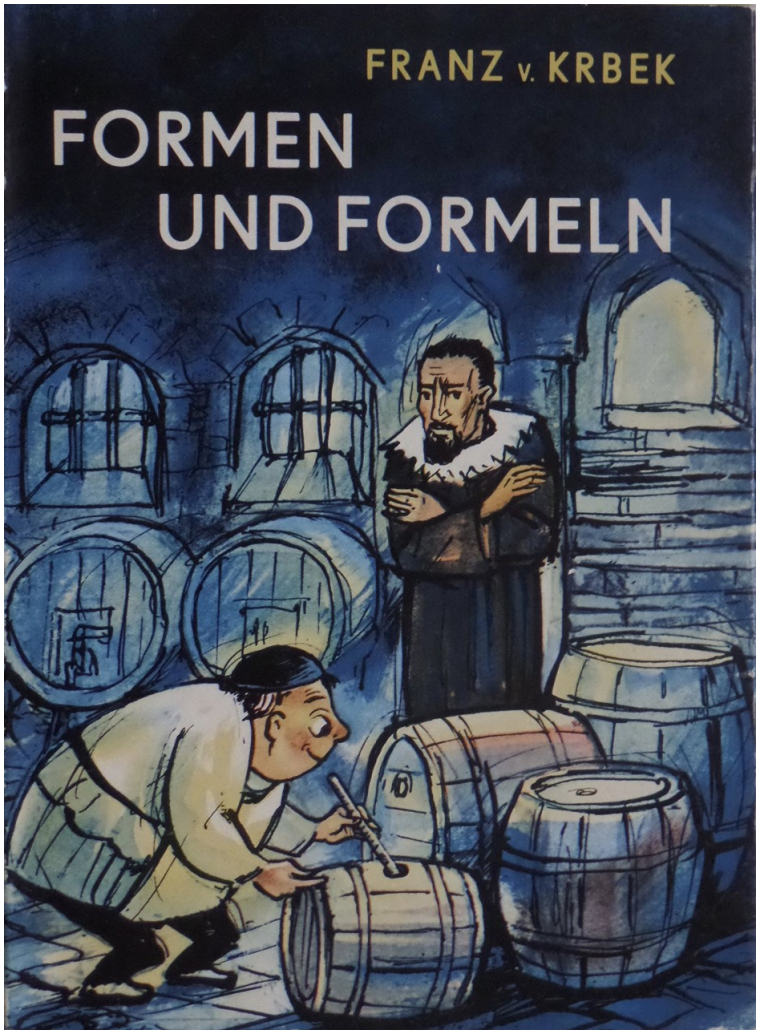
Formen und Formeln
von Prof. Dr. Franz von Krbek



mit 151 Textabbildungen
illustriert von Horst Röcke
1. Auflage

Printed in the German Democratic Republic

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

Vorwort

Die Natur ist ein Buch, das in der Sprache
der Mathematik geschrieben ist.
Galilei

Ohne Mathematik geht es also nicht!
Nun ist das matheFmatische Wissen inzwischen derart angewachsen, dass es der einzelne unmöglich zu beherrschen vermag. Notgedrungen wird er sich also Spezialisieren. Um aber seine Wahl vernünftig treffen zu können, hat er sich ein Grundwissen anzueignen, für das er volles Verständnis jedoch erst dann gewinnt, wenn er auch mit dem Wandel des grundlegenden Ideengutes vertraut ist. Ihm dabei zu helfen, ist Anliegen unserer Plaudereien.

Der Leser soll nicht nur Kenntnisse erlangen, sondern auch einen Blick für die fruchtbare Entwicklung der Mathematik. Damit gewinnt er erst den richtigen Abstand zu den immer neuen Ansätzen, die nötig sind, um diese Wissenschaft voranzutreiben. So begreift er das Wesen der Mathematik wie auch ihre Wirksamkeit zur Veränderung von Natur und Gesellschaft.

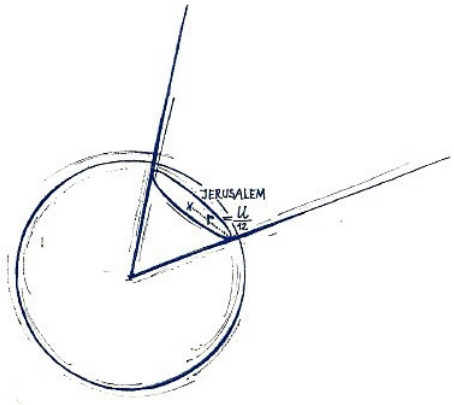
Der B. G. Teubner Verlagsgesellschaft in Leipzig möchte ich für die ansprechende Ausstattung meinen aufrichtigen Dank aussprechen. Besonderen Dank schulde ich aber dem Lektor, Herrn Wolfgang Arnold, für die unermüdliche Sorgfalt, mit der er auch dieses Buch betreute!

Greifswald, im Frühjahr 1967 Franz von Krbek

Inhaltsverzeichnis

1 Keplers Hobby

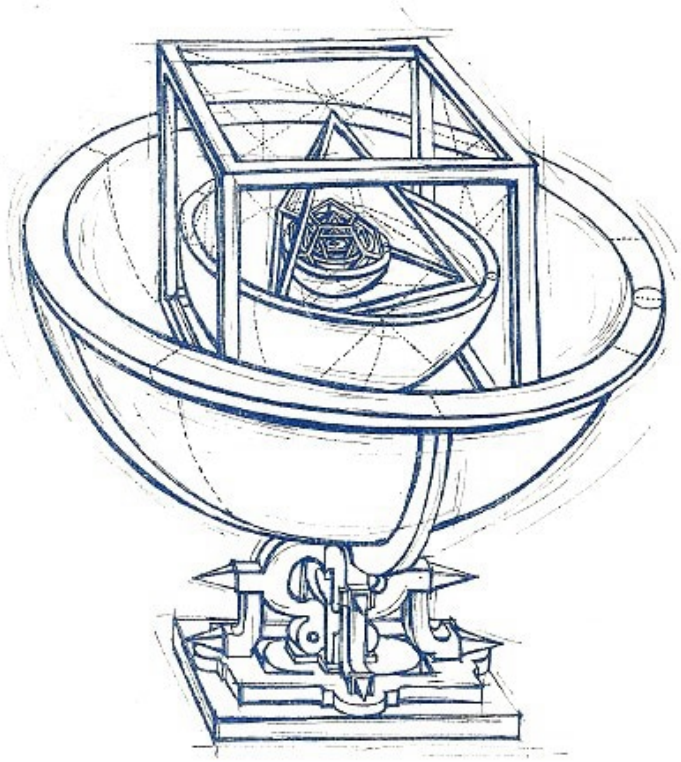
Um Keplers Leistungen gebührend zu würdigen, muss man sich den geistigen Tiefstand seiner Epoche vergegenwärtigen.
Am eindruckvollsten gelingt das an Hand der geradezu grotesk anmutenden Vorlesung seines Zeitgenossen und Brieffreundes Galilei, die dieser im Wintersemester 1587/88 an der Universität zu Pisa hielt. Auf Dante fußend bestimmte er doch allen Ernstes Ort und Gestalt der Unterwelt folgendermaßen:



Wenn man um Jerusalem einen Kreis vom Radius $\frac{1}{12}$ Erdumfang oder rund 3333 Kilometer schlägt und darüber einen Kegel mit der Spitze im Erdmittelpunkt errichtet, hat man die Hölle untergebracht. Auch die Größe des Höllenfürsten wurde von Galilei errechnet.
Während Dante bloß 3 Armlängen groß war, sollte Luzifer die stattliche Höhe von 2000 Armlängen besitzen. Beachtlich, vor allem der ganze Unfug, den der reife Galilei bestimmt als Jugendtorheit abtat, die er 24jährig beging.

So wie das Hobby des 24jährigen Galilei die Hölle war, war es für den 25jährigen Kepler der Himmel. In diesem Alter erschien sein "Mysterium cosmographicum" , das "Weltgeheimnis" , das bei seinem damaligen Vorgehen ewiges Geheimnis geblieben wäre.
Durch seinen Lehrer Maestlin in Tübingen lernte er das kopernikanische Weltbild kennen, insbesondere, dass sich die damals allein bekannten sieben Planeten in exzentrischen Kreisen um die Sonne bewegen. Der jugendliche Kepler erdachte dafür ein Modell, das vielleicht genial, aber bestimmt falsch war.

Wie auf dem Titelblatt seines Werkes zu ersehen ist, wurden Hohlkugeln den fünf regulären Körpern ein- und umgeschrieben, und zwar so, dass die Hohlkugeln und die regulären Körper abwechselnd ineinandergestellt waren. Die Dichte der Wandung der einzelnen Hohlkugeln wurde so gewählt, dass die entsprechende Planetenbahn gerade hineinpasste.

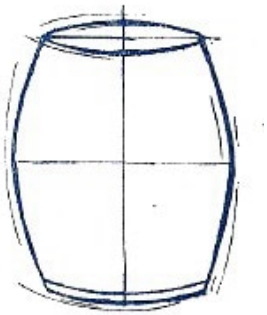


Beim Merkur wollte und wollte das nicht gelingen, so dass Kepler sein Prinzip "korrigieren" musste. Trotzdem hielt er an seiner Konstruktion fest, ja er nahm sie sehr lange Zeit übertrieben ernst. Doch sollte man Kepler dafür nicht allzu streng rügen, denn anderen erging es ähnlich.

Der Philosoph Hegel hat sich mehr als zweihundert Jahre später in Jena mit einer Arbeit habilitiert, in der er glaubte bewiesen zu haben, dass es nicht mehr als sieben Planeten geben kann, und die Astronomen tadelte, weil diese nach weiteren suchten.

Ganz anders sind die späteren Arbeiten von Kepler zu werten. Eine rein mathematische Schrift, mit der wir beginnen wollen, die "Inhaltslehre von Fässern" enthält Betrachtungen, die zwar nicht streng sind, aber zahlreiche wertvolle Anregungen zur Höheren Mathematik darstellen. Die Schrift geht über die Behandlung der Rauminhalte von Umdrehungskörpern weit hinaus.

Wie Kepler auf solche Betrachtungen verfiel, berichtet er 1614:



"Als ich im November des letzten Jahres meine Wiedervermählung feierte, zu einer Zeit, als an den Donauufern bei Linz die aus Niederösterreich herbeigeführten Weinfässer nach einer reichlichen Lese aufgestapelt und zu einem annehmbaren Preis zu kaufen waren, da war es die Pflicht des neuen Gatten und sorgenden Familienvaters, für sein Haus den nötigen Trank zu besorgen.

Als einige Fässer eingekellert waren, kam am vierten Tag der Verkäufer mit einer Messrute, mit der er alle Fässer, ohne Rücksicht auf ihre Form, ohne jede weitere Überlegung oder Rechnung, ihrem Inhalt nach bestimmte.

Die Visierrute wurde mit ihrer metallenen Spitze durch das Spundloch quer bis zu den Rändern der beiden Böden eingeführt, und als die beiden Längen gleich gefunden worden waren, ergab die Marke am Spundloch die Zahl der Eimer im Fass.
Ich wunderte mich, dass die Querlinie durch die Fasshälfte ein Maß für den Inhalt abgeben könne und bezweifelte die Richtigkeit der Methode, denn ein sehr niedriges Fass mit etwas breiteren Böden und daher sehr viel kleinerem Inhalt könnte dieselbe Visierlänge besitzen.

Es schien mir als Neuvermähltem nicht unzweckmäßig, ein neues Prinzip mathematischer Arbeiten, nämlich die Genauigkeit dieser bequemen und allgemein wichtigen Bestimmung nach geometrischen Grundsätzen zu erforschen und die etwa vorhandenen Gesetze ans Licht zu bringen."

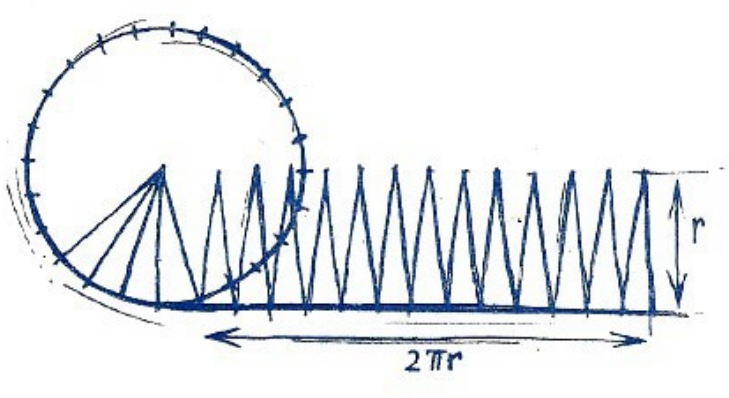
Somit diente Kepler doppelt, einmal seiner jungen Frau und zum anderen der Mathematik.

Gleich zu Anfang stellte er Überlegungen an, bei denen er sich auf Archimedes beruft, bei dem sie aber nicht stehen und auch nicht stehen können, weil sie Perlen keplerscher Erfindungsgabe sind.
Der gehässige Guldin will davon nur soviel wahrhaben, dass: "Kepler zwar den Archimedes

zitiert, wenn man aber dessen beide Bücher über die Kugel und den Zylinder durchblättert, so findet man darin nichts davon." Leider sind solche ungerechten Anwürfe kein Einzelfall.

Um einen historischen Zugang zu den genannten Überlegungen zu finden, beginnen wir (frei nach der Bibel) am besten so:
Am Anfang war auch diesmal das Wort. Es hieß indivisibile, unteilbar, und wurde vom nachmaligen Erzbischof von Canterbury, Bradwardine, im vierzehnten Jahrhundert geprägt. In seinen Schriften sucht man freilich vergeblich nach einer genauen Erklärung, trotzdem er Doctor profundus hieß. Damals war man im Verteilen solcher Ehrentitel freigebig und nannte, um nur zwei weitere Beispiele anzuführen, Thomas von Aquino Doctor universalis und Raimundus Lullus, der die Erfindung einer Denkmaschine anstrebte, ohne selber vernünftig gedacht zu haben, Doctor illuminatus.
Man muss an Indianer denken, die ihre Häuptlinge ehedem "Großer Bär" oder "Mächtiger Büffel" tauften.

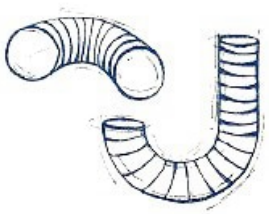
Suggestiver war es, von unendlich kleinen Größen zu reden, obschon diese einen ebenso wenig präzisen Sinn hatten wie die unteilbaren Größen des seligen Erzbischofs. Leider wird das auch heute noch übersehen, denn selbst in den besten Büchern über Mechanik werden virtuelle Verschiebungen als unendlich kleine Größen angesprochen, die es aber in der euklidischen Geometrie gar nicht geben kann.
Ihre Verwendung führte jedoch zu bedeutenden Erfolgen, mitunter freilich auch zu Misserfolgen, während die Höhere Mathematik, wie sie uns heute zu Gebote steht, nur Erfolge kennt. Allerdings wirkt sie recht nüchtern im Vergleich zum phantasievollen Voranstürmen ihrer Pioniere.



Genehmigen wir uns einige Kostproben keplerscher Erfindungskunst.
Man denke sich den Kreis in schmale Sektoren zerschnitten und diese zu einer Geraden ausgestreckt.
Es entsteht so ein kammförmiges Gebilde. Vermehrt man die Anzahl der Zacken ins Ungemessene, bis sie unendlich schmal ausfallen, dann bilden sie schließlich regelrechte Dreiecke, die über einer Strecke von der Länge des Kreisumfanges aufgebaut sind.

Die Lücken zwischen den einzelnen Zacken sind selbst zackenförmig und ergänzen den Kamm zu einem Rechteck von der Höhe des Kreishalbmessers. Die Fläche, die der Kamm selber einnimmt, beträgt die Hälfte der Rechtecksfläche.
Das führt unmittelbar auf die Formel für die Kreisfläche, zu der ja der Kamm wieder aufgerollt werden kann. Die Formel für den Umfang des Kreises wurde stillschweigend als bekannt vorausgesetzt.

Ein zweites Husarenstück bildet die Überlegung, mit der das Volumen eines Ringes, des Torus, bestimmt wird.
Der Ring werde durch senkrechte Ebenen, die strahlenförmig vom Mittelpunkt ausgehen, in dünne Scheiben geschnitten. Diese Scheiben fallen außen etwas dicker aus als innen.



Werden sie durch Anbringen von immer mehr Schnittebenen unendlich dünn, dann können sie durch überall gleich dicke Scheibchen ersetzt werden, denn was diesen außen an Dicke fehlt, haben sie innen zu viel. Sie können aufeinander getürmt werden und bilden dann einen Zylinder von einer Höhe, die dem Umfang der Mittellinie des Ringes gleich ist. Den Durchmesser des Zylinders bestimmt dagegen die Ringdicke.

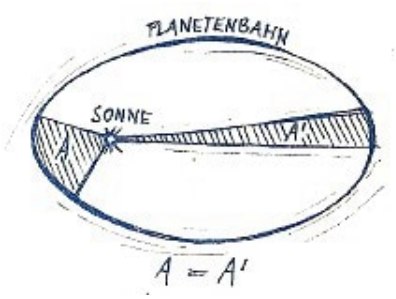
Die Bedeutung der "Inhaltslehre von Fässern" würdigte Laplace 1821 mit den Worten: "Kepler entwickelt in diesem Werk Auffassungen über das Unendliche, welche die Revolution der Geometrie gegen Ende des 17. Jahrhunderts beeinflussten; Fermat aber, den man als den eigentlichen Begründer der Differentialrechnung anzusehen hat, gründete seine schöne Methode, Extremwerte zu bestimmen, auf dieses Werk."

Ursprünglich wollte ja Kepler nur seinen Vorsatz ausführen und feststellen, wie weit die Messrute richtige Werte für die verschiedenen Fassinhalte liefern konnte. Die Lösung fand er in drei Tagen. Sie nahm sechs Seiten ein. Trotzdem fand er keinen Verleger dafür, so bekannt er auch schon damals war.

Damit blieb ihm nichts anderes übrig, als die inzwischen auf das Vielfache seines ursprünglichen Umfanges angewachsene Schrift auf eigene Kosten drucken zu lassen. Die Schrift erschien bei Plank, der eben eine Druckerei in Linz eröffnete, als dessen erster Druck.

Um die Leistungen von Kepler in der Astronomie zu würdigen, müssen wir etwas zurückgreifen. Die Zeiten, in denen er lebte, beherrschte noch die Inquisition, und so gab es Verfolgungen Andersgläubiger. Wegen einer Protestantenverfolgung musste Kepler Graz 1600 verlassen und wurde Adlatus von Tycho Brahe in Prag, dann ein Jahr später, nach dessen Tod, sein Nachfolger.

An Hand der sehr viel genaueren Beobachtungen seines Vorgängers konnte Kepler nach jahrelangen Rechnungen in der "Neuen Astronomie" 1609 die ersten zwei der nach ihm benannten Gesetze aufstellen. Das erste Gesetz betrifft die Bahnform und lautet:
Die Planeten bewegen sich auf Ellipsen, in deren einem Brennpunkt die Sonne steht.



Das zweite Gesetz, Flächensatz genannt, betrifft die Bewegung in der Bahn und lautet:
Die Planeten bewegen sich so, dass die Verbindungsgerade Sonne-Planet in gleichen Zeiten gleiche Flächen überstreicht. Eine Folge davon ist die veränderliche Geschwindigkeit, mit der sich die Planeten bewegen; in Sonnennähe schneller als in Sonnenferne.

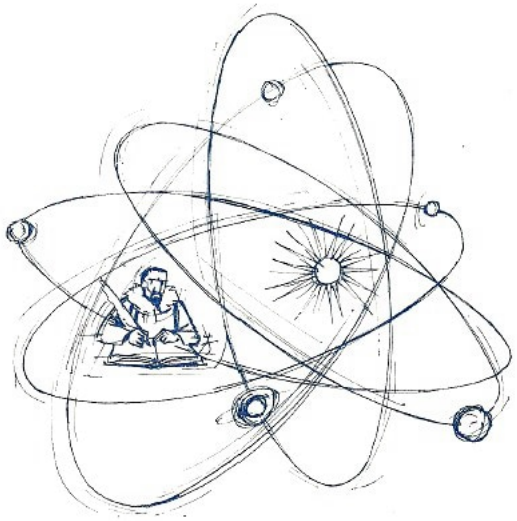
Zehn Jahre später, in den "Weltharmonien" , gelang ihm schließlich, das dritte nach ihm benannte Gesetz aufzufinden, auf der fortwährenden Suche nach dem Bau des Planetensystems, dem ja sein Erstlingswerk galt. Dieses dritte Gesetz stellt zwischen Bahngröße und Umlaufzeit eine Verbindung her und lautet:

Die Quadrate der Umlaufzeiten der Planeten verhalten sich wie die dritten Potenzen der großen Halbachsen ihrer Bahnen.

Wie neuartig den Zeitgenossen die Ellipsenbahnen der Wandelsterne vorkommen mussten, geht aus der Reaktion von Galilei hervor. In seiner Werbeschrift für Weltsysteme aus dem Jahr 1632, die ihm die Feindschaft des Papstes Urban VIII. und auf dessen Betreiben die Verurteilung durch die Inquisition einbrachte, nennt er Keplers Entdeckung Kinderei.
Dabei beruft er sich auf Aristoteles, der gelehrt hatte: "Es gibt keine ewige Bewegung außer der Kreisbewegung."

Das sagt derselbe Galilei, der durch seine Versuche die spekulative Mechanik des Aristoteles widerlegen konnte, ja die Physik endlich von der Dogmatik überhaupt erst säuberte! Nach allem vermochte sich Galilei doch noch nicht restlos freizudenken.

Wie gewaltig auch die Bedeutung der drei keplerschen Gesetze sein mochte, blieb daran doch unbefriedigend, dass sie nicht verrieten, warum sie gelten müssen. Diesen Grund hat erst Newton mit seinem Anziehungsgesetz aufgedeckt.



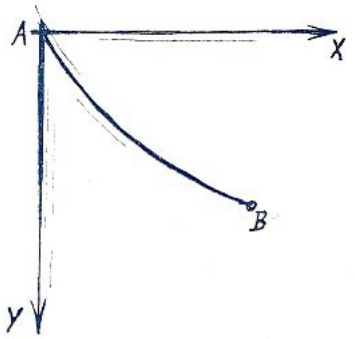
Heute können wir die drei Keplerschen Gesetze - das dritte sogar richtiggestellt, was für die Praxis allerdings nicht allzu viel bedeutet - aus der Formel für die Anziehung mit Hilfe der Höheren Mathematik in wenigen Zeilen herleiten.
Es ist noch von Interesse, auf die Vereinfachung hinzuweisen, welche die Höhere Mathematik im Laufe der Jahrhunderte erfahren hat, dank einer endgültigen Präzisierung ihrer Begriffsbildungen. Anfangs konnten sie nur führende Geister bewältigen, während sie heute bereits Anfängern gelehrt wird.

2 Verfeindete Brüder

Es ist bestimmt selten, dass in einer Familie in mehreren Generationen hintereinander bedeutende Mathematiker geboren werden. Die Schweizer Familie Bernoulli kann diesen Ruhm für sich in Anspruch nehmen. In der ersten Generation waren es die beiden Brüder Jakob und der um 13 Jahre jüngere Johann, in der zweiten Generation dessen Sohn Daniel.

Die darauffolgenden Generationen interessieren hier nicht weiter.
Bedauerlicherweise verfeindeten sich die beiden Brüder. Die Schuld daran lag bei Johann, der sich sogar mit dem eigenen Sohn überwarf, als dieser den großen Preis der Pariser Akademie vor dem Vater gewann.

Wenn auch der Charakter von Johann zu wünschen übrigließ, ein Genie war er ohne Zweifel. Wir wollen hier eine Überlegung von ihm analysieren, die er anstellte, um eine Aufgabe zu lösen, die er 1696 den Zeitgenossen vorlegte. Die Lösung von Johann ist genial, leistete aber gerade nur das Gewünschte, während Jakob eine Methode entwickelte, für welche die Aufgabe ein Spezialfall war.
Damit wurde Jakob der Begründer einer neuen Disziplin innerhalb der Höheren Mathematik, der Variationsrechnung. Ihr heutiges Gesicht verdankt sie allerdings Lagrange und Euler.



Johann fragte nach der Gestalt der Bahn, auf der ein Massenpunkt ohne Reibung unter alleiniger Einwirkung der Schwere in kürzester Zeit ankommt von einem Ort A der Ruhe zu einem Ort B, der mit A in derselben vertikalen Ebene und tiefer als A liegt.
Die Aufgabe wurde von Jakob, dann vom Marquis l'Hospital, Leibniz, Newton und sehr geistreich von Johann selbst gelöst.

Diesem stand keine fertige Methode zur Verfügung, so dass er auf heuristische Betrachtungen angewiesen war. Sein Ansatz bestand aus der Einsicht einer Analogie dieses mechanischen Problems mit einem Problem, das in der Optik bereits behandelt wurde, einer Analogie, die später von Hamilton und noch später von Schrödinger sehr vertieft wurde und letzteren zur Aufstellung der nach ihm benannten Fundamentalgleichung der Wellenmechanik führte.

Zunächst bemerkte Johann, dass die Geschwindigkeit des Massenpunktes zunimmt, wenn er auf irgendeiner Bahn tiefer gerät. Ähnliches liegt aber vor - und das war ein Genieblitz von Johann -, wenn Licht in optisch dünnere Medien gelangt. Das führte zu dem Ansatz, sich eine horizontale Schichtung zu denken, in der das Licht gemäß dem fermatschen Prinzip, das Johann ja geläufig war, seine Bahn in kürzester Zeit zurücklegt, und analog auch der Massenpunkt. Es geht jetzt darum, diesen Gedanken in Formeln umzusetzen.

Wir denken uns ein kartesisches Koordinatensystem mit dem Anfangspunkt in A und einer nach unten gerichteten y-Achse. Das Nullniveau der potentiellen Energie soll durch

B gehen.
Dann beträgt die Gesamtenergie eines Massenpunktes der Einheitsmasse in A $y_0 g$, mit y_0 die y-Koordinate von B und mit g die Schwerebeschleunigung bezeichnet. Die Gesamtenergie setzt sich aus potentieller und kinetischer Energie zusammen und bleibt konstant $= y_0 g$. Im Punkt (x, y) ist daher

$$(y_0 - y)g + \frac{1}{2}v^2 = y_0 g$$

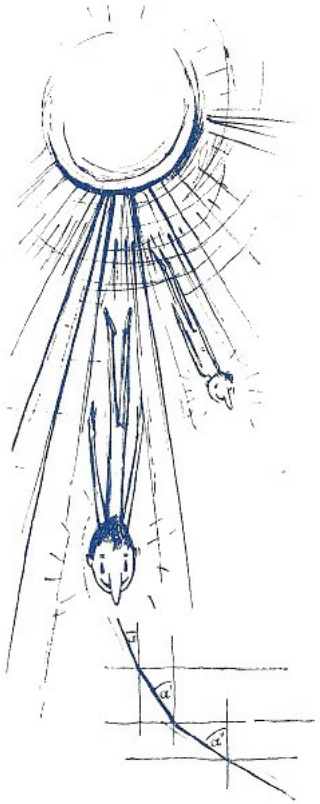
wobei v den Wert der Geschwindigkeit bezeichnet. Daraus folgt

$$v = \sqrt{2yg}$$

Die Geschwindigkeit hängt folglich nur von der Tiefe y ab, in der sich der Massenpunkt augenblicklich befindet, keineswegs aber von der Gestalt der Bahn, auf der er sich bewegt. Diese Feststellung bildet die Berechtigung zur Annahme,einer horizontalen Schichtung.

Greift man jetzt auf die Analogie mit dem Licht zurück, dann wird man homogene dünne Horizontalschichten annehmen, um das Gesetz von Snellius anwenden zu können. Bezeichnen wir die Lichtgeschwindigkeit in den aufeinanderfolgenden Schichten mit v, v', v'', ... und die Winkel, die der Lichtstrahl mit der Vertikalen bildet mit α, α', α'', ... dann gilt nach Snellius

$$\frac{\sin\alpha}{v} = \frac{\sin\alpha'}{v'} = \frac{\sin\alpha''}{v''} = ... = \text{ Konstante } k$$



Geht man nun zu "unendlich dünnen" Schichten über, das heißt aber zu einem Medium mit kontinuierlich veränderlichem v, dann folgt

$$\frac{\sin\alpha}{v} = k$$

in welcher Tiefe man sich auch befindet. Dieser gewagte Schritt lässt sich mit unseren heutigen Mitteln durchaus rechtfertigen, Johann aber dachte gar nicht daran, denn zu seiner Zeit bediente man sich fröhlich-frei der unendlich kleinen Größen, um Grenzübergänge vorzunehmen.
Die Tangente an die Bahnkurve schließt mit der Vertikalen den Winkel α, mit der Horizontalen dagegen den Winkel $\beta = 90° - \alpha$ ein, so dass $\cos\beta = \sin\alpha$ gilt. Es ist außerdem

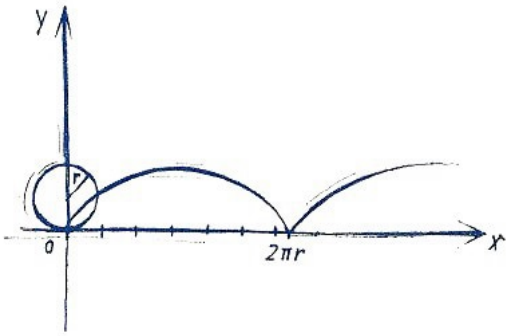
$$\cos^2\beta = \frac{1}{1+\tan^2\beta} \qquad \text{also} \qquad \sin^2\alpha = \frac{1}{1+\tan^2\beta}$$

Setzt man diesen Wert und außerdem für v^2 den Wert $2gy$ in die Gleichung $\sin^2\alpha = k^2v^2$ ein, dann folgt

$$y(1 + y'^2 = C$$

mit C die Konstante $\frac{1}{2gk^2}$ bezeichnet. Das ist aber die Differentialgleichung der Zykloide, die von irgendeinem Randpunkt eines Rades beim Abrollen beschrieben wird und die

Johann bekannt war. Die gesuchte Kurve, Brachistochrone genannt, ist also eine Zykloide.

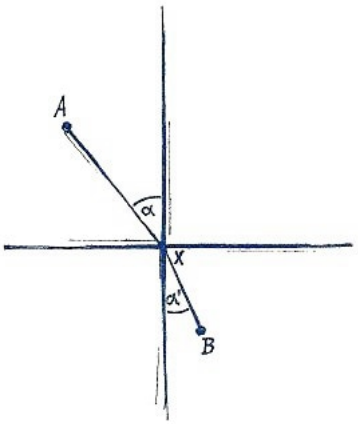


Wir sprachen vom Brechungsgesetz des Snellius, mit bürgerlichem Namen Willibrord Snell van Royen und dem Prinzip des Fermat. Das erste stammt aus dem Jahr 1618, das zweite aus dem Jahr 1657, und sie sind gleichwertig. Fermat knüpfte an ein Minimumprinzip des Heron an, fügte aber als neuen Gedanken hinzu, den Lichtweg durch die Zeit zu messen, in der er zurückgelegt wird.

Heron hatte das nicht nötig, weil er Spiegel untersuchte, bei denen ja der Lichtstrahl im Medium verbleibt. Das Prinzip von Fermat kann wie folgt formuliert werden:
Die Bahn eines Lichtstrahls zwischen zwei Punkten verläuft, dass das Licht auf seinem Weg die kürzeste Zeit benötigt.

Gehen wir von diesem Prinzip aus und betrachten den Fall zweier homogener Medien, die eine gemeinsame Ebene besitzen, und in denen sich das Licht mit der Geschwindigkeit v bzw. v' ausbreitet.

Vom Punkt A im ersten Medium soll ein Lichtstrahl durch diese Ebene hindurch zum Punkt B im zweiten Medium gelangen. Zunächst ist es klar, dass der Lichtweg in beiden Medien gerade verlaufen muss, weil die Gerade stets den kürzesten und darum schnellsten Weg darstellt.
Wir nehmen an, dass A und B in einer zur Trennebene der beiden Medien senkrechten Ebene liegen. Die Schnittgerade der beiden Ebenen sei g.



Ist X der Durchstoßungspunkt des Lichtweges mit der Trennebene, dann muss nach dem fermatschen Prinzip

$$\frac{AX}{v} + \frac{XB}{v'}$$

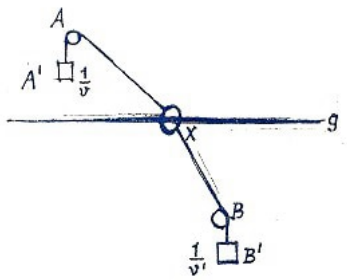
Minimum sein, weil Weg durch Geschwindigkeit das Maß für die Zeit ist, die das Licht braucht. Die Differentialrechnung beantwortet heute rein routinemäßig diese Forderung dahingehend, dass

$$\frac{\sin\alpha}{v} = \frac{\sin\alpha'}{v'}$$

ausfallen muss, wobei die Winkel, die AX bzw. XB mit der Normalen zu g bilden, mit α bzw. α' bezeichnet werden. Das aber ist das Brechungsgesetz des Snellius. Es sei noch bemerkt, dass vermutlich dieser Anlass Fermat zu seinen Entwicklungen in der Differentialrechnung führte.

Erfreulicherweise gelingt es, den Satz von Snellius aus dem fermatschen Prinzip auch elementar zu gewinnen, wenn man von einer mechanischen Interpretation ausgeht. Zum

Glück merkte das Fermat nicht, denn sonst hätte er sich um die Differentialrechnung nicht so bemüht!

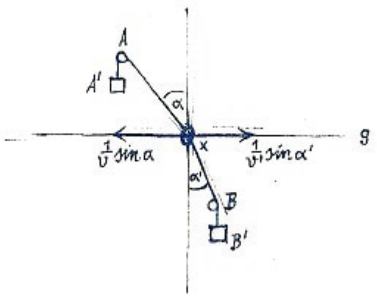


Wir denken uns die Gerade von vorhin als dünnen Stab, an dem ein Ring reibungslos entlanggleiten kann. An diesem Ring X seien zwei Schnüre befestigt, deren Länge sich nicht ändert, und die über je eine Rolle reibungslos gleiten, die in A und in B befestigt ist. Die beiden Schnüre seien am freien Ende in A' bzw. B' mit Gewichten von der Größe $\frac{1}{v}$ bzw. $\frac{1}{v'}$ belastet.

Dieses mechanische System befindet sich dann im Gleichgewicht, wenn die beiden Gewichte möglichst tief sinken, mit anderen Worten der Verlust an potentieller Energie

$$AA' \cdot \frac{1}{v} + BB' \cdot \frac{1}{v'}$$

maximal ausfällt.



Da die beiden Schnüre unausdehnbar sind, bleibt

$$(A'A + AX) \cdot \frac{1}{v} + (BB' + BX) \cdot \frac{1}{v'}$$

konstant, daher muss in Gleichgewichtslage

$$AX \cdot \frac{1}{v} + BX \cdot \frac{1}{v'}$$

minimal ausfallen. Formal liegt also das Prinzip von Fermat vor.
Infolge der mechanischen Deutung gewinnt man aber sofort als Ergebnis die Beziehung des Snellius durch die Bemerkung, dass im Gleichgewicht die Horizontalkomponenten der beiden Kräfte, die am Ring in X zerren, entgegengesetzt gleich sein müssen, formelmäßig

$$\frac{1}{v} \cdot \sin\alpha = \frac{1}{v'} \cdot \sin\alpha'$$

Es ist kein Einzelfall, dass neue Entdeckungen wie vorhin von Fermat, gemacht wurden, um ein Problem zu meistern, das auch ohne diese Entdeckungen lösbar war.

So sollte Archimedes einst im Auftrage seines königlichen Vetters feststellen, ob dessen Krone wirklich aus schierem Gold angefertigt wurde.
Er löste die Aufgabe durch die Entdeckung des Auftriebs: In Flüssigkeit getaucht, verliert jeder Körper soviel an Gewicht, wie die von ihm verdrängte Flüssigkeit wiegt.

Die Aufgabe lässt sich aber ohne Kenntnis dieses Gesetzes lösen. Man legt die Krone zunächst in einen vorher bis zum Rande mit Wasser gefüllten Zylinder. Nach Entfernen der Krone bestimmt man dann aus der Höhe des übriggebliebenen Wassers das Volumen der Krone. Hinterher hat man nur das Gewicht eines Klumpens aus reinem Gold vom gleichen Volumen zu wägen, um festzustellen, ob sich nicht unter einer Goldschicht etwa leichteres Blei befindet.

3 Lob der Trägheit

Sollte der Leser durch die Überschrift an das Werk "Lob der Torheit" des Erasmus von Rotterdam erinnert werden, so darf er nicht glauben, dass wir gegen träge Menschen losziehen wollen, nein, weder gegen faule, noch gegen geistig lahme.

Es handelt sich um die Trägheit, einen grundlegenden Begriff der Mathematischen Physik, oder vielmehr darum, dass man jahrhundertelang nicht in der Lage war, diesen Begriff genau zu umreißen.
Das war schon zu einem Skandal der Mathematischen Physik ausgeartet, der auch heute nicht dadurch behoben wird, dass ihn Autoren unserer Tage einfach übergehen.

Zum richtigen Verständnis des Sachverhaltes gehört allerdings ein Eingehen auf die Grundlegung der Mechanik. Deren einwandfreien Aufbau verdankt man also letztlich der Trägheit. Die Anfänge der Mechanik sind so alt wie der Mensch selber, dessen primitive Handwerkszeuge Vorläufer der modernen Maschinen waren. Die ersten Erfahrungen wurden mehr unbewusst, wohl zufällig gesammelt. Viel später erst dürfte ein Stadium in der Entwicklung eingetreten sein, das nach ordnenden Begriffen heischte, um dem Wust von Erfahrungen nicht zu erliegen. Mit solchen immer umfassenderen Begriffen beginnt dann die Ära der Wissenschaft. Sie versucht, die Welt in Gedanken nachzubilden, um Vorhersagen zu ermöglichen und aktiv in die Veränderungen der Gesellschaft einzugreifen. Damit wird Wissenschaft zum einsatzbereiten Wissen.

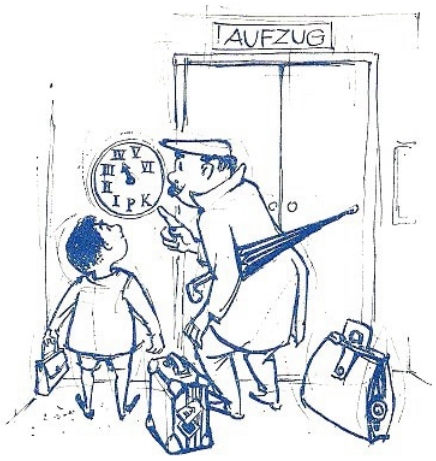
Die Wissenschaft kann nicht umhin, Begriffe einzuführen, denen keine Anschaulichkeit zukommt. Das wird vielmehr erst durch ein Zusammenspiel von solchen mit anderen, anschaulichen Begriffen herbeigeführt. Das liegt an der Struktur unseres Geistes, der keine unmittelbare Nachbildung des Naturgeschehens kennt. Infolgedessen verlegen wir das Geschehen in einen Raum und versehen es mit einer Zeit.

Bei der Zeit handelt es sich um einen grundlegenden Begriff, ohne den kein Bewegungsablauf verfolgt werden könnte. Seinem Wesen nach ist jedoch der Zeitbegriff mathematischer Art. Erst diese Feststellung erklärt die Sonderrolle, die die Zeit gegenüber anderen, physikalischen Größen spielt. Das Nichtbeachten davon führte zu einer Schwierigkeit in der Quantenmechanik, die einer ihrer bedeutendsten Förderer, von Neumann, als ihre Hauptschwäche bezeichnete.

Um den Zeitbegriff einzuführen, hat man von einem Fundamentalbegriff moderner Mathematik auszugehen. Es handelt sich um die umkehrbar eindeutige oder kürzer eineindeutige Abbildung. Liegen zwei Mengen vor, deren Elemente so gepaart werden können, dass kein Element leer ausgeht, dann liegt der Fall einer eineindeutigen Abbildung der beiden Mengen aufeinander vor.

Einmal im Besitz dieses reichlich abstrakten Begriffs von eineindeutigen Abbildungen, kann man diese auf den konkreten Fall anwenden, bei dem die eine Menge aus den Punk-

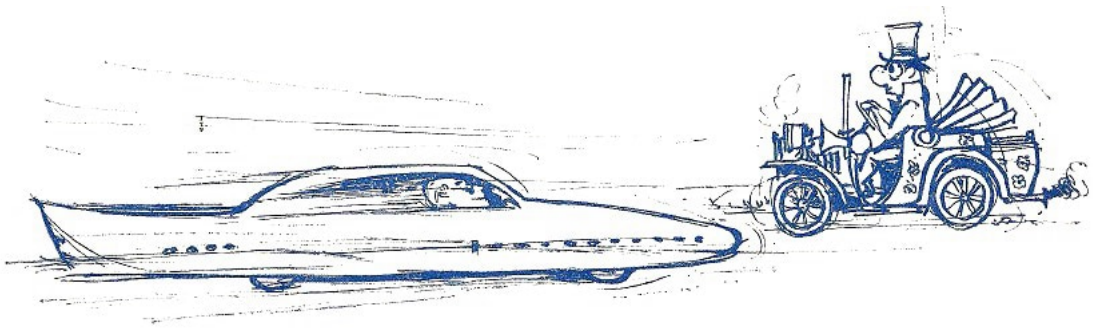
ten des Kreisbogens einer Uhr besteht, die andere aber aus den Punkten einer Strecke, die ein Fahrstuhl zurücklegt, der vom Erdgeschoss nach oben fährt.



Jedem Punkt des Kreisbogens auf der Uhr wird so ein Punkt auf der Strecke des Fahrstuhls eineindeutig zugeordnet. Die Zeigerstellung der Uhr auf dem Kreisbogen bestimmt auf diese Weise den augenblicklichen Ort des Fahrstuhls. Damit ist der Inhalt des Zeitbegriffs ausgeschöpft.
Die für die Mechanik entscheidende Wendung tritt nun bei der Frage auf, wieweit die geschilderte Abbildung zwischen Uhrzeigerbahn und Fahrstuhlstrecke vom Beobachter abhängt.
Die einfachste Annahme, mit der aber die klassische Mechanik auskommt, lautet: Die Abbildung ist unabhängig vom Beobachter.

Mit anderen Worten besagt das, zwei Beobachter nehmen immer dieselbe Abbildung wahr. Die Relativitätstheorie, schon die spezielle Relativitätstheorie, führt dagegen zur Auffassung, dass gegeneinander bewegte Beobachter verschiedene Abbildungen erleben. Das äußert sich darin, dass sie die Dauer eines Geschehens verschieden lang angeben, während zwei Beobachter in der klassischen Mechanik für die Dauer stets denselben Wert ermitteln werden. In der klassischen Mechanik ist die Zeit eben absolut, wie sich Newton ausdrückte.

Eine Schlüsselstellung in der Mechanik nimmt der Begriff der Kraft ein. Durch Kräfte werden Bewegungsverläufe bestimmt. Diese Kräfte können aber erst durch Rückschlüsse ermittelt werden.
Den ersten Ansatz dafür bildete die Annahme der klassischen Mechanik, dass zwischen Kraft und Beschleunigung Proportionalität besteht. Aus mathematischen Gründen folgt daraus sofort, dass zwei Beobachter, die sich gleichförmig gegeneinander bewegen, dieselben Kräfte wahrnehmen. Im Gegensatz dazu stellen zwei Beobachter, die sich beschleunigt gegeneinander bewegen, verschiedene Kräfte fest.



Wenn es daher gelingt, einen Beobachter - und damit alle zu ihm gleichförmig bewegten Beobachter - einwandfrei auszuzeichnen, dann sind für alle anderen Beobachter zusätzliche Kräfte anzusetzen. Man nennt sie Scheinkräfte, während die ausgezeichneten Beobachter mit nur eingeprägten oder wirklichen Kräften zu tun haben.

