Über Zahlen und Überzahlen
von Prof. Dr. Franz von Krbek

mit 106 Textabbildungen
illustriert von Horst Röcke
1. Auflage

Abschrift und LaTex-Satz: Steffen Polster 2020
https://mathematikalpha.de

Vorwort

Für jeden etwas sollten unsere Plaudereien bieten, die darum abwechselnd leicht und anspruchsvoll gehalten wurden.
Die Schulmathematik bildet dabei überall unsere Basis. Die Plaudereien sollten jedoch nicht allein neue Kenntnisse vermitteln, sondern dem Leser zu konzentriertem Denken verhelfen, das heißt aber, sein Abstraktionsvermögen entwickeln. Das ist eine Forderung, die das Zeitalter der Weltraumfahrt an uns stellt.

Die leitenden Gedanken wurden konsequent herausgearbeitet. Sie sollten den Leser zu selbstgestellten Fragen anregen, deren Behandlung vom Lernen zu eigenem Schaffen führt. Ein solches Training ist für die mathematische Ausbildung, auf die es heute immer mehr ankommt, von grundsätzlicher Bedeutung. Die Mathematik durchdringt und befruchtet ja zusehends Chemie, Biologie, Medizin und Ökonomie, lauter Disziplinen, die früher wenig mit ihr zu tun hatten.

Der Verlagsgesellschaft B.G. Teubner in Leipzig möchte ich für die ansprechende Ausstattung meinen aufrichtigen Dank aussprechen. Insbesondere aber gilt mein Dank meiner jungen Kollegin, Frau Dorothea Ziegler, für ihre unermüdliche Betreuung des Buches.

Im Sommer 1964 Franz von Krbek

Inhaltsverzeichnis

1 Mehr oder weniger - eine Vorstufe zur Zahl

Seit altersher ist es der Brauch, dass der neugewählte Rektor einer Universität sein Amt mit einer Rede aus seinem Fachgebiet antritt. In der Antrittsrede des ungarischen Mathematikers Friedrich Riesz in Szeged aus dem Jahr 1925 heißt es, dass der Professor Dirichlet aus Berlin 1843 gerade während der Ostermesse in der sixtinischen Kapelle in Rom die Tragweite des Schubfächerprinzips in der Zahlentheorie erkannte.
Woher Riesz das hat, weiß ich nicht. In den ausführlichen Briefen der Frau Dirichlet an die Schwester steht lediglich, dass,sobald die Kirchenmusik aussetzte, ein störendes Hustenkonzert einsetzte. Frau Dirichlet war übrigens Enkelin des Philosophen Mendelssohn und Schwester des Komponisten Mendelssohn-Bartholdy.

Was ist unter dem Schubfächerprinzip, von dem eben die Rede war, zu verstehen? Wenn es mehr Perlen als Schubfächer gibt, dann befinden sich in wenigstens einem der Schubfächer zwei Perlen, wenn diese auf die Schubfächer verteilt wurden.
Legt man etwa zunächst in jedes Schubfach eine Perle, dann bleibt mindestens eine Perle übrig, die in ein Schubfach gerät, in dem sich daraufhin zwei Perlen befinden. Wenn dagegen bei der anfänglichen Verteilung, bei der immer nur eine Perle in ein Schubfach gelegt wurde, ein oder mehrere Schubfächer leer ausgingen, dann bleiben erst recht Perlen übrig, und zwar mehr als leere Schubfächer da sind, da es ja mehr Perlen als Schubfächer gibt.
Setzt man die Verteilung der Perlen fort, dann kann man zunächst je eine Perle in die leeren Schubfächer tun und ist damit bei der ersten Etappe der anfangs behandelten Verteilung angelangt.

Dabei blieb mindestens eine Perle übrig, die folglich ein Schubfach mit der darin vorhandenen Perle teilen muss. Wie man es auch einrichtet, in wenigstens ein Schubfach geraten also mindestens zwei Perlen.

Der Mathematiker nennt den eben auseinandergesetzten Sachverhalt kurz Schubfächerprinzip. Es gestattet recht überraschende Antworten, so etwa auf die Frage, ob es zwei Berlinerinnen gibt, die gleich viele Haare haben. Eine Beantwortung wird hier dadurch

möglich, dass die Anzahl der Berlinerinnen größer ist als die größtmögliche Zahl von Haaren bei einer Frau. Die Berlinerinnen spielen dann in unserer Überlegung die Rolle von Perlen, die unterschiedlichen Anzahlen von Haaren - die eine Frau hat eben mehr Haare, die andere weniger - die Rolle von Schubfächern.
Daher lautet die Antwort: Es gibt bestimmt zwei Berlinerinnen mit gleich vielen Haaren! Freilich weiß man dann noch immer nicht, welche die beiden sind. Damit begegnet uns eine Überlegung, die zu den in der Mathematik häufig vorkommenden Existenzbeweisen gehört.

Das Schubfächerprinzip gestattet jedoch keine Antwort auf die Frage, ob es zwei Berlinerinnen gibt, die beide 100000 Haare haben. Man weiß lediglich, dass mindestens eine Haarzahl (aber nicht welche) wenigstens zweimal vertreten sein muss. Wollte man hingegen die beiden Berlinerinnen mit gleich vielen Haaren namhaft machen, dann hätte man ganz anders vorzugehen. Man müsste bei den einzelnen Berlinerinnen die Haare wirklich zählen, bis man eine findet, die ebenso viele Haare hat wie eine ihrer Vorgängerinnen.

Dass man schließlich auf eine solche Berlinerin stoßen muss, wissen wir schon. Diesmal kennt man dann die beiden Berlinerinnen mit gleich viel Haaren namentlich, im Gegensatz zu vorhin, und weiß auch, wieviel Haare sie haben.

Eine weitere Überraschung verschafft folgende Anwendung des Schubfächerprinzips. Mit etwa 60 Schriftzeichen (kleinen und großen Buchstaben, Punkt, Komma, dann dem die Worte trennende Zwischenraum usw.) kommt man in jedem Roman aus, hat man sich erst einmal zu einer bestimmten Sprache entschlossen.
Die einzelnen Worte darin mögen bis zu 50 Buchstaben ausschließlich enthalten, und es sollen nicht mehr als eine Million Worte vorkommen, denn kaum ein Roman enthält auch nur annähernd eine Million Worte. Wie großzügig wir dabei waren, erhellt die orientalische Märchensammlung "1001 Nacht". In der 445. Nacht heißt es, der Koran enthält 79439 Worte, die ihrerseits aus 323670 Schriftzeichen bestehen.

Um nun, wenn auch nur sehr roh, abzuschätzen, wie viele verschiedene Texte unter unseren Bedingungen möglich sind, bemerken wir zunächst, dass jeder der Romane aus höchstens 50 Millionen Schriftzeichen besteht. In jedem Roman folgt auf eines der 60 verschiedenen Schriftzeichen an erster Stelle ein zweites der 60 Schriftzeichen an zweiter Stelle. Das kann auf 60×60 verschiedene Weise geschehen.

An dritter Stelle steht wiederum eines der 60 Schriftzeichen, das auf jede der 60 mal 60 Möglichkeiten von vorhin folgen kann. Das ergibt dann $60 \times 60 \times 60$ verschiedene Möglichkeiten. So findet man schließlich, dass die Anzahl der verschiedenen Möglichkeiten, mit 60 Schriftzeichen 50 Millionen Stellen auszufüllen, durch ein Produkt von 50 Millionen Faktoren gegeben wird, in dem sie, alle Faktoren gleich 60 sind. Diese Anzahl lässt sich in Form einer Potenz als

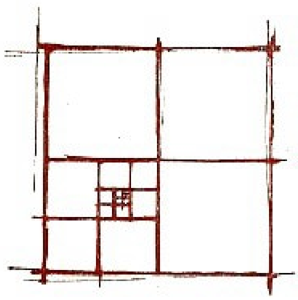
$$60^{50000000}$$

schreiben.

Darunter befinden sich aber Texte ganz ohne Sinn, die zum Beispiel aus der Wiederholung eines einzigen Buchstabens bestehen. Jedenfalls ist die Anzahl aller möglichen Romane zwar unvorstellbar groß, aber bestimmt endlich.

Diese endliche Höchstzahl möge jetzt die Rolle von Schubfächern spielen, die Romane aber die Rolle von Perlen. Dann lautet das Ergebnis, dass bei hinlänglich großer, aber immer noch endlicher Anzahl von Romanen zwei unbedingt gleich sein müssen; Titel und Text stimmen genau überein.
Das würde sicher zutreffen, wenn die Menschen lange Zeit genug Romane schreiben würden. Vielleicht nach Jahrtausenden erst, aber ganz bestimmt würde ein zweiter Autor auftreten, der, ohne seinen Vorgänger zu kennen, Wort für Wort denselben Text schreiben würde.

Dieses Paradoxon stammt in seinen Anfängen von Lasswitz und wurde 1914 von Hausdorff, einem bekannten Mathematiker, wiederentdeckt. Er verkannte jedoch dessen wahres Wesen und verquickte es mit dem Unendlichen, während wir davon keinen Gebrauch machten. Allerdings lässt sich das Schubfächerprinzip noch weiterbilden.
Gibt es unendlich viele Perlen, aber nur endlich viele Schubfächer, auf die diese Perlen zu verteilen sind, dann müssen sich in mindestens einem Schubfach unendlich viele Perlen befinden. Würde nämlich jedes der Schubfächer höchstens endlich viele Perlen enthalten, dann wäre die Gesamtzahl der Perlen selbst endlich.



Das so weiterentwickelte Schubfächerprinzip spielt in unserer zeitgenössischen Mathematik eine entscheidende Rolle.
Man betrachte ein Quadrat, in dem unendlich viel Punkte markiert sind. Vierteilt man das Quadrat, dann gibt es in mindestens einem der vier neuen Quadrate unendlich viel markierte Punkte.

Vierteilt man dieses und fährt dann fort, immer wieder jeweils das Quadrat mit unendlich vielen markierten Punkten zu vierteilen, dann gewinnt man eine Folge ineinandergeschachtelter Quadrate, von denen jedes unendlich viele markierte Punkte enthält.
Ein beliebig kleiner Kreis um den Punkt, auf den diese Folge von Quadraten zusammenschrumpft, enthält dann ganz im Inneren ein gewisses Quadrat sowie alle kleineren ineinandergeschachtelten Quadrate und damit auch alle markierten Punkte aus diesem Quadrat. Es handelt sich um einen Häufungspunkt der markierten Punkte, ein Begriff, der erst im vorigen Jahrhundert eingeführt wurde.

Zu der eben geschilderten Einsicht gelangte bereits Bolzano. Er bekleidete in der ersten Hälfte des vorigen Jahrhunderts eine Professur für Theologie in Prag. Wegen seiner liberalen Ansichten wurde er von Metternich aus dem Amt gejagt, so dass er auf dem kleinen Landsitz seines Bruders untertauchen musste.

Bolzano war ein passionierter Mathematiker; er entwickelte grundlegende Ideen, die leider erst nach seinem Tode bekannt und von anderen wiedergefunden wurden. So entdeckte Weierstraß den im vorhergehenden Absatz behandelten Satz in der zweiten Hälfte des vorigen Jahrhunderts erneut, so dass man vom Satz von Bolzano-Weierstraß spricht.

Im Gegensatz zu Bolzano nahm Weierstraß nachhaltigen Einfluss auf die zeitgenössische Mathematik. Jahrzehnte hindurch lehrte er an der Berliner Universität und setzte die seit Cauchy begonnene exakte Begründung der Analysis fort. Er verhalf dabei diesem Teilgebiet der Mathematik zu der heute üblichen Exaktheit und Strenge, ohne die kein Ergebnis als gesichert angesehen werden kann.
Zum Schluss, sei noch vermerkt, dass die Quelle unserer Überlegungen noch vor einem Abzählen, also vor Einführung des Zahlbegriffes, liegt: es genügte ein "mehr oder weniger".

2 Zauberer zürnen Zahlen

Märchen verraten geheime Wünsche. Seit jeher wollte man gut leben, doch nur den wenigsten war es vergönnt. Die anderen versuchten sich mit Hilfe ihrer Phantasie, so gut es ging, über das graue Dasein hinwegzutrösten. Sie erdachten Feen und Zauberer, denen die Natur gehorchte, sodass sie schier wunderbare Gaben zu verteilen vermochten.
Wer glaubt heute noch daran? Der Glaube schwindet, je besser der Mensch die Natur meistert. Das gelingt ihm mit Hilfe der Zahl als Vermittler.

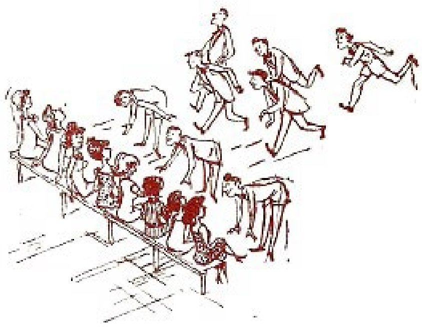
Sie erst gestattet es, Naturgesetze so zu formulieren, dass wir damit die Natur immer besser beherrschen. So lösten Zahlen die Zauberer ab.

Den Ausgang für unseren Zahlbegriff bildete das Abzählen. Der Mensch begann, die Gegenstände seiner Umgebung quantitativ zu erfassen. Naturvölker, an denen man sich orientieren kann, zählen noch heute an den Fingern oder mit Steinchen, Muscheln und dergleichen. Soll etwa eine Herde gezählt werden, dann legen sie für jedes Tier ein Steinchen hin. Die begriffliche Erfassung dieses Vorgehens bildet den Schlüssel zu unseren Überlegungen.

Allerdings zählt man heute so, dass man Zahlwörter mit den Gegenständen paart, die abgezählt werden sollen. Diese Paarbildung hat man wie folgt genau zu formulieren.
Aus einer Menge von Zahlwörtern und einer zweiten Menge von Gegenständen werden Paare gebildet, und zwar so, dass jedes einzelne Paar ein Zahlwort und einen Gegenstand enthält. Dabei möge jedes Zahlwort in genau einem Paar und jeder Gegenstand gleichfalls in nur einem Paar vorkommen.

Man denke an eine Tanzstunde mit gleich vielen Damen und Herren, also ohne Mauerblümchen.

Man kann noch weitergehen. Sieht man auch noch von der speziellen Beschaffenheit der beiden Mengen ab, dann liegt eine abstrakte Paarung vor, die in der Mathematik als umkehrbar eindeutige oder auch eineindeutige Abbildung der beiden Mengen aufeinander bezeichnet wird. Diese Allgemeinheit reicht dann aus, um mühelos zu erkennen, dass es sich dabei gar nicht mehr um endliche Mengen zu handeln braucht.



Gewiss, fertig wird man mit dem Abzählen nur im Fall von endlichen Mengen. Aber auch das nur im Prinzip. Niemand wird bezweifeln, dass man die Sandkörner der Sahara abzählen könnte, aber ebenso wenig wird man bezweifeln, dass keiner das wirklich schaffen würde.
Es genügt eben die Einsicht der prinzipiellen Möglichkeit dafür, die aber auch für unendliche Mengen besteht.

Dabei geht es augenscheinlich um eine Fortsetzung des Zählens über das Endliche hinaus. Die vorhandenen Zahlwörter reichen, dazu nicht aus. Um Abhilfe zu schaffen, braucht man nur zu bedenken, dass Zahlwörter nur Bezeichnungen sind, die für bestimmte Begriffe eingeführt und jeweils nach den Bedürfnissen erweitert werden können.
So hatten die Urvölker anfangs nur ganz wenige Zahlwörter: eins, zwei, viel. Mit den steigenden Anforderungen und wachsenden Kenntnissen entstanden nach und nach die uns bekannten Zahlwörter, von denen die Wörter Million, Billion usw. erst seit einigen Jahrhunderten verwendet werden. Um nun auch noch den unendlichen Mengen gerecht zu werden, führen wir einfach zusätzlich noch umfassendere Bezeichnungen ein.

Eine Möglichkeit dafür ist die Klasseneinteilung der Mengen. Zu einer Klasse gehören jeweils alle Mengen, die umkehrbar eindeutig aufeinander abgebildet werden können, die also äquivalent sind. Beispielsweise gehören alle Mengen aus 5 Elementen zu einer Klasse, da sich aus Elementen von jeweils zwei dieser Mengen in der von uns beschriebenen Weise Paare bilden lassen. Diese Klasse könnten wir etwa "fünf" nennen.

Man kann daraufhin die eben eingeführten Klassen als eine neue Art von Zahlwörtern betrachten. Sie wurden von Georg Cantor Kardinalzahlen oder auch Mächtigkeiten genannt. Die Kardinalzahlen von endlichen Mengen sind gleichbedeutend mit den geläufigen Zahlwörtern eins, zwei, drei usw., welche insgesamt die natürlichen Zahlen repräsentieren. Die Berechtigung für den Namen Kardinalzahl wird dadurch erbracht, dass man mit ihnen rechnen kann; doch gibt es zwischen den endlichen und den unendlichen Kardinalzahlen, die als eine Art Überzahlen aufgefasst werden können, einen ausschlaggebenden Unterschied.

Im Gegensatz zu den unendlichen Mengen gilt der Satz, dass keine endliche Menge einer ihrer echten Teilmengen äquivalent ist. Dafür ein neuer Beweis:

Wir nehmen an, eine bestimmte endliche Menge M wäre einer ihrer echten Teilmengen N äquivalent. Da N echte Teilmenge von M ist, kommt wenigstens ein Element a aus M in N nicht vor. Die Äquivalenz von M mit N wird durch Paare, bestehend aus je einem Element aus M und einem Element aus N, beschrieben. Von diesen Paaren lasse man dasjenige Paar weg, in dem das Element a vorkommt. Es sei das Paar (a, b).
Da wir vorher eine Äquivalenz von M und N angenommen hatten, drücken dann die übriggebliebenen Paare offenbar wieder eine Äquivalenz der um das Element a verringerten Menge M, die mit M' bezeichnet werde, mit einer Teilmenge N' von N aus, die aus N durch Streichen von b entstand. Da aber b in M' vorkommt, ist N' echte Teilmenge von M'. Nun streiche man das Paar, in dem b enthalten ist; es sei das Paar (b, c).

Damit entsteht aus M' durch Weglassen von b die Menge M'' und aus N' durch Weglassen von c die Menge N'', die wiederum echte Teilmenge von M'' ist. Nach unserer Annahme

müssen auch M'' und N'' wieder äquivalent sein.
Wenn wir dieses Verfahren fortsetzen, bleibt schließlich nach einem gewissen Schritt - da beide Mengen endlich sind - von der Untermenge nur ein Element übrig, während von der Menge M noch mehrere, mindestens zwei, vorhanden sind, weil N echte Teilmenge von M war.
So käme man schließlich zur unsinnigen Behauptung, dass eine Menge aus mehreren Elementen einer Menge mit einem einzelnen Element äquivalent ist.

Für unendliche Mengen trifft das Gegenteil zu. Jede unendliche Menge ist bestimmt einer echten Teilmenge äquivalent.
Zunächst betrachten wir die Menge aller natürlichen Zahlen. Eine Teilmenge davon bildet die Menge aller geraden Zahlen. Ordnet man den Zahlen 1, 2, 3, aus der ersten Menge der Reihe nach die Zahlen 2, 4, 6, aus der zweiten Menge zu, dann stellt man damit eine umkehrbar eindeutige Abbildung der beiden Mengen aufeinander her. Somit ist die Menge aller natürlichen Zahlen der echten Teilmenge aller geraden Zahlen äquivalent.

Liegt nun eine beliebige unendliche Menge M vor, dann wähle man daraus eine Folge F von Elementen. Da die Menge M unendlich ist, bricht F nicht ab. Die um F verringerte Menge M bezeichnen wir mit M'.
Behält man aus F nur jedes zweite Element wie vorhin im Falle von natürlichen Zahlen, dann ist diese Teilfolge F' der ursprünglichen Folge F äquivalent. Fügt man F' der Menge M' hinzu, dann wird die Äquivalenz zwischen M und der Menge M'', die aus M' durch Hinzufügen von F' entsteht, durch die Abbildung hergestellt, die jedem Element von M' dieses Element selbst, den Elementen aus F aber der Reihe nach die Elemente von F' zuordnet.
Damit ist die Äquivalenz von M mit der echten Teilmenge M'' nachgewiesen.

Der letzte Satz warnt uns davor, bei der Verallgemeinerung von Einsichten, seien sie auch an noch so zahlreichen Einzelfällen gewonnen, sorglos vorzugehen. Sicher wird im Alltag das Ganze stets größer als ein echter Teil ausfallen.
Für unendliche Mengen trifft das dagegen nicht zu, wie wir vorhin erkannt haben. Dadurch lernt man die Kühnheit zu würdigen, die dazu gehörte, unendliche Mengen systematisch in den Bereich mathematischer Überlegungen einzubeziehen; um so mehr, als dabei das Plausible oft falsch ist, während das Richtige paradox erscheint.

Diese Kühnheit besaß Georg Cantor, dessen Namen wir schon genannt haben.
Er lehrte an der Universität Halle. Seine bedeutendsten Arbeiten erschienen von 1873 bis 1880. Leider wurden seine Ideen in Deutschland zunächst in ihrer Bedeutung nicht erkannt und anerkannt, so dass er von 1880 an in den "Acta mathematica" veröffentlichen musste, in einer wissenschaftlichen Zeitschrift, die ein schwedischer Schüler von Weierstraß in diesem Jahr in Stockholm ins Leben rief.

Die Ideen von Cantor fielen dann in Frankreich auf fruchtbaren Boden, auf dem um die Jahrhundertwende die Theorie der reellen Funktionen gediehen ist. In ihren Hauptzügen wurde sie von Lebesgue entwickelt und ist für die heutige Mathematik unerlässlich.

3 Ein Übereinmaleins

Man kann auch ganz ohne Einmaleins rechnen, bloß geht das dann sehr umständlich vor sich. Man kann das im Rechen"buch" des Ahmes, der um 1650 v.u.Z. lebte, gut verfolgen.
Um ein Produkt wie $9 \cdot 13$ auszurechnen, verdoppelte der Ägypter der Pharaonenzeit den zweiten Faktor fortgesetzt, indem er die zu verdoppelnde Zahl jedesmal zu sich selbst addierte. So entstand das Schema

*1	13
2	26
4	52
*8	104

In der linken Spalte stehen die Angaben der aufeinanderfolgenden Verdopplungen, in der rechten die Ergebnisse. Die mit einem Stern versehenen Angaben in der linken Spalte ergeben zusammen den ersten Faktor 9, die Summe der entsprechenden Posten aus der rechten Spalte $13 + 104$ das Endergebnis 117.[1]

Das eben geschilderte Verfahren ägyptischer Rechner führt immer zum Ziele. Das liegt daran, dass jede natürliche Zahl eine Summe von Potenzen der Zahl 2 ist. Daher kann man die 2 zur Grundzahl wählen und so jede Zahl mit Hilfe von nur zwei Zittern an Stelle der zehn Ziffern unseres Dezimalsystems darstellen.

Ein Nachteil dieses dyadischen Systems besteht darin, dass verhältnismäßig kleine Zahlen recht viele Stellen erfordern. So muss man die im Dezimalsystem als 64 bezeichnete Zahl im dyadischen System als 1000000 schreiben, da ja

$$\begin{aligned} 64 =& 1 \cdot 2^6 + 0 \cdot 2^5 + 0 \cdot 2^4 + 0 \cdot 2^3 \\ &+ 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 \end{aligned}$$

ist. Allerdings könnte man dann neckischerweise vielen Jubilaren zum millionsten Geburtstag gratulieren.

Wenn man im dyadischen System viele Stellen braucht; dann liegt das an der Kleinheit

[1]Weitere Einzelheiten finden sich in meinem Buch "Eingefangenes Unendlich" (3. Aufl. 1962)

der Grundzahl. Die Babylonier hatten im Gegensatz dazu eine recht große Grundzahl, die 60.
In diesem Sexagesimalsystem brauchen selbst große Zahlen nur wenig Stellen, dafür aber bedarf es 60 Ziffern. Dementsprechend besteht dann das "kleine" Einmaleins aus 1770 Posten, im Gegensatz zu den 45 Posten im Dezimalsystem. Der geplagte Rechner in Babylon hatte auch seine Einmaleinstabelle stets zur Hand.
Das Sexagesimalsystem gestattet eine überraschende Anwendung. Wir haben schon gesehen, dass jeder Roman oder auch jedes Gedicht sich mit Hilfe von 60 Schriftzeichen, die nicht mehr als 50 Millionen Stellen ausfüllen mögen, hinschreiben lässt. Fasst man die Schriftzeichen als Ziffern im Sexagesimalsystem auf, dann besagt das soviel, dass sich jeder Roman als eine im Sexagesimalsystem geschriebene Zahl deuten lässt. Selbst das schönste Liebesgedicht bildet keine Ausnahme.

Man kann die Zahl noch in das Dezimalsystem umrechnen, um sie mit unseren gebräuchlichen Ziffern hinschreiben zu können.
Will man dann eine mit unseren Ziffern hingeschriebene Zahl wieder in einen Roman verwandeln, so braucht man sie nur in das Sexagesimalsystem umzurechnen. So wird aus jedem Roman eine Zahl, aber nicht umgekehrt, weil die geschilderte Verwandlung in einen Roman bei aufs Geratewohl hingeschriebenen Zahlen keinen sinnvollen Text zu liefern braucht.

Addieren lernt schon das Kind. Es lernt es mehr mechanisch, während, uns der Mechanismus selbst interessiert.
Man addiert 2 Äpfel zu 3 Äpfeln, indem man die 2 Äpfel zu den 3 Äpfeln legt. Allgemeiner gesprochen, gewinnt man die Summe von zwei natürlichen Zahlen m und n so, dass man zwei Mengen, die aus m bzw. n Elementen bestehen, vereinigt und dann die Gesamtzahl der Elemente bestimmt.
Dieses Verfahren bleibt nicht nur auf endliche Mengen beschränkt, denn wir wissen schon, dass bei unendlichen Mengen die Kardinalzahlen die Rolle von Anzahlen übernehmen.

Das ermöglicht uns, ganz allgemein die Summe von zwei Kardinalzahlen zu erklären. Es ist üblich, Kardinalzahlen durch kleine deutsche Buchstaben zu bezeichnen, so dass unsere beiden Kardinalzahlen $\mathfrak{m}$ und $\mathfrak{n}$ heißen mögen.

Da $\mathfrak{m}$ einer bestimmten Klasse von äquivalenten Mengen entspricht, wählen wir aus dieser Klasse irgendeine Menge; entsprechend verfahren wir bei $\mathfrak{n}$. Wir dürfen voraussetzen, dass die beiden ausgewählten Mengen kein gemeinsames Element besitzen, d.h., dass sie disjunkt sind.
Fasst man die Elemente aus den beiden Mengen zu einer einzigen Menge, der Vereinigungsmenge, zusammen, dann definiert man deren Kardinalzahl als die Summe von $\mathfrak{m}$ und $\mathfrak{n}$ und schreibt dafür $\mathfrak{m} + \mathfrak{n}$.

Zur Rechtfertigung dieser Erklärung müssen wir aber zeigen, dass es dabei nicht darauf ankommt, welche Mengen M bzw. N man aus den beiden Klassen mit den Kardinalzahlen $\mathfrak{m}$ bzw. $\mathfrak{n}$ ausgewählt hat.
Zunächst sind bei einer anderen Wahl M' bzw. N' der beiden Mengen M und M' bzw. N und N' äquivalent. Die mit dieser Aussage gleichbedeutende Paarbildung aus Elementen

von M und M' bzw. N und N' ergibt aber die Äquivalenz der Vereinigungsmengen M und N bzw. M' und N'. Man hat nur zu bedenken, dass M und N bzw. M' und N' disjunkt sind, so dass jedes Element der beiden Vereinigungsmengen nur in einem einzigen Paar vorkommt.

Man hat damit die Summenbildung auf unendliche Mengen ausgedehnt. Wie nicht anders zu erwarten, erlebt man bei unendlichen Mengen Überraschungen.
Die Vereinigungsmenge der Menge aller geraden Zahlen und der Menge aller ungeraden Zahlen enthält offenbar sämtliche natürlichen Zahlen. Nun wissen wir aber, dass die erste der zweiten Menge äquivalent ist. Ähnlich zeigt man, dass sie auch der dritten Menge äquivalent ist. Bezeichnet man, wie üblich, die Kardinalzahl der Menge aller natürlichen Zahlen mit $\mathfrak{a}$, dann folgt die paradoxe Gleichung $\mathfrak{a} + \mathfrak{a} = \mathfrak{a}$. Im Reich der natürlichen Zahlen hingegen ist null die einzige Zahl, die zu einer anderen Zahl addiert werden darf, ohne dass sich deren Wert ändert.

Man kann noch weitergehen und die Summe von beliebig vielen, auch unendlich vielen, Kardinalzahlen definieren. Das gelingt wiederum durch Vermittlung der Vereinigungsmenge.
Um dann zur Multiplikation vorzudringen, orientiert man sich daran, dass $3 \cdot 4$ als die Summe $4+4+4$ erklärt ist. So kam Cantor schließlich dazu, Summe, Produkt und Potenz von Kardinalzahlen zu erklären und eine regelrechte Arithmetik der Kardinalzahlen zu entwickeln.
Diese Arithmetik stimmt keineswegs in allen Einzelheiten mit der Schularithmetik überein, wie wir es an der Gleichung $\mathfrak{a} + \mathfrak{a} = \mathfrak{a}$ erkannten, ist aber durchaus folgerichtig und widerspruchsfrei.

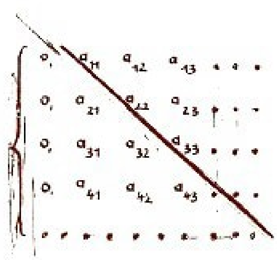
4 Irrlichter

Bis jetzt wissen wir noch nicht, ob nicht am Ende alle unendlichen Mengen äquivalent sind. Cantor führte scharfsinnig einen Nachweis, dass dem nicht so ist. Dieser Nachweis ging in die Geschichte der Mathematik als das Diagonalverfahren von Cantor ein. Man redet auch vom zweiten Diagonalverfahren, weil ein anderes Diagonalverfahren bereits auf Cauchy zurückgeht. Mit seinem Diagonalverfahren gelang Cantor der Nachweis, dass $\mathfrak{a}$ von der Kardinalzahl der Menge aller reellen Zahlen zwischen 0 und 1 verschieden ist, dass also die Menge der natürlichen Zahlen der Menge aller reellen Zahlen zwischen 0 und 1 nicht äquivalent ist.
Anders ausgedrückt heißt das, dass die letzte Menge nicht als Folge geschrieben werden kann. Den Beweis führen wir indirekt.

Trifft man die Übereinkunft, Ziffernfolgen mit lauter Neunen auszuschließen, dann lässt sich jede reelle Zahl auf nur eine Weise im Dezimalsystem schreiben. Ließe sich nun die Menge aller reellen Zahlen zwischen 0 und 1 als Folge schreiben, dann können wir sofort eine Zahl bestimmen, die ebenfalls zwischen 0 und 1 liegt, der Folge jedoch nicht angehört. Als ihre erste Dezimale wählen wir eine Ziffer, verschieden von 9 und der ersten Dezimale der ersten Zahl in der Folge, als ihre zweite Dezimale eine Ziffer, verschieden von 9 und der zweiten Dezimale der zweiten Zahl in der Folge.

Dieses Verfahren ist unbegrenzt fortsetzbar und liefert die gewünschte Zahl, weil ihre n-te Dezimale von der n-ten Dezimale der n-ten-Zahl in der Folge verschieden ausfällt, so dass sie keiner Zahl in der Folge gleich ist, die Ziffer 9 aber in ihr überhaupt nicht vorkommt.

Denkt man sich dabei die Zahlen der Folge untereinandergeschrieben, dann bilden die abgeänderten Ziffern in den einzelnen Zahlen eine Art Diagonale, daher die Bezeichnung Diagonalverfahren.
Die Kardinalzahl der Menge aller Zahlen zwischen 0 und 1 bezeichnet man mit $\mathfrak{c}$, dem Anfangsbuchstaben des lateinischen, hier verdeutscht geschriebenen Wortes Kontinuum.



Das Kontinuum enthält die Folge 0,1, 0,11, 0,111, ... als echte Teilmenge, kann aber dieser nach der eben geführten Überlegung nicht äquivalent sein. Das legt folgende Erklärung nahe. Ist die Menge M einer echten Teilmenge N' von N äquivalent, aber N weder M noch einer Teilmenge von M äquivalent, dann heißt die Kardinalzahl $\mathfrak{m}$ von M kleiner als die Kardinalzahl $\mathfrak{n}$ von N, in Zeichen $\mathfrak{m} < \mathfrak{n}$.
Dieser Sachverhalt liegt vor, wenn M die Menge aller natürlichen Zahlen und N das Kontinuum bezeichnet, so dass $\mathfrak{a} < \mathfrak{c}$ gilt. Ob nun zwischen den beiden Kardinalzahlen noch weitere Kardinalzahlen liegen, ist eine Frage, die schon Cantor vorgelegt hat und die als das sogenannte Kontinuumproblem bis heute unbeantwortet blieb.

Wir wollen noch einen Schritt weitergehen und nachweisen, dass es zu jeder Kardinalzahl eine noch größere Kardinalzahl gibt. Das besagt dann soviel, dass es unendlich viele verschiedene unendliche Kardinalzahlen gibt. Um den angekündigten Nachweis zu führen, treffen wir noch die Übereinkunft, zu den Teilmengen einer Menge M auch noch diese selbst und außerdem die leere Menge, die kein Element enthält, zu zählen.

$\mathfrak{M}$ bezeichne dann die Menge aller Teilmengen von M. Eine echte Teilmenge von $\mathfrak{M}$ bilden die Mengen, die aus je einem Element von M bestehen. Ihre Gesamtheit ist offensichtlich

M äquivalent. Dagegen ist es unmöglich, $\mathfrak{M}$ auf M oder eine Teilmenge von M umkehrbar eindeutig abzubilden. Das wollen wir indirekt beweisen.

Wir nehmen an, es existiert eine eineindeutige Zuordnung, durch die jedes Element m aus M einem Element N aus $\mathfrak{M}$, d.h. einer Teilmenge aus M, zugeordnet ist.
In dieser Zuordnung wird es Paare geben, in denen m in der ihm zugeordneten Teilmenge N enthalten ist, und solche, bei denen das nicht der Fall ist. Wir suchen nun alle m heraus, die in ihrem Partner N nicht enthalten sind. Diese m bilden wieder eine Teilmenge von M, also wieder ein Element von $\mathfrak{M}$. Diese Teilmenge, wir wollen sie N' nennen, kann nicht leer sein, denn sie enthält mindestens dasjenige m, das der leeren Menge zugeordnet ist. Dieses N' kann durch die obige Zuordnung nicht erfasst sein. Wir können das folgendermaßen nachweisen:

Angenommen, es existiere ein Element m', das der Untermenge N' zugeordnet sei. Dann darf m' in N' nicht enthalten sein, weil N' ja laut Erklärung aus den Elementen von M besteht, die bei der Paarung in ihrem Partner, einer Untermenge von M, nicht enthalten sind. Wenn aber m' in N' nicht enthalten ist, würde N' nicht die Gesamtheit aller Elemente sein, die bei der Paarung in ihrem Partner nicht enthalten sind.
Dieser Widerspruch zeigt, dass es unmöglich ist, unter den Elementen m von M einen Partner für N' zu finden.

Ähnlich folgt, dass $\mathfrak{M}$ auch keiner echten Teilmenge von M äquivalent ist. Folglich ist die Kardinalzahl von $\mathfrak{M}$ größer als die Kardinalzahl von M.
Wir wissen zwar, wann von zwei verschiedenen Kardinalzahlen die eine größer als die andere ist, wissen aber nicht, ob sie auch immer in unserem Sinne vergleichbar sind. Um das zu erschließen, stellte bereits Cantor folgende Betrachtung an, die jedoch die Vergleichbarkeit zunächst nur plausibel macht.

Man paare aus den vorliegenden Mengen M und N je ein Element nach Belieben, danach je ein Element aus den Restmengen und denke sich das so weit fortgesetzt, wie es geht. Dadurch wird mindestens eine der beiden Mengen M und N erschöpft. Sollte nämlich dieses Vorgehen mit den echten Teilmengen M' und N' abbrechen, dann könnte man ein weiteres Element aus M, das nicht zu M' gehört, mit einem Element aus N, das nicht zu N' gehört, paaren und wäre damit doch noch über die angeblich letzten Mengen M' und N' hinausgelangt.

