Autorenkollektiv

Streifzüge durch die Mathematik
Band 2

Übersetzung: Ludwig Boll, Horst Antelmann, Brigitte Mai
Illustrationen: Heinz Bormann
1966 Urania Verlag Leipzig / Jena / Berlin
MSB: Nr. 24
Abschrift und LaTex-Satz: 2022 https://mathematikalpha.de

Vorwort

Diese einführenden Bemerkungen sind vor allem für jene Leser gedacht, die den zweiten Band der "Streifzüge durch die Mathematik" zur Hand nehmen, ohne den ersten bereits gelesen zu haben.

Die "Streifzüge durch die Mathematik" sind die Übersetzung des mathematischen Teils einer zehnbändigen sowjetischen Jugendenzyklopädie. Nur an einigen Stellen wurden kleinere Änderungen oder Ergänzungen angebracht.

Gerade auf dem Gebiet der Mathematik fehlte in unserem Bücherangebot bisher eine Darstellung von Ideen, Methoden und Ergebnissen, die schon dem Schüler, dem Jugendlichen oder auch dem nicht speziell vorgebildeten Erwachsenen zugänglich ist. Das vorliegende Werk knüpft auf lebendige Art und Weise an den Schulstoff oder an den im täglichen Leben gewonnenen Erfahrungsschatz an; es zwingt und erzieht dabei zum selbständigen Nachdenken, stellt interessante Aufgaben und liefert Hinweise und Stoff zur weiteren Beschäftigung mit mathematischen Problemen.

Den Autoren kam es dabei unter anderem auch darauf an, den Schleier zu lüften, der die sogenannte höhere Mathematik von der Elementar- oder Schulmathematik trennt. Dabei waren sie auch bestrebt - ganz im Sinne von Richard Courant -, die Mathematik nicht als "abgeschlossenes System in sich ruhender Wahrheiten ohne Erinnerung an Herkunft und Ziel", sozusagen als Nabelschau eines weltentrückten Geistes, erscheinen zu lassen.
Sie legten vielmehr Wert darauf, die Tatbestände und Zusammenhänge der wirklichen Welt, die zu den Begriffsbildungen und Methoden der Mathematik geführt haben, klar hervortreten zu lassen. Nur wenn man sich darüber im klaren ist, dass auch die Mathematik eine - allerdings über viele Stufen der Abstraktion zustande gekommene - Widerspiegelung der Wirklichkeit ist, kann man verstehen, weshalb sie so vielfältige Anwendungen findet und im Zeitalter der technisch-wissenschaftlichen Revolution zur unmittelbaren Produktivkraft wird.

Wenn in diesem Zusammenhang von Mathematik die Rede ist, denkt jeder zuerst an die moderne maschinelle Rechentechnik, an die geradezu phantastisch anmutenden Geschwindigkeiten, mit denen die heutigen elektronischen Rechenautomaten außerordentlich komplizierte Aufgaben lösen können - Aufgaben, die man vor etwa fünfzig Jahren nicht einmal zu formulieren gewagt hätte.

Mit vollem Recht wird daher im ersten Band der "Streifzüge durch die Mathematik" sehr ausführlich auf das Rechnen, auf Rechenhilfsmittel und die Grundlagen der modernen Rechentechnik eingegangen. Neben der Welt der Zahlen hat aber seit den ältesten Zeiten auch die Welt der Figuren und Körper dem Menschen Aufgaben gestellt und ihn zum Nachdenken angeregt. Daher nimmt die Geometrie im ersten Band ebenfalls einen größeren Platz ein.

Die Mathematik ist aber als Wissenschaft von Zahl und Raum keineswegs erschöpfend bestimmt. Damit erfasst man nur einen Teil ihres Inhaltes, und nicht einmal den größten. Gerade in der heutigen Zeit gewinnt sie als Wissenschaft der Beschreibung von

Strukturen und Prozessen immer mehr an Bedeutung, obwohl auch auf diesem Gebiet die Wurzeln bis in die ältesten Zeiten zurückreichen.

"Struktur" und "Prozess" bilden gewissermaßen den Hauptinhalt dieses Bandes. Zunächst wird der für die Beschreibung von Prozessen in Natur und Technik und auch im Bereich der Gesellschaft, zum Beispiel in der Volkswirtschaft, fundamentale Begriff des funktionalen Zusammenhanges erarbeitet, nachdem die wichtigsten formalen Hilfsmittel dazu vermittelt worden sind. Dieser Funktionsbegriff wird an vielen Beispielen aus Natur und Technik erläutert und vertieft; dabei kommen die mächtigen Hilfsmittel der Analysis, die Begriffe Ableitung (Differentialquotient) und Integral, ins Spiel.

Diese ermöglichen es, Prozesse in ihrem Ablauf zu erfassen, zu beschreiben und, was die Hauptsache ist, ihren weiteren Verlauf im voraus zu berechnen. Naturgemäß stellen gerade diese Abschnitte einige Anforderungen an die Mitarbeit des Lesers. Als Lohn seiner Anstrengungen winkt ihm dafür aber das Verständnis der mathematischen Hilfsmittel, ohne die - um nur ein Beispiel zu nennen - von Mondlandung und Raumfahrt keine Rede sein könnte.

In die Wissenschaft von Strukturen wird der Leser über Vektorrechnung und Zahlentheorie eingeführt. Wichtige algebraische Strukturen lernt er an einfachen Beispielen kennen. In diesem Zusammenhang wird er auch mit den Anfangsgründen der Theorie der unendlichen Mengen bekannt gemacht.
Begriffsbildungen und Methoden der allgemeinen Mengenlehre sind ja inzwischen als ein Fundament der Mathematik erkannt werden, so dass gerade dieser Abschnitt auf ganz besonderes Interesse stoßen wird.

Das abschließende Kapitel befasst sich mit der Wahrscheinlichkeitsrechnung und gibt einen kurzen Ausblick auf die Informationstheorie, einen noch sehr jungen Zweig der Mathematik, der bereits reiche Früchte trägt. Hier bleibt es allerdings bei ersten Anregungen.
Wer mehr wissen will, muss zu weiterer Literatur greifen, auf die im Literaturverzeichnis hingewiesen wird.

Wer die "Streifzüge durch die Mathematik" gründlich gelesen hat, wird imstande sein, sich in der angegebenen Literatur zurechtzufinden und sie mit Erfolg zur Weiterbildung - und nicht zuletzt zum Vergnügen - zu benutzen.

Ludwig Boll

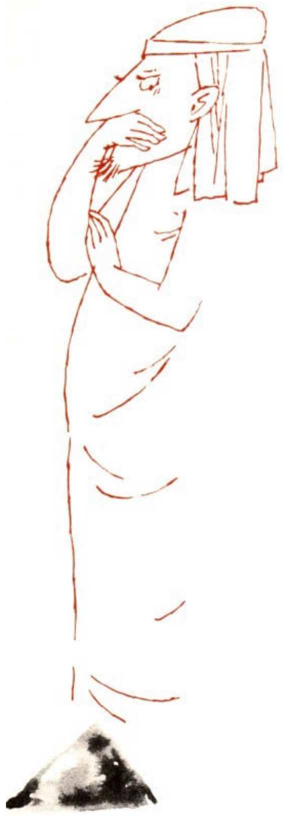
Anmerkung: Im Originaltext überlagern sich Grafiken, Abbildungen und Text teilweise so sehr, dass eine Übernahme in diese Abschrift nicht möglich war.
Daher wurden in dieser Abschrift einige Darstellungen neu gezeichnet, wobei die inhaltliche Aussage der Originalabbildung nicht verändert wurde.
Leider war es nicht möglich bei wenigen Zeichnungen eine äquivalente Abbildung zu erstellen, so dass diese Zeichnungen nur teilweise vorhanden sind oder fehlen.

Inhaltsverzeichnis

1 Gleichungen und Funktionen

1.1 Wie die Menschen lernten, Gleichungen zu lösen

I. J. Depman



Aus der Zeit zwischen dem Ende des zweiten Jahrtausends vor unserer Zeitrechnung und dem Beginn unserer Zeitrechnung stammt eines der ältesten mathematischen Schriftstücke: "Vorschrift zu gelangen zur Kenntnis aller dunklen Dinge ... aller Geheimnisse, die enthalten sind in den Gegenständen. Verfasst wurde dieses Buch im Jahre 33, Mesori Tag ... unter dem König von Ober- und Unterägypten Ra-a-us Leben gehend, nach dem Vorbild von alten Schriften, die verfertigt wurden in den Zeiten des Königs ...at durch den Schreiber Ahmes."
(Zitiert nach F. v. Krbek: Eingefangenes Unendlich, 2.Auflage, Seite 88.)

Ein Teil der Überschrift ist beschädigt, was hier durch Punkte angedeutet ist. Dieses Schriftstück ist eine ägyptische Aufgabensammlung, das "Rechenbuch des Ahmes" (Ahmes lebte um 1650 v.u.Z.). Es wird nach dem englischen Ägyptologen Alexander Henry Rhind, der dieses Schriftstück im Jahre 1858 erwarb, auch "Papyrus Rhind" genannt.

Der "Papyrus Rhind" enthält eine Reihe von Aufgaben, die wir heutzutage mit Hilfe von Gleichungen ersten Grades lösen würden. Ahmes löste sie mit der "Methode des falschen Ansatzes", die im Laufe mehrerer Jahrtausende bei verschiedenen Völkern zur Lösung solcher Aufgaben angewendet wurde. Eine ähnliche Methode wird auch heute noch im Arithmetikunterricht in der Schule benutzt.

Die Ägypter hatten um 2000 v.u.Z. zur Bezeichnung der Unbekannten ein besonderes Symbol und einen besonderen Namen.
Sie nannten sie "aha", was soviel wie eine Menge bedeutet und etwa mit "Haufen" übersetzt werden kann.
Als Beispiel wollen wir einmal die Aufgabe 24 aus dem "Papyrus Rhind" lösen; sie lautet:

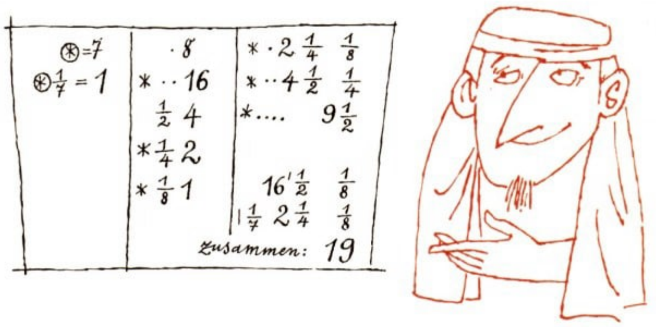
"Haufen. Sein siebter Teil, sein Ganzes. Was setzt zusammen 19."

Das bedeutet in unserer heutigen Redeweise:
Man bestimme diejenige Zahl, die, um ihren siebenten Teil vermehrt, 19 ergibt. Bei dieser Formulierung der Aufgabe wird stillschweigend angenommen, es gebe eine und nur eine solche Zahl. Vorsichtiger müsste man sagen: Man bestimme eine Zahl, die ...
Es ist also die Gleichung

$$x + \frac{x}{7} = 19$$

zu lösen. In der nachstehenden Zeichnung sehen wir, wie die alten Ägypter diese Aufgabe lösten.

Dabei wird der folgende Lösungsgang eingeschlagen:
Es wird versuchsweise angenommen, der "Haufen" sei gleich 7. Dann ist $\frac{1}{7}$ davon gleich 1 (siehe die erste Spalte). Die Sternchen sollen angeben, dass diese Zahlen für die Lösung der Aufgabe in Frage kommen.



Unter dieser Voraussetzung wäre die rechte Seite der zu lösenden Gleichung gleich 8. Deshalb steht oben in der zweiten Spalte die Zahl 8. Sie ist kleiner als die geforderte Zahl 19. Jetzt wird die Zahl 8 in Gedanken verdoppelt, und man erhält 16 (also wieder weniger als 19).

Eine weitere Verdoppelung würde die Zahl 32 ergeben, also größer als 19 sein. Deshalb wird 16 als Bestandteil der Lösung mit einem Stern versehen. Es fehlt nun noch 19 - 16 = 3. Man versucht jetzt, $\frac{1}{2}$ von 8, also 4 zu nehmen.
Dies kann aber kein weiterer Bestandteil der gesuchten Zahl sein, da nur noch 3 fehlt. Dann nimmt man $\frac{1}{4}$ und $\frac{1}{8}$ von 8, also 2 und 1 (siehe die zweite Spalte), und versieht sie mit einem Sternchen, da sie zusammen mit 16 die Zahl 19 bilden.

Somit ist in der zweiten Spalte angegeben, dass der ursprünglich angenommene Wert des "Haufens" $\left(2 + \frac{1}{4} + \frac{1}{8}\right)$ mal genommen werden muss, damit die Voraussetzung der Aufgabe erfüllt ist.
Es muss also 7 mit $2 + \frac{1}{4} + \frac{1}{8}$ multipliziert werden. (Die alten Ägypter und auch viele andere Völker schrieben nicht das Pluszeichen beim Addieren; wir wollen deshalb ebenfalls $2\frac{1}{4}$ statt $2 + \frac{1}{4}$ usw. schreiben.)

Es wird nun nicht 7 mal 2 $\frac{1}{4}$ $\frac{1}{8}$, sondern 2 $\frac{1}{4}$ $\frac{1}{8}$ mal 7 gerechnet. Dazu schreibt man in die dritte Spalte die Zahl 2 $\frac{1}{4}$ $\frac{1}{8}$, dann verdoppelt man sie (. .), und dann vervierfacht man sie (....)
Diese drei Zahlen addiert man und erhält 16 2 $\frac{1}{2}$ $\frac{1}{8}$.

Dazu fügt man $\frac{1}{7}$ des "Haufens" hinzu, also 2 $\frac{1}{4}$ $\frac{1}{8}$, und kommt so zu dem Ergebnis 19.

Wir erinnern in diesem Zusammenhang an die Ausführungen über die ägyptische Art des Multiplizierens und Dividierens durch Verdoppeln und Halbieren (vgl. dazu Band 1).

1.2 Die Methode des doppelten falschen Ansatzes

Später wurde dann von verschiedenen Völkern (in China und Indien) die Methode des doppelten falschen Ansatzes geschaffen.
In China war diese Methode unter der Bezeichnung "Überschuss - Mangel" bekannt.
Zu dieser Methode gaben die Araber eine Regel an, die sich bis ins 18.Jahrhundert erhalten hat und die unter anderem der russische Mathematiker Leonti Filippowitsch Magnizki (1669-1738) in seine "Arithmetik" aufnahm.
Dieses Buch wurde 1709 geschrieben und blieb bis zur Mitte des 18. Jahrhunderts das grundlegende Mathematiklehrbuch in Russland.

Magnizki ist sozusagen der russische Adam Ries. Wir wollen untersuchen, wie Magnizki diese Methode zur Lösung einer Aufgabe angewendet hat.

Aufgabe:
Ein Vater fragt den Lehrer seines Sohnes, wieviel Kinder er unterrichte. Der Lehrer antwortet: "Hätte ich noch einmal soviel Schüler, wie ich jetzt habe, und dann noch die Hälfte und dazu ein Viertel und dann noch Deinen Sohn, so wären es genau 100."

Wenn wir die Aufgabe lösen sollten, würden wir den Ansatz

$$x + x + \frac{x}{2} + \frac{x}{4} + 1 = 100$$

machen und diese Gleichung nach x auflösen.

Lösung nach Magnizki: Wir nehmen an, die Anzahl der Schüler sei 24. Dann hätten wir $24 + 24 + 12 + 6 + 1 = 67$, also $100 - 67 = 33$ weniger als angegeben.
Nun nehmen wir an, der Lehrer habe 32 Schüler. Dann ist $32+32+16+8+1 = 89$, das heißt $100-89 = 11$ weniger als 100. Auf der Grundlage der im Laufe von Jahrhunderten entwickelten Regel des doppelten falschen Ansatzes gibt Magnizki zur Bestimmung von x eine fertige Formel an:

$$x = \frac{32 \cdot 33 - 24 \cdot 11}{33 - 11} = 36$$

Ergibt sich bei dem einen Ansatz mehr und bei dem anderen weniger, als in der Aufgabe vorausgesetzt wurde, so müssen in dieser Formel statt der Minuszeichen Pluszeichen stehen.
Setzt man z. B. zuerst 60 an, so ergibt sich $60 + 60 + 30 + 15 + 1 = 166$, also $166 - 100 = 66$ ("Überschuss"). Nimmt man beim zweiten Ansatz die Zahl 20, so ist $20 + 20 + 10 + 5 + 1 = 56$ und es ist $100 - 56 = 44$ ("Mangel"). In diesem Fall muss man also folgendermaßen rechnen:

$$x = \frac{60 \cdot 44 + 20 \cdot 66}{66 + 44} = 36$$

Ob der Leser einmal versucht, einen Beweis für diese Regel zu finden ?

1.3 Die Einführung der „Unbekannten“

In der Entwicklung der Lehre von den Gleichungen ist die Einführung des Begriffs der "Unbekannten"und eines Symbols dafür von besonderer Wichtigkeit gewesen.

Diese Unbekannte war bei den alten Ägyptern das Wort "aha" mit einem bestimmten Hieroglyphenbild. Auch Babylonier, Chinesen, Inder, Griechen und andere Völker benutzten ein bestimmtes Symbol für die Unbekannte.
Bei den europäischen Völkern setzte sich eine systematische Bezeichnung der Unbekannten mit x, y, z usw. im Mittelalter durch. Diese Bezeichnung ist etwa seit dem 16. bzw. 17. Jahrhundert durch die Arbeiten des unter dem Namen Vieta bekannten französischen Mathematikers Francois Viète (1540-1603) und es französischen Philosophen, Mathematikers und Physikers René Descartes (1596-1650) allgemein üblich geworden.

Die Lösungsmethoden für Gleichungen ersten Grades, die sich auf arithmetische Operationen stützen, entwickelten sich bei den verschiedenen Völkern im Laufe von Jahrhunderten. Die grundlegenden Verfahren waren dabei das Hinüberschaffen eines Gliedes von der einen Seite einer Gleichung auf die andere unter Umkehrung des Vorzeichens und das Zusammenfassen gleichartiger Glieder.
Das erste Verfahren erforderte den Begriff der negativen Zahl, der sich aber erst viel später entwickelte. Infolgedessen gab man der Gleichung eine solche Form, dass alle ihre Glieder positiv waren, und schuf spezielle Regeln für die verschiedenen Arten von Gleichungen mit positiven Gliedern.

Schon der arabische Mathematiker Muhammad ibn Musa al-Huwarizmi (um 825) erklärte in seinem Buch "al-kitab a-muktasar fi hisab al-gabr wa-l-muqabala", dass die Lösung einer Gleichung ersten Grades auf zwei Operationen zurückgeführt werden kann: auf das Hinüberschaffen einzelner Glieder von einer Seite der Gleichung auf die andere (al-gabr) und das Zusammenfassen gleichartiger Glieder (wa-l-muqabala).

Das Wort "al-gabr" bedeutet soviel wie "Wiederherstellung", d.h. Umwandlung eines negativen Gliedes (das damals keinen Sinn besaß) in eine (positive) Zahl. Die Bezeichnung "al-gabr" kehrt in dem bei allen Völkern gebräuchlichen Wort "Algebra" wieder.
Al-Huwarizmi verwendete Gleichungen in großem Maße zur Lösung der verschiedensten praktischen Probleme. Dies hatte zur Folge, dass die europäischen Völker die elementare Algebra und die Arithmetik als die Lehre vom Auflösen von Gleichungen auffassten.

Das kommt zum Ausdruck bei dem englischen Mathematiker und Naturwissenschaftler Isaac Newton (1643-1727) und dem Schweizer Mathematiker Leonhard Euler (1707-1783). In der Schule wird die elementare Algebra auch heute noch von diesem Standpunkt aus behandelt.

Von den griechischen Mathematikern beherrschte anscheinend Diophant (3. Jh. u.Z.) die Methoden zur Lösung von Gleichungen ersten Grades. Leider ist ein Teil seines Buches über die Lösung dieser Gleichungen verlorengegangen. Die Art und Weise, Gleichungen anzuschreiben, die uns heute so naturgemäß und einfach vorkommt, kam erst im 17.Jahrhundert in Arbeiten von Vieta, Thomas Harriot (1560-1651) und Descartes auf.
Das Gleichheitszeichen und die Klammern kamen sogar erst im 18. Jahrhundert allge-

mein in Gebrauch. Klammern führte Euler ein.

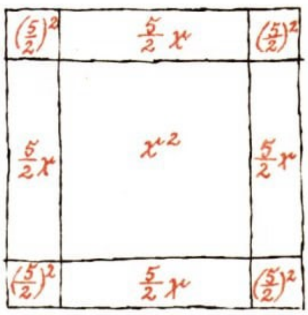
Lösungsmethoden für Systeme von Gleichungen ersten Grades entwickelten sich zuerst in Indien, China und bei den arabischen Völkern, in Europa erst vom 13. Jahrhundert an durch Leonardo Fibonacci von Pisa (1180?-1250?), Luca Pacioli (1445-1514) und den deutschen Mönch Michael Stifel (1487?-1547).
Zunächst gab es nur die Additions- und Subtraktionsmethode, dann aber kamen auch andere Methoden (Substitutionsmethode, Gleichsetzungsmethode) auf. Newton benutzte in seinen im Jahre 1707 herausgegebenen Vorlesungen schon alle diese Methoden.

1.4 Quadratische Gleichungen

Schon die Babylonier im 2. Jahrtausend vor unserer Zeitrechnung konnten Gleichungen zweiten Grades lösen, aber ihre Kenntnis von diesen Dingen hatte keinen Einfluss auf die Entwicklung der Wissenschaft in Europa. Ebenso waren auch die Leistungen andere Völker des Ostens in Europa lange unbekannt.

Die griechischen Mathematiker lösten quadratische Gleichungen; geometrisch: Euklid (365?-300? v. u. Z.) durch Konstruktion der mittleren Proportionale und äußere Teilung einer Strecke, Heron (um 100) und Diophant durch Methoden, die im wesentlichen mit unseren heutigen übereinstimmen.
Indische und chinesische Mathematiker betrachteten negative Wurzeln einer quadratischen Gleichung schon in den ersten Jahrhunderten unserer Zeitrechnung.
Aber noch der indische Mathematiker Bhaskara (1114- bis 1185?) hat negative Wurzeln nicht als Lösung einer Gleichung angesehen.

Verdienste bei der Untersuchung quadratischer Gleichungen hat sich auch der schon erwähnte arabische Mathematiker al-Huwarizmi erworben. Er leitete eine Lösungsformel her, die heute noch in vielen Lehrbüchern auftaucht. Al-Huwarizmi löste die Gleichung $x^2 + 10x = 39$ (Aufgabe 7 seines Buches) auf folgende Art:



Die gesuchte Größe x sei die Seite eines Quadrats. Dann konstruieren wir auf jeder Seite dieses Quadrats ein Rechteck, dessen Breite gleich dem vierten Teil des Koeffizienten von x, also gleich $\frac{10}{4} = \frac{5}{2}$ ist. Der Flächeninhalt der vier Rechtecke beträgt

$$4 \cdot \frac{5}{2}x = 10x$$

Der Flächeninhalt der so entstandenen kreuzförmigen Figur ist gleich $x^2 + 10x$, d. h. gleich der linken Seite der gegebenen Gleichung. Wir ergänzen nun diese Figur mit Hilfe kleiner Quadrate an den vier Ecken zu einem Quadrat mit der Seitenlänge $x + 5$.
Dieses große Quadrat hat also den Flächeninhalt $(x + 5)^2$. Ihn erhalten wir, wenn wir zu dem Flächeninhalt $x^2 + 10x$ der kreuzförmigen Figur den Flächeninhalt der vier Quadrate mit der Seitenlänge $\frac{5}{2}$, also $4\left(\frac{5}{2}\right)^2 = 25$ hinzufügen. Somit ist

$$(x+5)^2 = 39 + 25 = 64 \quad , \quad x + 5 = \pm 8$$
$$x = -5 \pm 8 \quad , \quad x_1 = 3, \quad x_2 = -13$$

Allgemein gilt für eine Gleichung $x^2 + px + q = 0$

$$\left(x + \frac{p}{2}\right)^2 = -q + \frac{p^2}{4}, \qquad x + \frac{p}{2} = \pm\sqrt{\left(\frac{p}{2}\right)^2 - q}$$
$$x_{1;2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

In Europa beherrschte Leonardo Fibonacci von Pisa zu Beginn des 13. Jahrhunderts Formeln zur Lösung quadratischer Gleichungen verschiedener Art. Auch die späteren Mathematiker konnten diese Gleichungen lösen.
Die allgemeine Formel stammt von Viéta, aber auch er erkannte nur positive Wurzeln an. Die italienischen Mathematiker des 16.Jahrhunderts, Geronimo Cardano (1501 bis 1576), Niccolo Tartaglia (1500?-1557), Ludovico Ferrari (1522 bis 1565) und Raffael Bombelli (um 1550), betrachteten neben den positiven Wurzeln nicht nur negative, sondern sogar schon komplexe Wurzeln als Lösungen. Seit dieser Zeit hat die Lösungsmethode für quadratische Gleichungen die heute übliche Form.

1.5 Gleichungen höheren als zweiten Grades

Auf Gleichungen dritten Grades stießen griechische Mathematiker wie Hippokrates (um 440 v.u.Z.), Archimedes (287?-212 v. u.Z.) u. a. bei der Lösung geometrischer Probleme, so bei der Verdoppelung des Würfels, dem sogenannten Delischen Problem, d. h. bei der Bestimmung der Kantenlänge des Würfels, dessen Volumen doppelt so groß ist wie das eines gegebenen Würfels, oder der Trisektion des Winkels, d.h. der Teilung eines gegebenen Winkels in drei gleiche.
Die rein geometrische Lösung (d.h. die Lösung mit Zirkel und Lineal erwies sich als unmöglich, als sich herausstellte, dass sich eine Wurzel einer kubischen Gleichung im allgemeinen nicht mit diesen Hilfsmitteln konstruieren lässt.

Diese Probleme wurden dann geometrisch mit Hilfe von Kurven (Hyperbel, Parabel) gelöst. Alle möglichen Fälle, die bei der Lösung kubischer Gleichungen mit geometrischen Methoden auftreten können, untersuchte der persische Mathematiker und Lyriker Umar ibn Ibrahim al-Hayyam (1044?-1123?).

$$x = \sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}$$

Algebraisch gelöst wurde die kubische Gleichung erst im 16. Jahrhundert durch die italienischen Mathematiker Scipione del Ferro (1403?-1526), Niccolo Tartaglia und Geronimo Cardano. Die Formel, die die Wurzeln der kubischen Gleichung $x^3 + px + q = 0$ durch die Koeffizienten p und q ausdrückt, wird Cardanische Formel genannt, obwohl Cardano nicht den Hauptanteil an ihrer Herleitung hatte.
Außerordentlich wichtig bei der Lösung einer allgemeinen kubischen Gleichung ist die Darstellung der Wurzeln in trigonometrischer Form.

Die algebraische Lösung von Gleichungen vierten Grades wurde im 16. Jahrhundert von Cardanos Schüler Ferrari gefunden. Eine besondere Lösungsmethode gab Euler im Jahre 1732 an.

Viele berühmte Mathematiker versuchten, auch Gleichungen fünften Grades algebraisch zu lösen, d.h., sie suchten Formeln, die die Wurzeln der Gleichung durch deren Koeffizienten ausdrücken. Aber diese Bemühungen waren erfolglos.
Viele Mathematiker, darunter der deutsche Gelehrte Gottfried Wilhelm Leibniz (1646-1716), Euler und der "Fürst der Mathematiker", Carl Friedrich Gauß (1777-1855), vermuteten, dass es für Gleichungen fünften und höheren Grades solche Formeln nicht gibt. Diese Vermutung wurde von dem norwegischen Mathematiker Niels Hendrik Abel (1802-1829) bewiesen.

Der französische Mathematiker Evariste Galois (1811-1832) entwickelte ein Verfahren, mit dessen Hilfe sich feststellen lässt, ob eine Gleichung algebraisch ("durch Radikale") lösbar ist oder nicht.
Der holländische Mathematiker Albert Girard (1595-1632) vermutete, dass eine Gleichung n-ten Grades n Wurzeln hat, wenn man auch die negativen und die komplexen Ausdrücke berücksichtigt. Girard selbst sah diese Ausdrücke nicht als Wurzeln der Gleichung an.

Descartes formulierte diesen Gedanken schon etwas entschiedener, aber erst Newton äußerte sich Ende des 17.Jahrhunderte mit voller Bestimmtheit. Euler behauptete im Jahre 1742, dass sich jedes Polynom n-ten Grades in Faktoren ersten und zweiten Grades mit reellen Koeffizienten zerlegen lässt.
Dies bedeutet mit anderen Worten, dass jede algebraische Gleichung mindestens eine Wurzel hat, die reell oder komplex sein kann. Streng bewiesen wurde diese Tatsache (Fundamentalsatz der Algebra) von dem 22 jährigen Gauß.

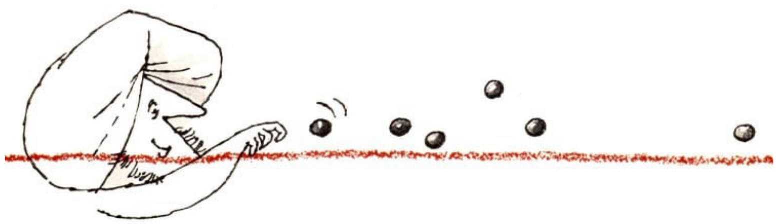
Zur näherungsweisen Berechnung der Wurzeln von Gleichungen höheren Grades gibt es viele Verfahren, die in der Praxis oft auch auf Gleichungen dritten und vierten Grades angewendet werden, da vielfach die Näherungswerte der Wurzeln völlig ausreichen und die Berechnung der exakten Werte schwierig ist.
Der persische Mathematiker al-Kasi (15. Jahrhundert) kannte schon eine gute Näherungsformel für die Wurzeln einer kubischen Gleichung, während Newton und andere Mathematiker Näherungsformeln für Gleichungen beliebigen Grades angaben. Ein sehr wichtiges Verfahren stammt von dem russischen Mathematiker Nikolai Iwanowitsch Lo-

batschewski (1792-1856).

Unabhängig von ihm fanden es der Franzose Pierre Dandelin (1794-1847) und der Schweizer Carl Heinrich Graeffe (179-1873), nach dem es in der deutschen Literatur meist benannt wird. Heute werden Gleichungen höheren Grades vielfach auf elektronischen Ziffernrechenautomaten oder mit Hilfe von Analogrechenmaschinen gelöst.

1.6 Wann entstanden die Fachausdrücke, die beim Studium von Gleichungen benutzt werden?

Abschließend wollen wir kurz die historischen Zeitabschnitte vermerken, in denen einige der heute geläufigen Fachausdrücke entstanden.
Heron (um 100) bezeichnete die Unbekannte mit r (tan); das bedeutet soviel wie "nichts". Der Inder Aryabhata (um 500 u.Z.) nannte die Unbekannte "gulika" (Kügelchen), zum Unterschied von der bekannten Größe "rupaka" (Münze von bekanntem Wert). Die späteren indischen Mathematiker gaben den Unbekannten Namen verschiedener Farben.



Diophant (3.Jh. u.Z.) bezeichnete die Unbekannte mit "Sigma" und den Ausdruck "ist gleich" mit ι (iota); das ist der Anfangsbuchstabe des griechischen Wortes "isosis" (Gleichung).
Die Römer schrieben hierfür "aequatio", und daraus entstand in den westeuropäischen Sprachen das Wort für "Gleichung", z. B. the equation (engl.), l'équation (franz.).

Unsere Bezeichnungsweise der Unbekannten durch den Buchstaben x ging wahrscheinlich aus dem italienischen Ausdruck "cosa" (Ding) hervor, der aus dem Arabischen entlehnt ist, wo das Wort für "Ding" mit einem unserem x ähnlichen Buchstaben anfängt. Descartes schließlich benutzte die letzten Buchstaben des Alphabets, zuerst besonders oft den Buchstaben z.

Gleichungen in der Schreibweise, dass sich auf der einen Seite alle Glieder befinden und auf der anderen Seite die Zahl 0 steht, traten im Jahrhundert auf (Harriot, Descartes), in einzelnen Fällen aber auch schon vorher (ein Beispiel für diese Schreibweise stammt sogar schon von den Babyloniern).
Statt xx, xxx, ... schrieb der Niederländer Simon Stevin (1548 bis 1620) einfach 2, 3, ... Descartes führte dafür die uns bekannten Abkürzungen x^2, x^3, ... ein. Der Ausdruck "Koeffizient" stammt von Vieta, "Grad einer Gleichung" von Descartes.

Die Bezeichnung "Wurzel einer Gleichung" benutzte Stifel (16. Jahrhundert), aber im Zusammenhang mit dem Wurzelziehen bei einer quadratischen Gleichung findet sich dieser Ausdruck schon bei Leonardo Fibonacci von Pisa (13. Jahrhundert).

Der Ausdruck "reziproke Gleichung" stammt von Euler, die Bezeichnung "Diskriminante" (das ist bei der quadratischen Gleichung $x^2 + px + q = 0$ der Ausdruck $p^2 - 4q$, von dem abhängt, ob die Gleichung reelle oder komplexe Wurzeln hat) von dem irischen Mathematiker William Rowan Hamilton (1805 bis 1865).

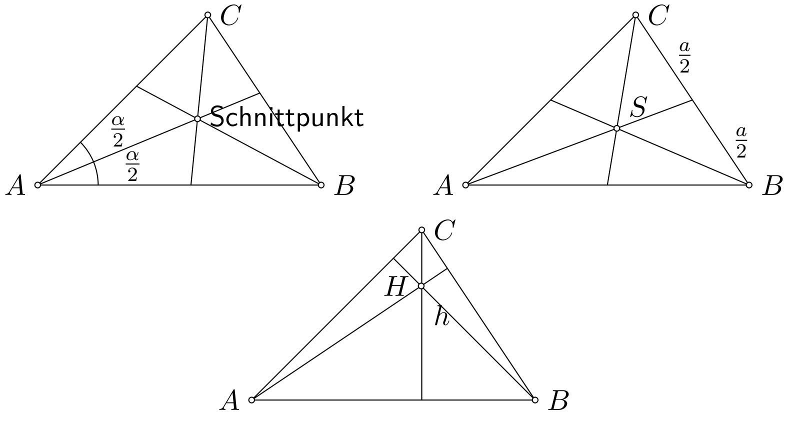
1.7 Was sind Koordinaten und wozu sind sie gut?

M. W. Potozki

Weshalb lassen sich manche geometrischen Probleme so schwer lösen?
Jeder weiß, mit welchen Schwierigkeiten man bei der Lösung vieler geometrischer Probleme zu kämpfen hat, beispielsweise bei geometrischen Konstruktionen, beim Aufsuchen von Bestimmungslinien oder beim Beweis von Sätzen. Worin bestehen nun diese Schwierigkeiten, und welche Ursache haben sie?

Die Schwierigkeiten bestehen u. a. darin, dass es in der Regel keine allgemeinen Verfahren gibt, mit deren Hilfe man solche Aufgaben lösen könnte.
Oft verlangt ein neues Problem auch eine besondere Lösungsmethode, die völlig anders sein kann als die Methoden, welche zur Lösung ähnlicher Probleme verwendet wurden. Wollen wir einen neuen Satz beweisen, so wissen wir vorher noch nicht, welche Hilfskonstruktionen wir benutzen müssen. Das Auffinden geeigneter Hilfskonstruktionen ist oft mit langem Suchen und allen möglichen Vermutungen verknüpft.

Sogar solche Behauptungen, die einander sehr ähnlich sind, müssen bisweilen mit völlig verschiedenen Methoden bewiesen werden. Wir brauchen nur an die Beweise für die Sätze über die drei merkwürdigen Punkte im Dreieck zu erinnern; diese Sätze lauten bekanntlich:[1]



1. Die drei Winkelhalbierenden eines Dreiecks schneiden einander in einem Punkt.
2. Die drei Seitenhalbierenden eines Dreiecks schneiden einander in einem Punkt.
3. Die Höhen eines Dreiecks schneiden einander in einem Punkt.

Diese drei Satz ähneln sich sehr. In jedem ist von drei Geraden die Rede, die einander in einem Punkt schneiden.

[1] Zeichnungen neu erstellt

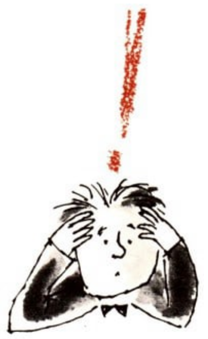
Natürlich handelt es in jedem Satz um andere Geraden; sonst wären es ja nicht drei verschiedene Sätze. Auch werden diese Sätze im Lehrbuch meist zusammen behandelt. Aber wie verschieden sind ihre Beweise!

Der erste Satz lässt sich ganz einfach beweisen: Die drei Winkelhalbierenden schneiden einander im Mittelpunkt des Inkreises. Jedoch ist es bei weitem nicht so einfach, die beiden anderen Sätze zu beweisen.

Dass es keine hinreichend allgemeinen Verfahren zur Lösung geometrischer Probleme und zum Beweis geometrischer Sätze gibt, wird oft als großer Mangel der Elementargeometrie empfunden. Dieser Mangel wird besonders dadurch fühlbar, dass in Naturwissenschaft und Technik wesentlich kompliziertere Kurven und Flächen auftreten, als wir sie von der Schule her kennen.
In der Technik begegnen uns z. B. solche Kurven wie Ellipse, Hyperbel, Parabel, Kettenlinie, Schraubenlinie, Kardioide usw. sowie Flächen wie Ellipsoide, Hyperboloide, Paraboloide, Pseudosphären und viele, viele andere.

Wir können uns leicht vorstellen, wieviel schwieriger es für die Geometer ist, Probleme für so komplizierte gekrümmte Flächen zu lösen, wenn sie mit heuristischen Ansätzen zum Ziel kommen und sozusagen jedes Problem einzeln lösen müssen.

Natürlich bemühten sich die Geometer, allgemeine Verfahren zur Lösung der verschiedenartigsten geometrischen Probleme zu finden. Sie wollten Verfahren haben, die für jede Aufgabe die Lösung liefern, ohne dass sie erst lange suchen mussten.
Aber gibt es überhaupt solche allgemeinen Verfahren? Ist es nicht sinnlos, nach ihnen zu suchen?
Es zeigt sich, dass es sie tatsächlich gibt. Um zu erfahren, wo sie etwa zu finden sind und wie sie aussehen können, verweilen wir etwas bei der Lösung arithmetischer und geometrischer Aufgaben, für die solche Verfahren existieren und dem Leser aus dem Schulunterricht gut bekannt sind.



Wir wissen, welchen Schwierigkeiten man schon bei nicht sehr verzwickten arithmetischen Aufgaben begegnet und wieviel einfacher es ist, diese Aufgaben algebraisch zu lösen. Woran liegt das, worin bestehen diese Vereinfachungen?

Die Schwierigkeiten bei der Lösung arithmetischer Aufgaben sind dieselben wie bei der Lösung geometrischer Aufgaben. Lesen wir die Aufgabenstellung durch, so erkennen wir aus dieser nicht ohne weiteres, welcher Weg zur Lösung einzuschlagen ist.

Oft muss man sogar zwei einander ganz ähnliche Probleme auf völlig verschiedene Weise lösen, und bei jedem neuen Problem muss man von neuem nachdenken, um das spezielle Lösungsverfahren zu finden.

Ganz anders verhält es sich dagegen in der Algebra. Dort wissen wir sofort, wie wir an die Lösung heranzugehen haben. Wenn wir die Unbekannten mit x, y, ... bezeichnet haben, können wir mit Hilfe der Voraussetzungen Gleichungen zwischen diesen Unbekannten aufstellen. Natürlich verlangt dies ebenfalls eine gewisse Fertigkeit, die jedoch schon von fast allen Schülern gemeistert wird. Diese Gleichungen lösen wir dann nach den bekannten Regeln der Algebra.

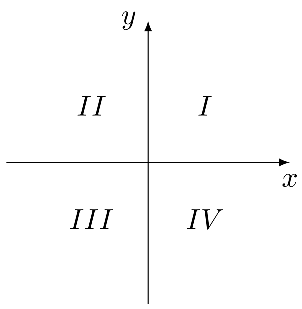
Somit gibt es in der Algebra für alle Probleme eine gemeinsame Lösungsmethode - nämlich das Aufstellen und Lösen von Gleichungen. Gerade dadurch ist es verhältnismäßig leicht, Probleme algebraisch zu lösen. Auch in der Geometrie werden zwar Gleichungen benutzt, aber bei weitem nicht in allen Fällen!
Dort, wo Gleichungen benutzt werden, vereinfacht dies sofort die Lösung. Jeder von uns weiß aus Erfahrung, wieviel einfacher Aufgaben sind, in denen etwas (mit Hilfe von Gleichungen) berechnet werden soll, als Aufgaben, bei denen etwas zu konstruieren oder zu beweisen ist.

Jedoch lassen sich mit Hilfe von Gleichungen nicht alle geometrischen Probleme lösen. Insbesondere können Aufgaben über Bestimmungslinien nicht immer algebraisch gelöst und geometrische Sätze nicht immer algebraisch bewiesen werden. Eben dies veranlasste, wie wir schon sagten, die Geometer, solche Lösungsverfahren zu suchen, die nach einer allgemeinen Regel zwar nicht auf alle, aber doch auf sehr viele geometrische Probleme angewendet werden können, sogar auf solche, bei denen neben den einfachsten geometrischen Kurven und Flächen auch kompliziertere auftreten, von denen vorhin schon die Rede war.
Diese Verfahren beruhen auf der Benutzung eines Koordinatensystems.

Bevor wir aber diese Verfahren erläutern können, müssen wir erst den Begriff des Koordinatensystems kennenlernen.

1.8 Das Koordinatensystem

Wir untersuchen hier nur Probleme aus der Planimetrie (der ebenen Geometrie) und werden deshalb nur ebene Koordinatensysteme betrachten.
Wir ziehen in der Ebene zwei zueinander senkrechte Geraden, der Einfachheit halber eine horizontale und eine vertikale, und bezeichnen ihren Schnittpunkt mit O. Der Punkt O heißt Koordinatenursprung oder Koordinatenanfangspunkt.
Eine bestimmte Strecke OE wählen wir als Einheit; das bedeutet, dass ihre Länge als Maßstabseinheit zum Messen aller Strecken dient, die wir auf den beiden Geraden abstecken. Der Einfachheit wegen wollen wir auf beiden Geraden den gleichen Maßstab wählen. Das braucht nicht immer so zu sein.



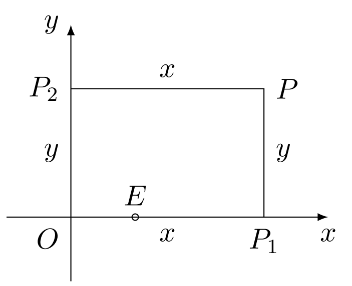
Da man auf jeder der beiden Geraden vom Punkt O aus in zwei Richtungen Strecken abstecken kann, legt man eine davon als positiv fest, und zwar gewöhnlich vom Punkt 0 aus nach rechts und nach oben. Die entgegengesetzten Richtungen, also vom Punkt O aus nach links und nach unten, werden als negativ angesehen.

Die beiden Geraden teilen die Ebene in vier Teile, die man Quadranten nennt. In der Abbildung sind sie mit römischen Ziffern gekennzeichnet. Die horizontale Gerade heißt

Abszissenachse (in der Zeichnung x-Achse), die vertikale Gerade Ordinatenachse (in der Zeichnung y-Achse). Beide Geraden zusammen bilden die Achsen eines Koordinatensystems.

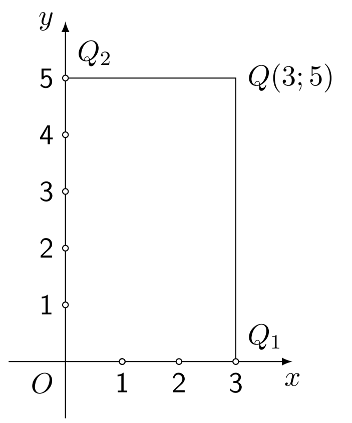
Dieses Koordinatensystem wird kartesisch genannt, nach dem französischen Mathematiker, Physiker und Philosophen René Descartes oder, latinisiert, Renatus Cartesius (1596-1650), der es als erster in seinem 1637 erschienenen Werk "La Géométrie" einführte.
Genauer müsste man es ein rechtwinkliges kartesisches Koordinatensystem nennen, da die beiden Geraden senkrecht aufeinanderstehen.

Das Koordinatensystem dient dazu, die Lage von Punkten, Strecken und verschiedenen geometrischen Figuren in der Ebene zu kennzeichnen. Gegeben sei in der Ebene, in der wir ein Koordinatensystem gezeichnet haben, ein Punkt P. Um die Lage von P im Koordinatensystem charakterisieren zu können, fällen wir vom Punkt P das Lot PP_1 auf die x-Achse und das Lot PP_2 auf die y-Achse.



Die Länge der Strecke OP_1 gemessen mit unserem angenommenen Maßstab und mit dem entsprechenden Vorzeichen versehen, heißt die Abszisse des Punktes P und wird in diesem x,y-Koordinatensystem mit x bezeichnet. Die Länge der Strecke OP_2, gemessen im gleichen Maßstab und ebenfalls mit dem entsprechenden Vorzeichen versehen, heißt die Ordinate des Punktes P und wird im x, y-Koordinatensystem mit y bezeichnet.

Die Vorzeichen von Abszisse und Ordinate werden folgendermaßen bestimmt: Liegt der Punkt P_1 rechts von O, also auf der positiven x-Achse, so versehen wir die Abszisse x mit dem positiven Vorzeichen; liegt er links von O, also auf der negativen x-Achse, so ist die Abszisse x negativ.
Liegt P_2 oberhalb von O, also auf der positiven y-Achse, so geben wir der Ordinate y das positive Vorzeichen; liegt P_2 unterhalb von O (auf der negativen y-Achse), so ist y negativ. Abszisse und Ordinate zusammen heißen die Koordinaten des Punktes P.

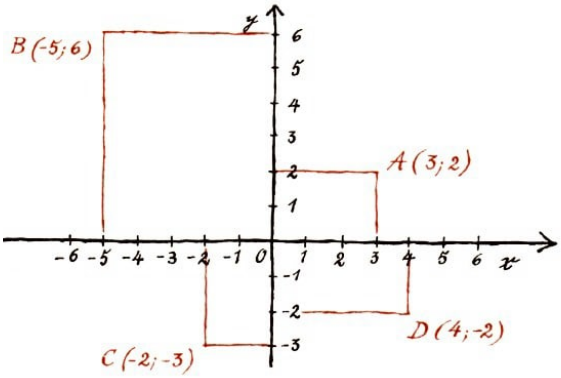


In diesem Beispiel gaben wir einen Punkt P vor und bestimmten seine Koordinaten. Es kann aber auch die umgekehrte Aufgabe gestellt sein, nämlich einen Punkt Q zu konstruieren, dessen Koordinaten gegeben sind.

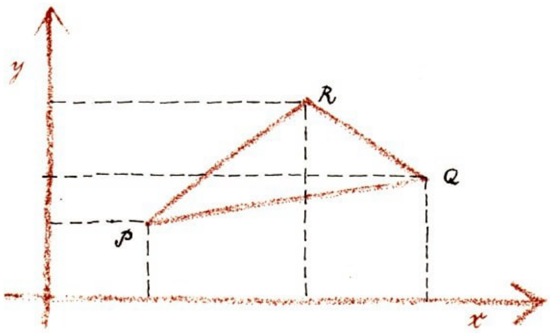
Gegeben sei also ein Koordinatensystem sowie die Abszisse $x = 3$ und die Ordinate $y = 5$ eines gewissen Punktes Q; das schreibt man in der Form $Q(x = 3; y = 5)$ oder, kürzer, $Q(3; 5)$. Die Abszisse steht immer an der ersten, die Ordinate an der zweiten Stelle.

Den Punkt Q konstruieren wir dann, indem wir die Strecke OQ_1 der Länge 3 auf der positiven x-Achse und die Strecke OQ_2 der Länge 5 auf der positiven y-Achse abstecken (3 und 5 sind ja positive Zahlen). In Q_1 errichten wir die Senkrechte auf der x-Achse,

in Q_2 die Senkrechte auf der y-Achse.
Der Schnittpunkt dieser beiden Senkrechten ist dann der Punkt Q.



In der vorliegenden Abbildung haben wir zum Beispiel die Punkte $A(3;2)$, $B(-5;6)$, $C(-2;-3)$ und $D(4;-2)$ konstruiert.
Wir sehen also, dass jeder Punkt der Ebene wohlbestimmte Koordinaten hat; umgekehrt definieren Abszisse und Ordinate zusammen genau einen Punkt der Ebene.

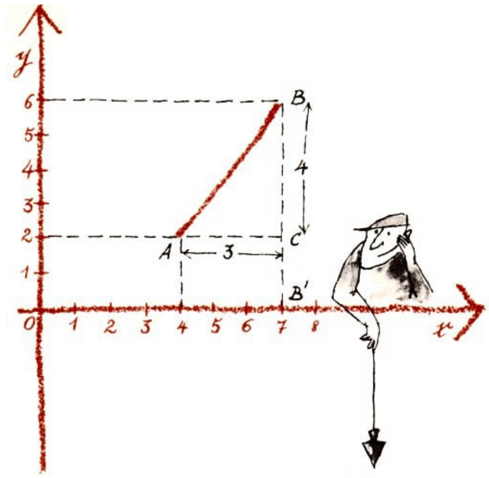


Ist in der Ebene ein Koordinatensystem gegeben, so können wir nicht nur die Lage von Punkten, sondern auch von beliebigen Figuren in der Ebene bestimmen. Wollen wir zum Beispiel die Lage eines Dreiecks PQR angeben, so genügt es, die Koordinaten seiner drei Ecken zu nennen.

Man wird nun natürlich folgendes vermuten: Kann man eine Figur, etwa ein Dreieck, durch Angabe der Koordinaten derjenigen Punkte beschreiben, aus denen sie besteht, so kann man auch alle Elemente der Figur bestimmen, insbesondere Seitenlängen, Winkel, Flächeninhalt usw.
Diese Vermutung wollen wir mit Hilfe des folgenden Beispiels bestätigen. Wir wollen zeigen, wie die Länge einer Strecke AB bestimmt wird, wenn die Koordinaten der Endpunkte A und B bekannt sind.

Beispiel:
Gegeben seien also wie in der Abbildung zwei Punkte A und B mit den Koordinaten $A(4;2)$ und $B(7;6)$, die Länge der Strecke AB sei zu berechnen.
Vom Punkt B fällen wir das Lot BB' auf die x-Achse, vom Punkt A das Lot auf die Strecke BB'. Den Fußpunkt dieses Lotes bildet der Punkt C auf BB'. Auf diese Weise erhalten wir ein rechtwinkliges Dreieck ABC. Nach dem Satz des Pythagoras ist bekanntlich

$$\overline{AB}^2 = \overline{AC}^2 + \overline{CB}^2$$



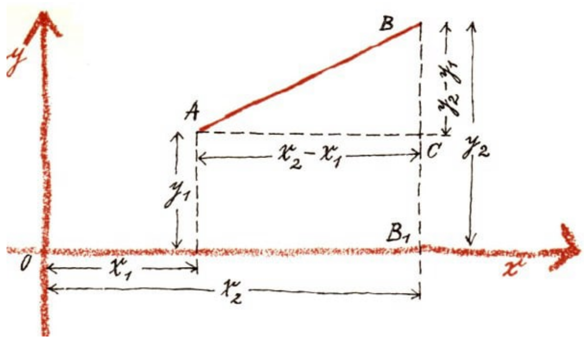
In unserem Fall gilt

$$\overline{AC} = 7 - 4 = 3 \qquad , \qquad \overline{CB} = 6 - 2 = 4$$

also ist

$$\overline{AB} = \sqrt{3^2 + 4^2} = 5$$

Wir können diese Formel zur Bestimmung der Länge einer Strecke auch in allgemeiner, algebraischer Gestalt angeben.



Dazu nehmen wir an, es seien zwei Punkte A und B durch ihre Koordinaten in allgemeiner Form gegeben, etwa durch $A(x_1, y_1)$ und $B(x_2, y_2)$. Dann bilden wir wie vorhin das rechtwinklige Dreieck ABC und erhalten

$$\overline{AC} = x_2 - x_1 \qquad , \qquad \overline{CB} = y_2 - y_1$$

Folglich ist nach dem Satz des Pythagoras

$$\overline{AB}^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$$

und schließlich ergibt sich die Länge der Strecke AB aus der Formel

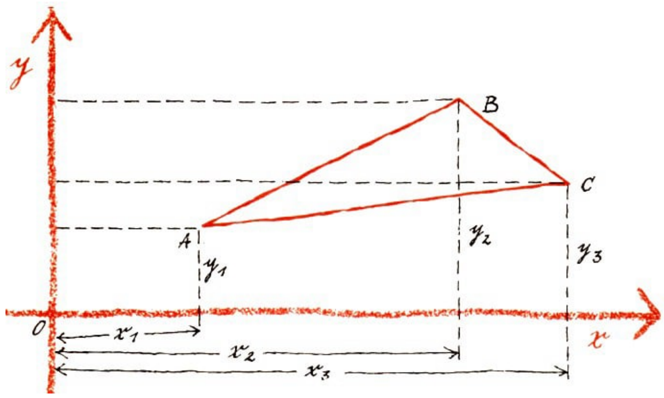
$$\overline{AB} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

Es lässt sich leicht zeigen, dass diese Formel bei beliebiger Lage der Punkte im Koordinatensystem gilt. Vor dem Wurzelzeichen steht stets das positive Vorzeichen, da die

Länge einer Strecke eine positive Größe ist.

Vergleichen wir die allgemeine Formel (1) mit der Formel aus dem vorhergehenden Beispiel, so sehen wir, dass diese ein Spezialfall der allgemeinen Formel (1) ist, und zwar für $x_1 = 4$, $y_1 = 2$, $x_2 = 7$, $y_2 = 6$.

Wir haben damit eine Methode gewonnen, mit deren Hilfe die Länge aller Strecken eines beliebigen Vielecks, z.B. der Seiten eines Dreiecks, eines Vierecks usw., berechnet werden kann.
An diesem Beispiel ist ersichtlich, dass wir mit Hilfe eines Koordinatensystems viele Aufgaben lösen können, sogar ohne die Figur selbst zu zeichnen. Es ist ganz klar, dass wir mit Hilfe der allgemeinen Formel (1) die Länge der Strecke zwischen den beiden Punkten $A(4;2)$ und $B(7;6)$ bestimmen können, ohne die Strecke AB zu zeichnen; wir brauchen dazu nur die Koordinaten der Punkte A und B in (1) einzusetzen. Dies ist besonders gut sichtbar im folgenden Beispiel.



Wir wollen ohne Beweis eine Formel angeben, mit deren Hilfe sich der Flächeninhalt eines Dreiecks bestimmen lässt, falls die Koordinaten der Ecken des Dreiecks bekannt sind.
Liegen die Ecken eines Dreiecks ABC in den Punkten mit den Koordinaten $A(x_1;y_1)$, $B(x_2;y_2)$ und $C(x_3;y_3)$, so hat die Formel für den Flächeninhalt F dieses Dreiecks die Gestalt

$$F = \frac{1}{2}|x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)|$$

oder, umgeformt,

$$F = \frac{1}{2}|(x_2 - x_1)(y_3 - y_1) + (x_3 - x_1)(y_2 - y_1)|$$

Die vertikalen Striche bedeuten, dass der absolute Betrag der zwischen ihnen stehenden Größe genommen werden soll.
Wir wollen nun an einem Zahlenbeispiel sehen, wie sich diese Formel anwenden lässt.

Beispiel:
Gegeben sei ein Dreieck ABC mit den Ecken $A(3;4)$, $B(5;7)$ und $C(9;6)$; man bestimme seinen Flächeninhalt F. In dieser Aufgabe ist

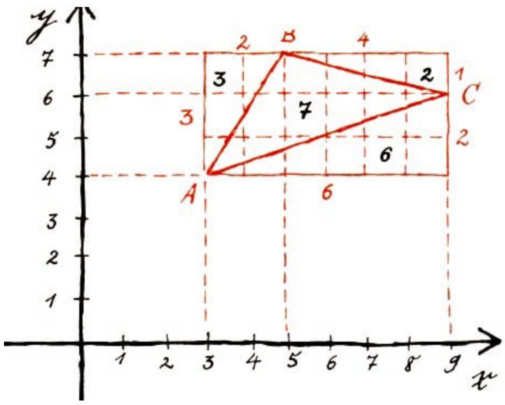
$$x_1 = 3, x_2 = 5, x_3 = 9, \qquad y_1 = 4, y_2 = 7, y_3 = 6$$

Setzen wir diese Werte in die obige Formel ein, so erhalten wir nach der ersten der angegebenen Formeln

$$F = \frac{1}{2}|3(7-6) + 5(6-4) + 9(4-7)| = \frac{1}{2}|3 + 10 - 27| = 7$$

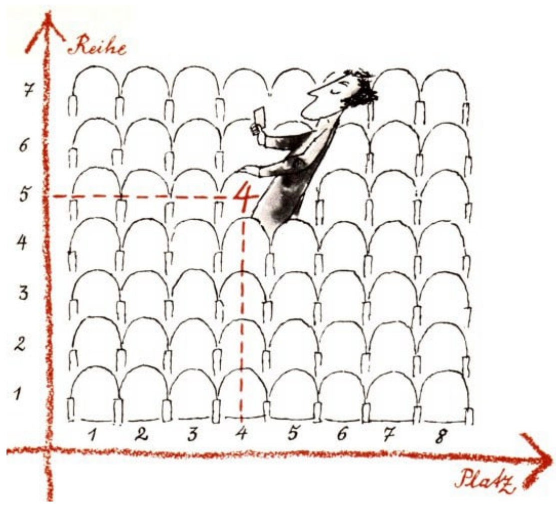
nach der zweiten

$$F = \frac{1}{2}|(5-3)(6-4) + (9-3)(7-4)| = \frac{1}{2}|-14| = 7$$

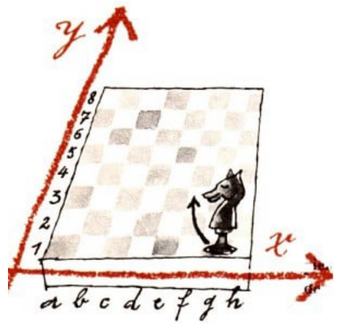


Wir haben also den Flächeninhalt des Dreiecks ohne jede Zeichnung bestimmt. Wenn wir uns jetzt eine Zeichnung anfertigen und das achsenparallele Rechteck, in dem das Dreieck liegt, in Quadrate mit der Seitenlänge 1 aufteilen, so sehen wir sofort, dass wir die Aufgabe richtig gelöst haben.

Auch hier können wir also mit Hilfe eines Koordinatensystems eine Aufgabe lösen, ohne die betreffende Figur erst zeichnen zu müssen. Das ist eine außerordentlich wichtige Eigenschaft des Koordinatensystems; wir brauchen die oft sehr komplizierten Figuren gar nicht erst zu zeichnen. Natürlich ist die geometrische Veranschaulichung meist sehr zweckmäßig und lehrreich.



Im täglichen Leben benutzen wir häufig ein Koordinatensystem, ohne uns dessen bewusst zu sein. Kaufen wir zum Beispiel eine Kinokarte, so lesen wir auf ihr die Nummer der Reihe und die des Platzes in dieser Reihe ab, etwa Reihe 5, Platz 4. Wie wir auf der Abbildung sehen, sind diese Zahlen eigentlich nichts anderes als die Koordinaten unseres Platzes.



Ein anderes Beispiel ist den Schachspielern vertraut. Bei Schachturnieren wird die Partie von dem Spieler aufgezeichnet, und Liebhaber des Schachspiels benutzen diese Aufzeichnungen, um die Partie nachzuspielen und zu analysieren. Worauf beruht nun diese Niederschrift?
Auf einem Koordinatensystem, das der Schachspieler in Gedanken auf dem Schachbrett eingeführt hat!

Notiert der Spieler den Zug Sg1-f3, so meint er, dass er seinen Springer von dem Feld mit den Koordinaten (g;1) auf das Feld mit den Koordinaten (f;3) zieht. Hier sind die Felder längs der Abszissenachse nicht mit Zahlen, sondern mit Buchstaben versehen.

Wir wollen nun wichtige Anwendungen des Koordinatensystems betrachten. Im Grunde genommen ist die Benutzung eines solchen Systems in zweifacher Hinsicht möglich. Erstens kann man mit seiner Hilfe verschiedene funktionale Abhängigkeiten, die in Mathematik, Naturwissenschaft und Technik sowie in anderen Wissenschaften auftreten, graphisch darstellen.
Zweitens lassen sich, wie wir schon zu Anfang unserer Betrachtungen erwähnten, mit seiner Hilfe gewisse allgemeine Verfahren zur Lösung der verschiedenartigsten geometrischen Probleme angeben. Wir werden dies nun an einigen Beispielen erläutern.

1.9 Das Zeichnen von Kurven

Jeder Leser wird sicher schon einmal in Mathematik oder Physik mit graphischen Darstellungen zu tun gehabt haben, und sicherlich kennt jeder die für eine gleichförmige Bewegung geltende Gleichung

$$s = vt$$

In dieser Gleichung bezeichnet s den Weg, den ein Körper in der Zeit t zurücklegt, und v die konstante Geschwindigkeit des Körpers. Die zu dieser Gleichung gehörende Kurve ist eine Gerade. Aus der Algebra führen wir folgendes Beispiel an:
Gegeben sei die Gleichung

$$y = x^2$$

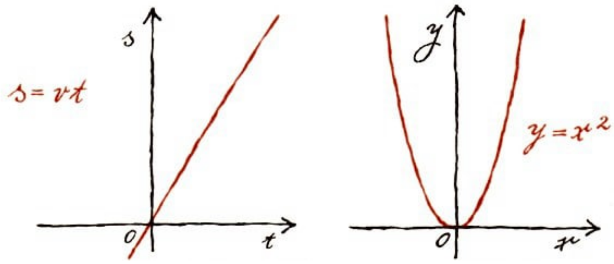
Wenn wir die zugehörige Kurve zeichnen wollen, geben wir zunächst verschiedene Werte von x vor und berechnen anhand der Gleichung die diesen x-Werten entsprechenden y-Werte:

x	...	-4	-3	-2	-1	0	1	2	3	4	...
y	...	16	9	4	1	0	1	4	9	16	...

Mit Hilfe dieser Wertetabelle können wir natürlich nur einzelne Punkte der Kurve zeichnen. Wenn wir aber der Abszisse x nicht nur ganzzahlige Werte erteilen, sondern auch

Werte zwischen diesen Zahlen, und dann die entsprechenden Werte der Ordinate y berechnen, erhalten wir immer dichter liegende Kurvenpunkte.
Verbinden wir die so konstruierten Punkte durch eine ununterbrochene Linie, so erhalten wir mit hinreichender Genauigkeit die Kurve zu unserer Gleichung. Die Kurve der Gleichung $y = x^2$ nennt man eine Parabel.

Wir sind uns sicher darüber einig, dass man sich eine Kurve manchmal leichter einprägen kann als eine Formel. Blicken wir auf die Zeichnung, so erfassen wir sofort, mit einem Blick, die Wechselbeziehung zwischen allen Werten der beiden Veränderlichen und sehen, welchen Einfluss die Änderung der einen Veränderlichen auf die andere hat. Es wäre oft viel schwieriger, diese Wechselbeziehung aus der Formel statt aus der Zeichnung abzulesen.



Diese beiden Beispiele zeigen, dass eine Gleichung zwischen zwei veränderlichen Größen eine Kurve in der Ebene definiert. Es gibt zwar auch Gleichungen, die keine Kurve definieren, aber diese sind selten, und wir wollen sie hier nicht betrachten.

1.10 Die Gleichung einer Kurve

Erst richtig bewusst wird uns der Wert des Koordinatensystems bei den folgenden Überlegungen. Wir sahen, dass wir eine Kurve zeichnen konnten, wenn uns deren Gleichung, welche zwei Veränderliche (etwa x und y) verknüpft, bekannt ist.
Daher kann man sagen, dass die Gleichung in ihrer algebraischen Gestalt sozusagen in "Kurzform" alle Eigenschaften der Kurve zum Ausdruck bringt, die wir auch an dem Kurvenbild selbst ablesen können. Das führt dazu, dass zum Studium einer Kurve sehr häufig nicht die Zeichnung, sondern die Gleichung der Kurve benutzt wird, und legt natürlich den folgenden interessanten Gedanken nahe:

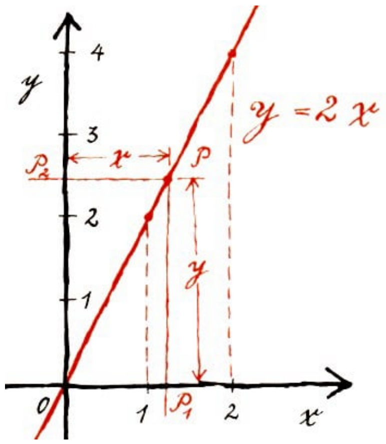
Wenn eine Gleichung die Eigenschaften der zu ihr gehörigen Kurve in sich vereint, ist es dann nicht möglich, auch das umgekehrte Problem zu lösen - nämlich anhand der Eigenschaften einer Kurve die entsprechende Gleichung aufzustellen?
Wir werden sehen, dass dies tatsächlich der Fall ist. Wir beginnen mit einem sehr einfachen Beispiel, um an ihm zu zeigen, wie die Gleichung einer Kurve gefunden werden kann. Dann gehen wir zu interessanteren und komplizierteren Fällen über und untersuchen, welche Schlussfolgerungen wir aus ihnen ziehen können.

Beispiel 1:
Man bestimme die Gleichung der Kurve, die im ersten und dritten Quadranten liegt und auf der jeder Punkt von der x-Achse doppelt so weit entfernt ist wie von der y-Achse.

Um die Gleichung dieser Kurve zu finden, zeichnen wir zunächst ein Koordinatensystem und einen beliebigen Punkt P der Kurve. Wir können annehmen, dass die Kurve durch Bewegung dieses Punktes P entsteht.
Dann formulieren wir die Eigenschaft, durch die die Kurve charakterisiert wird, d.h. die Eigenschaft, die jedem ihrer Punkte (und nur diesen) zukommt (also auch dem Punkt P während der ganzen Zeit seiner Bewegung), algebraisch, durch eine Formel.

Dazu fällen wir vom Punkt P die Lote auf die x-Achse und die y-Achse und bezeichnen ihre Fußpunkte mit P_1 bzw. P_2. Nun soll nach Voraussetzung die Strecke P_1P doppelt so lang sein wie die Strecke P_2P; wir schreiben also

$$\overline{P_1P} = 2\overline{P_2P}$$



Da $\overline{P_1P}$ die Ordinate y und $\overline{P_2P}$ die Abszisse x des Punktes P ist, erhalten wir die Gleichung $y = 2x$.

Wir können behaupten, dass dies die Gleichung der gegebenen Kurve ist, da man umgekehrt diese Kurve als graphische Darstellung der Gleichung erhält. Diese Kurve ist eine Gerade.
Gehen wir nämlich der Abszisse x beliebige Werte und berechnen wir die entsprechenden Werte der Ordinate y, so finden wir die Wertetabelle

x	...	-3	-2	-1	0	1	2	3	...
y	...	-6	-4	-2	0	2	4	6	...