Die Schar dieser gleichförmig gegeneinander bewegten Beobachter, die nur eingeprägte Kräfte erleben, repräsentiert die Inertialsysteme, und das Fundamentalproblem verlangt, diese Inertialsysteme aus allen überhaupt möglichen Systemen auszusieben.

Ehe wir zur Lösung schreiten, ist es aufschlussreich, missglückte Lösungsversuche Revue passieren zu lassen. Um deren verschiedene Formulierungen zu verstehen, müssen wir etwas zurückgreifen.

Gleichförmig gegeneinander bewegte Beobachter erleben dieselben Kräfte. Wenn folglich der zu mir gleichförmig bewegte Beobachter dieselben Kräfte wahrnimmt wie ich, dann bedeutet das, dass für seine Eigenbewegung keine Kraft verantwortlich gemacht werden darf, da diese von ihm aus betrachtet, unwirksam bliebe.
Das ist der präzisierte Inhalt des Trägheitsgesetzes von Galilei. Insbesondere besagt es, dass die Bewegung in Abwesenheit von Kräften gleichförmig verläuft. Man erkennt aber auch, dass es nicht herangezogen werden kann, um die Inertialsysteme auszusieben. Denn eingeprägte Kräfte, die durch Scheinkräfte gerade aufgehoben werden, könnten Inertialsysteme vortäuschen, solange man die beiden Kraftarten nicht auseinanderhalten kann.

In die philosophischen Spekulationen seiner Epoche verstrickt, glaubte Newton dadurch ein bestimmtes Inertialsystem nachweisen zu können, dass er die Fiktion eines absoluten Raumes in seine Mechanik einführte:

”Der absolute Raum bleibt vermöge seiner Natur und ohne Beziehung auf einen äußeren Gegenstand stets gleich und unbeweglich” formulierte er und weiter: ”Der Ort ist ein Teil des Raumes, welchen ein Körper einnimmt, und nach Verhältnis des Raumes absolut oder relativ” , schließlich: ”Die absolute Bewegung ist die Übertragung des Körpers von einem absoluten Ort nach einem anderen absoluten Ort.”

Nachzulesen in ”Mathematische Prinzipien der Naturlehre” in der Übersetzung von Wolfers aus dem Jahr 1872.

Der absolute Raum ist nach Newton als Inertialsystem anzusprechen. Doch besteht prinzipiell keine Möglichkeit, den absoluten Raum durch physikalische Messungen und aus diesen gezogenen Schlüssen zu ermitteln. Eine Fiktion aber, die sozusagen neben der Wirklichkeit, ohne ergründbaren Zusammenhang mit ihr, herläuft, ist als sinnwidrig abzulehnen.
Im Gegensatz hierzu wird man Newton zustimmen können, wenn man seine nachfolgenden Worte so auslegt, wie das unsere Bemerkung im Anschluss an das präzisierte Trägheitsgesetz verlangt.

”Die wahren Bewegungen der einzelnen Körper zu erkennen und von den scheinbaren scharf zu unterscheiden, ist übrigens sehr schwer, weil die. Teile jenes unbeweglichen Raumes, in denen sich die Körper wahrhaft bewegen, nicht sinnlich erkannt werden können. Die Sache ist jedoch nicht gänzlich hoffnungslos. Es ergeben sich nämlich die erforderlichen Hilfsmittel, teils aus den scheinbaren Bewegungen, welche die Unterschiede der wahren sind, teils aus den Kräften, welche den wahren Bewegungen als wirkende Ursachen zugrunde liegen.”

Man empfand die Forderung eines absoluten Raumes zwar unbefriedigend, konnte aber keinen rechten Verbesserungsvorschlag machen. Lange versuchte folgende Erklärung:

”Von einem bestimmten im Bezugssystem festen Raumpunkt aus schleudern wir nach drei nicht in einer Ebene liegenden Richtungen drei freie Massenpunkte. Sind deren Bahnen gerade Linien, so ist das Bezugssystem ein Inertialsystem.”
Gelegentlich redet man an Stelle von freien Massen punkten von solchen, die sich selbst überlassen sind. Wann aber sind Massenpunkte als frei anzusehen? Hinter dieser Frage versteckt sich diesmal die ganze Schwierigkeit. Sobald man eingeprägte von Scheinkräften unterscheiden kann und die ersten kennt, ist die Frage beantwortet, aber auch erst dann.

Neumann postulierte einen Alphakörper und hielt in einer weiteren Veröffentlichung seinen Standpunkt mit dem von Lange übereinstimmend. Auch erklärte er sein Vorgehen für rein hypothetisch, worin sein einziger Fortschritt über Lange hinaus besteht. Damit erübrigt sich eine weitere Kritik.

Der 1951 in München verunglückte Sommerfeld dagegen gibt offen zu: ”Praktisch verlassen wir uns hierbei auf die Astronomen, die uns im Fixsternhimmel hinreichend feste Achsenrichtungen und im mittleren Sonnentag eine hinreichend konstante Zeiteinheit liefern. Theoretisch sind wir aber leider zu einer unerfreulichen Tautologie gezwungen:
Dasjenige Bezugssystem ist das richtige, in dem für einen hinreichend kräftefreien Körper das galileische Trägheitsgesetz hinreichend genau gilt. Das Trägheitsgesetz wird dadurch zur formalen Identität degradiert; als positiver nicht formaler Inhalt desselben bleibt nur die Behauptung bestehen: Es gibt Bezugssysteme der geforderten Eigenschaft; ein solches wird, nach all unserer Erfahrung, approximiert durch astronomische Orts- und Zeitbestimmung.”

Eine aussichtslose Situation? Der Ausweg ist sofort da, wenn man sich entschließt, die Gravitation in die Grundlegung der Mechanik einzubeziehen. Sie ist nach der klassischen Mechanik eine Kraft, die untrennbar mit der Materie verknüpft ist. Daher können keine stichhaltigen Gründe geltend gemacht werden, um diesen entscheidenden Schritt zu verbieten.
Zunächst greife man auf den Ansatz von Newton zurück, der die Anziehung den Massen direkt und dem Quadrat ihrer gegenseitigen Entfernung indirekt proportional setzt. Meine Definition für Inertialsysteme lautet nun:

In solchen findet man für die Gravitation genau diesen Wert. Nach unseren früheren Ausführungen wird man dann in allen übrigen Systemen abweichende Werte für die Anziehung messen. Da weiter in genügend großer Entfernung von übrigen Körpern die Formel von Newton von jeher als hinreichend genau gültig angenommen wurde, wollen wir dort als einzige eingeprägte Kraft die Gravitation postulieren.

Unsere Einführung der Inertialsysteme legt zunächst ein erstes, rein statisches Gedankenexperiment nahe, um zu entscheiden, ob man sich in einem Inertialsystem befindet.
Man verbinde jede von zwei gleichen Massen, etwa zwei gleiche Kugeln aus homogenem Material, mit je einer von zwei Federn, die aus gleichem Material, gleich lang und drehbar so befestigt sind, dass sich die beiden Massen gerade gegenüberstehen. Durch die Anziehung werden die Federn etwas gespannt.

In Inertialsystemen bilden sie an jedem Ort zwei Strecken einer Geraden und sind gleich lang, sonst nicht. Denn eine Beschleunigung in Richtung der Verbindungsgeraden der beiden Massen würde die vordere Feder verlängern, die hintere verkürzen, eine dazu senkrechte Beschleunigung aber die Federn mit der Verbindungsgeraden einen Winkel bilden lassen. Selbstredend ist die Versuchsanordnung isoliert zu denken.

Selbst eine einzelne Feder würde die Inertialsysteme anzeigen, wenn man sie an einem Ende drehbar befestigt und ihre Länge in verschiedenen Lagen misst.
Wollte man jedoch ein solches Experiment als Erklärung für Inertialsysteme heranziehen, so wäre dagegen einzuwenden, dass sie nicht mehr mit Fundamentalbegriffen auskommt, sich vielmehr auf die Existenz von Federn berufen muss.

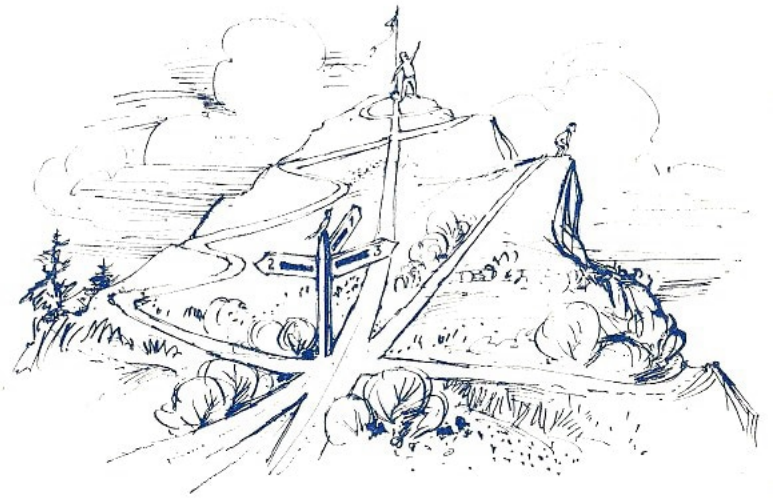
Die Forderung, das Gedankenexperiment führe an jedem Ort zu gleichem Ergebnis, ist wesentlich. Denn in einem abstürzenden Fahrstuhl würden unsere beiden Massen ein Inertialsystem anzeigen. Zu genau demselben Ergebnis würde aber jedes andere mechanische Experiment führen, und das mit Recht, denn innerhalb des Fahrstuhls herrschen dieselben Verhältnisse wie in Inertialsystemen. Doch sind letztere in jeder Richtung als unbegrenzt zu denken, im Gegensatz zum Fahrstuhl.
Das führt dazu, die Massen zu berücksichtigen, die das homogene Schwerefeld erzeugen, in dem der Fahrstuhl frei fällt. Man erkennt so, dass der Fahrstuhl doch kein Inertialsystem ist.

Ein weiteres Problem bildet es, Körper zu behandeln, die in ihrer Bewegungsfreiheit eingeschränkt sind. Die ihnen auferlegten Bedingungen, auch Bindungen genannt, sind durch Kräfte zu ersetzen, die Zwangskräfte oder Führungskräfte heißen. Lange Zeit hindurch berücksichtigte man ausschließlich nur Bindungen in der Art von glatten Flächen, wie die schiefe Ebene. Hertz wies dann vor ungefähr 70 Jahren nach, dass man damit nicht auskommt.

Johann Bernoulli machte den Anfang, indem er den Begriff von virtuellen Verschiebungen prägte, die mit den Bindungen verträglich sind. Mit Hilfe der virtuellen Verschiebungen gelang es dann d'Alembert, sein Prinzip aufzustellen, das erste allgemeine Prinzip der Mechanik, wenn man darunter die Zusammenfassung ihres Gehaltes in eine geeignete Aussage versteht.
Handelt es sich dabei um eine neue Annahme?

Ich konnte bereits vor Jahren zeigen, dass dies nicht der Fall ist, wenn auch vorher das Gegenteil davon behauptet wurde. Man diskutierte darüber viel, ohne zur Klarheit zu gelangen. Jacobi beklagt das mit den Worten:

"Die im obigen enthaltene Ausdehnung ist, wie sich von selbst versteht, nicht bewiesen, sondern nur als Behauptung historisch ausgesprochen. Dies ausdrücklich zu sagen, scheint nötig zu sein, denn obgleich Laplace diese Ausdehnung in der "Mécanique céleste" ebensowenig bewiesen hat, als es hier geschehen ist, sondern sie nur historisch behauptete, so hat man dies dennoch für einen Beweis gehalten.
Poinsot hat gegen diese Meinung eine eigene Abhandlung geschrieben und sagt darin sehr richtig, dass sich die Mathematiker häufig durch den sehr langen Weg täuschen lassen, den sie zurückgelegt haben, zuweilen aber auch durch den sehr kurzen.



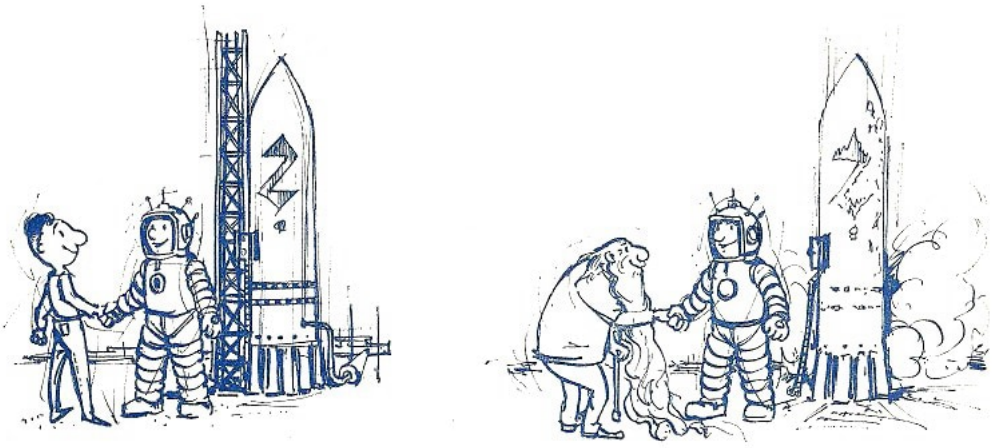
Durch den langen Weg lassen sie sich täuschen, wenn sie durch sehr weite Rechnungen endlich zu einer Identität kommen, dieselbe aber für einen Satz halten.
Ein Beispiel von dem Entgegengesetzten gibt unser Fall. Diese Ausdehnung zu beweisen, ist keineswegs unsere Absicht, wir wollen sie vielmehr als ein Prinzip ansehen, welches zu beweisen nicht nötig ist. Dies ist die Ansicht vieler Mathematiker, namentlich von Gauß."

Wie man aber beliebige Bindungen zu berücksichtigen hat, blieb bis vor 30 Jahren eine offene Frage. Vor allem die mathematische Behandlung ließ zu wünschen übrig.
Im Jahre 1941 zeigte ich dann, dass sich dafür ein von Hölder gegen Ende des vorigen Jahrhunderts entworfener Kalkül eignet, wenn man ihn vorher einwandfrei begründet. Diese Begründung hatte die Berichtigung von grundlegenden Formeln zur Folge. Das Endergebnis aber lautete, dass sich die klassische Mechanik als Ergebnis einer schrittweisen Erweiterung des zweiten Bewegungsgesetzes von Newton ergab, das uns bereits am Anfang als Proportionalität zwischen Beschleunigung und Kraft begegnete.

Die klassische Mechanik hielt daran fest, dass dieser Proportionalitätsfaktor, die Masse, von der Bewegung unbeeinflusst, also konstant bleibt. Versuche von Kaufmann um die Jahrhundertwende an schnell bewegten Elektronen führten aber dazu, dass man diese Konstanz aufgab. Damit erkaufte man eine Verfeinerung der klassischen Mechanik zur speziellen Relativitätstheorie.
Merkbar unterscheidet sich diese erst bei hohen Geschwindigkeiten von der newtonschen Mechanik, doch prinzipiell bedeutet sie einen entscheidenden Fortschritt. Denn sie führt zur Äquivalenz von Energie und Materie, die sich in der Kernphysik inzwischen bewahrheitete und im Uranmeiler die Quelle einer künftigen Energiewirtschaft zu werden verspricht.

Eine paradoxe Aussage der speziellen Relativitätstheorie betrifft das Messen von Länge und Dauer, die von einem bewegten Beobachter aus betrachtet, eine Kontraktion bzw. Dilatation zu erleiden scheinen.
Da Länge und Dauer durch Messen bestimmt werden, unterliegen sie neuartigen Formeln, den Lorentztransformationen, die den Übergang aus dem einen in das andere Inertialsystem regeln und so zur genannten Paradoxie führen.

Der folgende Trugschluss, der von dem Franzosen Langevin aus dem Jahre 1911 stammt, bildete zwar seither oftmals den Höhepunkt verschiedener utopischer Erzählungen, ist aber wie eben gesagt, leider ein Trugschluss.



In meinem Buch "Grundzüge der Mechanik" , dessen zweite Auflage 1961 erschien, und das in allen Einzelheiten mathematisch einwandfreie Darstellungen bringt, was besonders betont werden muss, wurde neben vielem anderen auch dargelegt, weshalb sich Langevin irrte.
Er meinte, wenn von Zwillingsbrüdern einer im Jünglingsalter in einem Raumschiff mit nahezu Lichtgeschwindigkeit erst von der Erde wegbraust, später aber wieder zurückkehrt, dann der auf der Erde Zurückgebliebene bereits ein Greis geworden sein sollte, während der Weltraumfahrer jugendlich blieb.

Zunächst ist für Abfahrt, Umkehr und Landung die spezielle Relativitätstheorie nicht zuständig. Den übrigen Verlauf der Reise kann man aber dann auch so ansehen, dass der eine der Zwillinge als Reisender mit der Erde wegbrauste, folglich mit gleichem Recht wie vorhin jung bleiben müsste, im Gegensatz zum Raumfahrer, der diesmal der Zurückgebliebene ist!
Wie sollte auch ein Ausflug in den Raum als Reise in die fernste Zukunft enden?

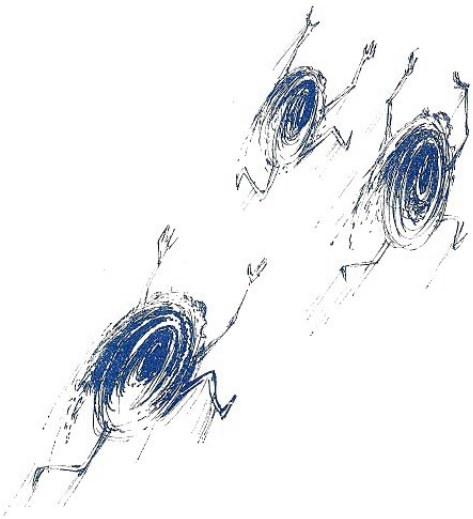
Blieb die spezielle Relativitätstheorie noch auf Inertialsysteme beschränkt, so befreite sie davon Einstein 1916, indem er in seiner allgemeinen Relativitätstheorie beliebige Koordinatensysteme zuließ.
Das gelang ihm dadurch, dass er die Gravitation für die Geometrie verantwortlich machte, die unsere Welt beherrscht und von der Schulgeometrie wesentlich abweicht. Eine solche mögliche, freilich unter bedenklichen Vereinfachungen eingeführte Geometrie führt auf das sich ausdehnende Universum, das endlich, aber selbstredend ohne Grenzen ist. Damit lässt sich die beobachtete Flucht der Spiralnebel erklären.

So erfolgreich auch die Relativitätstheorie war, sie ist auf den Makrokosmos beschränkt. Für den Mikrokosmos gelten andere Bewegungsgesetze, wie die Gleichung von Schrödinger, ja es sind völlig neue Begriffe nötig.
Das sind gewisse Operatoren und ihr Spektrum, die erst in jüngster Zeit von von Neumann und anderen Mathematikern untersucht wurden und noch werden. Doch scheint es, dass die Kernphysik, deren Ergründung heute das Bemühen der Forscher gilt, zu ihrer endgültigen Formulierung noch neuer Begriffe bedarf.

Es verlohnt noch, auf folgende, nur psychologisch zu verstehende Versuche hinzuweisen. Sie erklären sich aus der sozusagen Erstgeburtenrolle der Mechanik, die ganz naturgemäß der erste Zweig der Physik wurde. Die Vertrautheit mit der Bewegungslehre legte es daher nahe, neue Einsichten darauf zurückzuführen.

So versuchte vor einem Jahrhundert der Schotte Maxwell die nach ihm benannten Gleichungen für das elektromagnetische Geschehen durch mechanische Modelle zu erläutern. Warum ihm und später dem Wiener Boltzmann das nicht gelang und auch nicht gelingen konnte, erkennt man sofort, wenn man die Konstanz der Lichtgeschwindigkeit im feldlosen Vakuum berücksichtigt.

Erfahrungsgemäß ist sie vom Beobachter unabhängig. Ein mechanisches Modell würde dann verlangen, dass für den Wechsel $t' = t$ und $x' = x + vt$ der Inertialsysteme die Gleichung $x^2 = ct^2$ für die Lichtausbreitung sich in den gestrichenen Größen reproduziert, was aber keineswegs zutrifft.
Eine Auswirkung solcher misslungenen Erklärungsversuche war der Ausspruch eines Lehrers in meiner Jugend, man könne das Wesen der Elektrizität nicht begreifen.



4 Geisterreich

Unser Lebensraum ist dreidimensional. Verständlich also, wenn man sich zunächst damit begnügte, die Geometrie eines dreidimensionalen, Raumes zu entwickeln. Doch bereits im Altertum begann man, sich über eine vierte Dimension Gedanken zu machen. Der erste, der darauf verfiel, scheint Plato gewesen zu sein. In dessen "Staat" belehrt Sokrates, vielleicht auch darin ein Weiser, dass er den Schierlingsbecher seiner Xanthippe vorzog, Glaukon über die Eindrücke von Menschen, die, von jung auf gefesselt, mit dem Gesicht zur inneren Wand in einer Höhle leben.

Außerhalb der Höhle verläuft ein Weg. Ein Feuer werte die Schatten der Vorübergehenden durch den Eingang der Höhle auf die gegenüberliegende Wand, vor der die Gefangenen sitzen.

Einer von ihnen möge sich befreien. Erblickt er die Außenwelt zum erstenmal, so wird er sich in dieser zunächst kaum zurechtfinden. Denn statt zweidimensionaler Schatten steht er plötzlich dreidimensionalen Lebewesen gegenüber. Erst allmählich wird er sich in die neue Welt einleben.

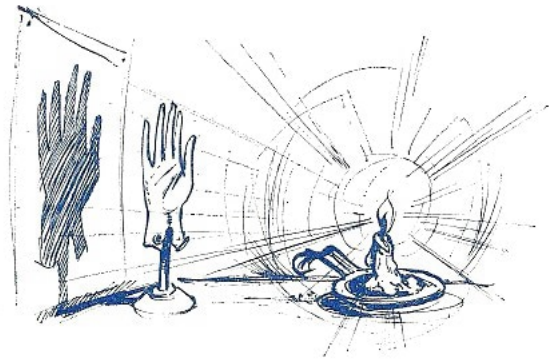
Sollte ihn dann ein grausames Missgeschick wieder an seinen alten Ort zurückverschlagen, wie könnte er dort seinen Leidensgenossen klarmachen, dass zu den vorüberziehenden Schatten Menschen gehören, die in der Sonne leben? Würde man ihn in der Höhle nicht für verrückt halten?
Sokrates schließt: "Dieses Gleichnis, mein lieber Glaukon, musst du seinem vollen Umfang nach mit den vorhergehenden Erörterungen in Verbindung bringen;

die durch das Gesicht uns erscheinende Raumwelt setze der Wohnstätte der Gefesselten gleich, den Lichtschein des Feuers aber in ihr der Kraft der Sonne; den Aufstieg nach oben und die Betrachtung der oberen Welt musst du der Erhebung der Seele in das Reich des nur Denkbaren vergleichen, wenn du eine richtige Vorstellung von meiner Meinung bekommen Willst, da du sie ja zu hören begehrt hast."

Der Astrophysiker Zöllner in Leipzig ergab sich der vierten Dimension dagegen bedin-

gungslos und glaubte fest an deren Vorhandensein. Er knüpfte daran in den siebziger Jahren des vorigen Jahrhunderts paradoxe Bemerkungen.
Ein rechter und ein linker Handschuh könnten nacheinander die "Schatten" desselben vierdimensionalen Körpers sein. Um das einzusehen, streife man eine Dimension ab. Denke ich mir den Scherenschnitt des rechten Handschuhs aus der Ebene gehoben und hinterher mit der Rückseite nach oben wieder auf die Ebene gelegt, dann verwandelte er sich in den Scherenschnitt eines linken Handschuhs.

Noch verblüffender wäre es, den Schatten eines vierdimensionalen Körpers in mehreren Dunkelkammern gleichzeitig aufzufangen. Dann würde "einer Vielheit einzelner Erscheinungen, die zu ein und derselben Gattung gehören, trotz ihrer individuellen Verschiedenheit nur ein einziges Objekt entsprechen."
Dieses Objekt wäre eine Art allgegenwärtiger Gottheit. Gewiss, wir könnten das Experiment auch in unserer Erfahrungswelt anstellen, doch gelten dort die verschiedenen Bilder, im Gegensatz zu den Schatten Platos, nicht mehr als Gegenstände.



Zöllner beschäftigte sich eingehend mit der vierten Dimension, die er später zur Erklärung spiritistischer Erscheinungen heranzog. Die Erklärungen waren schon richtig, bloß die Erscheinungen selber nicht. Zöllner wurde das Opfer eines amerikanischen Taschenspielers namens Slade, der Zöllners Hang zur Mystik geschickt ausbeutete und es verstand, mit dessen Ruf für sich Reklame zu machen.

Das gelang Slade um so leichter, als Zöllner stets bereit war, sich voll Gereiztheit auf die Seite einer vermeintlich unterdrückten guten Sache zu schlagen und dabei unterließ, diese vorher unbefangen zu prüfen.

Wie es kam, kann man den Erinnerungen des Göttinger Mathematikers Felix Klein entnehmen:

"Kurz vorher hatte ich Zöllner gelegentlich eines rein wissenschaftlichen Gesprächs von Resultaten erzählt, die ich über verknotete geschlossene Raumkurven gefunden und im Band 9 der 'Mathematischen Annalen' veröffentlicht hatte.
Es war dies die Tatsache, dass das Vorhandensein eines Knotens als eine wesentliche, das heißt gegen Verzerrungen invariante Eigenschaft einer geschlossenen Kurve nur insofern betrachtet werden kann, als man sich im dreidimensionalen Raum bewegt;

im vierdimensionalen Raume hingegen lässt sich eine geschlossene Kurve von einem solchen Knoten durch bloße Verzerrung befreien; der Knoten hört also auf, eine Eigenschaft der Analysis situs zu sein, wenn man die Betrachtung aus dem gewöhnlichen Raume heraushebt."

”Diese Betrachtung” , fährt Klein fort, ”nahm Zöllner mit einem mir unverständlichen Enthusiasmus auf. Er glaubte, ein Mittel in der Hand zu haben, die Existenz der vierten Dimension experimentell nachzuweisen und veranlasste Slade, die Beseitigung von Knoten geschlossener Schnüre in praxi vorzuführen.
Slade nahm die Anregung mit seinem gewöhnlichen ’we shall try it’ auf, und wirklich gelang es ihm, das Experiment binnen kurzem zu Zöllners Zufriedenheit durchzuführen.

Dass es sich bei diesem Versuch um eine versiegelte Schnur handelte, auf deren Siegelverschluss Zöllner beide Daumen pressen musste, möge nur beiläufig erwähnt werden.
Zöllner schloss aus diesem Experiment, dass es Medien gäbe, die in einer engeren Beziehung zur vierten Dimension stünden und die Kraft besäßen, Dinge unserer Körperwelt hinüber und herüber zu bewegen, so dass sie für unsere Sinne verschwinden und wiedererscheinen.”

Die Hexenmeister früherer Zeiten mögen vor Neid erblassen! Nicht, als hätte Slade mehr vermocht als sie ehedem, sondern wegen der ihm so freigiebig zugestandenen vierten Dimension. Ehe man an deren Existenz glaubt, müssten Beweise vorliegen und nicht zweifelhafte Manipulationen.
Zöllner ließ sich durch solche täuschen und stieß auf begreiflichen Widerstand. Er kämpfte fieberhaft um seine Überzeugung und veröffentlichte in seinen letzten Lebensjahren täglich über einen Bogen, das bedeutet 16 Druckseiten!
Seine heftige Natur konnte Widerspruch nicht gut vertragen, und er geriet in große Erregung, die sein Ende beschleunigt haben mochte. Mitten aus der Arbeit wurde er durch einen Gehirnschlag fortgerissen.

Das Verschwinden und Wiederauftauchen von Gegenständen kann wie folgt verständlich gemacht - aber nicht etwa vorgeführt! - werden.
Man denke sich Figuren in der Ebene. Hebt man sie aus dieser in die dritte Dimension, dann verschwinden sie für ein flaches Wesen, das in der Ebene lebt. Ähnlich würde für uns ein Körper verschwinden, wenn man ihn in eine vierte Dimension verfrachten könnte. An diese glauben allein Spiritisten, die sie mit Geistern bevölkern.

Durch die vierte Dimension könnte man Geld aus einer verschlossenen Kasse holen, einen linken Schuh in einen rechten verwandeln, das Innere eines Gummiballs in das Äußere umkrempeln. Man kann sich das sofort klarmachen, wenn man eine Dimension abstreift, wie es hier schon wiederholt geschah. Es gäbe keine Briefgeheimnisse, da man ja den Brief aus dem Umschlag nehmen könnte, ohne diesen zu öffnen. Und der Geburtshelfer würde das Kind ohne Geburtswehen aus dem Mutterleib holen. Das alles durch die vierte Dimension.

Die Physik fasst Raum und Zeit zu einer vierdimensionalen Welt zusammen. Den ersten Ansatz verdankt man Lagrange, der ihn in seiner ”Théorie des fonctions analytiques” 1813 machte.
Der Zeit kommt jedoch in dieser Welt eine Sonderstellung zu. Wir erleben den Augen-

blick, der einen Querschnitt durch die vierdimensionale Welt des Physikers bedeutet. Dieser Querschnitt stellt den Inbegriff aller gleichzeitigen Ereignisse dar, wobei die moderne Relativität der Gleichzeitigkeit nicht weiter beachtet werde. Man kann sich das mit Fechner, der darüber 1846 unter dem Pseudonym Dr. Mises mehr zum Spaß schrieb, wie folgt vorstellen.

”Eigentlich ist alles, was wir erleben werden, schon da, und was wir erlebt haben, ist noch da; unsere Fläche von drei Dimensionen, denn es hindert jetzt nichts, von einer solchen in Bezug zum Körperraum von vier Dimensionen zu sprechen, ist nur durch jenes schon durch und durch dieses noch nicht durch.
Wenn also ein Mensch zu Anfang Kind, zu Ende Greis, in der Mitte Mann ist, hat man sich vorzustellen, es erstrecke sich in die Richtung der vierten Dimension ein langer Balken hinein, von welchem Balken die drei Dimensionen im Fortschreiten immer so viel abschneiden, als in jedem Augenblicke in sie geht; das gibt dann den Menschen, der in diesem Augenblicke lebt.

Eine gleich wichtige Anwendung dieser Erfindung würde darin bestehen, dass uns die ganze Buchdruckerkunst hiermit erspart wäre.

Jedes Buch, was ein Autor schreibt, verlängert sich nämlich auch balkenförmig in die vierte Dimension hinein, da es ja doch nicht gleich, wenn es der Autor geschrieben hat, von der Erde verschwindet. Nach voriger Weise aber können wir beliebig viele Exemplare daraus schneiden, die überdies alle das Verdienst der Originalhandschrift des Verfassers haben. Freilich wird jedes dieser Exemplare wieder nur kurze Zeit dauern; aber was tut das bei Büchern, die ohnehin nur da sind, um neue Bücher danach zu schreiben ; sie würden ihren Zweck, diesen Platz zu machen, nur um so schneller erfüllen.”

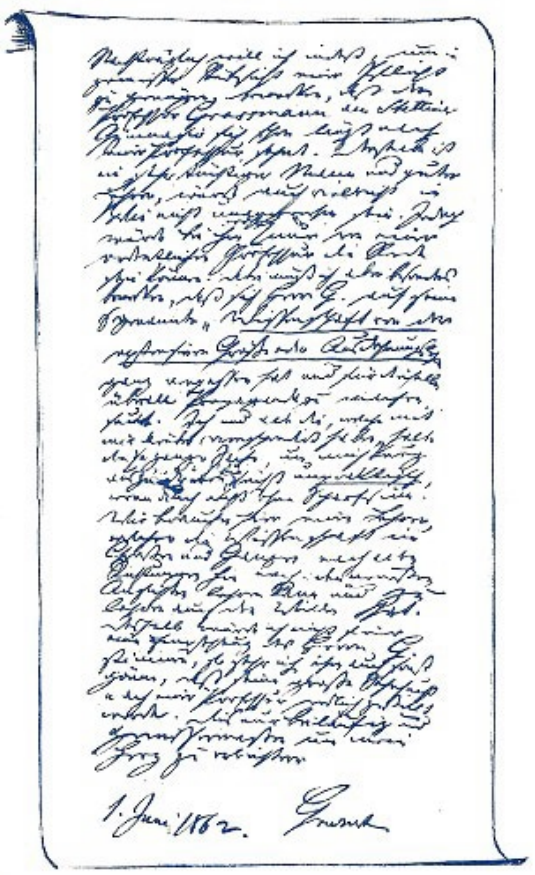
Nach dieser Abschweifung in die Satire erkennt Fechner von der höheren Warte einer vierdimensionalen Welt aus jede Bewegung als nur scheinbar:

”Natürlich, wenn sich etwas in unseren drei Dimensionen zu bewegen scheint, rührt dies nun auch bloß daher, dass der Balken, den es in den Raum der Vier hinausstreckt, schief gegen die drei Dimensionen gerichtet ist und daher beim Fortgange der Fläche von drei Dimensionen diese immer an anderen Stellen schneidet.
Je schiefer, um so schneller scheint die Bewegung. Ist diese Bewegung krummlinig, so rührt das bloß von einer krummen Gestalt des Balkens her.”

Eine systematische Behandlung gleich des n-dimensionalen Raumes verdankt man der ”Ausdehnungslehre” vom Jahre 1844 des Stettiner Gymnasiallehrers Graßmann.
Das Buch war originell aber wenig ansprechend geschrieben, so dass es vom Verfasser später vollständig umgearbeitet wurde. Es war ein verdienstvolles Werk, das Graßmann den berechtigten Wunsch hegen ließ, an einer Universität, und zwar in Greifswald, anzukommen. Als ich 1956 zur 500-Jahr-Feier dieser Universität einiges über die Geschichte

der Mathematik in Greifswald schrieb, veröffentlichte ich das Gutachten des Greifswalder Professors Grunert von 1862, das den verständlichen Wunsch von Graßmann für immer vereitelte. Hier der Wortlaut:

”Nachträglich will ich indess, um in gewisser Rücksicht meiner Pflicht zu genügen, bemerken, dass der Lehrer Graßmann sich schon längst nach einer Professur sehnt.
Derselbe ist ein sehr tüchtiger Mann und ein guter Lehrer, würde auch vielleicht in Berlin nicht unangenehm sein, jedoch würde bei ihm wohl nur von einer ordentlichen Professur die Rede sein können.



Dabei muss ich aber besonders bemerken, dass sich Herr G. auf seine sogenannte Wissenschaft von der extensiven Größe oder Ausdehnungslehre ganz versessen hat und für dieselbe überall Propaganda zu machen sucht.
Ich und auch die, welche mit mir darüber correspondirt haben, halten diese ganze Sache, um mich kurz auszudrücken, für höchst unpraktisch, wenn auch nicht ohne Scharfsinn.
Wir brauchen hier einen Lehrer, welcher die Wissenschaft im Großen und Ganzen nach allen Richtungen hin nach den neuesten Ansichten lehren kann und zu lehren auch den Willen hat.

Deshalb würde ich nicht für eine Empfehlung des Herrn G. stimmen, so sehr ich ihm auch sonst gönne, dass seine große Sehnsucht nach einer Professur endlich gestillt werde. Dies nur beiläufig und gewissermaßen um mein Herz zu erleichtern.”

Später, leider zu spät, revidierte Grunert seine Meinung.

Alle diese Räume besitzen euklidische Struktur. Dasselbe gilt auch noch von einem Raum von gleich unendlich viel Dimensionen, auf den der Göttinger Mathematiker Hilbert bei seinen Untersuchungen über Integralgleichungen im ersten Jahrzehnt unseres Jahrhunderts stieß. Dieser Raum bildet das geeignete mathematische Instrument für die Behandlung der Quantenmechanik.
Im Vergleich dazu ist die Relativitätstheorie relativ bescheiden, denn ihre Welt ist bloß vierdimensional, sie verlangt dagegen eine von der euklidischen grundverschiedene Metrik, die man dem 1866 verstorbenen Göttinger Mathematiker Riemann verdankt.

Wenn auch der Mathematiker kein Vorstellungsvermögen besitzt, das über unseren dreidimensionalen Lebensraum hinausreicht, ist er kraft seiner Schulung doch in der Lage, alle anderen Räume begrifflich zu erfassen.

5 Naiv oder nicht naiv?

Mit dieser Frage könnte ein Mathematiker einen Monolog eröffnen, der dem des Hamlet ähnelt. In ihm würde er sich dann über die Mengenlehre auslassen und das Gespenst einer Krise im Denken heraufbeschwören. In der Tat, vom naiven Standpunkt aus, ist man gegen Widersprüche nicht gefeit. Stellt man sich dagegen nicht auf den naiven Standpunkt, dann schüttet man bei der Axiomatisierung das Kind mit dem Badewasser aus. Fatal, aber wahr! Wie ist das zu verstehen?

Fundamental ist der Begriff der Menge, die der Begründer der systematischen Mengenlehre, Georg Cantor, wie folgt definiert:

Sie sei "eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die Elemente genannt werden) zu einem Ganzen."

Leider umfasst diese "naive" Definition auch widerspruchsvolle Mengen. Um das einzusehen, bemerken wir zunächst, dass eine Menge von ihren Elementen zu unterscheiden ist. Betrachtet man daraufhin die Menge aller Mengen, so ist das zwar nach Cantor erlaubt, aber widerspruchsvoll, denn sie müsste sich selber als Element enthalten, was unserer Bemerkung eben zuwiderläuft.

Als sich Ende des vorigen Jahrhunderts Widersprüche dieser Art mehrten, versuchte man, den Mengenbegriff einzuengen. Es wurden verschiedene kunstvolle Begründungen der Mengenlehre angepriesen, die aber bisher keinen durchschlagenden Erfolg hatten.

Klaua macht 1964 in seinem Buch "Allgemeine Mengenlehre" auf Seite 571 das Eingeständnis: "Die unerreichbaren Zahlen finden Anwendung bei der grundsätzlichen Fragestellung, ob die Cantorsche naive Mengenlehre durch ein Axiomensystem beschrieben werden kann, oder ob man, wie zunächst im allgemeinen vermutet wird, und wie wir es schon in der Einleitung dieses Buches erwähnt hatten, im Laufe der Entwicklung immer neue Axiome zu den bereits verwendeten Axiomen hinzunehmen muss. Auf Grund der Erfahrungen mit dem Existenzproblem für unerreichbare Zahlen kommen wir dabei im folgenden zu dem Schluss, die Cantorsche naive Mengenlehre wird wohl kaum jemals axiomatisch ausgeschöpft werden können."

Darum verbleiben wir lieber bei der naiven Definition von Mengen und entscheiden uns damit für die "naive" Mengenlehre.

Ein zweiter Fundamentalbegriff ist die Abbildung. Wird jedem Element aus einer Menge ein Element aus einer zweiten Menge, die mit der ersten Menge zusammenfallen darf, zugeordnet, dann liegt eine eindeutige Abbildung vor; ein Begriff, der in dieser Allgemeinheit erst von Dirichlet geprägt wurde.

Die Elemente der ersten Menge heißen Urbilder, die der zweiten Menge Bilder. Tritt dabei jedes Element der zweiten Menge als Bild auf, dann redet man von einer Abbildung der ersten Menge auf die zweite Menge, sonst aber in die zweite Menge. Ist die Abbildung

speziell so, dass verschiedenen Urbildern immer verschiedene Bilder entsprechen, dann liegt eine umkehrbar eindeutige Abbildung vor, die man auch eineindeutige Abbildung nennt. Die beiden Mengen heißen dann äquivalent, in Zeichen $\sim$. Damit sind wir in der Lage, das Abzählen begrifflich zu erfassen.

Beim Abzählen werden den abzuzählenden Gegenständen Zahlwörter zugeordnet. Die abzuzählenden Gegenstände bilden die Menge der Urbilder, die verwandten Zahlwörter die Menge der Bilder, und das Abzählen stellt eine umkehrbar eindeutige Abbildung der beiden Mengen aufeinander her.

Eine Verallgemeinerungsmöglichkeit leuchtet unmittelbar ein. Da es allein auf eine umkehrbar eindeutige Abbildung der beiden Mengen aufeinander ankommt, können wir von ihrer Endlichkeit absehen. So gelangt man zum Begriff der Kardinalzahl oder auch Mächtigkeit.
Sie kann endlich sein, dann handelt es sich um eine natürliche Zahl, sie kann aber auch unendlich sein, dann handelt es sich um eine transfinite Kardinalzahl.
Fasst man alle einander äquivalenten Mengen zu einer Klasse zusammen, dann kann die Klasse als Zahlwort gelten.

Betrachtet man die Menge aller natürlichen Zahlen, dann ist sie äquivalent der Menge aller geraden Zahlen. Man hat dafür der natürlichen Zahl n die gerade Zahl $2n$ zuzuordnen und zu bemerken, dass die Abbildung umkehrbar eindeutig ist. Paradox daran ist, dass eine Menge einer echten Teilmenge äquivalent sein kann, im Gegensatz zum Dogma, das Ganze ist größer als ein Teil.

Für endliche Mengen gilt allerdings das Dogma, so dass man die unendlichen von den endlichen Mengen gerade dadurch unterscheiden kann. Einen einfacheren Beweis als es der klassische ist, findet man in meinem Buch "Über Zahlen und Überzahlen" .

Sollen die Kardinalzahlen ihren Namen wirklich verdienen, dann müssen zwei Kardinalzahlen immer entweder gleich sein, oder die eine kleiner als die andere, mit einer geeigneten Erklärung für das kleiner-größer Verhältnis. Ein Nachweis für eine solche Vergleichbarkeit gelingt mittels einer ganz bestimmten Art von Mengen, den wohlgeordneten Mengen.

Eine Menge heißt wohlgeordnet, wenn für zwei verschiedene Elemente a und b immer entweder a "vor" b gilt, in Zeichen $a < b$, oder aber b "vor" a, in Zeichen $b < a$.
Ferner muss, mit c ein drittes Element bezeichnet, aus $a < b$ und $b < c$ immer $a < c$ folgen. Schließlich soll jede Teilmenge ein erstes Element besitzen, für das also "a vor allen übrigen Elementen der Teilmenge" gilt.

Zwei wohlgeordnete Mengen $M = \{a, b, ...\}$ und $M' = \{a', b', ...\}$ heißen ähnlich, wenn sie

umkehrbar eindeutig aufeinander abbildbar sind, so dass aus $a < b$ immer $a' < b'$ folgt, die Bilder durch aufgesetzte Striche bezeichnet. Ist insbesondere M' Teilmenge von M, dann kann niemals $a' < a$ sein. Den Beweisführen wir indirekt.

Unter den Elementen, deren Bild vor dem Element selbst liegt, würde es dann ein erstes Element a geben mit $a' < a$. Da a' zugleich Element aus M ist, liegt sein Bild $(a')'$ in M'. Infolge der Ähnlichkeit von M und M' folgt aus $a' < a$ aber $(a')' < a'$, so dass a doch nicht erstes Element aus M wäre mit einem Bild vor a.

Bezeichnet man daraufhin die Menge der Elemente, die in der wohlgeordneten Menge M vor deren Element a gelten mit M_a und nennt sie Abschnitt, dann folgt aus dem eben bewiesenen Satz, dass die Menge keinem ihrer Abschnitte M_a ähnlich sein kann. Sonst würde a als Bild ein Element aus M_a, das ja vor a gilt, entsprechen. Daraus folgt weiter, dass zwei verschiedene Abschnitte M_a und M_b nie ähnlich sind. Sonst würde, wenn etwa $a < b$ gilt, M_a Abschnitt von M_b sein, nach dem eben Bewiesenen also M_b unmöglich ähnlich.

Damit sind wir in der Lage, nachzuweisen, dass zwei wohlgeordnete Mengen $M = \{a, b, ...\}$ und $N = \{\alpha, \beta, ...\}$ entweder ähnlich sind, oder aber die eine einem Abschnitt der anderen ähnlich ist. Rein formal wären dazu vier Fälle zu unterscheiden, die sich offensichtlich gegenseitig ausschließen:

1. Zu jedem M_a gibt es einen ähnlichen Abschnitt N_α und umgekehrt.
2. Zu jedem M_a gibt es einen ähnlichen Abschnitt N_α, aber nicht umgekehrt.
3. Zu jedem N_α gibt es einen ähnlichen Abschnitt M_a, aber nicht umgekehrt.
4. Es gibt weder zu jedem M_a einen ähnlichen Abschnitt N_α, noch zu jedem N_α einen ähnlichen Abschnitt M_a.

