
82,06

Dabei werden die Ziffern des Multiplikators in umgekehrter Reihenfolge unter die Ziffern des Multiplikanden geschrieben. Wenn es nun darum geht, vier wesentliche Ziffern des Produktes zu bestimmen, so schreibt man die Ziffer höchster Ordnung des Multiplikators unter die vierte wesentliche Ziffer (von links) des Multiplikanden. Jedes Teilprodukt ergibt sich durch Multiplikation (der entsprechenden Ziffer des Multiplikators) mit jenen Ziffern des Multiplikanden, die über und links von der betreffenden Ziffer des Multiplikators stehen. So multipliziert man im Beispiel die Ziffer 2 mit der Zahl 2997, während man die Ziffer 7 nur mit der Zahl 299, die Ziffer 3 bereits nur mit 29 und die Ziffer 8 mit der Zahl 2 multipliziert. An die Stelle der vernachlässigten Ziffern des Multiplikanden tritt eine näherungsweise Korrektur. So stellen wir im Beispiel fest, daß bei der Berechnung des zweiten Teilproduktes die vernachlässigte Ziffer 7 bei der Multiplikation mit 7 ungefähr 5 Zehner ergibt; wir behalten dann diese Korrektur 5 im Sinn, multiplizieren 9 mit 7 und addieren zu diesem Produkt 63 die Korrektur 5, erhalten also 68; die Ziffer 8 schreiben wir unter die äußerste rechte Ziffer und behalten die 6 im Sinn; die weiteren Ziffern des Teilproduktes werden in der üblichen Weise gewonnen.

Auch die Stellung des Kommas kann man nach einer genauen Regel bestimmen, jedoch ist es einfacher, wenn man die Stelle des Kommas durch eine grobe Abschätzung gewinnt. Wenn wir also im Beispiel die Ziffern 8206 erhalten haben und wenn wir beachten, daß die Faktoren nahe an 30 bzw. 3 liegen, so stellen wir fest, daß als Produkt nur die Zahl 82,06 in Frage kommt, und nicht etwa 8,206 oder 820,6.

Wenn wir die Rechnung so führen, wie es eben angegeben wurde, so erhalten wir entweder gerade so viele Ziffern wie verlangt werden oder eine mehr. Im zweiten Fall lassen wir diese überschüssige Ziffer weg.

Die Rechtfertigung dieses Verfahrens bereitet keine Schwierigkeit. Man braucht nur die Teilprodukte der vollständigen Rechnung mit denen der abgekürzten Multiplikation zu vergleichen. Der Fehler des Resultates der abgekürzten Multiplikation übertrifft (bei genauen Faktoren) nicht die mit der Ziffernanzahl des Multiplikators multiplizierte halbe Einheit der letzten Ordnung des Produktes.

Die abgekürzte Division betrachten wir am Beispiel des auf vier wesentliche Ziffern zu bestimmenden Quotienten aus 81,3747 und 0,377264. Nachfolgend geben wir die Berechnung nach dem gewöhnlichen Verfahren (links) und nach dem abgekürzten Verfahren (rechts).

81,37 : 0,377264 = 215,6⋯		813747 : 377264 = 215,7	
75,45	28	= 215,7	7545
592	190		592
377	264		377
214	9260		215
188	6320		189
26	29400		26
22	63584		26
3	658160		0

Dabei kann man auch hier alle Ziffern rechts vom vertikalen Strich vernachlässigen; d. h., wir sondern im Divisor so viele wesentliche Ziffern ab, wie im Quotienten verlangt sind, also im betrachteten Beispiel vier wesentliche Ziffern, und beginnen die Division nach dem üblichen Verfahren. Dem Komma im Dividenden und im Divisor wenden wir zunächst keine Aufmerksamkeit zu. Der weitere Verlauf der Division unterscheidet sich dann von der üblichen Division darin, daß wir nach der Berechnung jeder Ziffer des Quotienten im Divisor eine (jeweils die letzte) Ziffer fortlassen und die nachfolgenden Ziffern des Dividenden nicht herunterziehen. Nachdem wir also im Beispiel 8137 durch 3772 geteilt haben und als erste Ziffer des Quotienten die Ziffer 2 gefunden haben, multiplizieren wir 3772 mit 2 und erhalten nach Korrektur der vernachlässigten Ziffer des Divisors als Produkt 7545 und als ersten Rest 592; jetzt lassen wir die letzte Ziffer des Divisors weg, teilen also 592 nur durch 377 und erhalten als zweite Ziffer des Quotienten die Ziffer 1, die wir mit 377 multiplizieren, wobei wir als zweiten Rest 215 bekommen; diesen teilen wir durch 37, erhalten als dritte Ziffer des Quotienten die Ziffer 5, deren Produkt mit 37 zusammen mit der Korrektur der vernachlässigten Ziffern des Divisors den Wert 189 ergibt, wobei wir als dritten Rest 26 erhalten; schließlich bleibt 26 durch 3 zu teilen; wenn wir hier im Quotienten die Ziffer 8 nehmen, so erhalten wir als Produkt zusammen mit der Korrektur der vernachlässigten Ziffern des Divisors den Wert 30 und als Rest −4; wenn wir dagegen nicht 8 sondern 7 nehmen, so liefert das Produkt zusammen mit der Korrektur gerade 26.

Damit ist die Ziffernfolge im Quotienten festgelegt. Es bleibt also nur noch die Stellung des Kommas zu bestimmen. Dies macht man wieder am besten durch eine grobe Abschätzung. So finden wir im Beispiel, daß der Quotient nahe an

$80 : 0{,}4 = 200$ liegen muß; daher setzen wir in der erhaltenen Ziffernfolge das Komma hinter die dritte wesentliche Ziffer und erhalten als Quotienten 215,7.

Die Regel der abgekürzten Division wird dadurch gerechtfertigt, daß man Schritt für Schritt die vollständigen Werte mit den abgekürzten Werten vergleicht.

Es verbleibt das abgekürzte Verfahren für die Berechnung der Quadratwurzel. Dieses beruht auf folgendem

Satz. *Wenn man nach der Berechnung der ersten n wesentlichen Ziffern einer Quadratwurzel den erhaltenen Rest durch den doppelten Näherungswert der Wurzel dividiert, so erhält man in den ersten $(n-1)$ Ziffern des Quotienten die $(n-1)$ folgenden Ziffern des Wertes der Wurzel.*

Zum Beweis dieses Satzes setzen wir voraus, daß der Radikand b einen ganzen Bestandteil mit $2n-1$ oder $2n$ wesentlichen Ziffern besitzt und daß die ersten n Ziffern der Wurzel gefunden seien, die für sich die Zahl a bilden mögen. Gefragt ist nach dem gebrochenen Bestandteil der Wurzel, den wir mit x bezeichnen wollen. Offensichtlich ist dann

$$\sqrt{b} = a + x, \quad b = a^2 + 2ax + x^2, \quad \frac{b-a^2}{2a} = x + \frac{x^2}{2a}.$$

Dabei ist die Differenz $b - a^2$ nichts anderes als der Rest, der nach der Berechnung der ersten n Ziffern der Wurzel verbleibt, und der Quotient $\frac{b-a^2}{2a}$ ist eben jener Quotient, von dem im obigen Satz die Rede ist. Daraus folgt, daß

$$x = \frac{b-a^2}{2a} - \frac{x^2}{2a}$$

ist. Wenn wir nun $x \approx \frac{b-a^2}{2a}$ setzen, so begehen wir den Fehler $\frac{x^2}{2a}$, für den wegen $x < 1$, $a \geq 10^{n-1}$

$$\frac{x^2}{2a} < 0{,}5 \cdot 10^{-(n-1)}$$

gilt. Falls wir also bei der Division von $b - a^2$ durch $2a$ nach $(n-1)$ Schritten, d. h. nachdem wir $(n-1)$ Dezimalstellen des Quotienten bestimmt haben, haltmachen und den gefundenen Wert in bekannter Weise runden, so kommt zu dem genannten Fehler noch ein Rundungsfehler hinzu, und der Gesamtfehler nähert sich im ungünstigsten Fall einer Einheit der Ordnung der $(n-1)$-ten Dezimalstelle, den er jedoch niemals vollständig erreicht.

Die Voraussetzung über die Kommastellung im Radikanden bedeutet keine Einschränkung der Allgemeinheit, da wir das Komma stets auf den betreffenden Platz bringen können, indem wir den Radikanden mit einer passenden geraden Potenz von 10 multiplizieren (bzw. durch eine passende Potenz dividieren) und den gefundenen Wurzelwert entsprechend durch die Potenz von 10 mit halbem Exponenten dividieren (bzw. mit ihr multiplizieren). In der Praxis braucht man diese Umformungen natürlich nicht vorzunehmen.

Wir wollen diese Regel an einem Beispiel demonstrieren: Es soll etwa $\sqrt{10}$ mit sieben wesentlichen Ziffern bestimmt werden. Nach dem üblichen Verfahren bestimmen wir zunächst die ersten vier wesentlichen Ziffern und danach durch Division des Restes durch den doppelten Näherungswert die drei weiteren. Zum Vergleich

geben wir daneben (rechts) die Berechnung aller sieben Ziffern nach dem üblichen Verfahren an.

$\sqrt{10} = 3{,}162277$

9	
61 | 100	
1 | 61	
626 | 3900	
6 | 3756	
6322 | 14400	
2 | 12644	
1756 : 6324	
1265	
491	
443	
48	
44	
4	

$\sqrt{10} = 3{,}162277\cdots$

9	
61 | 100	
1 | 61	
626 | 3900	
6 | 3756	
6322 | 14400	
2 | 12644	
63242 | 175600	
2 | 126484	
632447 | 4911600	
7 | 4427129	
6324547 | 48447100	
7 | 44271829	
	4175271

Den Rest 1756 haben wir als ganz angesehen und durch den doppelten Näherungswert der Wurzel dividiert, den wir gleichfalls als ganz angesehen haben. Die Ziffern des Quotienten haben wir einfach neben die bereits erhaltenen Ziffern der Wurzel geschrieben. In Wirklichkeit haben wir natürlich den Rest $1756 \cdot 10^{-6}$ durch den doppelten Näherungswert $6324 \cdot 10^{-3}$ dividiert und den erhaltenen Quotienten $0{,}277 \cdot 10^{-3}$ an die betreffenden Stellen geschrieben.

§ 15. Mathematische Tafeln

Unter allen Rechenhilfsmitteln besitzen gegenwärtig die mathematischen Tafeln wegen ihrer Verbreitung, Einfachheit und Bequemlichkeit die größte Bedeutung. Beim praktischen Rechnen werden laufend Tafeln verschiedener Art benötigt. Sie stellen ein schönes Rechenhilfsmittel dar, sind außerordentlich einfach in ihrer Ausstattung und Anwendung, allgemein zugänglich wegen ihres geringen Preises, garantieren in hohem Maße für die aufgenommenen Werte und ermöglichen eine große Zeit- und Kraftersparnis.

Die größte Verbreitung besitzen Tafeln, die eine Abhängigkeit zwischen zwei veränderlichen Größen angeben (zwischen Argumenten und Funktionswerten). Von dieser Art sind die Logarithmentafeln, die Tafeln der Quadrate und Kuben usw. Viel seltener sind Tafeln, in denen die Abhängigkeiten zwischen drei Veränderlichen aufgeführt sind (Tafeln von Funktionen zweier Argumente), wie z. B. die Tafeln irgendwelcher Produkte.

Für ein tiefergehendes Studium der Fragen, die mit der Ausstattung und Anwendung von mathematischen Tafeln zusammenhängen, ist eine besondere mathematische Disziplin — ein Zweig der Analysis — heranzuziehen, die „endliche Differenzenrechnung". Elementare Bemerkungen über Tafeln, die bereits den Schülern der höheren Klassen der Oberschulen zugänglich sind, finden sich in den „Er-

läuterungen" der (im Vergleich zu den früheren Auflagen wesentlich überarbeiteten) 19. Auflage der „Vierstelligen mathematischen Tafeln" von W. BRADIS (1948). Ausführliche Hinweise kann man auch den Büchern [2a] und [2b] entnehmen. Viele interessante und lehrreiche Beispiele und Übungen finden sich in dem Buch [5]; es wäre sehr wünschenswert, wenn sie in Verbindung mit der Behandlung von Tafeln auch in die Schulmathematik Eingang fänden.