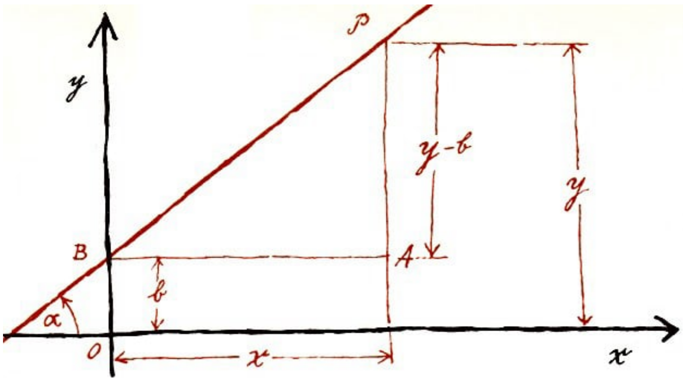
Da wir uns vorstellen können, diese Kurve sei durch die Bewegung des Punktes P entstanden, nennen wir den Punkt P den laufenden Punkt und seine Koordinaten x und y seine laufenden Koordinaten.

Während wir in Beispiel 1 die Gleichung einer Geraden durch den Koordinatenursprung herleiteten, wollen wir jetzt die Gleichung einer beliebig in der Ebene liegenden Geraden aufstellen. Dazu müssen wir eine beliebige Gerade vorgehen; wir bezeichnen den Punkt, in dem die Gerade die y-Achse schneidet, mit B und den Winkel, den sie mit der positiven Richtung der x-Achse einschließt, mit α. Das Problem kann dann folgendermaßen formuliert werden.

Beispiel 2:
Man bestimme die Gleichung der Geraden, die die y-Achse im Punkt $B(0; b)$ schneidet und mit der positiven Richtung der x-Achse den Winkel α bildet.
Wir betrachten wieder auf der Geraden den laufenden Punkt P mit den Koordinaten x und y. Die Bedingung, der der Punkt unterworfen ist, hat die Form

$$\overline{AP} : \overline{BA} = \tan\alpha$$

wie wir aus der Abbildung leicht ablesen können. Diese Beziehung ist stets erfüllt, solange der Punkt P auf der Geraden liegt. Dagegen gilt sie nicht mehr, wenn der Punkt P die Gerade verlässt.



Wir wollen jetzt die Größen, die in dieser Beziehung auftreten, durch die gegebenen Größen b und a, welche die Gerade charakterisieren, und durch die Koordinaten x und y von P ausdrücken.
Der Abbildung entnehmen wir, dass $\overline{AP} = y - b$ und $\overline{BA} = x$ ist. Die Beziehung lautet dann

$$\frac{y-b}{x} = \tan\alpha$$

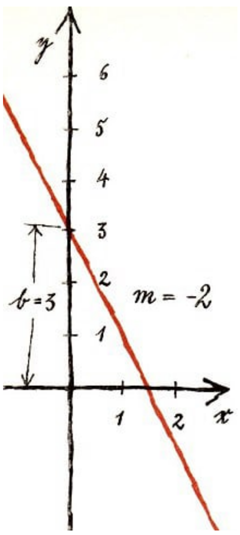
oder, nach y aufgelöst,

$$y = \tan\alpha \cdot x + b$$

Wenn wir nun für den Faktor $\tan\alpha$ den Buchstaben m einführen ergibt sich die gesuchte Geradengleichung in der Gestalt

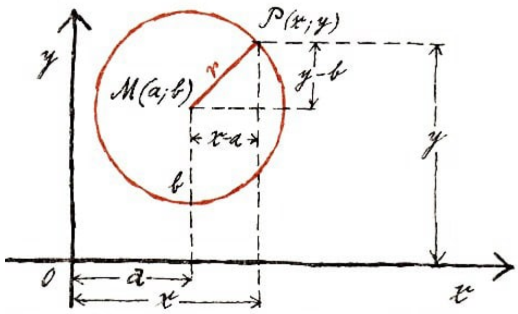
$$y = mx + b$$

Die Größe m heißt der Richtungskoeffizient der Geraden.



Aus der Herleitung der Gleichung ist ersichtlich, dass die Größen m und b beliebige, aber für eine bestimmte Gerade feste Zahlenwerte annehmen können. Liegt beispielsweise der Punkt B unterhalb der x-Achse, so ist b negativ. Bildet die Gerade mit der positiven Richtung der x-Achse einen stumpfen Winkel, so ist m negativ. Sowohl b als auch m können auch gleich Null sein.
Wir müssen noch darauf hinweisen, dass sich die Gleichung einer zur y-Achse parallelen Geraden nicht in der Form $y = mx+b$ schreiben lässt, da dann $m = \tan 90°$ ist und dieser Wert nicht existiert.
Nebenstehend ist die Gerade mit der Gleichung $y = -2x + 3$ gezeichnet.

Beispiel 3:
Man bestimme die Gleichung eines Kreises mit dem Mittelpunkt $M(a; b)$ und dem Radius r.
Wir müssen wieder die Eigenschaft, durch die der Kreis charakterisiert ist, algebraisch ausdrücken. Wie wir wissen, ist der Kreis die Menge aller Punkte, die von einem gegebenen Punkt, dem Mittelpunkt, einen gegebenen konstanten Abstand, Radius genannt, haben.

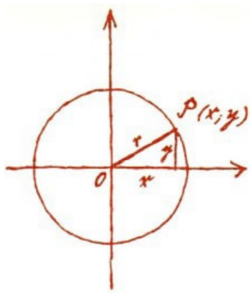


Der laufende Punkt P auf dem Kreis sei $P(x;y)$. Sein Abstand vom Mittelpunkt M des Kreises ist gleich dem Radius $MP = r$. Den Abstand MP können wir nach Formel (1) berechnen, wenn wir dort $x_1 = a$, $y_1 = b$, $x_2 = x$, $y_2 = y$ setzen. Wir finden

$$\overline{MP} = \sqrt{(x-a)^2 + (y-b)^2} \qquad \text{und damit} \qquad \sqrt{(x-a)^2 + (y-b)^2} = r$$

Wir quadrieren beide Seiten dieser Beziehung und gelangen so zu der Gleichung des Kreises:

$$(x-a)^2 + (y-b)^2 = r^2$$



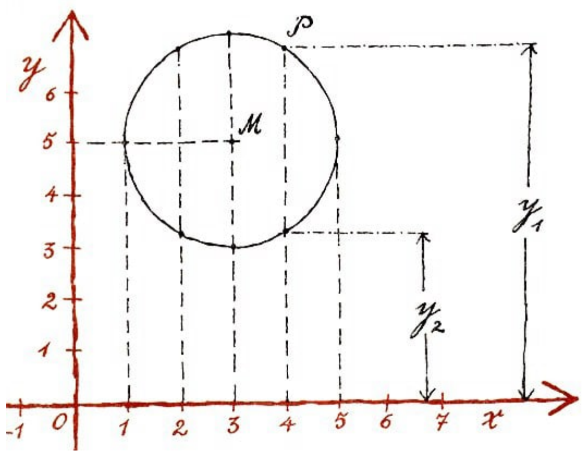
Liegt der Mittelpunkt des Kreises im Koordinatenursprung, so ist $a = 0$ und $b = 0$, und die Gleichung lautet

$$x^2 + y^2 = r^2$$

Nachstehend ist noch der Kreis mit der Gleichung

$$(x-3)^2 + (y-5)^2 = 4$$

graphisch dargestellt.



Nach diesen Beispielen mag sich der Leser vielleicht fragen:
Weshalb muss ich die Gleichungen der Geraden und des Kreises überhaupt herleiten, wenn ich diese Kurven auch ohne Gleichungen leicht zeichnen kann - die Gerade mit einem Lineal und den Kreis mit einem Zirkel?

Die Antwort auf diese Frage geben wir etwas später. Einstweilen erwähnen wir nur, dass das Ziel, das wir mit der Aufstellung dieser Gleichungen verfolgten, nicht darin besteht, eine Gerade oder einen Kreis zu konstruieren. Vielmehr befähigen uns diese Gleichungen, viele schwierige geometrische Aufgaben zu lösen, die mit diesen Kurven verknüpft sind.

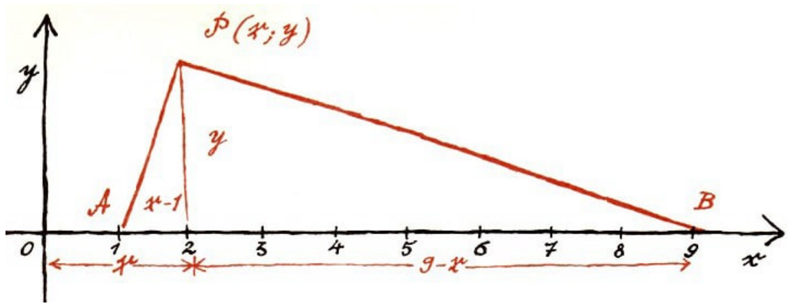
1.11 Die Lösung schwieriger geometrischer Probleme mit Hilfe eines Koordinatensystems

Durch die Herleitung der Gleichung einer Geraden und eines Kreises sind wir zu folgender wichtigen Schlussfolgerung gekommen: Wenn man die Gleichung einer Kurve aufstellen will, so muss man diese Kurve als Menge aller Punkte mit einer bestimmten Eigenschaft auffassen.
Die algebraische Schreibweise dieser Eigenschaft mit Hilfe der Koordinaten x und y des laufenden Punktes P führt uns auf die Kurvengleichung. Damit wissen wir also: Die Benutzung eines Koordinatensystems erlaubt uns erstens, die zu einer Gleichung gehörende Kurve graphisch darzustellen, wenn uns diese Gleichung bekannt ist, und zweitens, die Kurvengleichung anzugeben, wenn wir die Eigenschaft kennen, die die Kurve als Gesamtheit von Punkten besitzt.
Die zweite dieser Aussagen liefert uns gerade die allgemeinen Verfahren zur Lösung schwieriger geometrischer Aufgaben.
Wenn wir ein schwieriges Problem lösen oder einen Satz beweisen wollen, so müssen wir folgendermaßen vorgehen. Wir schreiben die Voraussetzungen der Aufgabe oder des Satzes in algebraischer Form, wie wir es in den vorhergehenden Beispielen taten. Damit erhalten wir die Gleichung der Kurve, von der in der Aufgabe die Rede ist. Manchmal können sich auch, je nach der Aufgabenstellung, mehrere Gleichungen ergeben. Jede dieser Gleichungen bringt dann alle Eigenschaften der durch sie bestimmten Kurve zum Ausdruck.

Damit haben wir die Lösung des geometrischen Problems auf die Behandlung von Gleichungen zurückgeführt, und dadurch wird, wie wir schon zu Beginn dieses Abschnitts sagten, die Lösung bedeutend vereinfacht.
Wir wollen nun zwei Probleme betrachten, die unsere Vermutung

Aufgabe:
Gegeben seien zwei Punkte mit den Koordinaten $A(1;0)$ und $B(9;0)$. Man untersuche, welche Kurve der Punkt beschreibt, der von Punkt B stets dreimal so weit entfernt ist; wie vom Punkt A.

Wollten wir versuchen, diese Aufgabe rein geometrisch, d.h. ohne Gleichungen zu lösen, so müssten wir lange nachdenken, da es schwierig ist, sofort zu sagen, auf welcher Kurve sich der Punkt bewegt. Benutzen wir jedoch die obigen Überlegungen, so erhalten wir sofort und fast mühelos die Lösung des Problems.



Wir wissen, dass die gesuchte Kurve dadurch charakterisiert ist, dass jeder Punkt P der Kurve von dem festen Punkt B dreimal so weit entfernt ist wie von dem festen Punkt A. Die Strecke PB ist also dreimal so lang wie die Strecke PA:

$$3\overline{PA} = \overline{PB} \tag{*}$$

Bezeichnen wir die Koordinaten von P wieder mit x und y und benutzen wir die Formel (1), so lassen sich die Abstände PA und PB durch die Koordinaten von A, B und P ausdrücken:

$$\overline{PA} = \sqrt{(x-1)^2 + y^2} \qquad , \qquad \overline{PA} = \sqrt{(x-9)^2 + y^2}$$

Damit ergibt sich aus (*)

$$3\sqrt{(x-1)^2 + y^2} = \sqrt{(x-9)^2 + y^2}$$

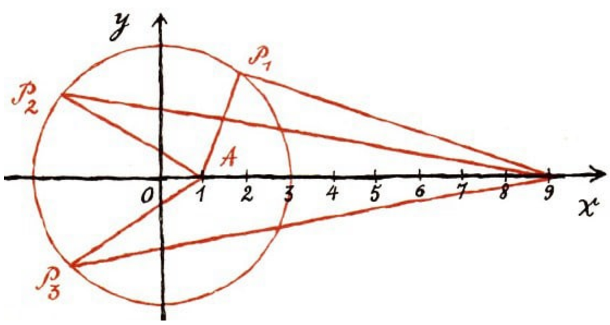
Zur Beseitigung der Wurzelzeichen quadrieren wir beide Seiten dieser Beziehung. Gleichzeitig rechnen wir die unter den Wurzelzeichen stehenden Klammern aus und erhalten

$$9(x^2 - 2x + 1 + y^2) = x^2 - 18x + 81 + y^2$$

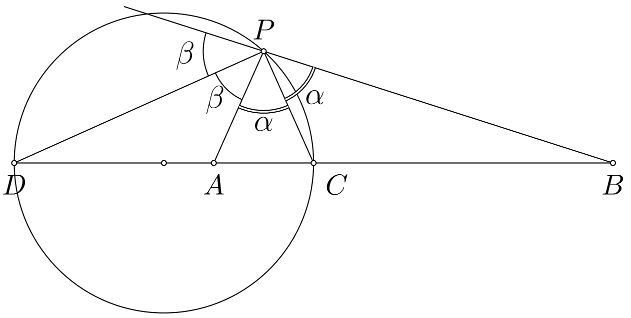
oder, vereinfacht,

$$x^2 + y^2 = 9$$

Wir haben damit die Gleichung des Kreises mit dem Mittelpunkt im Koordinatenursprung und dem Radius 3 erhalten. Das bedeutet, dass alle Punkte dieses Kreises, etwa P_1, P_2, P_3 den Bedingungen der Aufgabe genügen. Mit anderen Worten: Ein Punkt P, der bei seiner Bewegung stets dreimal so weit vom Punkt B entfernt ist wie vom Punkt A, muss sich auf diesem Kreis bewegen.



Wir sehen, wie leicht und fast "automatisch" wir die Aufgabe lösen konnten. Mit ähnlichen Verfahren lassen sich auch viele andere Probleme lösen.
Einem Leser, der diese Aufgabe gern ohne Koordinatensystem lösen möchte, wollen wir folgende kurze Erläuterung geben. Wie wir sehen, ist der Abstand zwischen den Punkten A und B gleich 8 Einheiten. Wir nehmen an, es existiere ein Punkt P, der den Bedingungen der Aufgabe genügt, für welchen also $\overline{PA} : \overline{PB} = 1 : 3$ ist.



Dann teilen wir die Strecke $\overline{AB}$ im Verhältnis 1 : 3. Das geschieht durch den Punkt C, der vom Punkt A genau 2 Einheiten entfernt ist; denn das Verhältnis $\overline{AC} : \overline{CB}$ ist gleich 2 : 6 = 1 : 3.
Nun verbinden wir den Punkt P mit A, B und C. Wir können behaupten, dass die Gerade PC die Winkelhalbierende von $\angle BPA$ im Dreieck ABP ist. Es gibt nämlich einen Satz, der besagt, dass die Halbierende eines Winkels im Dreieck die dem Winkel gegenüberliegende Seite in Teile zerlegt, die den anliegenden Seiten proportional sind.

Auch die Umkehrung dieses Satzes ist richtig:
Wenn eine Ecke eines Dreiecks mit demjenigen Punkt der gegenüberliegenden Seite verbunden wird, der diese Seite im Verhältnis der beiden anliegenden Seiten teilt, so ist diese Gerade die Winkelhalbierende. In unserem Fall ist

$$\overline{AP} : \overline{PB} = 1 : 3 \qquad , \qquad \overline{AC} : \overline{CB} = 1 : 3$$

Daraus folgt

$$\overline{AP} : \overline{PB} = \overline{AC} : \overline{CB}$$

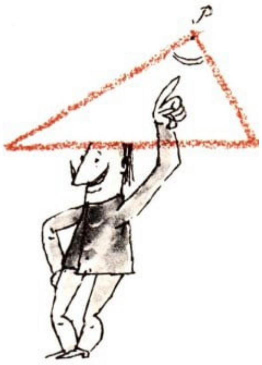
d.h., die Halbierende des Winkels BPA geht durch den Punkt C.
Ferner gilt der Satz, dass die Halbierende des Außenwinkels APQ des Dreiecks ABP die gegenüberliegende Seite, in unserem Fall also die Verlängerung der Seite BA, in Teile teilt (sogenannte äußere Teilung), die den anliegenden Seiten proportional sind,
Wenn in unserem Fall die Gerade PD den Außenwinkel halbiert, so liegt der Punkt D genau 4 Einheiten links vom Punkt A auf der Geraden durch A und B. Dann gilt

$$\overline{AD} : \overline{DB} = 4 : 12 = 1 : 3 \qquad , \qquad \overline{AP} : \overline{PB} = 1 : 3$$

also

$$\overline{AD} : \overline{DB} = \overline{AP} : \overline{PB}$$

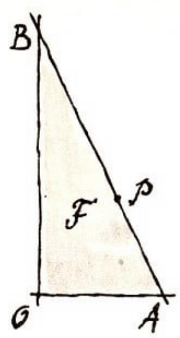
Wir können leicht feststellen, dass die beiden Winkelhalbierenden PC und PD einen rechten Winkel bilden. Es ist nämlich $2\alpha + 2\beta = 180°$, also $\alpha + \beta = 90°$. Das bedeutet, dass P auf einem Kreis (Thaleskreis) liegt, dessen Durchmesser die 6 Einheiten lange Strecke CD ist.



Man kann auch zeigen, dass umgekehrt jeder Punkt dieses Kreises die Bedingungen der Aufgabe erfüllt. Diese Konstruktion scheint nicht kompliziert zu sein, wenn man sie erst kennt. Aber zuerst einmal darauf zu kommen, ist bedeutend schwieriger!
Interessant ist vielleicht noch die Bemerkung, dass auch hierzu ein ganz allgemeiner bekannter Satz existiert.

Satz:
Die Menge aller Punkte P, deren Abstände von zwei gegebenen Punkten A und B im konstanten Verhältnis $m : n$ (mit $m \neq n$) stehen, ist ein Kreis.

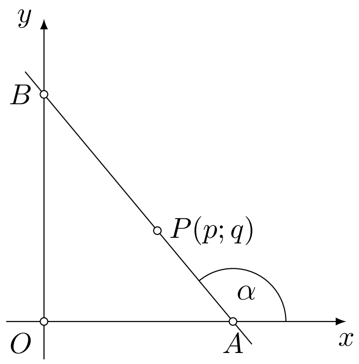
Dieser Kreis heißt der Kreis des Apollonius nach dem griechischen Mathematiker Apollonius von Pergae, der etwa bis 190 v.u.Z. lebte). Sein Durchmesser ist die Strecke, deren Endpunkte die Strecke AB innen und außen im Verhältnis $m : n$ teilen.
Ist $m = n$, so ist die Bestimmungskurve nicht ein Kreis, sondern die Gerade, die auf der Strecke AB senkrecht steht und durch deren Mittelpunkt geht.



Wir wollen nun noch eine weitere Aufgabe betrachten, deren rein geometrische Lösung außerordentlich schwierig ist, während wir mit Hilfe von Koordinaten mühelos zum Ziel kommen.

Aufgabe: Im Innern eines rechten Winkels sei ein Punkt P gegeben. Man ziehe durch P eine Gerade derart, dass das von ihr abgeschnittene rechtwinklige Dreieck OAB den Flächeninhalt F hat.

Wir wollen die Aufgabe gar nicht erst geometrisch lösen - wer Lust hat, möge es selbst versuchen - sondern sofort ein Koordinatensystem einführen.



Wir legen das Koordinatensystem so, dass sein Ursprung mit dem Scheitel O und die x- und die y-Achse mit den Schenkeln OA bzw. OB zusammenfallen. Die Lage des Punktes P kennzeichnen wir durch Koordinaten p und q, die wir als unbekannt annehmen.
Unsere Aufgabe besteht nun darin, durch den Punkt $P(p; q)$ eine Gerade zu legen, deren Winkel α mit der positiven Richtung der x-Achse so groß ist, dass der Flächeninhalt

des Dreiecks OAB gleich F ist, d.h., dass die Beziehung

$$\frac{\overline{OA} \cdot \overline{OB}}{2} = F \qquad \text{oder} \qquad \overline{OA} \cdot \overline{OB} = 2F \tag{1}$$

gilt.
Da wir die Aufgabe rechnerisch lösen wollen, müssen wir für die Gerade eine Gleichung finden, mit deren Hilfe wir sie dann zeichnen können.
Wir wissen schon, dass

$$y = mx + b \tag{2}$$

die Gleichung einer Geraden ist. In dieser Gleichung treten zwei konstante Größen m und b auf, die wir mit Hilfe der Aufgabenstellung bestimmen müssen. Wir finden die Größe b, wenn wir davon ausgehen, dass die Gerade durch den Punkt $P(p;q)$ verläuft. Die Größe m lässt sich aus der Beziehung (1) bestimmen, die den Flächeninhalt des Dreiecks OAB angibt.

Da die Gerade durch $P(p;q)$ geht, genügen die Koordinaten p und q dieses Punktes der Geradengleichung. Wenn wir also in (2) statt x die Größe p und statt y die Größe q einsetzen, ergibt sich

$$q = mp + b$$

Mit Hilfe dieser Beziehung können wir b durch m ausdrücken:

$$b = q - mp$$

Setzen wir nun diesen Wert in (2) ein, so finden wir

$$y = mx + q - mp \tag{3}$$

Nun müssen wir noch m bestimmen. Dazu gehen wir von der Bedingung aus, dass die Strecken OA und OB, die von der gesuchten Geraden auf den Koordinatenachsen abgeschnitten werden, der Beziehung (1) genügen müssen. Für die Strecke OB gilt

$$\overline{OB} = b = q - mp$$

Setzen wir nun, um $\overline{OA}$ zu finden, in die Gleichung (3) den Wert der Ordinate von A ein, also den Wert $y = 0$, so erhalten wir die Abszisse $x = \overline{OA}$ des Punktes A. Also ist

$$0 = mx + q - mp \qquad \text{und es folgt} \qquad \overline{OA} = x = \frac{mp - q}{m}$$

Damit haben wir die beiden Strecken OA und OB durch die eine unbekannte Größe m ausgedrückt. Nun setzen wir diese beiden Werte in (1) ein:

$$\frac{mp - q}{m}(q - mp) = 2F$$

Multiplizieren wir beide Seiten mit m, um den Nenner zu entfernen, und ziehen wir auf der linken Seite aus $q - mp$ den Faktor -1 heraus, so ergibt sich

$$-(mp - q)^2 = 2mF$$

Wir rechnen nun die Klammer aus, kehren alle Vorzeichen um und bringen das Glied $2mF$ von der rechten auf die linke Seite:

$$p^2m^2 - 2(pq - F)m + q^2 = 0$$

Dies ist eine quadratische Gleichung für die Unbekannte m. Alle anderen hier auftretenden Größen p, q und F sind gegeben. Lösen wir die Gleichung nach m auf, so erhalten wir

$$m = \frac{2(pq - F) \pm \sqrt{4(pq - F)^2 - 4p^2q^2}}{2p^2}$$

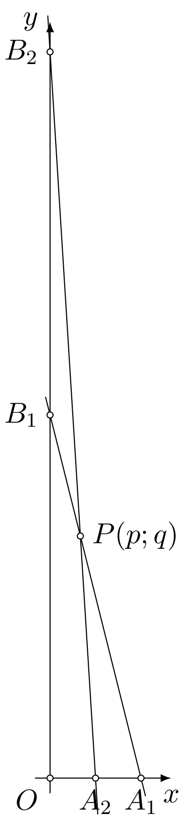
Hieraus folgt durch Vereinfachen des Radikanden schließlich

$$m = \frac{pq - F \pm \sqrt{F(F - 2pq)}}{p^2} \tag{4}$$

Damit ist die Aufgabe gelöst. Sind nämlich der Punkt P und der Flächeninhalt gegeben, sind also p, q und F bekannt, so berechnen wir mit Hilfe von (4) den Wert von m. Setzen wir diesen Wert in den Ausdruck

$$b = q - mp$$

ein, so können wir b bestimmen. Mit diesen beiden Werten von m und b gehen wir in die Gleichung (2) und erhalten so die gesuchte Geradengleichung. Dann ist es nicht mehr schwierig, diese Gerade zu konstruieren.
Nun haben wir aber zwei Werte für m erhalten. Welche Bewandtnis hat es damit? Zu dieser Überlegung wollen wir ein Zahlenbeispiel rechnen.



Aufgabe:
Gegeben sei der Punkt $P(1; 8)$. Man lege durch ihn eine Gerade, die mit den Koordinatenachsen ein Dreieck vom Flächeninhalt $F = 18$ Flächeneinheiten bildet. Wie lautet die Geradengleichung?
Wir benutzen die oben hergeleiteten Formeln und setzen $p = 1$, $q = 8$, $F = 18$ ein. Dann ist zunächst nach Formel (4)

$$m = -10 \pm 6 \qquad \text{oder} \qquad m_1 = -4; \qquad m_2 = -16$$

Zu jedem dieser m-Werte ergibt sich ein b-Wert:

$$b_1 = 8 - (-4) \cdot 1 = 12, \qquad b_2 = 8 - (-16) \cdot 1 = 24$$

Wir erhalten somit zwei Geraden, die den Bedingungen der Aufgabe genügen:

$$y = -4x + 12 \qquad , \qquad y = -16x + 24$$

Aus diesen Gleichungen ist abzulesen, dass die erste Gerade auf den Koordinatenachsen die Strecken $OA_1 = 3$ und $OB_1 = 12$, die zweite Gerade die Strecken $OA_2 = \frac{3}{2}$ und $OB_2 = 24$ abschneidet.

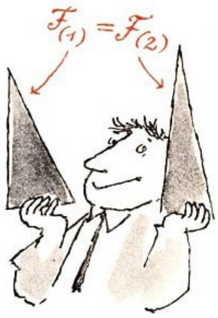
In beiden Fällen gilt für den Flächeninhalt der Dreiecke OA_1B_1 bzw. OA_2B_2:

$$F_{(1)} = \frac{\overline{OA_1} \cdot \overline{OB_1}}{2} = 18, \qquad F_{(1)} = \frac{\overline{OA_2} \cdot \overline{OB_2}}{2} = 18$$

Jedoch ist hiermit das Problem noch nicht bis zu Ende gelöst. Es ist sehr interessant, die erhaltenen Lösungen zu untersuchen. Wie wir der Formel (4) entnehmen, hat die Aufgabe zwei Lösungen, wenn

$$F - 2pq > 0$$

ist. (F ist eine positive Größe, daher kann der erste Faktor des Radikanden in dieser Ungleichung fortgelassen werden. Die Größen p und q sind ebenfalls positiv.) Dies war in unserem Zahlenbeispiel der Fall.



- Das Problem hat keine Lösung (d. h., es gibt keine Geraden, wie wir sie suchen), wenn

$$F - 2pq < 0$$

ist, und es hat nur eine Lösung (d.h., es gibt genau eine solche Gerade, wie wir sie suchen) im Falle $F - 2pq = 0$.

Uns interessiert nun, welche geometrische Bedeutung diese drei Bedingungen haben. In welchen Fällen gibt es zwei Lösungen, eine Lösung oder gar keine?

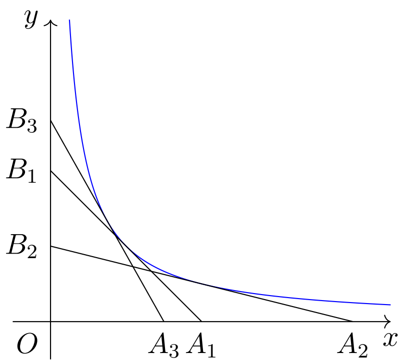
Wir wollen, ohne lange Beweise zu führen, hier nur das Wesentliche dieser Sache erklären.

Dem Leser ist vermutlich bekannt, dass die der Gleichung

$$y = \frac{a}{x} \qquad (a > 0)$$

entsprechende Kurve eine Hyperbel ist. Diese Hyperbel besitzt eine bemerkenswerte Eigenschaft. Legt man Tangenten an die Hyperbel in beliebigen Kurvenpunkten, so zeigt es sich, dass die Flächeninhalte F aller Dreiecke, die von den Tangenten mit den Koordinatenachsen gebildet werden, einander gleich sind, und zwar gleich $2a$:

$$F = F(\triangle OA_1B_1) = F(\triangle OA_2B_2) = F(\triangle OA_3B_3) = ... = 2a \tag{5}$$



Dabei ist a die in der Hyperbelgleichung auftretende Konstante.

Man kann auch die Umkehrung beweisen. Konstruieren wir Geraden, die mit den Koordinatenachsen Dreiecke $\triangle OA_1B_1, \triangle OA_2B_2, \triangle OA_3B_3$ gleichen Flächeninhaltes bilden, so sind diese Geraden Tangenten an die Hyperbel

$$y = \frac{a}{x}$$

Dabei ist die Größe a gleich dem halben Flächeninhalt eines Dreiecks, wie aus (5) folgt; denn ist $F = 2a$, so gilt

$$a = \frac{F}{2}$$

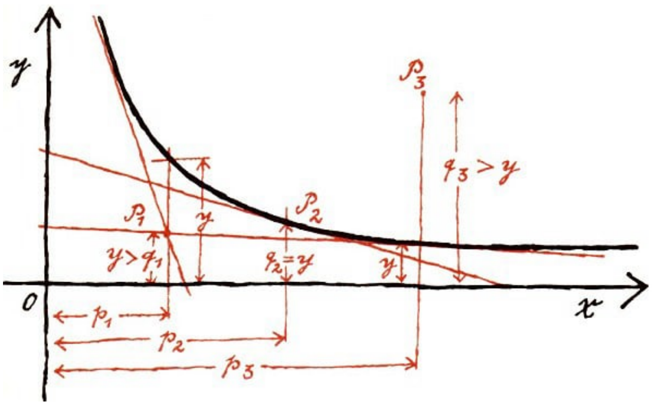
und die Hyperbelgleichung hat die Gestalt

$$y = \frac{F}{2x}$$

Hieraus erklärt sich auch der Sinn des Radikanden in der Formel (4), die den Wert m definiert.

In unserer Aufgabe sind zwei Bedingungen angegeben: Die gesuchte Gerade soll durch den Punkt $P(p;q)$ gehen und außerdem mit den Koordinatenachsen ein Dreieck OAB von gegebenem Flächeninhalt F bilden.
Wenn wir zeitweilig von der ersten Bedingung absehen, können wir sagen: Alle Geraden, die mit den Koordinatenachsen Dreiecke OAB von gegebenem Flächeninhalt F bilden, sind Tangenten an die Hyperbel
Nun ziehen wir die erste Bedingung wieder heran. Dann ist sofort klar: Liegt der Punkt $P(p;q)$, durch den die gesuchte Gerade gehen soll, zwischen den Koordinatenachsen und der Hyperbel (in der Abbildung der Punkt P_1), so verlaufen durch diesen Punkt zwei Tangenten an die Hyperbel

$$y = \frac{F}{2x}$$



Ist P ein Kurvenpunkt (in der Abbildung der Punkt P_2), so geht durch ihn nur eine Gerade, die die Hyperbel in eben diesem Punkt P_2 berührt. Liegt schließlich der Punkt P auf der anderen Seite der Hyperbel (in der Abbildung der Punkt P_3), so geht durch ihn überhaupt keine in Frage kommende Gerade, da man von ihm aus keine Tangenten an die Hyperbel legen kann.

Jetzt ist es nicht mehr schwer zu erklären, welche Bedeutung das Vorzeichen des Ausdrucks $F - pq$ hat, der unter dem Wurzelzeichen in (4) steht. Von seinem Vorzeichen hängt die Anzahl der Lösungen unseres Problems ab.
Liegt der Punkt P auf der Hyperbel, so erhalten wir, wenn wir seine Abszisse $x = p$ in die Hyperbelgleichung einsetzen, die Ordinate

$$y = \frac{F}{2p}$$

Andererseits muss die Ordinate q des Punktes $P_2(p;q)$, da er auf der Hyperbel liegt, gleich y sein; also gilt

$$y = \frac{F}{2p} = q, \qquad F = 2pq \qquad \text{oder} \qquad F - 2pq = 0$$

In diesem Fall erhalten wir eine einzige Lösung, d.h. eine Gerade.
Liegt der Punkt $P_1(p;q)$ zwischen den Koordinatenachsen und der Hyperbel, so ist seine Ordinate q kleiner als das aus der Hyperbelgleichung errechnete y, also

$$q < y = \frac{F}{2p}, \qquad F > 2pq \qquad \text{oder} \qquad F - 2pq > 0$$

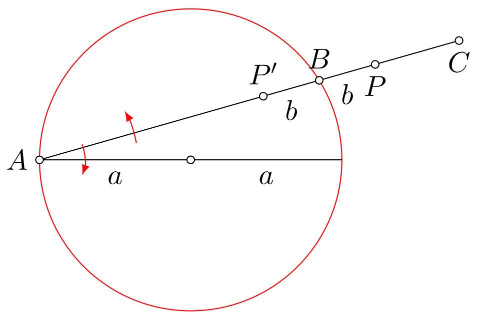
und es gibt zwei Lösungen des Problems.
Liegt schließlich der Punkt $P_3(p;q)$ auf der anderen Seite der Hyperbel, ist also q größer als das y aus der Hyperbelgleichung, so haben wir

$$q > y = \frac{F}{2p}, \qquad F < 2pq \qquad \text{oder} \qquad F - 2pq < 0$$

und das Problem hat keine Lösung.
Auf diese Weise haben wir die gleichen Bedingungen erhalten wie vorher, als wir den Radikanden der Gleichung (4) untersuchten.

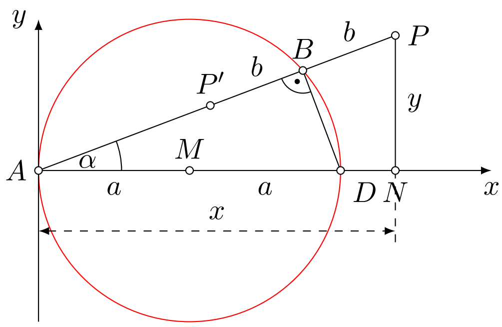
1.12 Die Pascalsche Schnecke

Wir wollen jetzt eine Kurve betrachten, die nach ihrem Entdecker, dem Vater des berühmten französischen Philosophen, Mathematikers und Physikers Blaise Pascal (1623-1662), Etienne Pascal, die Pascalsche Schnecke genannt wird. Wir werden zeigen, wie diese Kurve zur Lösung schwieriger geometrischer Probleme benutzt werden kann.



Die Pascalsche Schnecke konstruiert man folgendermaßen. Man nimmt einen Kreis vom Radius a, trägt auf diesem einen Punkt A ein und dreht um A einen Strahl AC.

Der zweite Schnittpunkt dieses Strahls mit dem Kreise sei der Punkt B. Von B aus steckt man nun auf dem Strahl nach beiden Seiten eine Strecke b ab. Die Menge der Endpunkte P und P' dieser Strecken ist dann die Pascalsche Schnecke.



Nun leiten wir die Gleichung der Pascalschen Schnecke her. Dazu bedienen wir uns wieder eines Koordinatensystems und ziehen einen Kreis mit dem Radius a um den Punkt $M(a;0)$. Den Punkt A, das Drehzentrum, legen wir in den Koordinatenursprung. Dann ziehen wir einen Strahl AC und stecken von seinem Schnittpunkt B aus nach beiden Seiten Strecken der Länge b ab. Die Endpunkte dieser Strecken seien P und P'.

Sie sind die laufenden Punkte der Pascalschen Schnecke. Die Koordinaten des Punktes P bezeichnen wir wieder mit x und y. Bei der Herleitung der Gleichung werden wir der Einfachheit halber nur den Punkt P betrachten. Es lässt sich nämlich zeigen, dass wir auch in diesem Fall die Gleichung der ganzen Pascalschen Schnecke erhalten.

Wir müssen nun, um die gesuchte Gleichung zu erhalten, die charakteristischen Eigenschaften der Pascalschen Schnecke durch die Koordinaten x und y des laufenden Punktes P algebraisch ausdrücken. Aus der Zeichnung ist ersichtlich, dass die Dreiecke ABD und ANP ähnlich sind, da die Winkel ABD und ANP rechte Winkel sind und der Winkel NAP beiden Dreiecken gemeinsam ist.

Die Dreiecke stimmen also in allen Winkeln überein. Daher gilt

$$\overline{AB} : \overline{AD} = \overline{AN} : \overline{AP}$$

Nun ist aber

$$\overline{AP} = \sqrt{x^2+y^2}, \qquad \overline{AN} = x, \qquad \overline{AD} = 2a$$
$$\overline{AB} = \overline{AP} - \overline{BP} = \overline{AP} - b = \sqrt{x^2+y^2} - b$$

Damit können wir

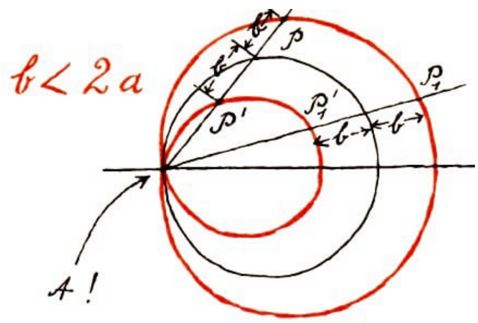
$$\frac{\sqrt{x^2+y^2}-b}{2a} = \frac{x}{\sqrt{x^2+y^2}}$$

schreiben. Zur Vereinfachung dieser Formel beseitigen wir zunächst die Nenner:

$$x^2+y^2-b\sqrt{x^2+y^2} = 2ax$$

Dann bringen wir das Glied mit der Wurzel nach rechts und das Glied $2ax$ nach links:

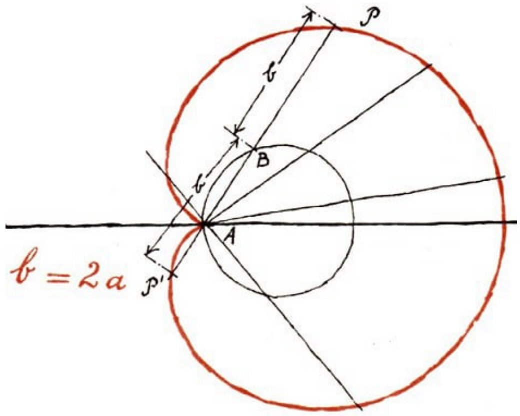
$$x^2+y^2-2ax = b\sqrt{x^2+y^2}$$

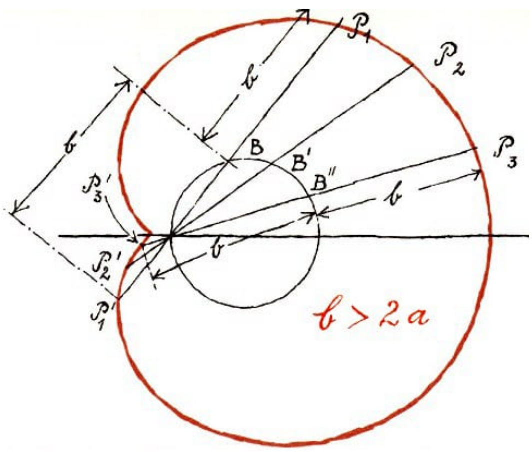


Wir quadrieren beide Seiten der Gleichung und finden

$$(x^2 + y^2 - 2ax)^2 = b^2(x^2 + y^2)$$

Dies ist die Gleichung der Pascalschen Schnecke. In den Abbildungen sind drei Pascalsche Schnecken, entsprechend den drei möglichen Fällen $b < 2a$, $b = 2a$ und $b > 2a$, gezeichnet. Zur Konstruktion wurden für a und b Zahlenwerte eingesetzt, die Gleichung nach y aufgelöst, dann eine Wertetabelle für x und y aufgestellt und danach jede Kurve punktweise gezeichnet.





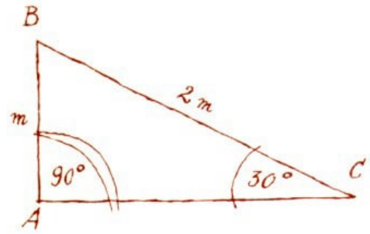
1.13 Dreiteilung eines Winkels mit Hilfe der Pascalschen Schnecke

Die Halbierung eines Winkels mit Hilfe der Winkelhalbierenden ist allen gut bekannt und leicht durchzuführen.
Dagegen ist die Teilung eines Winkels in drei gleiche Teile, kurz gesagt die Dreiteilung oder Trisektion eines Winkels, mit Zirkel und Lineal sehr kompliziert und nur in Spezialfällen möglich.

Schon die Geometer des Altertums versuchten, dieses Problem zu lösen, indem sie alle möglichen Konstruktionen mit Zirkel und Lineal probierten. Es wollte ihnen jedoch nicht gelingen, so ausgeklügelt und scharfsinnig diese Konstruktionen auch sein machten.
Schließlich kamen sie auf den Gedanken, dass solche Instrumente wie Zirkel und Lineal, mit deren Hilfe allein Kreise und Geraden konstruiert werden können, zur Lösung dieses Problems nicht genügen.

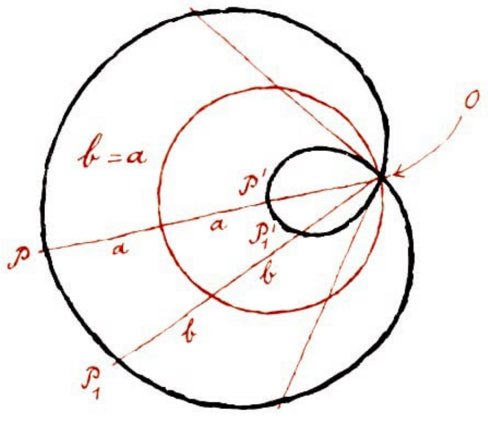
Daher schlugen sie vor, neue Instrumente, mit deren Hilfe kompliziertere Kurven gezeichnet werden können, zur Lösung des Problems zu benutzen. So entwickelte z.B. der griechische Geometer Nikomedes (um 180 v.u.Z.) den sogenannten Konchoidenzirkel, ein Instrument, mit dessen Hilfe eine Konchoide gezeichnet werden kann.
Damit kann ein beliebiger Winkel in drei gleiche Teile geteilt werden. Wir wollen jetzt zeigen, wie man das gleiche mit einer Pascalschen Schnecke erreichen kann, die entweder punktweise konstruiert oder mit einem Spezialinstrument gezeichnet wird.

Vorbereitend bemerken wir, dass einige bestimmte Winkel ganz leicht und ohne besondere Instrumente, nur mit Zirkel und Lineal, in drei gleiche Teile zerlegt werden können. Ein rechter Winkel wird z.B. folgendermaßen in drei Teile geteilt:



Man konstruiert ein rechtwinkliges Dreieck, dessen eine Kathete (etwa AB) halb so lang ist wie die Hypotenuse BC. Der dieser Kathete gegenüberliegende Winkel ist gleich $30°$, d.h. gleich einem Drittel des rechten Winkels.

Dass die Dreiteilung eines beliebigen Winkels allein mit Zirkel und Lineal nicht möglich ist, konnte erst im 19. Jahrhundert streng bewiesen werden. Dazu wurden ein Koordinatensystem und Kurvengleichungen benutzt.
Wir zeigen nun, wie die Dreiteilung mit Hilfe der Pascalschen Schnecke durchgeführt wird.



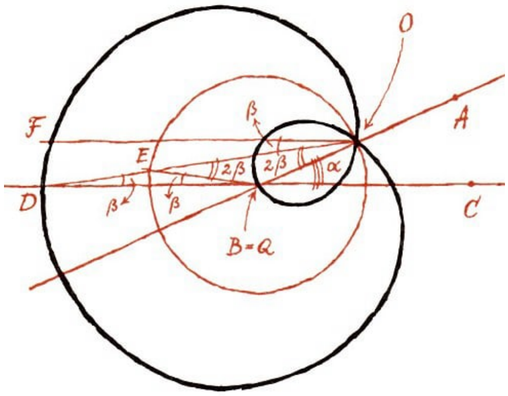
Gegeben sei ein beliebiger Winkel α und auf Transparentpapier eine Pascalsche Schnecke für $b = a$. Wir legen diese Schnecke so auf den Winkel α, dass der Punkt Q der Schnecke mit dem Scheitel B des Winkels zusammenfällt und die Strecke OQ (das ist der Radius des die Schnecke erzeugenden Kreises) mit dem Schenkel BA zur Deckung kommt.
Dann verlängern wir den Schenkel BC über den Scheitel B hinaus bis zum Schnittpunkt mit der Schnecke im Punkt D. Schließlich verbinden wir die Punkte O und D.

Wir können nun zeigen, dass der Winkel ODC der gesuchte ist, dass also $\angle ODC = \frac{\alpha}{3}$ gilt.
Dazu verbinden wir den Punkt E (den Schnittpunkt des erzeugenden Kreises mit der Strecke OD) mit dem Punkt B. Ferner ziehen wir durch den Punkt O die zu BC parallele Gerade OF.

Dann ist (Wechselwinkel an geschnittenen Parallelen)

$$\angle FOB = \angle ABC = \alpha$$

Den Winkel ODB bezeichnen wir mit β. Jetzt können wir aus der Abbildung folgendes ablesen:



1. $\angle EBD = \beta$, da auf Grund der Eigenschaft der Schnecke $\overline{BO} = \overline{BE} = \overline{ED} = a$ gilt und das Dreieck DEB gleichschenklig ist.
2. $\angle OEB = 2\beta$ als Außenwinkel des Dreiecks DEB.
3. $\angle BOE = 2\beta$ als Basiswinkel des gleichschenkligen Dreiecks OBE.
4. $\angle DOF = \angle ODC = \beta$ als Wechselwinkel an den parallelen Geraden OF und DC.
5. $\angle FOB = 3\beta = \angle OBC = \alpha$ als Wechselwinkel an den gleichen Parallelen.

Somit gilt $3\beta = \alpha$, also $\beta = \frac{\alpha}{3}$. Wir haben damit das Problem, den Winkel α in drei gleiche Teile zu teilen, mit Hilfe der Pascalschen Schnecke gelöst.

Zum Abschluss unserer Plauderei über das Koordinatensystem und seine Verwendung in der Geometrie wollen wir uns noch einmal alles ins Gedächtnis zurückrufen.
Wir sahen, dass es ein Koordinatensystem ermöglicht, alle möglichen funktionalen Abhängigkeiten graphisch darzustellen, damit also ihr Verständnis zu fördern und ihre Verwendung zu erleichtern. Anhand der gegebenen Eigenschaften einer Kurve ließ sich die Kurvengleichung bestimmen, und umgekehrt konnte man, wenn man die Gleichung kannte, die zugehörige Kurve zeichnen und deren Eigenschaften studieren.
Das Koordinatensystem ist also ein allgemeines Verfahren zur Lösung vieler geometrischer Probleme und zum Beweis von Sätzen, indem es sie auf Rechenprobleme zurückführt, die nach den üblichen Regeln der Algebra lösbar sind.

Das Koordinatensystem gibt uns z. B. in der Technik die Möglichkeit, spezielle Eigenschaften verschiedener Kurven auszunutzen.
Beispielsweise sei verlangt, die Festigkeit horizontal liegender Balken zu berechnen (etwa die der Balken einer Brücke. einer Zimmerdecke usw.). Beim Lösen dieser Aufgabe

benutzt der Ingenieur die Gleichung einer Kurve. Er geht dabei von der Belastung des Balkens aus und stellt dann nach den Gesetzen der Mechanik die Gleichung der Kurve auf, längs der sich der Balken durchbiegt.
Wenn er die Gleichung gefunden hat, kann er mühelos sämtliche notwendigen Abmessungen des Balkens errechnen.

Koordinatensysteme sind nicht nur für fast alle Zweige der Mathematik wichtig, sondern auch für alle Wissenschaften, in denen Mathematik benutzt wird. So stützen sich fast alle Berechnungen in Mechanik, Astronomie, Geodäsie und Geographie auf Koordinatensysteme.

Es sei nur noch erwähnt werden, dass die Längen- und Breitenkreise, die wir auf jeder Landkarte oder auf dem Globus sehen, ein System sogenannter krummliniger Koordinaten bilden. Die Lage eines Punktes auf der Erdkugel wird nämlich durch zwei Zahlen angegeben, durch Länge und Breite. d.h. durch zwei Koordinaten, und die Längen- und Breitenkreise selbst sind die Koordinatenlinien.
Es gibt also außer dem rechtwinkligen kartesischen Koordinatensystem noch andere Systeme, beispielsweise auch schiefwinklige.
Als sehr wichtiges ebenes Koordinatensystem wäre ferner das System der Polarkoordinaten zu nennen.

Auch im Raum kann man zur Lösung stereometrischer Aufgaben ein Koordinatensystem einführen. Räumliche Koordinatensysteme werden z. B. von den Kugelkoordinaten (oder sphärischen Koordinaten), von Zylinderkoordinaten usw. gebildet.
Ein räumliches kartesisches rechtwinkliges Koordinatensystem besteht aus den drei in einem Punkt aufeinander senkrecht stehenden Schnittgeraden von drei (Koordinaten-)Ebenen.

Fragen, die mit der Anwendung eines ebenen oder räumlichen Koordinatensystems zusammenhängen, werden in dem Teilgebiet der Mathematik untersucht, das den Namen "Analytische Geometrie"trägt. Hier haben wir uns auf die einfachsten Fragen beschränkt.
Für die Studenten der mathematisch-naturwissenschaftlichen Fakultät an Universitäten und Hochschulen werden darüber Vorlesungen gehalten., auch an Fachschulen und Oberschulen wird Analytische Geometrie gelehrt.

1.14 Funktionen in Natur und Technik

I. J. Wilenkin

Bei unseren bisherigen Betrachtungen stand das eigentlich Mathematische im Vordergrund, und die Beispiele aus dem täglichen Leben dienten mehr der Erläuterung der mathematischen Gedankengänge.
In dem vorliegenden Abschnitt folgen Beispiele, die zeigen, wie ein einziger mathematischer Begriff, nämlich der Begriff der Funktion, es ermöglicht, physikalische, chemische, technische und andere Vorgänge mathematisch zu erfassen und zu beschreiben. Dabei müssen zum Teil höhere mathematische Kenntnisse vorausgesetzt werden als in den

vorhergehenden Kapiteln.

Viele Leser werden daher manches von dem, was hier behandelt oder gestreift wird, erst später begreifen, wenn sie in der Schule trigonometrische Funktionen und die Exponentialfunktion kennengelernt haben.
Zum Verständnis der späteren Abschnitte brauchen sie den vorliegenden jedoch nicht, so dass sie die Lektüre dort, wo sie Schwierigkeiten haben, unterbrechen können. Aber selbst eine ganz oberflächliche Lektüre wird ahnen lassen, welches gewaltige Werkzeug der menschlichen Erkenntnis die Mathematik darstellt und wie vielfältig sie anzuwenden ist.

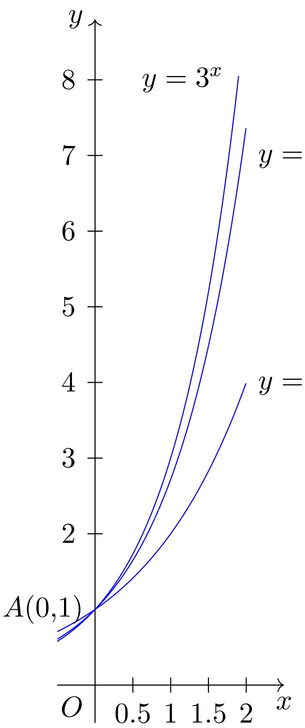
Einer der wichtigsten Begriffe in der Mathematik und ihren Anwendungen ist der Funktionsbegriff. Überall, wo es Größen gibt, die so miteinander verknüpft sind, dass sich eine von den Größen in Abhängigkeit von anderen ändert, haben wir es mit Funktionen zu tun.
Zum Beispiel ist der Flächeninhalt eines Quadrats eine Funktion der Seitenlänge; der von einem Körper im freien Fall zurückgelegte Weg ist eine Funktion der Zeit, die seit dem Beginn des Fallens verstrichen ist; die Stromstärke ist eine Funktion der Spannung und des im Stromkreis vorhandenen Widerstandes usw.
Meistens lässt sich die funktionale Abhängigkeit durch eine Formel ausdrücken. Man darf aber nicht denken, dass dies immer so ist:
Zum Beispiel ist die Lufttemperatur eine Funktion der Zeit, da sich die Temperatur im Laufe der Zeit ändert. aber es existiert keine Formel dafür.

Für viele Probleme der Physik genügen lineare und quadratische Funktionen. Beispielsweise wird die Geschwindigkeit bei einer gleichförmig beschleunigten Bewegung durch die lineare Funktion $v = at$ ausgedrückt. Der zurückgelegte Weg wird durch die quadratische Funktion $s = \frac{at^2}{2}$ beschrieben. Es gibt jedoch Fälle, in denen andere Funktionen notwendig sind.
Besonders oft aber begegnet man der Exponentialfunktion und den trigonometrischen Funktionen. Mit einigen dieser Fälle wollen wir uns jetzt beschäftigen.

1.15 Die Zahl e. Die natürlichen Logarithmen

Bei der Beschreibung physikalischer Gesetze, die mit der Exponentialfunktion verknüpft sind, benutzt man eine besondere Zahl, die man mit e bezeichnet. Sie lässt sich folgendermaßen bestimmen:

Wir stellen die Funktionen $y = a^x$ für verschiedene Werte der Basis a $(a > 1)$ graphisch dar. Je größer wir a wählen, desto steiler sind die Kurven.
Diese Kurven schneiden die y-Achse im Punkt $A(0,1)$ unter verschiedenen Winkeln. Der Winkel zwischen der positiven Richtung der y-Achse und der Kurve $y = 2^x$ ist annähernd gleich $55°15'$; für die Kurve $y = 3^x$ ist dieser Winkel etwa gleich $42°20'$.



Daher muss es zwischen 2 und 3 eine Zahl e geben, für die die Kurve $y = e^x$ die y-Achse genau unter dem Winkel $45°$ schneidet.

Genauere Rechnungen zeigen, dass diese Zahl e gleich 2,71828... ist. Die Logarithmen zur Basis e heißen natürliche Logarithmen; sie werden mit $\ln$ bezeichnet.

Kennen wir den dekadischen Logarithmus $\lg x$ einer Zahl x, so lässt sich ihr natürlicher Logarithmus $\ln x$ mit Hilfe der Formel

$$\ln x = \frac{\lg x}{M}$$

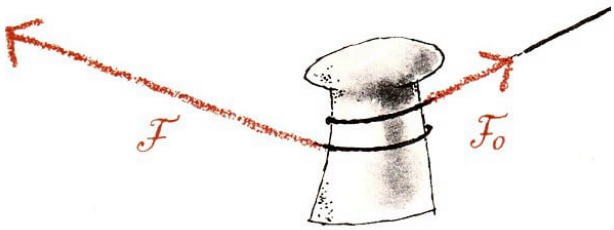
bestimmen. Die Zahl $M = 0{,}43429... = \lg e$ wird der Modul des dekadischen Logarithmensystems genannt.

1.16 Ein einzelner Mensch kann ein großes Schiff anhalten

Legt ein Schiff am Ufer an, so wird vom Schiff zum Kai ein Seil herübergeworfen. Hier wird es von einem Mann einige Male um einen Pfahl gewickelt, und das Schiff hält an. Wie kommt es nun, dass ein einziger Mensch das Schiff anhalten kann?
Das liegt an der Reibung. Wenn das Seil einmal um den Pfahl gewickelt wird, kann man mit; der Kraft F_0, die der Mensch aufbringt, durch die Reibung des Seils an dem Pfahl einer a-mal so großen Kraft F das Gleichgewicht halten. Das Verhältnis $F : F_0 = a$ hängt von dem Material des Seile und des Pfahls ab.
Ist z.B. das Seil aus Hanf und der Pfahl aus Eisen, so ist $a = 3{,}5$. Mit einer Kraft von 100 kp (Kilopond) kann man also, wenn man die Reibungskraft ausnutzt, eine Kraft von 350 kp im Gleichgewicht halten. Jede neue Wicklung des Seils um den Pfahl vergrößert das Kräfteverhältnis um den Faktor a.

Wenn wir also das Seil zweimal herumgewickelt haben ist das Kräfteverhältnis gleich a^2, bei dreimaligem Herumwickeln ist $F : F_0 = a^3$. Ist die Anzahl der Wicklungen x (x braucht keine ganze Zahl zu sein), so gilt $F : F_0 = ax$. Wickeln wir also ein Hanfseil zweimal um einen Eisenpfahl, so kann mit einer Kraft von 100 kp eine Kraft von etwa

1,2 t, bei dreimaligem Wickeln eine Kraft von etwa 4,2 t im Gleichgewicht gehalten werden.



1.17 Seile gleicher Zugfestigkeit

Ein Seil, an dem etwa in einem Schacht eine Last hängt, ist unzweckmäßig, wenn es über die ganze Länge von gleicher Dicke ist. Es hat nämlich außer der an ihm hängenden Last noch sein Eigengewicht zu tragen, denn an jedem Querschnitt des Seils hängen die Last sowie der unterhalb dieses Querschnitts befindliche Teil des Seils. Deshalb ist es besser, ein Seil zu nehmen, das so beschaffen ist, dass an jedem seiner Querschnitte die gleiche Kraft je Quadratzentimeter angreift. Berechnungen ergeben, dass sich dabei der Flächeninhalt des Seilquerschnitts nach dem folgenden Gesetz mit der Länge ändern muss:

$$A = A_0 e^{\frac{\gamma A_0 x}{G}}$$

In dieser Formel ist A_0 der Flächeninhalt des Querschnitts am unteren Seilende, A der Flächeninhalt des Querschnitts im Abstand x vom unteren Ende, γ die Wichte des Seilmaterials und G das Gewicht der zulässigen Last. Bei komplizierten Exponenten von e wählt man zweckmäßiger die Schreibweise $exp(...)$ also

$$A = A_0 \exp\left(\frac{\gamma A_0 x}{G}\right)$$

Ein Längsschnitt durch dieses Seil hat die in der Abbildung gezeigte Form. Solche Seile heißen Seile gleicher Zugfestigkeit. Sie haben ein kleineres Volumen als Seile mit konstantem Querschnitt, die für die gleiche Belastung vorgesehen sind und beanspruchen deshalb weniger Material.

1.18 Radioaktiver Zerfall

Wenn ein radioaktiver Stoff zerfällt, verringert sich seine Masse und nach einer gewissen Zeit ist nur noch die Hälfte der ursprünglichen Masse vorhanden. Dieses Zeitintervall t_0 heißt die Halbwertzeit des radioaktiven Stoffes. Vergeht ein weiterer Zeitabschnitt t_0 so ist von der übriggebliebenen Hälfte wieder die Hälfte zerfallen, so dass nur noch ein Viertel der ursprünglichen Masse vorhanden ist.

Die nach der Zeit t noch vorhandene Masse m des radioaktiven Stoffes ist gleich

$$m = m_0 \left(\frac{1}{2}\right)^{\frac{t}{t_0}}$$

wobei m_0 die ursprüngliche Masse des Stoffes ist. Je größer die Halbwertzeit ist, desto langsamer zerfällt der radioaktive Stoff.

Uran-238 hat z.B. eine Halbwertzeit von 4,5 Milliarden Jahren. Das bedeutet, dass seit Entstehung der Erde noch nicht einmal die Hälfte aller ursprünglich vorhandenen Uranvorkommen zerfallen ist. Dagegen hat Radium eine Halbwertzeit von etwa 1590 Jahren. Hätte also vor einer Million Jahren die ganze Erde aus Radium bestanden, so wäre heute von dem Radium nicht ein Atom mehr übrig.

Radium kommt nur deshalb noch vor, weil beim Zerfall des Urans stets neue Radiumatome entstehen.

1.19 Ein- und Ausschalten von Gleichstrom

Betätigen wir einen Lichtschalter, so flammt im gleichen Augenblick das elektrische Licht auf. Wir sind daran so gewöhnt, dass wir gar nicht darüber nachdenken, ob der Strom sofort seine volle Stärke erreicht. Wenn wir jedoch den in der untenstehenden Abbildung dargestellten elektrischen Stromkreis betrachten und den Strom einschalten, so sehen wir, dass die Stromstärke wegen der in dem Stromkreis vorhandenen Spule langsam wächst, da ein Selbstinduktionsstrom entsteht, der der Richtung des eingeschalteten Stroms entgegengesetzt ist.

Berechnungen zeigen, dass die Stromstärke I nach dem folgenden Gesetz von der Zeit t abhängt:

$$I = \frac{U_0}{R}\left(1 - e^{-\frac{R}{L}t}\right)$$

Hierbei ist R der Widerstand des Stromkreises, L der Koeffizient der Selbstinduktion, die sogenannte Induktivität der Spule und U_0 die Spannung.

Wir wollen nun untersuchen, wie sich die Stromstärke mit der Zeit ändert. Den Ausdruck $e^{-\frac{R}{L}t}$ kann man auch in der Form $\left(\frac{1}{e}\right)^{\frac{R}{L}t}$ schreiben. Nun ist $\frac{1}{e}$ kleiner als 1, und bekanntlich strebt die Exponentialfunktion mit wachsendem Exponenten gegen Null, sobald ihre Basis kleiner als 1 ist.

Daher wird der Subtrahend $e^{-\frac{R}{L}t}$ in der Formel für die Stromstärke mit wachsendem t immer kleiner und nähert sich dem Wert Null, so dass nach einer gewissen Zeit die Stromstärke fast gleich $\frac{U_0}{R}$ ist.

Dieser Wert ist uns schon vom Ohmschen Gesetz her bekannt. Die Differenz zwischen der tatsächlichen Stromstärke und dem Wert $\frac{U_0}{R}$ wird so klein, dass Sie mit keinem Gerät nachweisbar ist.

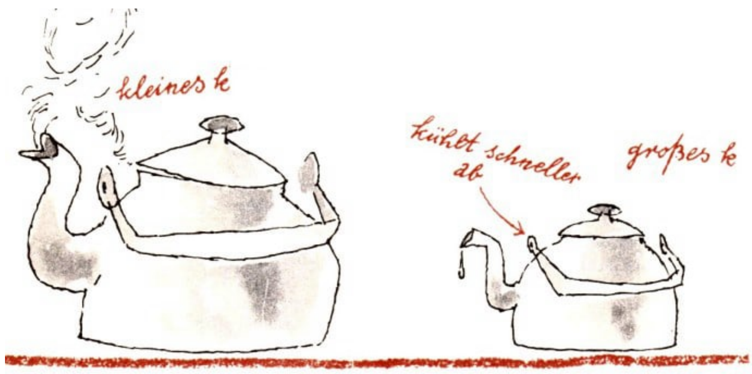
Dieselbe Erscheinung tritt auch beim Einschalten eines Stromkreises ohne Selbstinduktionsspule auf. Allerdings ist in diesem Fall die Induktivität L so klein, dass $e^{-\frac{R}{L}t}$ sich sehr rasch der Null nähert.
Das Ausschalten des Stroms ist ein dem Einschalten ähnlicher Prozess. Es wird durch die noch einfachere Formel

$$I = \frac{U_0}{R} e^{-\frac{R}{L}t}$$

beschrieben. Wir erwähnten schon, dass die Funktion $e^{-\frac{R}{L}t}$ mit wachsendem t abnehmend (d.h. von positiven Werten her) gegen Null strebt. Daher fließt nach einer gewissen Zeit nach dem Ausschalten kein Strom mehr im Stromkreis.

1.20 Abkühlung eines Teekessels

Wir alle haben schon bemerkt, dass ein Teekessel, wenn er vom Feuer genommen wird, zuerst schnell und dann immer langsamer abkühlt. Das liegt daran, dass die Abkühlungsgeschwindigkeit proportional der Differenz zwischen der Temperatur des Teekessels und der Temperatur der ihn umgebenden Luft ist.



Je kleiner also diese Differenz ist, desto langsamer kühlt der Teekessel ab. Ist die Temperatur des Teekessels am Anfang gleich T_0 und die der Luft gleich T_1, so ist die Temperatur T des Teekessels nach der Zeit t (in Sekunden) durch die Beziehung

$$T = (T_0 - T_1)e^{-kt} + T_1$$

gegeben. k ist eine Zahl, die von der Form und dem Material des Kessels sowie von der Wassermenge im Kessel abhängt.

1.21 Warum fällt ein Fallschirmspringer mit gleichbleibender Geschwindigkeit?

Wenn ein Körper im luftleeren Raum (Vakuum) fällt, nimmt seine Geschwindigkeit ständig zu. Beim Fall eines Körpers in Luft vergrößert sich zwar die Geschwindigkeit ebenfalls ständig, aber sie kann eine bestimmte Größe nicht überschreiten.

Wir wollen nun untersuchen, welchen Gesetzen ein Fallschirmspringer bei einem Sprung unterworfen ist. Nehmen wir an, dass der Luftwiderstand F proportional der Fallgeschwindigkeit v des Springers ist, also $F = kv$ gilt, so ist die Fallgeschwindigkeit nach der Zeit t (in Sekunden)

$$v = \frac{mg}{k}\left(1 - e^{-\frac{kt}{m}}\right)$$

m ist die Masse des Fallschirmspringers. Man beachte die Ähnlichkeit dieser Formel mit der Formel für die Stromstärke! Nach einer gewissen Zeit ist $e^{-\frac{kt}{m}}$ eine sehr kleine Zahl, und deshalb ist dann die Fallgeschwindigkeit fast gleich $\frac{mg}{k}$. Das bedeutet, dass der Fall fast gleichförmig ist. Der Proportionalitätsfaktor k hängt von den Abmessungen des Fallschirms ab. Die obige Formel ist nicht nur für einen Fallschirmspringer geeignet, sondern kann auch dazu benutzt werden, um festzustellen, wie schnell Regentropfen, Schneeflocken usw. fallen. Aus ihr ist abzulesen: Je kleiner der Quotient $\frac{mg}{k}$ ist, desto langsamer fällt der Körper.
Das erklärt, weshalb eine Schneeflocke langsamer fällt als ein Stein. Sie hat eine kleinere Masse, und der Flächeninhalt ihrer Oberfläche ist sehr groß. Die Luft kann ihr also einen beträchtlichen Widerstand entgegensetzen.

1.22 Wie misst man die Höhe mit Hilfe eines Barometers ?

Da wir wissen, dass der Luftdruck mit wachsender Höhe abnimmt, können wir diese Erscheinung zur Höhenmessung mit Hilfe eines Barometers benutzen. Für die Höhendifferenz zwischen zwei Punkten gilt bei konstanter Lufttemperatur die Formel

$$h_2 - h_1 = \frac{p_0(t + 273{,}1)}{\gamma_0} \ln \frac{p_1}{p_2}$$

Dabei ist p_1 bzw. p_2 der Luftdruck in der Höhe h_1 bzw. h_2; p_0 ist der Luftdruck in Höhe des Meeresspiegels. γ_0 bezeichnet das Gewicht eines Kubikmeters Luft bei der Temperatur von 0 °C und dem Druck p_0 und schließlich ist t die Lufttemperatur, gemessen in °C.