Im ersten Fall, da es nach dem zuletzt bewiesenen Satz zu M_a nur einen einzigen ähnlichen Abschnitt N_α geben kann, wird M durch die Zuordnung von a zu α umkehrbar eindeutig auf N abgebildet. Diese Abbildung ist ähnlich. Entsprechen nämlich den Elementen a und b mit $a < b$ dabei die Elemente α und β, dann folgt aus der Ähnlichkeit der Abschnitte M_b und N_β, dass der Abschnitt M_a von M_b auf einen Abschnitt N_{α^*} von N_β abgebildet und damit $\alpha = \alpha^*$ ist, was $\alpha < \beta$ besagt. Im ersten Fall sind also M und N ähnlich.

Im zweiten Fall gibt es Elemente α aus N so, dass N_α keinem Abschnitt von M ähnlich ist. Unter diesen Elementen α gibt es wegen der Wohlordnung ein erstes α^*. Dann gibt es zu jedem Abschnitt der wohlgeordneten Menge N_{α^*} einen ähnlichen Abschnitt von M und umgekehrt. Damit liegt der vorhergehende Fall vor, so dass M und N^*_α ähnlich sind, M also einem Abschnitt von N ähnlich ist.

Der dritte Fall wird durch Vertauschen von M mit N auf den zweiten Fall zurückgeführt. Im dritten Fall ist also N einem Abschnitt von M ähnlich.

Der vierte Fall schließlich kann überhaupt nicht eintreten, wie man mit Hilfe eines indirekten Beweises erkennt.
Denn wäre M_a der erste Abschnitt, zu dem es in N keinen ähnlichen Abschnitt gibt, und N_α der erste Abschnitt von, N, zu dem es keinen ähnlichen Abschnitt in M gibt, dann könnte jeder Abschnitt von M_{a^*} von M_a nur einem Abschnitt von N_α ähnlich sein.

Ebenso wäre jeder Abschnitt von N_α einem Abschnitt von M_a ähnlich. Für M_a und N_α würde somit der erste Fall vorliegen, so dass M_a entgegen der Annahme einem Abschnitt von N, nämlich N_α, ähnlich wäre.

Nimmt man an, dass jede Menge wohlgeordnet werden kann, dann schließt man wie folgt auf die Vergleichbarkeit der Kardinalzahlen: Es werden der erste und zweite bzw. der erste und dritte Fall von soeben zusammengefasst, in Zeichen $\leq$ bzw. $\geq$, wobei das Gleichheitszeichen der Äquivalenz vorbehalten bleibt.

Im Sinne dieser Erklärung braucht dann nur noch nachgewiesen zu werden, dass sich der zweite Fall mit Ungleichheitszeichen durch umordnen nicht in den dritten Fall verwandeln kann. Das aber folgt sofort aus einem dadurch nahegelegten Satz, der nach Felix Bernstein benannt wird, obschon ihn dieser erst 1897 bewiesen hat, während Dedekind bereits zehn Jahre vorher einen Beweis fand, der dazu noch den Vorzug besitzt, ohne natürliche Zahlen auszukommen.

Der Satz von Bernstein besagt, dass die Mengen M und N äquivalent sind, wenn M einer Teilmenge N^* von N, in Zeichen $N^* \subset N$, und N einer Teilmenge M^* von M äquivalent ist. Damit gleichwertig ist die Aussage, dass aus $M'' \subset M' \subset M$ und $M'' \sim M$ die Äquivalenz von M und M' folgt.
Aus $M \sim M''$ folgt nämlich $M' \sim M''' \subset M'' \subset M$, so dass M' einer Teilmenge von M äquivalent ist. Andererseits ist die Teilmenge M'' von M' laut Voraussetzung M äquivalent, so dass nach dem Satz von Bernstein $M \sim M'$ gilt.
Wenn umgekehrt aus $M'' \subset M' \subset M$ und $M'' \sim M$ die Äquivalenz $M' \sim M$ folgt, dann folgt im Falle $M^* \subset M$, $N^* \subset N$, $M^* \sim N$, $N^* \sim M$ aus $M^* \sim N$ die Äquivalenz $M^{**} \sim N^* \subset N \subset M^* \subset M$ mit $M^{**} \subset M$, daher $M^* \sim M$, wegen $M^* \sim N$ also schließlich $M \sim N$.

Es genügt daher die als gleichberechtigt erkannte Aussage zu beweisen.
Wir setzen $M'' = P$, $M' - M'' = Q$ und nennen die Teilmenge X von M normal, wenn sie bei der Abbildung f von M auf M'' sowohl ihr Bild $f(X)$ als auch Q enthält, $X \supset f(X) \cup Q$, mit $\cup$ die Vereinigungsmenge bezeichnet, die sowohl die Elemente von $f(X)$ als auch von Q enthält. M ist offensichtlich eine normale Menge.

Versteht man unter dem Durchschnitt von Mengen sämtliche ihnen allen gemeinsamen Elemente, dann ist der Durchschnitt X' aller normalen Mengen selber normal, und zwar die kleinste normale Menge. $Q \subset X'$ trifft nämlich offenkundig zu. Bezeichnet man den Durchschnitt mit $\cap$, dann folgt aus

$$X \cap Y \cap Z \cap \ldots \supset f(X) \cap f(Y) \cap f(Z) \cap \ldots = f(X \cap Y \cap Z \cap \ldots)$$

der noch fehlende Teil der Behauptung.

Die kleinste normale Menge X' ist nun augenscheinlich $f(X') \cup Q$. Aus $f(X') \subset P$ folgt weiter $P = f(X') \cup P^*$, daher

$$P \cup Q = f(X') \cup P^* \cup Q = X' \cup P^* \sim f(X') \cup P^* = P$$

oder anders geschrieben

$$M' = P \cup Q \sim P = M'' \sim M$$

Sofort erhebt sich die Frage, ob am Ende nicht alle transfiniten Kardinalzahlen gleich sind. Wir haben ja nur die drei Möglichkeiten nachgewiesen, ohne zu zeigen, dass jede von ihnen realisierbar ist. Die Frage ist zu verneinen, denn zu jeder Menge gibt es eine andere Menge von größerer Mächtigkeit.

Man hat dafür nur die Menge $\mathfrak{M}$ aller Teilmengen der vorliegenden Menge M zu bilden. Man bedenke dann zunächst, dass M auf einen echten Teil von $\mathfrak{M}$ eineindeutig abgebildet werden kann, indem man jedes Element m aus M mit dem Element aus $\mathfrak{M}$ paart, das gerade die Teilmenge mit dem einzigen Element m ist.
Bei dieser Paarung bleiben Elemente aus $\mathfrak{M}$ einsam, beispielsweise die leere Menge $\varnothing$, die kein Element enthält und als Teilmenge einer jeden Menge gilt. Vielleicht aber könnte M durch eine andere Paarung auf die vollständige Menge $\mathfrak{M}$ umkehrbar eindeutig abgebildet werden?

Es geht jetzt darum, diese Hoffnung zu zerstören. Wenn irgendeine Paarbildung vorliegt, so fasse man diejenigen m, denen Teilmengen zugeordnet sind, die m nicht enthalten, zu einer Menge N zusammen. Wir behaupten, dass N einsam bleibt. Sonst nämlich müsste N entweder das ihm zugeordnete m enthalten, was der Erklärung von N widerspräche, oder das ihm zugeordnete m nicht enthalten, entgegen der Vorschrift für die Bildung von N, nach der dann m in N aufzunehmen wäre.

Cantor hielt die Herstellbarkeit der Wohlordnung für eine Denknotwendigkeit. Um sie wenigstens plausibel zu machen, führte er eine Überlegung, deren Beweiskraft daran scheitert, dass sie eine im Endlichen verbindliche Schlussweise ohne Motivierung auf das Unendliche überträgt.

Im 93. Band der "Acta mathematica" nahm ich mir dann vor, die allgemeine Berechtigung der Schlussweise zu begründen. Dazu brauchte ich mich nur auf die vom naiven Standpunkt aus einleuchtende Forderung zu berufen, dass sich in jeder vorliegenden Menge ein Element auszeichnen lässt. Ist man allerdings ängstlich, dann muss man ein viel schärferes Postulat fordern. Dazu kann man nur sagen, dass dann der Teufel mit Beelzebub ausgetrieben wird.

Das Postulat stellte Levi bereits 1902 auf, doch blieb seine Veröffentlichung unbeachtet, im Gegensatz zur Mitteilung von Zermelo zwei Jahre später, der dazu von seinem Freund Erhard Schmidt angeregt wurde. Das heute nach Zermelo benannte Auswahlaxiom verlangt die prinzipielle Möglichkeit, aus jeder Gesamtheit von Mengen je ein Element gleichzeitig auszuwählen.

Abgesehen davon, dass führende Mathematiker wie Perron heute die äußerst weitgehende Forderung, die das Auswahlaxiom postuliert, mit Skepsis betrachten, erkennt man am Bemühen, gleichwertige Aussagen aufzufinden, wie wenig einleuchtend oder überzeugend die Forderung erachtet wird, ganz zu schweigen davon, dass sie anfangs bei bekannten Mathematikern groben Missverständnissen begegnete, wie man in der Polemik von Zermelo in den "Mathematischen Annalen" 1908 nachlesen kann.

Es seien hier wenigstens einige Aussagen angeführt, die dem Auswahlaxiom gleichwertig sind. Diese Aussagen sind zunächst das Axiom von Hausdorff-Birkhoff, dann das Axiom von Teichmüller-Tukey, schließlich das Axiom von Kuratowski-Zorn. Alle drei Axiome

werden heute nach ihren an zweiter Stelle genannten Wiederentdeckern benannt. Das wird sich wohl auch kaum noch ändern lassen!

Wir schließen mit dem Problem von Cantor, das bis heute ungelöst blieb und Kontinuumproblem heißt. Die Kardinalzahl der Menge aller rationalen Zahlen, die kleinste transfinite Kardinalzahl, ist kleiner als die Kardinalzahl der Menge aller reellen Zahlen. Cantor vermutete, dass es zwischen diesen beiden Kardinalzahlen keine Kardinalzahl gibt, die also größer als die erstgenannte, dagegen kleiner als die zweitgenannte Kardinalzahl wäre. Die Lösung dieses Problems erfordert wohl noch ausstehende Begriffsbildungen.

6 Linie als Regenschirm

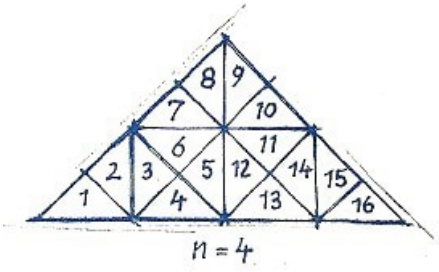
Was ist eine Linie? Eine Frage, auf die es viele Antworten gibt, ein Zeichen dafür, dass sie eine Frage ins Blaue war. In der Tat, trotz vernünftiger Antworten erlebt man Überraschungen, die schon fast Sensationen sind. Man verdankt sie unserer fortgeschrittenen Mathematik.
Im Altertum fehlten dazu noch Begriffe, die erst mühsam herausgearbeitet werden mussten. Ihre Vorläufer waren bloß verschwommene Vorstellungen, die man nicht recht fassen konnte. So die Definition von Euklid, dem Lehrmeister von zwei Jahrtausenden:
"Linie ist Länge ohne Breite."

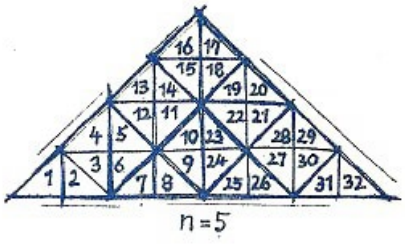
Aufgeweckte Kinder würden es so ähnlich sagen, weil sie über ein begriffliches Rüstzeug dazu noch ebenso wenig verfügen wie seinerzeit Euklid.
Die einfachsten Linien, wie Gerade, Kreis, Kegelschnitte und noch andere, waren schon im Altertum bekannt. Doch bildeten sie lauter Einzel- und keineswegs Spezialfälle. Ein einigermaßen umfassender Kurvenbegriff wurde nämlich erst durch die analytische Geometrie herbeigeführt, die von Descartes und Fermat voneinander unabhängig entwickelt wurde. Legt man dann,wie üblich, eine differenzierbare Funktion $f(x)$ zugrunde, so wird die Kurve durch die Abszisse x und die Ordinate $f(x)$ erklärt. Die vorhin genannten Linien oder Stücke von ihnen bilden jetzt Spezialfälle.

Um den Kurvenbegriff weiter zu verallgemeinern, greifen wir am besten auf die Anschauung zurück. Mit dem Bleistift fahre man über ein Blatt Papier. Es entsteht eine Zeichnung, die den Anlass zu unserer neuen Definition bildet.
In jedem Augenblick t befindet sich die Bleistiftspitze in einem Punkt mit den Koordinaten x und y, die demnach Funktionen von t sind, $x(t)$ und $y(t)$. So gelangen wir zur folgenden Definition:

Die Linie ist durch ein stetiges Funktionenpaar gegeben. t heißt Parameter und bleibt auf ein Intervall beschränkt. Man kann das noch so ausdrücken, dass die Linie stetiges Bild des Parameterintervalls ist.
Leider enthält diese Definition mehr als uns lieb ist. Der Italiener Peano konnte nämlich 1890 nachweisen, dass die ganze Fläche eines Quadrates im Sinne der soeben gegebenen Definition als Linie anzusprechen ist. Der auf einer Erholungsreise in Frankreich 1957 unerwartet verstorbene Professor Knapp stellte dazu 1917 eine Überlegung an, die nicht nur durchsichtiger ist, sondern auch weiter führt.

Offenbar genügt es, nachzuweisen, dass ein gleichschenkliges rechtwinkliges Dreieck als stetiges Bild der Einheitsstrecke aufgefasst werden kann. Man denke sich dazu das Dreieck durch das Lot auf die Hypotenuse in zwei kongruente und wieder gleichschenklige rechtwinklige Dreiecke zerlegt. Mit jedem der beiden Teildreiecke verfahre man ebenso. n-maliges Wiederholen führt auf 2^n kongruente Dreiecke, die dem ursprünglichen Dreieck ähnlich sind.





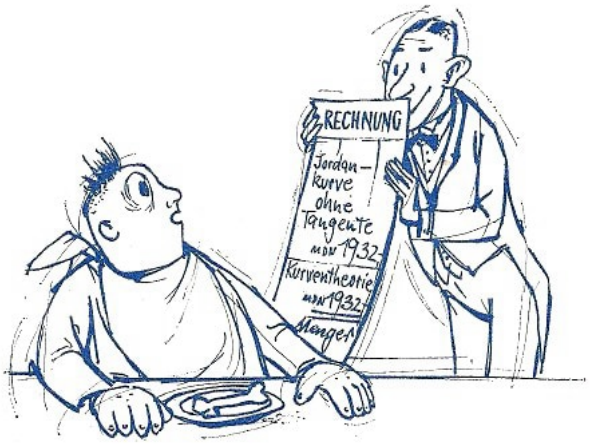
Wir nehmen an, es sei gelungen, die 2^n Dreiecke in eine bestimmte Reihenfolge zu bringen, in der aufeinanderfolgende Dreiecke eine gemeinsame Seite haben. Dann lassen sich die $2^n + 1$ Halb-Dreiecke der nächsten Unterteilung ebenfalls in eine solche Reihenfolge bringen. Denn jedes der 2^n Dreiecke hat entweder mit der Vorgängerin oder der Nachfolgerin eine Kathete gemeinsam.
Im ersten Fall sei dasjenige Halb-Dreieck, das an die Vorgängerin grenzt, das frühere, im zweiten Fall dasjenige der beiden Halbdreiecke, das an die Nachfolgerin grenzt, das spätere.

Da die 2^n Dreiecke der n-ten Unterteilung in eine bestimmte Reihenfolge gebracht werden konnten, entsprechen sie eindeutig den 2^n kongruenten Teilstrecken des Einheitsintervalls, für die ja eine natürliche Reihenfolge besteht.

Ineinandergeschachtelten Intervallen, die aufeinanderfolgenden Unterteilungen entnommen wurden, entsprechen ineinandergeschachtelte Dreiecke. Das genügt, um die Stetigkeit der augenscheinlich eindeutigen Abbildung einzusehen, weiter aber, dass jeder Punkt des ursprünglichen Dreiecks als Bildpunkt erscheint.

Eine Fläche als Linie anzusprechen widerstrebt aber unserer Anschauung. Die Wurzel des Übels liegt darin, dass die Abbildung der Einheitsstrecke auf das Dreieck zwar stetig, aber nicht umkehrbar eindeutig ist. Ist sie es dagegen, dann ist auch umgekehrt das Parameterintervall Bild der Linie, von dem wir zusätzlich annehmen, dass es ein stetiges Bild ist.

Diese Einengung des Kurvenbegriffs nahm erst der Franzose Jordan vor. Dadurch sicherte er das Erfülltsein einer Forderung, die an geschlossene Kurven zu stellen ist. Unsere Anschauung verlangt, dass diese, etwa dem Kreis ähnlich, ein Inneres und ein Äußeres haben.
Jordan konnte nachweisen, dass diese Forderung für seine Kurven wirklich erfüllt ist. Das bedeutet, dass die Kurve auf eine Gummihaut gezeichnet und diese verzerrt, das ursprüngliche Innere in das Innere der neuen Kurve übergeht.



Entsprechendes gilt dann für das Äußere. Das folgt einfach, weil die Verzerrung eine weitere umkehrbar eindeutige, in beiden Richtungen stetige Abbildung bedeutet.

Soweit ist alles gut und schön. Aber auch die Jordankurve präsentiert leider eine unliebsame Überraschung.
Es kann vorkommen, dass sie nirgends eine Tangente hat. Kein Wunder, wenn weitere Definitionen für Linien unternommen wurden, die sich aber samt und sonders nicht mit allen unseren Erwartungen decken. Eine eingehende Kritik solcher Definitionen findet man

in der "Kurventheorie" von Menger aus dem Jahre 1932.

Zum Schluss wollen wir noch eine Überraschung besprechen, die uns eine listig konstruierte Linie bereiten kann. Dazu gehen wir von der Knoppschen Kurve aus.
Für die Werte von t aus dem Einheitsintervall, $0 \leq t \leq 1$, nimmt das Funktionenpaar $x(t)$ und $y(t)$ alle Wertepaare an, die Punkten des ursprünglichen gleichschenkligen rechtwinkligen Dreiecks entsprechen. Wir betrachten jetzt eine Raumkurve, die wir durch

$$x(t), y(t), z(t) = t, \qquad 0 \leq t \leq 1$$

definieren. Infolge des dritten Koordinatenwertes entsprechen verschiedenen t-Werten verschiedene Kurvenpunkte, so dass die Abbildung der Einheitsstrecke auf die Raumkurve nicht nur stetig, sondern auch umkehrbar eindeutig ausfällt: es liegt ein schlichter Bogen vor.
Die Projektion dieses Bogens auf die (x, y)-Ebene, mit anderen Worten die Gesamtheit der Wertepaare $x(t)$, $y(t)$ für $0 \leq t \leq 1$, füllt aber ein Dreieck aus. Von unten gesehen, überdacht folglich unsere Linie ein ganzes Dreieck. Als Regenschirm könnte man sie trotzdem nicht gut patentieren lassen, da sie den Regen nicht abwehren würde.
Den Einwand aber, dass solche Paradoxa nur durch überaus komplizierte Konstruktionen zu erlangen sind, kann man wie folgt widerlegen:

Man betrachte zwei kongruente rechtwinklige Dreiecke mit je einer kurzen und einer langen Kathete. An der Hypotenuse zusammengefügt, bilden sie einen Streifen $ABCD$.

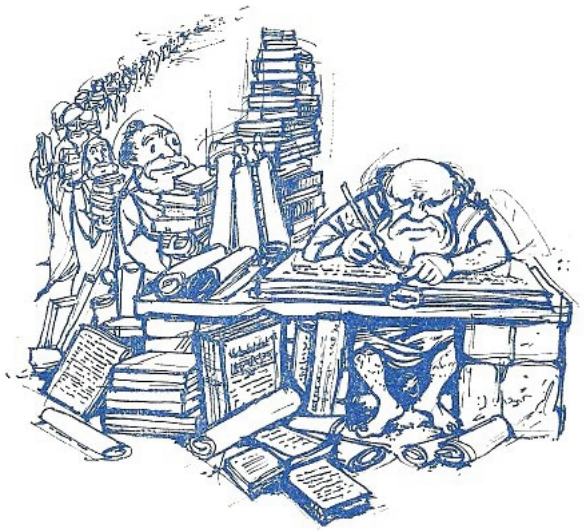
Denkt man sich die Dreiecke aus Papier angefertigt, dann besitzt jedes Dreieck zwei Flächenseiten, sagen wir eine weiße und eine schwarze. Man kann nun die beiden kurzen Katheten auf zweierlei Art zusammenfügen, wobei dann entweder A mit D und B mit C zusammenfällt, oder aber A mit C und B mit D.

Im ersten Fall grenzen dort dieselben Farben aneinander wie bei AC, im zweiten Fall aber nicht. Letzteres besagt, dass man beim Übertritt an der kurzen Kathete aus dem einen in das andere Dreieck plötzlich auf der anderen Flächenseite des Nachbardreiecks landet.

Wenn also eine Ameise vom Mittelpunkt M von AC aus eine Wanderung antritt, bei der sie vom Rand stets in derselben Entfernung bleibt, gelangt sie auf einem Weg von der Länge AD als Antipode nach M. Das bedeutet, dass die Fläche nur eine Seite besitzt. Sie entdeckte Möbius 1858 im Rahmen systematischer Untersuchungen von Polyedern, wobei Flächen aus Dreiecken gebildet wurden. Den Rand unserer einseitigen Fläche bildet eine einzige geschlossene Linie, $ADBC$, die aus den ursprünglichen Strecken AD und BC entsteht.
Nach einer paradoxen Linie und einer noch paradoxeren Fläche sei das vielleicht eindruckvollste Paradoxon der ganzen Mathematik überhaupt genannt, ein räumliches Paradoxon: Die Erde ist so in endlich viel Teile zerlegbar, dass diese Teile anders zusammengefügt, einen Raum von Erbsengröße gerade ausfüllen. Die Einzelheiten kann man in einer Veröffentlichung der beiden Polen Banach und Tarski in "Fundamente mathematicae" aus dem Jahr 1924 nachlesen.

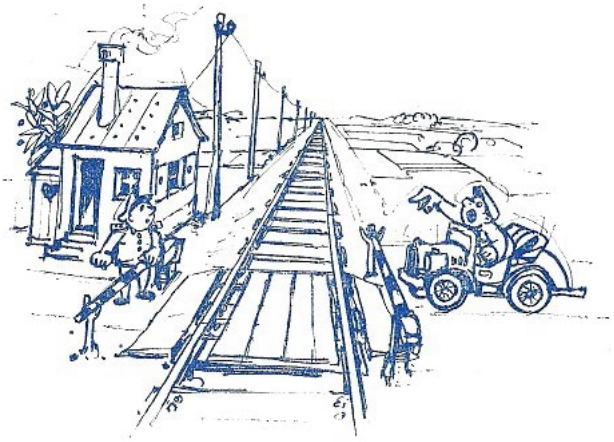
7 Das Gegenteil ist auch richtig

Am Anfang des 4. Jh. v. u. Z. fasste Euklid in seinen "Elementen" alle Erkenntnisse der zeitgenössischen Mathematik zusammen. Das Geringste aus dem Inhalt dieses Werkes stammt von ihm selbst, wohl aber die sehr hoch einzuschätzende Systematik. Er versuchte, von möglichst wenigen, mehr oder weniger plausiblen, Grundsätzen ausgehend, alle anderen Einsichten logisch schließend zu gewinnen.
Es war ein gewaltiges Unterfangen, das nicht restlos gelang und dessen geometrischer Teil erst rund 2000 Jahre später von dem 1942 verstorbenen Göttinger Professor David Hilbert zu Ende geführt wurde. In der heutigen Terminologie redet man von einem axiomatischen Aufbau der Geometrie, so wie man Euklids Stellung in Alexandria als Akademiemitglied bezeichnen könnte.



Einer der Grundsätze, das Parallelenpostulat des Euklid, erschien bereits seinen Zeitgenossen keineswegs einleuchtend.
Um Abhilfe zu schaffen, gingen sie davon aus, dass die Umkehrung des Parallelenpostulats einfach zu beweisen war. Warum sollte also das Postulat selbst nicht ebenfalls beweisbar sein, wenn auch wohl nicht einfach, wovon sie sich bald überzeugen mussten. Eine übereilte Überlegung, aber die Mathematiker waren damals naiv und blieben es auch noch lange, nicht so wie heute, wo sie durch tiefere Einsichten gewitzt wurden - wohlverstanden in der Mathematik - denn im Alltag tappt mancher von ihnen arglos umher.

Was wir sagten, wollen wir belegen. Der Neuplatoniker Proklos schrieb um 450 einen recht ausführlichen Kommentar von rund 300 Seiten zum ersten Buch des Euklid, das selbst nur rund 30 Seiten stark ist. Darin heißt es:
"Dieser Satz (das Parallelenpostulat) ist aus der Reihe der Forderungen (der Grundsätze) völlig zu streichen; denn er ist ein Lehrsatz mit vielen Schwierigkeiten, die zu lösen Ptolemaios in einem Buch sich als Aufgabe stellte."



Weiter heißt es dann:

"Von dieser Forderung (dem Parallelenpostulat) betonten wir, dass ihre Zugehörigkeit zu den ohne Beweis allgemein anerkannten Sätzen (den Grundsätzen) nicht von allen Seiten zugestanden wird."
Weiter aber:
"Es haben denn auch bereits einige andere diesen Satz (das Parallelenpostulat), den unser Autor (Euklid) als Forderung nahm, eher als Theorem bewertet und einen Beweis hierfür als erforderlich gehalten. Auch Ptolemaios scheint ihn zu beweisen in der Abhandlung über den Satz, dass Gerade, die von Winkeln ausgehen, die kleiner als zwei Rechte sind, wenn sie verlängert werden, zusammentreffen."

Nachdem Proklos diesen Beweisversuch analysiert, schließt er seine Erörterungen dazu mit den Worten: "Damit wollen wir unsere Bemerkungen zu Ptolemaios beenden. Denn die Darlegungen zeigen die Schwäche seines Beweises."

Beginnt man mit der Umkehrung des Parallelenpostulats und seines Beweises so, wie es von mir in den "Geometrischen Plaudereien" auf Seite 14 durchgeführt wurde, so erhält man als Ergebnis, dass mindestens eine Parallele existiert. Gibt es auch noch andere?
Das Parallelenpostulat verneint es. Es dauerte dann zwei Jahrtausende, ehe man eine Geometrie auf der gegenteiligen Annahme aufbaute, Welche recht erheblich von unserer Schulgeometrie abweicht.

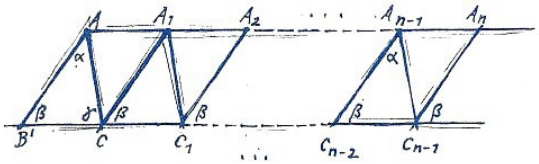
Zunächst wissen wir noch nicht, ob am Ende nicht auch noch eine gemischte Geometrie existiert, in der es für gewisse Paare von Punkt-Geraden nur eine einzige Parallele gibt, für andere Paare dagegen mehrere Parallelen. Wir werden bald sehen, dass dem nicht so ist und erkennen, dass das Parallelenpostulat gleichwertig der Aussage ist, die Winkelsumme in jedem Dreieck betrage zwei Rechte.
Beträgt sie dagegen weniger, dann gibt es z.B. aus A immer unendlich viele Parallelen zu a.

Man könnte daraufhin versuchen, experimentell zu ermitteln, ob in der Wirklichkeit der erste oder der zweite Fall vorliegt. Wegen der prinzipiellen Unvollkommenheit von Messungen wäre es aber aussichtslos, den Nachweis für die Gültigkeit der Schulgeometrie führen zu wollen. Wenn dagegen in einem Dreieck die Winkelsumme, sagen wir 175° betrüge, dann bestünde durchaus die Möglichkeit, experimentell nachzuweisen, dass sie weniger als zwei Rechte beträgt.

Es fehlte nicht an Versuchen, das Parallelenpostulat durch eine andere, plausiblere Forderung zu ersetzen. Wallis lieferte einen interessanten Beitrag dazu, in dem er davon ausging, dass es zu jeder Figur eine ähnliche gibt, die beliebig groß ausfallen darf. Bei Anerkennung dieser Voraussetzung, lässt sich dann darauf schließen, dass es durch einen Punkt außerhalb einer Geraden nur eine einzige Parallele zu der Geraden gibt.

So geistvoll diese Betrachtung auch war, geht sie doch an der Frage vorbei, ob das abgeänderte Parallelenpostulat zu einer neuen, widerspruchsfreien Geometrie führt. Wir werden uns damit bescheiden müssen, in dieser Richtung wenigstens einen ersten Anlauf zu nehmen.

Wir beweisen deshalb zwei Sätze, die den Schlüssel für unsere Ankündigung bilden. Beide stellte Legendre in verschiedenen der zahlreichen Auflagen ab 1794 seiner "Eléments de Géométrie" auf. Seine Haltung war dabei zunächst recht schwankend, ehe er sich zur Klarheit durchgerungen hat. Der erste Satz lautet:



Die Winkelsumme beträgt bei jedem Dreieck höchstens zwei Rechte. Der Beweis wird indirekt geführt. Dementsprechend sei

$$\alpha + \beta + \gamma > \pi$$

Wir verlängern die Seite BC über C hinaus wiederholt um sich selbst,

$$BC = CC_1 = C_1C_2 = ... = C_{n-2}C_{n-1}$$

In den Punkten $C, C_1, ..., C_{n-1}$ errichten wir den Winkel β und zeichnen

$$CA_1 = C_1A_2 = ... = C_{n-1}A_n = BA$$

Die Dreiecke CA_1C_1, $C_1A_2C_2$, ..., $C_{n-2}A_{n-1}C_{n-1}$ sind alle dem ursprünglichen Dreieck BAC kongruent. Daher gilt

$$\begin{aligned} C_1A_1 = ... = C_{n-1}A_{n-1} = CA \qquad & \text{und} \\ \angle CC_1A_1 = \angle C_1C_2A_2 = ... = \angle C_{n-2}C_{n-1}A_{n-1} = \gamma \qquad & \text{weiter} \\ \angle ACA_1 = \angle A_1C_1A_2 = ... = \angle A_{n-1}C_{n-1}A_n = \pi - \beta - \gamma & \end{aligned}$$

Daher sind auch die Dreiecke ACA_1, $A_1C_1A_2$, ..., $A_{n-1}C_{n-1}A_n$ einander kongruent, woraus folgt:

$$AA_1 = A_1A_2 = ... = A_{n-1}A_n$$

Nun ist nach unserer Annahme

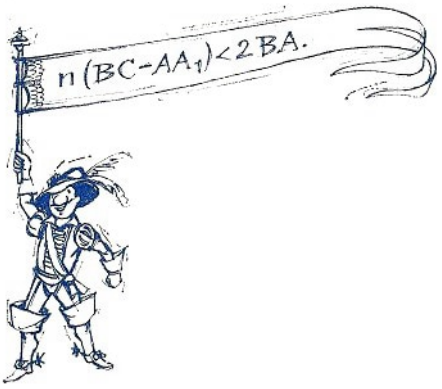
$$\alpha + \beta + \gamma > \pi \qquad \text{, also} \qquad \alpha > \pi - \beta - \gamma$$

In den beiden Dreiecken ABC und ACA_1, sind nach allem die Seiten BA und CA_1 gleich, AC eine gemeinsame Seite, der Winkel BAC größer als der Winkel ACA_1. Dem größeren Winkel liegt aber die größere Seite gegenüber, $AA_1 < BC$.

Der Streckenzug $BAA_1...A_nC_{n-1}$ ist länger als die Strecke BC_{n-1}, daher

$$BA + n \cdot AA_1 + A_nC_{n-1} = 2BA + nAA_1 > n \cdot BC$$

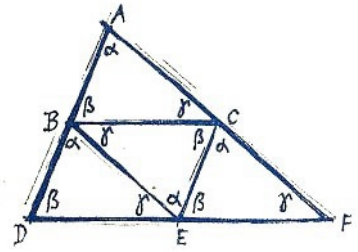
Das ist immer, wie groß auch die natürliche Zahl n sein mag, eine unsinnige Forderung. Daher war die Annahme $\alpha + \beta + \gamma > \pi$ falsch.



Sicherlich hatte auch Gauß als angehender Mathematiker die Abhandlungen von Legendre studiert. Wie das aber auch bei anderen bekannten Mathematikern der Fall war, hat er einfach vergessen, dass er diesen oder jenen Satz damals dort kennenlernte. So kommt es, dass in seinen Werken, Band VIII, Seite 190 (aus dem Nachlass abgedruckt), genau der erste Satz von Legendre und dessen Beweis mit dem Vermerk steht: gefunden 1828. Nov. 18.

Der zweite Satz von Legendre besagt, dass, wenn die Winkelsumme in einem einzigen

Dreieck zwei Rechte beträgt, sie dann in allen anderen Dreiecken ebenso groß ausfällt.



Zunächst behaupten wir, dass man vom Dreieck ABC ausgehend, ein neues Dreieck $AB'C'$ herstellen kann, dessen Seiten AB' und AC' beliebig lang ausfallen, wobei ABB' und ACC' auf je einer Geraden liegen.
Im Dreieck ABC sei die Winkelsumme gleich zwei Rechten.

Die Strecke AB über B hinaus bis D verdoppelt, errichten wir in B über BD den Winkel α und tragen am anderen Schenkel dieses Winkels die Strecke $BE = AC$ ab. Dann sind die Dreiecke ABC und BDE kongruent, folglich

$$\angle BDE = \angle ABC = \beta \qquad \text{und} \qquad \angle BED = \angle ACB = \gamma$$

Aus $\alpha + \beta + \gamma = \pi$ folgt

$$\angle EBC = \pi - \alpha - \beta = \gamma$$

so dass die beiden Dreiecke EBC und ACB ebenfalls kongruent sind. Daraus ergibt sich

$$EC = AB, \quad \angle BEC = \angle CAB = \alpha, \quad \angle BCE = \angle CBA = \beta$$

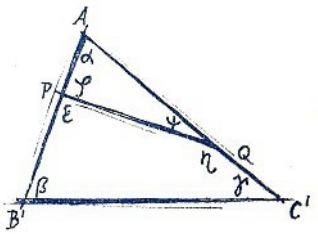
Wird AC über C hinaus bis F verdoppelt, erhält man

$$\angle ECF = \pi - \beta - \gamma = \alpha$$

Folglich sind auch die beiden Dreiecke CEF und ABC kongruent, und daher gilt

$$\angle CEF = \beta \qquad \text{und} \qquad \angle CFE = \gamma$$

Die drei Winkel bei E sind zusammen $\alpha + \beta + \gamma = \pi$, so dass die Punkte D, E, F auf einer Geraden liegen. Aus dem ursprünglichen Dreieck entstand so das Dreieck ADF mit denselben Winkeln und an den Schenkeln von α mit doppelt so langen Seiten. Das hinreichend oft wiederholt, ergibt unsere erste Behauptung.



Weiter behaupten wir, dass die Winkelsumme in jedem Dreieck, in dem ein Winkel gleich α ist, genau zwei Rechte beträgt. In einem Dreieck APQ mit dem Winkel α in A seien in P und Q die Winkel mit φ und ψ bezeichnet. Wie groß auch AP und AQ sind, kann man AB' und AC' so groß machen, dass $AP < AB'$ und $AQ < AC'$ ausfallen.

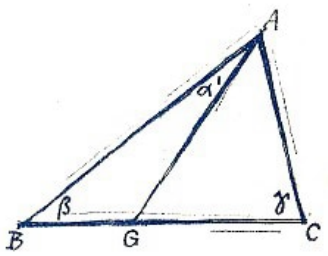
Dann liegen B' und C' in derselben Halbebene, die von der Geraden bestimmt wird, die durch die Punkte P und Q geht, so dass sich die Strecken PQ und $B'C'$ nicht schneiden.

Das Viereck $PQC'B'$ zerfällt folglich durch die Diagonale PC' in zwei Dreiecke. Die Winkelsumme im Viereck mit den Winkeln ϵ, η, β, γ ist gleich der Summe sämtlicher Winkel in den beiden Dreiecken $B'PQ$ und $B'QC'$, nach dem ersten Satz von Legendre also

$$\epsilon + \beta + \gamma + \eta \leq \pi + \pi = 2\pi$$

Im Dreieck $AB'C'$ beträgt die Winkelsumme $\alpha + \beta + \gamma = \pi$. Daraus folgt

$$\begin{aligned} \pi &= (\alpha + \beta + \gamma) + 0 + 0 = (\alpha + \beta + \gamma) + (\varphi + \epsilon - \pi) + (\psi + \varphi - \pi) \\ &= (\alpha + \varphi + \psi) + (\epsilon + \beta + \gamma + \eta - 2\pi) \leq \alpha + \varphi + \psi \end{aligned}$$



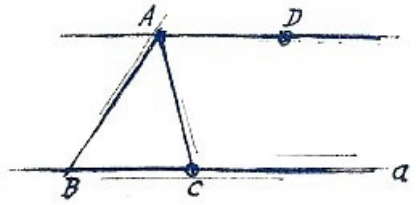
Nach dem oben angeführten Satz kann aber nur das Gleichheitszeichen gelten. Damit ist unsere zweite Behauptung ebenfalls bewiesen.

Hat nun ein beliebiges weiteres Dreieck die Winkel α', β', γ', dann ist nach diesem ersten Satz von Legendre $\alpha' + \beta' + \gamma' \leq \pi$, so dass im Hinblick auf $\alpha + \beta + \gamma = \pi$ nicht zugleich $\alpha < \alpha'$, $\beta < \beta'$, $\gamma < \gamma'$ gelten kann.

Es sei etwa $\alpha \leq \alpha'$. Sollte das Gleichheitszeichen gelten, dann wissen wir bereits, dass $\alpha' + \beta' + \gamma' = \pi$ sein muss. Betrachten wir also den Fall $\alpha' < \alpha$.

Im Dreieck ABC tragen wir an AB den Winkel α' an, dessen anderer Schenkel BC in G treffen möge. Die Winkelsumme im Dreieck ABC ist offensichtlich die um π verminderte Summe der Winkelsumme der beiden Dreiecke ABC und ACC, so dass die beiden letzten Winkelsummen zusammen 2π ausmachen. Nach dem ersten Satz von Legendre kann aber keine dieser beiden Winkelsummen den Betrag π übersteigen. Folglich haben beide Winkelsummen den Wert π.

Wenn aber in einem einzigen Dreieck mit (diesmal) dem Winkel α' die Winkelsumme zwei Rechte beträgt, dann gilt, wie wir schon wissen, dasselbe für alle Dreiecke, die einen Winkel von der Größe α' besitzen. Damit ist der Beweis des zweiten Satzes von Legendre beendet. Aus ihm folgt unmittelbar, dass die Winkelsumme entweder in allen Dreiecken zwei Rechte beträgt, oder aber in keinem. Nimmt man noch den ersten Satz von Legendre hinzu, dann folgt weiter, dass im zweiten Fall die Winkelsumme in allen Dreiecken kleiner ist als zwei Rechte.



Um jetzt die Winkelsumme in Dreiecken mit dem Parallelenpostulat in Zusammenhang zu bringen, machen wir zunächst die Annahme, dass in einem Einzelfall das Parallelenpostulat gelten soll.

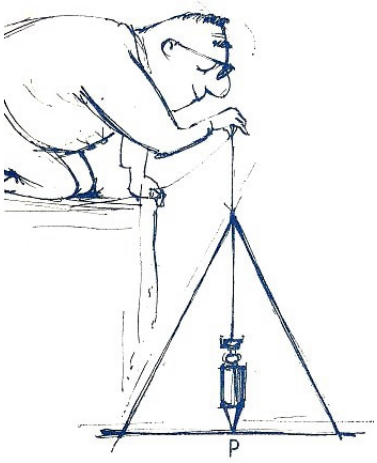
Das bedeutet, dass es zur Geraden a durch den außerhalb liegenden Punkt A eine einzige Parallele gibt. D sei ein weiterer Punkt auf der Parallelen, während B und C zwei Punkte auf a sein sollen. Wie wir wissen, wird die Gerade a von der Geraden AD nicht geschnitten, wenn

$$\angle CBA + \angle BAD = \pi$$

ausfällt. Da es durch A zu a eine einzige Parallele geben soll, gilt für diese die obige Gleichung, und weiterhin

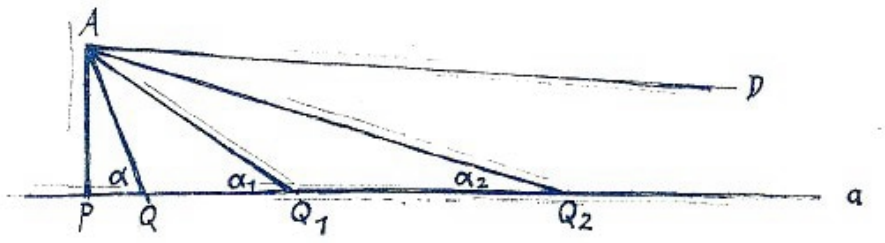
$$\angle BCA = \angle CAD$$

Aus den beiden letzten Gleichungen folgt aber, dass die Winkelsumme im Dreieck ABC zwei Rechte beträgt.
Gibt es dagegen in einem Einzelfall durch A zu a nicht nur eine Parallele, dann fälle man aus A auf a das Lot mit dem Fußpunkt P.



Ist dann AD eine zweite, von der Standardparallelen verschiedene Parallele zu a mit ei-

nem Winkel, der kleiner als ein rechter angenommen werden darf, $\angle PAD = \varphi < \frac{\pi}{2}$, dann wählen wir auf a einen zweiten Punkt Q, der mit D auf derselben Seite der Geraden durch die Punkte A und P liegt. Über PQ hinaus bestimmen wir den Punkt Q, so, dass $QQ_1 = AQ$ ausfällt,dann Q_2 so, dass $Q_1Q_2 = AQ_1$ wird und so fort.



Die Winkel AQP, AQ_1Q, AQ_2Q_1, ... der Reihe nach mit α, α_1, α_2, ..., bezeichnet, gilt in den aufeinanderfolgenden gleichschenkligen Dreiecken AQQ_1, AQ_1Q_2, ... nach dem ersten Satz von Legendre

$$\alpha_1 + \alpha_1 + (\pi - \alpha) \leq \pi \qquad \text{also} \qquad \alpha_1 \leq \frac{\alpha}{2}$$
$$\alpha_2 + \alpha_2 + (\pi - \alpha_1) \leq \pi \qquad \text{also} \qquad \alpha_2 \leq \frac{\alpha_1}{2} \leq \frac{\alpha}{4}$$

schließlich $\alpha_n \leq \dfrac{\alpha}{2^n}$.

Nun muss aber $\angle PAQ_n < \angle PAD$ ausfallen, da im Gegensatz zu AQ_n die Halbgerade AD die Gerade a laut Annahme nicht schneidet. Die Winkelsumme im Dreieck PAQ_n ist folglich kleiner als $\frac{\pi}{2} + \varphi + \frac{\alpha}{2^n}$, für hinreichend großes n wegen $\varphi < \frac{\pi}{2}$ also kleiner als π.

Zusammengefasst besteht demnach folgende Alternative:
Entweder beträgt die Winkelsumme in jedem Dreieck genau zwei Rechte und es gilt das Parallelenpostulat, oder die Winkelsumme in jedem Dreieck beträgt weniger als zwei Rechte und zu jeder Geraden gibt es durch jeden außerhalb liegenden Punkt unendlich viele Parallelen.