Im Schulunterricht haben sich am weitesten Logarithmentafeln und Tafeln der trigonometrischen Funktionen eingebürgert. Wir bemerken, daß für die meisten Aufgaben die Verwendung vierstelliger Tafeln ausreicht. Mit ihrer Hilfe kann man alle Resultate mit vier wesentlichen Ziffern erhalten (wobei allerdings häufig die letzte Ziffer nicht ganz zuverlässig ist). Wenn dem Schüler der Umgang mit vierstelligen Tafeln vollständig geläufig geworden ist, so ist es wünschenswert, daß er auch mit dem Gebrauch von vollständigeren Tafeln, etwa siebenstelligen Logarithmen, vertraut gemacht wird, da mitunter Aufgaben auftauchen (in der Schulpraxis freilich selten), welche eine Berechnung mit erhöhter Genauigkeit erfordern.

Neben den Logarithmentafeln gibt es zahlreiche Tafeln, deren Einführung in den beständigen Schulgebrauch dringend empfohlen werden muß. Hierzu gehören in erster Linie die Tafeln der Quadrate und Quadratwurzeln. Sie können in der Schule bereits vor allen anderen Tafeln benutzt werden und ermöglichen eine fühlbare Zeitersparnis (man hat doch sehr oft Zahlen ins Quadrat zu erheben und aus gegebenen Zahlen die Quadratwurzel zu ziehen!). Wir bemerken, daß man bei der Verwendung einer hinreichend vollständigen Tafel der Quadrate keine gesonderte Tafel für die Quadratwurzeln benötigt. Dasselbe gilt in gleichem Maße für die Tafeln der Kuben und der Kubikwurzeln. Überaus nützlich sind auch Tafeln für die Werte der trigonometrischen Funktionen. Sie machen in vielen Fällen die Reduktion auf eine Gestalt, die für das Logarithmieren geeignet ist, überflüssig und gestatten es, in der achten Klasse Dreiecksaufgaben zu behandeln, bevor die Logarithmen bekannt sind. Daneben können auch Reziprokentabellen der Zahlen und Tafeln für die Umrechnung vom Gradmaß in Bogenmaß Verwendung finden.

Falls eine gegebene Zahl wiederholt als Faktor oder als Divisor auftritt, empfiehlt es sich, eine Produkttafel selbst anzufertigen, die die Produkte der gegebenen Zahl mit den ersten zehn Zahlen enthält. Diese Tafeln erhält man schnell durch fortlaufende Addition der betreffenden Zahl. Hierbei ist die Addition bis zum zehnfachen Wert der gegebenen Zahl zu führen, der eine gute Kontrollmöglichkeit für die Richtigkeit der ganzen Tafel darstellt. Die Tafeln für den Kreisumfang und die Tafeln für die Umformung von Gradmaß in Bogenmaß sind nichts anderes als ausführliche Produkttafeln (nämlich der Zahlen π bzw. $\pi : 180$).

Die größte Schwierigkeit beim Gebrauch von Tafeln bietet die „Interpolation", d. h. die Berechnung von Funktionswerten für solche Argumentwerte, die zwischen zwei aufeinanderfolgenden Tafelwerten liegen, und der umgekehrte Prozeß. Dieses „Lesen zwischen den Zeilen der Tafel" wird für gewöhnlich so ausgeführt, daß man annimmt, daß sich die Funktionswerte zwischen den in der Tafel aufgeführten Argumentwerten gleichmäßig (linear) ändern. Diese „lineare" Interpolation ist aber nur dann zulässig, wenn sich zwischen den Tafelwerten, die äquidistanten Argumentwerten entsprechen, gleiche oder zumindest sich nur langsam ändernde Differenzen („Tafeldifferenzen") befinden. Nur ein volles Verständnis des Wesens der linearen Interpolation und der Bedingungen für ihre Zulässigkeit garantiert eine bewußte

und nicht nur mechanische Anwendung solcher überaus nützlicher und viel verwendeter Hilfsmittel für die lineare Interpolation wie der „proportionalen Teilung" und der „fertigen Interpolationstafeln". Man kann hier ergänzen, daß die Schüler durchaus nicht nur solche fertigen Interpolationstafeln verwenden müssen, sondern daß sie sich auch selbst derartige Tafeln zusammenstellen können.

§ 16. Graphische Berechnungen

In Fällen, in denen eine Genauigkeit von zwei bis drei wesentlichen Ziffern ausreicht, kann man oft die Berechnung des Resultates durch eine Konstruktion oder durch ein Ablesen aus einer fertigen Zeichnung ersetzen. Derartige graphische Methoden für die Lösung von Rechenaufgaben, die die Arbeit häufig um das zehn- oder hundertfache verkürzen, besitzen gegenwärtig weiteste Verbreitung in allen Zweigen der Technik und werden besonders dann angewendet, wenn es auf eine möglichst schnelle Ermittlung eines angenäherten Zahlenergebnisses ankommt. Die Oberschule würde sicherlich denjenigen ihrer Absolventen, die an Ingenieurschulen oder technischen Hochschulen studieren wollen, viel Mühe bei der Erlernung der graphischen Methoden ersparen, wenn sie die einfacheren graphischen Verfahren pflegen würde. Diese sind zum Teil selbst den Schülern der Grundschule verständlich. Allerdings sind die graphischen Berechnungen bis heute noch nicht vollständig methodisch verarbeitet, so daß in absehbarer Zeit kaum mit ihrer weitgehenden und planmäßigen Aufnahme in die Lehrpläne der Oberschule zu rechnen ist. Um so wichtiger ist es, dem Lehrer nachdrücklich zu empfehlen, daß er sich bei jeder Rechenaufgabe die Frage nach einer graphischen Lösung und deren Vorteilen gegenüber der üblichen rechnerischen Lösung vorlegen sollte. Besonders betonen wir hier nochmals, daß in jedem Falle die Vorzüge der graphischen Lösung zu klären sind: Im Prinzip kann man nämlich jede Rechenaufgabe graphisch lösen, jedoch bedeutet eine graphische Lösung nur manchmal einen Gewinn, während sie häufig gerade das Gegenteil bedeutet; wenn wir aber eine Aufgabe auf graphischem Wege lösen, die mit den üblichen Rechenmitteln viel einfacher und besser gelöst wird, so kompromittieren wir nur die graphische Methode in den Augen der Schüler.

Besonders auffallend sind die Vorteile der graphischen Methode bei sogenannten „Serienrechnungen", bei denen nach ein und derselben Formel, aber mit verschiedenen Ausgangszahlen, die Resultate zu bestimmen sind. Hierfür ein einfaches Beispiel:

Vorgelegt sei uns die Folge der Zahlen 18, 23, 38, 57, 85, 92, die etwa die Preise von verschiedenen Waren in D-Mark bedeuten. Es wird verlangt, diese Preise jeweils um 27% herabzusetzen. Bei der rechnerischen Lösung haben wir jede dieser Zahlen mit 0,73 zu multiplizieren. Falls uns eine Produkttafel oder ein anderes geeignetes Rechengerät zur Verfügung steht, werden wir die Rechnung unmittelbar ausführen. Sehr leicht läßt sich die Rechnung jedoch auch mittels eines graphischen Verfahrens durchführen. Dazu zeichnen wir (Abb. 7) auf ein Stück kariertes Papier, am besten Millimeterpapier, ein rechtwinkliges Dreieck mit den Katheten 100 mm und 73 mm, aus dem sich unmittelbar die erniedrigten Preise 13, 17, 28, 42, 62, 67 ablesen lassen (die Produkttafel liefert: 13,14; 16,79; 27,74; 41,61; 62,05; 67,16).

Auf Millimeterpapier lassen sich bequem graphische Lösungen von Aufgaben in linearen Funktionen $y = a\,x + b$ darstellen. Dagegen benutzt man die handelsüblichen „logarithmischen Papiere" für die graphische Lösung von Aufgaben, die mit Potenzfunktionen der Form $y = a\,x^{\alpha}$ zusammenhängen, wobei a und α beliebige reelle Zahlen sein können; auf gewissen derartigen Papieren sind die graphischen

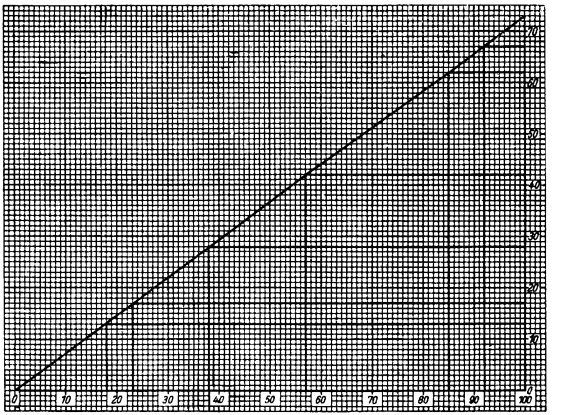


Abb. 7

Darstellungen dieser Funktionen gerade Linien, während auf anderen Papieren die Exponentialfunktionen $y = a\,b^{x}$ und die logarithmischen Funktionen $y = a \cdot \lg x + b$ als Geraden erscheinen (näheres findet man in dem Buch [2b]).

Das Bestreben, die Durchführung von Serienrechnungen der genannten Art zu erleichtern, hat zur Begründung eines ganzen neuen Zweiges der Mathematik, der sogenannten Nomographie geführt. Ihr sind eine Reihe von Büchern gewidmet[1]), unter denen für eine erste Einführung besonders das Buch [13] zu empfehlen ist; einige Beispiele, die besonders gut für die Schule geeignet sind, finden sich in dem Buch [2b]. Den Leser, der sich mit der Nomographie eingehender vertraut machen möchte, verweisen wir auf das Buch [14].

Wenn wir ein und dieselbe Aufgabe einmal rechnerisch und zum anderen graphisch lösen, gewöhnen wir die Schüler an eine brauchbare Selbstkontrolle, da die graphische Methode leicht grobe Rechenfehler aufdeckt. Dabei empfiehlt es sich sogar, mit der graphischen Lösung, die im allgemeinen weniger genaue Resultate als die Rechnung liefert, zu beginnen, weil sonst unwillkürlich auf Grund der Kenntnis des

[1]) Über die Anwendung der Nomographie auf die näherungsweise Auflösung von Gleichungen vgl. EdEM Bd. 2, A. P. Domorjad. Numerische und graphische Methoden für die Auflösung von Gleichungen.

genaueren Wertes des Resultates der graphisch gewonnene, weniger genaue Wert dem genaueren Wert „angepaßt" wird. In dieser Reihenfolge sind wir z. B. auch in § 3 (Seite 323) bei der Lösung der dort gestellten geometrischen Aufgabe vorgegangen.

§ 17. Der logarithmische Rechenschieber

Ohne Zweifel besteht für die Rationalisierung der Rechenarbeit in der Oberschule die dringendste Aufgabe gegenwärtig in der Vermittlung der Theorie und Praxis des logarithmischen Rechenschiebers. Dieses Recheninstrument ist einfach in seinem Aufbau, klein in seiner Dimension und deswegen handlich und erschwinglich im Preis. Mit seiner Hilfe kann man außerordentlich schnell Produkte, Quotienten, Potenzen, Wurzeln, Werte von Logarithmen und trigonometrischen Funktionen sowie eine große Zahl weiterer Rechenresultate nach komplizierteren Formeln ermitteln. Der Rechenschieber normaler Länge, d. h. mit einer 250 mm langen Skala, liefert Ergebnisse mit drei wesentlichen Ziffern; falls die erste wesentliche Ziffer eine 1 ist, erhält man sogar vier wesentliche Ziffern. Der Rechenschieber leistet allerdings keine Hilfe bei der Addition und der Subtraktion; deshalb ist es zweckmäßig, ihn mit dem Rechenbrett zu kombinieren. Der Rechenschieber liefert die Resultate im allgemeinen schneller als die üblichen Rechenmaschinen, so daß zumal bei Serienrechnungen der Geschwindigkeitsgewinn erheblich ist. So erhält man z. B. die Werte von $s = 4{,}9\,t^2$ für die Folge der Werte t von 3 bis 7 im Abstand 0,5 im Laufe von 30 bis 40 Sekunden:

$t =$	3	3,5	4	4,5	5	5,5	6	6,5	7
$s =$	44,1	60,0	78,4	99,2	122,4	148,0	176,4	207	240.

