
Erhard Quaisser, Hans-Jürgen Sprengel

Extrema

1986 Deutscher Verlag der Wissenschaften

MSB: Nr. 127

Abschrift und LaTeX-Satz: 2021

<https://mathematikalpha.de>

Vorwort

Anschließend soll das Buch auch noch gekauft werden, und die Rezensionen sollen gut ausfallen.
Aus einem Vorwort von W. Gilde

Es dem potentiellen Käufer zu erleichtern und auch dem Rezensenten, ist unser Vorwort gedacht. Das Thema "Extrema" kann sehr unterschiedliche Erwartungen auslösen, da es ja auch in der Tat die unterschiedlichsten Gebiete berührt und verschiedenartigste Methoden anspricht.

Unser Büchlein ist schon vom Umfang her viel zu klein, um die Vielfalt des Themas auch nur annähernd wiederzugeben.

Worauf haben wir uns beschränkt? Welche Gründe, welche Zielstellung gab es dafür?

Unter den Methoden haben wir uns auf elementare Lösungsmethoden beschränkt. Unser Ausblick auf Methoden der höheren Mathematik in den Kapiteln 5 und 6 knüpft bewusst an die Kenntnisse des "Elementaren" an.

Gerade die Methoden der höheren Mathematik sind umfangreich in der Literatur dargestellt und werden ausführlich in den unterschiedlichen Ausbildungsrichtungen gelehrt. Das führt allerdings auch dazu, dass der Blick für das Elementare verstellt und mitunter "mit Kanonen auf Spatzen geschossen" wird.

Auch sollte man bei allem Streben nach Allgemeinheit nie vergessen, "dass die bunte Vielfalt der individuellen Probleme für die Vitalität der Mathematik entscheidend ist" (Courant/Robbins).

Von solchen Überlegungen haben wir uns beim Schreiben der Kapitel 2, 3 und 4 leiten lassen. Die Beschränkung auf elementare Methoden lässt es auch zu, dass sich schon Schüler unserem Büchlein zuwenden können (und sollen!).

Schüler der 7. und 8. Klassen können durchaus das Wesentliche des Kapitels 2 erfassen, bereits von Schülern der 9. und 10. Klassen können alle anderen Kapitel bearbeitet werden. Das Büchlein könnte also einen Schüler über mehrere Jahre begleiten, es kann selbständig durchgearbeitet oder 4 Vorwort von Lehrern in der außerunterrichtlichen Tätigkeit verwendet werden.

Bei der Darstellung elementarer Methoden ging es uns nicht nur um eine Gegenüberstellung zu den Methoden der höheren Mathematik, unsere Ausführungen bezwecken auch, die verschiedenen elementaren Methoden untereinander zu vergleichen, daher findet man viele Aufgaben, die nach verschiedenen der vorgestellten Methoden gelöst werden.

Überhaupt ist der Einsatz von Aufgaben ein von uns sehr massiv und bewusst gewähltes Mittel, Kenntnisse zu vermitteln. Erfahrungen, die wir selbst in der außerunterrichtlichen Tätigkeit gesammelt haben, bestärkten uns darin, so vorzugehen.

Wenn wir auch kein Buch über die "Kunst des Aufgabenlösen" schreiben wollten, so sind doch unsere diesbezüglichen Erfahrungen mit eingeflossen. Von den verschiedenen Prinzipien und Regeln der Heuristik spielen insbesondere eine Rolle

- das Analogieprinzip (Bezugnahme auf ähnliche Aufgaben, die vorher gelöst wurden, als auch Fortsetzung durch weitere analog konstruierte Aufgaben),
- das Modellieren (d. h. das Finden eines geeigneten mathematischen Modells für das gegebene Problem),
- das plausible Schließen.

Neben der Einschränkung bezüglich der Methoden war außerdem eine Reduzierung der Stoffgebiete notwendig. So wird man leider beispielsweise aus der Zahlentheorie oder Kombinatorik keine Aufgaben finden. Und obwohl wir die Geometrie besonders betont haben, ist das von uns Gewählte nur ein kleiner Ausschnitt, Packungsprobleme etwa fehlen auch.

Bezüglich der der Praxis entnommenen Aufgabenstellung muss natürlich deutlich gesagt werden, dass es sich um sehr genau ausgewählte Probleme handelt, die gerade mit den von uns vorgestellten Methoden behandelt werden können. Sie dienen also dazu, die Verbreitung von Extremalproblemen in praktischen Aufgaben zu demonstrieren, dürfen aber keineswegs zu der Vorstellung verführen, dass den Praxisproblemen stets so leicht "beizukommen" ist.

Das Literaturverzeichnis verrät, wo wir Anregungen suchten und fanden und wo der Leser weitere Entdeckungen wird machen können. Für einzelne Aufgaben Quellen anzugeben oder auf Autorschaft zu bestehen erschien uns nicht sinnvoll, da sich viele Aufgaben kaum bis zu ihrer Primärquelle zurückverfolgen lassen. Auch Aufgaben, die man glaubt selbst erfunden zu haben, könnten sich als "längst bekannt" herausstellen.

Für die vielfältigen und kritischen Hinweise danken wir Herrn Prof. Dr. Helmut Bausch. Unser Dank gilt weiterhin Frau Helga Kirmse, die das Manuskript mit gewohnter Zuverlässigkeit schrieb, der Lektorin Erika Arndt für die gute Zusammenarbeit sowie dem VEB Druckhaus "Maxim Gorki" für die sorgfältige Arbeit.

Potsdam, im Frühjahr 1986

Erhard Quaisser, Hans-Jürgen Sprengel

Inhaltsverzeichnis

Vorwort	2
1 Überall und immer Extrema	5
1.1 Ist nicht alles ganz einfach?	5
1.2 Aus der Geschichte der Extremalaufgaben	6
2 Geometrische Extremwertaufgaben	9
2.1 Streckenzüge kürzester Länge in der Ebene	9
2.2 Kürzeste Streckenzüge und Linien im Raum	15
2.3 Extremale Flächen	20
2.4 Extremale Umfänge	25
3 Extrema von Funktionen	28
3.1 Funktionen einer reellen Veränderlichen und ihre Extremwerte	28
3.2 Funktionen zweiten und dritten Grades und ihre Extrema	29
3.3 Die Funktion $f(x) = x + \frac{d}{x}$ und ihre Extreme	35
3.4 Angenäherte Berechnung von Extremwerten	38
4 Ungleichungen und Extrema	43
4.1 Zusammenhänge zwischen Funktionen und Ungleichungen	43
4.2 Das Lösen von Extremalaufgaben unter Verwendung der Ungleichung über das arithmetische und geometrische Mittel	49
4.3 Variationen zum Thema	55
4.4 Die Methode der konkurrierenden Punkte	58
5 Ein Ausblick auf Methoden der höheren Mathematik	62
5.1 Extrema und infinitesimale Methoden	62
5.2 Das Problem der Brachystochrone	67
6 Ein Einblick in Optimierungsprobleme	70
6.1 Die neue Aufgabenstellung	70
6.2 Der lineare Fall ist der einfachste	71
6.3 Ein Beispiel der dynamischen Optimierung	76
7 Lösungen und Lösungshinweise	78
Literatur	90
Erklärung verwendeter Symbole	91

1 Überall und immer Extrema

1.1 Ist nicht alles ganz einfach?

Was wäre zu wählen:
 Der Beste unter den Schlechten zu sein
 oder der Schlechtesten unter den Besten?
 F. Hebbel

Wäre diese Frage leicht zu beantworten, dann hätte Friedrich Hebbel sie wohl kaum gestellt. Es ist auch keine Frage mathematischer Art, es geht um menschliche Eigenschaften und menschliches Verhalten. Wer möchte da einen sicheren Vergleich wagen.

Reelle Zahlen dagegen lassen sich mittels der Ordnungsrelationen $<$ oder \leq einfach vergleichen. In einer endlichen Menge von Zahlen findet man leicht die größte oder auch die kleinste.

(1.1) •¹ Man ermittle jeweils die kleinste und größte Zahl der folgenden Mengen:

a) $A = \{1; 0; 7; -3; 17; \pi; -23\}$,

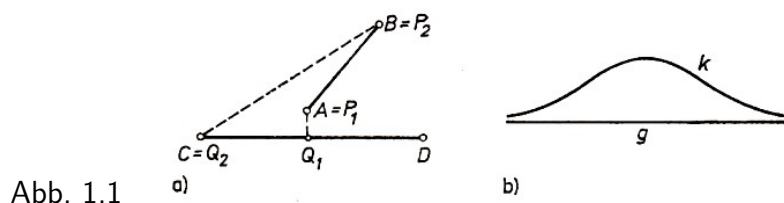
b) $B = \{4\pi; 12, 5; \sqrt{158}; \sqrt{156}\}$.

Im Fall a) "genügte ein Blick", bei b) auch, aber erst, nachdem wir die Zahlen durch Rechnung auf eine vergleichbare Form gebracht hatten. Selbstverständlich können wir dann auch die mittels reeller Zahlen (und Maßeinheiten) messbaren Größen - wie Längen, Flächen und Volumina in der Geometrie oder Kosten und Gewinne in der Ökonomie - miteinander vergleichen. Auch bei Entscheidungen von geringerer Bedeutung kann man davon Gebrauch machen:

Einer der vielen Günstlinge Katharinas II. glaubte, eine Bibliothek seinem Stande gemäß besitzen zu müssen und gab einem Buchhändler den Auftrag. "Was für Bücher befehlen Euer Exzellenz?"-

"Ja, das müssen Sie wissen, gerade so wie bei der Kaiserin, kleine Bücher oben, große unten!"²

Die bisher so einfach aussehende Aufgabe der Auswahl eines größten und kleinsten Elementes aus einer Menge wird schlagartig schwieriger, wenn es sich um eine Menge von unendlich vielen Elementen handelt. Nicht nur, dass wir mit dem bloßen "Überschauen" die gefragten Elemente natürlich nicht mehr herausfinden können, es entsteht eine prinzipiell neue Frage, die nach der Existenz.



(1.2) •* Man beweise, dass die Menge aller rationalen Zahlen r , für die $2 < r^2 \leq 3$ gilt, weder eine größte noch eine kleinste Zahl enthält.

Sind in der Ebene zwei Strecken \overline{AB} und \overline{CD} wie in Abb. 1.1a gegeben, so existiert ein Punktepaar (P_1, Q_1) , welches den kürzesten Abstand zwischen zwei Punkten P und Q mit $P \in \overline{AB}$ und $Q \in \overline{CD}$ aufzuweisen hat, und ein Punktepaar (P_2, Q_2) , welches den größten Abstand hat. (In unserem Beispiel sind diese Punktepaare sogar eindeutig bestimmt.)

¹Das Symbol • bedeutet, dass der Leser diese Aufgabe selbständig lösen soll. Sein Resultat kann er dann mit den Hinweisen in Kapitel 7 vergleichen.

²nach K. J. Weber, Demokritos, Berlin 1984, S. 61

Denken wir uns die in Abb. 1.1b angedeutete Kurve k und die Gerade g "bis ins Unendliche fortgesetzt", so ergibt sich ein Kurvenpaar, für welches weder ein kleinster noch ein größter Abstand zwischen ihren Punkten existiert.

Der folgende Trugschluss zeigt, dass wir schnell einem Fehler aufsitzen können, wenn wir die Existenz eines Extremums einfach voraussetzen.

Nehmen wir einmal (fälschlicherweise!) an, es existiere eine größte natürliche Zahl n . Dann lässt sich daraus herleiten, dass die größte natürliche Zahl die Zahl 1 ist:

Denn wäre $n > 1$, dann folgte nach Multiplikation mit n , dass $n^2 > n$ gilt, d. h., das Quadrat der natürlichen Zahl n (welches nach den Rechenregeln für natürliche Zahlen ja wieder eine natürliche Zahl ist) wäre noch größer als die größte Zahl. Das ist ein Widerspruch.

Da von den beiden "restlichen" Zahlen 0 und 1 die Zahl 1 die größere ist, ist folglich 1 die größte natürliche Zahl. Dieser Trugschluss sollte uns zur Vorsicht mahnen!

Wir werden öfter auf das Existenzproblem zurückkommen, allerdings in vielen Beispielen die Existenz auch als "gesichert" voraussetzen. Ein aufmerksames und kritisches Mitdenken unserer Leser ist uns sehr wichtig.

1.2 Aus der Geschichte der Extremalaufgaben

Die ersten uns überlieferten Extremwertaufgaben³ stammen aus den "Elementen" des Euklid⁴, und Euklid lebte bereits im 4. Jh. v. u. Z. in Alexandria. Dort hatten die Ptolemäer ein wissenschaftliches Zentrum mit einer berühmten Bibliothek einrichten lassen, es war die Zeit, in der sich der Typ des "berufsmäßigen Wissenschaftlers" entwickelte.

Weitere Extremwertuntersuchungen ergaben sich für die Mathematiker des Altertums im Zusammenhang mit dem isoperimetrischen Problem. Die Aufgabe besteht darin, mit einer Kurve vorgegebener Länge eine Fläche größtmöglichen Inhalts einzuschließen.

Der Sage nach soll Dido schon im 9. Jh. v. u. Z. bei der Gründung Karthagos eine Aufgabe aus diesem Problemkreis intuitiv gelöst haben:

Dido kaufte von den Einwohnern an der nordafrikanischen Küste ein Stück Land, "nicht größer als was eine Ochsenhaut umspannen kann". Sie zerschnitt die Ochsenhaut in feine schmale Streifen, aus denen sie eine lange Schnur drehte, mit dieser grenzte sie dann am Ufer des Mittelmeeres ihr erworbenes Land gegenüber dem anderen ab. Welche Form wählte sie? Den Halbkreis (mit seinem Mittelpunkt am Ufer)!

Um 180 v. u. Z. zeigte Zenodoros:

(1.3) Von allen regelmäßigen n -Ecken gleichen Umfangs besitzt das mit der größten Eckenanzahl den größten Flächeninhalt, und der Kreis übertrifft noch alle regelmäßigen n -Ecke gleichen Umfangs bezüglich der Fläche.⁵

Auch in der Natur bemerkte man das "Extremalprinzip". Pappos von Alexandria, der etwa um 320 u. Z. gelebt hat, verbindet in seinem Hauptwerk "Collectio"⁶, in dem er wesentliche Erkenntnisse seiner Zeit zusammenfasste, die Darstellung des isoperimetrischen Problems mit einem Exkurs über die Weisheit der Biene:

³Vgl. [7], S. 463.

⁴Euklid von Alexandria (etwa 365 bis etwa 300 v. u. Z.), griechischer Mathematiker.

⁵Vgl. [16], S. 147.

⁶nach [16], S. 204/205

"Es ist offensichtlich, dass Gott dem Menschen die beste und vollkommenste Vorstellung von der Weisheit im allgemeinen und der mathematischen Wissenschaft im speziellen verliehen hat, aber ein Teil dieser Dinge wies er gleichwohl einigen der vernunftlosen Tiere zu."

Nach Pappos trifft das insbesondere auf die Bienen zu, "... indem sie zunächst die Süßigkeiten der schönsten Blumen einsammeln, die auf der Erde wachsen, machen sie daraus, zur Aufnahme des Honigs, die Gefäße, die wir Honigwaben nennen, gleich alle und alle aneinander stoßend und von sechseckiger Form. Und dass sie dies ersonnen haben vermöge einer gewissen mathematischen Voraussicht, können wir auf die folgende Weise schlussfolgern.

Notwendigerweise mussten sie sich überlegen, dass die Figuren alle aneinanderstoßen, d. h., dass sie gemeinsame Seiten haben müssen, damit kein fremder Stoff in die Zwischenräume eindringen und so die Reinheit ihres Erzeugnisses entwürdigen könnte. Nun gibt es nur drei geradlinig begrenzte Figuren, die dieser Bedingung genügen, ich meine reguläre Figuren, die gleiche Seiten und gleiche Winkel besitzen; die Bienen nämlich würden keine Figuren haben wollen, die nicht einheitlich sind ...

Da es mithin drei Figuren gibt, die fähig sind, den gesamten Raum um einen Punkt auszufüllen, wählten die Bienen vermöge ihrer instinktiven Weisheit zur Konstruktion ihrer Honigwaben diejenige Figur, die die meisten Winkel hat, weil sie sich ausdachten, dass diese mehr Honig aufnehmen kann als jede der beiden anderen.

Mithin, die Bienen kennen genau die Tatsache - die ihnen selbst Nutzen bringt -, dass das Sechseck (der Fläche nach) größer ist als das Quadrat und das Dreieck und mehr Honig aufnehmen kann beim selben Aufwand an Material, das bei der Konstruktion der verschiedenen Figuren gebraucht wird. Wir jedoch, die wir einen größeren Teil Weisheit als die Bienen in Anspruch nehmen, wollen ein Problem von noch weiterer Ausdehnung untersuchen, das nämlich, dass von allen ebenen Figuren mit gleichen Seiten und gleichen Winkeln, die den gleichen Umfang haben, dasjenige mit der größeren Winkelzahl stets größer ist und dass die größte ebene Figur unter allen denjenigen, die einen mit den Polygonen gleichen Umfang besitzen, der Kreis ist."⁷

Die Mathematik des Mittelalters ist vergleichsweise arm an Untersuchungen über Extrema.

Die bedeutendsten Arbeiten über solche Probleme aus der Zeit vor der eigentlichen Erfindung der Differentialrechnung stammen von Pierre de Fermat (1601-1665). Er gehörte zu der um die damalige Zeit rasch zunehmenden Gruppe von Liebhabern und Amateuren, die sich außerhalb der Universitäten mit den Wissenschaften beschäftigten. Von Beruf war er Jurist, hatte aber wohl bei seinem Gerichtspräsidenten keinen sehr guten Stand, da dieser die durch das große Interesse an der Mathematik bewirkte Ablenkung von den Amtspflichten nicht gerade gern sah.

Die Nachwelt denkt anders darüber und wäre wohl froh, wenn sich Fermat noch mehr Zeit für die Mathematik genommen hätte und z. B. statt seiner berühmten Randnotiz zur Methode des Diophantos (bezüglich ganzzahliger Lösungen von Gleichungen) lieber etwas ausführlich gewesen wäre.

Auch zu seiner Studie "Über Maxima und Minima" erhielt Fermat Anregungen aus der antiken Mathematik. Neben interessanten Aufgaben mathematischer Art wird hier auch das Wirken von Extremalprinzipien in der Natur behauptet. Wir werden noch öfter auf das Fermatsche Prinzip der Optik zurückkommen:

⁷aus [16], S. 227/228

(1.4) Das Licht wählt zwischen zwei Punkten stets den Weg, auf dem es die minimale Zeit benötigt.⁸

Zur Lösung von Extremalaufgaben benutzte Fermat eine Methode, die schon in etwa der der Differentialrechnung entspricht. Während Fermat sich aber nur auf die Ermittlung von Extrema bzw. spezielle Funktionen (vgl. auch Abschnitt 5.1) bezieht, wird dann von Newton und Leibniz die Differentialrechnung als eine umfassende und weitreichende Methode entwickelt.

I. Newton (1643-1727) als der bedeutendste Physiker seiner Zeit sucht die Bewegung eines Teilchens zu beschreiben, für welches die Zeit und die Lage im Raum die variablen Größen sind. Newton spricht von den "fließenden Größen" und nennt daher seine von ihm entwickelte Methode "Fluxionsrechnung".

In seinem Werk "Philosophiae naturalis principia mathematica" (1687) wendet er die Fluxionsrechnung auf drei große Gebiete an, darunter die Ermittlung von Extremwerten.⁹

G. W. Leibniz (1646-1716), eines der letzten Universalgenies, war immer auf der Suche nach universellen Methoden. Sein Zugang zur Differentialrechnung war geometrischer Natur (vgl. Abschnitt 5.1), die von ihm eingeführten Symbole verwenden wir heute noch. Die Differentialrechnung erwies sich als außerordentlich fruchtbar für die Entwicklung der gesamten Mathematik, sie erwies sich auch als eine sehr weitreichende Methode bezüglich der Lösung von Extremalaufgaben, allerdings nicht als Universalmethode.

1696 stellte Johann Bernoulli¹⁰ das "Brachystochronenproblem", welches wiederum ganz neue Lösungsideen erforderte (vgl. Abschnitt 5.2).

In den dreißiger Jahren unseres Jahrhunderts traten verstärkt solche Extremalprobleme auf, die sich von den bis dahin bekannten vor allem durch die Vielzahl der Nebenbedingungen unterschieden, ihre Behandlung erforderte auch neue Methoden, wir sprechen von Optimierungsproblemen. Grundlegende Arbeiten dazu wurden u. a. von dem sowjetischen Mathematiker L. V. Kantorovic¹¹ veröffentlicht. Die Lösung von Optimierungsaufgaben bringt z. B. in der Wirtschaft großen Nutzen. Viele Optimierungsaufgaben werden im Alltag aber noch gefühlsmäßig behandelt, z. B.:

(1.5) Es ist ein Stundenplan so zu gestalten, dass alle Schüler pro Tag ohne Freistunden 6 Stunden unterrichtet werden und einige weitere Nebenbedingungen über die Reihenfolge von Stunden zu erfüllen sind (die man noch exakt formulieren müsste).

a) Wie muss dieser Plan aufgestellt werden, damit die Gesamtzahl der Freistunden aller Lehrer am kleinsten wird?

b) Wie muss dieser Plan aufgestellt werden, damit das Maximum der Freistunden eines jeden Lehrers minimal wird?

Nach diesem kurzen historischen Abriss ahnt man vielleicht auch, dass die Extremalaufgaben nicht nur als interessante Aufgaben innerhalb einer wirksamen Theorie auftreten, sondern dass sie ursprünglich als Probleme, die die Praxis stellte, vor solchen Theorien existierten und einen wesentlich fördernden Einfluss auf die Entwicklung entsprechender Theorien und Methoden hatten.

⁸Dieses ist die am häufigsten anzutreffende Formulierung, exakt muss allerdings von einer extremalen Zeit gesprochen werden. Man vgl. dazu etwa Grimsehl, Lehrbuch der Physik, Bd. III, Leipzig 1955, S. 144.

⁹Vgl. [17], S. 180.

¹⁰J. Bernoulli (1667-1748), Schweizer Mathematiker, wirkte in Genf, Paris, Groningen und Basel.

¹¹Leonid V. Kantorovic, geb. 1912, Prof. in Leningrad.

2 Geometrische Extremwertaufgaben

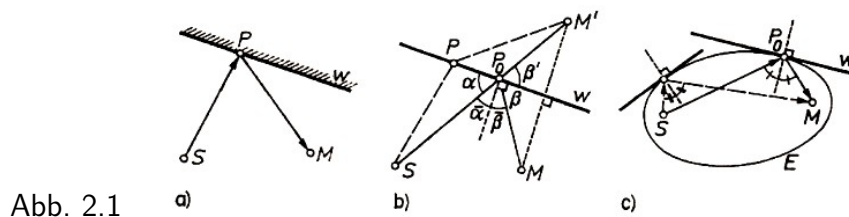
2.1 Streckenzüge kürzester Länge in der Ebene

Habe einen guten Gedanken,
man borgt dir zwanzig.
M. v. Ebner-Eschenbach

"Von allen Linienstücken, die gleiche Endpunkte haben, ist die gerade Linie die Kürzeste" formulierte Archimedes. Diese Erkenntnis finden wir auch im Fermatschen Prinzip der Optik wieder: Ist der Raum mit einem einheitlichen (homogenen) lichtdurchlässigen Stoff gefüllt, so breitet sich das Licht in diesem Raum geradlinig aus.

Man könnte zunächst der Meinung sein, dass sich mit diesem einfachen Grundprinzip nur sehr einfache Sachverhalte betrachten lassen. Die Behandlung der folgenden Probleme zeigt, dass man die obige Archimedische Erkenntnis günstig anwenden kann.

(2.1) Bei einem Staffelwettbewerb soll jeder Läufer vom Start S zu einer Wand w laufen, dort anschlagen, weiter zum Mal M und von dort zurück nach S (siehe Abb. 2.1 a). An welcher Stelle muss er die Wand w berühren, damit sein Gesamtweg möglichst klein wird?



Bei dieser Aufgabe können wir die Strecke \overline{MS} , deren Länge ja unabhängig von der Lage von P ist, aus den Überlegungen ausklammern. Es interessiert nur die Länge des Streckenzuges \overline{SPM} .

Wenn sich das Fermatsche Prinzip übertragen lässt, dann müsste der Läufer sich so bewegen, wie das Licht, das von S über eine spiegelnde Wand w nach M gelangt, d. h., der Läufer müsste sich bis zur Wand w geradlinig auf das durch Spiegelung an w erzeugte Bild M' von M zu bewegen!

Wir wollen diese Idee zu einem mathematisch einwandfreien Beweis gestalten:

Ist M' das Bild von M bei der Spiegelung an w , so gilt für jeden Punkt $P \in w$ auf Grund der Eigenschaften der Geradenspiegelung $|PM| = |PM'|$ und folglich $|SP| + |PM| = |SP| + |PM'|$.

Für die Streckenzüge SPM' sind aber S und M' feste Punkte. Nach Archimedes ist daher die "gerade Linie" von S nach M' über w die Kürzeste, d. h., der kürzeste Streckenzug geht über denjenigen Punkt $P_0 \in w$, der Schnittpunkt der Geraden $g_{SM'}$ mit w ist.

(Wir müssen also noch voraussetzen, dass die Wand w auch so lang ist, dass stets ein solcher Schnittpunkt existiert.)

Eine Begründung kann auch mit Hilfe der Dreiecksungleichung gegeben werden: Wäre P ein von P_0 verschiedener Punkt auf w , dann würden S, P, M' ein Dreieck bilden, und nach der Dreiecksungleichung wäre $|SP| + |PM'| > |SM'| = |SP_0| + |P_0M'|$ (Abb. 2.1b).

Im Punkt P_0 gilt übrigens wirklich das Reflexionsgesetz. Die Winkel mit den Größen α und β'

(siehe Abb. 2.1 b) sind Scheitelwinkel, also ist $\alpha = \beta'$. Nach den Eigenschaften der Geraden-
spiegelung ist β' gleich β ; also stimmen auch α und β und damit der "Einfallswinkel" $\bar{\alpha}$ und
der "Reflexionswinkel" $\bar{\beta}$ überein.

Wir wollen noch kurz bei den aufgezeigten geometrischen Sachverhalten verweilen.

Für alle von P_0 verschiedene Punkte P auf der Geraden w gilt $|SP| + |PM| > |SP_0| + |P_0M|$
(*).

Die Menge der Punkte Q , für die die Summe der Abstände zu S und M konstant und gleich
der Summe $|SP_0| + |P_0M|$ ist, bildet eine Ellipse E , die durch P_0 geht und bei der S und M
die Brennpunkte sind.¹² (Abb. 2.1 c).

Auf Grund der Ungleichung (*) ist P_0 der einzige gemeinsame Punkt der Geraden w mit der
Ellipse E , und damit ist w eine Tangente von E . Da nach unseren bisherigen Darlegungen das
Lot zu w durch P_0 den Winkel SP_0M halbiert, wird der Lichtstrahl von S nach P_0 so an der
Ellipse reflektiert, dass er durch M geht.

Man erkennt nun, dass dies für jeden von S ausgehenden Lichtstrahl gilt, wenn man die
Ellipsentangente durch denjenigen Punkt legt, bei dem der Strahl von S aus auf die Ellipse
trifft (Abb. 2.1 c). Damit ist die Bezeichnung "Brenn"-Punkt für S und M verständlich.

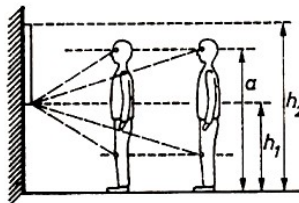


Abb. 2.2

Auch die nächste, unserem täglichen Leben entnommene Aufgabe hat mit Spiegeln zu tun. An
der Wand hängt ein Spiegel, in dem sich ein aufrecht stehender Mann mit der Augenhöhe von
 a cm betrachtet. Das untere Ende des Spiegels sei h_1 cm, das obere Ende h_2 cm über dem
Fußboden; entsprechend der Realität können wir $h_1 < a < h_2$ annehmen. Der Mann stellt
fest, dass er sich nicht ganz bis zu seinen Füßen betrachten kann.

(2.2) Bei welchem Abstand vom Spiegel kann er den größten Teil seines Körpers im Spiegel
sehen?

Eine Spiegelung zeigt, dass sich der Mann bei gleicher Augenhöhe unabhängig vom Abstand
zum Spiegel immer nur von einer bestimmten Höhe an sehen kann. Der nicht sichtbare Teil
hat die konstante Höhe $a - 2(a - h_1) = 2h_1 - a$ cm (Abb. 2.2).

Von diesem überraschenden Effekt kann man sich leicht durch einen Versuch überzeugen.

Bei der nächsten Aufgabenstellung helfen uns Spiegelungen an mehreren Geraden weiter.

(2.3) Im Innern eines spitzen Winkels mit dem Scheitel S und den Schenkeln p und q liege
ein Punkt R . Man konstruiere zwei Punkte $P_0 \in p$ und $Q_0 \in q$ derart, dass der Streckenzug
 $\overline{RP_0Q_0R}$ unter allen Streckenzügen \overline{RPQR} mit $P \in p$ und $Q \in q$ minimale Länge hat. (Dabei
gilt $S \notin p, q$; dies ist eine vielfach zweckmäßige und übliche Auffassung.)

Eine Lösung ergibt sich schnell, wenn wir entsprechend der Lösung von (2.1) zwei Reflexionen
nutzen:

Die Spiegelbilder R^p und R^q des Punktes R bei der Spiegelung an p bzw. q sind fest; also

¹²Zum Verständnis genügt, sich über die übliche Ortsdefinition einer Ellipse an Hand von Nachschlagewerken
oder Kompendien zu informieren, z. B. [5], S. 192ff.

ist die kürzeste Verbindung zwischen beiden Punkten die Strecke $R^p R^q$. Diese schneidet die Schenkel p in P_0 , und q in Q_0 . Dass $\overline{R^p P_0 Q_0 R^q}$ der gesuchte Streckenzug minimaler Länge ist, folgt nun wiederum aus der Invarianz der Länge beim Spiegeln. Doch halt! Ist unsere Lösung einwandfrei? Warum wurde in der Aufgabenstellung ein spitzer Winkel vorausgesetzt?

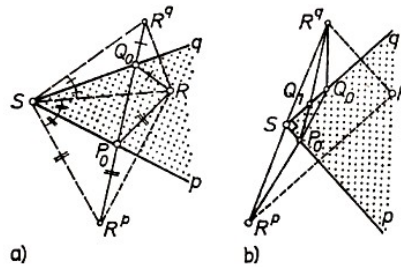


Abb. 2.3 a)

b)

Abb. 2.3a suggeriert uns zwar die Existenz der Punkte P_0 , und Q_0 bewiesen ist sie aber noch nicht!

Nun gilt aber wegen der Invarianz der Winkelgrößen bei Spiegelungen $|\angle SR^+, p| = |\angle SR^{p+}, p|$ und $|\angle SR^+, q| \leq |\angle SR^{q+}, q|$. Da $\angle p, q$ spitz ist, liegen nun p und q im Innern des Winkels $R^p S R^q$. Folglich schneidet die Strecke $R^p R^q$ tatsächlich die Halbgeraden p und q .

Wir sind auf diese Existenzfrage darum so ausführlich eingegangen, weil die Existenz eines Extremums keine Selbstverständlichkeit ist. Ein einfaches Beispiel mag das verdeutlichen.

(2.4) Gegeben sei eine Strecke \overline{AB} . Man ermittle eine Strecke \overline{PQ} größter Länge, deren Endpunkte im Innern von \overline{AB} liegen.

Eine solche Strecke gibt es nicht! Denn sind P_0, Q_0 verschiedene Punkte im Innern von \overline{AB} und liegt - was wir o. B. d. A. annehmen können - der Punkt P_0 zwischen A und Q_0 , dann gibt es einen Punkt P_1 zwischen A und P_0 , und es gilt $|P_1 Q_0| > |P_0 Q_0|$. Das heißt, zu jeder Strecke kann eine längere angegeben werden, deren Endpunkte ebenfalls innere Punkte von \overline{AB} sind.

In der Aufgabe (2.3) hatten wir zum Existenznachweis des Minimums die Spitzwinkligkeit benutzt. Damit ist zunächst nicht gesagt, dass diese Voraussetzung dafür notwendig ist.

(2.5) Man untersuche die Aufgabenstellung aus (2.3) unter der Voraussetzung, dass der Winkel $\angle p, q$ ein rechter oder stumpfer ist.

Zur Lösung sei zunächst $\angle p, q$ rechtwinklig. Dann liegt S in der Strecke $\overline{R^p R^q}$; der Punkt S ist sogar der Mittelpunkt dieser Strecke. Zu jedem Streckenzug $\overline{R^p P_0 Q_0 R^q}$ mit $P_0 \in p$ und $Q_0 \in q$ gibt es aber offenbar einen Punkt P_1 zwischen S und P_0 oder einen Punkt Q_1 zwischen S und Q_0 derart, dass der neue Streckenzug $\overline{R^p P_1 Q_0 R^q}$ oder $\overline{R^p P_0 Q_1 R^q}$ kürzer als der alte ist (Abb. 2.3b). Es gibt also keinen Streckenzug minimaler Länge.

Ist $\angle p, q$ stumpfwinklig, dann liegen R, p und q nicht im Innern des Winkels $\angle R^p S R^q$, und man erkennt in entsprechender Weise, dass kein Minimum existiert.

(2.6) • Im Innern eines Winkels $\angle p, q$ liegen voneinander verschiedene Punkte A und B . Man ermittle Punkte $P_0 \in p$ und $Q_0 \in q$ derart, dass der Streckenzug $\overline{A P_0 Q_0 B}$ unter allen Streckenzügen dieser Art minimale Länge hat.