Im 99. Band der Zeitschrift "Acta mathematica" führte ich aus, wie man diese Betrachtung zu einem Beweis ausbauen kann. Dazu stellte ich ein Postulat auf, das sehr viel weniger fordert als das Auswahlaxiom von Zermelo aus dem Jahre 1904.
Dieser verlangt die prinzipielle Möglichkeit, aus jeder Gesamtheit von Mengen je ein Element zugleich auszuwählen, Während ich die Auswahl nur für jede einzeln vorliegende Menge fordere. Es gibt zu denken, dass Zermelo seinerzeit von führenden Mathematikern schlankweg missverstanden wurde. Zuzugeben ist dagegen, dass seine Forderung von bedenklicher Allgemeinheit ist, ein Umstand, der zur Vorsicht mahnen sollte.

Allzugroße Allgemeinheit von Begriffsbildungen hat sich als Gefahrenquelle für die Widerspruchsfreiheit erwiesen, insbesondere bei Mengen. Ein bedenkenloser Gebrauch des Wortes "alle" führt hier sofort zu einem unlösbaren Widerspruch.
Man erkennt das, wenn man bedenkt, dass die Elemente von ihrer Zusammenfassung zur Menge verschieden sind, eine Menge also sich als Element nicht enthalten kann. Eine "Menge aller Mengen" müsste dagegen laut Erklärung sich als Element enthalten!

Schon dem Altertum waren Widersprüche dieser Art bekannt. So ist die Selbstbezichtigung eines Kreters, dass alle Kreter Lügner sind, widerspruchsvoll. Denn auf ihn selbst bezogen, besagt es, dass nicht wahr ist, was er behauptete; folglich lügt er selbst nicht, d.h., es sind doch nicht alle Kreter Lügner.

Damit wurden Probleme angeschnitten, die an den Fundamenten des Denkens rütteln. In unserem Jahrhundert wurden verschiedene, reichlich komplizierte Lösungsversuche unternommen, denen der Erfolg leider versagt blieb. Einige Forscher sahen darin eine Krise der Mathematik. Heute gibt es eine ausgedehnte eigene Literatur darüber und Forscher, die ihre ganze Kraft dafür einsetzen. Widerspruchsfreiheit wurde durch untragbare Einschränkungen erkauft.

Würde man damit Ernst machen, dann müssten wohlbewährte Teile der Mathematik geopfert werden. Eine verfahrene Situation, aus der man vorerst keinen Ausweg kennt, sehr zu unserem Kummer!

5 Euklids Kette

Bereits in den ersten Schuljahren lernt man dividieren. Man begnügt sich aber damit, dass der eingeschlagene Weg zum Ziele führt. Erst später wird eine Begründung dafür nachgeholt. Diese ist wie folgt zu führen.
n und N seien natürliche Zahlen, von denen die erste kleiner sei als die zweite, $n < N$. Man bilde die aufeinanderfolgenden Vielfachen von n, also n, $2n$, $3n$, ..., bis man an das erste, kn, gerät, das N mindestens gleich ist, $kn \geq N$.

Dann liegt also N zwischen $(k-1)n$ und kn, $(k-1)n < N \leq kn$. Daher ist der Unterschied zwischen N und $(k-1)n$ kleiner als der Unterschied zwischen kn und $(k-1)n$, in Zeichen $N-(k-1)n < kn-(k-1)n = n$.
Der nichtnegative ganzzahlige Unterschied $N-(k-1)n$ wird mit r bezeichnet und heißt Rest. Schreibt man für $k-1$ noch q, den Anfangsbuchstaben des Wortes Quotient, dann folgt

$$N = nq + r \qquad \text{mit} \qquad O \leq r < n$$

die Formel für die eindeutige Ausführbarkeit der Division.

Wir haben es vorhin, ohne uns weiter Gedanken zu machen, hingenommen, dass die Vielfachen von n schließlich N überbieten. Für Zahlen trifft das ja zu, es gibt aber auch solche Größen, für die das nicht mehr gilt. Eine Spielart solcher Größen bilden die Ausdrücke $ax + bx^2$. Zunächst wollen wir für sie die Gleichheit erklären und dann festlegen, welche von zwei verschiedenen solchen Größen die kleinere sein soll.

Gleich sind zwei Ausdrücke $ax + bx^2$ und $cx + dx^2$ genau dann, wenn $a = c$ und $b = d$ gilt. Ist dagegen $a < c$, dann gilt der erste Ausdruck kleiner als der der zweite, ungeachtet dessen, ob b kleiner, gleich oder größer als d ausfällt. Anders, wenn $a = c$ ist. Dann gilt der erste Ausdruck genau dann kleiner als der zweite, wenn $b < d$ ist. Auf diese Weise sind zwei beliebige Ausdrücke dieser Art vergleichbar.

Betrachtet man insbesondere die Werte $a = 0$ und $b = 1$, d.h. den Ausdruck x^2, dann kann man diese Größe beliebig vervielfachen, ohne jemals die Größe zu erreichen oder zu überbieten, die den Werten $c = 1$ und $d = 0$ entspricht, also dem Ausdruck x, ganz im Sinne unserer Festlegung.
Die Forderung, dass für die Größen aus einem System mit $g < g'$ zugleich stets $g' < kg$ mit einer geeigneten natürlichen Zahl k erfüllbar ist, braucht also keineswegs immer zuzutreffen, d.h., sie ist durchaus nicht selbstverständlich.
Wenn diese Forderung erfüllt ist, sagt man, das Größensystem genügt dem Axiom des Archimedes, obschon es richtiger wäre, dieses Axiom nach Eudoxos zu benennen, der es schon anderthalb Jahrhundert früher kannte.

Kehren wir zur Division von natürlichen Zahlen zurück. Für $n = 1$ ist $q = N$ und $r = 0$, weiter für $n = N$ ist $q = 1$ und $r = 0$. Das besagt, dass jede natürliche Zahl durch die

Einheit wie auch durch sich selbst teilbar ist. Besitzt sie keine weiteren Teiler, keine echten Teiler, dann heißt sie Primzahl. 7 ist eine Primzahl, 6 dagegen nicht, weil 2 und 3 Teiler von 6 sind.
Primzahlen können als Bausteine der natürlichen Zahlen aufgefasst werden. Jede natürliche Zahl lässt sich als Produkt von Primzahlen, den sogenannten Primfaktoren, schreiben. Liegt eine Zahl vor, die keine Primzahl ist, dann ist sie das Produkt von zwei Zahlen, die kleiner sind als die Ausgangszahl. Sind die beiden Faktoren Primzahlen, dann sind wir fertig. Sonst wiederholen wir die Überlegung an den einzelnen Faktoren.
Die Ausgangszahl erscheint dabei als Produkt von Faktoren, die bei jedem Schritt kleiner werden. Schließlich lassen sie sich nur noch in die Einheit und sich selbst zerlegen, d.h., sie sind Primzahlen.

Jede natürliche Zahl lässt sich auf nur eine Weise als Produkt von Primfaktoren schreiben, wenn man die Einheit nicht zu den Primzahlen rechnet. Für diese Eindeutigkeit der Zerlegung in Primfaktoren gab in der Hauptsache schon Euklid einen Beweis, wir aber wollen einen eigenen Beweis vortragen, der einfacher ist. Dabei berufen wir uns nur auf die anfangs behandelte Division mit Rest und führen den Beweis indirekt.

Würde es Zahlen mit verschiedenen Zerlegungen in Primfaktoren geben, dann sei m die kleinste unter diesen Zahlen. q bezeichne den kleinsten Primfaktor aus den Zerlegungen von m. Zwei verschiedene Zerlegungen von m können keinen gemeinsamen Primfaktor enthalten, weil man durch diesen gekürzt, zwei verschiedene Zerlegungen einer Zahl kleiner als m erhielte, im Widerspruch zur Bestimmung von m als kleinster Zahl mit verschiedenen Zerlegungen. In der einen Zerlegung von m darf also q nicht auftreten. Die einzelnen Primfaktoren darin durch q dividiert, ergeben Reste, die sämtlich kleiner als q ausfallen. Das würde dann so aussehen:

$$m = (n_1 q + r_1)(n_2 q + r_2)...(n_i q + r_i) = n_1 n_2 ... n_i q^i + ... + r_1 r_2 ... r_i$$

Da sowohl m als auch rechts jeder Summand bis auf den letzten durch q teilbar ist, muss q auch den letzten Summanden teilen, so dass es für den eine Zerlegung gibt, in der q als Primfaktor auftritt. Andererseits sind aber alle Reste kleiner als q, so dass in ihren Zerlegungen q als Primfaktor nicht auftreten kann, daher auch nicht im Produkt; das würde bedeuten, dass es für $r_1 r_2 ... r_i$ zwei verschiedene Zerlegungen gibt.
Da aber $r_1 r_2 ... r_i < m$ ist, steht das im Widerspruch zu unserer Voraussetzung, dass m die kleinste Zahl mit verschiedenen Zerlegungen ist.

Von grundsätzlicher Bedeutung für die Teilbarkeitslehre sind gewisse Divisionen mit Rest, die sich ähnlich den Gliedern einer Kette aneinanderreihen. Sie heißen Euklidischer Algorithmus.
Wir gehen von zwei natürlichen Zahlen n und N aus, von denen die erste kleiner sein möge als die zweite. Dann ist im Sinn der Division mit Rest

$$N = nq + n_1$$

wobei wir den Rest diesmal mit n_1 bezeichnen. Weiter ist

$$n = n_1 q_1 + n_2$$

n_2 ist der Rest bei dieser zweiten Division. Wird das fortgesetzt, nehmen die aufeinanderfolgenden Reste ständig ab. Da sie nichtnegative ganze Zahlen sind, muss schließlich die

Null als letzter Rest auftreten. Das möge bei der $(\nu+1)$-ten Division der Fall sein; nach dieser ergibt sich

$$n_{\nu-1} = n_\nu q_\nu + n_{\nu+1} \qquad \text{und} n_{\nu+1} = 0$$

Etwas ausführlicher lautet unsere Gleichungskette

$$\begin{aligned} N &= nq + n_1 \\ n &= n_1 q_1 + n_2 \\ &\dots \\ n_{\nu-2} &= n_{\nu-1} q_{\nu-1} + n_\nu \\ n_{\nu-1} &= n_\nu q_\nu \end{aligned}$$

Aus der letzten Gleichung kann man ablesen, dass n_ν Teiler von $n_{\nu-1}$ ist, aus der vorletzten Gleichung folgt deshalb, dass n_ν auch noch $n_{\nu-2}$ teilt. Fahren wir fort, von unten nach oben in der gleichen Weise von Gleichung zu Gleichung zu schließen, so folgt, dass n_ν Teiler von n und auch von N ist.

Auch die Umkehrung davon gilt, dass nämlich jeder gemeinsame Teiler von n und N zugleich n_ν teilt. Zunächst schließt man aus der ersten Gleichung, dass er n_1 teilt, dann aus der zweiten Gleichung, dass er auch noch n_2 teilt. Geht man Schritt für Schritt nach unten weiter, so folgt, dass er schließlich auch noch n_ν teilt.

Beide Einsichten vereint ergeben, dass n_ν größter gemeinsamer Teiler von n und N ist.

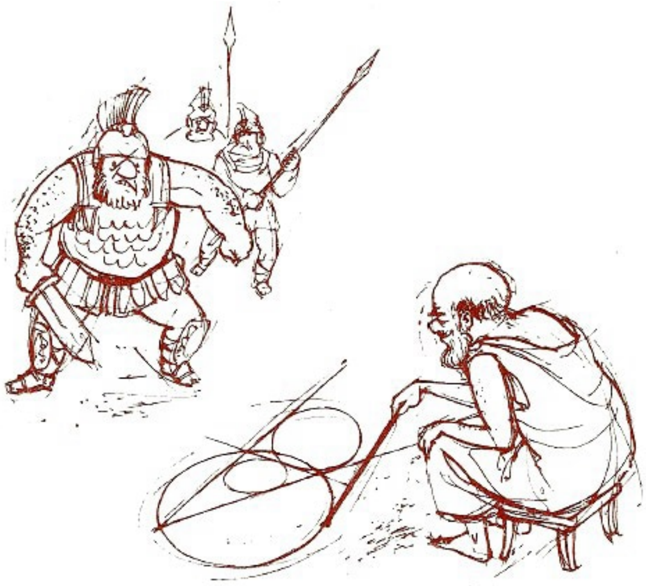
Wir nannten Eudoxos, Euklid und Archimedes. Diese waren die drei bedeutendsten griechischen Mathematiker. Eudoxos war ein Vorgänger von Euklid, diesem an Originalität weit Überlegen. Er wurde im Jahre 408 v.u.Z. geboren und war so arm, dass er als Schüler von Plato mit 23 Jahren im Hafen von Athen wohnte und bis zur Unterrichtsstätte täglich 8 Kilometer Fußmarsch zurückzulegen hatte. Nach dieser Zeit unternahm er ausgedehnte Studienreisen und gelangte so auch nach Ägypten, wo er 16 Monate zubrachte. Dann wurde er in seiner Heimat, der Insel Knidos, mit gesetzgeberischer Gewalt betraut.

Seiner Zeit war er in allem weit voraus und erkannte als erster den Betrug und Selbstbetrug chaldäischer Astrologen.

Sein Lebenswerk galt der Erforschung tiefliegender Fragen der Mathematik; er beschäftigte sich mit einem Gebiet, das wir heute zu den Stetigkeitsuntersuchungen rechnen würden. Er starb mit 53 Jahren.

Über das Leben des Euklid wissen wir wenig. Er wirkte um 300 v.u.Z. in Alexandrien, dem damaligen wissenschaftlichen Zentrum. Euklid versuchte, das mathematische Wissen seiner Zeit in seinem Werk ”Elemente” darzustellen.
Darin wollte er von einigen wenigen Grundwahrheiten, den Axiomen, ausgehen und daraus alle anderen Sätze folgern. Es ist ein imposantes Werk, dessen Verfasser als Lehrmeister von zwei Jahrtausenden galt; denn die ”Elemente” spielten bis Mitte des 19.Jh. die Rolle des Grundlehrgangs der Mathematik.

Archimedes kam im Jahre 212 v.u.Z. bei der Eroberung der Stadt durch die Römer um. Wie die Anekdote sagt, starb der Fünfundsiebzigjährige durch das Schwert eines plündernden römischen Soldaten.
Jahrelang könnte er seine Heimatstadt Syrakus durch Erfindung von Kriegsmaschinen gegen die Römer verteidigen. Diese Erfindungen waren Anwendungen seiner Kenntnisse der Mechanik. Zweifellos war Archimedes der hervorragendste Mathematiker der Antike; er war bereits bis zu Fragestellungen vorgedrungen, die wir heute zur Integralrechnung zählen.



6 Spröde Zahlen

Die wichtige Rolle der Primzahlen haben wir bereits erkannt. Zunächst kann man nach ihrer Anzahl fragen. Schon Euklid wusste, dass es unendlich viel Primzahlen gibt. Dafür genügt der Nachweis, dass es zu jeder Primzahl p eine größere Primzahl q gibt. Wir bilden das Produkt aller Primzahlen bis p inklusive und addieren die Einheit,

$$2 \cdot 3 \cdot 5 \cdot ... \cdot p + 1$$

Entweder ist die so erhaltene natürliche Zahl Q eine Primzahl und dann größer als p, oder sie ist eine zusammengesetzte Zahl. Dann ist sie das Produkt von mindestens zwei Primfaktoren q und q'.

Diese Primfaktoren sind aber von den Primzahlen $2, 3, 5, ..., p$ verschieden, weil Q durch irgendeine dieser Primzahlen dividiert, den Rest 1 gibt. Außer den Primzahlen $2, 3, 5, ..., p$ gibt es aber keine Primzahlen, die $\leq p$ ausfallen, daher ist $q > p$.
Man kann diese Aussage leicht verschärfen. Es gibt nicht nur unter allen natürlichen Zahlen unendlich viele Primzahlen, sondern auch unter den Zahlen, die durch 4 dividiert, den Rest 3 haben, also von der Form $4n + 3$ sind. Diesmal gehen wir vom Ausdruck

$$2^2 \cdot 3 \cdot 5 \cdot ... \cdot p - 1$$

aus, den man auch als

$$4(3 \cdot 5 \cdot ... \cdot p - 1) + 3 = 4n + 3$$

schreiben kann, der folglich zu der gewünschten Spielart von Zahlen gehört. An der ursprünglichen Schreibweise erkennt man unmittelbar, dass diese Zahl durch keine der Primzahlen $2, 3, 5, ..., p$ teilbar ist. Die weitere Überlegung verläuft dann wie vorhin.

Man kann noch weitergehen und mit Dirichlet behaupten, dass es unter den Zahlen von der Form $an+b$ oder aber $an-b$ unendlich viele Primzahlen gibt, wenn der größte gemeinsame Teiler von a und b die Einheit ist. Die Folgen der Form $an \pm b$ heißen arithmetische Folgen. Die beiden vorhin behandelten Fälle waren ganz spezielle arithmetische Folgen. Der Beweis für den Satz von Dirichlet erfolgt durch konsequente Anwendung der Analysis in der Zahlentheorie. Probleme dieser Art bilden ein besonderes Gebiet der Zahlentheorie, die analytische Zahlentheorie.

Wenn es unendlich viele Primzahlen gibt, dann entsteht die weitere Frage, wie man sie bestimmt. Eine naheliegende Art, im Prinzip alle Primzahlen zu bestimmen, führt zum Sieb des Eratosthenes.
Man denke sich die Folge der natürlichen Zahlen hingeschrieben und streiche, zunächst von 2 ausgehend, jede zweite Zahl, also 4, 6, 8, ..., dann von 3 ausgehend, jede dritte Zahl, also 6, 9, 12, ..., dann von 5 ausgehend, jede fünfte Zahl, also 10, 15, 20, usw.
Die 4 wurde als Ausgangszahl übergangen, weil sie als Vielfaches von 2 bereits gestrichen war. Bei diesem Vorgehen werden Zahlen mitunter wiederholt gestrichen, wie die 6. Übrig bleiben augenscheinlich nur die Primzahlen 2, 3, 5, ..., während die gestrichenen Zahlen Vielfache von Primzahlen, also zusammengesetzt sind.

Etwas anderes ist es, wenn man von einer vorgelegten Zahl n entscheiden soll, ob sie eine Primzahl ist oder nicht. Da sie im zweiten Fall das Produkt von mindestens zwei Faktoren ist, die beide größer als 1 ausfallen, muss wenigstens ein Faktor $\leq \sqrt{n}$ sein. Es genügt folglich, durch alle Primzahlen zu dividieren, die $\sqrt{n}$ nicht übersteigen. Verschwindet dabei kein Rest, dann ist n eine Primzahl, sonst nicht.

Der Ausdruck $4n+3$ liefert zwar, wie wir erkannten, unendlich viele Primzahlen, aber auch zusammengesetzte Zahlen; z.B. nimmt er für $n=3$ den Wert $15 = 3 \cdot 5$ an.
Goldbach erkannte 1732, dass - viel allgemeiner - ein Polynom mit ganzzahligen Koeffizienten nicht lauter Primzahlen als Werte annehmen kann. Es sei also $f(n)$ ein Polynom in n mit ganzzahligen Koeffizienten. Für natürliche Zahlen n nimmt es offenbar lauter ganzzahlige Werte an, die aber auch negativ ausfallen können.
Wir fordern, dass $f(n)$ für große Werte von n positiv sei. Dann muss der Koeffizient der höchsten Potenz von n selbst positiv sein, denn es ist

$$f(n) = a_k n^k + a_{k-1} n^{k-1} + ... + a_1 n + a_0 = n^k \left(a_k + \frac{a_{k-1}}{n} + ... + \frac{a_1}{n^{k-1}} + \frac{a_0}{n^k} \right)$$

Alle Summanden in der Klammer werden bis auf den ersten für hinreichend große n beliebig klein, so dass das Vorzeichen von a_k das Vorzeichen von $f(n)$ bestimmt; daher muss a_k positiv sein. Schließlich nimmt $f(n)$ mit n zu.
Wir wollen nun ein Polynom bilden, dessen Wert keine Primzahl ist. Dazu setzen wir $f(n) = N$ und betrachten $f(nN+n)$. Es gilt

$$f(n) < f(nN+n) = a_k(nN+n)^k + ... + a_1(nN+n) + a_0$$

Löst man die Klammern auf, dann entsteht eine Summe, die so umgeordnet werden kann, dass die letzten Glieder gerade die Summe

$$a_k n^{k-1} + ... + a_0$$

darstellen, alle Summanden davor aber durch N teilbar sind:

$$N \left(a_k n^k N^{k-1} + ... + a_1 n \right)$$

Da die vorhergehende Summe den Wert N besitzt, ergibt sich

$$f(nN+n) = N \left(a_k n^k N^{k-1} + ... + a_1 n \right) + N$$

so dass das Polynom für $nN + n$ durch N teilbar ist.

Auf der Suche nach Ausdrücken, die lauter Primzahlen liefern, vermutete Fermat 1637, dass

$$2^{2^n} + 1$$

stets eine Primzahl darstellt. Euler wies jedoch 1732 nach, dass für $n = 5$ dieser Ausdruck $641 \cdot 6700417$ ist. Heute weiß man, dass er auch noch für

$$n = 7, 8, 9, 11, 12, 15, 18, 23, 36, 88, 73$$

zusammengesetzte Zahlen darstellt. Offen ist die Frage, ob unter diesen Fermatschen Zahlen nur endlich viele Primzahlen sind.

Erst 1947 gelang Mills der Nachweis, dass es eine reelle Zahl a gibt, für die die größte ganze Zahl, die im Ausdruck

$$a^{3^n}$$

enthalten ist, immer eine Primzahl darstellt. Eine Variante dieses Problems erhält man, wenn man bei einem vorgelegten Ausdruck nach den Werten fragt, für die der Ausdruck eine Primzahl darstellt. Es möge sich um den Ausdruck

$$m^4 + 4n^4$$

handeln. Wir fragen, für welche natürlichen Zahlen m und n der Ausdruck eine Primzahl darstellt. Dafür formen wir den Ausdruck wie folgt um:

$$m^4 + 4n^4 = (m^2 + 2n^2)^2 - (2mn)^2 = (m^2 + 2n^2 + 2mn) \cdot (m^2 + 2n^2 - 2mn)$$

Der zweite Faktor rechts lässt sich weiter als $(m-n)^2 + n^2$ schreiben. Da der erste Faktor rechts stets größer als 1 ausfällt, kann unser Ausdruck nur dann eine Primzahl darstellen, wenn der zweite Faktor gleich 1 ausfällt, was nur für das Wertepaar $m = n = 1$ zutrifft. Mögen auch Kunstgriffe dieser Art in Einzelfällen zum Erfolg führen, so bedeuten sie für die Zahlentheorie selbst keine Bereicherung.

Da es, wie wir wissen, unendlich viele Primzahlen gibt, hat man nach Gesetzmäßigkeiten zu suchen, nach denen die Primzahlen aufeinanderfolgen. Zunächst stellen wir fest, dass es Lücken gibt, die beliebig groß ausfallen. Genauer formuliert, gibt es in der Folge der natürlichen Zahlen immer wieder beliebig lange Abschnitte, die aus lauter zusammengesetzten Zahlen bestehen.
Es seien wieder $2, 3, 5, ..., p$ alle Primzahlen bis p inklusive. Dann bilden die insgesamt $p - 1$ aufeinanderfolgenden Zahlen

$$2 \cdot 3 \cdot 5 \cdot ... \cdot p + k$$

einen solchen Abschnitt, wenn k nacheinander die Werte $2, 3, 4, ..., p$ annimmt. Der erste Summand ist nämlich durch jede der Primzahlen $2, 3, 5, ..., p$ teilbar. Der zweite Summand, k, ist aber entweder selbst eine Primzahl und dann, wegen $k \leq p$ eine der Primzahlen $2, 3, 5, ..., p$, oder zusammengesetzt und dann, ebenfalls wegen $\leq p$, ein Produkt aus Primfaktoren $\leq p$. Daher ist die Summe selbst jedenfalls durch eine der Primzahlen $2, 3, 5, ..., p$ teilbar.

Die nächste Frage lautet, ob es eine Formel gibt, die wenigstens angenähert die Anzahl der Primzahlen bis zu einer nach Belieben vorgegebenen natürlichen Zahl angibt. Bereits Legendre und Gauß hatten vermutet, dass die größte in $\frac{n}{\ln n}$ enthaltene natürliche Zahl mit zunehmendem n immer genauer diese Anzahl liefert, die Anzahl der Primzahlen also, die nicht größer als n sind.

Freilich erfolgt die Annäherung an die wirkliche Anzahl recht zögernd, wie aus dem Vergleich der wirklichen Werte für $n = 10^3, 10^6$ und 10^9: $168, 78498, 50847478$ mit den entsprechenden Werten aus der Formel $145, 72382, 48254942$ ersichtlich ist.

Den Nachweis für die Richtigkeit der, Vermutung von Legendre und Gauß erbrachten erst 1896 voneinander unabhängig der Franzose Hadamard und der Belgier de la Vallée Poussin. 1949 gelang schließlich Selberg ein Nachweis mit mehr elementaren Mitteln.

Wir wissen, dass eine natürliche Zahl entweder eine Primzahl oder ein Produkt von Primfaktoren ist. Man sollte sich aber auch für Darstellungen durch Summen interessieren, etwa durch Summen von Primzahlen. Es ist nun sehr wahrscheinlich, wenn es auch bisher noch immer nicht nachgewiesen werden konnte, dass sich jede gerade Zahl als Summe von zwei Primzahlen schreiben lässt.
Die Eindeutigkeit dafür gilt dann bestimmt nicht, denn es ist ja zum Beispiel $24 = 7+17 = 11 + 13$. Diese Vermutung geht auf Goldbach zurück; in einem Brief vom 7. 6.1742 teilte er sie Euler mit, der ihr dann obige prägnantere Fassung gab.

In diesem Abschnitt fielen gar viele Namen. Die beiden bedeutendsten seien dem Leser vorgestellt.

Euler war einer der größten Mathematiker des 18. Jahrhunderts und gewiss der produktivste. Er stammt aus Basel und wurde schon mit zwanzig Jahren 1727 an die Petersburger Akademie berufen. Übermäßiges angestrengtes Arbeiten führte 1735 zur Erblindung des rechten Auges. 1741 erhielt er eine Berufung als Direktor der mathematischen Klasse an die Berliner Akademie.
Die Spannungen mit dem selbstherrlichen Preußenkönig Friedrich II. führten dazu, dass er 1766 zum zweiten Male an die Petersburger Akademie ging. Dort war er bis zum Ende seines Lebens dank seines phänomenalen Gedächtnisses trotz völliger Erblindung weiterhin sehr produktiv tätig auf allen Gebieten der reinen und angewandten Mathematik.

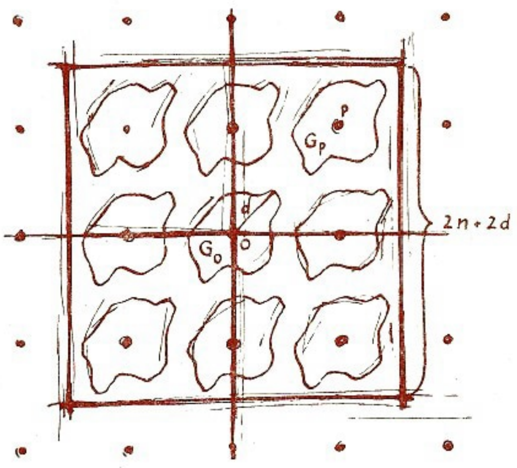
Die Begabung von Gauß erregte schon in seinen jüngsten Jahren Aufsehen. Mit noch nicht dreißig Jahren galt er schon als führender Mathematiker. Als der alte Lagrange in Paris einmal befragt wurde, wer der größte Mathematiker Deutschlands wäre, gab er den Doktorvater von Gauß an; als man ihn etwas verwundert an Gauß erinnerte, antwortete er, Gauß wäre der größte Mathematiker der ganzen Welt.
Gauß publizierte sehr bedachtsam und teilte manches aus Scheu vor Polemik nur brieflich mit. Das brachte ihn bei einigen fundamentalen Entdeckungen um die Priorität.

7 Geometrie der Zahlen

Im Jahre 1896 entwickelte der Göttinger Mathematiker Minkowski eine geometrische Methode, um Probleme der Zahlentheorie zu behandeln, bei denen andere Methoden nicht zum Ziele führten.

Wir gehen von einem kartesischen Koordinatensystem in der Ebene aus und betrachten die Punkte, deren beide Koordinaten ganze Zahlen sind. Sie bilden ein Gitter. Der Nullpunkt des Koordinatensystems möge im Gebiet G_0 liegen.
Unter einem Gebiet versteht man eine Punktmenge, für die folgende zwei Bedingungen erfüllt sind. Erstens kann um jeden Punkt von G_0 ein Kreis geschlagen werden, dessen ganze Fläche aus Punkten von G_0 besteht. Zweitens sollen zwei beliebige Punkte aus G_0 immer durch einen Streckenzug verbunden werden können, der ganz aus Punkten von G_0 besteht.

G_0 soll parallel zur x- und zur y-Achse verschoben werden, so dass der Punkt $(0{,}0) = O$ in den Gitterpunkt P übergeht. In der neuen Lage werde G_0 durch G_P bezeichnet.
Wenn P alle Gitterpunkte durchläuft und die G_P dabei paarweise disjunkt sind, mit anderen Worten, keine gemeinsamen Punkte haben, dann ist die Fläche von G_0 nicht größer als 1.

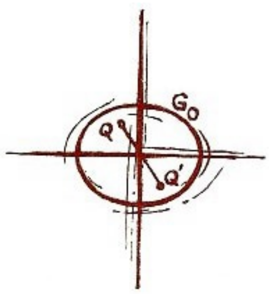


Um den Nachweis zu führen, bemerke man, dass es unter den Randpunkten von G_0 mindestens einen mit maximaler Entfernung d von 0 gibt.
Betrachtet man die Gebiete G_P, für die beide Koordinaten von P dem Absolutwert nach die natürliche Zahl n nicht übersteigen, dann liegen alle diese G_P augenscheinlich in einem achsenparallelen Quadrat mit dem Mittelpunkt O und der Seitenlänge $2(n + d)$.
Insgesamt sind es $(2n + 1)^2$ Gebiete G_P, das Gebiet G_0 dazugerechnet. Da die G_P sich nicht überlappen sollen, übersteigt die von ihnen überdeckte Fläche nicht die Fläche des großen Quadrats, d.h.

$$(2n + 1)^2 \cdot F \leq (2n + 2d)^2$$

den Flächeninhalt von G_0 dabei durch F bezeichnet. Nach kurzer Rechnung folgt daraus

$$F \leq \left(1 + \frac{d - \frac{1}{2}}{n + \frac{1}{2}}\right)^2$$

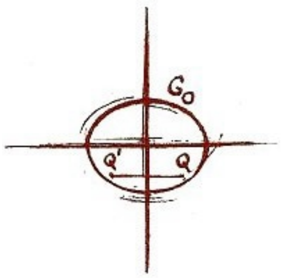


Mit zunehmendem n wird der zweite Summand in der Klammer beliebig klein, daher gilt $F \leq 1$.

Auf Grund dieser Einsicht sind wir in der Lage, einen Satz von Minkowski zu beweisen, der den Schlüssel zu seinen Überlegungen bildet. Um diesen Satz zu formulieren, ist G_0 noch zwei weiteren Einschränkungen zu unterwerfen.

Zunächst möge G_0 symmetrisch um 0 sein, das heißt, wenn Q ein beliebiger Punkt aus

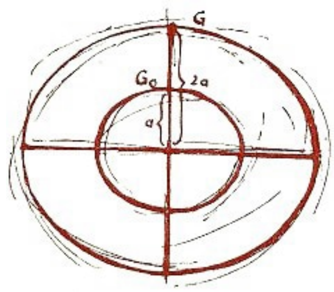
G_0 ist, soll der Punkt Q', der auf der Verbindungsgeraden QO auf der anderen Seite von 0 im gleichen Abstand wie Q liegt, ebenfalls G_O angehören.



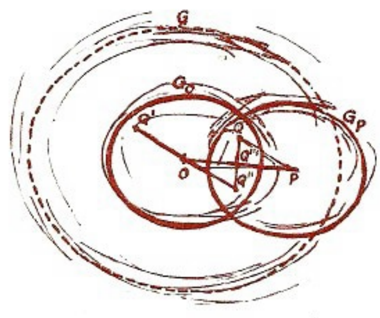
Q' heißt Spiegelbild von Q in bezug auf 0. Dann aber soll G_0 auch noch konvex sein, das heißt, wenn zwei beliebige Punkte Q und Q' dem Gebiet G_0 angehören, dann soll die ganze Strecke QQ' aus lauter Punkten von G_0 bestehen.
Für konvexe Gebiete G, die um 0 symmetrisch sind, gilt dann der grundlegende Satz von Minkowski, dass jedes solche Gebiet G, wenn nur sein Inhalt größer als 4 ausfällt, mindestens einen Gitterpunkt enthält.

Um den Nachweis zu führen, betrachte man das Gebiet G_0, das aus G entsteht, wenn man jeden Punkt von G nach O hin um die Hälfte seiner ursprünglichen Entfernung von O verschiebt. Dadurch schrumpft der Inhalt von G auf den vierten Teil zusammen. Man kann sich nämlich G mit beliebiger Genauigkeit aus gleichseitigen Dreiecken mit den Spitzen in O aufgebaut denken, deren Höhen wie Seiten sich beim Schrumpfen auf die Hälfte reduzieren.

Da der Inhalt von G_0 so nach größer ist als 1, muss es, wie wir vorhin erkannt haben, ein G_P geben, mit dem G_0 einen gemeinsamen Punkt Q besitzt. Da G_P aus G_0 durch Parallelverschiebung entsteht, muss der dem Punkt Q in G_P entsprechende Punkt Q' in G_0 liegen.
Deshalb ist die Strecke OQ' parallel der Strecke PQ. Da weiter G_0 symmetrisch um O ist, gehört das Spiegelbild Q'' von Q' in bezug auf O zu G_0.



G_0 ist aber auch konvex, so dass die ganze Strecke QQ'' zu G_0 gehört, insbesondere auch deren Mittelpunkt. Dieser Mittelpunkt Q''' ist zugleich Mittelpunkt der Strecke OP. Da G aus G_0 durch Verdoppeln aller Strecken OR hervorgeht, wenn R ein beliebiger Punkt aus G_0 ist, enthält G insbesondere den Punkt P, der aus OQ''' durch Verdoppeln entsteht.



8 Kettenbrüche

Der Algorithmus von Euklid verhilft zu einer Darstellung der Brüche in Form von Kettenbrüchen. Aus der ersten Gleichung des Algorithmus,

$$N = nq + n_1 \qquad \text{folgt} \qquad \frac{N}{n} = q + \frac{n_1}{n} = q + \frac{1}{\frac{n}{n_1}} \tag{*}$$

aus der zweiten Gleichung

$$n = n_1 q_1 + n_2 \qquad \text{weiter} \qquad \frac{n}{n_1} = q_1 + \frac{n_2}{n_1}$$

Diesen Wert für $\frac{n}{n_1}$ setzt man in die Gleichung (*) ein; dann lautet diese

$$\frac{N}{n} = q + \frac{1}{q_1 + \frac{n_2}{n_1}} = q + \frac{1}{q_1 + \frac{1}{\frac{n_1}{n_2}}} \tag{**}$$

Aus der dritten Gleichung im Algorithmus

$$n_1 = n_2 q_2 + n_3 \qquad \text{folgt} \qquad \frac{n_1}{n_2} = q_2 + \frac{n_3}{n_2}$$

Setzt man diesen Wert in die Gleichung (**) ein, so lautet diese

$$\frac{N}{n} = q + \frac{1}{q_1 + \frac{1}{q_2 + \frac{n_3}{n_2}}} = q + \frac{1}{q_1 + \frac{1}{q_2 + \frac{1}{\frac{n_2}{n_3}}}}$$

so weiterschließend, gewinnt man die Darstellung

$$\frac{N}{n} = q + \frac{1}{q_1 + \frac{1}{q_2 + \ddots + \frac{1}{q_\nu}}}$$

Es ist üblich, die Bruchstriche zu verkürzen,

$$\frac{N}{n} = q + \frac{1}{q_1 + \frac{1}{q_2 + \ddots + \frac{1}{q_\nu}}}$$

Ein solcher Kettenbruch gibt Anlass, Näherungsbrüche zu bilden

$$q + \frac{1}{q_1} = \frac{qq_1 + 1}{q}, \qquad q + \frac{1}{q_1 + \frac{1}{q_2}} = \frac{qq_1q_2 + q + q^2}{q_1q_2 + 1}, \ldots$$

die der Reihe nach durch

$$\frac{P_1}{Q_1}, \quad \frac{P_2}{Q_2}, \quad \ldots, \quad \frac{P_\nu}{Q_\nu} = \frac{N}{n}$$

bezeichnet werden. Wir wollen das Bildungsgesetz der Näherungsbrüche untersuchen.