Diese Formel gilt nur für nicht allzu große Höhen. Untersuchungen, die in der UdSSR anlässlich des Internationalen Geophysikalischen Jahres mit Hilfe von Raketen durchgeführt wurden, zeigten, dass der Luftdruck in großen Höhen anderen Gesetzen unterworfen ist.
Im allgemeinen hat jede physikalische Formel nur einen beschränkten Anwendungsbereich. Sie ist nur unter bestimmten Bedingungen richtig, und unter anderen Bedingungen versagt sie. Das liegt daran, dass bei der Herleitung einer physikalischen Formel gewisse Voraussetzungen gemacht werden. die nur näherungsweise richtig sind. Durch

sie wird das Problem nämlich vereinfacht. Wenn diese Voraussetzungen nicht mehr erfüllt sind, verliert die Formel ihre Gültigkeit.

1.23 Wieviel Treibstoff muss eine Rakete aufnehmen?

Viele schwierige mathematische Probleme sind in der Theorie der interplanetarischen Flüge zu lösen. Eines von ihnen ist das Problem, die Masse m_T des Treibstoffs zu bestimmen, welche not? wendig ist, um einer Rakete eine bestimmte Geschwindigkeit v zu erteilen, mit der sie die Erdatmosphäre verlassen kann. Diese Masse m_T hängt von der Masse m der Rakete (ohne Treibstoff) und von der Geschwindigkeit v_0 ab, mit der die Verbrennungsprodukte aus dem Triebwerk der Rakete ausgestoßen werden.

Diese Geschwindigkeit nennt man auch die Abgasgeschwindigkeit. Werden der Luftwiderstand und die Erdanziehung vernachlässigt, so kann man die Treibstoffmenge mit Hilfe der Formel

$$m_T = m\left(e^{\frac{v}{v_0}} - 1\right)$$

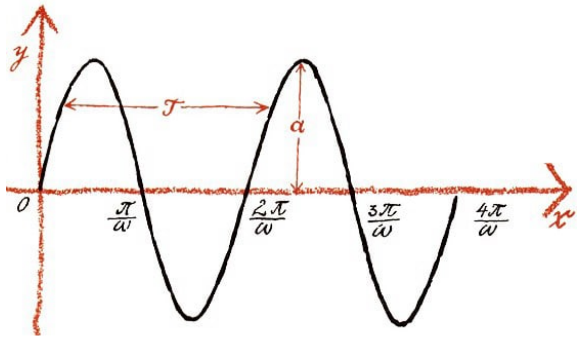
bestimmen. Diese Formel wird nach dem russischen Gelehrten Konstantin Eduardowitsch Ziolkowski (1857-1935) benannt.

Man braucht beispielsweise bei einer Abgasgeschwindigkeit von 2 km/s etwa 80 t Treibstoff, um einer Rakete mit einer Masse von 1,5 t eine Geschwindigkeit von 8 km/s zu erteilen. Hätte man eine Abgasgeschwindigkeit von 4 km/s, so brauchte man nur etwa 10 t Treibstoff. Je größer die Geschwindigkeit v_0 ist, desto kleiner ist $e^{\frac{v}{v_0}}$ und desto weniger Treibstoff wird benötigt.

1.24 Harmonische Schwingungen

Bis jetzt haben wir einige Beispiele aus Physik und Technik betrachtet, in denen die Exponentialfunktion auftritt. Nun wollen wir zu Beispielen übergeben, in denen wir trigonometrischen Funktionen begegnen. Wir beginnen mit den harmonischen Schwingungen.
Dazu betrachten wir eine an einer Schraubenfeder hängende Masse. Ziehen wir die Masse nach unten, so wird die Feder gedehnt, und je weiter die Masse nach unten gezogen wird, desto größer ist die Spannung der Feder. Lassen wir dann die Masse los, so bewegt sie sich auf Grund der Federkraft nach oben. Die Masse bleibt aber wegen ihrer Trägheit nicht in der Ruhelage (Gleichgewichtslage) stehen.
Sie bewegt sich so weit nach oben über die Ruhelage hinaus, wie sie vorher nach unten gezogen wurde. In Wirklichkeit bewegt sich die Masse wegen des Luftwiderstandes und der ungenügenden Elastizität der Feder nicht ganz so weit nach oben, aber das wollen wir jetzt nicht berücksichtigen. Geht die Masse nach oben, so wird die Feder gestaucht.

Dadurch wird die Geschwindigkeit, mit der sich die Masse bewegt, verringert und schließlich gleich Null.



In diesem Moment kehrt die Masse ihre Bewegungsrichtung um. Sie führt also Schwingungen aus, deren graphische Darstellung die in der Abbildung gezeigte Form hat. Diese Kurve heißt Sinuskurve. Die Auslenkungen der Masse aus der Ruhelage werden durch die Formel

$$s = \frac{v_0}{\omega} \sin \omega t$$

beschrieben. Dabei ist v_0 die Geschwindigkeit, mit welcher wir die Masse nach unten gezogen haben, und $\omega = \sqrt{\frac{k}{m}}$, wenn m die Masse und k die Federkonstante ist. Unter der Federkonstanten verstehen wir die Kraft, die notwendig ist, um die Feder um 1 Längeneinheit, also z.B. um 1 cm zu dehnen. Schwingungen, die dem Gesetz

$$s = a \sin \omega t \tag{1}$$

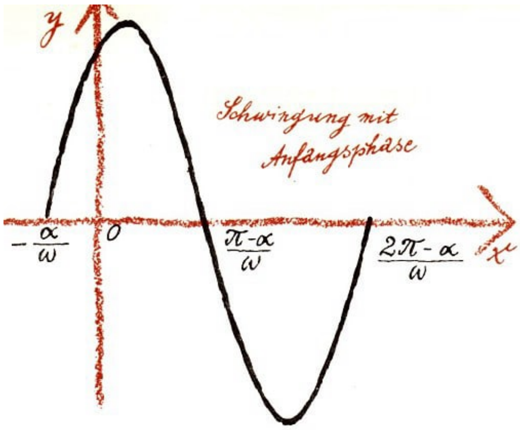
gehorchen, heißen Sinusschwingungen oder harmonische Schwingungen.

Wir erhalten eine Vorstellung von diesen Schwingungen, wenn wir einen Punkt, der sich gleichförmig auf einer Kreisbahn bewegt, beobachten, und zwar so, dass wir uns in der Ebene befinden, in der die Kreisbahn liegt.
Wir haben dann den Eindruck, der Punkt bewege sich nicht auf einer Kreisbahn, sondern laufe auf einer Geraden hin und her. In der Astronomie tritt diese Erscheinung bei der Beobachtung der Jupitermonde auf, wenn sich die Erde in der Bahnebene dieser Monde befindet.
Die Größe a, die Amplitude der Sinusschwingung, gibt die größten Ausschläge dieser Schwingungen an. Die Größe ω, die Kreisfrequenz der Schwingung, zeigt an, wieviel Schwingungen in 2π Sekunden ausgeführt werden.
Jeweils nach Ablauf der Zeit $\frac{2\pi}{\omega}$ (in Sekunden) kehrt die Masse in die Ausgangslage zurück. Daher ist die Periode einer Schwingung, die sogenannte Schwingungsdauer T, gleich $\frac{2\pi}{\omega}$.

Wenn wir die Masse zuerst um s_0 (in cm) aus ihrer Ruhelage entfernen und ihr dann die Geschwindigkeit v_0 erteilen, so schwingt sie nach dem komplizierteren Gesetz

$$s = a \sin(\omega t + \alpha) \tag{2}$$

Die Amplitude dieser Schwingung ist gleich $\sqrt{s_0^2 + \frac{v_0^2}{\omega^2}}$, und die Größe α genügt der Bedingung $\tan\alpha = \frac{s_0\omega}{v_0}$. Diese Schwingung unterscheidet sich von der Schwingung



$$s = a\sin\omega t$$

um den Summanden α im Argument des Sinus. Die Kurven dieser beiden Schwingungen sind hier angegeben. Die Kurve der Schwingung (2) ergibt sich aus der Kurve der Schwingung (I) durch Verschiebung um $\frac{\alpha}{\omega}$ nach links. Die Zahl α wird Anfangsphase oder Phasenverschiebung genannt.

1.25 Pendelschwingungen

Die Schwingungen eines Pendels folgen näherungsweise ebenfalls einem Sinusgesetz. Sind die Schwingungen klein, so ist der Pendelausschlag φ annähernd gleich

$$\sin\varphi_0 t\sqrt{\frac{g}{l}}$$

Dabei ist l die Länge des Pendels, g die Erdbeschleunigung und φ_0 der größte Auslenkwinkel. Je länger das Pendel ist, desto langsamer pendelt es hin und her.
Misst man an verschiedenen Punkten der Erdoberfläche die Schwingungsdauer eines Pendels bekannter Länge, so kann man dadurch die Erdbeschleunigung an diesen Stellen errechnen.

1.26 Entladung eines Kondensators

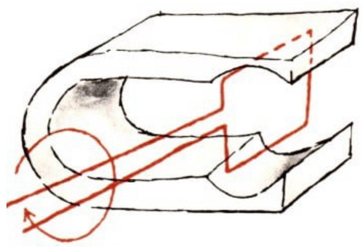
Nicht nur viele mechanische Schwingungen erfolgen nach einem Sinusgesetz. Auch in elektrischen Stromkreisen treten Sinusschwingungen auf.

Wie betrachten z.B. den hier gezeichneten Stromkreis. In ihm ändert sich die Stromstärke nach dem Sinusgesetz

$$I = I_0 \sin(\omega t + \alpha)$$

Die Frequenz ω dieser Schwingung ist gleich $\frac{1}{\sqrt{LC}}$, wenn C die Kapazität des Kondensators und L die Induktivität des Stromkreises ist. Diese Formel ist dem Gesetz für die elastischen Schwingungen einer Masse sehr ähnlich, nur steht hier statt der Federkonstanten eine Größe, die der Kapazität des Kondensators umgekehrt proportional ist, und statt der Masse die Induktivität der Spule.

1.27 Spannung eines Wechselstroms



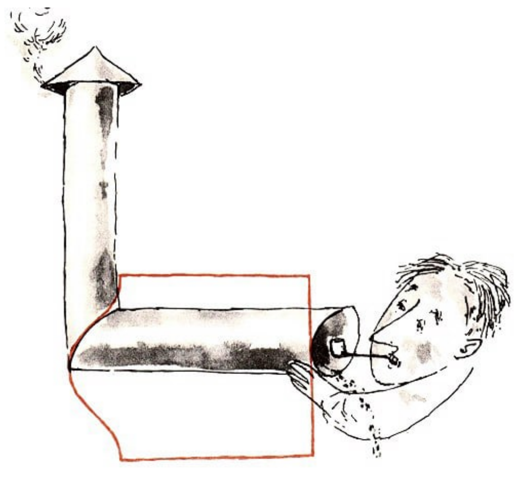
Wird eine Drahtschlinge in einem homogenen Magnetfeld gleichförmig gedreht, so entsteht in dem Draht ein Wechselstrom. Die Spannung dieses Stroms lässt sich durch

$$U = \omega B \cdot A \sin \omega t$$

ausdrücken: A ist der Flächeninhalt der Schlinge, B die magnetische Induktion, ω die Winkelgeschwindigkeit, mit der die Schlinge gedreht wird, und t die Zeit.

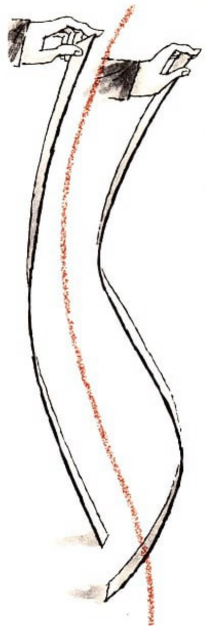
1.28 Wie setzt man zwei Rohre schräg zusammen?

Die bis jetzt gezeigten Beispiele könnten den Eindruck erwecken, dass Sinuskurven nur im Zusammenhang mit Schwingungen auftreten. Das ist aber durchaus nicht der Fall. Zum Beispiel findet man Sinuskurven, wenn man zwei zylindrische Rohre mit kreisförmigem Querschnitt unter einem bestimmten Winkel aneinandersetzt, wie das etwa bei einem Ofenrohr der Fall ist.



Dazu muss man das Rohr schräg durchschneiden. Die Schnittlinie ist eine Sinuskurve. Davon können wir uns überzeugen, wenn wir ein Stück Papier um eine Kerze wickeln, sie schräg durchschneiden und dann das Papier wieder abwickeln. Man kann deshalb, wenn man ein schräggeschnittenes Rohr erhalten will, zuerst das Blech an einem Ende in Form einer Sinuskurve zuschneiden und dann zu einem Rohr zusammenrollen.

1.29 Knickung eines Stabes



Auf Sinuskurven trifft man auch bei der Knickung eines Stabes unter Einwirkung einer vertikalen Kraft. Ist die Kraft zu klein, so tritt keine merkliche Verformung des Stabes auf. Erreicht aber die Kraft eine bestimmte Größe (die sogenannte Knicklast), so beginnt der Stab sich zu biegen, wobei seine Achse eine Sinuskurve bildet. Dieser Versuch lässt sich zum Beispiel sehr leicht mit Hilfe eines Metalllineals durchführen. Die Knicklast ist gleich

$$F = \frac{\pi^2 EJ}{l^2}$$

wobei l die Länge des Stabes ist, die Größe E vom Material und die Größe J vom Querschnitt des Stabes abhängt. Aus der Formel sieht man, dass die Knicklast um so kleiner ist, je länger der Stab ist. Auch dies lässt sich leicht mit Hilfe von Metalllinealen veranschaulichen.
Wird statt der durch die obige Formel gegebenen Kraft eine viermal so große Kraft aufgewendet, so verbiegt sich das Lineal nicht zu einer halben, sondern zu einer ganzen Sinusschwingung.

Die Formel für die Knicklast wurde von dem berühmten Mathematiker Leonhard Euler (1707-1783) angegeben.

1.30 Gedämpfte Schwingungen

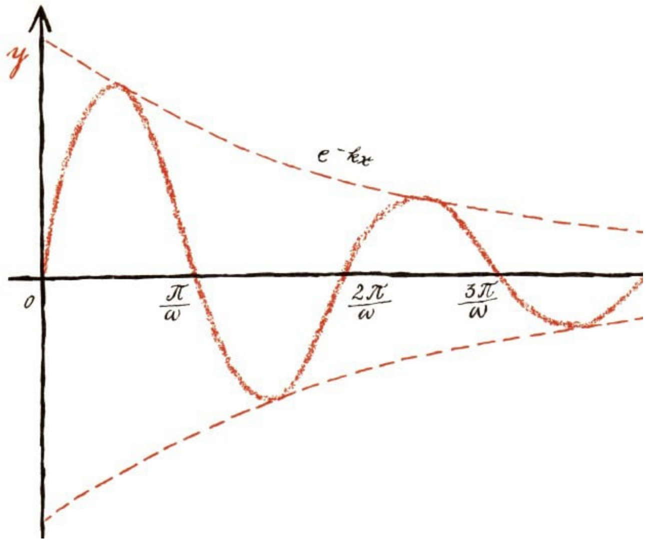
Bis jetzt haben wir, als wir die Schwingungen eines Pendels oder einer Masse an einer Schraubenfeder betrachteten, stets den Luftwiderstand vernachlässigt. In Wirklichkeit wird die Amplitude der Schwingungen wegen des Luftwiderstandes immer kleiner und kleiner, und die Schwingungen klingen ab. Der Weg, den ein Punkt bei einer solchen sogenannten gedämpften Schwingung zurücklegt, wird durch die Beziehung

$$s = ae^{-kt}\sin(\omega t + \alpha)$$

beschrieben. Da der Faktor e^{-kt} mit wachsendem t immer kleiner wird, nimmt die Amplitude immer mehr ab. Nach jeder vollen Schwingung verkleinert sich die Amplitude um den Faktor $e^{-\frac{2\pi k}{\omega}}$.
Die Zahl $\frac{2\pi k}{\omega}$ heißt das logarithmische Dekrement der gedämpften Schwingung. Je größer das logarithmische Dekrement ist, desto stärker wird die Schwingung gedämpft. Nach einiger Zeit ist sie so klein, dass der Körper praktisch zur Ruhe gekommen ist.

In der Abbildung ist eine gedämpfte Schwingung graphisch dargestellt.
Wenn der Widerstand des Mediums, das den schwingenden Körper umgibt, sehr groß ist (wenn z.B. ein Pendel nicht in Luft, sondern in Öl schwingt), gibt es keine Schwingung in diesem Sinne.
Nach Auslenkung des Pendels nähert es sich wieder langsam der Ruhelage, ohne über diese hinauszuschwingen. Eine solche Schwingung nennt man aperiodisch. In diesem

Fall lautet das Bewegungsgesetz



$$s = a_1 e^{-k_1 t} + a_2 e^{-k_2 t}$$

Die Größen a_1 und a_2 hängen von der Anfangslage und von der Anfangsgeschwindigkeit des Pendels ab. Auch elektrische Schwingungen sind wegen des Ohmschen Widerstandes im Stromkreis gedämpft.

1.31 Erzwungene Schwingungen

Wir betrachten wieder eine Masse an einer Schraubenfeder. Wird die Bewegung der Masse nicht gestört, so schwingt sie mit einer bestimmten Frequenz w. Diese Frequenz heißt die Eigenfrequenz der Schwingungen. Völlig anders sehen die Schwingungen aus, wenn wir die Masse erregen (in Schwingung versetzen).
Die erregende Kraft selbst ändere sich nach einer Sinusfunktion, d.h., sie ziehe die Masse abwechselnd nach oben und nach unten. Dann führt die Masse Schwingungen aus, die sich durch Überlagerung zweier Schwingungen ergeben.

Eine davon ist die Schwingung mit der Eigenfrequenz, die andere die mit der Frequenz der erregenden Kraft. Befindet sich die Masse zu Beginn der Schwingung in Ruhe und gehorcht die erregende Kraft dem Gesetz $F = a \sin \beta t$, so lautet das Bewegungsgesetz der Masse in diesem Fall

$$s = \frac{a}{\omega^2 - \beta^2} \left[\sin \beta t - \frac{\beta}{\omega} \sin \omega t \right]^2 \tag{1}$$

ω ist die Eigenfrequenz.
Die graphische Darstellung dieses Gesetzes ist überaus kompliziert. Das liegt daran, dass sich die Funktionen $\sin \beta t$ und $\sin \omega t$ mit verschiedener Frequenz ändern.

Dadurch sind die beiden Schwingungen, an denen die Masse beteiligt ist, zeitweise entgegengesetzt gerichtet (dann hemmen sie einander) und zeitweise gleich gerichtet (dann verstärken sie einander).

Die größte Amplitude der Schwingungen ist ungefähr gleich $\left|\dfrac{a}{\omega(\omega-\beta)}\right|$.
Wenn sich also β wenig von ω unterscheidet, d.h., wenn die Frequenz der erregenden Kraft wenig von der Eigenfrequenz abweicht, kann die Amplitude der Schwingungen sehr groß werden.
In dem Ausdruck $\left|\dfrac{a}{\omega(\omega-\beta)}\right|$ wird dann nämlich der Nenner sehr klein.
Ist $\omega = \beta$, d.h., erregen wir die Masse im Takt ihrer Eigenfrequenz, so ist die Formel (3) nicht mehr anwendbar.
In diesem Fall lautet das Bewegungsgesetz folgendermaßen:

$$s = \frac{a}{2\omega^2}\left[\sin\omega t - \omega t\cos\omega t\right]$$

Die Auslenkung der Masse nimmt mit der Zeit zu, und die Masse kann die Schraubenfeder zerreißen. Diese Erscheinung heißt Resonanz.

1.32 Überlagerung von Schwingungen

Manchmal ist ein Körper nicht nur an einer einzigen schwingenden Bewegung beteiligt, sondern an mehreren. Das ist zum Beispiel der Fall, wenn eine Masse m_1 an einer Schraubenfeder hängt und an diese Masse noch eine Feder mit einer Masse m_2 gehängt wird, wie es in der Abbildung dargestellt ist.
Wenn wir beide Federn dehnen und dann loslassen, sind die Schwingungen der Massen m_1 und m_2 überaus kompliziert. Beispielsweise werden die Schwingungen von m_1 erstens dadurch hervorgerufen, dass sich m_1 abwechselnd nach oben und unten bewegt, und zweitens dadurch, dass sich die Feder zwischen m_1 und m_2 dehnt bzw. zusammenzieht.
Wir sagen in diesem Fall, die Schwingung der Masse m_1 sei die Summe zweier Schwingungen, nämlich der Bewegung von m_1 und der Schwingung von m_2 in Bezug auf m_1.

Man kann auch andere Beispiele für die Überlagerung von Schwingungen nennen. Wenn ein Orchester spielt, ruft jedes Musikinstrument eine Luftschwingung hervor. Diese Schwingungen überlagern einander, und unser Ohr nimmt sie als einen einzigen Akkord auf.
Am häufigsten überlagern einander harmonische Schwingungen. Haben diese Schwingungen die gleiche Frequenz, so ist ihre Summe eine harmonische Schwingung der gleichen Frequenz.
Wenn sich z.B. die harmonischen Schwingungen

$$s_1 = a_1\sin(\omega t + \alpha_1) \qquad \text{und} \qquad s_2 = a_2\sin(\omega t + \alpha_2)$$

überlagern, so ist ihre Summe die Schwingung

$$s = a_1\sin(\omega t + \alpha_1) + a_2\sin(\omega t + \alpha_2)$$

Diese Schwingung lässt sich aber in der Gestalt

$$s = a\sin(\omega t + \alpha)$$

mit der Amplitude

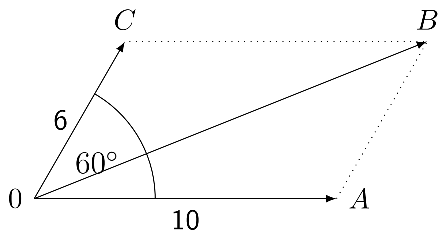
$$a = \sqrt{a_1^2 + a_2^2 + 2a_1a_2\cos(\alpha_1 - \alpha_2)}$$

darstellen. Die Anfangsphase ist dabei durch die Beziehung

$$\tan\alpha = \frac{a_1\sin\alpha_1 + a_2\sin\alpha_2}{a_1\cos\alpha_1 + a_2\cos\alpha_2}$$

bestimmt.
Statt so komplizierter Formeln kann man für die Überlagerung von Schwingungen auch eine einfache geometrische Regel benutzen.

Dazu benötigen wir Vektoren. (Näheres über Vektoren findet der Leser in dem Abschnitt "Geometrische Algebra") Es zeigt sich, dass nicht nur Kraft, Geschwindigkeit und Beschleunigung, sondern auch harmonische Schwingungen durch Vektoren dargestellt werden können.
Eine harmonische Schwingung mit der Amplitude a und der Anfangsphase α lässt sich durch einen Vektor der Länge a ausdrücken, der mit einer festen Achse den Winkel α bildet. Bei Überlagerung von Schwingungen addieren sich die entsprechenden Vektoren nach der Parallelogrammregel.



Hier ist z.B. die Überlagerung der Schwingungen

$$s_1 = 10\sin\omega t \qquad \text{und} \qquad s_2 = 6\sin\left(\omega t + \frac{\pi}{3}\right)$$

durch Vektoren dargestellt. Messen wir die Diagonale OB des Parallelogramms $OABC$, so stellen wir fest, dass die Amplitude der Summe dieser Schwingungen etwa gleich 14 ist.
Die Anfangsphase dieser Schwingung ist durch den Winkel AOB gegeben, also ungefähr 22° oder, gemessen im Bogenmaß, etwa gleich $0{,}37$.
Daher gilt

$$s = 10\sin\omega t + 6\sin\left(\omega t + \frac{\pi}{3}\right) \approx 14\sin(\omega t + 0{,}37)$$

Besonders einfach lassen sich Schwingungen mit gleicher Anfangsphase addieren. In diesem Fall haben nämlich beide Vektoren die gleiche Richtung. Als Summe erhalten wir einfach die Schwingung mit derselben Anfangsphase.

Die Amplitude der Summe dieser Schwingungen ist gleich der Summe aus den Amplituden der Summanden. Unterscheiden sich aber die Winkel α_1 und α_2 um π, also um $180°$, so ergibt sich bei Überlagerung eine Schwingung, deren Amplitude gleich der Differenz der Amplituden der Summanden ist.

Es kann sogar der Fall eintreten, dass sich die beiden Schwingungen gegenseitig aufheben. Diese Erscheinung wird in der Physik Interferenz genannt. Auf Grund der Interferenz kann ein Punkt, der von zwei Lichtquellen gleichphasig beleuchtet wird, unbeleuchtet erscheinen.

1.33 Schwebungen

Komplizierter ist es, wenn sich Schwingungen verschiedener Frequenz überlagern. Wir wollen zum Beispiel zwei Schwingungen

$$s_1 = a \sin \omega_1 t \quad , \quad s_2 = a \sin \omega_2 t$$

addieren, die die gleiche Amplitude, aber verschiedene Frequenzen haben. Das Resultat ist die Schwingung

$$s = a(\sin \omega_1 t + \sin \omega_2 t) \tag{4}$$

Diese Beziehung lässt sich auch in der Gestalt

$$s = 2a \sin \frac{\omega_1 + \omega_2}{2} t \cdot \cos \frac{\omega_1 - \omega_2}{2} t$$

schreiben. Die Faktoren $\sin \frac{\omega_1+\omega_2}{2} t$ und $\cos \frac{\omega_1-\omega_2}{2} t$ verhalten sich ganz unterschiedlich. Wenn sich die Frequenzen ω_1 und ω_2 wenig voneinander unterscheiden, ändert sich der Faktor $\cos \frac{\omega_1-\omega_2}{2} t$ langsam. Die Frequenz seiner Änderung ist gleich

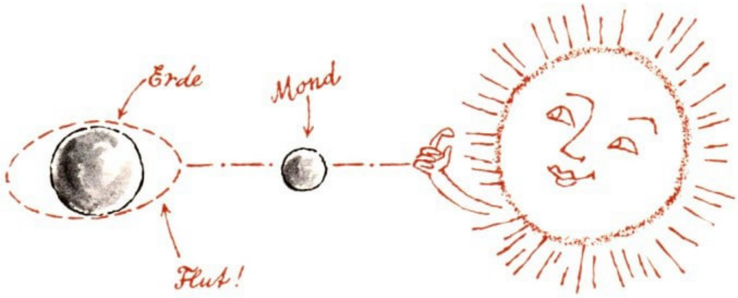
$$\frac{|\omega_1 - \omega_2|}{2}$$

Die Schwingung $\sin \frac{\omega_1+\omega_2}{2} t$ hat die dagegen große Frequenz $\frac{\omega_1+\omega_2}{2}$. Daher ist (4) eine Sinusschwingung mit langsam sich ändernder Amplitude.
Solche Schwingungen heißen Schwebungen.

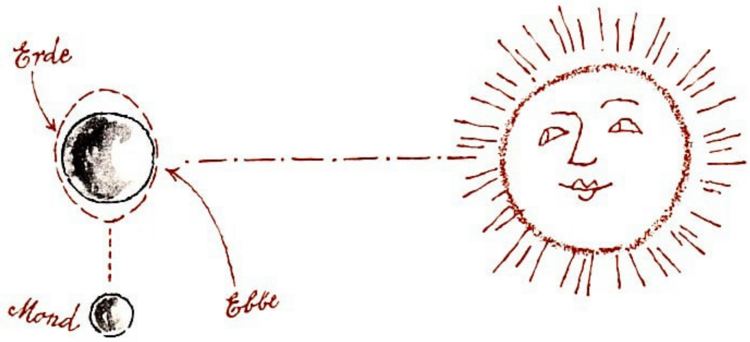
Spielen z.B. zwei Musikinstrumente nahe beieinander liegende Töne mit den Frequenzen ω_1 bzw. ω_2, so hören wir außer diesen beiden Tönen noch einen Ton mit der Frequenz $\frac{|\omega_1-\omega_2|}{2}$.
Komponisten müssen diese Erscheinung berücksichtigen. Aber nicht nur Komponisten, sondern auch Techniker haben mit Schwebungen zu tun. Beispielsweise muss bei der Projektierung eines Schiffes mit zwei Schrauben berücksichtigt werden, dass außer den Schwingungen mit den Frequenzen ω_1 und ω_2, die durch die Bewegung der Schrauben entstehen, noch eine Schwingung mit der Frequenz $\frac{|\omega_1-\omega_2|}{2}$ auftritt.

Stimmt diese Frequenz mit der Eigenfrequenz irgendeines Schiffs teils überein, so können erzwungene Schwingungen mit großer Amplitude entstehen. Schwebungen können auch beim Brennen einer Glühbirne beobachtet werden, die von zwei Wechselstromgeneratoren gespeist wird.

1.34 Ebbe und Flut



Ein sehr interessantes Beispiel für Schwebungen sind Ebbe und Flut. Auf Grund der Anziehung von Mond und Sonne ändert sich der Meeresspiegel der Ozeane ständig. Etwa alle zwölf Stunden erreicht der Meeresspiegel seinen höchsten und jeweils sechs Stunden später seinen niedrigsten Stand.
Jedoch haben die Schwingungen des Meeresspiegels, die vom Mond hervorgerufen werden, nicht die gleiche Schwingungsdauer wie diejenigen, die von der Sonne verursacht werden. Die zweite Schwingung hat eine Schwingungsdauer von 12 Stunden und 25 Minuten, während die erste nur eine von 12 Stunden hat.



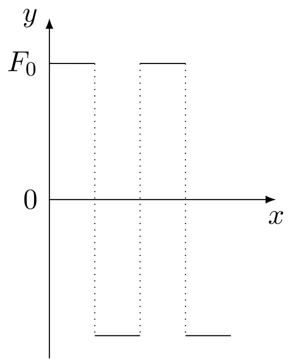
Durch Überlagerung dieser beiden Schwingungen mit fast gleicher Schwingungsdauer (also auch fast gleicher Frequenz) ergibt sich eine Schwebung.
Die größte bzw. kleinste Fluthöhe (bei Springflut bzw. Nippflut) wird dann erreicht, wenn Sonne, Erde und Mond wie in den Abbildungen zueinander stehen. Die größte Fluthöhe ist z.B. zwei- bis dreimal so hoch wie die kleinste.

1.35 Modulation von Schwingungen

Schwingungen mit periodisch sich ändernder Amplitude werden in der Radiotechnik benutzt. Rundfunksender strahlen elektromagnetische Schwingungen mit sehr hoher Frequenz aus (zwischen 150000 und 15 Millionen Schwingungen je Sekunde).
Die Amplitude dieser Schwingungen ändert sich z. B. mit der Frequenz der menschlichen Stimme, die einige hundert bis tausend Schwingungen je Sekunde hat. Die Änderung der Amplitude von Schwingungen heißt Amplitudenmodulation.

1.36 Zerlegung einer periodisch sich ändernden Kraft in harmonische Komponenten

Durch Zusammensetzen harmonischer Schwingungen lassen sich, wie wir wissen, Schwingungen von überaus komplizierter Form darstellen.
Oft hat man aber die umgekehrte Aufgabe zu lösen, nämlich komplizierte Schwingungen in einfache Sinusschwingungen zu zerlegen. Es wirke z.B. auf eine Masse eine erregende Kraft, die ihre Richtung periodisch ändert, Dabei bewege sich die Masse im Verlauf einer gewissen Zeit T mit der konstanten Kraft F, nach der einen Seite, und danach bewege sie sich in der gleichen Zeit und mit der gleichen Kraft in die entgegengesetzte Richtung. Wenn wir diesen Vorgang wiederholen, hat die Kurve der Kraftänderung die hier gezeigte.



Wie kann man nun den zeitlichen Verlauf dieser Kraft durch eine Formel ausdrücken? Jede periodisch sich ändernde Kraft kann mit beliebiger Genauigkeit als Summe mehrerer Kräfte dargestellt werden, die sich nach einem Sinusgesetz ändern. Beispielsweise kann die in der Abbildung gezeichnete Kraft F_0 näherungsweise als Summe dreier Kräfte geschrieben werden, die sich nach den folgenden Gesetzen ändern:

$$F_1 = \frac{4F_0}{\pi}\sin\frac{\pi t}{T}, \quad F_2 = \frac{4F_0}{3\pi}\sin\frac{3\pi t}{T}, \quad F_3 = \frac{4F_0}{5\pi}\sin\frac{5\pi t}{T}$$

Die graphische Darstellung der Summe dieser Kräfte ist der Abbildung sehr ähnlich. Wenn wir die in der Abbildung gezeichnete Kraft noch genauer angeben wollen, müssen wir zu der Kraft $F_1 + F_2 + F_3$ noch die Kraft $F_4 + F_5 + ... + F_n$ addieren; dabei ist

$$F_n = \frac{4F_0}{(2n-1)\pi}\sin\frac{(2n-1)\pi t}{T}$$

Wenn die erregende Kraft von einer Summe sinusförmiger Kräfte gebildet wird, können wir die Wirkung jedes einzelnen Summanden errechnen, wie wir es schon früher taten, und dann die Ergebnisse addieren. Diese Summe ist dann genau die Wirkung der gegebenen erregenden Kraft.

1.37 Spektralanalyse

Die Zerlegung periodischer Prozesse in eine Summe von sinusförmigen Prozessen wird nicht nur beim Studium mechanischer Schwingungen benutzt, sondern auch in vielen anderen Gebieten der Physik, der Technik und sogar der Kunst. Das einfachste Beispiel für eine solche Zerlegung liefert das Spektroskop.

Schicken wir durch dieses Gerät einen Strahl weißen Lichtes, so erhalten wir einen regenbogenfarbenen Streifen - das Spektrum. Jede Farbe des Spektrums entspricht einer sinusförmigen elektromagnetischen Schwingung bestimmter Frequenz.
Somit zerlegt ein Spektroskop eine komplizierte elektromagnetische Schwingung - nämlich das weiße Licht - in Sinusschwingungen.

Messen wir die Energie an verschiedenen Stellen des Spektrums, so können wir die Amplitude der einzelnen Sinusschwingungen berechnen, die im zerlegten Licht auftreten.
Die Zerlegung einer periodischen Schwingung in Sinusschwingungen wird Spektralanalyse genannt.

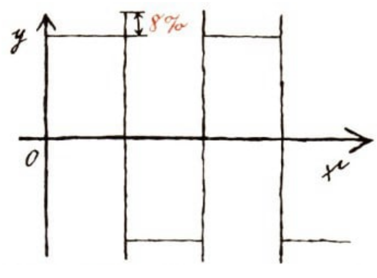
Sie ist z.B. wichtig für Chemie, Materialprüfung und Astronomie. Die Spektralanalyse lässt sich aber auch auf Töne und andere Schwingungen anwenden. Mit ihrer Hilfe kann man z.B. die Besonderheiten im Timbre der Stimme eines Sängers analysieren.

In der Technik wird die Spektralanalyse von Schwingungen dazu benutzt, um verschiedene Konstruktionen durchzurechnen. Beispielsweise kann es vorkommen, dass die Frequenz einer der sinusförmigen Komponenten der Schwingungen, die durch einen laufenden Motor im Flugzeug hervorgerufen werden, mit der Eigenfrequenz irgendeines Flugzeugteils übereinstimmt.

Dann können auf Grund der Resonanz durch das Laufen des Motors so starke Schwingungen dieses Flugzeugteils auftreten, dass es zerstört wird.

1.38 Eine Maschine entdeckte einen Satz

Zur Zerlegung periodischer Schwingungen in sinusförmige Komponenten lassen sich verschiedene Maschinen, sogenannte harmonische Analysatoren, verwenden. Es gibt auch Maschinen, die das umgekehrte Problem lösen, nämlich aus einzelnen sinusförmigen Komponenten die Gesamtschwingung zusammensetzen.
Einst wurden, um die Arbeit einer solchen Maschine zu überprüfen, die sinusförmigen Komponenten der in der Abbildung gezeigten Schwingung in die Maschine eingegeben und dann diese Komponenten addiert.



Nach der Addition zeichnete die Maschine jedoch nicht die erste Kurve, sondern die zweite, d.h. mit zusätzlichen Spitzen an den vertikalen Abschnitten. Zuerst schrieb man

diese Erscheinung einem Fehler der Maschine zu und suchte nach Mitteln und Wegen, um ihn zu korrigieren.
Dann bewies aber der amerikanische Mathematiker und Physiker Josiah Willard Gibbs (1839-1903), dass diese Spitzen an den Sprungstellen der Schwingung immer auftreten müssen.
Diese Erscheinung wurde nach Gibbs das Gibbssche Phänomen genannt, obwohl eigentlich eine Maschine sie entdeckte.

1.39 Störungen im transatlantischen Kabel

Als seinerzeit das erste Überseekabel durch den Atlantischen Ozean gelegt wurde, zeigte es sich, dass man mit Hilfe dieses Kabels keine Telegramme übermitteln konnte. Statt der Punkte und Striche kamen am anderen Ende des Kabels völlig unverständliche Signale an.
Mit der Untersuchung dieser Erscheinung wurde der englische Physiker und Mathematiker Lord Kelvin (1824-1907; bis zu seiner Erhebung in den Adelsstand im Jahre 1892 hieß er Sir William Thomson) beauftragt. Er zerlegte zunächst die Signale in sinusförmige Komponenten und untersuchte, wie diese Komponenten vom Kabel übertragen werden.

Dabei entdeckte er, dass Schwingungen verschiedener Frequenz auch verschieden übermittelt werden, und zwar die einen schneller, die anderen langsamer; die einen werden stärker abgeschwächt, die anderen weniger. Kommen also die Komponenten am anderen Ende des Kabels an, so hat ihre Summe gar keine Ähnlichkeit mehr mit den eingegebenen Signalen.
Lord Kelvin fand, wovon die Änderung der Geschwindigkeit und der Stärke der Sinusschwingungen abhängt. Ferner zeigte er, wie das Kabel beschaffen sein muss, damit es Schwingungen beliebiger Frequenz mit gleicher Geschwindigkeit und gleicher Abschwächung überträgt.

Nachdem man nach seinen Angaben ein Kabel hergestellt und durch den Atlantischen Ozean verlegt hatte, konnten die Signale tatsächlich ohne Verzerrung übertragen werden, und die transatlantische Verbindung war gesichert.

1.40 Radio und Stimmgabel

Manchmal trennen wir von einer Schwingung eine Komponente bestimmter Frequenz ab, ohne erst die Schwingung in einzelne Komponenten zu zerlegen. Dies tun wir nämlich dann, wenn wir ein Radiogerät auf eine bestimmte Frequenz einstellen.
Aus der Gesamtheit der elektromagnetischen Schwingungen, die von allen Rundfunkstationen erzeugt werden, sondern wir die Schwingung aus, die von dem gewünschten Sender ausgestrahlt wird. Genauso hören wir bei einer Stimmgabel den Ton. auf den sie festgelegt ist.

1.41 Schlussbemerkungen

Der Anwendungsbereich der Exponentialfunktion und der trigonometrischen Funktionen in Natur und Technik ist fast unbegrenzt.
Man fragt sich nun natürlich, was diese einzelnen Gebiete gemeinsam haben, z.B. die Reibung eines Seils an einem Pfahl, der radioaktive Zerfall, das Seil gleicher Zugfestigkeit usw. Oder man fragt sich, warum die elektromagnetischen Schwingungen den mechanischen so ähnlich sind oder weshalb man in den verschiedensten Zweigen von Wissenschaft und Technik so oft der Exponentialfunktion und den trigonometrischen Funktionen begegnet.

Es soll jetzt nicht auf diese Fragen geantwortet werden. Am Schluss des nächsten Kapitels, das einem der Gebiete der höheren Mathematik entnommen ist, werden wir darauf zurückkommen.

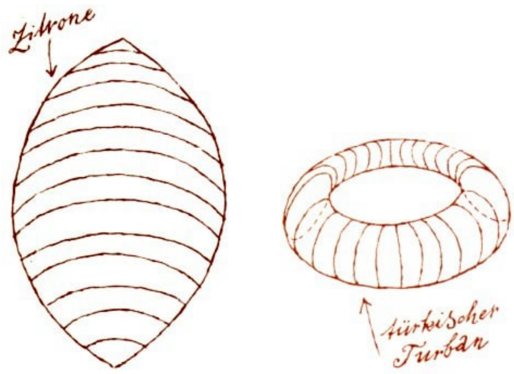
1.42 Integral und Ableitung - Die Keplersche Fassregel

W. G. Boltjanski, I. J. Wilenkin

Wenn Fässer reden könnten, so würden zweifellos viele von ihnen mit Vergnügen eine lehrreiche Geschichte von ihrem großen Verdienst um die Schöpfung der höheren Mathematik erzählen! Die Geschichte lautet so:

Im November 1613 feierte der kaiserliche Mathematiker und Astronom am österreichischen Hof, Johannes Kepler (1571-1630), Hochzeit, zu der er einige Fässer Wein besorgte.
Bei dem Kauf war Kepler sehr erstaunt, dass der Weinhändler den Rauminhalt der verschieden geformten Fässer auf ein und dieselbe Art bestimmte, und zwar maß er den Abstand des Spundlochs von der am weitesten entfernten Stelle des Fassbodens.

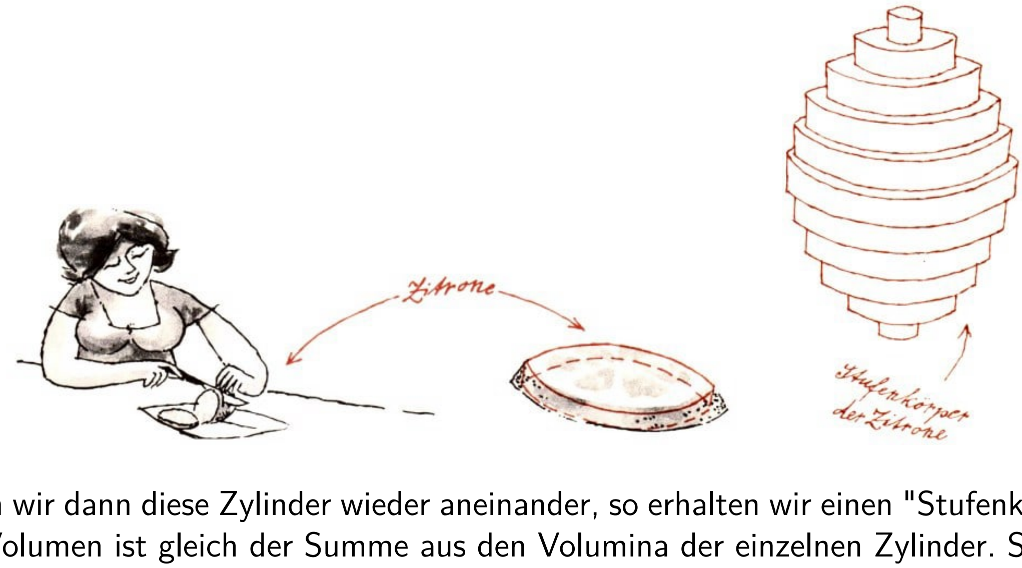
Bei dieser Art der Messung wurde aber die Form des Fasses völlig unberücksichtigt gelassen! Kepler erkannte sofort, dass da ein interessantes mathematisches Problem vorlag. Er dachte darüber nach und fand Formeln nicht nur für das Volumen von Weinfässern, sondern auch für das Volumen der verschiedenartigsten Körper, z.B. einer Zitrone, eines Apfels, einer Quitte und sogar eines türkischen Turbans. Für jeden dieser Körper ersann Kepler neue, zum Teil sehr scharfsinnige Methoden.



Heutzutage ist es bei vielen technischen Problemen notwendig, die Volumina verschiedener Körper (und sogar noch viel komplizierterer als der von Kepler untersuchten) zu berechnen.
Heute kann z.B. fast jeder Ingenieur das Volumen eines Schiffsrumpfes, eines Gasometers, eines Stausees usw. bestimmen. Einfache und allgemeine Methoden zur Lösung dieser Probleme liefert die höhere Mathematik.

1.43 Mathematik am Teetisch

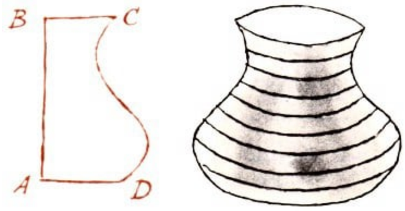
Damit wir eine Vorstellung von diesen allgemeinen Methoden erhalten, wollen wir das Volumen einer Zitrone (wenn auch nur näherungsweise) bestimmen. Keiner der in der Schule untersuchten Körper (Kugel, Zylinder, Kegel usw.) kann uns dabei helfen, denn keinem ist eine Zitrone ähnlich.
Jedoch kommt uns die Hausfrau zu Hilfe: Sie schneidet die Zitrone in dünne Scheiben. Schneiden wir nun noch den Rand der Scheiben glatt ab, so haben wir sie in flache Zylinder umgewandelt, deren Volumen leicht berechnet werden kann.



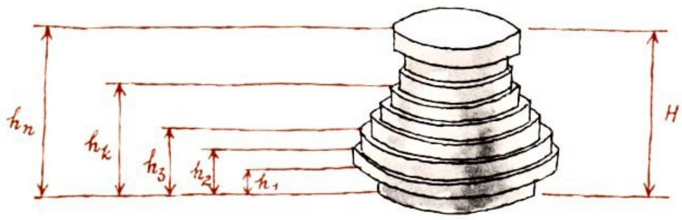
Setzen wir dann diese Zylinder wieder aneinander, so erhalten wir einen "Stufenkörper". Sein Volumen ist gleich der Summe aus den Volumina der einzelnen Zylinder. Sind die Scheiben sehr dünn, so unterscheidet sich das Volumen des Stufenkörpers wenig von dem Volumen der Zitrone, und je dünner wir die Scheiben schneiden, um so kleiner wird die Differenz.

1.44 Volumen eines Körpers

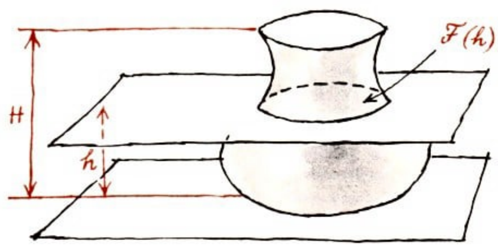
Das Verfahren, das wir zur Berechnung des Volumens einer Zitrone angewendet hatten, ist auch zur Berechnung des Volumens eines beliebigen Rotationskörpers (= Drehkörpers) geeignet.



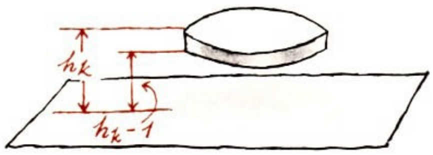
Wir nehmen an, eine Figur $ABCD$ rotiere um die Seite AB. Den so entstehenden Rotationskörper zerschneiden wir in dünne Scheiben und ersetzen jede Scheibe durch einen Zylinder. Dann können wir das Volumen des erhaltenen Stufenkörpers leicht berechnen.



Dazu müssen wir aber wissen, wie sich der Flächeninhalt des Querschnitts mit der Höhe ändert. Der Flächeninhalt des Querschnitts in der Höhe h sei gleich $F(h)$.
Wir nehmen außerdem an, der Körper sei durch Schnitte in den Höhen $h_0, h_1, ..., h_n$ über der Grundfläche in n Scheiben zerlegt.



(Die Grundfläche stimmt mit dem Schnitt in der Höhe h_0, die Deckfläche mit dem in der Höhe h_n überein, d.h., es ist $h_0 = 0, h_n = H$, wenn H die Höhe des Körpers ist.)



Der Flächeninhalt des Querschnitts in der Höhe h_k sei $F(h_k)$. Daher ist das Volumen des Zylinders, durch den wir die k-te Scheibe ersetzen, gleich $F(h_k) \cdot (h_k - k_{k-1})$, denn die Höhe des Zylinders ist gleich $h_k - h_{k-1}$.

Wenn wir die Volumina aller Zylinder addieren, erhalten wir das Volumen V_S des Stufenkörpers:

$$V_S = F(h_1) \cdot (h_1 - h_0) + F(h_2) \cdot (h_2 - h_1) + ... + F(h_n) \cdot (h_n - h_{n-1})$$

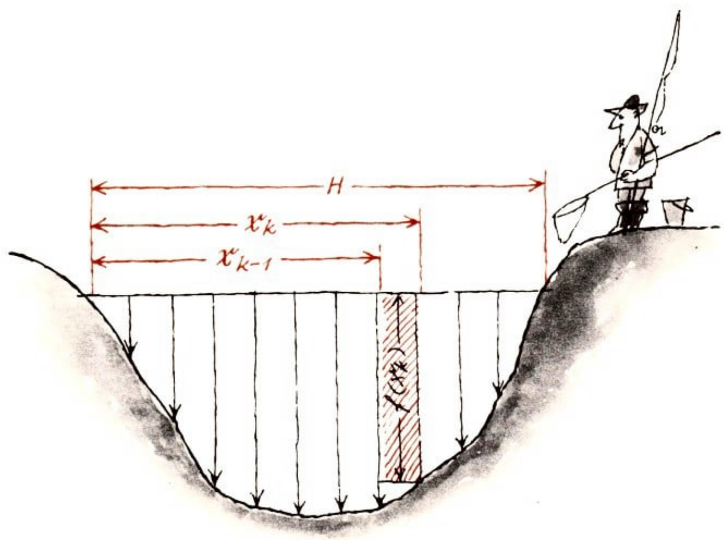
Je dünner die Scheiben sind, um so weniger unterscheidet sich V_S von dem Volumen des Rotationskörpers.
Auf dieselbe Art kann das Volumen jedes Körpers berechnet werden, wenn bekannt ist, wie sich der Flächeninhalt des Querschnitts mit der Höhe ändert.

Um das Volumen eines zu projektierenden Schiffes berechnen zu können, genügen z.B. Zeichnungen (in bestimmtem Maßstab) der Querschnitte des Schiffes. Anhand dieser Zeichnungen bestimmt man den Flächeninhalt jedes Querschnitts (wie man das macht, zeigen wir im nächsten Abschnitt) und errechnet dann mit Hilfe der obigen Formel näherungsweise das Volumen des Schiffes.
Natürlich lässt sich mit diesem Verfahren auch das Volumen von Gasometern, Stauseen usw. berechnen.

1.45 Vermessen eines Flusses

Bei der Projektierung von Wasserkraftwerken muss die Durchflussmenge des Flusswassers bekannt sein, d.h. die Wassermenge, die an einer bestimmten Stelle je Sekunde durchfließt.
Die Durchflussmenge ist gleich dem Produkt aus dem Flächeninhalt des Querschnitts an dieser Stelle und der Strömungsgeschwindigkeit.
Die Strömungsgeschwindigkeit ist einfach festzustellen, aber den Flächeninhalt des Querschnitts zu bestimmen ist schwieriger.
Jedoch auch hierbei hilft uns eine Zerlegung in Streifen. Jeder dieser Streifen kann durch ein Rechteck angenähert werden.

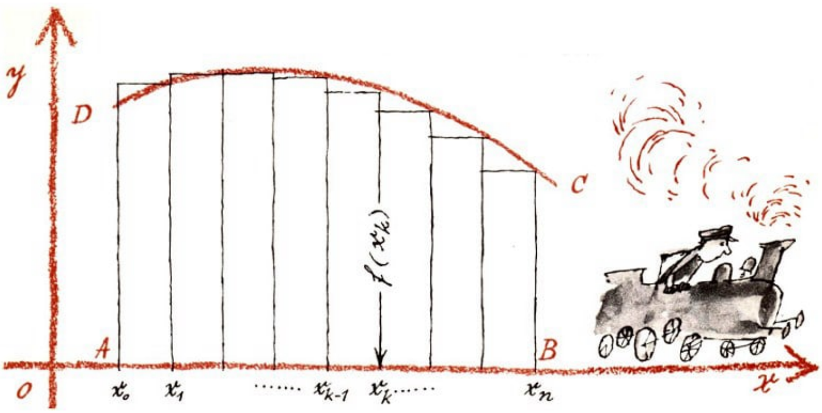
Addieren wir dann die Flächeninhalte dieser Rechtecke, so finden | wir einen Näherungswert für den Flächeninhalt des Querschnitts. h Je schmaler die Streifen sind, desto besser ist die Annäherung an den exakten Wert.



Gemessen werde die Flusstiefe in den Punkten mit den Abständen $x_0, x_1, ..., x_n$ vom Ufer (x_0 ist gleich Null, und x_n ist gleich der Flussbreite H). Im Abstand x_k vom Ufer sei die Flusstiefe gleich $f(x_k)$.
Dann ist der Flächeninhalt dieses Querschnitts annähernd gleich

$$F \approx f(x_1) \cdot (x_1 - x_0) + f(x_2) \cdot (x_2 - x_1) + ... + f(x_n) \cdot (x_n - x_{n-1})$$

Hat eine geometrische Figur $ABCD$ die hier gezeigte Form (eine solche Figur wollen wir krummliniges Trapez nennen) und ist die Höhe in einem Punkt mit der Abszisse x gleich $f(x)$, so können wir zur Berechnung des Flächeninhalts dieser Figur ebenfalls die obige Formel benutzen.



Je dichter die Punkte $x_0, x_1, ..., x_n$ auf der Strecke AB liegen, desto genauer ist der Wert, den wir mit Hilfe dieser Formel für den Flächeninhalt finden.

1.46 Autofahrt

Um den von einem Auto zurückgelegten Weg zu messen, ist ein spezielles Instrument in das Auto eingebaut. Arbeitet dieses Instrument aber nicht, so kann man den zurückgelegten Weg auch mit Hilfe des Geschwindigkeitsmessers, des sogenannten Tachometers, berechnen.

Dazu muss man das Tachometer zu den Zeitpunkten $t_0 = 0, t_1, t_2, ..., t_n = T$ ablesen. Würde sich das Auto vom Zeitpunkt t_{k-1} bis zum Zeitpunkt t_k gleichförmig mit der Geschwindigkeit $v(t_k)$ bewegen, die es in Wirklichkeit erst am Ende dieses Zeitintervalls hat, so hätte es in der Zeit $t_k - t_{k-1}$ den Weg $v(t_k) \cdot (t_k - t_{k-1})$ zurückgelegt.
Daher wäre der vom Auto in der ganzen Zeit von 0 bis T zurückgelegte Weg gleich

$$v(t_1) \cdot (t_1 - t_0) + v(t_2) \cdot (t_2 - t_1) + ...v(t_n) \cdot (t_n - t_{n-1})$$

Diese Formel liefert aber nur einen Näherungswert für den Weg, denn das Auto bewegt sich ja nicht immer gleichförmig, und selbst in kleinen Zeitintervallen kann es seine Geschwindigkeit mehrmals wechseln.
Je häufiger wir jedoch das Tachometer ablesen, d.h., je kleiner die Zeitintervalle zwischen den einzelnen Messungen sind, desto genauer gibt die obige Formel den Weg an, den das Auto zurückgelegt hat.

1.47 Arbeit einer Lokomotive

Um einen Zug zu bewegen, muss die Lokomotive Arbeit aufwenden. Diese Arbeit zu berechnen wäre sehr leicht, wenn die ganze vom Zug befahrene Strecke gleichmäßig beschaffen wäre.
Dann brauchte man nämlich nur die Zugkraft der Lokomotive mit der Länge des Weges zu multiplizieren.

Aber an einigen Stellen gehen die Bahngleise bergauf, an anderen bergab, oder die Gleise sind in besserem oder schlechterem Zustand. Deshalb muss die Lokomotive an der einen Stelle mehr Zugkraft aufwenden als an einer anderen.
Man kann auch sagen, die Zugkraft sei eine Funktion $P(l)$ des Weges l, den die Lokomotive seit Beginn ihrer Fahrt zurückgelegt hat. Um die Arbeit der Lokomotive zu bestimmen, müssen wir den ganzen Weg in kleine Teilstrecken zerlegen.

Die Teilpunkte mögen sich vom Anfangspunkt des Weges in den Abständen $l_0 = 0, l_1, ..., l_n = L$ befinden. L ist die Länge des ganzen Weges.
Auf jeder Teilstrecke zwischen zwei Punkten l_{k-1} und l_k können wir die Zugkraft näherungsweise als konstant annehmen und gleich $P(l_k)$ setzen. Folglich ist die auf dieser Teilstrecke aufgewendete Arbeit annähernd gleich $P(l_k) \cdot (l_k - l_{k-1})$, und für die ganze Arbeit W ergibt sich der Näherungswert

$$W \approx P(l_1) \cdot (l_1 - l_0) + P(l_2) \cdot (l_2 - l_1) + ... + P(l_n) \cdot (l_n - l_{n-1})$$

Auch hier wird der Wert für die Arbeit der Lokomotive um so genauer, je dichter wir die Punkte $l_0, l_1, ..., l_n$ wählen. Analog lässt sich auch in anderen Fällen die Arbeit einer veränderlichen Kraft errechnen.

1.48 Das Integral

Wir haben eben einige Beispiele aus verschiedenen Gebieten der Physik, aus der Technik und aus der Geometrie betrachtet. Trotz äußerer Unterschiede haben diese Beispiele eines gemeinsam.
Jedesmal erhielten wir für den Näherungsausdruck der gesuchten Größe (der Arbeit, des Volumens, des Flächeninhalts, des Weges usw.) eine Summe der Gestalt

$$f(x_1) \cdot (x_1 - x_0) + f(x_2) \cdot (x_2 - x_1) + ... + f(x_n) \cdot (x_n - x_{n-1})$$

Hierbei ist $f(x)$ eine in einem Intervall zwischen a und b gegebene Funktion, und $x_0 = a, x_1, ..., x_{n-1}, x_n = b$ sind Punkte auf dieser Strecke.

Beispielsweise war die Funktion $f(x)$ bei der Berechnung des Weges die Geschwindigkeit zur Zeit x (früher bezeichneten wir die Zeit mit t und nicht mit x, aber das ist unwesentlich), a war gleich 0 und b gleich der Zeit T (bei der Bewegung des Autos). Bei der Berechnung der Arbeit war $f(x)$ die Zugkraft der Lokomotive, und a und b bezeichneten den Anfangs- bzw. den Endpunkt des Weges usw.

Summen dieser Form findet man in der Mathematik und ihren Anwendungen sehr oft. Man nennt sie Integralsummen. Sie geben den Wert der gesuchten Größe nur näherungsweise an.
Legen wir aber die Punkte $x_0, x_1, ..., x_n$ im Intervall von a bis b immer dichter und dichter, so nähern sich die Integralsummen immer mehr einer gewissen Zahl, nämlich dem genauen Wert der gesuchten Größe.

Diese Zahl heißt das bestimmte Integral der Funktion $f(x)$ über das Intervall von a bis b, und man schreibt

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

Die Funktion $f(x)$ heißt der Integrand, x die Integrationsveränderliche oder Integrationsvariable, und a und b sind die Integrationsgrenzen. Also ist

$$\int\limits_a^b f(x)dx = \lim[f(x_1)\cdot(x_1-x_0)+f(x_2)\cdot(x_2-x_1)+...+f(x_n)\cdot(x_n-x_{n-1})]$$

Dabei ist der Grenzwert $\lim$ ist die Abkürzung des lateinischen Wortes limes) unter der Bedingung zu verstehen, dass die Anzahl der Teilintervalle unbeschränkt wächst und gleichzeitig die Länge jedes Teilintervalls immer kleiner und kleiner wird, d.h. gegen Null strebt.
Bei unseren Betrachtungen wollen wir uns auf solche Fälle beschränken, in denen dieser Grenzwert existiert. Man soll aber nicht glauben, das sei immer der Fall.

In der Bezeichnung $\int\limits_a^b f(x)dx$ ist die Erinnerung an die Integralsumme erhalten geblieben, aus der sich das Integral ergibt.
Früher wurde der Buchstabe S oft in der Form $\int$ geschrieben. Daher ist das Integralzeichen einfach der erste Buchstabe des lateinischen Wortes summa (Summe).
Hinter dem Integralzeichen wird angegeben, dass Ausdrücke der Form $f(x_k)\cdot(x_k-x_{k-1})$ summiert werden. Nur wird statt der Differenz $x_k - x_{k-1}$ das Zeichen dx geschrieben; d ist der erste Buchstabe des lateinischen Wortes differentia (Differenz).

Der Integralbegriff ist einer der grundlegenden Begriffe in der höheren Mathematik. Mit seiner Hilfe werden viele der früher erhaltenen Formeln wesentlich kürzer, und sie geben vor allen Dingen nicht einen angenäherten, sondern den exakten Wert an. Beispielsweise nimmt die Formel für das Volumen eines beliebigen Körpers die Gestalt

$$V = \int\limits_0^H F(h)dh \tag{1}$$

an. Dabei ist H die Höhe des Körpers und $F(h)$ der Flächeninhalt des Querschnitts, der parallel zur Grundfläche in der Höhe h liegt.
Die Formel für den Flächeninhalt der in der Abbildung gezeigten Figur lautet nun

$$F = \int_a^b f(x)dx \tag{2}$$

wobei $f(x)$ die Ordinate im Punkt x ist. Der Weg, der während eines Zeitintervalls von 0 bis T zurückgelegt wird, lässt sich durch die Geschwindigkeit $v(t)$ folgendermaßen ausdrücken:

$$s = \int_0^T v(t)dt \tag{3}$$

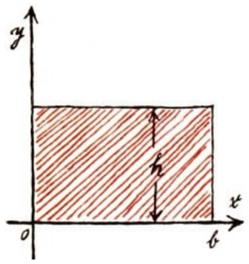
Die Arbeit W wird durch

$$W = \int_0^L P(l)dl \tag{4}$$

bestimmt, wobei L die Weglänge ist und $P(l)$ die Kraft angibt, durch die im Abstand l vom Beginn des Weges Arbeit geleistet wird.

1.49 Geometrische Berechnung von Integralen

Die Formeln (1) und (2) lassen sich zur Berechnung des Flächeninhalts bzw. des Volumens der verschiedensten Körper benutzen.
Da wir schon die Flächeninhalte und Volumina einfacher Körper von der Schule her kennen, können wir mit Hilfe dieser Werte einige einfache Integrale auswerten. Später werden wir zeigen, wie sich diese Integrale ohne geometrische Hilfsmittel unmittelbar berechnen lassen.



Die einfachste geometrische Formel für einen Flächeninhalt ist die für den Flächeninhalt eines Rechtecks: $F = hb$. Ein Rechteck kann man als krummliniges Trapez auffassen, dessen Höhe in allen Punkten gleich groß, in diesem Fall gleich h ist. Also lautet die Formel für den Flächeninhalt in Integralschreibweise

$$\int_0^b hdx \qquad (h \text{ konstant})$$

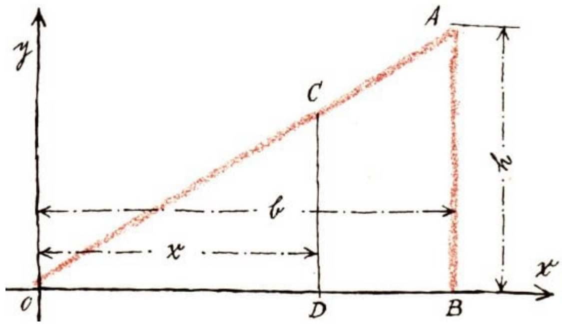
Damit haben wir die Beziehung

$$\int_0^b hdx = hb \qquad (h \text{ konstant})$$

bewiesen. Insbesondere ergibt sich für $h = 1$

$$\int_0^b dx = b \tag{5}$$

Wir erinnern uns nun an die Formel $F = 2hb$ für den Flächeninhalt eines rechtwinkligen Dreiecks; h und b sind die Katheten.



Aus der Abbildung ist ersichtlich, dass man das Dreieck als krummliniges Trapez auffassen kann, bei dem die Ordinate y im Punkt mit der Abszisse x gleich ist. Dies folgt aus der Ähnlichkeit der Dreiecke OAB und OCD. Daher lautet der Flächeninhalt des Dreiecks in Integralschreibweise

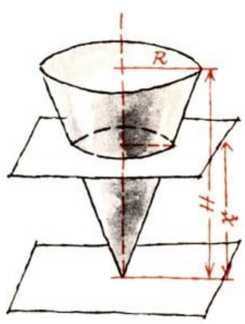
$$F = \int_0^b \frac{hx}{b} dx$$

Wir haben damit die Beziehung

$$\int_0^b \frac{hx}{b} dx = \frac{1}{2} hb$$

bewiesen. Ist das Dreieck OAB gleichschenklig, d.h., ist $h = b$, so erhalten wir

$$\int_0^b x dx = \frac{1}{2} b^2 \tag{6}$$



Zum Schluss wollen wir noch ein weiteres Beispiel betrachten, und zwar einen Kegel.
Seine Höhe sei H, der Radius seiner Grundfläche sei R. Wir stellen den Kegel so auf die Spitze, dass seine Achse genau senkrecht steht, und legen durch den Kegel im Abstand x von seiner Spitze eine zur Grundfläche des Kegels parallele Ebene.
Diese Ebene schneidet den Kegel in einem Kreis, dessen Radius gleich $\frac{Rx}{H}$ ist.

Der Flächeninhalt $F(x)$ des Schnittes ist gleich

$$\pi \left(\frac{Rx}{H}\right)^2 = \frac{\pi R^2}{H^2} x^2$$

Daher lässt sich das Volumen V des Kegels mit Hilfe der Formel (1) durch

$$V = \int_0^H \frac{\pi R^2}{H^2} x^2 dx$$

ausdrücken. Vergleichen wir diese Beziehung mit der aus der Schule bekannten Formel $V = \frac{1}{3}\pi R^2 H$ für das Volumen eines Kegels, so finden wir

$$\int_0^H \frac{\pi R^2}{H^2} x^2 dx = \frac{1}{3}\pi R^2 H$$

Diese Formel gilt für alle positiven Werte von R und H. Insbesondere erhalten wir für $R = \frac{1}{\sqrt{\pi}} b$, $H = b$ (also $\frac{\pi^2 R^2}{H^2} = 1$, $\pi R^2 H = b^3$)

$$\int_0^b x^2 dx = \frac{1}{3} b^3 \tag{7}$$

Die Formeln (5), (6) und (7) lassen sich offenbar zu einer einzigen Formel zusammenfassen:

$$\int_0^b x^n dx = \frac{1}{n+1} b^{n+1} \qquad (n = 0,1,2) \tag{8}$$

Diese Beziehung gilt aber, wie in der höheren Mathematik bewiesen wird, nicht nur für $n = 0, 1, 2$, sondern für alle positiven Werte des Exponenten n. So ist z.B.

$$\int_0^b x^{10} dx = \frac{1}{11} b^{11}, \qquad \int_0^b \sqrt{x} dx = \int_0^b x^{\frac{1}{2}} dx = \frac{2}{3} b^{\frac{3}{2}}$$

1.50 Integration von Polynomen

Nach diesen Vorbereitungen fällt uns die Integration eines beliebigen Polynoms nicht schwer. Zunächst wollen wir zwei einfache, aber sehr wichtige Bemerkungen vorausschicken.

Zwei Körper M_1 und M_2 mögen sich in der gleichen Richtung bewegen, und zwar so, dass die Geschwindigkeit von M_2 zu jeder Zeit k-mal so groß ist wie die von M_1. Dann ist klar, dass auch der von M_2 zurückgelegte Weg k-mal so groß ist wie der Weg, den der Körper M_1 in der gleichen Zeit durchläuft.
Wir wollen nun versuchen, diese Tatsache durch eine Formel auszudrücken.