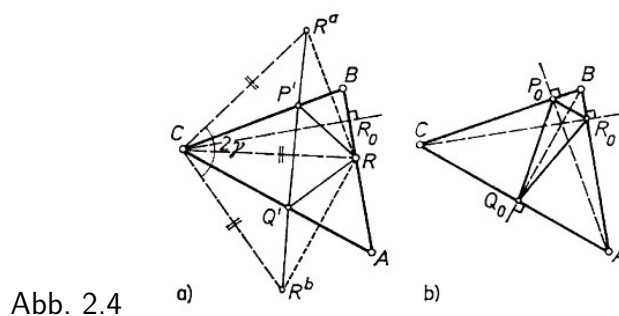
Für welche Größen des Winkels $\angle p, q$ gibt es einen minimalen Streckenzug? Ändern sich die Ergebnisse, wenn wir neben den Streckenzügen $\overline{A P Q B}$ die Streckenzüge $\overline{A Q P B}$ mit $P \in p$

und $Q \in q$ in die Betrachtung einbeziehen?

An die Aufgabe (2.3) lässt sich nun vorteilhaft eine Aufgabe anschließen, die in die Aufgaben-
gruppe über einbeschriebene Figuren mit extremalen Eigenschaften gehört.

(2.7) Gegeben ist ein spitzwinkliges Dreieck ABC . Diesem Dreieck ist ein Dreieck $P_0Q_0R_0$ so
einzubeschreiben, dass es unter allen einbeschriebenen Dreiecken PQR eines mit minimalem
Umfang ist. ("Einbeschrieben" bedeutet dabei, dass die Eckpunkte auf dem Rand des vorge-
gebenen Dreiecks liegen.)

Unter allen möglichen Lösungswegen ergibt sich einer aus den Überlegungen zu (2.3). Es sei
 R ein beliebiger, aber festgehaltener Punkt auf der Seite \overline{AB} (Abb. 2.4a). Wir konstruieren
zu diesem festen Punkt R zwei Punkte $P' \in CB^+$ und $Q' \in CA^+$ wie in der Aufgabe (2.3)
so, dass der Umfang des Dreiecks $RP'Q'$ möglichst klein wird.



Auf Grund der Spitzwinkligkeit des Dreiecks ABC sind P' und Q' innere Punkte von \overline{CB}
bzw. \overline{CA} . Jetzt müssen wir noch untersuchen, wie sich der Umfang aller solcher Dreiecke bei
Veränderung der Lage des Punktes R auf der Seite \overline{AB} ändert.

Der Umfang ist aber gerade die Länge der Strecke R^aR^b (Abb. 2.4a). Diese Strecke ist
eine Seite des Dreiecks R^aCR^b . Was verändert sich mit R an diesen Dreiecken, was bleibt
unverändert?

Alle diese Dreiecke sind gleichschenkelig, denn es gilt auf Grund der Spiegelungen $|CR^a| =$
 $|CR| = |CR^b|$; alle diese Dreiecke haben bei C einen gleichgroßen Winkel, nämlich einen mit
der Größe 2γ . (Siehe Lösung zu (2.3)!)
Für alle diese gleichschenkligen Dreiecke wird die Basis um so kleiner, je kleiner die Länge der
Schenkel, also $|CR|$ ist.

Die Länge von \overline{CR} ist aber genau dann minimal, wenn R der Fußpunkt des Lotes von C
auf die Seite \overline{AB} ist. Wegen der 2.2. Geometrische Extremwertaufgaben Spitzwinkligkeit ist
dieser Fußpunkt ein innerer Punkt der Strecke \overline{AB} .

Ein Dreieck $P_0Q_0R_0$ der gesuchten Art ergibt sich also genau dann, wenn R_0 Höhenfußpunkt
auf AB ist und die Punkte P_0 und Q_0 wie in (2.3) konstruiert werden.

Nun hätte man aber auch die gleiche Überlegung von einem zunächst fest gewählten Punkt
 $P \in \overline{BC}$ aus anstellen können und wäre zu dem Ergebnis gekommen, dass P_0 Höhenfußpunkt
auf \overline{BC} ist. Das gesuchte Dreieck ist also das Höhenfußpunktdreieck.¹³

Da in den Punkten P_0, Q_0 und R_0 das Reflexionsgesetz erfüllt ist und die Höhen im Dreieck
 ABC die Einfallslote darstellen, folgt aus unserer Überlegung noch:

¹³Die Aufgabe (2.7) stammt von dem Italiener J. F. Fagano, und die von uns angegebene Lösung entspricht
weitgehend einer von Frater Gabriel-Marie, dem Verfasser des bedeutenden Buches "Exercices de Geome-
trie".

(2.8) Die Höhen eines spitzwinkligen Dreiecks sind gleich den Winkelhalbierenden im zugehörigen Höhenfußpunktdreieck (Abb. 2.4 b).

Zur Lösung der nächsten Aufgaben, die wir dem Leser empfehlen, können die bisherigen Darlegungen sehr nützlich sein.

(2.9) • Man beschreibe einem Quadrat ein Quadrat minimalen Umfangs ein.

(2.10) • Im Innern eines Rechtecks $ABCD$ liege ein Punkt O . Man bestimme unter allen Streckenzügen $OPQRSO$ mit $P \in \overline{AB}$, $Q \in \overline{BC}$, $R \in \overline{CD}$, $S \in \overline{DA}$ einen solchen minimaler Länge.

Die Lösung der letzten Aufgabe führt übrigens auch auf eine Lösung folgender Billard-Aufgabe:

(2.10') • Auf einem rechteckigen Billardtisch liegt eine Kugel. Wie ist diese zu stoßen, damit sie jede Bande genau einmal berührt und danach zum Ausgangspunkt zurückrollt?

Neben den Geradenspiegelungen sind auch andere Bewegungen zur elementaren Lösung von Extremwertaufgaben über Streckenzügen nützlich. Das wollen wir an einigen Beispielen demonstrieren.

(2.11) Gegeben seien ein Streifen mit parallelen Randgeraden c und d sowie zwei Punkte A und B , die durch den Streifen voneinander getrennt liegen (Abb. 2.5). Man bestimme solche Punkte $C \in c$ und $D \in d$, für die $g_{CD} \perp c$ ist und der Streckenzug \overline{ACDB} kleinste Länge hat.

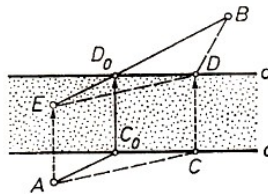


Abb. 2.5

Diese Aufgabenstellung kann man sich wie folgt veranschaulichen: Auf zwei Seiten eines breiten (geradlinig verlaufenden) Flusses liegen zwei Orte A und B . An welcher Stelle des Flusses ist senkrecht zu seinen Ufern eine Brücke anzulegen, damit die Orte auf dem kürzesten Wege verbunden werden können?

Zur Lösung betrachten wir zunächst zwei beliebige Punkte $C \in c$ und $D \in d$ mit $g_{CD} \perp c$ (Abb. 2.5). Gedanklich können wir die Reihenfolge der Teilwege \overline{AC} , \overline{CD} und \overline{DB} vertauschen; gehen wir also zuerst von A aus einen \overline{CD} entsprechenden Weg.

Das bedeutet, dass wir die Verschiebung \overrightarrow{CD} auf A anwenden; sie bildet A auf einen Punkt E ab (Abb. 2.5). Es ist $|AC| = |ED|$. Die Summe $|AC| + |DB|$ ist nun minimal genau dann, wenn das für $|ED| + |DB|$ zutrifft.

Der Punkt E liegt mit B auf verschiedenen Seiten von der Geraden d . Folglich schneidet d die Strecke \overline{EB} in einem Punkt D_0 , und $\overline{ED_0B}$ ist unter allen Streckenzügen \overline{EDB} mit $D \in d$ derjenige mit minimaler Länge.

Mit dem Bild C_0 von D_0 bei der Verschiebung \overrightarrow{DC} erhalten wir nun offensichtlich die Lösung $\overline{AC_0D_0B}$ der Aufgabe.

(2.12) Gegeben sei ein spitzwinkliges Dreieck ABC . Gesucht ist die Menge aller Punkte P im Innern des Dreiecks, für die die Abstandssumme $|AP| + |BP| + |CP|$ minimal wird.

Die größere Schwierigkeit gegenüber den vorangegangenen Problemstellungen besteht insbe-

sondere darin, dass es sich diesmal nicht um einen Streckenzug handelt; die Strecken gehen von einem Punkt aus. Um Anschluss an die Archimedische Formulierung zu gewinnen, wäre ein Streckenzug mit festen Endpunkten und gleicher Länge zu finden.

Wir müssten also z. B. versuchen, die Strecke \overline{AP} "zwischen \overline{BP} und \overline{CP} einzufügen". Das schafft eine Drehung um A mit 60° ! (Vgl. Abb. 2.6).

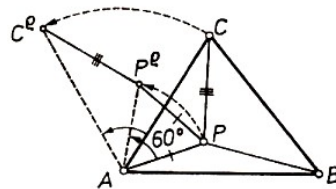


Abb. 2.6

Ist ρ eine Drehung um A mit 60° , dann gilt $|AP| = |PP^\rho|$, und damit ist die Strecke \overline{AP} "eingefügt". Der Streckenzug $\overline{BPP^\rho C^\rho}$ hat die Länge $|AP| + |BP| + |CP|$. Überdies sind seine Endpunkte B und C^ρ unabhängig von der Lage des Punktes P im Innern des Dreiecks. Also muss P so gewählt werden, dass P und P^ρ auf der Strecke BC^ρ liegen, um Minimalität zu erreichen.

Also muss notwendigerweise $|\angle C^\rho P^\rho P| = |\angle P^\rho P B| = 180^\circ$ und wegen der Gleichseitigkeit des Dreiecks APP^ρ dann $|\angle APB| = 120^\circ$ und $|\angle AP^\rho C^\rho| = |\angle APC| = 120^\circ$ und damit auch $|\angle PBC| = 120^\circ$ sein.

Wenn es also Punkte P der gesuchten Art gibt, dann sind es diejenigen, von denen aus jede der Dreiecksseiten unter einem Winkel von 120° erscheinen, bzw. diejenigen, die "nach den Ecken einen regelmäßigen Dreistrahl senden", wie es etwas altmodisch, aber anschaulich bei Sturm in [14], S. 58, heißt.

Konstruiert man über \overline{AC} und \overline{BC} nach außen jeweils ein gleichseitiges Dreieck und die dazugehörigen Umkreise, so muss P nach dem Peripheriewinkelsatz auf diesen Kreisen und außerdem im Innern des Dreiecks liegen. Ist der Winkel bei C - wie vorausgesetzt - ein spitzer Winkel, so schneiden sich (wie man sich leicht überlegt) die beiden Umkreise (außer in C noch) in einem Punkt P im Innern des Dreiecks.

Dieser Punkt P hat die oben beschriebenen Eigenschaften. Es existiert also ein solcher Punkt P , und zwar eindeutig.

(2.13) • a) Ändert sich die Lösung der Aufgabe (2.12), wenn wir uns nicht auf die inneren Punkte des Dreiecks beschränken?

b) Ändert sich die Lösung der Aufgabe (2.12), wenn wir die Voraussetzung der Spitzwinkligkeit fallen lassen?

Die Aufgabe (2.12) ist eine der "strapaziertesten" Extremwertaufgaben. Sie ist schon in der von Fermat 1629 veröffentlichten "Abhandlungen über Maxima und Minima" enthalten.

Über Martin Mersenne gelangte sie - ohne Lösung - in die Hände der Galilei-Schüler Cavalieri, Torricelli und Viviani. Nach Torricelli nennt man heute auch noch den Punkt P den Torricellischen Punkt.

Im 19. Jh. beschäftigte sich auch der berühmte Geometer Jakob Steiner (1796-1863) in seiner umfangreichen Abhandlung "Über Maximum und Minimum in der Ebene, auf der Kugelfläche und im Raum überhaupt" mit dieser Aufgabe. Es ist nicht überraschend, dass es sehr viele verschiedene Lösungen zu dieser Aufgabe gibt; in [12] findet man eine weitere (und ein räumliches Analogon), auf eine andere kommen wir später noch zu sprechen.

Wiederum regt eine solche Aufgabe an, ähnliche zu formulieren und zu untersuchen. Wenn wir nun in der Aufgabe (2.12) das Dreieck durch ein n -Eck oder die Summe durch eine Differenz ersetzen?

(2.14) • Gegeben ist in einer Ebene ε ein konvexes Viereck $ABCD$. Gesucht ist die Menge aller Punkte $P \in \varepsilon$, für die die Abstandssumme $|AP| + |BP| + |CP| + |DP|$ minimal wird.

Ersetzt man in (2.14) das Viereck durch ein konvexes n -Eck (mit $n \geq 5$), so lassen sich im allgemeinen nur Näherungslösungen ermitteln. (Vgl. auch Abschnitt 3.4.)

(2.15) • Es sei ABC ein gleichseitiges Dreieck. Gesucht ist die Menge derjenigen Punkte P , für die die Größe $s = |AP| + |BP| - |CP|$ minimal wird.

2.2 Kürzeste Streckenzüge und Linien im Raum

Fundament des zu Erreichenden
muss Erreichtes sein.
J. Brežan

Einige Problemstellungen und Aufgaben aus Abschnitt 2.1 lassen sich auf den Raum übertragen, und es zeigt sich, dass gewisse Fragen hinsichtlich kürzester Streckenzüge im Raum analog zu denen der ebenen Geometrie betrachtet werden können.

Ein Beispiel ist folgende Aufgabe

(2.16) Gegeben seien eine Ebene ε und zwei Punkte A und B , die im gleichen Halbraum bezüglich ε liegen. Man bestimme diejenigen Punkte $P \in \varepsilon$, für die die Abstandssumme $|AP| + |PB|$ minimal ist.

Eine dazu analoge Aufgabenstellung in der ebenen Geometrie ist (2.1). Dort erzielten wir eine Lösung mit Hilfe der Spiegelung an der Geraden w . Analog spiegeln wir hier nun den Punkt B an der Ebene ε . Das Bild B' und A liegen dann in verschiedenen Halbräumen bezüglich ε , und deshalb schneidet ε die Strecke AB' in einem Punkt P_0 (Abb. 2.7).

Für jeden Punkt $P \in \varepsilon$ gilt $|PB| = |PB'|$ und damit $|AP| + |PB| = |AP| + |PB'|$. Diese Längensumme ist nach der Dreiecksungleichung offensichtlich für den Fall $P = P_0$ und nur für diesen minimal.

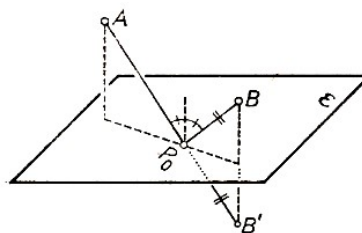


Abb. 2.7

Errichtet man auf ε in P_0 das Lot, so liegt dieses zusammen mit den Punkten B , B' und A in einer gemeinsamen Ebene, in der nun wieder das "Reflexionsgesetz" ersichtlich ist (Abb. 2.7).

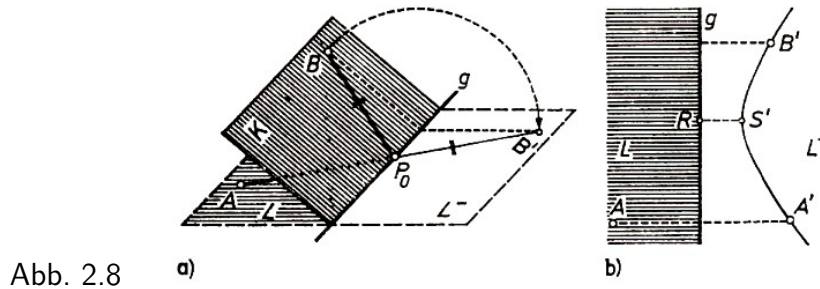
Eine neue Frage entsteht, wenn wir in der Aufgabe (2.16) die Ebene durch eine Gerade g ersetzen. Damit ein räumlicher Sachverhalt vorliegt, sollen A , B und die Gerade g nicht in einer gemeinsamen Ebene liegen. Um eine einfache Veranschaulichung zu ermöglichen, betrachten wir folgende konkrete Vorgabe

(2.17) Es sei $ABCDEFGH$ ein Würfel. Man bestimme diejenigen Punkte P auf der Geraden

$g = g_{EG}$, für die $|AP| + |PB|$ minimal wird.

Dem Leser wird empfohlen, sich einen Schrägriss des Würfels vorzugeben und darin die folgenden Überlegungen zur Lösung konstruktiv nachzuvollziehen. (Man kann dabei sein räumliches Vor- und Darstellungsvermögen überprüfen.)

Es liegt nahe, auch hier die Aufgabenstellung in geeigneter Weise auf die kürzeste Verbindung zweier Punkte zurückzuführen. Dies gelingt in diesem Fall aber nicht mit einer Spiegelung an der Geraden g . Hier hilft eine andere Bewegung des Raumes weiter. Die Gerade g bildet den Rand zweier Halbebenen L und K , die den Punkt A bzw. B enthält (Abb. 2.8 a).



Nun kann durch eine Drehung um die Gerade g die Halbebene K in die zu L entgegengesetzte Halbebene L^- übergeführt werden. Dabei ist $|PB| = |PB'|$ für das Bild B' von B und für jeden Punkt $P \in g$. Nun ist nach dem gleichen Schluss wie bei der Lösung von (2.1) der Schnittpunkt P_0 von g mit $\overline{AB'}$ der gesuchte Punkt.

Es lässt sich zeigen, dass die Winkelhalbierende w des Winkels $\angle AP_0B$ eine Senkrechte zu der Geraden g durch P_0 ist. Die zu w senkrechte Ebene η durch P_0 enthält dann die Gerade g . Folglich wird der Lichtstrahl von A aus nach P_0 so an der Ebene η reflektiert, dass er weiter durch B geht.

In der Praxis lässt sich dieser Reflexionseffekt z. B. gut an blanken Fahrradspeichen beobachten. Die Speiche besitzt auf Grund ihrer Dicke eine reflektierende Oberfläche, die ausreicht, die Reflexion einer (punktförmigen) Lichtquelle zu beobachten. Besonders deutlich sichtbar wird das bei einer Rotation des Rades.

Ausgehend von (2.17) stellt sich noch folgende Extremwertaufgabe

(2.18)* Es seien A, B zwei Punkte und g eine Gerade, die in keiner gemeinsamen Ebene liegen. Für welche Punkte C im Innern der Strecke \overline{AB} gibt es einen kürzesten Streckenzug von A über einen Punkt der Geraden g nach C ?

Wir gehen von den Überlegungen zur Lösung der Aufgabe (2.17) aus. (Siehe nochmals Abb. 2.8a.) Zur Bestimmung des kürzesten Streckenzuges von A über einen Punkt auf g nach einem beliebig vorgegebenen Punkt C auf der Verbindungsgeraden h der Punkte A und B ist der Punkt C durch eine Drehung um die Gerade g in die Halbebenen H^- überzuführen.

Die Menge der so erhaltenen Bildpunkte C' kann keine Gerade sein. Eine solche Gerade müsste entweder zu g parallel sein oder g schneiden. Im ersten Fall hätten alle Punkte von h den gleichen Abstand zu g , und im zweiten Fall gäbe es einen Punkt auf der Geraden h , der auf g liegt. Beide Fälle widersprechen der Voraussetzung, dass die Geraden g und $h = g_{AB}$ windschief zueinander sind.

Welche Gestalt hat nun die Menge der Bilder C' ? Bei der Drehung der zu g windschiefen

Geraden h um die Gerade g entsteht ein einschaliges Hyperboloid¹⁴ mit der Achse g . Der Schnitt dieser Fläche mit der Halbebene H^- ergibt einen Ast einer Hyperbel (Abb. 2.8 b).

Bekanntlich gibt es genau zwei Punkte $S \in h$ und $R \in g$, die den kürzesten Abstand je zweier Punkte dieser Geraden bilden.. (Siehe etwa [12]) Für sie ist $g_{RS} \perp g, h$. Demnach ist das Bild S' von S der Scheitel des Hyperbelbogens und R der Fußpunkt des Lotes von S' auf g (Abb. 2.8 b).

Die Aufgabe (2.18) besitzt also genau dann eine (und dann eindeutig bestimmte) Lösung, wenn unter den Punkten auf dem Hyperbelbogen zwischen den Bildern A' und B' von A und B ein Punkt mit dem kürzesten Abstand zu dem Punkt A existiert.

Beziehen wir die Aufgabe (2.18) auf die speziellen Vorgaben in (2.17), also $g = g_{EG}$ und $h = G_{AB}$, dann lässt sich auch ohne diese Betrachtungen eine Antwort rasch finden.

Man kann aber auch von ebenen Problemstellungen ausgehen und durch Analogiebetrachtungen zu räumlichen gelangen. Wir wollen einmal die Aufgabe (2.6) entsprechend variieren. Anstelle des Winkels $\angle p, q$ könnten zwei Halbebenen K und L mit gemeinsamer Randgeraden gewählt werden, die Punkte A und B wären durch zwei Punkte M und N im Raum zu ersetzen. Spezialisieren wir nun diese sehr allgemeine Situation wiederum, so gelangt man z. B. zu der folgenden Aufgabe.

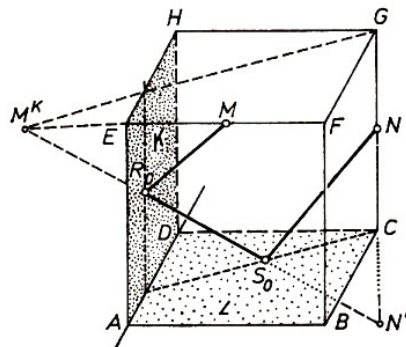


Abb. 2.9

(2.19) Es sei $ABCDEFGH$ ein Würfel, und K bzw. L seien diejenigen Halbebenen mit der Randgeraden g_{AD} , die den Punkt E bzw. B enthalten. Ferner seien M und N die Mittelpunkte der Strecken \overline{EF} bzw. \overline{CG} .

Man ermittle Punkte $R_0 \in K$ und $S_0 \in L$ derart, dass der Streckenzug $\overline{MR_0S_0N}$ unter allen Streckenzügen dieser Art minimale Länge hat. (Wir empfehlen dem Leser, die Überlegungen zur Lösung schrittweise in einem Schrägriss konstruktiv darzustellen.)

Zur Lösung benutzen wir Analogiebetrachtungen. Wie in der Lösung zu (2.6) spiegeln wir, jetzt aber M an der Ebene ε_{ADE} und N an der Ebene ε_{ADB} und erhalten die Bildpunkte M^K und N^L (Abb. 2.9).

Für alle Punkte $R \in K$ und $S \in L$ gelten $|MR| = |M^KR|$ sowie $|SN| = |SN^L|$. Schneidet die Strecke $\overline{M^KN^L}$ die Halbebenen K und L in Punkten R_0 und S_0 , dann ist nun offenbar $\overline{MR_0S_0N}$ der gewünschte Streckenzug. Man überzeugt sich selbst davon, dass diese Schnittpunkte hier existieren; sie lassen sich im Schrägriss leicht konstruieren (Abb. 2.9).

Wer Spaß daran findet, kann nun selbst räumliche Problemstellungen aus ebenen ableiten.

Ein weiterer Aufgabentyp entsteht, wenn man nach kürzesten Wegen auf Oberflächen von Körpern fragt, eine Problemstellung, die durchaus von großem praktischen Interesse sein kann.

¹⁴Näheres findet man in leicht zugänglichen Fachbüchern, etwa [5], S. 199 und S. 597

Möchte jemand auf unserer Erde von Berlin nach Hanoi gelangen, so interessiert kaum der kürzeste Weg durch unsere Erde hindurch, sondern wohl der auf der Erdoberfläche. Wir beginnen aber mit einer sehr viel einfacheren Aufgabe:

(2.20) Auf einem Würfel $ABCDEFGH$ bestimme man die kürzesten Wege von A nach G .

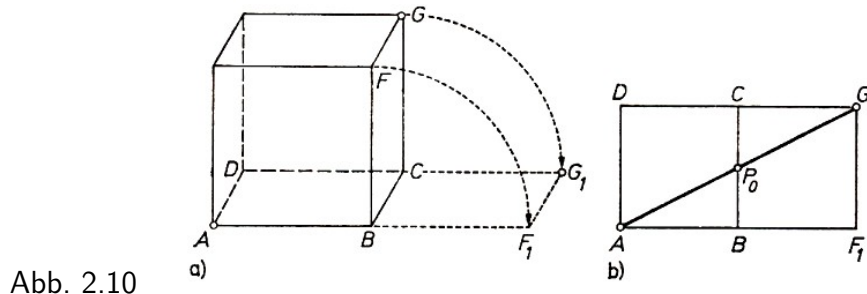


Abb. 2.10

Der geschlossene Streckenzug $BCDHEFB$ zerlegt die Oberfläche des Würfels so in zwei Teile, dass in jedem genau einer der Punkte A und G liegt (Abb. 2.10 a). Folglich enthält jeder Weg auf der Oberfläche von A nach G einen Punkt P dieses Streckenzuges.

Unter allen Wegen, die speziell über einen Punkt P der Strecke \overline{BC} verlaufen, ist der Streckenzug AP_0G mit dem Mittelpunkt P_0 von \overline{BC} der kürzeste. Die Seitenflächen $ABCD$ und $FBCG$ haben nämlich die Kante \overline{BC} gemeinsam, und beide lassen sich (durch eine Drehung um g_{BC}) verebnen (Abb. 2.10 b). Dabei bleiben Weglängen invariant.

Nun ergibt der Schnittpunkt P_0 der Geraden durch A und G_1 , mit der Strecke \overline{BC} , also der Mittelpunkt dieser Strecke tatsächlich den kürzesten Weg.

Er hat die Länge $a\sqrt{5}$, wenn a die Kantenlänge des Würfels ist. - Entsprechendes gilt für die Mitten der übrigen fünf Strecken von $\overline{BCDHEFB}$. Es gibt also sechs Streckenzüge als Kürzeste.

(2.21) • Man beantworte für einen Quader $ABCDEFGH$ mit den Kantenlängen $|AB| = a$, $|AD| = b$ und $|AE| = c$, für die $a < b < c$ gilt, die in (2.20) formulierte Frage.

(2.22) • Es sei $ABCDEFGH$ ein Würfel, und P sei

a) der Mittelpunkt der Seitenfläche $ABCD$ bzw.

b) der Mittelpunkt der Kante \overline{AB} .

Man bestimme diejenigen Punkte auf der Oberfläche des Würfels, die bezüglich der Längenmessung auf der Oberfläche am weitesten von P entfernt sind.

Geschenkpäckungen sind mitunter durch eine goldfarbene Gummischnur zur Zierde eingebunden. Eine solche Packung können wir als quaderförmig annehmen. Ferner nehmen wir an, dass der Gummifaden nur einmal mit sich geknotet ist, also im ganzen eine einzige Schleife bildet. Auf Grund seiner Elastizität wird der Gummifaden eine Lage minimaler Länge einnehmen. (Von Haftreibung wollen wir gänzlich absehen.)

Dann stellt sich hier offenbar wieder die Frage nach einem Streckenzug minimaler Länge.

(2.23) Welche Lagen kann ein geschlossener Gummifaden einnehmen, der um einen Würfel liegt, ohne sich zu überschneiden?

Der Würfel sei dabei als ideal glatt angenommen und die Elastizität des Gummifadens so groß, dass dieser sich sogar auf einen Punkt zusammenziehen kann.

Zunächst ist einsichtig, dass der Gummifaden nicht allein auf drei Würfelseiten, also auf drei paarweise benachbarten Seitenflächen liegen kann. Er würde sich auf die gemeinsame Würfelcke zusammenziehen.

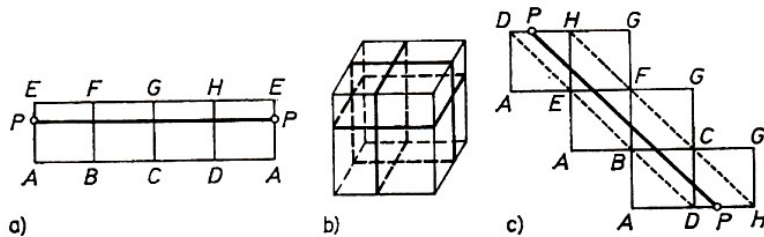


Abb. 2.11 a)

b)

c)

Soll der Faden über vier Würfelseiten laufen, so kommen offenbar nur solche Seiten in Betracht, die bei der Auffassung des Würfels als viereckiges gerades Prisma eine Mantelfläche darstellen. Eine solche Mantelfläche lässt sich in die Ebene abwickeln (Abb. 2.11 a). Der kürzeste Streckenzug auf dieser Mantelfläche von einem Punkt P einer Kante, etwa \overline{AE} , nach P zurück muss parallel zu den Kanten der Grund- und Deckfläche verlaufen.

Es gibt drei verschiedene Sorten einer derartigen Lage des Gummifadens. Insgesamt gibt es durch jeden (inneren) Punkt einer Seitenfläche des Würfels genau zwei derartige Lagen des Gummifadens (Abb. 2.11 b).

Dies sind nicht alle Lagen!

Da man einen Würfel $ABCDEFGH$ so mit einer Ebene schneiden kann, dass ein Sechseck entsteht - etwa durch die Mitten der Kanten \overline{BC} , \overline{CD} , \overline{DH} , \overline{HE} , \overline{EF} und \overline{FB} -, scheint eine Lage des Gummifadens mit minimaler Länge über alle Seitenflächen des Würfels möglich. In der Tat zeigt ein Netz des Würfels in der Art wie in der Abb. 2.11c, dass eine solche Lage gerade bei einem Verlauf des Gummifadens parallel zu Diagonalen der Seitenflächen eintritt. Demnach gibt es durch jeden (inneren) Punkt einer Seitenfläche genau zwei derartige Lagen des Gummifadens. Er hat dabei die Länge $3\sqrt{2}a$ ($> 4a$).

Damit ist die Aufgabe (2.23) vollständig gelöst.

(2.24) • Man löse die in (2.23) gestellte Aufgabe für einen Quader mit verschiedenen langen Kanten.

Die Aufgaben (2.20) bis (2.24) konnten durch Verebnung, also auf Grund der möglichen Abwicklung der Oberfläche des betrachteten Körpers in die Ebene einer einfachen Lösung zugeführt werden.

Neben den Polyedern gibt es noch andere Körper dieser Art, z. B. Kreiszyylinder oder Kreiskegel.

(2.25) • Es sei P ein Punkt auf der Mantelfläche eines geraden Kreiskegels. Man bestimme kürzeste Linien auf der Mantelfläche, die von P nach P führen, mit allen Mantellinien einen gemeinsamen Punkt besitzen und die Spitze S des Kegels nicht enthalten.

Welche Rolle spielt die Größe des Öffnungswinkels α des Kegels (d. h. die Größe des Winkels, den zwei gegenüberliegende Mantellinien bilden) bei der Lösung der Aufgabe?

Praktisch lässt sich hier mit einem Gummifaden eine Näherungslösung gewinnen.

Jetzt wenden wir uns der schon angedeuteten Frage nach den kürzesten Linien auf der Oberfläche der Erde zu. Ihre Beantwortung spielt z. B. für Schiffs- und Flugverbindungen eine große Rolle, nicht zuletzt aus ökonomischen Gründen.

In erster Näherung kann dabei die Oberfläche der Erde als Kugeloberfläche aufgefasst werden. Da diese sich aber nicht auf eine Ebene abwickeln lässt¹⁵, kann die obige Frage nicht wie bisher

¹⁵Ein Beweis dafür wird mit Mitteln und Methoden der Differentialgeometrie geführt.

auf die Frage nach kürzesten Streckenzügen zurückgeführt werden.

Die kürzesten Linien lassen sich dennoch leicht beschreiben:

Eine kürzeste Verbindung zwischen zwei Punkten P und Q auf der Kugeloberfläche ist ein Teil eines Großkreises der Kugel¹⁶, der P und Q enthält.¹⁷

Sind also P und Q zwei zueinander diametral gegenüberliegende Punkte, dann gibt es unendlich viele kürzeste Verbindungen von P nach Q . (Wir können P und Q etwa als Nord- und Südpol auffassen, und die Kürzesten sind dann die Meridianlinien.)

Andernfalls gibt es genau eine kürzeste Linie.

(2.26) • Es seien P und Q zwei Punkte auf der nördlichen Halbkugel. Man bestimme einen kürzesten Rundkurs eines Flugzeuges, der von P zur Äquatorlinie und weiter über Q zurück nach P führt. (Konkret könnten für P und Q z. B. die Städte Berlin und Hanoi sowie ihre ganzzahligen Längen- und Breitengrade gewählt werden.)

Zur Lösung von Aufgaben dieser Art und einer Reihe anderer Fragestellung auf der Kugelfläche ist die sphärische Trigonometrie entwickelt worden.

2.3 Extremale Flächen

Das ist ein weites Feld,
von dem ihr da redet ...

A. Stifter

Isoperimetrische Aufgaben

Unter isoperimetrischen ebenen Figuren versteht man solche mit gleichem Umfang. Eine isoperimetrische Aufgabe besteht im wesentlichen nun darin, unter allen Figuren, die gewissen Bedingungen genügen und den gleichen Umfang besitzen, diejenigen mit extremalem Flächeninhalt zu bestimmen.

Diese Aufgabenstellung kann in mancher Hinsicht noch modifiziert sein. Wir beginnen mit einem einfachen Beispiel.

(2.27) Man bestimme unter allen Rechtecken mit dem gleichen Umfang diejenigen mit dem größten Flächeninhalt.

Das Ergebnis lässt sich leicht erraten: das Quadrat. Einen Beweis für diese Behauptung wird aber vielleicht nicht jeder gleich bei der Hand haben.

Wir gehen von unserer Vermutung aus!

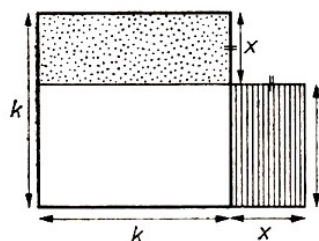


Abb. 2.12

¹⁶Die Großkreise der Kugel können einfach als diejenigen Kurven verstanden werden, die als Schnitt der Kugeloberfläche mit einer Ebene durch den Mittelpunkt der Kugel entstehen.