8 Größe eines Königs

In der Gesamtheit aller rationalen Zahlen kann man addieren, subtrahieren, multiplizieren und dividieren mit einer rationalen Zahl als Ergebnis. Die vier Grundrechnungsarten führen aus dieser Gesamtheit nicht hinaus. Dasselbe gilt für die Gesamtheit aller reellen Zahlen oder für die Gesamtheit aller komplexen Zahlen.
Allgemeiner kann man Gesamtheiten von Größen betrachten, für die vier Arten von Verknüpfungen von der formalen Beschaffenheit der vier Grundrechnungsarten erklärt sind.

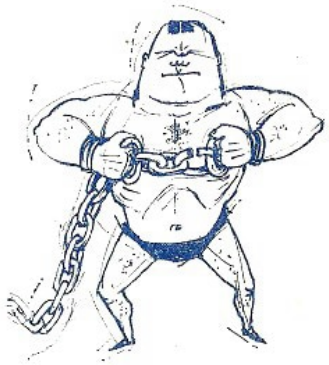
Ist das Ergebnis stets wieder eine Größe aus der Gesamtheit, dann nennt man diese einen Körper. Es handelt sich um einen Begriff, der bereits dem jugendlichen Begründer der modernen Algebra, Evariste Galois, bewusst war.
Jahrzehnte später hat dann der Berliner Mathematiker Kronecker dafür den treffenden Namen Rationalitätsbereich geprägt, dessen Verwendung aber im Abklingen begriffen ist.

Da bei den ganzen Zahlen die Division nicht aufzugehen braucht, bildet die Gesamtheit der ganzen Zahlen keinen Körper. Ähnlich braucht die Division bei Polynomen in einer Veränderlichen mit Koeffizienten aus einem Körper nicht aufzugehen. Die Analogie zwischen ganzen Zahlen und Polynomen ist aber noch viel enger. Um das zu zeigen, wollen wir auf die Division von Polynomen näher eingehen.

Sind $F(x)$ bzw. $f(x)$ zwei Polynome vom m-ten bzw. n-ten Grad mit Koeffizienten aus einem Körper K, und verschwindet $f(x)$ nicht identisch, dann gibt es genau ein Paar von Polynomen $q(x)$ und $r(x)$ mit, Koeffizienten aus K so, dass

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

gilt. Dabei ist der Grad von $r(x)$ kleiner als n.
Dieser Satz folgt bekanntlich durch Koeffizientenvergleich gleich hoher Potenzen auf den beiden Seiten. Verschwindet der Rest $r(x)$ identisch, dann heißt $F(x)$ durch $f(x)$ teilbar, ähnlich wie bei ganzen Zahlen. Da aber mit $F(x) = f(x)q(x)$ zugleich $F(x) = cf(x) - \frac{q(x)}{c}$ gilt, wenn c eine beliebige Größe aus K darstellt, hat man äquivalente Polynome wie $f(x)$ und $cf(x)$ zu identifizieren, um eine völlige Übereinstimmung mit der Teilbarkeitslehre bei ganzen Zahlen herzustellen.
Wir sprachen hier von identifizieren nach dem Sprachgebrauch früherer Algebraiker für die heutige Klasseneinteilung.



Man kann daraufhin, wie bei ganzen Zahlen, den euklidischen Algorithmus anwenden, um den größten gemeinsamen Teiler $d(x)$ oder kürzer d von zwei Polynomen $f_1(x)$ und $f_2(x)$, kürzer f_1 und f_2, zu ermitteln. So bildet man durch fortgesetzte Division

$$f_1 = f_2 q_1 + f_3$$
$$f_2 = f_3 q_2 + f_4 ...$$

Die Kette muss abbrechen, denn die Grade der Polynome $f_3, f_4, ...$ bilden eine abnehmende Folge nichtnegativer ganzer Zahlen. Es gibt also ein n so, dass

$$f_{n-1} = f_n q_{n-1} + f_{n+1} \qquad ; \qquad f_n = f_{n+1} q_n$$

gilt. Das Polynom f_{n+1} ist dann der größte gemeinsame Teiler von f_1 und f_2.
Jedes Polynom f_ν in dieser Kette kann nun in der Form

$$f_\nu = f_1 s_\nu + f_2 t_\nu$$

geschrieben werden, mit s_ν und t_ν Polynome mit Koeffizienten aus K bezeichnet. Zunächst gilt für $\nu = 1$ und $\nu = 2$

$$f_1 = f_1 \cdot 1 + f_2 \cdot 0 \qquad ; \qquad f_2 = f_1 \cdot 0 + f_2 \cdot 1$$

Um vollständige Induktion anwenden zu können, nehmen wir die Richtigkeit bis zu ν als erwiesen an. Dann gilt

$$\begin{aligned} f_{\nu+1} &= f_{\nu-1} + f_\nu q_{\nu-1} = (f_1 s_{\nu-1} + f_2 t_{\nu-1}) - (f_1 s_\nu + f_2 t_\nu) q_{\nu-1} \\ &= f_1(s_{\nu-1} - s_\nu q_{\nu-1}) + f_2(t_{\nu-1} - t_\nu q_{\nu-1}) = f_1 s_{\nu+1} + f_2 t_{\nu+1} \end{aligned}$$

Insbesondere kann also der größte gemeinsame Teiler $f_{n+1} = d$ in der Gestalt $f_1 s + f_2 t$ geschrieben werden.

Schließlich entsprechen Primzahlen irreduzible Polynome, die sich nicht als Produkt von Polynomen mindestens ersten Grades mit Koeffizienten aus K darstellen lassen. Ähnlich wie für ganze Zahlen folgt, dass ein Polynom entweder selbst irreduzibel oder aber das Produkt von irreduziblen Polynomen ist. Soviel genügt, um die Konstruktion von Gleichungswurzeln vorzunehmen. $f(x)$ sei ein Polynom n-ten Grades mit Koeffizienten aus K. Wenn nötig, dividiere man $f(x)$ durch den Koeffizienten des höchsten Gliedes. Die dann angesetzte Gleichung $f(x) = 0$ kann in der Form

$$x^n = -a_1 x^{n-1} - ... - a_n$$

geschrieben werden. Wir dürfen noch annehmen, dass unser Polynom $f(x)$ irreduzibel ist, weil wir die Betrachtung sonst an einem seiner irreduziblen Faktoren durchführen könnten.

Wir werden eine neue Größe J einführen, die zu den Größen aus K hinzugefügt bewirkt, dass in der so entstandenen Körpererweiterung die vier Verknüpfungen weiterhin vorgenommen werden können. Insbesondere haben dann die Ausdrücke J^n und $-a_1 J^{n-1} - ... - a_n$ einen Sinn.

Es wird sich zeigen, dass sie einander gleich sind. Das bedeutet aber, dass gilt

$$J^n + a_1 J^{n-1} + ... + a_n = 0$$

oder kürzer $f(J) = 0$. Die neue Größe J kann also als Wurzel der Gleichung $f(x) = 0$ angesprochen werden.

Im Anschluss an den 1913 verstorbenen ungarischen Mathematiker Julius König gelangen wir zu unserem Ziel wie folgt:

Wir betrachten alle n-gliedrigen Systeme $(g_0, g_1, ..., g_{n-1})$, $(h_0, h_1, ..., h_{n-1})$, ..., mit Größen g_0, g_1, ... aus K, die wir kürzer mit den entsprechenden großen Buchstaben G, H, ... bezeichnen. Um diese Systeme selbst als Größen aufzufassen, müssen wir zunächst die Gleichheit für sie erklären. Es soll $G = H$ genau dann gelten, wenn

$$g_0 = h_0, g_1 = h_1, ..., g_{n-1} = h_{n-1}$$

ist. Für die Summe $G + H$ und die Differenz $G - H$ definieren wir naheliegend

$$(g_0 + h_0, g_1 + h_1, ..., g_{n-1} + h_{n-1}) \qquad \text{bzw.} \qquad (g_0 - h_0, g_1 - h_1, ..., g_{n-1} - h_{n-1})$$

Offensichtlich ist die Addition kommutativ und assoziativ. Kunstvoll dagegen ist die Erklärung für das Produkt GH. Man bilde die Polynome

$$g(x) = g_0 + g_1 x + ... + g_{n-1}x^{n-1} \qquad \text{und} \qquad h(x) = h_0 + h_1 x + ... + h_{n-1}x^{n-1}$$

multipliziere sie miteinander, dividiere dieses Produkt durch $f(x)$ und erhalte

$$g(x)h(x) = f(x)q(x) + r(x)$$

Der Rest $r(x)$ hat einen niedrigeren Grad als $f(x)$, daher den Grad $n - 1$, also

$$r(x) = r_0 + r_1 x + ... + r_{n-1}x^{n-1}$$

Das System $R = (r_0, r_1, ..., r_{n-1})$ wird dann als das Produkt von G mit H erklärt, $GH = R$. Man kann sofort nachprüfen, dass diese Multiplikation kommutativ und assoziativ ist, und dass weiterhin auch das Distributivgesetz gilt.

$H \neq (0, 0, ..., 0)$ vorausgesetzt, verlangt die Division von G durch H das Auffinden eines Systems $U = (u_0, u_1, ..., u_{n-1})$ für das $G = HU$ gilt, oder ausführlich

$$h(x)u(x) = f(x)q(x) + g(x)$$

Da nach Voraussetzung $f(x)$ irreduzibel ist, ist der größte gemeinsame Teiler von $h(x)$ und $f(x)$ eine Konstante, die wir gleich eins setzen dürfen. Es gilt dann, wie wir anfangs sahen,

$$1 = h(x)s(x) + f(x)t(x)$$

Die vorhergehende Gleichung mit $s(x)$ multipliziert, folgt

$$h(x)s(x)u(x) = f(x)s(x)q(x) + g(x)s(x)$$

Wird für $h(x)s(x)$ der Ausdruck $1 - f(x)t(x)$ eingesetzt, ergibt sich

$$g(x)s(x) = f(x)[-s(x)q(x) - t(x)u(x)] + u(x)$$

Die Division von $g(x)s(x)$ durch $f(x)$ liefert aber den Ausdruck in der eckigen Klammer, die also ungeachtet dessen, dass $u(x)$ darin vorkommt, als bekannt anzusehen ist. Wird dieser Ausdruck mit $q'(x)$ bezeichnet, so gilt

$$u(x) = g(x)s(x) - f(x)q'(x)$$

Multipliziert man diese Gleichung mit $h(x)$, dann folgt

$$\begin{aligned} h(x)u(x) &= g(x)h(x)s(x) - f(x)h(x)q'(x) = g(x)[1 - f(x)t(x)] - f(x)h(x)q'(x) \\ &= g(x) + f(x)[-g(x)t(x) - h(x)q'(x)] \end{aligned}$$

Die eckige Klammer spielt jetzt die Rolle von $q(x)$ in der Ausgangsgleichung dieses Absatzes. Division durch $h(x)$ ergibt dann den gesuchten Quotienten $u(x)$.

Die n-gliedrigen Systeme bilden danach einen Körper K^*. Diejenigen unter ihnen, die außer an der ersten Stelle lauter Nullen enthalten, können wir den entsprechenden Größen k aus K gleichstellen,

$$k = (k, 0, ..., 0)$$

Nun führen die Verknüpfungsgesetze zu denselben Ergebnissen, gleichgültig ob man sie unmittelbar auf die Größen aus K oder aber auf ihre Systemdarstellung anwendet. So erkennt man, dass K^* den ursprünglichen Körper K umfasst, $K \subset K^*$.
Das System $(0, 1, 0, 0, ..., 0)$ werde mit J bezeichnet. Dann folgt

$$J^2 = (0, 0, 1, 0, ..., 0) \qquad ... \qquad J^{n-1} = (0, 0, 0, 0, ..., 1)$$

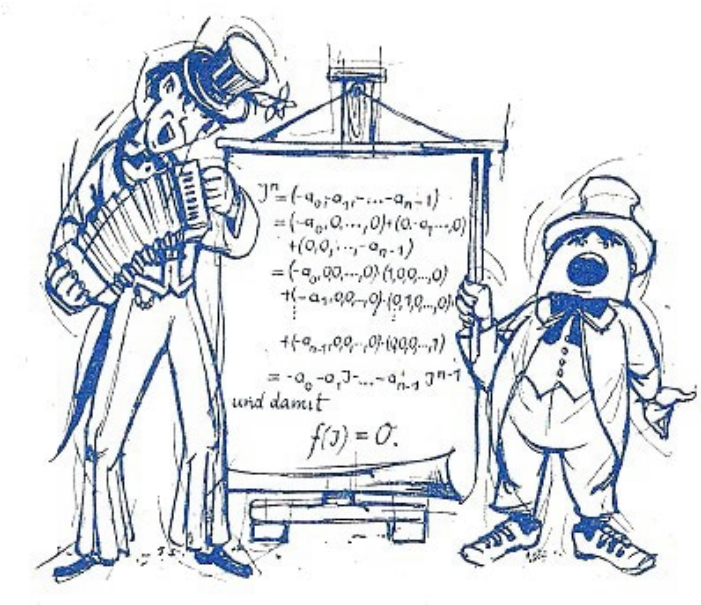
Den Größen J und J^{n-1} entsprechen danach die Polynome x und x^{n-1}. Um das Polynom zu finden, das $J^n = J \cdot J^{n-1}$ entspricht, hat man das Produkt der beiden Polynome x und $^{n-1}$ durch $f(x)$ zu dividieren. Der Rest ist dann das gesuchte Polynom. Nun ist

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

der Rest also

$$-a_0 - a_1 x - ... - a_{n-1} x^{n-1}$$

Dieses Polynom entspricht der Größe J^n. Im Sinne der Verknüpfungsgesetze ist aber



Die Größe J ist also eine Wurzel der Gleichung $f(x) = 0$.
Die Algebra kennt diese Konstruktion von Wurzeln unter dem Namen der symbolischen Adjunktion. Hier wurde jedoch die Konstruktion ohne Inanspruchnahme der algebraischen Systematik durchgeführt. Nun gilt

$$f(x) = (x - J) f^*(x)$$

wobei $f^*(x)$ ein Polynom $(n-1)$-ten Grades mit Koeffizienten aus K^* ist. Es besteht durchaus die Möglichkeit, dass $f^*(x)$ in Faktoren zerfällt, welche Polynome mit Koeffizienten aus K^* sind, dass es also in K^* nicht irreduzibel ist.

Wie dem auch sei, entweder an $f^*(x)$, wenn es irreduzibel ist, oder an einem der irreduziblen Faktoren wiederholen wir die Wurzelkonstruktion von vorhin, mit dem Ergebnis, dass von $f^*(x)$ ein linearer Faktor $x - J^*$ abgespaltet wird und J^* im Körper K^{**} enthalten ist, der K^* umfasst. Erneute Wiederholungen führen schließlich auf die Zerlegung von $f(x)$ in Linearfaktoren

$$f(x) = (x - J)(x - J^*)...(x - J^{**...*})$$

Die Körper $K \subset K^* \subset K^{**} \subset ...$ bilden eine aufsteigende Folge endlich vieler Glieder, und sämtliche Wurzeln J, J^*, ... gehören dem letzten dieser Körper an.

Mit diesem rein algebraisch gewonnenen Ergebnis kommt man im System der modernen Algebra aus. Früher nahm seine Stelle der Satz ein, dass $J, J^*, ...$ alle bereits in K liegen, wenn K der Körper der gewöhnlichen oder gaußschen komplexen Zahlen ist. Der Beweis dafür erfordert aber Stetigkeitsbetrachtungen, die der Algebra wesensfremd sind, so dass dieser sogenannte Fundamentalsatz der Algebra paradoxerweise gar kein Satz der Algebra ist.

Zum Schluss noch ein Beispiel, das in unserer Wurzelkonstruktion die Weiterführung eines Ansatzes von Hamilton erkennen lässt. Wir betrachten das Polynom

$$f(x) = x^2 + 1$$

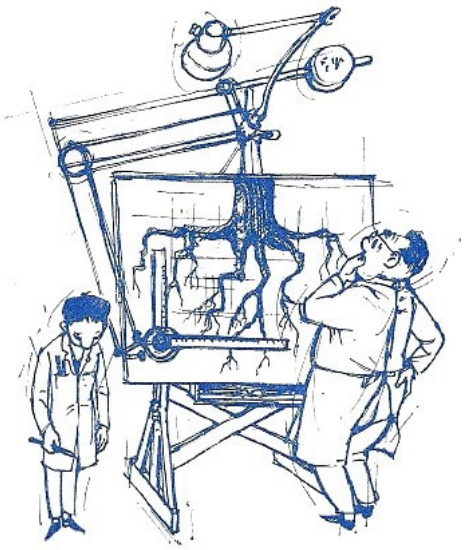
Im Körper K der reellen Zahlen ist es irreduzibel. K^* besteht diesmal aus allen zweigliedrigen Systemen

$$(g_0, g_1), (h_0, h_1), ...$$

dabei sind g_0, g_1, h_0, h_1 reelle Zahlen. Die Erklärungen für die Gleichheit und die vier Rechnungsarten stimmen im Ergebnis mit den Erklärungen von Hamilton überein. Das sei am Produkt erläutert:
Es ist

$$(g_0 + g_1x)(h_0 + h_1x) = g_1h_1(x^2 + 1) + (g_0h_0 - g_1h_1) + (g_0h_1 + g_1h_0)x$$

Wie man sieht, handelt es sich in der Tat um die Hamiltonsche Einführung der komplexen Zahlen.
Es sei aber noch vermerkt, dass Cauchy etwas später, etwa 1847, die komplexen Zahlen genauso einführte, wie unsere Wurzelkonstruktion für $f(x) = x^2 + 1$ verläuft.



9 Streckenrechnung

Bezeichnen x, y, z, u, komplexe Zahlen, dann kann man die Differenzen $y-x$, $z-y$, als gerichtete Strecken oder auch Vektoren in der Gaußschen Zahlenebene deuten, die vom Punkt X zu Y bzw. Y zu Z reichen. Dabei wurden die Punkte, die den komplexen Zahlen entsprechen, mit den gleichlautenden großen Buchstaben bezeichnet.
Zwei Strecken $y-x$ und $(y+u)-(x+u)$ gelten wegen

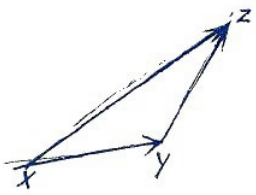
$$y-x=(y+u)-(x+u)$$

als gleich. Man sagt, die zweite Strecke entsteht aus der ersten durch eine Verschiebung. Weiter folgt aus

$$(z-y)+(y-x)=z-x$$

die Gültigkeit der üblichen Vektoraddition, in geometrischer Schreibweise, die beiden Summanden vorher vertauscht,

$$XY+YZ=XZ$$



Davon sei gleich eine Anwendung auf den Schwerpunkt gemacht. Der Schwerpunkt S der Punkte

$$P_1=a_1+b_1i,...,P_n=a_n+b_ni$$

wird durch

$$S=\frac{1}{n}(a_1+...+a_n)+\frac{1}{n}(b_1+...+b_n)i$$

bestimmt. Den Koordinatenanfangspunkt oder Nullpunkt üblicherweise mit 0 bezeichnet, ist

$$OP_1=(a_1+b_1i)-(0+0i),...,OP_n=(a_n+b_ni)-(0+0i)$$
$$OS=[\frac{1}{n}(a_1+...+a_n)+\frac{1}{n}(b_1+...+b_n)i]-(0+0i)$$

daher

$$OS=\frac{1}{n}(OP_1+...OP_n)$$

Für $OP_1=OS+SP_1$, ..., $OP_=OS+SP_n$ eingesetzt, hebt sich OS weg und man erhält

$$SP_1+...+SP_n=0$$

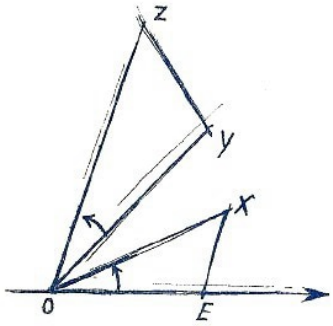
Um zwei Strecken miteinander zu multiplizieren, verwandelt man sie durch Verschiebung in Strecken mit dem Anfangspunkt O, also in OX und OY. Als Produkt gilt dann die Strecke OZ, wenn $z=xy$ ist. Um diese Strecke zu konstruieren, greift man auf die trigonometrische Darstellung von komplexen Zahlen zurück.
Aus

$$x=r(\cos\varphi+i\sin\varphi)\qquad;\qquad y=r'(\cos\varphi'+i\sin\varphi')$$

folgt

$$z=xy=x=rr'(\cos(\varphi+\varphi')+i\sin(\varphi+\varphi'))$$

Man gewinnt daher den Punkt Z, wenn man E auf der reellen Zahlenachse im Abstand eins vom Nullpunkt markiert und die ähnlichen Dreiecke OEX und OYZ konstruiert.

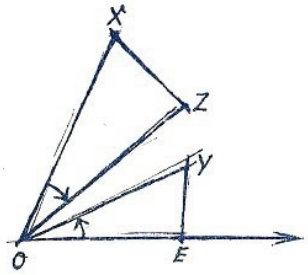


Die trigonometrische Darstellung kann in eleganterer Form geschrieben werden, mit der man auch besser rechnen kann. Dazu verhilft die Formel:

$$e^{i\varphi} = \cos\varphi + i\sin\varphi$$

die Euler 1748 mitteilte. Damit schreibt sich dann

$$x = re^{i\varphi}$$

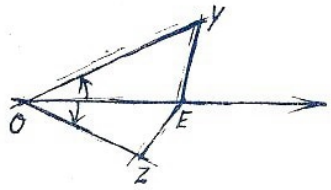


Subtraktion bzw. Division von Strecken können als Umkehroperationen von Addition bzw. Multiplikation eingeführt werden. Bei der Division sieht das so aus: Ist

$$x = re^{i\varphi} \qquad ; \qquad y = r'e^{i\varphi'}$$

dann wird

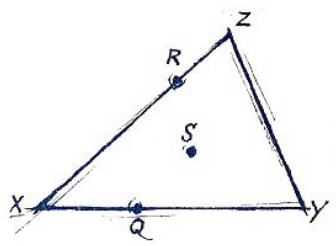
$$z = \frac{x}{y} = \frac{r}{r'}e^{i(\varphi-\varphi')}$$



Diese Formel führt zur folgenden Konstruktion der Strecke OZ, die Quotient der beiden Strecken OX und OY sein soll. Man konstruiert die beiden ähnlichen Dreiecke OYE und OXZ. Insbesondere gewinnt man die zu OY reziproke Strecke $\dfrac{1}{OY} = OZ$ mit Hilfe der Konstruktion der beiden ähnlichen Dreiecke OYE und GEZ.

Die so eingeführte Streckenrechnung erfüllt offensichtlich die formalen Rechengesetze, wie diese für reelle oder komplexe Zahlen gelten. Sie ist also kommutativ; assoziativ und distributiv für Addition und Multiplikation.

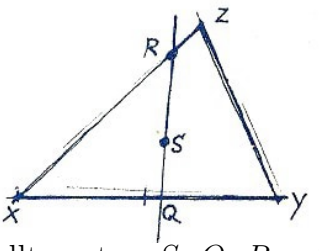
Wir zeigen jetzt, wie man mit Hilfe der Streckenrechnung Konstruktionsaufgaben lösen kann. Es soll das Dreieck XYZ gefunden werden, wenn der Schwerpunkt S der drei Eckpunkte und außerdem je ein Punkt Q bzw. R auf den Seiten XY bzw. XZ gegeben sind, die die Seiten im gegebenen Verhältnis η bzw. ξ teilen:



$$YQ : QX = \eta \qquad ; \qquad ZR : RX = \xi \qquad (*)$$

wobei η und ξ reelle Zahlen sind, beide verschieden von 0 und -1. Es ist

$$SY = SQ + QY = SQ + \eta XQ = SQ + \eta(XS + SQ) = (1+\eta)SQ + \eta XS \qquad \text{und}$$
$$SZ = SR + RZ = SR + \xi xr = SR + \xi(XS + SR) = (1+\xi)SR + \xi XS$$



Diese Ausdrücke in unsere Schwerpunktformel $SX + SY + SZ = 0$ eingesetzt, folgt

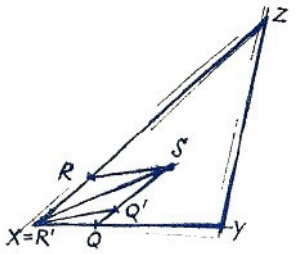
$$(\eta + \xi + 1)SX = (1+\eta)SQ + (1\xi)SR \qquad (**)$$

Sollten etwa S, Q, R auf einer Geraden liegen, dann muss $\eta + \xi - 1$ verschwinden, da

sonst nach $(**)$ auch noch X auf dieser Geraden liegen müsste, nach $(*)$ also auch noch Y und Z, während doch die Punkte X, Y, Z ein Dreieck bilden sollten.
Wenn man daher durch den Schwerpunkt S eine Gerade zieht, welche die Seiten XY bzw. XZ in den Punkten Q bzw. R schneidet, muss

$$\frac{YQ}{QX} + \frac{ZR}{RX} = 1$$

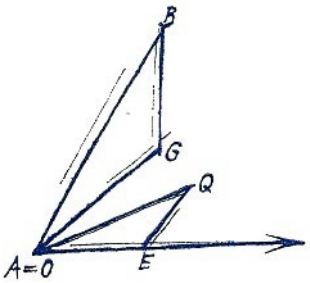
sein. Die Umkehrung dieses nebenbei gewonnenen Satzes gilt ebenfalls, denn aus $\eta+\xi = 1$ folgt aus der Gleichung $(**)$ sofort, dass das Verhältnis $SQ : SR$ einen reellen Wert besitzt, so dass die drei Punkte S, Q, R auf einer Geraden liegen.
Wenn dagegen S, Q, R nicht auf einer Geraden liegen, dann kann, wie zuletzt gezeigt, $\eta + \xi - 1$ nicht verschwinden, so dass die Gleichung $(**)$ die Strecke SX und damit den Punkt X bestimmt. Daraufhin können die Punkte Y und Z aus $(*)$ gefunden werden.



Als Beispiel sei $\eta = 2$ und $\xi = 3$ gewählt. Dann lautet die Gleichung $(**)$, wenn man sie noch durch 4 dividiert,

$$SX = \frac{3}{4}SQ + SR$$

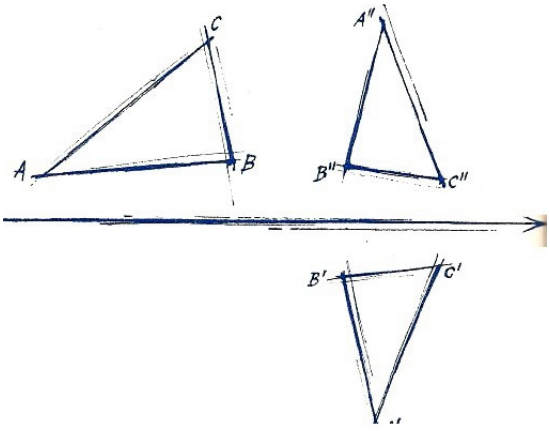
Man konstruiere $SQ' = \frac{3}{4}SQ$ und $Q'R' = SR$. Die neuen Strecken in die rechte Seite eingesetzt, folgt $SX = SQ' + Q'R' = SR'$, so dass die Punkte X und R' zusammenfallen. Da R' durch Konstruktion bekannt ist, ist damit X auch bekannt. Die Punkte Y und Z bestimmt man dann aus $(*)$ zu $YQ = 2QX$, $ZR = 3RX$.



Bei unserer nächsten Konstruktion brauchen wir eine Formel, welche die Ähnlichkeit symmetrisch gelegener Dreiecke ausdrückt. Wir erinnern daran, dass bei der Konstruktion des Streckenproduktes $AC \cdot AQ$, wenn man A durch Verschieben zum Nullpunkt macht, die beiden Dreiecke AQE und ABC gleichsinnig ähnlich ausfallen.
Wenn also zwischen vier Strecken die Proportion

$$\frac{AB}{AC} = \frac{A'B'}{A'C'}$$

besteht, bestimme man zunächst Q aus $AC \cdot AQ = AB$. Dann gilt auch $A'C' \cdot AQ = A'B'$. Denkt man sich noch den Punkt A' durch Verschiebung des Dreiecks $A'B'C'$ in die Lage A gebracht, dann folgt aus den letzten beiden Gleichungen



$$\triangle ABC \sim \triangle AQE \quad \text{und} \quad \triangle A'B'C' \sim \triangle AQE$$

Daher sind auch die Dreiecke ABC und $A'B'C'$ einander ähnlich. Die Umkehrung davon gilt ebenfalls, so dass die gleichsinnige Ähnlichkeit der beiden Dreiecke ABC und $A'B'C'$

durch die Formel

$$\frac{AB}{A'B'} = \frac{AC}{A'C'}$$

ausgedrückt werden kann.
Um daraus Formeln für die symmetrische Ähnlichkeit zu gewinnen, spiegeln wir das diesmal symmetrisch ähnliche Dreieck $A'B'C'$ an der reellen Achse. Für das so entstandene Dreieck $A''B''C''$ gilt

$$A''B'' = b'' - a'' = \overline{b'} - \overline{a'} = \overline{A'B'}$$

den zu P konjugierten Punkt sinngemäß mit $\overline{P}$ bezeichnet. Da das Dreieck $A''B''C''$ infolge der Spiegelung dem Dreieck ABC gleichsinnig ähnlich ist, gilt

$$\frac{AB}{\overline{A'B'}} = \frac{AC}{\overline{A'C'}}$$

Diese Formel bildet den Ausdruck für symmetrische Ähnlichkeit.
Die neue Konstruktionsaufgabe lautet nun: Zwei Strecken, die weder gleiche Länge, noch denselben Anfangs- oder Endpunkt haben, seien gegeben. Es wird ein Punkt X gesucht, so dass die Dreiecke ABX und $A'B'X$ symmetrisch ähnlich sind. Nach der letzten Formel gilt

$$AX : \overline{A'X'} = AB : \overline{A'B'} \qquad \text{daraus folgt} \qquad AX \cdot \overline{A'B'} = AB \cdot (\overline{A'A} + AX)$$

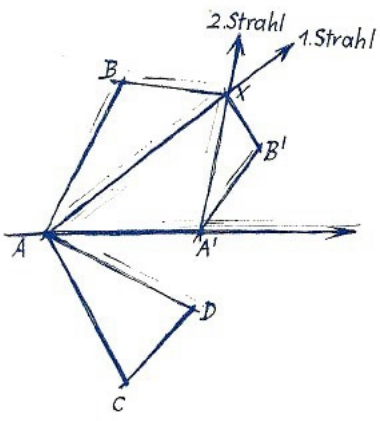
Die konjugierte zu dieser Gleichung lautet

$$\overline{AX} \cdot A'B' = \overline{AB} \cdot (A'A + AX)$$

Aus den letzten beiden Gleichungen berechnet sie AX zu

$$(AB \cdot \overline{AB} - A'B' \cdot \overline{A'B'}) \cdot AX = AB \cdot (AA' \cdot \overline{AB} + A'B' \cdot \overline{A'A'})$$

Wird die Länge einer Strecke PQ mit $|PQ|$ bezeichnet, so ist die Koeffizient von AX gleich $|AB|^2 - |A'B'|^2 \neq 0$, und die Aufgabe hat eine Lösung.



Um die Konstruktion auszuführen, nehmen wir, ohne Einschränkung der Allgemeinheit,

$$AA' = \overline{AA'} = 1$$

an. Dadurch verwandelt sich die letzte Gleichung in

$$(|AB|^2 - |A'B'|^2)AX = AB \cdot (\overline{AB} + A'B')$$

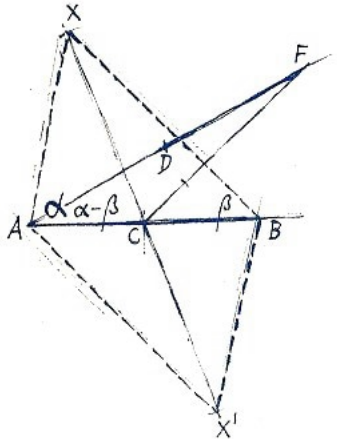
Konstruiert man nun $AC = \overline{AB}$, $A'B' = CD$, $AC + A'B' = AC + AD = AD$, dann findet man

$$(|AB|^2 - |A'B'|^2)AX = AB \cdot AD$$

Ist der Koeffizient von AX positiv, dann ist

$$\angle A'AX = \angle A'AB + \angle A'AC = \angle CAA' + \angle A'AD = \angle CAD$$

so dass $\angle CAD$ im Punkt A an die Seite AA' angetragen, mit dem zweiten Schenkel einen Strahl liefert, auf dem X liegt. Im Falle eines negativen Koeffizienten dagegen fällt der Winkel um 180° verschieden aus.
Um einen zweiten Strahl zu finden, auf dem X liegt, beachte man, dass infolge der symmetrischen Ähnlichkeit $\angle B'A'X = -\angle BAX = \angle XAB$ sein muss.
Man trägt also den bereits bekannten $\angle XAB$ im Punkt A' an die Seite $A'B'$ an und hat im zweiten Schenkel den zweiten Strahl, auf dem X liegt. Die sich kreuzenden Strahlen wurden in der Zeichnung mit Pfeilen versehen.



Nach den ersten beiden Konstruktionsaufgaben ersten Grades sei noch eine Konstruktionsaufgabe zweiten Grades behandelt.
Gegeben ist die Seite AB des Dreiecks ABX, das Produkt der beiden anderen Seiten, $|AX| \cdot |BX|$, und die Differenz $\alpha - \beta$ der anliegenden Winkel.

Zunächst konstruiert man $\angle BAD = \alpha - \beta$ und nimmt den Punkt D gleich so an, dass $|AD| \cdot |AB|$ gleich dem gegebenen Produkt der beiden Dreiecksseiten ausfällt. Das Produkt $AD \cdot BA$ ist eine Strecke, die mit der durch A und B gehenden reellen Achse

$$\angle BAD + 180° = \alpha - \beta + 180°$$

einschließt; für das Streckenprodukt $AX \cdot BX$ folgt ähnlich

$$\angle BAX + \angle ABX = \alpha + (180° - \beta)$$

so dass die beiden Winkel gleich sind. Da außerdem $|AD| \cdot |BA| = |AX| \cdot |BX|$ ist, folgt $AD \cdot BA = AX \cdot BX$.
Ist C Mittelpunkt von AB, also $AC = CB$, dann gilt $AX = AC + CX$, $BX = BC + CX = CX - AC$. Nach Multiplikation folgt

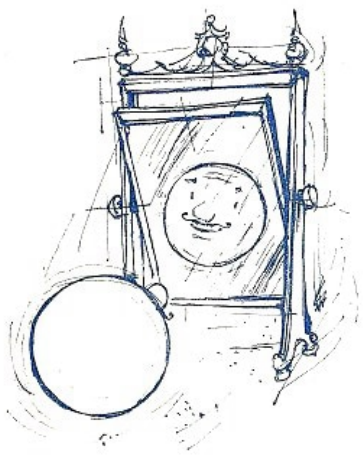
$$CX^2 - AC^2 = AX \cdot BX = AD \cdot BA = 2AD \cdot AC$$

und daraus, F auf dem Strahl durch AD so gewählt, dass $AF = 2AD$ ausfällt

$$CX^2 = 2AD \cdot CA + CA^2 = CA(CA + 2AD) = CA(CA + AF) = CA \cdot CF$$

Daher muss $2\angle ACX = \angle ACA + \angle ACF = \angle ACF$ sein. Dieser Forderung genügen zwei Winkel, nämlich $\angle ACX = \frac{1}{2}\angle ACF$ und $\angle ACX' = \frac{1}{2}\angle ACF + 180°$.
Die gesuchten Punkte X und X' liegen somit auf der Halbierungslinie von $\angle ACF$ zu beiden Seiten von C in einem Abstand, der die mittlere Proportionale von $|CA|$ und $|CH|$ ist.

10 Ein erfinderischer General



Die Spiegelung am Kreis, speziell am Einheitskreis, kann am elegantesten in der komplexen Zahlenebene erklärt werden.
Dabei bevorzugen wir den Einheitskreis lediglich, um Schreibarbeit zu sparen; die Ergebnisse gelten sinngemäß für Kreise mit beliebigem Radius. Es geht dann um eine Abbildung, die dem Punkt z den Punkt

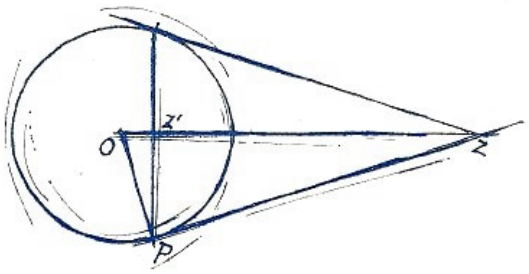
$$z' = \frac{1}{\overline{z}}$$

zuordnet. Erst die Konjugierte, dann die Reziproke gebildet, folgt daraus

$$z = \frac{1}{\overline{z'}}$$

in Worten: Das Bild des Bildes ist das Urbild bei der Spiegelung am Einheitskreis. Eine solche Abbildung heißt Involution.
Der trigonometrischen Darstellung kann man sofort die geometrische Konstruktion des Bildpunktes entnehmen.

Aus $z = re^{i\varphi}$ folgt $z' = \frac{1}{r}e^{i\varphi}$.



Die Berührungspunkte der beiden Tangenten aus z bilden daher eine Strecke, deren Schnittpunkt mit dem Strahl aus 0 durch z den Bildpunkt z' liefert. Das folgt sofort aus

$$\triangle 0z'P \sim \triangle OPZ$$

wenn man die trigonometrische Darstellung von z und z' berücksichtigt.

Da unsere Abbildung eine Involution darstellt, erkennt man weiter, dass das Innere des Einheitskreises bis auf den Mittelpunkt umkehrbar eindeutig auf das Äußere, und umgekehrt, abgebildet ist. Diese Abbildung lässt sich auch noch auf den Mittelpunkt, den Nullpunkt, ausdehnen, wenn man in der komplexen Zahlenebene einen einzigen unendlich fernen Punkt ∞ einführt, wie er sich in der Funktionentheorie bewährt hat. Dann werden die Punkte 0 und ∞ einander zugeordnet.

Die Spiegelung am Einheitskreis ist weiterhin eine Kreisverwandtschaft im Sinne von Möbius, die Kreise wieder in Kreise verwandelt, wenn man die Geraden zu den Kreisen zählt. Um die Berechtigung davon einzusehen, ziehen wir in die (x, y)-Ebene um. In dieser ist die Gleichung eines Kreises mit dem Mittelpunkt $M = (m, n)$ vom Radius r

$$(x - m)^2 + (y - n)^2 = r^2$$

Entwickelt lautet diese Gleichung

$$x^2 + y^2 - 2mx - 2ny + m^2 + n^2 - r^2 = 0$$

woraus nach Multiplikation mit a, wenn $-2ma = b$, $-2na = c$, $a(m^2+n^2-r^2) = d$ gesetzt wird, für den Kreis die Gleichung

$$a(x^2 + y^2) + bx + cy + d = 0$$

folgt. Für verschwindendes a geht aber diese Gleichung in die Gleichung einer Geraden über, so dass unsere Gleichung Geraden ebenso wie Kreise umfasst. Es ist daraufhin gestattet, aber auch im Sinne einer einheitlichen Terminologie erwünscht, die Geraden zu den Kreisen zu zählen.

Die Transformationsformeln schreiben sich jetzt, da $z = (x, y)$ und $z' = (x', y')$ auf demselben Halbstrahl mit dem Ursprung in 0 liegen, als

$$\frac{x}{y} = \frac{x'}{y'} \quad \text{daher} \quad x' = \lambda x, y' = \lambda y$$

Weiter ist

$$1 = |z||z'| = \sqrt{x^2 + y^2}\sqrt{x'^2 + y'^2} = \lambda(x^2 + y^2)$$

So folgt für λ der Wert $\dfrac{1}{x^2 + y^2}$ mit schließlich

$$x' = \frac{x}{x^2 + y^2} \quad ; \quad y' = \frac{y}{x^2 + y^2}$$

Da es sich um eine Involution handelt, gilt auch umgekehrt

$$x = \frac{x'}{x'^2 + y'^2} \qquad ; \qquad y = \frac{y'}{x'^2 + y'^2}$$

Geht man damit in die Kreisgleichung, dann wird

$$a\frac{x'^2 + y'^2}{(x'^2 + y'^2)^2} + b\frac{x}{x'^2 + y'^2} + c\frac{y'}{x'^2 + y'^2} + d = 0$$

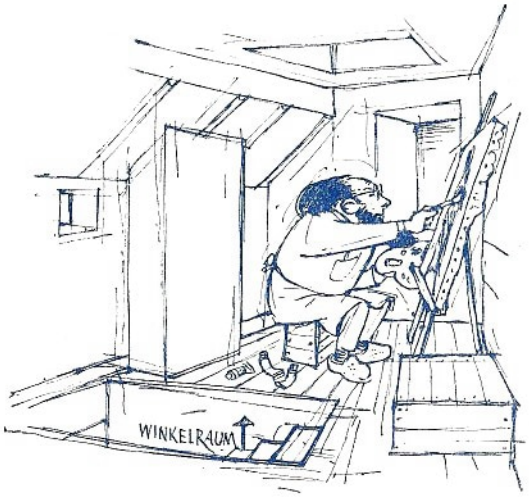
nach Multiplikation mit $x'^2 + y'^2$ also

$$a + bx' + cy' + d(x'^2 + y'^2) = 0$$

wiederum die Gleichung eines Kreises. Die Spiegelung am Einheitskreis ist also wirklich eine Kreisverwandtschaft.

Diese sei noch näher untersucht. Das Bild von einer Geraden, die nicht durch den Nullpunkt geht, $a = 0$, $d \neq 0$, ist ein wahrer Kreis, der durch den Nullpunkt geht, denn $x' = 0$, $y' = 0$ stellt jetzt eine Lösung unserer letzten Gleichung dar.
Das war zu erwarten, wenn der Punkt ∞ auf jeder Geraden liegen soll, weil dessen Bild, die 0, dem Bildkreis angehören muss. Da es sich weiter um eine Involution handelt, gilt auch das Umgekehrte.

Einem wirklichen Kreis, der nicht durch den Nullpunkt geht, $a \neq 0$, $d \neq 0$, entspricht wieder ein wirklicher Kreis.
Jede Gerade, die durch den Nullpunkt geht, $a = 0$, $d = 0$, entspricht sich selbst, allerdings nur als Ganzes und nicht etwa punktweise. Daraus folgt, dass ein Winkelraum mit dem Scheitel im Nullpunkt in sich übergeht.



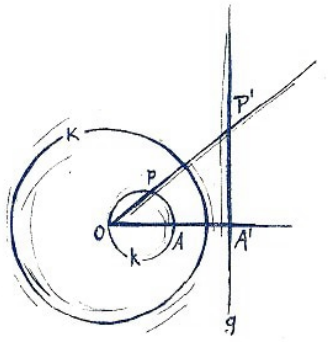
Ein Orthogonalkreis, der den Einheitskreis in zwei Punkten senkrecht schneidet, befindet sich im Winkelraum, den die aus 0 zum Orthogonalkreis gezogenen verlängerten Tangenten bilden.
Da er wieder in einen wahren Kreis übergeht, wie wir vorhin sahen, der im gleichen Winkelraum liegen muss und die beiden Schnittpunkte mit dem Einheitskreis unverändert bleiben, fällt der Bildkreis in diesem Fall mit dem Urbildkreis zusammen.

Es muss folglich für den Orthogonalkreis $a = d$ ausfallen. Das ist eine charakteristische Gleichheit.
Die letzten Einsichten haben wir aus den beiden Gleichungen

$$a(x^2 + y^2) + bx + cy + d = 0 \qquad \text{und} \qquad a + b'x + cy' + d(x'^2 + y'^2) = 0$$

gewonnen. Für eine Anwendung, die wir vorhaben, brauchen wir aber davon nur den Satz, dass einem Kreis, der durch den Nullpunkt geht, eine Gerade entspricht, die nicht durch den Nullpunkt geht. Dieser Satz lässt sich nun unmittelbar, ohne die beiden Gleichungen von vorhin heranzuziehen, wie folgt beweisen:

K sei der Einheitskreis, k ein Kreis vom Durchmesser $OA < 1$ durch 0, g eine Gerade, die auf der Verlängerung von OA über A hinaus im Abstand $0A' = \frac{1}{OA} > 1$ senkrecht steht. Diese Gerade ist dann das Bild von k.