Ist die Geschwindigkeit des Körpers M_1 zur Zeit t gleich $v(t)$, so ist die Geschwindigkeit von M_2 zur gleichen Zeit gleich k$v(t)$. Die Wege s_1 und s_2, die von M_1 bzw. M_2 im Zeitintervall von $t = 0$ bis $t = T$ zurückgelegt werden, sind

$$s_1 = \int_0^T v(t) dt \qquad , \qquad s_2 = \int_0^T kv(t) dt$$

Da nun der Weg des zweiten Körpers k-mal so lang ist wie der von M_1, also $s_2 = ks_2$ gilt, muss

$$\int_0^T kv(t)dt = k\int_0^T v(t)dt$$

sein. Wir sehen daran, dass ein konstanter Zahlenfaktor vor das Integralzeichen gezogen werden kann.

Nun bewege sich der Körper M_1 in einer Richtung, und auf seiner Oberfläche bewege sich der Körper M_2 in der gleichen Richtung.

Man stelle sich vor, ein Schiff fahre auf einem Fluss, und an Deck laufe ein Mensch in Fahrtrichtung. Die Geschwindigkeit des Körpers M_1 sei $v_1(t)$, die von M_2 auf dem Körper M_1 sei $v_2(t)$. Dann ist der Weg, den der Körper M_1 im Zeitintervall von $t = 0$ bis t = T zurücklegt, gleich

$$\int_0^T v_1(t)dt$$

Der Körper M_2 durchläuft auf der Oberfläche von M_1 in der gleichen Zeit den Weg

$$\int_0^T v_2(t)dt$$

Der Gesamtweg, den der Körper M_2 auf Grund seiner Eigenbewegung und der Bewegung des ihn tragenden Körpers M_1 im Raum zurücklegt, ist gleich

$$\int_0^T v_1(t)dt + \int_0^T v_2(t)dt$$

Da die Geschwindigkeit von M_2 im Raum offenbar gleich $v_1(t) + v_2(t)$ ist, muss andererseits der von M_2 im Raum zurückgelegte Weg gleich

$$\int_0^T [v_1(t) + v_2(t)]dt$$

sein. Diese beiden Werte müssen übereinstimmen. Also gilt

$$\int_0^T [v_1(t) + v_2(t)]dt = \int_0^T v_1(t)dt + \int_0^T v_2(t)dt$$

d.h., das Integral einer Summe zweier (oder mehrerer) Funktionen ist gleich der Summe der Integrale der einzelnen Summanden.
Wir wollen nun wieder zu der Integration von Polynomen zurückkehren und zum Beispiel das Integral

$$\int_0^b (x^2 - 3x + 5)dx$$

berechnen. Der Integrand ist die Summe $x^2 + (-3x) + 5$, so dass wir das Integral in drei Integrale zerlegen können:

$$\int_0^b x^2dx + \int_0^b (-3x)dx + \int_0^b 5dx$$

Im zweiten und dritten Integral können die Zahlenfaktoren vor das Integralzeichen gezogen werden, und wir gelangen leicht zu dem Resultat

$$\int_0^b (x^2 - 3x + 5)dx = \int_0^b x^2dx - 3\int_0^b xdx + 5\int_0^b dx = \frac{b^3}{3} - 3\frac{b^2}{2} + 5b$$

d.h., Polynome können gliedweise integriert werden. Ist allgemein

$$f(x) = a_0x^n + a_1x^{n-1} + ... + a_{n-1}x + a_n$$

ein Polynom n-ten Grades, so lautet die Formel zur Berechnung seines Integrals

$$\int_0^b = a_0\frac{b^{n+1}}{n+1} + a_1\frac{b^n}{n} + ... + a_{n-1}\frac{b^2}{2} + a_nb$$

1.51 Anwendung von Integralen

Wir wissen jetzt, wie Polynome integriert werden, und wollen zeigen, wie einfach sich einige von der Schule her bekannte Formeln mit Hilfe von bestimmten Integralen ergeben.
Wir leiten die Formel für den Weg her, den ein gleichförmig beschleunigter Körper zurücklegt. Ist die Anfangsgeschwindigkeit des Körpers zur Zeit $t = 0$ gleich v_0 und ist die Beschleunigung gleich a, so gilt für die Geschwindigkeit zu einer beliebigen Zeit t

$$v(t) = v_0 + at$$

Daher ist der Weg, den der Körper vom Anfang seiner Bewegung bis zur Zeit T durchlaufen hat, auf Grund von Formel (3) gleich

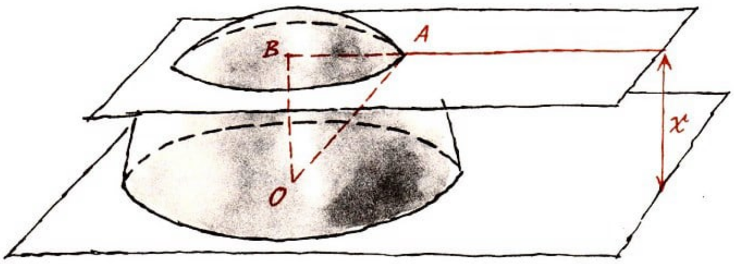
$$s = \int_0^T v(t)dt = \int_0^T (v_0 + at)dt = v_0T + \frac{aT^2}{2}$$

Ganz einfach lässt sich auch mit Hilfe von Integralen die Arbeit berechnen, die zur Dehnung einer Schraubenfeder aufgewendet werden muss. Nach dem Hookeschen Gesetz ist die Kraft $P(x)$, die zur Dehnung der Feder um die Größe x erforderlich ist, gleich kx, wobei k die Federkonstante bezeichnet. Daher ist die zur Verlängerung der Feder um die Strecke b notwendige Arbeit gleich

$$W = \int_0^b kx dx = k\frac{b^2}{2}$$

Da die aufgewandte Arbeit gleich der potentiellen Energie der gedehnten Feder ist, kann man diese Formel auch als Ausdruck für die potentielle Energie ansehen.

Wir wollen nun einige Formeln aus der Geometrie herleiten. Als erstes bestimmen wir das Volumen einer Kugel vom Radius R.



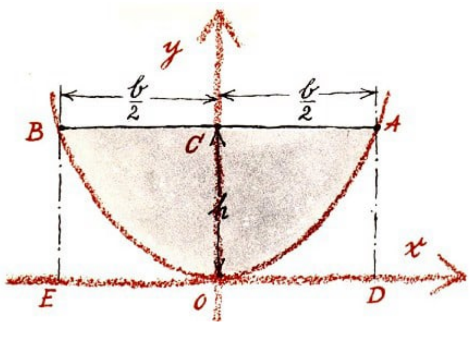
Dazu gehen wir so vor, dass wir das Volumen der Halbkugel berechnen und dann verdoppeln. Wir schneiden die Halbkugel mit einer zur Grundfläche parallelen Ebene. Diese Schnittebene habe von der Grundfläche den Abstand x. Die Schnittkurve ist ein Kreis vom Radius $AB = \sqrt{R^2 - x^2}$.
Das ergibt sich, wenn man den Satz des Pythagoras auf das Dreieck OAB anwendet. Damit errechnet sich der Flächeninhalt des Schnittkreises zu

$$\pi(\sqrt{R^2 - x^2})^2 = \pi R^2 - \pi x^2$$

und für das Volumen der Halbkugel (ihre Höhe ist gleich R) ergibt sich

$$V = \int_0^R (\pi R^2 - \pi x^2) dx = \pi R^2 R - \pi\frac{R^3}{3} = \frac{2}{3}\pi R^3$$

Folglich ist das Volumen der ganzen Kugel gleich $\frac{4}{3}\pi R^3$.

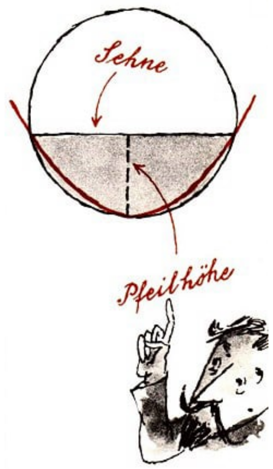


Mit Hilfe der Integralrechnung lassen sich aber auch solche Flächeninhalte und Volumina berechnen, die nicht allgemein bekannt sind. Wir wollen z.B. den Flächeninhalt eines Parabelsegments $AOBA$ bestimmen, bei dem die Sehne AB gleich b und die Pfeilhöhe OC gleich A ist. Die Parabelgleichung lautet $y = ax^2$.

Da die Ordinate AD im Punkt mit der Abszisse $x = \frac{b}{2}$ gleich h ist, gilt (auf Grund der Parabelgleichung) $h = \frac{ab^2}{4}$, also $a = \frac{4h}{b^2}$.
Somit ist das Parabelsegment von unten durch eine Parabel berandet, auf der ein Punkt mit der Abszisse x die Ordinate $y = \frac{4hx^2}{b^2}$ hat. Nun können wir leicht den Inhalt des Flächenstücks OAD berechnen. Er ist auf Grund der Formel (2) gleich

$$A = \int_0^{\frac{h}{2}} \frac{4hx^2}{b^2} dx = \frac{4h}{3b^2} \left(\frac{b}{2}\right)^3 = \frac{hb}{6}$$

Der Flächeninhalt des Parabelsegments ergibt sich, wenn man von dem Flächeninhalt des Rechtecks $ABED$, der gleich bh ist, den doppelten Inhalt der Fläche OAD subtrahiert. Das Ergebnis lautet $\frac{2bh}{3}$.



Ein Kreissegment mit nicht zu großem Zentriwinkel lässt sich näherungsweise durch ein Parabelsegment mit der gleichen Sehne und der gleichen Pfeilhöhe ersetzen. Deshalb gilt für den Flächeninhalt A eines Kreissegments die Näherungsformel

$$A \approx \frac{2}{3} bh$$

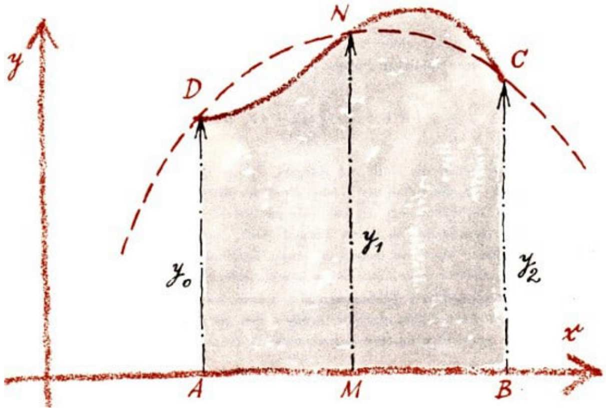
Ist beispielsweise der Zentriwinkel gleich $60°$, so ergibt sich nach dieser Formel $0{,}0893...R^2$. Der genaue Wert ist $0{,}0906...R^2$.
Also hat die Formel sogar für einen verhältnismäßig großen Zentriwinkel wie $60°$ noch eine Genauigkeit bis zu 1,5 %.

1.52 Eine Zauberformel

Dasselbe Verfahren, das wir zur angenäherten Berechnung des Flächeninhalts eines Kreissegments anwendeten, ist auch im Fall eines beliebigen krummlinigen Trapezes geeignet, das von oben durch eine Kurve CD mit der Gleichung $y = f(x)$, von unten durch die Strecke AB auf der x-Achse und an den Seiten von den Ordinaten AD und BC berandet wird.

Den Mittelpunkt der Strecke AB bezeichnen wir mit M, die Ordinate im Punkt M mit MN. Die Längen der Ordinaten AD, MN und BC seien y_0, y_1, bzw. y_2. Durch die Punkte D, N und C legen wir einen Parabelbogen (es gibt stets genau einen solchen Bogen).
Einfache Rechnungen mit den Formeln (5), (6) und (7) zeigen, dass der Inhalt der Fläche unterhalb dieses Parabelbogens zwischen den Ordinaten AD und BC gleich

$$\frac{b-a}{6}(y_0 + 4y_1 + y_2)$$



ist (a und b sind die Abszissen von A und B). Ohne allzu großen Fehler kann man annehmen, dieser Wert sei gleich dem Flächeninhalt F des krummlinigen Trapezes $ABCD$, also

$$F_{ABCD} \approx \frac{b-a}{6}(y_0 + 4y_1 + y_2)$$

Da der Flächeninhalt des krummlinigen Trapezes in Integralschreibweise gleich $\int\limits_a^b f(x)dx$ ist, gibt diese Formel einen Näherungswert dieses Integrals an:

$$\int\limits_a^b f(x)dx \approx \frac{b-a}{6}(y_0 + 4y_1 + y_2)$$

Dabei sind y_0, y_1 und y_2 die Werte der Funktion $f(x)$ in den Punkten mit den Abszissen $x = a$, $x = \frac{a+b}{2}$ bzw. $x = b$.
Das Volumen eines beliebigen Körpers kann näherungsweise nach einer ähnlichen Formel berechnet werden:

$$V \int\limits_0^H F(x)dx \approx \frac{H}{6}(F_0 + 4F_1 + F_2)$$

Dabei ist H die Höhe des Körpers, F_0 der Inhalt der Grundfläche, F_1 der Inhalt der Fläche, die in der Mitte zwischen Grund- und Deckfläche liegt und zu diesen parallel ist, und F_2, der Inhalt der Deckfläche. Zu dieser Formel greift man, wenn man das Volumen eines Baumstamms, eines Heuschobers, eines Fasses oder anderer mehr oder weniger komplizierter Figuren näherungsweise berechnen will.

Es ist bemerkenswert, dass diese Formel für alle in der Schule untersuchten Körper (Prisma, Zylinder, Pyramide, Pyramidenstumpf, Kegel, Kegelstumpf, Kugel, Kugelschicht, Kugelsegment) nicht einen Näherungswert, sondern den genauen Wert liefert. Wir empfehlen unseren Lesern, diese Behauptung einmal selbst nachzuprüfen.

1.53 Wie misst man die Geschwindigkeit eines Geschosses?

Wir haben früher schon oft von der Geschwindigkeit (etwa der eines Autos) gesprochen. Beispielsweise erhielten wir die Formel

$$s = \int\limits_0^T v(t)dt$$

in der $v(t)$ die Geschwindigkeit des betrachteten Körpers für den Zeitpunkt t bezeichnet. Wie lässt sich nun aber die Geschwindigkeit in diesem Augenblick messen?

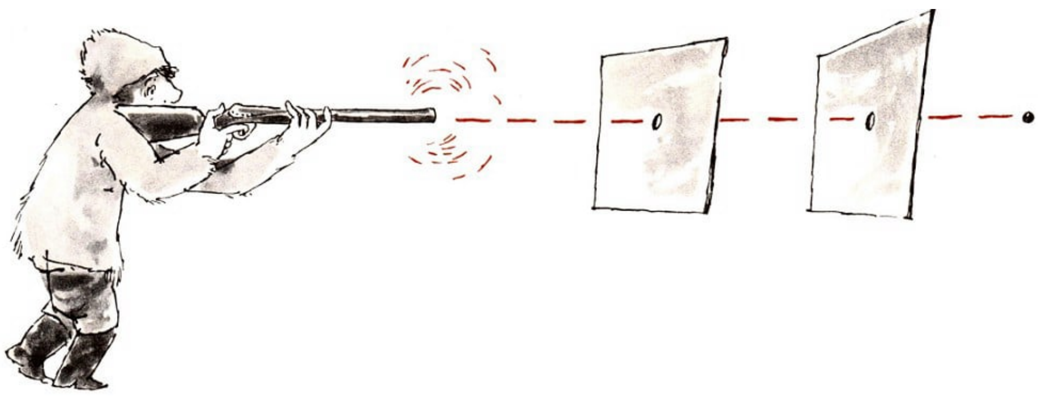
Sitzen wir in einem fahrenden Auto, so ist dies natürlich sehr einfach. Wir brauchen nur auf das Tachometer zu blicken und die Geschwindigkeit abzulesen.

Aber wie erkennen wir die Geschwindigkeit eines an uns vorbeifahrenden Autos oder etwa die eines Geschosses? Bekanntlich gibt es Geräte zur Längenmessung (Lineal, Bandmaß usw.), die wir nur an die zu messende Strecke anzulegen brauchen, um sofort den fraglichen Wert ablesen zu können. Oder es gibt Geräte, mit denen wir die Zeit messen (Uhr, Chronometer). Es gibt noch viele andere nützliche Geräte, aber kein Gerät, auf dem wir die Geschwindigkeit eines Körpers unmittelbar ablesen könnten, der sich an uns vorbeibewegt!

Wie sollte man auch ein solches Gerät an ein vorbeijagendes Auto oder ein fliegendes Geschoss anlegen!
Bis zu einem gewissen Grade können uns allerdings Geräte zur Längen- und Zeitmessung helfen. Mit ihnen messen wir den Weg, den z.B. ein Geschoss zurücklegt, und die Zeit, die es dazu benötigt. Dividieren wir dann den Weg durch die Zeit, so erhalten wir die Geschwindigkeit des Geschosses.

Aber auf diese Art finden wir nur die mittlere Geschwindigkeit, die nichts darüber aussagt, wie der Luftwiderstand die Bewegung nach und nach verlangsamt, so dass am Ende des Fluges die Geschwindigkeit kleiner ist als am Anfang. Wenn wir die Geschwindigkeit eines Geschosses an einem bestimmten Punkt seines Weges feststellen wollen, müssen wir also anders vorgehen.

An diesen Punkt stellen wir eine Wand aus einem dünnen Material. Diese Wand wird mit einer Uhr gekoppelt, welche den Zeitpunkt t_1 registriert, in dem das Geschoss die Wand durchschlägt. In bestimmtem Abstand von der ersten Wand wird eine zweite Wand angebracht, die ebenfalls mit einer Uhr gekoppelt ist, welche den Moment t_2 anzeigt, in dem das Geschoss diese Wand durchstößt.



Die erste Wand befinde sich im Abstand s_1 von der Abschussstelle, die zweite Wand im

Abstand s_2. Dann hat das Geschoss in der Zeit $t_2 - t_1$ den Weg $s_2 - s_1$ zurückgelegt. Das bedeutet, dass für die mittlere Geschwindigkeit v_m des Geschosses während dieser Zeit die Beziehung

$$v_m = \frac{s_2 - s_1}{t_2 - t_1}$$

gilt. Die Messung gibt aber keinen genauen Wert für die Geschwindigkeit im Augenblick t_1. Die Luft bremst nämlich das Geschoss, so dass es auf die zweite Wand mit einer etwas geringeren Geschwindigkeit trifft als auf die erste.

Um den Einfluss des Luftwiderstandes auf das Geschoss zu verringern, müssen wir die Wände näher zusammenstellen. Je näher sie aneinandergerückt werden, desto besser können wir die Geschwindigkeit des Geschosses im Augenblick t_1 feststellen.
Wir nehmen natürlich immer an, dass die Uhren ganz exakt gehen und das Metermaß einwandfrei ist.
Verkleinern wir den Abstand der beiden Wände mehr und mehr, so wird auch die Zeit, die das Geschoss zum Zurücklegen des Weges zwischen den Wänden braucht, immer kürzer. Wir können somit sagen, dass die Geschwindigkeit $v(t_1)$ des Geschosses im Augenblick t_1 gleich

$$v(t_1) = \lim_{t_2 \to t_1} \frac{s_2 - s_1}{t_2 - t_1}$$

ist. Der Grenzwert ist dabei so zu verstehen, dass der Wert s_2 gegen den Wert s_1 strebt (oder, was das gleiche ist, dass t_2 gegen t_1 strebt).

1.54 Geschwindigkeit des radioaktiven Zerfalls

Die verschiedenen radioaktiven Stoffe zerfallen verschieden schnell. Wie können wir nun sagen, ein Stoff zerfalle "langsamer" oder "schneller"?
Und wie lässt sich die "Zerfallsgeschwindigkeit" eines radioaktiven Stoffes in einem bestimmten Augenblick messen?

Die mittlere Zerfallsgeschwindigkeit für, sagen wir, ein Jahr auszurechnen ist leicht. Man misst die in einem Jahr zerfallene Masse in Gramm und dividiert sie durch die Anzahl der Sekunden, die ein Jahr hat. Dies ergibt die mittlere Zerfallsgeschwindigkeit, gemessen in Gramm je Sekunde.

Jedoch ist dieser Weg ungeeignet, wenn wir die Geschwindigkeit in einem bestimmten Moment wissen wollen; denn im Laufe eines Jahres nimmt ja die Masse des radioaktiven Stoffes ständig ab und zerfällt deshalb immer langsamer.
Daher müssen wir, wenn wir die Geschwindigkeit zu einem bestimmten Zeitpunkt genauer messen wollen, die mittlere Geschwindigkeit nicht für ein Jahr errechnen, sondern für einen Monat oder besser noch für einen Tag, eine Stunde, eine Minute usw.

Jedesmal müssen wir die während dieser Zeit zerfallene Stoffmasse (in Gramm) durch die Anzahl der Sekunden in dem entsprechenden Zeitintervall dividieren. Verkleinern wir ständig die Zeitintervalle zwischen zwei Messungen der Stoffmenge, so nähern wir uns einem gewissen Wert des Quotienten.
Diese Zahl ist genau die Zerfallsgeschwindigkeit zu einem bestimmten Zeitpunkt.

Zu diesen Überlegungen stellen wir nun die mathematischen Formeln auf. Wir nehmen an, zur Zeit t_1 sei die Masse des noch nicht zerfallenen radioaktiven Stoffes in unserem Reagenzglas gleich m_1. Nach einer gewissen Zeit, im Moment t_2, hat sich die Masse durch den radioaktiven Zerfall verkleinert und ist gleich m_2.
Somit hat die sich im Reagenzglas befindliche Masse während der Zeit $t_2 - t_1$ um $m_2 - m_1$ geändert. Die Änderung $m_2 - m_1$ ist negativ, da m_1 größer ist als m_2. Der Quotient

$$u_m = \frac{m_2 - m_1}{t_2 - t_1}$$

ist die mittlere Geschwindigkeit, mit der sich die Masse des radioaktiven Stoffes im Reagenzglas während des betrachteten Zeitintervalls geändert hat. Er gibt also die mittlere Zerfallsgeschwindigkeit an.
Je kleiner das Zeitintervall $t_2 - t_1$ ist, desto genauer drückt dieser Quotient die Zerfallsgeschwindigkeit in einem bestimmten Moment aus. Wir können also sagen, dass die Zerfallsgeschwindigkeit $u(t_1)$ im Augenblick t_1 gleich

$$u(t_1) = \lim_{t_2 \to t_1} \frac{m_2 - m_1}{t_2 - t_1}$$

ist. Dabei ist der Grenzwert so zu verstehen, dass t_2 gegen t_1 strebt, was durch die Schreibweise $t_2 \to t_1$, angedeutet ist, Auf dieselbe Art lässt sich auch die Geschwindigkeit einer chemischen Reaktion in einem bestimmten Augenblick errechnen,

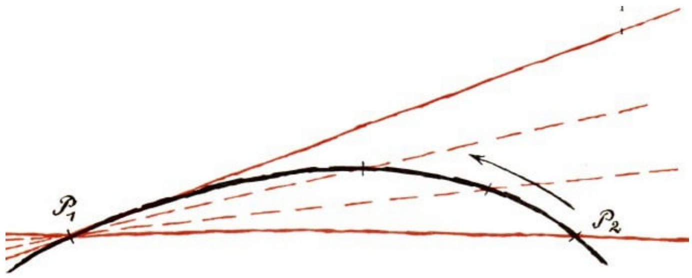
1.55 Wie zeichnet man eine Tangente?

Wenn uns jemand diese Frage stellt, werden wir uns natürlich sofort daran erinnern, wie wir in der Schule gelernt haben, eine Tangente an einen Kreis zu legen. Jetzt ist aber eine Tangente an eine beliebige Kurve gemeint und nicht nur an einen Kreis, und wie eine solche Tangente gezeichnet wird oder wie sie definiert ist, haben wir noch nicht gelernt.

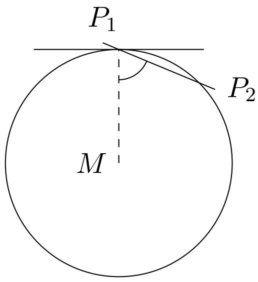
Eine Tangente an eine Kurve kann nicht einfach definiert werden, indem wir sagen, sie sei eine Gerade, die mit der Kurve nur einen einzigen Punkt gemeinsam hat.
Beispielsweise haben Parabel und Parabelachse zwar auch einen gemeinsamen Punkt, nämlich ihren Schnittpunkt, aber wir können nicht behaupten, dass diese Achse die Parabel berührt,

Was versteht man nun unter der Tangente an eine Kurve, und wie zeichnet man sie? Wir wollen versuchen, auf diese Fragen zu antworten.

Wir nehmen auf einer Kurve einen festen Punkt P_1 an und legen durch ihn eine Sekante P_1P_2. Wenn sich der Punkt P_2 längs der Kurve dem Punkt P_1 nähert, dreht sich die Sekante um den Punkt P_1 und nähert sich dabei mehr und mehr einer gewissen Geraden.

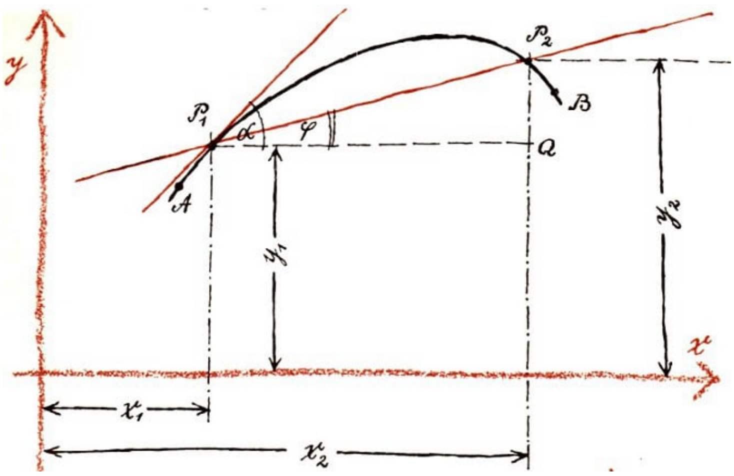


Diese Gerade ist die Tangente an die Kurve im Punkt P_1. Für den Kreis stimmt diese Definition mit der uns bekannten überein.
Wenn sich der Punkt P_2 dem Punkt P_1 nähert, strebt der Winkel $\angle MP_1P_22$ gegen einen rechten Winkel, und deshalb steht die Tangente an einen Kreis senkrecht auf dem Radius zum Berührungspunkt.



Die Tangente an eine Kurve im Punkt P_1 ist also die Gerade, der sich die Sekante P_1P_2 nähert, wenn der Punkt P_2 längs der Kurve gegen den Punkt P_1 strebt.

Jetzt ist es auch nicht mehr schwierig, die Lage der Tangente mit Hilfe einer Formel zu beschreiben. Wir nehmen dazu an, die Kurve AB sei die graphische Darstellung der Funktion $f(x)$. Die Ordinaten der Punkte P_1 und P_2 bezeichne wir mit y_1 bzw. y_2 die Abszissen mit x_1 bzw. x_2.
Betrachten wir dann das Dreieck P_1P_2Q mit der Hypotenuse P_1P_2 und den zu den Koordinatenachsen parallelen Katheten, so können wir leicht den Winkel φ zwischen der Sekante P_1P_2 und der positiven Richtung der x-Achse bestimmen.



Für ihn gilt die Beziehung

$$\tan\varphi = P_2Q : P_1Q$$

Aus der Abbildung lesen wir ab, dass

$$P_2Q = y_2 - y_1 \qquad , \qquad P_1Q = x_2 - x_1$$

ist. Damit ergibt sich

$$\tan\varphi = \frac{y_2 - y_1}{x_2 - x_1}$$

Lassen wir nun den Punkt P_2 längs der Kurve gegen den Punkt P_1 streben, so nähert sich die Sekante P_1P_2 der Tangente, indem sie sich um P_1 dreht. Wir erhalten beim Grenzübergang den Tangens des Winkels α, unter dem die Tangente die x-Achse schneidet:

$$\tan\alpha = \lim_{x_2 \to x_1} \frac{y_2 - y_1}{x_2 - x_1}$$

Der Grenzwert ist dabei so zu verstehen, dass der Punkt P_2 gegen den Punkt P_1, d.h., dass x_2 gegen x_1 strebt.

1.56 Die Ableitung

Wir haben eben einige Aufgaben aus Physik und Geometrie betrachtet. Obwohl sie sich äußerlich voneinander unterscheiden, haben sie doch etwas gemeinsam. In den ersten beiden Aufgaben bestand das Gemeinsame darin, dass wir die Geschwindigkeit suchten, mit der sich eine gewisse Größe ändert.
So ist die Geschwindigkeit, mit der sich ein Körper bewegt, die Änderung des von ihm zurückgelegten Weges in der Zeiteinheit, und die Zerfallsgeschwindigkeit eines radioaktiven Stoffes ist gleich der Änderung seiner Masse in der Zeiteinheit.

Aber auch im dritten Beispiel trat eine gewisse Geschwindigkeit auf: Der Tangens des Neigungswinkels der Tangente bezeichnet die Geschwindigkeit, mit der sich die Ordinate ändert, wenn sich der entsprechende Punkt auf der x-Achse bewegt. Der Quotient

$$\frac{y_2 - y_1}{x_2 - x_1}$$

ist nämlich die mittlere Geschwindigkeit, mit der sich die Ordinate ändert, wenn man vom Punkt P_1 zum Punkt P_2 übergeht; der Grenzwert dieses Quotienten, also $\tan\alpha$, ist die Geschwindigkeit, mit der sich die Ordinate in einem bestimmten Punkt (in diesem Fall im Berührungspunkt von Kurve und Tangente) ändert.

In allen drei Beispielen haben wir also die Geschwindigkeit in einem bestimmten Punkt betrachtet. Damit ist erklärt, weshalb wir bei der Bestimmung dieser auf den ersten Blick so verschiedenen Größen auf sehr ähnliche Formeln stießen.

Rein mathematisch lässt sich die Geschwindigkeit einer Änderung folgendermaßen definieren. Gegeben sei eine Funktion $y = f(x)$.
Ihre Werte in den Punkten P_1 und P_2 bezeichnen wir mit y_1 bzw. y_2. Dann gibt die Differenz $y_2 - y_1$ an, wie sich der Wert der Funktion beim Übergang von x_1 nach x_2 ändert. Der Quotient

$$\frac{y_2 - y_1}{x_2 - x_1}$$

ist gleich der mittleren Geschwindigkeit, mit der sich $y = f(x)$ im Intervall von x_1 bis x_2 ändert. Verkleinern wir dieses Intervall, indem wir x_2 gegen x_1 streben lassen, so erhalten wir die Geschwindigkeit der Änderung der Funktion im Punkt x_1. Sie ist gleich

$$\lim_{x_2 \to x_1} \frac{y_2 - y_1}{x_2 - x_1}$$

Diese Geschwindigkeit der Änderung heißt die Ableitung der Funktion $y = f(x)$ nach dem Argument x an der Stelle x_1; wir wollen sie mit $f'(x_1)$ bezeichnen.

Bei dieser Bezeichnung wird ausdrücklich angegeben, in welchem Punkt die Geschwindigkeit der Änderung (d.h. die Ableitung) betrachtet wird. Es gibt auch noch andere Bezeichnungen für die Ableitung. Darauf wollen wir hier aber nicht eingehen.
Die Ableitung $f'(x)$ kann natürlich in verschiedenen Punkten berechnet werden, so dass auch sie eine Funktion von x ist. Damit ist klar, dass die vorhin betrachteten Beispiele aus Physik und Geometrie mit Hilfe der Ableitung formuliert werden können:

Die Geschwindigkeit $v(t)$ einer Bewegung ist gleich der Ableitung des Weges $s(t)$ nach der Zeit t:

$$v(t) = s'(t) \tag{9}$$

Die Geschwindigkeit $u(t)$ des radioaktiven Zerfalls ist gleich der Ableitung der Masse $m(t)$ des radioaktiven Stoffes nach der Zeit t:

$$u(t) = m'(t) \tag{10}$$

Schließlich ist der Tangens des Neigungswinkels x der Tangente an eine Kurve $y = f(x)$ im Punkt mit der Abszisse x gleich der Ableitung der Funktion $f(x)$:

$$\tan \alpha = f'(x) \tag{11}$$

1.57 Die Ableitung eines Polynoms

Aus den obigen Überlegungen folgt, dass viele Probleme aus Mathematik, Physik und anderen Wissenschaften gelöst werden, indem die Ableitung einer Funktion bestimmt wird.
Das Bestimmen der Ableitung wird Differentiation oder Differenzieren genannt.

Wir wollen jetzt an einem Beispiel zeigen, wie man die Ableitung unmittelbar berechnen kann. Wir nehmen dazu die Funktion $y = x^3$. Der zu untersuchende Quotient hat hier die Gestalt

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{x_2^3 - x_1^3}{x_2 - x_1} = \frac{(x_2 - x_1)(x_2^2 + x_2 x_1 + x_1^2)}{x_2 - x_1} = x_2^2 + x_2 x_1 - x_1^2$$

Lassen wir nun x_2 gegen x_1 streben, so nähert sich der letzte Ausdruck dem Wert $x_1^2 + x_2^2 + x_1^2 = 3x_1^2$. Also hat die Ableitung der Funktion $y = x^3$ im Punkt $x = x_1$ den Wert $3x_1^2$. Man schreibt dafür auch

$$(x^3)'\Big|_{x=x_1} = 3x_1^2$$

Da dieses Ergebnis aber nicht nur für den Punkt x_1, sondern für alle Punkte gilt, schreiben wir

$$(x^3)' = 3x^2$$

Auf dieselbe Art finden wir sofort die Ableitungen der Funktionen $y = x^2$ und $y = x$. Sie lauten

$$(x^2)' = 2x \qquad , \qquad (x)' = 1$$

Allgemein gilt für die Ableitung der Funktion $y = x^n$ die Formel

$$(x^n)' = n \cdot x^{n-1} \tag{12}$$

Ist n eine natürliche (d.h. ganze positive) Zahl, so lässt sich diese Formel auf die gleiche Art nachprüfen, wie wir vorhin die Ableitung von x^3 berechnet haben.
In der höheren Mathematik wird bewiesen, dass die Beziehung (12) für jedes (also nicht nur für ein natürliches) n gültig ist. Wir erwähnen noch, dass die Ableitung der Zahl 1 (oder überhaupt einer konstanten Größe) gleich Null ist. Dies folgt leicht aus (12), ist aber auch ohnehin klar, da die Geschwindigkeit, mit der sich eine konstante Größe ändert, offenbar gleich Null ist.

Wir wollen noch darauf hinweisen, dass die Ableitung die beiden einfachen, aber wichtigen Eigenschaften besitzt:
Ein konstanter Faktor kann vor das Differentiationszeichen gezogen werden:

$$[kf(x)]' = k \cdot f'(x)$$

Die Ableitung einer Summe zweier (oder mehrerer) Funktionen ist gleich der Summe aus den Ableitungen der einzelnen Summanden:

$$[f_1(x) + f_2(x)]' = f_1'(x) + f_2'(x)$$

Die Gültigkeit dieser Regeln können wir mit Hilfe von (9) leicht nachprüfen, und zwar genauso, wie wir es schon bei den Integralen getan haben.
Jetzt können wir die Ableitung beliebiger Polynome bestimmen. Beispielsweise ist

$$\begin{aligned}(3x^3 - 2x^2 + 5x - 4)' &= (3x^3)' + (-2x^2)' + (5x)' + (-4)' \\ &= 3(x^3)' - 2(x^2)' + 5(x)' - 4(1)' = 3 \cdot 3x^2 - 2 \cdot 2x + 5 \cdot 1 - 4 \cdot 0 \\ &= 9x^2 - 4x + 5\end{aligned}$$

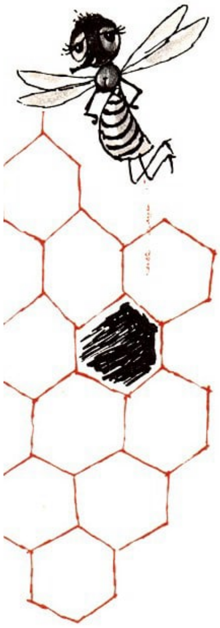
Allgemein hat ein Polynom a.

$$f(x) = a_0 x^n + a_1 x^{n-1} + ... + a_{n-1}x + a_n$$

vom Grade n die Ableitung

$$f(x)' = na_0 x^{n-1} + (n-1)a_1 x^{n-2} + ... + a_{n-1}$$

1.58 Bienen als Mathematiker

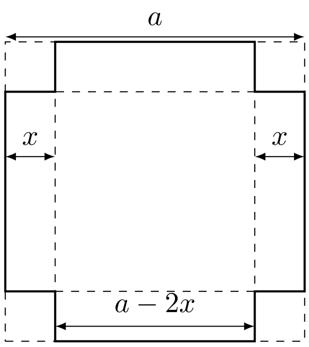


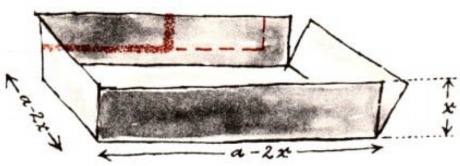
Der berühmte russische Mathematiker Pafnuti Lwowitsch Tschebyschew (1821-1894) schrieb in seiner Arbeit "Das Zeichnen geographischer Karten", dass diejenigen wissenschaftlichen Methoden besonders wichtig sind, mit denen Aufgaben gelöst werden könne, die für die praktische Tätigkeit des Menschen große Bedeutung haben.
Es ist z.B. wichtig, wie man disponieren muss, um mit seinen Mitteln einen möglichst großen Nutzen zu erzielen. So bemüht sich beispielsweise ein Metallarbeiter, aus dem vorhandenen Metallstück so viele Einzelstücke wie möglich zu erhalten; ein Zuschneider in einer Schuhfabrik versucht, aus dem Stück Leder möglichst viele Schuhe herauszuwirtschaften; ein Technologe ist bemüht, in einer Fabrik die Maschinen so aufzustellen, dass die Bearbeitung der einzelnen Werkstücke in möglichst kurzer Zeit erfolgt.

Aber nicht nur den Menschen gelingt es, solche Probleme zu lösen. Die Bienen lösen unbewusst eines dieser Probleme - sie sind bemüht, den Waben eine solche Form zu geben, dass bei gegebenem Volumen der Waben so wenig Wachs wie möglich verbraucht wird.
Obwohl die Bienen nichts von Mathematik wissen, können sie das Problem exakt lösen, und zwar mit Hilfe ihres Instinkts.

Der Mensch unterscheidet sich von den Bienen dadurch, dass er nicht nach seinem Instinkt, sondern nach seinem Verstand handelt. Wenigstens hat er dazu die Möglichkeit, und er sollte sie nutzen!
Er kann also mit Überlegung vorgehen, so dass sich der schlechteste Baumeister vor der besten Biene dadurch auszeichnet, dass er die Zelle in seinem Kopf gebaut hat, bevor er sie in Wachs baut.
Um eine Vorstellung davon zu bekommen, wie Mathematiker an solche Probleme herangehen, wollen wir eines von ihnen betrachten.

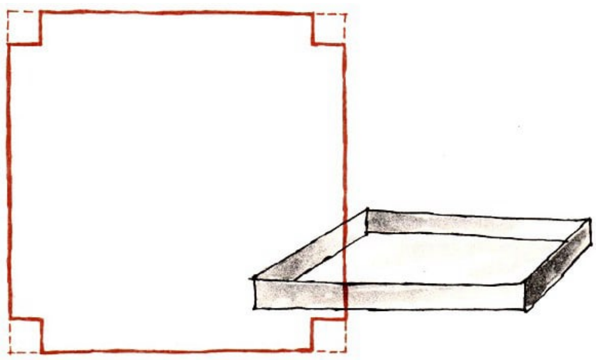
1.59 Wie bastelt man eine Schachtel mit möglichst großem Volumen?





Vor uns liege ein quadratisches Stück Pappe. Die Seitenlänge des Quadrats sei a. Aus diesem Stück Pappe wollen wir eine Schachtel ohne Deckel anfertigen, und zwar soll das Volumen der Schachtel möglichst groß sein.
Aus dem Stück Pappe schneiden wir an den Ecken kleine Quadrate aus und knicken die Pappe längs der gestrichelten Linien.

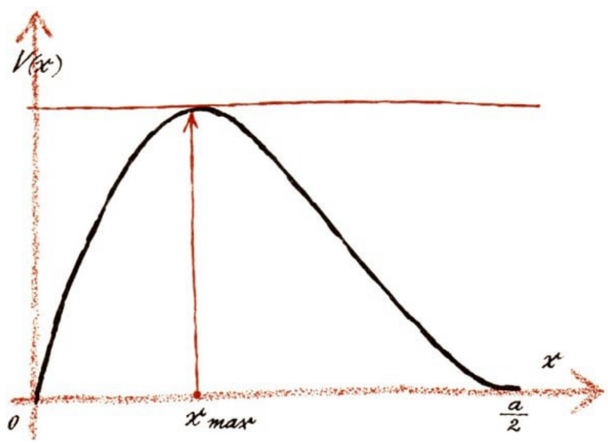
So erhalten wir eine Schachtel. Aber hat sie das größtmögliche Volumen? Das ist von der Größe der an den Ecken ausgeschnittenen Quadrate abhängig. Sind diese Quadrate sehr klein, so ist die Schachtel sehr flach und kann nicht viel aufnehmen. Das gleiche ist der Fall, wenn diese Quadrate sehr groß sind, die Schachtel also sehr hoch ist.



Wir wollen nun versuchen festzustellen, wie groß $x_{\max}$ sein muss, damit die Schachtel das größtmögliche Volumen $V(x_{\max})$ hat. Ein Blick auf die Abbildungen zeigt, dass

$$V(x) = x(a-2x)^2 = 4x^3 - 4ax^2 + a^2x$$

ist. Weiter unten ist diese Funktion graphisch dargestellt. Dabei haben wir die Kurve nur zwischen x_0 und $x = \frac{a}{2}$ gezeichnet, denn der von uns gesuchte Wert $x_{\max}$ muss dazwischen liegen. Wir können nämlich offenbar aus dem Stück Pappe, das die Seitenlänge a hat, an den Ecken keine Quadrate ausschneiden, deren Seiten länger als $\frac{a}{2}$ sind.



Die graphische Darstellung zeigt, dass in dem Punkt, in dem das Volumen seinen größten Wert erreicht, die Tangente an die Kurve horizontal, also parallel zur x-Achse verläuft, d.h., der Winkel zwischen der Tangente und der positiven Richtung der x-Achse gleich Null ist.

Das bedeutet, dass die Ableitung von $V(x)$ in diesem Punkt gleich Null ist. Somit müssen wir, um den Wert $x_{\max}$ zu finden, bei dem das Volumen der Schachtel am größten ist, diejenigen Werte von x bestimmen, für die die Ableitung von $V(x)$ gleich Null ist.
Unter diesen Werten muss sich $x_{\max}$ befinden, Wir differenzieren $V(x)$ nach x und erhalten

$$V'(x) = 12x^2 - 8ax + a^2$$

Dann setzen wir

$$12x^2 - 8ax + a^2 = 0$$

oder gleichbedeutend

$$x^2 - \frac{2}{3}ax + \frac{a^2}{12} = 0$$

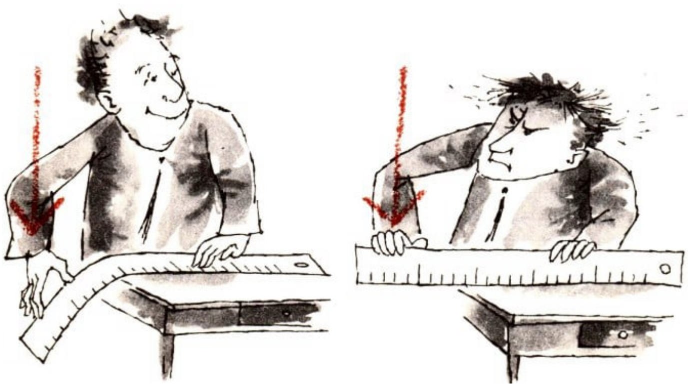
und bestimmen die beiden Wurzeln dieser Gleichung:

$$x_{1;2} = \frac{a}{3} \pm \sqrt{\frac{a^2}{9} - \frac{a^2}{12}} = \frac{a}{3} \pm \frac{a}{6}; \qquad x_1 = \frac{a}{2}, \quad x_2 = \frac{a}{6}$$

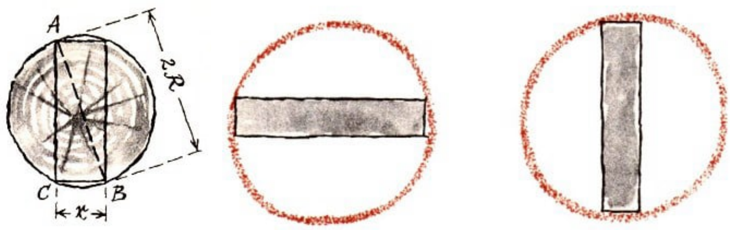
Selbstverständlich kommt die Wurzel $x_1 = \frac{a}{2}$ nicht in Frage, denn schneiden wir Quadrate mit dieser Seitenlänge aus, so bleibt von dem Stück Pappe nichts übrig. Also ist $x_{\max} = \frac{a}{6}$. Schneiden wir kleine Quadrate mit dieser Seitenlänge aus, so erhalten wir eine Schachtel mit dem Volumen $V_{\max} = \frac{2a^3}{27}$.
Aus dem gegebenen Stück Pappe eine Schachtel mit noch größerem Volumen zu konstruieren ist unmöglich.

1.60 Balken größter Festigkeit

Das Grundelement jeder Baukonstruktion ist der Balken. Die Festigkeit eines Balkens hängt von der Form seines Querschnitts ab. Dabei spielt die Höhe des Querschnitts eine bedeutend größere Rolle als die Breite.



Wir alle wissen, wieviel schwieriger es ist, ein auf der Kante stehendes als ein auf der Seite liegendes Lineal zu verbiegen. Berechnungen zeigen, dass die Festigkeit eines Balkens mit rechteckigem Querschnitt proportional der Breite a und dem Quadrat der Höhe h des Balkens ist.
Sie ist also gleich kah^2, wobei der Proportionalitätsfaktor k von der Länge des Balkens, seinem Material usw. abhängt.



Holzbalken werden im allgemeinen aus runden Stämmen angefertigt. Das Problem besteht dabei darin, wie man einen Baumstamm vom Radius R zuschneiden muss, um einen Balken größter Festigkeit zu erhalten.
In der Abbildung ist ein Stamm im Querschnitt zu sehen. Wie wir schon erwähnten, ist die Festigkeit des Balkens eine Funktion seiner Breite. Wählen wir eine zu große Breite (etwa fast den Durchmesser des Stammes), so hat der Balken eine zu kleine Höhe und damit eine zu geringe Festigkeit.

Dieselbe Wirkung erzielen wir, wenn wir die Breite zu klein ansetzen. Wir nehmen wieder die Mathematik zu Hilfe, um herauszufinden, in welchem Verhältnis Breite und Höhe stehen müssen, damit die Festigkeit des Balkens am größten ist. Dazu fassen wir die Festigkeit als Funktion der Breite x des Balkens auf.

Aus dem Dreieck ABC ersehen wir, dass ein Balken mit der Breite x die Höhe $\sqrt{4R^2 - x^2}$ hat (Satz des Pythagoras!), wenn R der Radius des Baumstammes ist. Deshalb ist die Festigkeit y dieses Balkens gleich

$$y = kx(4R^2 - x^2) = 4R^2kx - kx^3$$

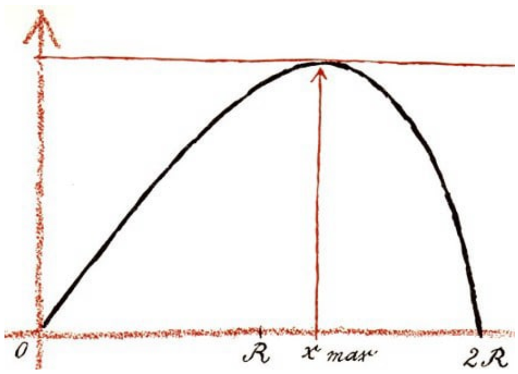
Die Ableitung y' dieser Funktion ist $y' = 4R^2k - 3kx^2$. Sie verschwindet (d.h. wird Null) für

$$x_{1;2} = \pm\frac{2R\sqrt{3}}{3}$$

Da die Breite eine positive Zahl sein muss, erhalten wir den Balken größter Festigkeit, wenn die Breite gleich $a = \frac{2R\sqrt{3}}{3}$ ist. Die Höhe dieses Balkens lässt sich mit Hilfe der Formel $h = \sqrt{4r^2 - a^2}$ bestimmen:

$$h = \sqrt{4R^2 - \left(\frac{2R\sqrt{3}}{3}\right)^2} = \frac{2R\sqrt{6}}{3}$$

Der Quotient $\frac{h}{a}$ ist gleich $\sqrt{2} \approx \frac{7}{5}$.



1.61 Hauptsatz der Differential- und Integralrechnung

Zwischen Differentiation und Integration besteht ein enger Zusammenhang: Die Formel (3) zeigt, dass sich der Weg durch Integration der Geschwindigkeit in einem bestimmten Punkt ergibt, und die Formel (9) besagt, dass wir die Geschwindigkeit aus dem Weg durch Differentiation erhalten.

Das bringt uns natürlich auf den Gedanken, ob nicht auch Differentiation und Integration so ähnlich miteinander verknüpft sind wie etwa Addition und Subtraktion, Multiplikation und Division, Potenzieren und Radizieren, d.h., ob nicht auch diese Operationen zueinander invers sind. Das ist tatsächlich der Fall.

Benutzen wir zum Beispiel die Beziehung $v(t) = s'(t)$, so können wir der Formel (3) die Gestalt

$$s = \int\limits_0^T s'(t)dt$$

geben. Hierbei ist s der Weg. den der Körper seit dem Zeitpunkt $t = 0$ durchlaufen hat.
Es kann nun vorkommen, dass der zurückgelegte Weg nicht vom Moment $t = 0$, sondern von einem noch früheren Zeitpunkt an gerechnet wird (bei einem Auto zum Beispiel nicht vom Beginn einer Fahrt, sondern etwa vom Zeitpunkt der Auslieferung des Autos aus der Fabrik).

Dann ist es zweckmäßiger, den Weg s als Differenz $s(T) - s(0)$ zu schreiben, nämlich als Differenz aus dem zur Zeit $t = T$ und dem zur Zeit $t = 0$ zurückgelegten Weg. Die Gleichung (3) geht dann über in

$$s(T) - s(0) = \int\limits_0^T s'(t)dt$$

Für zwei beliebige Zeitpunkte $t = a$ und $t = b$ gilt die Beziehung

$$s(b) - s(a) = \int\limits_a^b s'(t)dt$$

Allgemein genügt jede (differenzierbare) Funktion $F(x)$ dem sogenannten Hauptsatz der Differential- und Integralrechnung:

$$F(b) - F(a) = \int\limits_a^b F'(x)dx$$

Diese Formel wird manchmal auch nach dem berühmten englischen Mathematiker und Naturwissenschaftler Isaac Newton (1643-1772) und dem deutschen Gelehrten Gottfried Wilhelm Leibniz (1646-1716) als Newton-Leibnizsche Formel bezeichnet.
Beide fanden sie fast gleichzeitig Ende des 17. Jahrhunderts, etwa 70 Jahre, nachdem Kepler 1615 seine Arbeit über die "Neue Stereometrie der Fässer" geschrieben hatte.

Es muss noch erwähnt werden, dass dieser Hauptsatz in geometrischer Form schon von Newtons Lehrer Isaac Barrow (1630-1677) im Jahre 1670 ausgesprochen worden war. Er zeigte, dass die Berechnung von Flächeninhalten die zur Konstruktion von Tangenten inverse Operation ist.

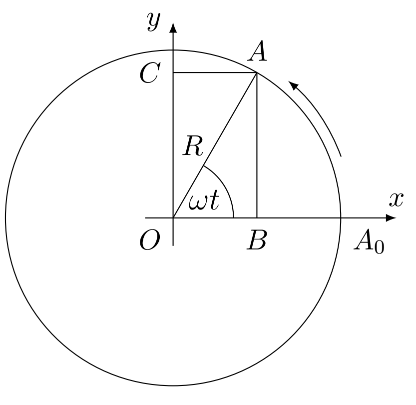
Die Bedeutung des Hauptsatzes besteht in folgendem: Kennen wir eine Funktion $F(x)$, deren Ableitung gleich $f(x)$ ist, für die also $F'(x) = f(x)$ gilt, so lässt sich das Integral $\int\limits_a^b f(x)dx$ ganz leicht berechnen - es ist einfach gleich der Differenz der Werte von $F(x)$ in den Punkten $x = b$ und $x = a$.
Jede Funktion $F(x)$, für die $F'(x) = f(x)$ ist, heißt Stammfunktion der Funktion $f(x)$. Ist also $F(x)$ eine Stammfunktion von $f(x)$, so ist $f(x)$ die Ableitung von $F(x)$.

Die Berechnung von Integralen lässt sich damit im Grunde genommen auf die Bestimmung von Stammfunktionen zurückführen.
Die Bestimmung einer Stammfunktion ist das zur Differentiation inverse Problem. Je mehr Funktionen wir differenzieren, desto mehr Stammfunktionen lernen wir damit kennen und desto mehr Integrale können wir berechnen.

Bis jetzt sind wir nur imstande, Polynome zu differenzieren. Das genügt aber schon, um beliebige Polynome zu integrieren (ohne zu den oben angewendeten geometrischen Methoden greifen zu müssen). In vielen Problemen stoßen wir jedoch auf andere Funktionen als Polynome.
Deshalb wollen wir uns jetzt einmal ansehen, wie die trigonometrischen Funktionen und die Exponentialfunktion differenziert werden.

1.62 Die Ableitungen von Sinus und Cosinus

Die Ableitungen der trigonometrischen Funktionen lassen sich am leichtesten berechnen, wenn man von physikalischen Überlegungen ausgeht. Wir betrachten einen Punkt, der sich mit der Geschwindigkeit ωR auf einem Kreis vom Radius R bewegt.

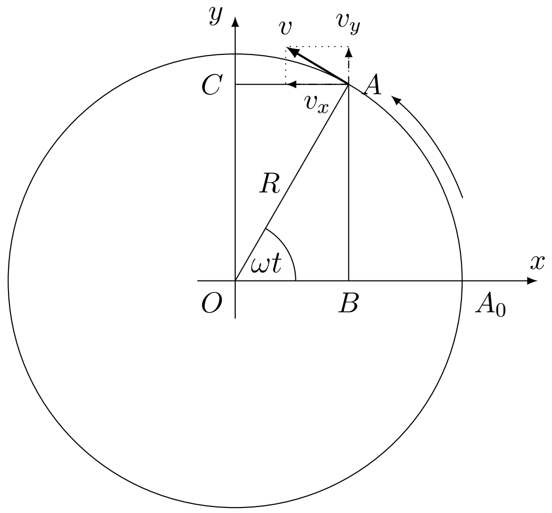


Der Punkt möge sich zur Zeit $t = 0$ in der Lage A_0 befunden haben. Nach der Zeit t (in Sekunden) hat der Punkt einen Weg der Länge ωRt zurückgelegt und befindet sich in A.
Der Bogen A_0A hat die Länge ωRt. Das bedeutet, dass der Winkel AOA_0 gemessen

im Bogenmaß, gleich ωt ist. Daher sind die Koordinaten des Punktes A gleich

$$x = R\cos\omega t \qquad , \qquad y = R\sin\omega t$$

wie man am Dreieck AOB leicht ablesen kann. Mit anderen Worten, die Projektion B des Punktes A auf die x-Achse bewegt sich nach dem Gesetz $x = R\cos\omega t$, die Projektion C von A auf die y-Achse nach dem Gesetz $y = R\sin\omega t$.



Wir wollen nun die Geschwindigkeiten dieser Bewegungen bestimmen. Dazu zerlegen wir die Geschwindigkeit des Punktes A in zwei Komponenten, in eine horizontale und eine vertikale. Der Vektor der Geschwindigkeit v von A (er hat die Länge ωR) liegt auf der Tangente an den Kreis im Punkt A.
Er bildet deshalb mit der positiven Richtung der x-Achse den Winkel $\omega t + \frac{\pi}{2}$, mit der positiven Richtung der y-Achse den Winkel ωt. Folglich ist seine Projektion auf die x-Achse (d.h. die Geschwindigkeit Punktes B) gleich

$$v_x = \omega R\cos\left(\omega t + \frac{\pi}{2}\right) = -\omega R\sin\omega t$$

und seine Projektion auf die y-Achse (d.h. die Geschwindigkeit des Punktes C) ist gleich

$$v_y = \omega R\cos\omega t$$

Damit haben wir gezeigt, dass die Geschwindigkeit der Bewegung, die nach dem Gesetz $x = R\cos\omega t$ vor sich geht, gleich $v_x = -\omega R\sin\omega t$ ist. Da die Geschwindigkeit gleich der Ableitung des Weges nach der Zeit ist, muss

$$(R\cos\omega t)' = -\omega R\sin\omega t$$

oder für $R = 1$

$$(cos\omega t)' = -\omega\sin\omega t \tag{13}$$

gelten. Genauso lässt sich, wenn man die Bewegung von C betrachtet, zeigen, dass

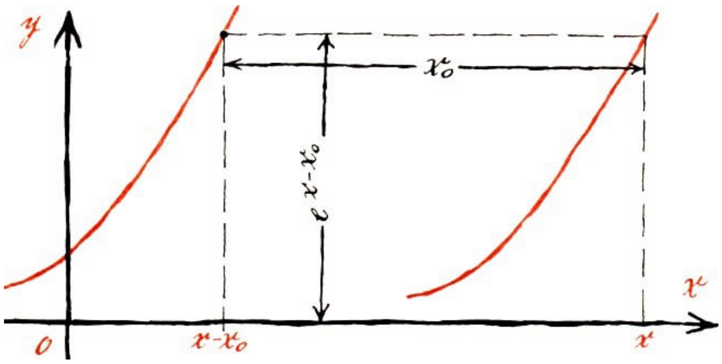
$$(\sin\omega t)' = \omega\cos\omega t \tag{14}$$

gilt. Insbesondere ergibt sich für $\omega = 1$

$$(\cos t)' = -\sin t \qquad , \qquad (\sin t)' = \cos t$$

1.63 Die Ableitung der Exponentialfunktion

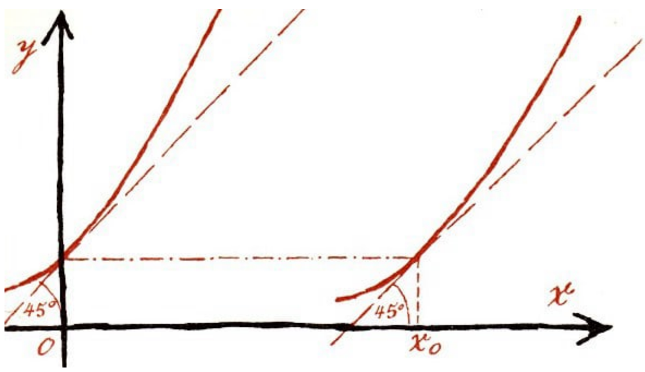
Wir wollen nun die Exponentialfunktion $y = e^x$ differenzieren.
Uns ist schon bekannt, dass die Tangente an die Kurve $y = e^x$ im Schnittpunkt dieser Kurve und der y-Achse mit der positiven Richtung beider Koordinatenachsen einen Winkel von $45°$ bildet.



Wenn wir uns an die geometrische Bedeutung der Ableitung erinnern, so können wir sagen, dass die Ableitung der Funktion $y = e^x$ im Punkt $x = 0$ gleich $\tan 45°$, also gleich 1 ist:

$$(e^x)'|_{x=0} = 1$$

Um die Ableitung der Funktion $y = e^x$ in einem beliebigen Punkt x_0, zu bestimmen, verschieben wir die Kurve um die Strecke x_0, parallel zur x-Achse.
Nach dieser Verschiebung ist die Ordinate im Punkt x nicht gleich e^x, sondern gleich e^{x-x_0}, d.h., die verschobene Kurve ist die graphische Darstellung der Funktion $y = e^{x-x_0}$. Bei der Verschiebung geht die Tangente an die Kurve $y = e^x$ im Punkt $x = 0$ in die Tangente an die verschobene Kurve $y = e^{x-x_0}$ im Punkt $x = x_0$ über. Also schließt auch die Tangente an die Kurve $y = e^{x-x_0}$ im Punkt x_0, mit der positiven Richtung der x-Achse einen Winkel von $45°$ ein.



Damit ist

$$(e^{x-x_0})'\Big|_{x=x_0} = 1$$

Jetzt finden wir sofort die Ableitung der Funktion $y = e^x$ im Punkt $x = x_0$. Da der konstante Faktor e^{x_0} vor das Differentiationszeichen gezogen werden kann, erhalten wir

$$(e^x)'|_{x=x_0} = (e^{x_0}e^{x-x_0})'\Big|_{x=x_0} = e^{x_0}(e^{x-x_0})'\Big|_{x=x_0} = e^{x_0} \cdot 1 = e^{x_0}$$

Damit ist bewiesen, dass die Ableitung der Funktion $y = e^x$ im Punkt $x = x_0$ gleich e^{x_0} ist. Da x_0 aber ein beliebiger Punkt war, gilt für alle Punkte der Kurve die Beziehung

$$(e^x)' = e^x$$

Dass auch die Formel

$$(e^{cx})' = ce^{cx} \tag{15}$$

gilt, ist leicht zu beweisen.

1.64 Der radioaktive Zerfall

Viele physikalische Gesetze verknüpfen eine gewisse Größe und die Geschwindigkeit ihrer Änderung miteinander. Wir wollen beispielsweise wieder den radioaktiven Zerfall betrachten.
Die Zerfallsgeschwindigkeit ist um so größer, je mehr radioaktiver Stoff vorhanden ist. Das ist auch verständlich.

Zerfallen nämlich von jedem Gramm radioaktiven Stoffes je Sekunde 0,0001 g, so zerfallen von 2 g des Stoffes je Sekunde 0,0002 g, von 7 g je Sekunde 0,0007 g usw. Mit anderen Worten, die Zerfallsgeschwindigkeit [wir wollen sie wieder mit u bezeichnen, vgl. Formel (10)] ist der Masse m des vorhandenen radioaktiven Stoffes direkt proportional:

$$u = -km \tag{16}$$

Hierbei ist k ein positiver Proportionalitätsfaktor, die sogenannte Zerfallskonstante, und das Minuszeichen steht deshalb, weil die Menge beim Zerfall abnimmt, d.h. die Geschwindigkeit negativ ist. Dieses Gesetz, das die Masse des radioaktiven Stoffes und die Zerfallsgeschwindigkeit miteinander verknüpft, gilt aber nur, wenn die Masse des radioaktiven Stoffes nicht zu groß ist und keine Kettenreaktion hervorrufen kann.

Auf den ersten Blick scheint es, dass wir mit Hilfe der Gleichung (16) gar nichts berechnen können, denn sie ist eine Gleichung in den beiden Unbekannten u und m (die Zerfallskonstante k wird für jeden radioaktiven Stoff aus Versuchen bestimmt), und um zwei Unbekannte finden zu können, benötigen wir stets zwei Gleichungen.
Jedoch ist die zweite Gleichung leicht gefunden.
Da u die Geschwindigkeit ist, mit der sich die Masse m ändert, gilt offenbar $u = m'$. Somit lässt sich das Gesetz des radioaktiven Zerfalls, d.h. die Formel (16), in der Gestalt

$$m' = -km \tag{17}$$

schreiben. Wir haben damit eine Gleichung für die Unbekannte m erhalten. Nur haben wir eine solche Gleichung in der Schule noch nicht gelöst. Diese Gleichung verknüpft die Größe m mit der Geschwindigkeit der Änderung von m, also mit ihrer eigenen Ableitung. Gleichungen, in denen außer der gesuchten Größe auch noch Ableitungen dieser Größe auftreten, heißen Differentialgleichungen.

Wir können leicht nachprüfen, dass die Funktion $m = Ce^{-kt}$ (C eine beliebige Konstante) eine Lösung der Differentialgleichung (17) ist, d.h., wenn wir diese Funktion in (17) für m einsetzen, so wird (17) eine Identität:

$$m' = (Ce^{-kt})' = -Cke^{-kt} = -km$$

Man kann zeigen, dass es keine anderen Funktionen außer der Funktion $m(t) = Ce^{-kt}$ gibt, die der Gleichung (17) genügen, d.h., jede Lösung von (17) hat die Form $m(t) = Ce^{-kt}$.
Dies ist das Gesetz, nach dem die Masse eines radioaktiven Stoffes im Laufe der Zeit abnimmt.

Eine Frage blieb bis jetzt unbeantwortet: Wie groß ist die Konstante C?
Diese Frage ist leicht zu beantworten. Aus der Beziehung $m(t) = Ce^{-kt}$ finden wir, wenn wir $t = 0$ setzen, dass die Masse des radioaktiven Stoffes zur Zeit $t = 0$ (also zu Beginn des Zerfalls) gleich $m(0) = Ce^0 = C$ ist.
Also ist die Konstante C gleich der Anfangsmasse, die wir mit m_0 bezeichnen wollen. Ersetzen wir nun C durch m_0, so erhalten wir das Gesetz des radioaktiven Zerfalls in der endgültigen Form

$$m(t) = m_0 e^{-kt} \tag{18}$$

Wir werden nun ermitteln, wieviel Jahre ein radioaktiver Stoff braucht, um sich auf die Hälfte zu verringern. Dieses Zeitintervall (die sogenannte Halbwertzeit) wollen wir mit T_0 bezeichnen. Diese Zahl T_0 lässt sich aus der Beziehung

$$e^{-kT_0} = \frac{1}{2}$$

ermitteln, die wir aus der Formel (18) ablesen, wenn wir dort $t = T_0$ setzen und $m(T_0) = \frac{1}{2}m_0$, berücksichtigen. Um nun T_0 zu berechnen, logarithmieren wir diese Beziehung zur Basis e und finden

$$-kT_0 = \ln\frac{1}{2} \qquad , \qquad T_0 = \frac{1}{k}\ln 2 \approx \frac{0{,}69}{k}$$

Hieran erkennen wir, dass die Halbwertzeit T_0 nicht von der Menge des radioaktiven Stoffes, sondern nur von k abhängt, also nur davon, welcher Stoff betrachtet wird.
Wie wir schon erwähnten, ist die Halbwertzeit von Radium etwa 1590 Jahre, die von Uran-238 gleich 4,5 Milliarden Jahre.
Dagegen ist die Halbwertzeit von Thorium C' (einem Poloniumisotop) gleich 0,0000003 Sekunden.

Mit Hilfe von T_0 kann das Gesetz (18) in der Form

$$m(t) = m_0\left(e^{-kT_0}\right)^{\frac{t}{T_0}} = m_0\left(\frac{1}{2}\right)^{\frac{t}{T_0}}$$

geschrieben werden, in der es in der Physik häufig verwendet wird.