¹⁷Dieser Sachverhalt ist anschaulich klar und einfach; ein strenger Beweis wird im Rahmen der Differentialgeometrie geführt.

Der als konstant vorausgesetzte Umfang sei $4k$. Dann existiert genau ein Quadrat mit diesem Umfang. Wir zeigen, dass jedes Rechteck, welches den gleichen Umfang hat, aber kein Quadrat ist, einen kleineren Flächeninhalt als dieses Quadrat hat. Betrachten wir dazu Abb. 2.12!

Verlängern wir zwei der Quadratseiten um x , so müssen die beiden anderen um x verringert werden. Der Flächeninhalt des Quadrates nimmt um $x \cdot k$ ab und um $x \cdot l$ zu. Nun ist l aber kleiner k , also würde sich insgesamt der Flächeninhalt verringern.

Später lernen wir weitere Beweismöglichkeiten kennen. Gerade diese Aufgabe wird uns ein treuer Begleiter durch verschiedene Kapitel bleiben.

(2.28) • Man zeige: Unter allen Dreiecken mit gegebenem Umfang u und gegebener Seitenlänge $c = |AB|$ hat das gleichschenklige Dreieck den größten Flächeninhalt.

Die grundlegende Voraussetzung, dass gleicher Umfang besteht, kann weitgehend modifiziert vorliegen.

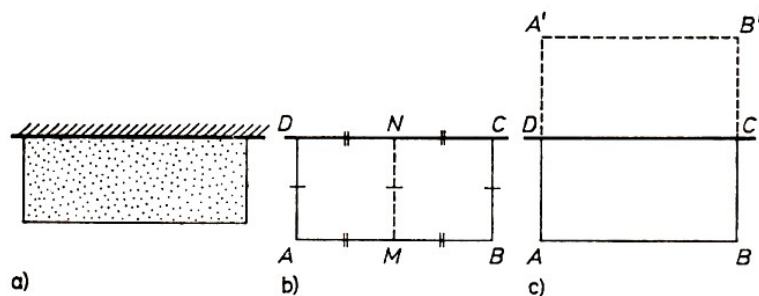


Abb. 2.13 a)

(2.29) An einer (hinreichend langen) Mauer soll ein rechteckiger Garten angelegt werden. Zu einer Einzäunung stehen 32 m Maschendraht zur Verfügung. Wie erzielt man eine möglichst große Gartenfläche? (Siehe Abb. 2.13a.)

Diese Aufgabe lässt sich auf zweierlei Weise auf die in (2.27) vorliegende Fragestellung zurückführen. Halbieren wir die rechteckige Gartenfläche wie in Abb. 2.13 b) angegeben, so hat das Rechtecke $BCNM$ offenbar den Umfang

$$2|BC| + 2\left(\frac{1}{2}|AB|\right) = |AB| + 2|BC|$$

also gleich der konstanten Maschendrahtlänge.

Nach (2.27) hat nun das Rechteck $BCNM$ und damit das Rechteck $ABCD$ genau dann maximalen Flächeninhalt, wenn $BCNM$ ein Quadrat und damit $|AB| = 2|BC|$ ist. Ein anderer Beweis ergibt sich, wenn man die "Gartenfläche" an der "Mauer" spiegelt (vgl. Abb. 2.13 c).

Das entstehende Rechteck $ABB'A'$ hat dann den Umfang von 64 m und nach (2.27) genau dann maximalen Flächeninhalt, wenn es ein Quadrat ist, d. h., wenn $|AB| = 2|BC|$ gilt. Die maximale Gartenfläche hat also eine Größe von $16 \text{ m} \cdot 8 \text{ m} = 128 \text{ m}^2$.

Ebenfalls anschaulich sehr einleuchtend erscheint

(2.30) Unter allen ebenen konvexen Figuren mit konstantem Umfang besitzt der Kreis den größten Flächeninhalt.

Doch ein Beweis ist keinesfalls so einfach zu führen. Erste Beweise hat erst Jacob Steiner gegeben. Wir müssen hier auf nähere Beweisüberlegungen verzichten.

Unter Verwendung von (2.30) löse man

(2.31) • Es sei $\angle p, q$ ein rechter Winkel. Wie ist ein (nicht notwendig geradliniger) Schnitt vorgegebener Länge von einem Schenkel durch das Innere des Winkels bis zum anderen Schenkel zu führen, damit ein möglichst großes Flächenstück abgetrennt wird?

Man könnte eine Vielzahl von Aufgaben angeben, die zu der von uns angesprochenen Fragestellung gehören. Wir haben uns auf einige wenige beschränken müssen. In der abschließenden Aufgabe ist nicht nur der Umfang konstant vorgegeben, sondern auch seine Teile (Seiten), aus denen er sich zusammensetzt.

(2.31)* • Gegeben sei ein konvexes Viereck $ABCD$ mit den Seitenlängen $a = 4$, $b = 3$, $c = 3$ und $d = 2$. Es ist der größtmögliche Flächeninhalt zu bestimmen, den dieses Viereck haben kann.

Eine Umkehrung der Aufgabenstellung

Es sei \mathfrak{M} eine vorgegebene Menge von ebenen Figuren, die einen Flächeninhalt und einen Umfang besitzen. Ferner enthalte \mathfrak{M} mit jeder Figur auch alle dazu ähnlichen. Ein Beispiel ist die Menge aller Rechtecke.

Bisher bestand die grundlegende Aufgabe darin, in der Menge $\mathfrak{M}(U)$ von Figuren aus \mathfrak{M} , die den gleichen Umfang U besitzen, diejenigen mit dem maximalen Flächeninhalt zu bestimmen. Es sei A_{\max} der maximale Flächeninhalt und $\mathfrak{M}_{\max}(U)$ die Menge von Figuren aus $\mathfrak{M}(U)$ mit diesem Flächeninhalt.

Umgekehrt kann nun bezüglich der gleichen Menge \mathfrak{M} die Menge $\mathfrak{M}(A)$ von Figuren aus \mathfrak{M} betrachtet werden, die den gleichen Flächeninhalt A besitzen, und hier nach denjenigen gefragt werden, die den kleinsten Umfang U_{\min} besitzen; ihre Menge sei entsprechend mit $\mathfrak{M}_{\min}(A)$ bezeichnet. Wählen wir als vorgegebenen konstanten Flächeninhalt gerade $A = A_{\max}$, so gilt folgende generelle Aussage:

(2.32) Es ist $\mathfrak{M}_{\min}(A) = \mathfrak{M}_{\max}(U)$ und $U_{\min} = U$.

Ein Beweis kann leicht indirekt geführt werden. Gäbe es unter allen Figuren aus \mathfrak{M} mit dem Flächeninhalt A_{\max} eine solche Figur F , die einen Umfang $u < U_{\min}$ hätte, dann gäbe es eine Zahl $k > 1$ mit $u \cdot k = U$ und eine zur Figur F ähnliche Figur F' mit dem Ähnlichkeitsfaktor k (etwa durch eine zentrische Streckung mit dem Faktor k erzeugt), die auch in \mathfrak{M} liegt und den Umfang U besitzt (also zu $\mathfrak{M}(U)$ gehört), deren Flächeninhalt aber $k^2 \cdot A_{\max} > A_{\max}$ ist. Dies widerspricht aber der Voraussetzung, dass A_{\max} der größte Flächeninhalt aller Figuren aus $\mathfrak{M}(U)$ ist.

Man bezeichnet zwei durch (2.32) in Beziehung gesetzte Aufgabenstellungen als zueinander dual.

Mit Hilfe von (2.32) folgen aus (2.27) bzw. (2.30) unmittelbar:

(2.33) Unter allen Rechtecken mit gleichem Flächeninhalt besitzen die Quadrate den kleinsten Umfang.

(2.34) Unter allen konvexen Figuren mit gleichem Flächeninhalt besitzt der Kreis den kleinsten Umfang.

(2.35) • Unter allen Dreiecken mit gleichem Flächeninhalt besitzen die gleichseitigen Dreiecke den kleinsten Umfang.

Wir sehen also, dass bei solchen Aufgaben die regelmäßigen Figuren favorisiert sind. Das hilft

uns mitunter, eine Vermutung zu finden, doch Vorsicht ist geboten!

Extremale Flächeninhalte

Die Bedingungen, unter denen ebene Figuren mit extremalem Flächeninhalt gesucht werden, können auch von ganz anderer Art sein:

(2.36) Es sei $CDEF$ ein Quadrat. Wir legen durch E alle Geraden g , die sowohl die Halbgerade CD^+ in einem Punkt A als auch die Halbgerade CF^+ in einem Punkt B schneiden (Abb. 2.14 a). Für welche dieser Geraden hat das diesbezügliche Dreieck ABC den kleinsten Flächeninhalt?

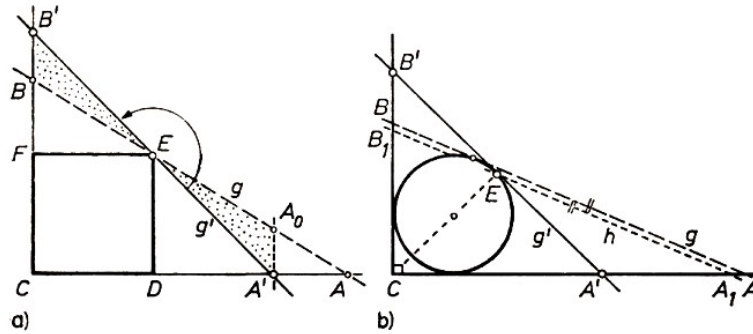


Abb. 2.14

Zunächst gibt es (genau) eine Gerade g' durch E , die ein gleichschenkliges Dreieck $A'B'C'$ (mit $|CA'| = |CB'|$) liefert.

Wir zeigen nun, dass jedes andere Dreieck ABC einen größeren Flächeninhalt besitzt.

Ohne Beschränkung der Allgemeinheit können wir dazu voraussetzen, dass A' zwischen D und A liegt. Dann liegt B zwischen F und B' (Abb. 2.14 a).

Nun schneidet die Parallele zur Geraden g_{CF} durch A' die Gerade $g = g_{EA}$ in einem Punkt A_0 zwischen E und A . Bei der Drehung um E mit 180° geht das Dreieck $EA'A_0$ in das Dreieck $EB'B$ über. Wegen

$$\begin{aligned} A(DAE) &= A(DA'E) + A(EA'A_0) + A(AA'A_0) \\ A(FBE) &= A(FB'E) - A(EB'B) \quad \text{und} \quad A(DA'E) = A(FB'E) \end{aligned}$$

ist damit

$$A(DAE) + A(FBE) > A(DA'E) + A(FB'E)$$

und schließlich $A(ABC) > A(A'B'C)$.

Die Vorgabe in der Aufgabenstellung (2.36) können wir auch so sehen, dass in einen rechten Winkel mit dem Scheitel C ein Quadrat derart eingeschoben werden ist, dass zwei seiner Seiten auf den Schenkeln des Winkels liegen.

Wir tauschen nun das Quadrat durch einen Kreis aus und formulieren eine (2.36) entsprechende Aufgabe:

(2.37) Gegeben seien ein rechter Winkel mit dem Scheitel C und ein Kreis k , der beide Schenkel des Winkels berührt. Für welche Geraden, die den Kreis berühren und die Schenkel des Winkels in Punkten A und B schneiden, hat das Dreieck ABC den kleinsten Flächeninhalt?

Zur Lösung kann vorteilhaft das Ergebnis von (2.36) verwendet werden. Zunächst gibt es (genau) eine Tangente g' an den Kreis, die beim Schnitt mit den Schenkeln ein gleichschenkliges Dreieck $A'B'C'$ (mit $|CA'| = |CB'|$) liefert (Abb. 2.14b).

Der Kreisberührungspunkt sei E . Jede andere Kreistangente g , die die Schenkel in A und B schneidet, geht nicht durch E . Folglich schneidet die Parallele h durch E zu dieser Tangente g den Kreis, und damit hat das durch sie gebildete Dreieck A_1B_1C einen kleineren Flächeninhalt als $A(ABC)$.

Nach (2.36) ist aber außerdem $A(A_1B_1C) > A(A'B'C)$. Also liefert die Tangente g' das Minimum.

(2.38) • Man löse Aufgabe (2.37) für den Fall, dass der Winkel mit dem Scheitel C' spitz bzw. stumpf ist.

Häufig auch von praktischem Interesse (Materialausnutzung !) sind folgende Aufgabenstellungen: Gegeben ist eine Figur, und in diese soll eine Figur vorgegebener Art mit möglichst großem Flächeninhalt gelegt werden. Wir beginnen mit einem einfachen Beispiel.

(2.39) Einem rechtwinkligen Dreieck ABC mit rechtem Winkel bei A ist ein Rechteck $PQRA$ mit maximalem Flächeninhalt einzubeschreiben, von dem zwei Seiten auf den Katheten des Dreiecks liegen.

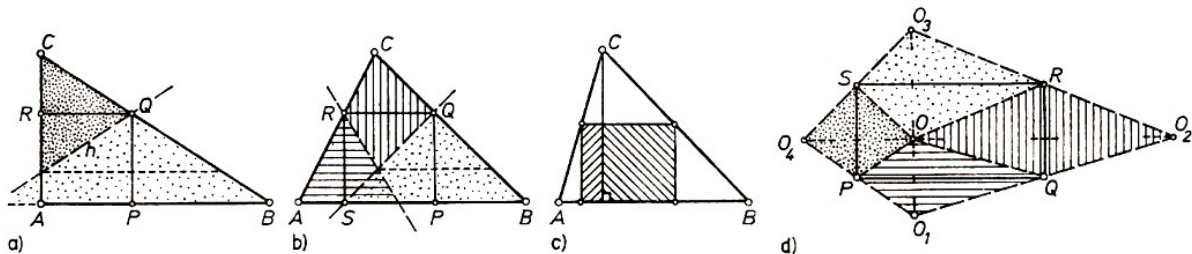


Abb. 2.15

Es liegt die Vermutung nahe, dass das Maximum genau dann vorliegt, wenn P , Q und R die Mittelpunkte von \overline{AB} , \overline{BC} bzw. \overline{CA} sind. Unter den verschiedenen Beweismöglichkeiten wählen wir die folgende:

Man spiegele für ein beliebiges Rechteck $PQRA$ mit $P \in \overline{AB}$, $Q \in \overline{BC}$ und $R \in \overline{CA}$ die Gerade g_{BC} an der Geraden g_{QP} (Abb. 2.15a). Wegen $g_{QR} \perp g_{QP}$ liefert die Spiegelung an g_{QR} das gleiche Bild h von g_{BC} . Die Gerade h zerlegt das Rechteck $APQR$, und die Spiegelung der Teile an g_{QP} bzw. g_{QR} zeigt unmittelbar, dass $A(APQR) \leq \frac{1}{2}A(ABC)$ gilt und Gleichheit dann und nur dann besteht, wenn h durch A geht, wenn also Q der Mittelpunkt von \overline{BC} ist.

Auf diese Weise, nämlich durch Spiegelungen an Rechteckseiten, erkennt man leicht die gegenüber (2.39) allgemeinere Aussage:

(2.40) Unter allen Rechtecken, die einem Dreieck ABC mit $\alpha, \beta \leq 90^\circ$ eingeschrieben sind und die mit einer Seite auf \overline{AB} liegen, besitzt dasjenige den größten Flächeninhalt, das die Mittelpunkte von \overline{BC} und \overline{CA} als Ecken besitzt (vgl. Abb. 2.15 b).

Dieser Satz ergibt sich auch mit Hilfe von (2.39).

Unter den Voraussetzungen $\alpha, \beta < 90^\circ$ zerlegt das Lot von C auf g_{AB} das Dreieck ABC in zwei rechtwinklige Dreiecke, und die dazu nach (2.39) gegebenen maximalen Rechtecke ergeben zusammen ein dann maximales Rechteck der vorgegebenen Lage (Abb. 2.15c).

Da dieses maximale Rechteck einen halb so großen Flächeninhalt wie das zugrunde liegende Dreieck besitzt, enthält jedes spitzwinklige Dreieck genau drei, jedes rechtwinklige Dreieck

genau zwei und jedes stumpfwinklige Dreieck genau ein maximales Rechteck, das mit einer Seite auf einer Dreiecksseite liegt.

Damit bleibt zunächst offen, ob bei einer anderen Lage des Rechtecks im Dreieck ABC der Inhalt $\frac{1}{2}A(ABC)$ oder vielleicht sogar noch ein größerer erzielt werden könnte. Es gilt jedoch:

(2.41) Einem Dreieck kann nur dann ein Rechteck mit maximalem Inhalt einbeschrieben werden, wenn eine Rechteckseite auf einer Seite des Dreiecks liegt.

Dies ist an Hand der bereits oben benutzten Spiegelungen an den Rechteckseiten einsichtig. Ist $PQRS$ ein Rechteck und O ein Punkt in seinem Inneren, dann ergeben die Spiegelungen von O an den Rechteckseiten ein Viereck $O_1O_2O_3O_4$ (Abb. 2.15 d), das das vorgegebene Viereck umschreibt und einen doppelt so großen Flächeninhalt besitzt.

Jedes Dreieck mit der Ecke O_1 , auf dessen Seiten P und Q liegen und das das Viereck $PQRS$ umschreibt, muss einen größeren Inhalt als $2A(PQRS)$ besitzen.

Liegt dagegen O auf dem Rand des Rechtecks $PQRS$, so ergeben die Spiegelungen ein doppelt so großes Dreieck, bei dem auf (wenigstens) einer Seite eine Seite des Vierecks liegt.

(2.42) • Einem konvexen Viereck $ABCD$ ist ein Parallelogramm mit maximalem Inhalt einzubeschreiben, dessen Seiten parallel zu den Diagonalen des Vierecks sind.

(2.43) • Man zeige: a) Die Dreiecke, die einem Kreis einbeschrieben sind und den größten Flächeninhalt besitzen, sind gleichseitig.

b) Die konvexen Vierecke, die einem Kreis einbeschrieben sind und den größten Flächeninhalt besitzen, sind Quadrate.

(2.44) • Man zeige: a) Unter allen Dreiecken, die Tangentendreieck eines vorgegebenen Kreises sind, sind die gleichseitigen diejenigen mit minimalem Inhalt.

b) Unter allen konvexen Vierecken, die Tangentenviereck eines vorgegebenen Kreises sind, sind die Quadrate diejenigen mit kleinstem Inhalt.

Zum Beweis von (2.44 a) kann (2.38) benutzt werden!

(2.45) • Es seien ABC ein Dreieck und P, Q, R innere Punkte der Seiten $\overline{AB}, \overline{BC}$ bzw. \overline{CA} . Man zeige, dass es unter allen Dreiecken PQR weder eines mit maximalem noch eines mit minimalem Inhalt gibt.

2.4 Extremale Umfänge

Wir setzen jetzt ein Thema fort, welches mit den Aufgaben (2.1), (2.3), (2.5), (2.7), (2.9) und (2.10) schon in Abschnitt 2.1 eine große Rolle spielte.

(2.46) Man bestimme unter allen Dreiecken, die einen vorgegebenen Kreis umschreiben, diejenigen mit dem kleinsten Umfang.

Eine Lösung erhält man leicht mit Hilfe von (2.44 a). Ist $U = a+b+c$ der Umfang des betrachteten Dreiecks und r der Radius des vorgegebenen Kreises, so ist $A = \frac{1}{2}r(a+b+c) = \frac{1}{2}rU$ der Inhalt des Dreiecks. Demnach ist U minimal genau dann, wenn A minimal ist. Und das ist nach (2.44 a) für gleichseitige Dreiecke der Fall.

(2.47) Man bestimme unter allen Dreiecken, deren Ecken auf einem vorgegebenen Kreis liegen, diejenigen mit dem größten Umfang.

Zur Lösung wählen wir vorerst zwei Punkte A und B auf dem Kreis k und bestimmen dazu

die Punkte C auf dem Kreis, für die das Dreieck ABC den größten Umfang besitzt. Wir beschränken uns dabei zunächst auf eine Seite bezüglich der Geraden g_{AB} . Trägt man $|BC|$ auf der Halbgeraden CA^- von C aus ab, so erhält man einen Punkt B' , für den $|\angle BB'A| = \frac{\gamma}{2}$ ist (Abb. 2.16).

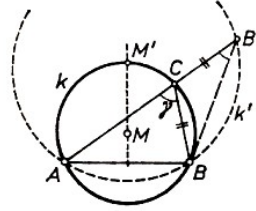


Abb. 2.16

Da für den Schnittpunkt M' der Mittelsenkrechten von \overline{AB} mit dem Kreis k die Beziehung $|\angle AM'B| = \gamma$ gilt, liegt B' auf dem Kreis k' um M' durch A (und B). Nun ist die Summe $|AC| + |CB| = |AB'|$ offenbar genau dann maximal, wenn B' auf dem Kreis k' dem Punkt A diametral gegenüber liegt, d. h., wenn $C = M'$ und damit $|AC| = |BC|$ ist.

Hinsichtlich des Maximums ist von beiden Seiten bezüglich g_{AB} offenbar nur eine solche von Interesse, für die C und M nicht auf verschiedenen Seiten liegen. Da für das Maximum bei einer festgehaltenen Seite nur ein gleichschenkliges Dreieck in Frage kommt, haben die gleichseitigen (und zueinander kongruenten!) Dreiecke den größten Umfang.

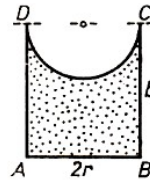


Abb. 2.17

(2.48)¹⁸ Abb. 2.17 zeigt ein Flächenstück, das aus der Fläche des Rechtecks $ABCD$ mit den Seitenlängen $|AB| = |CD| = 2r$ und $|BC| = |AD| = b$ mit $b > r$ durch Herausschneiden einer Halbkreisscheibe mit dem Durchmesser \overline{CD} entstanden ist.

Man denke sich nun eine positive reelle Zahl A beliebig gegeben. Dann sind alle geordneten Paare (r, b) positiver reeller Zahlen mit $r < b$ zu ermitteln, für die das entsprechende Flächenstück den Inhalt A und dabei möglichst kleinen Umfang hat.

Ist $b > r > 0$, so ergibt sich eine Figur wie in Abb. 2.17. Für ihren Umfang U und ihren Flächeninhalt A gilt

$$U = 2r + 2b + \pi r \quad \text{und} \quad A = 2br - \frac{1}{2}\pi r^2 \quad (1,2)$$

Daraus folgt $A = (U - 2r - \pi r)r - \frac{1}{2}\pi r^2$ und weiter

$$r^2 \left(\frac{3}{2}\pi + 2 \right) - Ur + A = 0 \quad (3)$$

Diese quadratische Gleichung in r besitzt nach Voraussetzung eine Lösung. Dann muss notwendig ihre Diskriminante nichtnegativ sein, d. h., es gilt

$$\frac{U^2}{(3\pi + 4)^2} - \frac{2A}{3\pi + 4} \geq 0$$

¹⁸Aufgabe 10 10 43 B der OJM

und folglich

$$U \geq \sqrt{2A(3\pi + 4)} \quad (4)$$

Das Gleichheitszeichen in (4) gilt wegen (3) genau dann, wenn $r = \frac{U}{3\pi+4}$ ist, und unter Beachtung von $U = \sqrt{2A(3\pi + 4)}$ folgt aus dieser Gleichung

$$r = \sqrt{\frac{2A}{3\pi + 4}} \quad (5)$$

und schließlich

$$b = \sqrt{\frac{2A}{3\pi + 4}}(\pi + 1) = r(\pi + 1) \quad (6)$$

Wählt man also zu vorgegebenem $A (> 0)$ die Zahlen r und b gemäß (5) und (6), so ergibt sich wegen $b > r$ eine Figur, wie sie in der Aufgabe beschrieben ist. Zwischen deren Inhalt A und Umfang U besteht die Relation $U = \sqrt{2A(3\pi + 4)}$. Daher genügt das Paar

$$(r, b) = \left(\sqrt{\frac{2A}{3\pi + 4}}, \sqrt{\frac{2A}{3\pi + 4}}(\pi + 1) \right)$$

den Bedingungen der Aufgabe, und dies ist das einzige dieser Art.

Oder ist es einfacher, die zu dieser Aufgabe duale zu lösen?

Sehr ähnlich wird die Lösung zur folgenden Aufgabe sein:

(2.49)¹⁹ • Dirk erklärt Jürgen den Nutzen der Differentialrechnung an Hand der Lösung der folgenden Aufgabe:

Es sei $ABCDE$ ein ebenes konvexes Fünfeck derart, dass A, B, C, E die Eckpunkte eines Rechtecks und C, D, E die Eckpunkte eines gleichseitigen Dreiecks bilden.

Als Flächeninhalt des Fünfecks $ABCDE$ werde nun ein geeigneter Wert A vorgeschrieben.

Man ermittle, ob unter allen diesen Fünfecken eines von kleinstem Umfang U existiert. Ist das der Fall, so berechne man für alle derartigen Fünfecke minimalen Umfangs den Wert $a : b$, wobei $|AB| = a$ und $|BC| = b$ bedeutet.

Am nächsten Tage teilt Jürgen Dirk mit, dass er eine Lösung dieser Aufgabe ohne Verwendung der Differentialrechnung gefunden habe.

Man gebe eine Lösung an, die Jürgen gefunden haben könnte.

Zum Abschluss noch eine Aufgabenstellung, in der der Umfang nur diskrete Werte annehmen kann.

(2.50) • Man bestimme den minimalen und maximalen Umfang aller sogenannten "Pentominos", d. h. aller einfachen n -Ecke, die sich aus fünf Einheitsquadraten zusammensetzen lassen.

¹⁹Aufgabe 11 10 43 B der OJM

3 Extrema von Funktionen

3.1 Funktionen einer reellen Veränderlichen und ihre Extremwerte

Kommt ein Einfall völlig zwanglos,
ist er meistens belanglos;
jedes gute Resultat
braucht den ganzen Apparat.
W. Dege

Bei den im Kapitel 2 betrachteten Aufgaben veränderte sich mit der Lage von Punkten oder der Form von Figuren die Maßzahl von Längen, Flächen oder Volumina. Könnten wir nun diese Abhängigkeit durch eine Funktion erfassen, so ließe sich die spezielle gesuchte Situation, nämlich die, für die die Maßzahl extremal wird, aus den Extremwerten der Funktion ablesen, falls man diese kennt!

Wir untersuchen das einmal an einem der einfachsten Beispiele, dem der Aufgabe (2.27).

Die Fläche A der Rechtecke lässt sich mittels der Seitenlängen x und z durch die Funktion zweier(!) Veränderlicher x und z mit $x > 0$ und $z > 0$ als $A(x, z) = x \cdot z$ angeben.

Die Nebenbedingung $u = 2x + 2z = \text{const}$ gibt uns aber die Möglichkeit, eine der beiden Variablen durch die andere zu ersetzen (und damit zu eliminieren):

$$z = \frac{u}{2} - x \quad \text{und} \quad A(x) = x \left(\frac{u}{2} - x \right)$$

Es ist also der größte Funktionswert der Funktion $A(x)$ für $x > 0$ gesucht! Wie findet man aber die Extremwerte einer Funktion?

Zunächst wollen wir exakt festlegen, worüber wir sprechen, wir definieren:

(3.1) Für alle Wertepaare (x, y) einer Menge f von reellen Zahlen x, y gelte: Wenn $(x, y_1) \in f$ und $(x, y_2) \in f$, dann folgt $y_1 = y_2$. Eine solche Menge f nennt man eine reellwertige Funktion einer reellen Veränderlichen.²⁰

Zur Beschreibung einer solchen Menge dient uns häufig eine Funktionsgleichung $y = f(x)$. Die Menge aller reellen Zahlen x , für die ein Funktionswert $f(x)$ erklärt ist, heißt Definitionsbereich $D(f)$ der Funktion f . Die Menge aller y , für die ein $x \in D(f)$ so existiert, dass $y = f(x)$ gilt, nennt man den Wertevorrat $W(f)$ der Funktion f .

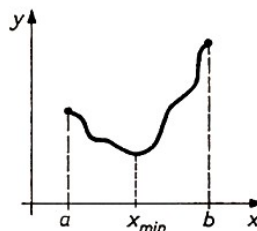


Abb. 3.1

(Mitunter sprechen wir dann auch kurz von der Funktion $y = f(x)$.)

Wir suchen unter allen $x \in D(f)$ diejenigen, für welche $f(x) \in W(f)$ maximal oder minimal wird. Solche Extremwerte nennen wir globale Extremwerte der Funktion f .

²⁰Zur ausführlichen Einführung des Funktionsbegriffes vgl. etwa [13].

(3.2) Existiert ein $x_0 \in D(f)$ so, dass $f(x_0) \geq f(x)$ (bzw. $f(x_0) \leq f(x)$) für alle $x \in D(f)$ gilt, so heißt $f(x_0)$ das globale Maximum (bzw. das globale Minimum) der Funktion f . (Im weiteren sprechen wir aber abkürzend vom Maximum oder Minimum.)

Bei vielen Anwendungen ist wichtig, dass $D(f)$ durch die in der Aufgabenstellung enthaltenen Bedingungen bestimmt wird.

Betrachten wir die in der Abb. 3.1 veranschaulichte Funktion f mit $D(f) = [a, b]$, so entnehmen wir dieser graphischen Darstellung, dass f für $x_0 = b$ das globale Maximum und für $x_0 = x_{\min}$ das globale Minimum annimmt.

In den Teilintervallen $a \leq x \leq x_{\min}$ und $x_{\min} \leq x \leq b$ verläuft die dargestellte Funktion monoton. Es ist zu vermuten, dass das Monotonieverhalten in einem engen Zusammenhang mit den Extremwerten steht.

(3.3) Sind $a, b \in D(f)$ und gilt für alle $x_1, x_2 \in D(f)$ mit $a \leq x_1 < x_2 \leq b$ die Ungleichung $f(x_1) < f(x_2)$ (bzw. $f(x_1) > f(x_2)$), so nennt man die Funktion f auf dem Intervall $a \leq x \leq b$ streng monoton wachsend (bzw. streng monoton fallend).

(3.4) • Man beweise: Jede auf einem abgeschlossenen Intervall $a \leq x \leq b$ definierte und dort streng monotone Funktion nimmt ihre Extremwerte in den Randpunkten des Intervalls an. Ist dagegen eine streng monotone Funktion auf einem offenen Intervall $a < x < b$ definiert, so hat sie keine Extremwerte.

(3.5) • Man ermittle den größtmöglichen Definitionsbereich der Funktion f mit $f(x) = \sqrt{\frac{1}{x} - 1}$ und zeige, dass f zwar ein Minimum, aber kein Maximum besitzt.

Die linearen Funktionen f mit $f(x) = mx + n$ haben also keine Extrema oder nehmen diese, falls ihr Definitionsbereich auf ein abgeschlossenes Intervall $a \leq x \leq b$ eingeschränkt wird, in den Punkten a bzw. b an.

3.2 Funktionen zweiten und dritten Grades und ihre Extrema

Unser einführendes Beispiel (und viele andere auch) führt auf eine quadratische Funktion.

(3.6) Satz. Eine für alle reellen Zahlen x definierte quadratische Funktion f (also $D(f) = \mathbb{R}$) mit $f(x) = ax^2 + px + q$ ($a \neq 0$), nimmt genau an der Stelle $x_0 = -\frac{p}{2a}$ einen Extremwert an, und zwar für $a > 0$ ein Minimum und für $a < 0$ ein Maximum.

Beweis. Es gilt

$$ax^2 + px + q = a \left(x^2 + \frac{p}{a}x + \frac{q}{a} \right) = a \left[\left(x + \frac{p}{a} \right)^2 - \frac{p^2}{4a^2} + \frac{q}{a} \right]$$

Da $-\frac{p^2}{4a^2} + \frac{q}{a}$ eine Konstante ist, wird der Ausdruck

$$\left(x + \frac{p}{a} \right)^2 - \frac{p^2}{4a^2} + \frac{q}{a}$$

genau dann minimal, wenn $\left(x + \frac{p}{a} \right)^2$ minimal wird, d. h., wenn $x_0 = -\frac{p}{2a}$ gilt. Ist $a > 0$, so ist demnach auch $f(x)$ an dieser Stelle minimal, für $a < 0$ dagegen maximal.

Wir möchten nachdrücklich darauf aufmerksam machen, dass der Satz (3.6) für eine auf einem

endlichen Intervall definierte quadratische Funktion nicht gelten muss.

(3.6') Wird der Definitionsbereich der quadratischen Funktion aus (3.6) auf ein abgeschlossenes Intervall $c \leq x \leq d$ eingeschränkt und gilt $c \leq x_0 \leq d$ mit $x_0 = -\frac{p}{a}$, so existiert in x_0 ein Minimum (bzw. Maximum), aber außerdem ein Maximum (bzw. Minimum) in einem der Randpunkte des Intervalls.

Wir kommen auf unser einführendes Beispiel (2.27) zurück.

Die Funktion $A(x) = x\left(\frac{u}{2} - x\right)$ wird durch Ausmultiplizieren in die "Normalform" $A(x) = -x^2 - \frac{u}{2}x$ übergeführt. Aus dem Text zu (2.27) ergibt sich der eingeschränkte Definitionsbereich $0 \leq x \leq u$. Aus dem Satz (3.6) und der Anmerkung (3.6') folgt (wie wir bereits anderweitig ermittelt hatten), dass für $x_0 = \frac{u}{4}$ der Flächeninhalt maximal wird.

(3.7) Man beschreibe einem Quadrat ein Quadrat minimalen Umfangs ein.

Diese bereits unter (2.9) formulierte Aufgabe wollen wir jetzt mit den bereitgestellten Hilfsmitteln lösen.

In Abb. 3.2 ist die Seitenlänge des vorgegebenen Quadrates mit q bezeichnet.