Zunächst bemerken wir: Wenn man in dem Näherungsbruch P_{k_1}/Q_{k_1} für q_{k_1} den Ausdruck

$q_{k-1} + \dfrac{1}{q_k}$ einsetzt, entsteht gerade der nächstfolgende Näherungsbruch. Das auf $k = 3$ angewandt, folgt

$$\frac{P_3}{Q_3} = \frac{qq_1\left(q_2 + \frac{1}{q_3}\right) + q + \left(q_2 + \frac{1}{q_3}\right)}{q_1\left(q_2 + \frac{1}{q_3}\right) + 1} = \frac{q_3(qq_1q_2 + q + q_2) + (qq_1 + 1)}{q_3(q_1q_2 + 1) + q_1} = \frac{q_3P_2 + P_1}{q_3Q_2 + Q_1}$$

Ähnlich vorgehend, findet man allgemein

$$P_k = q_kP_{k-1} + P_{k-2} \qquad \text{und} \qquad Q_k = q_kQ_{k-1} + Q_{k-2}$$

Damit folgt dann weiter

$$\begin{aligned} P_kQ_{k-1} - P_{k-1}Q_k &= (q_kP_{k-1} + P_{k-2})Q_{k-1} - P_{k-1}(q_kQ_{k-1} + Q_{k-2}) \\ &= -(P_{k-1}Q_{k-2} - P_{k-2}Q_{k-1}) \end{aligned}$$

In gleicher Weise ergibt sich

$$\begin{aligned} P_kQ_{k-1} - P_{k-1}Q_k &= -(P_{k-1}Q_{k-2} - P_{k-2}Q_{k-1}) \\ &= -[(q_{k-1}P_{k-2} + P_{k-3})Q_{k-2} - P_{k-2}(q_{k-1}Q_{k-2} + Q_{k-3})] \\ &= +(P_{k-2}Q_{k-3} - P_{k-3}Q_{k-2}) \end{aligned}$$

Diesen Schluss kann man nun so lange wiederholen, bis man

$$P_kQ_{k-1} - P_{k-1}Q_k = (-1)^k(P_2Q_1 - P_1Q_2)$$

erhält. Setzt man für P_1, P_2, Q_1, Q_2 die entsprechenden Werte in den q, q_1, q_2 ein, so ergibt sich

$$P_2Q_1 - P_1Q_2 = -1 \qquad \text{und damit} \qquad P_kQ_{k-1} - P_{k-1}Q_k = (-1)^{k-1}$$

Nun nehmen wir an, N und n seien teilerfremd. Das lässt sich ohne weiteres erreichen, indem man N und n durch den größten gemeinsamen Teiler dividiert. Dann ist $N = P_\nu$, und $n = Q_\nu$, und wir können schreiben

$$NQ_{\nu-1} - nP_{\nu-1} = (-1)^{\nu-1}$$

Dieses Ergebnis kann man verwenden, um ganzzahlige Lösungen unbestimmter Gleichungen mit zwei Unbekannten zu erhalten, denn es sind

$$Q_{\nu-1} = x' \qquad \text{und} \qquad P_{\nu-1} = y'$$

ganzzahlige Lösungen der Gleichung

$$Nx' - ny' = (-1)^{\nu-1} \tag{*}$$

Um folglich ein ganzzahliges Lösungspaar x und y der unbestimmten Gleichung

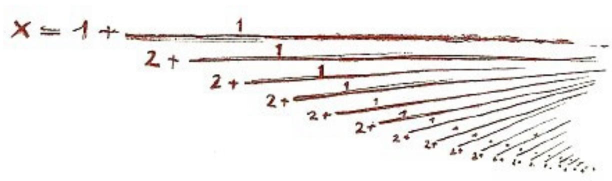
$$ax + by = c \tag{**}$$

zu gewinnen, wenn a, b und c beliebige ganze teilerfremde Zahlen sind, setze man $|a| = N$, $|b| = n$.

Danach führe man mit diesen Zahlen den Euklidischen Algorithmus durch. Man erhält dabei ν sowie die Werte $Q_{\nu-1} = x'$ und $P_{\nu-1} = y'$. Nun kann man x und y mit Hilfe der folgenden Gleichungen ausrechnen:

$$x = (-1)^{\nu-1} c x' \frac{|a|}{a} \qquad ; \qquad y = (-1)^{\nu} c y' \frac{|b|}{b}$$

denn diese Werte in die Gleichung (**) eingesetzt, führen auf Gleichung (*) zurück. Sind a, b und c nicht teilerfremd, so dividiert man sie vor dem geschilderten Verfahren durch den gemeinsamen Faktor. Haben nur a und b einen gemeinsamen Faktor, der in c nicht enthalten ist, dann ist die Gleichung widersprüchlich und besitzt keine Lösung.



Man erkennt so, wie ein Kettenbruch zu einem ganzzahligen Lösungspaar der allgemeinen unbestimmten Gleichung führt, die nach Diophant bezeichnet wird, obschon dieser auch Lösungen in rationalen Zahlen zuließ.
Wir bemerken noch, dass mit dem gefundenen Lösungspaar x und y zugleich alle Paare $x + gb$ und $y - ga$ lauter ganzzahlige Lösungen darstellen, wenn g eine beliebige ganze Zahl bezeichnet, wie man sich davon durch Einsetzen sofort überzeugt.

Kettenbrüche, sogar unendliche Kettenbrüche, wurden bereits im Altertum betrachtet. Insbesondere beschäftigte die Pythagoräer der unendliche Kettenbruch

$$1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2} \ddots}}$$

Selbstverständlich war man zu dieser Zeit noch nicht in der Lage, eine befriedigende Erklärung zu geben, was man darunter zu verstehen hat, denn für einen Konvergenzbegriff, der dafür erforderlich ist, war die Zeit noch nicht reif. Wir brauchen nur an die verfehlten Gedankengänge eines Zenon zu erinnern, der dem Achilles absprach, eine Schildkröte einholen zu können, obschon die Praxis schon damals das Gegenteil lehrte.

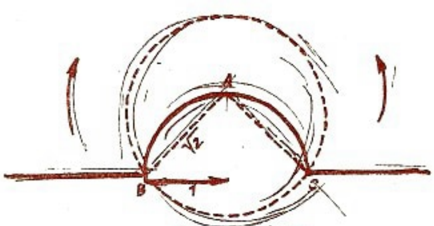
Die Pythagoräer gingen formal vor und gaben für den obengenannten unendlichen Kettenbruch einen Wert an, den er bei richtiger Sinngebung tatsächlich besitzt. Die formale Überlegung verläuft wie folgt. Wenn man den fraglichen Werl mit x bezeichnet, ist rein optisch

$$x = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \ddots}}} = 1 + \frac{1}{1 + 1 + \frac{1}{2 + \frac{1}{2 + \ddots}}} = 1 + \frac{1}{1+x}$$

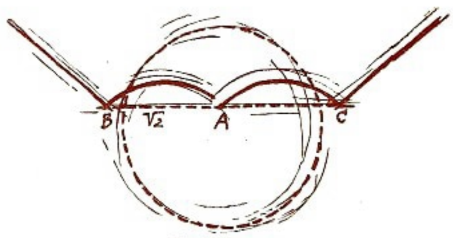
und damit $x^2 = 2$ woraus $x = \sqrt{2}$ folgt.

An diesem Wert hatten die Pythagoräer ein besonderes Interesse, weil er nach dem Pythagoras die Länge der Hypotenuse in einem gleichseitigen rechtwinkligen Dreieck darstellt, dessen beide Katheten die Länge 1 besitzen.

Weiter spielt $\sqrt{2}$ eine Rolle bei dem Trick, eine Münze durch ein kleineres Loch in einem Blatt Papier zu praktizieren. Zunächst faltet man das Papier so, dass das Loch zum Halbkreis an der Faltungskante wird. Die beiden Kantenstücke dreht man dann um den Scheitelpunkt A des Halbkreises in der Pfeilrichtung, bis die Sehnen AB und AC in einer Geraden liegen. Die Öffnung BC wurde dabei um das $\sqrt{2}$-fache vergrößert.



Ausgangslage



Endlage

9 Ganz reell

Außer den ganzen Zahlen und den Brüchen, mit denen wir bisher zu tun hatten, gibt es noch die irrationalen Zahlen. Zusammen mit den vorher genannten beiden Zahlenarten bilden sie die Gesamtheit der reellen Zahlen. Um diese zu gewinnen, ging man bisher schrittweise vor.

Von den natürlichen Zahlen ausgehend schaffte man nacheinander die immer umfassenderen Gesamtheiten der ganzen Zahlen, dann der Brüche, schließlich der reellen Zahlen. Wir schlagen einen neuen Weg ein, um eine unmittelbare Definition aller reellen Zahlen geben zu können. So ist die Überschrift zu deuten.

Jede reelle Zahl lässt sich im Dezimalsystem darstellen. Für diese Darstellbarkeit kann die Eindeutigkeit erzwungen werden auf Grund der Übereinkunft, eine Folge aus lauter Neunen durch Nullen zu ersetzen und eine Eins davor zu setzen, so dass an Stelle von $1{,}2\overline{9}$ dann 1,3 zu schreiben ist.
Nach dieser Vereinbarung wollen wir nun, auf unseren Kenntnissen über das Rechnen im Dezimalsystem aufbauend, unter Verwendung der natürlichen Zahlen alle reellen Zahlen auf einmal einführen, sozusagen in einem einzigen Schöpfungsakt. Wir definieren:

Reelle Zahl heißt jede nicht abbrechende Folge aus den Ziffern 0 bis 9, die mit einem der beiden Vorzeichen Plus oder Minus und außerdem mit einem Komma versehen ist, das einen endlichen Anfangsabschnitt abgrenzt.
Es sei gestattet, der Ziffernfolge endlich viele Nullen voranzustellen; weiterhin ist eine Folge von lauter Neunen durch eine Folge aus lauter Nullen zu ersetzen, vor die dann noch eine Eins kommt. Damit ist entschieden, wann zwei Ziffernfolgen dieselbe reelle Zahl darstellen können.

Liegen zwei reelle Zahlen vor, dann können wir nötigenfalls durch Voranstellen von Nullen erreichen, dass die beiden Abschnitte vor dem Komma gleich lang ausfallen. Unter dieser Voraussetzung definieren wir die Summe von zwei reellen Zahlen gleichen Vorzeichens wie folgt:
Das gemeinsame Vorzeichen gilt als Vorzeichen der Summe. Lauten die Ziffernfolgen der beiden Summanden

$$a_1 a_2 ... a_\nu,\, a_{\nu+1} ... \qquad \text{und} \qquad b_1 b_2 ... b_\nu,\, b_{\nu+1} ...$$

dann bilde man die Abschnitte

$$\frac{a_1}{b_1}, \frac{a_1 a_2}{b_1 b_2}, \frac{a_1 a_2 a_3}{b_1 b_2 b_3}, ...$$

und addiere bei jedem so, als wären es zwei Zahlen im Dezimalsystem. Jeder Abschnitt liefert so eine abbrechende Ziffernfolge, die Partialsumme heißen möge. Wir wollen nun sehen, wie weit die aufeinanderfolgenden Partialsummen in ihren Ziffern miteinander übereinstimmen.

Ist die letzte Ziffer c_n in der n-ten Partialsumme verschieden von 9, dann kann sie in der $(n+1)$-ten Partialsumme durch Übertrag aus der $(n+1)$-ten Stelle um 1 auf eine 9 erhöht werden, während die letzte Ziffer c_{n+1} in der $(n+1)$-ten Partialsumme dann höchstens eine 8 sein kann. In der $(n+2)$-ten Partialsumme kann sich zwar diese 8 auf eine 9 erhöhen, aber die unmittelbar davorstehende Ziffer bleibt davon bereits unberührt, und so weiter.

Wir sehen so, dass die vor der Ziffer c_n stehenden Ziffern in allen auf die n-te Partialsumme folgenden Partialsummen erhalten bleiben. So bewirkt jede von 9 verschiedene Ziffer, die in einer Partialsumme auftritt, dass die vor ihr stehenden Ziffern in allen darauffolgenden Partialsummen die gleichen sind. Damit stellen sie den Anfang in der Ziffernfolge der Summe dar.
Nun kann sich zweierlei ereignen. Entweder treten in den Ziffernfolgen der sukzessiven Partialsummen immer wieder von 9 verschiedene Ziffern auf. Dann bestimmen sie abschnittsweise die Ziffernfolge in der Summe. Oder aber es tritt in den Partialsummen schließlich immer nur die 9 hinzu; auch in diesem Fall ist dann die Ziffernfolge in der Summe durch die Partialsummen eindeutig bestimmt. Allerdings ist das nur im Prinzip richtig, denn in der Praxis kann man auf unüberwindliche Schwierigkeiten stoßen. Diese können aber genauso beim üblichen Gebrauch des Dezimalsystems auftreten, wie das Beispiel +4,5454...546... +5,4545...456... zeigt, in dem die ersten Partialsummen

$$9; \quad 9{,}9; \quad 9{,}99; \quad ...; \quad 9{,}99...9; \quad 9{,}99...99; \quad 10{,}00...002$$

sind. Wenn die jeweils ersten drei Punkte in den Summanden 1000 Stellen andeuten, dann sieht man, dass über 1000 Partialsummen zu berechnen sind, ehe man die allererste Ziffer der Summe gewonnen hat. Wir vermerken noch, dass die ν-te Partialsumme den Platz des Kommas in der Summe bestimmt, die insgesamt durch unsere wohlbestimmte Ziffernfolge, das Komma und das Vorzeichen der beiden Summanden definiert ist.

Die Addition von reellen Zahlen, die wir jetzt kürzer durch große Buchstaben bezeichnen, möge durch das Zeichen $\oplus$ angezeigt werden. Es wird sich als nützlich erweisen, noch gewisse unendliche Summen einzuführen.
Mit der Folge von reellen Zahlen A, B, C, ... bilde man nacheinander die Summen

$$A \oplus B, \qquad A \oplus B \oplus C, \qquad ...$$

Die einzelnen Summen bilden eine Folge von Ziffernfolgen. Es kann zutreffen, dass an einer gewissen festen Stelle vor oder hinter dem Komma von einer Ziffernfolge an in allen darauffolgenden Ziffernfolgen dieselbe "feste" Ziffer steht. Trifft das für jede Stelle vor oder hinter dem Komma zu, dann bilden die "festen" Ziffern eine Ziffernfolge, die zur Definition der unendlichen Summe

$$A \oplus B \oplus C \oplus ...$$

verhilft.
Scheint auch unser Vorgehen bei der Addition umständlich zu sein, so werden doch nur geläufige Kenntnisse verwendet, die aber dabei plastischer als sonst in Erscheinung treten. Das gilt auch für die Multiplikation, der wir uns jetzt zuwenden wollen.
Dabei werden wir, gerade weil es sich um eine fortlaufende Anwendung geläufiger Kenntnisse handelt, uns mit Hinweisen begnügen, die den Leser in die Lage versetzen, die Beweise selbst auszuführen.

Das Produkt aus zwei reellen Zahlen erhalte das Vorzeichen Plus, wenn beide Faktoren dasselbe Vorzeichen haben, und das Vorzeichen Minus, wenn das nicht der Fall ist. Als Produkt einer beliebigen reellen Zahl A mit einer zweiten B^*, in der aber höchstens eine einzige Ziffer b von null verschieden ist, erklären wir die reelle Zahl, die man wie folgt gewinnt:

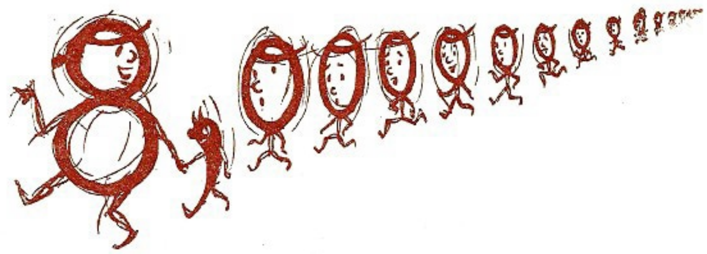
Zunächst legt man das Vorzeichen in der vereinbarten Weise fest. Dann ist in A das Komma um $m-1$ Stellen nach rechts bzw. um m Stellen nach links zu rücken, je nachdem, ob b in B^* an m-ter Stelle links bzw. rechts vom Komma steht. So erhält man eine neue reelle Zahl A', die dazu verhilft, das Produkt AB^* zu erklären.
Im Falle $b > 1$ wird es als die b-gliedrige Summe $A' \oplus ... \oplus A'$ definiert, im Falle $b = 1$ als A', schließlich im Falle $b = 0$ als $0{,}\overline{0}$.

Liegt eine beliebige reelle Zahl B vor, dann stellen wir damit eine Folge $B_1^*, B_2^*, ...$ von reellen Zahlen vom eben geschilderten speziellen Aufbau her, indem in B mit Ausnahme der n-ten Ziffer alle anderen Ziffern durch Nullen ersetzt werden.
Nun multipliziere man A in der eben beschriebenen Art nacheinander mit den reellen Zahlen der Folge, um hinterher die unendliche Summe

$$AB_1^* \oplus AB_2^* \oplus ...$$

zu bilden. Sie ist sinnvoll, weil das Komma in den sukzessiven Summanden immer weiter nach links rückt, so dass die Einflussnahme auf vorhergehende Ziffern von Summand zu Summand schwächer wird, ähnlich wie in der Diskussion zur Addition. Daher können wir diese unendliche Summe als das Produkt AB ansprechen.

Nun lassen sich die Kommutativität und Assoziativität von Addition und Multiplikation nachweisen, sowie die Distributivität, die Addition und Multiplikation miteinander verknüpft. Subtraktion und Division können wie üblich als Umkehroperationen eingeführt werden.



Reelle Zahlen, in denen nach dem Komma lauter Nullen stehen, sind mit den ganzen Zahlen zu identifizieren; diejenigen von unseren reellen Zahlen, die aus einer, von endlich vielen ersten Ziffern abgesehen, periodischen Ziffernfolge bestehen, stimmen mit Brüchen überein, so dass man schließlich erkennt, dass es sich bei unseren reellen Zahlen um die üblichen reellen Zahlen handelt.
Daher sind die neuen Rechenoperationen lediglich Verallgemeinerungen, für die man die üblichen Zeichen verwenden darf. Darauf wollen wir nicht näher eingehen, sondern eine fundamentale Eigenschaft unserer reellen Zahlen nachweisen.

Dazu müssen wir unsere reellen Zahlen erst ordnen, indem wir für sie die Beziehung kleiner-größer einführen. Um zwei reelle Zahlen zu vergleichen, schreibe man beide ohne $\overline{9}$ und mit gleichvielen Ziffern vor dem Komma. Letzteres besagt, dass nötigenfalls bei der einen Nullen voranzustellen sind.
Liegen dann zwei solche Ziffernfolgen $a_1a_2...$ und $b_1b_2...$ vor, dann möge die zur ersten gehörige positive reelle Zahl kleiner als die zur zweiten gehörige positive reelle Zahl gelten, wenn zwar $a_1 = b_1$, $a_2 = b_2$, ..., $a_n = b_n$, aber $a_{n+1} < b_{n+1}$ ist.
Für negative reelle Zahlen gelte das Umgekehrte, während eine negative Zahl stets kleiner als eine positive Zahl sei, bis auf eine Ausnahme: Ein genaueres Eingehen auf die Subtraktion führt dazu, $+0{,}\overline{0}$ und $-0{,}\overline{0}$ gleichzusetzen.

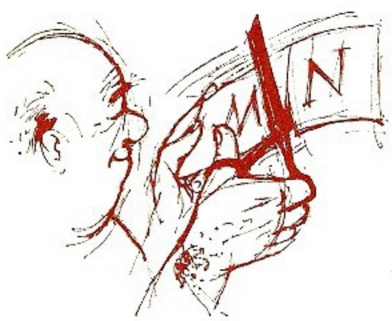
Aus der Erklärung folgt sofort, dass, wenn A kleiner als B und B kleiner als C ist, dann immer A kleiner als C ausfällt, in Zeichen:, aus $A < B$ und $B < C$ folgt $A < C$, das heißt die kleiner-größer-Beziehung ist transitiv.

Der Relation kleiner-größer entspricht auf der Geraden die links-rechts-Beziehung ihrer Punkte. Diese Übereinstimmung lässt erwarten, Eigenschaften der Geraden bei den reellen Zahlen wiederzufinden. Eine fundamentale Eigenschaft der Geraden ist ihre Lückenlosigkeit.
Jeder Punkt P auf ihr führt eine Aufteilung ihrer Punkte in zwei Mengen herbei.
Die eine Menge M enthält die Punkte, die links von P liegen, die andere Menge N die Punkte, die rechts von P liegen. Der Punkt P selbst kann nach Belieben zu M oder N gerechnet werden. Jeder Punkt aus M liegt augenscheinlich links von jedem Punkt aus N.
Davon gilt aber auch die Umkehrung. Liegt nämlich eine Einteilung der Punkte der Geraden in zwei Mengen vor, so dass jeder Punkt der einen Menge M links von jedem Punkt der anderen Menge N liegt, dann gibt es immer einen einzigen Punkt P, der diese Einteilung herbeiführt. Das lässt sich nun auf die Gesamtheit der reellen Zahlen übertragen.

Nehmen wir an, es läge eine Aufteilung aller reellen Zahlen in zwei Mengen vor, so dass jede reelle Zahl aus der einen Menge M kleiner ist als jede Zahl aus der zweiten Menge N.
Dann wollen wir eine reelle Zahl nachweisen, die größer ist als jede von ihr verschiedene Zahl aus M und kleiner als jede von ihr verschiedene Zahl aus N.
Um das einzusehen, betrachte man irgendeine Zahl

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

aus N. Der Kürze halber nehmen wir an, dass es in M positive Zahlen gibt. Dann gelingt es, eine größte Zahl für M wie folgt anzugeben.



Ihre erste Ziffer c_1 sei die größte unter allen Ziffern, die in positiven Zahlen aus M an n-ter Stelle vor dem Komma vorkommen. Ihre zweite Ziffer c_2 sei die größte unter allen Ziffern an $(n-1)$-ter Stelle vor dem Komma, die in denjenigen positiven Zahlen aus M vorkommen, in denen an n-ter Stelle vor dem Komma die Ziffer c_1 steht, und so weiter.
Die Zahl $+c_1 c_2 ... c_n, c_{n+1} ...$ ist zufolge ihrer Konstruktion offensichtlich größer als jede von ihr verschiedene Zahl aus M und kleiner als jede von ihr verschiedene Zahl aus N. Man sagt, sie erzeuge den Schnitt $(M|N)$.

Mit reellen Zahlen rechnete man von jeher, ohne insbesondere die irrationalen Zahlen einwandfrei erklärt zu haben. Das geschah erst im Jahre 1872. Cantor und Dedekind veröffentlichten in diesem Jahr voneinander unabhängig ihre Definitionen.
Der erstgenannte führte die irrationalen Zahlen mit Hilfe von konvergenten Folgen ein, der zweite mit Hilfe der heute nach ihm benannten Schnitte, auf deren Idee er nach eigener Angabe bereits am 24.11.1858 verfiel.

Für beide Forscher aber war eine vorhergehende Einführung der rationalen Zahlen unumgänglich notwendig, während wir alle reellen Zahlen zugleich in unmittelbarem Anschluss an die natürlichen Zahlen einzuführen vermochten.

Dedekind war Professor in Braunschweig. Ihm verdankt die moderne Mathematik entscheidende Beiträge, und zwar zunächst zur Algebra, dann zur Mengenlehre. Mit Cantor, dem wir schon begegneten, stand er im Gedankenaustausch und förderte dessen Ausbau der Mengenlehre wesentlich durch kritische Hinweise.

Wie tief Dedekind auch in dieser Materie vorgedrungen ist, erkennt man daran, dass er für einen Satz, den Cantor selbst nur vermutete, und den zu beweisen erst seinem Schüler Felix Bernstein 1897 gelang, bereits am 11.7.1887 einen überlegenen Beweis fand, der im Gegensatz zu dem vorher genannten Beweis keinen Gebrauch von den natürlichen Zahlen machen muss.

10 Libussa verteilt Körbe

Libussa war Fürstin der Böhmen und Tochter einer Fee. Sie war klug und schön, wie es Sagen vorsehen. Die Dame Libussa war also eine gute Partie, so dass sich viele Freier einfanden.
Zwei Ritter unter diesen bedrängten sie gar arg, doch sie hatte schon ihren Auserwählten. Um Fehden aus dem Weg zu gehen, verfiel sie auf die schlaue Idee, ihren Herzallerliebsten Primislas in einem Quiz-Wettbewerb um ihre Hand gewinnen zu lassen. Wir möchten nicht so ungalant sein und annehmen, dass sie ihre Preisfrage mit ihrem Auserwählten vorher schon abgesprochen hat.

Nach dem Berichterstatter Musäus verlief das geistige Turnier wie folgt:

"Da nun die Reihe an sie, Libussa, kam, eins aufzugeben, berief sie den Fürsten Wladomir, den Ritter Mizisla und den Junker Primislas zu sich und sprach: Ihr wackern Gesellen, jetzt schickt euch an, ein von mir aufgegebenes Rätsel zu lösen, damit offenbar werde, wer unter euch der Weiseste und Verständigste sei.
Ich habe euch allen dreien eine Spende zugedacht aus diesem Körbchen, gefüllt von Pflaumen, die ich gepflückt habe in meinem Garten. Einer unter euch soll die Hälfte davon haben und eine darüber, der andere soll wieder die Hälfte haben und eine darüber, der dritte soll nochmals die Hälfte haben und drei darüber. So sich nun befindet, dass der Korb ausgeleert ist, sagt mir an, wieviel Pflaumen jetzt innen sind ?"

Programmgemäß versagten die zwei ungestümen Ritter, während Primislas die richtige Antwort fand - oder sie zumindest gab. Wir wollen die Antwort selbst finden und setzen zunächst x für die Gesamtzahl der Pflaumen im Korb. Die erste Entnahme besteht dann aus

$$\frac{x}{2} + 1$$

Pflaumen, so dass noch

$$x - \left(\frac{x}{2} + 1\right) = \frac{x}{2} - 1$$

Pflaumen im Korb verbleiben. Die zweite Entnahme besteht dann aus

$$\frac{1}{2}\left(\frac{x}{2} - 1\right) = \frac{x}{4} + \frac{1}{2}$$

Pflaumen, so dass noch

$$\left(\frac{x}{2} - 1\right) - \left(\frac{x}{4} + \frac{1}{2}\right) = \frac{x}{4} - \frac{3}{2}$$

Pflaumen im Korb verbleiben. Die letzte Entnahme besteht schließlich aus

$$\frac{1}{2}\left(\frac{x}{4} - \frac{3}{2}\right) + 3 = \frac{x}{8} + \frac{9}{4}$$

Pflaumen. Die drei Entnahmen sollten den Korb entleert haben, so dass

$$\left(\frac{x}{2} + 1\right) + \left(\frac{x}{4} + \frac{1}{2}\right) + \left(\frac{x}{8} + \frac{9}{4}\right) = x$$

sein muss. Aus dieser linearen Gleichung ergibt sich für die Gesamtzahl der Pflaumen im Korb der Wert $x = 30$.

Unser Gewährsmahn Musäus schließt seinen Bericht mit den Worten: "Niemand konnte begreifen, wie der menschliche Witz auf der einen Seite eine gemeine Zahl so rätselhaft in Worte verschränken und auf der anderen mit solcher Zuverlässigkeit sie aus dieser kunstreichen Verborgenheit herauszuklauben vermöge.

Den ledigen Korb verlieh das Fräulein den beiden Rittern, die ihrer Liebe nicht teilhaft werden konnten, zum Andenken an die erloschene Liebschaft. Daher kommt die Gewohnheit, dass man von einem zurückgewiesenen Freier sagt, er habe von seinem Liebchen einen Korb bekommen, bis auf den heutigen Tag."

Die allgemeine lineare Gleichung lautet:

$$ax + b = 0$$

mit $a \neq 0$ und besitzt die einzige Lösung

$$x = -\frac{b}{a}$$

Sind dabei a und b beide ganze Zahlen, dann kann die Lösung durchaus ein echter Bruch sein. Sind weiter a und b selbst Brüche, dann ist die Lösung ebenfalls immer ein Bruch oder eine ganze Zahl. Man erkennt so, dass lineare Gleichungen wohl aus der Gesamtheit der ganzen Zahlen, nicht aber aus der Gesamtheit der Brüche herausführen.

’ Anders bei quadratischen Gleichungen. Bereits die Gleichung

$$x^2 - 2 = 0$$

besitzt keine rationale Lösung. Wäre der Bruch p/q eine Lösung, dann können wir ihn uns gleich in gekürzter Form denken, so dass p und q teilerfremd sind. Wenn nun p/q Lösung wäre, müsste

$$p^2 = 2q^2$$

sein. Da die rechte Seite eine gerade Zahl ist, müsste es die linke Seite ebenfalls sein, was nur dann möglich wäre, wenn p selbst gerade wäre.
Dann wäre aber p^2 durch 4 teilbar, folglich auch die rechte Seite, $2q^2$. Es müsste daher q^2 und damit q selbst gerade sein. Demnach wären p und q beide gerade Zahlen, während sie doch teilerfremd sein sollten.

Die quadratische Gleichung führt also aus der Gesamtheit der rationalen Zahlen heraus, indem die Lösung $\sqrt{2}$ keine rationale Zahl ist. Die quadratische Gleichung führt aber auch aus der umfassenderen Gesamtheit aller reellen Zahlen, zu denen $\sqrt{2}$ gehört, heraus, wie das Beispiel

$$x^2 + 1 = 0$$

zeigt. Keine reelle Zahl kann dieser Gleichung genügen, weil das Quadrat einer reellen Zahl niemals negativ ist, was jedoch erforderlich wäre, damit die linke Seite gleich null ist.

Das veranlasst eine erneute Erweiterung des Zahlbegriffes über die reellen Zahlen hinaus. Dabei ging man nur zögernd vor, wie allein der Name imaginäre Zahlen zeigt.
Eine erste Fühlungnahme mit den komplexen Zahlen, um die es geht, verzeichnet die Stellungnahme von Cardano 1545 in seinem Buch, das er großsprecherisch "Ars magna" nannte. Er geht von der Aufgabe aus, zwei Zahlen x und y zu bestimmen, deren Summe 10, deren Produkt 40 ist:

$$x + y = 10 \qquad \text{und} \qquad xy = 40$$

Aus der ersten Gleichung erhalten wir für y den Wert $10 - x$. Setzen wir ihn in die zweite Gleichung ein, dann folgt die quadratische Gleichung

$$x^2 - 10x + 40 = 0$$

Rechnet man mit den Größen $5 + \sqrt{-15}$ und $5 - \sqrt{-15}$ formal so wie mit reellen Zahlen, dann stellen sie Lösungen der quadratischen Gleichung und zugleich der Aufgabe dar, von der wir ausgingen. Nach dieser Bemerkung ist aber Cardano auf diese Zahlen nicht weiter eingegangen.

Der Italiener Bombelli hatte dann in einem Buch aus dem Jahre 1572 an Ausdrücken der Form

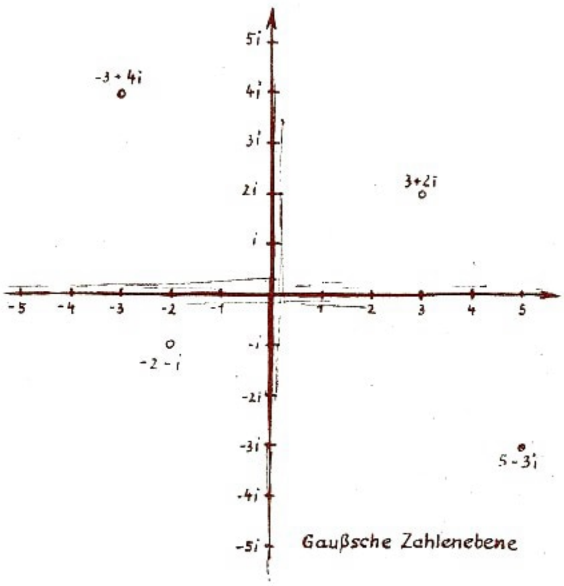
$$\sqrt[3]{a + \sqrt{-b}} + \sqrt[3]{a - \sqrt{-b}}$$

Rechnungen mit komplexen Größen durchgeführt. Dazu bemerkt er selbst: "Ein ausschweifender Gedanke nach der Meinung vieler. Ich selbst war eine zeitlang derselben Ansicht.

Die Suche schien mir auf Sophismen mehr als auf Wahrheit zu beruhen, aber ich suchte solange, bis ich den Beweis fand."

Als erster stieß Wessel zu der Auffassung vor, dass die komplexen Größen wirkliche Zahlen sind. Seine Arbeit, die 1797 in den Berichten der dänischen Akademie erschien, blieb jedoch unbeachtet.

Erst Gauß gelang es 1831 dank seiner Autorität bei den Zeitgenossen Gehör zu finden. Heute wird die Interpretation der komplexen Zahlen als Punkte einer Ebene nach ihm benannt. Zur Einführung der komplexen Zahlen schlagen wir hier einen Weg ein, den Hamilton vier Jahre später fand.



Wir schaffen eine neue Spielart von Größen dadurch, dass wir Paare aus reellen Zahlen bilden. Letztere bezeichnen wir durch kleine lateinische Buchstaben. Zwei Paare (a, b) und (c, d) Sollen genau dann gleich sein, das heißt

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

wenn für die vier reellen Zahlen a, b, c und d die Gleichungen $a = c$ und $b = d$ gelten.

Die Addition wird durch

$$(a,b) + (c,d) = (a + c, b + d)$$

erklärt. Augenscheinlich spielt die Größe (0,0) dabei die Rolle der Null. Aus der Definition der Addition folgt unmittelbar, dass sie kommutativ und assoziativ ist.
Die Subtraktion als Umkehroperation der Addition führt auf

$$(a,b) - (c,d) = (a - c, b - d)$$

Die Multiplikation wird durch die etwas verwickeltere Formel

$$(a, b) \cdot (c, d) = (ac - bd, ad + bc)$$

erklärt. Augenscheinlich spielt die Größe (1,0) die Rolle der Eins. Die Kommutativität folgt daraus unmittelbar und nach kurzer Rechnung auch die Assoziativität. Es ist einerseits

$$[(a,b) + (c,d)] \cdot (e,f) = (a + c,b + d) \cdot (e,f) = ((a + c)e - (b + d)f, (a + c)f + (b + d)e)$$

andererseits

$$\begin{aligned}(a,b) \cdot (e,f) + (c,d) \cdot (e,f) &= (ae - bf, af + be) + (ce - df, cf + de)\\ &= ((a + c)e - (b + d)f, (a + c)f + (b + d)e)\end{aligned}$$

Aus der Gleichheit der rechten Seiten folgt daher das Distributivgesetz

$$[(a,b) + (c,d)] \cdot (e,f) = (a,b) \cdot (e,f) + (c,d) \cdot (e,f)$$

Die Formel für die Subtraktion konnte man ohne Rechnung aus der Forderung ableiten, dass es sich um die Umkehroperation der Addition handelt. Bei der Division, der Umkehroperation der Multiplikation, ist dies nicht so leicht möglich.

Um die gesuchte Formel zu finden, setzen wir für den Quotienten (x, y) des Zählers (a, b) und des Nenners (c, d) die Gleichung

$$(c,d) \cdot (x,y) = (a,b)$$

an. Nach den Erklärungen für die Multiplikation und die Gleichheit von Zahlenpaaren ist sie gleichwertig mit den beiden Gleichungen

$$cx - dy = a \qquad \text{und} \qquad cy + dx = b$$

Daraus folgt

$$x = \frac{ac + bc}{c^2 + d^2} \qquad \text{und} \qquad y = \frac{bc - ad}{c^2 + d^2}$$

so dass der Quotient durch

$$(x,y) = \left(\frac{ac + bc}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2} \right)$$

bestimmt ist.