1.65 Die Exponentialfunktion in Natur und Technik

In der Natur laufen unzählige Prozesse ab, die alle durch eine Differentialgleichung der Form (17) beschrieben werden. Das Gemeinsame dieser Prozesse besteht darin, dass die Geschwindigkeit, mit der sich die zu betrachtende Größe y ändert, direkt proportional dieser Größe in einem bestimmten Augenblick ist, d.h., es ist

$$y' = cy \tag{19}$$

Der Proportionalitätsfaktor c ist positiv oder negativ, je nachdem, ob sich y mit der Zeit vergrößert oder verkleinert. Die Differentialgleichung (19) hat genau dieselbe Form wie (17), nur dass hier c statt $-k$ steht und die Unbekannte mit y statt mit m bezeichnet ist.
Also muss auch die Lösung von (19) dieselbe Form wie die von (17) haben:

$$y(t) = y_0 e^{ct}$$

Dabei ist y_0 der Wert von y zur Zeit $t = 0$.
Jetzt ist uns auch verständlich, weshalb wir in Natur und Technik so oft Größen begegnen, die einem Exponentialgesetz folgen (siehe den vorhergehenden Abschnitt "Funktionen in Natur und Technik"). Alle diese Größen genügen Differentialgleichungen der Form (19).

1.66 Leverrier und Adams entdecken einen neuen Planeten

Nach dem zweiten Newtonschen Bewegungsgesetz, dem Grundgesetz der Dynamik, ist die Kraft F gleich dem Produkt aus der Masse m und der Beschleunigung a:

$$F = ma$$

Die Beschleunigung a eines geradlinig sich bewegenden Körpers ist gleich der "Geschwindigkeit der Änderung der Geschwindigkeit", also gleich der Ableitung der Geschwindigkeit v:

$$a = v'$$

Nun ist aber die Geschwindigkeit selbst die Ableitung des zurückgelegten Weges nach der Zeit: $v = s'$. Das bedeutet also, dass wir die Funktion $s(t)$ zweimal nach t differenzieren müssen, um auf die Beschleunigung zu kommen. Die Beschleunigung ist also die zweite Ableitung des Weges nach der Zeit.
Man schreibt dafür

$$a(t) = s''(t)$$

Damit können wir dem zweiten Newtonschen Bewegungsgesetz die Gestalt

$$F = ms''$$

geben. Die Kraft F hängt von mehreren Größen ab: von der Zeit, von der Geschwindigkeit und dem Ort, an dem sich der sich bewegende Körper im Raum befindet.

Auf einen Fallschirmspringer, der mit geöffnetem Schirm zur Erde sinkt, wirken z.B. die Schwerkraft mg und die Kraft des Luftwiderstandes, die man proportional zur Fallgeschwindigkeit, also gleich $-kv$ ansetzen kann. Damit ist die auf den Fallschirmspringer wirkende Gesamtkraft gleich

$$F = mg - kv = mg - ks'$$

Folglich wird die Bewegung des Fallschirmspringers durch die Differentialgleichung

$$ms'' = mg - ks'$$

beschrieben.
Eine andere Gestalt hat die Bewegungsgleichung einer Rakete, die auf Grund der Trägheit nach der vollständigen Verbrennung des Treibstoffes senkrecht aufsteigt. Die Anziehungskraft zwischen Rakete und Erde ist umgekehrt proportional dem Quadrat des Abstandes zwischen Rakete und Erdmittelpunkt, also gleich

$$F = -\frac{k}{s^2}$$

Hierbei nehmen wir an, die Rakete habe schon die Erdatmosphäre verlassen, so dass wir die auf sie wirkende Kraft des Luftwiderstandes nicht zu berücksichtigen brauchen. Die Bewegung der Rakete gehorcht also der Differentialgleichung

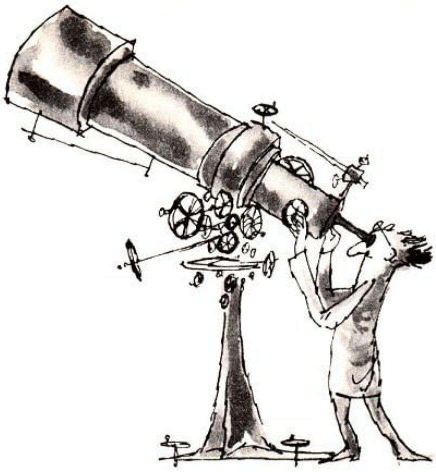
$$ms'' = -\frac{k}{s^2}$$

m ist die Masse der Rakete. Durch diese Gleichung all auch der senkrechte Fall eines Meteoriten auf die Erde bis; zu seinem Eintauchen in die Atmosphäre beschrieben.

Das zweite Newtonsche Bewegungsgesetz gestattet es, die verschiedenartigsten Bewegungen der Körper durch Differentialgleichungen zu beschreiben. So kann man z.B. Differentialgleichungen für die Bewegung des Kolbens einer Dampfmaschine, eines Schiffes auf dem Meer, eines Planeten um die Sonne oder eines künstlichen Satelliten um die Erde aufstellen.

Die Differentialgleichungen für die Bewegung der Planeten und ihrer Satelliten sind überaus kompliziert, da die Planeten nicht nur von der Sonne angezogen werden, sondern sich auch gegenseitig anziehen (Mehrkörperproblem !).
Durch Lösung dieser Differentialgleichungen sind Wissenschaftler in der Lage, den zeitlichen Verlauf der Planetenbewegung sowie Sonnen- und Mondfinsternisse vorherzubestimmen.

Als sich einst beim Planeten Uranus Abweichungen von seiner vorher berechneten Bahn zeigten, zweifelten die Gelehrten nicht im geringsten an der "Richtigkeit" der Mathematik.



Der französische Astronom Joseph Leverrier (1811-1877) und der englische Astronom John Couch Adams (1819-1892) machten gleichzeitig und voneinander unabhängig die kühne Voraussage, die Störung in der Bahn des Uranus werde durch die Anziehung eines noch nicht entdeckten Planeten hervorgerufen.

Mit Hilfe von Differentialgleichungen hatten sie die Lage dieses neuen Planeten berechnet und genau angegeben, an welcher Stelle des Himmels er zu suchen sei. Genau an dieser Stelle fand der deutsche Astronom Johann Gottfried Galle (1812 bis 1910) im Jahre 1846 diesen neuen Planeten, der den Namen Neptun erhielt.

1.67 Die Schwingungsgleichung

In vielen Fällen schwingt ein Körper um seine Ruhelage unter Einwirkung einer Kraft, die der Auslenkung des Körpers aus seiner Ruhelage proportional ist und danach trachtet, den Körper in seine Ruhelage zurückzubringen.

Dies ist z.B. bei einer an einer Schraubenfeder hängenden Masse der Fall.
Die Kraft lässt sich also durch

$$F = -ks$$

ausdrücken; s ist die Auslenkung des Körpers aus seiner Ruhelage. k die Federkonstante. Daher hat (auf Grund des zweiten Newtonschen Gesetzes) die Bewegungsgleichung des Körpers die Form

$$ms'' = -ks$$

oder, wenn wir $\frac{k}{m} = \omega^2$ setzen,

$$s'' = -\omega^2 s$$

Diese Gleichung nennt man die Gleichung der (ungedämpften) harmonischen Schwingung, da die Funktion

$$s = C_1 \cos \omega t + C_2 \sin \omega t \tag{20}$$

für alle C_1 und C_2 eine Lösung der Gleichung ist. Auf Grund von (13) und (14) ist nämlich die Geschwindigkeit eines sich nach dem Gesetz (20) bewegenden Körpers gleich

$$v = s' = -C_1 \omega \sin \omega t + C_2 \omega \cos \omega t$$

Differenzieren wir sie noch einmal, so finden wir seine Beschleunigung:

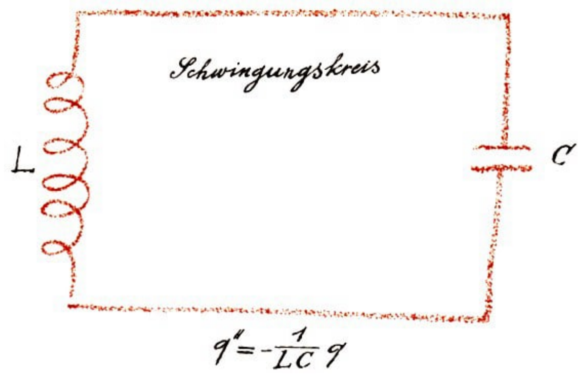
$$a = s'' = -C_1 \omega^2 \cos \omega t - C_2 \omega^2 \sin \omega t = -\omega^2 (C_1 \cos \omega t + C_2 \sin \omega t) = -\omega^2 s$$

Also genügt die Funktion (20) tatsächlich der Gleichung

$$s'' = -\omega^2 s$$

Man kann auch zeigen, dass jede Lösung dieser Gleichung die Form (20) hat. Also ruft die Kraft, die der Auslenkung eines Körpers aus der Ruhelage proportional ist und die den Körper in seine Ruhelage zurückbringen möchte, harmonische Schwingungen dieses Körpers mit der Frequenz ω hervor, wobei $\omega^2 = \frac{k}{m}$, m die Masse des Körpers und k ein Proportionalitätsfaktor ist.

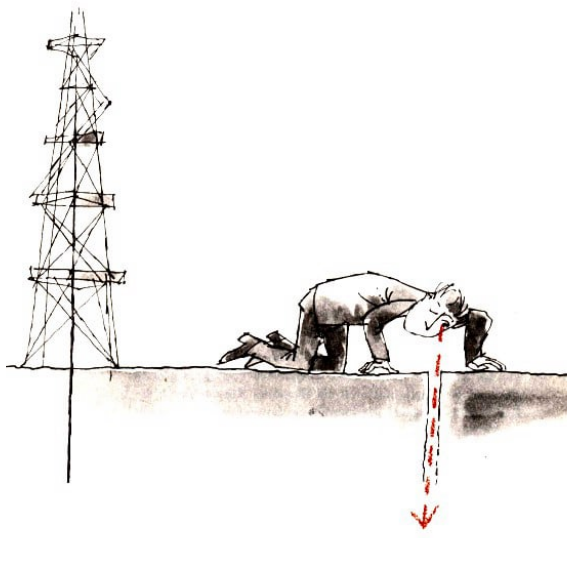
Für die Schwingungen in einem elektrischen Stromkreis gibt es ein analoges Gesetz, nur muss die Masse des Körpers durch die Induktivität der Spule, der vom Körper zurückgelegte Weg durch die Spannung am Kondensator, die Geschwindigkeit des Körpers durch die Stromstärke usw. ersetzt werden.



Da die Gesetze, denen diese Erscheinungen gehorchen, völlig analog sind, lassen sich auch die in beiden Fällen auftretenden Schwingungen durch ein und dieselben Formeln beschreiben.
Gedämpfte Schwingungen treten auf, wenn außer der Kraft, die den Körper in seine Ruhelage zieht, noch der Widerstand des Mediums, der der Geschwindigkeit des Körpers proportional ist, oder der Widerstand des elektrischen Stromkreises wirkt.

1.68 Modellierung (Simulation)

Die Tatsache, dass sich die verschiedensten Erscheinungen durch dieselben Differentialgleichungen beschreiben lassen, wird in der Praxis oft ausgenutzt. Sie gestattet es, eine bestimmte Erscheinung durch Beobachtung einer anderen zu studieren, wenn beide denselben Differentialgleichungen gehorchen. Es soll beispielsweise festgestellt werden, wie sich das Erdöl unter der Erde im Bereich eines Bohrlochs bewegt.



Nun lässt sich die Bewegung des Erdöls unter der Erde schlecht beobachten. Da aber die Flüssigkeitsbewegung durch dieselbe Differentialgleichung beschrieben wird wie der Strom in einem elektrischen Stromkreis, kann man einen Stromkreis als Modell für die Flüssigkeitsbewegung benutzen.

Misst man Spannung und Stromstärke im Stromkreis, so lässt sich z.B. errechnen, wo man am zweckmäßigsten das Bohrloch anbringt, wo man Wasser, Gas oder Luft in die Erde hineinpumpen muss, um die Förderung des Erdöls zu erleichtern, usw.

Diese Untersuchung einer bestimmten Erscheinung mit Hilfe einer anderen, die derselben Differentialgleichung genügt, heißt Modellierung oder Simulation und wird auf den verschiedensten Gebieten oft angewendet. Insbesondere liegt sie der Konstruktion und Anwendung von Analogrechenmaschinen zugrunde. Der Leser vergleiche dazu den entsprechenden Abschnitt im Band 1.

2 Algebra des Endlichen und Unendlichen

2.1 Geometrische Algebra

A. M. Lopschiz

Es mag auf den ersten Blick verwunderlich erscheinen, dass gerade die Algebra, in der doch, wie jeder weiß, nur Formeln vorkommen und kaum Zeichnungen, geometrisch sein soll.
Und doch gibt es bereits seit mehr als hundert Jahren eine derartige Algebra. Die Mathematiker entwickelten sie entsprechend den Bedürfnissen der Physik und der Mathematik selbst. Die Grundformeln der geometrischen Algebra sind uns fast alle von der Schule her geläufig, nur haben jetzt die in den Formeln auftretenden Buchstaben eine andere Bedeutung als damals:

Sie stehen nicht mehr für Zahlen, sondern für "gerichtete Strecken", wie sie beispielsweise in der Physik zur Darstellung von Kräften, Geschwindigkeiten oder Beschleunigungen verwendet werden. Die Mathematiker bezeichnen solche gerichteten Strecken als Vektoren.

In diesem Abschnitt werden wir die Grundbegriffe und Grundregeln dieser Algebra (der sogenannten Vektoralgebra) besprechen und zeigen, wie man sie zur Lösung von verschiedenartigen geometrischen und physikalischen Aufgaben benutzen kann, welche mit den Hilfsmitteln der Schulgeometrie allein nur schwer zu lösen wären.

2.2 Die Geometrie hilft der Physik

Wie stellt man Kräfte dar? Wir erinnern daran, dass eine in einem Punkt P eines festen Körpers angreifende Kraft zeichnerisch durch eine gerichtete Strecke PA mit der Spitze im Punkt A dargestellt wird.

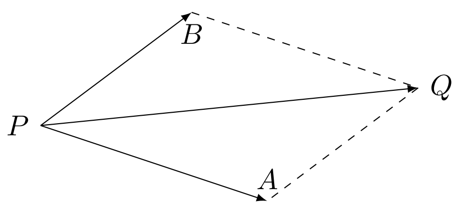
Die Richtung dieser Strecke, d.h. die Richtung des Strahls PP', auf dem die Strecke PA liegt, stimmt mit der Richtung der dargestellten Kraft überein.

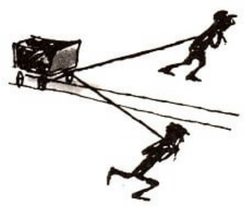
Die Länge der Strecke PA wählt man so, dass in ihr gleichviel Längeneinheiten enthalten sind, etwa Strecken der Länge 1 cm, wie Einheiten der Kraft in der darzustellenden Kraft, gemessen etwa in Kilopond.

Um anzudeuten, dass PA eine gerichtete Strecke ist, schreibt man gewöhnlich $\overrightarrow{PA}$. Der Pfeil über den Buchstaben soll andeuten, dass es nicht nur auf die Länge der gerichteten Strecke ankommt (diese bezeichnet man entweder mit $\overline{PA}$ oder auch mit $\overline{AP}$), sondern auch auf ihre Richtung.
Eine gerichtete Strecke $\overrightarrow{PA}$ nennt man einen Vektor. Der Punkt P, d.h. der Angriffspunkt der Kraft, wird Anfangspunkt des Vektors, der Punkt A sein Endpunkt.

2.3 Die resultierende Kraft

Wenn in einem Punkt P zwei Kräfte in verschiedenen Richtungen angreifen, so ist ihre gemeinsame Wirkung, wie aus der Physik bekannt ist, gleich der Wirkung einer einzigen Kraft, die ebenfalls im Punkt P angreift.





Die dieser resultierenden Kraft ist nun keineswegs gleich der Summe der Größen der beiden ursprünglichen Kräfte, sondern kleiner als diese Summe. Wie kann man nun Richtung und Größe der resultierenden Kraft exakt angeben?
Hier helfen glücklicher Weise geometrische Vorstellungen.

Zahlreiche physikalische Erfahrungen bestätigen die Richtigkeit des folgenden Gesetzes, das sich gerade deshalb so einfach aussprechen lässt, weil wir die Kräfte in der oben beschriebenen Weise geometrisch darstellen können:

Die Resultierende zweier durch die Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ dargestellten Kräfte wird dargestellt durch den Vektor $\overrightarrow{PQ}$, der in dem aus $\overrightarrow{PA}$ und $\overrightarrow{PB}$ konstruierten Parallelogramm die Diagonale bildet. Die exakte Bestimmung des resultierenden Vektors zweier Kräfte wird somit zu einer einfachen geometrischen Operation.
Wie nennt man sie? Die Operation, mit deren Hilfe aus gegebenen Kraftkomponenten die Resultierende gebildet wird, heißt Zusammensetzen von Kräften oder Addition von Kräften. Die Resultierende selbst bezeichnet man als die Summe der beiden Kräfte.

Dabei darf man allerdings nicht vergessen, dass für die Addition von Kräften (bzw. für die Addition von Vektoren) ganz andere Regeln gelten als für die Addition von Zahlen. Man bezeichnet sie deshalb oft auch als geometrische Addition.

Wie schreibt man die Addition von Vektoren? Natürlich verwendet man die in der Algebra und in der Arithmetik gebräuchlichen Zeichen und schreibt:

$$\overrightarrow{PA} + \overrightarrow{PB} = \overrightarrow{PQ} \tag{1}$$

ohne dabei jedoch zu vergessen, dass das Zeichen "$+$" in dieser Formel nicht dasselbe bedeutet, wie etwa in dem Ausdruck $3 + 5 = 8$.

Sind zwei Kraftkomponenten gegeben, so entsteht eine resultierende Kraft. Diese hängt jedoch nicht davon ab, in welcher Reihenfolge wir diese Kräfte aufschreiben. Daher können wir anstelle von Gleichung (1) auch

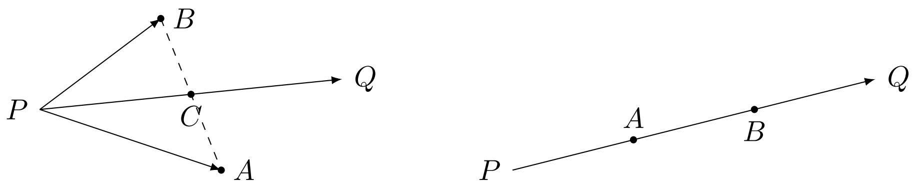
$$\overrightarrow{PB} + \overrightarrow{PA} = \overrightarrow{PQ} \tag{1}$$

schreiben. Mit anderen Worten, es gilt die Beziehung

$$\overrightarrow{PA} + \overrightarrow{PB} = \overrightarrow{PB} + \overrightarrow{PA}$$

Wir sind zu einer Formel gelangt, die sich äußerlich nicht von dem entsprechenden algebraischen Ausdruck $a + b = b + a$ unterscheidet. Bekanntlich nennt man diese Beziehung das Kommutativgesetz der Addition.
Die Ähnlichkeit dieser beiden Formeln beruht darauf, dass wir zur Bezeichnung der geometrischen Addition dasselbe Zeichen verwandten, das zur Darstellung der Addition in der Arithmetik und in der Algebra benutzt wird.

2.4 Die Additionsregel für Vektoren



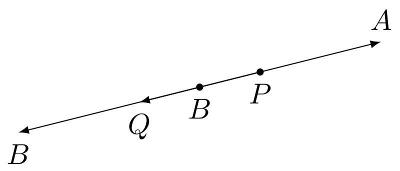
Um die Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ zu addieren, verfahren wir folgendermaßen:
1. Wir bestimmen zunächst den Mittelpunkt C der Strecke AB, welche die Endpunkte der zu addierenden Vektoren verbindet.
2. Wir konstruieren den Punkt Q, der in Bezug auf den Punkt C zum Punkt P symmetrisch liegt, d.h., wir bestimmen auf der Geraden PC den Punkt Q, für den der Punkt C Mittelpunkt der Strecke PQ ist.
Im Ergebnis dieser Konstruktion erhalten wir die Diagonale PQ des von den gegebenen Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ aufgespannten Parallelogramms. Daher ist

$$\overrightarrow{PA} + \overrightarrow{PB} = \overrightarrow{PQ}$$

Das Viereck $PAQB$ ist tatsächlich ein Parallelogramm; denn nach unserer Konstruktion halbieren die Diagonalen einander im Punkt C.

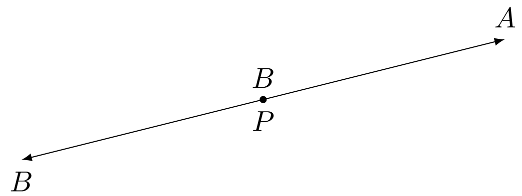
Der besondere Wert dieser Mittelpunktsregel (wie man sie auch nennt) besteht darin, dass man mit ihrer Hilfe die Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ auch dann addieren kann, wenn sie beide auf einer Geraden liegen.
Aus derartigen Vektoren lässt sich natürlich kein Parallelogramm konstruieren.
Ganz offensichtlich ergibt sich aus der Mittelpunktsregel folgendes (rechte obere Abbildung):

1. Das Ergebnis der Addition von zwei gleichgerichteten Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ ist der Vektor $\overrightarrow{PQ}$, der dieselbe Richtung hat wie die beiden zu addierenden Vektoren und dessen Länge gleich der Summe der Längen dieser Vektoren ist.



2. Das Ergebnis der Addition von zwei entgegengesetzt gerichteten Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ ist der Vektor $\overrightarrow{PQ}$, der die Richtung des größeren (längeren) Summanden hat, und dessen Länge gleich der Differenz der Längen der zu addierenden Vektoren ist.

Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$, die in einer Geraden liegen, nennt man kollinear, unabhängig davon, ob sie gleich oder entgegengesetzt gerichtet sind. Die Additionsregeln für solche Vektoren entsprechen genau der physikalischen Erfahrung bei der Bestimmung der Resultierenden zweier kollinearer Kräfte.



Besondere Aufmerksamkeit verdient der Fall, dass die zu addierenden Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ nicht nur entgegengesetzt gerichtet, sondern auch gleich lang sind. Solche Vektoren, die also gleich groß und entgegengesetzt gerichtet sind, heißen entgegengesetzt gleich.
Wir wenden die Mittelpunktsregel an:

1. Wir konstruieren den Mittelpunkt C der Strecke AB; er fällt offenbar mit dem Punkt P zusammen.
2. Wir konstruieren den Punkt Q, der in Bezug auf den Punkt C (hier also bezüglich des Punktes P) zum Punkte P symmetrisch liegt; wie wir sehen, stimmt auch der Punkt Q mit dem Punkt P überein.

Sind also die Vektoren $\overrightarrow{PA}$ und $\overrightarrow{PB}$ entgegengesetzt gleich, so ist

$$\overrightarrow{PA} + \overrightarrow{PB} = \overrightarrow{PP}$$

Wir haben einen Vektor besonderer Art erhalten: einen Vektor, bei dem Anfangspunkt und Endpunkt zusammenfallen.
Vielleicht erscheint es dem Leser auf den ersten Blick sonderbar, dass das ein Vektor sein soll. Er hat ja keine Richtung.

Wollte man aber derartige "besondere" Vektoren ausschließen, so hieße das, dass man nicht in jedem Fall zwei Vektoren addieren könnte, und diese Ausnahme wäre natürlich unzweckmäßig.

Wir wollen also annehmen, dass zwei beliebige Punkte P und Q, unabhängig davon, ob sie verschieden sind oder nicht, eine gerichtete Strecke $\overrightarrow{PQ}$ bestimmen.

Sind die Punkte P und Q verschieden, so muss selbstverständlich gesagt werden, welcher der beiden als der erste (Anfangs-) Punkt und welcher als der zweite (End-) Punkt anzusehen ist. Dieses Punktepaar muss, wie wir sagen wollen, geordnet sein. Fallen aber die Punkte P und Q zusammen, dann braucht auf die Reihenfolge der Punkte nicht geachtet zu werden.

Um zu sehen, wie man mit dem besonderen Vektor zu rechnen hat, prüfe man nach, dass die Beziehung

$$\overrightarrow{PA} + \overrightarrow{PP} = \overrightarrow{PA} \tag{2}$$

gilt. Zum Beweis braucht man nur die Addition nach der Mittelpunktsregel auszuführen. Diese Gleichung erinnert natürlich an die algebraische Gleichung $a + 0 = a$ und überzeugt uns davon, dass der besondere Vektor $\overrightarrow{PP}$ bei der Vektoraddition dieselbe Rolle spielt wie die Zahl Null in der Arithmetik. Deshalb nennt man den besonderen Vektor $\overrightarrow{PP}$ auch den Nullvektor. Man bezeichnet ihn mit $\overrightarrow{0}$.
Die Gleichung (2) kann man also auch in der Gestalt

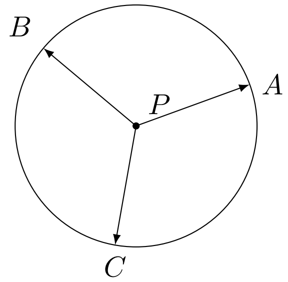
$$\overrightarrow{PA} + \overrightarrow{0} = \overrightarrow{PA}$$

schreiben.

2.5 Anwendungen aus der Geometrie

1. Bekanntlich teilt der Schnittpunkt S der Seitenhalbierenden in einem Dreieck ABC jede dieser Linien im Verhältnis 2:1.
Es seien in diesem Punkt P drei Kräfte angetragen, die durch die Vektoren $\overrightarrow{PA}$, $\overrightarrow{PB}$ und $\overrightarrow{PC}$ dargestellt werden. Man bestimme die Resultierende dieses Kräftesystems. (Wir greifen hier etwas vor, indem wir schon von der Summe dreier Vektoren reden).

Lösung: Die Summe der drei Kräfte ist gleich $\overrightarrow{PP}$; folglich hält dieses Kräftesystem dem Punkt P im Gleichgewicht.



2. Die Punkte A_1, A_2 und A_3 sollen einen Kreis mit dem Mittelpunkt C in drei gleich große Bogen teilen. Man beweise die Beziehung

$$(\overrightarrow{CA_1} + \overrightarrow{CA_2}) + \overrightarrow{CA_3} = \overrightarrow{CC}$$

3. Die Punkte A und C seien gegenüberliegende Ecken in einem Parallelogramm $ABCD$. Man beweise, dass bei beliebiger Wahl des Punktes P

$$\overrightarrow{PA} + \overrightarrow{PC} = \overrightarrow{PB} + \overrightarrow{PD}$$

gilt.

4. Drei voneinander verschiedene, nicht kollineare Punkte A,B,C und ein von diesen verschiedener vierter Punkt D bilden ein Viereck $ABCD$. Wie muss es beschaffen sein, damit bei beliebiger Wahl des Punktes P die Gleichung

$$\overrightarrow{PA} + \overrightarrow{PC} = \overrightarrow{PB} + \overrightarrow{PD}$$

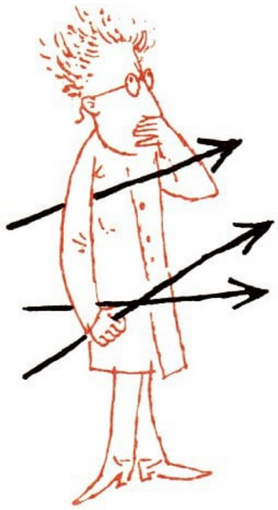
erfüllt ist?

Lösung: Es muss ein Parallelogramm sein mit den gegenüberliegenden Ecken A und C.

5. Durch einen beliebigen Punkt P im Inneren eines Kreises mit dem Mittelpunkt O werden die aufeinander senkrecht stehenden Sehnen AB und CD gezogen. Man bestimme die Resultierende von zwei Kräften, die im Punkte P angreifen und durch die Vektoren $\overrightarrow{PM}$ und $\overrightarrow{PN}$ dargestellt werden, wenn M und N die Mittelpunkte der Sehnen AB bzw. CD sind.

Lösung: Die Resultierende ist der Vektor $\overrightarrow{PO}$.

2.6 Die Geometrie schafft eine neue Algebra



Was ist die Summe dreier Vektoren? Die Antwort scheint einfach zu sein. Um drei Vektoren zu addieren, genügt es, zunächst zwei zu addieren und dann zur Summe den dritten hinzuzufügen.

Hier erhebt sich jedoch die Frage: Welche zwei der drei Vektoren soll man zuerst addieren? Soll man als Summe der drei Vektoren $\overrightarrow{PA}$, $\overrightarrow{PB}$ und $\overrightarrow{PC}$ den Vektor $(\overrightarrow{PA} + \overrightarrow{PB}) + \overrightarrow{PC}$ oder den Vektor $(\overrightarrow{PB} + \overrightarrow{PC}) + \overrightarrow{PA}$ oder etwa den Vektor $(\overrightarrow{PC} + \overrightarrow{PA}) + \overrightarrow{PB}$ ansehen?

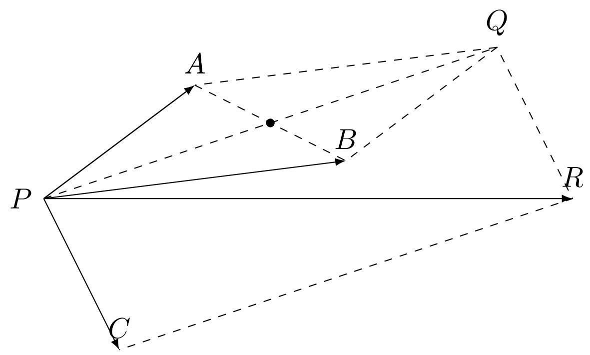
Das gleiche Problem hat der Leser bereits in der Arithmetik kennengelernt: Wie bildet man die Summe dreier Zahlen a, b, c? Ist sie gleich $(a+b)+c$ oder gleich $(b+c)+a$ oder gleich $(c+a)+b$? Die Antwort auf diese Frage ist wohlbekannt. Für die Summe dreier Zahlen gilt die Formel

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

das sogenannte Assoziativgesetz der Addition.
Ist vielleicht auch die Vektoraddition assoziativ, d.h., gilt die Formel

$$(\overrightarrow{PA} + \overrightarrow{PB}) + \overrightarrow{PC} = \overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC}) \tag{3}$$

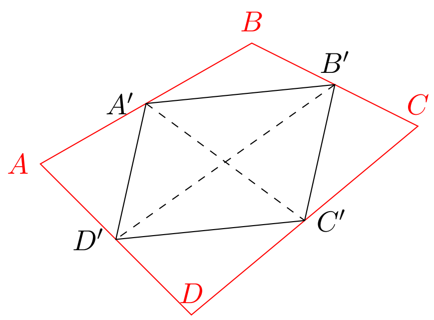
Prüft man diese Beziehung an einem Beispiel nach, so findet man sie bestätigt. Da man jedoch eine mathematische Aussage nicht durch Beispiele beweisen kann, muss man sich nach einem echten Beweis umsehen.

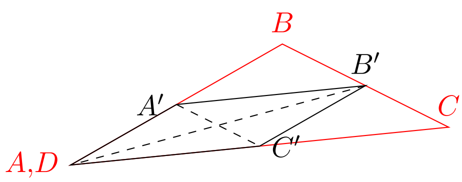


Zu diesem Zweck brauchen wir einen einfachen geometrischen Satz, der sich leicht beweisen lässt:

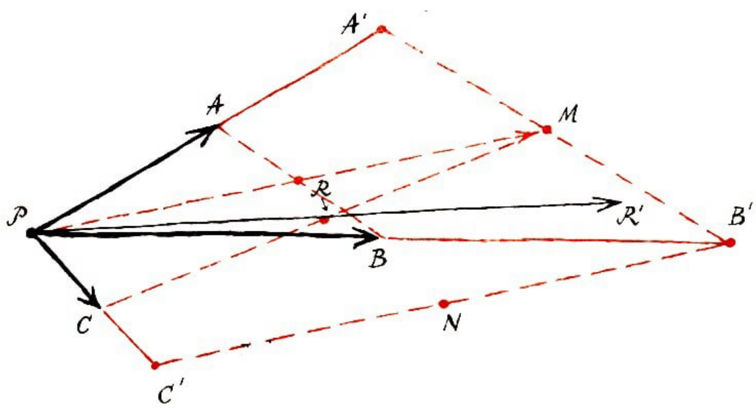
Sind A', B', C' und D' die Mittelpunkte der Seiten AB, BC, CD und DA eines beliebigen Vierecks, so fällt der Mittelpunkt der Strecke $A'C'$ mit dem Mittelpunkt der Strecke $B'D'$ zusammen.

Bei Beweis dieses Hilfssatzes muss man je nach der Lage der vier beliebigen Ausgangspunkte A,B,C und D verschiedene Fälle unterscheiden.
Zwei Fälle sind in der Abbildung wiedergegeben.





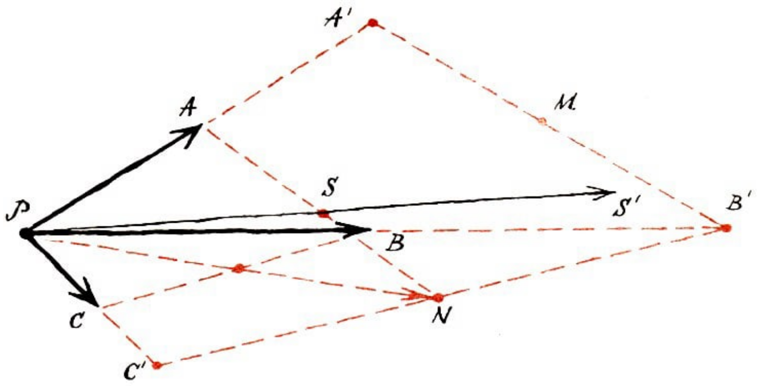
Wir wenden uns jetzt dem Beweis der Formel (3) zu. Sind die drei Vektoren $\overrightarrow{PA}$, $\overrightarrow{PB}$ und $\overrightarrow{PC}$ gegeben, so konstruieren wir die Punkte A', B' und C', die in Bezug auf A, B und C zu P symmetrisch liegen, und bezeichnen die Mittelpunkte der Seiten $A'B'$ und $B'C'$ im Viereck $PA'B'C'$ mit M und N. Jetzt benutzen wir unseren Hilfssatz für das Viereck $PA'B'C'$.



Danach stimmt der Mittelpunkt der Strecke AB mit dem Mittelpunkt der Strecke PM überein. Nach der Mittelpunktsregel ist daher

$$\overrightarrow{PA} + \overrightarrow{PB} = \overrightarrow{PM}$$

Es sei nun R der Mittelpunkt der Strecke MC und R' der Punkt, der in Bezug auf R zu P symmetrisch liegt.



Dann ist

$$(\overrightarrow{PA} + \overrightarrow{PB}) + \overrightarrow{PC} = \overrightarrow{PM} + \overrightarrow{PC} = \overrightarrow{PR}$$

Genauso erhalten wir $\overrightarrow{PB} + \overrightarrow{PC} = \overrightarrow{PN}$; hieraus folgt aber

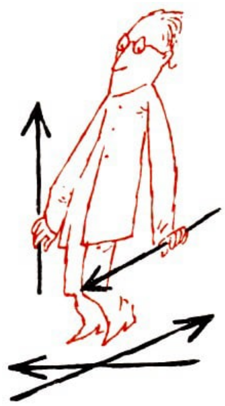
$$\overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC}) = \overrightarrow{PA} + \overrightarrow{PN} + \overrightarrow{PS'}$$

wobei S' der Punkt ist, der zu P in Bezug auf den Mittelpunkt S der Strecke AN symmetrisch liegt.
Wenden wir unseren Hilfssatz auf das Viereck $PA'B'C'$ an, so bemerken wir, dass die Punkte R' und S' zusammenfallen; also stimmen $\overrightarrow{PR'}$ und $\overrightarrow{PS'}$ überein. Somit ist

$$(\overrightarrow{PA} + \overrightarrow{PB}) + \overrightarrow{PC} = \overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC})$$

Deshalb kann man die Klammern weglassen und die Summe dreier Vektoren einfach in der Gestalt $\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC}$ schreiben.

2.6.1 Die Summe von mehr als drei Vektoren



Um vier Vektoren zu addieren, bildet man die Summe von drei Vektoren und addiert zur Summe den vierten Vektor. Das Ergebnis der Addition hängt nicht davon ab, welche Vektoren zuerst addiert werden. Das kann man jetzt schon rein algebraisch beweisen.
Als Beispiel beweisen wir die Beziehung

$$(\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC}) + \overrightarrow{PD} = \overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC} + \overrightarrow{PD})$$

In der Tat, es ist

$$\begin{aligned}(\overrightarrow{PA} + \overrightarrow{PB} + \overrightarrow{PC}) + \overrightarrow{PD} &= \{\overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC})\} + \overrightarrow{PD} = \overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC}) + \overrightarrow{PD} \\ &= \overrightarrow{PA} + \{(\overrightarrow{PB} + \overrightarrow{PC}) + \overrightarrow{PD}\} = \overrightarrow{PA} + (\overrightarrow{PB} + \overrightarrow{PC} + \overrightarrow{PD})\end{aligned}$$

Man prüft leicht nach, dass das Assoziativgesetz auch für eine beliebige Anzahl von Vektorsummanden gilt. Deshalb darf man die Klammern weglassen und

$$\overrightarrow{PA_1} + \overrightarrow{PA_2} + ... + \overrightarrow{PA_{n-1}} + \overrightarrow{PA_n}$$

schreiben.

2.6.2 Multiplikation eines Vektors mit einer positiven ganzen Zahl

So wie wir es von der Algebra in der Schule her gewohnt sind, wollen wir im weiteren die folgenden Bezeichnungen verwenden:

$$\begin{gathered} \overrightarrow{PA} + \overrightarrow{PA} \equiv 2 \cdot \overrightarrow{PA} \\ \overrightarrow{PA} + \overrightarrow{PA} + \overrightarrow{PA} \equiv 3 \cdot \overrightarrow{PA} \end{gathered} \tag{4}$$

Das Zeichen $\equiv$ soll bedeuten, dass die rechte Seite eine Abkürzung für die linke ist.
Entsprechend wird man natürlich sagen, der Vektor $n \cdot \overrightarrow{PA}$ sei das Ergebnis der Multiplikation der Zahl n mit dem Vektor $\overrightarrow{PA}$. Dafür schreibt man auch $\overrightarrow{PA} \cdot n$.
Unter Benutzung des Assoziativgesetzes der Addition und der Bezeichnungsweise (4) beweist man leicht die folgenden Formeln, die den aus der Schulalgebra bekannten Formeln auffallend ähneln:

$$m\overrightarrow{PA} + m\overrightarrow{PB} = m(\overrightarrow{PA} + \overrightarrow{PB}) \tag{5}$$

$$m\overrightarrow{PA} + n\overrightarrow{PA} = (m+n)\overrightarrow{PA} \tag{6}$$

$$m(n \cdot \overrightarrow{PA}) = (mn)\overrightarrow{PA} \tag{7}$$

für ganzzahlige m und n.

2.6.3 Multiplikation eines Vektors mit einer positiven rationalen Zahl $\frac{m}{n}$

Wir wollen noch eine weitere in der Algebra gebräuchliche Bezeichnung übernehmen und auch in unserer "Vektoralgebra"

$$\overrightarrow{PA} = \frac{m}{n} \cdot \overrightarrow{PQ} \qquad \text{bzw.} \qquad \overrightarrow{PA} = \overrightarrow{PQ} \cdot \frac{m}{n}$$

schreiben, wenn

$$n \cdot \overrightarrow{PA} = m \cdot \overrightarrow{PQ}$$

ist. Hieraus folgt, dass (wie man leicht beweist) auch für die Multiplikation eines Vektors mit einer positiven rationalen Zahl die folgenden Formeln gelten:

$$\frac{m}{n}(\overrightarrow{PA} + \overrightarrow{PB}) = \frac{m}{n}\overrightarrow{PA} + \frac{m}{n}\overrightarrow{PB} \tag{8}$$

$$\left(\frac{m_1}{n_1} + \frac{m_2}{n_2}\right)\overrightarrow{PA} = \frac{m_1}{n_1}\overrightarrow{PA} + \frac{m_2}{n_2}\overrightarrow{PA} \tag{9}$$

$$\frac{m_1}{n_1}\left(\frac{m_2}{n_2} \cdot \overrightarrow{PA}\right) = \left(\frac{m_1}{n_1} \cdot \frac{m_2}{n_2}\right)\overrightarrow{PA} \tag{10}$$

Diese Formeln enthalten die bereits bewiesenen Formeln (5), (6) und (7) als Spezialfälle.

2.6.4 Multiplikation eines Vektors mit der Zahl -1

Wieder benutzen wir Bekanntes aus der Schulalgebra. Wir erinnern daran, dass

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

ist. Jetzt setzen wir natürlich

$$(-1)\overrightarrow{PA} + \overrightarrow{PA} = \overrightarrow{PP}$$

d. h., wir setzen fest:
Der Vektor $(-1)\overrightarrow{PA}$ sei derjenige Vektor, welcher ebenso groß wie der Vektor $\overrightarrow{PA}$, aber entgegengesetzt gerichtet ist.

2.6.5 Multiplikation eines Vektors mit einer negativen rationalen Zahl

Diese Rechenoperation verstehen wir folgendermaßen:

$$\overrightarrow{PA} \cdot \left(-\frac{m}{n}\right) = \left(\overrightarrow{PA} \cdot \frac{m}{n}\right) \cdot (-1)$$

Wir setzen also fest, der Vektor $-\frac{m}{n}\overrightarrow{PA}$ sei der zum Vektor $\frac{m}{b}\overrightarrow{PA}$ entgegengesetzt gerichtete gleich große Vektor. Wie zweckmäßig diese Regel ist, erkennt man beispielsweise daran, dass aus ihr, wie man beweisen kann, die Richtigkeit der Formeln (7), (8) und (9) auch für den Fall folgt, dass einer der darin auftretenden Zahlenfaktoren negativ wird oder sogar verschwindet.
Man muss nur noch festsetzen, dass

$$0 \cdot \overrightarrow{PA} = \overrightarrow{PP} = \overrightarrow{0}$$

ist. An dieser Vereinbarung wollen wir auch im weiteren festhalten.

2.6.6 Subtraktion von Vektoren

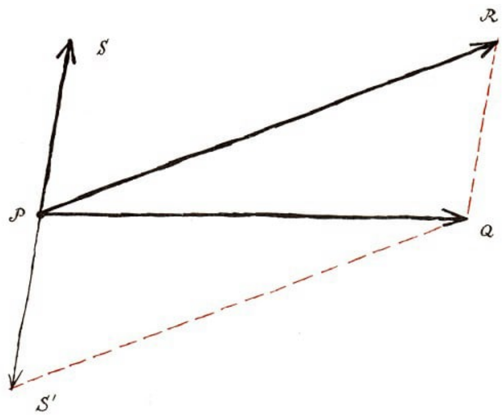
Diese Rechenoperation definieren wir als die Umkehrung der Vektoraddition (man sagt auch: die zur Addition inverse Operation).
Der Vektor $\overrightarrow{PQ}$ heißt die Differenz der beiden Vektoren $\overrightarrow{PR}$ und $\overrightarrow{PS}$, in Zeichen

$$\overrightarrow{PQ} = \overrightarrow{PR} - \overrightarrow{PS}$$

wenn der Vektorminuend $\overrightarrow{PR}$ gleich der Summe aus der Vektordifferenz $\overrightarrow{PQ}$ und dem Vektorsubtrahenden $\overrightarrow{PS}$ ist, d.h., wenn,

$$\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{PS}$$

gilt. Sehr wichtig ist, dass man sich folgende Regel klarmacht und einprägt: Die Bestimmung der Differenz $\overrightarrow{PR} - \overrightarrow{PS}$ ist gleichwertig mit der Bestimmung der Summe aus dem Minuenden $\overrightarrow{PR}$



und dem zum Subtrahenden $\overrightarrow{PS}$ entgegengesetzt gleichen Vektor, d.h., es ist

$$\overrightarrow{PR} - \overrightarrow{PS} = \overrightarrow{PR} + (-1) \cdot \overrightarrow{PS} \tag{11}$$

Es ist nämlich

$$|\overrightarrow{PR} + (-1) \cdot \overrightarrow{PS}| + \overrightarrow{PS} = \overrightarrow{PR} + |(-1) \cdot \overrightarrow{PS} + \overrightarrow{PS}| = \overrightarrow{PR} + \overrightarrow{PP} = \overrightarrow{PR}$$

Diese Ersetzung der Subtraktion durch die Addition ermöglicht es uns, die wichtigen Formeln für die Subtraktion einer Summe und einer Differenz leicht herzuleiten:

$$\overrightarrow{PA} - (\overrightarrow{PB} + \overrightarrow{PC}) = \overrightarrow{PS} + (-1) \cdot \overrightarrow{PB} + (-1) \cdot \overrightarrow{PC} \tag{12}$$
$$\overrightarrow{PA} - (\overrightarrow{PB} - \overrightarrow{PC}) = \overrightarrow{PA} + (-1) \cdot \overrightarrow{PB} + \overrightarrow{PC} \tag{13}$$

Wir wollen die zweite Formel beweisen.

$$\begin{aligned}\overrightarrow{PA} - (\overrightarrow{PB} - \overrightarrow{PC}) &= \overrightarrow{PA} + (-1) \cdot [\overrightarrow{PB} + (-1) \cdot \overrightarrow{PC}] \\ &= \overrightarrow{PA} + (-1) \cdot \overrightarrow{PB} + (-1) \cdot (-1) \cdot \overrightarrow{PC} = \overrightarrow{PA} + (-1) \cdot \overrightarrow{PB} + \overrightarrow{PC}\end{aligned}$$

Wir übernehmen noch etwas aus der Schulalgebra. Im Algebraunterricht taucht hin und wieder die Frage auf: Kann man sagen, die Zahl $-a$ sei eine negative Zahl?
Die Antwort ist klar: Das kann man nicht sagen.

Es ist nämlich $-a = (^-1)a$. Ist also a eine negative Zahl, so ist $-a$ positiv; nur dann, wenn a eine positive Zahl ist, ist $-a$ negativ. Die Zahlen a und $-a$ nennt man bekanntlich entgegengesetzt gleich.
Wir wollen auch das aus der Schulalgebra übernehmen und sagen, es sei

$$(-1)\overrightarrow{PA} = -\overrightarrow{PA}$$

den Vektor $-\overrightarrow{PA}$ nennen wir zum Vektor $\overrightarrow{PA}$ entgegengesetzt gerichtet. Keinen dieser beiden Vektoren kann man als negativ bezeichnen, da wir auch den Begriff eines "positiven Vektors" nicht eingeführt haben (und das auch nicht tun werden).

Mit Hilfe dieser Bezeichnungen können wir die Formeln (11), (12) und (13) in eine

Form umschreiben, die jedem Schüler vertraut ist, der die Regeln für das Auflösen von Klammern kennt, vor denen ein Minuszeichen steht:

$$\overrightarrow{PR} - \overrightarrow{PS} = \overrightarrow{PR} + (-\overrightarrow{PS})$$
$$\overrightarrow{PA} - (\overrightarrow{PB} + \overrightarrow{PC}) = \overrightarrow{PA} - \overrightarrow{PB} - \overrightarrow{PC}$$
$$\overrightarrow{PA} - (\overrightarrow{PB} - \overrightarrow{PC}) = \overrightarrow{PA} - \overrightarrow{PB} + \overrightarrow{PC}$$

Diese Übereinstimmung mit den aus der Schulalgebra bekannten Formeln wird noch größer, wenn wir jeden Vektor kurz mit einem einzigen Buchstaben bezeichnen, z.B.:

$$\overrightarrow{PA} = \overrightarrow{a}, \qquad \overrightarrow{PB} = \overrightarrow{b}, \qquad \overrightarrow{PC} = \overrightarrow{c}$$

Die Striche über den Buchstaben (a, b, c) sollen den Leser daran erinnern, dass er es mit Vektoren zu tun hat und nicht mit Zahlen. Der Einfachheit halber werden auch die folgenden Bezeichnungen für Vektoren benutzt:

$$\overrightarrow{PA} = \mathbf{a}, \qquad \overrightarrow{PB} = \mathbf{b}, \qquad \overrightarrow{PC} = \mathbf{c}$$

(sogenannte halbfette Buchstaben) oder

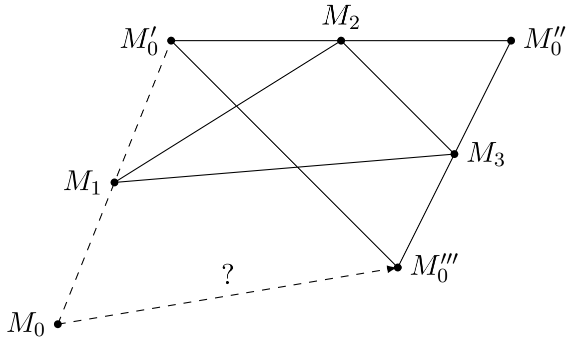
$$\overrightarrow{PA} = \mathfrak{a}, \qquad \overrightarrow{PB} = \mathfrak{b}, \qquad \overrightarrow{PC} = \mathfrak{c}$$

(sogenannte Frakturbuchstaben).

2.7 Die Vektoralgebra hilft der Geometrie

Aus der Geometrie ist eine neue Algebra, die Vektoralgebra, hervorgegangen. Diese wird in der Geometrie zur Lösung von Aufgaben herangezogen. Wir wollen einige Beispiele dafür angeben.

Beispiel:
Es sei ein beliebiges Dreieck $M_1M_2M_3$ gegeben. Wir wählen einen Punkt M_0 und konstruieren den Punkt M_0', der in Bezug auf M_1 zu M_0 symmetrisch liegt.
Danach konstruieren wir den Punkt M_0'' der zu M_0' bezüglich M_2 symmetrisch liegt, und schließlich den zu M_0'' bezüglich M_3 symmetrischen Punkt M_0'''.
Sehen wir uns die fertige Zeichnung an, so stellen wir fest, dass die Punkte M_0 und M_0''' nicht übereinstimmen.



So ergibt sich das Problem: Kann man zu drei gegebenen Punkten M_1, M_2 und M_3 einen vierten Punkt M_0 so wählen, dass der bei der oben angegebenen Konstruktion entstehende Punkt M_0''' mit dem Punkt M_0 zusammenfällt?

Der Leser möge versuchen, die Aufgabe selbständig zu lösen, ehe er die nachstehende Lösung zur Kenntnis nimmt.

Lösung: Nehmen wir an, der gesuchte Punkt M_0 sei gefunden!
Wir wählen einen beliebigen Punkt P und zeichnen die Vektoren $\overrightarrow{PM_1}$, $\overrightarrow{PM_2}$ und $\overrightarrow{PM_3}$, ferner die Vektoren $\overrightarrow{PM_0}$, $\overrightarrow{PM_0'}$, $\overrightarrow{PM_0''}$ und $\overrightarrow{PM_0'''}$, die wir zur Abkürzung mit $\mathfrak{r}_1$, $\mathfrak{r}_2$, $\mathfrak{r}_3$ bzw. $\mathfrak{r}_0$, $\mathfrak{r}'_0$, $\mathfrak{r}''_0$, $\mathfrak{r}'''_0$ bezeichnen. (Man fertige eine Zeichnung an!)

Der Aufgabenstellung entsprechend ist M_1 Mittelpunkt der Strecke M_0M_0', daraus folgt (nach der Mittelpunktsregel für die Vektoraddition

$$\overrightarrow{PM_0} + \overrightarrow{PM_0'} = 2\overrightarrow{PM_1} \qquad \text{oder kürzer} \qquad \mathfrak{r}_0 + \mathfrak{r}'_0 = 2\mathfrak{r}_1$$

(Wir hatten ja $\overrightarrow{PM_0} = \mathfrak{r}_0$, $\overrightarrow{PM_0'} = \mathfrak{r}'_0$, $\overrightarrow{PM_1} = \mathfrak{r}_1$ gesetzt.) Wir erhalten also:

$$\mathfrak{r}'_0 = 2\mathfrak{r}_1 - \mathfrak{r}_0 \tag{A}$$

Berücksichtigen wir nun, dass nach den Voraussetzungen der Aufgabe M_0'' bezüglich M_2 zu M_0' symmetrisch liegt, und wiederholen wir die Überlegungen, die uns zur Gleichung (A) geführt haben, so gelangen wir zu der Gleichung

$$\mathfrak{r}''_0 = 2\mathfrak{r}_2 - \mathfrak{r}'_0$$

Ersetzen wir in dieser Gleichung $\mathfrak{r}'_0$ durch den Ausdruck (A), dann ergibt sich

$$\mathfrak{r}''_0 = 2\mathfrak{r}_2 - (2\mathfrak{r}_1 - \mathfrak{r}_0)$$

bzw., wenn wir die Klammern auflösen,

$$\mathfrak{r}''_0 = 2\mathfrak{r}_2 - 2\mathfrak{r}_1 + \mathfrak{r}_0 \tag{B}$$

Berücksichtigen wir schließlich noch die Symmetrie von M_0''' und M_0'' bezüglich M_3, so erhalten wir ganz analog wie oben:

$$\mathfrak{r}'''_0 = 2\mathfrak{r}_3 - \mathfrak{r}''_0$$

Setzen wir hier für $\mathfrak{r}''_0$ den Ausdruck (B) ein und lösen gleich die Klammern auf (nach den Regeln der Vektoralgebra), so folgt

$$\mathfrak{r}'''_0 = 2\mathfrak{r}_3 - 2\mathfrak{r}_2 + 2\mathfrak{r}_1 - \mathfrak{r}_0$$

Jetzt lässt sich aber die Bedingung dafür, dass die Punkte M_0''' und M_0 zusammenfallen, bereits ganz leicht angeben. Diese Punkte fallen zusammen, wenn der Vektor $\overrightarrow{PM_0'''} = \mathfrak{r}'''_0$ mit dem Vektor $\overrightarrow{PM_0} = \mathfrak{r}_0$ übereinstimmt, d.h., wenn die Beziehung

$$\mathfrak{r}_0 = 2\mathfrak{r}_3 - 2\mathfrak{r}_2 + 2\mathfrak{r}_1 - \mathfrak{r}_0$$

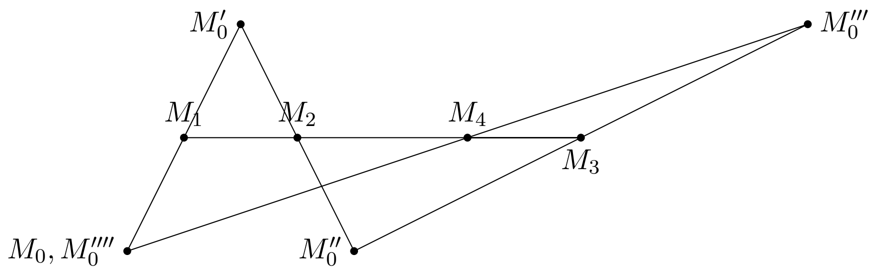
gilt. Nach einfachen Umformungen erhalten wir

$$\mathfrak{r}_3 + \mathfrak{r}_1 = \mathfrak{r}_2 + \mathfrak{r}_0 \qquad \text{bzw.} \qquad \overrightarrow{PM_3} + \overrightarrow{PM_1} = \overrightarrow{PM_2} + \overrightarrow{PM_0}$$

Diese Gleichung gilt aber dann und nur dann (oder, wie die Mathematiker sagen, genau dann), wenn der Mittelpunkt der Strecke M_3M_1 mit dem Mittelpunkt der Strecke M_2M_4 zusammenfällt, d.h., wenn die vier Punkte M_1, M_2, M_3 und M_4 ein Parallelogramm bilden, bei dem die Punkte M_1 und M_3 gegenüberliegende Ecken sind (vgl. Aufgabe 4).

Wenn also M_0''' und M_0 zusammenfallen sollen, müssen wir den Punkt M_0 so wählen, dass er die vierte Ecke des Parallelogramms ist, dessen übrige Ecken die Punkte M_1, M_2 und M_3 sind. Ferner müssen M_1 und M_3 gegenüberliegende Ecken sein.

Jetzt liegt natürlich der Gedanke nahe, die neuen vektoriellen Methoden zur Lösung einer ähnlichen, aber etwas komplizierteren Aufgabe anzuwenden.



Beispiel:
Zu vier vorgegebenen Punkten M_1, M_2, M_3, M_4 ist ein Punkt M_0 so zu bestimmen, dass bei aufeinanderfolgenden "Spiegelungen" dieses Punktes M_0 in Bezug auf die Punkte M_1, M_2, M_3 und M_4 ein Punkt M_0'''' entsteht, der mit dem Ausgangspunkt M_0 zusammenfällt.

Unter aufeinanderfolgenden Spiegelungen verstehen wir dabei folgendes: Zunächst konstruieren wir den Punkt M_0', der in Bezug auf M_1 zu M_0 symmetrisch liegt, dann den Punkt M_0'', der in Bezug auf M_2 zu M_0' symmetrisch ist usw.

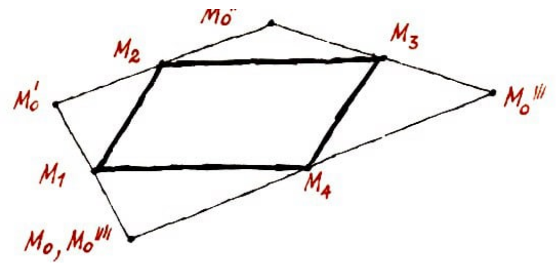
Wie man mühelos nachprüft (das sei dem Leser überlassen), ist

$$\overrightarrow{PM_0''''} = 2\overrightarrow{PM_4} - 2\overrightarrow{PM_3} + 2\overrightarrow{PM_2} - 2\overrightarrow{PM_1} + \overrightarrow{PM_0} \qquad (14)$$

Aus dieser Formel folgt, dass die Punkte M_0'''' und M_0 nur dann zusammenfallen, d.h. die Vektoren $\overrightarrow{PM_0''''}$ und $\overrightarrow{PM_0}$ übereinstimmen, wenn

$$\overrightarrow{PM_4} - \overrightarrow{PM_3} + \overrightarrow{PM_2} - \overrightarrow{PM_1} = \overrightarrow{0} \qquad (15)$$

ist.
Wenn also die Punkte M_1, M_2, M_3 und M_4 nicht auf einer Geraden liegen, so müssen sie ein Parallelogramm bilden, in dem M_1 und M_3 gegenüberliegende Ecken sind.



Liegen die Punkte jedoch auf einer Geraden, dann müssen sie so angeordnet sein, dass der Mittelpunkt der Strecke M_1M_3 auch der Mittelpunkt der Strecke M_2M_4 ist.
Wenn die Punkte M_1, M_2, M_3 und M_4 jedoch anders liegen, dann hat unsere Aufgabe keine Lösung. Haben aber die Punkte M_1, M_2, M_3 und M_4 die angegebene Lage, so geht die Gleichung (14) auf Grund von Gleichung (15) in $\overrightarrow{PM_0''''} = \overrightarrow{PM_0}$ über.
Daraus folgt, dass in diesem Falle nach einer viermaligen Spiegelung jeder beliebige Punkt M_0 in einen Punkt M_0'''' übergeht, der mit M_0 übereinstimmt.

Wir möchten darauf hinweisen, dass der Punkt M_0 hier sogar außerhalb der Ebene liegen kann, in welcher die vier Ausgangspunkte M_1, M_2, M_3 und M_4 liegen.
Vielleicht möchte der Leser jetzt noch wissen, welches Ergebnis man erhält, wenn nicht drei oder vier, sondern fünf, sechs, ... Punkte gegeben sind. Wir wollen dafür keine vollständige Lösung angeben, sondern nur bemerken, dass beispielsweise bei sechs Punkten $M_1, M_2, M_3, M_4, M_5, M_6$ ein nach 6 Spiegelungen wieder in sich übergehender Punkt M_0 nur dann existiert, wenn der Schnittpunkt der Seitenhalbierenden des Dreiecks $M_1M_3M_5$ mit dem Schnittpunkt der Seitenhalbierenden des Dreiecks $M_2M_4M_6$ zusammenfällt.

Wir stellen jetzt einige Aufgaben, bei deren Lösung sich der Leser davon überzeugen kann, ob er gelernt hat, die Vektoralgebra zur Lösung geometrischer Aufgaben zu verwenden.

1. Wir spiegeln einen beliebigen Punkt M_0 nacheinander in Bezug auf die Ecken eines Dreiecks $M_1M_2M_3$. Den dabei entstehenden Punkt M_0'''' (der, wie oben gezeigt wurde, nicht notwendig mit dem Punkt M_0 zusammenfällt) spiegeln wir wieder in Bezug auf die Punkte M_1, M_2 und M_3. Man beweise, dass der Punkt M_0 im Ergebnis dieser sechs aufeinanderfolgenden Spiegelungen in sich übergeht.

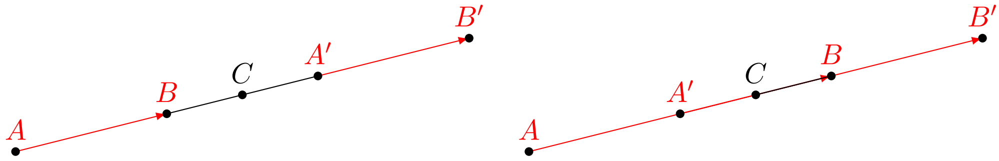
2. Wir spiegeln einen beliebigen Punkt P_0 nacheinander in Bezug auf die Ecken einer Pyramide $M_1M_2M_3M_4$. Den dabei entstehenden Punkt P_1 spiegeln wir wieder bezüglich der Eckpunkte M_1, M_2, M_3 und M_4 und erhalten einen Punkt P_2.
Diesen Punkt spiegeln wir nacheinander an den Ecken M_1, M_2, M_3 und M_4. Den dabei entstehenden Punkt P_3 spiegeln wir wieder in der angegebenen Weise usw.

Man beweise, dass die Punkte $P, P_1, P_2, P_3, P_4, ...$ auf einer Geraden liegen, und zwar so, dass jeder dieser Punkte (beginnend mit P_1) Mittelpunkt der Strecke ist, die durch den vorhergehenden und den nachfolgenden Punkt bestimmt ist.

2.8 Gleichheit von Vektoren

Wir wollen vereinbaren, den Vektor $\overrightarrow{AB}$ und den Vektor $\overrightarrow{A'B'}$ gleich zu nennen, wenn der Mittelpunkt der Strecke AB' mit dem Mittelpunkt der Strecke BA' zusammenfällt.

Liegen dabei die Punkte A, B, A' und B' nicht sämtlich auf einer Geraden, so ist das Viereck $ABB'A'$ offenbar ein Parallelogramm (als Viereck, dessen Diagonalen einander halbieren), in dem die Punkte A und B' gegenüberliegende Ecken sind.
Folglich ist die Länge der Strecke AB gleich der Länge der Strecke $A'B'$ und die Geraden AB und $A'B'$ sind einander parallel.

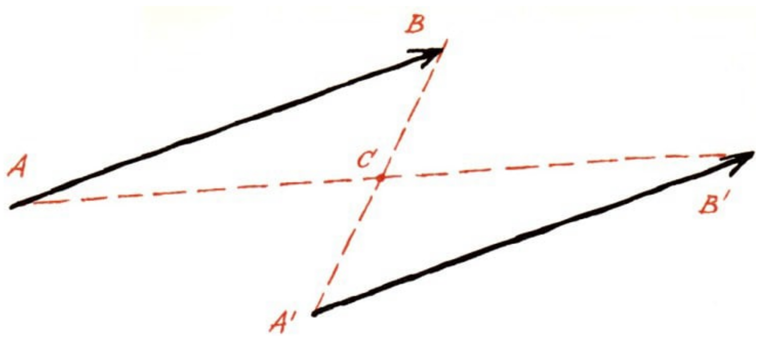


Rein anschaulich ist klar, dass dabei die Vektoren AB und $A'B'$ gleich gerichtet sind. Man kann nun die folgenden Sätze beweisen:

1. Wenn der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A'B'}$ geometrisch gleich ist, dann ist der Vektor $\overrightarrow{A'B'}$ dem Vektor $\overrightarrow{AB}$ geometrisch gleich.

2. Der Vektor $\overrightarrow{AB}$ ist dem Vektor $\overrightarrow{AB}$, d.h. sich selbst, geometrisch gleich.

3. Wenn der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A'B'}$ geometrisch gleich ist und der Vektor $\overrightarrow{A'B'}$ einem Vektor $\overrightarrow{A''B''}$ geometrisch gleich ist, dann ist der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A''B''}$ geometrisch gleich.



Die Richtigkeit der ersten beiden Sätze folgt unmittelbar aus der Definition der geometrischen Gleichheit von Vektoren. Zum Beweis des dritten Satzes bemerken wir folgendes:
Ist der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A'B'}$ geometrisch gleich, so gilt, wie man leicht sieht, die folgende Beziehung:

$$\overrightarrow{PA} + \overrightarrow{PB'} = \overrightarrow{PB} + \overrightarrow{PA'} \tag{16}$$

wobei P ein beliebiger Punkt ist. Es gilt aber auch die Umkehrung:
Wenn die Gleichung (16) erfüllt ist, so ist der Vektor $\overrightarrow{AB}$ geometrisch gleich dem Vektor $\overrightarrow{A'B'}$; denn Gleichung (16) gilt offenbar nur dann, wenn der Mittelpunkt der Strecke

AB' auch Mittelpunkt der Strecke BA' ist.

Da der Vektor $\overrightarrow{A'B'}$ dem Vektor $\overrightarrow{A''B''}$ geometrisch gleich ist, gilt die Beziehung

$$\overrightarrow{PA'} + \overrightarrow{PB''} = \overrightarrow{PB'} + \overrightarrow{PA''} \tag{17}$$

Durch Addition der Gleichungen (16) und (17) folgt (nach Vereinfachung)

$$\overrightarrow{PA} + \overrightarrow{PB''} = \overrightarrow{PB} + \overrightarrow{PA''}$$

und hieraus folgt (auf Grund der eben bewiesenen Umkehrung), dass der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A'B'}$ geometrisch gleich ist.
Der Leser hat natürlich bemerkt, dass die drei Sätze über die geometrische Gleichheit von Vektoren genauso aussehen wie die aus der Arithmetik bekannten Eigenschaften der Gleichheit von Zahlen:

1. Wenn $\alpha = \beta$ ist, dann ist $\beta = \alpha$.
2. $\alpha = \alpha$.
3. Wenn $\alpha = \beta$ ist und $\beta = \gamma$, dann ist $\alpha = \gamma$.

Um anzudeuten, dass der Vektor $\overrightarrow{AB}$ dem Vektor $\overrightarrow{A'B'}$ geometrisch gleich ist, benutzt man die Schreibweise

$$\overrightarrow{AB} = \overrightarrow{A'B'}$$

Unsere drei Sätze über die geometrische Gleichheit von Vektoren nehmen dann folgende Gestalt an:

1. Wenn $\overrightarrow{AB} = \overrightarrow{A'B'}$ ist, dann ist $\overrightarrow{A'B'} = \overrightarrow{AB}$.
2. $\overrightarrow{AB} = \overrightarrow{AB}$.
3. Wenn $\overrightarrow{AB} = \overrightarrow{A'B'}$ ist und $\overrightarrow{A'B'} = \overrightarrow{A''B''}$, dann ist $\overrightarrow{AB} = \overrightarrow{A''B''}$.

Gewöhnlich sagt man der Kürze halber, die Vektoren $\overrightarrow{AB}$ und $\overrightarrow{A'B'}$ seien gleich, wenn sie in dem oben definierten Sinne geometrisch gleich sind. Allerdings darf man dabei nicht vergessen, dass das Wort "gleich" hier eine ganz andere Bedeutung hat als in der Arithmetik.
Zwei (geometrisch) gleiche Vektoren stimmen also in Länge und Richtung überein, ihre Anfangs- und Endpunkte brauchen aber nicht übereinzustimmen. Unmissverständlicher wäre es natürlich, wenn man immer "geometrisch gleich" sagen würde.