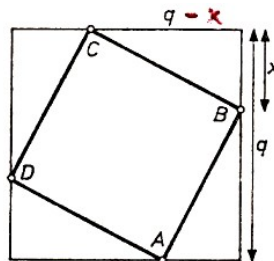


Abb. 3.2

Soll das einbeschriebene Viereck $ABCD$ auch ein Quadrat sein, so muss jede Seite durch die Punkte A , B , C bzw. D in Abschnitte der Länge x und $q - x$ geteilt werden.

Aus dem Satz des Pythagoras ergibt sich für den Umfang des einbeschriebenen Quadrates

$$U(x) = 4\sqrt{(q-x)^2 + x^2} \quad \text{mit} \quad 0 \leq x \leq q$$

Da $U(x) \geq 0$ gilt, wird $U(x)$ genau dann minimal, wenn $y(x) := \frac{U^2(x)}{16}$ minimal wird. Es gilt $y(x) = 2x^2 - 2qx + q^2$.

Die Anwendung unserer theoretischen Einsichten liefert uns ein Minimum an der Stelle $x_0 = \frac{q}{2}$, d. h., die Punkte A , B , C bzw. D müssen die Seitenmittelpunkte des vorgegebenen Quadrats sein.

Auch bei diesem zweiten Beispiel war es nicht schwierig, die "zugehörige" quadratische Funktion zu finden.

(3.8) • Man löse mit Hilfe von Satz (3.6) a) die Aufgabe (2.39), b) das zu (2.48) duale Problem.

Weitere einfache Beispiele, die sich auf die Anwendung von Satz (3.6) zurückführen lassen, findet man in [9], S. 11-18.

Die folgenden Beispiele sind u. a. auch darum wichtig, weil nicht gleich ersichtlich ist, dass unser einfaches mathematisches Modell zutreffend ist; erst eine geeignete Formalisierung wird das jeweils zeigen.

(3.9) Ein Betrieb bietet auf dem Weltmarkt ein Produkt zum Stückpreis A an. Er stellt fest,

dass er bei einem bestimmten Preisnachlass von ε (pro Stück) 100 Stück von dem Produkt mehr absetzt. Welche höchste Gesamteinnahme kann der Betrieb erzielen, wenn man annimmt, dass die Anzahl der mehr verkauften Stücke proportional zum Preisnachlass wächst?

Zur Lösung der Aufgabe mittels einer Funktion müssen wir zunächst eine Variable einführen, die die Veränderung geeignet erfasst. Der Betrieb will ja feststellen, wie oft er den Preisnachlass vervielfachen muss, um eine Höchsteinnahme zu erzielen. Wir führen daher $x \geq 0$ als Vielfaches des Preisnachlasses ein.

Vor der Einführung des Preisnachlasses wurden s Stück mit der Einnahme $s \cdot A$ verkauft. Nach der Einführung des (vervielfachten) Preisnachlasses $x \cdot \varepsilon$ werden (laut Annahme der Aufgabenstellung) $s + 100 \cdot X$ Stück verkauft, jedes Stück erzielt aber nur den Preis $A - \varepsilon x$. Die Gesamteinnahme $g(x)$ beträgt also $(s + 100 - x)(A - \varepsilon x)$. Das Maximum von $g(x)$ ergibt sich nach (3.6) - man rechne das bitte nach - für

$$x_0 = \frac{100A - \varepsilon s}{200\varepsilon}$$

es ergibt sich

$$g(x_0) = \frac{(100A + \varepsilon s)^2}{400\varepsilon}$$

Akzeptieren wir diese Werte als Ergebnis der gestellten Aufgabe, so haben wir "stillschweigend" angenommen, dass auch $x \geq 0$ gilt. Das ist aber nur für $\varepsilon \leq \frac{100A}{s}$ der Fall.

Da wir ja für die Konstanten A , s und ε keine Zahlen vorgegeben haben, können wir nicht feststellen, ob tatsächlich $x_0 \geq 0$ gilt.

Ergibt sich also in einem konkreten Fall $x_0 < 0$ nach der obigen Formel, so bedeutet das, dass der Betrieb zwar bei einem Preisnachlass einen Mehrverkauf registrieren kann, aber auch weniger Einnahmen erzielt. Die Gesamteinnahme ist in diesem Fall durch einen "Mehrverkauf durch Preisnachlass" nicht zu erhöhen.

Das Maximum der quadratischen Funktion mit $x \geq 0$ liegt in dem Randpunkt $x_0 = 0$. (Man zeichne sich einmal die Funktion $g(x)$ für zwei solcher Art prinzipiell verschiedene Fälle !).

Ebenfalls mit dem einfachen Mittel der quadratischen Funktion lässt sich ein Spezialfall der "Methode der Kleinsten Quadrate" behandeln. Diese Methode wird oft in der Ausgleichs- und Fehlerrechnung gebraucht. Wir wollen das an einem Beispiel erläutern:

Unter Verwendung des Ohmschen Gesetzes $U = R \cdot I$ soll durch Messung von U und I der Wert eines unbekanntes Widerstandes R ermittelt werden. Da solche Messungen von U und I stets ungenau sind, wird man mehrere Messungen durchführen, die dann allerdings bei der Rechnung $\frac{U}{I}$ verschiedene Werte für R ergeben. Wie wäre hier sinnvoll ein Ausgleich zu schaffen?

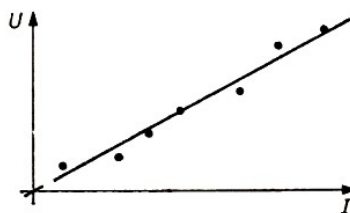


Abb. 3.3

Trägt man die verschiedenen Messergebnisse (I_i, U_i) ($i = 1, \dots, n$) in ein Koordinatensystem ein, so würden sie im Idealfall nach dem Ohmschen Gesetz auf einer Geraden liegen, im Realfall wird das aber nicht der Fall sein (vgl. Abb. 3.3).

Eine sinnvolle Forderung ist beispielsweise, die Gerade so zu legen, dass die Gesamtabweichung aller Punkte von der Geraden minimal wird. Unter der Gesamtabweichung wird man die Summe aller Abweichungen verstehen; aber wie wollen wir die Abweichung definieren?

Man könnte die Ordinatendifferenz wählen, die aber wäre vorzeichenbehaftet, und Abweichungen "nach oben" und "nach unten" würden sich gegenseitig aufheben und in der Gesamtabweichung fälschlicherweise eine große Genauigkeit vortäuschen.

Wählt man statt dessen den Betrag der Ordinatendifferenzen, könnte man diesen Mangel beseitigen, Beträge erfordern aber häufig viele Fallunterscheidungen beim weiteren Arbeiten. Numerisch als sehr bequem erweist es sich, für die Abweichung das Quadrat der Ordinatendifferenz einzuführen.

In unserem Beispiel heißt das, R ist so zu wählen, dass

$$f(R) = \sum_{i=1}^n (R \cdot I_i - U_i)^2$$

den minimalen Wert annimmt (n Anzahl der Messpunkte)!

Einfache Umformungen ergeben

$$f(R) = \left(\sum_{i=1}^n I_i^2 \right) R^2 - 2 \left(\sum_{i=1}^n I_i U_i \right) R + \sum_{i=1}^n U_i^2$$

Das ist aber eine quadratische Funktion mit

$$a = \sum_{i=1}^n I_i^2 \quad , \quad p = -2 \sum_{i=1}^n I_i U_i$$

deren Minimum nach (3.6) bei

$$R_0 = -\frac{p}{2a} = \frac{\sum_{i=1}^n I_i U_i}{\sum_{i=1}^n I_i^2}$$

liegt.

Statt $U = R \cdot I$ wäre auch die Schreibweise $y = f(x) = mx$ mit dem entsprechend formulierten Resultat möglich. Das Spezielle des von uns betrachteten Falles besteht darin, dass der Ausgleich durch eine Gerade erfolgt und überdies schon ein Punkt dieser Geraden (nämlich der Ursprung des Koordinatensystems) festliegt.

Erstmalig wurde die "Methode der kleinsten Quadrate" von A. M. Legendre²¹ bei der Berechnung von Kometenbahnen benutzt, C. F. Gauß²² begründete sie und begann mit vertiefenden mathematischen Betrachtungen, heute gibt es allein zu dieser Methode umfangreiche Bücher.²³

Die folgende Aufgabe führt uns auf ganzrationale Funktionen höheren Grades.

(3.10) a) Gegeben sei durch $y = f(x) = kx^2$ mit $k > 0$ und $-t \leq x \leq t$ ein "Parabelsegment" (vgl. Abb. 3.4). Man beschreibe diesem Parabelsegment entsprechend der Abb. 3.4 ein

²¹Adrien-Marie Legendre (1752-1833), französischer Mathematiker.

²²Carl Friedrich Gauß (1777-1855), Professor in Göttingen, bedeutendster Mathematiker der Neuzeit.

²³Beispielsweise J. W. Linnik; Methode der kleinsten Quadrate in moderner Darstellung, Deutscher Verlag der Wissenschaften, Berlin 1961.

Rechteck so ein, dass es unter allen einbeschriebenen Rechtecken den maximalen Flächeninhalt hat.

b) Bei Rotation des Parabelsegmentes (aus Teil a) um die y -Achse entsteht ein Rotationsparaboloid. Man beschreibe diesem analog einen Zylinder maximalen Volumens ein (siehe die Umschlagzeichnung).

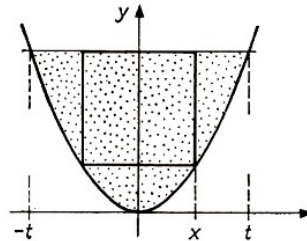


Abb. 3.4

Welches Problem wird wohl die größeren Schwierigkeiten bereiten, das ebene oder das räumliche?

Wir betrachten zunächst beide Probleme parallel. Eine Seite des Rechtecks bzw. der Durchmesser des Zylinders habe die Länge $2x$ mit $0 \leq x \leq t$. Die Höhe des Parabelsegmentes beträgt kt^2 , die Höhe des Zylinders (bzw. zweite Seite des Rechtecks) hat dann die Länge $kt^2 - kx^2 = k(t^2 - x^2)$. diesem analog einen Zylinder maximalen Volumens ein (siehe die Umschlagzeichnung).

Zur Lösung der Aufgabe a) ist also das Maximum von $f_e(x) = 2x - k(t^2 - x^2)$ und für die der Aufgabe b) das von $f_r(x) = \pi x^2 \cdot k(t^2 - x^2)$ zu bestimmen.

Die Existenz eines Maximums ist in beiden Fällen leicht einsichtig, da die Funktionswerte in den Endpunkten 0 bzw. t des Definitionsbereiches verschwinden und die Funktionswerte sich stetig ändern. Welches der Maxima ist aber einfacher zu ermitteln?

$f_e(x)$ ist eine Funktion dritten Grades und $f_r(x)$ sogar eine vierten Grades (man multipliziere aus!). Aber mit der Substitution $z =: x^2$ erhalten wir für $f_r(x)$ eine Funktion $f_r(z)$, die quadratisch in z ist und bei Anwendung von (3.6) $z_0 = \frac{t^2}{2}$ ergibt.

Wegen $x_0 > 0$ und $t > 0$ folgt daraus, dass $f_r(x)$ das Maximum genau dann annimmt, wenn der Zylinder einen Grundkreisradius von $x_0 = \frac{\sqrt{2}}{2}t$ - womit auch $0 \leq x \leq t$ erfüllt ist - und eine Höhe $\frac{kt^2}{2}$ hat.

Um den schwierigeren Teil a.) der Aufgabe zu lösen, beschäftigen wir uns zunächst allgemein mit der Extremwertermittlung von Funktionen f der Gestalt $f(x) = ax^3 + bx^2 + cx + d$ und $D(f) = \mathbb{R}$.

Wir betrachten zunächst die Nullstellen dieser Funktion.

Unter den Nullstellen der Funktion f verstehen wir diejenigen reellen Zahlen x_0 , für die $f(x_0) = 0$ gilt. Aus dem Fundamentalsatz der Algebra folgt, dass f höchstens drei Nullstellen haben kann. Wir wollen hier als plausibel akzeptieren, dass immer eine Darstellung

$$ax^3 + bx^2 + cx + d = a(x - r)(x^2 + px + q) \quad (*)$$

existiert.

Aus (*) folgt, dass f stets eine reelle Nullstelle ($x_0 = r$) hat und weitere, falls reelle Lösungen von $x^2 + px + q = 0$ existieren. Der Spezialfall $x^2 + px + q = (x - s)^2$ mit der "doppelten" Nullstelle $x_0 = s$ wird für uns eine wichtige Rolle spielen.

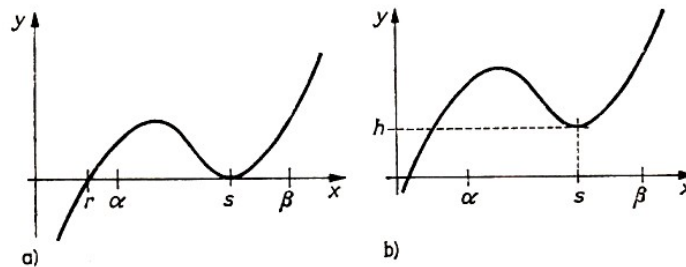


Abb. 3.5

Wir betrachten den Graphen der Funktion f mit

$$f(x) = a(x - r)(x - s)^2, \quad a > 0, \quad r < s$$

Der in der Abb. 3.5a skizzierte Graph ergibt sich aus folgenden Überlegungen:

- (1) $f(r) = 0$ und $f(s) = 0$,
- (2) für alle $x > r$ gilt $f(x) > 0$, und für alle $x < r$ gilt $f(x) < 0$,
- (3) verändert sich x "wenig", dann auch $f(x)$, d. h., die Funktion ist für alle x stetig.²⁴

Dann hat aber nach der Definition (3.2) die auf das Intervall $[\alpha, \beta]$ eingeschränkte Funktion f an der Stelle s ein Minimum, wenn $s > \alpha > r$ und $\beta > s$ gilt.

Hat nun umgekehrt die Funktion f an der Stelle s im Inneren des Intervalls $[\alpha, \beta]$ ein Minimum, so lässt sich durch eine Verschiebung längs der Ordinatenachse um $-h$ (vgl. Abb. 3.5 b) unser gerade diskutierter Fall erzeugen. Analoges gilt für ein Maximum. Wenn aus diesen Überlegungen eine Methode zur Ermittlung der Extrema abgeleitet werden soll, so muss insbesondere gezeigt werden, wie s und h zu ermitteln sind. Es soll

$$\begin{aligned} ax^3 + bx^2 + cx + d - h &= a(x - r)(x - s)^2 \\ ax^3 + bx^2 + cx + d - h &= ax^3 - a(r + 2s)x^2 + as(s + 2r)x - ars^2 \end{aligned} \quad (**)$$

gelten. Da (**) für alle $x \in D(f)$ gelten soll, müssen die Koeffizienten entsprechender Potenzen von x übereinstimmen (Methode des Koeffizientenvergleichs):

$$b = -a(r + 2s), \quad c = as(s + 2r), \quad d - h = -ars^2$$

wobei a, b, c und d vorgegeben sind. Aus den ersten beiden Gleichungen sind r und s zu ermitteln, aus der letzten dann h . Wir wollen das an unserem Beispiel (3.8 a) zeigen.

Wir suchen laut Aufgabenstellung das Maximum der Funktion

$$f_e(x) = 2kx(t^2 - x^2) = -2kx^3 + 0x^2 + 2kt^2x + 0$$

im Intervall $0 \leq x \leq t$. Um eine bessere Anpassung an die obige Darstellung zu erreichen, ermitteln wir das Maximum der Funktion $-f_e(x) = +2kx^3 - 2kt^2x$. Der Koeffizientenvergleich ergibt

$$0 = -2k(r + 2s), \quad -2kt^2 = +2ks(s + 2r), \quad h = 2krs^2$$

Da $k \neq 0$ gilt, folgt aus der ersten Gleichung $r = -2s$ und damit aus der zweiten $2kt^2 = 6ks^2$, d. h. $s = \frac{1}{\sqrt{3}}$ (mit $0 \leq s \leq t$).

Damit ist die Extremalstelle s der Funktion $-f_e(x)$ gefunden!

²⁴Vgl. [13], S. 53-56.

Mit s berechnet man r und damit schließlich h . Das Maximum von $f_e(x)$ ist dann $-h = \frac{4kt^3}{3\sqrt{3}}$. Diesen Wert ermitteln wir nochmals zur Kontrolle aus

$$f_e(s) = 2k \frac{1}{\sqrt{3}} \left(t^2 - \frac{t^2}{3} \right) = \frac{4kt^3}{3\sqrt{3}}$$

Zur Festigung dieser Methode wende man sie selbständig auf die nächste Aufgabe an.

(3.11) • Die Tragkraft eines Balkens von der Breite x und der Höhe y ist proportional dem Produkt xy^2 . Mit welchen Abmessungen muss man demnach einen Balken maximaler Tragkraft aus einem Baumstamm ausschneiden, dessen Durchmesser d als konstant angenommen wird?

Auch die in [9] angegebenen (und dort anders gelösten) Aufgaben 14 bis 18 kann man nach der oben beschriebenen Methode lösen.

3.3 Die Funktion $f(x) = x + \frac{d}{x}$ und ihre Extreme

Schlag die Tür nicht ein -
sie lässt sich leicht mit dem Schlüssel öffnen.
Sprichwort aus Dagestan

Bei der Fülle von Extremalproblemen ergeben sich natürlich auch sehr viele verschiedene Funktionen, mit denen man die Probleme beschreiben kann. Wir wollen aus dieser Fülle noch eine Funktion herausgreifen und an verschiedenartigen Beispielen zeigen, welche Bedeutung ihr zukommt.

(3.12) Für alle von 0 verschiedenen reellen Zahlen x ist durch $f(x) = x + \frac{d}{x}$ mit positivem d eine Funktion f definiert.

Wir betrachten zunächst einige Eigenschaften dieser Funktion:

Ersetzt man x durch $-x$, so gilt $f(-x) = -f(x)$, d. h., die Funktion ist ungerade (vgl. [13]), es genügt daher im folgenden die Betrachtung von f für $x > 0$. Die graphische Darstellung von f lässt sich aus denen der Funktionen $g(x) = x$ und $h(x) = \frac{d}{x}$ - zumindest angenähert - durch Addition der Funktionswerte leicht ermitteln (vgl. Abb. 3.6).

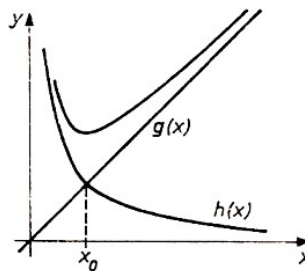


Abb. 3.6

Aus dieser graphischen Darstellung ist die Vermutung ersichtlich, dass die Funktion f für $x > 0$ genau an der Stelle x_0 das Minimum besitzt, an der sich die Graphen von $h(x)$ und $g(x)$ schneiden.

Wir werden diese Vermutung später beweisen. Zunächst wollen wir aber an Aufgaben zeigen, dass verschiedenartigste Problemstellungen durch diese Funktion erfasst werden. Wir lösen nochmals - jetzt eben mit anderen Mitteln - die uns schon bekannten Aufgaben (2.36) und (2.37):

Auf die Problemstellung (2.36) kann man gut den Strahlensatz anwenden. Setzt man für die Länge der Quadratseite o. B. d. A. $|CD| = 1$ und für die sich in Abhängigkeit von der Lage der Geraden g verändernde Streckenlänge $|DA| = x$, so gilt $x : (1+x) = 1 : |BC|$ bzw. $|BC| = \frac{1+x}{x}$. (Man überprüfe diese Angaben im Vergleich mit Abb. 2.14a).

Den Flächeninhalt des Dreiecks ABC können wir jetzt als Funktion von x angeben:

$$A(x) = \frac{1}{2}|AC| \cdot |BC| = \frac{1}{2} \frac{(1+x)^2}{x} = \frac{1}{2} \left(x + \frac{1}{x} \right) + 1$$

$A(x)$ ist genau dann minimal, wenn der Wert der Funktion $f(x) = x + \frac{1}{x}$ minimal ist. Nach unserer Vermutung müsste das für $x_0 = 1$ (in Übereinstimmung mit dem früheren Ergebnis) der Fall sein.

Im Beispiel (2.37) setzen wir o. B. d. A. $r = 1$ und entsprechend zu der Lösung von (2.37) $|AD| = x$. (Vgl. Abb. 2.14b.) Um $|BF|$ zu ermitteln, benutzen wir den Satz des Pythagoras. Dabei ist zu beachten, dass $|BE| = |BF|$ und $|AB| = |AD| = x$ gilt:

$$(1+x)^2 + (1+|BF|)^2 = (x+|BF|)^2 \quad , \quad |BF| = \frac{x+1}{x-1}$$

Die Fläche des Dreiecks ABC denken wir uns zusammengesetzt aus den Flächen des Quadrates $CEMF$ und der Dreiecke AEM , AMG , MGB und BMF . Damit erhalten wir schließlich

$$A(x) = 1 + x + \frac{x+1}{x-1}$$

Um eine Funktion vom Typ (3.12) zu erkennen, führen wir noch eine Variablensubstitution aus: $z = x - 1$. (Aus Abb. 2.14 b ist ersichtlich, dass auch z positiv ist.) Damit ergibt sich

$$A(z) = 2 + z + \frac{z+2}{z} = z + \frac{2}{z} + 3$$

$A(z)$ wird genau dann minimal, wenn $z + \frac{2}{z}$ minimal wird, wir haben also wieder den Anschluss an (3.12) hergestellt.

(3.13) • Man löse (2.48) unter Verwendung der Funktion (3.12).

(3.14) Elektrische Leistung: Die elektromotorische Kraft einer Stromquelle sei e , ihr innerer Widerstand R_i . Wie groß muss der äußere Widerstand x gewählt werden, damit die von der Stromquelle abgegebene Leistung N maximal wird?

Es gilt $e = I(R_i + x)$ und $N = U \cdot I$, wenn I die Stromstärke und U die Spannung im äußeren Stromkreis sind.²⁵

Daraus erhält man unter Einbeziehung des Ohmschen Gesetzes

$$N = N(x) = \frac{e^2 x}{(R_i + x)^2}$$

Da e^2 konstant ist, nimmt $N(x)$ mit $x > 0$ genau an den Stellen einen maximalen Wert an, an denen $\frac{x}{(R_i+x)^2}$ maximal bzw der Kehrwert $\frac{(R_i+x)^2}{x}$ minimal wird!

Der Kehrwert

$$\frac{(R_i + x)^2}{x} = \frac{R_i^2}{x} + 2R_i + x$$

²⁵Vgl. W. H. Westphal, Physik, Berlin 1939, S. 303.

wird aber genau dann minimal, wenn $f(x) = x + \frac{R_i^2}{x}$ minimal wird, da $2R_i$ eine konstante Größe ist. Damit haben wir die Lösung der Aufgabe auf die Betrachtung der Funktion (3.12) mit $d := R_i^2$ zurückgeführt.

(3.15) Lagerhaltungskosten: Eine Verkaufseinrichtung hat nach ihren Erfahrungen für einen bestimmten Artikel einen Bedarf von 160 Stück pro Monat. Der Artikel muss angeliefert werden.

Die Transportkosten betragen pro Lieferung (unabhängig von der gelieferten Anzahl!) 30 M. Für die angelieferten Artikel entstehen aber in der Verkaufseinrichtung Lagerkosten von 1,50 M pro Stück und Monat.

Welche Anzahl x des Artikels muss in welchen Abständen geliefert werden, damit die Gesamtkosten möglichst niedrig sind? Es wird vorausgesetzt, dass der Bestand des Artikels linear abnimmt.

Besteht eine Lieferung aus x Stück, so sind $z = \frac{160}{x}$ Lieferungen pro Monat nötig, um den Bedarf zu decken, die Transportkosten pro Stück betragen $\frac{30}{x}$ M. Will man die Lagerkosten pro Stück berechnen, so steht man vor der Schwierigkeit, dass man ja die Lagerzeit eines einzelnen Stückes nicht kennt.

Da aber vorausgesetzt ist, dass der Bestand des Artikels linear abnimmt, kann man für jedes Stück die gleiche durchschnittliche Lagerzeit annehmen, diese ist halb so groß wie die Zeit zwischen zwei Lieferungen, also $\frac{1}{2z}$ Monate. Da pro Monat Lagerkosten von 1,50 M pro Stück entstehen, sind bei z Lieferungen im Durchschnitt $\frac{0,75}{z}$ M pro Stück zu bezahlen.

Für ein Stück entstehen folglich in Abhängigkeit von x Gesamtkosten in Höhe von (in M):

$$k(x) = \frac{30}{x} + \frac{0,75}{160}x$$

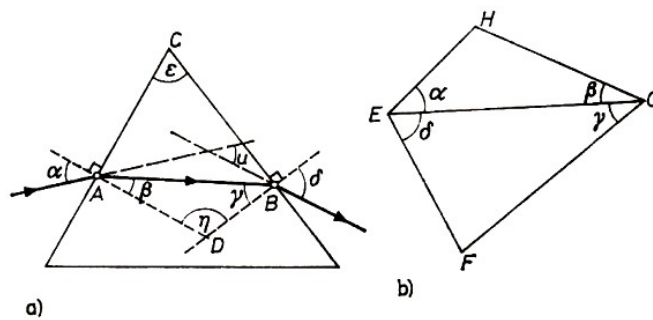


Abb. 3.7 a)

Das Minimum dieser Funktion $k(x)$ ist an der gleichen Stelle gelegen wie das der Funktion

$$f(x) = \frac{160 \cdot 4}{3}k(x) = \frac{6400}{x} + x$$

Bei einem Vergleich von f mit f aus (3.12) ergibt sich $d = 6400$, es ist aber auch zu beachten, dass f nur für positive natürliche Zahlen x erklärt ist.

(3.16)* Lichtablenkung durch ein Prisma: Geht ein Lichtstrahl durch ein Prisma (mit dem Brechungsindex n), so wird er an den Trennflächen nach dem Snelliusschen Brechungsgesetz abgelenkt (vgl. Abb. 3.7a). Bei welchem Strahldurchgang ist die Ablenkung u am kleinsten?

Das Snelliussche Brechungsgesetz (welches wir später noch aus dem Extremalprinzip des Lichtes herleiten werden) lautet mit den Winkelbezeichnungen der Abb. 3.7a

$$\sin \alpha = n \sin \beta \quad \text{bzw.} \quad \sin \delta = n \sin \gamma \quad (1)$$

Da die Einfallslotte g_{AD} und g_{BD} nach Definition senkrecht auf den Trennflächen stehen, gilt $\eta = 180^\circ - \varepsilon$, und der Winkel u , der die Ablenkung beschreibt, ergibt sich zu $u = \delta + \alpha - \varepsilon$. Da ε ein konstanter Winkel ist, wird u genau dann minimal, wenn $\alpha + \delta$ minimal wird. Dieses Extremalproblem versuchen wir mittels einer trigonometrischen Funktion zu lösen.

Die Winkel α und δ sind wegen ihrer physikalischen Bedeutung spitz, also gilt $0 < \alpha + \delta < 180^\circ$. In diesem Intervall ist die cos-Funktion monoton fallend. Folglich ist $\alpha + \delta$ genau dann minimal, wenn $\cos(\alpha + \delta)$ maximal ist.

Wir verwenden den Kosinussatz, um $\cos(\alpha + \delta)$ zu berechnen. Dafür benötigen wir eine Hilfsabbildung, in der die Winkel α und δ und auch β und γ geeignet eingetragen sind (vgl. Abb. 3.7 b).

Wählen wir $|EH| = 1$ und $|EF| = x$, so ergibt sich aus (1) mit dem Sinussatz $|HG| = n$ und $|FG| = x \cdot n$.

Unter zweimaliger Verwendung des Kosinussatzes erhalten wir schließlich

$$\begin{aligned}\cos(\alpha + \delta) &= \frac{1 + x^2 - |HF|^2}{2x} \\ \cos(\alpha + \delta) &= \frac{1}{2x} [1 + x^2 - n^2 - n^2 x^2 + 2n^2 x \cos(\gamma + \beta)] \\ \cos(\alpha + \delta) &= \frac{1}{2x} (1 - n^2)(1 + x^2) + n^2 \cos \varepsilon\end{aligned}\quad (2)$$

Da $n^2 \cos \varepsilon$ konstant ist und $n > 1$ gilt, folgt aus (2), dass $\cos(\alpha + \delta)$ genau dann maximal wird, wenn $f(x) = \frac{1}{x} + x$ minimal wird.

Nach diesen - z. T. recht anspruchsvollen - Beispielen fällt uns vielleicht ein weiteres ein, welches wir schon betrachtet haben und das uns nun auch den Beweis unserer Vermutung liefern wird, wir meinen (2.27).

Setzen wir in $a + b = U$ für b den Wert $\frac{A}{a}$ ein, so erhalten wir gerade, dass $a + \frac{A}{a}$ zu minimieren ist. (2.27) liefert aber das Minimum $a_0 = b_0 = \sqrt{A}$ für $A > 0$. Diese geometrische Interpretation ist dabei immer möglich.

Ein zweiter Beweis arithmetischer Art ist auch sehr einfach. Wir ergänzen mittels einer Konstanten

$$f(x) - 2\sqrt{d} = x + \frac{d}{x} - 2\sqrt{d} = \left(\sqrt{x} - \sqrt{\frac{d}{x}} \right)^2 \geq 0$$

Der kleinste Funktionswert ist 0, und er wird genau für $x_0 = \sqrt{d}$ angenommen. Außerdem ergibt sich sofort $f(x_0) = 2\sqrt{d}$.

(3.17) • Man ermittle für die drei vorangegangenen Beispiele die Lösungen.

3.4 Angenäherte Berechnung von Extremwerten

Das Unzulängliche ist produktiv.
J. W. v. Goethe

Einerseits haben wir in den Abschnitten 3.2 und 3.3 gezeigt, welche Vielfalt von Extremalproblemen schon mit recht einfachen Funktionen zu lösen sind, andererseits muss man erwarten, dass diese Funktionen keinesfalls ausreichen werden. Wir betrachten ein einfaches Beispiel:

(3.18) An einem Fluss g soll für die Versorgung der drei Städte A, B, C ein Hafen H eingerichtet werden. (Die Lagebeziehungen zwischen g, A, B und C sind in Abb. 3.8a (S. 64) angegeben.)

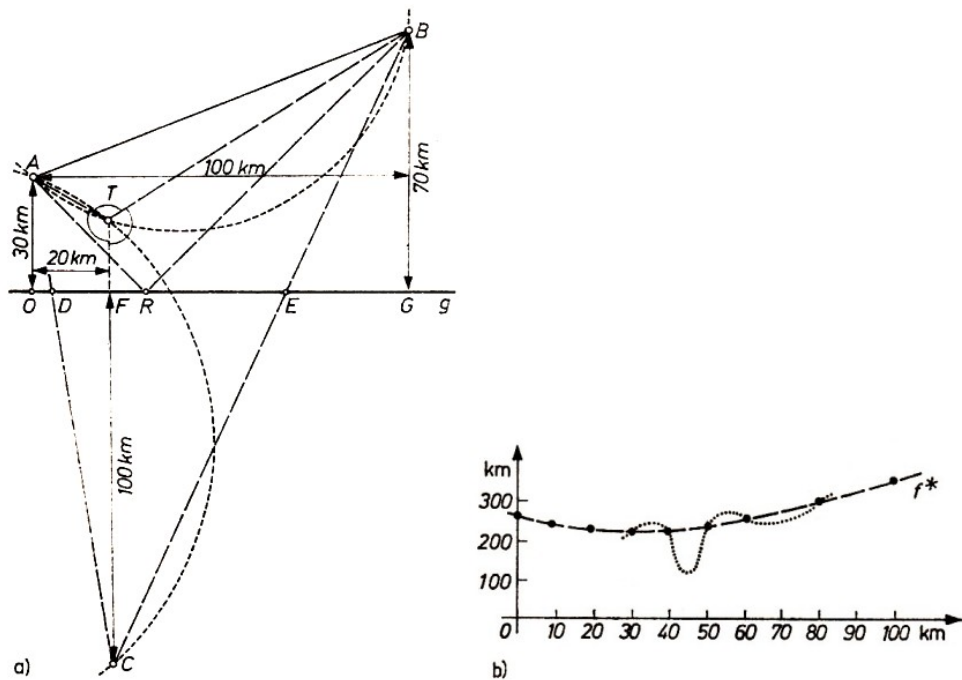


Abb. 3.8 a)

b)

An welcher Stelle muss H liegen, damit das "Wegenetz" $\overline{AH} + \overline{BH} + \overline{CH}$ möglichst kurz wird?

Diese Aufgabe erinnert uns an schon betrachtete. Würden wir A vernachlässigen, so wäre $H = E$, bei Vernachlässigung von B dagegen $H = D$ zu wählen (vgl. Abb. 3.8 a).

Berücksichtigen wir \overline{CH} nicht, so ergibt sich $H = R$ nach (2.1). Der Torricellische Punkt T (vgl. (2.12)) würde das Problem lösen, wenn überdies $T \in g$ gelten würde. Keiner der Punkte D, E, R oder T löst unser Problem (3.18)!

Versuchen wir, es mittels einer Funktion f zu erfassen: Es sei O der Fußpunkt des Lotes von A auf g , $x = |\overline{OH}|$ [in km]. Dann ergibt sich aus dem Satz des Pythagoras

$$f(x) = |\overline{AH}| + |\overline{BH}| + |\overline{CH}| = \sqrt{x^2 + 900} + \sqrt{(x - 20)^2 + 10000} + \sqrt{(100 - x)^2 + 4900} \quad (*)$$

Von der Funktion f mit der durch (*) gegebenen Funktionsgleichung suchen wir die Minimalstelle wo im Intervall $0 \leq x \leq 100$. Die Funktion f ist für alle reellen Zahlen in diesem Intervall definiert.

Jetzt nehmen wir an, dass ein solches x_0 in diesem Intervall eindeutig existiert, das erscheint zwar nach unseren Vorbetrachtungen einleuchtend, ist aber nicht bewiesen. Allerdings steht man gerade bei praxisbezogenen Aufgaben öfter vor dieser Situation, zunächst die Existenz und Eindeutigkeit eines Extremums annehmen zu müssen.