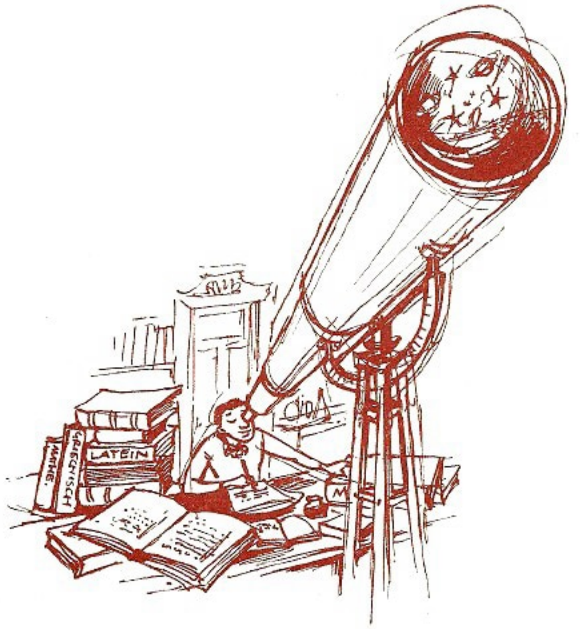
Nachdem für die neuen Größen die vier Rechnungsarten erklärt wurden, die denselben formalen Regeln genügen wie bei den reellen Zahlen, wollen wir versuchen, die reellen Zahlen mit Spezialfällen der neuen Größen zu identifizieren. Das gelingt auf Grund der Darstellung

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

Trifft man die Übereinkunft, Größen wie $(a, 0)$ mit der entsprechenden reellen Zahl a und $(0, 1)$ mit der imaginären Einheit i zu identifizieren, dann geht die Gleichung

$$(0,1) \cdot (0,1) = (-1,0) \qquad \text{in} \qquad i^2 = -1$$

über, während die Gleichung (*) die Darstellung $a + ib$ für die Größe (a, b) gibt, die sie als komplexe Zahl erkennen lässt.



Hamilton, dessen Methode wir anwandten, wurde 1827 mit 22 Jahren Direktor der Sternwarte in Dublin.
Er war außerordentlich vielseitig, beherrschte zum Beispiel schon im Knabenalter 13 Sprachen und war sein Leben lang schriftstellerisch tätig.
Er führte eine weitere wichtige Spielart von Zahlen ein, die Quaternionen, auf die allerdings bereits Euler vor ihm gestoßen ist. Hamilton und noch mehr seine Nachfolger überschätzten jedoch die Bedeutung der Quaternionen und machten die Quaternionenlehre zu einer Art von Kult.
Weitere wichtige Beiträge leistete Hamilton auch noch auf dem Gebiet der Mechanik und Optik.

11 Cardano greift ein

Um die quadratische Gleichung

$$ax^2 + bx + c = 0 \qquad \text{mit} \qquad a \neq 0$$

oder anders geschrieben

$$x^2 + \frac{b}{a}x + \frac{c}{a} = 0$$

aufzulösen, subtrahiere man beiderseits $\frac{c}{a}$ und addiere hinterher zu beiden Seiten $\left(\frac{b}{2a}\right)^2$.
So entsteht die Gleichung

$$x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 = \left(\frac{b}{2a}\right)^2 - \frac{c}{a}$$

Ihre linke Seite ist das Quadrat des linearen Ausdrucks $x + \frac{b}{2a}$, so dass die Gleichung die Gestalt

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

annimmt, wenn man die beiden Glieder rechts noch auf gleichen Nenner bringt. Zieht man in der letzten Gleichung die Quadratwurzel, ergibt sich

$$x + \frac{b}{2a} = \frac{\sqrt{b^2 - 4ac}}{2a} \qquad \text{woraus} \qquad x = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

folgt.
Da eine Quadratwurzel zweier Werte fähig ist, stellt der Ausdruck rechts zwei Werte dar, die nur zusammenfallen, wenn der Radikand verschwindet. Steht unter dem Wurzelzeichen speziell eine nichtnegative reelle Zahl, dann unterscheiden sich die beiden Werte lediglich im Vorzeichen der Wurzel voneinander:

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \qquad , \qquad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

Unser Vorgehen gilt aber ganz allgemein, auch für komplexe Koeffizienten a, b und c, weil ja - wie wir gezeigt haben - die Rechenregeln ihre Gültigkeit auch dann noch behalten.
Die Auflösung der quadratischen Gleichung, selbstverständlich nur mit reellen Koeffizienten, war schon im Altertum bekannt. Die Auflösung von Gleichungen dritten und vierten Grades dagegen gelang erst im 16.Jahrhundert in Italien.
Um die kubische Gleichung auflösen zu können, weisen wir zunächst nach, dass eine Kubikwurzel drei verschiedene Werte besitzt. Dazu gehen wir von der Identität

$$x^3 - 1 = (x - 1) \cdot (x^2 + x + 1)$$

aus. Von ihrer Richtigkeit überzeugt man sich durch Ausmultiplizieren auf der rechten Seite. Die Werte, die, für x eingesetzt, die linke Seite zum Verschwinden bringen, sind die dritten Einheitswurzeln. An Hand der rechten Seite erkennt man aber, dass dies der Wert 1 und die beiden Lösungen der quadratischen Gleichung

$$x^2 + x + 1 = 0$$

sind, die nach der eben abgeleiteten Formel

$$\frac{-1+\sqrt{-3}}{2} \qquad \text{und} \qquad \frac{-1-\sqrt{-3}}{2}$$

lauten. Bezeichnen wir den ersten der beiden beiden Werte mit ρ, folgt

$$\rho^2 = \frac{1-2\sqrt{-3}-3}{4} = \frac{-1-\sqrt{-3}}{2}$$

so dass $1, \rho, \rho^2$ die dritten Einheitswurzeln sind.
Ist nun a eine komplexe Zahl und α eine beliebige Kubikwurzel aus a, $\alpha^3 = a$, dann sind die von α und voneinander verschiedenen Werte $\rho\alpha$ und $\rho^2\alpha$ ebenfalls Kubikwurzeln aus a, weil

$$(\rho\alpha)^3 = \rho^3 \cdot \alpha^3 = 1 \cdot a = a \qquad \text{und} \qquad (\rho^2\alpha)^3 = (\rho^2)^3 \cdot \alpha^3 = (\rho^3)^2 \cdot \alpha^3 = 1 \cdot a = a$$

ist.
Die allgemeinste kubische Gleichung kann zunächst dadurch, dass man durch den Koeffizienten des höchsten Gliedes dividiert, in der nur scheinbar weniger allgemeinen Gestalt

$$x^3 + ax^2 + bx + c = 0$$

geschrieben werden. Setzt man $x = y - \frac{a}{3}$, dann folgt nach kurzer Rechnung

$$y^3 + \left(b - \frac{a^3}{3}\right) y + \left(c - \frac{ab}{3} + \frac{2a3}{27}\right) = 0$$

eine Gleichung, in der das quadratische Glied fehlt. Wir schreiben

$$b - \frac{a^3}{3} = p \qquad \text{und} \qquad c - \frac{ab}{3} + \frac{2a^3}{27} = q$$

Damit lautet die kubische Gleichung

$$y^3 + py + q = 0$$

Mit Hudde 1659 setzen wir darin $y = u + v$, so dass die Gleichung in

$$u^3 + v^3 + (3uv + p) \cdot (u + v) + q = 0$$

übergeht. Diese Gleichung ist aber augenscheinlich erfüllt, wenn u und v den Forderungen

$$u^3 + v^3 = -q \qquad \text{und} \qquad 3uv = -p$$

genügen. Aus der Identität

$$(u^3 - v^3)^2 = (u^3 + v^3)^2 - 4u^3v^3$$

folgt dann

$$u^3 - v^3 = \sqrt{q^2 + \frac{4p^3}{27}} = 2\sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$$

Addiert man $u^3 + v^3 = -q$ hinzu, dann ergibt sich

$$u^3 = -\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3} \qquad \text{und} \qquad v^3 = -\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}$$

Die rechten Seiten davon genügen zunächst der Forderung $u^3 + v^3 = -q$. Weiterhin ergibt das Produkt u^3v^3 ausgerechnet

$$\begin{aligned} u^3v^3 &= \left[-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}\right]\left[-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}\right] \\ &= \left(-\frac{q}{2}\right)^2 - \left(\sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}\right)^2 = -\left(\frac{p}{3}\right)^3 \end{aligned}$$

Daher ist das Produkt uv einem der Werte $-\frac{p}{3}, -\frac{\rho p}{3}, -\frac{\rho^2 p}{3}$ gleich.

Hat man also den Wert von u als dritte Wurzel nach Belieben gewählt, dann kann man durch passende Wahl des Wertes von v, das selbst eine dritte Wurzel ist, erreichen, dass die Gleichung

$$uv = -\frac{p}{3}$$

gilt und damit auch noch die zweite Forderung erfüllt ist. Mit dem auf die Wahl von u abgestimmten Wert von v ist dann $y = u + v$. Mit den Werten u und v erfüllen aber zugleich noch ρu und $\rho^2 v$, weiter aber $\rho^2 u$ und ρv dieselben Forderungen, so dass

$$y = \rho u + \rho^2 v \qquad \text{und} \qquad y = \rho^2 u + \rho v$$

zwei weitere Lösungen der Gleichung $y^3 + py + q = 0$ darstellen. Um damit die Lösungen der ursprünglichen Gleichung

$$x^3 + ax^2 + bx + c = 0$$

zu gewinnen, hat man in den Ausdrücken für y, die aus p und q aufgebaut sind, diese Größen durch ihre Ausdrücke aus den Koeffizienten a, b und c zu ersetzen und die Beziehung $x = y - \frac{a}{3}$ heranzuziehen.

Cardano, nach dem die Auflösungsformeln für die kubische Gleichung benannt sind, fand sie nicht selbst, wenn er dazu auch einen bedeutsamen Beitrag lieferte. Er, der Professor der Mathematik, erfuhr die Auflösung von dem italienischen Büchsenmacher Tartaglia, der von seinem ebenbürtigen Vorgänger Scipione dal Ferro nichts wusste und die Auflösung der kubischen Gleichung selbständig entdeckt hatte.
So ist die Personalbezeichnung der Formeln ungerecht, was leider keinen Einzelfall darstellt.

Um auch noch auf die Person von Cardano mit ein paar Worten einzugehen, sei auf seine Autobiographie hingewiesen, in der er sich schonungslos als voller Widersprüche zeichnet. Lasterhaft und fromm, gelehrt und abergläubisch, war er Philosoph und Mediziner, Mathematiker und Astrolog. Die Sterne verrieten ihm angeblich seine genaue Sterbestunde.

Er starb zur festgesetzten Zeit 75jährig. Sein Lebenswille erlag der Eitelkeit, so dass er wohl nachhalf, um termingemäß abzuleben.

Ähnlich unausgeglichen war sein Schüler Ferrari, der ebenfalls ein gewaltsames Ende nahm, als er von der eigenen Schwester vergiftet wurde. Ferrari gelang die Auflösung der biquadratischen Gleichung, die wir heute durch den Ansatz y= $u+v+w$ bewerkstelligen.

12 Revolutionierte Algebra

Als Lösungen der quadratischen und der kubischen Gleichung ergeben sich Ausdrücke, die - aufgebaut aus den Gleichungskoeffizienten - unter Anwendung der vier Rechnungsarten und des Wurzelausziehens gewonnen wurden. Ausdrücke dieser Art heißen algebraisch. Als Lösungen der biquadratischen Gleichung erhielt man ebenfalls algebraische Ausdrücke.
Als man aber versuchte, auch für die Lösungen der Gleichung fünften Grades algebraische Ausdrücke aufzufinden, wollte das nicht gelingen. Sehr viel später erst entdeckten der Italiener Ruffini und der Norweger Abel voneinander unabhängig, dass dies auch nicht gelingen konnte!

Das bedeutete eine Wende in der algebraischen Problematik. Einerseits konnte Gauß 1799 in seiner Doktorarbeit nachweisen, dass jede Gleichung, gleich welchen Grades, mindestens eine Lösung in reellen oder komplexen Zahlen besitzt, andererseits wurde wenig später von Ruffini und Abel der Nachweis geführt, dass diese Lösungen für Gleichungen höheren als vierten Grades nicht algebraisch sind.
Dazu ist noch die Bemerkung zu machen, dass dabei allgemeine Gleichungen gemeint wurden, denn die Lösungen einer speziellen Gleichung wie

$$x^n - 1 = 0$$

lauten ja $x = \sqrt[n]{1}$, sind also algebraisch.

Diese Überlegungen führten schließlich zur Frage nach einer Charakterisierung von Gleichungen, der man entnehmen kann, ob sie algebraische Lösungen haben oder nicht. Eine solche Charakterisierung gelang dem unvergleichlichen Genie des Franzosen Galois, der darüber hinausging und sich nicht mehr auf den algebraischen Aufbau der Lösungen beschränkte.
Zur Präzisierung des Aufbaus verwandte er den für die ganze Mathematik zentralen Gruppenbegriff. Galois erkannte die entscheidende Rolle, die eine Gruppe von Permutationen bei der Auflösungstheorie algebraischer Gleichungen spielt. Er ordnet jeder algebraischen Gleichung eine Permutationsgruppe zu und kann aus der Struktur der Gruppe zurückschließen auf die Struktur der Lösungen; insbesondere kann er entscheiden, ob die Lösungen algebraisch sind.

Wir möchten es bei diesen orientierenden Bemerkungen bewenden lassen und uns einer Frage zuwenden, die unseren Mitteln zugänglich ist, der Frage nach der größtmöglichen Anzahl von Lösungen der allgemeinen Gleichung n-ten Grades. Ohne Beschränkung der Allgemeinheit kann diese in der Form

$$x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n = 0$$

geschrieben werden, weil man ja mit dem Koeffizienten des höchsten Gliedes durchdividieren kann. Wir beginnen mit der Identität

$$y^n - z^n = (y - z) \cdot (y^{n-1} + y^{n-2}z + ... + yz^{n-2} + z^{n-1}) \qquad (*)$$

von deren Richtigkeit man sich durch Ausmultiplizieren der rechten Seite überzeugt, auf der sich alle Glieder bis auf das erste und das letzte wegheben.
Nun sei α eine Lösung (oder Wurzel) der obengenannten Gleichung n-ten Grades, deren linke Seite mit $f(x)$ bezeichnet werde. Es gilt also $f(\alpha) = 0$, so dass

$$f(x) - f(\alpha) = f(x)$$

ist. Die linke Seite lässt sich noch in der Form

$$x^n - \alpha^n + a_1(x^{n-1} - \alpha^{n-1}) + ... + a_{n-1}(x - \alpha)$$

schreiben. Wendet man die Identität (*) auf die einzelnen Klammern an, so lässt sich die linke Seite als $(x - \alpha) \cdot f^*(x)$ schreiben, wobei $f^*(x)$ ein Polynom von niedrigerem als n-ten Grad ist. Damit folgt

$$f(x) = (x - \alpha) \cdot f^*(x)$$

Wenn dann β eine weitere, von α verschiedene Wurzel von $f(x)$ darstellt, muss β zugleich eine Wurzel von $f^*(x)$ sein, damit die rechte Seite verschwinden kann. Aus

$$f^*(x) = (x - \beta) \cdot f^{**}(x)$$

folgt weiter

$$f(x) = (x - \alpha) \cdot (x - \beta) \cdot f^{**}(x)$$

So weiterschließend erkennt man, dass jede Wurzel von $f(x)$ einen linearen Faktor erzeugt. Das Produkt von mehr als n linearen Faktoren würde aber ein Polynom von höherem als n-ten Grad ergeben. Daher kann eine Gleichung n-ten Grades nicht mehr als n Wurzeln besitzen.

Ungleich schwieriger ist der Nachweis, dass jede Gleichung eine Wurzel besitzt. Im Laufe von anderthalb Jahrhunderten wurden verschiedene Beweise dafür erbracht. Gauß allein lieferte vier Beweise. Der kürzeste Beweis macht von der Funktionentheorie Gebrauch und übersteigt damit unsere Mittel.
Erst recht trifft das für die Theorie von Galois zu. Sie ist in der klassischen "Algebra" (3.Aufl. 1951) von Perron meisterhaft dargestellt.

Die folgerichtige Weiterentwicklung dieser Ideen führte dann zu Abstraktionen, die einen Wesenszug der heutigen Algebra ausmachen. Wieder begegnet uns dabei Dedekind, dann Emmi Noether, eine Frau, die in der Emigration starb, und der jüngst verstorbene Hamburger Professor Artin. Ihren Niederschlag fand die gewandelte Disziplin der Algebra im gleichnamigen Buch des Holländers v.d. Waerden (5. Aufl.1955/9).

Die ganze Entwicklung leitete jedoch das siebente Kapitel in der Zahlentheorie "Disquisitiones arithmeticae" von Gauß aus dem Jahre 1801 ein. Darin wird das zweitausend Jahre alte Problem gelöst, welche regulären Polygone mit Zirkel und Lineal konstruierbar sind. Die dort entwickelte Methode gab einen entscheidenden Anstoß für die Untersuchungen von Abel und Galois, wie die beiden betonten.

Das Leben von Abel war tragisch, denn ein Ruf an die Berliner Universität, der ihm endlich Anerkennung und Entlastung von materiellen Sorgen gebracht hätte, kam zu spät an, wenige Tage, nachdem er einer Tbc erlegen war. Heute steht ein Denkmal in seiner Heimat, das eher einen jungen Herkules als den schmächtigen Abel darstellt.

Weit tragischer noch waren Leben und Tod von Galois.
Noch als Schüler schickte er Abhandlungen an die Französische Akademie, die bis auf eine unbeantwortet blieben. An dieser einen Arbeit nörgelten die beiden Gutachter herum, ohne sie zu verstehen. Sie bemängelten zwar mit Recht die viel zu knappe Form, erkannten aber nicht die weittragende Bedeutung der Arbeit, deren Inhalt das Kernstück der modernen Algebra bildet.

In der Aufnahmeprüfung für die École Polytechnique von Paris fiel Galois durch. Nach allem klagte er über eine Verschwörung gegen das Genie.
Kaum 21jährig fiel er dann in einem Pistolenduell, das ihm nach einer Lesart aufgezwungen wurde, um ihn als Exponenten der republikanischen Partei zu beseitigen.
Seine mathematische Bedeutung "entdeckte" man erst ein Jahrzehnt nach seinem Tode: 1846 wurden seine Abhandlungen durch Liouville herausgegeben. Die Galoissche Gleichungstheorie bildet ein Glanzstück moderner Mathematik.

13 Rechnen mit Polynomen

Polynome können als Größen aufgefasst werden, mit denen man ähnlich wie mit ganzen Zahlen rechnet. Dafür muss erst erklärt werden, wann zwei Polynome gleich sind. Die Polynome

$$f(x) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n \qquad \text{und} \qquad g(x) = b_0x^m + b_1x^{m-1} + ... + b_{m-1}x + b_m$$

gelten als gleich, wenn $n = m$ ist und die Koeffizienten gleichhoher Potenzen von x paarweise gleich sind:

$$a = k = b_k \qquad \text{für} \quad k = 0, 1, ..., n$$

Da $0 \cdot x^k = 0$ ist, verändert ein Hinzufügen von Potenzen von x, die mit der Null multipliziert auftreten, nach der Erklärung für die Gleichheit das Polynom nicht.
Um die Addition zu definieren, darf man sich daher auf Polynome vom gleichen Grad beschränken. Für zwei solche Polynome $f(x)$ und $g(x)$ gilt das Polynom

$$(a_0 + b_0)x^n + (a_1 + b_1)x^{n-1} + ... + (a_{n-1} + b_{n-1})x + (a_n + b_n)$$

als die Summe $f(x) + g(x)$ der beiden Polynome. Das Ersetzen des Pluszeichens durch das Minuszeichen führt zur Definition der Subtraktion.

Unter dem Produkt $f(x) \cdot g(x)$ der beiden Polynome wiederum gleichen Grades $f(x)$ und $g(x)$ versteht man das Polynom

$$\begin{aligned} a_nb_nx^{2n} + (a_{n-1}b_n + a_nb_{n-1}x^{2n-1} + ... + (a_0b_n + a_1b_{n-1} + ... + a_{n-1}b_1 + a_nb_0)x^n \\ + ... + (a_0b_1 + a_1b_0)x + a_0b_0 \end{aligned}$$

Der Koeffizient der k-ten Potenz von x im Produkt $f(x) \cdot g(x)$ besteht aus der Summe aller Produkte $a_\mu b_\nu$ mit $0 \leq \mu \leq n$, $0 \leq \nu \leq n$, $\mu + \nu = k$, $k = 0, 1, ..., 2n$.
Der Fall $n + m$ ist darin enthalten, weil man die ersten Koeffizienten in einem der beiden Faktoren nach Bedarf gleich null setzen darf, und es bleibt nur hinzuzufügen, dass der Grad von $f(x) \cdot g(x)$ dann gleich $n + m$ ist.

Sind der Zähler $f(x)$ und der Nenner $g(x)$, dann gibt es wie bei den ganzen Zahlen genau einen Quotienten $q(x)$ und einen Rest $r(x)$.
Dabei möchten wir den Grad n von $f(x)$ größer als den Grad m von $q(x)$ annehmen. Der Grad von $q(x)$ beträgt dann $n - m$, während der Grad von $r(x)$ kleiner als der Grad des Nenners ausfällt. Um dann die Formel

$$f(x) = q(x) \cdot g(x) + r(x)$$

zu begründen, setzen wir

$$q(x) = q_0x^{n-m} + ... + q_{n-m} \qquad \text{und} \qquad r(x) = r_0x^{m-1} + ... + r_{m-1}$$

Damit wird die zu beweisende Formel gleichbedeutend mit den Gleichungen

$$\begin{aligned} a_0 &= b_0q_0 \\ a_1 &= b_0q_1 + b_1q_0 \\ &\vdots \\ a_{n-m} &= b_0q_{n-m} + ... + b_{n-m}q_0 \\ a_{n-m+1} &= b_1q_{n-m} + ... + b_{n-m+1}q_0 + r_0 \end{aligned}$$

$$\begin{aligned} a_{n-m+2} &= b_2 q_{n-m} + ... + b_{n-m+2} q_0 + r_1 \\ &\vdots \\ a_n &= b_m q_{n-m} + r_{m-1} \end{aligned}$$

Aus den ersten $n-m+1$ Gleichungen bestimmt man nacheinander die Koeffizienten von $q(x)$ und aus den darauffolgenden Gleichungen nacheinander die Koeffizienten von $r(x)$.

Die vier Grundrechnungsarten für Polynome gewinnt man nach allem so, dass man mit den Koeffizienten und der Unbestimmten wie mit ganzen Zahlen rechnet.
Dabei können die Koeffizienten selbst rationale, reelle oder komplexe Zahlen sein. Daraus erkennen wir, dass wir von der Beschaffenheit der Koeffizienten ganz absehen können, wenn nur für diese und die Unbestimmte die formalen Rechengesetze gelten. Dann aber müssen die vier Grundrechnungsarten für die Polynome selbst so erklärt werden, wie es eben geschehen ist.

Hat man einmal diesen abstrakten Standpunkt eingenommen, dann ist es angebracht, den Begriff des Körpers, der Galois bereits geläufig war, einzuführen.
Eine Menge von Elementen heißt Körper, wenn für die Elemente die vier Grundrechnungsarten definiert sind und ihre Anwendung auf die Körperelemente immer wieder Körperelemente liefert.
Die vier Grundrechnungsarten führen also nicht aus dem Körper hinaus. So bildet die Menge der rationalen Zahlen einen Körper, die Menge der ganzen Zahlen dagegen nicht.

Eine Abschwächung des Körperbegriffs ist der Ring. Die drei ersten Rechnungsarten führen aus dem Ring nicht hinaus; dagegen braucht die Division kein Ringelement zu liefern. Also ist jeder Körper ein Ring, nicht aber umgekehrt, denn die ganzen Zahlen bilden einen Ring. Die Gesamtheit von Polynomen mit Koeffizienten aus einem Körper bildet ebenfalls einen Ring, aber keinen Körper.

Der Umstand, dass der Grad des Restpolynoms kleiner ist als der Grad des Nenners, führt bei wiederholter Division - ähnlich wie bei ganzen Zahlen - zu einem Euklidischen Algorithmus und erlaubt, den größten gemeinsamen Teiler von zwei Polynomen zu bestimmen.
Dabei heißt das Polynom $g(x)$ Teiler von $f(x)$, wenn der Rest $r(x)$ verschwindet. Es ist aber zu beachten, dass mit $g(x)$ zugleich $c \cdot g(x)$ das Polynom $f(x)$ teilt, wenn c ein beliebiges Element aus dem Koeffizientenkörper ist. Zwei Polynome, die sich nur durch einen konstanten Faktor aus dem Koeffizientenkörper unterscheiden, heißen äquivalent.

In der Teilbarkeitslehre für Polynome entsprechen den Primzahlen die irreduziblen Polynome $p(x)$.
Sie lassen nur die Teiler c und $c \cdot p(x)$ zu, wobei c wieder ein Element aus dem Koeffizientenkörper bezeichnet. Es bleibt jedoch zu beachten, dass ein und dasselbe Polynom in einem Koeffizientenkörper irreduzibel, in einem anderen aber reduzibel, das heißt nicht irreduzibel, sein kann, wie etwa $x^2 + 1$.
Im Körper der reellen Zahlen ist dieses Polynom irreduzibel, denn die Zerlegung

$$x^2 + 1 = (x + a) \cdot (x + b)$$

in zwei Faktoren würde $a \cdot b = 1$ und $a + b = 0$ verlangen, woraus

$$b = \frac{1}{a} \qquad \text{und damit} \qquad a + \frac{1}{a} = 0$$

folgt. Multipliziert man in der letzten Gleichung beide Seiten mit a, so müsste $a^2 + 1 = 0$ sein, was jedoch für keine reelle Zahl a zutriftt.
Im Körper der komplexen Zahlen dagegen gilt die Zerlegung

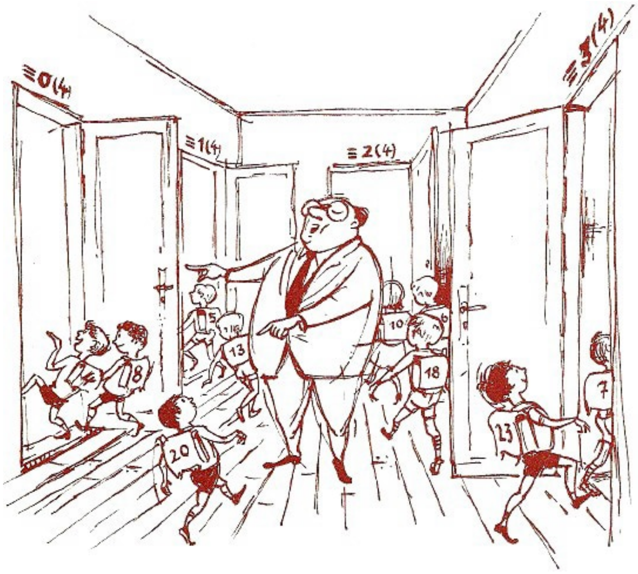
$$x^2 + 1 = (x + i) \cdot (x - i)$$

Aus der Gesamtheit äquivalenter Polynome lässt sich ein einziges durch die Forderung bestimmen, dass der Koeffizient des höchsten Gliedes die Einheit sei. Man spricht dann von einem normierten Polynom.
Für normierte Polynome gilt ein Analogon zur Darstellung von ganzen Zahlen als Produkt von Primzahlen: Jedes normierte Polynom lässt eine einzige Darstellung als Produkt von normierten irreduziblen Polynomen zu.

14 Kongruenzen

Wir sind so weit, Ringe und Körper zu konstruieren, die aus nur endlich vielen Elementen bestehen. Unser Vorgehen ist typisch für die moderne Algebra und kann als ein Keim dafür gelten.

m sei eine beliebige aber festgehaltene natürliche Zahl. Wir teilen die ganzen Zahlen in Klassen ein. Alle ganzen Zahlen, die bei Division durch m denselben Rest lassen, mögen zu einer Klasse zusammengefasst werden. Es gibt m Klassen entsprechend den Resten $0,1,...,m-1$, die als Restklassen bezeichnet werden. Für diese neuen Größen sollen nun die Grundrechnungsarten erklärt werden.



Zunächst gehören zwei ganze Zahlen, deren Differenz durch m teilbar ist, derselben Restklasse an. Es seien $a = r + km$ und $a' = r' + k'm$ die beiden Zahlen; dabei sind mit k und k' ganze Zahlen, mit r und r' Divisionsreste nach m bezeichnet. Aus $0 \leq r < m$ und $0 \leq r' < m$ folgt

$$0 \leq |r - r'| < m$$

im Falle $r - r' < 0$ gilt

$$0 < m + (r - r') < m$$

Wir dürfen $r + km > r' + k'm$ annehmen, woraus

$$(r + km) - (r' + k'm) = (r - r') + (k - k')m > 0$$

folgt. Der Rest nach m ist dann entweder $r - r'$ oder $m + (r - r')$, je nachdem $r - r' > 0$ oder < 0 ist. Daher kann m die Differenz $a + a'$ allein im Fall $r = r'$ teilen.
Nach dieser Feststellung sind wir in der Lage, eine Addition für die Restklassen A und B zu erklären. Dazu wähle man eine Zahl a bzw. b aus A bzw. B und addiere diese. Die Restklasse, der $a + b$ angehört, gelte als Summe der beiden Restklassen A und B. Zu dieser Definition ist man berechtigt, denn sind a' bzw. b' zwei weitere Zahlen aus A bzw. B, dann ist $a' = a + k_1m$ und $b' = b + k_2m$ und damit

$$(a' + b') - (a + b) = (a' - a) + (b' - b) = k_1m + k_2m = (k_1 + k_2)m$$

so dass nach dem vorhergehenden $a + b$ und $a' + b'$ derselben Restklasse angehören. Diese ist also trotz der Willkür in der Wahl von a und b eindeutig bestimmt.

Die Differenz von zwei Restklassen A und B wird als die Restklasse definiert, der $a-b$ angehört, und die Berechtigung dieser Erklärung ähnlich wie vorhin nachgewiesen.
Das Produkt von zwei Restklassen definiert man als die Restklasse, der $a \cdot b$ angehört. Man beachte dazu, dass mit $a' = a + k_1 m$ und $b' = b + k_2 m$ sofort

$$a'b' - ab = (a + k_1 m)(b + k_2 m) - ab = m(bk_1 + ak_2 + k_1 k_2 m)$$

folgt, so dass sich ab und $a'b'$ in derselben Restklasse befinden.
Addition, Subtraktion und Multiplikation von Restklassen genügen denselben formalen Regeln, wie sie für gewöhnliche Zahlen gelten, wie man sich mühelos Überzeugen kann. Daher bilden die Restklassen einen Ring, der aus insgesamt endlich vielen Elementen besteht.

Die gewonnenen Einsichten verhelfen zu einem Kalkül, der in der Teilbarkeitslehre mit Erfolg angewandt wird. m heißt dann Modul, und zwei Elemente a und a' aus derselben Restklasse kongruent nach dem Modul, in Zeichen

$$a \equiv a' \quad (m)$$

Ist weiter $b \equiv b' \quad (m)$, dann können wir unsere Einsichten in die Formeln

$$a + b \equiv a' + b' \quad (m)$$
$$a - b \equiv a' - b' \quad (m)$$
$$ab \equiv a'b' \quad (m)$$

kleiden, die jedoch den entscheidenden Schritt, Restklassen als neue Größen aufzufassen, nicht so recht verdeutlichen.

Namen und Begriff der Kongruenzen prägte bereits Goldbach 1732 in einer Abhandlung, die in den Berichten der Petersburger Akademie erschien. Bürgerrechte in der Mathematik verschaffte den Kongruenzen aber erst Gauß, der sie im ersten Kapitel seiner schon genannten Zahlentheorie systematisch behandelte.

Wir wollen mit Hilfe der Restklassen einen Satz von Fermat aus dem Jahre 1640 beweisen. Sind a und b natürliche Zahlen, dann gewinnt man eine Zerlegung ihres Produktes ab in Primfaktoren dadurch, dass man die Primfaktorzerlegungen von a und b miteinander multipliziert.
Wegen der Eindeutigkeit gewinnt man aber so "die" Primfaktorenzerlegung von ab. Daher können darin nur Primzahlen auftreten, die in mindestens einer der beiden Zahlen a und b vorkommen.

p sei eine Primzahl, welche die natürliche Zahl a nicht teilt, und r eine natürliche Zahl $0 < r < p$. Dann kann p das Produkt ra nicht teilen, weil p weder unter den Primfaktoren von r, noch unter den Primfaktoren von a vorkommt. Gilt daher für die natürlichen Zahlen r und r'

$$0 < r < r' < p$$

dann können ra und $r'a$ nach dem Modul p nicht kongruent sein, weil sonst $r'a - ra = (r - r')a$ mit $0 < r' - r < p$ durch p teilbar sein müsste. Die Divisionsreste der Zahlen $a, 2a, ..., (p-1)a$ nach dem Modul p sind folglich, von der Reihenfolge abgesehen, $1, 2, ..., p-1$, so dass nach dem Multiplikationssatz für Kongruenzen

$$a \cdot 2a \cdot ... \cdot (p-1)a \equiv 1 \cdot 2 \cdot ... \cdot (p-1) \quad (p)$$

gilt, mit anderen Worten

$$(a^{p-1} - 1) \cdot [1 \cdot 2 \cdot ... \cdot (p-1)]$$

durch p teilbar ist. Da das Produkt in der eckigen Klammer den Primfaktor p nicht enthält, muss

$$a^{p-1} - 1$$

durch p teilbar sein, in Zeichen

$$a^{p-1} \equiv 1 \quad (p)$$

Der eben bewiesene Satz heißt der ”kleine Fermat”, während der ”große Fermat” die auch heute noch nicht in vollem Umfang bewiesene Behauptung ist, dass die Gleichung

$$x^n + y^n = z^n$$

in ganzen Zahlen x, y und z nicht gelöst werden kann, wenn n eine natürliche Zahl größer als 2 bedeutet.
Fermat schrieb an den Rand eines Buches, er habe einen wunderbaren Beweis dafür gefunden. Nachdem ein Preis für einen Beweis ausgesetzt wurde, ging eine Lawine von Beweisversuchen los, die allesamt falsch waren. 1925 sah ich Kostproben davon in einem besonderen Fach des Mathematischen Instituts in Göttingen, über dem das von den Apothekern verwendete Kennzeichen für Gift prangte.

Das Rechnen mit Kongruenzen verhilft verhältnismäßig einfach zur Einsicht von Euler, dass die Fermatsche Zahl $2^{2^5} + 1$ keine Primzahl ist.
Allerdings wäre man ohne Eulers Vorarbeit kaum auf den Ansatz verfallen, von dem wir ausgehen. Es gilt

$$641 = 5^4 + 2^4 \qquad \text{und} \qquad 641 = 5 \cdot 2^7 + 1$$

Geht man zu Kongruenzen nach dem Modul 641 über, dann folgt aus der ersten Gleichung nach Multiplikation mit -2^{28}

$$0 \equiv -5^4 \cdot 2^{28} - 2^{32} \quad (641) \qquad (*)$$

Die zweite Gleichung verwandelt sich zunächst in die Kongruenz

$$0 \equiv 5 \cdot 2^7 + 1 \quad (641) \qquad \text{woraus} \qquad 5 \cdot 2^7 \equiv -1 \quad (641)$$

in die vierte Potenz erhoben

$$5^4 \cdot 2^{28} \equiv 1 \quad (641) \qquad (**)$$

folgt. Die beiden Kongruenzen (*) und (**) addiert ergeben

$$-2^{32} \equiv 1 \quad (641)$$

Da aber $32 = 2^5$ ist, besagt die letzte Kongruenz, dass $2^{2^5} + 1$ durch 641 teilbar ist.