Um den Unterschied in den Begriffsinhalten des hier in zwei Bedeutungen gebrauchten Wortes "gleich" hervorzuheben, bemerken wir folgendes.
Wenn in der Arithmetik eine reelle Zahl α von einer Zahl β verschieden, also α ungleich β, in Zeichen $\alpha \neq \beta$, ist, dann ist α entweder größer als β oder kleiner als β.

In der Vektoralgebra folgt jedoch aus der Ungleichheit der Vektoren $\overrightarrow{AB}$ und $\overrightarrow{A'B'}$

keineswegs, dass der Vektor $\overrightarrow{AB}$ größer oder kleiner ist als der Vektor $\overrightarrow{A'B'}$, weil in der Vektoralgebra die Beziehung "ein Vektor $\overrightarrow{AB}$ ist größer (bzw. kleiner) als ein Vektor $\overrightarrow{A'B'}$" nicht erklärt ist.

Dem widerspricht nicht, dass die Aussage "die Länge des Vektors $\overrightarrow{AB}$ ist größer als die Länge des Vektors $\overrightarrow{A'B'}$" einen völlig bestimmten Sinn hat. Die Länge des Vektors ist ja eine wohlbestimmte Zahl.

Wie konstruiert man einen Vektor, der einem gegebenen Vektor gleich ist?
Es sei $\overrightarrow{AB}$ der vorgegebene Vektor und M ein beliebiger Punkt. Wir konstruieren die Summe $\overrightarrow{AB}$ und $\overrightarrow{AM}$ (man fertige eine Zeichnung an). Es sei

$$\overrightarrow{AB} + \overrightarrow{AM} = \overrightarrow{AN} \tag{17a}$$

Beachten wir, dass der Mittelpunkt der Strecke AN mit dem Mittelpunkt der Strecke BM zusammenfällt, so ergibt sich, dass die Vektoren $\overrightarrow{AB}$ und $\overrightarrow{MN}$ geometrisch gleich sind, es ist also $\overrightarrow{AB} = \overrightarrow{MN}$.
Damit ist der gesuchte Vektor konstruiert.

Wie man leicht sieht, ist diese Konstruktion eindeutig, d.h., es gibt nur einen einzigen Vektor, dessen Anfangspunkt im Punkte M liegt, und der den Vektor $\overrightarrow{AB}$ geometrisch gleich ist.
Gäbe es nämlich einen von N verschiedenen Punkt N' und wäre auch $\overrightarrow{AB} = \overrightarrow{MN'}$, dann müsste die Gleichung

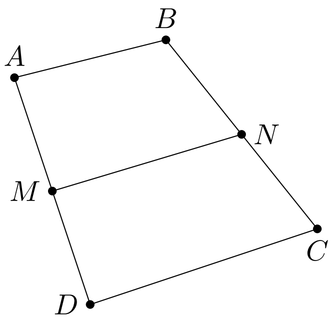
$$\overrightarrow{AB} + \overrightarrow{AM} = \overrightarrow{AN'}$$

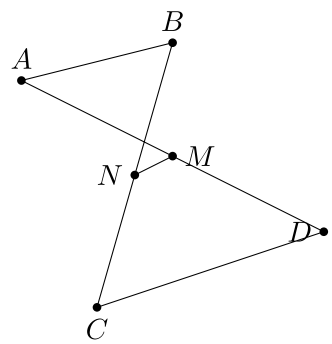
gelten, die aber der Gleichung (17a) widerspricht.

2.9 Die Mittellinien im Viereck

Als Mittellinie eines Trapezes bezeichnet man bekanntlich die Strecke, welche die Mittelpunkte der nichtparallelen Seiten verbindet! In einem beliebigen Viereck $ABCD$ bezeichnet;man als Mittellinien diejenigen Strecken, welche die Mittelpunkt gegenüberliegender Seiten verbinden (etwa von AD und BC).
Der von uns schon benutzte Satz kann jetzt folgendermaßen formuliert werden:

Die beiden Mittellinien in einem Viereck halbieren einander.





Aufgabe:
Es sei $ABCD$ ein ganz beliebiges Viereck. Seine Seiten dürfen sich sogar überschneiden, oder es darf windschief sein (d.h., seine Eckpunkte brauchen nicht in einer Ebene zu liegen).

Wir wählen irgend zwei Punkte A' und D' und konstruieren die Vektoren $\overrightarrow{A'B'}$ und $\overrightarrow{C'D'}$, die den Vektoren $\overrightarrow{AB}$ bzw. $\overrightarrow{CD}$ gleich sind.
Dann ziehen wir die Mittellinie MN, die im Viereck $ABCD$ die Mittelpunkte der Seiten AD und BC verbindet, sowie im Viereck $A'B'C'D'$ die Mittellinie $M'N'$, welche die Mittelpunkte der Seiten $A'D'$ und $B'C'$ verbindet.

Man beweise, dass die Mittellinien MN und $M'N'$ gleich lang sind und auf parallelen Geraden (bzw. auf ein und derselben Geraden) liegen.

Zum Beweis braucht man nur nachzuprüfen, dass die Vektoren $\overrightarrow{MN}$ und $\overrightarrow{M'N'}$ geometrisch gleich sind, d.h., dass

$$\overrightarrow{PM} + \overrightarrow{PN'} = \overrightarrow{PN} + \overrightarrow{PM'} \tag{18}$$

gilt, wobei P ein beliebiger Punkt ist (deshalb ist der Punkt P nicht eingezeichnet).
Wir beachten, dass nach Voraussetzung die Beziehungen

$$\overrightarrow{PM} = \frac{1}{2}(\overrightarrow{PA} + \overrightarrow{PD}) \quad ; \quad \overrightarrow{PN} = \frac{1}{2}(\overrightarrow{PB} + \overrightarrow{PC})$$
$$\overrightarrow{PM'} = \frac{1}{2}(\overrightarrow{PA'} + \overrightarrow{PD'}) \quad ; \quad \overrightarrow{PN'} = \frac{1}{2}(\overrightarrow{PB'} + \overrightarrow{PC'})$$

gelten. Daher kann man die dem Beweis zugrunde liegende Gleichung (18) nach Multiplikation mit 2 auch folgendermaßen schreiben:

$$\overrightarrow{PA} + \overrightarrow{PD} + \overrightarrow{PB'} + \overrightarrow{PC'} = \overrightarrow{PB} + \overrightarrow{PC} + \overrightarrow{PA'} + \overrightarrow{PD'} \tag{19}$$

Nun ist aber nach Voraussetzung

$$\overrightarrow{AB} = \overrightarrow{A'B'} \qquad \text{und} \qquad \overrightarrow{DC} = \overrightarrow{D'C'}$$

also erhalten wir:

$$\overrightarrow{PA} + \overrightarrow{PB'} = \overrightarrow{PB} + \overrightarrow{PA'} \quad , \quad \overrightarrow{PD} + \overrightarrow{PC'} = \overrightarrow{PC} + \overrightarrow{PD'}$$

Addition dieser beiden Gleichungen liefert (19).

Weitere Sätze der Vektoralgebra sowie Beispiele für Anwendungen zur Lösung geometrischer Aufgaben bzw. zum Beweis geometrischer Sätze kann man der Spezialliteratur entnehmen.
Zur Übung mögen nachstehende Aufgaben dienen.

1. Die Punkte $A_0, A_1, A_2, A_3, A_4, ...$ liegen so auf einer Geraden, dass sie jeweils Mittelpunkt der Strecke sind, die zwischen dem vorhergehenden und dem nachfolgenden Punkt liegt. Man beweise die Beziehungen

$$\overrightarrow{PA_1} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot 1, \qquad \overrightarrow{PA_2} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot 2,$$

$$\overrightarrow{PA_3} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot 3 \text{ usw.}$$

2. Auf der Strecke A_0A_1 liegen Punkte $B_0, B_1, B_2, ..., B_{n-1}, B_n$ so, dass B_0 gleich A_0 und B_n gleich A_1 ist, während jeder der übrigen Punkte Mittelpunkt derjenigen Strecke ist, welche vom vorhergehenden und vom nachfolgenden Punkt bestimmt wird.
Man beweise die Beziehungen.

$$\overrightarrow{PB_1} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{1}{n}, \qquad \overrightarrow{PB_2} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{2}{n}$$

$$\overrightarrow{PB_3} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{3}{n} \text{ usw.}$$

3. Wie liegen die Punkte $C_1, C_2, C_3, ...$, wenn die nachstehenden Beziehungen bekannt sind:

$$\overrightarrow{PC_1} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{1}{n}, \qquad \overrightarrow{PC_2} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{2}{n}$$

$$\overrightarrow{PC_3} = \overrightarrow{PA_0} + (\overrightarrow{PA_1} - \overrightarrow{PA_0}) \cdot \frac{3}{n} \text{ usw.}$$

4. Man beweise: Ist der Punkt C der Schnittpunkt der Seitenhalbierenden eines Dreiecks $A_1A_2A_3$, so gilt

$$\overrightarrow{PC} = \frac{1}{3}(\overrightarrow{PA_1} + \overrightarrow{PA_2} + \overrightarrow{PA_3})$$

wobei P ein beliebiger Punkt ist. Der Punkt C heißt der Schwerpunkt des Systems der drei Punkte A_1, A_2 und A_3.

5. Wir verbinden jeden der vier Punkte A_1, A_2, A_3 und A_4 mit dem Schwerpunkt des Systems der übrigen drei Punkte. Man beweise:
Die dabei entstehenden vier Geraden schneiden einander in einem Punkt C, für den

$$\overrightarrow{PC} = \frac{1}{4}(\overrightarrow{PA_1} + \overrightarrow{PA_2} + \overrightarrow{PA_3} + \overrightarrow{PA_4})$$

gilt, wobei P ein beliebiger Punkt ist. Der Punkt C heißt der Schwerpunkt des Systems der vier Punkte A_1, A_2, A_3 und A_4. Er teilt jede der Strecken im Verhältnis $1 : 3$.

6. Man verallgemeinere die vorhergehende Aufgabe auf den Fall, dass fünf Punkte, sechs Punkte usw. gegeben sind.

2.10 Zahlen und Rechenoperationen

W. G. Boltjanski

Auf die Frage, womit sich die Algebra beschäftigt, würde man wahrscheinlich ganz verschiedene Antworten erhalten.
Man könnte z.B. erfahren, dass man in der Algebra Gleichungen löst - und mit Hilfe der Gleichungen praktische Aufgaben. Und das wäre gar nicht falsch. Nehmen doch die mit der Lösung von Gleichungen der verschiedenartigsten Formen zusammenhängenden Probleme im Algebra-Unterricht recht breiten Raum ein.

Und dennoch ist diese Antwort unvollständig. Will man nämlich etwa die Gleichung

$$\frac{1}{x^2-x}+\frac{2x}{x^2-2x+1}=\frac{9}{x}$$

lösen, so muss man dabei die Faktorzerlegung beherrschen, um die Brüche auf den Hauptnenner bringen zu können. Man muss Klammern auflösen und gleiche Glieder zusammenfassen können und anderes mehr. Mit anderen Worten, man muss verschiedene (Rechen-)Operationen richtig ausführen können.

Die zur Lösung einer Gleichung notwendigen (Rechen-)Operationen werden an Zahlen ausgeführt. Daran ändert auch der Umstand nichts, dass in der Gleichung Buchstaben (sogenannte Variable) vorkommen, die für bekannte oder unbekannte Größen stehen. In der Algebra dienen die Buchstaben zur Bezeichnung von Zahlen.

Um aber die Operationen mit den Zahlen ausführen zu können, muss man die Regeln für diese Operationen, d.h. die Rechengesetze, kennen. Mit der Untersuchung dieser Regeln befasst sich ein bedeutender Teil des Algebra-Unterrichts. Beispiele dafür sind das Distributivgesetz

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

auf dem das Auflösen von Klammern und das Ausklammern beruht, die Potenzgesetze

$$a^b\cdot a^c=a^{b+c},\qquad (a^b)^c=a^{bc},\qquad \text{usw.}$$

Die beiden Hauptgegenstände des Algebra-Unterrichtes sind die Lehre von den Gleichungen und die Lehre von den Eigenschaften der Rechenoperationen.
Zwar behandelt man heute in der (Schul-)Algebra ein weiteres Gebiet - die Lehre von den Funktionen und ihren grafischen Darstellungen, aber dieses Gebiet gehört im Grunde nicht zur Algebra und wird vielleicht in Zukunft nicht mehr im Algebra-Unterricht behandelt werden, sondern in einem besonderen Fach, der Funktionenlehre.

Die Unterteilung des Problemkreises der Algebra in zwei Gebiete - einerseits in die Theorie der Polynome und Gleichungen, andererseits in die Lehre von den Eigenschaften der Rechenoperationen - spielt nicht nur im Algebra-Unterricht in der Schule eine Rolle, sondern auch in den Vorlesungen an den Universitäten, ja sogar in den Arbeiten der heutigen Wissenschaftler.
Selbstverständlich hängen diese beiden Gebiete eng miteinander zusammen.

Wieviel Arten von Zahlen braucht man zur Lösung von Gleichungen?
Die einfachsten Gleichungen sind die Gleichungen ersten Grades (die "linearen Gleichungen") mit einer Unbekannten, Die Zahlen, die uns am geläufigsten sind, sind die ganzen positiven Zahlen, mit denen wir seit frühester Kindheit durch das Abzählen von Gegenständen vertraut sind.

Würden wir nur diejenigen Gleichungen ersten Grades mit einer Unbekannten untersuchen, deren Koeffizienten ganze positive Zahlen sind, so könnten wir mit ihrer Hilfe eine Vielzahl von Aufgaben lösen, wie sie im Unterricht der ersten Schulklassen gestellt werden. Dort stellt man die Aufgaben so, dass auch die Lösungen ganze positive Zahlen sind.
Andere Zahlen kennen die Schüler auf dieser Stufe nicht.

Kann man nun daraus ableiten, dass man außer den ganzen positiven Zahlen überhaupt keine anderen Zahlen mehr zu kennen braucht, um jede beliebige Gleichung lösen zu können? Natürlich nicht! Betrachten wir beispielsweise die Gleichung

$$ax + b = c \tag{1}$$

deren Koeffizienten a, b und c ganze positive Zahlen sind. Auf diese Gestalt kann man jede Gleichung ersten Grades mit ganzen positiven Koeffizienten bringen. Zu ihrer Lösung bringt man zuerst die Zahl b auf die rechte Seite der Gleichung, d.h., man bringt die Gleichung auf die Form $ax = c - b$, und dividiert dann die Differenz $c - b$ durch a; als Ergebnis erhält man

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

Wenn man also jede Gleichung mit ganzen positiven Koeffizienten lösen will, dann muss man zuerst einmal die Differenz $c-b$ ausrechnen können. Das zwingt zur Einführung der negativen (ganzen) Zahlen, weil für $b > c$ die größere Zahl von der kleineren Zahl abgezogen werden muss.
Zweitens muss man die Differenz $c - b$ durch a dividieren können. Damit gelangt man zur Einführung der rationalen Zahlen. Die Menge der rationalen Zahlen umfasst die positiven und die negativen ganzen Zahlen und die positiven und die negativen Brüche sowie die - bisher noch nicht erwähnte - Null.

Zur zahlenmäßigen ("numerischen") Lösung beliebiger Gleichungen ersten Grades mit ganzen positiven Koeffizienten braucht man somit als "Zahlenvorrat" alle rationalen Zahlen.
Das war einer der Gründe für die Einführung der gebrochenen und der negativen Zahlen in der Mathematik. Hat man alle rationalen Zahlen zur Verfügung, so kann man nicht nur Gleichungen ersten Grades mit ganzen positiven Koeffizienten lösen, sondern Gleichungen ersten Grades mit beliebigen rationalen Koeffizienten.
Ausgenommen bleibt nur der Fall, dass der Koeffizient a in der Gleichung (1) gleich Null ist; aber dann kommt ja in dieser "Gleichung" die Unbekannte x gar nicht vor. Mit anderen Worten, zur Lösung von Gleichungen ersten Grades mit rationalen Koeffizienten braucht man nur die rationalen Zahlen zu kennen.

Man überlegt sich leicht, woraus sich diese (für die Lösung von Gleichungen ersten Grades) angenehmen Eigenschaften der rationalen Zahlen erklären. Zur Lösung von Gleichungen ersten Grades bzw. von Systemen solcher Gleichungen braucht man nämlich nur die vier Grundrechenarten ausführen zu können: die Addition, die Subtraktion, die Multiplikation und die Division (selbst verständlich mit Ausnahme der nicht erklärten Division durch Null), und in der Menge der rationalen Zahlen sind diese Operationen immer ausführbar.

Eine Menge von Zahlen, in welcher die vier Grundrechenarten (mit Ausnahme der Division durch Null) unbeschränkt ausführbar sind, nennen die Mathematiker einen Zahlkörper.
Die Menge der rationalen Zahlen ist somit ein Zahlkörper. Die Menge aller ganzen positiven Zahlen ist dagegen kein Zahlkörper: Subtraktion und Division sind in ihr nicht immer ausführbar (z.B. sind die Aufgaben $2 - 5$ und $3 : 7$ nicht ausführbar, wenn man nur die ganzen positiven Zahlen kennt).

Um die Subtraktion immer ausführen zu können, muss man zu den ganzen positiven Zahlen noch die Zahl "Null" und alle ganzen negativen Zahlen hinzunehmen. Wir erhalten so die Menge aller ganzen Zahlen. Aber auch sie ist kein Zahlkörper.
Zwar kann man in ihr stets die Addition, die Subtraktion und die Multiplikation ausführen, jedoch ist die Division nicht immer ausführbar. So ist z.B. die Operation $2 : 7$ nicht realisierbar, wenn man nur die ganzen Zahlen zur Verfügung hat. Nehmen wir nun zu den ganzen Zahlen noch die gebrochenen Zahlen hinzu, so erhalten wir die Menge aller rationalen Zahlen, die bereits ein Zahlkörper ist.
Mehr darüber, was eine Menge ist, erfahren Sie in dem Abschnitt "Der Mengenbegriff".

2.11 Neue Zahlen werden gebraucht

Nun darf man jedoch nicht etwa annehmen, der Körper der rationalen Zahlen enthalte bereits soviel Zahlen, dass sich in diesem Bereich zu jeder algebraischen Gleichung eine Wurzel finden ließe. Was eine algebraische Gleichung ist, setzen wir aus der Schule als bekannt voraus. Gibt es doch außer den Gleichungen ersten Grades noch quadratische Gleichungen und Gleichungen höheren Grades.
Die Gleichung

$$x^2 - 3 = 0 \tag{2}$$

besitzt beispielsweise im Körper der rationalen Zahlen keine Lösung. Es gibt nämlich keine rationale Zahl, deren Quadrat gleich 3 wäre.
Wenn wir erreichen wollen, dass diese Gleichung eine Lösung hat, dann müssen wir den Körper der rationalen Zahlen erweitern um eine Zahl, deren Quadrat gleich 3 ist. (Wir stellen uns hier auf einen formalen, algebraischen Standpunkt. Die Zulässigkeit dieses Verfahrens muss natürlich bewiesen werden.)

Man bezeichnet diese Zahl mit $\sqrt{3}$. Nimmt man jedoch zum Körper der rationalen Zahlen nur die eine neue Zahl $\sqrt{3}$ hinzu, so ist damit noch nicht viel erreicht; denn man erhält dadurch noch keinen Zahlkörper.
Wollen wir zu einem Zahlkörper gelangen, so müssen wir neben der Zahl $\sqrt{3}$ auch die Zahlen $2\sqrt{3}$, $\frac{1}{3}\sqrt{3}$, $-\frac{1}{3}\sqrt{3}$ usw. zum Körper der rationalen Zahlen hinzufügen, d.h. alle Zahlen der Gestalt $b\sqrt{3}$, wobei b eine beliebige rationale Zahl ist.

Im Körper der rationalen Zahlen muss ja die Multiplikation immer ausführbar sein.

Außerdem müssen wir aber auch noch die Zahlen $1-\sqrt{3}$, $5+\frac{1}{3}\sqrt{3}$, $\frac{7}{11}+\frac{1}{5}\sqrt{3}$ usw. hinnehmen, d.h. alle Zahlen der Gestalt $a+b\sqrt{3}$, wobei a und b beliebige rationale Zahlen sind; müssen doch im Körper der rationalen Zahlen die Addition und die Subtraktion immer ausführbar sein!

Interessanterweise braucht man nun nichts weiter hinzuzunehmen. Die Zahlen der Gestalt $a+b\sqrt{3}$ mit beliebigen rationalen Zahlen a und b, bilden einen Zahlkörper, d.h., man kann in der Menge aller dieser Zahlen die vier Grundrechenarten ausführen.

Die Division lässt sich beispielsweise so ausführen, dass im Nenner keine Irrationalität mehr vorkommt:

$$\frac{\frac{3}{2}+\sqrt{3}}{1-\frac{1}{2}\sqrt{3}} = \frac{\left(\frac{3}{2}\sqrt{3}\right)\left(1+\frac{1}{2}\sqrt{3}\right)}{\left(1-\frac{1}{2}\sqrt{3}\right)\left(1+\frac{1}{2}\sqrt{3}\right)} = \left(\frac{3}{2}+\sqrt{3}\right)\left(1+\frac{1}{2}\sqrt{3}\right) : \frac{1}{4} = 12+7\sqrt{3}$$

Zum Körper der Zahlen $a+b\sqrt{3}$ gelangten wir, ausgehend vom Körper der rationalen Zahlen, indem wir diesen um eine Menge neuer Zahlen erweiterten. Er heißt deshalb eine Erweiterung des Körpers der rationalen Zahlen. Dieser Körper hat viele interessante Eigenschaften.

Sehen wir uns einmal nur die Teilmenge der "ganzen" Zahlen der Gestalt $a+b\sqrt{3}$ an, d.h. die Menge der Zahlen $a+b\sqrt{3}$ mit ganzzahligen Koeffizienten a und b. Wir können diese Zahlen (wie die gewöhnlichen zusammengesetzten ganzen Zahlen) in Faktoren zerlegen.
Überraschenderweise lässt unter diesen Zahlen (d.h. unter den Zahlen $a+b\sqrt{3}$ mit ganzzahligen Koeffizienten a und b) auch die Eins eine Faktorzerlegung zu! Hier sind Beispiele für derartige Zerlegungen:

$$1=(2-\sqrt{3})(2+\sqrt{3}) \qquad , \qquad 1=(7-4\sqrt{3})(7+4\sqrt{3})$$

Jetzt wird auch klar, dass die Frage: "Hat die Gleichung $x^2-3=0$ eine Wurzel?" ohne den Hinweis darauf, in welchem Körper wir ihre Wurzeln bestimmen sollen, sinnlos ist. Wollten wir die Wurzeln im Körper der rationalen Zahlen suchen, dann wäre die Antwort verneinend: Es gibt keine rationale Zahl, deren Quadrat gleich 3 ist. Im Körper der Zahlen der Gestalt $a+b\sqrt{3}$ (a, b rational) dagegen ist die Antwort positiv: Die Gleichung $x^2-3=0$ besitzt die Wurzeln $x_1=\sqrt{3}$ und $x_2=-\sqrt{3}$.

Genauso muss man zur Beantwortung der Frage: "Kann man ein gegebenes Polynom in Faktoren zerlegen?" wissen, in welchem Körper die Koeffizienten der Faktorpolynome liegen sollen.
So ist z.B. im Körper der rationalen Zahlen das Polynom x^2-3 nicht in Faktoren zerlegbar. Es ist, wie man auch sagt, irreduzibel.

Im Körper der Zahlen $a + b\sqrt{3}$ dagegen zerfällt dieses Polynom in Faktoren:

$$x^2 - 3 = (x - \sqrt{3})(x + \sqrt{3})$$

2.12 Verdoppelung des Würfelvolumens

Ausgehend von der Gleichung (2) gelangten wir zum Körper der Zahlen der Gestalt $a+b\sqrt{3}$. Wir wollen jetzt eine weitere quadratische Gleichung untersuchen, nämlich

$$x^2 - 5 = 0 \qquad (3)$$

Man kann beweisen, dass diese Gleichung nicht nur im Körper der rationalen Zahlen keine Wurzel besitzt, sondern auch im Körper der Zahlen $a+b\sqrt{3}$. Um einen Körper zu erhalten, in dem die Gleichung (3) Wurzeln hat, muss man zum Körper der rationalen Zahlen die Zahl $\sqrt{5}$ hinzunehmen, ferner alle Zahlen der Gestalt $a + b\sqrt{5}$ mit beliebigen rationalen Zahlen a und b.

Wenn man zu einem Zahlkörper gelangen will, in dem sowohl die Gleichung (2) als auch die Gleichung (3) lösbar ist, so muss man zum Körper der rationalen Zahlen die beiden Zahlen $\sqrt{3}$ und $\sqrt{5}$ sowie alle Zahlen der Gestalt $a + b\sqrt{+c}\sqrt{5} + d\sqrt{15}$ hinzunehmen, wobei a, b, c und d beliebige rationale Zahlen sind.

Der Leser möge selbst nachprüfen, dass diese Zahlen tatsächlich einen Zahlkörper bilden (d.h., dass in der Menge dieser Zahlen die vier Grundrechenarten ausführbar sind).

Eine bedeutende Vergrößerung des "Zahlenvorrats" erreichen wir, wenn wir alle aus den rationalen Zahlen mit Hilfe der vier Grundrechenarten und des Quadratwurzelziehens entstehenden Zahlen betrachten. Die sich so ergebenden Zahlen nennen wir quadratische Irrationalitäten. Darunter verstehen wir solche Zahlen wie

$$1+\sqrt{3}, \quad 2+\frac{2}{3}\sqrt{5}-\frac{1}{11}\sqrt{17}, \quad \sqrt{1+\sqrt{2}}, \quad \sqrt[4]{3} = \sqrt{\sqrt{3}}, \quad \sqrt[4]{1+\sqrt[8]{\sqrt{3}+\sqrt{11}}}, \text{ u.a.m.}$$

Die Menge aller quadratischen Irrationalitäten ist ein Zahlkörper.
Außerdem ist die Quadratwurzel aus einer beliebigen positiven Zahl, die eine quadratische Irrationalität ist, wieder eine quadratische Irrationalität. Daraus folgt, dass die Wurzeln jeder quadratischen Gleichung, deren Koeffizienten quadratische Irrationalitäten sind, ebenfalls quadratische Irrationalitäten sind.
(Damit diese Wurzeln reell sind, müssen wir noch voraussetzen, dass die Diskriminante der quadratischen Gleichung positiv ist.)

Zur Lösung quadratischer Gleichungen braucht man nämlich nur die vier Grundrechenarten und das Quadratwurzelziehen, und diese Operationen sind im Körper der quadratischen Irrationalitäten stets ausführbar. Somit kann man im Körper der quadratischen Irrationalitäten nicht nur die vier Grundrechenarten ausführen (d.h. Gleichungen ersten

Grades lösen), sondern auch Quadratwurzeln ziehen, d.h. auch quadratische Gleichungen (mit positiver Diskriminante, falls die Wurzeln reell sein sollen) lösen.

Zwischen dem Körper der quadratischen Irrationalitäten und den geometrischen Konstruktionen mit Zirkel und Lineal besteht ein enger Zusammenhang.
Es sei etwa eine Strecke a gegeben und die Aufgabe gestellt, ausgehend von dieser Strecke mit Zirkel und Lineal eine geometrische Konstruktion auszuführen. Beispielsweise sei das regelmäßige Zehneck mit der Seite a zu konstruieren oder ein gleichseitiges Dreieck, dessen Seitenhalbierende gleich a ist, usf.

Wenn wir eine Konstruktionsaufgabe mit Zirkel und Lineal lösen, dann wiederholen wir (in bestimmter Reihenfolge) immer nur drei Operationen:

1. Wenn schon vier Punkte M, N, P und Q konstruiert sind, ist der Schnittpunkt der Geraden MN und PQ zu bestimmen. Diese Operation wird mit dem Lineal ausgeführt.

2. Drei Punkte M, N und O und die Strecke l sind bereits konstruiert. Auf der Geraden MN ist der Punkt A zu bestimmen, derart, dass die Länge der Strecke AO gleich l ist. Dazu braucht man ein Lineal (um die Gerade MN zu ziehen) und einen Zirkel (um einen Kreis um O zu schlagen).

3. Zwei Punkte Q_1 und Q_2 sowie zwei Strecken l_1 und l_2 sind bereits konstruiert. Zu bestimmen ist ein Punkt A, dessen Abstand vom Punkte Q_1 bzw. Q_2 gleich l_1 bzw. gleich l_2 ist. Diese Operation wird mit dem Zirkel ausgeführt.

Mit Hilfe der drei genannten Operationen können wir immer neue Punkte und Strecken konstruieren, bis wir die verlangte Figur fertiggestellt haben.
Wie einfache Rechnungen zeigen (man braucht dazu nur den Satz des Pythagoras, die Heronische Formel und andere Sätze aus der Schulgeometrie), haben alle bei der Ausführung dieser drei Operationen entstehenden Strecken Längen, die sich durch die Längen der Ausgangsstrecken mit Hilfe der vier Grundrechenarten und des Quadratwurzelziehens ausdrücken lassen.

Wenn wir die Länge der ursprünglich gegebenen Strecke a als Einheit nehmen, sind die Längen aller Strecken, die wir mit Zirkel und Lineal konstruieren können, quadratische Irrationalitäten. Es gilt aber auch die Umkehrung:
Jede Strecke, deren Länge gleich einer quadratischen Irrationalität ist, kann mit Zirkel und Lineal konstruiert werden.

Man kann also an Hand von Berechnungen feststellen, ob eine Aufgabe mit Zirkel und Lineal zu lösen ist oder nicht. Wenn die Rechnung ergibt, dass die Längen aller Strecken, die man zur Lösung der Aufgabe konstruieren muss, quadratische Irrationalitäten sind, dann ist die Aufgabe mit Zirkel und Lineal lösbar.
Zeigt die Rechnung jedoch, dass die Aufgabe auf die Konstruktion einer Strecke führt, deren Länge keine quadratische Irrationalität ist, dann lässt sich diese Aufgabe nicht mit Zirkel und Lineal lösen.

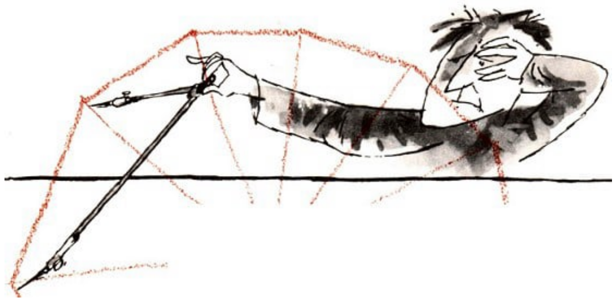
Wenn wir etwa ein regelmäßiges Fünfeck mit der Seitenlänge l konstruieren wollen, so müssen wir dazu die Länge der Diagonale im Fünfeck kennen (dann lassen sich die

Eckpunkte mit dem Zirkel markieren).
Wie die Rechnung zeigt, ist die Länge der Diagonale in einem derartigen Fünfeck gleich $\frac{1+\sqrt{5}}{2}$. Weil diese Zahl eine quadratische Irrationalität ist, kann man das regelmäßige Fünfeck mit Zirkel und Lineal konstruieren.

Der große deutsche Mathematiker C.F.Gauß (1777-1855) hat bewiesen, dass ein regelmäßiges p-Eck (p eine Primzahl) mit vorgegebener Seitenlänge a dann und nur dann konstruiert werden kann, wenn p die Gestalt

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

hat (n eine natürliche Zahl). Beispielsweise kann man ein regelmäßiges Siebzehneck mit Zirkel und Lineal konstruieren; denn es ist $17 = 2^{2^2} + 1$. Dagegen ist das regelmäßige Elfeck mit Zirkel und Lineal nicht konstruierbar.



In der Antike hatten viele Mathematiker versucht, das folgende "Delische Problem" zu lösen. Das Orakel in Delos hat der Überlieferung zufolge diese Aufgabe gestellt:

Gegeben ist ein Würfel mit der Kantenlänge a. Man Konstruiere den Würfel, dessen Volumen doppelt so groß ist wie das des gegebenen Würfels.
Wir wollen berechnen, welche Strecke man zur Lösung dieses Problems konstruieren müsste.

Als Einheit nehmen wir die Länge der Strecke a und bezeichnen die Kantenlänge des gesuchten Würfels mit x. Das Volumen des gegebenen Würfels ist dann gleich Eins, das Volumen des gesuchten Würfels gleich x^3.
Nach den Bedingungen der Aufgabe ist

$$x^3 = 2$$

und das bedeutet, dass $x = \sqrt[3]{2}$ ist. Die Zahl $\sqrt[3]{2}$ ist aber keine quadratische Irrationalität. Daher ist die Aufgabe, das Volumen Würfels zu verdoppeln, mit Zirkel und Lineal nicht lösbar.

2.13 Algebraische und nichtalgebraische (transzendente) Zahlen

Bisher haben wir uns für die Lösung von linearen Gleichungen bzw. von quadratischen Gleichungen interessiert. Wir wollen uns nun Gleichungen höheren Grades zuwenden.

Nimmt man zum Körper der rationalen Zahlen alle diejenigen reellen Zahlen hinzu, welche Wurzeln aller möglichen algebraischen Gleichungen (beliebigen Grades) mit ganzen Koeffizienten sind, so erhält man eine neue Erweiterung des Körpers der rationalen Zahlen.

Die zu diesem Körper gehörenden Zahlen heißen algebraische Zahlen. Speziell sind die quadratischen Irrationalitäten algebraische Zahlen (wie überhaupt alle reellen Zahlen, die sich durch Radikale ausdrücken lassen, z.B. $\sqrt[3]{2}$).

Selbst wenn wir den Körper aller algebraischen Zahlen bilden, erhalten wir noch nicht alle reellen Zahlen, d.h. alle Zahlen, welche wir als Punkte auf der Zahlengeraden abbilden.
Überdies gibt es unter den reellen Zahlen viel mehr nichtalgebraische Zahlen als algebraische, obwohl es sowohl unendlich viele algebraische als auch unendlich viele nichtalgebraische Zahlen gibt (vgl. dazu den Abschnitt "Der Mengenbegriff").

Um dies zu veranschaulichen, denken wir uns die Zahlengerade weiß gefärbt und danach die den algebraischen Zahlen entsprechenden Punkte mit schwarzer Farbe gekennzeichnet. Dann bleiben weiße "Lücken" übrig (nichtgeschwärzte Punkte), und zwar kann man diese Lücken überall entdecken, in jedem beliebigen Intervall, und sei es noch so klein.

Die nichtalgebraischen Zahlen werden auch transzendente Zahlen genannt. Von dem französischen Mathematiker Joseph Liouville (1809-1882) stammt ein Verfahren, mit dem man beliebig viele transzendente Zahlen angeben kann.
Dagegen ist es ein außerordentlich schwieriges und in den meisten Fällen bisher ungelöstes Problem, festzustellen, ob eine vorgegebene Zahl algebraisch oder transzendent ist.

Der französische Mathematiker Charles Hermite (1822-1901) zeigte, dass die Zahl e transzendent ist. Der deutsche Mathematiker Ferdinand Lindemann (1852-1939) hat 1882 bewiesen, dass die Zahl π transzendent ist, d.h. nicht Wurzel einer algebraischen Gleichung mit ganzen Koeffizienten ist.
Daraus folgt insbesondere, dass π keine quadratische Irrationalität ist, und damit ist bewiesen, dass die "Quadratur des Kreises", d.h. die Verwandlung eines Kreises in ein inhaltsgleiches Quadrat, mit Zirkel und Lineal nicht ausführbar ist.

Der sowjetische Mathematiker A.O. Gelfond hat folgendes bewiesen:
Sind a und b positive algebraische Zahlen, so ist die Zahl $\log_a b$ entweder rational oder transzendent, keinesfalls aber algebraisch und nicht rational.

So sind beispielsweise die Zahlen $\log_2 3$, $\log_{10} 5$, $\log_{10} 27$ usw. transzendent. Bis auf die Logarithmen der Zahlen 1, 10, 100, 1000, ..., $\frac{1}{10}$, $\frac{1}{100}$, $\frac{1}{1000}$, ... sind die Briggschen Logarithmen aller Zahlen transzendent.

Die Menge aller reellen Zahlen ist ebenfalls ein Zahlkörper; denn man kann die reellen Zahlen addieren, subtrahieren, multiplizieren und dividieren.

2.14 Die Gesamtheit der Zahlen

Selbst die Menge aller reellen Zahlen erweist sich nicht als ausreichend, um jede beliebige algebraische Gleichung mit ganzen Koeffizienten zu lösen. Die Gleichung

$$x^2 + 1 = 0 \tag{4}$$

besitzt im Körper der reellen Zahlen keine Wurzeln; man kann diese Gleichung ja auch in der Form $x^2 = -1$ schreiben.
Das Quadrat jeder reellen (positiven oder negativen) Zahl ist aber eine nichtnegative Zahl. Um die Gleichung (4) lösen zu können, muss man zum Körper der reellen Zahlen die neue ("imaginäre") Zahl $\sqrt{-1}$ hinzunehmen.

Zusammen mit dieser neuen Zahl $\sqrt{-1}$ muss man natürlich auch alle Zahlen der Gestalt $a + b\sqrt{-1}$ hinzunehmen, wobei a und b beliebige reelle Zahlen sind. Ist $b = 0$, so sieht man die beiden Zahlen $a + 0 \cdot \sqrt{-1}$ und a als gleich an.
Zahlen der Gestalt $a+b\sqrt{-1}$ heißen komplexe Zahlen. Sie werden nach den bekannten Regeln (wie Polynome) addiert, subtrahiert und multipliziert. Die Division führt man aus, indem man Zähler und Nenner mit der konjugiert-komplexen Zahl multipliziert:

$$\frac{2+\sqrt{-1}}{3-\sqrt{-1}} = \frac{(2+\sqrt{-1})(3+\sqrt{-1})}{(3-\sqrt{-1})(3+\sqrt{-1})} = \frac{5+5\sqrt{-1}}{3^2-(\sqrt{-1})^2)} = \frac{1}{2} + \frac{1}{2}\sqrt{-1}$$

Mit den komplexen Zahlen lassen sich somit die Grundrechenarten ausführen, d.h., die komplexen Zahlen bilden einen Körper.
Der Körper der komplexen Zahlen ist eine Erweiterung des Körpers der reellen Zahlen, und in diesem Erweiterungskörper besitzt die Gleichung (4) bereits Wurzeln, nämlich

$$x_1 = \sqrt{-1} \qquad \text{und} \qquad x_2 = -\sqrt{-1}$$

Eine bemerkenswerte Eigenschaft des Körpers der komplexen Zahlen besteht darin, dass er algebraisch abgeschlossen ist. Das bedeutet, im Körper der komplexen Zahlen hat jede algebraische Gleichung (beliebigen Grades) mit reellen oder komplexen Koeffizienten Wurzeln. Das ist der berühmte "Fundamentalsatz der Algebra", für den als erster der französische Mathematiker d'Alembert Mitte des 18. Jahrhunderts einen wenn auch noch unvollständigen Beweis angab.
Er wurde erstmalig vollständig und in aller Strenge im Jahre 1799 von C.F. Gauß bewiesen.

Aus diesem Satz folgt, dass jedes Polynom in x mit komplexen (speziell mit reellen) Koeffizienten in Faktoren ersten Grades (mit komplexen Koeffizienten) zerlegt werden kann. Beispielsweise ist

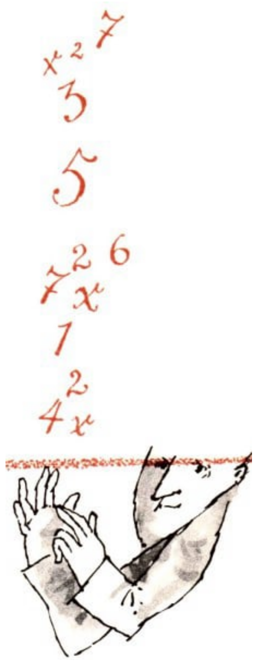
$$x^2 + 1 = (x + \sqrt{-1})(x - \sqrt{-1})$$

Der Körper der reellen Zahlen ist nicht algebraisch abgeschlossen; er wird es erst durch Hinzunahme der komplexen Zahlen.
Dieser Umstand zeigt deutlich, dass die Einführung der komplexen Zahlen in die Mathematik kein Zufall war, sondern ein gesetzmäßiges Ergebnis der Entwicklung der Theorie

der algebraischen Gleichungen.

Gegenwärtig werden die komplexen Zahlen nicht nur zur Lösung von Gleichungen benutzt, sondern auch in verschiedenen anderen Gebieten der Mathematik. Auch in der Elektrotechnik und in der Aerodynamik werden sie verwendet.
Für $\sqrt{-1}$ schreibt man meist den Buchstaben i; diese Zahl wird imaginäre Einheit genannt.

2.15 Drei statt vier Rechenoperationen



Die arithmetischen Operationen werden in der Algebra nicht nur auf Zahlen, sondern auch auf Buchstabenausdrücke, z.B. auf Polynome, angewandt.
Wir werden nur Polynome in der einen Variablen x mit reellen Koeffizienten untersuchen, etwa $x^2 - 5$, $x^3 + 2x^2 - x$, $\frac{1}{4}x^4 + \sqrt{2}$ usw.
Man kann auch die Zahlen als Polynome ansehen (als Polynome nullten Grades, d.h als Polynome, die den Buchstaben x überhaupt nicht enthalten).

Derartige Polynome kann man addieren, subtrahieren und multiplizieren, das Ergebnis ist stets wieder ein Polynom. Die Division von Polynomen durch Polynome ist jedoch im allgemeinen nicht ohne Rest möglich.

So geht z.B. das Polynom $x + 1$ nicht ohne Rest in $x^2 + 3$ auf (d.h., es gibt kein Polynom, das, mit $x + 1$ multipliziert, das Polynom $x^2 + 3$ ergäbe).
In der Menge aller Polynome sind somit stets drei Operationen ausführbar: Addition, Subtraktion und Multiplikation.

Dagegen ist die vierte Rechenoperation (die Division) nicht immer ausführbar. Die Rechengesetze für Polynome sind übrigens dieselben wie beim Zahlenrechnen.

Beispielsweise gilt das Distributivgesetz

$$(a + b)c = ac + bc$$

sowohl für Zahlen als auch für Polynome. Desgleichen gelten die beiden Kommutativgesetze

$$ab = ba \qquad , \qquad a + b = b + a$$

die Regeln für das Auflösen von Klammern usw.
Eine Menge von mathematischen Objekten (etwa von Zahlen oder Polynomen) heißt ein Ring, wenn sich in ihr die drei Rechenoperationen Addition, Subtraktion und Multiplikation ausführen lassen und dabei den üblichen Rechengesetzen (dem Distributivgesetz, dem Kommutativgesetz u.a.) genügen.

Die Menge der Polynome in x ist somit ein Ring ("Polynomring" oder "Ring der Polynome" genannt).

Gelegentlich verzichtet man auf die Forderung, dass die Multiplikation kommutativ ist. Dann spricht man von einem Schiefring.
Um Missverständnisse zu vermeiden, sagt man vorsichtshalber statt "Ring" oft "kommutativer Ring".

Selbstverständlich ist jeder Zahlkörper auch ein Ring; in ihm sind doch Addition, Subtraktion und Multiplikation stets ausführbar (außerdem noch die Division).
Nicht jeder Ring ist jedoch ein Körper, weil in einem Ring die Division im allgemeinen nicht ausführbar zu sein braucht. Ein Beispiel für einen Ring, der kein Körper ist (d.h. für einen Ring, in welchem die Division nicht ausführbar ist), ist gerade der Ring der Polynome.

Als weiteres Beispiel für einen Ring, der ebenfalls kein Körper ist, führen wir die Menge aller ganzen Zahlen an: Addition, Subtraktion und Multiplikation sind immer, die Division dagegen ist nicht immer ausführbar.
Die geraden Zahlen bilden ebenfalls einen Ring, denn Summe, Differenz und Produkt von geraden Zahlen sind wieder gerade Zahlen.
Dagegen ist die Menge aller ungeraden Zahlen kein Ring, weil schon die Addition und Subtraktion aus dieser Menge herausführen: Summe und Differenz ungerader Zahlen sind gerade.

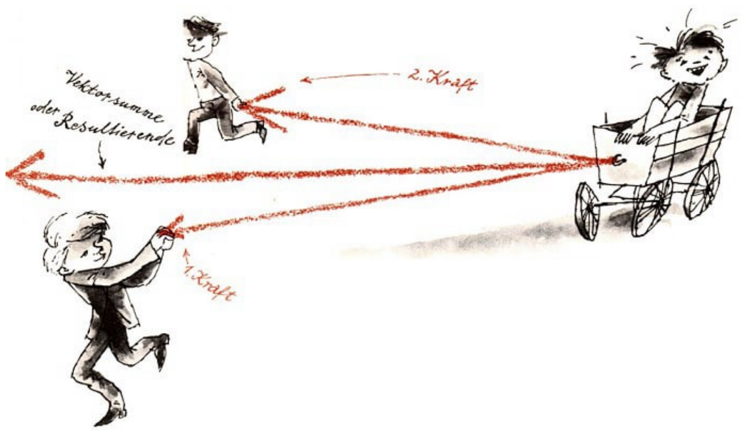
2.16 Mit welchen Objekten kann man rechnen?

An der Addition, Subtraktion und Multiplikation von Polynomen haben wir zum ersten Male gesehen, dass man die algebraischen Operationen nicht nur an Zahlen ausführen kann, sondern auch an anderen mathematischen Objekten, z.B. Polynomen.
Selbst Schüler finden heute nichts Ungewöhnliches mehr daran, dass man die algebraischen Rechenoperationen außer an Zahlen auch an anderen Objekten durchführen kann.

Einiges über Addition und Subtraktion von Vektoren findet man z.B. in dem Abschnitt "Geometrische Algebra". In dem Abschnitt "Ungewöhnliche Arithmetik" werden Addition, Subtraktion und Multiplikation von Restklassen beschrieben und in dem Abschnitt "Der Begriff der geometrischen Abbildung" die Addition von Transformationen.

Was haben nun diese auf den ersten Blick völlig verschiedenen Operationen, die wir alle mit demselben Ausdruck "Addition" (Addition von Zahlen, von Vektoren, von Transformationen) bezeichnen, gemein?
Selbstverständlich hat es keine Bedeutung, dass wir für sie den gemeinsamen Namen Addition benutzen. Man könnte dafür auch irgendwelche anderen Bezeichnungen einführen.

So könnte man statt von der Summe zweier Vektoren von ihrer Resultierenden (oder Resultante) sprechen. Diese Bezeichnung verwendet man in der Mechanik für die "Summe" von Kräften. Das Gemeinsame aller dieser Operationen besteht darin, dass ihre Eigenschaften an die der gewöhnlichen Addition von Zahlen erinnern.



Als Beispiel wollen wir einmal die Vektoraddition untersuchen, die schon eingehender betrachtet wurde. Wie wir wissen, ist sie assoziativ:
Für je drei beliebige Vektoren $\mathfrak{a}$, $\mathfrak{b}$ und $\mathfrak{c}$ gilt die Beziehung

$$\mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{c}$$

Diese Eigenschaft der Vektoraddition weist auf eine Verwandtschaft mit der Addition von Zahlen hin.
Ferner kommt unter den Vektoren, genau wie unter den Zahlen, eine Null (der Nullvektor $\mathfrak{o}$) vor, so dass (wie bei Zahlen) die Beziehung

$$\mathfrak{a} + \mathfrak{o} = \mathfrak{a}$$

für jeden Vektor $\mathfrak{a}$ erfüllt ist.
Wie bei den Zahlen gibt es zu jedem Vektor $\mathfrak{a}$ einen entgegengesetzt gleichen Vektor, der mit $-\mathfrak{a}$ bezeichnet wird und für den die Beziehung

$$\mathfrak{a} + (-\mathfrak{a}) = \mathfrak{o}$$

gilt. Hierdurch ist es möglich, die Subtraktion von Vektoren zu definieren: Die Differenz $\mathfrak{a} - \mathfrak{b}$ ist der Vektor $\mathfrak{a} + (-\mathfrak{b})$.
Alle vom Zahlenrechnen her bekannten Eigenschaften von Gleichungen bleiben auch für Vektoren gültig. So kann man beispielsweise Glieder von einer Seite einer Gleichung auf die andere bringen, wobei sich die Vorzeichen dieser Glieder umkehren.
Hieraus folgt, dass für irgend zwei Vektoren $\mathfrak{a}$ und $\mathfrak{b}$ die Gleichung

$$\mathfrak{x} + \mathfrak{b} = \mathfrak{a}$$

genau eine Lösung besitzt, nämlich

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

Außerdem wollen wir noch darauf hinweisen, dass die Vektoraddition kommutativ ist:

$$\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$$

Daraus geht hervor, dass die Vektoraddition und die Addition von Zahlen tatsächlich vieles gemeinsam haben.
Genauso ließen sich viele Eigenschaften anführen, die der Addition von Zahlen und der Addition von geometrischen Transformationen gemeinsam sind. Ebenso gibt es Analogien zwischen der Addition von Zahlen und der Addition von Restklassen usw. Es gibt noch sehr viele Beispiele für mathematische Objekte, deren "Addition" ähnliche Eigenschaften besitzt.

2.17 Der Begriff der Gruppe

Diese Analogien führen auf einen wichtigen mathematischen Begriff - den Begriff der Gruppe. Mathematische Objekte (etwa Vektoren) bilden eine Gruppe, wenn für sie zwei zueinander inverse Operationen definiert sind (die man gewöhnlich "Addition" und "Subtraktion" nennt), deren Eigenschaften an die der Addition und Subtraktion von Zahlen erinnern. Ehe wir den Gruppenbegriff genauer definieren, gehen wir etwas ausführlicher auf seine Bedeutung ein.

Was hat man damit erreicht, dass man Mengen von mathematischen Objekten, die ähnlichen Rechenregeln unterliegen, unter der einheitlichen Bezeichnung "Gruppe" zusammenfasst?
Das hat folgenden Vorteil: Hat man irgendeinen Satz bewiesen, der für jede Gruppe richtig ist, so gilt er für jedes der oben betrachteten Beispiele von Gruppen (für Zahlen, Restklassen, geometrische Transformationen, Vektoren) sowie für eine ganze Anzahl weiterer Beispiele, auf die wir nicht eingegangen sind.

Auf diese Weise kann man die Regeln für die Addition und die Subtraktion von Polynomen, die Vorzeichenregeln beim Auflösen von Klammern usw. für alle diese Beispiele zugleich beweisen, ohne im einzelnen für Vektoren, für geometrische Transformationen, für Restklassen usw. spezielle Beweise führen zu müssen.
Mit anderen Worten, jeder allgemeine Satz über Gruppen ermöglicht es uns, Eigenschaften festzustellen, die sowohl Zahlen und Vektoren als auch geometrischen Transformationen usw. zukommen.

Das gilt natürlich ebenso für die anderen, schon erwähnten "algebraischen Strukturen" Ringe und Körper. Bis zu einem gewissen Grade lässt sich das am Beispiel der Lösung von quadratischen Gleichungen illustrieren:

Wenn man die allgemeine Lösungsformel kennt, kann man eine ganze Reihe von quadratischen Gleichungen mit Zahlenkoeffizienten lösen. Genauso verhält es sich in unserem Falle: Die Aufstellung einer allgemeinen mathematischen Theorie - der Gruppentheorie - ermöglicht es, bei einer Vielzahl von mathematischen Problemen zu wichtigen Ergebnissen zu gelangen: in der Geometrie (Vektoren, geometrische Transformationen u.a.m.), in der Zahlentheorie, in der Algebra usf., ja sogar in einigen Zweigen der modernen Physik.

Wir wollen jetzt definieren, was eine Gruppe ist. Zunächst ist eine Gruppe eine Menge (vgl. dazu den Abschnitt "Der Mengenbegriff") von irgendwelchen Elementen, z.B. die

Menge aller Zahlen, die Menge aller Vektoren, aller geometrischen Transformationen usw. Ferner ist für die Elemente dieser Menge (sagen wir, für Vektoren) eine Verknüpfung (eine Rechenoperation) erklärt, welche man "Addition" nennen kann und die je zwei Elementen a und b aus der betrachteten Menge ein drittes Element dieser Menge, ihre "Summe" $a+b$, zuordnet.

Wenn diese Verknüpfung bestimmte Eigenschaften hat (die sie der Addition von Zahlen vergleichbar macht), so heißt die betrachtete Menge von Elementen, in der diese Verknüpfung definiert ist, eine Gruppe.

Welche Forderungen stellt man nun an die Verknüpfung (Addition), damit die betrachtete Menge wirklich eine Gruppe ist?
Ehe wir sie aufführen, möchten wir noch bemerken, dass die Addition nicht kommutativ zu sein braucht. Es wird nicht verlangt, dass für je zwei beliebige Elemente a und b die Beziehung

$$a+b=b+a$$

gilt.

Die Summen $a+b$ und $b+a$ brauchen nicht übereinzustimmen. Beispielsweise ist die Addition bestimmter geometrischer Transformationen nicht kommutativ.
Dagegen müssen die folgenden drei Bedingungen erfüllt sein:

1. In der betrachteten Menge gibt es ein Nullelement, das man mit 0 bezeichnet, welches man zu jedem beliebigen Element a addieren kann, ohne dass das Element a in ein anderes übergeht:

$$a+0=a \qquad , \qquad 0+a=a$$

2. Es gibt zu jedem Element a ein inverses Element $-a$ derart, dass die Summe von a und $-a$ gleich dem Nullelement ist:

$$a+(-a)=0 \qquad , \qquad (-a)+a=0$$

Bemerkung zu 1. und 2.: Ist die Addition kommutativ, so ist automatisch $a+0=0+a$, und von den beiden Gleichungen $a+0=a$ und $0+a=a$ braucht nur eine einzige angegeben zu werden. Wenn die Addition jedoch nicht kommutativ ist, müssen beide Gleichungen angegeben werden.

3. Die Addition ist assoziativ, d.h., für je drei Elemente a, b und c gilt die Beziehung

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

Definition:
Eine Gruppe ist eine Menge, in der eine Verknüpfung definiert ist, die den Bedingungen 1 bis 3 (den sogenannten Gruppenaxiomen) genügt. Ist diese Verknüpfung außerdem noch kommutativ, d.h., gilt für beliebige Elemente a und b die Beziehung $a + b = b + a$, so nennt man die betrachtete Gruppe kommutativ (oder abelsch, nach dem norwegischen Mathematiker Niels Hendrik Abel, 1802-1829).

Eine Gruppe ist also durch ihre Elemente allein noch nicht bestimmt.
Man muss auch noch die Verknüpfung angeben, welche zwischen den Elementen erklärt ist. Man sagt auch, dass der Menge durch die Verknüpfung eine Gruppenstruktur aufgeprägt wird. Im Grunde ist hier die Struktur wichtiger als die Natur der Elemente, aus denen die Gruppe besteht.

Es wäre ja denkbar, dass man in ein und derselben Menge verschiedene Verknüpfungen, die den Gruppenaxiomen genügen, einführen kann, und dann ergeben sich verschiedene Gruppen.

2.18 Beispiele für Gruppen

Einfachste Beispiele für Gruppen haben wir bereits angeführt. Wenn wir die Menge der reellen Zahlen mit der gewöhnlichen Addition als Verknüpfung betrachten, dann erhalten wir eine Gruppe (bekanntlich sind die Bedingungen 1 bis 3 für Zahlen erfüllt). Diese Gruppe - das Zahlenkontinuum - ist offenbar kommutativ.

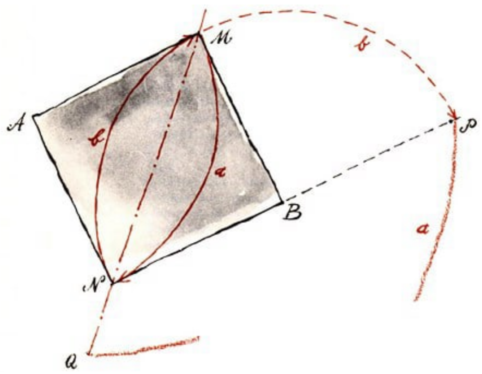
Die Menge aller Vektoren in der Ebene bzw. im Raum mit der üblichen Addition (nach der Parallelogrammregel) ist eine Gruppe, die ebenfalls kommutativ ist. Die erste dieser beiden Gruppen (die Vektoren der Ebene) heißt die zweidimensionale Vektorgruppe, die zweite (die Vektoren im Raum) die dreidimensionale Vektorgruppe.

Es gibt Gruppen, die nur aus endlich vielen Elementen bestehen.
Ein Beispiel bietet die Gruppe der Restklassen modulo m. Unter einer Restklasse modulo m versteht man die Menge aller Zahlen, die bei der Division durch m denselben Rest lassen.

Beispielsweise besteht eine Restklasse modulo 5 aus den Zahlen ..., -6, -1, 4, 9, ..., eine andere aus den Zahlen ..., -7, -2, 3, 8, ...
Die Elemente der Restklassengruppe modulo m sind diese Restklassen, die man sich meist durch den kleinsten nichtnegativen Rest "repräsentiert" denkt; die Verknüpfung ist die gewöhnliche Addition (vgl. dazu den Abschnitt "Eine ungewöhnliche Arithmetik").
Diese (offensichtlich kommutative) Gruppe umfasst m Elemente, die "Repräsentanten". So besteht die Gruppe der Restklassen modulo 3 aus den drei Elementen 0, 1, 2 (den Repräsentanten der Menge der Zahlen, die bei der Division durch 3 die Reste 0 bzw. 1 bzw. 2 lassen) und die Gruppe der Restklassen modulo 10 aus den zehn Elementen 0, 1, 2, ..., 9 (den Repräsentanten der Menge der Zahlen, die bei der Division durch 10 die Reste 0 bzw. 1 bzw. ... bzw. 9 lassen).
Die Elemente einer Restklassengruppe bilden ein "vollständiges Restsystem".

Die Parallelverschiebungen und Drehungen in der Ebene bilden ebenfalls eine Gruppe - sie heißt die Gruppe der Bewegungen in der Ebene.
Das Nullelement dieser Gruppe ist die identische Transformation der Ebene, d.h. diejenige "Transformation", die alle Punkte der Ebene unverändert lässt. Diese Transformation kann man als "Drehung um den Winkel Null" oder als "Parallelverschiebung um die Strecke Null" bezeichnen.

Das zu einer Drehung inverse Element ist die Drehung um den gleichen Winkel in entgegengesetzter Richtung, das inverse Element zu einer Parallelverschiebung ist die Verschiebung um die gleiche Strecke in entgegengesetzter Richtung.
Interessanterweise ist die Gruppe der Bewegungen in der Ebene nichtkommutativ.



Zum Beweis dieser Tatsache betrachten wir in der Ebene ein Quadrat $AMBN$ und bezeichnen mit a die Drehung der Ebene um den Punkt A in Uhrzeigerrichtung um $90°$, mit b die Drehung der Ebene um $90°$ um den Punkt B in Uhrzeigerrichtung.
Wir werden zeigen, dass die Summe der Bewegungen a und b von der Reihenfolge der Summanden abhängt, d.h., dass $a+b \neq b+a$ ist. Dabei versteht man unter der Summe zweier Bewegungen die Hintereinanderausführung.

Bei der Bewegung a geht der Punkt M in den Punkt N über. Führen wir danach die Bewegung b aus, so bilden wir den Punkt N in den Punkt M ab. Die Bewegung $a + b$ (zuerst a, dann b) führt also den Punkt M wieder in den Punkt M über.

Untersuchen wir jetzt die Bewegung $b + a$ (zuerst b, dann a).
In der Abbildung[2] ist der Punkt, in welcher der Punkt M bei der Bewegung b übergeht, mit dem Buchstaben P bezeichnet; die Bewegung a bildet den Punkt P in den Punkt Q ab.
Während also die Bewegung $a + b$ den Punkt M wieder in den Punkt M überführt, geht der Punkt M bei der Bewegung $b + a$ in einen anderen Punkt Q über. Somit sind die Bewegungen $a + b$ und $b + a$ verschieden.

Ein weiteres einfaches Beispiel für eine Gruppe finden wir in der Menge der ganzen Zahlen (der positiven und der negativen und der Null) mit der üblichen Addition als Verknüpfung.
Die Menge der Zahlen 0, 1, 2, 3, ... (die ganzen nichtnegativen Zahlen), versehen mit der üblichen Addition, ist jedoch keine Gruppe: Sie genügt nicht der Bedingung 2; denn

[2] Die Abbildung ist leider nicht vollständig in diese Anschrift übernehmbar.

sie enthält nicht die zu den positiven Zahlen inversen Elemente.

Ein einfaches, auf den ersten Blick vielleicht etwas sonderbar anmutendes Beispiel für eine Gruppe ist die Menge aller positiven reellen Zahlen, in der als "Addition" die gewöhnliche Multiplikation von Zahlen erklärt ist.
Die Benennung der Verknüpfung, ob sie nun Addition oder Multiplikation heißt, ist nebensächlich; es kommt nur darauf an, dass die Gruppenaxiome 1 bis 3 erfüllt sind. Diese Bedingungen sind hier aber tatsächlich erfüllt.
Das "Nullelement" ist die Zahl 1: $a \cdot 1 = a$ und $1 \cdot a = a$.
(Wir behalten den Malpunkt zur Bezeichnung der Verknüpfung bei, obwohl wir sie "Addition" nennen.) Das zur Zahl a inverse Element ist die Zahl $\frac{1}{2}$:

$$a \cdot \frac{1}{a} = 1 \qquad \text{und} \qquad \frac{1}{a} \cdot a = 1$$

Außerdem ist diese "Addition" (d.h. die gewöhnliche Multiplikation von Zahlen) bekanntlich assoziativ:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$

Wir haben somit tatsächlich eine Gruppe vor uns. Dieses Beispiel zeigt die große Ähnlichkeit der beiden Operationen Addition und Multiplikation. Sie kommt übrigens auch beim Rechnen mit Logarithmen zum Ausdruck.

Lange Zeit hindurch bezeichnete man aus historischen Gründen jede Operation, die den Gruppenaxiomen 1 bis 3 genügt, als "Multiplikation". In den letzten Jahren beginnt sich jedoch auch die Bezeichnung "Addition" für die Gruppenoperation durchzusetzen, insbesondere bei kommutativen Gruppen.
Das unterstreicht noch einmal, dass es nicht auf die Bezeichnung der Verknüpfung, sondern auf ihre Eigenschaften ankommt. Als neutrale Bezeichnung verwendet man das Wort "Verknüpfung" (Gruppenverknüpfung).

Wird die Verknüpfung "Multiplikation" genannt, so ist statt "Nullelement" die Bezeichnung "Einselement" oder "Einheitselement" üblich. Oft spricht man auch einfachheitshalber vom "neutralen" Element.

2.19 Permutationsgruppen

Der Gruppenbegriff entstand im 18. Jahrhundert bei der Behandlung von Gleichungen höheren Grades im Zusammenhang mit der Untersuchung der sogenannten Permutationen, deren große Bedeutung von dem französischen Mathematiker Lagrange (1736 bis 1813) erkannt wurde.

Es seien n beliebige Dinge $A_1, A_2, ..., A_n$ gegeben. (Diese "Dinge" waren bei den ursprünglichen Untersuchungen die Wurzeln von Gleichungen n-ten Grades.)

Unter einer Permutation versteht man eine Änderung der Reihenfolge der Elemente $A_1, A_2, ..., A_n$, also eine Tätigkeit. Ersetzt man beispielsweise das Element A_1 durch das Element A_3, das Element A_2 durch A_1 und das Element A_3 durch A_2, so erhält man eine Permutation der Elemente A_1, A_2 und A_3. Diese Permutation schreibt man in der Form

$$\begin{pmatrix} A_1 & A_2 & A_3 \\ A_3 & A_1 & A_2 \end{pmatrix}$$

oder, noch einfacher, da es ja nur auf die Reihenfolge ankommt, in der Gestalt

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

In der oberen Zeile stehen die Nummern der Ausgangselemente, in der unteren die Nummern der Elemente, mit welchen sie vertauscht werden. In der gleichen Weise werden auch die Permutationen beliebig vieler Elemente aufgeschrieben. So besagt beispielsweise der Ausdruck

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}$$

dass die Elemente A_1, A_2, A_3 und A_4 durch A_4, A_3, A_1 und A_2 ersetzt wurden.

Die Hintereinanderausführung von Permutationen bezeichnet man als Multiplikation. Bei dem Beispiel der Bewegungsgruppe hatten wir die entsprechende Operation Addition genannt.
Wenn bei der ersten Permutation das Element A_1 in das Element A_3 übergeht und die zweite Permutation das Element A_3 durch A_4 ersetzt, so wird durch das Produkt dieser beiden Permutationen das Element A_1 unmittelbar durch das Element A_4 ersetzt.

Man kann selbstverständlich nur Permutationen multiplizieren, die sich auf die gleichen Dinge beziehen. Als Beispiel für die Multiplikation von Permutationen sei angeführt:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix}$$

Hier steht die zuerst auszuführende Permutation an erster Stelle von links. Der Leser möge sich an diesem Beispiel davon überzeugen, dass die Multiplikation von Permutationen nicht kommutativ ist.

Die Permutationen einer bestimmten Menge von Dingen, z.B. von A_1, A_2, A_3, A_4 und A_5, bilden bezüglich der oben angegebenen Multiplikation eine nicht-abelsche Gruppe. Die in der Permutationsgruppe untersuchte Verknüpfung könnte auch Addition genannt werden, doch hat sich für diese Operation die früher entstandene Bezeichnung "Multiplikation" erhalten. Das ändert natürlich nichts an der Verknüpfung selbst.

Das Nullelement (hier meist "Einselement" genannt) ist diejenige Permutation, die alle Elemente an ihrem Platz lässt:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

Die inverse Permutation wird folgendermaßen definiert:
Wenn eine Permutation das Element A_1 in das Element A_3 überführt, so ersetzt die inverse Permutation das Element A_3 durch das Element A_1. Zwei zueinander inverse Permutationen sind z.B.

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 1 & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix}$$

Die Permutationen einer bestimmten Anzahl von Dingen (etwa der Dinge $A_1, A_2, ..., A_n$ bilden eine Gruppe.
Vielleicht ist der Leser imstande, diese Behauptung selbständig zu beweisen.

Wie im Deutschen vielfach üblich, bezeichnet man mit dem Wort "Permutation" nicht nur die Tätigkeit des Umordnens von Elementen, sondern auch das Resultat einer solchen Tätigkeit, in diesem Falle also eine bestimmte Anordnung der Elemente. In dieser Bedeutung wird das Wort in der Kombinatorik, der Lehre von den Anordnungen (Kombinationen), verwendet. Erfahrungsgemäß bereitet diese doppelte Bedeutung des Wortes dem Anfänger Schwierigkeiten.
Derselbe Doppelsinn haftet beispielsweise dem Wort "Rechnung"an: Der Oberkellner führt eine Rechnung aus (rechnet) und präsentiert dem Gast die Rechnung.

Ebenso ist es mit dem Wort Lösung, von seiner dritten Bedeutung in der Chemie ganz abgesehen.

Im Zusammenhang mit der Entstehung der Gruppentheorie ist vor allem ein Name zu nennen: Evariste Galois. Dieser berühmte französische Mathematiker wurde 1811 geboren und kam im Alter von 2] Jahren, im Jahre 1832, in einem Duell ums Leben.
Trotz der Kürze seines Lebens übte er auf die Entwicklung der Mathematik einen großen Einfluss aus.