Bei der Arbeit von zwei Personen, von denen die eine den Rechenschieber handhabt und die andere die erhaltenen Werte nach Diktat aufschreibt, beschleunigt sich die Arbeit nochmals um ungefähr 30%.

Wie schon erwähnt, liefert der Rechenschieber die Resultate nur mit einer beschränkten Genauigkeit, so daß er damit z. B. für Finanzrechnungen ungeeignet wird, da man es bei ihnen häufig mit Zahlen aus 6, 7, 8 und mehr Ziffern zu tun hat und die Rechnung auf den Pfennig genau stimmen muß, wie groß auch immer die Anzahl der Ausgangszahlen sein mag. Anders steht es aber z. B. bei technischen Berechnungen, bei denen man in der überwiegenden Mehrheit der Fälle nur an den ersten drei wesentlichen Ziffern, ja mitunter nur an den ersten beiden interessiert ist. Hierbei ist die mit dem Rechenschieber erhaltene, beschränkte Genauigkeit der Resultate sogar oft geradezu vorteilhaft, da alle (oder fast alle) Ziffern, die zu vernachlässigen sind (weil sie ungenau sind), von selbst fortfallen. Es ist deshalb ganz natürlich, daß der Rechenschieber eine weite Verbreitung gefunden hat: Ohne ihn kommt heute kein Ingenieur, kein Techniker und kein Student der technischen Hochschule mehr aus.

In der Oberschule wird gegenwärtig der Rechenschieber nicht behandelt, obwohl es eine Zeit gab, in der die Vermittlung seines Aufbaues und seines Gebrauchs zum üblichen Schulpensum gehörte.[1]) Das ist daraus zu erklären, daß einmal bis vor

[1]) Dies bezieht sich naturgemäß auf die Schule in der Sowjetunion. — *Anm. d. wissenschaftl. Red.*

kurzem nicht genügend billige und hinsichtlich der Präzision befriedigende Rechenschieber hergestellt wurden, und zum anderen dadurch, daß die Mehrzahl der Mathematiklehrer nicht hinreichend geschickt mit dem Rechenschieber umzugehen vermochte. Auch an den pädagogischen Hochschulen ist bis in letzter Zeit dem Studium des Rechenschiebers nicht mehr genügend Beachtung geschenkt worden. Der erste Umstand beginnt nach und nach zu schwinden: In den Läden erscheinen allmählich wieder hinreichend gute und nicht teure Rechenschieber. Um so bedauerlicher ist der zweite Grund. Falls jedoch ein Lehrer die Arbeitsweise des Rechenschiebers gut beherrscht und seinen Wert richtig einschätzt, so sollte er interessierten Schülern vorschlagen, daß sie sich mit ihm in zusätzlicher Arbeit, etwa im Rahmen einer Arbeitsgemeinschaft, mit dem Studium des Rechenschiebers befassen; er wird dann sicher eine Gruppe von Schülern finden, die sich am eigenen Rechenschieber die Fertigkeit in seinem Gebrauch erfolgreich aneignen werden.

Heute gibt es eine ganze Anzahl von Lehrbüchern, in denen die Theorie des logarithmischen Rechenschiebers entwickelt wird, und welche Anweisungen für ein schnelles Erlernen seines Gebrauches geben; Kapitel über den Rechenschieber finden sich auch in fast allen Büchern über praktisches Zahlenrechnen. Wir geben im folgenden den Anfang des Kapitels über den logarithmischen Rechenschieber aus dem Buch [2a] des Verfassers wieder, in welchem die Grundprinzipien dieses Instrumentes und die ersten Anweisungen für seinen Gebrauch entwickelt werden.

Wenn uns zwei gewöhnliche Millimeterlineale von z. B. je 300 mm Länge zur Verfügung stehen, so können wir uns mit ihrer Hilfe mühelos ein Gerät für die mechanische Addition und Subtraktion von ganzen Zahlen, die selbst und deren Resultat unter 300 liegen, herstellen. Dazu legen wir die Lineale so aneinander, daß sich die Kanten mit der Stricheinteilung decken und ändern die Numerierung auf dem oberen Lineal so, daß wir 0 an Stelle von 30, 1 an Stelle von 29 usw. schreiben. Auf diese Weise erhalten wir zwei Millimeterskalen, wie sie in Abbildung 8 A dargestellt sind. Obgleich die Zahlen nur an die Striche gesetzt sind, die ganzen Zentimetern entsprechen, so können wir doch von der Zahl sprechen, die durch einen beliebigen Strich, ja durch einen beliebigen Punkt der Skala bezeichnet wird. So

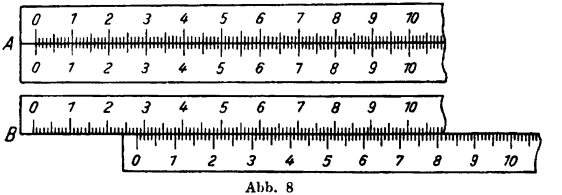


Abb. 8

bezeichnet der halblange Strich zwischen den Strichen mit den Zahlen 2 und 3 die Zahl 2,5, der nächste Strich rechts die Zahl 2,6, der folgende die Zahl 2,7 usw. Der Punkt in der Mitte zwischen den Strichen mit den Zahlen 2,6 und 2,7 bezeichnet dann die Zahl 2,65 u. dgl. m. Auf diese Weise bezeichnet jeder Punkt der Skala eine gewisse reelle Zahl, nämlich die Zahl, die (in Zentimetern) den Abstand des Punktes vom Anfangspunkt der Skala angibt. Im folgenden werden wir statt von „dem Punkt für die Zahl a" auch kurz von „der Zahl a" sprechen.

Verschieben wir nun eine der beiden Skalen (etwa die untere) so, daß ihr Anfang z. B. der Zahl 2,8 der anderen Skala gegenübersteht (siehe Abb. 8B), so steht jeder Zahl a der unteren Skala die Zahl $a + 2{,}8$ der oberen Skala gegenüber, d. h., wir haben (in den Grenzen der Skalen) die Addition von 2,8 vollzogen; wenn wir also z. B. auf der unteren Skala die Zahl 4,6 aufsuchen, so lesen wir auf der oberen Skala 7,4, also die Summe $4{,}6 + 2{,}8$ ab; nehmen wir auf der unteren Skala die Zahl 5,9, so finden wir auf der oberen Skala die Zahl $8{,}7 = 5{,}9 + 2{,}8$ usw.

Weiter können wir auch subtrahieren. Wenn wir z. B. von der Zahl 8,2 die Zahl 3,7 abziehen wollen, so brauchen wir nur die Zahl 8,2 der oberen Skala der Zahl 3,7 der unteren Skala gegenüberzustellen. Dann erscheint der Anfang der unteren Skala gegenüber der Zahl 4,5 der oberen Skala, und diese ist gerade die Differenz $8{,}2 - 3{,}7$.

Wir haben auf diese Weise zwei Regeln gewonnen: 1). Um die Summe $c = a + b$ zu finden, muß man die Zahl a der einen Skala dem Anfang der anderen Skala gegenüberstellen und kann dann gegenüber der Zahl b der zweiten Skala auf der ersten Skala die Zahl c ablesen. 2). Um die Differenz $d = a - b$ zu finden, muß man die Zahl a der ersten Skala der Zahl b der zweiten Skala gegenüberstellen und kann dann gegenüber dem Anfang der zweiten Skala auf der ersten Skala die Zahl d ablesen.

Da die Addition und die Subtraktion für sich allein so einfach sind, kommt dem eben geschilderten Instrument, das wir als „metrischen Rechenschieber" ansprechen könnten, kaum irgendeine Bedeutung zu. Jedoch führt uns sein Prinzip, die Gegenüberstellung zweier Skalen, zu einer Reihe von anderen Formen des Rechenschiebers. Unter diesen spielt eine hervorragende Rolle ein Instrument, das den Namen „logarithmischer Rechenschieber" trägt und das sich in der Praxis als außerordentlich wertvoll erwiesen hat, so daß es im Laufe der letzten Jahrzehnte weite Verbreitung überall dort gefunden hat, wo numerische Berechnungen irgendwelcher Art auszuführen sind.

Beim metrischen Rechenschieber haben wir zwei metrische Skalen gegenübergestellt, bei denen die Entfernung jedes Punktes der Skala vom Anfang der Skala proportional der durch diesen Punkt bezeichneten Zahl ist. Wenn wir nun unter Beibehaltung der zahlenmäßigen Wertung der Striche die Striche so auf der Skala verschieben, daß die Entfernung der Striche vom Anfang der Skala proportional den Logarithmen (bezüglich einer gegebenen Basis) der durch die Striche bezeichneten Zahlen wird, so erhalten wir eine sogenannte *logarithmische Skala*. Falls wir mit $\bar{a}$ den Abstand (in Millimetern) des Striches mit der Zahl a vom Anfang der Skala bezeichnen, so erhalten wir als „Gleichung der logarithmischen Skala": $\bar{a} = m \cdot \lg a$, wobei m ein Proportionalitätsfaktor ist, den wir den „Modul" der vorliegenden Skala nennen wollen. Für beliebiges m finden wir offenbar die Zahl 1 am Anfang der Skala, da $m \lg 1 = 0$ ist, die Zahl 10 im Abstand m mm vom Anfang der Skala, die Zahl 100 im Abstand von $2m$ mm vom Anfang der Skala usw. In Abb. 9 A ist eine logarithmische Skala mit dem Modul $m = 100$ und Strichen für die Zahlen von 1 bis 10 dargestellt. Für $m = 100$ befindet sich z. B. der Strich mit der Zahl 1,5 in einer Entfernung von $100 \cdot \lg 1{,}5 = 17{,}6$ mm, der Strich mit der Zahl 2 in einer Entfernung von $100 \cdot \lg 2 = 30{,}1$ mm usw. Wir haben in der Zeichnung die Striche nach oben und nach unten fortgesetzt, so daß man, wenn man diese doppelte Skala entlang ihrer Achse aufschneidet, zwei spiegelbildliche logarithmische Skalen gleichen

Moduls erhält. Wenn wir nun z. B. die untere Skala so verschieben, daß ihr Anfang, also ihre Zahl 1, unter der Zahl 2 der anderen Skala zu liegen kommt (Abb. 9 B), so stellen wir fest, daß über jeder Zahl a der unteren Skala auf der oberen Skala

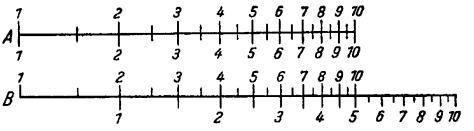


Abb. 9

die Zahl $b = 2a$ erscheint, d. h., wir haben (in den Grenzen der Skala) die Multiplikation mit 2 vollzogen.

Man kann die genannte Tatsache leicht begründen: Wenn sich z. B. die Zahl a der oberen Skala und der Anfang der unteren Skala gegenüberstehen, so erscheint über gewissen Zahlen b der unteren Skala eine gewisse Zahl c der oberen Skala (Abb. 10). Bezeichnen wir nun entsprechend wie oben mit $\bar{a}$, $\bar{b}$, $\bar{c}$ den Abstand der Zahlen a, b, c vom Anfangspunkt der jeweiligen Skala, so erhalten wir: $\bar{a} + \bar{b} = \bar{c}$. Dann sind aber die Zahlen selbst durch die Gleichung $m \lg a + m \lg b = m \lg c$ oder $\lg a + \lg b = \lg c$ oder $ab = c$ verknüpft. Wenn umgekehrt gegenüber der Zahl a der oberen Skala die Zahl b der unteren Skala liegt und dem Anfang der unteren

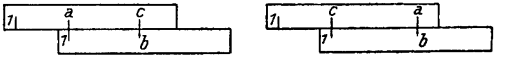


Abb. 10

Skala die Zahl c der oberen Skala gegenüberliegt (Abb. 10), so besteht zwischen den Abständen die Beziehung $\bar{a} - \bar{b} = \bar{c}$ und mithin zwischen den Zahlen die Beziehung

$$m \lg a - m \lg b = m \lg c \text{ oder } \frac{a}{b} = c.$$

Wir kommen also auf diese Weise zu den Regeln über die Ausführung der Multiplikation und Division auf einem Paar logarithmischer Skalen, die vollständig den obigen Regeln über die Addition und Subtraktion auf einem Paar metrischer Skalen entsprechen: 1.) Um das Produkt $ab = c$ zu finden, muß man die Zahl a der einen Skala dem Anfang der anderen Skala gegenüberstellen und kann dann gegenüber der Zahl b der zweiten Skala die Zahl c ablesen. 2). Um den Quotienten $\frac{a}{b} = c$ zu finden, muß man die Zahl a der ersten Skala der Zahl b der zweiten Skala gegenüberstellen und kann dann gegenüber dem Anfang der zweiten Skala die Zahl c ablesen.