Anschließend eine interessante Eigenschaft der Fermatschen Zahlen. Wir schicken die Bemerkung voraus, dass nach der Summenformel für die geometrische Progression

$$1 + 2 + 2^2 + ... + 2^{n-1} = 2^n - 1$$

ist. Für die beiden ersten Primzahlen ist $p_1 = 2 = 2^{2^0}$ und $p_2 = 3 < 2^{2^1}$. Wir nehmen an, dass für die darauffolgenden Primzahlen bis zur n-ten

$$p_k \leq 2^{2^{k-1}} \qquad \text{mit} \qquad k = 3, 4, ..., n$$

gilt. Die $(n+1)$-te Primzahl p_{n+1} ist aber, wie wir schon wissen, nach Euklid $\leq p_1 \cdot p_2 \cdot ... \cdot p_n + 1$, daher weiter

$$< 2^{2^0} \cdot 2^{2^1} \cdot ... \cdot 2^{2^{n-1}} = 2^{2^0+2^1+...+2^{n-1}} + 1$$

so dass also erst recht

$$p_{n+1} < 2^{2^n} + 1$$

sein muss: Die Folge der Fermatschen Zahlen majorisiert die aufsteigende Folge der Primzahlen. Daher majorisiert sie erst recht die aufeinanderfolgenden Lücken bei den Primzahlen, so roh auch diese Abschätzung sonst ist.

Zum Schluss noch die Konstruktion eines Körpers, der aus endlich vielen Elementen besteht. Solche Körper heißen Galois-Felder. Sind r und r' zwei nichtverschwindende Divisionsreste nach der Primzahl p, dann ist rr' durch p nicht teilbar, weil keiner der beiden Faktoren durch p teilbar ist. Da die Differenz von irgend zwei der Produkte

$$r \cdot 1, r \cdot 2, ..., r \cdot (p-1)$$

von der Form rr' ist, sind sie paarweise inkongruent modulo p. Daher sind ihre Divisionsreste nach p, von der Reihenfolge abgesehen, gleich $1, 2, ..., p-1$.

Bezeichnen wir irgendeinen der Reste aus dieser Zeile mit r'', gibt es also unter den Produkten der vorletzten Formelzeile ein ganz bestimmtes, wir wollen es rx nennen, das kongruent r'' modulo p ist, in Zeichen

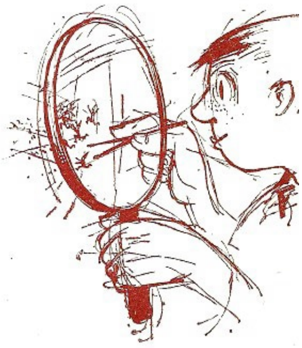
$$rx \equiv r'' \quad (p)$$

Für die entsprechenden Restklassen bedeutet das die Gleichung

$$R \cdot X = R''$$

Da r und r'' zwei beliebige unter den nichtverschwindenden Divisionsresten nach p waren, erkennt man, dass die Division immer ausführbar und daher der Ring der Restklassen nach einem Primzahlmodul p sogar ein Körper ist, der aus genau p Elementen besteht.

15 Grenzen



Im Rechnen mit Kongruenzen lernten wir soeben einen Kalkül kennen, den sein Initiator Gauß nicht nur eingeführt, sondern zugleich streng begründet hat. Die exakte Begründung ist keineswegs selbstverständlich, wie das Beispiel eines unvergleichlich wichtigeren Kalküls zeigt.

Wir meinen die Infinitesimalrechnung. Ihre Prinzipien wurden voneinander unabhängig und ziemlich gleichzeitig von Leibniz und Newton aufgestellt. Dabei verdankt man unsere heutigen Bezeichnungen Leibniz, der sie einführte, ohne selbst in der Lage zu sein, die eigenen Entwicklungen streng zu begründen.

Das gelang auch mehreren auf ihn folgenden Generationen von Mathematikern nicht.
Eine erste strenge Begründung glückte erst Cauchy 1821. Bis dahin half man sich, so gut es eben ging, durch Berufung auf unendlich kleine Größen, brave Heinzelmännchen, die letzten Endes unfassbar blieben. Die Erfolge häuften sich aber so, dass bereits d'Alembert den Ausspruch wagte: "Nur Mut, und der Glaube stellt sich ein!"

Aus den Aufzeichnungen von Leibniz wissen wir es auf den Tag genau, wann er der höheren Mathematik ihr endgültiges Gesicht gab.
Es war der 29. Oktober 1675, als er den Satz niederschrieb, in dem die beiden Wahrzeichen der Infinitesimalrechnung, die Zeichen für das Differential und das Integral, vorkommen. Leibniz war sich seiner Leistung voll bewusst, denn er schrieb:

"Bei der Bezeichnung ist darauf zu achten, dass sie für das Erfinden bequem sind. Dies ist am meisten der Fall, so oft sie das innere Wesen der Sache mit Wenigem ausdrücken und gleichsam abbilden.
So wird nämlich auf wunderbare Weise die Denkarbeit vermindert. Von solcher Beschaffenheit sind aber die Bezeichnungen, die ich angewandt habe, und durch die ich oft die schwierigsten Probleme auf wenigen Zeilen löse."

Auch der Beitrag von Cauchy zur Analysis ist recht hoch einzuschätzen. Welche Anstrengungen im Umdenken er erforderte, erkennt man daraus, dass Cauchy selbst nicht alle Feinheiten durchschaute, die aus seinen Präzisierungen von Fundamenten der Analysis folgten. Er erkannte nicht die Wichtigkeit des Sachverhaltes, den wir heute als gleichmäßige Stetigkeit bezeichnen, mit der wir uns noch auseinanderzusetzen haben.

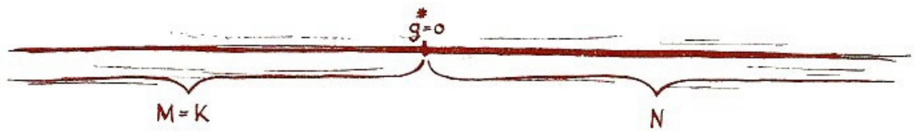
Abel erkannte dies, doch in seiner Unerfahrenheit berichtigte er Cauchys Vorgehen zu dessen Verdruss so nebenbei in einer Fußnote in einer seiner sieben Veröffentlichungen, die allesamt im ersten Jahrgang des 1826 gegründeten "Journals für die reine und angewandte Mathematik" des Berliner Baumeisters Crelle erschienen.

Es war dann noch immer ein weiter Weg, ehe man die Analysis so streng und zugleich durchsichtig darzustellen vermochte, wie es heute geschieht. Unsere nachfolgenden Ausführungen gelten alle dieser Darstellung. Dazu müssen wir zunächst die Begriffe Supremum und Infimum einführen. Diese Bezeichnungen fangen an, die älteren Benennungen obere und untere Grenze zu verdrängen. Dem sollen einige Erklärungen vorangehen.

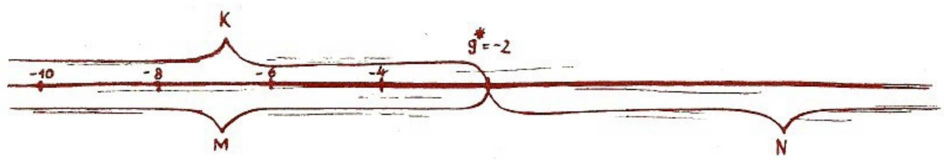
Wir betrachten Mengen von reellen Zahlen. Gibt es zu einer Menge M eine Zahl, die

größer ist als jedes Element von M, dann heißt M nach oben beschränkt. Gibt es zu M eine andere Zahl, die kleiner ist als jedes Element von M, dann heißt M nach unten beschränkt. Schließlich heißt M schlichtweg beschränkt, wenn M sowohl nach unten als auch nach oben beschränkt ist.
Es gibt Mengen, die nach oben nicht beschränkt sind, etwa die Menge der natürlichen Zahlen, Mengen, die nach unten nicht beschränkt sind, etwa die Menge der negativen ganzen Zahlen, und Mengen, die weder nach oben noch nach unten beschränkt sind, etwa die Menge aller ganzen Zahlen.

Ist K nach oben beschränkt und g eine Zahl größer oder nicht kleiner als jedes Element von K, dann heißt g eine obere Schranke von K. Offensichtlich ist dann jede Zahl, die größer als g ausfällt, selbst obere Schranke von K.
Daher bilden die oberen Schranken von K eine Menge N, die mit den übrigen reellen Zahlen, deren Menge mit M bezeichnet werde, einen Schnitt $(M|N)$ bildet. Zu jeder Zahl aus M gibt es ein Element von K, das mindestens so groß ist wie die Zahl selbst.
Die Zahl g, die den Schnitt $(M|N)$ erzeugt, ist die kleinste obere Schranke und heißt Supremum von K, abgekürzt $\sup K$. Kein Element von K kann größer als g^* sein, weil rechts von g^* lauter obere Schranken von K sind. Das Supremum von K kann einem Element von K gleich sein.
Trifft das nicht zu, dann muss in beliebiger Nähe links vom Supremum wenigstens ein Element von K liegen. Dann ist das Supremum Häufungspunkt von K. Schließlich ist auch der Fall noch möglich, dass das Supremum Häufungspunkt und zugleich eine Zahl aus K ist.
Wir wollen nun einige Beispiele betrachten. Die Elemente von K seien alle negativen reellen Zahlen. N umfasst dann alle positiven Zahlen und die Null, denn jede dieser Zahlen ist obere Schranke von K. M und K sind in diesem Falle identisch.



Den Schnitt $(M|N)$ bildet die Null, die selbst nicht zu K gehört, aber Häufungspunkt von K ist.



Enthält K alle geraden negativen Zahlen, so ist $\sup K = -2$ und gehört zu K, denn alle anderen Elemente von K sind kleiner als -2, während N alle reellen Zahlen enthält, die größer sind als jedes Element von K.
Für Mengen K, die nach unten beschränkt sind, ergibt sich eine sinngemäße Definition des Infimums, abgekürzt $\inf K$.

Ist K eine beschränkte unendliche Menge, dann wissen wir aus der ersten Plauderei, dass K mindestens einen Häufungspunkt besitzt, der aber kein Element von K zu Sein braucht. Diese Aussage lässt sich dahingehend verschärfen, dass es einen Häufungspunkt von K gibt, den Limes superior, abgekürzt $\limsup K$ oder auch $\overline{\lim} K$, so dass rechts von $\overline{\lim} K$ kein weiterer Häufungspunkt von K liegt.

N sei nämlich die Menge aller reellen Zahlen, von denen rechts immer nur höchstens endlich viele Elemente von K liegen, M die Menge aller übrigen reellen Zahlen. Der Schnitt $(M|N)$ wird gerade vom Limes superior l^* von K erzeugt. Ist nämlich l' eine Zahl links von l^*, dann gehört l' zu M, so dass rechts von l' unendlich viele Elemente von K liegen.
Ist dagegen l'' eine Zahl rechts von l^*, dann gibt es höchstens endlich viele größere Zahlen aus K. Das besagt, dass l^* selbst Häufungspunkt von K ist, und kein Häufungspunkt rechts von l^* liegen kann.

Entsprechend weist man die Existenz eines Limes inferior, abgekürzt $\underline{\lim} K$ nach, der selbst Häufungspunkt von K ist, und links von dem es keinen Häufungspunkt von K geben kann. Der Limes superior und der Limes inferior können zusammenfallen.

Als Beispiel für K sei die Folge $(-1)^n \cdot \dfrac{n-1}{n}$ genannt, für die $\overline{\lim} K = +1$ und $\underline{\lim} K = -1$ ist, wie sich auch aus der Skizze erkennen lässt.

Die beschränkte Menge K sei insbesondere eine Folge von reellen Zahlen. Der Spezialfall, in dem die beiden äußeren Häufungsstellen, der Limes inferior und der Limes superior, zusammenfallen, verdient besonderes Interesse. Dann gibt es nur einen einzigen Häufungspunkt, $l = \overline{\lim} K = \underline{\lim} K$, und die Folge heißt konvergent.
Dafür ist zunächst notwendig, dass, wie klein auch das Intervall J gewählt werde, das in seinem Inneren l enthält, außerhalb dieses Intervalls höchstens endlich viele Elemente von K liegen. Sonst würden in einem zweiten Intervall, das disjunkt zu J ist, unendlich viele Punkte von K liegen, die folglich einen Häufungspunkt besäßen, und zwar außerhalb von J. Damit hätte K zwei Häufungspunkte, entgegen der Voraussetzung.

Es ist zweckmäßig, eine auf Kowalewski zurückgehende Redeweise einzuführen. Besitzen alle Elemente einer Menge bis auf höchstens endlich viele eine Eigenschaft, dann sagt man, dass fast alle Elemente diese Eigenschaft besitzen. Damit können wir die vorhin als notwendig erkannte Bedingung so ausdrücken, dass J fast alle Punkte von K enthält. Diese Bedingung ist aber auch hinreichend. Denn außerhalb J kann dann kein Häufungspunkt von K liegen, innerhalb J muss dagegen mindestens ein Häufungspunkt liegen. Allein l kann jedoch Häufungspunkt sein, weil ja das Intervall J beliebig klein gewählt werden darf, wenn es nur l enthält.

Folgen mit nur einem einzigen Häufungspunkt nannten wir konvergente Folgen. Nach dem eben Bewiesenen kann man sie auch noch durch die Forderung erklären, dass fast alle ihre Elemente in beliebiger Nähe des Häufungspunktes liegen, der dann Limes oder Grenzwert der Folge heißt, mit anderen Worten, dass das Intervall J, das l im Inneren enthält und beliebig klein sein darf, alle Elemente der Folge mit höchstens endlich vielen Ausnahmen enthält.
Die höhere Mathematik, deren Boden wir soeben betreten haben, hat dafür eine einfache Schreibweise. Sie bedient sich der Epsilontik, die so benannt wurde nach der konsequenten Anwendung des griechischen Buchstabens ε.

Bezeichnen wir die Elemente der Folge mit $a_1, a_2, ..., a_n, ...$, so schreibt man den eben geschilderten Sachverhalt in der Form

$$|a_n - l| < \varepsilon \qquad \text{für} \qquad n > N$$

Die Ungleichungen sind dabei so zu verstehen, dass es zu beliebig kleinem positiven ε stets eine natürliche Zahl N gibt, so dass für jedes $n > N$ die Beziehung $|a_n - l| < \varepsilon$ gilt, weil ja fast alle a_k in beliebiger Nähe, also näher als ε, bei l liegen.
Um die Abhängigkeit der natürlichen Zahl N von ε anzudeuten, schreibt man gelegentlich $N(\varepsilon)$. Der Grenzwert l wird mit

$$\lim_{n\to\infty} a_n$$

bezeichnet, der ganze Sachverhalt mit

$$a_n \to l \qquad \text{für} \qquad n \to \infty$$

16 Folgen

Konvergente Folgen nennt man noch Fundamentalfolgen oder auch Cauchy-Folgen, weil Cauchy als erster die Begriffe präzisierte. Geht man von der Definition;

$$|a_n - I| < \varepsilon \qquad \text{für} \qquad n > N$$

aus, dann folgt sofort, dass die Menge $\{a_k\}$, deren Elemente die Glieder der Folge sind, beschränkt ist. Denn fast alle Elemente der Folge liegen im Sinne der Ungleichung im Intervall $J = (I - \varepsilon, I + \varepsilon)$, während die übrigen endlich vielen Elemente offensichtlich in einem endlichen Intervall untergebracht werden können.

Bei systematischen Betrachtungen über Fundamentalfolgen verdienen die Nullfolgen ein besonderes Augenmerk. Nullfolgen sind Fundamentalfolgen mit dem Grenzwert null.
Jede Fundamentalfolge lässt sich als Summe von zwei konvergenten Folgen schreiben, von denen die eine aus lauter gleichen Elementen vom Wert I besteht, die andere aber eine Nullfolge ist. Das folgt aus der Darstellung

$$a_n = I + (a_n - I)$$

Die Folge $(a_1 - I), (a_2 - I), ..., (a_n - I), ...$ ist eine Nullfolge,

$$a_n - I \to 0$$

denn diese Aussage ist gleichbedeutend mit der Behauptung

$$|(a_n - I) - 0| < \varepsilon \qquad \text{für} n > N$$

die wegen $a_n \to I$ zutrifft.
Multipliziert man die Glieder einer Nullfolge $a_1, a_2, ...$ mit den entsprechenden Gliedern einer zweiten Folge $b_1, b_2, ...$, die nicht konvergent zu sein braucht, wenn sie nur beschränkt ist, $|b_n| < k$, dann ist auch die Folge

$$a_1b_1, a_2b_2, ..., a_nb_b, ...$$

eine Nullfolge. Es gilt nämlich zunächst die Abschätzung $|a_nb_n| < |a_n| \cdot k$.
Wählt man als beliebig kleine Größe statt ε diesmal $\frac{\varepsilon}{k}$, dann folgt weiter aus

$$|a_n| < \frac{\varepsilon}{k} \quad \text{für} \quad n > N\left(\frac{\varepsilon}{k}\right) \qquad \to \qquad |a_n| \cdot k < \frac{\varepsilon}{k} \cdot k = \varepsilon \quad \text{für} \quad n > N\left(\frac{\varepsilon}{k}\right)$$

Fassen wir die beiden Ungleichungen zusammen, so ergibt sich

$$|a_nb_n| < \varepsilon \qquad \text{für} \qquad n > N$$

das heißt, dass a_nb_n wirklich das allgemeine Glied einer Nullfolge ist.
Weiter ist die Summe von zwei Nullfolgen $a_1, a_2, ...$ und $a'_1, a'_2, ...$ ebenfalls eine Nullfolge, weil aus

$$|a_n| < \frac{\varepsilon}{2} \quad \text{für} \quad n > N\left(\frac{\varepsilon}{2}\right) \qquad \text{und} \qquad |a'_n| < \frac{\varepsilon}{2} \quad \text{für} \quad n > N'\left(\frac{\varepsilon}{2}\right)$$

sofort

$$|a_n + a'_n| \leq |a_n| + |a'_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{für} \qquad n > \max\{N\left(\frac{\varepsilon}{2}\right), N'\left(\frac{\varepsilon}{2}\right)\}$$

folgt.
Die letzten zwei Sätze gestatten das Rechnen mit Fundamentalfolgen zu begründen. Bezeichnen jetzt $a_1, a_2, ...$ und $a_1', a_2', ...$ nicht mehr die speziellen Nullfolgen von vorhin, sondern zwei beliebige konvergente Folgen mit den Grenzwerten I und I', dann gelten die Sätze

$$a_n + a_n' \to I + I'; \qquad a_n - a_n' \to I - I'; \qquad a_n \cdot a_n' \to I \cdot I'; \qquad \frac{a_n}{a_n'} \to \frac{I}{I'}$$

Für die letzte Gleichung muss allerdings die Einschränkung gemacht werden, dass keiner der Nenner verschwindet.
Aus der ersten Gleichung folgt, dass $(a_n+a_n')-(I+I')$ das allgemeine Glied einer Nullfolge ist; denn '

$$(a_n + a_n') - (I + I') = (a_n - I) + (a_n' - I')$$

und rechts steht die Summe von zwei Nullfolgen. Ähnlich führt man den Nachweis für die Subtraktion. Um den Multiplikationssatz nachzuweisen, schreibe man

$$a_n a_n' - II' = (a_n - I)a_n' + (a_n' - I')I$$

und bemerke, dass rechts die Summe von zwei Nullfolgen steht. Schließlich folgt die Regel für die Division aus der Umformung

$$\frac{a_n}{a_n'} - \frac{I}{I'} = \frac{(a_n - I)I' - (a_n' - I')I}{a_n' I'}$$

durch wiederholte Anwendung der beiden Sätze über Nullfolgen, die zur Einsicht führen, dass die rechte Seite tatsächlich eine Nullfolge darstellt.

Abschließend sei noch eine Kennzeichnung von Fundamentalfolgen, das Konvergenzkriterium von Cauchy, nachgewiesen. Aus der Konvergenz der Folge $a_1, a_2, ...$ gegen den Grenzwert I folgt für $n > N\left(\frac{\varepsilon}{2}\right)$ und $m > N\left(\frac{\varepsilon}{2}\right)$ die Abschätzung

$$|a_n - a_m| = |(a_n - I) - (a_m - I)| \leq |a_n - I| + |a_m - I| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Umgekehrt ist letztere Bedingung auch hinreichend für die Konvergenz der Folge $a_1, a_2, ...$.
Man wähle nämlich für ε der Reihe nach die Werte $\frac{1}{2}$, $\frac{1}{2^2}$, ..., $\frac{1}{2^k}$, ... Nach Voraussetzung gehört dann zu $\frac{1}{2^k}$ ein N_k, so dass

$$|a_n - a_m| < \frac{1}{2^k} \qquad \text{ist für} \qquad n > N_k \quad \text{und} \quad m > N_k$$

Für die sukzessiven Werte von k wähle man $m_1 < m_2 < ...$, so dass dabei auch noch die Forderung $m_k > N_k$ erfüllt ist. Dann gilt

$$|a_n - a_m| < \frac{1}{2^k} \qquad \text{für} \qquad n > m_k$$

Zunächst enthält das Intervall $\left(a_{m_1} - \frac{1}{2}, a_{m_1} + \frac{1}{2}\right)$ alle a_n mit $n > m_1$, insbesondere also a_{m_2}. Der Durchschnitt J' dieses Intervalls mit $\left(a_{m_2} - \frac{1}{2^2}, a_{m_2} + \frac{1}{2^2}\right)$ enthält alle a_n mit $n > m_2 > m_1$, insbesondere also a_{m_3}. Der Durchschnitt J'' von J' mit $\left(a_{m_3} - \frac{1}{2^3}, a_{m_3} + \frac{1}{2^3}\right)$

enthält alle a_n mit $n > m_3$, insbesondere also a_{m_4}.

So fortfahrend gewinnt man eine Intervallschachtelung $\left(a_{m_1} - \frac{1}{2}, a_{m_1} + \frac{1}{2}\right)$, J', J'', ... mit Intervallen, die auf einen Punkt I zusammenschrumpfen.
I ist dann der Grenzwert der Folge $a_1, a_2, \ldots$ Zunächst liegt I in jedem der sukzessiven Intervalle, daher auch im k-ten Intervall von der Länge $\frac{1}{2^{k-1}}$, in dem alle a_n mit $n > m_k$ liegen. Der Abstand aller dieser a_n von I beträgt daher weniger als die Intervalllänge

$$|a_n - I| < \frac{1}{2^{k-1}} \qquad \text{für} \qquad n > m_k$$

17 Wie die Welt nicht erschaffen wurde

Konvergente Folgen verhelfen dazu, unendliche Summen einzuführen. Liegt die Folge $a_1, a_2, ...$ vor, dann bilde man die sukzessiven Partialsummen

$$\begin{aligned} a_1 &= s_1 \\ a_1 + a_2 &= s_2 \\ a_1 + a_2 + a_3 &= s_3 \\ &\vdots \\ a_1 + a_2 + ... + a_n &= s_n \\ &\vdots \end{aligned}$$

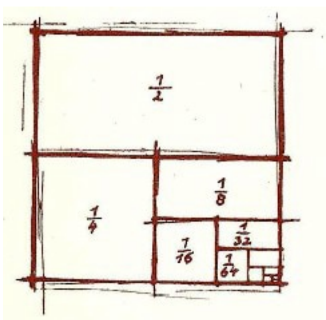
Konvergiert die Folge dieser Partialsummen, die man auch Reihe nennt und in vereinfachter Form $a_1 + a_2 + ...$ schreibt, gegen den Grenzwert s, dann heißt s Wert der unendlichen Summe aus den Zahlen der Folge $a_1, a_2, ...$:

$$s = a_1 + a_2 + a_3 + ...$$

Wir wollen uns dies an einem Beispiel veranschaulichen. Ein Quadrat mit dem Flächeninhalt 1 werde fortgesetzt halbiert. Es entstehen Flächenstücke der Größe

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, ...$$

Die Summe dieser unendlich vielen Flächenstücke erhalten wir als Grenzwert der Folge von Partialsummen:



$$\begin{aligned} s_1 &= \frac{1}{2} \\ s_2 &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} = \frac{2^2 - 1}{2^2} = 1 - \frac{1}{2^2} \\ s_3 &= \frac{3}{4} + \frac{1}{8} = \frac{7}{8} = \frac{2^3 - 1}{2^3} = 1 - \frac{1}{2^3} \\ &\vdots \\ s_n &= 1 - \frac{1}{2^n} \end{aligned}$$

$$\text{Grenzwert: } s = \lim_{n\to\infty} \left(1 - \frac{1}{2^n}\right) = 1$$

Die Summe der Flächenstücke stimmt also mit der Fläche des Quadrates überein, was ja auch unserer Anschauung entspricht. Das gleiche Ergebnis erhalten wir mit Hilfe der bereits aus der Schule bekannten Summenformel für die geometrische Folge

$$a + ax + ax^2 + ... = \frac{a}{1-x} \qquad ; \qquad s = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1$$

Das Nichtbeachten der Konvergenzforderung trieb in früheren Zeiten mitunter skurrile Blüten. Als die Begriffe noch nicht geklärt waren, wurden führende Forscher durch ihren Genius vor groben Fehlern bewahrt, nicht aber die anderen Gelehrten.

Der Mönch Guido Grandi, Zeitgenosse von Leibnitz, beachtete nicht, dass die geometrische Folge nur für $|x| < 1$ die Summe

$$1 + x + x^2 + x^3 + ... = \frac{a}{1-x}$$

liefert. Er setzte für x den Wert -1 ein und erhielt

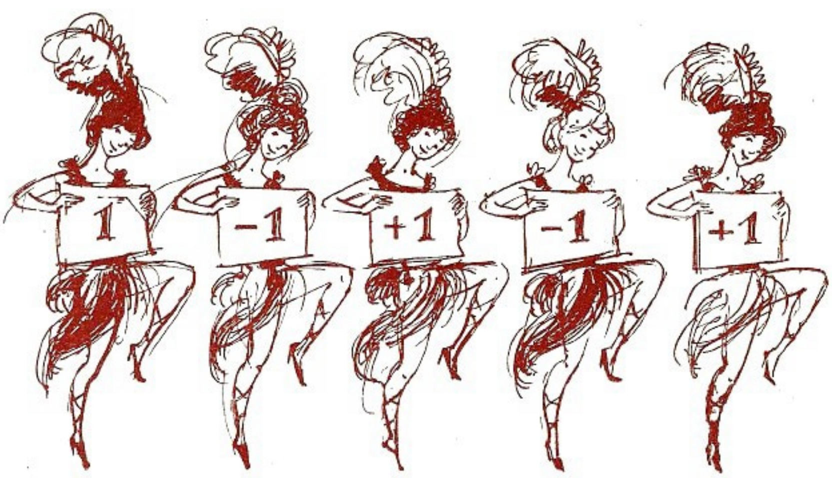
$$1 - 1 + 1 - 1 + ... = \frac{1}{2}$$

Nun darf man in einer unendlichen Summe, falls, sie sinnvoll ist, Summanden nach Belieben zusammenfassen.

Dadurch entsteht nämlich eine Teilfolge der ursprünglichen Partialsummen, die nunmehr Folge der Partialsummen der abgeänderten unendlichen Summe ist, und die Teilfolge konvergiert ja gegen den Grenzwert der ursprünglichen Folge.
Grandi fasste die Summanden paarweise zusammen,

$$(1 - 1) + (1 - 1) + ... = 0 + 0 + ...$$

und meinte, weil der Wert $\frac{1}{2}$ der unendlichen Summe sich nicht geändert habe, also $0 + 0 + ... = \frac{1}{2}$ gelten sollte, den göttlichen Schöpfungsakt aus dem Nichts erklärt zu haben.



Die Partialsummen der Reihe

$$1 - 1 + 1 - 1 + ...$$

nehmen abwechselnd die Werte 1 und 0 an, bilden also keine konvergente Folge. Um die Erklärung von unendlichen Summen auch auf solche Fälle auszudehnen, bilde man nach Hölder aus der Folge der Partialsummen $s_1, s_2, ...$ die neue Folge

$$\sigma_1 = s1; \qquad \sigma_2 = \frac{s_1 + s_2}{2}; \qquad ...; \qquad \sigma_n = \frac{s_1 + s_2 + ... + s_n}{n}; \qquad ...$$

Konvergiert die Folge $\sigma_1, \sigma_2, ...$, dann heißt ihr Grenzwert σ verallgemeinerte Summe der Reihe $a_1 + a_2 + ...$
Diese Verallgemeinerung umfasst die frühere Definition als Spezialfall. Um das einzusehen, formt man σ_n wie folgt um:

$$\begin{aligned}\sigma_n &= \frac{s_1 + ... + s_m}{n} + \frac{s_{m+1} + ... + s_n}{n} \\ &= \frac{s_1 + ... + s_m}{n} + \frac{(s_{m+1} - s) + ... + (s_n - s)}{n} + \frac{n-m}{n}s \\ &= \frac{s_1 + ... + s_m}{n} + \frac{(s_{m+1} - s) + ... + (s_n - s)}{n} - \frac{m}{n}s + s\end{aligned}$$

Konvergiert nun s_k gegen den Grenzwert s, dann muss für $l > N$

$$|s_l - s| < \frac{\varepsilon}{3}$$

sein, insbesondere für $l = m+1, ..., n$. Aus der letzten Gleichung folgt damit

$$|\sigma_n - s| < \left|\frac{s_1 + ... + s_m}{n}\right| + \frac{n-m}{n} \cdot \frac{\varepsilon}{3} + \frac{m}{n}|s| < \left|\frac{s_1 + ... + s_m}{n}\right| + \frac{\varepsilon}{3} + \frac{m}{n}|s|$$

Nun werde n außerdem so groß gewählt, dass der erste und der dritte Summand rechts kleiner als $\frac{\varepsilon}{3}$ ausfallen. Für so große n gilt dann $|\sigma_n - s| < \varepsilon$, daher konvergiert für $s_k \to s$ wie behauptet σ_n gegen denselben Grenzwert $\sigma_n \to s$.

Die Bildung der arithmetischen Mittel σ_n von den Partialsummen s_n heißt Summationsverfahren. Wir wollen sehen, zu welchem Ergebnis das im Falle von $1 - 1 + 1 - 1 + ...$ führt. Die Partialsummen mit ungeradem Index besitzen den Wert 1, die mit geradem Index den Wert 0. Daher ist

$$\begin{aligned}\sigma_{2n} &= \frac{s_1 + ... + s_{2n}}{2n} = \frac{n \cdot 1 + n \cdot 0}{2n} = \frac{1}{2} \qquad \text{und} \\ \sigma_{2n+1} &= \frac{(n+1) \cdot 1 + n \cdot 0}{2n+1} = \frac{n + \frac{1}{2}}{2n+1} + \frac{\frac{1}{2}}{2n+1} = \frac{1}{2} + \frac{1}{4n+2}\end{aligned}$$

Die Folge $\sigma_1, \sigma_2, ...$ konvergiert folglich gegen den Grenzwert $\frac{1}{2}$, obschon die $s_1, s_2, ...$ selbst noch nicht konvergieren.

Außer der angegebenen wurden noch andere Summationsverfahren ersonnen, die immer mehr Folgen $a_1, a_2, ...$ verallgemeinerte Summen zuweisen konnten. Das von uns vorhin erörterte Summationsverfahren aus dem Jahre 1882 erlaubte Lebesgue 1905 einen grundlegenden Satz von abschließendem Charakter in der Theorie der Fourierreihen aufzustellen, einer Theorie, die auch für Physiker von grundsätzlicher Bedeutung ist.
Weiterhin wollen wir uns jedoch nur mit den zuerst eingeführten unendlichen Summen beschäftigen.

Eine berühmte Reihe, die einen eigenen Namen führt, die harmonische Reihe,

$$1 + \frac{1}{2} + \frac{1}{3} + ...$$

besitzt keine Summe, wie wir diese anfangs definiert haben, denn ihre Partialsummen werden ständig größer, nehmen monoton zu, $s_1 < s_2 < s_3 < ...$, und sind fast alle größer als eine beliebig vorgegebene noch so große Zahl g.
Um das einzusehen, schätzen wir die speziellen Partialsummen s_{2^k} ab. Es gilt

$$\begin{aligned}s_{2^{k+1}} &= s_{2^k} + \frac{1}{2^k+1} + \frac{1}{2^k+2} + ... + \frac{1}{2^{k+1}} > s_{2^k} + \frac{1}{2^{k+1}} + \frac{1}{2^{k+1}} + ... + \frac{1}{2^{k+1}} \\ &= s_{2^k} + \left(2^{k+1} - 2^k\right) \cdot \frac{1}{2^{k+1}} = s_{2^k} + \frac{1}{2}\end{aligned}$$

Damit folgt

$$s_{2^{k+m}} > s_{2^{k+m-1}} + \frac{1}{2} > s_{2^{k+m-2}} + \frac{1}{2} + \frac{1}{2} > ... > s_{2^k} + \frac{m}{2}$$

Ist m groß genug, dann steht rechts eine Zahl größer als g, wie behauptet. Infolge der Monotonie der Folge der Partialsummen $s_1, s_2, ...$ ist damit unsere Behauptung, die als $s_n \to \infty$ geschrieben wird, in vollem Umfang bewiesen. Man sagt, die harmonische Reihe divergiert.

Es sei noch eine bemerkenswerte Eigenschaft der harmonischen Reihe behandelt. Keiner ihrer Abschnitte, die aus endlich vielen aufeinanderfolgenden Summanden bestehen,

$$\frac{1}{m} + \frac{1}{m+1} + ... + \frac{1}{m+n}$$

kann eine ganze Zahl sein, mit Ausnahme von $s_1 = 1$.
Zunächst liegt zwischen zwei natürlichen Zahlen, die durch dieselbe Potenz 2^k teilbar sind,

$$2^k \cdot t < 2^k \cdot v$$

wo t und v ungerade Zahlen bezeichnen, eine weitere natürliche Zahl, die durch eine höhere Potenz von 2 teilbar ist. Infolge $t < v$ liegt nämlich zwischen den ungeraden Zahlen t und v mindestens eine gerade Zahl, $2u$. Daher gilt

$$2^k \cdot t < 2^k \cdot 2u = 2^{k+1} \cdot u < 2^k \cdot v$$

Infolgedessen befindet sich unter den Zahlen m bis $m+n$ eine einzige, m', die durch eine maximale Potenz von 2 teilbar ist.
Sollte es nämlich zwei geben, dann müsste zwischen den beiden eine weitere Zahl liegen, die durch eine noch höhere Potenz von 2 teilbar wäre. 2^k sei diese maximale Potenz. Der größte gemeinsame Nenner der Stammbrüche $\frac{1}{m}$ bis $\frac{1}{m+n}$ ist dann ein Produkt von 2^k und einer ungeraden Zahl q. Nun kann jede der Zahlen m bis $m+n$ in der Form $2^j \cdot r$ geschrieben werden, wobei j einer der Werte 0 bis k und r eine ungerade Zahl ist.
Bringt man die Stammbrüche $\frac{1}{m}$ bis $\frac{1}{m+n}$ auf den gleichen gemeinsamen Nenner, so liefert jeder einen Beitrag

$$\frac{2^{k-j} \cdot \frac{q}{r}}{2^k \cdot q}$$

Dabei fällt $\frac{q}{r}$ ungerade aus, weil q ungerade ist. j nimmt allein für m' den Wert k an, sonst aber einen kleineren Wert. Daher sind die Zähler lauter gerade Zahlen bis auf den Beitrag von m'.
Der Gesamtzähler wird folglich ungerade, während der gemeinsame Nenner eine gerade Zahl ist, weshalb der Quotient unmöglich eine ganze Zahl sein kann.

Wir wollen uns der Frage zuwenden, wann die unendliche Summe $a_1 + a_2 + ...$ kommutativ ist.
Das ist genau dann der Fall, wenn die aus den Absolutwerten der Folge gebildete unendliche Reihe eine Summe besitzt, wie wir sie anfangs für beliebige Folgen definierten. Die Reihe selbst heißt dann absolut konvergent.
Zunächst folgt aus der absoluten Konvergenz die Konvergenz. In Worten klingt das trivial, ist es aber nicht. Man kann so schließen. Werden die Partialsummen der Folge der Absolutbeträge $|a_1|$, $|a_2|$, ... mit s'_1, s'_2, ... bezeichnet, ist die Konvergenz der letzten Folge nach dem Konvergenzkriterium von Cauchy mit

$$|s'_{k+m} - s'_k| < \varepsilon \qquad \text{für} \qquad k > N \quad \text{und} \quad k+m > N$$

gleichbedeutend. Die Differenz

$$s'_{k+m} - s'_k < |a_{k+1}| + ... + |a_{k+m}|$$

majorisiert aber die Summe $a_{k+1} + ... + a_{k+m}$, die ihrerseits gleich $s_{k+m} - s_k$ ist. Erneut im Sinne des Konvergenzkriteriums von Cauchy ist aber die Ungleichung

$$|s_{k+m} - s_k| < \varepsilon$$

gleichbedeutend mit der Konvergenz der Partialsummen $s_1, s_2, ...$. Die Umkehrung braucht nicht zu stimmen, wie das Verhalten der alternierenden Reihe

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

verrät. Für diese ist

$$s_{k+m} = \frac{(-1)^k}{k+1} + \frac{(-1)^{k+1}}{k+2} + ... + \frac{(-1)^{k+m-1}}{k+m} = (-1)^k \left[\frac{1}{k+1} - \frac{1}{k+2} + ... \pm \frac{1}{k+m}\right]$$

Da der Wert des Ausdrucks in der eckigen Klammer bestimmt kleiner als $\frac{1}{k+1}$ ist, gilt

$$|s_{k+m} - s_k| < \frac{1}{k+1}$$

so dass im Sinne des Konvergenzkriteriums von Cauchy unsere alternierende Reihe konvergiert. Sie konvergiert dagegen nicht absolut, denn die Absolutbeträge bilden die harmonische Reihe, von der wir bereits wissen, dass sie divergiert.