Wir können das Minimum angenähert aus dem Graphen f^* von $f(x)$ ablesen. Zur Ermittlung von f^* ermitteln wir Funktionswerte an einigen Stellen x ; und tragen diese in ein Koordinatensystem ein (vgl. Abb. 3.8 b).

In Abb. 3.8 b haben wir aus den Punkten \bullet auf den Verlauf (gestrichelte Linie) von f^* geschlossen; einen Verlauf, wie ihn etwa die punktierte Linie charakterisiert, haben wir damit ausgeschlossen. Wir nehmen also an, dass der Graph f^* "parabelähnlich" verläuft, was für die

Anwendung der folgenden Näherungsmethode von grundsätzlicher Bedeutung ist. Allgemeiner und exakter sollten wir formulieren:

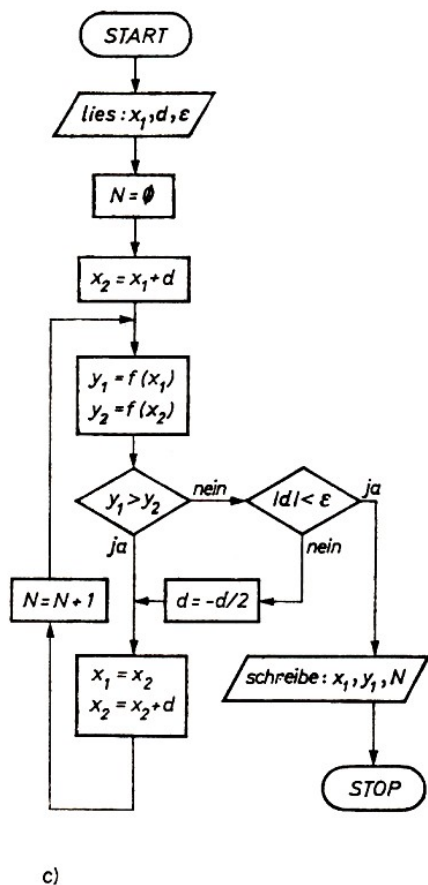


Abb. 3.8 c

Eine Parabel ist eindeutig durch drei voneinander verschiedene Punkte, die auf dieser Parabel liegen, bestimmt. Sind $(x_1, y_1), (x_2, y_2), (x_3, y_3)$ diese Punkte, so muss

$$ax_1^2 + bx_1 + c = y_1, \quad ax_2^2 + bx_2 + c = y_2, \quad ax_3^2 + bx_3 + c = y_3$$

gelten. Subtrahiert man die erste Gleichung sowohl von der zweiten als auch von der dritten, so erhält man

$$\begin{aligned} a(x_2^2 - x_1^2) + b(x_2 - x_1) &= y_2 - y_1 \\ a(x_3^2 - x_1^2) + b(x_3 - x_1) &= y_3 - y_1 \end{aligned}$$

Das ist ein lineares Gleichungssystem, aus dem sich eindeutig a und b und damit $\tilde{x}_0 = -\frac{b}{2a}$ ermitteln lassen. In Abhängigkeit von den benutzten Umformungen erhält man eine Darstellung für x_0 in Abhängigkeit von den Koordinaten der drei die Parabel bestimmenden Punkte. Für numerische Berechnungen (z. B. mit einem Taschenrechner) ist die folgende Form günstig:

$$2\tilde{x}_0 = x_1 + x_2 + \frac{x_3 - x_2}{1 - \frac{x_2 - x_1}{x_3 - x_1} \cdot \frac{y_3 - y_1}{y_2 - y_1}} \quad (0)$$

Verwenden wir in unserem Beispiel $x_1 = 0, x_2 = 50$ und $x_3 = 100$, so ergibt sich $\tilde{x}_0 \approx 29,5$. Die nächste Näherungparabel wird man demnach etwa durch $x_1 = 20, x_2 = 30$ und $x_3 = 40$ bestimmen. (0) liefert uns $\tilde{x}_0 \approx 26,8$. Da bei jedem Schritt die gleiche Formel (0) benutzt wird, kann man sich wohl vorstellen, dass dieses Verfahren auf einem Computer sehr schnell

den gesuchten Wert mit beliebiger Genauigkeit ermittelt.

Die Realisierung auf einem Rechner ist auch wesentlich für die

3. Möglichkeit: Die Beschreibung des Verfahrens geben wir durch den in Abb. 3.8 c dargestellten Programmablaufplan (PAP).

Vertieft man sich in diesen PAP, so erkennt man seine Wirkungsweise. So wie in den vorherigen Verfahren wird auch hier vorausgesetzt, dass ein Minimum an genau einer Stelle x_0 existiert. Mit dem durch den PAP beschriebenen Algorithmus wird diese Stelle x_0 mit vorgegebener Genauigkeit ε berechnet.

Der Startwert x_1 wird kleiner als x_0 gewählt, z. B. $x_1 = 0$. Mit der Schrittweite d läuft man nun auf die Minimalstelle zu bis man - durch Vergleich der Funktionswerte - bemerkt, dass man diese bereits überschritten hat.

Jetzt läuft man mit der halben Schrittweite in die entgegengesetzte Richtung ($d = -\frac{d}{2}$). Dieses "Pendeln" um die Minimalstelle x_0 wird so lange fortgesetzt, bis die Schrittweite dem Betrag nach kleiner als die vorgegebene Zahl ε ist.

Man kann diesen PAP mit einem Taschenrechner abarbeiten. Wir haben ein entsprechendes BASIC-Programm für einen Kleinstrechner geschrieben.

Mit den Startwerten $x_1 = 0$ und $d = 5$ und $\varepsilon = 0,5$ (bzw. $\varepsilon = 0,1$) erreicht man mit 15 (bzw. 20) Zyklusdurchläufen (Zählvariable N) jeweils nach einigen Sekunden Rechenzeit $x_0 = 26 \pm 0,5$ (bzw. $x_0 = 26,4 \pm 0,1$).

In beiden Fällen erhält man den zugehörigen Funktionswert $y_0 \approx 241,74$. Diese damit erreichte Genauigkeit ist bezüglich des gestellten praktischen Problems (3.18) unangemessen hoch. Richtet man den Hafen in etwa 26,5 km von O entfernt ein, so wird das Wegenetz mit etwa 241,74 km minimale Länge haben. Der Leser vergleiche dieses Ergebnis mit seinen Erwartungen an Hand von Abb. 3.8 a.

Wohlgermerkt ging es uns bei den obigen Darlegungen in erster Linie um die Demonstration prinzipieller Potenzen, die die numerischen Verfahren anbieten. Bei unseren drei Verfahren handelte es sich stets um ein "durch die Anschauung gestütztes Abtasten".

Man wird ahnen, dass die "höhere Mathematik" vielfältigere und effektivere Methoden zur Verfügung stellt. Diese Hinweise sind auch bei dem abschließenden Beispiel zu beachten. Wir wählen eine Aufgabe, deren "duale Formulierung" wir später als Aufgabe (4.10) lösen werden:

(3.19) Man ermittle unter allen Quadern mit dem Umfang 4 diejenigen mit dem maximalen Volumen.

Die Kantenlängen der Quader seien mit x, y und z bezeichnet. Dann gilt $V = x \cdot y \cdot z$ und $x + y + z = 1$. Daraus folgt

$$V = V(x, y) = xy(1 - x - y)$$

für alle x, y mit $x + y < 1$.

Ein Maximum kann man nun schrittweise angenähert ermitteln, indem man abwechselnd x und y fest wählt.

Wir beginnen mit $x = 0,5$: $V(0,5, y) = 0,5y(0,5 - y)$ ist die Funktionsgleichung einer Parabel, deren Maximum nach (3.6) bei $y_0 = 0,25$ liegt.

Wir wählen folglich $y_0 = 0,25$ und betrachten $V(x, 0,25) = 0,35x(0,75 - x)$, das ist eine quadratische Funktion mit dem Maximum bei $x_0 = \frac{3}{8}$.

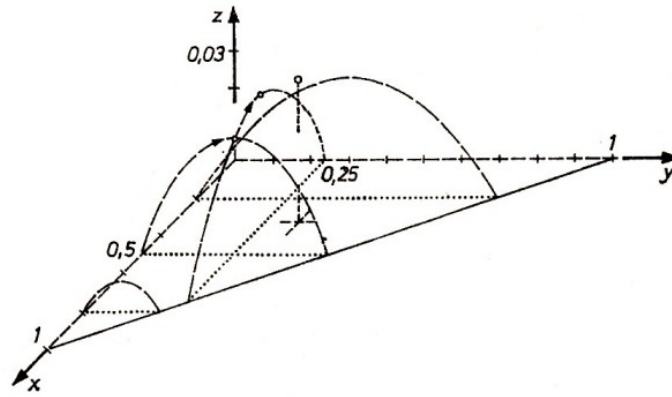


Abb. 3.9

$V\left(\frac{3}{8}, y\right)$ hat ein Maximum bei $y_1 = \frac{5}{16}$.

So fortfahrend erhält man schließlich $x_n \approx y_n \approx \frac{1}{3}$. Daraus folgt $z_n \approx \frac{1}{3}$. Unter den Quadern ist also (in Näherung) der Würfel ein solcher mit maximalem Volumen.

Veranschaulicht man sich die Funktion $V(x, y)$ in einem räumlichen Koordinatensystem, so kann man diesen Näherungsprozess gut verfolgen.

Aus der Darstellung der Funktion folgt auch die Existenz und Einzigkeit des Maximum (vgl. Abb. 3.9).

4 Ungleichungen und Extrema

4.1 Zusammenhänge zwischen Funktionen und Ungleichungen

Nicht überall, wo Wasser ist, sind Frösche,
aber wo man Frösche hört, ist Wasser.
J. W. v. Goethe

Nimmt eine Funktion an einer Stelle einen Extremwert an, so lässt sich das wie in (3.2) mittels einer Ungleichung beschreiben. Ordnet man umgekehrt einer Ungleichung eine Funktion zu, so muss man mit Schlussfolgerungen vorsichtiger sein.

Jede Ungleichung in einer Variablen x über einer Teilmenge U der reellen Zahlen \mathbb{R} , kann man so umformen, dass sie die Gestalt

$$T(x) \geq 0 \quad (*) \quad \text{oder} \quad T(x) > 0 \quad (**)$$

annimmt. $T(x)$ ist dabei ein Term in x , und $(*)$ bzw. $(**)$ wird immer nur von einer bestimmten Menge $L \subseteq U \subseteq \mathbb{R}$ erfüllt. L ist die Lösungsmenge der Ungleichung. (Bei Beweisaufgaben wird häufig verlangt zu zeigen, dass eine vorgegebene Ungleichung für alle $x \in K$ mit vorgegebenem K gilt. Dann muss natürlich $K \subseteq L$ gelten.)

Nun kann man zunächst dem Term $T(x)$ eine Funktion mit der Funktionsgleichung $y = T(x)$ und dem Definitionsbereich $D(T) = L$ zuordnen. Dann sind nach $(*)$ bzw. $(**)$ alle Funktionswerte nichtnegativ bzw. positiv.

Kann man aber daraus schließen, dass die Funktion ein Minimum besitzt und dieses den Wert 0 hat? Im Fall $(**)$ werden wir sicher nicht so leicht zu diesem Fehlschluss verleitet, denn hier besagt doch schon die Form der Ungleichung, dass der Wert 0 im Wertevorrat der zugeordneten Funktion gar nicht auftritt. Ein einfaches Gegenbeispiel ist auch schnell zur Hand:

Für alle positiven x gilt $\frac{1}{x} > 0$, die Funktion mit $y = \frac{1}{x}$ nimmt zwar beliebig kleine Funktionswerte an, niemals aber den Wert 0.²⁶

Leichter wird man im Fall $(*)$ zu Fehlschlüssen verleitet. Man beachte aber, dass man häufig für $(**)$ auch $(*)$ setzen kann. Wir hätten oben auch schreiben können: Für alle positiven x gilt $\frac{1}{x} \geq 0$.

Wird aber im Fall $(*)$ der Wert 0 für wenigstens ein x angenommen, so besitzt die Funktion T ein Minimum $T(x_{\min}) = 0$.

Aus $(*)$ auf die Existenz eines Minimums zu schließen, ist also nur zulässig, wenn man gezeigt hat, dass ein x_0 mit $T(x_{\min}) = 0$ existiert. Dieser gut zu beachtende Zusammenhang ist auch einer der Gründe dafür, dass man in Aufgaben über Ungleichungen vom Typ $(*)$ häufig ausdrücklich fordert:

"Man ermittle alle x , für die das Gleichheitszeichen gilt!" oder "Man beweise, dass das Gleichheitszeichen angenommen wird!"

Den soeben untersuchten Zusammenhang haben wir schon in Kapitel 3 benutzt: Aus der offensichtlich für alle reellen Zahlen x gültigen Ungleichung $(x - 1)^2 \geq 0$ folgt $x + \frac{1}{x} - 2 \geq 0$ für alle positiven x , und das Gleichheitszeichen gilt (genau) für $x = 1$.

²⁶In Lehrbüchern der höheren Mathematik führt man den Begriff des Infimums ein. In unserem Beispiel gilt $\inf \frac{1}{x} = 0$ für $x > 0$.

Die Funktion $f(x) = x + \frac{1}{x} - 2$ mit $D(f) = \mathbb{R}_+^*$ hat also (genau) an der Stelle $x_0 = 1$ das Minimum $f(1) = 0$.

Alle diese Überlegungen kann man leicht auf den Fall des Maximums übertragen.

(4.1) • Man ermittle alle Extrema der Funktion f mit $f(x) = \frac{x^2}{x^4 + 1}$ und $D(f) = \mathbb{R}$.

Bei diesen Beispielen waren uns Ungleichungen für die Ermittlung der Extrema einer Funktion nützlich. Kann man auch in der umgekehrten Richtung den Zusammenhang ausnutzen? Natürlich! Wir wollen zur Illustration eine sehr wichtige Ungleichung herleiten:

Wir knüpfen an unsere Ausführung über die Methode der kleinsten Quadrate an. Ersetzen wir dort die Symbole I_i durch x_i , U_i durch y_i und R durch m , so hat das gar keinen Einfluss auf die Herleitung, man sieht auch, dass man x_i und y_i nicht auf die positiven reellen Zahlen einschränken muss.

Die in m quadratische Funktion f mit $f(m) = \sum_{i=1}^n (mx_i - y_i)^2$ hat nur nichtnegative Funktionswerte, und ihr Minimum liegt bei $m_0 = \frac{\sum x_i y_i}{\sum x_i^2}$.

Es gilt natürlich auch $f(m_0) \geq 0$, bzw. - man setze ein und rechne einmal nach -

$$-\frac{(\sum x_i y_i)^2}{\sum x_i^2} + \sum y_i^2 \geq 0$$

Da $\sum x_i^2 > 0$ gilt, ist diese Ungleichung äquivalent mit

$$\sum x_i^2 \cdot \sum y_i^2 \geq \left(\sum x_i y_i\right)^2$$

Das Gleichheitszeichen gilt genau dann, wenn ein m so existiert, dass $mx_i = y_i$ für alle i gilt. (Das ist aus der obigen Darstellung von $f(m)$ ersichtlich.) Wir fassen unser Resultat zusammen:

(4.2) Die Cauchy-Schwarzsche Ungleichung: Für n Paare (x_i, y_i) beliebiger reeller Zahlen gilt

$$\sum_{i=1}^n x_i^2 \cdot \sum_{i=1}^n y_i^2 \geq \left(\sum_{i=1}^n x_i y_i\right)^2$$

Das Gleichheitszeichen in dieser Ungleichung gilt genau dann, wenn alle n Paare die Gestalt (x_i, mx_i) mit festem m haben.

Diese Art, Eigenschaften von Funktionen zur Ableitung von Ungleichungen auszunutzen, ist möglicherweise überraschend, aber sehr wichtig und anregend.

Im folgenden gehen wir dazu von konvexen oder konkaven Funktionen aus.

(4.3) Eine Funktion $y = f(x)$ heißt ("von unten gesehen"²⁷) konvex (bzw. konkav) im Intervall $[a, b]$ genau dann, wenn bezüglich des Graphen f^* von f im Intervall für jede beliebige Sehne \overline{AB} mit $A, B \in f^*$ der Bogen $\widehat{AB} \subseteq f^*$ unterhalb (bzw. oberhalb) oder auf der Sehne liegt.

²⁷Diese wenig exakt klingende Formulierung wurde von uns hinzugefügt, da in den verschiedenen Veröffentlichungen die Definitionen unterschiedlich, d. h. gerade entgegengesetzt sind. Die hier gewählte stimmt z. B. mit der in [5] gegebenen überein. Im Mathematischen Wörterbuch, Bd. I, S. 996, findet man den Begriff "Konvex von einer Seite": Verläuft eine ebene Kurve ... in einer Umgebung des Berührungspunktes einer Tangente ... ganz auf deren einer Seite, so heißt die Kurve von dieser Seite konkav, von der entgegengesetzten Seite konvex.

Insbesondere heißt dabei die Funktion streng konvex bzw. streng konkav, wenn die Bögen \widehat{AB} ganz unterhalb (bzw. oberhalb) der zugehörigen Sehne \overline{AB} liegen. (In Abb. 4.1a wurde eine streng konkave Funktion dargestellt.)

Der Begriff wurde von uns aus anschaulichen Gründen mit geometrischen Mitteln erklärt.²⁸ Aus der Definition folgt leicht folgende Eigenschaft.

(4.4) Ist $y = f(x)$ eine im Intervall $[a, b]$ streng konkave Funktion, so gilt für alle reelle Zahlen x_1, x_2 mit $a < x_1, x_2 < b$ und $x_1 \neq x_2$ die Ungleichung

$$\frac{f(x_1) + f(x_2)}{2} < f\left(\frac{x_1 + x_2}{2}\right) \quad (1)$$

Wie man an Hand der Abb. 4.1 b sofort erkennt, hat der Mittelpunkt M der Sehne \overline{AB} die Koordinaten $x_M = \frac{x_1 + x_2}{2}$, $y_M = \frac{f(x_1) + f(x_2)}{2}$, und y_M ist in der Tat nach Voraussetzung kleiner als die y -Koordinate $f\left(\frac{x_1 + x_2}{2}\right)$ desjenigen Punktes N auf dem Graphen der Funktion, der die gleiche x -Koordinate wie der Punkt M hat.

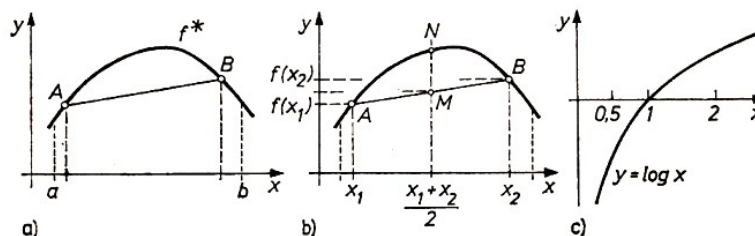


Abb. 4.1

(Für eine streng konvexe Funktion kehrt sich das Relationszeichen in (1) um.)

Ein Beispiel für eine streng konkave Funktion ist die Logarithmusfunktion $y = \log x$, und zwar im ganzen Definitionsbereich $0 < x < \infty$ (Abb. 4.1 c).

Nach (4.4) gilt also für alle voneinander verschiedenen reellen Zahlen $x_1, x_2 > 0$ die Ungleichung

$$\log \sqrt{x_1 x_2} = \frac{\log x_1 + \log x_2}{2} < \log \frac{x_1 + x_2}{2}$$

und auf Grund des Monotonieverhaltens der Logarithmusfunktion folgt daraus die Ungleichung

$$\sqrt{x_1 x_2} < \frac{x_1 + x_2}{2} \quad (2)$$

für alle reellen Zahlen $x_1, x_2 > 0$ mit $x_1 \neq x_2$.

Dies ist die bekannte Aussage, dass das geometrische Mittel zweier verschiedener positiver reeller Zahlen stets kleiner als ihr arithmetisches Mittel ist.

Der Satz (4.4) lässt sich verallgemeinern:

(4.5) Ist $y = f(x)$ eine im Intervall $[a, b]$ streng konkave Funktion und sind x_1, x_2, \dots, x_n beliebige reelle Zahlen aus dem Intervall, die nicht alle gleich sind, dann gilt

$$\frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n} < f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \quad (3)$$

Diese Ungleichung (3) heißt (spezielle) Jensensche Ungleichung. (Für $n = 2$ ergibt sich (1).)

²⁸Andere Möglichkeiten der Begriffserklärung sind die mittels der Differentialrechnung oder eben der von uns nun abzuleitenden Ungleichungen.

Zum Beweis benutzen wir aus der analytischen Geometrie die Aussage, dass die Koordinaten x_S, y_S des Schwerpunktes von n (Massen-) Punkten $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n)$ bestimmt sind durch²⁹

$$x_S = \frac{x_1 + x_2 + \dots + x_n}{n}, \quad y_S = \frac{y_1 + y_2 + \dots + y_n}{n}$$

Die behauptete Ungleichung (3) ist aus geometrischer Sicht nun leicht zu begründen. Wir können o. B. d. A. davon ausgehen, dass $x_1 \leq x_2 \leq \dots \leq x_n$ ist und dass hier wenigstens drei verschiedene reelle Zahlen vorliegen.

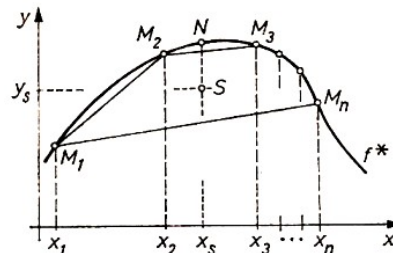


Abb. 4.2

Die Punkte $M_1(x_1, y_1), M_2(x_2, y_2), \dots, M_n(x_n, y_n)$ mit $y_i = f(x_i)$ ($i = 1, \dots, n$) liegen auf dem Graph der Funktion und bilden nach Voraussetzung die Ecken eines konvexen Vielecks (Abb. 4.2). Der Schwerpunkt $S(x_s, y_s)$ dieser Figur liegt in ihrem Innern und demnach unterhalb desjenigen Punktes N des Graphen, der die gleiche x -Koordinate $x_s = \frac{x_1 + x_2 + \dots + x_n}{n}$ wie der Schwerpunkt S besitzt; also gilt

$$y_S < f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$$

d. h.

$$\frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n} < f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$$

Ist die Funktion $y = f(x)$ streng konvex - wie z. B. $y = x^2$ im Intervall $-\infty < x < \infty$ oder $y = \tan x$ im Intervall $0 < x < \frac{\pi}{2}$ -, dann ist nach den bisherigen Überlegungen

$$\frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n} > f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \quad (3')$$

als Jensensche Ungleichung zu verstehen, wobei wieder wenigstens zwei der Zahlen x_1, \dots, x_n voneinander verschieden sein müssen.

Wählen wir wieder als Beispiel die streng konkave Logarithmusfunktion, so erhalten wir aus dem Satz (4.5) als Folgerung:

$$\frac{\log x_1 + \log x_2 + \dots + \log x_n}{n} < \log\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right)$$

und damit

$$\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} < \frac{x_1 + x_2 + \dots + x_n}{n} \quad (4)$$

für alle positiven reellen Zahlen x_1, \dots, x_n von denen wenigstens zwei voneinander verschieden sind.³⁰

²⁹Siehe dazu etwa [5], S. 319.

³⁰Eine andere Herleitung von (4) findet man z. B. in [9], S. 21 ff.

Dies ist der Satz vom geometrischen und arithmetischen Mittel; die Ungleichung (2) ist ein Spezialfall von (4).

Eine weitere nützliche Ungleichung folgt aus (4.5) mit Hilfe der Potenzfunktion $y = x^m$ ($m \geq 2$, ganzzahlig). Diese Funktion ist für $0 \leq x < \infty$ streng konvex. Also gilt für alle reellen Zahlen $x_1, \dots, x_n \geq 0$, von denen wenigstens zwei voneinander verschieden sind, die Ungleichung

$$\frac{x_1^m + x_2^m + \dots + x_n^m}{n} > \left(\frac{x_1 + x_2 + \dots + x_n}{n} \right)^m \quad (5)$$

bzw.

$$\frac{x_1 + x_2 + \dots + x_n}{n} < \sqrt[m]{\frac{x_1 + x_2 + \dots + x_n}{n}} \quad (5')$$

Dem auf der rechten Seite von (5') stehenden Term nennt man das " m -te Potenzmittel der Zahlen x_1, \dots, x_n ".

Aus unseren bisherigen Betrachtungen folgt auch:

(4.6) Die Ungleichungen (1), (2), (3), (3'), (4), (5) und (5') sind genau dann durch Gleichungen zu ersetzen, wenn alle x_i paarweise gleich sind, bzw. in den Fällen (1), (3) und (3') auch dann, wenn $f(x)$ eine lineare Funktion ist (die also weder streng konkav noch streng konvex ist).

Die Sätze (4.4) und (4.5) wie auch die Ungleichungen (2), (4) und (5) im besonderen finden vielfältige Anwendungen. Wir können das hier nur an einigen wenigen Aufgaben zeigen. Zunächst stehen trigonometrische Funktionen im Mittelpunkt, im nächsten Kapitel dann speziell die Ungleichung (4).

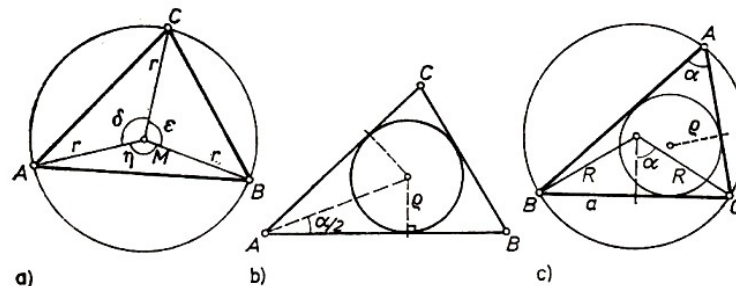


Abb. 4.3

Wir beginnen mit der Lösung der uns schon bekannten Aufgabe (2.43a). Zunächst überlege sich der Leser, dass der Flächeninhalt des einbeschriebenen Dreiecks sicher noch vergrößert werden könnte, wenn der Mittelpunkt M des Umkreises nicht im Inneren des Dreiecks liegt. Danach kann man mit den Bezeichnungen der Abb. 4.3a und unter Verwendung des "Flächensatzes" den Inhalt eines einbeschriebenen Dreiecks angeben:

$$A_{\Delta} = \frac{1}{2}r^2 \sin \delta + \frac{1}{2}r^2 \sin \varepsilon + \frac{1}{2}r^2 \sin \eta$$

$$A_{\Delta} = \frac{1}{2}r^2(\sin \delta + \sin \varepsilon + \sin \eta)$$

mit $\delta + \varepsilon + \eta = 2\pi$. Der Inhalt A_{Δ} wird nun genau dann maximal, wenn $s = \sin \delta + \sin \varepsilon + \sin \eta$ maximal wird. Da nach den Vorüberlegungen jeder der Winkel δ, ε und η kleiner als π und die Funktion $y = \sin x$ im Intervall $[0, \pi]$ streng konkav ist, gilt nach (3) und (4.6)

$$\frac{1}{3}(\sin \delta + \sin \varepsilon + \sin \eta) = \sin \frac{\delta + \varepsilon + \eta}{3}$$

d. h.

$$s \leq 3 \cdot \sin \frac{2\pi}{3} = \frac{3}{2}\sqrt{3}$$

Wesentlich ist, dass das Maximum für s und damit für A_{Δ} genau dann angenommen wird, wenn $\delta = \varepsilon = \eta = 120^\circ$ gilt, d. h., wenn das Dreieck gleichseitig ist.

(4.7) Es sei A_{Δ} der Flächeninhalt eines Dreiecks und A_k der Inhalt des zugehörigen Inkreises. Man bestimme das Maximum des Verhältnisses $q = \frac{A_k}{A_{\Delta}}$. Für welche Dreiecke liegt dieses Extremum vor?

Zur Lösung gehen wir von bekannten trigonometrischen Sachverhalten aus. Entsprechend den in Abb. 4.3b angegebenen Bezeichnungen gilt $A_k = \pi\rho^2$ sowie

$$A_{\Delta} = 2 \left(\frac{1}{2}\rho^2 \cot \frac{\alpha}{2} + \frac{1}{2}\rho^2 \cot \frac{\beta}{2} + \frac{1}{2}\rho^2 \cot \frac{\gamma}{2} \right) = \rho^2 \left(\cot \frac{\alpha}{2} + \cot \frac{\beta}{2} + \cot \frac{\gamma}{2} \right)$$

und damit

$$q = \frac{\pi}{\cot \frac{\alpha}{2} + \cot \frac{\beta}{2} + \cot \frac{\gamma}{2}}$$

Nun ist q maximal genau dann, wenn $q^* := \cot \frac{\alpha}{2} + \cot \frac{\beta}{2} + \cot \frac{\gamma}{2}$ minimal ist.

Die Winkelgrößen $\frac{\alpha}{2}$, $\frac{\beta}{2}$ und $\frac{\gamma}{2}$ liegen zwischen 0 und $\frac{\pi}{2}$, und in dem Intervall $(0, \frac{\pi}{2})$ ist die Funktion $y = \cot x$ streng konvex. Nach (3') gilt dann

$$\frac{q^*}{3} = \frac{\cot \frac{\alpha}{2} + \cot \frac{\beta}{2} + \cot \frac{\gamma}{2}}{3} > \cot \left(\frac{\frac{\alpha}{2} + \frac{\beta}{2} + \frac{\gamma}{2}}{3} \right) = \cot \frac{\pi}{6} = \sqrt{3}$$

wenn wenigstens zwei der drei Winkelgrößen α, β, γ voneinander verschieden sind. Nach (4.6) gilt $\frac{q^*}{3} = \sqrt{3}$ nur für $\alpha = \beta = \gamma = \frac{\pi}{3}$. Folglich ist $q_{\max} = \frac{\pi}{3\sqrt{3}} \approx 0,60$; und dieses Maximum liegt für die gleichseitigen Dreiecke und nur für diese vor.

(4.8) Es seien α, β, γ die Größen der Innenwinkel eines Dreiecks. Für welche Dreiecke ist

$$\frac{\sin \alpha + \sin \beta + \sin \gamma}{\sin \alpha \cdot \sin \beta \cdot \sin \gamma}$$

maximal?

Wegen $\sin \alpha, \sin \beta, \sin \gamma > 0$ gilt nach (4) zunächst

$$\sin \alpha \sin \beta \sin \gamma \leq \left(\frac{\sin \alpha + \sin \beta + \sin \gamma}{3} \right)^3$$

wobei das Gleichheitszeichen wegen (4.6) für $\sin \alpha = \sin \beta = \sin \gamma$ steht. Also ist

$$\frac{\sin \alpha + \sin \beta + \sin \gamma}{\sin \alpha \cdot \sin \beta \cdot \sin \gamma} \geq \frac{27}{(\sin \alpha + \sin \beta + \sin \gamma)^2}$$

Da die Funktion $y = \sin x$ im Intervall $0 < x < \pi$ streng konkav ist, gilt

$$\sin \alpha + \sin \beta + \sin \gamma \leq 3 \cdot \sin \left(\frac{\alpha + \beta + \gamma}{3} \right) = 3 \cdot \sin \frac{\pi}{3} = 3 \cdot \frac{\sqrt{3}}{2}$$

wobei das Gleichheitszeichen hier nur für $\alpha = \beta = \gamma = \frac{\pi}{3}$ steht. Folglich ist

$$\frac{\sin \alpha + \sin \beta + \sin \gamma}{\sin \alpha \cdot \sin \beta \cdot \sin \gamma} \leq \frac{27}{\left(3 \cdot \frac{\sqrt{3}}{2}\right)^2} = 4$$

und Gleichheit besteht nur für die gleichseitigen Dreiecke.

Wir verwenden das gewonnene Resultat weiter:

Sind R und ρ die Radien des Um- bzw. Inkreises eines Dreiecks, so gilt für den Flächeninhalt A des Dreiecks bekanntlich (Abb. 4.3 c):

$$A = \frac{1}{2}ab \sin \gamma = \frac{1}{2}(2R \sin \alpha)(2R \sin \beta) \sin \gamma = 2R^2 \sin \alpha \sin \beta \sin \gamma$$

sowie

$$A = \frac{1}{2}\rho(a + b + c) = \frac{\rho}{2}2R(\sin \alpha + \sin \beta + \sin \gamma)$$

Daraus folgt

$$\frac{R}{\rho} = \frac{1}{2} \frac{\sin \alpha + \sin \beta + \sin \gamma}{\sin \alpha \cdot \sin \beta \cdot \sin \gamma}$$

Das Ergebnis aus (4.8) hat dann zur Folge, dass

$$\frac{R}{\rho} \geq 2$$

ist und Gleichheit nur für gleichseitige Dreiecke besteht.

Diese Extremalaussage für Dreiecke kann auch elementargeometrisch gewonnen werden, wenn man von einem Dreieck ABC zu den Mitten A', B', C' der Seiten $\overline{BC}, \overline{CA}, \overline{AB}$ übergeht, die ein zum Dreieck ABC ähnliches Dreieck (mit dem Ähnlichkeitsfaktor $\frac{1}{2}$ bilden (vgl. [12])).

(4.9) • Es sei A_{Δ} der Flächeninhalt eines Dreiecks und A_k der Inhalt des zugehörigen Umkreises. Man bestimme für alle Dreiecke das Minimum des Verhältnisses $p = \frac{A_k}{A_{\Delta}}$. Für welche Dreiecke liegt dieses Extremum vor?

Kombiniert man die Ergebnisse von (4.7) und (4.9), so kann man wiederum die Ungleichung $R \geq 2\rho$ ableiten.

4.2 Das Lösen von Extremalaufgaben unter Verwendung der Ungleichung über das arithmetische und geometrische Mittel

Wenn man dem Faden nachgeht,
findet man auch den Knäuel.
M. de Cervantes

Genau besehen haben wir die Ungleichung (2) in Abschnitt 4.1 schon benutzt, z. B. bei der Lösung von (2.27) oder dem zugehörigen dualen Problem (2.33).