Wir sehen also, daß man auf einem Paar logarithmischer Skalen die Multiplikation und die Division von Zahlen mit der gleichen Leichtigkeit ausführen kann,

mit der man auf einem Paar metrischer Skalen die Addition und die Subtraktion ausführt. Das so erhaltene Recheninstrument ist im wesentlichen der logarithmische Rechenschieber.

Nachdem man das Prinzip des logarithmischen Rechenschiebers verstanden hat, bereitet es keinerlei Schwierigkeiten, sich in alle Einzelheiten des Aufbaus und des Gebrauches eines Rechenschiebers fabrikmäßiger Herstellung hineinzufinden. Er unterscheidet sich von dem eben beschriebenen selbstverfertigten darin, daß seine Skalen sehr sorgfältig ausgeführt sind und eine große Anzahl von Strichen enthalten, und zweitens darin, daß er außer dem beschriebenen Paar logarithmischer Skalen noch eine Reihe von weiteren Skalen besitzt. Mit ihrer Hilfe kann man außer der Multiplikation und Division noch eine große Anzahl von weiteren mathematischen Operationen ausführen: das Potenzieren, das Radizieren, die Lösung von Dreiecksaufgaben, das Bestimmen von Logarithmen und Numeri usw.

Es werden Rechenschieber verschiedener Größe (Taschenformat 125 mm und 250 mm, Tischformat 500 mm, Demonstrationsmodelle bis zu 2 m) und Ausführungen hergestellt. Rechenschieber dienen einerseits allgemeinen Zwecken, sind für die Ausführung der Multiplikation, Division usw. bestimmt, andererseits aber auch speziellen Zwecken, sind also z. B. besonders für Elektrotechniker, Wärmetechniker usw. gedacht. Größte Verbreitung besitzen gegenwärtig die normalen Rechenschieber mit einer Länge von 250 mm und 6 oder 7 Skalen. Mit ihrer Hilfe kann man die Resultate gewöhnlich mit drei, zum Teil mit vier wesentlichen Ziffern bestimmen. Für die meisten technischen Berechnungen reicht diese Genauigkeit durchaus hin. Bei geschickter Handhabung ermöglicht der Rechenschieber bei der Durchführung der Rechnungen einen erheblichen Zeitgewinn, liefert zuverlässige Resultate und erspart einen wesentlichen Kraftaufwand. Daher ist es ganz natürlich, daß der Rechenschieber eine so große Verbreitung gefunden hat: weder ein Ingenieur, noch ein Techniker, noch ein Student der technischen Hochschule kann mehr ohne dieses Rechenhilfsmittel auskommen. Es stimmt allerdings, daß die Arbeit mit dem Rechenschieber hohe Anforderungen an die Augen stellt; Personen mit schwacher Sehkraft kann er deshalb nicht empfohlen werden (eine einfache Kurzsichtigkeit dagegen behindert die Arbeit mit dem Schieber nicht). Aus diesem Grunde ist auch die Benutzung des Rechenschiebers bei schlechter Beleuchtung unzulässig.

Das Studium des Rechenschiebers vollzieht sich in zwei Schritten: Im ersten Schritt erlernt man die Technik der Durchführung der Multiplikation, der Division, des Potenzierens mit den Exponenten 2 und 3, der Berechnung von Quadrat- und Kubikwurzeln, der Lösung von Verhältnisaufgaben und der Berechnung von Werten, die gegebenen Werten direkt oder umgekehrt proportional sind. Im zweiten Schritt schließen sich eine Reihe von komplizierteren Operationen an, insbesondere trigonometrische Berechnungen. Für die vollständige Beherrschung der im ersten Schritt zu erlernenden Operationen benötigt man etwa 15 bis 20 Arbeitsstunden und für die Beherrschung der im zweiten Schritt zu erlernenden Operationen etwa die gleiche Zeit. Dieser beträchtliche Zeitaufwand wird jedoch bald durch die Ökonomie kompensiert, die die Verwendung des Rechenschiebers jedem Rechner einbringt.

Hierzu noch ein Zitat aus dem Buch „Meine Erinnerungen" von A. N. Krylow (Verl. d. Akad. d. Wiss. d. UdSSR, 1945, S. 116):

„Der von mir im Jahre 1892 gehaltenen Vorlesung über die Theorie des Schiffes schickte ich einen Überblick über die Grundprinzipien der Näherungsrechnung im

allgemeinen und ihrer Anwendungen auf das Schiff im besonderen voraus. Dabei erhob ich zum obersten Prinzip, daß alle Berechnungen mit der Genauigkeit zu führen sind, die für die Praxis notwendig ist: jede unrichtige Ziffer stellt einen Fehler dar, jede überflüssige Ziffer einen halben Fehler. In welchem Maße die damalige Rechenpraxis unvollkommen war, zeigte ich an einer Reihe von Beispielen, in denen 90% der angegebenen Ziffern überflüssig waren. Bei einer Berechnung, die im Zeichensaal des Marinetechnischen Komitees durchgeführt wurde, erreichte diese unnütze Arbeit sogar 97%. In meiner langjährigen Praxis gelangte ich zu der Überzeugung, daß ein zur Routine gewordener Unsinn sich um so schwieriger beseitigen läßt, je absurder er ist."

Diese überflüssigen Ziffern, die nach dem Ausdruck von A. N. KRYLOW einen halben Fehler darstellen, kann man radikal beseitigen, wenn man laufend die Regeln der Ziffernzählung (§§ 10 bis 13) benutzt. Ein überaus wirksames Kampfmittel gegen die überflüssigen Ziffern stellt außerdem die breite Verwendung des Rechenschiebers dar.

§ 18. Über den Rechenunterricht in den verschiedenen Schuljahren

In einer erläuternden Bemerkung zu den heute gültigen Unterrichtsplänen für Mathematik in den Oberschulen der RSFSR können wir folgendes lesen: Die Verbindung von Theorie und Praxis ist im Verlaufe des Mathematikunterrichts durch geeignete Aufgaben herzustellen. Diese müssen eine gewisse Vorbereitung für die Lösung wirklicher praktischer Fragen darstellen. Zum anderen muß sie verwirklicht werden durch die Ausübung eigener praktischer Arbeiten der Schüler, in denen die mathematischen Kenntnisse verwertet werden. Diese Arbeiten und Übungen müssen im organischen Zusammenhang mit dem jeweiligen Stoff stehen und dürfen nicht der Arbeitsweise der mathematischen Wissenschaften entgegengesetzt sein (Ausgabe 1950, S. 5).

Ein Lehrer, der bestrebt ist, diese wichtige Forderung ernsthaft zu erfüllen und der die Lösung von Aufgaben lehren will, die wirklich in der Landwirtschaft und den verschiedenen Zweigen der Technik vorkommen, also Aufgaben, wie sie in der erläuternden Bemerkung gemeint sind, stößt auf zweierlei Schwierigkeiten. Die erste besteht darin, daß derartige Aufgaben im allgemeinen ziemlich umfangreiche Zahlenrechnungen erfordern, die viel Zeit rauben und keinerlei Bildungswert besitzen, da die Fertigkeit in der rationellen Ausführung solcher Rechnungen schnell erworben ist. Die zweite Schwierigkeit besteht darin, daß alle bei der Lösung von „realen" Aufgaben auftretenden Zahlen zwangsläufig Näherungszahlen sind. Genaue Ausgangszahlen treten nur ausnahmsweise auf; angenäherte Ausgangszahlen dagegen führen wieder nur auf angenäherte Resultate. Unausbleiblich erhebt sich daher die Frage: Welche Ziffern des Resultats verdienen Vertrauen? Ein Runden ohne eine feste Stütze in Form einer Methode der genauen Fehlerrechnung oder auch nur in Form der Regeln der Ziffernzählung birgt die Gefahr in sich, daß man entweder Zahlen mit falschen Ziffern angibt („Täuschungen" durch Vorspiegelung einer „Genauigkeitsillusion") oder daß man richtige Ziffern fortläßt, d. h. ungerechtfertigt die Genauigkeit der Resultate vermindert. Auf Grund einer

jahrhundertealten Tradition ignoriert die Schulmathematik den angenäherten Charakter der meisten Zahlenergebnisse für reale Größen. Hierin liegt eine der Ursachen dafür, daß die Schulmathematik so lebensfremd ist.

Wie kann man nun diese beiden Schwierigkeiten beseitigen? Notwendig ist zunächst eine Rationalisierung der Rechenarbeit der Schüler. Sie besteht erstens in einer ökonomischen Verwendung der verschiedensten Hilfsmittel und Verfahren, die das Operieren an mehrstelligen Zahlen erleichtern und beschleunigen (Kopfrechnen, Verbesserung der Technik des schriftlichen Rechnens, weitgehende Verwendung von Recheninstrumenten, Tafeln und graphischen Darstellungen usw.) und zweitens in der Aneignung der Elemente der Theorie der Näherungsrechnung. Um die herrschende Lage von Grund auf zu verbessern, muß man einen gewissen Umbau der Lehrpläne vornehmen. Jedoch ist es auch im Rahmen der geltenden Pläne durchaus möglich, den Schülern viele wertvolle, beständige und bewußte Fähigkeiten zu vermitteln.

Dazu ist es notwendig, daß man sich besonders dem Kopfrechnen und der Technik des schriftlichen Rechnens eingehend in der fünften Klasse widmet und daß man für die Erhaltung und Erweiterung der erworbenen Fähigkeiten auch in allen folgenden Jahren Sorge trägt. Das Rechnen auf dem Rechenbrett wird im Programm bereits für die unteren Klassen empfohlen, und es ist sehr bedauerlich, daß die Mehrzahl der Lehrer es vernachlässigt. Das Rechnen auf dem Rechenbrett kann man als einen ersten Schritt der Mechanisierung der Rechenarbeit ansehen; es erleichtert die Addition und Subtraktion mehrstelliger Zahlen ganz erheblich. Einen zweiten Schritt stellt die Einführung der NEPERschen Streifen dar; sie erleichtern im selben Maße die Multiplikation und Division mehrstelliger Zahlen. Im derzeit gültigen Unterrichtsplan der Oberschulen werden die NEPERschen Streifen nicht erwähnt. Man kann daher ihre Herstellung und Verwendung nur als zusätzliche und nicht obligatorische Betätigung empfehlen (besonders als Aufgabe für bessere Schüler; Zirkelarbeit). Den folgenden Schritt, den man als Übergang von der kleinen Mechanisierung zur großen Mechanisierung der Rechenarbeit ansehen kann, stellt die Beherrschung der Arbeit an der eigentlichen Rechenmaschine dar. Dies ist ein schönes Thema für eine Zirkelarbeit, die allerdings davon abhängt, daß es dem Lehrer gelingt, sich eine Rechenmaschine zu verschaffen, und er selbst damit arbeiten lernt.

Noch wichtiger ist die Einführung des logarithmischen Rechenschiebers in den Schulunterricht. Sie ist ganz natürlich verbunden mit dem Unterricht über Logarithmen im neunten Schuljahr. Da die Jugend an einer Vielzahl von Beispielen die Verwendung des Rechenschiebers bei Arbeiten in vielen Fachgebieten sieht, wird sie gern bereit sein, sich einen Rechenschieber zu verschaffen und an einem Zirkel teilzunehmen, um ihn kennen zu lernen. Beim späteren Übergang an die technischen Hochschulen werden diejenigen, die in der Schule den Rechenschieber kennengelernt haben, einen erheblichen Vorteil gegenüber den anderen Studenten besitzen. Aber auch in der Schule erspart der Rechenschieber bei der Durchführung von Berechnungen, die in der Mathematik und in der Physik auftreten, viel Zeit. Was die organisatorische Verbesserung des Mathematikunterrichts anbetrifft, so ist keine Angelegenheit so entscheidend wichtig und doch so leicht lösbar wie die Forderung nach der dauernden Einführung des Rechenschiebers in den Schulgebrauch.