Liegt eine monoton zunehmende Folge vor, die beschränkt ist, dann konvergiert sie gegen ihr Supremum. Denn laut Definition muss links vom Supremum in beliebiger Nähe ein Glied der Folge liegen, wegen der Monotonie also auch alle darauffolgenden Glieder.
Diese Bemerkung lässt erkennen, dass eine absolut konvergente Reihe, wie man auch die Reihenfolge der Summanden darin abändert, immer absolut konvergiert und damit schlechthin konvergiert. Wir behaupten nun, dass sich auch die Reihensumme bei der Umordnung nicht geändert hat.

Die Folge $a_1, a_2, ...$ werde nach Belieben umgeordnet, und die Partialsummen der umgeordneten Folge mögen mit $s'_1, s'_2, ...$ bezeichnet werden. Wie groß auch n gewählt wurde, gibt es eine erste unter den letztgenannten Partialsummen, s'_ν, die a_1 bis a_n enthält. Das trifft dann auch für die nachfolgenden Partialsummen $s'_{\nu+1}, ...$ zu. In der Differenz $s'_j - s_n$, wo $j \geq \nu$ ist, heben sich also die a_1 bis a_n alle weg.
Nun sei die Reihe $a_1 + a_2 + ...$ absolut konvergent. Nach dem Konvergenzkriterium von Cauchy ist das gleichbedeutend mit

$$|a_{k+1}| + ... + |a_{k+m}| < \frac{\varepsilon}{2} \qquad \text{für} \qquad k > N$$

Wenn dann k auch noch größer als ν ist, folgt auf Grund der letzten Ungleichung erst recht

$$|s'_k - s_n| < \frac{\varepsilon}{2}$$

Da außerdem

$$|s_n - l| < \frac{\varepsilon}{2} \qquad \text{für} \qquad n > N'$$

ist, folgt weiter

$$|s'_k - l| = |(s'_k - sn) + (s_n - l)| \leq |s'_k - s_n| + |s_n - l| < \varepsilon \qquad \text{für} \qquad k > \max(\nu, N, N')$$

was mit $s'_k \to l$ gleichbedeutend ist.

Wir wollen noch nachweisen, dass in einer konvergenten aber nicht absolut konvergenten Reihe $a_1 + a_2 + ...$ sowohl die aus den positiven Gliedern bestehende Reihe als auch die aus den Beträgen der negativen Glieder bestehende Reihe divergieren. Wir schicken dazu die Bemerkung voraus, dass für positive a

$$a = \frac{|a| + a}{2}$$

ist, während die rechte Seite dieses Ausdrucks für negative a verschwindet. Für letztere stellt dagegen der Ausdruck $\frac{|a| - a}{2}$ den Betrag von a dar, während dieser Ausdruck für positive a verschwindet. Damit schreibt sich die Reihe der positiven Glieder als

$$\frac{|a_1| + a_1}{2} + \frac{|a_2| + a_2}{2} + ...$$

und die Reihe von den Beträgen der negativen Glieder als

$$\frac{|a_1| - a_1}{2} + \frac{|a_2| - a_2}{2} + ...$$

Zunächst können nicht beide Reihen konvergieren, weil sonst auch ihre Summe $|a_1| + |a_2| + ...$ konvergieren würde, entgegen der Voraussetzung. Weiter kann nicht die erste Reihe divergieren, die zweite aber konvergieren, weil sonst ihre Differenz, $a_1 + a_2 + ...$, nicht konvergieren könnte. Ebenfalls kann nicht die erste Reihe konvergieren, die zweite divergieren. Es bleibt folglich nur das behauptete Verhalten der beiden Teilreihen denkbar.

Mit Hilfe dieser Einsicht konnten Dirichlet und Riemann nachweisen, dass jede konvergente aber nicht absolut konvergente Reihe so umgeordnet werden kann, dass die neue Reihensumme eine beliebig vorgegebene Zahl wird. Darauf wollen wir nicht weiter eingehen, wohl aber auf die Leistungen von Lebesgue und von Riemann.

Die Analysis, insbesondere Funktionentheorie, und die Differentialgeometrie verdanken Riemann schöpferische Gedanken von großer Tragweite. Sein Beitrag zur Funktionentheorie war jedoch für die Zeitgenossen sehr schwer verständlich; viele der von ihm aufgeworfenen Fragen, die ganze Theorien andeuten, sind bis heute noch unentschieden. Hinzu kam, dass er in einer grundlegenden Überlegung Supremum dem Maximum gleichsetzte, ein Fehler, der infolge der Kritik von Weierstraß das Wirken seines Gedankengutes beeinträchtigte.
Die ganze Laufbahn von Riemann spielte sich in Göttingen ab, vom Studium bis Zur Professur. Er war immer leidend und starb 40jährig 1866 auf einer Erholungsreise in Italien.

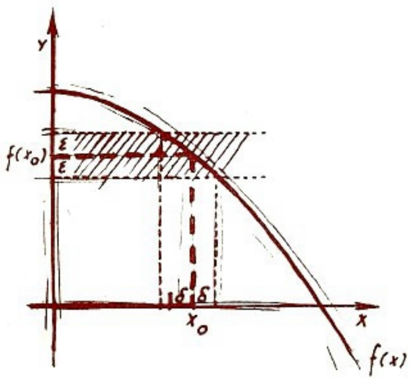
Lebesgue ist der eigentliche Vater unserer Theorie der reellen Funktionen. Wenn auch an dem nach ihm benannten Maßbegriff Borel entscheidend beteiligt war, prägte Lebesgue das ebenfalls nach ihm benannte Integral, das den Zentralbegriff der neuen Theorie bildet und weitestgehende Konsequenzen hat, die er selber zu ziehen vermochte.

Er führte es in seiner Doktorarbeit 1902 ein, die er - für einen Franzosen ungewöhnlich spät - mit 27 Jahren schrieb. Dafür bedeutet aber die Arbeit einen Meilenstein in der modernen Mathematik.

18 Stetigkeit verpflichtet

”Ein steter Tropfen höhlt den Stein” heißt es zu unrecht, weil ein stetiger Tropfen kein Tropfen wäre, sondern ein Strahl. Damit möchte ich aber nicht gleich gegen alle Sprichwörter Sturm laufen, denn es gibt darunter auch solche, die des Pudels Kern treffen, wie ”Hunde, die bellen, beißen nicht”. Sicher nicht, denn solange sie hellen, können sie unmöglich zubeißen.
Kein noch so gelungenes Sprichwort enthebt aber den Mathematiker der Notwendigkeit, einwandfreie Erklärungen abzugeben.

Was versteht er also insbesondere unter stetig? Die Frage ließe sich mit Hilfe von konvergenten Folgen beantworten. Da wir jedoch keine Systematik treiben wollen, geben wir mit der Epsilontik eine unmittelbare Antwort.

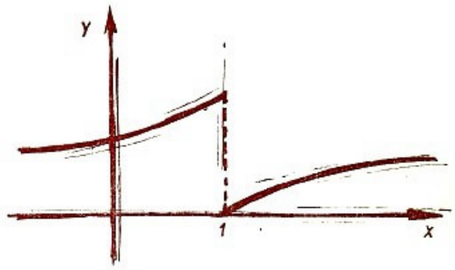


Die Funktion $f(x)$ sei in der, Umgebung der Stelle x_0 erklärt. Sie heißt stetig in x_0, wenn

$$|f(x) - f(x_0)| < \varepsilon \qquad \text{für} \qquad |x - x_0| < \delta$$

ist. Zu beliebig vorgegebenem positivem ε soll es also immer eine positive Zahl δ geben, so dass für alle Punkte x, die näher als δ bei x_0 liegen, der Funktionswert $f(x)$ von dem Funktionswert im Punkt x_0 um weniger als ε abweicht.

Bei einer stetigen Funktion lässt sich also zu jedem vorgegebenen ε-Intervall um $f(x_0)$ ein δ-Intervall um x_0 finden, so dass für alle x innerhalb dieses δ-Intervalls die zugehörigen Funktionswerte $f(x)$ in das vorgegebene ε-Intervall fallen.



Geometrisch bedeutet diese Aussage, dass die Funktion in der Nähe der Stelle x_0 innerhalb eines beliebig schmalen, zur x-Achse parallelen Streifens verläuft, wenn man nur ein genügend kleines Intervall um x_0 nimmt.
Das ist durchaus nicht immer der Fall. Bei der Funktion

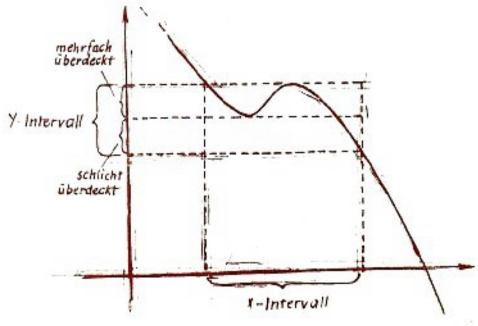
$$f(x) = \frac{1}{1 + e^{\frac{1}{x-1}}}$$

die nebenstehende Gestalt hat, lässt sich für $x_0 = 1$ um $f(x_0)$ kein beliebig schmaler ε-Streifen angeben, für das sich ein δ-Intervall mit den genannten Bedingungen finden ließe.

Die allgemeine Erklärung für die Stetigkeit setzt unseren abstrakten Funktionsbegriff voraus, der auf Dirichlet zurückgeht.
Die Grundlage bildet die Abbildung einer Menge X auf eine zweite Menge Y. Das ist so zu verstehen, dass jedem Element x von X ein Element y von Y, das Bild, zugeordnet ist. Wir arbeiten hier mit dem Spezialfall, dass X ein Intervall ist.

Aus der Stetigkeit der Funktion $f(x)$ in jedem Punkt des Intervalls, kurz, aus der Stetigkeit

von $f(x)$ in J, werden wir folgern können, dass Y selbst ein Intervall ist. Dabei kann es vorkommen, dass in verschiedenen Punkten von J derselbe Funktionswert auftritt, so dass das Y-Intervall also nicht schlicht, sondern mehrfach überdeckt wird.
Dazu möchten wir noch bemerken, dass man vor dem Auftreten Dirichlets unter Funktionen nur algebraische, später auch noch transzendente Ausdrücke verstanden hat.



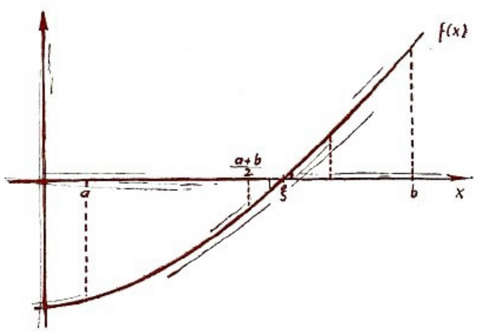
Die Funktion $f(x)$ sei im Intervall $J = [a, b]$ definiert und in jedem Punkt von J stetig, kurz, in J stetig. Nimmt diese Funktion in den beiden Endpunkten a und b Werte verschiedenen Vorzeichens an, dann betrachte man die beiden Intervallhälften

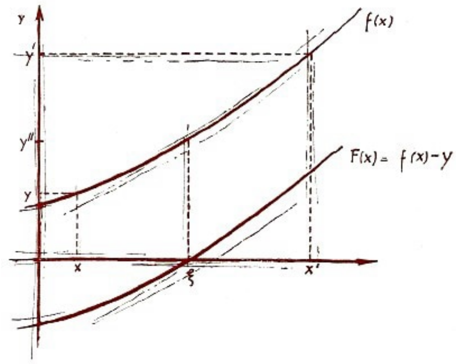
$$\left[a, \frac{a+b}{2}\right] \text{ und } \left[\frac{a+b}{2}, b\right].$$

Verschwindet $f(x)$ in $\frac{a+b}{2}$ nicht, dann nimmt $f(x)$ in einer der beiden Intervallhälften in den Endpunkten wieder Werte verschiedenen Vorzeichens an.

Man halbiere dann diese Intervallhälfte erneut und wiederhole das. So gewinnt man eine Intervallschachtelung, die auf einen Punkt ξ zusammenschrumpft, es sei denn, die Wiederholung scheitert daran, dass die Funktion in einem der Teilpunkte verschwindet.
War das nicht der Fall, dann muss sie in ξ verschwinden. Sonst müsste sie in hinreichender Nähe von ξ infolge der Stetigkeit lauter Werte vom selben Vorzeichen wie in ξ haben, entgegen der Konstruktion von ξ. Da $f(x)$ laut Voraussetzung in a und b nicht verschwindet, muss ξ im Inneren von J liegen.

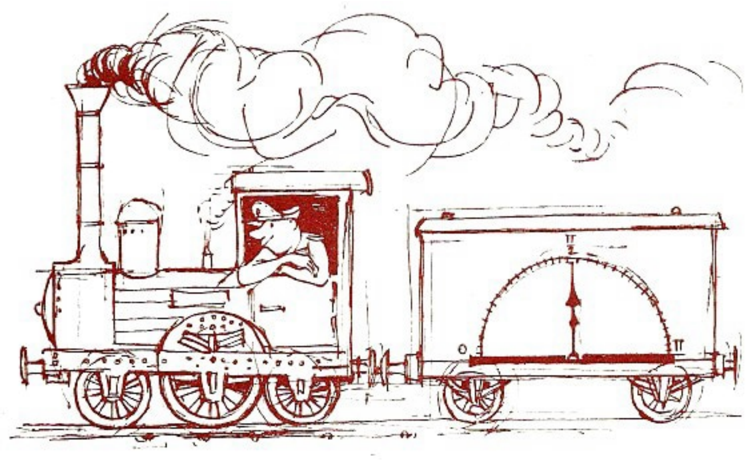
Nun möge $f(x)$ in einem Intervall stetig sein und in den beiden Endpunkten die voneinander verschiedenen Werte y und y' annehmen. y'' sei eine Zahl, die zwischen y und y' liegt.
Die Funktion $F(x) = f(x) - y''$ ist dann ebenfalls stetig und nimmt in den beiden Intervallendpunkten Werte verschiedenen Vorzeichens an. Daher muss $F(x)$ in einem inneren Punkt des Intervalls verschwinden: $F(\xi) = f(\xi) - y'' = 0$.





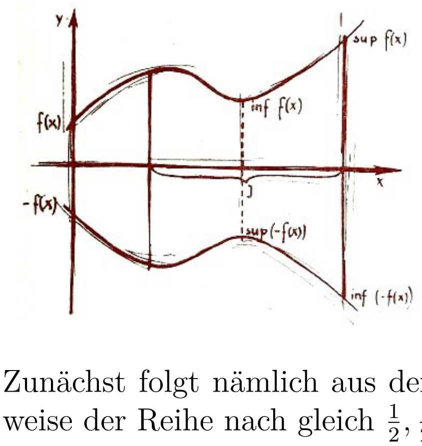
Die stetige Funktion $f(x)$ nimmt also jeden Zwischenwert y'' an.
Dieser Zwischenwertsatz gestattet eine überraschende Anwendung in der Mechanik. In einem geschlossenen Wagen sei zuunterst ein Zeiger um eine horizontale Achse senkrecht zur Fahrtrichtung reibungslos beweglich. Bei willkürlichem Fahrplan wirken auf den Zeiger beim Anfahren und Bremsen Kräfte ein, weiter die ganze Zeit hindurch die Schwere.

Gibt es immer eine Anfangsstellung des Zeigers, so dass dieser bei der Ankunft senkrecht steht?



Die Endlage des Zeigers wird durch den Fahrplan als stetige Funktion der Anfangslage bestimmt. Denn der Zeiger bewegt sich während der Fahrt kontinuierlich so, dass für hinreichend kurze Zeit die Änderung immer beliebig klein ausfällt.
Dieser stetige Anschluss verbietet aber einen ersten Augenblick, wo die Endlage sich sprunghaft änderte. Den Anfangslagen 0 bzw. π entsprechen die Endlagen 0 bzw. π. Nach dem Zwischenwertsatz gibt es daher eine Anfangslage, der die Endlage $\frac{\pi}{2}$ entspricht. Damit ist die Frage bejaht!

Der Wertevorrat einer stetigen Funktion bildet eine beschränkte Menge. Sonst müsste der Wertevorrat in mindestens einer der beiden Intervallhälften ebenfalls unbeschränkt sein, dann nach erneuter Halbierung dieser Intervallhälfte in mindestens einem der nun vorliegenden Intervallviertel, was bei Wiederholung zu einer Intervallschachtelung führt, die auf einen Punkt x_0 zusammenschrumpft. Infolge der Stetigkeit der Funktion in diesem Punkt müssten aber die Funktionswerte beliebig wenig vom Funktionswert in x_0 abweichen, im Widerspruch zur Konstruktion, die zu x_0 führte.



Der Wertevorrat der in J stetigen Funktion $f(x)$ besitzt also ein endliches Supremum. Entsprechendes gilt für das Infimum. Das folgt aus der Bemerkung, dass das Supremum von $-f(x)$ das mit einem Minuszeichen versehene Infimum von $f(x)$ liefert, so dass beide zugleich endlich ausfallen.
Die stetige Funktion nimmt ihr Supremum s an einer Stelle wirklich an, so dass man vom Maximum reden darf, genauer vom absoluten Maximum, das sich auf das ganze Intervall bezieht.

Zunächst folgt nämlich aus der Erklärung des Supremums, dass, wenn man ε beispielsweise der Reihe nach gleich $\frac{1}{2}, \frac{1}{2^2}, \ldots$ wählt, es Punkte $x_1, x_2, \ldots$ gibt, so dass

$$|f(x_n) - s| < \frac{1}{2^n}$$

gilt. ξ sei ein Häufungspunkt von $\{x_1, x_2, \ldots\}$. Dann ist $f(\xi) = s$. Dazu betrachte man die gegen ξ konvergierende Teilfolge $x_{k_1}, x_{k_2}, \ldots$ Da es sich um eine Teilfolge handelt, ist

$k_m \geq m$. Wegen der Stetigkeit in ξ gilt

$$|f(x_{k_m}) - f(\xi)| < \varepsilon \qquad \text{für} \qquad m > N$$

Aus

$$|f(x_{k_m}) - s| < \frac{1}{2^m} \qquad \text{folgt damit}$$

$$|f(\xi) - s| = |[f(\xi) - f(x_{k_m})] + [f(x_{k_m}) - s]| \leq |f(\xi) - f(x_{k_m})| + |f(x_{k_m}) - s| < \varepsilon + \frac{1}{2^m}$$

Da links eine feste Zahl steht, rechts aber eine beliebig kleine Zahl, muss die linke Seite verschwinden.

J sei ein abgeschlossenes Intervall, das heißt, die beiden Endpunkte gehören mit zu J. Dann ist eine in J stetige Funktion $f(x)$ in J gleichmäßig stetig. Darunter versteht man folgendes.
Wie klein auch ein positives ε vorgegeben wird, immer gibt es ein ebenfalls positives δ dazu, so dass

$$|f(x_1) - f(x_2)| < \varepsilon$$

ist für die beiden Punkte x_1 und x_2, wenn diese nur um weniger als δ voneinander entfernt sind, wo sie sonst auch in J liegen mögen. Den Beweis führt man indirekt:
Wäre dem nicht so, gäbe es ein festes ε, für das

$$|f(x'_n) - f(x''_n)| \geq \varepsilon$$

wäre für Punktepaare mit beliebig kleinem Abstand $|x'_n - x''_n| < \frac{1}{2^n}$.

Da die x'_n alle in J liegen, haben sie einen Häufungspunkt ξ, gegen den eine Teilfolge der x'_n konvergiert. Die entsprechende Teilfolge der x''_n konvergiert dann ebenfalls gegen ξ. Die vorletzte Ungleichung widerspricht aber der Stetigkeit in ξ, die vorausgesetzt war.

Bei der gleichmäßigen Stetigkeit handelt es sich also um eine globale Aussage, während die vorher besprochene Stetigkeit eine lokale Aussage ist.
1.5ex] Immer wieder wandten wir hier Intervallschachtelungen an.
Dieses Beweiselement gestattet eine Vereinheitlichung unserer Überlegungen durch Aufstellen eines abstrakten Schemas, wie ich in der Zeitschrift "Collectanea mathematica", die in Barcelona erscheint, 1960 ausgeführt habe. Unsere einzelnen Sätze lassen sich dann jeweils auf nur wenigen Zeilen begründen, dank der höheren Abstraktionsstufe, von der aus sie gewonnen werden.

Zum Schluss wenden wir uns konvergenten Reihen mit veränderlichen Gliedern zu. Die Funktionen $f_1(x)$, $f_2(x)$, seien im selben Intervall J definiert, und die Reihe

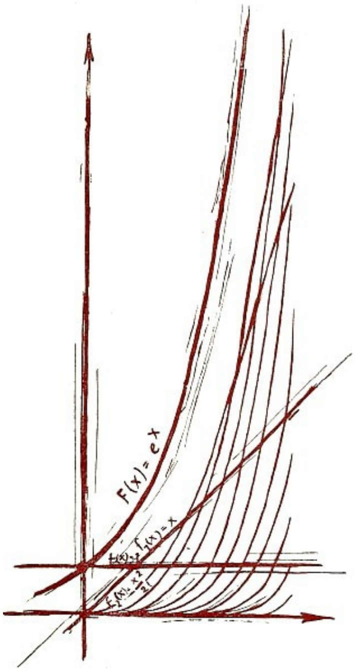
$$f_1(x) + f_2(x) + ...$$

möge in jedem Punkt von J konvergieren. Die Reihensumme ist dann eine Funktion $F(x)$ in J, die Grenzfunktion heißt. Sind nun die Funktionen $f_1(x)$, $f_2(x)$, ... alle stetig, dann braucht das für die Grenzfunktion selbst keineswegs zuzutreffen.
Eine hinreichende Bedingung für ihre Stetigkeit ist aber die gleichmäßige Konvergenz der Reihe

$$|s_n(x) - F(x)| < \varepsilon \qquad \text{für} \qquad n > N$$

und beliebiges x aus J, wenn man unter $s_n(x)$ die Partialsumme $f_1(x)+...+f_n(x)$ versteht. Wir gehen von der Ungleichung

$$\begin{aligned}|F(x) - F(x_0)| &= |[F(x) - s_n(x)] + [s_n(x) - s_n(x_0)] + [s_n(x_0) - F(x)]| \\ &\leq |F(x) - s_n(x)| + |s_n(x) - s_n(x_0)| + |s_n(x_0) - F(x)|\end{aligned}$$



aus.

Der erste und der dritte Summand sind für hinreichend großes n und beliebiges x bzw. x_0 aus J infolge gleichmäßiger Konvergenz beliebig klein, diesmal $< \frac{\varepsilon}{3}$. Weiter ist $s_n(x)$ als Summe von endlich vielen stetigen Funktionen selbst stetig.
Das besagt, dass für $|x - x_0| < \delta$ der mittlere Summand beliebig klein, $< \frac{\varepsilon}{3}$, ausfällt. Daher ist die rechte Seite der Ungleichung $< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$, also erst recht

$$|F(x) - F(x_0)| < \varepsilon \qquad \text{für} \qquad |x - x_0| < \delta$$

so dass die Grenzfunktion $F(x)$ in jedem Punkt x_0 von J stetig ist.

Als Beispiel sei hier die Reihe

$$F(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + ...$$

angeführt. Jede der Funktionen $f_n(x) = \frac{x^n}{n!}$ ist stetig. F(x) ist in jedem endlichen Intervall gleichmäßig konvergent und damit auch stetig.

19 Monotone Überraschungen

Die Funktion $f(x)$ heißt in einem Intervall streng monoton wachsend, wenn für zwei beliebige Punkte $x_1 < x_2$ aus dem Intervall $f(x_1) < f(x_2)$ ausfällt.

Wenn an Stelle des Ungleichheitszeichens nur $\leq$ gilt, fällt das Beiwort streng weg. Monotone Abnahme wird durch $f(x_1) > f(x_2)$ bzw. $f(x_1) \geq f(x_2)$ für $x_1 < x_2$ erklärt.

Wir hatten als abgeschlossenes Intervall $[a, b]$ ein Intervall bezeichnet, dem die beiden Endpunkte mit angehören. Nun wollen wir halboffene und offene Intervalle definieren. Die halboffenen Intervalle $[a, b)$ bzw. $(a, b]$ entstehen aus $[a, b]$ durch Entfernen des rechten bzw. linken Endpunktes, während das Entfernen beider Endpunkte zum offenen Intervall (a, b) führt.
Mit diesen Bezeichnungen gilt folgender Satz für monoton zunehmende Funktionen $f(x)$. Ist x ein beliebiger Punkt von $[a, b]$, dann ist bei Annäherung an x von links das Supremum s von $f(x)$ in $[a, x)$ linksseitiger Grenzwert, der mit $f(x-0)$ bezeichnet wird.
Das folgt sofort aus der Erklärung des Supremums. Es muss ja für beliebiges positives ε ein x_0 in $[a, x)$, also links von x geben, so dass

$$0 \leq |s - f(x_0)| < \varepsilon$$

ausfällt. Wegen der monotonen Zunahme der Funktion gilt für alle ξ aus $[x_0,x)$

$$f(x_0) \leq f(\xi)$$

und mit Hinblick auf $f(\xi) \leq s$ daher schließlich

$$0 \leq |s - f(\xi)| < \varepsilon$$

für alle ξ aus $[x_0,x)$.
Infolge monotoner Zunahme ist $f(x)$ obere Schranke für $\{f(\xi)\}$. Da $s = f(x-0)$ die kleinste obere Schranke für diese Werte darstellt, folgt $f(x-0) \leq f(x)$.

Ähnlich wie vorhin findet man, dass das Infimum aller Funktionswerte aus dem Intervall $(x, b]$ rechtsseitiger Grenzwert ist bei Annäherung an x von rechts her. Daher ist für dieses Infimum die Bezeichnung $f(x+0)$ berechtigt, und es gilt weiter $f(x) \leq f(x-0)$.

Eine Folge der beiden letzten Ungleichungen bildet der Satz, dass eine monoton zunehmende Funktion $f(x)$ an der Stelle x genau dann stetig ausfällt, wenn

$$f(x-0) = f(x+0)$$

gilt. Wenn diese Gleichung nicht besteht, spricht man von einem Sprung $f(x+0) - f(x-0)$ an der Stelle x.

In unserem Beispiel $f(x) = \dfrac{1}{1+e^{\frac{1}{x-1}}}$ tritt an der Stelle $x = 1$ ein solcher Sprung auf, denn der linksseitige Grenzwert ist an dieser Stelle verschieden vom rechtsseitigen Grenzwert.

Liegen x' bzw. x'' links bzw. rechts von x, $x' < x < x''$, dann ist

$$f(x') \leq f(x-0) \leq f(x+0) \leq f(x'') \qquad \text{woraus} \qquad f(x+0) - f(x-0) \leq f(x'') - f(x')$$

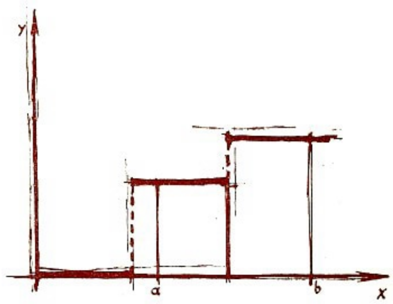
folgt.
Damit lässt sich der Nachweis führen, dass die monoton zunehmende Funktion $f(x)$ höchstens abzählbar viele Sprungstellen in $[a,b]$ besitzt.

Bezeichnen wir $f(b) - f(a)$ mit d und setzen in der letzten Ungleichung $x' = a$ und $x'' = b$, so erhalten wir

$$f(x+0) - f(x-0) \leq d$$

das heißt, dass in $[a,b]$ offensichtlich keine Sprungstelle mit einem größeren Sprung als d existiert und aus demselben Grunde höchstens eine mit dem Sprung d.

In letzterem Falle handelt es sich um eine Treppenfunktion, die von a bis zur Sprungstelle x den konstanten Wert $f(a)$, von x bis zu b den konstanten Wert $f(b)$ besitzt. Liegt nicht dieser Fall vor, dann seien $x_1 < ... < x_n$ Sprungstellen mit den Sprüngen $d_1, ..., d_n$. Dann gibt es immer aufeinanderfolgende Stellen $x'_1 < ... < x'_{n+1}$, so dass für $k = 1, ..., n-1$



$$x'_k < x_k < x'_{k+1} < x_{k+1}$$

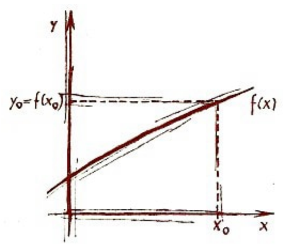
gilt. Dann ist $d_k \leq f(x_{k+1}) - f(x_k)$, daher

$$\begin{aligned} d_1 + ... + d_n &\leq [f(x_2) - f(x_1)] + [f(x_3) - f(x_2)] + ... + [f(x_{n+1} - f(x_n)] \\ &= f(x_{n+1}) - f(x_1) \leq f(b) - f(a) = d \end{aligned}$$

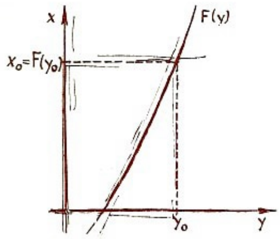
Daraus folgt, dass es höchstens n Sprünge geben kann, die nicht kleiner als $\frac{d}{n}$ ausfallen. Man betrachte nun der Reihe nach die Sprünge, deren Werte $d, \frac{d}{2}, \frac{d}{3}, ...$ übersteigen. Schreibt man ihre Sprungstellen für jedes n in beliebiger Reihenfolge hin, so entsteht eine Folge, die alle Sprungstellen enthält. Freilich kann nicht behauptet werden, dass die in der Folge aufeinanderfolgenden Sprungstellen auch auf der Zahlengeraden aufeinander folgen.
Die entsprechenden Sätze und Beweise gelten für monoton abnehmende Funktionen ebenfalls.

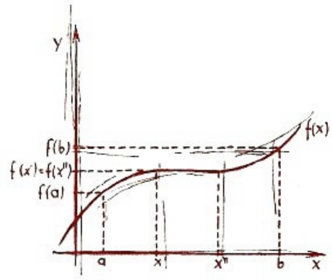
Für streng monotone Funktionen lässt sich die Umkehrfunktion recht einfach finden. Denn zu verschiedenen Werten des Arguments gehören dann verschiedene Funktionswerte, in der Abbildung also, die der Funktion entspricht, zu jedem Bild nur ein Urbild:
Die Abbildung ist umkehrbar eindeutig. Ordinaten aber, deren Werte innerhalb von Sprungstellen der ursprünglichen Funktion liegen, sind durch die Umkehrung noch keine Abszissenwerte zugeordnet. Ebenso ist bei monotonen Funktionen, die nicht streng monoton sind, die Umkehrfunktion noch nicht bestimmt.
Dem kann durch eine entsprechende Definition abgeholfen werden, die durch die Anschauung nahegelegt wird, deren Tragweite aber übersehen wurde, wie ich das im "Archiv der Mathematik" 1958 ausführte.

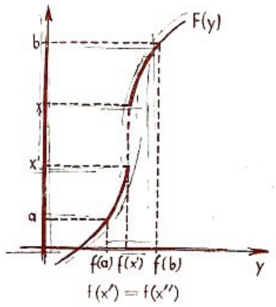
Eine monotone, jedoch nicht streng monotone Funktion besitzt wenigstens ein Intervall, in dem ihr Wert konstant ist. Dann muss erst erklärt werden, welches der unendlich vielen Urbilder des Konstanzintervalls als Wert der Umkehrfunktion gelten soll.



Um das auszuführen, betrachten wir eine monoton zunehmende Funktion $f(x)$, die im Intervall $[a, b]$ definiert sei.
Für $a \leq x' < x'' \leq b$ gilt also immer $f(x') \leq f(x'')$. Die Umkehrfunktion $F(y)$, die im Intervall $[f(a), f(b)]$ definiert ist, führen wir ein durch die Forderung, dass $F(y) =$ Supremum aller x, für die $f(x) \leq y$ ausfällt.



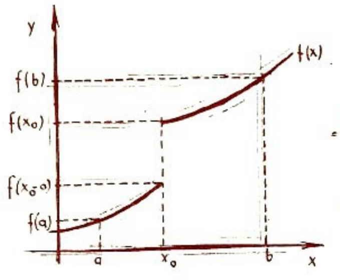


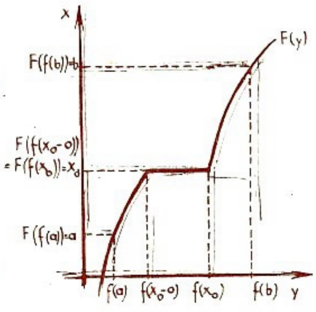


Aus unserer Erklärung geht unmittelbar hervor, dass die Umkehrfunktion selbst eine monoton zunehmende Funktion im Intervall $[f(a), f(b)]$ darstellt.

Für Konstanzintervalle besagt unsere Forderung folgendes. Ist $f(x)$ in (x', x'') konstant gleich c, für $x > x''$ aber $f(x) > f(x'')$, dann gilt $F(c) = x''$. Ist für $x < x'$ weiter $f(x) < f(x')$, dann fällt für $y < c \quad F(y) < x'$ aus.
Ein Konstanzintervall von $f(x)$ ist folglich nach unserer Definition in c Sprungstelle der Umkehrfunktion mit einem Sprung, der die Länge $x'' - x'$ des Konstanzintervalls besitzt.

Nun sei x_0 Sprungstelle von $f(x)$, speziell $f(x_0-0) < f(x_0)$. Dann besitzt für $y = f(x_0-0)$ die Umkehrfunktion den Wert x_0, für $y = f(x_0)$ ebenfalls den Wert x_0. Da aber die Umkehrfunktion monoton ist, folgt daraus, dass $[[f(x_0 - 0), f(x_0)]$ Konstanzintervall von $F(y)$ ist.
Für $f(x_0) < f(x_0+0)$ folgt ähnlich, dass $[f(x_0), f(x_0+0)]$ Konstanzintervall der Umkehrfunktion ist.



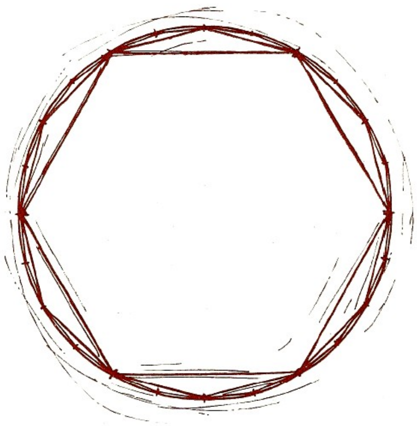


Die beiden letzten Einsichten lassen schon vermuten, dass die Umkehrfunktion einer monotonen Funktion $f(x)$ genau dann stetig ausfällt, wenn $f(x)$ streng monoton ist, also keine Konstanzintervalle besitzt. Dann ist ja, wie wir schon wissen, $F[f(x)] = x$. Daher nimmt $F(y)$ alle Werte aus $[a, b]$ an.

Da die Umkehrfunktion selbst monoton zunimmt, $F(y-0) \leq F(y) \leq F(y+0)$ kann keines der beiden Ungleichheitszeichen gelten. Sonst müsste die monoton zunehmende Umkehrfunktion $F(y)$ die Werte aus $(F(y - 0), F(y))$ bzw. $(F(y), F(y * 0))$ auslassen, die jedoch dem Intervall $[a, b]$ angehören. Dabei mutet es paradox an, dass eine stetige monotone Funktion eine unstetige Umkehrfunktion haben kann, wenn sie nicht streng monoton ist, während eine nicht stetige, aber streng monotone Funktion eine stetige Umkehrfunktion besitzt.