Die fundamentalen Untersuchungen von Galois beziehen sich auf die moderne Algebra, in der ein Teilgebiet nach ihm benannt ist (Galoistheorie). In seinen Schriften hat Galois die Grundlagen der Theorie der algebraischen Gleichungen und der Theorie der Körpererweiterungen geschaffen. Er untersuchte verschiedene Gruppen und Körper, in der Hauptsache die Permutationsgruppen. Mit Hilfe seiner Gruppentheorie erzielte Galois zahlreiche bemerkenswerte Ergebnisse in Bezug auf die Lösung von Gleichungen höheren Grades.

Lange Zeit fanden die Arbeiten Galois' keine Anerkennung, obwohl einige schon zu seinen Lebzeiten gedruckt wurden.

Erst 14 Jahre nach dem Tode von Galois sichtete der französische Mathematiker Liouville (1809-1882) die Arbeiten Galois' und gab sie heraus. In den siebziger Jahren des 19. Jahrhunderts waren die Ideen und Methoden Galois' weit verbreitet.

Etwa zu dieser Zeit setzte auch die rasche Entwicklung der allgemeinen Gruppentheorie ein [Jordan (1838-1922), Burnside (1852-1927) u.a.], die später nicht nur die ganze Mathematik beeinflusste, sondern auch auf einigen Gebieten der Physik zur Anwendung kam (z.B. in der Quantenmechanik und in der Kristallografie) und in viele andere Wissenschaften Eingang fand.
Den Gruppenbegriff benutzt man heute fast überall in der Mathematik.

Zur Entwicklung der Gruppentheorie haben auch zahlreiche russische und sowjetische Mathematiker beigetragen, insbesondere O.J. Schmidt (1891-1956) und A.G.Kurosch, in Deutschland vor allem G.Frobenius (1849-1917) und I.Schur (1875-1941). In Deutschland wurde auf dem Gebiet der Algebra sehr viel gearbeitet.

Die Theorie der algebraischen Zahlkörper wurde vor allem von David Hilbert (1862-1943) weiterentwickelt. In der Körpertheorie muss man insbesondere das klassische Werk von Steinitz erwähnen. Richard Dedekind (1831-1916) und Emmy Noether (1882-1935) leisteten bahnbrechende Arbeit in der Theorie der Ringe.

2.20 Eine ungewöhnliche Arithmetik - Arithmetik der Restklassen

W. G. Boltjanski und B. A. Kordemski

Ein Zählmechanismus, der nur bis 1000 zählen kann, ist wohlbekannt, nämlich der elektrische Zähler, der die verbrauchte Elektroenergie in Kilowattstunden (kWh) registriert. In der Abbildung zeigt der Zähler 905,73 an.
Seit dem Augenblick, als der Zähler auf Null stand, sind also 905,73 kWh elektrischer Energie verbraucht worden. Nach einiger Zeit wird der Zähler den Wert 999,99 kWh anzeigen. Die Zahl 1000 dagegen wird man am Zähler nie ablesen können; für 1000 erscheint wieder die 0, denn so ist der Zähler konstruiert.

Angenommen, im Fensterchen des Zählers stünde die Zahl 016,09, und vor einem Monat sei der Wert 880,12 abgelesen worden. Wir wollen die Zehntel und Hundertstel in beiden Zahlen unberücksichtigt lassen und unter diesen vereinfachenden Annahmen die Frage beantworten: Wieviel Elektroenergie wurde im letzten Monat verbraucht?

Da wir die Konstruktion des Zählers kennen, werden wir selbstverständlich folgendermaßen rechnen: $1016-880 = 136$ (in kWh). Nun kann der Zähler selbst nicht weiter als bis 1000 zählen; wenn auch wir das nicht könnten, würden wir zu einer recht eigenartigen

Arithmetik gelangen:

$$016 - 880 = 136 \qquad \text{oder} \qquad 880 + 136 = 16$$

Man kann diesen Sachverhalt auch so ausdrücken: Der Zähler kann zwar addieren (z.B. zählt er 136 zu 880 hinzu), aber er zeigt nur die letzten drei Ziffern der Summe an (die Zehntel und Hundertstel rechnen wir nicht mit), selbst wenn die Summe größer wird als Tausend.
Hat also der Zähler im Vormonat eine bestimmte Zahl angezeigt und wird im laufenden Monat so viel Energie verbraucht, dass der Gesamtbetrag größer ist als 999, so zeigt der Zähler nicht die Summe selbst an, sondern den Rest, den man bei Division dieser Summe durch 1000 erhält.

Daher können wir nicht ausrechnen, wieviel Elektroenergie in der Zeit von der Aufstellung des Zählers vor mehreren Jahren, als er auf 000 stand, bis heute, wenn beispielsweise der Zählerstand 016 ist, verbraucht wurde. Wir können nur sagen, dass 16 kWh mehr als ein ganzzahliges Vielfaches von tausend Kilowattstunden verbraucht worden sind.
Wieviel Tausende es aber waren, können wir nicht feststellen. Das müssen die Elektrizitätswerke beachten, wenn sie nicht monatlich, sondern, wie es heute vielfach geschieht, nur in größeren Zeitabständen ablesen lassen.

Ein solcher Zähler kann nicht nur addieren, sondern auch multiplizieren.
Stellen wir uns einmal vor, jemand würde an jedem Tag zur gleichen Zeit die elektrische Beleuchtung und andere elektrische Geräte anschalten und später wieder abschalten, so dass in jedem Monat die gleiche Energiemenge verbraucht wird, etwa 136 kWh. Wie groß ist der Energieverbrauch in einem Jahr?

Die Antwort lautet natürlich: 136 kWh $\times$ 12 $=$ 1632 kWh. Der Zähler gibt aber eine andere Antwort. War beispielsweise der Zählerstand Anfang des Jahres 016, so wird am Ende des Jahres, nach einem Energieverbrauch von 1632 kWh, nicht die Zahl 1648, sondern die Zahl 648 angezeigt, so dass der Zählerstand also nur um 632 vergrößert wurde.
Vom "Standpunkt" des Zählers aus ist also

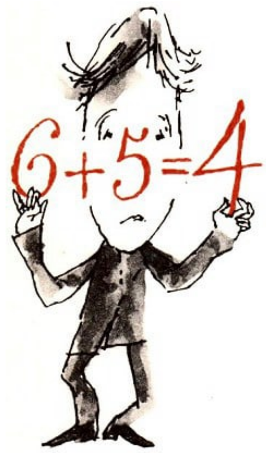
$$136 \cdot 12 = 632$$

Auch hier gibt der Zähler nicht das Produkt selbst an (das wäre 1632), sondern nur den Rest, der sich bei Division dieser Zahl durch 1000 ergibt, d.h. die Zahl 632.

Die Arithmetik des Zählers ist also eine Arithmetik der Reste bei der Division durch 1000. In dieser Arithmetik gibt es nur 999 ganze Zahlen und die Null. Summe und Produkt zweier Zahlen sind niemals größer als 999, und bei der Subtraktion ergeben sich nie negative Zahlen.
Wenn wir eine solche Arithmetik der Reste genauer untersuchen wollen, ist es natürlich bequemer, zunächst mit weniger und kleineren Zahlen zu rechnen. Dazu denken wir uns beispielsweise einen Apparat, der nur bis 6 zählen kann, also folgendermaßen zählt:

$0,1,2,3,4,5,6$ und danach wieder 0 (statt der Zahl 7), 1 (statt der Zahl 8), 2 (statt der Zahl 9) usw., also einen Mechanismus, der anstelle einer Zahl den Rest bei der Division durch 7 anzeigt.

Bei der Addition der Zahlen 6 und 5 ergibt sich dann nicht 11, sondern 4; daher schreiben wir $6+5=4$. Desgleichen ist $4+3=0$.
Ferner ist $5\cdot 6 = 2$, denn 2 ist der Rest bei der Division von 30 durch 7. Da wir addieren und multiplizieren können, wollen wir schließlich auch noch versuchen, die inversen Operationen, Subtraktion und Division, auszuführen.



Beispielsweise ist $1-5=3$ (weil $3+5=1$ ist) und $5:3=4$ (denn es ist $3\cdot 4=5$).

In einer derartigen Arithmetik, die man "Arithmetik der Reste bei der Division durch 7" nennen könnte, ist die Menge der Zahlen endlich - es sind insgesamt 7. Diese sieben Zahlen (d.h. die Reste bei der Division durch 7, mit welchen so gerechnet wird, wie wir es oben erklärt haben) heißen Repräsentanten (d.h. Vertreter) der Restklassen modulo 7 oder kurz die Reste modulo 7. Jede dieser Zahlen vertritt (ist "Repräsentant" für) die Menge der Zahlen, die bei Division durch 7 eben diesen Rest lassen.
So vertritt die Null die durch 7 teilbaren Zahlen ... $-14,-7,0,7,14$, ..., die Zahl 1 die Zahlen ... $-13,-6,1,8,15$, ... usw.

Der Zählerstand (abgesehen von den Zehnteln und Hundertsteln) ist also jeweils ein Rest, und zwar modulo 1000.
Selbstverständlich kann man auch Reste modulo irgendeiner anderen Zahl betrachten; so zeigt z.B. der große Zeiger einer Uhr modulo 60 an (denn alle 60 Minuten beginnt der Zeiger von vorn zu zählen: eine Minute, zwei Minuten ...).

Um Verwechslungen mit der gewöhnlichen Arithmetik zu vermeiden, benutzt man in der Restklassenarithmetik, um die Gleichheit zweier Restklassen auszudrücken, statt des Gleichheitszeichens $=$ das Zeichen $\equiv$ und gibt außerdem den Divisor m an, auf den diese Reste bezogen sind:

$$a \equiv b(\text{mod } m)$$

Dies liest man: a ist kongruent b modulo m. Statt "bei dem Divisor m" sagt man also "modulo m".

Diese Bezeichnungsweise stammt von Gauß. Sie ist so zweckmäßig und hat sich als so sehr fruchtbar für die Weiterentwicklung der Zahlentheorie erwiesen, dass Edmund Landau (1877-1938, Prof. in Göttingen) zu sagen pflegte:

"Wenn Gauß nichts anderes getan hätte als diese Bezeichnung einzuführen, wäre er schon ein großer Mathematiker gewesen."

Hier zeigt also ein mathematisches Beispiel, wie wichtig es sein kann, einen bestimmten Inhalt in der richtigen Form wiederzugeben. Unsere obigen Zahlenbeispiele schreibt man also unmissverständlich in der Form:

$$6 + 5 \equiv 4 (\text{mod } 7) \qquad \text{bzw.} \qquad 5 \cdot 6 \equiv 2 (\text{mod } 7)$$

Man liest das folgendermaßen: $6+5$ ist kongruent 4 modulo 7 bzw. $5 \cdot 6$ ist kongruent 2 modulo 7. Der Kürze halber verzichten wir hier aber auf diese korrekte Schreibweise und bleiben bei dem einfachen Gleichheitszeichen, raten aber dem Leser, alle vorkommenden Gleichungen zwischen Resten zur Übung auch in der Form aufzuschreiben, wie sie in der Mathematik üblich ist.

Kehren wir wieder zu den Resten modulo 7 zurück. Wie wir gesehen haben, lassen sich die Reste ganz einfach addieren und multiplizieren: Man addiert bzw. multipliziert sie wie gewöhnliche Zahlen, nimmt aber nicht das erhaltene Resultat als Ergebnis, sondern dessen Rest bei der Division durch 7.
Berechnet man alle möglichen Summen und Produkte der Reste, dann kann man für sie "Additionstabellen" und "Multiplikationstabellen" aufstellen. Nachstehend finden Sie diese beiden Tabellen für die Reste modulo 7.

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

×	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6
2	0	2	4	6	1	3	5
3	0	3	6	2	5	1	4
4	0	4	1	5	2	6	3
5	0	5	3	1	6	4	2
6	0	6	5	4	3	2	1

In den Tabellen stehen in der obersten Zeile und in der linken Spalte (beide sind durch eine Linie abgeteilt) die Summanden bzw. Faktoren, im übrigen Teil stehen die Summen bzw. die Produkte der entsprechenden Zahlen.

Wollen wir beispielsweise das Produkt $4 \cdot 5$ bestimmen, so suchen wir in der Multiplikationstabelle die Zeile, vor welcher der Faktor 4 steht, und die Spalte, über welcher der Faktor 5 steht. Im Schnittpunkt der entsprechenden Zeile und Spalte finden wir das gesuchte Produkt $4 \cdot 5 = 6$.

Man kann die oben angegebenen Tabellen auch zur Subtraktion und Division benutzen. So wollen wir einmal $5 : 2$ bestimmen. Das bedeutet, wir suchen denjenigen Rest x, für den $2x = 5$ gilt.
Dazu müssen wir offenbar die Zeile in der Multiplikationstabelle aufsuchen, welche dem Multiplikanden 2 entspricht, und dann nachsehen, wo in dieser Zeile das Produkt 5 steht:

×							6
2	0	2	4	6	1	3	5

Nun braucht man nur noch auf die oberste Zeile zu schauen, um den Faktor x zu finden, für den das Produkt $2 \cdot x$ gleich 5 ist.
Das ist die Zahl 6; also ist $2 \cdot 6 = 5$. Daraus folgt: $5 : 2 = 6$.

In der aus der Multiplikationstabelle herausgeschriebenen Zeile, die dem Multiplikanden 2 entspricht, kommen nun alle Reste (von 0 bis 6) einmal vor. Das bedeutet, dass jeder Rest durch 2 teilbar ist. Das gilt auch für alle anderen Zeilen der Multiplikationstabelle, mit Ausnahme der Zeile natürlich, die dem Muliplikanden 0 entspricht.
Daraus folgt: In der Arithmetik der Reste modulo 7 ist die Division immer ausführbar, mit Ausnahme der Division durch Null. Mit anderen Worten, in dieser Arithmetik gibt es keine gebrochenen Zahlen.

Dasselbe gilt auch für die Reste modulo 5 oder für die Reste modulo 13, allgemein für die Reste modulo einer Primzahl. Ist der Restklassenmodul keine Primzahl, so ist die Division nicht immer möglich. Beispielsweise kann man in der Arithmetik der Reste modulo 10 nicht den Rest 1 durch den Rest 5 dividieren.
Die Multiplikation mit 5 ergibt entweder 0 oder 5. Für jeden Rest x modulo 10 gilt somit eine der beiden Gleichungen: $5 \cdot x = 0$ oder $5 \cdot x = 5$; daher ist die Gleichung $5 \cdot x = 1$ unlösbar (d.h., sie besitzt keine Wurzeln).

2.21 Kann man in der Restklassenarithmetik alle quadratischen Gleichungen lösen?

Wir untersuchen hier nur solche quadratischen Gleichungen $ax^2 + bx + c = 0$, deren Koeffizienten Reste modulo 5 sind.
Ferner soll $a \neq 0$ sein, weil sonst die Gleichung nicht mehr quadratisch ist.

Unter diesen Einschränkungen für die Koeffizienten a, b und c ist die Anzahl der möglichen quadratischen Gleichungen nicht mehr allzu groß. Sie fänden alle auf einer Heftseite Platz, und jeder Schüler könnte ruhig erklären, er würde alle solchen quadratischen Gleichungen lösen.
Bei der Lösung einer solchen Gleichung ist zu beachten, dass ihre Wurzeln ebenfalls nur Reste modulo 5 sein können, also die Zahlen 0, 1, 2, 3 und 4. Welche dieser Zahlen einer bestimmten Gleichung genügt, stellt man am einfachsten durch Probieren fest, indem man sie nacheinander in die entsprechende Gleichung einsetzt.

Man kann auch so vorgehen, dass man die linke Seite der Gleichung zu einem vollständigen Quadrat ergänzt. Ist z.B. die Gleichung $x^2 + x + 3 = 0$ gegeben, so erhält man durch Addition von 1 die Gleichung $x^2 + x + 4 = 1$ oder $x^2 + 2 \cdot 3x + 3^2 = 1$, also $(x + 3)^2 = 1$.[3]

Sehen wir uns nun die Quadratzahltabelle modulo 5 an:

$$0^2 = 0, \quad 1^2 = 1, \quad 2^2 = 4, \quad 3^2 = 4, \quad 4^2 = 1$$

Hieraus ergibt sich sofort, dass die 1 auf der rechten Seite der Gleichung $(x + 3)^2 = 1$ entweder das Quadrat von 1 oder das Quadrat von 4 ist.
Daraus ergeben sich die Lösungen: $x_1 + 3 = 1$; $x_2 + 3 = 4$, also $x_1 = 3$; $x_2 = 1$. (Wieder ist zu beachten, dass $1 = 6$ und $4 = 3^2$ modulo 5 ist.)

[3] Anmerkung: Diese fehlerhafte Schlussfolgerung ist im Original zu finden.

Aus der Quadratzahltafel ist ersichtlich, dass es keinen Rest modulo 5 gibt, dessen Quadrat gleich 2 oder 3 wäre. Eine gewisse Teilmenge der Menge der quadratischen Gleichungen hat also keine Wurzeln. Unlösbar ist z.B. die Gleichung $x^2 = 2$.
Nun liegt der Gedanke nahe, zur Menge der Reste modulo 5 so viele neue "Reste" hinzuzunehmen, wie notwendig sind, um alle zu untersuchenden quadratischen Gleichungen lösen zu können.

Tatsächlich brauchen wir nur den irrationalen Rest $\sqrt{2}$ hinzuzunehmen, dessen Quadrat nach Definition gleich 2 ist. Dann ist die Gleichung $x^2 = 2$ lösbar, und ihre Lösung lautet $x = \sqrt{2}$.
Bei einer solchen künstlichen Erweiterung darf allerdings nicht vergessen werden, dass auch die Bedingungen unserer Arithmetik erfüllt sein müssen: Wenn wir die vier Grundrechenarten in der erweiterten Menge der "Reste" ausführen, dürfen sich wieder nur Zahlen aus dieser Menge ergeben.

Daher ist $\sqrt{2}$ nicht der einzige "irrationale Rest", den wir zu den fünf Resten 0, 1, 2, 3 und 4 hinzunehmen müssen. Es müssen insgesamt die folgenden "irrationalen Reste" hinzugenommen werden:

$$\begin{array}{lllll}
\sqrt{2} & 1+\sqrt{2} & 2+\sqrt{2} & 3+\sqrt{2} & 4+\sqrt{2} \\
2\sqrt{2} & 1+2\sqrt{2} & 2+2\sqrt{2} & 3+2\sqrt{2} & 4+2\sqrt{2} \\
3\sqrt{2} & 1+3\sqrt{2} & 2+3\sqrt{2} & 3+3\sqrt{2} & 4+3\sqrt{2} \\
4\sqrt{2} & 1+4\sqrt{2} & 2+4\sqrt{2} & 3+4\sqrt{2} & 4+4\sqrt{2}
\end{array}$$

Zusammen mit den Resten 0, 1, 2, 3 und 4 gelangt man zu einer aus 25 Elementen bestehenden Menge.
In dieser aus 25 Elementen bestehenden Menge ist außer der Null jedes Element eine ganzzahlige Potenz des Elementes $2 + \sqrt{2}$ Beispielsweise ist

$$\begin{aligned}
(2+\sqrt{2})^2 &= 1+4\sqrt{2} \\
(2+\sqrt{2})^3 &= (1+4\sqrt{2})(2+\sqrt{2}) = 4\sqrt{2} \\
(2+\sqrt{2})^4 &= 4\sqrt{2}(2+\sqrt{2}) = 3+3\sqrt{2} \\
(2+\sqrt{2})^5 &= (3+3\sqrt{2})(2+\sqrt{2}) = 2+4\sqrt{2} \\
&\dots \\
(2+\sqrt{2})^{24} &= 1
\end{aligned}$$

2.22 Die Restklassenarithmetik hilft der gewöhnlichen Arithmetik

Wir wollen einmal annehmen, es sei der Rest zu bestimmen, den man bei der Division der Zahl $13 \cdot 10^4 + 702^3 \cdot 61$ durch 7 erhält.
Es hat zunächst den Anschein, als müsse man dazu erst einmal die Potenzen ausrechnen, die Multiplikationen und die Addition ausführen und dann das Resultat durch 7 dividieren.

Nun erkennt man jedoch leicht, dass man den gesuchten Rest nicht nur erhält, wenn

man diese Rechenoperationen mit den Zahlen 13, 10, 702 und 61 ausführt und dann dividiert, sondern auch schon dann, wenn man sie mit den Resten dieser Zahlen modulo 7 ausführt.
Dabei muss man selbstverständlich nach den Regeln der Restklassenarithmetik verfahren, d.h. die obigen Additions- und Multiplikationstabellen benutzen. Nun gilt die Tabelle

Zahl	13	10	702	61
Rest modulo 7	6	3	2	5

Ferner ist $10^4 = 3^4 = 3 \cdot 3 \cdot 3 \cdot 3 = 4$, $702^3 = 2 \cdot 2 \cdot 2 = 1$ und $6 \cdot 4 + 1 \cdot 5 = 1$. Damit ist der gesuchte Rest gefunden. Allgemein gilt der folgende Satz:

Ergibt sich eine Zahl durch Kombinationen von Addition, Subtraktion und Multiplikation mehrerer Zahlen, so ist der Rest, der sich bei Division dieser Zahl durch m ergibt, gleich derselben arithmetischen Kombination der Reste der gegebenen Zahlen bei Division durch m.

Im Lichte dieser Erkenntnis wollen wir jetzt einige der bekannten Teilbarkeitskriterien untersuchen und neue Kriterien herleiten.

2.23 Teilbarkeitskriterien

Teilbarkeit durch 9.
Wenn man wissen will, ob eine Zahl durch 9 teilbar ist, untersucht man die Summe der Ziffern (die Quersumme) dieser Zahl: Ist die Quersumme durch 9 teilbar, so ist es auch die Zahl selbst, ist die Quersumme dagegen nicht durch 9 teilbar, so ist auch die Zahl nicht durch 9 teilbar.

Beispielsweise ist die Zahl 157608 durch 9 teilbar (ihre Quersumme ist gleich 27), die Zahl 123576 dagegen nicht. Worauf beruht dieses wohlbekannte Teilbarkeitskriterium?

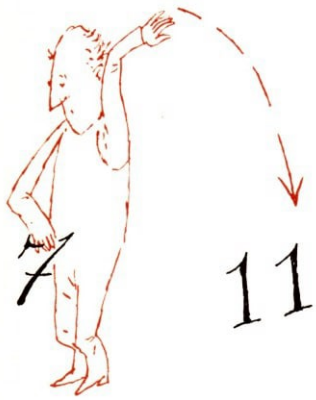
Wie wir wissen, ist die Ziffernreihe 157608 im Dezimalsystem eine Abkürzung für

$$1 \cdot 10^5 + 5 \cdot 10^4 + 7 \cdot 10^3 + 6 \cdot 10^2 + 0 \cdot 10 + 8 \cdot 1 \tag{1}$$

Die Zahl 10 ergibt bei Division durch 9 den Rest 1. Die Potenzen der Zahl 10 $(10^2, 10^3, ...)$ ergeben bei der Division durch 9 ebenfalls den Rest 1. Wir wollen nun in der Summe (1) jede der Zahlen $10^5, 10^4, 10^3, 10^2$ und 10 durch ihren Rest bei der Division durch 9 ersetzen, d.h. durch 1. Dann erhalten wir die Summe

$$1 + 5 + 7 + 6 + 0 + 8$$

d.h. die Quersumme der betrachteten Zahl. Bei Division dieser Summe durch 9 muss sich der gleiche Rest ergeben, wie bei der Division der Zahl 157608. Daraus folgt das oben formulierte Teilbarkeitskriterium für 9.



Teilbarkeit durch 11.
Um ein Kriterium für die Teilbarkeit durch 11 zu finden, müssen wir die Reste der Zahlen $10, 10^2, 10^3, 10^4$ usw. bei der Division durch 11 bestimmen. Das macht man am einfachsten so, dass man 10 nicht als Zahl, sondern als Rest modulo 11 ansieht. Da die Division von 10 durch 11 den Rest 10 ergibt, erhalten wir modulo 11:

$$10^1 = 10$$
$$10^2 = 10 \cdot 10 = 1$$
$$10^3 = 10 \cdot 10^2 = 10$$
$$10^4 = 10 \cdot 10^3 = 1$$
$$10^5 = 10 \cdot 10^4 = 10 \qquad \text{usw.}$$

Wie man sieht, sind die Reste gerader Potenzen gleich Eins, die Reste ungerader dagegen gleich Zehn. Es ergibt sich also, dass die Ziffern an den geradzahligen Stellen (von rechts an gerechnet) mit Eins multipliziert werden müssen, die Ziffern an den ungeradzahligen Stellen dagegen mit Zehn.

Anstelle des Faktors 10 verwendet man besser -1 (weil $10 = 11 - 1$ ist). Damit erhalten wir folgendes Teilbarkeitskriterium:

Wenn die Differenz zwischen der Summe der Ziffern an den geradzahligen Stellen und der Summe der Ziffern an den ungeradzahligen Stellen durch 11 teilbar ist, so ist auch die Zahl selbst durch 11 teilbar; ist dagegen 11 kein Teiler dieser Differenz, so ist auch die Zahl selbst nicht durch 11 teilbar.

Teilbarkeit durch 7.
Wir wollen nun ein Kriterium für die Teilbarkeit durch 7 bestimmen. Wieder betrachten wir die Zahl 157608.
Wie bei der Untersuchung der Teilbarkeit durch 9 können wir auch hier in der Summe (1) jede der Zahlen $10^5, 10^4, 10^3, 10^2$ und 10 durch ihren Rest bei der Division durch 7 ersetzen. Dann erhalten wir anstelle der Zahl 157608 eine beträchtlich kleinere Zahl, die bei der Division durch 7 denselben Rest lässt, und können so leicht feststellen, ob die ursprüngliche Zahl durch 7 teilbar ist.

Wenn wir nämlich mit den Resten modulo 7 rechnen, so gelangen wir stets wieder zu einem Rest modulo 7, d.h., wir erhalten höchstens sechs. Daher ist klar, dass es sich mit den Resten leichter rechnen lässt, um so mehr, als wir ja für die Reste eine Multiplikationstabelle haben!
Die Division von 10 durch 7 ergibt den Rest 3; daraus folgt

der Rest von 10 ist $3^1 = 3$
der Rest von 10^2 ist $3^2 = 3 \cdot 3 = 2$
der Rest von 10^3 ist $3^3 = 3^2 \cdot 3 = 6$
der Rest von 10^4 ist $3^4 = 3^3 \cdot 3 = 4$
der Rest von 10^5 ist $3^5 = 3^4 \cdot 3 = 5$

Ersetzen wir nun die Zahlen 10, 10^2, 10^3, 10^4 und 10^5 durch 3, 2, 6, 4 und 5, so erhalten wir für die Summe (1) die Zahl

$$1 \cdot 5 + 5 \cdot 4 + 7 \cdot 6 + 6 \cdot 2 + 0 \cdot 3 + 8 \cdot 1 = 87$$

Diese Zahl kann man aber leicht "im Kopf" durch 7 dividieren; der Rest ist gleich 3. Derselbe Rest ergibt sich also auch, wenn 157608 durch 7 dividiert wird.

Wir erhalten also folgendes Kriterium für die Teilbarkeit durch 7:
Man multipliziere die erste, zweite, dritte ... (von rechts) Ziffer mit den Zahlen 1, 3, 2, 6, 4, 5, ... und addiere die erhaltenen Produkte. Ist diese Summe durch 7 teilbar, so ist auch die Zahl selbst durch 7 teilbar, ist 7 dagegen kein Teiler der Summe, dann ist auch die Zahl nicht durch 7 teilbar.

Noch einfacher ist es allerdings, die Ziffern (von rechts nach links) nicht mit den Zahlen 1, 3, 2, 6, 4, 5, ... zu multiplizieren, sondern mit den Zahlen 1, 3, 2, -1, -3, -2,....
Lässt man in der Gleichung $6 = 7 - 1$ die Zahl 7 weg, so kann man anstelle von 6 die Zahl -1 einsetzen. Ebenso kann man für 4 die Zahl -3 und für 5 die Zahl -2 schreiben.

Man gelangt dadurch bei der Multiplikation zu kleineren Zahlen. Will man beispielsweise feststellen, ob die Zahl 25518 durch 7 teilbar ist, so bildet man die algebraische Summe:

$$8 \cdot 1 + 1 \cdot 3 + 5 \cdot 2 - 5 \cdot 1 - 2 \cdot 3 = 10$$

die Division von 25518 durch 7 ergibt den Rest 3.

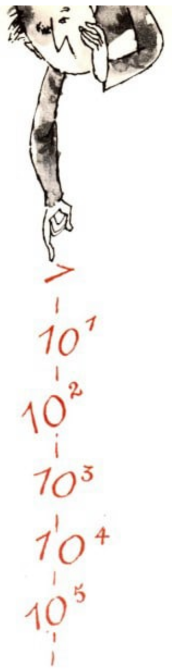
Der Leser möge folgendes Kriterium für die Teilbarkeit einer mehrstelligen Zahl durch 17 nachprüfen (indem er modulo 17 rechnet):
Die Ziffern sind (wieder von rechts nach links gerechnet) mit folgenden Zahlen zu multiplizieren: 1, 10, 15, 14, 4, 6, ...
bzw. (wenn man einige Reste durch negative Reste ersetzt) mit den Zahlen 1, -7, -2, -3, 4, 6, ...

2.24 Der Fermatsche Satz

Auf den ersten Blick haben die angegebenen Teilbarkeitskriterien (beispielsweise für 7 und für 17) einen Mangel:
Falls die gegebene Zahl sehr groß ist, etwa mehrere Dutzend Stellen hat, dann muss man, um diese Teilbarkeitskriterien anwenden zu können, erst zusätzliche Berechnungen ausführen, wenn man feststellen will, mit welchen Faktoren die weiteren Ziffern multipliziert werden müssen.
Man muss also nicht nur die Reste bei der Division der Zahlen 10, 10^2, 10^3, 10^4 und 10^5 (sagen wir, durch 7) berechnen, sondern auch die der weiteren Potenzen 10^6, 10^7, ...

Je größer die Anzahl der Ziffern einer Zahl ist, um so mehr Potenzen von 10 müssen auf ihre Reste bei der Division (durch 7) untersucht werden.
Als Beispiel berechnen wir einmal die Reste, die sich bei der Division der Zahlen 1, 10, $10^2, 10^3$ usw. bis 10^{20} durch 7 ergeben.

Wie wir wissen, muss dazu die Zahl 3 in verschiedene Potenzen erhoben werden; denn 3 ist der Rest von 10 modulo 7. Wenn wir den Rest 3 bereits in die k-te Potenz erhoben haben, so brauchen wir zur Berechnung der $(k+1)$-ten Potenz dieses Ergebnis nur noch einmal mit 3 zu multiplizieren.



Mit anderen Worten, jeder folgende Rest ergibt sich aus dem vorhergehenden durch Multiplikation mit 3. Wenn man das weiß und die Multiplikationstabelle für die Reste modulo 7 zur Verfügung hat, so kann man die Reste leicht bestimmen, welche bei der Division der Zahlen 1, 10, 10^2, 10^3, ..., 10^{20} durch 7 auftreten,

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

Es fällt sofort auf, das sich diese Reste periodisch wiederholen (eine Periode besteht aus sechs Zahlen). Das ist nun eine allgemeine Gesetzmäßigkeit:

Welchen Rest (nach einem bestimmten Modul) wir auch wählen, wenn wir ihn in die erste, zweite, dritte usw. Potenz erheben, so gelangen wir zu Ergebnissen, die sich periodisch wiederholen.
Es ist leicht einzusehen, warum das so ist. In der Folge (2) erhält man beispielsweise jeden Rest durch Multiplikation des vorhergehenden mit 3. Da es aber nur endlich viele Reste nach einem bestimmten Modul gibt, muss nach einer bestimmten Anzahl von Schritten notwendigerweise einmal ein Rest auftreten, der in der Folge bereits vorgekommen ist.

Steht z.B. an der siebenten Stelle der gleiche Rest wie an der ersten, so steht an der achten Stelle der gleiche Rest wie an der zweiten:
Beide ergeben sich durch Multiplikation des an der siebenten bzw. an der ersten Stelle stehenden Restes mit 3. Außerdem steht an der neunten Stelle der gleiche Rest wie an der dritten. Von der siebenten Stelle an werden sich somit alle Reste wiederholen, die von der ersten Stelle bis zur siebenten vorkommen.

Also wiederholen sich die bei der Division der Zahlen 1, 10, 10^2, 10^3, ... durch eine beliebige Zahl erhaltenen Reste periodisch, so dass man nur diese Periode zu kennen braucht, um das Teilbarkeitskriterium auch für beliebig große Zahlen anwenden zu können.

Jetzt entsteht natürlich die Frage: Wie lang ist denn diese Periode? Nach wieviel Zahlen beginnen sich beispielsweise die Reste bei der Division der Zahlen 1, 10^1, 10^2, 10^3, ... durch 17 zu wiederholen?
Die Antwort auf diese Frage gibt der folgende Satz, der von P.Fermat (1601-1665), einem bedeutenden französischen Mathematiker, stammt.

Satz:

Ist p eine Primzahl und a ein Rest modulo p, so wiederholen sich die Reste

$$a, a^2, a^3, a^4, a^5, ... \tag{3}$$

periodisch, und die Länge der Periode (d.h. die Anzahl der Glieder dieser Folge, nach welchen die Wiederholung beginnt) ist ein Teiler der Zahl $p-1$.

Die Zahl $p-1$ ist also entweder selbst die Länge der Periode, oder sie ist gleich der Länge mehrerer Perioden. Folglich muss sich nach jeweils $p-1$ Gliedern alles wiederholen, und es ist

$$a^n = a^{n+(p-1)}$$

Insbesondere erhalten wir für $n=1$:

Ist p eine Primzahl und a ein Rest modulo p, so gilt die Beziehung

$$a = a^p$$

Man kann sie auch anders formulieren:
Ist p eine Primzahl und a eine beliebige Zahl, so ergeben die Zahlen a und a^p bei Division durch p den gleichen Rest, d.h., die Zahl p ist Teiler der Differenz $a^p - a$.

Gerade in dieser Form wird der Fermatsche Satz häufig angegeben.
Wir wollen diesen Satz zur Lösung der folgenden Aufgabe benutzen: Welcher Rest ergibt sich bei Division von $222^{555} + 555^{222}$ durch 7?

Zuerst ersetzen wir die Grundzahlen 222 und 555 durch ihre Reste bei der Division durch 7, d.h., wir ersetzen die gegebene Zahl durch die folgende:

$$5^{555} + 2^{222}$$

Dieser Wert ist selbst Lösung unserer Aufgabe, wenn wir die Grundzahlen 5 und 2 als Reste modulo 7 ansehen. Nun ist 7 eine Primzahl, also wiederholen sich nach dem Fermatschen Satz in den Folgen 2^1, 2^2, 2^3, ... und 5^1, 5^2, 5^3, ... nach jeweils 6 Gliedern die Reste. Die Zahl 555 ergibt bei Division durch 6 den Rest 3, die Zahl 222 ist durch 6 teilbar. Daher ist

$$5^{555} = 5^3 \quad , \quad 2^{222} = 2^0 \quad \text{also}$$

$$5^{555} = 5^3 = 5 \cdot 5 \cdot 5 = 4 \cdot 5 = 6 \quad ; \quad 2^{222} = 2^0 = 1$$

Somit ist $5^{555} + 2^{222} = 6 + 1 = 0$, d.h., die Zahl $222^{555} + 555^{222}$ ist durch 7 teilbar.

2.25 Ein Geheimnis der Schnellrechner

Einer Ihrer Freunde denkt sich eine ganze Zahl, erhebt sie in die dritte Potenz und sagt seinen Freunden das Ergebnis: 19034163.
Sie sollen die gedachte Zahl finden, d.h. aus der Zahl 19034163 die Kubikwurzel ziehen: $\sqrt[3]{19034163}$.
Nach kurzem Überlegen haben Sie die Lösung gefunden: 267.

Ihr Freund staunt. Das Spiel wird mehrmals wiederholt, immer wieder ist aus einer achtstelligen Zahl die Kubikwurzel zu ziehen.
Jedesmal lösen Sie die Aufgabe rasch und richtig. Wie machen Sie das? Es ist gar nicht schwer, hinter den Trick zu kommen.

Dazu braucht man sich nur die letzte Ziffer der dritten Potenz jeder der Zahlen zwischen 1 und 10 (in der gewöhnlichen Arithmetik) zu merken:
$1^3 = 1, 2^3 = 8, 3^3 = 27; 4^3 = 64, 5^3 = 125, 6^3 = 216, 7^3 = 343, 8^3 = 512, 9^3 = 729, 10^3 = 1000$.

Außerdem muss man noch die Tabelle für die dritten Potenzen dieser Zahlen in der Arithmetik der Reste modulo 11 auswendig können (oder zur Hand haben).

Tabelle der Reste modulo 11 und ihrer Kuben

Reste	0	1	2	3	4	5	6	7	8	9	10
Kuben	0	1	8	5	9	4	7	2	6	3	10

Nun überlegt man folgendermaßen: 100^3 ist eine Million, 1000^3 sind tausend Millionen. Daher ist $\sqrt[3]{19034163}$ größer als hundert, aber kleiner als tausend, also eine dreistellige Zahl.
Die letzte Ziffer dieser dreistelligen Zahl ist 7 (das stellen wir an der letzten Ziffer des Radikanden fest), die erste ist 2 (das überlegt man folgendermaßen: $200^3 = 8$ Millionen, $300^3 = 27$ Millionen, und der Radikand liegt zwischen diesen beiden Werten).
Nur die mittlere Ziffer x muss in der Zahl noch bestimmt werden, wir können also $2x7$ schreiben.

Wir ersetzen dazu den Radikanden durch seinen Rest modulo 11.
Den gleichen Rest ergibt bei Division durch 11 die Differenz zwischen der Summe der Ziffern an den ungeradzahligen Stellen (von rechts nach links gezählt) und die Summe der Ziffern an den geradzahligen Stellen: $(3 + 1 + 3 + 9) - (6 + 4 + 1) = 5$.
Bei Division der Zahl 19034163 durch 11 erhält man also den Rest 5.
Wie wir der Tabelle entnehmen, ist der Rest 5 die dritte Potenz des Restes 3. Dividiert man die gesuchte Zahl $2x7$ (mit der unbekannten Zehnerziffer x) durch 11, so erhält man den Rest 3. Dadurch lässt sich x leicht bestimmen:

$$3 = (7 + 2) - x \qquad ; \qquad x = 6$$

Also lautet das Ergebnis: $\sqrt[3]{19034163} = 267$.

Als weiteres Beispiel berechne man $\sqrt[3]{573856191}$.
Die gesuchte Zahl ist dreistellig, die erste Ziffer ist 8, die letzte 1. Der Rest bei der Division des Radikanden durch 11 ist gleich dem der Differenz 15 - 30. Diese Differenz ist aber in der Arithmetik der Reste modulo 11 gleich 7; das heißt, auch der Rest ist gleich 7.
Der Rest 7 ist (laut Tabelle) die dritte Potenz des Restes 6. Damit der Rest bei der Division der gesuchten Zahl durch 11 gleich 6 ist, muss die mittlere Ziffer gleich $(8 + 1) - 6 = 3$ sein. Die gesuchte Zahl ist also 831.

2.26 Der Mengenbegriff - Endliche und unendliche Mengen

P. S. Alexandrow

Man bezeichnet gewöhnlich die Arithmetik als die Wissenschaft von den Zahlen. Die Zahlen in des Wortes einfachster Bedeutung, die sogenannten natürliche Zahlen

$$1, 2, 3, 4, 5, ...$$

sind Antworten auf die Frage "wieviel": Wieviel Schüler sind in der Klasse? Wieviel Bücher liegen auf dem Tisch? Wieviel Gänse schwimmen auf dem Teich?

Jedesmal, wenn wir fragen, "Wieviel Dinge sind das?", müssen wir diese Dinge übersehen können, sie in ihrer Gesamtheit erfassen.
In diesem Sinne sprechen wir von der Gesamtheit aller Schüler, die in eine Klasse gehen, von der Gesamtheit der Bücher, die auf einem Tisch liegen, von der Gesamtheit der Gänse, die auf einem Teich schwimmen. Jede natürliche Zahl gibt eine Anzahl von Dingen (Lebewesen oder Gegenständen) an, die eine Gesamtheit bilden. Manchmal lassen sich diese Dinge leicht abzählen, beispielsweise wenn es sich um die Anzahl der Bücher handelt, die auf einem Tisch liegen, oder um die Anzahl der Schüler, die in einem Klassenzimmer sitzen.

Sehr viel schwerer sind dagegen Fragen der folgenden Art zu beantworten:
"Wieviel Wale schwimmen zu einem bestimmten Zeitpunkt in den Weltmeeren?" oder "Wieviel Igel leben in den Wäldern um Moskau?"

Und noch viel schwerer kann man genau sagen, wieviel Moleküle in einem Glas Wasser enthalten sind oder wieviel Sterne unsere Milchstraße umfasst. Wir sind jedoch in jedem dieser Fälle sicher, dass es sich um endlich viele handelt, selbst wenn die Anzahl vielleicht sehr groß ist oder beim gegenwärtigen Stand unseres Wissens nicht genau angegeben werden kann.

In der Mathematik untersucht man nun nicht nur endliche, sondern auch unendliche Gesamtheiten. Ein ganz einfaches Beispiel einer solchen ist die Gesamtheit, oder wie man gewöhnlich sagt, die Menge aller natürlichen Zahlen:

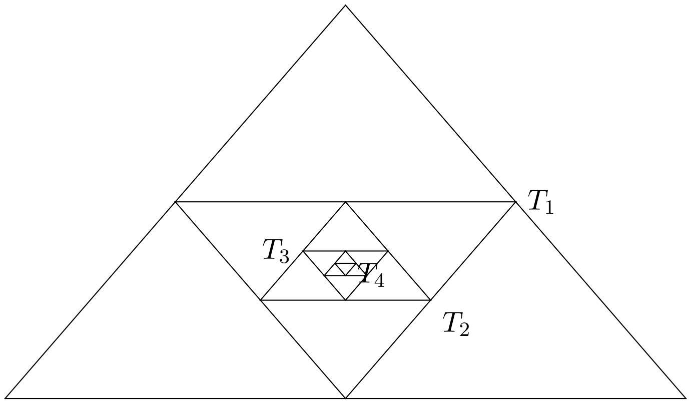
$$1,2,3,4,5,...$$

Wir haben schon gesagt, dass jede natürliche Zahl eine Anzahl von Dingen ist, die eine endliche Gesamtheit, eine endliche Menge bilden. Die Menge aller natürlichen Zahlen ist aber schon keine endliche Menge mehr.
Auf die Frage: "Wieviel natürliche Zahlen gibt es?" muss man antworten, dass es ihrer unendlich viele sind.

Wie groß wir uns eine Gesamtheit natürlicher Zahlen auch denken mögen, stets gibt es natürliche Zahlen, die nicht in der Menge der gedachten Zahlen enthalten sind.

In der Mathematik stoßen wir ständig auf Beispiele für unendliche Mengen. Betrachten wir etwa ein gleichseitiges Dreieck T_1. dem wir ein gleichseitiges Dreieck T_2 einbeschreiben, dessen Ecken die Mittelpunkte der Seiten des Dreiecks T_1 sind. Auf die gleiche

Art beschreiben wir dem Dreieck T_2 ein gleichseitiges Dreieck T_3, ein, T_3 das Dreieck T_4 usw.



Diese Konstruktion führt zu einer unendlichen Menge von gleichseitigen Dreiecken:

$$T_1, T_2, T_3, T_4, T_5, ..., T_n, ... \qquad (1)$$

Um so mehr ist die Menge aller in einer Ebene liegenden gleichseitigen Dreiecke unendlich.
Dieser Satz ist irgendwie doppelsinnig. Das Wort "um so mehr" kann einerseits im Sinne von "erst recht" aufgefasst werden: Da die Menge der ineinandergeschachtelten gleichseitigen Dreiecke unendlich ist, ist auch die Menge aller gleichseitigen Dreiecke in einer Ebene unendlich.

Andererseits kann das Wort "mehr" als Komparativ auffassen, als Ausdruck dafür, dass die Menge aller gleichseitigen Dreiecke in der gegebenen Ebene umfassender, auf irgendeine Art "unendlicher" ist als die unendliche Menge der von uns konstruierten Dreiecke x

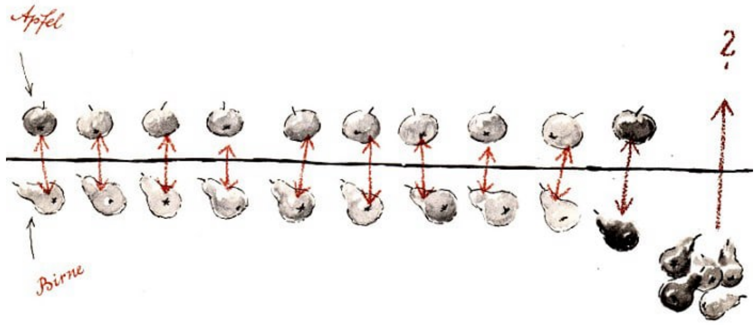
$$T_1, T_2, T_3, T_4, T_5, ..., T_n, ... \qquad (1)$$

Hiermit sind wir auf ein interessantes Problem gestoßen, das die Mathematiker wegen einer (übrigens nur scheinbaren) Paradoxie lange Zeit abgeschreckt hatte:
Gibt es sozusagen verschiedene "Stufen"des Unendlichen? Kann man unendliche Mengen quantitativ charakterisieren, so dass man von einer Menge sagen kann, sie sei "umfassender unendlich" als eine andere? Oder ist die Aussage, eine gegebene Menge sei unendlich, in dem Sinne endgültig, dass eine weitere quantitative Unterteilung oder Abstufung unmöglich ist ?

Als erster hat der bekannte tschechische Mathematiker und Philosoph Bernard Bolzano (1781-1848) diese Frage zu beantworten versucht, aber er konnte noch nicht alle damit zusammenhängenden Schwierigkeiten bewältigen. Wir wollen nun darlegen, worin diese Schwierigkeiten bestehen und wie das Problem gelöst wird.

2.27 Umkehrbar eindeutige Zuordnung zweier Mengen

Wir stellen uns zwei endliche Mengen vor, sagen wir einen Korb Äpfel und einen Korb Birnen. Wenn wir feststellen wollen, ob wir mehr Äpfel oder mehr Birnen haben, können wir die Früchte in jedem Korb zählen (was die einfachste Lösung sein wird). Wir erhalten dann zwei Zahlen, und wenn wir sie vergleichen, erhalten wir die Antwort auf unsere Frage.



Haben wir aber zwei unendliche Mengen vor uns, so können wir dieses Verfahren nicht benutzen, um festzustellen, welche der beiden Mengen "unendlicher" ist. Das geht einfach deshalb nicht, weil eine unendliche Menge nicht "gezählt" werden kann. Jedenfalls wissen wir nicht, wie wir das machen könnten.
Deshalb wollen wir versuchen, die Frage, ob wir mehr Äpfel oder mehr Birnen haben, einmal nicht durch Abzählen zu beantworten, sondern ohne Benutzung des Zahlbegriffes. Dazu bietet sich der folgende Weg an.

Wir legen unsere Äpfel etwa auf einem Tisch aus und versuchen nun, jedem Apfel eine Birne gegenüberzulegen. Dann gibt es drei Möglichkeiten.

Erste Möglichkeit: Es liegt tatsächlich jedem Apfel eine Birne gegenüber, und es wurden nicht nur alle Äpfel, sondern auch alle Birnen ausgelegt. In diesem Falle haben wir offenbar genauso viele Äpfel wie Birnen.

Zweite Möglichkeit: Jedem Apfel liegt eine Birne gegenüber, doch im Korbe liegen noch weitere Birnen, die dabei übriggeblieben sind. In diesem Falle haben wir mehr Birnen als Äpfel.

Schließlich gibt es noch eine dritte Möglichkeit: Es gelingt uns nicht, jedem Apfel eine Birne gegenüberzulegen - die Birnen reichen dazu nicht aus. Dann haben wir offenbar weniger Birnen als Äpfel.

Wie man sieht, konnten wir die beiden Mengen, den Korb Äpfel und den Korb Birnen, miteinander vergleichen, ohne genau zählen zu müssen, wieviel Früchte von jeder Sorte vorhanden sind. Wir haben vielmehr festgestellt, ob es von der einen Sorte mehr gibt bzw. ob von jeder Sorte gleich viel vorhanden sind.

Diesen Vergleich haben wir ausgeführt, indem wir, wie man sagt, eine umkehrbar eindeutige Zuordnung zwischen der einen Menge und der anderen bzw. einem Teil dieser anderen Menge herstellten. Vielleicht reicht dieses eine Beispiel noch nicht aus, um klar zu machen, was eine umkehrbar eindeutige Zuordnung zwischen zwei Mengen ist. Deshalb wollen wir noch einige weitere Beispiele dafür anführen.

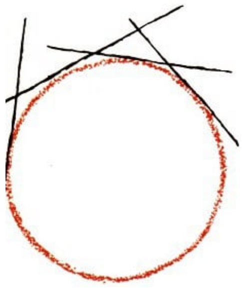
Wenn man ein Konzert besuchen will, braucht man eine Eintrittskarte. Es sind also zwei Mengen vorhanden: die Menge der Leute, die zum Konzert gehen wollen und die wir mit A bezeichnen, und die Menge der Eintrittskarten, die wir mit B bezeichnen. Es sind nun verschiedene Fälle zu unterscheiden.

Erster Fall (zwar nicht sehr wahrscheinlich, aber mathematisch der einfachste): Alle, die ins Konzert gehen möchten, haben Karten bekommen, und alle Eintrittskarten wurden verkauft. Dann entspricht jedem Element der Menge A (d.h. jedem Menschen, der ins Konzert gehen will) ein bestimmtes Element der Menge B (nämlich die von dem Betreffenden gekaufte Eintrittskarte).
Dabei wird jedes Element der Menge B einem einzigen Element der Menge A zugeordnet (d.h. derjenigen Person, welche diese Eintrittskarte gekauft hat). Zwischen der Menge A und der Menge B wurde eine umkehrbar eindeutige (dafür sagt man auch kurz "eineindeutige") Zuordnung hergestellt. Man spricht auch von einer umkehrbar eindeutigen Abbildung der einen Menge auf die andere.

Es kann jedoch vorkommen, dass alle, die ins Konzert gehen möchten, Karten gekauft haben und dass an der Kasse noch Eintrittskarten übriggeblieben sind. Dann haben wir zwar wieder eine umkehrbar eindeutige Abbildung der Menge A vor uns, doch jetzt nicht auf die ganze Menge B, sondern auf einen gewissen Teil von ihr, und zwar auf denjenigen Teil, oder, wie man sagt, auf diejenige Teilmenge von B, die aus den verkauften Eintrittskarten besteht.

Schließlich kann es noch vorkommen, dass nicht alle Leute, die ins Konzert gehen wollten, eine Eintrittskarte bekommen haben. Dann bezeichnen wir die Menge aller derjenigen, die nicht nur ins Konzert gehen wollen, sondern dafür auch eine Karte bekommen haben, mit A'. Die Menge A' ist auf die Menge B umkehrbar eindeutig abgebildet.

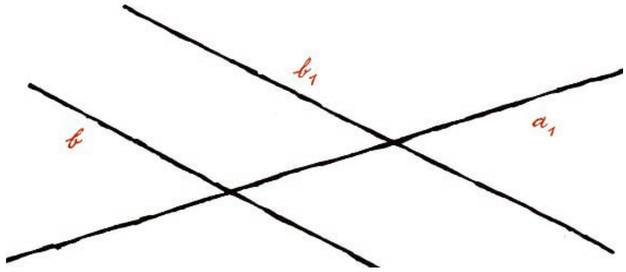
In der Mathematik gibt es zahlreiche Beispiele für umkehrbar eindeutige Zuordnungen. So kann man beispielsweise jedem Eckpunkt eines Dreiecks oder eines Tetraeders die diesem Eckpunkt gegenüberliegende Seite bzw. Seitenfläche zuordnen.
Dann ist also eine umkehrbar eindeutige Zuordnung zwischen der Menge aller Eckpunkte eines Dreiecks (Tetraeders) und der Menge seiner Seiten (Seitenflächen) hergestellt.
Die Menge aller Seiten eines regelmäßigen Polygons lässt sich umkehrbar eindeutig der Menge aller Lote zuordnen, die vom Mittelpunkt des regelmäßigen Polygons auf seine Seiten gefällt werden. Der Menge aller Seitenflächen einer Pyramide ist umkehrbar eindeutig die Menge der auf diese Seitenflächen gefällten Höhen dieser Pyramide zugeordnet usw.



Es ist nun von ganz wesentlicher Bedeutung, dass sich auch unendliche Mengen einander

umkehrbar eindeutig zuordnen lassen.
Dafür einige Beispiele:

1. Es sei A die Menge der Punkte einer Kreislinie und B die Menge aller Geraden, welche Tangenten an diesen Kreis sind. Zwischen den Mengen A und B wird eine umkehrbar eindeutige Zuordnung hergestellt, wenn wir jedem Punkt des Kreises die Tangente in diesem Punkt zuordnen.
Dann ist jedem Element der Menge A ein einziges Element der Menge B zugeordnet (statt "ist zugeordnet" sagt man auch "entspricht"), und jedes Element der Menge B (d.h. jede Tangente) ist dabei genau einem Element der Menge A zugeordnet, nämlich dem Berührungspunkt der gegebenen Tangente.



2. Wir betrachten einmal zwei einander schneidende Geraden a_1 und b_1. Die Menge aller Punkte auf der Geraden a_1 bezeichnen wir mit A und mit B die Menge, welche aus der Geraden b_1 und aus allen Geraden besteht, die zu b_1 parallel sind.
Jedem Element b der Menge B (d.h. jeder Geraden b, die zur Geraden b_1 parallel ist oder mit ihr zusammenfällt) entspricht ein einziges Element der Menge A, nämlich derjenige Punkt auf der Geraden a_1, durch welchen die Gerade b hindurchgeht.

3. Als drittes Beispiel möge die von uns schon untersuchte Menge von gleichseitigen Dreiecken $T_1, T_2, ..., T_n, ...$ dienen.
Jedes dieser Dreiecke, mit Ausnahme des ersten, war dem vorhergehenden einbeschrieben. Die Menge aller dieser Dreiecke sei mit X bezeichnet. Jedes Dreieck hatte eine bestimmte natürliche Zahl n als Nummer erhalten.
Die Nummer des Dreiecks T_n ist die natürliche Zahl n. Dadurch ist offenbar eine umkehrbar eindeutige Zuordnung zwischen der Menge X unserer Dreiecke und der Menge der natürlichen Zahlen hergestellt worden.

2.28 Abzählbare Mengen

Dieses Beispiel lässt sich verallgemeinern. Wenn sich alle Elemente einer Menge X mit Hilfe der natürlichen Zahlen durchnummerieren lassen, so dass jede natürliche Zahl genau einem Element der Menge X zugeordnet ist, dann ist durch diese Nummerierung eine umkehrbar eindeutige Zuordnung zwischen der gegebenen Menge X und der Menge der natürlichen Zahlen hergestellt.

Entsprechend kann man jede umkehrbar eindeutige Zuordnung zwischen einer Menge

X und der Menge der natürlichen Zahlen als Nummerierung (Aufzählung) der Elemente der Menge X mit Hilfe der natürlichen Zahlen ansehen. Wir schreiben einfach die jedem Element der Menge X entsprechende natürliche Zahl als Index an das Element.

Damit sind wir auf einen äußerst wichtigen Begriff gestoßen. Die Herstellung einer umkehrbar eindeutigen Zuordnung zwischen einer Menge X und der Menge der natürlichen Zahlen ist nämlich eine direkte Übertragung des Zählens der Elemente einer endlichen Menge (etwa eines Korbes Äpfel oder einer Schar von Gänsen) auf den Bereich unendlicher Mengen, die mit den natürlichen Zahlen vollzogen wird.
Nur bei endlichen Mengen kommen wir beim Zählen ihrer Elemente mit endlich vielen Zahlen aus (wir zählen: eins, zwei, drei, ... bis zu der Zahl, die uns angibt, wieviel Äpfel im Korb und wieviel Gänse in der Schar sind). In dem Beispiel der Menge der Dreiecke $T_1, T_2, ..., T_n$ ... wie überhaupt bei jeder beliebigen Menge X, die der Menge der natürlichen Zahlen umkehrbar eindeutig zugeordnet werden kann, müssen wir als Nummern sämtliche natürlichen Zahlen benutzen.

Jetzt taucht aber eine entscheidende und für die ganze Mengenlehre grundlegende Frage auf: Ist es immer möglich, die Elemente einer unendlichen Menge stets so mit natürlichen Zahlen zu belegen, sie zu nummerieren, dass jedes Element der Menge eine bestimmte Nummer erhält?

Mit anderen Worten: Kann man zwischen jeder unendlichen Menge und der Menge der natürlichen Zahlen eine umkehrbar eindeutige Zuordnung herstellen?

Diese Frage muss verneint werden, und wir wollen uns davon überzeugen, dass dem so ist. Dazu bedarf es jedoch einiger Vorbereitungen.
Zuerst wollen wir für diejenigen Mengen, die sich der Menge der natürlichen Zahlen umkehrbar eindeutig zuordnen lassen, eine besondere Bezeichnung einführen. Solche Mengen nennt man verständlicherweise abzählbar:

Eine abzählbare Menge ist eine Menge, die mit Hilfe der natürlichen Zahlen abgezählt werden kann.

Unsere Aufgabe besteht nun darin, zu zeigen, dass es nichtabzählbare Mengen gibt, d.h. solche Mengen, die der Menge der natürlichen Zahlen nicht umkehrbar eindeutig zugeordnet werden können.

2.29 Die Menge aller rationalen Zahlen ist abzählbar

Auf der Suche nach einer nichtabzählbaren Menge wenden wir uns zunächst der Menge der rationalen Zahlen zu.
Diese Menge besteht bekanntlich aus allen ganzen und allen gebrochenen Zahlen. Wir wollen sehen, ob sich alle rationalen Zahlen mit Hilfe der natürlichen Zahlen durchnummerieren lassen.

Der Einfachheit halber untersuchen wir erst einmal alle positiven rationalen Zahlen und versuchen, sie auf irgendeine Art zu nummerieren. Dabei stoßen wir sogleich auf eine Schwierigkeit:

Offenbar gibt es unter den positiven rationalen Zahlen keine kleinste Zahl, während bei den natürlichen Zahlen die Eins diese Rolle spielt.

Wie klein wir auch eine positive rationale Zahl r wählen mögen, die Zahl $\frac{1}{2}r$ ist wieder eine positive rationale Zahl, und sie ist kleiner als r.
Selbst wenn wir nun annehmen wollten, wir hätten diese Schwierigkeit überwunden und irgendeine rationale Zahl r_1, gefunden, die wir als erste bezeichnen könnten, so stünden wir vor einer weiteren Schwierigkeit:

Welche rationale Zahl ist die zweite, d.h., welche rationale Zahl folgt beim Zählen unmittelbar auf die Zahl r_1? Wie wir nämlich auch eine rationale Zahl $r_2 > r_1$ wählen, stets gibt es unendlich viele rationale Zahlen, die größer sind als r_1 und kleiner als r_2, beispielsweise die Zahlen

$$r_3 = \frac{1}{2}(r_1 + r_2) \qquad , \qquad r_4 = \frac{1}{2}(r_1 + r_3), ...$$

Es gibt also unter allen rationalen Zahlen, die größer sind als unser r_1, keine kleinste. Welche der auf r_1 folgenden Zahlen soll man nun als die erste bezeichnen?

Diese Schwierigkeit ist jedoch nur scheinbar. Sie zeigt nämlich nur, dass es nicht möglich ist, die rationalen Zahlen mit Hilfe der natürlichen Zahlen so zu nummerieren, dass wachsenden Nummern wachsende Zahlen entsprechen.

Man muss also versuchen, die rationalen Zahlen anders zu nummerieren. Man wird darauf verzichten müssen, der Größe nach zu nummerieren. Dann braucht also nicht die unmittelbar auf r_1 folgende Zahl r_2 die nächstgrößere Zahl, d.h. die kleinste unter allen auf r_1 folgenden Zahlen zu sein. Dann lässt sich für die rationalen Zahlen tatsächlich eine Nummerierung angeben.

Bekanntlich lässt sich jede positive rationale Zahl eindeutig als nichtkürzbarer Bruch $\frac{p}{q}$ schreiben (die ganzen Zahlen n werden dann in der Form $\frac{n}{1}$ geschrieben, und ein solcher "Bruch" wird ebenfalls als nichtkürzbar angesehen).
Unter der Höhe eines Bruches $\frac{p}{q}$ versteht man die natürliche Zahl $p + q$, und unter der Höhe einer rationalen Zahl die Höhe des eindeutig bestimmten nichtkürzbaren Bruches, der die gegebene rationale Zahl darstellt.

Die Höhe der rationalen Zahl $\frac{1}{2}$ ist also 3, während die Höhen der sie darstellenden Brüche $\frac{2}{4}, \frac{3}{6}, \frac{4}{8}$, ... die Zahlen 6, 9, 12, ... sind. Nun wollen wir feststellen, wieviel rationale Zahlen es zu jeder Höhe gibt.
Die Höhe 1 kann keiner positiven rationalen Zahl zukommen; denn die Höhe $p + q$ des entsprechenden unkürzbaren Bruches $\frac{p}{q}$ ist wegen $p = 1$ und $q = 1$ mindestens 2.

Die Höhe 2 kommt offenbar nur der rationalen Zahl $\frac{1}{2} = 1$ zu, da es keine anderen additiven Zerlegungen der Zahl 2 in natürliche Zahlen gibt. Da die Zahl 3 nur in der

Form $3 = 1 + 2$ und $3 = 2 + 1$ in natürliche Zahlen additiv zerlegt werden kann, haben nur die rationalen Zahlen $\frac{1}{2}$ und $\frac{2}{1} = 2$ die Höhe 3.
Die Höhe 4 haben die Brüche $\frac{1}{3}, \frac{2}{2}$ und $\frac{3}{1}$; darunter sind aber nur $\frac{1}{3}$ und $\frac{3}{1}$ unkürzbar. Die Höhe 4 kommt also den rationalen Zahlen $\frac{1}{3}$ und 3 zu.

Für die Höhe 5 kommen die Brüche $\frac{1}{4}$, $\frac{2}{3}$, $\frac{3}{2}$ und $\frac{4}{1}$ in Frage, unter denen sich kein kürzbarer Bruch befindet, so dass es 4 rationale Zahlen mit der Höhe 5 gibt.
Die Höhe 6 haben die Brüche $\frac{1}{5}, \frac{2}{4}, \frac{3}{3}, \frac{4}{2}, \frac{5}{1}$ von denen nur der erste und der letzte unkürzbar sind. Folglich entsprechen der Höhe 6 die rationalen Zahlen $\frac{1}{5}$ und 5.

Durch entsprechende Überlegungen können wir uns also zunächst davon überzeugen, dass es zu jeder natürlichen Zahl $h > 1$ nur endlich viele rationale Zahlen mit der Höhe A gibt. Offenbar kommt die Höhe h den Brüchen

$$\frac{1}{h-1}, \frac{2}{h-2}, ..., \frac{h-1}{1}$$

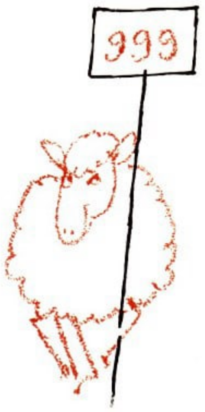
zu, und das sind $(h^{-}1)$ Stück. Unter diesen Brüchen können kürzbare vorkommen; die übrigen sind dann gerade die rationalen Zahlen mit der Höhe h.

Nach diesen Vorbereitungen bietet die Nummerierung der positiven rationalen Zahlen keinerlei Schwierigkeiten mehr.
Wir beginnen mit der kleinsten Höhe 2, vergrößern die Höhe jeweils um Eins und zählen die (stets endlich vielen) rationalen Zahlen ab, die der gegebenen Höhe entsprechen.

Die Zahl $1 = r_1$ erhält somit die Nummer 1. Als nächstes kommen die beiden Zahlen $r_2 = \frac{1}{2}$ und $r_3 = 2$ mit der Höhe 3, danach die beiden Zahlen $r_4 = \frac{1}{3}$ und $r_5 = 3$ mit der Höhe 4; nun folgen die vier Zahlen $r_6 = \frac{1}{4}$, $r_7 = \frac{2}{3}$, $r_8 = \frac{3}{2}$ und $r_9 = 4$ mit der Höhe 5, dann die Zahlen $r_{10} = \frac{1}{5}$, $r_{11} = 5$ mit der Höhe 6 usw.
Wenn wir die Anzahl der rationalen Zahlen der Höhe h mit n_h bezeichnen, können wir unser Ergebnis zusammenfassen:

Höhe	Zahl mit dieser Höhe
2	r_1 $(n_2 = 1)$
3	r_2, r_3 $(n_3 = 2)$
4	r_4, r_5 $(n_4 = 2)$
5	r_6, r_7, r_8, r_9 $(n_5 = 4)$
6	r_{10}, r_{11} $(n_6 = 2)$
7	$r_{12}, r_{13}, r_{14}, r_{15}, r_{16}, r_{17}$ $(n_7 = 6)$
8	$r_{18}, r_{19}, r_{20}, r_{21}$ $(n_8 = 4)$
...	...

Da jede positive rationale Zahl eine natürliche Zahl h als Höhe besitzt, muss sie in der dieser Höhe entsprechenden Zeile vorkommen. Sie erhält also eine bestimmte Nummer, die nicht größer ist als die Zahl $n_2 + n_3 + ... + n_{h-1} + n_h$.
Damit ist gezeigt: Die Menge der positiven rationalen Zahlen ist abzählbar. Daraus folgt sofort, dass die Menge aller rationalen Zahlen ebenfalls abzählbar ist.



Viel schwieriger ist es herauszufinden, welche Nummer eine bestimmte positive rationale Zahl bei unserer Nummerierung erhält.
Übrigens kann man in ähnlicher Weise zeigen, dass auch die Menge aller algebraischen Zahlen abzählbar ist. Wie wir schon ausführten, bezeichnet man eine Zahl als algebraisch, wenn sie Nullstelle eines Polynoms mit ganzen Koeffizienten ist.

Man kann nun auch für diese Polynome eine "Höhe" einführen und zeigen, dass es zu jeder Höhe nur endlich viele Polynome gibt, die jeweils natürlich nur endlich viele Nullstellen haben. Damit ist dann gezeigt, dass die Menge der algebraischen Zahlen - abzählbar ist.

2.30 Die Menge der reellen Zahlen ist nicht abzählbar

Nach diesem überraschenden Ergebnis könnte man die Frage stellen, ob es überhaupt nichtabzählbare (man sagt auch "überabzählbare") Mengen gibt.
Es zeigt sich, dass die Menge aller reellen Zahlen nicht abzählbar ist. Diese bemerkenswerte Tatsache wurde - ebenso wie übrigens die Abzählbarkeit der Menge der rationalen Zahlen - zuerst von dem berühmten deutschen Mathematiker Georg Cantor (1845-1918), dem Begründer der modernen Mengenlehre, bewiesen.