In viel zu geringem Maße werden in der Schule auch die als Hilfsmittel für das Rechnen doch so gut geeigneten mathematischen Tafeln benutzt. Eine Tafel kann jedesmal dann benutzt werden, wenn die grundlegenden Verfahren für die Durchführung der jeweiligen Grundrechenart früher erlernt sind und in irgendeinem späteren Zusammenhang Aufgaben auftreten, die die mehrmalige Anwendung dieser Grundrechenart erfordern. Die ersten Tafeln, die die Schüler kennenlernen können, sind die Produkttafeln. Schon eine kleine, selbst verfertigte Produkttafel (eine gegebene Zahl wird mit allen einstelligen Zahlen multipliziert) stellt eine wirkliche Hilfe für die Multiplikation und Division dar. Weiterhin sind die Tafeln der Quadrate und der Quadratwurzeln zu empfehlen. Sie können in der sechsten und siebenten Klasse verwendet werden. Die Tafeln der Kuben, mittels derer sich auch Kubikwurzeln bestimmen lassen, können als wesentliches Hilfsmittel für die Lösung vieler guter geometrischer Aufgaben mit durchaus realem Gehalt benutzt werden. Breite Möglichkeiten eröffnet auch die konsequente Verwendung von Tafeln für den Kreisumfang, für die Kreisfläche, für das Bogenmaß, für die reziproken Werte der Zahlen usw. Ein erfolgreiches Studium der Propädeutik der Trigonometrie in der achten Klasse ist undenkbar ohne den Gebrauch der üblichen Sinus- und Tangenstafeln. Den Tafeln für die Logarithmen der Zahlen und der trigonometrischen Funktionen werden in der Schule augenblicklich genügend Beachtung geschenkt.

Äußerst wichtig ist es jedoch, daß von vornherein die Tafeln nicht mechanisch, sondern vollkommen bewußt angewendet werden. Dazu muß man bei den Schülern vor allem eine Vorstellung davon wecken, wie eine Tafel zusammengestellt wird; hierfür ist es gut, wenn man einmal probeweise einige Tafelwerte nachrechnet (bei rationeller Arbeitsverteilung nimmt diese Rechnung den Charakter einer Kontrolle der fertigen, gedruckten Tafel oder eines gewissen Teiles derselben an). Hierbei gelangt man zu einer vollständig bewußten Anwendung der Interpolation und wird erst danach die Hilfsmittel der linearen Interpolation (proportionale Teilung, fertige Korrekturtafeln) vollständig verstehen.

In diesem Zusammenhang erhebt sich die Frage, ob die Einführung in die Tafeln nicht zu viel Zeit erfordert, an der es dem Mathematiklehrer häufig fehlt? Hierzu kann man nur sagen, daß es bei der Einführung der Tafeln ähnlich steht wie in allen anderen Fällen der Rationalisierung der Arbeit: Die Zeit, die man für die Einführung einer neuen Tafel opfern muß, macht sich dank der Ökonomie, die durch diese Tafel erzielt wird, später durch einen entsprechenden Gewinn an Zeit und eine erhebliche Kraftersparnis bezahlt. Die Berechnung einer Quadratwurzel auf vier wesentliche Ziffern, für die man bei Anwendung der üblichen schriftlichen Rechenverfahren zwei bis drei Minuten benötigt, kann mittels einer vierstelligen Tafel der Quadrate oder der Quadratwurzeln in 10 bis 15 Sekunden geleistet werden, und auch Fehler sind dabei in viel geringerem Maße zu befürchten als bei der üblichen Berechnung. Wieviele derartiger Berechnungen können aber durch die konsequente Benutzung von Tafeln erleichtert werden!

Häufig wird empfohlen, in der Schule auch die „abgekürzten Verfahren" für die Ausführung der Grundrechenarten zu lehren (siehe oben S. 377). Hier kann man jedoch an der Zweckmäßigkeit zweifeln. Der Grund dafür ist, daß diese Rechenverfahren erst voll zum Tragen kommen, wenn man es mit Zahlen einer großen Ziffernanzahl zu tun hat; diese kommen jedoch bei praktischen Rechnungen ziemlich selten vor. Bei Rechnungen mit 2-, 3-, 4-stelligen Zahlen halten die Verfahren der

abgekürzten Durchführung der Grundrechenarten einem Vergleich mit den verschiedenen Rechenhilfsmitteln (Tafeln, Recheninstrumenten, graphischen Darstellungen usw.) nicht stand. Diese Einstellung zu den abgekürzten Verfahren bezieht sich jedoch nicht auf die Näherungsformeln (§ 14) und auf die speziellen Verfahren für die Ausführung der arithmetischen Grundoperationen. Von den letzteren haben wir in § 2 gesprochen.

Von der Unterschätzung der graphischen Methoden in der Schule war schon oben die Rede (§ 16). Graphische Berechnungen können buchstäblich in allen Klassen durchgeführt werden, und es ist durchaus nützlich, daß man sich bei jeder Aufgabe die Frage nach der Zweckmäßigkeit der Verwendung eines graphischen Verfahrens vorlegt. Ein Beispiel einer Aufgabe, für deren Lösung man vorteilhaft eine graphische Darstellung auf einem Stück Millimeterpapier benutzt, ist folgende: Wieviel Prozent ihrer Gesamtsumme stellen die und die vorgegebenen Zahlen dar? Auch wenn man die Sätze über die Ähnlichkeit von geometrischen Figuren noch nicht kennt, begreift man leicht die Möglichkeit einer proportionalen Veränderung einer Strecke durch Parallelverschiebung einer Seite eines geeigneten Dreiecks.

Bei der Einführung in die Elemente der Näherungsrechnung kann man z. B. folgendermaßen vorgehen: In der fünften Klasse taucht bei der Wiederholung der Maße unvermeidlich die Frage nach dem Messen selbst auf. Hier kann man ganz natürlich den Begriff des Näherungswertes und des Meßresultates einführen und ebenfalls den Schülern brauchbare Methoden für die Behandlung von Meßresultaten entwickeln (vgl. § 7). Weiterhin treten dann ganz natürlich die Regeln der Ziffernzählung (§ 10) in Erscheinung. Ihre Zweckmäßigkeit kann man an einfachen, dem Verständnis der Schüler der fünften Klasse entsprechenden Beispielen erläutern, wobei man die unbekannten Ziffern jeweils durch Fragezeichen ersetzt. Eine wesentliche Bedeutung hierfür besitzen die Einsicht des Lehrers in den Wert dieser Regeln und die Forderung nach ihrer beständigen Anwendung. Die Regeln der Ziffernzählung, die allein dem Verlangen entspringen, nur vertrauenswürdige Ziffern hinzuschreiben (Prinzip von A. N. Krylow), bilden den ersten wesentlichen Kreis von Kenntnissen in der Näherungsrechnung.

Einen zweiten Kreis bildet das einfachste Verfahren der genauen Fehlerrechnung, die Methode der Wertschranken (§ 8). Prinzipiell läßt sich auch diese Methode vollständig in der fünften Klasse bewältigen, doch ist es wohl besser, wenn man sie auf die sechste Klasse verschiebt und später im Zusammenhang mit dem Studium der Ungleichungen auf sie zurückkommt.

Der dritte Kreis von Kenntnissen aus der Näherungsrechnung betrifft das Verfahren der Schranken des absoluten und des relativen Fehlers. Es ist wünschenswert, daß man dieses Verfahren in der siebenten Klasse behandelt, wobei man sich allerdings mit ungefähren Vorstellungen dieser Schranken begnügen muß. Jedoch muß ein Absolvent der Grundschule durchaus in der Lage sein, den Sinn von Ausdrücken der Form $134 (\pm 1)$ mm und $5{,}4 (\pm 2\%)$ zu verstehen, und in solcher Form seine Meßresultate (und die nach den Methoden der Schranken gewonnenen Rechenergebnisse) anzugeben. Die Frage nach den Sätzen über Fehlerschranken der Resultate der Grundrechenarten, die im bisherigen Programm nicht vorgesehen sind, wird bereits dringlich angesichts der Anwendungen dieser Sätze im Physikunterricht der höheren Klassen der Oberschule. Bei der Ableitung dieser Sätze kann man nochmals auf die Regeln der Ziffernzählung zurückkommen und eine neue Begründung für sie geben, indem man den Zusammenhang zwischen der Anzahl der

Dezimalstellen und den Schranken des absoluten Fehlers herstellt und entsprechend auf die Zusammenhänge zwischen der Anzahl der wesentlichen Ziffern und den Schranken des relativen Fehlers hinweist. (Vgl. S. 361 und den Artikel [6]). Wenn die Verfahren der Schranken gründlich erlernt sind, lassen sich die Sätze über Schranken des absoluten und des relativen Fehlers sehr leicht beweisen. Eine gewisse Schwierigkeit tritt lediglich beim Fortlassen der Zahlen zweiter Größenordnung auf, die sich jedoch leicht beheben läßt, wenn man die Schüler vorher mit den Näherungsformeln bekannt macht.

Bei jeder Aufgabe muß man zwischen Planen und Ausführen unterscheiden. Die Ausführung jeder mathematischen Aufgabe erfordert nun mehr oder weniger erhebliche Zahlenrechnungen. Dieser angewandte Teil der Mathematik wird im heutigen Schulunterricht erst in dritter oder noch späterer Linie behandelt. Dies führt dazu, daß der ganze Komplex der mathematischen Kenntnisse und Fertigkeiten, die der Schüler in der Oberschule erhält, einen im hohen Grade formalen Charakter trägt.

Zur Beseitigung dieses Übelstandes ist es erforderlich, daß der Lehrer selbst hinreichende theoretische und praktische Rechenfertigkeiten besitzt, zumindest in dem Maße, wie sie in den vorangehenden Darlegungen angeführt wurden.

LITERATUR

[1] Берман, Г. Н., Приёмы быстрого счёта (Berman, G. N., Schnellrechenverfahren), Moskau-Leningrad 1947.

[2a] Брадис, В. М., Средства и способы элементарных вычислений (Bradis, W. M., Hilfsmittel und Verfahren für das elementare Rechnen), Moskau 1948 (3. Aufl. 1954).

[2b] Брадис, В. М., Теория и практика вычислений (Bradis, W. M., Theorie und Praxis des Rechnens), 5. Aufl., Moskau 1937.

[3] Филиппов, А. О., Четыре арифметических действия (Filippow, A. O., Die vier Grundrechenarten), 1909.

[4] Чуканцов, С. М., Больше внимания технике арифметических вычислений (Tschukanzow, S. M., Mehr Beachtung der arithmetischen Rechentechnik), Mathematik in der Schule Nr. 4 (1948).

[5] Гончаров, В. Л., Арифметические упражнения и функциональная пропедевтика (Gontscharow, W. L., Arithmetische Aufgaben und Funktionalpropädeutik), Moskau 1947.

[6] Александров, П. С., и А. Н. Колмогоров, Свойства неравенств и понятие о приближённых вычислениях (Alexandroff, P. S., und A. N. Kolmogoroff, Eigenschaften von Ungleichungen und die Behandlung von näherungsweisen Berechnungen), Mathematik in der Schule Nr. 2 (1942).

[7] Романовский, В. И., Основные задачи теории ошибок (Romanowski, W. I., Grundlegende Probleme der Fehlerrechnung), Moskau 1947.

[8] Гнеденко, Б. В., и А. Я. Хинчин, Элементарное введение в теорию вероятностей (Gnedenko, B. W., und A. J. Chintschin, Elementare Einführung in die Wahrscheinlichkeitsrechnung), 2. Aufl., Moskau 1950.
Deutsche Übersetzung: Gnedenko, B. W., und A. J. Chintschin, Elementare Einführung in die Wahrscheinlichkeitsrechnung, 7. Aufl., Berlin 1967.

[9] Кавун, И. Н., Приближённые вычисления (Kawun, I. N., Näherungsrechnung), 1923.

[10] Крылов, А. Н., Лекции о приближённых вычислениях (Krylow, A. N., Vorlesungen über Näherungsrechnung), 5. Aufl., Moskau 1950.