Um das einzusehen, sei P ein beliebiger Punkt auf k und P' der Punkt, in dem die Verlängerung von OP über P hinaus g schneidet. Die rechtwinkligen Dreiecke OPA und $OA'P'$ sind ähnlich, denn in 0 haben sie einen Winkel gemeinsam. Daher ist

$$\frac{OP}{OA} = \frac{OA'}{OP'}, \text{ woraus } OP \cdot OP' = OA \cdot OA' = 1$$

folgt, womit alles bewiesen ist.

Damit sind wir in der Lage, einen Gelenkmechanismus anzugeben, der eine kreisförmige in eine geradlinige Bewegung verwandelt. Im Maschinenbau kennt man dieses Problem als Geradführung.
Nach vielen vergeblichen Versuchen glaubte man schließlich Mitte des vorigen Jahrhunderts, dass es unlösbar ist, bis dann 1864 der französische General Peaucellier die, richtiger eine, Lösung fand. Einiges aus der Geschichte des Problems und eine andere Lösung findet man in meinem Buch "Geometrische Plaudereien".

Der Inversor genannte Gelenkmechanismus von Peaucellier besteht aus einem Gelenkrhombus $PQP'R$ von der Seitenlänge a, an dessen gegenüberliegenden Ecken Q und R zwei gleichlange Stangen von der Länge c gelenkig befestigt sind, die sich im festen Punkt 0 treffen, in dem sie drehbar angebracht sind.

Dabei ist $c > a$. Da gilt

$$\triangle 0QP \cong \triangle ORP$$

liegt P stets auf der Winkelhalbierenden von $\angle QOR$, desgleichen ebenfalls P' wegen

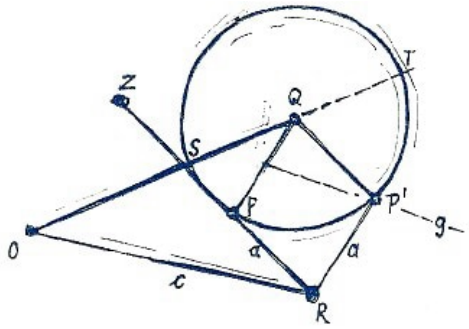
$$\triangle OQP' \cong \triangle ORP'$$

so dass 0, P, P' stets in einer Geraden liegen. Um Q den Kreis mit dem Radius a geschlagen, liegen P und P' auf diesem Kreis. Die Stange OQ bzw. ihre Verlängerung schneiden diesen Kreis in S bzw. T. Für die beiden Sekanten OPP' und OST gilt aber

$$OP \cdot OP' = OS \cdot OT = (c-a)(c+a) = c^2 - a^2$$

so dass das Produkt $OP \cdot OP'$ konstant bleibt. Fasst man c bzw. a als Hypotenuse bzw. Kathete eines rechtwinkligen Dreiecks auf, dessen zweite Kathete als Maßeinheit dient, dann folgt $OP \cdot OP' = 1$, das heißt P und P' gehen durch Spiegelung am Einheitskreis um 0 auseinander hervor.

Um dann diesen aus 6 Stangen bestehenden Apparat zur Geradführung auszubauen, bringt man noch eine siebente Stange ZP von der Länge ZO an, die sich im festen Punkt Z drehen kann und in P gelenkig ist. Danach bewegt sich P auf einem Kreis, der durch 0 geht. Infolgedessen muss sich P' auf einer Geraden g bewegen, die auf OZ senkrecht steht.
Damit ist der Inversor im Prinzip gebrauchsfertig, nur kann P nicht über 0 hinaus.



11 Möbius überführt

Die Transformation

$$w = \frac{az+b}{cz+d}$$

die z in w überführt, heißt lineare Transformation $w = I(z)$. Gelegentlich wird sie auch nach dem 1868 verstorbenen Leipziger Mathematiker Möbius benannt, der sie systematisch untersuchte.

Die rechte Seite kann für $c \neq 0$ in der Form

$$-\frac{ad-bc}{c} \cdot \frac{1}{cz+d} + \frac{a}{c}$$

geschrieben werden, aus der man ersieht, dass die Abbildung für $ad - bc = 0$ ausartet, indem sie dann der ganzen z-Ebene den einzigen Punkt $\frac{a}{b}$ in der w-Ebene zuordnet.
Für $c = 0$ wird die Transformation $\frac{a}{d}z + \frac{b}{d}$, wobei sinngemäß $d \neq 0$ anzunehmen ist, damit nicht der ganze Nenner $cz + d$ der ursprünglichen Transformation verschwindet.

Die allgemeinste lineare Transformation kann durch drei spezielle Transformationen erzeugt werden, die hintereinander auszuführen sind, nämlich

$$\mathfrak{z} = cz + d; \qquad \mathfrak{w} = \frac{1}{\mathfrak{z}}; \qquad w = a'\mathfrak{w} + b'$$

mit $a' = -\frac{ad-bc}{c}$ und $b' = \frac{a}{c}$.
Die erste und dritte heißen ganze lineare Transformationen. Eine ganze lineare Transformation $w = az + b$ kann wiederum in drei Etappen durchgeführt werden. Für $a = re^{i\varphi}$ wird sie erzeugt, wenn man der Reihe nach

$$\mathfrak{z} = e^{i\varphi}z; \qquad \mathfrak{w} = r\mathfrak{z}; \qquad w = \mathfrak{w} + b$$

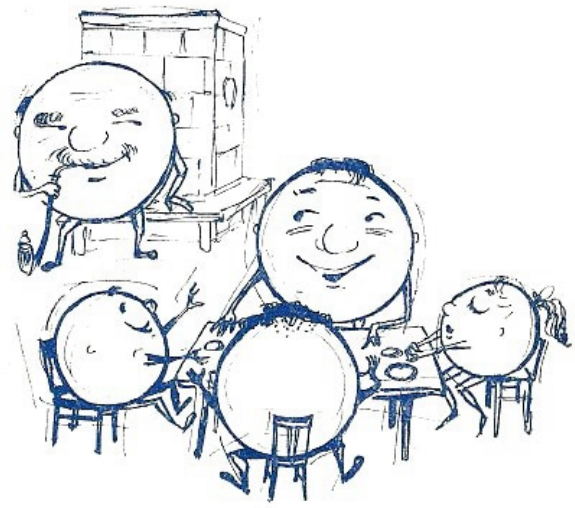
setzt.
Die erste Etappe bedeutet eine Drehung der starren Ebene um den Winkel φ. Dabei gehen Geraden wieder in Geraden und wahre Kreise wieder in wahre Kreise über. Die zweite Etappe bedeutet eine Streckung im Verhältnis r. Die Entfernung von zwei Punkten $\mathfrak{z}' = rz'$ und $\mathfrak{z}'' = rz''$ wird dabei nach

$$|\mathfrak{z}'' - \mathfrak{z}'| = |rz'' - rz'| = r|z'' - z'|$$

im selben Verhältnis gestreckt. Daher gehen wahre Kreise in wahre Kreise über. Aus $\mathfrak{x} + i\mathfrak{y} = \mathfrak{z} = rz = r(x + iy)$ und $kx + my + n = 0$ folgt weiter

$$0 = krx + mry + rn = k\mathfrak{x} + m\mathfrak{y} + rn$$

so dass Geraden in Geraden übergehen.
Schließlich bedeutet $w = \mathfrak{w} + b$ eine Translation der starren Ebene, bei der offensichtlich Geraden in Geraden und wahre Kreise in wahre Kreise übergehen.



Die Transformation $\mathfrak{w} = \frac{1}{\mathfrak{z}}$ geht aus $\mathfrak{u} = \frac{1}{\bar{\mathfrak{z}}}$ durch Spiegelung an der reellen Zahlenachse hervor. Da dabei wahre Kreise und Geraden wieder in wahre Kreise und Geraden übergehen, die Transformation $\mathfrak{u} = \frac{1}{\bar{\mathfrak{z}}}$ aber, wie wir von früher her wissen, eine Kreisverwandtschaft darstellt, so folgt, dass $\mathfrak{w} = \frac{1}{\mathfrak{z}}$ selbst eine Kreisverwandtschaft ist.

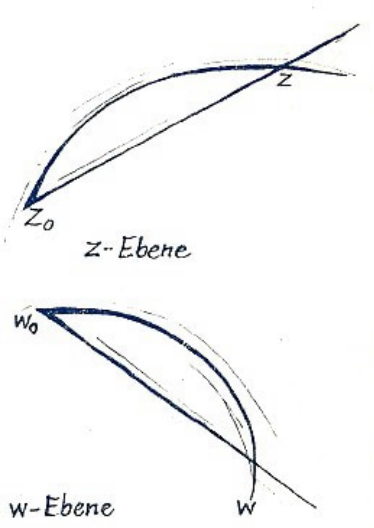
Da nach allem die allgemeine lineare Transformation aus hintereinander ausgeführten Kreisverwandtschaften entsteht, ist sie selber eine Kreisverwandtschaft.

Durch Auflösung nach z entsteht aus der allgemeinen linearen Transformation wiederum eine lineare Transformation,

$$z = \frac{-dw + b}{cw - a}$$

die als $z = I^{-1}(w)$ geschrieben wird. Infolge der Auflösbarkeit nach z bildet also eine lineare Transformation die komplexe Zahlenebene umkehrbar eindeutig in Form einer Kreisverwandtschaft auf sich ab.

Es gilt aber noch mehr. Wir betrachten einen Kreisbogen in der z-Ebene mit dem Anfangspunkt z_0 und den ihm entsprechenden Kreisbogen in der w-Ebene mit dem Anfangspunkt w_0. Zieht man Sekanten aus z_0 nach z bzw. aus w_0 nach w, dann gilt



$$\frac{w - w_0}{z - z_0} \to \left.\frac{dw}{dz}\right|_{z=z_0} = \frac{ad - bc}{(cz_0 + d)^2}$$

Für $z_0 \neq -\frac{d}{c}$ ist die Ableitung endlich.

Die linke Seite ist eine komplexe Zahl, deren Argument gegen das Argument ϑ der festen festen komplexen Zahl $\left.\frac{dw}{dz}\right|_{z=z_0}$ konvergieren muss.

Das Argument der linken Seite ist aber die Differenz der Argumente des Zählers und des Nenners, das sind aber das Argument Φ der Sekante in der w-Ebene und das Argument φ der Sekante in der z-Ebene. Es findet also die Konvergenz $\Phi - \varphi \to \vartheta$ statt.

Nun ist die Grenzlage der Sekanten die entsprechende Tangente, die in der w-Ebene den Winkel Φ_0 und in der z-Ebene den Winkel φ_0 mit der reellen Achse bilden möge, so dass gilt $\Phi \to \Phi_0$ und $\varphi \to \varphi_0$. Daraus folgt

$$\Phi_0 - \varphi_0 = \vartheta \qquad \text{und} \qquad \Phi_0 = \varphi_0 + \vartheta$$

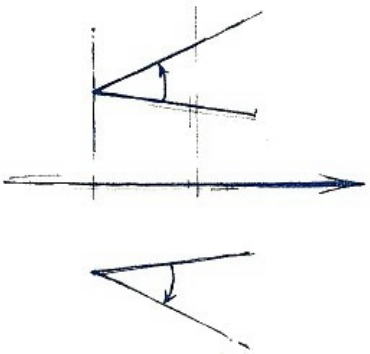
Bilden zwei Kreisbogen in der z-Ebene den Winkel α miteinander, mit anderen Worten, die Tangenten an die, Kreisbogen im Punkt z_0 schließen den Winkel α ein, dann ist dieser gleich der Differenz der beiden Winkel φ' und φ'', welche die beiden Tangenten mit der

reellen Zahlenachse bilden: $\alpha = \varphi'' - \varphi'$.

Durch die lineare Abbildung verwandeln sich aber, wie wir eben gesehen haben, φ' bzw. φ'' in $\Phi' = \varphi' + \vartheta$ bzw. $\Phi'' = \varphi'' + \vartheta$. Daher ist

$$\Phi'' - \Phi' = (\varphi'' + \vartheta) - (\varphi' + \vartheta) = \varphi'' - \varphi' = \alpha$$

Lineare Transformationen sind also winkeltreu.



Da die Spiegelung am Einheitskreis $z' = \dfrac{1}{z}$ aus der speziellen linearen Transformation $w = \dfrac{1}{\overline{z}}$ durch Spiegelung an der reellen Achse hervorgeht, $z' = \overline{w}$, folgt schließlich, dass die Spiegelung am Einheitskreis eine winkeltreue Abbildung mit Umlegung der Winkel ist.
Wird der Punkt z durch die Abbildung nicht verändert, dann heißt er Fixpunkt. Aus $z = \dfrac{az+b}{cz+d}$ folgt aber $cz^2 - (a-d)z - b = 0$, was besagt, dass eine lineare Transformation höchstens zwei Fixpunkte hat, es sei denn, dass $b = c = 0$ und $a = d$ ist.
Dann handelt es sich um die identische Abbildung, die jeden Punkt fest lässt. Wenn danach eine lineare Abbildung drei Punkte fest lässt, muss sie die identische Abbildung sein.

Zwei lineare Transformationen hintereinander ausgeführt, ergeben wieder eine lineare Transformation. Die beiden linearen Transformationen seien

$$w = I_1(z) = \frac{az+b}{cz+d} \qquad \text{und} \qquad u = I_2(z) = \frac{a'z+b'}{c'z+d'}$$

Dann ist

$$u = \frac{(aa'+bc')z + (ab'+bd')}{(ca'+dc')z + (cb'+dd')}$$

Die lineare Transformation, die z erst in w und w dann in u überführt, bezeichnet man mit $I_2 I_1(z)$.

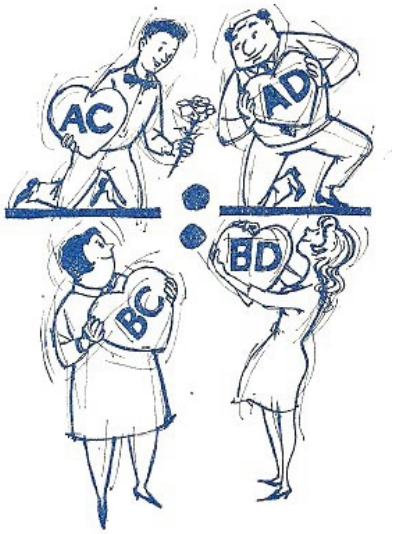
Zum Schluss wollen wir noch nachweisen, dass das Doppelverhältnis von vier Punkten A, B, C, D

$$\frac{AC}{BC} : \frac{AD}{BD}$$

bei linearen Transformationen erhalten bleibt. Den vier Punkten mögen der Reihe nach die komplexen Zahlen z_1, z_1, z_3, z_4 entsprechen. Dann ist der Wert des Doppelverhältnisses

$$\frac{z_3 - z_1}{z_3 - z_2} : \frac{z_4 - z_1}{z_4 - z_2}$$

Darin wollen wir z_4 durch z ersetzen und den so erhaltenen Ausdruck mit $D(z_1, z_2, z_3, z)$ bezeichnen.



Zunächst behaupten wir, dass die Gleichung

$$D(w_1, w_2, w_3, w) = D(z_1, z_2, z_3, z)$$

eine lineare Transformation $w = I(z)$ darstellt. Die rechte wie die linke Seite sind nämlich lineare Transformationen $I_1(z)$ bzw. $I_2(z)$, für die $I_2(w) = I_1(z)$ gelten soll, woraus, auf beide Seiten I_2^{-1} angewandt,

$$w = I_2^{-1} I_1(z)$$

und damit die Behauptung folgt.

Sowohl I_1 als auch I_2 führen die Punkte z_1, z_2, z_3 bzw. w_1, w_2, w_3 der Reihe nach in ∞, 0, 1 über, so dass $I(z) = I_2^{-1} I_1(z)$ die Punkte z_1, z_2, z_3 der Reihe nach in w_1, w_2, w_3 verwandelt.

Wäre nun $I'(z)$ eine zweite lineare Transformation, die z_1, z_2, z_3 ebenfalls in w_1, w_2, w_3 überführt, dann würde $I'^{-1} I(z)$ drei Fixpunkte z_1, z_2, z_3 besitzen, wäre also die Identität, woraus $I'(z) = I(z)$ folgt.
Ist daher $I(z_4) = w_4$, dann muss wegen der Gleichwertigkeit von $w = I(z)$ und

$$D(w_1, w_2, w_3, w) = D(z_1, z_2, z_3, z) \qquad \text{die Gleichung} \qquad D(w_1, w_2, w_3, w_4) = D(z_1, z_2, z_3, z_4)$$

gelten. Das Doppelverhältnis erwies sich damit linearen Transformationen gegenüber invariant.
Es besitzt genau dann einen reellen Wert, wenn die vier Punkte entweder auf einer Geraden oder auf einem Kreis liegen. Denn

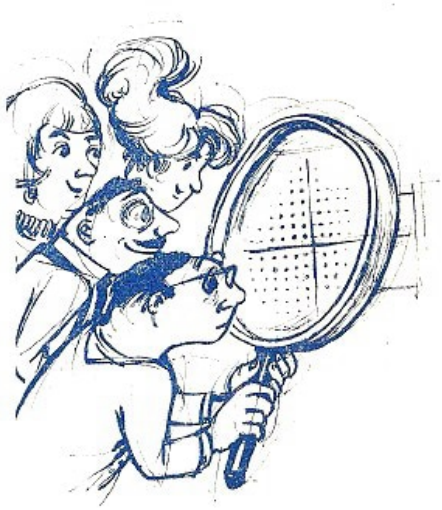
$$\arg \frac{z_3 - z_1}{z_3 - z_2} \qquad \text{bzw.} \qquad \arg \frac{z_4 - z_1}{z_4 - z_2}$$

ist der Winkel, den die Vektoren $z_1 z_3$ und $z_2 z_3$ bzw. $z_1 z_4$ und $z_2 z_4$ miteinander einschließen. Im Falle einer Geraden verschwinden diese Winkel, im Falle eines Kreises beträgt ihre Differenz nach dem Peripheriewinkelsatz ein Vielfaches von 2π.
Der Wert des Doppelverhältnisses in trigonometrischer Darstellung hat also offensichtlich in beiden Fällen einen verschwindenden Imaginärteil.

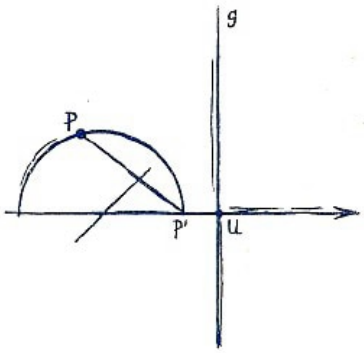
12 Halbkreise als Geraden

Es geht diesmal darum, in der euklidischen Ebene Gebilde anzugeben, die mit Recht als Punkte, Geraden usw. anzusprechen sind, weil sie mit Ausnahme des Parallelenpostulates sämtlichen Axiomen genügen, die für die "gewöhnlichen" Punkte, Geraden usw. gelten. Damit hätten wir dann ein Modell der nichteuklidischen Geometrie, das deren Widerspruchslosigkeit garantiert, weil ein Widerspruch in diesem Modell zugleich einen Widerspruch in der euklidischen Geometrie bedeuten würde.

Der Franzose Poincaré konstruierte 1882 ein Modell in der vertrauten euklidischen Halbebene. Bei seiner Erörterung machen wir von unseren früheren Ausführungen über Spiegelungen und Doppelverhältnisse Gebrauch.
Da es sich dabei um Gebilde der euklidischen Halbebene handelt, werden wir zu unterscheiden haben, ob das Gebilde als Bestandteil der euklidischen Halbebene oder als Punkt, Gerade usw. des Modells betrachtet wird.
Das kann durch ein vorgesetztes E bzw. N geschehen, so dass etwa N-Gerade, wie wir sie gleich definieren werden, einen E-Halbkreis bedeuten kann.



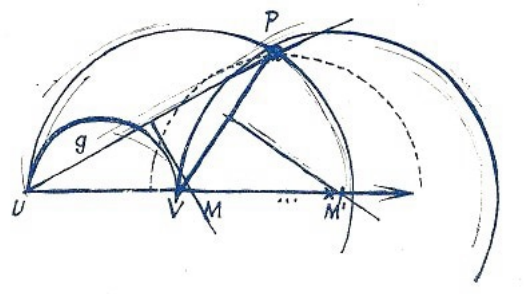
Wir betrachten die gaußsche Zahlenebene. Ihre Punkte mit positivem Imaginärteil sind unsere N-Punkte. Die N-Punkte von E-Geraden, die senkrecht auf der reellen Achse stehen, sowie die N-Punkte von E-Kreisen mit dem Mittelpunkt auf der reellen Achse sind unsere N-Geraden. Die E-Kreise, deren Teil eine N-Gerade bildet, sind also zur reellen Achse orthogonal.



Man sieht sofort, dass es zu einer N-Geraden g und einem nicht daraufliegenden N-Punkt unendlich viel N-Geraden gibt, die g nicht schneiden. Ist nämlich g Teil einer E-Geraden, dann sind es die N-Geraden, die man gewinnt, wenn im Schnittpunkt des Mittellotes auf der Strecke PP' ein E-Kreis errichtet wird.
Dabei ist P' ein beliebiger E-Punkt auf der E-Strecke, die durch den E-Fußpunkt von P und den Schnittpunkt der E-Trägergeraden von g mit der reellen Achse begrenzt wird.

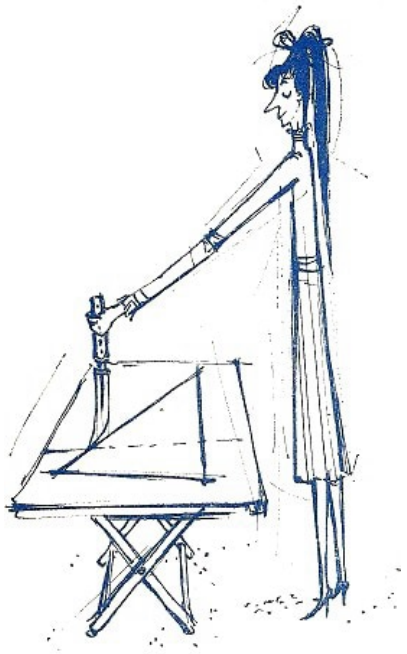
Wenn dagegen g Teil eines E-Kreises ist, dann bestimmen die Mittellote der E-Strecken PU bzw. PV durch ihre E-Schnittpunkte mit der reellen Achse die Mittelpunkte M und M' von zwei E-Kreisen, die N-Parallelen zu g liefern.
Dazwischen verlaufen alle übrigen N-Parallelen, die auf E-Kreisen liegen mit Mittelpunkten auf der Strecke MM'.

Zunächst überprüfen wir die Axiome der Verknüpfung. Auf g liegen unendlich viel N-Punkte und durch einen N-Punkt gehen unendlich viel N-Geraden. Auch das dritte Verknüpfungsaxiom ist erfüllt, denn durch zwei N-Punkte geht stets genau eine N-Gerade. Sie ist entweder die E-Trägergerade der beiden Punkte, oder sie wird mit Hilfe der uns schon geläufigen Konstruktion eines Mittellotes bestimmt.



Die E-Anordnung von Punkten übernehmen wir für die N-Anordnung. Dann folgt unmittelbar, dass alle Anordnungsaxiome erfüllt sind, insbesondere auch das Axiom von Pasch:
Wird eine Seite eines Dreiecks von einer Geraden geschnitten, die durch keinen Eckpunkt geht, dann schneidet die Gerade genau eine der beiden übrigen Seiten.

Als N-Winkel zweier N-Geraden und ihre Breite übernehmen wir den E-Winkel und seine Breite, unter dem sich die betreffenden E-Geraden und E-Halbkreise schneiden.



Das Spiegeln einer N-Geraden an einer N-Geraden bedeutet eine Spiegelung entweder an einer E-Geraden oder an einem E-Kreis. Dabei bleiben die Winkel erhalten, so dass unsere rechten Winkel wieder in rechte Winkel übergehen. Das Ergebnis des Spiegelns ist folglich wieder eine N-Gerade.
Damit sind wir in der Lage, Kongruenzen zu definieren. In der euklidischen Ebene heißen zwei Strecken kongruent, wenn sie durch Bewegen der starren Ebene zur Deckung gebracht werden können. Die Rolle dieser Bewegungen übernehmen in unserem Modell die Spiegelungen, sei es an einem E-Kreis oder an einer E-Geraden.

Wie steckt man nun auf einem vorgegebenen N-Strahl s von dessen Anfangspunkt A' die N-Strecke AB ab?
Zunächst kann A durch Spiegelung in A' überführt werden. Dabei sind zwei Fälle zu unterscheiden. Besitzen A und A' von der reellen Achse denselben E-Abstand, dann spiegelt man am E-Mittellot der E-Strecke AA'. Andernfalls schneidet die E—Gerade durch A und A' die reelle Achse im Punkt M.
Schlägt man um M einen E-Kreis mit dem Radius $\sqrt{MA \cdot MA'}$, dann geht durch Spiegelung an diesem A in A' und der Punkt B in B'' über. A liegt dabei auf s, B'' aber nicht.
Eine weitere Spiegelung an der Winkelhalbierenden des Winkels mit den Schenkeln s und der N-Geraden durch A' und B'' lässt A' unverändert, während sie B'' in B' auf s überführt.

Es fragt sich nur, ob man durch andere Spiegelungen nicht eine davon verschiedene Lösung erhalten könnte. Dem ist nicht so.
Liegt zunächst die N-Strecke AB auf einem E-Kreis, der die reelle Achse in den Punkten U und V schneidet, dann betrachte man das Doppelverhältnis $D(U, V, A, B)$. Bei Spiegelungen bleibt es, wie wir von früher her wissen, invariant. Dadurch ist aber B' eindeutig festgelegt.
Liegt dagegen die N-Strecke AB auf einer E-Geraden, welche die reelle Achse im Punkt U schneidet, dann betrachtet man das einfache Verhältnis $AU : BU$, das aus $D(U, V, A, B)$ im Sinne der Auffassung von Geraden als Kreisen mit unendlichem Radius für $V \to \infty$

hervorgeht, weil dann $\frac{VA}{VB} \to 1$.

Bezeichnet man die Kongruenz von N-Strecken durch Gleichheitszeichen (für N-Winkel besteht ja tatsächliche Gleichheit), dann lässt sich der erste N-Kongruenzsatz wie folgt einsehen.
Es sei $AB = A'B'$, $AC = A'C'$, $\angle BAC = \angle B'A'C'$.
Man führt zunächst die N-Strecke AB durch Spiegelungen in die N-Strecke $A'B'$ über. Da die Winkelbreiten dabei erhalten bleiben, kommt die N-Strecke AC auf den N-Strahl $A'C'$ zu liegen, oder sie liegt auf der anderen Seite von $A'B'$.

Im letzteren Fall führt sie eine weitere Spiegelung an $A'B'$ auf den N-Strahl $A'C'$. Wegen $AC = A'C'$ koinzidieren dann die Punkte C und C'. Die Dreiecke ABC und $A'B'C'$ erweisen sich damit in allen übrigen Stücken ebenfalls kongruent.
Um dabei eine Ähnlichkeit mit der euklidischen Geometrie herzustellen, sei darauf hingewiesen, dass deren Bewegungen - das Aufeinanderlegen - durch Spiegelungen erzeugt werden können.

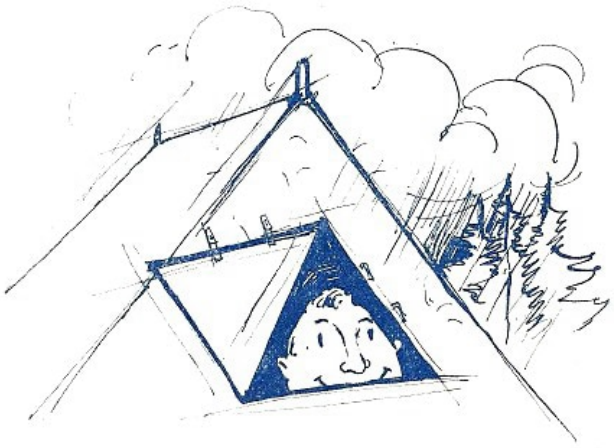
Das N-Winkelmaß haben wir bereits eingeführt. Es ist das E-Maß des Winkels. Es fehlt aber noch ein N-Längenmaß für N-Strecken. Zunächst muss es bei N-Bewegungen, also bei E—Spiegelungen, unverändert bleiben. Es muss aber auch noch additiv sein. Darunter ist folgendes zu verstehen.
Liegen zunächst die N-Punkte A, B, C auf einem E-Kreis, der die reelle Achse in den Punkten U und V schneidet, in der N-Reihenfolge, die ja zugleich E-Reihenfolge ist, U, A, B, C, V, dann muss die Summe der N-Längen der N-Strecken AB und BC gleich der N-Länge der N-Strecke AC sein. Nun genügt das Doppelverhältnis $D(U, V, A, B)$ zwar der ersten Forderung, nicht aber der zweiten, es sei denn wir nehmen den Logarithmus $\ln D(U, V, A, B)$.
Dann ist

$$\begin{aligned}\ln D(U,V,A,B) + \ln D(U,V,B,C) &= \left(\ln \frac{UA}{VA} + \ln \frac{UB}{VB}\right) + \left(\ln \frac{UB}{VB} - \ln \frac{UC}{VC}\right) \\ &= \ln \frac{UA}{VA} - \ln \frac{UC}{VC} = \ln D(U,V,A,B)\end{aligned}$$

Handelt es sich dagegen um N-Punkte auf einer E-Geraden, welche die reelle Achse im Punkt U schneidet, und ist die Reihenfolge der Punkte darauf U, A, B, C, dann gilt

$$\ln \frac{UA}{UB} + \ln \frac{UB}{UC} = (\ln UA - \ln UB) + (\ln UB - \ln UC) = \ln \frac{UA}{UC}$$

Unsere N-Länge erwies sich damit in beiden Fällen additiv. So viel möge genügen, um den Sinn der scheinbar unsinnigen Überschrift, welche über dieser Plauderei steht, aufzuhellen.



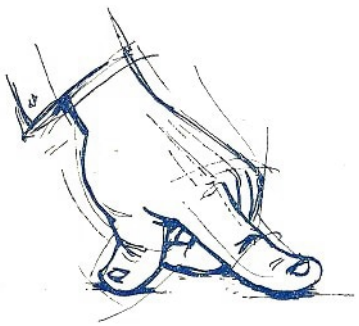
13 Makellos

Nachdem Klein 1871 ein Modell der hyperbolischen Geometrie innerhalb der projektiven Geometrie und Poincaré innerhalb der euklidischen Geometrie konstruiert hatte, war es eine folgerichtige Entwicklung, wenn Hilbert um die Jahrhundertwende ein Modell der euklidischen Geometrie innerhalb der Arithmetik angab.

Bei einem Modell geht es richtig gesehen nicht so sehr darum, eine andere Interpretation für Grundbegriffe und Relationen zu schaffen, als vielmehr um die Einsicht, dass ein Widerspruch, der bei der ursprünglichen Interpretation auftreten würde, sich auf die neue Interpretation überträgt und damit zu einem Widerspruch im System führt, in dem die neue Interpretation erfolgte.

Bei dem Modell von Poincaré würde das besagen:
Wenn die hyperbolische Geometrie nicht widerspruchsfrei ist, dann gibt es in der euklidischen Geometrie Widersprüche. Und weiter bei dem arithmetischen Modell der euklidischen Geometrie, die Hilbert angab, würden Widersprüche in der euklidischen Geometrie Widersprüche in der Arithmetik aufzeigen. Soweit die Widerspruchsfreiheit der Arithmetik feststeht, ist damit zugleich die Widerspruchsfreiheit der euklidischen Geometrie gesichert.
Denkt man weiterhin an das Modell von Poincaré, dann ist zugleich die Widerspruchsfreiheit der hyperbolischen Geometrie gewährleistet. Wie man sieht, geht es jedesmal um die Zurückführung auf die Widerspruchsfreiheit einer bereits als makellos geltenden Disziplin.

Die analytische Geometrie weist den Weg zum arithmetischen Modell der euklidischen Geometrie. Dementsprechend nennen wir ein geordnetes Paar (x, y) von reellen Zahlen einen Punkt. Für $x \neq y$ ist also $(x, y) \neq (y, x)$. x und y heißen Koordinaten des Punktes, der gelegentlich durch einen großen Buchstaben bezeichnet wird.



Die Gesamtheit aller Punkte, deren Koordinaten die lineare Gleichung

$$px + qy + r = 0$$

erfüllen, heißt eine Gerade, wenn p und q nicht gleichzeitig verschwinden. Dieselben Punkte genügen auch noch der Gleichung $cpx + cqy + cr = 0$, mit c eine beliebige reelle Zahl bezeichnet, die hier die Rolle einer multiplikativen Konstanten spielt.

Genügen die Koordinaten eines Punktes der Gleichung einer Geraden, dann sagt man, der Punkt liegt auf dieser Geraden, oder auch, diese Gerade geht durch den Punkt.
Durch geeignete Wahl von p, q, r kann man erreichen, dass unendlich viel Geraden durch einen vorgegebenen Punkt gehen. Außerdem besitzt die Gleichung einer Geraden unendlich viel Lösungen (x, y), so dass unendlich viel Punkte auf einer vorgegebenen Geraden liegen. Weiter gibt es auch unendlich viel Punkte, die nicht auf ihr liegen, deren Koordinaten also die linke Seite nicht verschwinden lassen.
Die ersten beiden Axiome der Verknüpfung in der Terminologie von Hilbert sind damit erfüllt.

Das dritte Axiom dieser Gruppe verlangt, dass durch zwei verschiedene Punkte (x_1, y_1) und (x_2, y_2) genau eine Gerade geht. Das trifft ebenfalls zu, denn diese Gerade ist bestimmt durch

$$p = y_1 - y_2, \qquad q = x_1, x_2, \qquad r = x_1y_2 - x_2y_1$$

Aus $px_1 + qy_1 + r = 0$ folgt nämlich

$$q = \frac{-r - px_1}{y_1}$$

Wird dieser Ausdruck in $px_2 + qy_2 + r = 0$ eingesetzt, so ergibt sich

$$p(x_1y_2 - x_2y_1) = r(y_1 - y_2)$$

Nun ist die Gleichung der Geraden nur bis auf eine multiplikative Konstante bestimmt, so dass man eine der Koeffizienten p, q, r nach Belieben vorschreiben darf. Wir entscheiden uns für r, und zwar soll r den Wert $x_1y_2 - x_2y_1$ haben. Dann folgt aus der letzten Gleichung $p = y_1 - y_2$. Ähnlich gewinnt man $q = x_1 - x_2$.Aus

$$p(x_0 + qt) + q(y_0 - pt) + r = px_0 + qy_0 + r$$

folgt, dass mit (x_0, y_0) zugleich $(x_0 + qt, y_0 - pt)$ auf der Geraden liegt. Der Parameter t darf dabei jeden reellen Wert annehmen, $-\infty < t < \infty$.
Liegt $x_*, y_*)$ ebenfalls auf der Geraden, und zieht man die Gleichung $px_* + q_y * + r = 0$ von der Gleichung $px_0 + qy_0 + r = 0$ ab, dann hebt sich r weg, und es gilt $-p(x_0 - x_*) = q(y_0 - y_*)$. Setzt man für

$$\frac{x_0 - x_*}{q} = t_* \qquad \text{dann folgt} \qquad \frac{y_0 - y_*}{-p} = t_*$$

Daher ist

$$x_* = x_0 + qt_* \qquad ; \qquad y_* = y_0 - pt_*$$

oder in Worten: Die Parameterdarstellung erfasst alle Punkte der Geraden.
Da in den beiden Parameterdarstellungen

$$x = x_0 + qt, \qquad y = y_= - pt \qquad \text{und}$$

$$x = x_* + q\tau, \qquad y = y_* + p\tau$$

derselben Geraden $t = t_* - \tau$ ist, entsprechen sich $t_1 < t_2 < t_3$ und $\tau_1 < \tau_2 < \tau_3$ gegenseitig, so dass man die Punkte mit Hilfe einer beliebigen der Parameterdarstellungen ordnen kann und die Relation "zwischen" erhalten bleibt.

Die Punkte mit $t > 0$ bilden einen, die Punkte mit $t < 0$ einen zweiten Strahl, die beide durch den Punkt (x_0, y_0) erzeugt werden.
Die Gerade $px + py + r = 0$ bestimmt zwei "Halbebenen". In der einen ist die linke Seite positiv, in der anderen negativ. Nach dem Zwischenwertsatz für stetige Funktionen folgt sofort, dass die Verbindungsgerade von zwei Punkten aus je einer Halbebene die Gerade schneiden muss.

Damit sind wir bei der zweiten Gruppe von Axiomen angelangt, bei denen der Anordnung, die auf Pasch zurückgehen. Obschon sie zum Aufbau der Geometrie unentbehrlich sind, sucht man sie bei Euklid vergebens.
Sind (x_1, y_1) und (x_2, y_2) zwei Punkte auf einer Geraden, dann folgt aus

$$px_1 + qy_1 + r = 0 \qquad \text{und} \qquad px_2 + qy_2 + r = 0$$

wenn die erste Gleichung mit λ, die zweite mit $1 - \lambda$ multipliziert und zur ersten addiert wird, die Gleichung

$$p[\lambda x_1 + (1 - \lambda)x_2] + q[\lambda y_1 + (1 - \lambda)y_2] + r = 0$$

aus der hervorgeht, dass für $-\infty < \lambda < \infty$ jeder Punkt

$$(\lambda x_1 + (1 - \lambda)x_2; \lambda y_1 + (1 - \lambda)y_2)$$

auf der Verbindungsgeraden von (x_1, y_1) und (x_2, y_2) liegt. Jeder Punkt (x, y) der Verbindungsgeraden kann aber auch wirklich so dargestellt werden. Man hat nur für $\frac{x-x_2}{x_1-x_2}$ λ zu setzen. Daraus folgt für x der Wert

$$\lambda x_1 + (1 - \lambda)x_2$$

Wird dieser Wert für x in die Gleichung der Geraden eingesetzt, so folgt durch Vergleich mit der letzten Gleichung

$$qy = a[\lambda y_1 + (1 - \lambda)y_2] \qquad \text{so dass} \qquad y = \lambda y_1 + (1 - \lambda)y_2$$

gilt. Insbesondere liefern die Werte von λ zwischen 0 und 1 die Strecke mit den Endpunkten (x_1, y_1) und (x_2, y_2). Denn mit

$$x + qt_1 = \lambda x_1 + (1 - \lambda)x_2 \qquad \text{folgt} \qquad t = (1 - \lambda)\frac{x_2 - x_1}{q}$$

Für die rechte Seite wollen wir $t(\lambda)$ schreiben. Nach unserer Definition der Relation "zwischen" kommt es bei der Strecke mit den Endpunkten (x_1, y_1) und (x_2, y_2) auf diejenigen Werte von 1 an, die zwischen $t(1) = 0$ und $t(0) = \dfrac{x_2 - x_1}{q}$ liegen. Diese erhält man aber, wenn man für λ alle Werte zwischen 0 und 1 einsetzt.

Werden die Punkte (x_1, y_1) und (x_2, y_2) mit P_1 und P_2 bezeichnet, dann schreibt man für die Gerade, die diese beiden Punkte miteinander verbindet, gebildet von den Punkten

$$(\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2) \qquad -\infty < \lambda < \infty$$

abgekürzt

$$\lambda P_1 + (1 - \lambda)P_2 \qquad -\infty < \lambda < \infty$$

Eines der Anordnungsaxiome trägt den Namen von Pasch. Um die Gültigkeit dieses Axioms in unserem arithmetischen Modell nachzuweisen, sei $px + qy + r = 0$ eine Gerade g, die durch keinen der Punkte $A = (a, a')$, $B = (b, b')$, $C = (c, c')$ geht, so dass keiner der Ausdrücke

$$pa + qa' + r \qquad ; \qquad pb + qb' + r \qquad ; \qquad pc + qc' + r$$

verschwindet. Ihre Werte seien der Reihe nach mit w, w', w'' bezeichnet. Die Schnittpunkte von g mit den Verbindungsgeraden AB, BC, CA seien, soweit vorhanden,

$$\lambda A + (1 - \lambda)B; \qquad \mu B + (1 - \mu)C; \qquad \nu C + (1 - \nu)A$$

Werden die Gleichungen

$$w = pa + qa' + r \qquad ; \qquad w' = pb + qb' + r$$

mit λ bzw. $1 - \lambda$ multipliziert und addiert, so folgt

$$\lambda w + (1 - \lambda)w' = p[\lambda a + (1 - \lambda)b] + q[\lambda a' + (1 - \lambda)b'] + r$$

Die rechte Seite verschwindet aber, da der Punkt

$$(\lambda a + (1 - \lambda)b, \lambda a' + (1 - \lambda)b')$$

auf g liegt. So erhalten wir schließlich

$$\lambda w + (1 - \lambda)w' = 0$$

Ähnlich gewinnt man die Gleichungen, die eben von unserer Sekretärin abgeschrieben werden:
Nun geht das Axiom von Pasch von der Annahme aus, dass g eine der drei Verbindungsstrecken AB, BC, CA trifft. Es sei dies die Strecke AB.

Das Axiom behauptet dann, dass genau eine der beiden anderen Strecken, also entweder BC oder CA, von g ebenfalls getroffen wird. Dabei geht g durch keine der Ecken A, B, C.
Ist zunächst $w' = w''$, dann kann

$$\mu w' + (1 - \mu)w'' = w' \neq 0$$

für kein μ verschwinden, dagegen genügt dann $\nu = 1 - \lambda$ der Gleichung

$$\nu w'' + (1 - \nu)w = (1 - \lambda)w' + \lambda w = 0$$

Weil dabei mit $0 < \lambda < 1$ zugleich $0 < \nu = 1 - \lambda < 1$ ausfällt, schneidet g die Strecke CA. Für $w = w''$ würde g die Strecke BC schneiden.

Wenn aber w, w', w'' paarweise verschieden sind, dann haben die mit ihnen angesetzten Gleichungen alle drei die Lösungen

$$\lambda = \frac{w'}{w' - w} \qquad ; \qquad \mu = \frac{w''}{w'' - w'} \qquad ; \qquad \nu = \frac{w}{w - w''}$$

Da weiter nach unserer Annahme $0 < \lambda < 1$ sein soll, trifft zugleich $0 < 1 - \lambda < 1$, so dass der Bruch $\dfrac{1-\lambda}{\lambda}$ positiv ausfällt. Nun ist

$$\frac{1-\lambda}{\lambda} \cdot \frac{1-\mu}{\mu} \cdot \frac{1-\nu}{\nu} = \left(-\frac{w}{w'}\right)\left(-\frac{w}{w''}\right)\left(-\frac{w''}{w}\right) = -1$$

Damit die linke Seite negativ ist, muss entweder der zweite oder der dritte Faktor positiv sein, da der erste Faktor bestimmt positiv ist.

Das besagt aber, dass eine der beiden Zahlen μ oder ν zwischen 0 und 1 liegt, die andere jedoch nicht. Damit ist das Axiom von Pasch aus dem Jahr 1882 in vollem Umfang bewiesen.
Da dieses Axiom auch für die nichteuklidische Geometrie unerlässlich ist, war das Vorgehen ihrer Entdecker um 1830 nicht einwandfrei. Selbst ein Gauß war dabei nach Zeugnis der Jahreszahlen nicht unfehlbar.

Das letzte der Anordnungsaxiome verlangt, dass zwei aneinanderliegende Strecken auf einer Geraden wieder eine Strecke bilden. Das folgt sofort aus der Parameterdarstellung der Geraden.