Interpretiert man in der Ungleichung $\frac{x_1 + x_2}{2} \geq \sqrt{x_1 x_2}$ die linke Seite als ein Viertel des Umfangs eines Rechtecks mit den Seitenlängen x_1 und x_2 , dann stellt die rechte Seite die Wurzel aus dem Flächeninhalt dieses Rechteckes dar. Ist nun eine dieser Seiten konstant, so ist der Term der anderen Seite durch diese Konstante (nach unten oder oben) beschränkt,

das Gleichheitszeichen gilt jeweils für das Extremum. Man beachte, wie sich die Dualität der Problemstellungen in der Ungleichung widerspiegelt!

Eine analoge Aufgabenstellung für den Raum erfordert kaum zusätzliche Überlegungen.



Abb. 4.4

(4.10) Quaderförmige Pakete, die ein vorgegebenes Volumen V haben, sollen mit Bindfaden (wie in Abb. 4.4 angegeben) verschnürt werden. Welche Gestalt müssen die Pakete haben, damit möglichst wenig Bindfaden verbraucht wird?

Die Kantenlängen des Quaders seien x_1, x_2, x_3 . Die benötigte Bindfadenlänge ist dann $4(x_1 + x_2 + x_3)$, sie wird genau dann minimal, wenn $\frac{x_1+x_2+x_3}{3}$ minimal wird.

Nach der Ungleichung (4) aus Abschnitt 4.1 gilt aber

$$\frac{x_1 + x_2 + x_3}{3} \geq \sqrt[3]{x_1 \cdot x_2 \cdot x_3}$$

und das Gleichheitszeichen gilt für den Fall $x_1 = x_2 = x_3$.

Unabhängig von der Wahl der Belegung der Variablen gilt stets $x_1 \cdot x_2 \cdot x_3 = V$, also kann die linke Seite nie kleiner als $\sqrt[3]{V}$ werden, das Minimum wird für $x_1 = x_2 = x_3$, d. h. für eine Würfelform angenommen.

Möchte man für Dreiecke entsprechende Resultate (wie in (2.28) bzw. (2.35)) erzielen, so ist eine Formel nötig, die den Flächeninhalt in Abhängigkeit von allen drei Seiten angibt, das ist gerade die Heronsche Formel

$$A = \sqrt{s(s-a)(s-b)(s-c)} \quad \text{mit} \quad s = \frac{1}{2}(a+b+c) \quad (*)$$

Ist nun beispielsweise A unter der Annahme $s = \text{const}$ zu maximieren, so ergibt sich die Lösung etwa wie folgt: (*) ist äquivalent mit

$$\sqrt[3]{\frac{A^2}{s}} = \sqrt[3]{(s-a)(s-b)(s-c)}$$

Aus der Ungleichung (4) ergibt sich mit $x_1 = s-a, x_2 = s-b, x_3 = s-c$

$$\sqrt[3]{(s-a)(s-b)(s-c)} \leq \frac{s-a+s-b+s-c}{3} = \frac{1}{3}s$$

das arithmetische Mittel erweist sich somit wieder als konstant. Folglich wird das geometrische Mittel und damit auch A maximal genau für $s-a = s-b = s-c$, d.h. für $a = b = c$.

Aus

$$\sqrt[3]{\frac{A^2}{s}} \leq \frac{1}{3}s \quad \text{folgt} \quad A \leq \sqrt{3} \left(\frac{s}{3}\right)^2 \quad (**)$$

Hätten wir statt dieser beiden Beispiele die zugehörigen dualen betrachtet, so hätte das offensichtlich keinen wesentlichen Unterschied gemacht.

Die nächste Aufgabe schließt an das letzte Beispiel an.

(4.11) Wir betrachten alle Tetraeder mit den Kantenlängen a, b, c, d, e, f , für die

$$a^2 + b^2 + c^2 + d^2 + e^2 + f^2 = q = \text{const} \quad (0)$$

gilt. Man ermittle unter diesen Tetraedern alle diejenigen, für die der Oberflächeninhalt maximal wird, und gebe diesen maximalen Inhalt an.³¹

Es seien x, y, z die Seitenlängen irgendeines der vier Dreiecke der Tetraederoberfläche. Wegen (**) gilt für diese Dreiecksfläche

$$A \leq \sqrt{3} \frac{s^2}{9} = \frac{\sqrt{3}}{4} \left(\frac{x+y+z}{3} \right)^2$$

Jetzt benutzen wir, um Anschluss an (0) zu erhalten, die Ungleichung (5). Wir erhalten

$$A \leq \frac{\sqrt{3}}{4} \left(\frac{x+y+z}{3} \right)^2 \leq \frac{\sqrt{3}}{4} \frac{x^2 + y^2 + z^2}{3}$$

Das Gleichheitszeichen gilt an beiden Stellen genau für $x = y = z$!

Addieren wir die vier Ungleichungen (die sich für die vier Dreiecke der Tetraederoberfläche ergeben) mit den entsprechenden Bezeichnungen, so erhalten wir für die Tetraederoberfläche

$$O \leq \frac{\sqrt{3}}{4} \cdot \frac{1}{3} (2a^2 + 2b^2 + \dots + 2f^2) = \frac{\sqrt{3}}{6} q$$

Das Gleichheitszeichen gilt nur, wenn es in jeder der Ungleichungen gilt; bei dem gesuchten Tetraeder handelt es sich also um ein regelmäßiges Tetraeder mit $a = \frac{\sqrt{q}}{6}$ seine Oberfläche beträgt $O_{\max} = \frac{\sqrt{3}}{6} q$.

(4.12) •* Lässt sich die gleiche Frage wie in (4.11) für Oktaeder (konvexe Polyeder mit acht dreieckigen Seitenflächen und $\frac{8 \cdot 3}{2} = 12$ Kanten) und für Ikosaeder (konvexe Polyeder mit 20 dreieckigen Seitenflächen und $\frac{20 \cdot 3}{2} = 30$ Kanten) stellen? Zu welchem Resultat gelangt man hier gegebenenfalls?

Setzt man in (2) $x_1 = x$ und $x_2 = \frac{d}{x}$, so erhält man $x + \frac{d}{x} \geq 2\sqrt{d}$, wobei das Gleichheitszeichen nur für $x_1 = x_2$, d. h. für $x = \sqrt{d}$ gilt.

Damit haben wir wiederum ein bekanntes Resultat gefunden, nämlich das, auf dem der gesamte Abschnitt 3.3 aufbaute.

Lässt sich nun von der nur geringfügig (?) anderen Funktion f mit

$$f(x) = x + \frac{a}{x^2}, \quad x > 0, \quad a = \text{const und } a > 0$$

das Minimum - falls überhaupt eines existiert - ermitteln?

Der Versuch, in (2) $x_1 = x$ und $x_2 = \frac{a}{x^2}$ zu setzen, liefert $x + \frac{a}{x^2} \geq 2\sqrt{\frac{a}{x}}$ und schlägt damit fehl. Warum?

Mit der Veränderung des Arguments x ändert sich nicht nur der Wert der linken Seite der erhaltenen Ungleichung, sondern auch der der rechten Seite. Damit liefert also keine der Seiten

³¹Diese Aufgabe wurde aus OJM 211246 A abgeleitet.

eine feste Schranke für die andere. Unsere mehrfach mit Erfolg praktizierte Methode ist wohl doch nicht so universell anwendbar.

Vielleicht lässt sie sich noch einmal retten, analysieren wir den Fehlschlag!

Das Scheitern des Versuchs ist doch wesentlich daran erkennbar, dass das geometrische Mittel keine konstante Größe darstellt. Das liegt aber z. B. daran, dass in dem Produkt $x \cdot \frac{a}{x^2}$ der Faktor x nur einmal statt zweimal vorkommt.

Könnten wir erreichen, dass er ein zweites Mal auftritt? Dann müsste x auch im arithmetischen Mittel zweimal als Summand enthalten sein. Was noch nicht ist, kann in dem Fall geschaffen werden: $x = \frac{1}{2}x + \frac{1}{2}x$ ist die zunächst als Trick erscheinende Lösung des Problems. Es gilt unter Verwendung von (4) mit $n = 3$

$$x + \frac{a}{x^2} = \frac{1}{2}x + \frac{1}{2}x + \frac{a}{x^2} \geq 3 \cdot \sqrt[3]{\frac{x}{2} \cdot \frac{x}{2} \cdot \frac{a}{x^2}} = 3 \cdot \sqrt[3]{\frac{a}{4}}$$

wobei das Gleichheitszeichen nach (4.6) nur für $\frac{1}{2}x = \frac{1}{2}x = \frac{a}{x^2}$, d. h. für $x_0 = \sqrt[3]{2a}$ gilt. Der minimale Funktionswert ist

$$f(x_0) = \sqrt[3]{2a} + \frac{a}{\sqrt[3]{4a^2}} = 3 \cdot \sqrt[3]{\frac{a}{4}}$$

denn für alle anderen x - das beweist unsere Herleitung - gilt $f(x) > f(x_0)$.

Wir verallgemeinern dieses gewonnene Resultat:

(4.13) Gegeben sei eine Funktion f mit der Funktionsgleichung $f(x) = \lambda x + \frac{a}{x^m}$, $m \in \mathbb{N}^*$, und dem Definitionsbereich $D(f) = \mathbb{R}_+^*$. Es seien λ und a fest vorgegebene positive Konstanten. Man ermittle das Minimum dieser Funktion.

Die Lösung finden wir im Prinzip wie oben. Wir zerlegen λx in m Summanden $\frac{\lambda}{m}x$. In (4) ist dann $n = m + 1$, $x_1 = x_2 = \dots = x_{n-1} = \frac{\lambda}{m}x$ und $x_n = \frac{a}{x^m}$. Damit ergibt sich aus (4)

$$\lambda x + \frac{a}{x^m} \geq (m+1) \sqrt[3]{\left(\frac{\lambda}{m}\right)^m \cdot a}$$

und das Gleichheitszeichen gilt nach (4.6) genau für $\frac{\lambda}{m}x = \frac{a}{x^m}$, d. h. für $x_0 = \sqrt[3]{\frac{am}{\lambda}}$. Der minimale Funktionswert ist

$$f(x_0) = (m+1) \sqrt[3]{\left(\frac{\lambda}{m}\right)^m \cdot a}$$

(Zur Kontrolle setze man $\lambda = 1$ und $m = 2$.)

Mit diesen Vorbereitungen können wir eine recht anspruchsvolle Aufgabe lösen:

(4.14)* Gegeben sei eine Funktion f von drei reellen Veränderlichen x, y, z mit der Funktionsgleichung

$$f(x, y, z) = \frac{a}{x^m} + \frac{b}{y^n} + \frac{c}{z^p}$$

a, b, c seien fest vorgegebene positive Konstanten, m, n und p seien positive natürliche Zahlen. Unter den Nebenbedingungen, dass x, y, z positiv sind und die Summe $x + y + z$ einen konstanten Wert k hat, ermittle man das Minimum von $f(x, y, z)$.

Bei den Versuchen, eine Lösung zu finden, muss man sich natürlich insbesondere fragen, wie die Nebenbedingungen $x + y + z = \text{const}$ einzubeziehen ist. Wir benutzen einen Trick, der der Euler-Lagrangeschen Multiplikatorenmethode der höheren Mathematik entlehnt ist.

Ist λ eine (positive) Konstante, so ist auch $\lambda(x + y + z)$ konstant, und die Funktion

$$g(x, y, z) = \lambda(x + y + z) + f(x, y, z)$$

nimmt genau dann ihr Minimum an, wenn das für $f(x, y, z)$ der Fall ist. Es gilt aber

$$g(x, y, z) = \lambda x + \frac{a}{x^m} + \lambda y + \frac{b}{y^n} + \lambda z + \frac{c}{z^p}$$

Außerdem sind alle Voraussetzungen erfüllt, um das Ergebnis aus (4.13) anzuwenden.

Die Funktion $g(x, y, z)$ und damit auch $f(x, y, z)$ wird demnach genau für das Tripel (x_0, y_0, z_0) mit

$$x_0 = \sqrt[m+1]{\frac{am}{\lambda}}, \quad y_0 = \sqrt[n+1]{\frac{bn}{\lambda}}, \quad z_0 = \sqrt[p+1]{\frac{cp}{\lambda}}$$

minimal, und $f(x_0, y_0, z_0)$ ist der zugehörige minimale Funktionswert. (Einschränkend ist anzumerken, dass unser Ergebnis statt der vorgegebenen Konstanten k die Konstante λ enthält. Der Zusammenhang zwischen beiden wird durch die Gleichung

$$\sqrt[m+1]{\frac{am}{\lambda}} + \sqrt[n+1]{\frac{bn}{\lambda}} + \sqrt[p+1]{\frac{cp}{\lambda}} = k$$

angegeben.)

Wir kehren jetzt zu der Art und Weise zurück, in der wir die Aufgabe (4.13) gelöst haben. Dieser "Trick einer geschickten Zerlegung" wird sich wohl noch öfter anwenden lassen. Wir wählen als Beispiel die Aufgabe (3.11). Laut Aufgabenstellung ist die Größe auf unter der Nebenbedingung $x^2 + y^2 = d^2 = \text{const}$ zu maximieren. Die Versuche, (4.6) direkt anzuwenden, scheitern:

$$\frac{x^2 + y^2}{2} \geq \sqrt{x^2 y^2} = xy \quad \text{oder} \quad \sqrt[3]{xyy} \leq \frac{x + 2y}{3}$$

Es wird aber xy^2 genau dann maximal, wenn $(xy^2)^2 = x^2 y^4$ maximal wird. Dieses Produkt orientiert auf ein geometrisches Mittel, im zugeordneten arithmetischen Mittel müssten x^2 und y^2 stehen. Da aber y im Produkt in einer höheren Potenz vorkommt, müsste sich y^2 aus mehreren Summanden zusammensetzen.

Wenn man dem Faden nachgeht, findet man auch den Knäuel! Wir finden

$$\sqrt[3]{\frac{1}{4}x^2 y^4} = \sqrt[3]{x^2 \cdot \frac{1}{2}y^2 \cdot \frac{1}{2}y^2} \leq \frac{x^2 + \frac{1}{2}y^2 + \frac{1}{2}y^2}{3} = \frac{x^2 + y^2}{3}$$

d. h., $\sqrt[3]{\frac{1}{4}x^2 y^4}$ wird genau dann maximal, wenn $x^2 = \frac{1}{2}y^2$ bzw. $y = \sqrt{2}x$ gilt. Für diesen maximalen Wert ist

$$\sqrt[3]{\frac{1}{4}x^2 y^4} = \frac{d^2}{3} \quad \text{bzw.} \quad x^2 y^4 = \frac{4d^6}{27}$$

ablesbar.

In [9], S. 24-28, und [12], S. 74/75 beispielsweise, findet man weitere Aufgaben, die auf diese Art gelöst werden, höchste Zeit, aus dem Trick eine Methode zu machen:

(4.15) Das Produkt $p = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ sei für alle n -Tupel positiver reeller Zahlen x_i konstant.

Die k_i ($i = 1, 2, \dots, n$) seien von 0 verschiedene natürliche Zahlen. Unter all diesen n -Tupeln ist dasjenige zu ermitteln, für welches die Summe

$$s = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

minimal wird. Die a_i ($i = 1, 2, \dots, n$) sind dabei positive reelle Zahlen.

Die Lösung dieser Aufgabe umfasst die von vielen speziellen Aufgaben. Wir nehmen die geschickte Zerlegung ein für alle Male vor: Betrachten wir die Darstellung

$$s = k_1 \cdot \frac{a_1}{k_1}x_1 + k_2 \cdot \frac{a_2}{k_2}x_2 + \dots + k_n \cdot \frac{a_n}{k_n}x_n$$

mit $t := k_1 + k_2 + \dots + k_n$, so lässt sich s auffassen als Summe von insgesamt t Summanden der Gestalt $\frac{a_i}{k_i}x_i$ ($i = 1, 2, \dots, n$).

s ist gleichzeitig zu interpretieren als das t -fache des arithmetischen Mittels dieser t Summanden. Das zugehörige geometrische Mittel enthält dann insgesamt t Faktoren der Gestalt $\frac{a_i}{k_i}x_i$ jeder dieser Faktoren tritt k_i -mal auf. Nach (4) und (4.6) gilt

$$\begin{aligned} s &\geq t \sqrt[t]{\left(\frac{a_1}{k_1}\right)^{k_1} x_1^{k_1} \left(\frac{a_2}{k_2}\right)^{k_2} x_2^{k_2} \dots \left(\frac{a_n}{k_n}\right)^{k_n} x_n^{k_n}} \\ s &\geq t \sqrt[t]{\left(\frac{a_1}{k_1}\right)^{k_1} \cdot \left(\frac{a_2}{k_2}\right)^{k_2} \dots \left(\frac{a_n}{k_n}\right)^{k_n}} \cdot \sqrt[t]{x_1^{k_1} \cdot x_2^{k_2} \dots x_n^{k_n}} \end{aligned}$$

Mit der als gültig vorausgesetzten Bedingung heißt das

$$s \geq t \sqrt[t]{\left(\frac{a_1}{k_1}\right)^{k_1} \cdot \left(\frac{a_2}{k_2}\right)^{k_2} \dots \left(\frac{a_n}{k_n}\right)^{k_n}} \cdot \sqrt[t]{p}$$

Die rechte Seite dieser Ungleichung ist also eine von den x_i unabhängige Konstante. s kann den Wert k dieser Konstanten nicht unterschreiten; nach (4.6) nimmt sie diesen Wert k genau dann an, wenn alle t Summanden untereinander gleich sind.

Da jeweils k_i Summanden nach Konstruktion einander gleich sind, bleibt die Bedingung

$$\frac{a_1}{k_1}x_1 = \frac{a_2}{k_2}x_2 = \dots = \frac{a_n}{k_n}x_n \quad (*)$$

zu erfüllen. Gilt (*), so folgt für jedes i ($i = 1, 2, \dots, n$)

$$s = t \frac{a_i}{k_i}x_i = t \sqrt[t]{\left(\frac{a_1}{k_1}\right)^{k_1} \cdot \left(\frac{a_2}{k_2}\right)^{k_2} \dots \left(\frac{a_n}{k_n}\right)^{k_n}} \cdot p$$

d. h.

$$x_i = \frac{k_i}{a_i} \sqrt[t]{\left(\frac{a_1}{k_1}\right)^{k_1} \cdot \left(\frac{a_2}{k_2}\right)^{k_2} \dots \left(\frac{a_n}{k_n}\right)^{k_n}} \cdot p \quad (**)$$

Mit (**) sind alle x_i des n -Tupels (x_1, x_2, \dots, x_n) eindeutig bestimmt. Für diese x_i gilt (*), und damit ist das Minimum von s ermittelt, und zwar eindeutig.

(4.16) • Gegeben sei die Bedingung, dass die Summe

$$s = a_1x_1 + a_2x_2 + \dots + a_nx_n$$

für alle n -Tupel positiver reeller Zahlen x_i konstant ist. Die a_i ($i = 1, 2, \dots, n$) seien positive reelle Zahlen. Unter all diesen n -Tupeln ist dasjenige zu ermitteln, für welches das Produkt

$p = x_1^{k_1} x_2^{k_2} \dots x_n^{k_n}$ maximal wird. Die k_i ($i = 1, 2, \dots, n$) sind dabei von 0 verschiedene natürliche Zahlen.

Anmerkungen. Die Aussagen (4.15) und (4.16) lassen sich auch auf den Fall übertragen, dass die k_i positive rationale Zahlen sind. Einen solchen Fall kann man nämlich durch geeignetes Potenzieren auf (4.15) oder (4.16) führen.

Weiterhin ist zu beachten, dass ein Teil der x_i auch durch $\frac{1}{x_i}$, x_i^2 u. a. ersetzbar ist bzw. dass solche Terme für die obigen x_i stehen.

(4.17) • Man stelle fest, inwieweit sich die vorher betrachteten Aufgaben des Abschnitts 4.2 als Spezialfälle der allgemeinen Probleme (4.15) bzw. (4.16) erweisen.

(4.18) • Man löse die Aufgaben (2.29), (2.39) und (2.48) unter Verwendung von (4.15) bzw. (4.16).

(4.19) • Man formuliere ein zu (2.27) räumliches Analogon und beweise es.

(4.20) • Man beweise : Unter allen geraden Doppelkegeln, die ein vorgegebenes Volumen haben, kommt der kleinste Oberflächeninhalt denjenigen Kegeln zu, deren Mantellinien von dem Berührungskreis der dem Doppelkegel einbeschriebenen Kugel im Verhältnis 2 : 1 geteilt werden.

(4.21) • Man ermittle alle Punkte im Innern eines gegebenen Dreiecks (bzw. Tetraeders), für die das Produkt ihrer Abstände von den Seiten (bzw. Flächen) maximal wird.

4.3 Variationen zum Thema

Ein Prinzip, das unbedingt alles erklären will,
erklärt in Wirklichkeit nichts.

A. Daltschew

Natürlich haben wir uns im letzten Abschnitt bemüht, einen Trick durch eine Methode zu ersetzen. Andererseits ist es aber verfehlt, nun wiederum alles "über einen Leisten schlagen" zu wollen.

Gerade die virtuose Anwendung der zur Verfügung stehenden Mittel macht den Meister aus. Erproben wir unsere Meisterschaft an einer Aufgabe!

(4.22) x_1, x_2, x_3, x_4 seien reelle Zahlen, und es gelte $0 < x_1 < x_2 < x_3 < x_4$.

a) Man beweise, dass dann die folgende Ungleichung gilt:

$$\sqrt{(x_4 - x_3)^2 + 1} + \sqrt{(x_3 - x_2)^2 + 1} + \sqrt{(x_2 - x_1)^2 + 1} \geq \sqrt{(x_4 - x_1)^2 + 3^2}$$

b) In welchen Fällen gilt das Gleichheitszeichen?

Zunächst betrachte man sorgfältig die Terme in (*). Welche Besonderheit fällt auf? Es gilt

$$(x_4 - x_3) + (x_3 - x_2) + (x_2 - x_1) = x_4 - x_1 \quad (**)$$

Die Summanden auf der linken Seite von (*) können wir als Funktionswerte $f(z)$ der Funktion f mit $f(z) = \sqrt{z^2 + 1}$ auffassen und erhalten wegen (**) an Stelle von (*) dann

$$f(z_4) + f(z_3) + f(z_2) \geq \sqrt{(z_4 + z_3 + z_2)^2 + 3^2}$$

Und da uns auf der rechten Seite im Radikand eine 1 statt der 3^2 lieber wäre, dividieren wir die gesamte Ungleichung durch 3:

$$\frac{f(z_4) + f(z_3) + f(z_2)}{3} \geq \sqrt{\left(\frac{z_4 + z_3 + z_2}{3}\right)^2 + 1} = f\left(\frac{z_4 + z_3 + z_2}{3}\right) \quad (***)$$

Da mit $z_i = x_i - x_{i-1}$ ($i = 4, 3, 2$) die Ungleichung (***) äquivalent mit (*) ist, wäre Teil a) schon bewiesen, wenn wir nachgewiesen hätten, dass auf f die Jensensche Ungleichung anwendbar ist! Dann müsste nach (4.5) bei Vergleich von (3) und (3') mit (***) die Funktion f mit $f(z) = \sqrt{z^2 + 1}$ für $z \geq 0$ konvex sein. Trifft das zu?

Wir können uns nicht nur von unserem Wunsch leiten lassen.

Die Funktion $g(z) = z^2 + 1$ ist konvex, muss dann aber auch $f(z) = \sqrt{g(z)}$ konvex sein? Wer das leichten Herzens bejaht, der sollte auch folgende Überlegung bedenken: $h(z) = \sqrt{z + 1}$ ist konkav, vielleicht auch $f(z) = h(z^2)$?

Hilft uns die graphische Darstellung bei der Entscheidungsfindung? (Wohlgermerkt, wir sprechen von einer Entscheidungshilfe, nicht von einem Beweis!)

Für $z \geq 0$ gilt

$$z = \sqrt{z^2} < \sqrt{z^2 + 1} \leq \sqrt{z^2 + 2z + 1} = z + 1$$

Das bedeutet, dass der Graph der Funktion f für alle z zwischen den beiden $y = z$ und $y = z + 1$ zugeordneten (zueinander parallelen) Geraden verläuft (vgl. Abb. 4.5a). Da ist wahrlich wenig Platz, ein ausgeprägtes Verhalten zu registrieren!

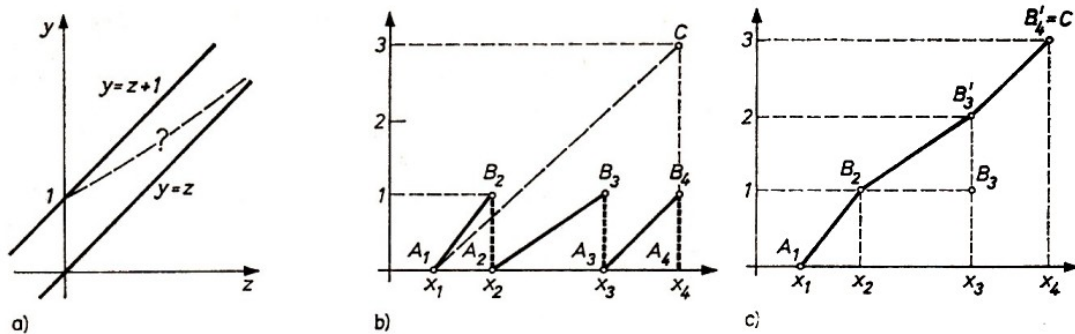


Abb. 4.5

Dieses Beispiel macht uns nachdrücklich auf eine Problematik aufmerksam, die wir bisher umgangen haben. Bei Verwendung der von uns unter geometrischen Gesichtspunkten vorgenommenen Definition (4.3) nahmen wir an, von einem Graphen f^* zu wissen, ob ihm die Eigenschaft "konvex" oder "konkav" zukommt. Und bei den bisher benutzten Funktionen ($f(x) = \lg x$, $f(x) = x^m$, $m \in \mathbb{N}$, $f(x) = \tan x$, $f(x) = \sqrt{x}$ - jeweils für "geeignete" Intervalle) haben wir diese Kenntnis einfach vorausgesetzt. Wie lässt sich aber die "Konvexität" beweisen?

Betrachten wir nochmals die Definition (4.3) und den Satz (4.4): $\frac{x_1+x_2}{2}$ und $\frac{f(x_1)+f(x_2)}{2}$ sind stets die Koordinaten des Mittelpunktes der "zugehörigen" Sehne. Gilt nun für alle Paare (x_1, x_2) mit $a \leq x_i \leq b$ und $x_1 < x_2$ die Ungleichung

$$\frac{f(x_1) + f(x_2)}{2} > f\left(\frac{x_1 + x_2}{2}\right) \quad (1')$$

dann besagt das doch gerade, dass jeder Sehnenmittelpunkt oberhalb des Graphen f^* liegt. Dann muss aber auch die gesamte Sehne oberhalb von f^* liegen. Denn würde ein Teil einer

Sehne unterhalb von f^* verlaufen, so existierte auch eine Sehne, deren Mittelpunkt unterhalb von f^* liegen würde. Damit ist gezeigt, dass aus der Gültigkeit von (1') die strenge Konvexität folgt (bzw. aus (1) in (4.4) die Eigenschaft "streng konkav").

Wir wenden diese Erkenntnis auf unser Beispiel an: Angenommen, es sei der Graph f^* von $f(z) = \sqrt{z^2 + 1}$ nicht streng konvex, dann existiert wenigstens ein Paar (z_1, z_2) mit $z_1 < z_2$ und

$$\left(\frac{z_1 + z_2}{2}\right) \geq \frac{f(z_1) + f(z_2)}{2} \quad \text{bzw.} \quad \sqrt{\left(\frac{z_1 + z_2}{2}\right)^2 + 1} \geq \frac{\sqrt{z_1^2 + 1} + \sqrt{z_2^2 + 1}}{2}$$

bzw.

$$\sqrt{(z_1 + z_2)^2 + 4} \geq \sqrt{z_1^2 + 1} + \sqrt{z_2^2 + 1}$$

Da beide Seiten dieser letzten Ungleichung positiv sind, dürfen wir quadrieren:

$$\begin{aligned} (z_1 + z_2)^2 + 4 &\geq z_1^2 + 1 + 2\sqrt{(z_1^2 + 1)(z_2^2 + 1)} + z_2^2 + 1 \\ z_1 z_2 + 1 &\geq \sqrt{(z_1^2 + 1)(z_2^2 + 1)} \\ (z_1 z_2 + 1)^2 &\geq (z_1^2 + 1)(z_2^2 + 1) \\ 2z_1 z_2 &\geq z_1^2 + z_2^2 \end{aligned}$$

bzw. $0 \geq (z_1 - z_2)^2$. Das ist aber ein Widerspruch zu der Voraussetzung, dass z_1, z_2 verschiedene reelle Zahlen sind. Folglich gilt für alle Paare (z_1, z_2) mit $z_1 < z_2$ die Ungleichung (1'), die strenge Konvexität ist nachgewiesen, die Aufgabe (4.22a) ist nach den vorbereitenden Betrachtungen gelöst.

Hier könnte jemand einwenden, dass an Stelle des Konvexitätsnachweises in der "gleichen Art" auch sofort (*) bewiesen werden könnte. Man versuche es! Das Auftreten des dritten Summanden in (*) erhöht den Aufwand beträchtlich. Weiterhin spricht für unser Vorgehen, dass es auch ohne Aufwandsvergrößerung auf den Fall n ($n \geq 3$) reeller Zahlen $0 < x_1 < x_2 < \dots < x_n$ und damit n Summanden in (*) anwendbar ist.

Da $f(z)$ streng konvex ist, kann in (1') das Relationszeichen " $<$ " nur dann durch " $=$ " ersetzt werden, wenn $z_2 = z_3 = z_4$ bzw. $x_2 - x_1 = x_3 - x_2 = x_4 - x_3$ gilt, d. h., wenn die x_i äquidistant liegen. Damit ist auch Teil b) gelöst.

Haben wir nicht trotz der Verwendung graphischer Darstellungen die elementaren geometrischen Methoden aus dem Auge verloren? Einseitigkeit ist vieler Fehler Anfang!

$\sqrt{t^2 + 1}^2$ kann auch die Aufforderung sein, an den Satz des Pythagoras zu denken. Tragen wir doch einfach einmal alle zugehörigen Dreiecke in ein Koordinatensystem ein (vgl. Abb. 4.5 b). (*) entspricht dann die Ungleichung

$$l := |A_1 B_2| + |A_2 B_3| + |A_3 B_4| \geq |A_1 C|$$

Erinnern wir uns an das Kapitel 2.1!

Ein Streckenzug $\overline{A_1 B_2 B_3 B_4}$ mit $B_4 = C$ wäre zu finden. Wegen $|A_2 B_2| = |A_3 B_3| = |A_4 B_4| = 1$ ist das aber leicht zu erreichen (vgl. Abb. 4.5c).

Der Streckenzug $\overline{A_1 B_2 B_3 B_4}$ hat nun nach Konstruktion die Länge 1, und nach Abschnitt 2.1 ist die geradlinige Verbindung von A_1 mit C die Kürzeste. Damit ist die Ungleichung $l \geq |A_1 C|$ bewiesen.

Das Gleichheitszeichen gilt nur, wenn B_2 und B'_3 in der Strecke $\overline{A_1C}$ liegen. In diesem Fall müssen die x_i voneinander gleichen Abstand haben. (Strahlensatz !).

(4.23) •³² Man beweise, dass für alle reellen Zahlen a, b und c

$$\sqrt{(a+c)^2 + b^2} + \sqrt{(a-c)^2 + b^2} \geq 2\sqrt{a^2 + b^2}$$

gilt.

(4.24) • Man beweise, dass für alle reellen Zahlen x_1, x_2, \dots, x_n und y_1, y_2, \dots, y_n

$$\sum_{i=1}^{n-1} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \geq \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2}$$

gilt.

4.4 Die Methode der konkurrierenden Punkte

Da gehen sie in tiefem Schweigen,
wohin? Das wird sich später zeigen.
W. Busch

Wollen wir Extremalaufgaben mit Hilfe einer Funktion einer reellen Veränderlichen lösen, so müssen wir zumindest eine zuzuordnende Funktion finden. Das Extremalproblem ist dann in Abhängigkeit von einer Veränderlichen formuliert. (Wir erinnern an die Beispiele in Kapitel 3!) Enthielt das ursprüngliche Problem mehrere Veränderliche, so mussten diese bis auf eine mittels der in der Aufgabe enthaltenen Nebenbedingungen eliminiert werden. In Abschnitt 4.2 sahen wir, wie man u. U. durch geschickte Anwendung von Ungleichungen mit mehreren Variablen arbeiten kann.

Das Geschick bestand darin, die Nebenbedingungen in die Ungleichung "einzubauen". In diesem Abschnitt werden jetzt Variable dadurch eingeführt, dass zwei Punkte in ihrer Beziehung zueinander betrachtet werden.

Der Vergleich des "Angebots", welches jeder dieser Punkte bezüglich des Extremums macht, lässt sich mit einer Ungleichung erfassen. Es werden also verschiedene der bereits betrachteten Methoden zur Anwendung kommen, in ihrer Gesamtheit aber doch etwas Neues ergeben.

Wir wollen diese Methode der konkurrierenden Punkte³³ zunächst an einer uns schon bekannten Aufgabe, nämlich (2.12), demonstrieren.