[11] Безикович, Я. С., Приближённые вычисления (Besikowitsch, J. S., Näherungsrechnung), 5. Aufl., Moskau 1941.

[12] Франк, М. Л., Элементарные приближённые вычисления (Frank, M. L., Elementare Näherungsrechnung), Moskau-Leningrad 1932.

[13] Глаголев, А. А., Номография для школьника (Glagoljew, A. A., Nomographie für Schüler), Moskau 1935.

[14] Глаголев, Н. А., Теоретические основы номографии (Glagoljew, N. A., Theoretische Grundlagen der Nomographie), 2. Aufl., Moskau-Leningrad 1936.

[15] Панов, Д. Ю., Счётная линейка (Panow, D. J., Der Rechenschieber), 6. Aufl., Moskau 1949.
Deutsche Übersetzung: Panow, D. J., Der Rechenstab, Leipzig 1959.

[16] Семендяев, К. А., Счётная линейка (Semendjajew, K. A., Der Rechenschieber), 2. Aufl., Moskau 1950.

Der deutsche Leser sei noch verwiesen auf:

Archangelski, N. A., und B. I. Saizew, Automatische Ziffernrechenmaschinen, Berlin 1960 (Übersetzung aus dem Russischen).

Baule, B., Die Mathematik des Naturforschers und Ingenieurs, 11. Band: Ausgleichs- und Näherungsrechnung, 8. Aufl., Leipzig 1966.

Biermann, O., Vorlesungen über mathematische Näherungsmethoden, Braunschweig 1905.

Bojko, J., Lehrbuch der Rechenvorteile, Leipzig und Berlin 1920. Aus Natur und Geisteswelt Nr. 739.

Bruns, H., Grundlinien des wissenschaftlichen Rechnens, Leipzig 1903.

Czuber, E., Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik und Lebensversicherung, I. Band, 4. Aufl., Leipzig und Berlin 1924; 6. Aufl. 1941.

DELAUNAY, B. N., Kurzer Kurs über mathematische Maschinen, Teil I (russ.), Moskau-Leningrad 1952.
FRICKE, H. W., Der Rechenschieber, Leipzig 1952.
HÄNSEL, H., Grundzüge der Fehlerrechnung, 3. Aufl., Berlin 1967.
HAPPACH, V., Ausgleichsrechnung, 2. Aufl., Leipzig 1950.
HEINRICH, H., Einführung in die praktische Analysis, Teil I. Leipzig 1963.
HELMERT, F. R., Die Ausgleichsrechnung nach der Methode der kleinsten Quadrate. 2. Aufl., Leipzig und Berlin 1924.
KÄMMERER, W., Ziffernrechenautomaten, 3. Aufl., Berlin 1963.
KITOW, A. I., und N. A. KRINIZKI, Elektronische Digitalrechner und Programmierung, Leipzig 1962 (Übersetzung aus dem Russischen).
KÖRWIEN, H., Graphisches Rechnen (Nomographie), 5. Aufl., Leipzig 1949.
LINNIK, J. W., Die Methode der kleinsten Quadrate in moderner Darstellung. Berlin 1961 (Übersetzung aus dem Russischen).
LUCKEY, P., Nomographie, 2. Aufl., Leipzig und Berlin 1927; 7. Aufl. 1953.
MAENNCHEN, PH., Geheimnisse der Rechenkünstler, 5. Aufl., Leipzig 1951.
MEHMKE, R., Leitfaden zum graphischen Rechnen, 2. Aufl., Leipzig und Wien 1924.
–, Numerisches Rechnen, Encyklopädie der mathematischen Wissenschaften, Band I, Teil 2: Arithmetik und Algebra, Leipzig 1900–1904.
MEYER ZUR CAPELLEN, W Leitfaden der Nomographie, Berlin-Göttingen-Heidelberg 1953.
MILLER, M., Rechenvorteile, 3. Aufl., Leipzig 1964.
OBÁDOVICS, J. GY., Taschenbuch der Elementarmathematik, 2. Aufl., Leipzig 1964 (Übersetzung aus dem Ungarischen).
PENTKOWSKI, M. W., Nomographie, Berlin 1953 (Übersetzung aus dem Russischen).
POLOSHI, G. N. (Hrsg.), Mathematisches Praktikum, Leipzig 1963 (Übersetzung aus dem Russischen).
ROHRBERG, A., Theorie und Praxis des logarithmischen Rechenschiebers, 10. Aufl., Leipzig 1952.
RUNGE, C., und H. KÖNIG, Vorlesungen über numerisches Rechnen, Berlin 1924.
VON SANDEN, H., Praktische Analysis, 2. Aufl., Leipzig und Berlin 1923.
–, Mathematisches Praktikum I, Leipzig und Berlin 1927.
–, Praktische Mathematik, 3. Aufl., Leipzig 1953.
SCHRÖDER, K. (Hrsg.), Mathematik für die Praxis I, II, III, 3. Aufl., Berlin 1966
SCHRUTKA, L., Zahlenrechnen, Leipzig und Berlin 1923.
–, Theorie und Praxis des logarithmischen Rechenschiebers, 2. Aufl., Leipzig 1929; 3. Aufl., Wien 1943.
SCHULZ, G., Formelsammlung zur praktischen Mathematik, Berlin und Leipzig 1937.
SCHWERDT, H., Lehrbuch der Nomographie, Berlin 1924.
–, Die Anwendung der Nomographie in der Mathematik, Berlin 1931.
TUKATSCHINSKI, M. S., Maschinen als Mathematiker, Berlin 1960 (Übersetzung aus dem Russischen).
WERKMEISTER, P., Praktisches Zahlenrechnen, 3. Aufl., Berlin 1945.
WILLERS, FR. A., Methoden der praktischen Analysis, 3. Aufl., Berlin 1957.

NAMENVERZEICHNIS

Abu Zaqarijjā Alhassar 53
Alexander de Villa Dei 41
Alexandroff, P. S. 65, 341
Alhwārāzmi, Muhammad Ibn Mūsā 41
Alkāšī, Gījat Eddīn 56
Alnasawī 53, 56
Appolonius 2, 26
Archimedes 3, 16, 17, 23, 26, 346, 347
Argand, J.-R. 198
Aristoteles 4, 14
Āryabhata 34, 134

Beda Venerabilis 13
Beyer, J. H. 57
Bézout, ét. 304
Bhāskara 134
Bobynin, W. W. 22, 27, 40, 43, 50, 53
Boetius 40
Bombelli, R. 198
Borel, E. 294
Boyer, Ch. 22, 23
Bradis, W. 381
Brahmagupta 134
Bubnow, N. M. 22, 35, 40, 50
Bürgi, J. 57

Cajori, F. 30, 32, 36
Cantor, G. 17, 164, 179, 307
Cantor, M. 23, 25, 29
Cardano, G. 56, 198
Césaro, E. 215
Chintschin, A. J. 164, 237
Chiss, T. 23
Chuquet, N. 28
Clavius, Chr. 57

Dedekind, R. 124, 164, 197
Delaunay, B. N. 227
Diophant 14, 25, 52
Dirichlet, P. G. L. 239, 297, 298, 300, 301
Dzierzynski, F. E. 326

Engels, F. 6, 7, 8, 63, 102
Eratosthenes 241
Euklid 163, 230, 233
Euler, L. 227, 234, 248, 249

Faddejew, D. K. 204
Fermat, P. 249
Fibonacci 42
Finaeus, Orontius 56
Frobenius, G. 218

Galois, E. 86
Gandz 56
Gauss, C. F. 198, 229, 238, 241, 339
Gelfond, A. O. 311
Gerbert 5, 41
Gödel, K. 131
Goldbach, Chr. 240
Gontscharow, W. L. 377
Grammateus 56

Hadamard, J. 239
Hamilton, W. R. 211
Hammurabi 32
Hankel, G. 23
Hermite, Ch. 310, 311
Herodian 19
Herodot 10
Heron von Alexandrien 52
Hilbert, D. 311
Hurwitz, W. A. 294, 311
Huygens, Chr. 272

Janowskaja, S. A. 52
Janschul, I. N. 335
Johannes de Sacrobosco (Johann Halifax) 41, 43
Johannes von Sevilla 41, 56

Karpinsky, L. C. 35, 42
Kepler, J. 57
Kewitsch, G. 29
Kiseljow, A. P. 337
Knopp, K. 237
Kolmogoroff, A. N. 341
Kronecker, L. 6, 301
Krylow, A. N. 363, 367, 389
Kudrjawzew, L. D. 8
Kusmin, R. O. 204
Kyrillus 26

Landau, E. 164
Laplace, P. S. de 3
Legendre, A.-M. 229, 238, 291, 292
Leibniz, G. W. 8
Leonardo von Pisa 42, 53
Light, H. 57
Lindemann, F. 311
Liouville, J. 304, 305, 307
Ljusternik, L. A. 333
Lobatschewski, N. I. 70
Löffler, E. 29, 30
Lomonossow, M. W. 45
Lusin, N. N. 65

Magnizki, L. 5, 43—45, 54, 58
v. Mangoldt, H. 237
Marx, K. 7
Maximus Planudes 26
Miklucho-Maklai, N. N. 12, 37
Minkowski, H. 241, 301, 302

Neper (Napier), J. 57, 326
Nesselmann, G. H. F. 55
Neugebauer, O. 30—32, 52
Norton, R. 57

Ocreat, N. 42
Odhner, W. T. 327, 332
Odo von Cluny 48
Okunjew, L. J. 83

Pascal, B. 4
Peano, G. 112, 114
Perron, O. 164
Péter, R. 124
Proschko, W. M. 336
Proskurjakow, I. W. 197
Ptolemaeus, C. 23, 55

Regiomontanus 56
Riemann, B. 239
Rudio, F. 347

Schnirelmann, L. G. 240
Semendjajew, K. A. 336

Smith, D. E. 35, 42
Spasski, I. G. 45
Stevin, Simon 56, 57
Subbotin, M. F. 373

Tannery, P. 23
Taylor 7
Thureau-Dangin, F. 32
Timtschenko, I. J. 30, 32, 36
Tolstow, S. P. 40
Tschebyschew, P. L. 227, 238—240, 242, 301, 332
Tscherepnin, L. W. 49

Vahlen, Th. 294
Vieta, F. 134
Vogel, K. 52
van der Waerden, B. L. 83
Wallis, J. 19
Weierstrass, K. 164, 197
Wessel, C. 198
Wesselowski, I. N. 32, 49, 51
Widmann J, 134
Willers, Fr. A. 335
Wilson, J. 257
Winogradow, I. M. 239, 240
Woepcke, Fr. 29, 40
Wygodski, M. J. 25, 26, 31, 32

Zassenhaus, H. 239
Zeuthen, H. G. 37

SACHVERZEICHNIS

Abacisten 41—43
Abakus 35—38, 41, 42, 53
Abbildung 70, 71
—, Bild eines Elementes bei einer 70
—, eindeutige 70
—, eineindeutige 70, 71, 85
—, entgegengesetzte 71
—, identische 71
—, inverse 71
—, umkehrbar eindeutige 70, 71, 85
—, Urbild eines Elementes bei einer 70
—, volles Urbild eines Elementes bei einer 70
abelsche Gruppe 85, 87
abgekürzte Verfahren für die Grundrechenarten 377 bis 380, 391
abgeschlossenes Intervall 67, 71, 72
Abiponer 6, 9
abrunden 338
Abschnitt der Folge der natürlichen Zahlen 73—75, 82
absoluter Betrag 105, 108 bis 111, 146, 158, 186, 205, 206, 209, 210
— Fehler 339, 341
absolut prim 228
Abweichung vom Mittel 343
abzählbar 76—78
Abzählbarkeit der Menge der algebraischen Zahlen 308
Abzählbarkeit der Menge der rationalen Zahlen 72, 73
Addition auf der Sprossenradmaschine 329
— von ganzen Zahlen 140
— — komplexen Zahlen 202
— — natürlichen Zahlen 114—117
— — rationalen Zahlen 151
— — reellen Zahlen 181
additive Gruppe 95
additives Bezeichnungsprinzip 6, 16, 17, 19, 22, 28, 36, 55, 59
Ägypter 8, 17—19, 22, 36, 46, 47, 50—53
Ähnlichkeit von geordneten Mengen 81, 82
Akkader 31, 32, 47
Algebra, hyperkomplexe 212, 215—221
algebraische Kongruenz 251 bis 259
— Operation 64, 83, 104
— Zahl 303—306
Algorithmiker 41, 43
algorithmischer Bruch 47, 54
algorithmische Zahl 5, 6, 14, 15, 17, 28, 33, 34, 46
Algorithmus, das Wort 41, 43
—, Euklidischer 259—264, 266
alphabetisches Bezeichnungssystem 21—28, 43, 47, 52, 59
angeordneter Körper 105 bis 111
— Ring 105—111
Anordnung der ganzen Zahlen 143
— — rationalen Zahlen 156
— — reellen Zahlen 181
Anzahl der Elemente einer endlichen Menge 74, 75, 82
Apices 41
Approximation, diophantische 297—302
— einer Zahl durch Näherungsbrüche 287—291
Äquivalenzklasse 136, 137
Äquivalenzrelationen 136, 137
Äquivalenz von Mengen 71 bis 74, 81
Aramäer 34
Arbeit 7
archimedisch angeordneter Körper 111
— — Ring 111
archimedische Anordnung 111, 122, 148, 160, 174 bis 177, 179, 184
Archimedisches Axiom 111, 122, 164, 184, 196
Arcus einer komplexen Zahl 205, 206
Argument einer komplexen Zahl 205, 206
arithmetische Progression 239