Der Abstand von zwei Punkten (x_1, y_1) und (x_2, y_2) wird durch die nichtnegative Quadratwurzel aus

$$(x_1 - x_2)^2 + (y_1 - y_2)^2$$

definiert. Er heißt auch Länge der Strecke mit den Endpunkten (x_1, y_1) und (x_2, y_2). Wird die Parameterdarstellung der Geraden etwas abgeändert als

$$(x_0 + at, y_0 + bt)$$

geschrieben, dann beträgt der Abstand der beiden Punkte A und A', die den Parameterwerten t und t' entsprechen,

$$\sqrt{(at - at')^2 + (bt - bt')^2} = (t' - t)\sqrt{a^2 + b^2}$$

Ist daher A'' ein weiterer Punkt auf der Geraden, der dem Parameterwert t'' entspricht, dann ist nach

$$(t' - t)\sqrt{a^2 + b^2} + (t'' - t')\sqrt{a^2 + b^2} = (t'' - t)\sqrt{a^2 + b^2}$$

die Länge von AA'' die Summe der Längen von AA' und $A'A''$. Zwei Strahlen s und s'

$$(x_0 + at, y_0 + bt) \qquad \text{und} \qquad (x_0 + at', y_0 + bt') \quad t > 0$$

mit dem gemeinsamen Anfangspunkt (x_0, y_0) bilden einen Winkel (ss') oder auch $(s's)$. Sind g bzw. g' die Geraden, denen s bzw. s' angehören, dann bilden die Punkte, die auf derselben Seite von g liegen wie s' und zugleich auf derselben Seite von g' wie s, das Innere des Winkels (ss').

Da (x_0+at, y_0+bt) mit beliebigem t alle Punkte der Geraden g liefert, folgt aus $x-x_0 = at$ und $y - y_0 = bt$ nach Multiplikation mit b bzw. $-a$ und Addition als Gleichung für g

$$b(x - x_0) - a(y - y_0) = 0$$

und entsprechend als Gleichung für g'

$$b'(x - x_0) - a'(y - y_0) = 0$$

Die Punkte $(x_+a't, y_0 + b't)$ in die vorletzte Gleichung eingesetzt, ergeben das richtige Vorzeichen der Halbebene, in der s' liegt, so dass $ba't - ab't$ und $b(x - x_0) - a(y - y_0)$ dasselbe Vorzeichen haben. Da t dabei positiv ist, muss der Quotient

$$\frac{b(x - x_0) - a(y - y_0)}{ba' - ab'} = \sigma$$

positiv sein. Ähnlich folgt, dass auch der Quotient

$$\frac{b'(x - x_0) - a'(y - y_0)}{b'a - a'b} = \rho$$

positiv ist. Für das Innere des Winkels müssen beide Quotienten zugleich positiv ausfallen.

Aus der vorletzten Gleichung folgt

$$y - y_0 = \frac{b(x - x_0) - \sigma(ba' - ab')}{a}$$

Wird das in die letzte Gleichung eingesetzt, erhalten wir

$$ab'(x - x_0) - a'b(x - x_0) + \sigma a'(ba' - ab') = \rho a(b'a - ab')$$

also

$$-(ba' - ab')(x - x_0) + \sigma a'(ba' - ab') = \rho a(b'a - ab')$$

und nach Kürzen durch $ba'' - ab'$ folglich

$$x = x_0 + a\rho + a'\sigma, \quad \rho < 0, \sigma > 0 \qquad (*)$$

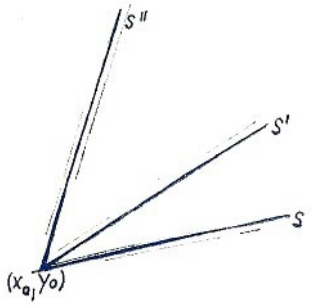
entsprechend auch

$$y = y_0 + b\rho + b'\sigma, \quad \rho < 0, \sigma > 0 \qquad (*)$$

Umgekehrt liegt für alle $\rho > 0$, $\sigma > 0$ jeder dieser Punkte im Inneren von (ss'), weil dann beide Quotienten

$$\frac{b(x - x_0) - a(y - y_0)}{ba' - a'b} = \frac{(ba - ab)\rho + (ba' - a'b)\sigma}{ba' + a'b} = \sigma$$
$$\frac{b'(x - x_0) - a'(y - y_0)}{ba' - a'b} = \frac{(b'a - a'b)\rho + (b'a' - a'b')\sigma}{ba' + a'b} = \rho$$

positiv ausfallen. Das Formelpaar $(*)$ stellt also genau das Innere des Winkels (ss') dar.



Wir betrachten noch einen dritten Strahl s''

$$(x_0 + a''t, y_0 + b''t) \qquad t > 0$$

mit demselben Anfangspunkt (x_0, y_0) und nehmen an, dass s' im Inneren des Winkels (ss'') verläuft. Dann müssen sich, wie wir eben erkannten, alle Punkte $(x_0 + a't, y_0 + b't)$ mit $t > 0$, also insbesondere für $t = 1$ der Punkt $(x_0 + a', y_= + b')$, in der Form $(x_0 + a\rho + a''\sigma, y_0 + b\rho + b'\sigma)$ mit geeigneten positiven ρ^* und σ^* darstellen lassen. Das besagt aber

$a' = a\rho^* + a''\sigma^*$ und $b' = b\rho^* + b''\sigma^*$. Nun stellen

$$\begin{aligned} (x_0 + at, y_0 + bt) \qquad &\text{und} \qquad (x_0 + a\rho^* t, y_+ b\rho^* t) \quad t > 0 \\ (x_0 + a''t, y_0 + b''t) \qquad &\text{und} \qquad (x_0 + a''\sigma^* t, y_+ b''\sigma^* t) \quad t > 0 \end{aligned}$$

denselben Strahl s bzw. s'' dar, so dass wir statt a und b bzw. a'' und b'' auch $a\rho^*$ und $b\rho^*$ bzw. $a''\sigma^*$ und $b''\sigma^*$ verwenden dürfen. Wir wollen für diese neuen Werte die alten Bezeichnungen a und b bzw. a'' und b'' gebrauchen.
Das Innere von (ss') besteht dann aus allen Punkten mit den Koordinaten

$$x_0 + a\kappa + (a + a'')\lambda, \quad y_0 + b\kappa + (b + b'')\lambda; \qquad \kappa > 0, \lambda > 0 \qquad (+)$$

das Innere von $(s's'')$ aus allen Punkten mit den Koordinaten

$$x_0 + (a + a'')\mu + a''\nu, \quad y_0 + (b + b'')\mu + b''\nu; \qquad \mu > 0, \nu > 0 \qquad (++)$$

das Innere von (ss'') aus allen Punkten mit den Koordinaten

$$x_0 + a\rho + a''\sigma, \quad y_0 + b\rho + b''\sigma; \qquad \rho > 0, \sigma > 0 \qquad (+++)$$

Wie man unmittelbar erkennt, befinden sich alle Punkte aus $(+)$ und $(++)$ unter denen von $(+++)$, zu denen noch die Punkte auf s' gehören, die aus den Punkten von $(+++)$ durch $\rho = \sigma$ hervorgehen.

Damit ist auch das Axiom erfüllt, das verlangt, dass zwei Winkel aneinandergelegt einen Winkel erzeugen, dessen Inneres aus dem Inneren der beiden erzeugenden Winkel und ihrem gemeinsamen Schenkel besteht.

Um ein Winkelmaß für (ss') einzuführen, erinnern wir uns an die Ungleichung von Cauchy

$$(aa' + bb')^2 \geq (a^2 + b^2)(a'^2 + b'^2)$$

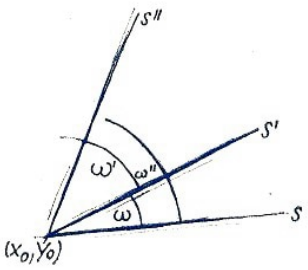
Daraus folgt mit den positiven Quadratwurzeln sofort

$$-1 \leq \frac{aa' + bb'}{\sqrt{a^2 + b^2}\sqrt{a'^2 + b'^2}} \leq +1$$

so dass man die Bruch $\cos\omega$ gleichsetzen und damit für (ss') ein Winkelmaß ω einführen kann. Es folgt dann zunächst

$$\frac{aa' + bb'}{\sqrt{a^2 + b^2}\sqrt{a'^2 + b'^2}} = \sin\omega$$

weil diese Aussage mit der Gleichung $\cos^2\omega + \sin^2\omega = 1$ gleichwertig ist, die für die beiden Quotientenausdrücke zutrifft.
Da auf Grund unserer Erklärung Winkel nie größer als π ausfallen, ist das im Ausdruck für den Sinus dadurch zu berücksichtigen, dass der Zähler als Absolutwert eingesetzt wird.



Der endgültige Ausdruck lautet daher

$$\sin\omega = \frac{|ab' - a'b|}{\sqrt{a^2 + b^2}\sqrt{a'^2 + b'^2}}$$

Sind s, s', s'' die drei Strahlen von vorhin und bezeichnen ω, ω', ω'' die Maße der Winkel (ss'), $(s's'')$, $(ss'') = (s''s)$, dann gilt

$$\cos\omega = \frac{a(a+a'') + b(b+b'')}{\sqrt{a^2+b^2}\sqrt{(a+a'')^2 + (b+b'')^2}} \quad ; \quad \sin\omega = \frac{|ab'' - a''b|}{\sqrt{a^2+b^2}\sqrt{(a+a'')^2 + (b+b'')^2}}$$

$$\cos\omega' = \frac{(a+a'')a'' + (b+b'')b''}{\sqrt{a''^2+b''^2}\sqrt{(a+a'')^2 + (b+b'')^2}} \quad ; \quad \sin\omega' = \frac{|ab'' - a''b|}{\sqrt{a''^2+b''^2}\sqrt{(a+a'')^2 + (b+b'')^2}}$$

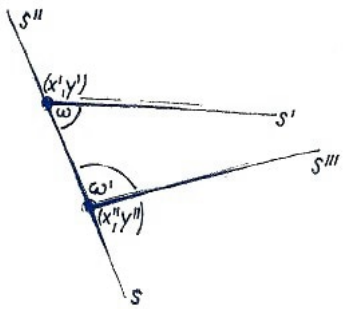
$$\cos\omega'' = \frac{a'' + b''}{\sqrt{a''^2+b''^2}\sqrt{a^2+b^2}} \quad ; \quad \sin\omega'' = \frac{|ab'' - a''b|}{\sqrt{a''^2+b''^2}\sqrt{a^2+b^2}}$$

Berechnet man damit die Ausdrücke $\sin\omega\cos\omega' + \cos\omega\sin\omega'$ und $\cos\omega\cos\omega' - \sin\omega\sin\omega'$ dann ergibt sich deren Gleichheit mit den Ausdrücken für $\sin\omega''$ und $\cos\omega''$. Mit Hinblick auf die Additionstheoreme für den Sinus und Cosinus besagt das aber

$$\sin(\omega + \omega') = \sin\omega'' \qquad \text{und} \qquad \cos(\omega + \omega') = \cos\omega''$$

woraus $\omega + \omega' = \omega''$ folgt, oder in Worten: Die Maße aneinanderliegender Winkel addieren sich.

Es wären nur noch einige wenige Axiome nachzuprüfen.
Über diese Axiome selbst orientiert man sich am besten an deren Zusammenstellung am Schluss dieses Buches oder in dem Buch von Perron "Nichteuklidische Elementargeometrie der Ebene", in welchem vor allem die hyperbolische Geometrie behandelt wird. Bis auf eines der ausstehenden Axiome erfordert das Nachprüfen lediglich Rechnungen, die aus der analytischen Geometrie hinlänglich bekannt sein dürften.



Allein das Parallelenaxiom werde hier noch behandelt:
(x', y') und (x'', y'') seien zwei Punkte. Der Strahl s möge vom ersten Punkt ausgehend, den zweiten Punkt treffen:

$$s = (x' + (x'' - x')t, y' + (y'' - y')t) \quad t > 0$$

der Strahl s' im ersten Punkt beginnen:

$$s' = (x' + a't, y' + bt) \quad t > 0$$

der Strahl s'' vom zweiten Punkt ausgehend, den ersten Punkt treffen:

$$s'' = (x'' + (x' - x'')t, y'' + (y' - y'')t) \quad t > 0$$

schließlich der Strahl s''' im zweiten Punkt beginnen:

$$s''' = (x'' + a''t, y'' + b''t) \quad t > 0$$

Die Strahlen s' und s''' sollen auf derselben Seite der Verbindungsgeraden durch (x', y') und (x'', y'') mit der Gleichung

$$(y' - y'')x + (x'' - x')y + x'y'' - x''y' = 0$$

verlaufen. Damit wird verlangt (für x und y die Punkte von s' und s'' eingesetzt), dass die linke Seite

$$(y' - y'')(x' + a't) + (x'' - x')(y' + b't) + x'y'' - x''y'$$
$$\text{bzw.} \qquad (y' - y'')(x'' + a''t) + (x'' - x')(y'' + b''t) + x'y'' - x''y'$$

für alle positiven t dasselbe Vorzeichen haben muss. Ausmultipliziert verwandeln sich diese Ausdrücke in

$$[a'(y' - y'') + b'(x'' - x')]t \qquad \text{bzw.} \qquad [a''(y' - y'') + b''(x'' - x')]t$$

Die mit p' und p'' bezeichneten eckigen Klammern sollen also dasselbe Vorzeichen haben. Bezeichnet man die Winkel (ss') bzw. $(s''s''')$ mit ω bzw. ω', so gilt

$$\cos\omega = \frac{(x'' - x')a' + (y'' - y')b'}{\sqrt{a'^2 + b'^2}\sqrt{(x'' - x')^2 + (y'' - y')^2}}$$
$$\cos\omega' = \frac{a''(x' - x'') + b''(y' - y'')}{\sqrt{a''^2 + b''^2}\sqrt{(x' - x'')^2 + (y' - y'')^2}}$$

Zu beweisen haben wir: Für $\omega + \omega' < \pi$ schneiden sich die Strahlen s' und s'''.
Nun folgt für $\alpha < \pi$ und $\beta < \pi$ aus $\alpha < \beta$ die Ungleichung $\cos\alpha > \cos\beta$. Die Ungleichung $\omega + \omega' < \pi$ in der Form $\omega < \pi - \omega'$ geschrieben, liefert

$$\cos\omega > \cos(\pi - \omega') = -\cos\omega'$$

womit die Forderung $\omega + \omega' < \pi$ die Gestalt $\cos\omega + \cos\omega' > 0$ angenommen hat.
Zunächst behaupten wir, dass $a'b'' - a''b'$ nicht verschwinden kann. Sonst wäre $a'' = a'q$ und $b'' = b'q$, und weil p' und p'' dasselbe Vorzeichen haben, wegen $\dfrac{p''}{p'} = q$ der Faktor q positiv. Werden die Werte $a'q$ bzw. $b'q$ für a'' und b'' in die Formeln für $\cos\omega'$ eingesetzt, würde sich

$$\cos\omega' = -\cos\omega$$

also $\cos\omega + \cos\omega' = 0$ ergeben, im Widerspruch zu $\cos\omega + \cos\omega' > 0$.
Die beiden Gleichungen, durch die p' und p'' eingeführt wurden, können dann nach $x' - x''$ und $y' - y''$ mit dem Ergebnis

$$x' - x'' = \frac{a''p' - a'p''}{a'b'' - a''b'} \qquad , \qquad y' - y'' = \frac{b''p' - b'p''}{a'b'' - a''b'}$$

aufgelöst werden. Durch Einsetzen dieser Ausdrücke in die Formeln für $\cos\omega$ und $\cos\omega'$, folgt nach umständlicher, aber elementarer Rechnung

$$\cos\omega + \cos\omega' = \frac{p'}{a'b'' - a''b'} \cdot \frac{\sqrt{a''^2 + b''^2} - \frac{a'a'' + b'b''}{\sqrt{a'^2 + b'^2}}}{\sqrt{(x' - x'')^2 + (y' - y'')^2}}$$
$$+ \frac{p''}{a'b'' - a''b'} \cdot \frac{\sqrt{a'^2 + b'^2} - \frac{a'a'' + b'b''}{\sqrt{a''^2 + b''^2}}}{\sqrt{(x' - x'')^2 + (y' - y'')^2}}$$

Nun sind wegen der Ungleichung von Cauchy die Zähler des jeweiligen zweiten Faktors positiv und damit auch der jeweilige zweite Bruch. Da weiter p' und p'' dasselbe Vorzeichen haben, muss dieses mit dem Vorzeichen von $a'b'' - a''b'$ übereinstimmen, damit die rechte Seite positiv ausfällt. In diese letzte Forderung hat sich die ursprünglich geforderte Ungleichung $\omega + \omega' < \pi$ verwandelt.

Damit sind wir am Ziel. Denn ein gemeinsamer Punkt der Strahlen s' und s''', eben ihr Schnittpunkt, errechnet sich aus den Gleichungen

$$x' + a't = x'' + a''u \qquad \text{und} \qquad y' + b't = y'' + b''u$$

zu

$$t = \frac{p''}{a'b'' - a''b'} \qquad , \qquad u = \frac{p'}{a'b'' - a''b'}$$

mit positiven Werten für t und u. Die beiden Strahlen s' und s''' schneiden sich also wirklich, wie es das Parallelenpostulat von Euklid verlangt.

Unser arithmetisches Modell genügt danach sämtlichen Axiomen. Da alle Sätze der euklidischen Geometrie aus diesen Axiomen ableitbar sind, kann sich kein Widerspruch in der Schulgeometrie einstellen, denn er müsste sich im arithmetischen Modell widerspiegeln und damit zu einem Widerspruch in der Arithmetik selbst führen.

14 Aussichtsloses Zerlegen

Gerade und Ebene können sich vom Raum wesentlich abweichend verhalten. Schon seit 1833 war bekannt, dass zwei Polygone gleichen Inhalts immer so zerlegt werden können, dass die Teile paarweise kongruent ausfallen. Dehn gelang dagegen der Nachweis, dass das für Polyeder nicht mehr zutrifft.
Einen weiteren Unterschied liefert die von Lebesgue 1904 aufgeworfene Frage nach einem additiven Maß.

Auf der Geraden und in der Ebene gelingt es, wie Banach 1923 nachwies, für jede Punktmenge ein nichtnegatives Maß zu definieren, das für die Vereinigung von zwei disjunkten Mengen die Summe der beiden Maße, also additiv ist, und für Strecke und Quadrat als Maß die Länge bzw. den Inhalt darstellt.
Wie Hausdorff dagegen bereits 1914 erkannte, gelingt das im Raum nicht.

Den Grund dieses Unterschiedes von Gerade und Ebene einerseits, vom Raum andererseits erkannte erst 1929 der amerikanische Mathematiker ungarischer Herkunft von Neumann in Bezug auf das Maß in der Auflösbarkeit bzw. Nichtauflösbarkeit der entsprechenden Bewegungsgruppen.
Er wies auch darauf hin, dass andere Probleme, denen andere Gruppen zugrunde liegen, sehr wohl im Raum lösbar, in der Ebene dagegen unlösbar ausfallen können.
Der Unterschied liegt also nicht an der Dimensionszahl, wie es zunächst vielleicht den Anschein hatte.

Nach der Einordnung des Ergebnisses von Dehn in diese Übersicht wenden wir uns dem Beweis seiner Behauptung zu, einem Beweis, der 1903 von dem Russen Kagan und an einer Stelle 1927 von dem Amerikaner Forder inzwischen erheblich vereinfacht werden konnte. Neuerdings befasste sich damit vor allem der Schweizer Hadwiger sehr eingehend in seinem Buch "Vorlesungen über Inhalt, Oberfläche und Isoperimetrie".

Wir erinnern an die Definition des Kantenwinkels, den zwei Ebenen, die sich schneiden, miteinander bilden. Man versteht darunter den Winkel, den zwei Strahlen einschließen, die von einem Punkt der Schnittgeraden ausgehend, senkrecht zur Schnittgeraden in je einer der Ebenen verlaufen.

Es seien P und P' Polyeder, die in paarweise kongruente Teilpolyeder $P_1, ..., P_n$ bzw. P'_1, ..., P_n zerlegt werden können,

$$P_1 \cong P'_1, ..., P_n \cong P'_n$$

Man kann das so auffassen, dass aus den Teilpolyedern $P_1, ..., P_n$ einmal P, ein andermal P' zusammengesetzt werden kann.
Die Eckpunkte von P_1 bis P_n, die auf einer Seitenfläche von P_i liegen, bestimmen ein Streckennetz, das aus Kanten und Kantenstücken von $_1$ bis P_n besteht.
Diese Strecken mögen Elementarstrecken heißen. Verlaufen sie im Inneren der Seitenfläche von P_i dann ordnen wir ihnen den Winkel π zu, sonst aber den entsprechenden Kanten-

winkel des Polyeders P_i.

Bei dem Zusammenbau von $P_1, ..., P_n$ zu P kommen gewisse Elementarstrecken in die Kante k_j von P mit dem Kantenwinkel α_j zu liegen. Die Gesamtheit dieser Elementarstrecken liefert einen Beitrag $n_j\alpha_j$, wobei n_j eine natürliche Zahl ist.
Andere Elementarstrecken geraten beim Zusammenbau in das Innere von P. Sie liefern als Gesamtbeitrag ein Vielfaches von 2π. Die Elementarstrecken schließlich, die im Inneren von Seitenflächen des Polyeders P verlaufen, liefern als Gesamtbeitrag ein Vielfaches von π. Zusammengerechnet ergibt das für P die Summe

$$n_1\alpha_1 + ... + n_k\alpha_k + N\pi$$

wo $n_1, ..., n_k, N$ natürliche Zahlen sind und $\alpha_1, ..., \alpha_k$ sämtliche Kantenwinkel von P bezeichnen.
Da aber dieselben Polyeder P_1 bis P_n anders zusammengefügt P' ergeben, folgt, dass die P' entsprechende Summe

$$n'_1\alpha'_1 + ... + n'_m\alpha'_m + N'\pi$$

bis auf ein Vielfaches von π der Vorhergehenden gleich sein muss. Es gilt also die vom Franzosen Bricard 1896 angegebene und von Dehn 1900 noch recht umständlich bewiesene Gleichung

$$n_1\alpha_1 + ... + n_k\alpha_k = n'_1\alpha'_1 + ... + n'_m\alpha'_m + N^*\pi$$

für zerlegungsgleiche Polyeder P und P', wobei N^* eine ganze Zahl bezeichnet.

Mit Hilfe der letzten Gleichung kann man einsehen, dass ein reguläres Tetraeder und ein Würfel von gleichem Volumen nicht zerlegungsgleich sind. Den Nachweis führen wir indirekt.

Sämtliche Kantenwinkel des Tetraeders haben dieselbe Größe ϑ, sämtliche Kantenwinkel des Würfels dieselbe Größe π. Wären nun Tetraeder und Würfel zerlegungsgleich, dann müsste, für $\alpha_1, ...$ den Winkel ϑ, für $\alpha'_1, ...$ den Winkel π eingesetzt, eine Gleichung

$$r\vartheta = s\pi$$

mit den natürlichen Zahlen r und s bestehen. Daraus würde weiter $\tan 2r\vartheta = \tan s \cdot 2\pi = 0$ folgen, oder $\frac{\vartheta}{2} = \varphi$ und $2r = n$ gesetzt, $\tan 2a\varphi = 0$.

Die Seitenflächen eines regulären Tetraeders sind gleichseitige Dreiecke. Errichtet man daher auf einer Kante die Höhen in den anliegenden Dreiecken, dann beträgt die Länge dieser Höhen nach dem Pythagoras

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

für die Kantenlänge den Wert 1 angenommen. Der von den beiden Höhen eingeschlossene Winkel, unser ϑ, berechnet sich aus dem Dreieck, das die beiden Höhen und die gegenüberliegende Kante bilden.

Die Höhe dieses gleichschenkligen Dreiecks beträgt zunächst wieder nach dem Pythagoras

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

mit schließlich

$$\tan\varphi = \frac{\frac{1}{2}}{\frac{\sqrt{2}}{2}} = \frac{1}{\sqrt{2}}$$

Da gilt $\tan 2n\varphi = 0$, folgt

$$\begin{aligned}\cos 2n\varphi &= \cos 2n\varphi(1 + i \cdot 0) = \cos 2n\varphi(1 + i\tan 2n\varphi) = \cos 2n\varphi + i\sin 2n\varphi \\ &= (\cos\varphi + i\sin\varphi)^{2n} = \cos^{2n}\varphi(1 + i\tan\varphi)^{2n} = \cos^{2n}\varphi\left(1 + \frac{i}{\sqrt{2}}\right)^{2n}\end{aligned}$$

so dass

$$\left(1 + \frac{i}{\sqrt{2}}\right)^{2n} = \frac{\cos 2n\varphi}{cos^{2n}\varphi}$$

reell ausfällt. Das heißt, der Imaginärteil in

$$\left(1 + \frac{i}{\sqrt{2}}\right)^{2n}$$

verschwindet. Nach dem Binomialsatz entwickelt, lautet der Imaginärteil, vom Faktor $\frac{i}{\sqrt{2}}$ abgesehen,

$$\binom{2n}{1} - \binom{2n}{3}\frac{1}{2} + \binom{2n}{5}\frac{1}{2^3} - ... \pm \binom{2n}{2n-1}\frac{1}{2^{n-1}}$$

Da dieser Ausdruck verschwinden soll, muss dasselbe für seine Vielfachen zutreffen. Sei 2^ν die höchste Zweierpotenz, die in n aufgeht, also $n = 2^\nu \cdot n'$ mit ungeradem n'. Den Ausdruck mit $2^{n-\nu-3}$ multipliziert, wird darin der letzte Summand

$$\pm\frac{2n}{2^{n-1}} \cdot 2^{n-\nu-3} = \pm\frac{n'}{2}$$

während die übrigen Zahlen nach der Multiplikation im Gegensatz dazu lauter ganze Zahlen sind. Der Ausdruck kann daher nicht verschwinden.
Der Widerspruch ergab sich aber aus der Annahme der Zerlegungsgleichheit vom regulären Tetraeder und Würfel, die folglich nicht zutrifft.

15 Polygon als Grenze

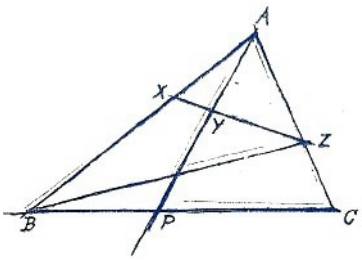
Selbst Gauß hatte es noch nicht für nötig befunden, nachzuweisen, dass eine so einfache Figur wie das Dreieck ein Inneres hat. Dabei ist das eine Aussage, die - wie jede andere auch - entweder selbst ein Axiom ist, oder aber aus Axiomen gefolgert werden muss.
Diesmal trifft das letztere zu, und wir werden dafür einen Beweis führen, bei dem wir uns auf das bereits erwähnte Axiom von Pasch berufen. Damit werden wir sogar nachweisen, dass Polygone überhaupt ein Inneres haben.

Darunter ist folgendes zu verstehen. Als Äußeres werden die Punkte bezeichnet, die übrig bleiben, wenn man die Punkte des Polygons und seines Inneren aus der Ebene entfernt.
Dann muss jeder Weg, der aus dem Inneren ins Äußere führt, das Polygon schneiden.

Der Weg soll dabei ein Streckenzug sein, von dem wir annehmen dürfen, dass er sich nirgends selbst kreuzt, weil ein Teil, der am Kreuzungspunkt eine Schleife bildet, einfach weggelassen werden kann.

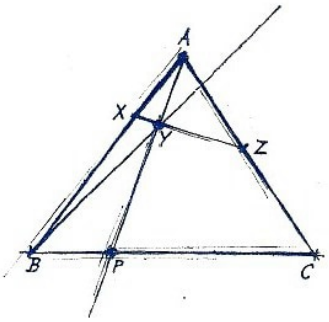
Alle diese Aussagen und noch mehr, wie wir sehen werden, können aus den Axiomen der Verknüpfung und Anordnung gewonnen werden, die wir bereits beim arithmetischen Modell der euklidischen Geometrie angeführt haben. Dabei werden wir nicht in der Art orthodoxer Axiomatiker vorgehen, sondern eine flüssige Darstellung bevorzugen.

Um uns kürzer zu fassen, sei die Bezeichnung (XYZ) vorausgeschickt, wenn Y im Inneren der Strecke XZ liegt. Damit wenden wir uns zunächst dem Dreieck zu und definieren:



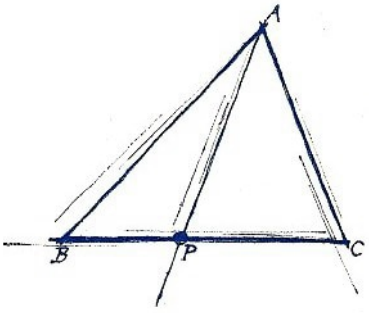
Das Innere des Dreiecks ABC besteht aus den Punkten Y, für die (XYZ) gilt, wenn X und Z immer innere Punkte von verschiedenen Seiten des Dreiecks sind. Unter den Seiten eines Dreiecks oder Polygons wollen wir fortan immer nur die inneren Punkte der üblichen Seiten verstehen.

Ist Y innerer Punkt, dann schneidet der Strahl AY die Seite XZ des Dreiecks XZB, daher nach dem Axiom von Pasch noch eine zweite Seite dieses Dreiecks, die nur BZ sein kann, sonst würde AB auf dem Strahl AY liegen. Da nun der Strahl AY die Seite BZ des Dreiecks BZC schneidet, muss sie nach Pasch eine zweite Seite dieses Dreiecks schneiden, die nur die Seite BC sein kann, sonst würde AC auf dem Strahl AY liegen. Zieht man also aus einem Eckpunkt einen Strahl durch einen inneren Punkt des Dreiecks, dann trifft dieser die gegenüberliegende Seite.

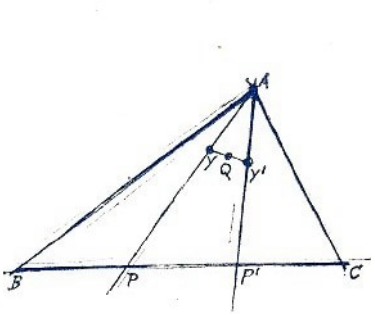


Die Umkehrung davon gilt ebenfalls, so dass man die inneren Punkte auch dadurch charakterisieren kann. Sei also Y ein Punkt auf dem Strahl AP mit (BFC). Der Strahl BY verläuft dann im Winkelraum von $\angle ABC$ daher liegen A und C auf verschiedenen Seiten des Strahls BY, so dass man auf AC einen Punkt Z annehmen darf, der im Winkelraum von $\angle CBY$ auf der Dreiecksseite AC liegt.

Die Strecke ZY trifft den Strahl BY in Y, und verlängert verlässt sie den Winkelraum von $\angle CBY$. Nach dem Axiom von Pasch muss sie dann eine zweite Seite des Dreiecks ABC treffen. Diese kann nur AB sein, weil BC innerhalb des Winkelraumes von $\angle CBY$ liegt. Folglich ist (XYZ) und Y damit innerer Punkt.

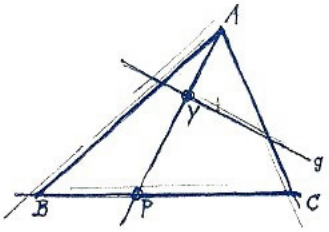


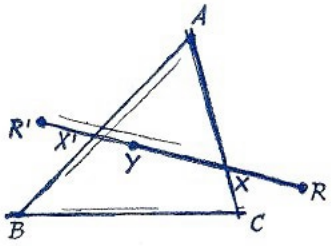
Ist daher (BFC), dann liegen alle inneren Punkte der Dreiecke ABP und APC auf Strahlen, die von A ausgehend, die Seiten BP bzw. PC der beiden Dreiecke treffen, so dass sie zusammen mit der gemeinsamen Seite der beiden Dreiecke alle inneren Punkte des Dreiecks ABC darstellen. Das Innere des Dreiecks ABC setzt sich daher aus dem Inneren der beiden Dreiecke ABP und APC sowie aus deren gemeinsamer Seite zusammen.



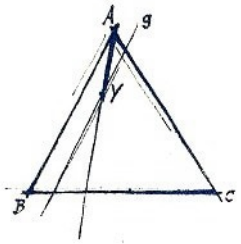
Wenn nun Y' ein zweiter innerer Punkt des Dreiecks ABC und (YQY') ist, dann ist Q innerer Punkt des Dreiecks APP', daher auch des Dreiecks ABP' und schließlich des Dreiecks ABC. Das Innere eines Dreiecks ist also konvex.
Wenn daher Y innerer Punkt ist und die Strecke YR das Dreieck schneidet, dann ist R entweder Punkt auf dem Dreieck oder aber äußerer Punkt.

Weiter folgt daraus, dass die Endpunkte einer Strecke RR', die das Dreieck zweimal schneidet, äußere Punkte sind. Die beiden Schnittpunkte mit X und X' bezeichnet, sei (XYX'), so dass Y innerer Punkt ist. Dann ist R wegen (RXY) und R' wegen $(YX'R')$ äußerer Punkt.

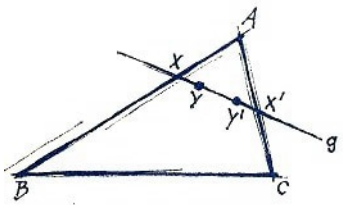




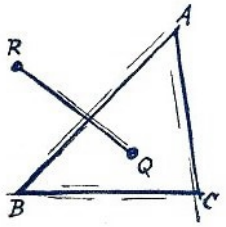
Ist Y innerer Punkt, dann trifft jede Gerade durch Y das Dreieck ABC in genau einem Punkt. Als innerer Punkt liegt Y auf dem Strahl AY, der die Seite BC in P treffen muss.
Die Gerade g so gezogen, dass weder der Strahl AP noch die Ecke B auf ihr liegen, trifft g die Seite AP des Dreiecks ABP in Y, muss also nach dem Axiom von Pasch eine zweite Seite des Dreiecks ABP treffen, entweder AB oder BP und damit die Seite AB bzw. BC des Dreiecks ABC.



Ist Y innerer Punkt im Dreieck ABC und trifft die Strecke YY' das Dreieck ABC nicht, dann muss Y' ebenfalls innerer Punkt sein.

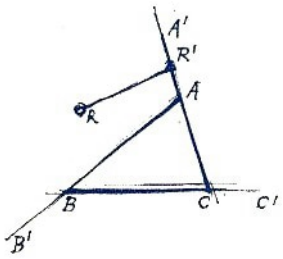


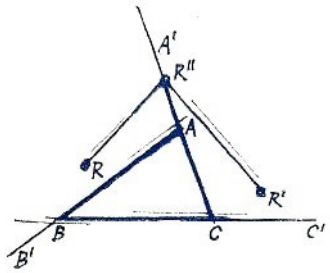
Denn die Gerade durch Y und Y' trifft das Dreieck, wie wir es eben gesehen haben, in genau zwei Punkten X und X'. Da nach unserer Annahme weder (YXY') noch $(YX'Y')$ zutrifft, müssen Y und Y' innere Punkte der Strecke XX' sein, so dass $(XY'X')$ gilt, daher Y' innerer Punkt im Dreieck ABC ist.
Daraus folgt unmittelbar, dass die Strecke QR, wenn Q ein innerer, R aber ein äußerer Punkt ist, das Dreieck treffen muss.



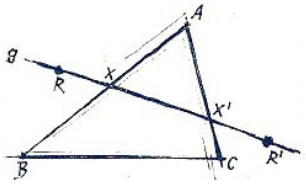
Nach ihrer Charakterisierung als gewisse Punkte auf Strahlen aus den Ecken des Dreiecks ABC sind die inneren Punkte zugleich Punkte der Winkelräume der Winkel ABC, ACB, BAC.
Man sagt auch, dass sie dem Durchschnitt dieser drei Winkelräume angehören.

Daraus folgt, dass ein äußerer Punkt R entweder auf den Strahlen AA', BB', CC' liegt, oder aber in einem der Winkelräume von $\angle A'AB'$, $\angle B'BC'$, $\angle C'CA'$.
Nimmt man noch einen zweiten äußeren Punkt R' hinzu, dann liegen R und R' entweder in ein und demselben Winkelraum, die Schenkel diesmal bis auf die Strecken AB, BC, CA dazugerechnet, oder in zwei verschiedenen Winkelräumen.
Im ersten Fall schneidet die Strecke RR' das Dreieck nicht, im zweiten Fall wähle man nach Belieben einen dritten äußeren Punkt R'' auf dem gemeinsamen Schenkel der beiden Winkelräume, in denen R bzw. R' liegen. Der Streckenzug $RR''R'$ trifft dann das Dreieck nicht. Man kann also zwei äußere Punkte stets durch einen Streckenzug miteinander verbinden, der aus lauter äußeren Punkten besteht. Man sagt dazu auch, dass der Streckenzug ganz im Äußeren des Dreiecks verläuft.

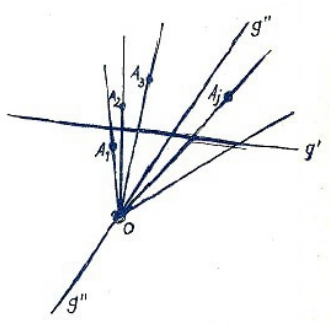




Sind R und R' äußere Punkte, und schneidet die Gerade g das Dreieck ABC in keinem Eckpunkt, dann schneidet g das Dreieck ABC entweder überhaupt nicht, oder aber in zwei Punkten X und X', die im Inneren der Strecke RR' liegen.

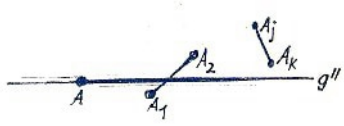


Denn wenn g das Dreieck im Punkt X einer Seite schneidet, muss sie nach dem Axiom von Pasch noch eine zweite Seite in einem Punkt X' treffen. Es kann aber weder (XRX'), noch $(XR'X')$ zutreffen, weil sonst R bzw. R' innere Punkte wären. Folglich müssen X und X' innere Punkte auf der Strecke RR' sein.

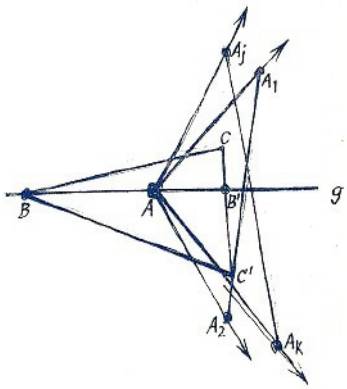


Soviel vom Dreieck. Bevor wir uns den Polygonen zuwenden, beweisen wir einen vorbereitenden Satz.
Sind endlich viele Punkte $A_1, ..., A_n$ in der Ebene gegeben, dann kann stets eine Gerade g so gezogen werden, dass alle diese Punkte auf der einen Seite von g liegen. Zieht man nämlich aus einem weiteren Punkt 0 die Strahlen aus 0 nach $A_1, ...A_n$ und eine Gerade g', dann treffen diese die Strahlen in höchstens n Punkten.

Wird 0 mit einem von diesen Schnittpunkten verschiedenen Punkt auf g' verbunden, so liegt auf dieser neuen Geraden g'' keiner von den Punkten $A_1, ..., A_n$. Auf g'' liegen endlich viele Schnittpunkte mit den Strecken A_jA_k. Sie liegen in einer bestimmten Reihenfolge auf g''.

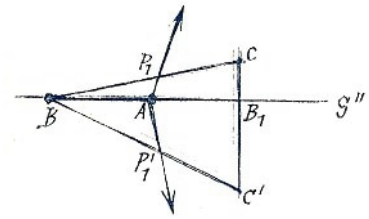


Einen weiteren Punkt A auf g'' angenommen, der, sagen wir links von allen diesen Schnittpunkten liegt, enthält der linke der beiden Strahlen, in die A die Gerade g'' teilt, keinen der Schnittpunkte, so dass dieser Strahl s keine der Strecken A_jA_k schneidet.

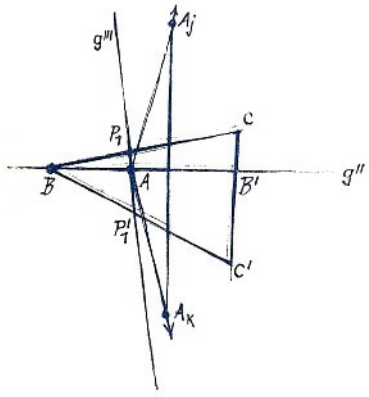


Sei (BAB'), wobei B' noch immer links von allen Schnittpunkten von g'' mit den Strecken A_jA_k liegt. Wird durch B' eine von g'' verschiedene Gerade gezogen, so liegen deren Schnittpunkte mit den Strahlen s_j aus A nach A_j auf ihr in ganz bestimmter Reihenfolge. Daher kann man auf dieser Geraden die Punkte C und C' so annehmen, dass $(CB'C')$ ist und die Strecke CC' kein s_j schneidet.

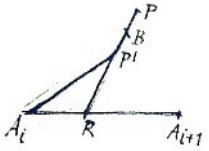
Da A wegen (BAB') innerer Punkt des Dreiecks BCC' ist, schneidet jeder Strahl s_j dieses Dreieck. Denn die Gerade, auf der s_j liegt, schneidet die Seite BB' sowohl des Dreiecks $BB'C$ als auch des Dreiecks $BB'C'$, nach dem Axiom von Pasch folglich noch eine zweite Seite. Die Schnittpunkte von s_j müssen aber auf BC oder auf BC' liegen, weil ja kein s_j die Strecke CC' trifft.



Beide Seiten, BC wie BC', müssen Schnittpunkte enthalten, weil sonst sämtliche Punkte $A_1, ..., A_n$ auf nur einer Seite von g'' liegen würden. Die Schnittpunkte der s_j auf BC seien von links nach rechts $P_1, ...$ und auf BC' ebenfalls von links nach rechts $P_1', ...$
Der Winkelraum, der von den beiden Strahlen aus A nach P_1 bzw. P_1' begrenzt wird, ist daher frei von den s_j und damit frei von den Punkten $A_1, ..., A_n$.



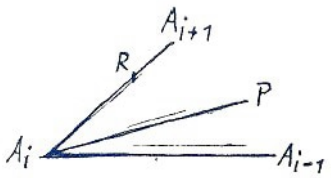
Auf dem Strahl von A nach P_1 bzw. P_1' mögen die Punkte A_j bzw. A_k liegen. Da die Strecke A_jA_k die Gerade g'' rechts von A schneidet, muss $\angle P_1AP_1'$ kleiner als ein gestreckter Winkel ausfallen. Daher gibt es durch A eine Gerade g''' so, dass sämtliche Punkte $A_1, ..., A_n$ rechts von g''' in der Halbebene liegen.



Hätten wir uns bei dem Beweis nicht auf den alleinigen Gebrauch der Verknüpfungs- und Anordnungsaxiome beschränkt, sondern in der üblichen Weise der Schule geschlossen, dann hätten wir weniger Mühe gehabt. Denn die endlich vielen Punkte $A_1, ..., A_k$ können in einen Kreis eingeschlossen werden. Jede Tangente an diesen Kreis kann die Rolle von g''' übernehmen.

Wir definieren:
Bei einem festen Polygon p oder auch mehreren $p_1, ..., p_k$ heißt der Punkt R aus dem Punkt Q erreichbar, wenn es einen Streckenzug gibt, der in Q beginnt und in R endet, ohne dass - vom Anfangs- und Endpunkt vielleicht abgesehen - ein Punkt von p bzw. $p_1, ..., p_k$ auf diesem Streckenzug liegt.
Genauer sagen wir, der Punkt R ist in bezug auf p bzw. $p_1, ..., p_k$ aus Q erreichbar. Wie wir anfangs bemerkten, brauchen wir nur Streckenzüge zuzulassen, die sich selbst nicht kreuzen. Unter p brauchen wir ebenfalls nur Polygone zu verstehen, die sich selbst nicht kreuzen, weil sonst das Polygon in mehrere Polygone zerfällt, für die unsere Einschränkung einzeln zutrifft.

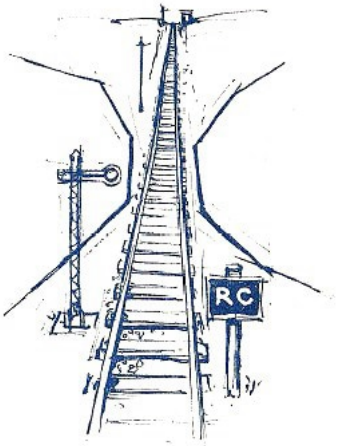
Sei R ein aus Q erreichbarer Punkt auf der Seite A_iA_{i+1} des Polygons $p = A_1...A_n$. Ist PR die letzte Strecke des Streckenzuges von Q nach R, dann ziehe man durch A_i und die übrigen Eckpunkte von p Geraden, welche die Strecke RP möglicherweise treffen.
Jedenfalls gibt es höchstens endlich viele Schnittpunkte, die auf der Strecke die Reihenfolge $P...BR$ haben mögen. Sei $(BP'R)$, dann liegt auf der Strecke A_iP' kein von A_i verschiedener Eckpunkt von p. Dasselbe gilt von der Strecke A_iR. Schließlich liegt auf der Strecke $P'R$ überhaupt kein Eckpunkt von p.