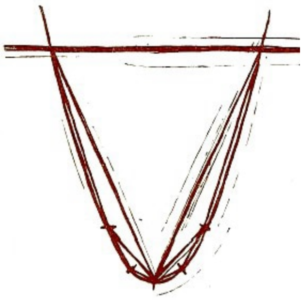
Für die Umkehrfunktionen von monoton abnehmenden Funktionen gelten entsprechende Sätze und Beweise.

20 Inhalt und Integral



Die Anfänge der Integralrechnung reichen ins Altertum zurück.
Die Exhaustionsmethode lässt sich als Integration deuten. Sie diente zur Inhaltsbestimmung von Figuren, deren Rand kein Polygon ist. Obwohl diese Methode ihren Namen erst im 17. Jahrhundert erhielt, wurde sie bereits im Altertum, zum Beispiel von Archimedes, angewandt.

Um etwa den Inhalt von Kreis und Parabelsegment zu bestimmen, verfuhr er wie folgt.



Er ging in beiden Fällen von einem Vieleck aus, dessen Eckpunkte auf dem Rand lagen. Bei dem Kreis war es ein eingeschriebenes reguläres Sechseck, bei dem, Parabelsegment ein Dreieck.
Dann baute er an jeder Seite ein Dreieck an, dessen dritte Ecke auf dem Rand lag. Setzte er das immer weiter fort, gewann er Polygone, die Über ihre Vorgänger hinauswuchsen und sich dem Rand immer mehr anschmiegten.

Das Supremum der Polygoninhalte galt als Inhalt der Ausgangsfigur. Den Nachweis für die Berechtigung dieses Verfahrens zu führen, blieb allerdings dem 19.Jahrhundert vorbehalten.

Erst in der zweiten Hälfte des vorigen Jahrhunderts wurden einwandfreie Erklärungen abgegeben für Bogenlänge und Inhalt. Es stellte sich später heraus, dass der Inhaltsbegriff Abstufungen kennt, die für die zeitgenössische Analysis unumgänglich erforderlich sind. Wir meinen damit das Integral von Lebesgue. Doch beschränken wir uns hier auf die klassischen Integrale, die bereits von größter Fruchtbarkeit sind. Man gewinnt sie als verschiedene Anwendungen eines Lemmas, das ich in meiner Monographie über "Die Grundlagen der Quantenmechanik und ihre Mathematik" aus dem Jahre 1936 aufstellte.

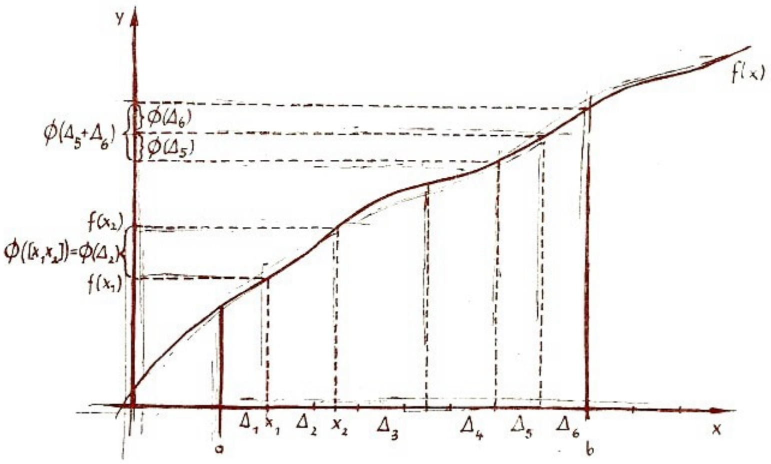
Wie ich in meinem Buch "Grundzüge der Mechanik" (2.Aufl. 1961) ausführte, erfasst das Lemma auch das Integral von Lebesgue, doch bedarf es dann des Maßbegriffes, der uns nicht zur Verfügung steht und auf den wir hier nicht eingehen können.[a]
Um den Begriff des klassischen Integrals einzuführen, müssen wir erst einige Vorbemerkungen einfügen.

[a]Eine lehrbuchmäßige Darstellung findet der Leser bei Natanson in "Theorie der Funktionen einer reellen Veränderlichen" (2. Auflage 1961).

$\{\Delta\}$ sei die Gesamtheit aller Teilintervalle von $[a, b]$. Φ bezeichne eine Abbildung dieser Gesamtheit auf eine Teilmenge der reellen Zahlen. Nach Dirichlet ist Φ als Funktion, genauer als Intervallfunktion anzusprechen.
Zunächst geht die Klasse der additiven Intervallfunktionen nicht über die bisher betrachteten Funktionen hinaus. Man nennt Φ eine additive Intervallfunktion, wenn für zwei angrenzende Teilintervalle von $[a, b]$, Δ_1 und Δ_2, immer die Gleichung

$$\Phi(\Delta_1 + \Delta_2) = \Phi(\Delta_1) + \Phi(\Delta_2)$$

besteht, mit $\Delta_1 + \Delta_2$ das Teilintervall von $[a,b]$ bezeichnet, das durch die Vereinigung von Δ_1 und Δ_2 entsteht.



Eine solche additive Intervallfunktion gewinnt man mit Hilfe einer beliebigen Funktion $f(x)$, die im selben Intervall $[a, b]$ erklärt ist, wenn

$$f(x_2) - f(x_1) = \Phi([x_1, x_2])$$

gesetzt wird. Zum Unterschied zur Intervallfunktion nennt man $f(x)$ in diesem Zusammenhang Punktfunktion. Ist f monoton, dann besitzt Φ lauter Werte vom selben Vorzeichen. Die Umkehrung davon gilt ebenfalls. Ist Φ eine additive Intervallfunktion, so gibt es eine bis auf eine additive Konstante eindeutig bestimmte erzeugende Punktfunktion $f(x)$, so dass zwischen Φ und f die Gleichung von vorhin besteht. Man definiere die Punktfunktion $f(x)$ durch

$$f(x) = \Phi([a,x])$$

Dann folgt aus der Additivität von Φ für $\xi > x$

$$f(\xi) = \Phi([a,\xi]) = \Phi([a,x]) + \Phi([x,\xi])$$

woraus sich

$$\Phi([x,\xi]) = \Phi([a,\xi]) - \Phi([a,x]) = f(\xi) - f(x)$$

ergibt. Ist dann $F(x)$ eine zweite Punktfunktion, für die gleichfalls

$$\Phi([x,\xi]) = F(\xi) - F(x)$$

gelten soll, dann folgt, $x = a$ gesetzt,

$$\Phi([a,\xi]) = F(\xi) - F(a) = f(\xi) - f(a)$$

und damit

$$F(\xi) = f(\xi) + [F(a) - f(a)]$$

Die beiden Punktfunktionen F und f unterscheiden sich also durch die additive Konstante $[F(a) - f(a)]$ voneinander.
Die Intervallfunktion Φ heißt stetig, wenn $|\Phi(\Delta)| < \varepsilon$ gilt für alle Teilintervalle von $[a,b]$ von hinreichend kleiner Länge, also für $\Delta = [x_1,x_2]$ mit $x_2 - x_1 < \delta$. Gleichwertig damit ist für additive Intervallfunktionen die Forderung, dass die erzeugende Punktfunktion stetig sei.

Nunmehr sind wir in der Lage, unser Lemma zu formulieren.
Dafür spielt eine andere Kategorie von Intervallfunktionen, nicht die additiven, eine Rolle. Die stetige Intervallfunktion Φ möge für angrenzende Intervalle Δ_1 und Δ_2 stets die Ungleichung

$$\Phi(\Delta_1 + \Delta_2) \geq \Phi(\Delta_1) + \Phi(\Delta_2)$$

erfüllen.
$\mathfrak{Z}$ bezeichne eine Zerlegung von $[a, b]$ in endlich viele Teilintervalle $\Delta_1, ..., \Delta_n$. Unter diesen befindet sich mindestens eines von maximaler Länge λ. Diese maximale Länge kann infolge der Stetigkeit von Φ so gewählt werden, dass $|\Phi(\Delta_k)| < \varepsilon$ ist für $\lambda < \delta(\varepsilon)$ und $k = 1, ..., n$.

Jede Zerlegung $\mathfrak{Z}$ gibt Anlass, eine dazugehörige Summe

$$\Phi(\Delta_1) + ... + \Phi(\Delta_n)$$

zu bilden, die durch $S_{\mathfrak{Z}}$ bezeichnet werde. Wenn das Infimum J von $\{S_{\mathfrak{Z}}\}$, der Menge der Summen, die zu allen möglichen Zerlegungen von $[a, b]$ gehören, endlich ausfällt, dann behauptet das Lemma

$$0 \leq S_{\mathfrak{Z}} - J < \varepsilon \qquad \text{für} \qquad \lambda < \delta^*$$

anders ausgedrückt $S_{\mathfrak{Z}} \to J$ für $\lambda \to 0$.

Um den Nachweis zu führen, erinnern wir an die Definition des Infimums. Danach muss es ein $S_{\mathfrak{Z}^*}$, geben, so dass

$$0 \leq S_{\mathfrak{Z}^*} - J < \frac{\varepsilon}{2}$$

ausfällt. Die Zerlegung $\mathfrak{Z}^*$ bestehe aus m Teilintervallen von $[a, b]$.
Entsteht nun $\mathfrak{Z}''$ aus $\mathfrak{Z}'$ dadurch, dass Teilintervalle von $\mathfrak{Z}'$ weiter zerlegt werden, dann folgt aus der für Φ postulierten Ungleichung

$$S_{\mathfrak{Z}''} \leq S_{\mathfrak{Z}'}$$

$S_{\mathfrak{Z}}$ sei nun eine Zerlegung mit $\lambda < \delta\left(\frac{\varepsilon}{6m}\right)$. Wir vereinen die Teilpunkte von $\mathfrak{Z}$ und $\mathfrak{Z}^*$. Es entsteht eine Zerlegung $\mathfrak{Z}^{**}$, für die $S_{\mathfrak{Z}^{**}} \leq S_{\mathfrak{Z}^*}$ ist, woraus $S_{\mathfrak{Z}^{**}} - J \leq S_{\mathfrak{Z}^*} - J$ folgt, so dass

$$0 \leq S_{\mathfrak{Z}^{**}} - J < \frac{\varepsilon}{2}$$

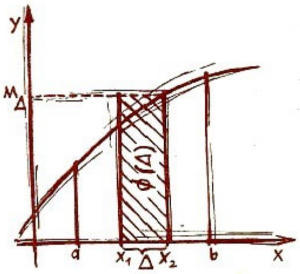
gilt.
Lässt man die $m - 1$ Teilpunkte der Zerlegung $\mathfrak{Z}^*$ in $\mathfrak{Z}^{**}$ weg, dann gewinnt man selbstverständlich die Zerlegung $\mathfrak{Z}$ zurück.
In $\mathfrak{Z}$ sind Teilintervalle aus $\mathfrak{Z}^{**}$, die einen der $m - 1$ Teilpunkte enthalten, vereint. Die übrigen Teilintervalle von $\mathfrak{Z}^{**}$ und $\mathfrak{Z}$ sind dieselben. Bis auf höchstens $3(m - 1)$ Teilintervalle fallen die Zerlegungen $\mathfrak{Z}^{**}$ und $\mathfrak{Z}$ daher zusammen.

In $S_3 - S_{3^{**}}$ heben sich folglich nur die zu diesen höchstens $3(m-1)$ Teilintervallen gehörenden Summanden nicht weg. Das führt zur Ungleichung

$$0 \leq S_3 - S_{3^{**}} < 3(m-1)\frac{\varepsilon}{6m} < \frac{\varepsilon}{2}$$

Mit $0 \leq S_{3^{**}} - J < \frac{\varepsilon}{2}$ zusammen folgt damit

$$0 \leq S_3 - J = (S_3 - S_{3^{**}} + (S_{3^{**}} - J) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$



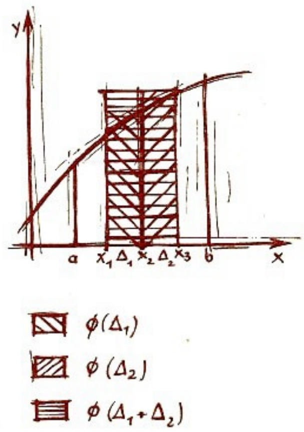
Für das δ^* in unserer Aussage können wir danach

$$\delta^* = \delta\left(\frac{\varepsilon}{6m}\right)$$

setzen. Wird an Stelle der ursprünglichen die Ungleichung

$$\Phi(\Delta_1 + \Delta_2) \leq \Phi(\Delta_1) + \Phi(\Delta_2)$$

gefordert, dann konvergieren die Summen S_3 mit $\lambda \to 0$ gegen das Supremum von $\{S_3\}$, falls dieses endlich ausfällt.



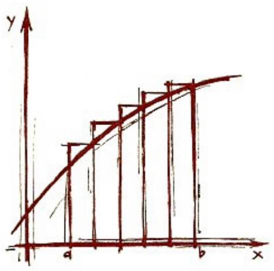
Die bisherigen Ergebnisse wollen wir nun verwenden, um den Begriff des Integrals zu erläutern.
Wir betrachten eine in $[a, b]$ beschränkte Funktion $f(x)$. Mit M_Δ bezeichnen wir das Supremum von $f(x)$ in $\Delta = [x_1, x_2]$. Die Intervallfunktion $\Phi(\Delta)$ wird diesmal durch

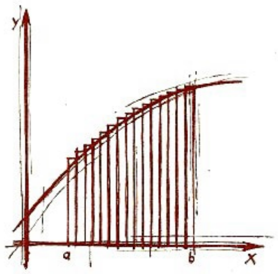
$$\Phi(\Delta) = (x_2 - x_1)M_\Delta$$

definiert.

Wegen $M_\Delta \leq M_{[a,b]}$ ist Φ stetig und genügt für $\Delta_1 = [x_1, x_2]$, $\Delta_2 = [x_2, x_3]$ der Ungleichung

$$\begin{aligned}\Phi(\Delta_1 + \Delta_2) &= (x_3 - x_1)M_{\Delta_1+\Delta_2} = (x_3 - x_2)M_{\Delta_1+\Delta_2} - (x_2 - x_1)M_{\Delta_1+\Delta_2} \\ &\geq (x_3 - x_2)M_{\Delta_2} + (x_2 - x_1)M_{\Delta_1} = \Phi(\Delta_2) + \Phi(\Delta_1)\end{aligned}$$

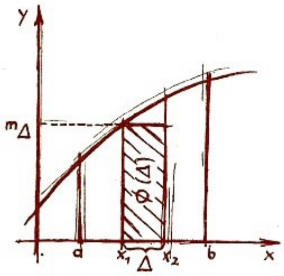




Die Intervalle von $\mathfrak{Z}$ seien $\Delta_1 = [a, x_1]$, $\Delta_2 = [x_1, x_2]$, ..., $\Delta_n = [x_{m-1}, b]$. Dann ist S_3 diesmal der Ausdruck

$$(x_1 - a)M_{\Delta_1} + (x_2 - x_1)M_{\Delta_2} + ... + (b - x_{n-1})M_{\Delta_n}$$

der $\geq (b-a)m$ ausfällt, wobei mit m das Infimum von f in $[a, b]$ bezeichnet wird. Das endliche Infimum J, gegen den diese Ausdrücke im Sinne unseres Lemmas konvergieren, heißt das obere Integral von Darboux.

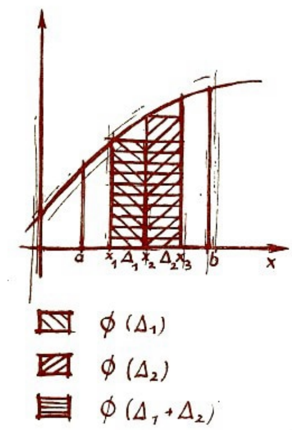


Verwendet man an Stelle des Supremums M_Δ von $f(x)$ in Δ das Infimum m_Δ, und setzt

$$\Phi([x_1,x_2]) = (x_2 - x_1)m_{[x_1,x_2]}$$

dann folgt sofort die Stetigkeit auch dieser neuen Intervallfunktion, die jetzt der Ungleichung

$$\Phi(\Delta_1 + \Delta_2) \leq \Phi(\Delta_1) + \Phi(\Delta_2)$$



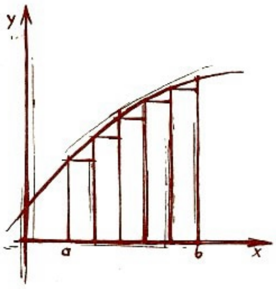
genügt. Im Sinne unseres Lemmas gilt also die Konvergenz der Summen

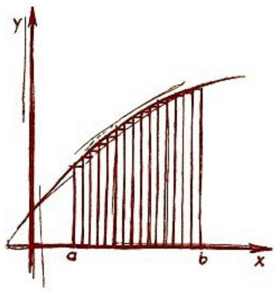
$$(x_1 - a)m_{\Delta_1} + ... + (b - x_{n-1})m_{\Delta_n}$$

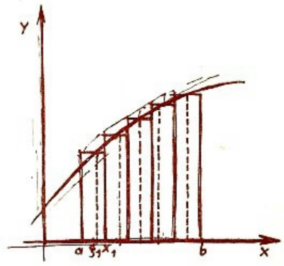
gegen das Supremum aller dieser Summen, das unteres Integral von Darboux heißt. Aus $m_\Delta \leq M_\Delta$ folgt

$$(x_1-a)m_{\Delta_1}+...+(b-x_{n-1})m_{\Delta_n} \leq (x_1-a)M_{\Delta_1}+...+(b-x_{n-1})M_{\Delta_n}$$

Daher ist das untere Integral von Darboux nicht größer als das obere Integral.
Im Falle der Gleichheit heißt die Funktion $f(x)$ integrabel, genauer nach Riemann integrierbar.







Riemann selber betrachtete allerdings die Summen

$$(x_1 - a)f(\xi_1) + ... + (b - x_{n-1})f(\xi_n) \qquad \text{wo} \qquad x_k \leq \xi_k \leq x_{k+1}$$

ist. Aus

$$m_{[x_k,x_{k+1}]} \leq f(\xi) \leq M_{[x_k,x_{k+1}]}$$

folgt dann sofort, dass auch diese Summen gegen den gemeinsamen Wert der beiden Darbouxschen Integrale konvergieren. Dieser gemeinsame Wert wird mit

$$\int\limits_a^b f(x)dx$$

bezeichnet, eine Bezeichnung, die aus vorhergehenden Zeiten übernommen wurde.

Hinreichend für die Integrierbarkeit im Sinne von Riemann, aber keineswegs notwendig, ist die Stetigkeit des Integranden $f(x)$. Da f dann in $[a,b]$ gleichmäßig stetig ausfällt, kann

λ so klein gewählt werden, dass $0 \leq M_{\Delta_k} - m_{\Delta_k}$ für alle Intervalle Δ_1, ..., Δ_k zugleich beliebig klein, $< \frac{\varepsilon}{b-a}$, ausfällt. Dann ist

$$\begin{aligned}
&[M_{\Delta_1}(x_1 - a) + ... + M_{\Delta_n}(b - x_{n-1})] - [m_{\Delta_1}(x_1 - a) + ... + m_{\Delta_n}(b - x_{n-1})] \\
&\qquad = (M_{\Delta_1} - m_{\Delta_1})(x_1 - a) + ... + (M_{\Delta_n} - m_{\Delta_n})(b - x_{n-1}) \\
&\qquad < \frac{\varepsilon}{b-a}[(x_1 - a) + ... + (b - x_{n-1})] = \varepsilon
\end{aligned}$$

Das obere Integral von Darboux ist als Infimum nicht größer als die Summe in der ersten eckigen Klammer links, das untere Integral als Supremum nicht kleiner als die Summe in der zweiten eckigen Klammer links vom Gleichheitszeichen, die Differenz der beiden Darbouxschen Integrale folglich ebenfalls kleiner als ε.
Da diese Differenz eine feste Zahl ist, ε dagegen beliebig klein, muss die Differenz verschwinden, so dass die Werte der beiden Darbouxschen Integrale zusammenfallen. Aus

$$(b - a)m \leq (x_1 - a)m_{\Delta_1} + ... + (b - x_{n-1})m_{\Delta_n}$$

folgt, dass das Supremum aller möglichen Summen von rechts erst recht größer als die linke Seite ausfällt. Entsprechend folgt, wenn mit M das Supremum von $f(x)$ in $[a, b]$ bezeichnet wird, aus

$$(x_1 - a)M_{\Delta_1} + ... + (b - x_{n-1})M_{\Delta_n} \leq (b - a)M$$

dass das Infimum aller möglichen Summen von links erst recht kleiner als die rechte Seite ausfällt.

Für integrable Funktionen gilt daher

$$(b - a)m \leq \int_a^b f(x)dx \leq (b - a)M$$

so dass mit einem geeigneten Zwischenwert $m \leq \mu \leq M$ die Gleichung

$$\int_a^b f(x)dx = (b - a)\mu$$

besteht, die erster Mittelwertsatz der Integralrechnung heißt.
Ist der Integrand insbesondere eine stetige Funktion, dann ist diese gewiss in $[a, b]$ beschränkt und nimmt jeden Zwischenwert an, so dass dann

$$\int_a^b f(x)dx = (b - a)f(\xi) \qquad \text{mit} \qquad a \leq \xi \leq b \qquad \text{gilt.}$$

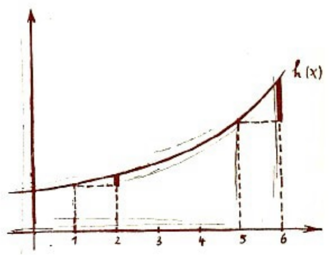
Die von Riemann betrachteten Summen stellen augenscheinlich den Gesamtinhalt der eingezeichneten Rechtecke in unserer Figur dar. Die Konvergenz führt dazu, den Inhalt der Fläche, die von der Kurve $f(x)$, der Strecke $[a, b]$ auf der Abszissenachse und von den Ordinaten $f(a)$ und $f(b)$ begrenzt ist, durch das bestimmte Integral

$$\int_a^b f(x)dx$$

zu definieren. Dieser Flächeninhalt ist schon recht allgemein, aber nicht allgemein genug.

Der Holländer Stieltjes nahm 1894 eine Weiterbildung des Integrals von Riemann vor. Man gelangt ganz natürlich dazu, wenn man zunächst bemerkt, dass die Länge von Intervallen eine additive und stetige Intervallfunktion ist. Mit Hilfe jeder anderen, ebenfalls stetigen und additiven Intervallfunktion $\phi(\Delta)$, die allein nichtnegativer Werte fähig ist, führe man ein neues Längenmaß ein, allerdings nur unter Verzicht auf die Maßgleichheit kongruenter Intervalle.

Wie wir schon wissen, wird ϕ durch eine monoton zunehmende Funktion $h(x)$ erzeugt. Für die in $[a, b]$ beschränkte Funktion $f(x)$ definiere man damit $\phi(\Delta)$ als $M_\Delta \phi(\Delta)$ bzw. $m_\Delta \phi(\Delta)$.
Die so eingeführten Intervallfunktionen Φ erweisen sich dann als stetig, und es gelten für sie die im Lemma geforderten Ungleichungen, so dass die Summen



$$M_{\Delta_1}\phi(\Delta_1) + ... + M_{\Delta_n}\phi(\Delta_n) \qquad \text{bzw.} \qquad m_{\Delta_1}\phi(\Delta_1) + ... + m_{\Delta_n}\phi(\Delta_n)$$

gegen das Infimum bzw. Supremum aller Summen der ersten bzw. zweiten Art konvergieren. Fallen aber die Werte vom Infimum und Supremum zusammen, dann redet man vom Stieltjes-Integral, das mit

$$\int_a^b f(x)dh(x)$$

bezeichnet wird. Das Riemann-Integral entspricht dem Spezialfall $h(x) = x$.

Als letzte Anwendung unseres Lemmas behandeln wir die Bogenlänge ebener Kurven. Die Kurve sei durch zwei Funktionen $f(x)$ und $g(x)$ beschrieben, die im Intervall $[a, b]$ stetig sind. Mit

$$a < x_1 < ... < x_{n-1} < b$$

entsteht ein der Kurve eingeschriebener Sehnenzug, dessen Ecken $(x(t_k), y(t_k))$ auf der Kurve liegen. Die Länge dieses Sehnenzuges beträgt

$$\sqrt{[f(x_1) - f(a)]^2 + [g(x_1) - g(a)]^2} + ... + \sqrt{[f(b) - f(x_{n-1})]^2 + [g(b) - g(x_{n-1})]^2}$$

Wir machen mit Jordan die Annahme, dass alle möglichen Summen, auch Schwankungssummen genannt,

$$|f(x_1) - f(a)| + ... + |f(b) - f(x_{n-1})| \quad \text{und} \quad |g(x_1) - g(a)| + ... + |g(b) - g(x_{n-1})|$$

unter einer festen endlichen Schranke K liegen. Aus

$$\sqrt{[f(x_{k+1}) - f(x_k)]^2 + [g(x_{k+1}) - g(x_k)]^2} \leq |f(x_{k+1}) - f(x_k)| + |g(x_{k+1}) - g(x_k)|$$

folgt dann, dass die Länge aller; möglichen eingeschriebenen Sehnenzüge kürzer als $2K$ ist. Die Ungleichung selber besagt, dass die Kathetensumme größer als die Hypotenuse ausfällt.

Das endliche Supremum aller eingeschriebenen Sehnenzüge wird als Bogenlänge definiert. Die beiden Funktionen $f(x)$ und $g(x)$ heißen Funktionen von beschränkter Schwankung.

Wir setzen für $a \leq x' < x'' \leq b$

$$\Phi([x',x'']) = \sqrt{[f(x'') - f(x')]^2 + [g(x'') - g(x')]^2}$$

Dann ist Φ stetig und genügt der Ungleichung $\Phi(\Delta_1 + \Delta_2) \leq \Phi(\Delta_1) + \Phi(\Delta_2)$.

Ausgeschrieben bildet diese die arithmetische Fassung des elementargeometrischen Satzes, dass im Dreieck keine Seite größer ausfällt als die Summe der beiden anderen Seiten. Im Sinne unseres Lemmas konvergieren daher die Längen der eingeschriebenen Sehnenzüge gegen die Bogenlänge, wie es die Anschauung fordert.

Zum Schluss werde noch die Struktur der Funktionen von beschränkter Schwankung untersucht. Aus der Erklärung folgt sofort, dass $f(x)$ im Teilintervall $[a,b]$ ebenfalls von beschränkter Schwankung ist. Das Supremum aller Schwankungssummen des Intervalls $[a,x]$ werde mit $s(x)$ bezeichnet. Dann ist $s(x)$ eine monoton zunehmende Funktion, denn für $x < \xi$ gilt

$$s(x) \leq s(\xi)$$

weil es zu jeder Schwankungssumme im Intervall $[a,x]$ eine andere in $[a,\xi]$ gibt, die alle Teilpunkte der ersten Summe enthält, folglich größer ausfällt.
Wir betrachten spezielle Schwankungssummen im Intervall $[a,\xi]$, die zwischen x und ξ keinen weiteren Teilpunkt enthalten. Deren Supremum ist offensichtlich

$$s(x) + |f(\xi) - f(x)|$$

Diese speziellen Summen bilden eine Teilmenge aller Schwankungssummen des Intervalls $[a,\xi]$, daher ist

$$s(x) + |f(\xi) - f(x)| \leq s(\xi)$$

Setzt man jetzt $s(x) - f(x) = t(x)$, dann ist $t(x)$ eine monoton zunehmende Funktion, denn

$$t(\xi) - t(x) = [s(\xi) - s(x)] - [f(\xi) - f(x)]$$

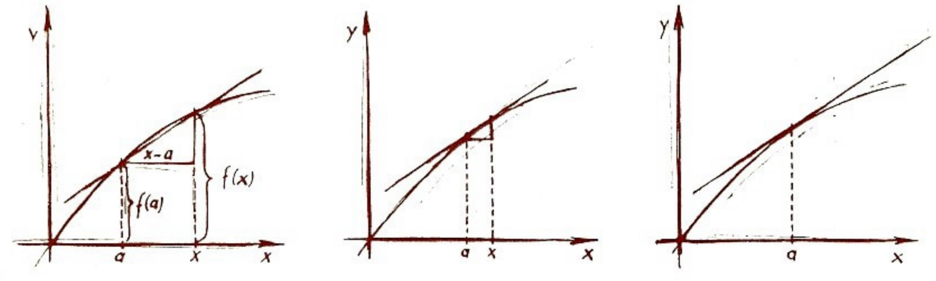
ist infolge der letzten Ungleichung nicht negativ. So gewinnt man für $f(x)$ die Darstellung

$$f(x) = s(x) - t(x)$$

Die Funktionen von beschränkter Schwankung können also als Differenz von zwei monoton zunehmenden Funktionen geschrieben werden. Die Umkehrung davon ist ebenfalls richtig, doch wollen wir den einfachen Beweis dafür übergehen.

21 Gegenspielerin Ableitung

In jeder Bewegung steckt eine Ableitung: die Geschwindigkeit. Da man im Altertum noch nicht differenzieren konnte, bewegte man sich ohne Erklärung. Aber man bewegte sich.



Die Ableitung $f'(a)$ von der Funktion $f(x)$ an der Stelle a wird als Grenzwert des Differenzenquotienten

$$\frac{f(x)-f(a)}{x-a}$$

für $x \to a$ definiert, falls dieser Grenzwert existiert.

Früher machte man sich von dieser Existenz ganz falsche Vorstellungen. Ampére, einer der 40 Unsterblichen der Französischen Akademie, meinte 1806 bewiesen zu haben, dass eine stetige Funktion immer differenzierbar ist. Doch 1875 wurde das erste auf Weierstraß zurückgehende Beispiel einer stetigen Funktion veröffentlicht, die nirgends eine Ableitung besitzt. Allerdings konstruierte bereits 40 Jahre vorher Bolzano ein solches Beispiel, das aber erst nach seinem Tode bekannt wurde.

Um eine stetige Funktion zu konstruieren, die nirgends differenzierbar ist, beschränken wir x zunächst auf das Intervall [0,1]. Die Dezimalbruchschreibweise von

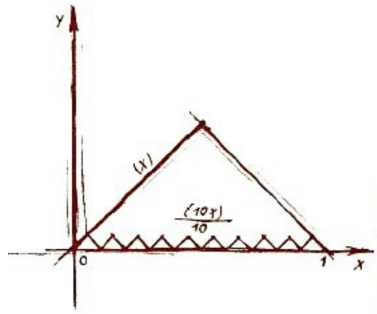
$$x = 0, a_1 a_2 a_3 ...$$

werde durch das Verbot eindeutig gemacht, keine unendlichen Folgen aus lauter Neunen zu verwenden. (r) möge den Abstand der reellen Zahl r von der nächstgelegenen ganzen Zahl bezeichnen. Mit dieser Bezeichnung gilt für beliebige x und $n = 0{,}1{,}2{,}...$

$$(10^n x) = 0{,}a_{n+1}a_{n+2}... \quad \text{bzw.} = 1 - 0{,}a_{n+1}a_{n+2}...$$

je nach dem $0{,}a_{n+1}a_{n+2}...$ den Wert $\frac{1}{2}$ nicht übersteigt oder übersteigt. Wir vermerken dazu, dass die Änderung der Ziffer a_k sich auf die Werte von $(10^k x)$, $(10^{k+1}x)$, ... offensichtlich nicht mehr auswirkt.

Wir definieren eine zum festen Wert x gehörige Nullfolge $h_1, h_2, ...$ durch die Forderung, dass für $a_m = 4$ oder 9 der Wert von $h_m = -10^{-m}$ sei, für alle anderen Ziffern aber $= 10^{-m}$. Diese Wahl ist so getroffen, dass $(10^n[x+h_m])$ und $(10^n x)$ beide zugleich $\leq \frac{1}{2}$ oder $> \frac{1}{2}$ ausfallen, so dass für $n < m$

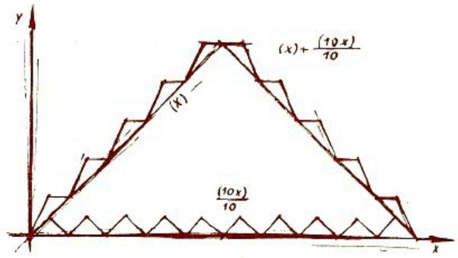


$(10^n[x+h_m]) = 0{,}a_{n+1}a_{n+2}... + 10^n h_m = (10^x) + 10^n h_m$
bzw. $= 1 - [0{,}a_{n+1}a_{n+2}... + 10^n h_m] = (10^x) - 10^n h_m$ ist.

Für $n \geq m$ gilt, wie wir schon wissen, das Gleichheitszeichen:

$$(10^x[x + h_m]) = (10^n h_m)$$

Nach diesen Vorbereitungen bilden wir die Reihe



$$f(x) = (x) + \frac{(10x)}{10} + \frac{(10^2 x)}{10^2} + ...$$

die durch die geometrische Folge mit dem Quotienten $\frac{1}{10}$ majorisiert wird und daher gleichmäßig konvergiert. Da die einzelnen Glieder der Reihe stetige Funktionen sind, erweist sich die Grenzfunktion $f(x)$ als stetig.

Der Differenzenquotient

$$\frac{f(x + h_m) - f(x)}{h_m}$$

lautet als Reihe

$$\sum_{n=0}^{\infty} \frac{(10^n[x + h_m]) - (10^n x)}{10^n h_m}$$

Alle Glieder, vom m-ten angefangen, verschwinden, wie wir schon wissen, während jedes der ersten $m - 1$ Glieder den Beitrag ± 1 liefert. Jeder neue Summand ändert so die ganzzahlige Reihensumme, wenn sie gerade ist, in eine ungerade Zahl, und wenn sie ungerade ist, in eine gerade Zahl ab.
Nun sind für $m = 1$ die Reihensummen, nämlich 1 oder -1, und m beide übereinstimmend ungerade, daher müssen fortan m und die Reihensumme für alle m beide zugleich gerade oder beide zugleich ungerade ausfallen.
Da m als Index der Nullfolge h_m abwechselnd gerade und ungerade ist, folgt, dass der Differenzenquotient, eben die abwechselnd gerade und ungerade Reihensumme, keinem Grenzwert zustreben kann. Die überall stetige Funktion $f(x)$ ist nirgends differenzierbar.

Aus der Stetigkeit folgt also die Differenzierbarkeit nicht, wohl aber umgekehrt aus der Differenzierbarkeit die Stetigkeit. Die Differenzierbarkeit in a besagt nämlich, dass in

$$\frac{f(x) - f(a)}{x - a} = f'(a) + \varepsilon(x)$$

die Funktion $\varepsilon(x)$ mit x zugleich gegen Null konvergieren muss. Nun ist

$$f(x) = f(a) + (x - a)f'(a) + (x - a)\varepsilon(x)$$

Der zweite und dritte Summand konvergieren aber mit $x \to a$ gegen Null, daher gilt $f(x) \to f(a)$ für $x \to a$, die Funktion ist in a stetig.

Schreibt man den Differenzenquotienten wie vorhin als Summe $f'(a) + \varepsilon(x)$, so folgt

$$f(x) - f(a) = (x - a)[f'(a) + \varepsilon(x)]$$

Verschwindet die Ableitung an der Stelle a nicht, dann lässt sich δ so klein wählen, dass in $[a - \delta, a + \delta]$ überall $|\varepsilon(x)| < \frac{1}{2}|f'(a)|$ ausfällt. Die eckige Klammer besitzt dann in diesem ganzen Intervall das Vorzeichen von $f'(a)$.

Befindet sich x links von a, so ist $x - a$ negativ, rechts von a dagegen positiv. Dieser Vorzeichenwechsel der rechten Seite bedeutet aber, wenn die linke Seite der Gleichung betrachtet wird, dass im Falle $f'(a) > 0$ die Funktionswerte links von a kleiner, rechts von a größer ausfallen als $f(a)$.
In diesem Fall passiert also die Funktion die Stelle a wachsend, im Falle $f'(a) < 0$ dagegen fallend. Daher kann eine Funktion an einer Stelle kein Extremum haben, wenn sie dort eine von Null verschiedene Ableitung besitzt.