Aruntier 9, 10
assoziatives Gesetz 84, 86, 87, 90, 91, 116, 120, 195
aufrunden 338
Auswertung von Meßresultaten 342—346
Axiom 101
—, Archimedisches 111, 122, 164, 184, 196
— der vollständigen Induktion 112
— — Vollständigkeit 172, 196
axiomatisch 194
axiomatischer Aufbau der Mathematik 101, 102
Axiomensystem für die ganzen Zahlen 197, 198
— — — natürlichen Zahlen 112—114
— — — rationalen Zahlen 197
— — — reellen Zahlen 195 bis 198
Azteken 15, 20

Babylonier 28—33, 47, 54, 55
Basis eines Vektorraumes 212
benachbarte Elemente in einer geordneten Menge 80
— natürliche Zahlen 122
beschränkt 146, 167, 197
—, nach oben 146, 167
—, nach unten 146, 167
beste Näherung erster Art 290, 291
— — zweiter Art 290, 291
Betrag einer komplexen Zahl 205, 206, 209, 210
Beweis durch vollständige Induktion 113, 122, 123, 147
Bezeichnungsprinzip, additives 6, 16, 17, 19, 22, 28, 36, 55, 59
—, divisorisches 21
—, multiplikatives 16, 17, 21, 25, 34, 36—38, 60
—, substraktives 6, 16, 17, 19, 21, 59
Bezeichnungssystem, alphabetisches 21—28, 43, 47, 52, 59
Bezeichnungssystem, nichtpositionelles 5, 6, 17—28, 30—32, 34
—, positionelles 3—5, 21, 22, 25, 26, 28—45, 59
Bild eines Elementes bei einer Abbildung 70
Bogenmaß eines Winkels 188
Bruch, algorithmischer 47, 54
—, Dezimal- 56—58
—, dyadischer 50, 51
—, geschichtliche Entstehung 46—58
—, Individual- 47, 49
—, Sexagesimal- 26, 42, 53 bis 58
—, Zwölfer- 42, 53
Bruchrechnung der Ägypter 8, 51, 52
—, alt-russische 8, 51
Bruchstrich 52, 53

Cauchy-Folge 170—175, 210
Charakteristik eines Körpers 97, 98, 108
Chinesen 20

Dedekindscher Schnitt 197
Definition durch vollständige Induktion 114, 115, 118, 119, 123—125
Definitionsbereich 69
Dewangari-Ziffern 40
Dezimalbruch 56—58, 266, 274, 286, 287
—, endlicher 76, 77, 272
—, unendlicher 76, 77
Dezimalsystem 5, 12, 30, 31, 34, 39—45, 55
Differenzentafel 334, 335
Differenz von Gruppenelementen 90
— — Mengen 67
— — natürlichen Zahlen 128
— — reellen Funktionen 188
— — Ringelelementen 92 bis 94
diophantische Approximation 297—302
— Gleichung 253—255, 259, 264
Dirichletsche Funktion 69
Dirichletscher Primzahlsatz 239, 240
distributives Gesetz 91—93, 119, 120, 195, 211
divergente Folge 167
Division auf der Sprossenradmaschine 330, 331
— in einer Gruppe 89
— — einem Körper 96
— mit dem Rechenschieber 387
— von natürlichen Zahlen 129
Divisionsalgebra 212, 213, 215—221
divisorisches Bezeichnungsprinzip 21
Dreiecksungleichung 108, 109
Durchschnitt 67
— eines Mengensystems 68
Dutzend-Rechnung 13
dyadischer Bruch 50, 51
dyadisches Zahlsystem 8, 19

Eindeutige Abbildung 70
eineindeutige Abbildung 70, 71, 85
Einermenge 66, 81
Eingeborene der Fidschi-Inseln 13, 16
— — Halbinsel Prince York 13
— — Inseln der Torres-Straße 8, 10, 12, 13, 16
— — Salomon-Inseln 13
— von Florida 13
— — Neu-Guinea 12, 37
Einselement einer Gruppe 87, 88
— eines Körpers 96
Einstellwerk 327
Elementbeziehung 65, 66
Elferprobe 244, 322
endliche Menge 67, 73—78, 81, 82
endlicher Dezimalbruch 76, 77, 272
— Kettenbruch 266—274, 284, 285
— systematischer Bruch 275, 279
entgegengesetzte Abbildung 71
entgegengesetztes Element 89, 92

erstes Element einer geordneten Menge 80
Erweiterung eines Körpers 99, 100
Erweiterungsprinzip 134 bis 136
Etruaker 21
Euklidischer Algorithmus 259—264, 266
— — für Polynome 262, 263
Eulersche Funktion 248 bis 251

Fehler, absoluter 339, 341
— beim Runden 338, 363, 364
—, maximaler 356—361
—, mittlerer 343
—, mittlerer quadratischer 345, 363—366
—, relativer 340, 341
—, systematischer 343
—, zufälliger 343
Fehlerrechnung 346—375, 392, 393
— nach der Methode der Fehlerschranken 350 bis 356, 393
— — — — — — Wertschranken 346—350, 353, 355, 356, 392
Fehlerverteilung 367, 368
Fermatscher Satz, großer 249
— —, kleiner 249
Fingerrechnung 4, 5, 13
Folge 166—175
—, divergente 167
—, Grenzwert einer 167 bis 175, 210
—, konvergente 167—175
—, stationäre 174, 184
Fundamentalfolge 170—175, 210
Fünfer-System 5, 6, 12, 33
Funktion 69, 70, 83
—, Definitionsbereich einer 69
—, Dirichletsche 69
—, ganze rationale 188
—, gebrochene rationale 188
—, reelle 187
—, stetige 187—191
—, Wertevorrat einer 69

ganze Zahl 64, 134—148, 197, 198
— komplexe Zahl 91
— rationale Funktion 188
Gaußscher Zahlkörper 220
gebrochene rationale Funktion 188
Geld-Gewichts-System, babylonisches 30—32
—, römisches 49
gemischtperiodischer Bruch 278, 279
Genauigkeitsverlust bei Subtraktion 372
genetisch 194
geometrische Darstellung der komplexen Zahlen 204
geordnete Menge 79—82
geordnete Mengen, Gleichheit von 80
— —, Ähnlichkeit von 81, 82
gerade Zahl 240
Gleichheit von Mengen 66
— — geordneten Mengen 80
Gleichmächtigkeit von Mengen 71—74
Glied einer Folge 166
Goldbachsches Problem 240
Grad, das Wort 55
graphisches Rechnen 382 bis 384, 390, 391
Grenzwert einer Folge 167 bis 175, 210
Griechen 16, 17, 19, 22—26, 36, 52—55
größter gemeinsamer Teiler 230, 232, 242, 259—264
— — — von Polynomen 263, 264
Grundbeziehung 101
Gruppe 64, 83—90
—, additive Schreibweise 89
—, kommutative (abelsche) 85, 87
—, multiplikative Schreibweise 89
—, Permutations- 86
—, Transformations- 86
Gubär-Ziffern 39, 40

halboffenes Intervall 67
Häufungspunkt 197
Höchstfehler 356—361
Höhe einer algebraischen Gleichung 308
— — rationalen Zahl 72, 73
hyperkomplexe Algebra 212, 215—221
— Zahl 211—221

identische Abbildung 71
Inder 33—35, 37—40, 53, 55
Indianer 13—15, 20
Individualbruch 47, 49
Inka 10
inkommensurable Strecken 162, 172, 175, 176
Integritätsbereich 93
Interpolation 381, 382, 391
Intervall, abgeschlossenes 67, 71, 72
—, halboffenes 67
—, offenes 67, 71, 72
inverse Abbildung 71
inverses Element in einer Gruppe 87, 88
— — — einem Körper 96
irrationale Zahl 64, 76, 186, 279, 284, 285, 297—302
irreduzibel 262
Isomorphie 64, 102—105

Jeton 35, 36, 38
Joruba 14, 16, 19, 37

k-adischer Bruch 274—279
Kardinalzahl 118
Kategorizität 131, 196
Kauri-Muscheln 14, 37
Kettenbruch 266—274, 280 bis 296
—, endlicher 266—274, 284, 285
—, unendlicher 280—285
Kharosthi-Ziffern 20, 21, 34, 37
Klamathen 15
klassifizierendes Verb 15
kleinstes gemeinsames Vielfaches 232
Knotenzahl 5, 14, 15, 17, 22, 28, 33, 34, 47, 59
Komma 57
kommensurable Strecken 162, 175, 176
kommutative Gruppe 85, 87
— Operation 84

kommutatives Gesetz 84, 86, 87, 90, 91, 116, 120, 195
komplexe Zahl 64, 198—210
— —, ganze 91
— —, rationale 91
Komponenten eines Vektors 212
Kongruenz 241—259
—, algebraische 251—259
konjugierte Quaternion 214
konjugiert komplexe Zahl 209
konstruktiv 194
konvergente Folge 167—175
Kopfrechnen 319—321, 390
Körper 64, 95—100, 103—105
—, angeordneter 105—111
—, archimedisch angeordneter 111
— der rationalen Funktionen 173, 174
— — — — Zahlen 100, 111, 148—163, 197
— — reellen Zahlen 111, 176 bis 198, 213—221
— der komplexen Zahlen 198—204, 213—221
—, endlicher 97
—, stetiger 175, 176
—, vollständiger 172—176
Kosinus 188—190
Kotangens 188
k-stellige Näherungszahl 340

leere Menge 66, 71, 73, 79, 81
letztes Element einer geordneten Menge 80, 81
linear abhängig 215
Lochkarte 334
logarithmischer Rechenschieber 386—389
logarithmische Skala 386
logarithmisches Papier 383
Löschvorrichtung 328

Mächtigkeit 71
Maori 16
Maßsystem, ägyptisches 49, 50
—, alt-russisches 49, 50, 56
—, dezimales 56, 58
mathematische Tafeln 380 bis 382, 390, 391
Maya 20, 33
Median 343
Menge 64—82
—, abzählbares 76—78, 308
— der ganzen Zahlen 85,90,
— — geraden Zahlen 66, 67, 72, 76, 91
— — komplexen Zahlen 76, 85, 91
— — natürlichen Zahlen 66, 67, 72, 75, 76, 91
— — Primzahlen 72, 76, 233
— — rationalen Zahlen 67, 72, 73, 76, 80, 85, 91
— — reellen Zahlen 67, 72, 76, 77, 85, 91
— — ungeraden Zahlen 72, 76
—, endliche 67, 73—78, 81, 82
—, geordnete 79—82
—, leere 66, 71, 73, 79, 81
—, nichtabzählbare 76, 77
—, überabzählbare 76, 77
—, unendliche 66, 67, 73—78
Mengen, Äquivalenz von 71 bis 74, 81
—, Gleichheit von 66
—, Gleichmächtigkeit von 71 bis 74, 81
Mengendifferenz 67
Mengensystem 68
Methode der Wertschranken 337, 346
metrischer Rechenschieber 285, 286
Minute, das Wort 55
mittlerer Fehler 343
— quadratischer Fehler 345, 363—366
Modell 131—133
Modul 89
Moivresche Formel 207
Monotoniegesetz 106, 107, 121, 122
Multiplikation auf der Sprossenradmaschine 329, 330
— mit dem Rechenschieber 387
— den Neperschen Streifen 326
— über kreuz 320
— von ganzen Zahlen 140
Multiplikation von komplexen Zahlen 202, 206, 207
— — natürlichen Zahlen 118—120
— — rationalen Zahlen 151
— — reellen Zahlen 181
multiplikative Gruppe 96, 96
multiplikatives Bezeichnungsprinzip 16, 17, 21, 25, 34, 36—38, 60