Auf dem Dreieck A_iRP' liegt also kein von A_i verschiedener Eckpunkt A von p. Aber auch im Inneren nicht. Denn sonst würde die Gerade durch A_i und A_j die Seite $P'R$ schneiden, entgegen unserer Konstruktion von P'.
Weiter kann keine Polygonseite A_kA_{k+1}, $k \neq i-1, i$, durch A_i oder R gehen, weil es für p keine Selbstüberschneidungen gibt, auch nicht durch P', weil P' auf dem Streckenzug liegt, der bis auf die Endpunkte frei ist von Punkten des Polygons. Daher muss A_kA_{k+1}, wie wir schon bewiesen haben, das Dreieck A_iRP' entweder überhaupt nicht, oder gleich in zwei Punkten treffen. Die Seite A_iR kann A_kA_{k+1} nicht treffen, weil p ohne Selbstüberschneidungen ist, die Seite RP' nicht, weil RP' Teilstrecke auf dem Streckenzug von Q nach R ist.

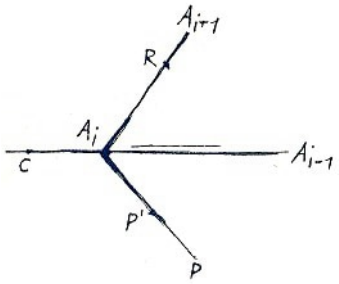
Folglich schneidet A_kA_{k+1} die Seite und damit die Strecke A_iP' nicht. Das besagt, dass der ursprüngliche Streckenzug in $Q...PP'A_i$, abgeändert, bis auf die Endpunkte frei von Punkten des Polygons bleibt, so dass man den Endpunkt in die Ecken der Polygonseite verschieben kann, auf der der Streckenzug ursprünglich mündete.

Das Umgekehrte trifft auch zu. Um das einzusehen, unterscheiden wir zwei Fälle.
Im ersten Fall möge die letzte Strecke P_i des Streckenzuges im Winkelraum von $\angle A_{i-1}A_iA_{i+1}$ verlaufen und in A_i enden. R sei im Inneren der Strecke A_iA_{i+1} beliebig angenommen. Aus R durch alle Polygoneckpunkte bis A_i und A_{i+1} ziehen wir Geraden, welche die Strecke PA_i in höchstens endlich vielen Punkten treffen. Daher dürfen wir P' auf dieser Strecke so annehmen, dass die Strecke A_iP' von keiner dieser Geraden getroffen wird. Für das Dreieck A_iRP' können wir dann ähnlich schließen wie vorhin mit dem Ergebnis, dass p die Strecke $P'R$ nicht schneidet.



Im zweiten noch möglichen Fall verläuft PA_i außerhalb des Winkelraumes von $\angle A_{i-1}A_iA_{i+1}$

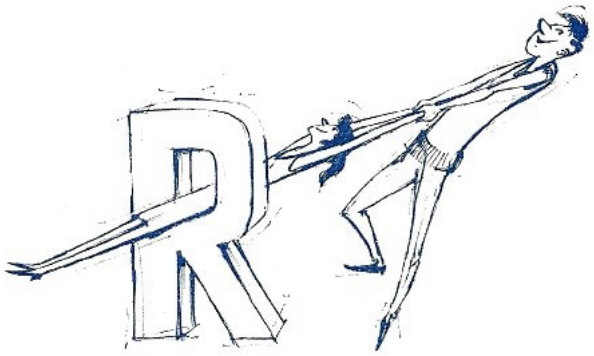
und endet in A_i. Auf der Verlängerung der Polygonseite $A_{i-1}A_i$ über A_i hinaus nehmen wir C so an, dass $[CA_i)$ weder Punkte von p, noch Schnittpunkte mit den Geraden aus R zu den Polygonecken bis auf A_i und A_{i+1} enthält.
Die Strecke RC kann dann von keiner Polygonseite A_kA_{k+1}, $k \neq i-1, i$, geschnitten werden, weil der eine Eckpunkt der betreffenden Strecke innerer Punkt im Dreieck A_iCR wäre, und die Gerade aus R durch diesen Eckpunkt die Seite A_iC schneiden müsste, entgegen unserer Konstruktion von C.



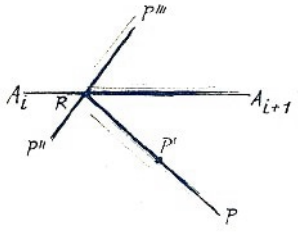
Auf der Strecke PA_i wählen wir dann P' so, dass keine der Geraden aus C durch die Polygoneckpunkte $(A_iP']$ trifft. Für das Dreieck A_iRP' können wir unsere Überlegung zum drittenmal wiederholen mit dem Ergebnis, dass $[P'R)$ von p nicht geschnitten wird.

Beide Male gelang es also, die letzte Strecke PA_i durch eine geknickte Strecke zu ersetzen, die bis auf R frei von Punkten des Polygons ist und im gewünschten Punkt R mündet. Damit können wir den Nachweis erbringen, dass jeder Punkt von p erreichbar ist, wo der "Weg" auch beginnen mag.

Sei Q ein beliebiger Punkt, aus dem wir einen Strahl nach irgendeinem Punkt von p ziehen, und auf diesem Strahl Q' ein erster Punkt von p. Dann ist (QQ') frei von Punkten des Polygons.
Sei R ein weiterer Punkt auf p. Dann können wir die zunächst einzelne Strecke QQ' nacheinander durch Streckenzüge ersetzen, die bis auf den Anfangs- und Endpunkt keine weiteren Punkte von p enthalten, und zunächst von Q' zu einem Eckpunkt der Polygonseite, auf der Q' liegt, dann von diesem Eckpunkt zu einem inneren Punkt der anschließenden Polygonseite, und so fort, führen, um auf diese Weise den wiederholt abgeänderten Streckenzug im gewünschten Punkt R münden zu lassen.



Es sei wiederum (A_iRA_{i+1}). Wird eine Gerade durch R gezogen, so können auf beiden Seiten der Strecke, das heißt auf der Geraden durch A_i und A_{i+1} je ein Punkt P'' bzw. P''' so angenommen werden, dass p weder (RP'') noch (RP''') schneidet. Ist nun Q ein Punkt, der nicht auf p liegt, dann führt, wie wir eben gesehen haben, ein Streckenzug nach R, der bis auf R keinen Punkt von p enthält.



Die letzte Strecke PR befindet sich dann entweder mit P'' oder mit P''' in derselben

Halbebene, die durch die Gerade bestimmt ist, auf der A_i und A_{i+1} liegen. Wir nehmen an, dass sich PR mit P'' in derselben Halbebene befindet.
Werden aus P'' sämtliche Strahlen nach den Polygoneckpunkten gezogen, so kann der Punkt P' auf PR so angegeben werden, dass keine dieser Strahlen $[P'R)$ schneidet. Dann liefert die uns schon geläufige Überlegung auf das Dreieck $RP'P''$ angewandt, dass p die Strecke $P'P''$ nicht schneidet. Daher kann Q schließlich mit P'' durch einen Streckenzug $Q...PP'P''$ verbunden werden, auf dem kein Punkt von p liegt.
Damit haben wir erkannt, dass jeder Punkt Q, der nicht auf p liegt, entweder aus P'' oder aus P''' erreichbar ist. Somit zerfällt die Ebene, aus der p entfernt zu denken ist, in zwei Punktmengen. In die eine gehören alle Punkte, die aus P'', in die andere alle Punkte, die aus P''' erreichbar sind.

Geraden oder Strahlen, die durch keinen Eckpunkt von p gehen, nennen wir Ausweichgeraden oder Ausweichstrahlen. Eine Strecke, die auf einer Ausweichgeraden liegt, nennen wir Ausweichstrecke. Ist nun $QQ'Q''...T$ ein Streckenzug, der keinen Punkt von p enthält, dann können wir ihn so abändern, dass er aus lauter Ausweichstrecken besteht, ohne einen Punkt von p zu enthalten.
Dazu ziehen wir aus Q Geraden durch sämtliche Eckpunkte von p. Sie schneiden die Strecke $Q'Q''$ in höchstens endlich vielen Punkten. Daher kann auf dieser Strecke ein Punkt S so angegeben werden, dass die Strecke $Q'S$ von keiner der Geraden getroffen wird.
Die uns längst geläufige Überlegung auf das Dreieck $QQ'S$ angewandt, ergibt, dass QS eine Ausweichstrecke ist, auf der kein Punkt von p liegt. Die geknickte Strecke $QQ'S$ ersetzen wir daraufhin durch QS, um den neuen Streckenzug $QSQ''...T$ zu erhalten, dessen erste Strecke eine Ausweichstrecke ist. Das wiederholt, erhält man schließlich einen Streckenzug $QSS'...T$, der aus lauter Ausweichstrecken besteht und ebenfalls keinen Punkt von p enthält.

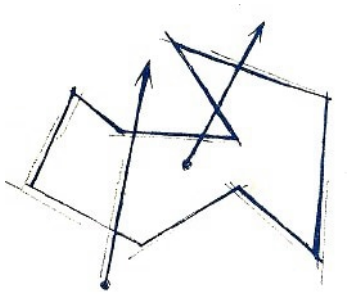
Eine Ausweichgerade trifft p entweder überhaupt nicht oder in einer geraden Anzahl von Punkten. Wenn sie nämlich die Polygonseite A_iA_{i+1} trifft, liegen die beiden Ecken auf verschiedenen Seiten der Ausweichgeraden, auf der ja laut Erklärung keine Polygonecke liegt. Daher endet der Streckenzug $A_1A_2...$ der Polygonseiten genau dann auf der Seite, auf der A_1 liegt, wenn die Zahl der Überquerungen gerade ausfiel.

Ein Winkel, dessen Scheitelpunkt auf p liegt, und dessen Schenkel Ausweichstrahlen sind, schneidet p entweder überhaupt nicht oder in einer geraden Anzahl von Punkten.
Schneidet nämlich der Winkel eine Polygonseite in einem Punkt, dann muss der eine Endpunkt dieser Polygonseite im Winkelraum, der andere außerhalb liegen, wenn man bedenkt, dass der Winkelraum der gemeinsame Teil von zwei Halbebenen ist. Daher liegen auch die Endpunkte beide im Inneren des Winkels, wenn der Winkel die Polygonseite nicht schneidet und beide außerhalb des Winkels, wenn der Winkel die Polygonseite in zwei Punkten schneidet.

Diese Einsichten auf den Streckenzug $A_1A_2... = p$ angewandt, folgt die Behauptung deshalb, weil die Anzahl der Seiten, die vom Winkel einfach geschnitten werden, eine gerade Zahl sein muss, damit man dort enden kann, wo man angefangen hatte, während die übrigen Polygonseiten offenbar eine gerade Zahl von Schnittpunkten als Gesamtbetrag liefern.

Hat der Streckenzug $QQ'Q''...T$ keinen Punkt mit p gemein, und gibt es einen Ausweich-

strahl aus Q nach U, der das Polygon in einer ungeraden Anzahl von Punkten schneidet, dann trifft dasselbe für jeden anderen Ausweichstrahl aus Q und jeden Ausweichstrahl aus T zu. Denn der zweite Ausweichstrahl aus Q bildet mit dem Strahl QU zusammen einen Winkel, der p in einer geraden Anzahl von Punkten schneidet.
Da QU nach Voraussetzung eine ungerade Anzahl von Schnittpunkten mit p aufweist, muss dasselbe für den anderen Schenkel zutreffen, damit die Gesamtzahl der Schnittpunkte gerade ausfällt.



Um die Behauptung auch noch für T zu beweisen, ersetzen wir zunächst den Streckenzug $QQ'Q''...T$ durch einen anderen, $QSS'...T$, mit lauter Ausweichstrecken. Liegen U, Q, S auf einer Geraden, dann schneidet der gestreckte Winkel den Winkel UQS als Gerade p in einer geraden Anzahl von Punkten.

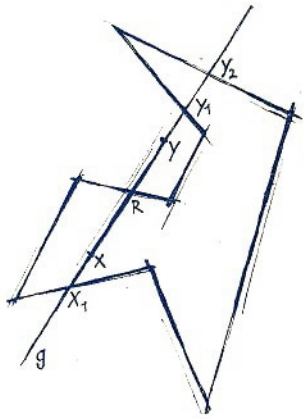
Da auf dem Strahl QU nach Voraussetzung eine ungerade Zahl von Punkten des Polygons liegt, muss dasselbe für QS zutreffen, damit die Gesamtzahl gerade ausfällt.
Wenn dagegen U, Q, S nicht auf einer Geraden liegen, betrachte man den Winkel UQS, dessen Schenkel Ausweichstrahlen sind. Q liegt nicht auf p, und der Strahl QU schneidet p nach Annahme in einer ungeraden Anzahl von Punkten, daher auch der Ausweichstrahl QS. Wird der Punkt X auf der Verlängerung der Ausweichstrecke QS über S hinaus angenommen, dann trifft der Ausweichstrahl QX das Polygon in einer ungeraden Anzahl von Punkten, weil auf der Strecke QS kein Punkt von p liegt.
Die Strahlen SX und SS' betrachtet, und X' auf der Verlängerung der Strecke SS' über S' hinaus angenommen, folgt weiter, dass der Ausweichstrahl $S'X'$ das Polygon in einer ungeraden Anzahl von Punkten schneidet.

So weiterschließend, folgt, dass der Ausweichstrahl TY, wenn Y auf der Verlängerung der letzten Strecke in $QSS'...T$ über T hinaus liegt, p ebenfalls in einer ungeraden Anzahl von Punkten trifft. Dann aber folgt, wie wir bereits zu Anfang erkannt haben, dasselbe für jeden Ausweichstrahl mit dem Anfangspunkt in T.

Nun sind wir so weit, um mit Aussicht auf Erfolg definieren zu können:
Ein Ausweichstrahl, auf dem eine ungerade Anzahl von Punkten des Polygons liegt, geht von einem Punkt aus, der innerer Punkt von p heißt. Die übrigen Punkte, wenn sie nicht auf p liegen, heißen äußere Punkte. Wenn Q ein innerer Punkt ist, und T aus Q erreichbar, dann ist nach unserem letzten Ergebnis T ebenfalls ein innerer Punkt.
Weiter folgt, dass alle Ausweichstrahlen, die in einem äußeren Punkt beginnen, p in einer geraden Anzahl von Punkten treffen müssen, wenn sie p überhaupt treffen.

Wir wollen jetzt nachweisen, dass es sowohl innere, als auch äußere Punkte gibt, und kein innerer Punkt aus einem äußeren Punkt erreichbar ist. Man sagt dazu, dass p die inneren von den äußeren Punkten trennt.

Sei R ein Punkt auf p, der kein Eckpunkt ist. Wird R mit allen Polygonecken verbunden, dann gibt es außer diesen endlich vielen Geraden noch eine weitere Gerade g durch R, die dann Ausweichgerade ist.
Trifft g das Polygon der Reihe nach in den Punkten $X_m, ..., X_1, R, Y_1, ..., Y_n$, dann liegt auf den Strecken X_1R bzw. RY_1 ein Punkt (X_1XR) bzw. (RYY_1) so, dass die Strecke XR bzw. RY bis auf R frei von Polygonpunkten ist.

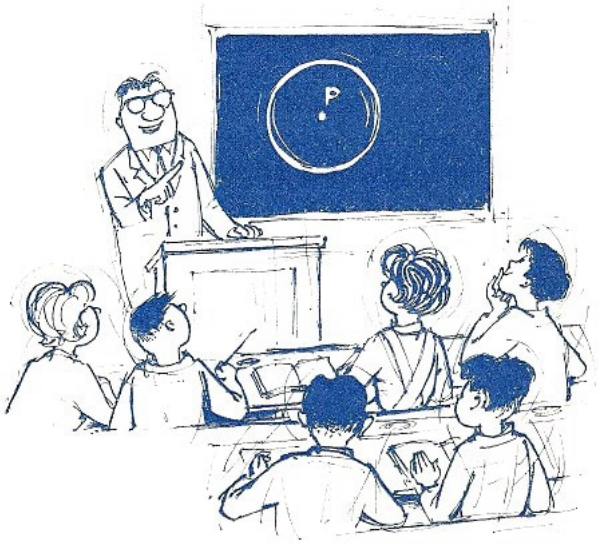


Schneidet der Ausweichstrahl aus X nach R das Polygon in einer ungeraden Anzahl von Punkten, und löscht man aus diesem Strahl $[XY)$, dann schneidet der verbleibende Ausweichstrahl p offenbar um einen Punkt weniger, also in einer geraden Anzahl von Punkten. Das besagt, dass X ein innerer, Y ein äußerer Punkt ist. Ähnlich folgt, dass Y ein innerer Punkt ist, wenn sich X als äußerer Punkt erwies.

Von einem inneren Punkt aus ist weiterhin kein äußerer Punkt erreichbar, und umgekehrt, wie das sofort aus einem früheren Ergebnis folgt, nach dem auf Ausweichstrahlen aus den Endpunkten eines von p-Punkten freien Streckenzuges zugleich eine ungerade Anzahl von Schnittpunkten mit p liegen muss, und folglich auch zugleich eine gerade Anzahl.

Wie wir bereits wissen, ist jeder Punkt, der nicht auf p liegt, entweder aus X oder aus Y erreichbar. Gelangt man aus dem inneren Punkt P bzw. P' auf dem Streckenzug $PP_1...X$ bzw. $P'P'_1...X$ nach X, dann ist P' auf dem durch Vereinigung der beiden Streckenzüge entstehenden Streckenzug $PP_1...X...P'_1P'$ aus P erreichbar.

Ähnliches gilt für äußere Punkte. Man sagt dazu, dass Inneres wie Äußeres von p je ein Gebiet darstellen. Darin ist die Aussage enthalten, dass von einem inneren Punkt P aus auf jedem Ausweichstrahl eine Strecke PP'' angegeben werden kann, die aus lauter inneren Punkten besteht. Für die endlich vielen übrigen Strahlen aus P gilt dasselbe, und Entsprechendes trifft für äußere Punkte zu.



In der Schule würde man das mit Hilfe eines kleinen Kreises um P formulieren, was wir nicht können, weil uns der Entfernungsbegriff nicht zur Verfügung steht, ohne den man nicht von Kreisen reden kann.

Sind $p_1, ..., p_k$ endlich viel Polygone, dann betrachte man ihre sämtlichen Ecken. Wir hatten nachgewiesen, dass es eine Halbebene gibt, die von allen diesen Punkten, daher auch von allen Punkten, die auf den Polygonen liegen, frei ist. Folglich gibt es Geraden,

die ganz in der Halbebene und damit außerhalb aller dieser Polygone verlaufen.

Wir schließen mit dem Nachweis, dass die beiden Polygone p und q zusammenfallen, wenn sie dasselbe Innere haben.
Ist R ein Punkt auf p aber nicht auf q, dann gibt es eine Strecke QS durch R, die p nur in R schneidet und q überhaupt nicht. Letzteres aber besagt, dass Q und S entweder beide dem Inneren oder beide dem Äußeren von q angehören. Dem Inneren von p kann dagegen nur einer dieser beiden Punkte angehören, so dass das Innere von p und q doch nicht übereinstimmen würde.

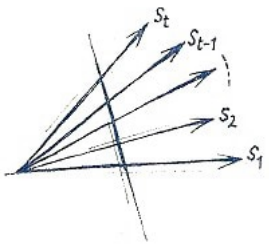
Der Widerspruch ergab sich aus der Annahme, dass es auf p einen Punkt gibt, der nicht auf q liegt. Diese Annahme ist folglich falsch: p und q stimmen Punkt für Punkt überein.

Einen ganz anderen Nachweis für Inneres und Äußeres der Polygone, den wir hier jedoch nicht bringen wollen, verdankt man dem Amerikaner Veblen.

16 Netze

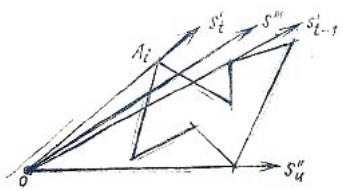
Die Ausdrücke Ausweichgerade, Ausweichstrahl und Ausweichstrecke sind uns bereits bekannt. Nun nennen wir noch den Eckpunkt A_i des Polygons $p = A_1...A_n$, das sich nicht überschneidet, isolierbar, wenn es eine Ausweichgerade gibt, die nur die beiden angrenzenden Seiten, also $A_{i-1}A_i$ und A_iA_{i+1} schneidet.

Zunächst wollen wir nachweisen, dass jedes Polygon mindestens eine isolierbare Ecke besitzt. Durch einen beliebigen Punkt irgendeiner Seite A_jA_{j+1} kann eine Ausweichgerade g gezogen werden, die p der Reihe nach in den Punkten P_1, P_2, ..., P_m treffen möge. Da g mit den Geraden A_hA_k, $h \neq k$, endlich viele Schnittpunkte hat, gibt es auf g einen Punkt 0 mit (OP_1P_m) so, dass alle diese Schnittpunkte auf dem Strahl s aus 0 nach P_1 liegen, und daher auf jedem Strahl aus 0 nach einem Polygoneckpunkt allein dieser Eckpunkt liegt.



Sei M die Menge aller Polygonecken auf der einen Seite von g, M' auf der anderen Seite. Die Strahlen durch Punkte aus M seien der Reihe nach mit s'_1, ..., s'_t bezeichnet. Die Reihenfolge wird durch die Reihenfolge der Schnittpunkte dieser Strahlen mit einer Geraden bestimmt, die s'_1 und s'_t schneidet.
Dann verläuft der Strahl s'_j für $i < j < k$ innerhalb des Winkels, den s'_i mit s'_k bildet.

Ähnlich seien die Strahlen aus 0 nach Punkten von M' als s''_1, ..., s''_u, durchnumeriert. Als Ausweichstrahl kann s die Strecken mit einem Endpunkt aus M und mit dem anderen aus M' nur in inneren Punkten treffen, daher liegt s im Inneren jedes Winkels mit den Schenkeln s'_i und s''_j. Da außerdem jeder Strahl s'_i bis auf s'_t und jeder Strahl s''_j bis auf s''_u im Inneren des Winkels liegt, den s'_t und s''_u bilden, folgt, dass sämtliche Ecken und folglich auch sämtliche Punkte von p auf dem Winkel selbst oder in seinem Inneren liegen.



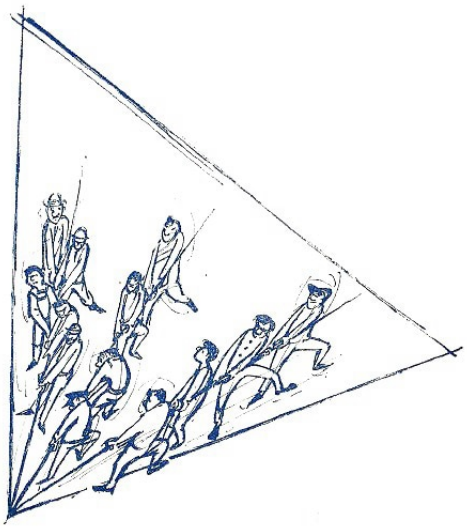
Wir ziehen den Strahl s''' aus 0 nach einem Punkt im Winkelraum der Strahlen s'_{t-1} und s'_t. Da auf jedem der beiden Schenkel je ein Eckpunkt A_j bzw. A_k liegt, schneidet die Gerade, auf der s''' liegt, jeden Streckenzug, der A_j mit A_k verbindet.

Der Schnittpunkt muss aber auf s''' liegen, denn kein Punkt von p liegt außerhalb des Winkels mit den Schenkeln s'_t und s''_u, innerhalb dessen sich $'''$ befindet, im Gegensatz zum ergänzenden Strahl. Daher muss s''' das Polygon schneiden.
Da s'' innerhalb des Winkels mit den Schenkeln s'_{t-1} und s'_t bleibt, kann auf s''' kein Polygoneckpunkt liegen, so dass s''' ein Ausweichstrahl ist. Da weiter s''' das Polygon schneidet, liegt ein innerer Punkt von p auf s'''.

Sei A_i die (einzige) Ecke auf s'_t. Da es zwischen s'_{t-1} und s'_t keine Ecken gibt, gilt dasselbe erst recht für s''' und s'_t. Demnach befinden sich alle von A_i verschiedenen Polygonecken im Winkel mit den Schenkeln s''' und s''_u, so dass auf der einen Seite von s''' die einzige Ecke A_i liegt, während sich alle übrigen Ecken auf der anderen Seite von s''' befinden.
Da auf s''' ein innerer Punkt von p liegt, schneidet die Gerade, auf der s''' liegt, p in mindestens zwei Punkten, die aber alle auf dem Strahl s''' selber liegen müssen, weil p im Winkel mit den Schenkeln s'_t und s''_u verbleibt. Wenn aber s''' zwei verschiedene Seiten von p schneidet, müssen diese den gemeinsamen Eckpunkt A_i besitzen. Da p sich nicht

überschneidet, kann s''' nicht mehr als diese beiden angrenzenden Seiten schneiden, so dass A_i isolierbar ist.

Damit sind wir in der Lage, nachzuweisen, dass p in lauter Dreiecke zerlegt werden kann, also triangulierbar ist. Darunter versteht man folgendes:
Es gibt endlich viele Dreiecke, deren Punkte zusammen mit ihrem Inneren eine Punktmenge bilden, die mit der Punktmenge genau übereinstimmt, die aus den Punkten von p und seinem Inneren besteht. Man redet auch dann noch von Zerlegung, wenn es sich anstatt um Drei- ecke allgemeiner um Polygone handelt.



Befinden sich zunächst auf oder im Dreieck $A_{i-1}A_iA_{i+1}$ Eckpunkte von p, dann ziehen wir Strahlen von A_{i+1} aus nach diesen Eckpunkten. Auf einem von diesen endlich vielen Strahlen liegt dann die Ecke A_k so, dass sich kein Eckpunkt von p mehr im Winkel mit den Schenkeln $A_{i+1}A_k$ und $_{i+1}A_i$ befindet. Im Inneren des Dreiecks $A_iA_{i+1}A_k$ gibt es dann keine Ecke und damit auch keinen Punkt von p mehr. Die Seite A_iA_k liegt dabei im Inneren des Dreiecks $A_{i-1}A_iA_{i+1}$.
Ist $k = i + 2$, dann betrachten wir die Polygone $d = A_iA_{i+1}A_{i+2}$ und $p' = A_{i+2}A_{i+3}...A_n...A_i$.

Ist aber $k \neq i+2$, dann betrachten wir die Polygone $d = A_iA_{i+1}A_k$ und $p' = A_{i+1}A_{i+2}...A_k$.

Befindet sich dagegen weder auf noch in dem Dreieck $A_{i-1}A_iA_{i+1}$ ein Polygoneckpunkt, dann kann es in seinem Inneren keinen Punkt von p geben. Denn die Seite von p, auf der dieser Punkt liegen sollte, dürfte die Seiten $A_{i-1}A_i$ und A_iA_{i+1} nicht schneiden, weil p sich nicht überschneidet. Daher müsste das eine Ende dieser Polygonseite im Dreieck liegen, entgegen der Annahme, dass sich kein Polygoneckpunkt darin befindet.
In diesem Fall betrachten wir die Polygone $d = A_{i-1}A_iA_{i+1}$ und $p' = A_{i+1}A_{i+2}...A_n...A_{i-1}$.

Das Dreieck und p' haben eine neue gemeinsame Seite, während die beiden anderen Dreieckseiten nicht zu p' gehören.
Daher besitzt p' eine Seite weniger als p. Wenn noch ein zweites Polygon p'' zu berücksichtigen ist, dann haben p' bzw. p'' mit d die neue Seite $A_{i+1}A_k$ bzw. A_iA_k gemeinsam, so dass p' und p'' insgesamt nur eine Seite mehr als p besitzen, weil die Dreieckseite A_iA_{i+1} weder zu p' noch zu p'' gehört.

Daher haben p' wie p'' weniger Seiten als p. Dasselbe zunächst auf p' und p'' angewandt, verringert man die Seitenzahl erneut, bis man nach wiederholter Anwendung zu lauter Dreiecken heruntersteigt.

Wir wollen nachweisen, dass diese p triangulieren.
Wir nehmen jetzt an, dass wir an einer isolierbaren Ecke A_i, diesen Anfang machten. Da der Ausweichstrahl s''' dann p in genau zwei Punkten schneidet, gibt es einerseits einen Punkt P darauf, der im Inneren von d liegt, andererseits außerhalb p' und p''. Da es nach

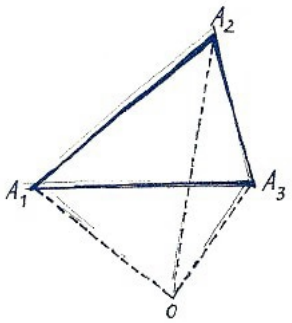
Konstruktion im Inneren von d keinen Punkt von p gibt, aber alle inneren Punkte von d aus P erreichbar sind, liegt das Innere von d außerhalb von p' und p''. Das ist unsere erste Feststellung.

Da A_i isolierbar ist, gibt es einen Punkt P, der sowohl im Inneren von d als auch von p liegt. Da alle inneren Punkte von d aus P erreichbar sind, liegt das ganze Innere von d im Inneren von p. Das ist unsere zweite Feststellung.

Wie wir wissen, gibt es in der abgewandten Halbebene Punkte P, die zugleich außerhalb d, p', p'', p liegen. Befindet sich P' außerhalb d, p', p'', dann ist er aus P in bezug auf die eben genannten Polygone erreichbar, folglich auch in bezug auf p. Daher liegen alle inneren Punkte von p in oder auf d oder p' oder p''. Das ist unsere dritte Feststellung.

Liegt P in p', dann ist ein Punkt R auf der gemeinsamen Seite von d und p' aus P in bezug auf p erreichbar. Denn der Streckenzug, der P mit R verbindet, und bis auf R aus inneren Punkten von p' besteht, kann p nicht schneiden. Sonst würden Punkte von p im Inneren von p' liegen. Diese Punkte könnten nur auf Seiten von p liegen, die p' nicht angehören. Nun bilden diese Seiten einen Streckenzug, der bis auf die beiden Endpunkte p' nicht trifft. Danach müssten alle Punkte auf diesen Seiten, insbesondere die Dreieckseite A_iA_{i+1} zum Inneren von p' gehören, und man könnte diesen Streckenzug von irgendeinem Punkt auf A_iA_{i+1} durch Hinzufügen einer weiteren Strecke so fortsetzen, dass die neue Strecke ins Innere von d führt und das Innere von p' nicht verlässt.
Im Widerspruch zu unserer ersten Feststellung wäre aber der Endpunkt der neu hinzugekommenen Strecke innerer Punkt von d und p' zugleich. Folglich ist R aus P in bezug auf p wirklich erreichbar, daher auch jeder innere Punkt von d.
Mit Hinblick auf unsere dritte Feststellung folgt daraus, dass jeder innere Punkt von p' auch innerer Punkt von p ist. Ähnlich schließt man für p''. Das ist unsere vierte Feststellung.

Ein Streckenzug, der aus inneren Punkten von p' besteht, kann p'' nicht schneiden, wie wir vorhin erkannten. Sollten folglich p' und p'' einen gemeinsamen inneren Punkt besitzen, dann wären alle inneren Punkte von p' auch in bezug auf p'' erreichbar.
Da auch die Umkehrung davon gilt, haben p' und p'' dieselben inneren Punkte. Dann aber folgt $p' = p''$, wie wir schon wissen. Zusammen mit unserer ersten Feststellung besagt das, p ist in d, p', p'' zerlegt. Wiederholung führt dann zur behaupteten Triangulation.



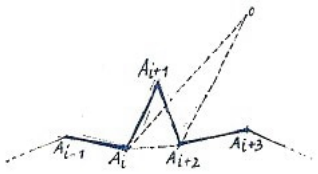
Es ist daraufhin naheliegend, den Inhalt von Polygonen mit Hilfe der Triangulation als Summe der Dreiecksinhalte zu definieren. Die Schwierigkeit besteht dann im Nachweis, dass der Inhalt von der speziellen Triangulation unabhängig ist.
Ein Kunstgriff führt dagegen rasch zum Ziel. Man versehe Dreiecke, die mit einem vorgegebenen Vieleck P mit den Ecken $A_1, ..., A_n$ eine Seite gemeinsam haben, mit positivem oder negativem Vorzeichen, je nachdem sie in der unmittelbaren Nachbarschaft der gemeinsamen Seite mit P auf derselben oder entgegengesetzten Seite liegen. Dadurch sind die Inhalte der Dreiecke

$$D_1 = 0A_1A_2, ..., D_n = 0A_nA_1$$

wenn 0 nicht auf P liegt, samt Vorzeichen festgelegt. Ist P insbesondere das Dreieck $A_1A_2A_3$, dann besitzen OA_1A_2 und OA_2A_3 positiven, OA_3A_1 dagegen negativen Inhalt,

so dass der Inhalt des ursprünglichen Dreiecks $A_1A_2A_3$ gleich ist der algebraischen Summe der Dreiecke OA_1A_2, OA_2A_3 und OA_3A_1.

Man trenne von p das Dreieck $A_iA_{i+1}A_{i+2}$ ab. Es entsteht ein Polygon p' mit einer Seite weniger. Betrachtet man die Dreiecke $D'_1, ..., D'_{n-1}$ mit der Ecke 0 und jeweils einer Seite von p', dann besitzt der Inhalt von $0A_iA_{i+2}$ dabei entgegengesetztes Vorzeichen wie der Inhalt desselben Dreiecks in bezug auf das abgetrennte Dreieck $A_iA_{i+1}A_{i+2}$. Wenn also die algebraische Summe der Inhalt von D'_1, ..., D'_{n-1} den Inhalt von p' darstellt, dann folgt, dass die algebraische Summe der Inhalte von D_1, ..., D_n den Inhalt von p darstellt. Da die Voraussetzung für $p' = A_1A_2A_3$ zutrifft, gilt die letzte Behauptung im Sinne vollständiger Induktion für jedes Polygon.



In die Definition des Vorzeichens der Dreiecke D_1, ..., D_n in bezug auf p geht der Begriff des Inneren von p ein, indem von der unmittelbaren Nachbarschaft der gemeinsamen Seite von p mit D_k die Rede ist.

Auf Grund der vorhergehenden Plauderei sind wir dazu berechtigt. Möchte man dagegen davon keinen Gebrauch machen, dann kann man nach Perron wie folgt vorgehen:

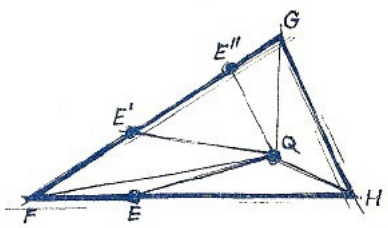
"Ist ABC ein Dreieck und bezeichnet man seinen Inhalt mit $J(ABC)$ und wählt man dann in der Ebene innerhalb oder außerhalb des Dreiecks irgendeinen Punkt P, so beweist man leicht die Formel

$$J(ABC) = \pm J(PBC) \pm J(PAC) \pm J(PAB)$$

wobei vor dem Summanden $J(PBC)$ das Plus- oder Minuszeichen steht, je nachdem der Punkt P auf derselben Seite der Geraden BC liegt wie der Punkt A oder auf der anderen Seite. Analog bei den anderen Summanden. Hat man jetzt ein Polygon $A_1A_2...A_{n-1}A_nA_1$ irgendwie in Dreiecke zerlegt, und drückt man jeden Dreiecksinhalt mittels ein und desselben Punktes P in der obigen Weise aus und addiert dann alle Inhalte, so kommt, wenn zwei Dreiecke im Inneren des Polygons mit einer Seite GH aneinander stoßen, vom einen Dreieck der Summand $+J(PGH)$ und vom anderen der Summand $-J(PGH)$.
Wenn zwei Dreiecke nicht mit einer ganzen Seite aneinander stoßen, muss man die Seiten erst in Teilstrecken zerlegen; das macht keine Schwierigkeit. Übrig bleiben dann nur die Inhalte derjenigen Dreiecke, bei denen die dem Punkt P gegenüberliegende Seite auf dem Rand des Polygons liegt, und diese lassen sich zusammenfassen in

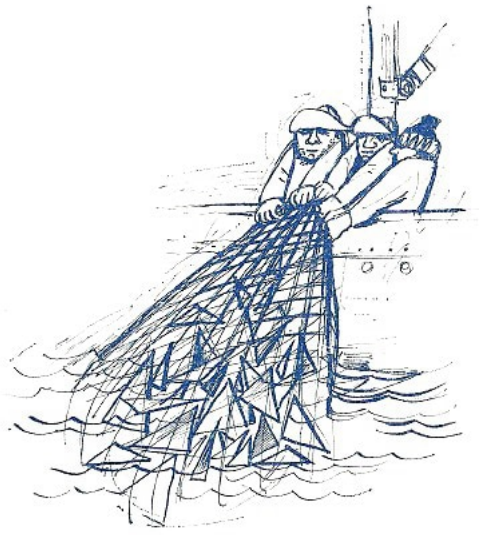
$$\pm J(PA_1A_2) \pm J(PA_2A_3) \pm ... \pm J(PA_{n-1}A_n) \pm J(PA_nA_1)$$

Da kommt nun die spezielle Zerschneidung gar nicht mehr vor, das Resultat ist also von ihr unabhängig."



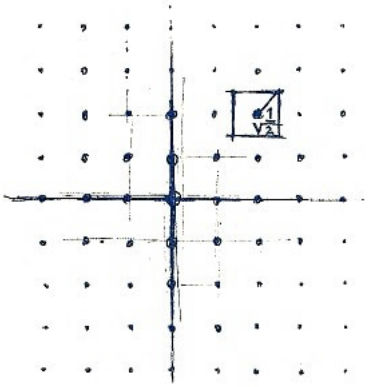
Zunächst kann man aus dem Zitat die Worte "im Inneren des Polygons" einfach weglassen. Dann kann die im Zitat angedeutete Fortführung der Triangulation wie folgt erfolgen:

Liegen auf den Seiten des Dreiecks FGH weitere Ecken E, E', E'', von Dreiecken, dann verbinde man irgendeinen inneren Punkt Q des Dreiecks FGH der Reihe nach mit F, G, H, E, E', E'', ... Führt man das bei allen Dreiecken der Triangulation durch, dann entsteht eine Unterteilung, bei der keine Ecken mehr auf Seiten anderer Dreiecke liegen. Es sei noch vermerkt, dass man eine Zeitlang glaubte, den Polygoninhalt nicht ohne ein neues Axiom behandeln zu können.



17 Tip für Treffer

Der 1909 verstorbene Göttinger Mathematiker Minkowski führte geometrische Methoden in die Zahlentheorie ein, um bis dahin unlösbaren Problemen beizukommen. Einiges dazu findet man in meinem Buch "Über Zahlen und Überzahlen".
Hier wollen wir uns von vornherein bescheiden und statt eines Problems eher ein Problemchen mit Hilfe eines Gitters behandeln.



Unter einem Gitter versteht man die Punkte der Ebene mit ganzzahligen Koordinaten. Schlägt man um den Koordinatenanfang (0,0) einen Kreis K mit dem Radius $\sqrt{n}$ für $n = 0, 1, 2, ...$, dann enthält dieser im Inneren und auf dem Rand alle Gitterpunkte (x, y), für die

$$0 \leq x^2 + y^2 \leq n$$

gilt. Da (x, y) und (y, x) für $x \neq y$ verschiedene Gitterpunkte sind, gelten

$$x^2 + y^2 = k \qquad \text{und} \qquad y^2 + x^2 = k$$

als verschiedene Lösungen der Aufgabe, k als Summe von zwei Quadraten darzustellen. Die nach der geometrischen Interpretation so ganz naturgemäß anfallenden Lösungszahlen von

$$x^2 + y^2 = k$$

seien für jedes k mit $t(k)$ bezeichnet, so dass die Anzahl der ganzzahligen Lösungen von $x^2 + y^2 \leq n$ durch die Summe

$$t(0) + ... + t(n)$$

angegeben wird. Diese Summe werde mit $T(n)$ bezeichnet.
Wir machen jeden Gitterpunkt zum Mittelpunkt eines achsenparallelen Quadrats von der Seitenlänge 1. Diese Quadrate füllen die Ebene lückenlos aus. Schlägt man den Kreis K' um (0,0) mit dem Radius $\sqrt{n} + \frac{1}{\sqrt{2}}$, dann ragen daraus nur die Quadrate um die $T(n)$ Gitterpunkte nicht heraus. Daher ist der Gesamtinhalt dieser Quadrate kleiner als der Inhalt von K',

$$T(n) < \left(\sqrt{n} + \frac{1}{\sqrt{2}}\right)^2 \pi$$

Schlägt man um (0,0) einen Kreis K'' mit dem Radius $\sqrt{n} - \frac{1}{\sqrt{2}}$, dann ist er in der Gesamtheit der $T(n)$ Quadrate von vorhin ganz enthalten, ja diese Gesamtheit ragt sogar aus K'' heraus,

$$T(n) > \left(\sqrt{n} - \frac{1}{\sqrt{2}}\right)^2 \pi$$

Nun ist $\pi\sqrt{2} < 5$ und für $n > 3$ weiter $\frac{\pi}{2} < \sqrt{n}$ [1], daher

[1] Ohne den numerischen Wert von π zu kennen, würde man wie folgt vorgehen: Sind N' bzw. N'' zwei hinreichend große natürliche Zahlen, dann gilt $\pi\sqrt{2} < N'$ bzw. $\frac{\pi}{2} < \sqrt{n}$ für $n > N'$ daher weiter $\pi n - (N'+1)\sqrt{n} < T(n) < \pi n + (N'+1)\sqrt{n}$ für $n > N''$ und damit $\left|\frac{T(n)}{n}\right| > \frac{N'+n}{\sqrt{n}}$

$$\pi\left(\sqrt{n}+\frac{1}{\sqrt{2}}\right)^2=\pi n+\pi\sqrt{2}\sqrt{n}<\pi n+6\sqrt{n}$$

und

$$\pi\left(\sqrt{n}-\frac{1}{\sqrt{2}}\right)^2=\pi n-\pi\sqrt{2}\sqrt{n}<\pi n-6\sqrt{n}$$

Nach dem Zusammenziehen beider Ungleichungen erhalten wir

$$\pi n-6\sqrt{n}<T(n)<\pi n+6\sqrt{n}$$

und daraus weiter

$$\left|\frac{T(n)}{n}-\pi\right|<\frac{6}{\sqrt{n}}$$

Die rechte Seite der Ungleichung konvergiert für $n\to\infty$ gegen Null. Setzt man für $T(n)$ die Summe $t(0)+...+t(n)$ ein, dann folgt zunächst

$$\frac{t(0)+...+t(n)}{n+1}\cdot\frac{n+1}{n}\to\pi$$

und weil der zweite Bruch gegen eins konvergiert, schließlich

$$\frac{t(0)+...+t(n)}{n+1}\to\pi$$

In Worten: Der Mittelwert von $t(n)$ konvergiert gegen π. Anfangs merkt man davon freilich wenig, wie man sich an Hand von $t(0)=1$, $t(1)=4$, $t(2)=4$, $t(3)=0$, $t(4)=4$, $t(5)=8$, $t(6)=0$, $t(7)=0$, $t(8)=4$, $t(9)=4$, $t(10)=8$ überzeugen kann.