Wir geben hier den Beweis von Cantor wieder und zeigen, dass bereits die Menge der reellen Zahlen im Intervall $(0;1)$, d.h. die Menge der reellen Zahlen zwischen Null und Eins, nicht abzählbar ist.
Jede derartige reelle Zahl kann als unendlicher Dezimalbruch mit einer Null vor dem Komma geschrieben werden. Diese Darstellung ist für alle reellen Zahlen eindeutig, mit Ausnahme derjenigen Zahlen, die durch endliche Dezimalbrüche ausdrückbar sind.

Jede solche Zahl, wie etwa 0,2476622021711, kann auf zwei Arten als unendlicher Dezimalbruch geschrieben werden: in der Form

0,2476622021711000000000... und in der Form
0,2476622021710999999999...

In der einen Darstellung treten von einer gewissen Stelle an nur noch Nullen auf, in der anderen nur noch Neunen. Wenn wir nun vereinbaren, die endlichen Dezimalbrüche nicht als solche mit der Periode 9 zu schreiben, so wird es für jede reelle Zahl nur eine einzige Darstellung als unendlicher Dezimalbruch geben.

Unsere Behauptung, die Menge der reellen Zahlen sei nicht abzählbar, werden wir indirekt beweisen. Wir nehmen also an, die Menge der reellen Zahlen x aus dem Intervall $(0;1)$ sei abzählbar, d.h., man könnte sie mit Hilfe der natürlichen Zahlen nummerieren. Dann könnte man die Gesamtheit der reellen Zahlen x aus dem Intervall $(0,1)$ in Form einer Folge schreiben:

$$x_1, x_2, ...$$

Die Entwicklung der Zahl x_n in einen unendlichen Dezimalbruch schreiben wir in der Gestalt

$$x_n = 0{,}a_1^{(n)} a_2^{(n)} a_3^{(n)} a_4^{(n)} ... a_n^{(n)} ...$$

wobei $a_1^{(n)}, a_2^{(n)}, a_3^{(n)}, ...$ die aufeinanderfolgenden Dezimalstellen von x_n sind. Auf Grund unserer Verabredung können nicht von einer bestimmten Stelle an nur noch Neunen vorkommen.
Alle reellen Zahlen x des Intervalls $(0; 1)$ seien also aufgeschrieben:

$$(I) = \begin{cases} x_1 = 0{,}a_1^{(1)} a_2^{(1)} a_3^{(1)} a_4^{(1)} a_5^{(1)} ... a_n^{(1)} ... \\ x_2 = 0{,}a_1^{(2)} a_2^{(2)} a_3^{(2)} a_4^{(2)} a_5^{(2)} ... a_n^{(2)} ... \\ x_3 = 0{,}a_1^{(3)} a_2^{(3)} a_3^{(3)} a_4^{(3)} a_5^{(3)} ... a_n^{(3)} ... \\ x_4 = 0{,}a_1^{(4)} a_2^{(4)} a_3^{(4)} a_4^{(4)} a_5^{(4)} ... a_n^{(4)} ... \\ ... \\ x_n = 0{,}a_1^{(n)} a_2^{(n)} a_3^{(n)} a_4^{(n)} a_5^{(n)} ... a_n^{(n)} ... \\ ... \end{cases}$$

Wenn also die Menge der reellen Zahlen zwischen Null und Eins abzählbar wäre, müsste jede reelle Zahl aus diesem Intervall unter den x_n vorkommen, also an einer bestimmten Stelle dieser nach unten unbeschränkt fortgesetzten Tabelle stehen.

Wir werden unsere Annahme dadurch zu einem Widerspruch führen, dass wir eine reelle Zahl xi zwischen 0 und 1 angeben, die bestimmt nicht in der Tabelle (I) enthalten ist. Dazu greifen wir die in der Diagonale der Tabelle (I) stehenden Ziffern heraus, also

$$a_1^{(1)}, a_2^{(2)}, a_3^{(3)}, a_4^{(4)}, ..., a_n^{(n)}, ...$$

und wählen zu jedem n eine natürliche Zahl b_n, die nicht größer als 8 und von der Zahl $a_n^{(n)}$ verschieden ist (für $a_n^{(n)} < 8$ setzen wir etwa $b_n = a_n^{(n)} + 1$, für $a_n^{(n)} = 8$ dagegen $b_n = 7$) . Nun betrachten wir den unendlichen Dezimalbruch

$$0, b_1 b_2 b_3 b_4 b_5 ... b_n ...$$

Er enthält keine Neun und stellt eine Zahl ξ dar, die zwischen 0 und 1 liegt und ganz sicher von jeder der Zahlen $x_1, x_2, x_3, ..., x_n, ...$ verschieden ist.
Wäre etwa

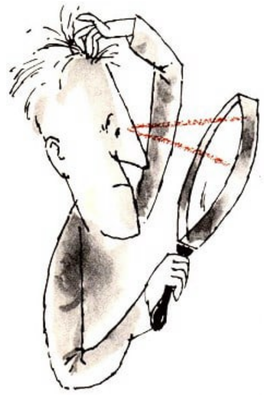
$$\xi = x_m = 0{,}a_1^{(m)} a_2^{(m)} a_3^{(m)} a_4^{(m)} ... a_m^{(m)} ...$$

so müsste an der m-ten Dezimalstelle der Zahl ξ die Ziffer $a_m^{(m)}$ stehen, während in Wirklichkeit aber $b_m \neq a_m^{(m)}$ ist.

Mit diesem Widerspruch ist der Satz bewiesen.
Diese Schlussweise ist unter der Bezeichnung "Cantorsches Diagonalverfahren" bekannt geworden.

2.31 Mächtigkeit einer Menge

Über dieses Ergebnis müssen wir gründlich nachdenken und daraus einige Schlüsse ziehen. Begonnen hatten wir mit dem Begriff der umkehrbar eindeutigen Zuordnung von zwei Mengen. Dass eine solche Zuordnung möglich war, war (für endliche Mengen) gleichbedeutend damit, dass die beiden Mengen gleich viele Elemente enthielten.



Das weist uns auf eine Möglichkeit hin, auch für zwei unendliche Mengen den Begriff der "Gleichzahligkeit" einer, quantitativen Äquivalenz einzuführen. Wir sagen, dass zwei (endliche oder unendliche) Mengen äquivalent sind oder dass sie die gleiche Mächtigkeit haben, wenn zwischen ihnen eine umkehrbar eindeutige Zuordnung hergestellt werden kann.
Der Begriff der "Gleichmächtigkeit" besagt für endliche Mengen, dass sie aus gleich vielen Elementen bestehen.

Ferner wollen wir sagen, dass eine Menge A eine höhere Mächtigkeit besitzt als die Menge B, wenn man die Menge B umkehrbar eindeutig auf eine echte Teilmenge von A abbilden kann, während sich A nicht auf eine Teilmenge von B abbilden lässt.

Jetzt können wir sagen, die abzählbaren Mengen seien diejenigen Mengen, welche der Menge der natürlichen Zahlen äquivalent (gleichmächtig) sind.
Außerdem gibt es noch nichtabzählbare Mengen, beispielsweise die Menge aller reellen Zahlen zwischen 0 und 1. Für diese Menge verwendet man gewöhnlich den Begriff Intervall.
Alle reellen Zahlen x, die der Ungleichung $a < x < b$ genügen, bilden auf der Zahlengeraden ein Intervall $(a; b)$, in unserem Falle also das Intervall $(0; 1)$.

Wie man leicht sieht, kann man als nichtabzählbare Menge schließlich auch ganz allgemein die Menge der reellen Zahlen in jedem Intervall bezeichnen; denn jedes Intervall auf der Zahlengeraden kann durch eine Ähnlichkeitstransformation (Stauchung oder Streckung) auf das Intervall $(0; 1)$ umkehrbar eindeutig abgebildet werden.

Um sich davon zu überzeugen, dass jede nichtabzählbare Menge eine höhere Mächtigkeit hat als jede abzählbare Menge (alle abzählbaren Mengen haben offenbar die gleiche Mächtigkeit), muss man die folgenden beiden Sätze beweisen:

1. Jede Teilmenge einer abzählbaren Menge ist endlich oder abzählbar.
2. Jede unendliche Menge (also insbesondere jede nichtabzählbare) enthält eine abzählbare Menge.

Beweis der ersten Behauptung.
Es sei X eine abzählbare Menge und X_0 eine Teilmenge von X. Die Elemente der Menge X können mit Hilfe der natürlichen Zahlen nummeriert werden, d.h., sie lassen sich als Folge schreiben:

$$x_1, x_2, x_3, ..., x_n, ... \qquad (2)$$

Unter diesen Elementen befinden sich auch die Elemente der Menge X_0. Es seien dies die Elemente (nach wachsenden Indizes aus der Folge (2) herausgegriffen)

$$x_{n_1}, x_{n_2}, x_{n_3}, ..., x_{n_k}, ... \qquad (3)$$

Nun sind zwei Fälle möglich:
a) Die Folge (3) bricht nach dem k-ten Schritt ab; dann besteht die Menge X_0 aus den endlich vielen Elementen $x_{n_1}, x_{n_2}, x_{n_3}, ..., x_{n_k}$

b) Wir gelangen zu einer unendlichen Folge $x_{n_1}, x_{n_2}, x_{n_3}, ..., x_{n_k}, ...$. Setzen wir nun

$$y_1 = x_{n_1}, y_2 = x_{n_2}, ..., y_k = y_1 = x_{n_k}, ...$$

so erhalten wir die Folge der Elemente von X_0 in der Form

$$y_1, y_2, ..., y_k, ...$$

und man erkennt sofort, dass X_0 eine abzählbare Menge ist.

Beweis der zweiten Behauptung:
Es sei X eine unendliche Menge. Dann wählen wir in X irgendein Element x_1 aus. Nun gibt es in X sicher Elemente, die von x_1 verschieden sind (denn sonst bestünde X nur aus diesem einen Element, wäre also endlich).
Wir nehmen eines dieser Elemente und bezeichnen es mit x_2.

Die Elemente x_1 und x_2 schöpfen aber die unendliche Menge X nicht aus. Daher gibt es in X ein Element x_3, das sowohl von x_1 als auch von x_2 verschieden ist. Setzen wir diesen Auswahlprozess fort, so gelangen wir zu einer abzählbaren Menge

$$x_1, x_2, x_3, ..., x_n, ...$$

die in X enthalten ist.

Die zu Anfang unserer Ausführungen gestellte Frage, ob es unendliche Mengen verschiedener "Abstufungen des Unendlichen" (d.h. verschiedener Mächtigkeit) gibt, können wir jetzt bejahen.
Es gibt Mengen reeller Zahlen, die verschiedene Mächtigkeiten haben: die nichtabzählbare Menge der reellen Zahlen eines beliebigen Intervalles einerseits und irgendeine abzählbare Menge reeller Zahlen (z.B. die Menge der positiven rationalen Zahlen) andererseits.

Zu diesem Ergebnis gelangten wir über den Begriff der umkehrbar eindeutigen Zuordnung, auf dem der quantitative Vergleich unendlicher Mengen beruht. Man darf nun aber nicht annehmen, es gebe keine Unterschiede zwischen den umkehrbar eindeutigen Zuordnungen von unendlichen Mengen einerseits und den umkehrbar eindeutigen Zuordnungen von endlichen Mengen andererseits.

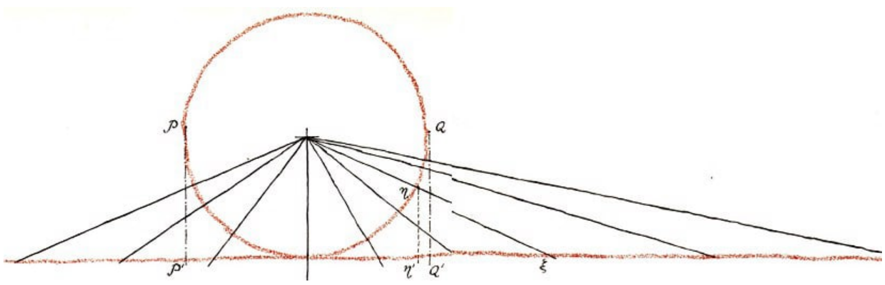
Offenbar kann man keine endliche Menge auf eine ihrer echten Teilmengen umkehrbar eindeutig abbilden (ein Teil ist niemals das im Bereich der unendlichen-Mengen durchaus möglich ist.

Wir hatten gesehen, dass jede unendliche Teilmenge einer abzählbaren Menge abzählbar ist, d.h., jede abzählbare Menge kann auf jede ihrer echten unendlichen Teilmenge eineindeutig abgebildet werden. Schreiben wir beispielsweise unter die natürlichen Zahlen der Reihe nach alle geraden Zahlen

1	2	3	4	5	6	7	8	9	10	...
2	4	6	8	10	12	14	16	18	20	...

so erhalten wir eine umkehrbar eindeutige Zuordnung zwischen der Menge der natürlichen Zahlen und einer ihrer Teilmengen, nämlich der Menge der geraden Zahlen.

Ein anderes Beispiel: Zwischen der Menge aller reellen Zahlen (der ganzen Zahlengeraden) und jedem beliebigen Intervall derselben lässt sich eine umkehrbar eindeutige Zuordnung herstellen.
Um zu einer solchen Zuordnung zu gelangen, kann man folgendermaßen vorgehen.



In der Ebene zeichnet man einen Kreis, der oberhalb der Abszissenachse liegt und diese berührt. Wir betrachten den unteren Halbkreisbogen PQ, ohne diese Endpunkte P und Q. Jetzt stellen wir eine umkehrbar eindeutige Zuordnung zwischen den Punkten auf dem Halbkreis PQ und den Punkten auf der Zahlengeraden her, indem wir jedem Punkt ξ auf der Zahlengeraden denjenigen Punkt η auf dem Halbkreis zuordnen, in dem der Halbkreis von dem Strahl geschnitten wird, der den Mittelpunkt des Kreises mit ξ verbindet.

Jetzt projizieren wir den Halbkreis PQ auf das Intervall $P'Q'$ der Abszissenachse und ordnen dem Punkte η auf dem Halbkreis seine Projektion η' zu.

Jetzt entspricht jedem Punkt ξ der Geraden ein Punkt η' im Intervall $P'Q'$. Diese Zuordnung ist eine umkehrbar eindeutige Abbildung der ganzen Zahlengeraden auf das Intervall $P'Q'$.

Man kann noch viele andere, auf den ersten Blick paradox erscheinende Sätze über die Mächtigkeiten verschiedener Mengen beweisen. Nur einen einzigen wollen wir erwähnen: Die Punkte einer Geraden lassen sich umkehrbar eindeutig auf die Punkte der Ebene abbilden.

Abschließend sei noch folgendes bemerkt. In der Mathematik sind die sogenannten Zahlenmengen von außerordentlicher Bedeutung.
Das sind Mengen, deren Elemente reelle Zahlen sind. Alle bisher bekannten Zahlenmengen sind entweder abzählbar oder haben die gleiche Mächtigkeit wie die Zahlengerade

(das sogenannte Kontinuum). In diesem Zusammenhang entstand die Hypothese, jede nichtabzählbare Zahlenmenge habe die gleiche Mächtigkeit wie die Zahlengerade.
Diese Vermutung stammt bereits von Cantor und ist als "Kontinuumshypothese" bekannt.

Sie ist bisher weder bewiesen noch widerlegt worden, was offenbar auf die großen Schwierigkeiten zurückzuführen ist, die bei der Untersuchung beliebiger Zahlenmengen auftreten. Diese und andere Schwierigkeiten, die im Zusammenhang mit der Entwicklung der Mengenlehre auftraten, führten zur Ausarbeitung einer besonderen Disziplin der Mathematik, die unter der Bezeichnung "Grundlagen der Mathematik" bekannt ist und in enger Beziehung zur mathematischen Logik steht.
Leider können wir an dieser Stelle darauf nicht näher eingehen.

Es war unser Anliegen, von einigen einfachsten Begriffen aus einem umfangreichen und erst etwa hundert Jahre alten Gebiet der Mathematik, der Mengenlehre, einige Vorstellungen zu vermitteln.
Die Mengenlehre ist zu einem der Fundamente der modernen Mathematik geworden.
Wer sich gründlicher damit befassen will, sei auf die Spezialliteratur verwiesen.

3 Einführung in die Wahrscheinlichkeitsrechnung

3.1 Die Wissenschaft vom Zufall

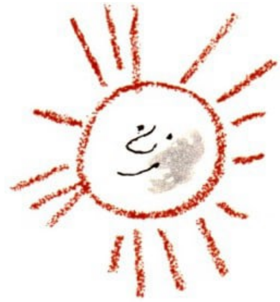
A. J. Chintschin, A. M. Jaglom

Als zufällig erscheint uns, was den gewohnten Ablauf von Ereignissen stört. Dafür einige Beispiele: Sie haben es eilig, zur Schule zu kommen, und sitzen im Autobus. Aber da hat der Autobus eine Panne, und Sie kommen zu spät zum Unterricht. Ein zufälliges Ereignis hat Ihren Tagesablauf, Ihr geregeltes Leben beeinflusst.
Oder nehmen wir ein weiteres Beispiel. Sie hatten beschlossen, am nächsten schulfreien Tag zu wandern. Doch das Wetter verschlechterte sich so sehr, dass man nicht aus dem Hause gehen konnte. Wieder ist durch das unvorhergesehene Eintreten eines Zufalls der normale Gang der Ereignisse in Ihrem Leben gestört worden.

Natürlich sind zufällige Ereignisse nicht immer unerfreulich. Stellen Sie beispielsweise fest, dass sich die Nummer Ihres Loses unter den Gewinnzahlen einer Lotterie befindet, so bringt Ihnen in diesem Falle ein zufälliges Ereignis, das Auslosen der Gewinnzahlen, Nutzen und Freude. Sie können auch einen Freund auf der Straße treffen, ohne mit ihm verabredet gewesen zu sein.

Allen diesen zufälligen Ereignissen ist jedoch gemeinsam, dass sie Abweichungen vom normalen, planmäßigen Ablauf der Erscheinungen mit sich bringen. Deshalb sagt man mitunter auch: Von Zufälligkeiten abgesehen ...

Auch im gesellschaftlichen Leben können zufällige Ereignisse eintreten. So kann etwa in einem bestimmten Gebiet durch eine überraschende Dürre die Ernte einer landwirtschaftlichen Kultur vernichtet worden sein. Das ist eine schwerwiegende Zufälligkeit.
Es gibt deren noch viele andere, die von der Wissenschaft bis jetzt noch nicht vorhergesagt werden können. Ein Erdbeben beispielsweise, das einen ganzen Bezirk vernichten kann, ist vorläufig noch ein nicht im voraus erkennbarer Zufall. Wir sprechen von "Naturgewalten", die die Wissenschaft noch nicht beherrscht.
Die Wissenschaft bemüht sich jedoch, den Menschen in seinem Kampf gegen solche Zufälle zu wappnen.



Mit welchen Mitteln kann man nun beispielsweise den Zufälligkeiten der Witterung begegnen?
Man untersucht die Wachstumsbedingungen der Pflanzen in der betreffenden Gegend möglichst allseitig und pflanzt entsprechend den Erfordernissen dürre- bzw. frostbeständige Sorten an. Dadurch verringert sich der Einfluss von Zufälligkeiten der Witterung auf die Ernte.

Man kann auch versuchen vorherzubestimmen, ob etwa im Frühjahr starke Nachtfröste drohen, und in diesem Falle Vorbereitungen treffen, um die Kulturen beispielsweise durch Abdecken oder durch Räuchern zu schützen.

Schließlich kann man auch versuchen, die allgemeinen Gesetze zu ermitteln, nach denen der Wechsel von gutem und schlechtem Wetter erfolgt, und in Übereinstimmung mit diesen Gesetzen Vorsichtsmaßnahmen ergreifen.
Auch in solchen Fällen, in welchen ein Zufall sich günstig auswirkt, ist es zweckmäßig, die Gesetzmäßigkeiten zu untersuchen, nach denen diese zufälligen Erscheinungen ablaufen, damit der Nutzen möglichst groß wird.

3.2 Gesetzmäßigkeit und Zufall

Auf den ersten Blick mag es scheinen, als ob zufällige Ereignisse keinerlei Gesetzmäßigkeiten gehorchen würden. Gerade deshalb heißen sie ja zufällige Ereignisse.
Bei näherem Hinsehen gelangt man jedoch zu der Einsicht, dass auch zufällige Erscheinungen oft nicht völlig chaotisch ablaufen. In zahlreichen Fällen lassen sich auch hier bestimmte Gesetzmäßigkeiten beobachten. Diese Gesetzmäßigkeiten sind jedoch nicht mit den gewöhnlichen physikalischen Gesetzen vergleichbar; sie sind ganz anderer Natur.
Trotzdem kann man sagen, dass es für zufällige Erscheinungen Gesetzmäßigkeiten gibt.

Wir wollen ein Beispiel betrachten, das so einfach ist, dass jeder die weiter unten beschriebenen Erscheinungen selbst beobachten kann. An diesem einfachen Beispiel lassen sich bereits allgemeine Gesetze des Zufalls feststellen.

Wir stellen uns vor, dass wir eine Münze auf den Tisch werfen. Dann sehen wir nach, welche Seite nach oben zeigt, Zahl oder Wappen. Das "Spiel" wird nach den älteren Münzen auch "Kopf oder Adler" genannt. Wirft man die Münze nur ein einziges Mal, so kann man nicht vorhersagen, welche Seite oben liegen wird.
Wenn man aber die Münze hundertmal hintereinander wirft, so kann man schon viel mehr darüber aussagen, wie oft die eine und wie oft die andere Seite der Münze oben liegt.

Man kann mit Sicherheit behaupten, dass das Wappen nicht bloß ein- oder zweimal erscheinen wird, sondern häufiger, und auch nicht 98- oder 99 mal, sondern seltener.
Münzen sind doch wenigstens im allgemeinen gleichmäßig gearbeitet, so dass keine der beiden Seiten vor der anderen ausgezeichnet ist. Naturgemäß darf man daher erwarten, dass die Münze ungefähr bei der Hälfte aller Würfe mit dem Wappen nach oben und in der Hälfte aller Würfe mit der Zahl nach oben zu liegen kommt.
Das Wappen wird somit wahrscheinlich etwa 50 mal oben liegen. Selbstverständlich kann man nicht behaupten, dass das genau 50 mal geschieht.

Mit einiger Sicherheit lässt sich allerdings vorhersagen, dass das Wappen mindestens 40 mal und höchstens 60 mal oben liegen wird. Man kann das leicht nachprüfen und sich davon überzeugen, dass das immer so ist.

Man kann das Ergebnis sogar noch genauer vorhersagen. Es wird kein besonderes Risiko sein, zu behaupten, das Wappen werde mindestens 43 mal und höchstens 57 mal oben liegen und die Zahl dementsprechend höchstens 57 mal und mindestens 43 mal.
Was man aber nicht vorhersagen kann, ist die genaue Zahl der Fälle, in denen das Wappen oben liegt. Sie hängt vom Zufall ab. Allerdings hat, wie wir sehen werden, auch der Zufall seine Gesetze.

In unserem Beispiel führt er dazu, dass sich die Anzahl der Fälle, in denen das Wappen oben liegt, bei verschiedenen "Versuchsreihen" von je hundert Würfen in den Grenzen von 43 bis 57 bewegen wird.
Fast niemals wird sich mehr als 57 mal oder weniger als 43 mal "Wappen" ergeben. Genauer gesagt kann man ausrechnen, dass nur in ungefähr 15% der Fälle das Wappen mehr als 57 mal oder weniger als 43 mal oben liegen wird, wenn man sehr viele Male je 100 mal die Münze wirft.

Wir bringen ein weiteres Beispiel. Denken wir an einen guten Schützen, von dem man sagt, er erziele 93% Treffer. Was heißt das?

Das bedeutet, dass von je 100 Schüssen, die er auf eine Zielscheibe abgibt, ungefähr 93 treffen und etwa 7 vorbeigehen. Unter diesen Umständen kann man auch von den nächsten 100 Schüssen, die der Schütze abgeben wird, voraussagen, dass davon etwa 93 ihr Ziel erreichen werden.
Mit Sicherheit kann man behaupten, die Trefferzahl werde größer als 80 sein. Selbstverständlich unterliegt der Prozentsatz der Treffer mit der Zeit gewissen Schwankungen.

So kann sich der Schütze im Schießen vervollkommnen oder, umgekehrt, aus der Übung kommen. Das gleiche kann man von einem Jäger sagen.

Wir wollen noch ein drittes Beispiel für eine Gesetzmäßigkeit angeben, der Ereignisse unterworfen sind, die auf den ersten Blick völlig zufällig erscheinen. Wenn eine Familie ein Kind erwartet, kann niemand voraussagen, ob es ein Junge oder ein Mädchen wird. Selbst wenn zu erwarten ist, dass eine Familie mehrere, sagen wir zwei, drei oder vier Kinder haben wird, lässt sich von vornherein nicht sagen, wieviel davon Jungen sein werden. In allen Ländern und bei allen Völkern sind jedoch im Durchschnitt von 1000

Neugeborenen 511 Jungen und 489 Mädchen.

Die auffallende Konstanz dieses Verhältnisses wurde von zahlreichen Gelehrten bemerkt, darunter von einem der Begründer der Wahrscheinlichkeitsrechnung, dem französischen Mathematiker Pierre Simon Laplace (1749-1827).
In den Fällen, in denen uns also die Anzahl der Jungen bei einer sehr großen Zahl von Geburten interessiert, können wir diese Zahl mit einem hohen Genauigkeitsgrad sicher voraussagen, wobei wir uns nur ganz selten irren werden.

Bei der Durchsicht des Geburtenregisters der Stadt Paris für die Jahre 1745-1784 stellte Laplace fest, dass das Verhältnis der Anzahl der Knaben zur Gesamtzahl der Neugeborenen ungefähr 0,510 betrug, d.h. etwas kleiner war als 0,511.
Obwohl diese Differenz sehr klein war, schloss Laplace daraus auf einen besonderen Grund, der die Anzahl der Mädchen vergrößert haben musste. Die Anzahl der Geburten in diesen 39 Jahren war auch zu damaliger Zeit in Paris schon sehr hoch, so dass selbst eine so kleine Abweichung vom normalen Verhältnis nicht durch Wirkung des Zufalls erklärt werden konnte.

Und tatsächlich, Laplace fand den Grund für die Abweichung. In der Zahl der in Paris gemeldeten Kinder war die Anzahl derjenigen Kinder enthalten, die in einem besonderen Waisenhaus (für Findelkinder) abgegeben worden waren, dem einzigen in ganz Frankreich.
Da die französischen Bauern jedoch Söhne als zukünftige Arbeitskräfte mehr schätzten als Töchter, wurden häufiger Mädchen ausgesetzt als Jungen.

Als Laplace die Anzahl der Findelkinder (von denen viele tatsächlich nicht in Paris geboren waren) von der Anzahl der Neugeborenen abzog, gelangte er auch für Paris zu dem normalen Verhältnis der Anzahl der Jungen zur Anzahl der Mädchen.
Damit hatte er gezeigt, dass man bei einer sehr großen Zahl von Geburten die Gesamtzahl der Knaben mit einer Genauigkeit von 0,1% voraussagen kann. Jede größere Abweichung muss besondere Gründe haben-

Jetzt wollen wir versuchen, die hier besprochenen Gesetzmäßigkeiten in mathematischer Sprache auszudrücken.
Wir untersuchen ein Ereignis A, das in jedem einzelnen Falle eintreten oder auch nicht eintreten kann. Es mögen N Versuche durchgeführt werden, bei denen in M Fällen das Ereignis A eintrete, in $N - M$ Fällen dagegen nicht.

Die Häufigkeit des Ereignisses A (d.h. das Verhältnis der Zahl der Versuche, in denen das Ereignis A eintritt, zur Gesamtzahl der Versuche) ist hier gleich $\frac{M}{N}$.
Für große Werte von N erweist sich die Häufigkeit eines Ereignisses als nahezu konstant. So liegt beispielsweise beim Werfen einer Münze die Häufigkeit $\frac{M}{N}$ des Falles, dass das Wappen oben liegt, stets um $\frac{1}{2}$ und bei Geburten liegt die Häufigkeit $\frac{M}{N}$ der Geburt eines Jungen nahe bei 0,511.

Es ist nun ein Grundgesetz der zufälligen Erscheinungen, dass die Häufigkeit eines bestimmten Ereignisses bei einer sehr großen Anzahl von Versuchen (Fällen) konstant ist. Je größer die Anzahl der Versuche, desto kleiner ist eine zufällige Abweichung der

Häufigkeit von ihrem durchschnittlichen Wert.
Dieses Gesetz wird auf vielen Gebieten der Volkswirtschaft praktisch angewandt.

Man kann sogar behaupten, dass es keinen einzigen Zweig der Volkswirtschaft gibt, in dem nicht dieses Gesetz und andere Gesetze der zufälligen Erscheinungen, von denen im folgenden noch die Rede sein wird, benutzt würden.

Als Beispiel dafür möge die Lebensversicherung dienen. Nehmen wir einmal an, ein Mann im Alter von 55 Jahren möchte eine Lebensversicherung abschließen. Um nun den Betrag, den er monatlich einzuzahlen hat, mit der Versicherungssumme, die den Hinterbliebenen im Falle seines Todes zu zahlen ist, abstimmen zu können, muss man berechnen, wie groß die Lebenserwartung eines 55 jährigen ist, d.h. ermitteln, wieviel Menschen, die heute 55 Jahre alt sind, in 5 oder 10 Jahren noch leben und wieviel während dieser Zeit sterben.

Hier scheint jede Voraussicht unmöglich; denn keiner weiß, wann er sterben wird. Es zeigt sich jedoch, dass man bei einer großen Anzahl von Fällen mit Bestimmtheit voraussagen kann, wieviel Menschen mindestens 60 oder mindestens 65 Jahre alt werden.

Wenn man die Gesamtheit aller Versicherten berücksichtigt, kann man sehr genaue Vorhersagen machen. Diese werden um so genauer sein, je größer die Anzahl der Versicherten ist.
Ganz entsprechend kann man im voraus berechnen, wieviel Feuerwachen eine Stadt braucht oder welche Getreidevorräte ein Staat oder ein Bezirk für den Fall einer Missernte zur Verfügung haben sollte.

3.3 Die Wahrscheinlichkeit eines Ereignisses

Die Zahl, um welche die bei großen Versuchsreihen gleichbleibende Häufigkeit eines bestimmten zufälligen Ereignisses schwankt und der sie sich annähert, heißt die Wahrscheinlichkeit dieses Ereignisses.
Mit der Untersuchung von Wahrscheinlichkeiten und ihren Gesetzmäßigkeiten befasst sich die Wahrscheinlichkeitsrechnung.

In einigen Fällen lässt sich die Wahrscheinlichkeit eines Ereignisses leicht im voraus bestimmen, ohne dass man erst irgendwelche Versuche durchführen muss.
Wie groß ist beispielsweise beim Werfen einer Münze die Wahrscheinlichkeit dafür, dass das Wappen oben liegt?

Offenbar ist die Chance dafür, dass dieses Ereignis eintritt, genauso groß wie die, dass es nicht eintritt. Daher ist die Wahrscheinlichkeit für das Eintreten des betrachteten Ereignisses A hier gleich $\frac{1}{2}$. Man schreibt das so:

$$P(A) = \frac{1}{2}$$

Der Buchstabe P ist der Anfangsbuchstabe des französischen Wortes probabilité (= Wahrscheinlichkeit).

Das bedeutet, dass bei einer großen Anzahl von Würfen ungefähr in der Hälfte aller Fälle "Wappen" geworfen wird. Was bedeutet

$$P(A) = 0{,}003?$$

Nun, diese Schreibweise besagt, dass bei einer großen Anzahl von Versuchen das Ereignis A etwa in drei von tausend Fällen eintreten wird.
Die Aussage

$$P(A) = 1$$

bedeutet, dass das Ereignis A bestimmt eintritt. So ist z.B. die Wahrscheinlichkeit dafür, dass morgen früh die Sonne aufgehen wird, gleich Eins. Dagegen besagt die Gleichung $P(A) = 0$, dass das Ereignis A unmöglich ist.
Beispielsweise ist die Wahrscheinlichkeit dafür, dass eine auf den Tisch geworfene Münze auf der Kante stehen bleibt, gleich Null.

Im folgenden werden wir einige Aufgaben kennenlernen, mit denen sich die Wahrscheinlichkeitsrechnung beschäftigt.
Die ersten Beispiele mögen vielleicht etwas sonderbar anmuten. Es sind nämlich Aufgaben über die Gewinnaussichten im Würfelspiel.

Man sollte sich dabei aber vor Augen halten, dass sich die Wahrscheinlichkeitsrechnung im 17. Jahrhundert gerade bei der Lösung von Problemen zu entwickeln begann, die beim Würfelspiel auftauchten. Manche Mathematiker bemerken in diesem Zusammenhang scherzhaft, dass aus einem so einfachen Spiel wie dem Würfeln eine umfangreiche und tiefgründige, für die praktische Tätigkeit des Menschen so wichtige Wissenschaft wie die Wahrscheinlichkeitsrechnung entstand, während das geistvolle Schachspiel in der Geschichte der Wissenschaften keine Rolle gespielt habe.
Bei genauerem Hinschauen zeigt sich allerdings, dass auch beim Schachspiel interessante mathematische Fragestellungen auftauchen.

3.4 Welche Probleme förderten die Wahrscheinlichkeitsrechnung?

Wir wollen jetzt einige der einfachen Aufgaben bringen, die bereits im 17. Jahrhundert gelöst wurden.

Man würfelt mit einem Spielwürfel, d.h. mit einem Würfel, dessen sechs Seitenflächen mit den Zahlen 1 bis 6 nummeriert sind. Wie groß ist die Wahrscheinlichkeit dafür, dass man eine 5 wirft?

Der Würfel kann auf 6 verschiedene Seitenflächen fallen. Da alle Seiten gleichberechtigt sind und der Würfel als homogen, nicht als gefälscht vorausgesetzt wird, darf man annehmen, dass alle Seiten gleich oft oben erscheinen, also jede Seite in einem Sechstel der Fälle drankommt. Demnach ist

$$P(5) = \frac{1}{6}$$

Hierbei ist $P(5)$ die Wahrscheinlichkeit dafür, dass eine Fünf gewürfelt wird.
Wie groß ist die Wahrscheinlichkeit dafür, dass eine gerade Augenzahl gewürfelt wird? Hier kann eine 2, eine 4 oder eine 6 fallen. Es gibt also drei Möglichkeiten. Deshalb ist

$$P(\text{gerade Augenzahl}) = \frac{3}{6} = \frac{1}{2}$$

Wie groß ist die Wahrscheinlichkeit dafür, gleichzeitig mit zwei Würfeln Fünfen zu würfeln?
Beim Würfeln mit zwei Würfeln gibt es 36 mögliche Fälle: (1; 1), (1; 2), (1; 3), ..., (1; 6), (2; 1), (2; 2), ..., (2; 6) usf. bis (6; 1), (6; 2), ..., (6; 6). (Hier gibt die erste Ziffer die mit dem ersten Würfel geworfene Augenzahl an, die zweite die Augenzahl des zweiten.) Alle Fälle sind völlig gleichberechtigt; es ist also

$$P(5;5) = \frac{1}{36}$$

Dabei ergab sich offenbar

$$P(5;5) = P(5) \cdot P(5)$$

Diese Beziehung ist ein Spezialfall des wichtigen Multiplikationssatzes für Wahrscheinlichkeiten, auf den wir hier aber nicht weiter eingehen können.
Wie groß ist die Wahrscheinlichkeit dafür, dass die Summe der mit zwei Würfeln geworfenen Augenzahlen gleich 8 wird? Offenbar sind folgende Fälle möglich, in denen die Augensumme gleich 8 ist: (2; 6), (3; 5), (4; 4), (5; 3), (6; 2).

Es gibt also insgesamt fünf Kombinationen, die unsere Bedingung erfüllen. Bei der Lösung der vorigen Aufgabe hatten wir bemerkt, dass jede dieser Kombinationen bei 36 aller Würfe auftreten kann.
Die Bedingung, dass die Augensumme 8 beträgt, wird also ungefähr in fünf von jeweils 36 Würfen erfüllt. Somit ist die gesuchte Wahrscheinlichkeit gleich $\frac{5}{36}$.

Jetzt bringen wir ein etwas komplizierteres historisches Beispiel, das zeigt, wie im 17. Jahrhundert die praktischen Probleme der Würfelspieler die Entwicklung der Wahrscheinlichkeitsrechnung förderten, d.h. letztlich dazu beitrugen, die Probleme der Gesetzmäßigkeiten des Zufalls wissenschaftlich zu formulieren.

Ein französischer Adliger, Chevalier de Mère, war ein leidenschaftlicher Würfelspieler. Er versuchte alles, um durch das Spiel reich zu werden, und dachte sich zu diesem Zweck verschiedene komplizierte Regeln aus, die ihm, so hoffte er, zu seinem Ziel verhelfen würden. Damals wollten viele durch Glücksspiele reich werden; das griff wie eine Seuche um sich.

Chevalier de Mère hatte sich beispielsweise folgendes Spielchen überlegt.
Es sollte viermal hintereinander mit einem Würfel gewürfelt werden. Er wettete, dass dabei wenigstens einmal eine 6 fallen würde. Anderenfalls, wenn also unter den vier Würfen keine 6 vorkam, sollte sein Gegenspieler gewonnen haben.

Damals war der genaue Wert der Wahrscheinlichkeit dafür, dass unter diesen Bedingungen eine 6 fällt, unbekannt, obwohl klar war, dass er bei $\frac{1}{2}$ liegt.
De Mère nahm an, dass er häufiger gewinnen würde als verlieren. Vorsichtshalber wandte er sich jedoch an einen seiner Bekannten, einen der hervorragendsten Mathematiker des 17. Jahrhunderts, Blaise Pascal (1623-1662), mit der Bitte, auszurechnen, wie groß die Wahrscheinlichkeit eines Gewinnes in dem von ihm erdachten Spiel sei.

Pascal überlegte folgendermaßen.
Bei jedem einzelnen Wurf ist die Wahrscheinlichkeit für das Würfeln einer Sechs gleich $\frac{1}{6}$. Die Wahrscheinlichkeit dafür, dass keine Sechs fällt, ist $\frac{5}{6}$.

Nun werde zweimal gewürfelt. Der Versuch, der im zweimaligen Würfeln besteht, werde mehrfach, etwa N-mal, wiederholt. Dann fällt in ungefähr $\frac{5}{6}$ dieser N Fälle beim ersten Würfeln keine Sechs. In ungefähr $\frac{5}{6}$ dieser $\frac{5}{6}\cdot N$ Fälle, d.h. insgesamt in

$$\frac{5}{6}\left(\frac{5}{6}\cdot N\right)=\left(\frac{5}{6}\right)^2\cdot N$$

Fällen, wird auch beim zweiten Würfeln keine Sechs geworfen. Somit ist die Wahrscheinlichkeit dafür, dass bei zweimaligem Würfeln keine Sechs fällt, gleich

$$\left(\frac{5}{6}\right)^2=\frac{25}{36}$$

Genauso zeigt man, dass die Wahrscheinlichkeit dafür, bei dreimaligem Würfeln keine Sechs zu werfen, gleich

$$\left(\frac{5}{6}\right)^3=\frac{125}{216}$$

ist. Die Wahrscheinlichkeit dafür, dass bei viermaligem Würfeln keine Sechs fällt, ist schließlich gleich

$$\left(\frac{5}{6}\right)^4=\frac{625}{1296}$$

Die Wahrscheinlichkeit, zu verlieren, war also für den Chevalier de Mère gleich $\frac{625}{1296}$, d.h. kleiner als $\frac{1}{2}$; die Wahrscheinlichkeit, zu gewinnen, war also größer als $\frac{1}{2}$.
Somit hatte de Mère in dem Spiel die besseren Aussichten, zu gewinnen. Bei oftmaliger Wiederholung des Spiels war er mit großer Sicherheit der Gewinner.

Und in der Tat, je öfter de Mère spielte, um so mehr gewann er. Chevalier de Mère war sehr zufrieden und dachte, er habe den Weg zum Reichtum entdeckt. Mit der Zeit kamen allerdings die anderen Spieler darauf, dass dieses Spiel für sie unvorteilhaft war, und sie spielten es nicht mehr mit ihm.
Jetzt musste sich de Mère neue Regeln ausdenken. Mit 2 Würfeln sollte 24 mal gewürfelt werden, und er wettete, dass dabei wenigstens ein Paar Fünfen fallen würden.

De Mère rechnete damit, auch bei diesem Spiel häufiger zu gewinnen als zu verlieren, und glaubte, auf den Rat seines Freundes Pascal verzichten zu können.
Doch diesmal irrte er sich. Die Wahrscheinlichkeit dafür, dass mit zwei Würfeln ein Paar Fünfen geworfen wird, ist, wie wir wissen, gleich $\frac{1}{36}$.
Die Wahrscheinlichkeit dafür, dass nicht zwei Fünfen fallen, ist also gleich $\frac{35}{36}$. Dementsprechend ist die Wahrscheinlichkeit dafür, dass bei 24 maligem Würfeln mit zwei Würfeln keine zwei Fünfen fallen, gleich

$$\left(\frac{35}{36}\right)^{24}$$

und diese Zahl ist größer als $\frac{1}{2}$. Folglich war für de Mère die Wahrscheinlichkeit, zu verlieren, größer als $\frac{1}{2}$.

Es war also zu erwarten, dass er um so mehr verlieren würde, je öfter er spielte. Und so geschah es denn auch. Je mehr er spielte, um so ärmer wurde er, so dass er schließlich als Bettler endete.

Das Interessanteste an dieser historischen Anekdote ist zweifelsohne, dass sich aus solchen eigenartigen "praktischen Problemen" eine Theorie zur Berechnung von zufälligen Ereignissen entwickelte.
Die Gelehrten des 17. und 18. Jahrhunderts erblickten in diesen Beispielen nichts als "amüsante Möglichkeiten" für die Anwendung mathematischer Kenntnisse auf Erscheinungen, die von untergeordneter Bedeutung waren. Ein würfelnder Glücksritter verdiente es wahrlich nicht, dass man zu seiner Unterstützung eine besondere Wissenschaft entwickelte.

3.5 Die Wahrscheinlichkeitsrechnung half auch im Kriege

Es zeigte sich aber, dass das Anwendungsgebiet der Wahrscheinlichkeitsrechnung erstaunlich groß ist.
Die Wahrscheinlichkeitsrechnung befasst sich mit der Untersuchung aller Massenerscheinungen, d.h. aller sich häufig wiederholenden Zufälle, ob sie nun im alltäglichen Leben, in der Wissenschaft oder in der Technik auftreten.
Wir wollen ein Beispiel aus dem militärischen Bereich anführen.

Bekanntlich ist es sehr schwer, ja nahezu aussichtslos, mit einem Gewehrschuss ein Flugzeug abzuschießen. Das Flugzeug muss ja nicht nur irgendwo, sondern an einer entscheidenden Stelle getroffen werden. Daher ist die Wahrscheinlichkeit dafür, dass ein Flugzeug mit einem einzigen Gewehrschuss abgeschossen wird, in der Tat ganz gering. Der Schütze wird höchstwahrscheinlich das Flugzeug überhaupt nicht treffen.

Völlig anders liegen die Verhältnisse bei ganzen Salven. Mit Hilfe der Wahrscheinlichkeitsrechnung kann man ausrechnen, wieviel Schüsse zur selben Zeit auf ein gegnerisches Flugzeug abgefeuert werden müssen, damit eine Chance besteht, es abzuschießen. Bei dieser Rechnung verfahren wir folgendermaßen.

Wir nehmen an, die Wahrscheinlichkeit, ein Flugzeug mit einem einzigen Schuss abzuschießen, sei 0,004. Mit anderen Worten, die Wahrscheinlichkeit dafür, dass ein Schütze das Flugzeug nicht trifft, ist 0,996.
Selbst wenn ein Schütze viele Male auf das Flugzeug schießt, hat er nur ganz geringe Chancen, es zu treffen.

Jetzt wollen wir annehmen, dass gleichzeitig 500 Schützen auf das Flugzeug schießen. Die Wahrscheinlichkeit dafür, dass keiner der 500 Schützen das Flugzeug trifft, ergibt sich zu

$$(0.996)^{500}$$

(Wir erinnern uns, wie wir die Gewinn- bzw. Verlustchancen bei den Spielen des Chevalier de Mère berechnet haben.) Mit Hilfe einer Logarithmentafel finden wir leicht, dass

$$(0.996)^{500} \approx 0{,}14$$

ist. Das ist also die Wahrscheinlichkeit dafür, dass von 500 gleichzeitig abgefeuerten Gewehrschüssen kein einziger das Flugzeug trifft.
Mit anderen Worten, die Wahrscheinlichkeit dafür, dass das Flugzeug getroffen wird, ist 0,86. Man kann also sagen, das Flugzeug werde bei einer Salve von 500 Schüssen mit ziemlicher Sicherheit abgeschossen.
Die Chance, dass das Flugzeug dabei zum Absturz gebracht wird, ist sogar viel größer als die, dass es diesem Abwehrfeuer heil entkommt.

In Wirklichkeit ist die Wahrscheinlichkeit dafür, dass ein einzelner Schuss das Flugzeug trifft, kleiner als 0,004. Man nimmt im allgemeinen an, dass sie bei 0,001 liegt. Allerdings ist auch in diesem Falle die Wahrscheinlichkeit dafür, dass das Flugzeug bei einer Salve von 500 Schüssen getroffen wird, gleich 0,54. Die Chance, dass es zum Absturz gebracht wird, ist also immer noch größer als die, dass es entkommt.

Die Wahrscheinlichkeitsrechnung wurde während des zweiten Weltkrieges auch angewendet, um die zweckmäßigsten Methoden bei der Suche nach feindlichen Flugzeugen oder Unterseebooten herauszufinden oder um Routen zu ermitteln, auf denen eine Begegnung mit feindlichen Kräften ausgeschlossen war.
Eine typische Aufgabe dieser Art ist die Festlegung des Kurses für einen Konvoi von Handelsschiffen über den Ozean, auf dem feindliche U-Boote operieren. Wenn man die Konvois aus relativ vielen Schiffen zusammenstellt, kann man mit relativ wenigen Konvois auskommen, was die Wahrscheinlichkeit für eine Begegnung mit feindlichen Unterseeboten verkleinert.

Andererseits werden dann natürlich bei einem Zusammentreffen die Verluste größer.

Die Methoden der Wahrscheinlichkeitsrechnung ermöglichten es, optimale Pläne aufzustellen, nach denen sich unter den Bedingungen des Krieges der Schiffsverkehr abspielte. Die günstigsten Längen der Konvois und die Häufigkeit ihrer Fahrten wurden berechnet. Aufgaben dieser Art waren in einem derartigen Umfang zu lösen, dass bei den Stäben Spezialgruppen gebildet werden mussten, die diese Probleme der Wahrscheinlichkeitsrechnung zu bearbeiten hatten.

Ähnliche Methoden wurden nach dem Kriege auch bei der Lösung zahlreicher volkswirtschaftlicher Probleme angewendet.
Sie bilden den Inhalt einer neuen, inhaltsreichen Disziplin, der sogenannten Unternehmensforschung, die sich vor unseren Augen zu einer selbständigen Wissenschaft entwickelt.

3.6 Aus der Geschichte der Wahrscheinlichkeitsrechnung

Im 17. Jahrhundert beschäftigten sich solche hervorragenden Mathematiker wie die Franzosen B. Pascal und P. Fermat sowie der Holländer Ch.Huygens mit Wahrscheinlichkeitsrechnung.
Wie wir gesehen haben, wurden die ersten Ergebnisse im Zusammenhang mit dem Studium von Glücksspielen erzielt.

Schon gegen Ende des 17. Jahrhunderts jedoch begann man, die Wahrscheinlichkeitsrechnung bei der Versicherung von Schiffen gegen Unglücksfälle zu verwenden. Als Begründer der Versicherungsmathematik wird Jan de Witt (1625-1672) angesehen, der zeitweilig Präsident der Republik der Niederlande war.
So berechnete man die Chancen dafür, dass ein Schiff unversehrt in den heimatlichen Hafen zurückkehrte, dass es nicht im Sturm unterging und nicht von Piraten gekapert wurde, dass die Ladung nicht nass wurde, nicht verbrannte, nicht faulte usw.

Auf Grund dieser Berechnungen konnten die Versicherungsprämien festgelegt werden, die an die Versicherungen gezahlt werden mussten, ebenso wie die Summen, die den Reedern oder den Besitzern der Ladung im Schadensfalle zu erstatten waren, ohne dass die Versicherungen Einbußen erlitten. Das waren ja Unternehmen, die Profit machen wollten.

Sehr viel hat der Schweizer Mathematiker Jakob Bernoulli (1654-1705) für die Entwicklung der Wahrscheinlichkeitsrechnung getan. Gleich nach ihm müssen die Mitglieder der französischen Akademie P.S. Laplace und S.D. Poisson (1781-1840) sowie der berühmte deutsche Mathematiker C.F. Gauß genannt werden, die Ende des 18. und Anfang des 19. Jahrhunderts gearbeitet haben.

Trotz alledem kann man für die zweite Hälfte des 18. und die erste Hälfte des 19. Jahrhunderts von der Wahrscheinlichkeitsrechnung sagen, dass sie "auf der Stelle trat".
Zu einem großen Fortschritt in der Wahrscheinlichkeitsrechnung haben Mitte des 19. Jahrhunderts die Arbeiten des bedeutenden russischen Mathematikers P.L. Tscheby-

schew (1821-1894) geführt. Er fand neue Wege zur Lösung von Aufgaben, die schon früher formuliert worden waren, und er vermochte eine große Gruppe junger Wissenschaftler um sich zu scharen, von denen in der Folgezeit einige weltberühmt wurden.

Hier sind vor allem A.A. Markow (1856-1922) und A.M. Ljapunow (1857-1918) zu nennen. Etwas später trat S.N. Bernstein, der vor kurzem seinen 90. Geburtstag feierte, mit seinen Arbeiten hervor.
Die sowjetische Schule der Wahrscheinlichkeitsrechnung unter Leitung der Akademiemitglieder A.N. Kolmogorow und B.W. Gnedenko bewahrt die Tradition der vorrevolutionären russischen Mathematik und entwickelt sie weiter.

In Deutschland ist vor allem R. v. Mises (1883-1953) zu nennen, der wesentliche Ergebnisse erzielte. Wichtige Resultate wurden von der französischen Schule (M.Frechet, P.Levy u.a.), in Polen (H.Steinhaus, M.Fisz u.a.), in Ungarn (A.Rényi u.a.) und in anderen Ländern gewonnen.

In der jüngsten Zeit wurden Probleme mit wahrscheinlichkeitstheoretischen Methoden angepackt, die man früher der "klassischen, exakten" Mathematik vorbehalten glaubte. Diese Methoden sind unter der Bezeichnung "Monte-Carlo-Methoden" bekannt. Sie wurden in Zusammenhang mit der Verbreitung und dem Einsatz von programmgesteuerten Rechenautomaten fruchtbar.

3.7 Welche Aufgaben löst man heute mit Methoden der Wahrscheinlichkeitsrechnung?

Im Laufe des 19. Jahrhunderts gelangten viele Physiker dazu, die Ergebnisse der Wahrscheinlichkeitsrechnung auch auf physikalische Erscheinungen anzuwenden.
In der zweiten Hälfte dieses Jahrhunderts entwickelten der Engländer J.C. Maxwell (1831 bis 1879), der Österreicher L. Boltzmann (1844-1906) und der Amerikaner J. W. Gibbs (1839-1903) die sogenannte "statistische Physik", einen Zweig der Physik, welcher große Gesamtheiten von Atomen oder Molekülen eines beliebigen Stoffes speziell vom Standpunkt der Wahrscheinlichkeitsrechnung untersucht.

Seit der Mitte des 19. Jahrhunderts machte die Wahrscheinlichkeitsrechnung gewaltige Fortschritte und eroberte immer neue Anwendungsgebiete. Auch die Wissenschaft des Zufälligen selbst erhielt neue Impulse.
In dem Maße, in dem zufällige Massenerscheinungen immer tiefgehender studiert wurden, entdeckte man auch neue Gesetzmäßigkeiten des Zufalls.

Von Jahr zu Jahr wächst die praktische Bedeutung der Wahrscheinlichkeitsrechnung, jenes Zweiges der Mathematik, der die Massenerscheinungen untersucht.
In letzter Zeit hat die Wahrscheinlichkeitsrechnung beispielsweise in der Astronomie wichtige Anwendungen gefunden. Die Bedeutung der Wahrscheinlichkeitsrechnung für die Astronomie ist seither außerordentlich gestiegen, da die Astronomen nicht mehr nur einzelne Planeten und Sterne untersuchen, sondern begannen, sich für ganze Sternkomplexe, Sternhaufen und außergalaktische Nebel oder Galaxien, die aus unvorstellbar vielen Sternen bestehen, zu interessieren.

Während der letzten Zeit wurde in der Astronomie auch eine Richtung entwickelt, die mit den Methoden der Wahrscheinlichkeitsrechnung die Gesamtmasse derjenigen Galaxien untersucht, die dem gewaltigen Sternensystem ähneln, zu dem unsere Sonne mit allen Planeten gehört. Die riesige Anzahl der Galaxien im Universum ermöglicht es, sie mit den Methoden zu behandeln, die sich auf Massenerscheinungen beziehen.

Die Freisetzung der Atomenergie, die Herstellung von Atombomben und später der Bau von Atomkraftwerken, Atommotoren usw. stellten die Wahrscheinlichkeitsrechnung vor neue, schwierige und wichtige Probleme.
In jeder Atomanlage spielen Kernreaktionen die Hauptrolle, also Kernspaltungen oder Kernfusionen, die bei zufälligen Zusammenstößen von Kernen mit Elementarteilchen (in erster Linie Neutronen) oder von Kernen untereinander ausgelöst werden.
Hierbei ist es äußerst wichtig, den Ablauf aller Prozesse in einer Anlage ganz genau vorauszuberechnen.

Der kleinste Fehler kann ja zu einer unabsehbaren Katastrophe führen. Auch hier kommt die Wahrscheinlichkeitsrechnung zu Hilfe. Ungeachtet des zufälligen Charakters der einzelnen Zusammenstöße der Atome ermöglicht sie es, den Ablauf der Veränderungen in der Gesamtmasse der Atome mit größter, Genauigkeit im voraus,zu bestimmen.

Heute hat der handwerkliche Betrieb überall nur noch ganz geringe Bedeutung. Dagegen wächst die Bedeutung der Großproduktion ständig.
Daher wird in der Wirtschaft die Theorie der Massenerscheinungen, die eine wissenschaftliche Produktionslenkung erlaubt, immer wichtiger. Man braucht wissenschaftlich begründete Methoden zur Qualitätskontrolle, Methoden zur Bestimmung der Ausschussquote durch Untersuchung von Stichproben der Produktion. Kein großer Betrieb kann dabei auf die Anwendung der Wahrscheinlichkeitsrechnung verzichten.

Eine einzelne Fabrik ist jedoch oft nicht in der Lage, die bei der wissenschaftlichen Behandlung ihrer praktischen Probleme auftretenden Schwierigkeiten zu überwinden. Mathematiker, die speziell in Wahrscheinlichkeitsrechnung ausgebildet sind, arbeiten an der Schaffung einer allgemeinen Theorie der Massenerscheinungen, die auch alle in der Wirtschaft entstehenden konkreten Aufgaben umfasst.
Heutzutage wenden fast alle Zweige der Großindustrie Ergebnisse aus der Wahrscheinlichkeitsrechnung an.

Oder nehmen wir eine so weitverbreitete Einrichtung wie das Telefon. Dass man das Telefonnetz automatisieren kann, beruht auf der Wahrscheinlichkeitsrechnung. Um die Rolle der Wahrscheinlichkeitsrechnung beim Telefonieren zu beleuchten, betrachten wir ein ganz spezielles Beispiel.

Jeder hat wohl schon einmal die Zeitansage angerufen, eine automatische Anlage in der Telefonzentrale, die sofort sagt, wie spät es ist.
Die Wahrscheinlichkeitsrechnung ermöglicht es, ungefähr die Anzahl der Teilnehmer zu berechnen, die bei der Zeitansage gleichzeitig um Auskunft anrufen. Obwohl es beispielsweise in Berlin sehr viele Leute gibt, die zur selben Minute oder sogar zur selben Sekunde über das Telefon die Uhrzeit erfahren wollen, treten bei der Zeitansage

praktisch niemals Wartezeiten auf.
Wann man auch anrufen mag, die Zeitansage ist niemals besetzt.

Das beruht darauf, dass man mit Hilfe der Wahrscheinlichkeitsrechnung ermitteln kann, wieviel Leute anrufen und zu welcher Zeit das geschieht.
Davon ausgehend, kann man die Minimalzahl der Automaten bestimmen, die zur Befriedigung der Ansprüche von Millionen Einwohnern einer Stadt notwendig sind.

3.8 Wahrscheinlichkeitsrechnung und Informationstheorie

Die Bedeutung der Wahrscheinlichkeitsrechnung für die Praxis wächst in unseren Tagen besonders stürmisch. Neue große Teildisziplinen entstehen und finden sogleich zahlreiche Anwendungen.
Bald wird es keinen Techniker und keinen Ingenieur mehr geben, der ohne Kenntnisse auf diesem Gebiet auskommen kann.
Kenntnisse in der Wahrscheinlichkeitsrechnung werden in Zukunft zur Allgemeinbildung gehören. Die Anzahl der technischen Lehranstalten, an denen Wahrscheinlichkeitsrechnung gelehrt wird, wird ständig größer. In einigen Staaten, z.B. in Japan, gehören Elemente der Wahrscheinlichkeitsrechnung bereits zum Stoff der Oberschulen.

Wir haben schon auf eine wichtige neue Richtung in der Wissenschaft hingewiesen, die eng mit der Wahrscheinlichkeitsrechnung zusammenhängt, und zwar die sogenannte Unternehmensforschung.

Zum Schluss wollen wir noch auf ein ganz junges, aber äußerst wichtiges Teilgebiet der Wahrscheinlichkeitsrechnung eingehen, auf die Informationstheorie.

Die Wahrscheinlichkeitsrechnung beschäftigt sich mit zufälligen Ereignissen, deren Ausgang uns vor ihrem Stattfinden unbekannt ist. Wenn wir erfahren, wie ein Ereignis ausgegangen ist, haben wir also unser Wissen vermehrt, haben wir, wie man auch sagt, Information gewonnen, Information geliefert erhalten.
Dabei werden offenbar unsere Kenntnisse in verschiedenen Fällen ganz verschieden zunehmen. Zu wissen, auf welche Seite eine Münze beim einmaligen Werfen auf den Tisch gefallen ist, ist weniger, als wenn man weiß, welches Ergebnis zwei Würfe mit derselben Münze liefern.

Zu wissen, ob es in Moskau bereits an den Novemberfesttagen schneien wird, muss als beträchtlich wertvoller angesehen werden als zu wissen, dass es in den Winterferien schneien wird.

Das können wir ohne jede Nachprüfung als völlig sicher ansehen. In den Jahren 1947/48 gab der amerikanische Mathematiker und Ingenieur Claude Shannon eine Methode an, die es gestattet, die "Informationsmenge", die in der Angabe des Resultats eines Versuches mit zufälligem Ausgang enthalten ist, quantitativ zu erfassen, zu messen. Das erwies sich für viele Aufgaben in Naturwissenschaft und Technik als sehr nützlich.

Die Shannonschen Arbeiten stehen in unmittelbarem Zusammenhang mit Problemen der Telegrafie. Ein Telegrafenapparat dient zur Übertragung bestimmter Informationen. Man wird sich naturgemäß die Frage vorlegen, wie diese Information am wirtschaftlichsten zu übertragen ist. Bekanntlich wird die Übertragung oft mit Hilfe des Morsealphabetes durchgeführt, bei dem die Buchstaben in Form bestimmter Kombinationen von Punkten und Strichen geschrieben werden.

So wird beispielsweise der Buchstabe "e" im Morsealphabet durch einen Punkt "." dargestellt, während dem Buchstaben "o" drei Striche entsprechen "—".
Es gibt gute Gründe dafür, dass der Buchstabe "e" im Morsealphabet kürzer ist als der Buchstabe "o". In normalen Texten tritt e viel öfter auf als o. Daher ist es so wichtig, für e eine kurze Bezeichnung zu haben.

Somit wurde bereits bei der Aufstellung des Morsealphabetes berücksichtigt, dass in Texten ("in Nachrichten") die einzelnen Buchstaben mit verschiedenen Häufigkeiten (Wahrscheinlichkeiten) auftreten.
Gerade diese Tatsache wird benutzt, um die mit Hilfe dieses Alphabets "kodierten" und zu übertragenden Nachrichten in möglichst kurzer Form darzustellen.

Dieses Verfahren lässt sich wesentlich verbessern, wenn man berücksichtigt, dass die Wahrscheinlichkeit, mit der ein Buchstabe in einer Nachricht vorkommt, wesentlich von dem vorhergehenden Buchstaben abhängt. Ist beispielsweise der vorhergehende Buchstabe ein von i verschiedener Vokal, so ist die Wahrscheinlichkeit größer, dass der Buchstabe selbst ein Konsonant ist.

War der vorhergehende Buchstabe aber ein c, so folgt mit ziemlicher Sicherheit ein h oder ein k.
Ähnliche Abzählungen, die unmittelbar mit der Wahrscheinlichkeitsrechnung zu tun haben, führte seinerzeit bereits der berühmte russische Mathematiker A.A. Markow durch, der zu diesem Zweck Abschnitte aus "Eugen Onegin" und aus anderen Werken der klassischen russischen Literatur einer sorgfältigen Analyse unterzog.

Doch erst Shannon verknüpfte diese Berechnungen mit dem allgemeinen Problem der "Informationsmenge", die man über einen Telegrafen in der Zeiteinheit übertragen kann. Shannon leitete daraus praktische Regeln ab, die es ermöglichen, die vorteilhaftesten Verfahren zur Kodierung und Übertragung von Nachrichten zu bestimmen.

Die von Shannon angegebene Möglichkeit einer Abschätzung der "Informationsmenge"

erwies sich nicht nur für die Zwecke der Telegrafie als sehr wertvoll, sondern auch in allen Fällen, bei denen wir es mit irgendeiner Form der Übertragung oder Speicherung von bestimmten Daten zu tun haben.
Darunter fallen sowohl die automatische Regelung des Zugverkehrs und die modernen Rechenautomaten, die sehr komplizierte Aufgaben selbständig lösen. als auch die Übertragung bestimmter Signale durch die menschlichen Nerven, von den Sinnesorganen zur Gehirnrinde und vom Gehirn zu den Muskeln.

Die Informationstheorie gehört heute zu den mathematischen Grundlagen der Kybernetik, einer wissenschaftlichen Richtung, die eine ganze Reihe von Wissenschaften vereint und einen breiten Kreis von Erscheinungen untersucht, die mit Technik, Physik, Biologie und Ökonomie zusammenhängen.

3.9 Literaturhinweise

Zum Abschnitt Gleichungen und Funktionen

Alexandrow, P. S.: Einführung in die Gruppentheorie. Übersetzung aus dem Russischen. Berlin 1965

Alexandrow, A. D.: Kurven und Flächen. Übersetzung aus dem Russischen. Berlin 1959

Boltjanski, W.G.: Differentialrechnung einmal anders. Übersetzung aus dem Russischen. Berlin 1956

Breidenbach, W.: Die Dreiteilung des Winkels. Leipzig 1951

Enzyklopädie der Elementarmathematik, Band III: Analysis. Übersetzung aus dem Russischen. Berlin 1958

Juschkewitsch, A. P.: Geschichte der Mathematik im Mittelalter. Übersetzung aus dem Russischen. Leipzig 1964

Kurosch, A.G.: Algebraische Gleichungen beliebigen Grades. Übersetzung aus dem Russischen. Berlin 1965

Markuschewitsch, A. I.: Flächeninhalte und Logarithmen. Übersetzung aus dem Russischen. Berlin 1965

Markuschewitsch, A. I.: Komplexe Zahlen und konforme Abbildungen. Übersetzung aus dem Russischen. Berlin 1965

Natanson, I. P.: Summierung unendlich kleiner Größen. Übersetzung aus dem Russischen. Berlin 1955

Natanson, I. P.: Einfachste Maxima- und Minimaaufgaben. Übersetzung aus dem Russischen. Berlin 1965

Schafarewitsch, I. R.: Über die Auflösung von Gleichungen höheren Grades. Übersetzung aus dem Russischen. Berlin 1965

Scherwatow, W. G.: Hyperbelfunktionen. Übersetzung aus dem Russischen. Berlin 1956

Smirnow, W.I.: Lehrgang der höheren Mathematik. Teil I, II. Übersetzung aus dem Russischen. Berlin 1964

Zum Abschnitt Algebra des Endlichen und Unendlichen

Alexandrow, P. S.: Einführung in die Mengenlehre und die Theorie der reellen Funktionen. Übersetzung aus dem Russischen. Berlin 1965

Enzyklopädie der Elementarmathematik, Band I: Arithmetik. Übersetzung aus dem Russischen. Berlin 1966.

Enzyklopädie der Elementarmathematik, Band II: Algebra. Übersetzung aus dem Russischen. Berlin 1966

Görke, L.: Mengen, Relationen, Funktionen. Berlin 1965

Hasse, M.: Grundbegriffe der Mengenlehre und Logik. Leipzig 1965

Haupt, D.: Mengenlehre leicht verständlich. Leipzig 1965

Infeld, L.: Wen die Götter lieben. Übersetzung aus dem Englischen. Wien 1954 und Berlin 1957

Jung, H. W. E.: Einführung in die Zahlentheorie. Leipzig 1950

Kolosow, A. A.: Kreuz und quer durch die Mathematik. Übersetzung aus dem Russischen. Berlin 1963

Kordemski, B. A.: Köpfchen, Köpfchen. Übersetzung aus dem Russischen. Leipzig/Jena/Berlin 1965

Korowkin, P. P.: Ungleichungen. Übersetzung aus dem Russischen. Berlin 1965

Krbek, F. v.: Über Zahlen und Überzahlen. Leipzig 1964

Leman, A. und B. Schoeneberg: Vom periodischen Dezimalbruch zur Zahlentheorie. Leipzig 1952

Maennchen, Ph.: Geheimnisse der Rechenkünstler. Leipzig und Berlin 1924

Peter, R.: Das Spiel mit dem Unendlichen. Leipzig 1963

Sominski, I. S.: Methode der vollständigen Induktion. Übersetzung aus dem Russischen. Berlin 1965

Trachtenbrot, B. A.: Wieso können Automaten rechnen? Übersetzung aus dem Russischen. Berlin 1966

Worobjew, N.N.: Die Fibonaccischen Zahlen. Übersetzung aus dem Russischen. Berlin 1954

Zum Abschnitt Einführung in die Wahrscheinlichkeitsrechnung

Dynkin, E.B. und W.A.Uspenski: Mathematische Unterhaltungen, Teil III (Aufgaben aus der Wahrscheinlichkeitsrechnung). Übersetzung aus dem Russischen. Berlin 1965

Gnedenko, B. W., und A. J. Chintschin: Elementare Einführung in die Wahrscheinlichkeitsrechnung. Übersetzung aus dem Russischen. Berlin 1966

Jaglom, A. M. und I. M. Jaglom: Wahrscheinlichkeit und Information. Übersetzung aus dem Russischen. Berlin 1965

Poletajew, I. A.: Kybernetik. Übersetzung aus dem Russischen. Berlin 1964

Storm, R.: Wahrscheinlichkeitsrechnung, mathematische Statistik und statistische Qualitätskontrolle. Leipzig 1965