Nachfolger 80, 112, 121
n-adisch rationale Zahl 160 bis 162
Näherungsbruch 268—272, 280, 287—296
Näherungsbrüche, Approximation durch 287—291
Näherungsformeln 375, 376, 392
Näherungsgleichung 338, 339
Näherungswert 336—375, 392
Näherungszahl, k-stellige 340
Nasik, Höhleninschriften von 20, 34
natürliche Zahl 3, 64, 73, 112—133
negativ 105—111
negative Zahl 64, 134, 135, 136
Neger 13, 14, 36, 37
Nepersche Streifen 326, 390
Neunerprobe 244, 322
nichtabzählbare Menge 76, 77
Nichtabzählbarkeit der Menge der reellen Zahlen 76, 77, 309
nichtperiodischer systematischer Bruch 279
nicht-positionelles Bezeichnungssystem 5, 6, 17—28, 30—32, 34
Nomographie 383
Norm einer Quaternion 214
Null, das Wort 43
—, Entstehung der 35, 38, 60
—, Fehlen der 6, 26, 29
—, Zeichen für die 26, 34, 38, 42, 55
Nullelement einer Gruppe 89
— eines Ringes 92, 93

Nullfolge 182
Nullteiler 93, 96, 97, 107

obere Schranke 339, 347
Oberkörper 99, 100
Obermenge 66, 75
—, echte 66
Odhnersches Rad 328
Operation, algebraische 64, 83, 104
—, kommutative 84
—, Umkehrungen einer 84
Ordinalzahl 17, 118
Ordnung 80
— der ganzen Zahlen 146
— — natürlichen Zahlen 120—123
— — rationalen Zahlen 158
— — reellen Zahlen 184, 186
—, entgegengesetzte 80
Ordnungstypus 81

Papyrus, Moskauer 22
—, Rhind 22, 52
Peanosches Axiomensystem 112—114
periodischer systematischer Bruch 277—279
Permutation 82
Permutationsgruppe 86
Planimeter 333
Polynesier 9, 10
Positionsprinzip 3, 22, 23, 28, 29, 39, 40, 42, 60
Polynom 188, 211, 215, 261 bis 265
Polynome, Ring der 91, 111
Positionssystem 3—5, 21, 22, 25, 26, 28—45, 59
— der Babylonier 28—33, 54
— — Inder 33—35, 38, 53, 56
— — Maya 33
— mit beliebiger Basis 4
—, modernes dezimales 3, 5, 39—45, 58, 59
positiv 105—111
positive ganze Zahl 146
Potenz eines Gruppenelementes 88
— von natürlichen Zahlen 127
Primkörper 100
Primzahl 163, 228—231, 233 bis 240, 245, 246, 257
Primzahlsatz 238, 239
—, Dirichletscher 239, 240
Primzahlverteilung 234—240
Primzahlzerlegung 228—231
Primzahlzwillinge 237
Primzerlegung von Polynomen 264
Prinzip der größten Zahl 123, 147
— — kleinsten Zahl 122, 147
— von A. N. Krylow 363 bis 367
Produkt von reellen Funktionen 188
Pythagoreer 14

Quaternion 64, 211—221
Quipu 10, 11
Quotientenwerk 328
Quotient von Körperelementen 97
— — reellen Funktionen 188

rationale Funktion 95
— Zahl 64, 76, 148—163, 197, 279, 284, 285
rechenanalytische Maschine 333, 334
Rechenbrett 325, 390
Rechenfehler 323
Rechenmaschine 8, 327—336, 390
Rechenkontrolle 322, 323
Rechenschema 323
Rechenschieber, logarithmischer 386—390
—, metrischer 385, 386
Rechnungen mit Meßresultaten 336, 337
reduziertes Restsystem 248
reelle Funktion 187
— Zahl 58, 64, 67, 76, 176 bis 198
Regel der geraden Ziffer 338
Regeln der Ziffernzählung 357—361, 369—374, 392
rein imaginäre Zahl 204
reinperiodischer systematischer Bruch 278
Rektifikation eines Kreisbogens 176
Restgleichheit 242
Restklasse 99, 246—251
Restklassenkörper 100, 246
Restklassenring 99, 246
Restsystem, reduziertes 248
—, vollständige 247, 248, 250
Resultatwerk 328
Ring 64, 90—94, 100, 103 bis 105
—, angeordneter 105—111
—, archimedisch angeordneter 111
— der ganzen Zahlen 111, 134—148, 197, 198
— — Polynome 91, 111
—, endlicher 97
—, nichtkommutativer 211
Römer 36, 47, 49, 50, 53
runden, 338, 369
—, nach oben 338
—, nach unten 338

Schiefkörper 211, 212
— der Quaternionen 213 bis 215, 218—221
Schlitten 327
Schnitt in einer geordneten Menge 197
Schnittzahl 197
Schranke des absoluten Fehlers 339, 341, 350
— — relativen Fehlers 340, 341, 351, 352
—, obere 339, 347
—, untere 339, 347
schriftliches Rechnen 321 bis 324, 390
Sekunde, das Wort 55
Sexagesimal-Bruch 26, 42, 53—58
Sexagesimalsystem der Babylonier 28—33, 54
Sinus 188—190, 192, 193
Sprossenrad 328
Sprossenradmaschine 327 bis 331
stationäre Folge 174, 184
stetige Funktion 187—191
stetiger Körper 175, 176
Strecken, inkommensurable 162, 172, 175, 176
—, kommensurable 162, 175, 176
Streckenverhältnis 162, 164, 165, 171, 172, 175, 176

Strukturkonstanten 213
Subtraktion auf der Sprossenradmaschine 329
— in einer Gruppe 90
— — einem Ring 92
— von natürlichen Zahlen 128
subtraktives Bezeichnungsprinzip 6, 16, 17, 19, 21, 59
Sumerer 29—32, 47
Summe von reellen Funktionen 188
systematischer Bruch 266, 274—279
— —, endlicher 275, 279
— —, gemischtperiodischer 278, 279
— —, nichtperiodischer 279
— —, periodischer 277—279
— —, reinperiodischer 278
— —, unendlicher 275
— Fehler 343

Tangens 188
Tasmanier 6, 12
Teilbarkeit 228—233
— von Polynomen 262
teilerfremd 229, 230
— für Polynome 262
Teilkörper 99, 100, 104, 105
Teilmenge 66, 75, 76
—, echte 66, 71, 75, 76
Teilnenner eines Kettenbruches 267
Transformationsgruppe einer Menge 86
transzendente Zahl 304—311;
— Zahlen von Liouville 305—307
Transzendenz von e 310, 311
— — π 310, 311
Triebwerk 328
trigonometrische Darstellung der komplexen Zahlen 205—207
Tschimsian 14

überabzählbare Menge 76, 77
Überschlagrechnung 320 bis 322
Umdrehungszählwerk 328
Umfang des Kreises 176
umkehrbar eindeutige Abbildung 70, 71, 85
Umkehroperation 84
Unabhängigkeit eines Axiomensystems 132, 133, 196
unendliche Menge 66, 67, 73—78
unendlicher Dezimalbruch 76, 77
— Kettenbruch 280—285
— systematischer Bruch 275
unmittelbarer Nachfolger 80, 112, 113, 121
— Vorgänger 80, 113
untere Schranke 339, 347
Untergruppe 90
Unterkörper 99, 100, 104, 105
Untermenge 66, 75, 76
—, echte 66, 71, 75, 76
Unterring 94, 104, 105
Urbild eines Elementes bei einer Abbildung 70
—, volles 70

Vektor 211, 212
Vektorraum 211—213
Vereinigung 67
— eines Mengensystems 68
Vielfaches einer natürlichen Zahl 127
— eines Gruppenelementes 89
— — Ringelementes 92, 93
volles Urbild 70
vollständige Induktion, Axiom der 112
— —, Beweis durch 113, 122, 123, 147
— —, Definition durch 114, 115, 123—125
vollständiger Körper 172 bis 176, 210
vollständiges Restsystem 247, 248, 250
Vollständigkeit eines Axiomensystems 131, 132, 196
Vollständigkeitsaxiom 172, 196
Vorgänger 80, 113

Wertevorrat einer Funktion 69
wesentliche Ziffer 338, 360, 361
Widerspruchsfreiheit eines Axiomensystems 130, 131, 196
Wilsonscher Satz 257
Wurzel 162, 163, 165, 166, 171, 172, 198, 199, 207 bis 209, 303
Wurzelziehen auf der Sprossenradmaschine 331, 332

Zahl, algebraische 303—306
— als Eigenschaft 9
— — Stückanzahl 13, 14, 36
—, der Begriff 6, 7, 63, 64
—, ganze 7, 64, 134—148, 197, 198
—, ganze komplexe 91
—, hyperkomplexe 211—221
—, irrationale 64, 76, 186, 279, 284, 285, 297—302
—, komplexe 64, 198—210
— mit algorithmischer Bezeichnung 5, 6, 14, 15, 17, 28, 33, 34, 46, 47, 54
— — individueller Bezeichnung 5, 14, 15, 17, 22, 28, 33, 34, 47, 49, 59
—, natürliche 3, 64, 73, 112 bis 133
—, rationale 64, 76, 148—163, 197, 279, 284, 285
—, rationale komplexe 91
—, reelle 58, 64, 67, 76, 176 bis 198
—, rein imaginäre 204
—, transzendente 304—311
—, Versinnbildlichung durch Körperteile oder dergleichen 10, 12
Zählmethoden der Abiponer 9
— — Aruntier 10
— — Eingeborenen der Fidschi-Inseln 13
— — — — Halbinsel Prince York 13
— — — — Inseln der Torres-Straße 8, 10, 12, 13
— — — von Florida 13
— — — — Neu-Guinea 12, 37
— — Polynesier 10
— — Tschimsian 14
— kanadischer Indianer 13, 14

Zahlsystem, dyadisches, 8, 19
—, dezimales 5, 12, 30, 31, 34, 39—45, 55
Zahlwörter 4, 5, 27, 28, 47
Zahlzeichen, ägyptische 17 bis 19, 21, 22, 28, 46, 49
—, algorithmische 5, 6, 14, 15, 17, 28, 33, 34, 46
—, arabische 24, 39
—, arabisch-indische 40—43, 45
—, armenische 24
—, assyrisch-babylonische 18, 28—33, 47, 49
—, attische (herodianische) 18, 19, 23
—, chinesische 20, 21
— der Azteken 15, 20, 21
Zahlzeichen der Höhleninschriften von Nasik 30, 34
— — Maya 20
—, georgische 24
—, glagolitische 21, 22, 24, 27
—, hebräische 24
—, individuelle 4—6, 14, 15, 17, 22, 28, 33, 34, 47, 59
—, ionische 19, 21—26, 47, 52
—, Kharosthi-Ziffern 20, 21, 34, 37
—, kretische 21
—, kyrillische 21, 22, 24, 26 bis 28, 43, 45
—, palmyrenische 18, 19
—, phönizische 18, 19
—, römische 6, 18, 23, 27, 41 bis 43, 53
—, singhalesische 22
Zahlzeichen, syrische 18
Ziffer, das Wort 43
—n, arabisch-indische 40 bis 43, 45
—, römische 6, 18, 23, 27, 41 bis 43, 54
—, wesentliche 338, 360, 361
—, zuverlässige 340
Ziffernprinzip 22
zufälliger Fehler 343
zusammengesetzte Zahl 228, 234
Zusatzziffer 369
zuverlässige Ziffer 340
Zwanziger-System 5, 12, 21, 33
Zwischenwertsatz 190, 191
Zwölfer-Brüche, römische 42, 53