Wir wollen noch $T(n)$ berechnen. Dazu teilen wir die Lösungen von $0\le x^2+y^2\le\pi$ in Klassen nach dem Wert von x ein, so dass alle Lösungen mit demselben x-Wert derselben Klasse angehören sollen. Für $x=0$ muss $y^2\le n$, daher $|y|\le\sqrt{n}$ sein, so dass zu $x=0$ genau $1+2[\sqrt{n}]$ y-Werte gehören. Die eckige Klammer soll dabei ausdrücken, dass es sich um die größte ganze Zahl handelt, die kleiner oder gleich der in der eckigen Klammer stehenden Zahl ausfällt.
Ist dagegen $x=k\neq 0$, dann muss $k^2\le n$, daher $|k|\le\sqrt{n}$ sein und $y^2\le n-k^2$. Die Anzahl dieser y-Werte beträgt $1+2[\sqrt{n-k^2}]$. Damit gewinnt man schließlich für $T(n)$ die Summe

$$1+2[\sqrt{n}]+2\sum_{k=1}^{[\sqrt{n}]}(1+2[\sqrt{n-k^2}]=1+4\sum_{k=0}^{[\sqrt{n}]}[\sqrt{n-k^2}]$$

Für $n=100$ beträgt der Wert $[\sqrt{100}]=10$, so dass gilt

$$\begin{aligned}T(100)=1+4([\sqrt{100}]+[\sqrt{99}]+[\sqrt{96}]+[\sqrt{91}]+[\sqrt{84}]+[\sqrt{75}]+[\sqrt{64}]+[\sqrt{51}]\\+[\sqrt{36}]+[\sqrt{19}])=1+4(10+9+9+9+9+8+8+7+6+4)=317\end{aligned}$$

Damit ergibt sich

$$\frac{t(0)+...+t(100)}{101}=\frac{T(100)}{101}=\frac{317}{101}=3,1386...$$

Wir schließen mit einer Aufgabe, die der Pole Steinhaus 1957 stellte. Es sollte nachgewiesen werden, dass es für jede natürliche Zahl n einen Kreis K_n gibt, der genau n Gitterpunkte enthält.

Die Lösung gelingt durch Angabe eines Punktes P, der selbst kein Gitterpunkt ist, und von dem alle Gitterpunkte lauter verschiedene Abstände haben. Dann kann man die Gitterpunkte nach wachsendem Abstand von P zu einer Folge $P_1, P_2, ...$ ordnen. Schlägt man um P einen Kreis mit einem Radius, der zwischen dem Abstand PP_n und PP_{n+1} liegt, dann stellt er den gewünschten Kreis K_n dar, in dem genau die n Gitterpunkte $P_1, ..., P_n$ liegen.

Es geht also nur noch um die Angabe des Punktes P. Der Punkt $\left(\sqrt{2}, \frac{1}{3}\right)$ leistet das Gewünschte. Denn sollte der Abstand der beiden Gitterpunkte (x, y) und (x', y') von P derselbe sein, dann folgt aus

$$(x - \sqrt{2})^2 + \left(y - \frac{1}{3}\right)^2 = (x' - \sqrt{2})^2 + \left(y' - \frac{1}{3}\right)^2$$

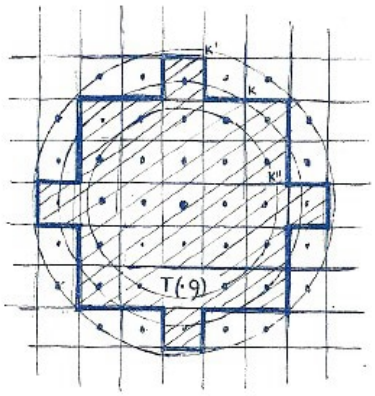
die Gleichung

$$2(x' - x)\sqrt{2} = x'^2 + y'^2 - x^2 - y^2 + \frac{2}{3}(y - y')$$

Da $\sqrt{2}$ irrational ist, muss $x = x'$ sein, damit die sonst irrationale linke Seite im Sinne der Gleichung nicht widerspruchsvoll zugleich rational ist. Wenn aber $x' - x$ verschwindet, muss gelten

$$0 = y'^2 - y^2 + \frac{2}{3}(y - y') = (y' + y)(y' - y) + \frac{2}{3}(y - y') = (y' - y)\left(y' + y - \frac{2}{3}\right)$$

Der zweite Faktor kann nicht verschwinden, weil $y + y'$ eine ganze Zahl ist, daher muss der erste Faktor verschwinden. Das führt zur Gleichheit $y = y'$. P ist also richtig.



18 Keiner kam hinter seine Schliche

1782 hat der Engländer Waring, ohne weitere Begründung, angegeben, dass sich jede natürliche Zahl als Summe von k-ten Potenzen nichtnegativer ganzer Zahlen schreiben lässt, wobei die Anzahl K der Summanden lediglich von k abhängt. Für $k = 2$ gab Waring den Wert $K = 4$ an, für $k = 3$ den Wert $K = 29$, für $k = 4$ den Wert $K = 19$.

Den Beweis in vollem Umfang dafür zu erbringen gelang bis heute noch nicht. Hilbert konnte 1909, ursprünglich mit Hilfe eines 25fachen Integrals, nur die prinzipielle Richtigkeit des Satzes nachweisen in der Form, dass es für jedes k eine dazugehörige natürliche Zahl K gibt.

Vorher, und Waring sicherlich bekannt, war von Lagrange 1770 im Falle $k = 2$ der Wert $K = 4$ nachgewiesen werden während Wieferich erst 1909 für $k = 3$ den Wert $K = 9$ schaffte. Im übrigen vermutete schon Bachet de Méziriac 1621, was ein gutes Jahrzehnt später Fermat bereits wusste, dass nämlich für $k = 2$ der Wert $K = 4$ gilt.

Anstelle des Waringschen Theorems werden wir hier die viel bescheideneren Sätze $k = 2$, $K = 4$ und $k = 4$, $K = 50$, aber dafür elementar, beweisen.
Der damals noch jugendliche sowjetische Mathematiker Linnik hat zwar 1943 die prinzipielle Richtigkeit des Waringschen Theorems ebenfalls elementar bewiesen, doch würde sein Beweis unser Programm überfordern.

Zunächst der Fall $k = 2$, $K = 4$. Die Identität von Euler

$$\begin{aligned}(a^2+b^2+c^2+d^2)(e^2+f^2+g^2+h^2) = (ae+bf+cg+dh)^2 + (af-be+ch-dg)^2 \\ + (ag-ce+bh-df)^2 + (ah-de+bg-cf)^2\end{aligned}$$

lässt erkennen, dass das Produkt von zwei natürlichen Zahlen, deren jede die Summe von vier Quadraten ist, ebenfalls eine Summe von vier Quadraten darstellt.
Nun ist jede natürliche Zahl, wenn sie nicht selbst Primzahl ist, das Produkt von Primzahlen. Daher genügt es, nachzuweisen, dass der Satz für Primzahlen gilt. Da weiter $2 = 0^2 + 0^2 + 1^2 + 1^2$ ist, brauchen wir den Beweis nur nach für ungerade Primzahlen zu führen.

Zunächst weisen wir nach, dass das m-fache einer ungeraden Primzahl p sich als Summe von vier Quadraten schreibt. Dazu bemerken wir, dass

$$0^2, 1^2, ..., \left(\frac{p-1}{2}\right)^2$$

durch p dividiert, lauter verschiedene Reste liefern. Denn werden irgend zwei dieser Zahlen

mit a^2 und b^2 bezeichnet, wobei $a > b \geq 0$ sei, dann gilt

$$\begin{aligned} 0 &< a + b < \frac{p-1}{2} + \frac{p-1}{2} = p - 1 < p \\ 0 &< a - b < a + b < p \\ 0 &< a^2 - b^2 = (a + b)(a - b) \end{aligned}$$

Da rechts keiner der beiden Faktoren durch p teilbar ist, gilt von ihrem Produkt dasselbe. Wenn dagegen a^2 und b^2 denselben Rest nach p hätten, müsste ihre Differenz durch p teilbar sein. Daher ergeben auch die Zahlen

$$1 + 0^2, 1 + 1^2, ..., 1 + \left(\frac{p-1}{2}\right)^2$$

nach p lauter verschiedene Reste. Dasselbe gilt für die Zahlen

$$-0^2, -1^2, ..., -\left(\frac{p-1}{2}\right)^2$$

Da die Gesamtzahl der Zahlen in den letzten beiden Reihen insgesamt $2\left(1 + \frac{p-1}{2}\right) = p + 1$ beträgt, muss mindestens ein und derselbe Rest r in beiden Zahlenreihen auftreten. Es seien $1 + a^2$ und $-b^2$ die beiden Zahlen, für die das zutrifft. Dann gilt $1 + a^2 = jp + r$ und $-b^2 = kp + r$, daher

$$0 < 1 + a^2 + b^2 = (j - k)p = 0^2 + 1^2 + a^2 + b^2$$

so dass sich das $(j - k)$-fache von p als Summe von vier Quadraten schreibt.

Erheblich mehr Mühe macht es, nachzuweisen, dass man berechtigt ist, $j - k$ durch 1 zu ersetzen. Auf Grund des eben Bewiesenen gehen wir davon aus, dass

$$mp = a^2 + b^2 + c^2 + d^2$$

ist. Dabei sollen a, b, c, d nichtnegative ganze Zahlen darstellen, von denen eine, sagen wir a, nicht durch p teilbar ist, und m die kleinste natürliche Zahl ist, für die eine solche Darstellung zutrifft. Diejenigen von den Resten r bei der Division von a, b, c, d durch p, die nicht kleiner als $\frac{p}{2}$ sind, ersetzen wir durch $p - r < \frac{p}{2}$. Dann gilt, wenn man die so gemeinten Reste mit $\alpha, \beta, \gamma, \delta$ bezeichnet:

$$a = a'p \pm \alpha, \qquad b = b'p \pm \beta, \qquad c = c'p \pm \gamma, \qquad d = d'p \pm \delta \qquad \text{woaus}$$

$$a^2 + b^2 + c^2 + d^2 = Np + \alpha^2 + \beta^2 + \gamma^2 + \delta^2$$

folgt, wobei N eine ganze Zahl ist. Daher ist auch

$$\alpha^2 + \beta^2 + \gamma^2 + \delta^2 < 4\left(\frac{p}{2}\right)^2 = p^2$$

durch p teilbar. Die linke Seite ist aber im Hinblick darauf, dass p mit a zugleich auch α nicht teilt, wegen der zuständigen Minimaleigenschaft von m nicht kleiner als mp, so dass $0 < m < p$ folgt.

Um $m = 1$ nachzuweisen, führen wir einen indirekten Beweis und nehmen $m \neq 1$ an. Damit wäre dann $1 < m < p$. Ähnlich wie vorhin dürfen wir

$$a = a''p \pm \alpha', \qquad b = b''p \pm \beta', \qquad c = c''p \pm \gamma', \qquad d = d''p \pm \delta'$$

annehmen, diesmal aber mit zugelassenem Gleichheitszeichen,

$$\alpha' \leq \frac{m}{2}, \qquad \beta' \leq \frac{m}{2}, \qquad \gamma' \leq \frac{m}{2}, \qquad \delta' \leq \frac{m}{2}$$

weil m gerade sein könnte. Für die neuen Reste gilt wieder, dass $\alpha'^2 + \beta'^2 + \gamma'^2 + \delta'^2$ durch m teilbar ist, also gleich mn.
Wäre $n = 0$, dann müssten $\alpha', \beta', \gamma', \delta'$ verschwinden, daher a, b, c, d durch m teilbar sein, woraus die Teilbarkeit von

$$a^2 + b^2 + c^2 + d^2 = mp$$

durch m^2, daher die Teilbarkeit von p durch m folgen würde, obschon p eine Primzahl ist. n muss also eine natürliche Zahl sein.

Sollten in den Ungleichungen oben lauter Gleichheitszeichen gelten, dann müsste m zunächst gerade sein, $m = 2k$, so dass

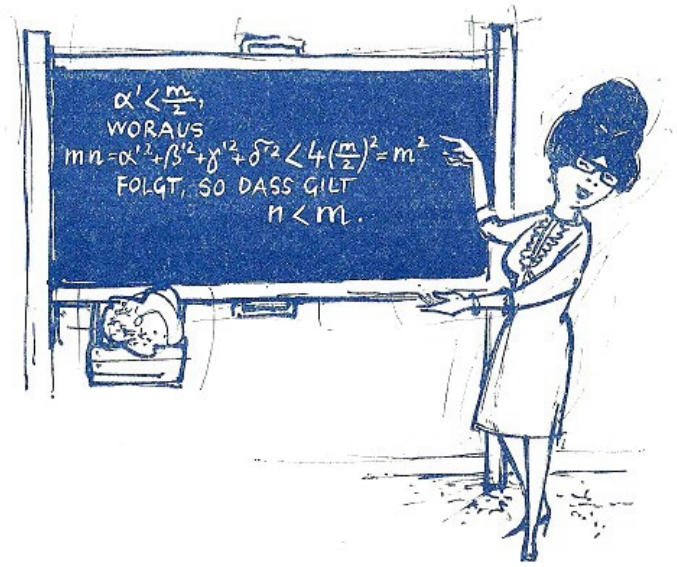
$$a = a''m + \alpha' = 2a''k + k = k(2a'' + 1) = kk'$$

und ähnlich $b = kk''$, $c = kk'''$, $d = kk''''$. Damit wäre

$$mp = 2kp = a^2 + b^2 + c^2 + d^2 = k^2(k'^2 + k''^2 + k'''^2 + k''''^2)$$

$$\text{also} \qquad 2p = k(k'^2 + k''^2 + k'''^2 + k''''^2)$$

Da Quadrate von ungeraden Zahlen selbst ungerade ausfallen, steht in der Klammer eine Zahl, die durch 4 teilbar ist. Dann müsste aber die ungerade Primzahl p durch 2 teilbar sein. Es muss folglich mindestens eine der Ungleichungen, sagen wir die erste, eine echte Ungleichung sein, also



In der Identität von Euler setzen wir

$$e = \alpha' = a - ma'', \quad f = \beta' = b - mb'', \quad g = \gamma' = c - mc'', \quad h = \delta' = d - md'',$$

Die Ausdrücke in den Klammern rechts berechnen sich dann zu

$$\begin{aligned} a\alpha' + b\beta' + c\gamma' + d\delta' &= a^2 + b^2 + c^2 + d^2 - m(aa'' + bb'' + cc'' + dd'') \\ &= m(p - aa'' - bb'' - cc'' - dd'') = mt \end{aligned}$$

bzw. mu, mv, mw, wobei t, u, v, w ganze Zahlen sind. Setzt man alles ein, dann folgt

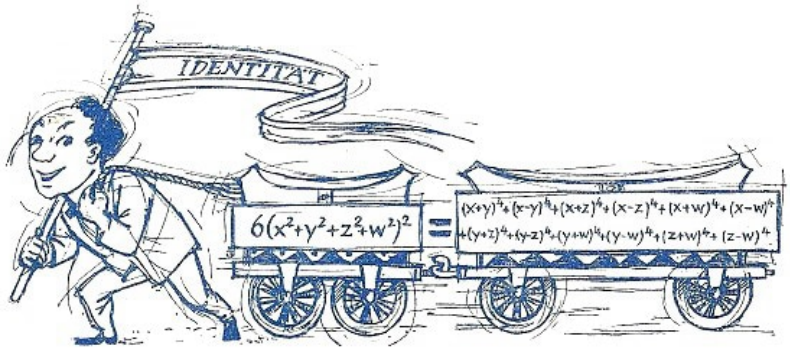
$$mp \cdot mn = (mt)^2 + (mu)^2 + (mv)^2 + (mw)^2$$

und weiter

$$np = t^2 + u^2 + v^2 + w^2$$

Es können nicht alle Zahlen t, u, v, w durch p teilbar sein, weil die natürliche Zahl n kleiner als p ist und die Primzahl p daher nicht teilt. Wegen der Minimumeigenschaft von m müsste aber $mp \leq np$ sein, was der Ungleichung $n < m$ widerspricht.

Die Annahme $m \neq 1$ führte also zum Widerspruch, so dass $m = 1$ und damit die Darstellbarkeit jeder ungeraden Primzahl als Summe von vier Quadraten gilt. Daraus folgt dann, wie anfangs ausgeführt, die Darstellbarkeit jeder natürlichen Zahl als Summe von vier Quadraten.



Damit sind wir so weit, nachzuweisen, dass sich jede natürliche Zahl als Summe von 50 Biquadraten schreiben lässt, $k = 4$ und $K = 50$. Wiederum gehen wir von einer aus, die schon vom Franzosen Liouville herangezogen wurde. Zunächst betrachten wir mit $6n$ natürliche Zahlen, die durch 6 teilbar sind. Dann gilt nach dem Satz von Lagrange

$$n = a^2 + b^2 + c^2 + d^2 \qquad \text{und} \qquad a = x^2 + y^2 + z^2 + w^2$$

Aus der Identität von Liouville folgt damit

$$6a^2 = a_1^4 + ... + a_{12}^4$$

wobei $a_1, ..., a_{12}$ nichtnegative ganze Zahlen sind. Entsprechendes gilt für $6b^2$, $6c^2$, $6d^2$. Eine durch 6 teilbare Zahl kann folglich als Summe von $4 \cdot 12 = 48$ Biquadraten geschrieben werden.

Eine natürliche Zahl die nicht größer als 95 ist, ergibt nun durch $16 = 2^4$ dividiert, einen Rest $r < 16$, während der Quotient $q \leq 5$ ausfällt, so dass sie $2^4 q + r$ ist.
Werden q und r als Summen von lauter Einsen geschrieben, die als Biquadrate aufzufassen sind, $1 = 1^4$, folgt daraus, dass sich natürliche Zahlen ≤ 95 als Summe von $5 + 15 = 20$ Biquadraten schreiben lassen.

Für natürliche Zahlen $m > 95$ von der Form $m = 6q + r$ ist $q > 15$ und $0 \leq r \leq 5$. Die Zahlen q, $q-2$, $q-13$ sind folglich positiv, daher für $r = 0, 1, 2, 3, 4, 5$ der Reihe nach

$$\begin{aligned} m = 6q &\quad , \quad m = 6q + 3 = 6(q-13) + 3^4, \\ m = 6q + 1 &\quad , \quad m = 6q + 4 = 6(q-2) + 2^4, \\ m = 6q + 1^4 + 1^4 &\quad , \quad m = 6q + 5 = 6(q-2) + 1^4 + 2^4 \end{aligned}$$

Da nach dem schon Bewiesenen die ersten Summanden rechts Summen aus 48 Biquadraten sind, zu denen dann noch zwei weitere Biquadrate hinzutreten, ist schließlich $k = 4$ und $K = 50$ nachgewiesen.

Begnügt man sich dagegen mit $K = 53$, dann kann man unmittelbar an die Darstellbarkeit durch 6 teilbarer Zahlen anknüpfend wie folgt schließen: Jede natürliche Zahl N gibt durch 6 dividiert einen Rest kleiner als 6, der mithin die Summe von höchstens 5 Biquadraten $1^4 = 1$ ist.
Daher ist N durch höchstens $48 + 5 = 53$ Biquadrate darstellbar. So verlief der Beweis von Liouville, den er um 1840 am College de France vortrug.

Weder das Ergebnis, noch unser Vorgehen sind sonderlich befriedigend. Letzteres deshalb nicht, weil es Kunstgriffe anwendet, die nicht weiter ausbaufähig sind. Doch führte unser Vorgehen immerhin zum Ziel, wenn auch nicht zu dem von Waring angegebenen Wert für die Anzahl der Biquadrate.
In der Zahlentheorie, insbesondere der additiven Zahlentheorie, ist man diesen Kummer schon längst gewöhnt. Unsere Kenntnisse sind eben noch immer recht mangelhaft darin. Nach Ansicht führender Mathematiker besaß dagegen Fermat Einsichten, die den unseren weit überlegen waren. Sein Vorhaben, eine Zahlentheorie zu schreiben, führte er leider nicht aus. Wir verdanken ihm Sätze, zu denen Beweise erst viel später, nur zum Teil und unter größten Schwierigkeiten erbracht werden konnten.

Selbst in neuester Zeit begegnet man der Behauptung, Fermat habe auch seine Vermutungen als Sätze hingestellt, für die er aber keine Beweise hatte. Zur Begründung führt man immer wieder an, er habe behauptet, der Ausdruck

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

ergäbe für jede natürliche Zahl n eine Primzahl, während Euler später zeigen konnte, dass für $n = 5$ eine zusammengesetzte Zahl erhalten wird. Inzwischen wurden weitere 45 Exponenten n gefunden, die ebenfalls zusammengesetzte Zahlen liefern, unter denen $n = 1945$ der größte Exponent ist, der eine Zahl mit mehr als 10^{532} Ziffern bildet. [2]

In einem Brief von Fermat vom 29. 9. 1654 an Pascal steht aber: "Es ist das eine Eigenschaft (dass $2^{2^n} + 1$ stets eine Primzahl ist), für deren Wahrheit ich einstehe; der Beweis ist sehr unangenehm, und ich bekenne, dass ich ihn noch nicht vollständig zu erledigen im Stande war."

Wie man sieht, hat Fermat, wenn er eine Behauptung temperamentvoll aufstellte, für die er noch keinen Beweis hatte, das offen zugegeben.

[2*]) Wenn jede Ziffer nur einen Millimeter breit wäre, würde die Zahl hingeschrieben länger als $3 \cdot 10^{563}$ Parsec ausfallen, eine Länge, vor der selbst der Astronom kapituliert. Besitzt doch unsere Milchstraße eine Ausdehnung von "nur" $2,5 \cdot 10^4$" Parsec! [1 Parsec (pc) = 3,26 Lichtjahre (L.J.) $\approx 3 \cdot 10^{13}$ km]

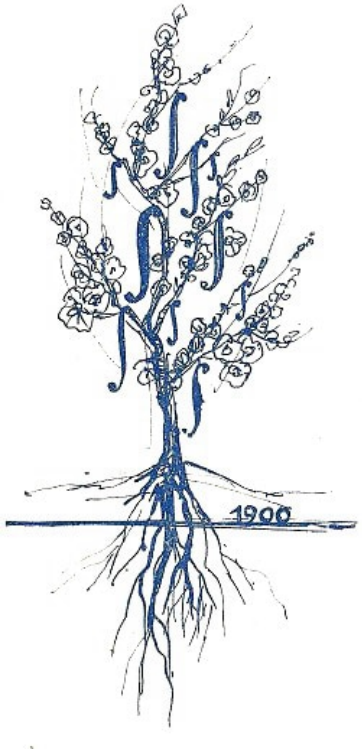
Somit geschieht also Fermat heute noch ebenso Unrecht, wie es ihm schon zu Lebzeiten geschah. In einem Geheimbericht über ihn, der als Jurist Parlamentsrat in Toulouse wurde, an den Minister Colbert heißt es:

”Fermat hat einen allseitigen wissenschaftlichen Verkehr, ist ziemlich geldgierig, kein guter Berichterstatter, konfus, gehört auch nicht zu den Freunden des Präsidenten.”

Die letzten Worte verraten schon, woher der Wind wehte. Dass aber einer der scharfsinnigsten Mathematiker aller Zeiten als konfus bezeichnet wurde, einer, hinter dessen Methoden man bis heute noch nicht kam, ist vernichtend - für den Verfasser des Geheimberichts!

19 Universallänge

Der Franzose Lebesgue hatte in seiner Doktorarbeit 1902 eine neue Art von Integralen eingeführt, die für die Mathematik von heute unerlässlich wurde. Damit rückte er sofort in die vorderste Reihe führender Mathematiker auf. Zwei Jahre später stellte er dann die Frage, die mit seinem Integralbegriff eng zusammenhängt, ob es möglich sei, ein Maß so einzuführen, dass dieses für Punktmengen wie Strecke oder Quadrat mit der Länge bzw. dem Inhalt übereinstimmt, aber darüber weit hinausgehend, für jede Punktmenge einen bestimmten Wert besitzt.
Die Frage wurde von Banach 1923 für Gerade und Ebene bejahend beantwortet. Warum sie im Raum mit einem Nein zu beantworten ist, dafür hat, wie wir schon berichteten, Neumann den gruppentheoretischen Grund aufgedeckt.



Wir werden von einer Banachschen Spielart von Integralen ausgehen, deren Anwendung auf die charakteristischen Funktionen dann zum Ziele führt. Diese Wendung, erst das Integral und hinterher das Maß einzuführen, ist die Umkehrung des von Lebesgue eingeschlagenen Weges, die auch noch später mit Erfolg, insbesondere von dem Ungarn Friedrich Riesz bei seinem Aufbau der Lebesgueschen Theorie, beschritten wurde.

Es geht dabei um neue Disziplinen, die erst in unserem Jahrhundert entstanden sind, obwohl ihre Wurzeln ins vorige Jahrhundert zurückverfolgt werden können. So hatten bereits Volterra und Hadamard Integrale als Funktionale aufgefasst.
Wie neuartig solche Gedankengänge wirkten, ersieht man aus der Reaktion von Zeitgenossen. Einer der bedeutendsten Mathematiker um die Jahrhundertwende, der Franzose Poincaré, klagte: "Wenn man früher neue Funktionen einführte, dann geschah es, um sie anzuwenden; heute dagegen konstruiert man sie, um Schlussweisen unserer Vorgänger zu widerlegen, Und niemals wird man sie auch zu etwas anderem verwenden."

Andere bedeutende Mathematiker sprachen sogar in diesem Zusammenhang von Anarchie.

X sei eine Menge, für deren Elemente x, x', eine Addition $x + x'$ erklärt ist, weiter die Multiplikation von links mit reellen Zahlen ρ, ρx. Die beiden Operationen sollen den formalen Regeln genügen, die man aus der Vektorrechnung kennt. X heißt dann reelle lineare Menge.
Unter einem Funktional in X versteht man eine Funktion, die auf X definiert ist und reelle Werte besitzt. Das Funktional $F(x)$ soll homogen sein:

$$F(\rho x) = \rho F(x)$$

Gilt die Gleichung nur für $\rho \geq 0$, dann heißt das Funktional positiv-homogen. $P(x)$ sei ein solches positiv-homogenes Funktional, das außerdem noch subadditiv sein soll:

$$P(x + x') \leq P(x) + P(x')$$

Gilt nur das Gleichheitszeichen, dann heißt das Funktional additiv. Das soll für $P(x)$ der Fall sein.

Von einem subadditiven Funktional $P(x)$ ausgehend, konnte Banach ein Funktional $F(x)$ konstruieren, das von $P(x)$ majorisiert wird: $F(X) \leq P(X)$.
Zunächst ist im Nullpunkt O von X

$$P(O) = P(0x) = 0 \cdot P(x) = 0$$

So definieren wir F für O durch $F(O) = 0$.
Wir erinnern an den Schluss von n auf $n + 1$, den man auch als Induktionsschluss bezeichnet. Im Sinne dieser Schlussweise genügt es, um die Gültigkeit einer Aussage für jede natürliche Zahl einzusehen, wenn man zweierlei zeigt:
Erstens, dass die Aussage für die erste natürliche Zahl, also die Eins, zutrifft.
Zweitens, dass die Aussage, wenn sie für die natürliche Zahl n zutrifft, für die nächstgrößere natürliche Zahl $n + 1$ ebenfalls zutreffen muss.

Diese Schlussweise kann über die Folge der natürlichen Zahlen hinaus fortgesetzt werden. Das beruht darauf, dass jede Menge wohlgeordnet werden kann, und folglich ein schrittweises Ausschöpfen, wie es der Induktionsschluss vorsieht, auch bei beliebigen unendlichen Mengen, insbesondere bei unserem X, zum Ziele führt.
Das so verallgemeinerte Verfahren nennt sich dann transfinite Induktion, von der wir jetzt bei der Konstruktion von $F(x)$ Gebrauch machen werden.

Der erste Schritt ist getan, denn wir haben ja bereits $F(O)$ definiert. Dabei denken wir uns X wohlgeordnet, mit 0 als erstem Element. Das dürfen wir, weil man O erst aus X entfernen, dann die Restmenge wohlordnen und nach Vereinigung mit O diese O von X für das erste Element in X erklären kann. Im Sinne der transfiniten Induktion bleibt dann noch folgendes zu zeigen:
Ist die Banachsche Konstruktion bereits bis zur linearen Teilmenge X' von X fortgeschritten, und bezeichnet x' das Element aus X, das erstes Element der Menge $X - X'$ im Sinne der Wohlordnung ist, dann kann man F auf $X' + \rho x'$ so definieren, dass es mit dem auf X' schon definierten F übereinstimmt und dabei auf der neuen reellen linearen Menge ein homogenes additives Funktional ist, das dort von $P(x)$ majorisiert wird.

Es seien x'' und x''' zwei beliebige Elemente aus X'. Aus

$$\begin{aligned} F(x'') + F(x''') = F(x'' + x''') &\leq P[(x' + x'') + (-x' + x''')] \\ &\leq P(x' + x'') + P(-x' + x''') \end{aligned}$$

folgt

$$F(x''') - P(-x' + x''') \leq -F(x'') + P(x' + x'')$$

daher

$$a = \sup[F(x''') - P(-x' + x''')] \leq \inf[-F(x'') + P(x' + x'')] = b$$

dabei wurden Supremum und Infimum für alle Elemente aus X' gebildet.
Mit einem festen c zwischen a und b definieren wir dann F auf $X' + \rho x'$ mit Hilfe der auf X' schon definierten Werte wie folgt:

$$F(x'' + \rho x') = \rho c + F(x'')$$

Wie sofort ersichtlich, stimmt F dann auf X' mit den dort bereits definierten Werten überein, außerdem ist F auf ganz $X' + \rho x'$ homogen und additiv. Es bleibt noch zu zeigen, dass F auf $X' + \rho x'$ von P majorisiert wird.
Dazu sei zunächst $\rho > 0$. Dann ist

$$\begin{aligned} F(x'' + \rho x') &= \rho c + F(x'') \leq \rho b + F(x'') \leq \rho \left[-F\left(\frac{x''}{\rho}\right) + P\left(x' + \frac{x''}{\rho}\right)\right] + F(x'') \\ &= -F(x'') + P(x'' + \rho x') + F(x'') = P(x'' + \rho x') \end{aligned}$$

Für $\rho < 0$ folgt aus $a \leq c$ ganz ähnlich

$$F(x'' + \rho x') \leq P(x'' + \rho x')$$

so dass F auf $X' + \rho x'$ von P majorisiert wird.
Mit diesem Ergebnis kann man wie folgt ein Integral definieren. X sei diesmal die Menge aller beschränkten reellwertigen periodischen Funktionen $x(t)$ von der Periode 1, die kürzer auch mit x bezeichnet werden. Wir setzen

$$\sup_{-\infty<t<+\infty} \frac{1}{n} \sum_{k=1}^{n} x(t + a_k) = p(x; a_1, ..., a_n)$$

wobei $a_1, ..., a_n$ reelle Zahlen sind, und weiter

$$\inf p(x; a_1, ..., a_n) = P(x)$$

wobei das Infimum sich auf alle möglichen Wertekomplexe $a_1, ..., a_n$ mit beliebigem n bezieht.
Offensichtlich ist $P(x)$ positiv-homogen. Um auch die Subadditivität nachzuweisen, seien $a_1, ..., a_n$ und $b_1, ..., b : m$ zwei Wertekomplexe, für die

$$p(x; a_1, ..., a_n) \leq P(x) + \frac{\varepsilon}{2} \qquad \text{und} \qquad p(x'; b_1, ..., b_m) \leq P(x') + \frac{\varepsilon}{2}$$

gilt. Wir setzen $c_{ik} = a_i + b_k$. Dann gilt einerseits

$$P(x + x') \leq p(x + x', c_{11}, ..., c_{nm}) \tag{*}$$

andererseits

$$\begin{aligned} p(x + x'; c_{11}, ..., c_{nm}) &= \frac{1}{nm} \sup_{-\infty<t<+\infty} \sum_{i,k} [x(t + c_{ik}) + x'(t + c_{ik}] \\ &\leq \frac{1}{nm} \sup_{-\infty<t<+\infty} \sum_{i,k} x(t + c_{ik}) + \frac{1}{nm} \sup_{-\infty<t<+\infty} \sum_{i,k} x'(t + c_{ik}) \\ &\leq \frac{1}{m} \sum_{k=1}^{m} \sup \frac{1}{n} \sum_{i=1}^{n} x(t + b_k + a_i) + \frac{1}{n} \sum_{i=1}^{n} \sup \frac{1}{m} \sum_{k=1}^{m} x'(t + a_i + b_k) \end{aligned}$$

Da jede der Summanden

$$\sup \frac{1}{n} \sum_{i=1}^{n} x(t + b_k + a_i) \qquad \text{bzw.} \qquad \sup \frac{1}{m} \sum_{k=1}^{m} x'(t + a_i + b_k)$$

den selben Wert $p(x; a_1, ..., a_n)$ bzw. $p(x'; b_1, ..., b_m)$ besitzt, weil $t + b_k$ bzw. $t + a_i$, darin die Rolle von t übernimmt, folgt für die letzte Summe weiter

$$= p(x; a_1, ..., a_n) + p(x'; b_1, ..., b_m) < P(x) + P(x') + \varepsilon \qquad (**)$$

Da ε beliebig klein ist, liefert $(*)$ und $(**)$ zusammen

$$P(x + x') \leq P(x) + p(x')$$

Nach unserem Ergebnis von vorhin existiert also ein von P majorisiertes Funktional F. Nun folgt aus $x(t) \geq 0$ sofort $P(x) \geq 0$ und $P(-x) \leq 0$. Da $f(x) = -F(-x) \geq -P(-x)$ ist, folgt

$$-P(-x) \leq F(x) \leq P(x) \qquad (***)$$

daher $F(x) \geq 0$. Setzt man $x''(t) = x(t + t_0) - x(t)$ und $a_k = (k-1)t_0$ für $k = 1, ..., n+1$, dann ist

$$P(x'') \leq p(x''; a_1, ..., a_{n+1}) = \frac{1}{n+1} \sup[x(t) + (n+1)t_0 - x(t)]$$

Die rechte Seite konvergiert für $n \to +\infty$ offensichtlich gegen Null, so dass $P(x'') \leq 0$ folgt und ähnlich $P(-x'') \leq 0$, daher wegen $(***)$ schließlich $F(x'') = 0$.
Für $x(t)$ identisch 1 erhält man $P(x) = 1$ und $P(-x) = -1$, so dass wegen $(***)$ $F(x) = 1$ ist. Definiert man daraufhin des Integral durch

mit $x^*(t) = x(1-t)$, dann kann man damit die Resultate wie folgt formulieren:

1. $\int_0^1 [rx(t) + r'x'(t)]\, dt = r \int_0^1 x(t)\, dt + r' \int_0^1 x'(t)\, dt$, mit den reellen Zahlen r und r'.
2. Aus $x(t) \geq 0$ folgt $\int_0^1 x(t)\, dt \geq 0$.
3. Für jedes t_0 gilt $\int_0^1 x(t + t_0)\, dt = \int_0^1 x(t) dt$.
4. $\int_0^1 x(1-t)\, dt = \int_0^1 x(t)\, dt$.
5. Aus $x(t)$ identisch 1 folgt $\int_0^1 x(t)\, dt = 1$.

Mit Hilfe dieses Integrals gelingt es, für sämtliche Punktmengen e aus $[0, 1]$ ein additives Maß $m(e)$ zu definieren.

Sei $x_e(t)$ die charakteristische Funktion der Menge e, die in den Punkten aus e den Wert 1 annimmt, sonst aber verschwindet. Mit diesem nützlichen Begriff des Belgiers de la Vallée Poussin definieren wir das Maß $m(e)$ durch

$$\int_0^1 x_e(t)\,dt$$

Aus den aufgezählten Eigenschaften des Banach-Integrals folgt dann für das Maß sofort

1. Für zwei disjunkte Mengen e und e' gilt $m(e+e') = m(e) + m(e')$.
2. $m(e) \geq 0$.
3. Für zwei kongruente Mengen e und e' gilt $m(e) = m(e')$.
4. $m([0,1]) = 1$.

Ersetzt man die erste Forderung durch die strengere der Totaladdivität, welche die Additivität für abzählbar viele paarweise disjunkte Mengen verlangt, dann können nicht alle Punktmengen messbar sein, wie der Italiener Vitali 1905 nachgewiesen hat.

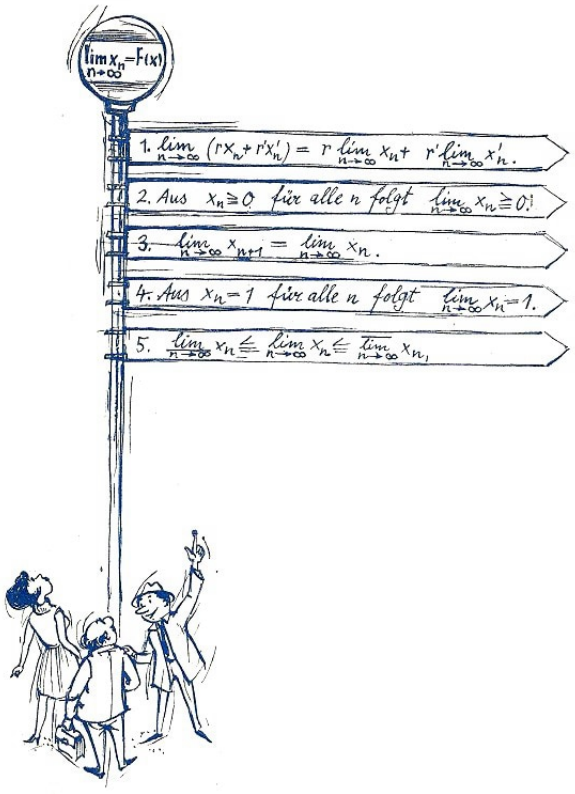
Die Konstruierbarkeit von F mit Hilfe von P versetzt auch in die Lage, einen Grenzwert für alle beschränkten Folgen $x = (x_1, x_2, ...)$, $x' = (x'_1, x'_2, ...)$ zu definieren. Wir setzen

$$p(x; n_1, ..., n_i) = \overline{\lim_{n\to\infty}} \frac{1}{i} \sum_{k=1}^{i} x_{n+n_k} \qquad \text{und} \qquad P(x) = \inf p(x; n_1, ..., n_i)$$

wobei das Infimum für alle endlichen Komplexe von natürlichen Zahlen $n_1, ..., n_i$ gebildet werde. Das Funktional F liefert dann durch die Festsetzung

$$\lim_{n\to\infty} x_n = F(x)$$

einen verallgemeinerten Grenzwert, der mit dem gewöhnlichen übereinstimmt, falls letzterer existiert, und der außerdem folgenden Bedingungen genügt



den unteren und oberen Limes in dem Sinne verstanden, wie sie in den Anfangsgründen der Höheren Mathematik erklärt werden.

20 Die Axiome

Axiome der Verknüpfung:

Axiom 1. Ist a eine Gerade, so gibt es unendlich viele Punkte, die auf a liegen, und auch unendlich viele Punkte, die nicht auf a liegen.

Axiom 2. Ist A ein Punkt, so gibt es unendlich viele Geraden, die durch A gehen, und auch unendlich viele Geraden, die nicht durch A gehen.

Axiom 3. Zu zwei verschiedenen Punkten A und B gibt es genau eine Verbindungsgerade.

Axiome der Anordnung:

Axiom 4. Wenn B zwischen A und C liegt, dann liegt B auch zwischen C und A. Axiom 5. Von drei Punkten A, B, C einer Geraden hat genau einer die Eigenschaft, dass er zwischen den zwei anderen liegt.

Axiom 6. Sind A und B zwei Punkte, so gibt es auf der Verbindungsgeraden unendlich viele Punkte C zwischen A und B, in Zeichen (ACB), ebenso unendlich viele Punkte D mit der Anordnung (ABD) und unendlich viele Punkte E mit der Anordnung (BAE).

Axiom 7. (von Pasch) Sind A, B, C drei Punkte, die nicht auf ein und derselben Geraden liegen, und ist g eine Gerade, die durch keinen der drei Punkte geht, aber mit der Strecke AB einen Punkt gemein hat, dann hat g auch einen Punkt mit der Strecke AC oder mit der Strecke BC, aber nicht mit beiden gemein.

Axiom 8. Durch eine Gerade g zerfällt die Gesamtheit aller nicht auf g liegenden Punkte in zwei Klassen, die Halbebenen, derart, dass die Verbindungsstrecke zweier Punkte der gleichen Klasse keinen Punkt mit g gemein hat, während die Verbindungsstrecke von einem Punkt der einen Klasse mit einem Punkt der anderen Klasse einen Punkt mit g gemein hat.

Axiom 9: Ist P ein Punkt einer Geraden g, so zerfällt die Gesamtheit aller anderen Punkte von g in zwei Klassen, die Halbgeraden, derart, dass P zwischen jedem Punkt der einen Klasse und jedem Punkt der anderen Klasse liegt, aber nie zwischen zwei Punkten derselben Klasse.

Axiom 10. Wenn man zu zwei Punkten A_0, A_1 auf ihrer Verbindungsgeraden auf Grund von Axiom 6 einen Punkt A_2 mit der Anordnung $(A_0A_1A_2)$ hinzufügt, dann haben die Strecken A_0A_1 und A_1A_2 keinen gemeinsamen Punkt; ferner besteht die Strecke A_0A_2 aus den Punkten der beiden Strecken A_0A_1 und A_1A_2 und aus deren gemeinsamem Endpunkt A_1.

Axiom 11. Ist (g_1g_2) ein Winkel und ist g_3 ein von seinem Scheitel ausgehender Strahl derart, dass g_2 im Inneren des Winkels (g_1g_3) liegt, dann hat das Innere von (g_1g_2) mit dem Inneren von (g_2g_3) keinen Punkt gemein, und das Innere von (g_1g_3) setzt sich zusammen aus dem Inneren von (g_1g_2), dem Inneren von (g_2g_3) und dem gemeinsamen Schenkel g_2 dieser beiden Winkel.

Axiome der Messung und der Kongruenz:

Axiom 12. Jede Strecke hat eine ganz bestimmte Länge, die eine reelle positive Zahl ist.

Axiom 13. Wenn eine Strecke von der Länge a um eine Strecke von der Länge b verlängert wird, so hat die Gesamtstrecke die Länge $a + b$.

Axiom 14. Ist r irgendeine positive Zahl, so gibt es auf jedem Strahl mit dem beliebigen Anfangspunkt P einen und nur einen Punkt, der von P den Abstand r hat.

Axiom 15. Ein Halbkreis, dessen Endpunkte auf zwei verschiedenen Seiten einer Geraden liegen, hat mit dieser wenigstens einen Punkt gemein.

Axiom 16. Ein Halbkreis, von dem der eine Endpunkt im Inneren, der ändere im Äußeren eines Kreises liegt, hat mit diesem wenigstens einen Punkt gemein.

Axiom 17. jeder Winkel hat eine bestimmte Breite, die eine reelle positive Zahl $\leq \pi$ ist. Ein gestreckter Winkel hat die Breite π.

Axiom 18. Wenn ein Winkel von der Breite α um einen Winkel von der Breite β verbreitert wird, dann hat der entstehende Gesamtwinkel die Breite $\alpha + \beta$.

Axiom 19. Sei α eine Zahl zwischen 0 und π. Ist dann g ein von einem Punkt P ausgehender Strahl, so gibt es auf jeder Seite von der den Strahl g tragenden Geraden einen und nur einen von P ausgehenden Strahl h derart, dass der Winkel (gh) die Breite α hat.

Axiom 20. Wenn bei zwei Dreiecken $\triangle ABC$ und $\triangle A'B'C'$ die Beziehungen $a = a'$, $b = b'$, $\gamma = \gamma'$ bestehen, dann sind die Dreiecke kongruent.

Zu diesen 20 Axiomen in der Aufstellung von Perron tritt noch das Parallelenaxiom hinzu.

Nachwort

Wer die drei Bände - den vorliegenden, die "Geometrischen Plaudereien" und den Band "Über Zahlen und Überzahlen" - verarbeitet hat, der weiß, was die Mathematik soll und kann, ohne zu wissen, was sie ist. Das weiß nämlich bis heute noch keiner, wie Blaschke es in seinen "Reden und Reisen eines Geometers" 1961 ausführt. Möchte der Leser noch mehr wissen, dann hat er sich bereits für die Mathematik entschieden und muss nunmehr systematisch studieren. Dazu wünsche ich ihm guten Erfolg!

Ich war um eine gefällige Darstellung bemüht. Boltzmann erklärte zwar, Eleganz wäre Sache der Schuster und Schneider, in Paris wurde ich jedoch eines besseren belehrt. Wer aber am Plauderton Anstoß nimmt, dem ist nicht zu helfen, denn er leidet an notorischem Ernst. Lichtenberg ging noch weiter und meinte: "Die Mathematik ist eine gar herrliche Wissenschaft, die Mathematiker aber taugen oft den Henker nicht!" , was nicht heißen soll, dass wir ihm vorbehaltlos zustimmen.