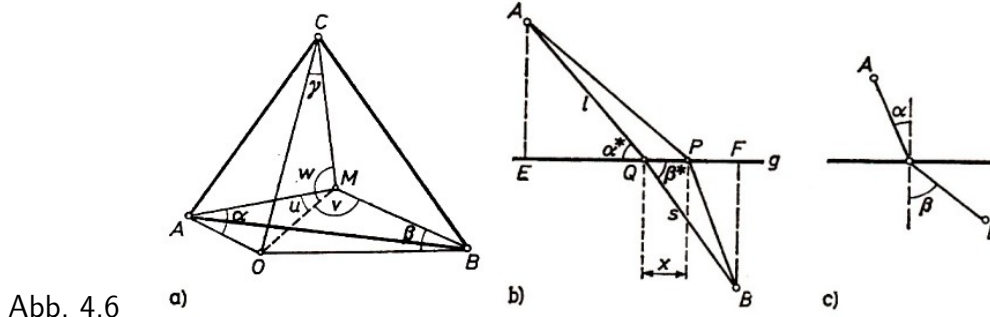


Abb. 4.6

³²Aufgabe 19 10 36 der OJM

³³Das ist unseres Wissens keine in der Literatur dafür übliche Bezeichnung.

Um die Punkte minimaler Abstandssumme von den Eckpunkten (wir wissen schon, dass es genau einen gibt) zu finden, lassen wir zwei Punkte M und O zueinander in Konkurrenz treten. Ihre Lage in der Ebene bezüglich der Punkte A , B und C beschreiben wir durch Winkel, wie das in Abb. 4.6a angegeben ist. Ohne Beschränkung der Allgemeinheit liege O im Innern des Winkels $\angle AMB$. Mit Hilfe trigonometrischer Berechnungen erhält man leicht

$$|MA| = |OA| \cos \alpha + |MO| \cos u \quad (1)$$

$$|MB| = |OB| \cos \beta + |MO| \cos v \quad (2)$$

$$|MC| = |OC| \cos \gamma + |MO| \cos w \quad (3)$$

Addiert man die Gleichungen (1), (2) und (3), so ergibt sich

$$\begin{aligned} |MA| + |MB| + |MC| &= |OA| \cos \alpha + |OB| \cos \beta + |OC| \cos \gamma \\ &\quad + |MO|(\cos u + \cos v + \cos w). \end{aligned} \quad (4)$$

Wir suchen nun für den Punkt M eine ausgezeichnete Lage. Gibt es einen Punkt M so, dass für jeden Punkt O ($\neq M$) die Klammer $\cos u + \cos v + \cos w$ gleich 0 ist? Für einen solchen Punkt M würde dann nämlich aus (4)

$$|MA| + |MB| + |MC| = |OA| \cos \alpha + |OB| \cos \beta + |OC| \cos \gamma$$

folgen, und da der Kosinus eines Winkels dem Betrag nach höchstens 1 ist, würde für alle Punkte O

$$|MA| + |MB| + |MC| \leq |OA| + |OB| + |OC|$$

und, da man den Fall $\alpha = \beta = \gamma = 90^\circ$ noch ausschließen kann, die Minimalitätseigenschaft

$$|MA| + |MB| + |MC| < |OA| + |OB| + |OC| \quad (*)$$

gelten.

Wenden wir uns also der obigen Frage zu. Bei der von uns o. B. d. A. angenommenen Lage von O gilt mit $x := |\angle AMB|$ und $y := |\angle AMC|$, dass $v = x - u$ und $w = y + u$ ist, folglich

$$\cos u + \cos v + \cos w = \cos u + \cos(x - u) + \cos(y + u) \quad (5)$$

Mit Additionstheoremen ergibt sich aus (5)

$$\begin{aligned} \cos u + \cos v + \cos w &= \cos u + \cos x \cos u + \sin x \sin u + \cos y \cos u + \sin y \sin u \\ \cos u + \cos v + \cos w &= \cos u(1 + \cos x + \cos y) + \sin u(\sin x + \sin y) \end{aligned} \quad (6)$$

Soll nun immer $\cos u + \cos v + \cos w = 0$ gelten, so ist das nur möglich, wenn $1 + \cos x + \cos y = 0$ und $\sin x - \sin y = 0$ gilt. Denn sonst könnte man ja mit O den Wert des Winkels von u so ändern, dass die rechte Seite von (6) verschiedene Werte annimmt.

Aus $\sin x = \sin 3y$ folgt $x = y$; oder $x = 180^\circ - y$. Gleichzeitig soll aber

$$\cos x + \cos y = -1 \quad (7)$$

gelten, d.h. $x > 90^\circ$ und $y > 90^\circ$; es kann also nur $x = y$ gelten und wegen (7) $x = y = 120^\circ$. Damit existiert tatsächlich genau ein solcher Punkt M , für den (*) gilt. Seine Lage ist genau so fixiert, wie uns das schon bekannt ist.

Möchte man eine Methode mit Hilfe von Beispielen erläutern, so wird ein Beispiel sicher nicht

ausreichen. Unser zweites Beispiel kann man in verschiedenen Einkleidungen anbieten.

(4.25) In Abb. 4.6 b stelle die Gerade g die Trennlinie zwischen dem Festland und dem Wasser dar. Ein Sportler soll in möglichst kurzer Zeit von einem Startpunkt A auf dem Festland die Insel B im Wasser erreichen. Seine Geschwindigkeit auf dem Festland sei v_l , im Wasser v_s . Es gelte $v_l > v_s$.

An welchem Punkt auf g muss er vom Laufen zum Schwimmen wechseln?

Da er schneller läuft als schwimmt, wird der Sportler sicherlich nicht einen insgesamt geradlinigen Weg von A nach B planen. Aber wird nun der Punkt Q oder der Punkt P günstiger sein?

Es konkurrieren also wieder zwei Punkte miteinander, ihr Abstand sei $|x| \neq 0$. Liegt P rechts von Q , so sei x positiv, sonst negativ. Wir berechnen nun die Wege $|AP|$ und $|PB|$ unter Einbeziehung von konstanten Größen und der Größen l , s und x :

$$\begin{aligned} |AP|^2 &= |EP|^2 + |AE|^2 = (|EQ| + x)^2 + l^2 - |EQ|^2 = l^2 + 2|EQ|x + x^2 \\ &= \left(l + \frac{|EQ|x}{l} \right)^2 + x^2 \left(1 - \left(\frac{|EQ|}{l} \right)^2 \right) \end{aligned} \quad (1)$$

Entsprechend ergibt sich

$$|PB|^2 = \left(s - \frac{|QF|x}{s} \right)^2 + x^2 \left(1 - \left(\frac{|QF|}{s} \right)^2 \right) \quad (2)$$

Da $|EQ| < l$ und $|QF| < s$ für $Q \neq E, F$ gilt, folgt aus (1), (2)

$$|AP| > l + \frac{|EQ|x}{l} \quad \text{und} \quad |PB| > s - \frac{|QF|x}{l}$$

Für die zu minimierende Gesamtzeit t folgt daraus

$$t = \frac{|AP|}{v_l} + \frac{|PB|}{v_s} > \frac{l}{v_l} + \frac{s}{v_s} + \left(\frac{|EQ|}{l \cdot v_l} - \frac{|QF|}{s \cdot v_s} \right) x \quad (3)$$

Jetzt stellen wir wieder die Frage, ob sich die Lage des einen Punktes so fixieren lässt, dass die Ungleichung (3) unabhängig von der Lagebeziehung zwischen P und Q , also unabhängig von x wird. Es müsste der Punkt dass $\frac{|EQ|}{l \cdot v_l} - \frac{|QF|}{s \cdot v_s} = 0$ gilt. Ist das stets möglich?

Unsere Bedingungsgleichung ist für $Q \neq F$ äquivalent mit $v_l : v_s = \frac{|EQ|}{l} : \frac{|QF|}{s}$, d.h.

$$v_l : v_s = \cos \alpha^* : \cos \beta^* \quad (4)$$

Wird Q so gewählt, dass (4) gilt, dann folgt aus (3), dass die Gesamtzeit t auf einem Weg über $P \neq Q$ stets größer als $t_Q = \frac{l}{v_l} + \frac{s}{v_s}$ ist. \overline{AQB} ist dann der Streckenzug, der in der kürzesten Zeit zurückgelegt wird.

Für vorgegebenes v_l und v_s existiert nun genau ein Q im Inneren der Strecke \overline{EF} , so dass (4) gilt: Für den angenommenen Fall $E = Q$ gilt $\alpha^* = 90^\circ$ und damit $\cos \alpha^* = 0$, $\cos \beta^* \neq 0$.

Bewegt sich Q von E nach F , so wird α^* kleiner und β^* größer, d. h., der Zähler des Bruches $\frac{\cos \alpha^*}{\cos \beta^*}$ wächst, und der Nenner wird kleiner. Nähert sich Q dem Punkt F , so wird $\cos \beta^*$ beliebig klein, $\frac{\cos \alpha^*}{\cos \beta^*}$ ist folglich eine monoton wachsende und sich stetig ändernde Größe zwischen 0

und ∞ . Der vorgegebene Wert $\frac{v_l}{v_s}$ wird an genau einer Stelle angenommen.

Das betrachtete Problem lässt sich auch physikalisch beschreiben: Gehen wir von dem Prinzip (1.4) aus, so muss sich für die Lichtausbreitung beim Übergang von einem Medium in ein anderes ebenfalls die Brechung des Lichtstrahles ergeben.

Für das Verhältnis der Lichtgeschwindigkeiten in den beiden Medien schreiben wir das Brechungsverhältnis n , an Stelle von α^* und β^* wählen wir den Einfallswinkel α und den Brechungswinkel β (vgl. Abb. 4.6 c). Statt (4) erhalten wir das Snelliussche Brechungsgesetz)³⁴

$$\sin \alpha : \sin \beta = n$$

Der Holländer Willebrord Snell van Royen fand das Gesetz um 1618. Fermat bewies es, ausgehend von dem Prinzip (1.4), in etwa in der oben angegebenen Art.

³⁴Vgl. Grimsehl, Lehrbuch der Physik, Bd. III, S. 31.

5 Ein Ausblick auf Methoden der höheren Mathematik

5.1 Extrema und infinitesimale Methoden

Ein Genre, meine ich, hat seine eigene Nase.
 Man soll ihr folgen und nicht versuchen,
 sie mit Gewalt irgendwo hineinzustecken.
 P. Hacks

Wie bereits mehrfach erklärt und praktiziert, wollen wir elementare Vorgehensweisen in den Mittelpunkt unserer Betrachtungen stellen. In diesem Abschnitt möchten wir aber einen kurzen Einblick in Möglichkeiten geben, die die Differentialrechnung eröffnet. Mittel und Methoden aus diesem Bereich werden traditionell zur "höheren Mathematik" gezählt. Wir wollen an Hand geometrisch anschaulicher Betrachtungen eine Vorstellung von den grundlegenden Ideen vermitteln.

Bereits in Kapitel 3 haben wir gesehen, dass sich gegebenenfalls Extremalaufgaben durch Funktionen erfassen und einer Lösung zuführen lassen. In diesem Fall werden die ursprünglichen Sachverhalte und Zusammenhänge durch die Eigenschaften einer reellwertigen Funktion wiedergespiegelt. Uns interessierten speziell die Extremwerte.

Es sei nun f eine über einem Intervall $a < x < b$ definierte Funktion. Folgende Erklärung ist anschaulich sofort einsichtig (Abb. 5.1a):

(5.1) Die Funktion f besitzt im Intervall (a, b) an der Stelle x_0 ein relatives Maximum bzw. Minimum genau dann, wenn für jede beliebig kleine Zahl $\varepsilon > 0$ gilt, dass

$$f(x) < f(x_0) \quad \text{für alle} \quad x_0 - \varepsilon < x < x_0 + \varepsilon \quad \text{mit} \quad x \neq x_0$$

bzw.

$$f(x) > f(x_0) \quad \text{für alle} \quad x_0 - \varepsilon < x < x_0 + \varepsilon \quad \text{mit} \quad x \neq x_0$$

ist.

Dieser Begriff des relativen Extremums ist von dem in (3.2) eingeführten Begriff des absoluten Extremums zu unterscheiden.

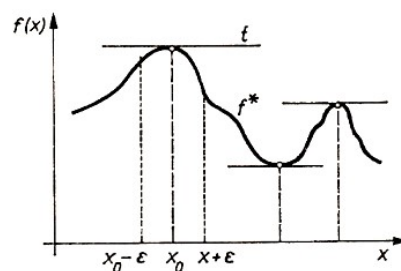


Abb. 5.1

Wie man an Hand der Abb. 5.1 erkennt, ist selbstverständlich nicht jedes relative Maximum (bzw. Minimum) ein globales Maximum (bzw. Minimum).

Für eine sehr große Klasse von Funktionen ist die Differentialrechnung ein ausgezeichnetes Instrument, die relativen Extrema zu ermitteln. Diese Formulierung lässt aber auch schon ahnen, dass diese Methode nicht auf jede Funktion anwendbar sein wird.

Das "Zünglein an der Waage" ist die Existenz einer Tangente in einem Punkt an den Graphen f^* . Wir kennen den Begriff "Tangente an einem Kreis im Punkt P_0 ", lässt sich dieser Begriff

auf andere Kurven übertragen?

Diese Frage ist so leicht nicht zu beantworten, hängt eben unlösbar mit dem Komplex zusammen, den wir gerade erst erklären wollen.

Ist t die Tangente in P_0 , an den Kreis k , so ist uns geläufig, dass sich eine Sehne g_{P_0P} oder $g_{P^*P_0}$ der Lage dieser Tangente t um so stärker nähert, je dichter P bzw. P^* bei P_0 liegen (vgl. Abb. 5.2a). Lässt sich nun für eine beliebige Kurve f^* die Tangente an f^* in P_0 auch als "Grenzlage" einer Folge von Sehnen festlegen?

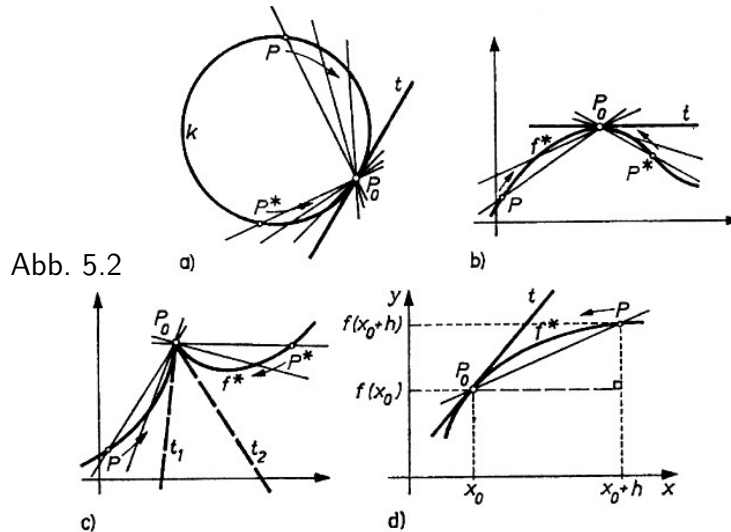


Abb. 5.2

Im Falle des Graphen einer Funktion dritten Grades (vgl. Abb. 5.2 b) scheint es so. In dem durch Abb. 5.20 gegebenen Beispiel dagegen ist wohl auf diese Art keine Tangente im Punkt P_0 eindeutig zu bestimmen.

Es ist nach diesen etwas vagen Vorbetrachtungen anzunehmen, dass für eine Klasse von Funktionen f in jedem Punkt $P_0 \in f^*$ eine Tangente an f^* existiert, solche Funktionen nennen wir differenzierbare Funktionen.

Wenn aber an f^* Tangenten existieren, dann können wir die relativen Extreme. folgendermaßen charakterisieren:

(5.2) Besitzt der Graph f^* der Funktion f im Punkt (x_0, y_0) eine waagrecht (also zur x -Achse parallel) verlaufende Tangente t und ist in jedem beliebig kleinen Intervall $(x_0 - \varepsilon, x_0 + \varepsilon)$ der Graph f eine streng konvexe (bzw. streng konkave) Kurve, so hat die Funktion f an der Stelle c_0 ein relatives Minimum (bzw. Maximum).

Nun bleibt zu klären, wie man zu einer vorgegebenen Funktion f an der Stelle x_0 die Tangente (an den Graphen f^*) - ihre 102 5. Ein Ausblick auf Methoden der höheren Mathematik Existenz vorausgesetzt - ermitteln kann. Wir betrachten neben dem zu x_0 gehörigen Punkt P_0 von f^* (also dem Punkt mit den Koordinaten x_0 und $f(x_0)$) einen benachbarten Punkt P des Graphen f^* mit den Koordinaten $x_0 + h$ und $f(x_0 + h)$ mit $h > 0$ (Abb. 5.2 d).

Die Verbindungsgerade von P_0 und P - eine Sekante von f^* - hat dann den Anstieg

$$\frac{f(x_0 + h) - f(x_0)}{(x_0 + h) - x_0} = \frac{f(x_0 + h) - f(x_0)}{h} \quad (*)$$

Mit kleiner werdendem h nähert sich der Punkt P dem Punkt P_0 , die Sekante nähert sich der Tangente t und damit der Anstieg der Sekante dem der Tangente!

Die neue Qualität dieser Art von Betrachtungen drückt sich darin aus, dass h "infinitesimal

klein wird", "gegen 0 geht". Solche Art der Betrachtungen wurde schon von Fermat (intuitiv?) benutzt und von Newton und Leibniz zur Differentialrechnung entwickelt. In der exakten Terminologie der Differentialrechnung ist zu formulieren:

Hat der Quotient (*) für jede Nullfolge³⁵, die h durchläuft, den gleichen Grenzwert, dann hat f^* an der Stelle x_0 eine Tangente, die diesen Grenzwert als Anstieg besitzt. Man sagt dann, dass die Funktion f an der Stelle x_0 differenzierbar ist, und bezeichnet den Grenzwert mit $f'(x_0)$; in kurzer mathematischer Bezeichnung ist also

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

$f'(x_0)$ heißt der Differentialquotient oder die Ableitung der Funktion f an der Stelle x_0 .

Wir betrachten dazu ein Beispiel, eine Funktion, die bereits in Kapitel 3 näher untersucht wurde, und werden sehen, dass man in einfachen Fällen auch ohne umfassende Kenntnisse aller Begriffe schon Wesentliches erfassen kann.

(5.3) Man bestimme für die Funktion f mit $f(x) = x^2$ und $D(f) = \mathbb{R}$ an der Stelle x_0 den Anstieg der Tangente, falls er existiert.

Wir bilden zunächst den Quotienten (*)

$$\frac{f(x_0 + h) - f(x_0)}{h} = \frac{(x_0 + h)^2 - x_0^2}{h} = \frac{2x_0h + h^2}{h} = 2x_0 + h$$

Wählen wir nun h beliebig klein, so nähern wir uns beliebig gut dem Wert $2x_0$. Exakt ausgedrückt:

Für jede Nullfolge von h hat $2x_0 + h$ den gleichen Grenzwert, nämlich $2x_0$; also ist

$$f'(x_0) = 2x_0$$

Damit ist gezeigt, dass der Graph f^* der Funktion f mit $f(x) = x^2$ an jeder Stelle x_0 eine Tangente besitzt und diese Tangente den Anstieg $2x_0$ hat.

(5.3') Für eine lineare Funktion f mit $f(x) = mx + n$ ergibt sich offensichtlich $f'(x_0) = m$.

(5.4) • Man zeige, dass für die Funktion f mit $f(x) = x^3$ an jeder Stelle x_0 die Ableitung $f'(x_0)$ gleich $3x_0^2$ ist.

Man kann in gleicher Weise wie zur Lösung von (5.3) noch zeigen:

(5.5) • Sind die Funktionen f und g an der Stelle x_0 differenzierbar, so auch die Funktion p und k mit $p(x) = f(x) + g(x)$ und $k(x) = a f(x)$ (mit $a \neq 0$), und es gilt³⁶

$$p'(x_0) = f'(x_0) + g'(x_0) \quad , \quad k'(x_0) = a \cdot f'(x_0)$$

Jetzt wenden wir uns der Betrachtung der relativen Extrem zu.

Verknüpfen wir unsere Erkenntnisse mit (5.2), so folgt, dass für die Existenz eines relativen

³⁵Hinsichtlich einer präzisen Fassung der Begriffe "Nullfolge" und "Grenzwert" sei auf leicht zugängliche Fachliteratur, etwa [5], verwiesen.

³⁶Weitere Regeln für das Ableiten findet der Leser bereits in allen Nachschlagebüchern, die Anfänge der Differentialrechnung enthalten (u. a. auch in Tafelwerken für Mathematik).

Minimums oder Maximums an der Stelle wo die Bedingung $f'(x_0) = 0$ (vorausgesetzt, $f'(x_0)$ existiert) notwendig ist.

Betrachten wir beispielsweise $f(x) = x^3$ an der Stelle $x_0 = 0$ (Abb. 5.3), dann erkennt man, dass diese Bedingung nicht hinreichend ist.

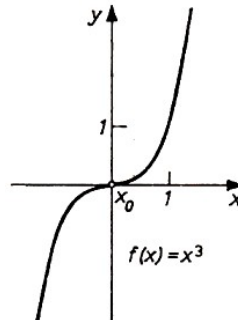


Abb. 5.3

Weiß man aber bereits, dass in einem Intervall (a, b) genau ein (relatives) Extremum existiert, und ergibt die Lösung von $f'(x_0) = 0$ genau ein $x_0 \in (a, b)$, so reicht die erste Ableitung zur Ermittlung der Extremstelle aus. Wir illustrieren dieses Verfahren einmal an der Funktion f mit

$$f(x) = ax^3 + bx^2 + cx + d$$

aus Abschnitt 3.2.

Unter Verwendung von (5.3), (5.3'), (5.4) und (5.5) ergibt sich sofort

$$f'(x_0) = 3ax_0^2 + 2bx_0 + c$$

Die quadratische Gleichung $3ax_0^2 + 2bx_0 + c = 0$ hat aber höchstens zwei reelle Lösungen. Damit hat eine ganzrationale Funktion dritten Grades höchstens zwei relative Extrema. Sind x_0 und \bar{x}_0 die reellen Lösungen obiger Gleichung, so können nur x_0 und \bar{x}_0 Extremstellen sein. Gilt $x_0 \in (a, b)$ und $\bar{x}_0 \notin (a, b)$ und ist bekannt, dass in (a, b) ein Extremum liegt, so liegt es an der Stelle x_0 . Für die Lösung von (3.10a) war ein x mit $0 < x < t$ so gesucht, dass $f_e(x) = -2kx^3 + 2kt^2x$ maximal wird. Die Existenz einer Extremalstelle war einsichtig. Es gilt

$$f'_e(x) = -6kx^2 + 2kt^2$$

$f'_e(x_0) = -6kx_0^2 + 2kt^2 = 0$ ergibt $x_0 = \frac{t}{\sqrt{3}}$ und $\bar{x}_0 = -\frac{t}{\sqrt{3}}$, aber nur $x_0 = \frac{t}{\sqrt{3}}$ liegt im betrachteten Intervall und muss folglich die Extremalstelle sein.

Auch auf die in Abschnitt 3.3 betrachtete Funktion $f(x) = x + \frac{d}{x}$ lässt sich die neue Methode gut anwenden:

Wegen (5.3') und (5.5) benötigen wir nur noch den Differentialquotienten für $g(x) = \frac{d}{x}$. Nach (*) bilden wir

$$\frac{g(x_0 + h) - g(x_0)}{h} = \frac{\frac{d}{x_0+h} - \frac{d}{x_0}}{h} = \frac{\frac{-dh}{(x_0+h)x_0}}{h} = \frac{-d}{(x_0 + h)x_0}$$

Wiederum ist die Existenz des Grenzwertes leicht einzusehen:

$$g'(x_0) = -\frac{d}{x_0^2}$$

Damit gilt $f'(x_0) = 1 - \frac{d}{x_0^2}$ und $f'(x_0) = 0$ genau dann, wenn $x_0 = \pm\sqrt{d}$ ist. Das in Abschnitt 3.3 für $x_0 > 0$ gesuchte Minimum muss also an der Stelle $x_0 = \sqrt{d}$ liegen.

Diese einfachen Beispiele lassen schon vermuten, welche außerordentlich zugkräftige Methode für die Lösung von Extremwertaufgaben mit der Differentialrechnung zur Verfügung steht. Hinzu kommt, dass bei der Einbeziehung "höherer Ableitungen" auch hinreichende Bedingungen für die Existenz der Extrema angebar sind. Außerdem lässt sich diese Methode für den Fall mehrerer Veränderlicher modifizieren.

Trotzdem sollte man keine der Methoden voreilig zur universellen Methode krönen. Wer aufmerksam tiefer eindringt, wird auch Grenzen sehen. Überdies hängt es stark von der konkreten Aufgabe ab, ob eine elementare Methode oder eine aus der "höheren Mathematik" die elegantere ist. Man löse einmal die Aufgabe (2.1) mittels der Differentialrechnung, was durchaus möglich ist!

Jedes Genre hat eben seine eigene Nase, man muss sie nicht überall hineinstecken!

Bedeutende Arbeiten über Extremwerte aus der Zeit vor der eigentlichen Erfindung der Differentialrechnung stammen von Fermat. In einer 1629 vorgelegten Sammlung von Abhandlungen über Maxima und Minima³⁷ und in späteren Arbeiten gibt er eine allgemeine Methode für die Bestimmung von Extremwerten an.

Da sie eine erstaunliche Ähnlichkeit mit den vorangegangenen einführenden Betrachtungen zur Differentialrechnung hat, wollen wir hier eines der Beispiele mitteilen:

Eine Strecke B ist so zu teilen, dass das Rechteck $A \cdot (B - A)$ ein Maximum wird. (In funktionaler und heutiger Beschreibung würde diese Aufgabe lauten:

Zu einer vorgegebenen Zahl $b > 0$ bestimme man das Maximum der Funktion $f(x) = x(b - x)$ (mit $0 < x < b$). Diese Aufgabe haben wir zu Beginn von Kapitel 3 - ausgehend von einer geometrischen Fragestellung - bereits behandelt.)

Fermat setzt nun für den einen Teil von B (nämlich A) den Ausdruck $A + E$ und hat für den zweiten Teil ($B - A$) dann $B - A - E$ zu setzen.

"... das aus den beiden Abschnitten gebildete Rechteck (ist) gleich $BA - A^2 + BE - 2AE - E^2$, dies ist näherungsweise gleichzusetzen dem obigen Rechteck $BA - A^2$. Nach Wegfall der gemeinsamen Glieder erhält man $BE \approx 2AE + E^2$.

Wird alles durch E dividiert, so bleibt $B \approx 2A + E$. Wird E gestrichen, so ergibt sich $B = 2A$. Also ist zur Lösung der Aufgabe B zu halbieren." ([17], S. 175.)

In Bezug zu unseren einführenden Betrachtungen über Differentialrechnung bildet also Fermat

$$f(x + h) = (x + h)(b - (x + h)) = bx - x^2 + bh - 2xh - h^2$$

setzt dies näherungsweise gleich

$$f(x) = x(b - x) = bx - x^2$$

setzt also

$$f(x + h) - f(x) = bh - 2xh - h^2$$

näherungsweise gleich 0 und dividiert durch h :

$$0 \approx \frac{f(x + h) - f(x)}{h} = b - 2x - h$$

³⁷In deutscher Sprache herausgegeben in "Ostwalds Klassiker der exakten Wissenschaften", Nr. 238 (Leipzig 1934).

und erhält daraus für $h = 0$ die Lösung $x = \frac{b}{2}$.

Sein Verfahren besteht allgemeiner gesagt darin, dass er $f(x + h)$ und $f(h)$ näherungsweise gleichsetzt und nach Beseitigung der gemeinsamen Glieder auf beiden Seiten durch h bzw. eine Potenz von h dividiert, bis auf einer Seite kein Glied mehr h als Faktor enthält. Danach werden alle Glieder gestrichen, die noch irgendwie h enthalten. Die übrigbleibende Gleichung ergibt den gesuchten x -Wert.

Es ist bemerkenswert, dass Fermat in entsprechender Weise das Tangentenproblem behandelt, also die Aufgabe, an eine Kurve in einem beliebigen Punkt die Tangente zu legen. Seine Methode reichte jedoch nur für eine gewisse Klasse von Kurven (Funktionen) aus.

5.2 Das Problem der Brachystochrone

Das Problem der Brachystochrone hat Johann Bernoulli 1696 gestellt, und es besteht in folgender Aufgabe:

(5.6) Ein Massepunkt gleitet ohne Reibung längs einer Kurve - allein auf Grund der Schwerkraft - von einem Punkt A zu einem tiefer gelegenen Punkt B (Abb. 5.4a). Für welche Kurve, die A und B verbindet, wird die Laufzeit am kürzesten?

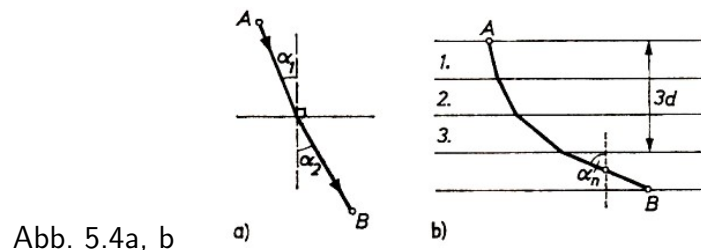


Abb. 5.4a, b

Dies ist offensichtlich auch eine Extremwertaufgabe. Aber sie ist im Vergleich zu denen, die man mit Hilfe der Differentialrechnung lösen kann, von anderer Art.

Während vorher die Größe, deren Maximum oder Minimum man bestimmen wollte, von einer (oder mehreren) Variablen abhing, sich also durch eine Funktion darstellen ließ, hängt hier die betrachtete Größe, die Laufzeit, von dem ganzen Verlauf der Kurven ab. Dagegen ist das bereits angesprochene isoperimetrische Problem (vgl. Abschnitt 2.3) von ähnlicher Art, denn es sucht z. B. unter allen einfach geschlossenen, ebenen (und konvexen) Kurven mit gleicher Länge nach denjenigen, die die größte Fläche umfassen.

Man sieht ein, dass das gleitende Teilchen auf verschiedenen Bahnkurven durchaus verschiedene Zeiten benötigen wird. Doch welche leistet das Gewünschte. Diese Aufgabe war eine Herausforderung zur damaligen Zeit (und von J. Bernoulli, der eine Lösung gefunden hatte, wohl auch so angelegt).

Im folgenden geben wir den Grundgedanken einer originellen Lösung von Jakob Bernoulli³⁸ wieder.

Aus der Mechanik ist bekannt, dass die Geschwindigkeit v eines Massepunktes, der von A aus längs irgendeiner Kurve nach unten fällt, proportional zu \sqrt{h} ist, wobei h die Fallhöhe angibt.³⁹ Bei geeigneter Wahl der physikalischen Einheiten können wir $v = \sqrt{h}$ setzen.

³⁸Jakob Bernoulli (1655-1705), Schweizer Mathematiker und Physiker, wirkte in Basel, ältester Bruder von Johann Bernoulli.

³⁹Für die Fallgeschwindigkeit gilt bekanntlich $v = g \cdot t = \sqrt{2g \cdot h}$, wobei g die Erdbeschleunigung ist.

Jakob Bernoulli dachte sich den Raum zwischen A und B in viele dünne horizontale Schichten mit der Dicke d zerlegt (Abb. 5.4b).

Es wird nun angenommen, dass sich die Geschwindigkeit des Massepunktes nicht kontinuierlich, sondern nur an den Schichtgrenzen sprunghaft ändert; in der ersten Schicht sei die Geschwindigkeit $v_1 = \sqrt{d}$, in der zweiten Schicht dann $v_2 = \sqrt{2d}$ usw.

In jeder Schicht selbst verläuft die Bahn des Teilchens geradlinig.

Die gesamte Bahnkurve von A nach B ist dann ein Streckenzug. Das Originelle an Jakob Bernoullis Lösung bestand darin, diese Bahn mit dem Weg zu vergleichen, den das Licht von A nach B nehmen würde, wenn es sich auf Grund verschiedener Brechungsindizes in den einzelnen homogenen Schichten mit unterschiedlicher Geschwindigkeit bewegen würde. Das Licht bewegt sich nach dem Snelliusschen Brechungsgesetz, wie wir das in Abschnitt 4.4 abgeleitet haben. Es gilt z. B.

$$\sin \alpha_1 : v_1 = \sin \alpha_2 : v_2$$

wobei v_1 und v_2 die Lichtgeschwindigkeiten in den Schichten 1 und 2 sind. Allgemein gilt

$$\frac{\sin \alpha_1}{\sqrt{d}} = \frac{\sin \alpha_2}{\sqrt{2d}} = \dots = \frac{\sin \alpha_n}{\sqrt{nd}} = \dots \quad (*)$$

wobei α_n der Winkel zwischen der Strecke (des Polygons) in der n -ten Schicht und der Senkrechten zu der Schichtgrenze ist. ("Einfallswinkel" beim Übergang von der n -ten zur $(n+1)$ -ten Schicht).

Für die weiteren Konsequenzen wird von der Vorstellung ausgegangen, dass sich, wenn die Schichtdicke d immer kleiner und die Anzahl der Schichten folglich immer größer gewählt wird, die Lösung des angenäherten Problems der Lösung des ursprünglichen Problems beliebig nahekommmt.

Da sich bei dieser Grenzwertbetrachtung an der Gültigkeit von (*) nichts ändert, hat Bernoulli auf folgende Eigenschaft der Lösungskurve f^* geschlossen:

Ist P irgendein Punkt der Kurve f^* , h sein Höhenunterschied zu A und α der Winkel zwischen der Tangente an f^* durch P und der Vertikalen, so ist $\frac{\sin \alpha}{\sqrt{h}}$ eine Konstante.

Eine Kurve mit einer derartigen Eigenschaft war den Mathematikern damals bereits bekannt; sie ist eine Zykloide.

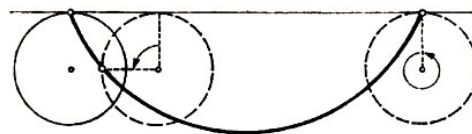


Abb. 5.4c c)

Man kann sie recht einfach mechanisch erzeugen und entsprechend geometrisch beschreiben. Lässt man einen Kreis auf einer Geraden (ohne zu gleiten) abrollen, so beschreibt ein Punkt auf dem Kreisrand eine solche Kurve (Abb. 5.4 c); sie wird deshalb treffend auch Rollkurve genannt.