So viel verhilft bereits, die für die Physik wichtige Methode der kleinsten Quadrate zu begründen. Die Erfahrung lehrt, dass noch so sorgfältig vorgenommene Messungen mehr oder weniger voneinander abweichende Messwerte $x_1, ..., x_n$ liefern.
Man muss sich folglich auf ein Prinzip einigen, das eine Zahl x als abschließendes Messergebnis aus den $x_1, ..., x_n$ berechnen lässt. Das kann nur mit Hilfe einer plausiblen Konvention gelingen. Diese lautet:
Als Messergebnis gelte die Zahl x, die den Ausdruck

$$f(x) = (x - x_1)^2 + ...(x - x_n)^2$$

zum Minimum macht. Da ein Minimum nur dort zu erwarten ist, wo die Ableitung verschwindet, $f'(x) = 0$ ist, folgt

$$x = \frac{x_1 + ... + x_n}{n}$$

Besitzt $f(x)$ in einem ganzen Intervall $[a - d, a + d]$ eine Ableitung, die dazu noch stetig ist, dann wähle man $d' < d$ so klein, dass die Ableitung in $[a - d', a + d']$ überall das Vorzeichen von $f'(a)$ hat, $f'(a)$ selbst von Null verschieden vorausgesetzt. Wendet man das Ergebnis des vorletzten Absatzes auf die einzelnen Punkte von $[a - d', a + d']$ an, so folgt, dass $f(x)$ in diesem Intervall streng monoton ist.

Diese Einsicht werde auf die Funktion $f(x) = \dfrac{x}{\ln x}$ angewandt, die uns in der Zahlentheorie bereits begegnete. Die Ableitung

$$f'(x) = \frac{1}{\ln x} - \left(\frac{1}{\ln x}\right)^2$$

ist für $x > e$ positiv, weil dann $\ln x > 1$ ausfällt. Daher wächst $f(x)$ streng monoton. Sind also $2 < a < b$ natürliche Zahlen, dann gilt

$$\frac{a}{\ln a} < \frac{b}{\ln b}$$

Beide Seiten mit $\ln a \cdot \ln b$ multipliziert, folgt $a \ln b < b \ln a$. Da die Exponentialfunktion mit dem Exponenten wächst, folgt weiter

$$e^{a \ln b} < e^{b \ln a}$$

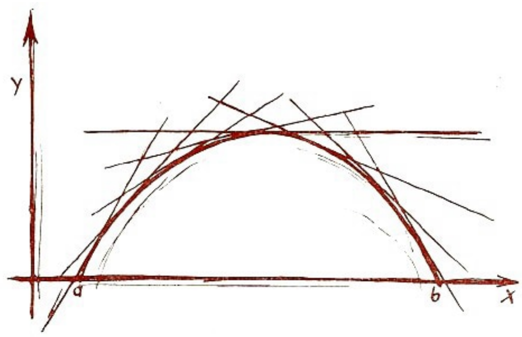
das heißt aber $b^a < a^b$ eine Aussage über natürliche Zahlen, in der nichts mehr auf den Umweg über die Ableitung hinweist.
Wählt man insbesondere $a = 3$ und $b = 27$, dann ist $27^3 < 3^{27}$. Das Potenzieren ist also nicht kommutativ. Die letzte Ungleichung, in der Form

$$(3^3)3 < 3^{3^3}$$

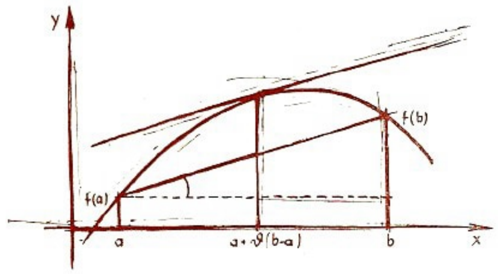
geschrieben, lässt erkennen, dass das Potenzieren auch nicht assoziativ ist.

Denkt man an die geometrische Bedeutung der Ableitung, dann legt die Anschauung einen Satz nahe, der für diesen Ideenkreis grundlegend ist.
Eine Kurve möge die Abszissenachse in den Punkten a und b schneiden. Verfolgt man die Richtungen, welche die Tangente bei einer Wanderung entlang des Kurvenbogens zwischen den Stellen a und b einnimmt, dann vermutet man sofort, dass die Tangente dabei auch parallel zur Abszissenachse wird. Das erkannte bereits der Franzose Rolle 1690.

Rein arithmetisch formuliert lautet sein Satz wie folgt.
$f(x)$ sei in $[a, b]$ überall differenzierbar und $f(a) = f(b) = 0$. Dann gibt es im Inneren von $[a, b]$ eine Stelle, wo die Ableitung von $f(x)$ verschwindet.
Da eine in $[a, b]$ überall differenzierbare Funktion $f(x)$ dort überall stetig ist, besitzt sie ein Maximum M und ein Minimum m.



Vom trivialen Fall $M = m = 0$ abgesehen, in dem $f(x)$ und folglich auch $f'(x)$ identisch verschwindet, muss mindestens M oder m verschieden von null ausfallen, sagen wir M. Da die in $[a, b]$ stetige Funktion $f(x)$ den Wert M in einem Punkt annimmt, kann es sich nur um einen inneren Punkt handeln, weil f in a und in b verschwindet. In diesem inneren Punkt muss dann die Ableitung verschwinden.



Der Satz von Rolle ist dem Mittelwertsatz der Differentialrechnung gleichwertig. Fasst man eine Gerade durch die Punkte $[a, f(a)]$ und $[b, f(b)]$, also die Sekante, als neue Abszissenachse auf, so ist nach Rolle die Tangente in einem inneren Punkt $a + \vartheta(b - a)$ mit $0 < \vartheta < 1$ des Intervalls $[a,b]$ parallel der Sekante.

Der Tangens des Winkels, den diese mit der ursprünglichen Abszissenachse bildet, beträgt

$$\frac{f(b) - f(a)}{b - a}$$

Diesen Wert hat man auf Grund der Parallelität der Ableitung

$$f'[a + \vartheta(b - a)]$$

gleichzusetzen, woraus

$$f(b) - f(a) = (b - a)f'[a + \vartheta(b - a)]$$

folgt.
Ein bemerkenswerter Satz von Darboux aus dem Jahre 1875 besagt, dass die Ableitung jeden Zwischenwert annimmt. Bemerkenswert ist der Satz deshalb, weil er unabhängig davon gilt, ob die Ableitung selbst stetig ist oder nicht. Dass der zweite Fall wirklich vorkommt, zeigt das Beispiel der Funktion

$$f(x) = x^2 \sin\frac{\pi}{x}$$

deren Definitionsbereich durch die Festsetzung $f(0) = 0$ ergänzt werde. Für $x \neq 0$ ist

$$f'(x) = 2x \sin\frac{\pi}{x} - \pi \cos\frac{\pi}{x}$$

während für $x = 0$ der Differenzenquotient

$$\frac{f(h) - f(0)}{h} = h \sin\frac{\pi}{h}$$

mit h zusammen gegen null strebt. Die damit überall erklärte Funktion $f'(x)$ ist an der Stelle $x = 0$ nicht stetig, weil für $x = \frac{1}{n}$ mit $n = 1,2,...$ der Wert von $f'(x)$ abwechselnd π und $-\pi$ ist.

Um den angekündigten Satz von Darboux zu beweisen, sei $f(x)$ in $[a, b]$ überall differenzierbar und c eine Zahl, die zwischen $f'(a)$ und $f'(b)$ liegt. Die Funktion

$$f^*(x) = f(x) - cx$$

besitzt dann eine Ableitung, die in a und in b Werte verschiedenen Vorzeichens annimmt, sagen wir $f'^*(a) > 0$ und $f'^*(b) < 0$. Daher muss f^* selbst in a nach rechts hin wachsen, in b von links her abnehmen. Das Maximum von $f^*(x)$ in $[a, b]$ muss daher größer ausfallen als die beiden Werte $f^*(a)$ und $f^*(b)$. Infolgedessen kann $f^*(x)$ sein Maximum nur in einem inneren Punkt von $[a, b]$ annehmen, so dass dort

$$f'(\xi) - c = 0$$

ist: Die Abbildung nimmt jeden Zwischenwert an.
Wir betrachten das Integral

$$\int_a^x f(t)dt$$

einer in $[a, b]$ stetigen Funktion $f(x)$. Es stellt eine Stammfunktion $F(x)$ von $f(x)$ dar; das heißt, in jedem Punkt x aus $[a, b]$ stimmt die Ableitung von $F(x)$ mit dem Integranden $f(x)$ überein. Der Differenzenquotient

$$\frac{F(x+h) - F(x)}{h}$$

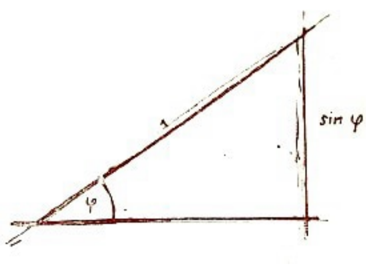
ist nämlich gleich

$$\frac{1}{h}\int_x^{x+h} f(t)dt$$

und nach dem ersten Mittelwertsatz der Integralrechnung, angewandt auf die stetige Funktion f, weiter gleich

$$\frac{1}{h} \cdot h \cdot f(x + \vartheta h)$$

Wegen der Stetigkeit von f folgt daraus, dass F in x eine Ableitung hat, die den Wert $f(x)$ besitzt.



Zum Schluss noch eine Anwendung der Differential- und Integralrechnung auf die Einführung von trigonometrischen Funktionen. Diese werden Üblicherweise geometrisch erklärt, der Sinus etwa als gegenüberliegende Kathete im rechtwinkligen Dreieck mit einer Hypotenuse von der Länge eins.

In der höheren Mathematik dagegen kommt es auf ihr analytisches Verhalten an. Diese Inkonsequenz kann dadurch vermieden werden, dass man die trigonometrischen Funktionen analytisch einführt. Das soll hier mit dem Sinus geschehen.

Im Integral

$$x = F(y) = \int_0^y \frac{dt}{+\sqrt{1-t^2}}$$

ist der Radikand im Integranden positiv, $F(y)$ daher in $[0, \vartheta]$, mit $0 \leq \vartheta < 1$, streng monoton zunehmend. Daher existiert eine eindeutige Umkehrfunktion $y = f(x)$ im Intervall $[F(0) = 0, F(\vartheta)]$. Da F Stammfunktion des Integranden ist, gilt für $0 \leq y < 1$

$$\frac{dx}{dy} = \frac{dF}{dy} = \frac{1}{+\sqrt{1-y^2}}$$

Die Ableitung von f als Umkehrfunktion von F ist, wie aus den Anfängen der Differentialrechnung bekannt, der reziproke Wert der Ableitung von F, also

$$\frac{dy}{dx} = \frac{df}{dx} = +\sqrt{1-y^2}$$

Identifiziert man $y = f(x)$ mit $\sin x$, dann erkennt man darin die bekannte Formel

$$\frac{d\sin x}{dx} = +\sqrt{1-\sin^2 x} = \cos x$$

wieder.

22 Ungleichungen

In der Analysis spielen Ungleichungen eine entscheidende Rolle. Eine fundamentale Ungleichung geht auf Cauchy zurück. Zunächst möchten wir dem Ursprung dieser Ungleichung nachgehen, um die reizvolle Entfaltung einer fruchtbaren Idee zu verfolgen.
Dazu stellen wir zunächst den Kosinus durch einen Ausdruck dar, der aus den Anfängen der Vektorrechnung fast unmittelbar folgt. Dort wird gezeigt, dass das innere Produkt von zwei Vektoren $\mathfrak{a} = (a_1{,}a_2)$ und $\mathfrak{b} = (b_1, b_2)$,

$$a_1b_1 + a_2b_2$$

gleich dem Produkt

$$|\mathfrak{a}| \cdot |\mathfrak{b}| \cos\vartheta$$

ist, mit ϑ den Winkel bezeichnet, den die beiden Vektoren einschließen. Die Gleichung

$$\cos\vartheta = \frac{a_1b_1 + a_2b_2}{|\mathfrak{a}| \cdot |\mathfrak{b}|}$$

erheben wir ins Quadrat, wobei $|\mathfrak{a}|^2 = a_1^2 + a_2^2$ und $|\mathfrak{b}|^2 = b_1^2 + b_2^2$ ist, und erhalten

$$\cos^2\vartheta = \frac{(a_1b_1 + a_2b_2)^2}{(a_1^2 + a_2^2)(b_1^2 + b_2^2)}$$

Da die linke Seite einen Wert zwischen null und eins hat, kann rechts der Zähler nicht größer als der Nenner ausfallen:

$$(a_1b_1 + a_2b_2)^2 \leq (a_1^2 + a_2^2) \cdot (b_1^2 + b_2^2)$$

Diese Ungleichung ist nun verallgemeinerungsfähig. Um das zu erkennen, betrachte man den Ausdruck

$$(a_1x + b_1)^2 + (a_2x + b_2)^2 + ... + (a_nx + b_n)^2$$

der als Summe von Quadraten offensichtlich keine negativen Werte annimmt. Entwickelt kann er in der Gestalt

$$x^2 + 2Bx + C$$

geschrieben werden mit

$$\begin{aligned} A &= a_1^2 + a_2^2 + ... + a_n^2 \\ B &= a_1b_1 + a_2b_2 + ... + a_nb_n \\ C &= b_1^2 + b_2^2 + ... + b_n^2 \end{aligned}$$

So wird man dazu geführt, sich mit dem letzten, in x quadratischen Ausdruck zu befassen, der niemals negativ wird. Wir behaupten, dass daraus die Ungleichung

$$0 \geq B^2 - AC$$

folgt. Zunächst bemerken wir, dass im Falle $A = 0$ auch noch B verschwinden muss, so dass dann die behauptete Ungleichung trivial wird. Mit $A = 0$ geht nämlich der quadratische in den linearen Ausdruck $2Bx + C$ über. Dieser würde aber, wenn B nicht verschwindet, für negative und positive x Werte verschiedenen Vorzeichens annehmen, entgegen der Voraussetzung.

Im Falle $A \neq 0$ aber kann der quadratische Ausdruck auch in der Gestalt

$$\frac{1}{A}(Ax+B)^2 - \left(\frac{B^2}{A} - C\right)$$

geschrieben werden. Soll dieser Ausdruck nie negativ sein, dann folgt

$$\frac{1}{A}(Ax+B)^2 \geq \frac{B^2}{A} - C$$

Nun ist unser spezielles A als Summe von Quadraten von vornherein nicht negativ, doch können wir weiterschließen, ohne uns darauf zu berufen. Wäre nämlich A negativ, dann würde die linke Seite in der letzten Ungleichung für große x negative Werte annehmen, die unter jeden vorgegebenen Wert, also auch unter den Wert der rechten Seite, herabsinken. Mit einem positiven A darf man aber multiplizieren:

$$(Ax+B)^2 \geq B^2 - AC$$

Daraus folgt für $x = -\frac{B}{A}$ die Behauptung.

Für A, B und C setzen wir die obigen Ausdrücke wieder ein und gewinnen damit die Ungleichung von Cauchy:

$$(a_1b_1 + a_2b_2 + ... + a_nb_n)^2 \leq (a_1^2 + a_2^2 + ... + a_n^2) \cdot (b_1^2 + b_2^2 + ... + b_n^2)$$

Setzt man darin für $b_1 = b_2 = ... = b_n = 1$, dann folgt

$$(a_1b_1 + a_2b_2 + ... + a_nb_n)^2 \leq n(a_1^2 + a_2^2 + ... + a_n^2)$$

Eine zweite wichtige Ungleichung vergleicht das arithmetische mit dem geometrischen Mittel von nichtnegativen Zahlen. Unter den verschiedenen Beweisen dafür wählten wir den aus, dessen reizvoller Gedankengang auch noch in anderen Überlegungen vorkommt.

Aus $(X_1 - X_2)^2 \geq 0$ folgt

$$X_1^2 + X_2^2 \geq 2X_1X_2$$

Das Gleichheitszeichen gilt nur im Fall $X_1 = X_2$. Wird für X_1^2 bzw. X_2^2 jetzt x_1 bzw. x_2 gesetzt, dann geht die letzte Ungleichung in

$$\frac{x_1 + x_2}{2} \leq \sqrt{x_1x_2}$$

über. Setzt man in dieser Ungleichung für x_1 bzw. x_2 weiter $\frac{x_1+x_2}{2}$ bzw. $\frac{x_3+x_4}{2}$ ein, dann folgt

$$\frac{\frac{x_1+x_2}{2} + \frac{x_3+x_4}{2}}{2} \geq \sqrt{\frac{x_1+x_2}{2} \cdot \frac{x_3+x_4}{2}} \geq \sqrt{\sqrt{x_1x_2}\sqrt{x_3x_4}}$$

also

$$\frac{x_1 + x_2 + x_3 + x_4}{4} \geq \sqrt[4]{x_1x_2x_3x_4}$$

Wird dieser Vorgang wiederholt, so folgt

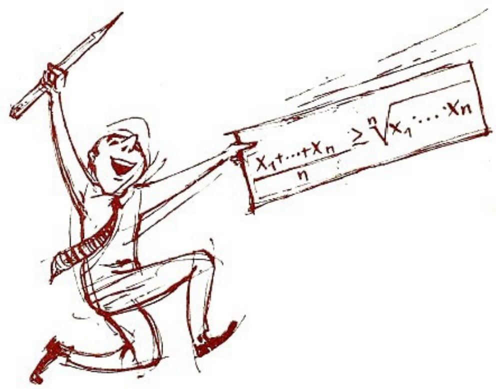
$$\frac{x_1 + x_2 + ... + x_{2^m}}{2^m} \geq \sqrt[2^m]{x_1x_2...x_{2^m}}$$

Die Ungleichung gilt aber nicht nur für die sukzessiven Potenzen von 2, sondern für jede natürliche Zahl. Das folgert man durch eine überraschende Wendung, durch den Schluss von n auf $n-1$. Gelingt es nachzuweisen, dass die Ungleichung, wenn sie für eine natürliche Zahl richtig ist, immer auch noch für die nächst niedrige natürliche Zahl gilt, dann folgt sofort, dass sie für jede natürliche Zahl n gelten muss. Denn wählt man m so groß, dass $n < 2^m$ ausfällt, gilt die Ungleichung ja für 2^m.

Daraufhin gilt sie auch für $2^m - 1$, dann für $(2^m - 1) - 1$ und so fort. So heruntersteigend landet man schließlich bei n.
Fermat hat als erster vom Schluss von n auf $n - 1$ systematischen Gebrauch gemacht, während sein Zeitgenosse Pascal sich des Schlusses von n auf $n + 1$ bediente.

Es bleibt daher nur noch der Nachweis zu führen, dass aus

$$\frac{x_1 + x_2 + ... + x_n}{n} \geq \sqrt[n]{x_1 x_2 ... x_n}$$



die entsprechende Ungleichung für die ersten $n - 1$ Zahlen folgt. Dazu setzen wir in der letzten Ungleichung für x_n den Ausdruck

$$\frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}$$

ein und erhalten so zunächst

$$\frac{x_1 + x_2 + ... + \frac{x_1+x_2+...+x_{n-1}}{n-1}}{n} \geq \sqrt[n]{x_1 x_2 ... x_{n-1} \frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}}$$

Die linke Seite ist

$$= \frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}$$

die rechte Seite

$$\sqrt[n]{x_1 x_2 ... x_{n-1}} \cdot \sqrt[n]{\frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}}$$

Die beiden Ausdrücke in die n-te Potenz erhoben und hinterher durch

$$\frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}$$

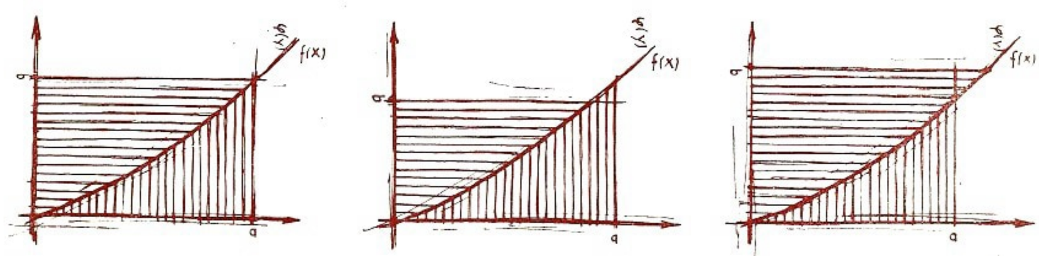
dividiert ergeben

$$\left(\frac{x_1 + x_2 + ... + x_{n-1}}{n - 1}\right)^{n-1} \geq x_1 x_2 ... x_{n-1}$$

woraus man schließlich durch Ausziehen der $(n-1)$-ten Wurzel die gewünschte Ungleichung erhält.

Eine dritte Ungleichung betrifft monotone Funktionen. Sie wurde vom Engländer Young 1912 aufgestellt, der sie allgemeiner als wir formulierte, dafür aber mit erheblichem Mehraufwand.

Es sei $f(x)$ eine streng monoton zunehmende stetige Funktion im Intervall $[0, a]$. Die Umkehrfunktion ist dann im Intervall $[f(0), f(A)]$ eine stetige und streng monoton zunehmende Funktion, wie wir schon wissen.
Fassen wir die $y = f(x)$ als Kurve auf, so tritt von der Ordinatenachse als Basis betrachtet die Umkehrfunktion $x = \phi(y)$ in Erscheinung. Es sei $f(0) = 0$. Liegen a bzw. b im Definitionsintervall von $f(x)$ bzw. $\phi(y)$, dann überdecken diesen senkrecht bzw. waagerecht schraffierten Flächen das eingezeichnete Rechteck von der Basis a und der Höhe b.



Diese Überdeckung erfolgt für $f(a) = b$ einfach, sonst aber ragt sie zum Teil Über das Rechteck hinaus.
Um das formelmäßig auszudrücken, bemerken wir, dass der Flächeninhalt des Rechtecks ab beträgt, und dass die senkrecht schraffierten Flächen den Inhalt $\int\limits_0^a f(x)dx$, die waagerecht schraffierten dagegen den Inhalt $\int\limits_0^b \phi(y)dy$ haben.
Die vorhin angeführte Überdeckung des Rechtecks lautet damit als Formel

$$ab \leq \int\limits_0^a f(x)dx + \int\limits_0^b \phi(y)dy$$

Wir wollen diese Einsicht benutzen, um daraus eine Ungleichung herzuleiten, die für die moderne Analysis von grundsätzlicher Bedeutung ist. Dazu wähle man $f(x) = x^{p-1} = y$ mit $p > 1$. Dann lautet die Umkehrfunktion $\phi(y) = x = \sqrt[p-1]{y}$.
Die Stammfunktionen von $f(x)$ bzw. $\phi(y)$ sind

$$\frac{x^p}{p} \qquad \text{bzw.} \qquad \frac{y^{\frac{1}{p-1}+1}}{\frac{1}{p-1}+1} = \frac{y^{\frac{p}{p-1}}}{\frac{p}{p-1}}$$

Der letzte Ausdruck kann noch, wenn für $\frac{p}{p-1}$ wie üblich q gesetzt wird, als $\frac{y^q}{q}$ geschrieben werden. Damit verwandelt sich die Ungleichung von Young in

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

23 Zu neuen Ufern

Die Analysis entwickelte sich in unserem Jahrhundert in einem Umfang, der frühere Zeiten weit in den Schatten stellte. Das hat verschiedene Ursachen?
Die Quellen flossen immer ergiebiger. Zunächst wurden Maß und Integral so erweitert, wie das für den Fortschritt unumgänglich erforderlich war. Dann erfuhr der Raumbegriff verschiedenartige Verallgemeinerungen, die unserer heutigen mathematischen und physikalischen Problematik zugrunde liegen.

In diese stürmische Entwicklung griffen verständlicherweise außer den Mathematikern auch noch die Physiker ein. Da sie jedoch, wie kaum anders zu erwarten, von der physikalischen Problemstellung ausgingen, waren ihre, Methoden manchmal mathematisch nicht genügend untermauert und führten gelegentlich zu physikalischen Scheinproblemen. Eine Zeitlang rechneten führende Physiker mit unbeschränkten Matrizen in einer Weise, die mathematisch unhaltbar war, wie v. Neumann bereits 1927 nachwies.

Eine andere Methode, die vom Engländer Dirac entwickelt wurde, erlebte durch den verallgemeinerten Funktionsbegriff von Laurent Schwartz zwei Jahrzehnte später ihre nachträgliche Rechtfertigung. In wenigen Jahrzehnten entstand so das Lehrgebäude der Funktionalanalysis, des mächtigsten mathematischen Instruments unserer Tage. Erfreulicherweise wirkt sich die Funktionalanalysis auch auf die zeitgenössische Physik in immer mehr steigendem Maße aus.

Ein recht allgemeiner Raumbegriff ist der metrische Raum. Er ist eine Menge, für deren Elementepaare x und y, die dann sinngemäß Punktepaare heißen, ein Abstand $\rho(x, y)$ definiert ist. Der Wert von ρ soll dabei für $x \neq y$ positiv ausfallen, für $x = y$ verschwinden. Weiter soll ρ symmetrisch sein, $\rho(x,y) = \rho(y,x)$, und schließlich der Dreiecksungleichung

$$\rho(x,z) \leq \rho(x,y) + \rho(y,z)$$

genügen. In der Ebene der Schulgeometrie ist für $P = (x, y)$ und $P' = (x', y')$

$$\rho(P, P') = \sqrt{(x - x')^2 + (y - y')^2}$$

ρ verhilft zur Einführung von Fundamentalfolgen $x_1, x_2, ...$ im metrischen Raum. Es sind Folgen, die dem Kriterium

$$\rho(x_m, x_n) < \varepsilon \qquad \text{für} \qquad m > N(\varepsilon), n > N(\varepsilon)$$

genügen. Für reelle Zahlen $x_1, x_2, ...$ war das mit $\rho(x_m, x_n) = |x_m - x_n|$ das Kriterium von Cauchy, weshalb Fundamentalfolgen auch Cauchy-Folgen heißen.
Eine Folge $x_1, x_2, ...$, die im Sinne von $\rho(x_n,x) \to 0$ gegen den Punkt x konvergiert, $x_n \to x$, bildet eine Fundamentalfolge. Das folgt aus der Dreiecksungleichung

$$\rho(x_m,x_n) \leq \rho(x_m,x) + \rho(x,x_n)$$

weil die rechte Seite für hinreichend große m und n beliebig klein ausfällt.

Gehören alle Punkte, gegen die Fundamentalfolgen aus Punkten eines metrischen Raumes konvergieren, diesem an, dann heißt der metrische Raum vollständig.
Uns interessieren hier kontrahierende Abbildungen K des metrischen Raumes in sich. Darunter versteht man Abbildungen, die jedem Punkt x des Raumes einen Punkt Kx desselben Raumes zuordnen, so dass

$$\rho(Kx,Ky) \leq q\rho(x,r) \qquad \text{mit} \qquad 0 < q < 1$$

ist.
Für kontrahierende Abbildungen in vollständigen metrischen Räumen hat der Pole Banach in seiner Doktorarbeit, abgedruckt 1922 in der Zeitschrift "Fundamenta mathematicae", einen Fixpunktsatz von großer Tragweite veröffentlicht.

Danach gibt es für K genau einen Fixpunkt x^*, für den also

$$Kx^* = x^*$$

gilt. Um den Nachweis dafür zu führen, gehen wir von einem beliebigen Punkt x_0 aus und bilden die Folge $x_1 = Kx_0, x_2 = Kx_1, \ldots$. Da K eine kontrahierende Abbildung ist, gilt

$$\rho(x_{n+1}, x_n) = \rho(Kx_n, Kx_{n-1}) \leq q\rho(x_n, x_{n-1})$$

Wiederholt man das bei $g(x_n, x_{n-1})$, so folgt

$$\rho(x_n, x_{n-1}) \leq q\rho(x_{n-1}, x_{n-2})$$

und weiter

$$\rho(x_n, x_{n-1}) \leq q^n \rho(x_1, x_0) = cq^n$$

wenn noch $\rho(x_1, x_0)$ mit c bezeichnet wird. Die wiederholte Anwendung der Dreiecksungleichung führt zur Abschätzung

$$\begin{aligned}\rho(x_m, x_n) &\leq \rho(x_m, x_{m+1}) + \rho(x_{m+1}, x_n) \leq \rho(x_m, x_{m+1}) + \rho(x_{m+1}, x_{m+2}) + \rho(x_{m+2}, x_n) \\ &\leq \ldots \leq \rho(x_m, x_{m+1}) + \ldots + \rho(x_{n-1}, x_n)\end{aligned}$$

Die rechte Seite ist aber

$$\leq cq^m + \ldots + cq^{n-1} \leq cq^m(1 + q + \ldots) = \frac{cq^m}{1-q}$$

so dass schließlich folgt, dass für hinreichend große m und n $\rho(x_m, x_n)$ beliebig klein ausfällt.
Da also die $x_0, x_1, \ldots$ eine Fundamentalfolge bilden, gibt es wegen der Vollständigkeit des Raumes darin einen Punkt x^*, gegen den die Folge konvergiert, $x_n \to x^*$. Aus

$$\rho(x_{n+1}, Kx^*) = \rho(Kx_n, Kx^*) \leq q\rho(x_n, x^*) \qquad \text{und} \qquad \rho(x_n, x^*) \to 0$$

folgt

$$\rho(x_{n+1}, Kx^*) \to 0 \qquad \text{also auch} \qquad \rho(x_n, Kx^*) \to 0$$

Die Dreiecksungleichung

$$\rho(x^*, Kx^*) \leq \rho(x^*, x_n) + \rho(x_n, Kx^*)$$

lässt erkennen, dass die linke Seite als fester Wert verschwinden muss, weil die rechte Seite für hinreichend große n beliebig klein wird. Die beiden Punkte x^* und Kx^* mit verschwindendem Abstand müssen aber zusammenfallen:

$$x^* = Kx^*$$

Es kann für K keinen zweiten Fixpunkt x^{**} geben, denn dann müsste

$$\rho(x^*,x^{**}) = \rho(Kx^*,Kx^{**}) \leq q\rho(x^*,x^{**})$$

sein, das heißt die positive Zahl links wäre kleiner als sie ist.

Verschiedene Interpretationen dieses Fixpunktsatzes ergeben Überraschend elegant die Existenz von Lösungen für gewöhnliche Differentialgleichungen, für lineare Integralgleichungen und anderes mehr. Um das auszuführen, betrachte man die Gesamtheit aller stetigen Funktionen über dem Intervall $[a,b]$.
Diese Menge wird durch eine geeignete Abstandsdefinition zum metrischen Raum.
Sind $f(x)$ und $g(x)$ zwei in $[a, b]$ stetige Funktionen, dann ist

$$|f(x) - g(x)|$$

ebenfalls stetig in $[a, b]$, besitzt daher dort ein Maximum. Dieses Maximum wird zum Abstand $\rho(f,g)$ der beiden Funktionen erklärt.
Die in $[a,b]$ stetigen Funktionen übernehmen damit die Rolle von Punkten des so definierten metrischen Raumes C. Dafür muss aber erst nachgewiesen werden, dass $\rho(f,g)$ den Forderungen genügt, die an den Abstand in metrischen Räumen gestellt wurden. Nichttrivial ist dabei allein die Dreiecksungleichung, die wie folgt begründet wird.

Aus

$$|f(x) - h(x)| \leq |f(x) - g(x)| + |g(x) - h(x)|$$

folgt, wenn die rechte Seite dadurch vergrößert wird, dass die beiden Summanden durch die Maximalwerte ersetzt werden, die sie annehmen können,

$$|f(x) - h(x)| \leq \rho(f,g) + \rho(g,h)$$

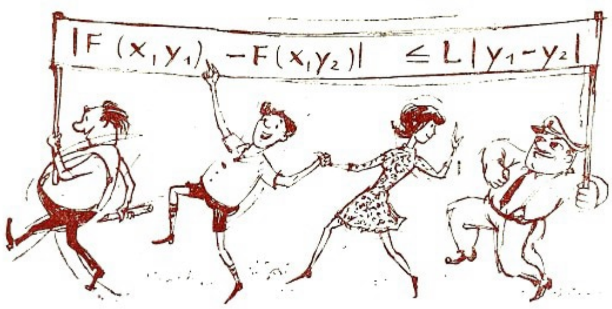
Rechts steht eine feste Zahl, die von der linken Seite nicht überboten werden kann, also auch nicht vom Maximalwert der linken Seite,

$$\rho(f,h) \leq \rho(f,g) + \rho(g,h)$$

Nun betrachten wir eine Funktion $F(x,y)$, die in der ganzen x,y-Ebene stetig sein und der Lipschitzbedingung

$$|F(x,y_1) - F(x,y_2)| \leq L|y_1 - y_2|$$

genügen soll, mit L eine positive Konstante bezeichnet.



Wir behaupten, dass die Abbildung K, die den "Punkt" $f(x)$ des C-Raumes in den "Punkt" $g(x)$ überführt, wobei $g(x)$ durch

$$g(x) = y_0 + \int_{x_0}^{x} F[t, f(t)]dt$$

berechnet werde, eine kontrahierende Abbildung ist. Dafür wählen wir δ so klein, dass $L\delta < 1$ ausfällt, und setzen $a = x_0 - \delta$, $b = x_0 + \delta$. Dann gilt für $Kf = g$ und $Kh = j$, wenn die Maxima für das Intervall $[x_0 - \delta, x_0 + \delta]$ gemeint sind,

$$\rho(g,j) = \max|g(x) - j(x)| = \max|\int_{x_0}^{x}[F(t,f(t)) - F(t,h(t))]dt|$$
$$\leq \max\int_{x_0}^{x} L[f(t) - h(t)|dt \leq L\delta \cdot \rho(f,h)$$

Die Rolle von q übernimmt hier $L\delta$.
Im Sinne des Fixpunktsatzes gibt es folglich genau einen Fixpunkt f mit $Kf = f$, ausgeschrieben

$$f(x) = y_0 + \int_{x_0}^{x} F(t, f(t))dt$$

Nach Differentiation ergibt sich, wenn noch $f(x) = y$ gesetzt wird,

$$y' = F(x,y)$$

Damit ist der Nachweis für die Existenz und Eindeutigkeit der Lösung, die durch den Punkt x_0, y_0 geht, für diese Differentialgleichung erbracht, ein Nachweis, der sonst umständliche Rechnungen verlangt.

Um die Lösung unserer Differentialgleichung zu approximieren, greifen wir auf die Folge $x_0, x_1, ... \to x^*$ zurück, die wir bei dem Beweis des Fixpunktsatzes betrachtet haben. Aus der Abschätzung dort

$$\rho(x_m,x_n) \leq \frac{cq^m}{1-q}$$

folgt, wenn wir bei festgehaltenem m den Grenzübergang $n \to \infty$ vollziehen, mit Hinblick auf

$$\rho(x_m, x^*) \leq \rho(x_m,x_n) + \rho(x_n, x^*)$$

und den Umstand, dass der zweite Summand beliebig klein wird, die Abschätzung

$$\rho(x_m, x^*) \leq \frac{cq^m}{1-q}$$

Daher ist

$$\rho(x_m,x^*) = \max|f_m(x) - f(x) \leq |\frac{L^m\delta^m}{1-L\delta}|\max f_1(x) - f_0(x)|$$

mit $f_0(x)$ eine beliebige in $[x_0 - \delta, x_0 + \delta]$ stetige Funktion und mit $f_1 = Kf_0$, $f_2 = Kf_1$, ... bezeichnet.
Ist die rechte Seite der Differentialgleichung nur im Rechteck

$$x_0 - a \leq x \leq x_0 + a \qquad \text{und} \qquad y_0 - b \leq y \leq y_0 + b$$

stetig, dann bezeichne m das Maximum von $|F|$ in diesem Rechteck. Unser Existenzbeweis lässt sich dann mit einem Wert $\delta < \min\left(a, \frac{b}{m}, \frac{1}{L}\right)$ durchführen.

Nachwort

Wir hatten hier Streifzüge unternommen, die durch charakteristische Gedankengänge moderner Mathematik führten. Vom Zwang einer Systematik befreit, besitzen sie den Vorzug von Unmittelbarkeit. Diese regt dann zum eingehenderen Studium der Mathematik an, das ein Gebot der Stunde geworden ist.

Die Zeiten, da es noch möglich war, die gesamte zeitgenössische Mathematik zu beherrschen, sind endgültig vorbei. Der Stoff ist derart angewachsen, dass man nur noch einzelne mathematische Disziplinen mit Erfolg bearbeiten kann. Die Wahl hat man nach der eigenen Veranlagung zu treffen.

Dieser Wahl muss aber ein gewissenhaftes Studium der Grundlagen vorausgehen. Gerade dazu sollten unsere Streifzüge wie auch mein Buch "Geometrische Plaudereien" anregen. Der Plauderton ist dabei geeignet, ein erstes Eindringen zu erleichtern und auf ein systematisches Studium vorzubereiten.