Das Ergebnis war damals überraschend, denn eine solche Kurve stand auf Grund ihres Zusammenhangs mit mechanischen Problemen im Mittelpunkt des Interesses, insbesondere spielte sie bei der Konstruktion eines idealen Pendels eine große Rolle (Huygens hatte entdeckt, dass ein Massepunkt, der unter der Schwerkraft ohne Reibung längs einer Zykloide schwingt, eine von der Amplitude dieser Bewegung unabhängige Schwingungsdauer besitzt.)

Bernoullis Überlegung - wie schon die von Fermat eingangs vorgestellte - ist aus heutiger Sicht keineswegs streng.

Sie enthält eine Reihe von Annahmen, die zu rechtfertigen, zu begründen sind; so z. B., dass überhaupt eine Lösungskurve f^* existiert und dass sich das Polygon der wirklichen Lösung nähern muss. Doch derartige Überlegungen sind einleuchtend, anregend und mathematisch sinnvoll, und gerade das waren und sind Impulse für die Weiterentwicklung der Mathematik.

Die Mathematiker Euler⁴⁰ und Lagrange⁴¹ entwickelten eine allgemeine Methode für die Lösung von Extremalaufgaben, bei denen die Veränderliche nicht eine einzelne numerische Variable, sondern eine ganze Funktion (oder ein System von Funktionen) ist.

Diese neue Methode wird Variationsrechnung genannt. Wesentlichen Anteil beim Auf- und Ausbau dieser Disziplin haben Weierstrass (1815-1897), Ostrogradski (1801-1861) und Caratheodory (1873-1950).

Diese Theorie liefert stets zum Ziele führende Verfahren. Näheres würde bei weitem den angelegten Rahmen dieses Büchleins sprengen.

⁴⁰Leonhard Euler (1707-1783), Schweizer Mathematiker und Physiker, wirkte in Berlin und St. Petersburg (heute Leningrad).

⁴¹Louis Joseph Lagrange (1736-1813), französischer Mathematiker, wirkte in Turin, Berlin und Paris.

6 Ein Einblick in Optimierungsprobleme

Vielleicht lässt sich eine weitläufige Materie durch nichts kürzer erschöpfen als durch Beispiele.
W. L. Wekherlin

6.1 Die neue Aufgabenstellung

Optimierung, Optimum, optimal sind wieder neue Wörter, bringen sie aber auch einen neuen Inhalt? Sehen wir in Wörterbüchern nach, so finden wir für Optimum etwa "das Beste", "das Günstigste". Steckt dahinter - mathematisch gesehen - nicht doch wieder ein Maximum oder Minimum?

Das schon, und die Trennung zwischen Extremalproblemen und Optimierungsaufgaben kann nicht scharf vollzogen werden. Allerdings unterscheiden sich die für die Optimierung typischen Aufgabenstellungen, die ja auch die Entwicklung dieses Teilgebietes der Mathematik wesentlich auslösten, deutlich von unseren bisher betrachteten Extremalproblemen.

Wir wollen das an einem Beispiel zeigen.

Eines der ersten Probleme, welches in der bisher kurzen Geschichte der mathematischen Optimierung gelöst wurde, ist das von Hitchcock⁴² (1941) und Kantorovic (1942) unabhängig voneinander formulierte Transportproblem:

(6.1) n Lieferanten L_i mit der Lieferkapazität $a_i \geq 0$ ($i = 1, \dots, n$) liefern an k Verbraucher V_j mit dem Bedarf $b_j \geq 0$ ($j = 1, \dots, k$) ein und dasselbe Produkt P .

Der Transport einer Einheit des Produktes P vom Lieferanten L_i zum Verbraucher V_j kostet c_{ij} Geldeinheiten. Die Mengen x_{ij} des Produktes P , die von L_i nach V_j transportiert werden, sind so zu ermitteln, dass die Gesamtkosten für alle Transporte möglichst gering werden.

Die Lieferanten könnten z. B. Braunkohletagebaue sein, die Verbraucher Kraftwerke, wir haben die Situation in Abb. 6.1 angedeutet.

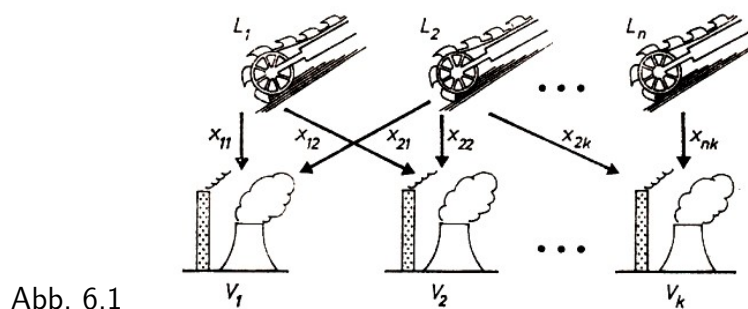


Abb. 6.1

Dieser deutliche Bezug zu einem praktischen Problem, zu ökonomischen Interessen ist ein wesentliches Charakteristikum für Optimierungsaufgaben. Die mathematische Optimierung ist ein sehr schönes Beispiel dafür, wie ganz konkrete Bedürfnisse der Praxis die Mathematiker herausfordern, neue Methoden zu entwickeln.

Bei dem Transportproblem wird das Minimum für die Kosten gesucht, das Vorgehen ist also das nach dem Prinzip des geringsten Aufwandes, bei anderen Optimierungsaufgaben, die auf dem Prinzip des größten Effektes (der größten Produktivität) beruhen, ist das Maximum einer

⁴²Frank Lauren Hitchcock (1875-1957), amerikanischer Mathematiker.

Größe gesucht. Bei geeigneten Festlegungen sind die nach dem einen oder anderen Prinzip formulierten Aufgabenstellungen wiederum zueinander dual (vgl. Abschnitt 2.3).

Wir wollen nun die Aufgabe (6.1) mit mathematischen Mitteln beschreiben:

Die Gesamtkosten y ergeben sich als Summe aus allen Einzelkosten

$$c_{ij}x_{ij} \quad (i = 1, \dots, n; j = 1, \dots, k) \quad \text{d. h.} \quad y = \sum_{i=1}^n \sum_{j=1}^k c_{ij}x_{ij}$$

Dabei gilt $x_{ij} \geq 0$, die Nichtnegativitätsbedingung. Die Funktion $y = f(x_{ij})$ nennen wir die Zielfunktion.

Im Zusammenhang mit der Kopplung an die Praxisbedürfnisse ist einsichtig, dass die Anzahl der Variablen, von denen der Funktionswert y abhängt, sehr groß ist, dies ist ein weiteres Charakteristikum. Wir haben zwar auch in unseren vorangegangenen Kapiteln den Fall mehrerer Variablen betrachtet (z. B. in (4.5) oder (4.14)), dort aber erwies sich dieser Fall nicht gerade als der typische.

Ein weiteres Charakteristikum sind die Nebenbedingungen. Auch Extremalaufgaben enthalten Nebenbedingungen, und zwar als typischen Bestandteil!

Man denke nur an die Ermittlung des minimalen Umfangs bei vorgegebenem Inhalt oder die Einbeschreibung einer Figur mit extremalen Eigenschaften in eine andere vorgegebene Figur. Nebenbedingungen in Form von Gleichungen dienen häufig dazu, die ursprüngliche Anzahl der Variablen zu reduzieren. Wie lassen sich die Nebenbedingungen für das Transportproblem (6.1) formulieren?

Jeder Lieferant L_i kann nicht mehr von dem Produkt P abgeben, als seine Lieferkapazität a_i angibt. Ein Lieferant L_i liefert an die k Verbraucher V_j insgesamt die Menge $\sum_{j=1}^k x_{ij}$. Es muss also berücksichtigt werden, dass für jedes i ($i = 1, \dots, n$)

$$\sum_{j=1}^k x_{ij} \leq a_i$$

gilt; das ist ein System von n Ungleichungen!

Aus dem Anspruch der Verbraucher V_j , die ihren Bedarf b_j exakt erfüllt sehen möchten, ergibt sich das System der k Gleichungen

$$\sum_{i=1}^n x_{ij} = b_j$$

Die Nebenbedingungen haben also die Struktur vieler Gleichungen bzw. Ungleichungen, ein für die meisten Optimierungsaufgaben typisches Erscheinungsbild.

Alle diese Besonderheiten erforderten auch besondere Lösungsmethoden. In den nächsten beiden Abschnitten wollen wir eine erste Vorstellung von solchen Methoden vermitteln.

6.2 Der lineare Fall ist der einfachste

Bisher sind für die Lösung einiger Aufgabentypen Verfahren entwickelt worden. Wegen der im allgemeinen sehr großen Anzahl von Variablen ist die Anwendung dieser Verfahren an die moderne Rechentechnik gebunden. Wir wollen aber nun hier nicht solche Verfahren vorstellen,

geschweige denn anwenden, sondern nur grundlegende Ideen erläutern, auf denen diese Verfahren beruhen. Merkwürdigerweise eignen sich dafür Beispielaufgaben, die nur wenige Variable benutzen.

(6.2) Innerhalb einer festgelegten Zeit t stellt ein Betrieb die beiden Produkte P_1 und P_2 aus den Grundstoffen M und M^* her. Für die Herstellung des Produktes P_1 wird außerdem eine Spezialmaschine benötigt, die in der Zeit t höchstens acht Einheiten des Produktes erzeugen kann.

Die weiteren Produktionsbedingungen werden durch die folgende Tabelle und deren Erläuterung beschrieben:

Produkt	prod. Einheiten	Preis	Grundstoff	
			M	M^*
P_1	x_1	1,2	2,5	3
P_2	x_2	1	3,5	2
	vorhandene Menge		35	30

Der Preis ist dabei in Geldeinheiten pro produzierter Einheit angegeben. Die Spalten "Grundstoff" geben an, wieviel eines jeden Grundstoffes (in ein und derselben Mengeneinheit) für die Erzeugung einer Einheit der Produkte benötigt werden und wieviel von jedem Grundstoff zur Verfügung steht.

Wie viele Einheiten sind von jedem der Produkte zu erzeugen, damit der beim Verkauf aller Produkte erzielte Betrag am größten wird?

Zur Lösung dieser Aufgabe suchen wir zunächst ein diesem Sachverhalt äquivalentes mathematisches Modell. Der beim Verkauf erzielte Betrag sei y . Dann gilt:

$$y = f(x_1, x_2) = 1,2x_1 + x_2 \quad (1)$$

ist die Zielfunktion, und (x_1, x_2) ist so zu ermitteln, dass y maximal wird. Die Nebenbedingungen lauten wegen der in der Aufgabenstellung vorgegebenen Beschränkungen

$$2,5x_1 + 3,5x_2 \leq 35 \quad (2)$$

$$3x_1 + 2x_2 \leq 30 \quad (3)$$

$$x_1 \leq 8 \quad (4)$$

Da die produzierten Einheiten keine negativen Größen sein können, muss auch die Nichtnegativitätsbedingung

$$x_i \geq 0 \quad (i = 1, 2) \quad (5)$$

erfüllt sein. Unter allen Wertepaaren (x_1, x_2) reeller Zahlen, die die Bedingungen (2) bis (5) erfüllen, sind diejenigen gesucht, für die der Wert von y aus (1) maximal wird.

Da es sich nur um zwei Variable handelt, können wir diese Aufgabenstellung graphisch bezüglich eines x_1, x_2 -Koordinatensystems veranschaulichen.

Wir tragen zunächst die der Gleichung $2,5x_1 + 3,5x_2 = 35$ entsprechende Gerade g ein. Alle Punkte (x_1, x_2) , die mit ihren Koordinaten die Ungleichung (2) erfüllen, liegen dann in der durch g und den Ursprung bestimmten Halbebene $H_{(2)}$; entsprechend ordnen wir der Ungleichung (3) bzw. (4) die Halbebenen $H_{(3)}$ bzw. $H_{(4)}$ zu (vgl. Abb. 6.2).

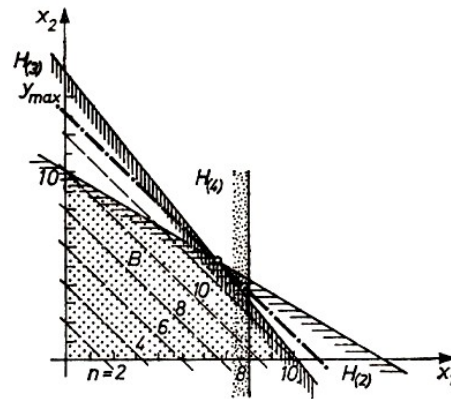


Abb. 6.2

Punkte, deren Koordinaten (2), (3) und (4) gleichzeitig erfüllen, liegen im Durchschnitt dieser drei Halbebenen. Die Ungleichung (5) besagt, dass nur die Punkte im ersten Quadranten in Betracht kommen.

In unserem Beispiel ergibt sich, dass nur solche (x_1, x_2) die Funktion (1) maximieren können, die auf dem Rand oder im Inneren eines konvexen Fünfecks liegen; diese Punktmenge B nennen wir daher Zulässigkeitsbereich.

Jetzt betrachten wir alle (x_1, x_2) , für die die Funktion (1) einen vorgegebenen Wert n annimmt, also $1, 2x_1 + x_2 = n$ gilt. Die zugehörigen Punkte liegen auf der Geraden $x_2 = -1, 2x_1 + n$. Für jedes n ergibt sich eine solche Gerade, die parallel zu jeder anderen dieser Art ist (Abb. 6.2).

Verschieben wir also eine solche Gerade in eine zu ihr parallele, so verändert sich das zugehörige n , d. h. der zu maximierende Funktionswert. Folglich verschieben wir so, dass n wächst, allerdings muss auf der betrachteten Geraden immer mindestens ein Punkt von B liegen.

In unserem Beispiel liefert die Strich-Punkt-Gerade (in Abb. 6.2) das Optimum. Sowohl der maximale Funktionswert y_{\max} als auch die zu produzierenden Einheiten x_1, x_2 lassen sich aus der graphischen Darstellung ablesen.

Natürlich lassen sich x_1 und x_2 auch als Schnittpunktskoordinaten von Geraden berechnen und daraus schließlich y_{\max} . Es ergibt sich $x_1 \approx 6,36$, $x_2 \approx 5,45$, $y_{\max} \approx 13,09$. Sowohl die graphische Darstellung als auch die rechnerische Kontrolle ergibt, dass in diesem Fall die Realisierung der optimalen Lösung die vorhandene Menge beider Grundstoffe aufbraucht, die Spezialmaschine aber nicht voll auslastet.

Was kann man aus diesem sehr speziellen Beispiel für den allgemeinen Fall an Erkenntnissen gewinnen?

Jede der Nebenbedingungen und auch die Nichtnegativitätsbedingung wird durch eine lineare Ungleichung beschrieben und daher in der Veranschaulichung durch eine Halbebene wiedergegeben, für den Fall von drei Variablen x_1, x_2, x_3 würde einer solchen Ungleichung ein Halbraum entsprechen.

Ist der Durchschnitt aller Halbebenen (bzw. Halbräume) eine nichtleere und beschränkte Menge, so lässt sich zeigen, dass diese Menge - der Zulässigkeitsbereich - ein konvexes n -Eck (bzw. konvexes Polyeder) ist.

Ist die Zielfunktion linear, so wird sie durch eine Geradenschar (bzw. Ebenenschar) charakterisiert. Aus dieser Schar wird nun diejenige Gerade (bzw. Ebene) ausgewählt, für die $y = f(x_1, x_2)$ (bzw. $y = f(x_1, x_2, x_3)$) maximal wird.

Diese Gerade (bzw. Ebene) hat aber wegen der Konvexität des Zulässigkeitsbereiches B mit

B nur Randpunkte gemeinsam. Daraus folgt, dass man bei Kenntnis aller Eckpunkte von B die optimale Lösung ermitteln kann, denn die Anzahl der Eckpunkte ist endlich!

Das alles ist keine exakte Herleitung, aber bei einer solchen geometrischen Veranschaulichung gut "einsichtig". Auch für mehr Variable, also höhere Dimensionen, ergibt sich aus der Linearität der Nebenbedingungen die Konvexität von B und damit aus der Linearität der Zielfunktion die Erkenntnis, dass die Lösung unter den endlich vielen Eckpunkten von B zu suchen ist. (Das muss übrigens nicht bedeuten, dass die Lösung durch genau einen Eckpunkt repräsentiert wird. Man vergleiche dazu die Aufgabe 6.3b) !)

Das alles lässt sich zwar nicht mehr n -dimensional vorstellen, aber exakt herleiten (vgl. etwa [4]). Auf diesen grundsätzlichen Überlegungen basieren auch Verfahren zur Berechnung der Lösung. Am bekanntesten ist das Simplexverfahren.

Bei diesem Verfahren beginnt man mit einem Startelement, welches einem Eckpunkt von B entspricht. Ein - rein rechnerisch durchzuführender - Simplexschritt erzeugt aus diesem Startelement ein neues, welches wieder einem Eckpunkt von B entspricht.

Bei einem solchen Schritt wächst aber, solange die optimale Lösung noch nicht erreicht ist, auch der Wert der Zielfunktion. Mit diesem Verfahren steuert man also zielstrebig auf die Lösung zu und erreicht sie nach endlich vielen Schritten, ohne alle der zwar endlich vielen, meistens aber doch sehr vielen Eckpunkte von B berücksichtigen zu müssen.

Ist die Zielfunktion einer Optimierungsaufgabe linear und werden überdies die Nebenbedingungen durch ein System linearer (Un)gleichungen beschrieben, so sprechen wir von einer linearen Optimierungsaufgabe. Sind die Variablen dieser Aufgabe Variable über der Menge der nichtnegativen reellen Zahlen, so kann man eine solche Aufgabe prinzipiell nach dem Simplexverfahren lösen.

(6.3) • Man verändere die Aufgabe (6.2) wie folgt:

a) An Stelle der produzierten Einheiten x_1, x_2 seien Stückzahlen x_1, x_2 gesetzt, d. h., es seien nur natürliche Zahlen für x_1, x_2 zulässig ("ganzzahlige" Optimierung). Man zeige, dass es falsch ist, von der für (6.2) gefundenen Lösung auf das "nächstgelegene" ganzzahlige Paar (6, 5) als Lösung zu schließen!

b) Man löse die Aufgabe (6.2) für den Fall, dass lediglich der Preis 1,2 in 1,5 abgeändert wird!

Gibt man die "Linearität" auf, so bereiten die entstehenden Aufgaben zumeist beträchtliche Schwierigkeiten. Man ist also gut beraten, seine Schritte in das nichtlineare Gebiet behutsam zu setzen.

Bedenken wir nochmals, wie sich die Linearität auswirkte. Neben der Einfachheit der Operationen mit linearen Termen war es insbesondere die Konvexität von B und die "Lagebeziehung" zwischen B und einer die Zielfunktion charakterisierenden Kurvenschar, die Lösungsprinzipien suggerierte.

Liegen nichtlineare Nebenbedingungen vor, die aber wiederum einen konvexen Zulässigkeitsbereich erzeugen, und eine lineare Zielfunktion, so wird man vielleicht ähnlich wie in (6.2) vorgehen können.

(6.4) • Man ermittle den maximalen Funktionswert der Funktion $y = f(x_1, x_2) = a \cdot x_1 + x_2$ unter den Bedingungen $x_1^2 + x_2^2 \leq 8$, $0 \leq x_1 \leq 2,5$ und $0 \leq x_2 \leq 2,5$ für $a = 0, 1; 1; 10$.

Das folgende Beispiel wird man kaum in Lehrbüchern der Optimierung finden. Wir haben es dennoch ausgewählt. Es ist nämlich das uns bereits bekannte Beispiel (2.27), und es bietet einen sehr einfachen Fall einer nichtlinearen Zielfunktion an.

Man sieht daran vielleicht auch, wie fließend die Grenzen zwischen den Teilgebieten sind.

(6.5) Unter allen Rechtecken mit den Seitenlängen x_1 und x_2 , deren Umfang einen vorgegebenen Wert $2c$ nicht überschreitet, ist dasjenige mit dem maximalen Flächeninhalt zu ermitteln.

Zielfunktion: $y = f(x_1, x_2) = x_1 \cdot x_2$

Nichtnegativitätsbedingung: $x_1 \geq 0, x_2 \geq 0$;

Nebenbedingung: $x_1 + x_2 \leq c$

Ersichtlich ist, dass die Zielfunktion nichtlinear ist und die geringe Veränderung des Textes in (6.5) gegenüber den ursprünglichen Aufgabenstellungen (2.27) so gewählt wurde, dass sich für die Nebenbedingung wieder eine Ungleichung ergibt.

Wir gehen prinzipiell wie in Beispiel (6.2) vor. Die die Zielfunktion charakterisierende Kurvenschar ist jetzt aber eine Schar von Hyperbeln ($x_2 = \frac{n}{x_1}$), für deren Schnittpunkte S mit der Geraden $x_2 = x_1$ die Gleichung $|OS| = n \cdot \sqrt{2}$ gilt (vgl. Abb. 6.3). Da der Graph einer solchen Hyperbel (von unten) konvex ist (vgl. (4.3)), lässt sich unser obiges Vorgehen der Verschiebung wiederum anwenden.

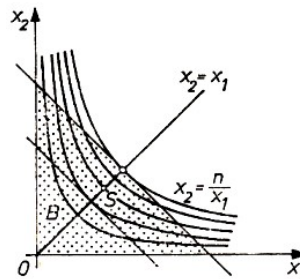


Abb. 6.3

Unter obiger Nebenbedingung wird n für diejenige Hyperbel maximal, die mit der Geraden $x_1 + x_2 = c$ genau einen Punkt gemeinsam hat. Dieser aus der graphischen Darstellung ablesbare Punkt lässt sich auch rechnerisch ermitteln:

Aus $x_1 + x_2 = c$ und $x_2 = \frac{n}{x_1}$ folgt $x_1 + \frac{n}{x_1} = c$ bzw. $x_1^2 - cx_1 + n = 0$.

Diese quadratische Gleichung hat genau dann genau eine Nullstelle, wenn $\frac{c^2}{4} - n = 0$ gilt, d. h. $n = \frac{c^2}{4}$ und damit $x_1 = \frac{c}{2}$ und weiter $x_2 = \frac{c}{2}$. Für das gesuchte Rechteck gilt also $x_1 = x_2$.

Auch die folgende Aufgabe, die der 18. Olympiade Junger Mathematiker der DDR entnommen ist, lässt sich nach den bisher vermittelten Methoden lösen.

(6.6) • Es sei M die Menge aller Tripel (x_1, x_2, x_3) von reellen Zahlen, für die die folgenden Ungleichungen erfüllt sind:

$$\begin{aligned} 55x_1 + x_3 &\leq 54, & 55x_2 + x_3 &\leq 54, & 55x_1 - 4x_3 &\geq 4 \\ 55x_2 - 4x_3 &\geq 4, & x_3 &\geq -1 \end{aligned}$$

Man untersuche, ob für die Funktion

$$y = f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$$

ein Tripel $(\bar{x}_1, \bar{x}_2, \bar{x}_3) \in M$ so existiert, dass für alle Tripel $(x_1, x_2, x_3) \in M$ die Ungleichung

$$f(\bar{x}_1, \bar{x}_2, \bar{x}_3) \geq f(x_1, x_2, x_3)$$

gilt. Ist dies der Fall, so ermittle man hierzu $f(\bar{x}_1, \bar{x}_2, \bar{x}_3)$.

6.3 Ein Beispiel der dynamischen Optimierung

Betrachten wir die Produktion in einem Betrieb als einen in der Zeit t ablaufenden Prozess, so spielte dieser Prozesscharakter allerdings in Beispiel (6.2) keine Rolle.

Bei der Beobachtung der Produktion über einen längeren Zeitraum wird dagegen der Prozesscharakter sehr deutlich. Die zu einem frühen Zeitpunkt erzielten Resultate haben wesentlichen Einfluss auf spätere Situationen. Ein Betrieb plant seine Produktion z. B. für mehrere Jahre, er legt aber nach jedem Jahr eine Zwischenbilanz vor.

Schon diese Einteilung legt es nahe, die Produktion als einen Mehrstufenprozess anzusehen und die Optimierung des gesamten Prozesses auf die der einzelnen Stufen zurückzuführen. Grundsätzlich gilt dabei, dass keine Stufe isoliert betrachtet werden darf.

Erzielt ein Betrieb z. B. im ersten Jahr einen sehr großen Gewinn dadurch, dass er alle Mittel für den Kauf von Rohstoffen einsetzt und die vorhandenen Maschinen - ohne Rücksicht auf Verschleiß - mit voller Leistung fahren lässt, so wird möglicherweise im zweiten Jahr wegen ausfallender Maschinen, die wegen fehlender Finanzen nicht erneuert werden können, die Leistung extrem niedrig sein. Vernünftig erscheint folgendes Prinzip:

Ein Prozess wird in endlich viele Stufen unterteilt, und jede Stufe wird unter Berücksichtigung der nachfolgenden Stufen optimiert. Dieses Prinzip nennen wir dynamische Optimierung. Dann existiert unter allen Stufen genau eine, die ohne Blick in die Zukunft optimiert werden darf, nämlich die letzte! Die dynamische Optimierung erfolgt daher immer entgegengesetzt dem Zeitablauf.

Um eine allererste Vorstellung zu vermitteln, betrachten wir ein Beispiel zum sogenannten Verteilungsproblem.⁴³

(6.7) Für zwei Industriezweige I und II sind die Investitionen für drei Jahre, unter folgenden Annahmen zu planen:

- Werden die Mittel x im Zweig I investiert, so bringen sie jährlich den Gewinn $f_I(x) = x^2$, und durch die Ausgaben reduzieren sich die ursprünglichen Mittel auf $0,75x$.
- Werden die Mittel x im Zweig II investiert, so bringen sie jährlich den Gewinn $f_{II}(x) = 2x^2$, und durch die Ausgaben reduzieren sich die ursprünglichen Mittel auf $0,3x$.
- Der Anfangsbestand z_0 an Mitteln ist unter diesen Bedingungen für jedes Jahr der Planperiode auf die Zweige I und II so aufzuteilen, dass der Gesamtgewinn möglichst groß wird.

Zur Lösung dieses Problems gehen wir von der letzten Stufe aus. Nach dem zweiten Jahr sei der Bestand an Mitteln z_2 . Die Mittel, die davon im dritten Jahr im Zweig I eingesetzt werden, seien x_3 , für den Zweig II stehen dann noch die Mittel $z_2 - x_3$ zur Verfügung. Es muss natürlich

$$0 \leq x_3 \leq z_2 \quad (1)$$

gelten. Für den Gesamtgewinn des dritten Jahres ergibt sich nach den Voraussetzungen

$$G_3 = x_3^2 + 2(z_2 - x_3)^2 = 3x_3^2 + 4z_2x_3 + 2z_2^2$$

d. h., G_3 ist in Abhängigkeit von x_3 über dem abgeschlossenen Intervall (1) eine Parabel mit positivem Leitkoeffizienten.

⁴³Vgl. [15], S. 61ff.

Deren Maximum liegt aber nach (3.6') auf dem Rand des Intervalls (1)! Welcher Randpunkt ergibt den größeren Wert?

Dies ist der für $x_3 = 0$! Folglich werden alle Mittel z_2 im dritten Jahr im Zweig II investiert! Der Gesamtgewinn des dritten Jahres ist damit $2z_2^2$.

Welche Mittel z_2 stehen aber nach dem zweiten Jahr zur Verfügung? Sind x_2 und $z_1 - x_2$ die im zweiten Jahr investierten Mittel, so haben diese sich auf $z_2 = 0,75x_2 + 0,3(z_1 - x_2)$ reduziert. (z_1 ist dabei der Bestand an Mitteln nach dem ersten Jahr.)

Mit Hilfe dieser Überlegungen können wir nun den Gesamtgewinn $G_{2,3}$ der letzten beiden Jahre maximieren:

$$G_{2,3} = x_2^2 + 2(z_1 - x_2)^2 + 2z_2^2$$
$$G_{2,3} = 3,505x_2^2 - 3,46z_1x_2 + 2,18z_1^2$$

Wiederum unter Bezug auf (3.6') muss man sich für $x_2 = 0$ oder $x_2 = z_1$ entscheiden. Setzt man diese beiden Möglichkeiten ein, so ergibt sich das Maximum $2,18z_1^2$ für $x_2 = 0$.

Für das erste Jahr ergibt sich analog

$$z_1 = 0,453x_1 + 0,3z_0 \quad \text{und}$$

$$G_{1,2,3} = x_1^2 + 2(z_0 - x_1)^2 + 2,18z_1^2 \quad \text{bzw.} \quad G_{1,2,3} \approx 3,44x_1^2 - 3,41z_0x_1 + 2,20z_0^2$$

Die Entscheidung nach (3.6') liefert jetzt $x_1 = z_0$. Damit haben wir die optimale Investitionsstrategie gefunden!

Im ersten Jahr müssen alle Mittel im Zweig I investiert werden, im zweiten und dritten Jahr wird nur im Zweig II investiert.

Der Gesamtgewinn beträgt $\approx 2,23z_0^2$, und nach drei Jahren steht noch ein Rest von $\approx 0,07z_0$ an Investitionsmitteln zur Verfügung.

(6.8) • Wie lautet die optimale Investitionsstrategie, wenn man bei dem Problem (6.7) für m Jahre mit $m > 3$ zu planen hat?

Mit unseren acht "Beispielen" haben wir die "weitläufige Materie" der Optimierung nun keineswegs erschöpft, aber hoffentlich einen ersten Eindruck von der "Weitläufigkeit" vermittelt.

7 Lösungen und Lösungshinweise

Es lohnt nicht, einen Fehler zu machen,
wenn keiner ihn bemerkt.
P. Tille

(1.1) a) Kleinste Zahl: -23; größte Zahl: 17 ; b) Kleinste Zahl: $\sqrt{156}$; größte Zahl: $\sqrt{158}$.

(1.2) Angenommen, es existiert eine rationale Zahl r_{\min} , so dass für alle r mit $2 < r^2 \leq 3$ auch $2 < r_{\min}^2 \leq r^2$ gilt. Zwischen den zwei rationalen Zahlen 2 und r_{\min}^2 liegen aber beliebig viele rationale Zahlen, unter diesen befinden sich auch Quadrate rationaler Zahlen, z. B.

$$r_{\min}^2 - x(2r_{\min} - x) = (r_{\min} - x)^2$$

(für hinreichend kleines positives rationales x). Das ergibt aber einen Widerspruch zur Annahme, dass r_{\min} die kleinste derartige Zahl ist.

Es gibt keine rationale Zahl r mit $r^2 = 3$. Gäbe es also eine größte rationale Zahl r_{\max} mit $r_{\max} \leq 3$, dann müsste $r_{\max} < 3$ gelten. Diese Annahme kann man analog zu obigen Ausführungen zum Widerspruch führen.

(2.6) Man verfähre analog zu (2.3). Statt R sowohl an p als auch an q zu spiegeln, spiegle man A an p und B an q . Die Bilder seien A' und B' , und S sei wiederum der Scheitel des Winkels $\angle p, q$.

Dann gibt es nach den gleichen Überlegungen wie zu (2.3) (genau) eine Lösung $P_0 \in p$, $Q_0 \in q$, wenn (und nur wenn) die Größe des Winkels $\angle A'SB'$ kleiner als 180° und p, q im Innern von $\angle : A'SB'$ liegen. Damit die Aufgabe also stets lösbar ist, muss $|\angle p, q| = 60^\circ$ sein (vgl. dazu Abb. 7.1).

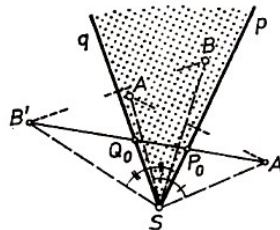


Abb. 7.1

Von den Streckenzügen \overline{APQB} und $\overline{AQP'B}$ ist im allgemeinen - falls sie beide existieren - einer kürzer als der andere. Dann ist der Streckenzug minimaler Länge eindeutig bestimmt. Gilt $B \in g_{SA}$, so sind dagegen diese Streckenzüge von gleicher Länge; die Aufgabe hat dann für spitze Winkel $\angle p, q$ genau zwei Lösungen, für nichtspitze Winkel $\angle p, q$ keine Lösung.

(2.9) Es sei $PQRS$ ein dem Quadrat $ABCD$ entsprechend Abb. 7.2 einbeschriebenes Quadrat. Dann geht bei der Drehung um den Mittelpunkt M des Quadrats $ABCD$ mit 90° das Dreieck SAP in das Dreieck PBQ über, und damit ist PMQ ein gleichschenkliges Dreieck mit rechtem Winkel bei M .

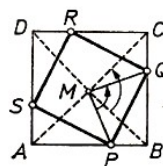


Abb. 7.2

Nun ist offenbar der Umfang des Quadrats $PQRS$ genau dann minimal, wenn $|PQ|$ minimal und damit nach dem Vorangegangenen die Länge $|MP|$ minimal ist. Unter allen Punkten P aus \overline{AB} leistet dies der Mittelpunkt dieser Seite. (Entsprechende Lagen haben dann Q , R und S .)

Eine weitere Lösung ergibt die bisherige Methode mit Geradenspiegelungen.

(2.10) Man spiegle O , O' , O'' und O''' an den durch die Rechtecksseiten bestimmten Geraden entsprechend Abb. 7.3 und verbinde zunächst O^{IV} mit O . Wegen der Längeninvarianz der Spiegelung und der Gültigkeit des Reflexionsgesetzes gilt:

$$\begin{aligned} |OP| &= |OP'|, & |OP| + |PQ| &= |QO''| \\ |OP| + |PQ| + |QR| &= |RO''|, & |OP| + |PQ| + |QR| + |RS| &= |SO^{IV}| \end{aligned}$$

S muss auf OO^{IV} liegen. Überdies ist dann $g_{SO} \parallel g_{OP}$ und damit $O \in \overline{SP}$. Dies führt zur Lösung von (2.10'). (Auf eine Determination gehen wir hier nicht ein.)

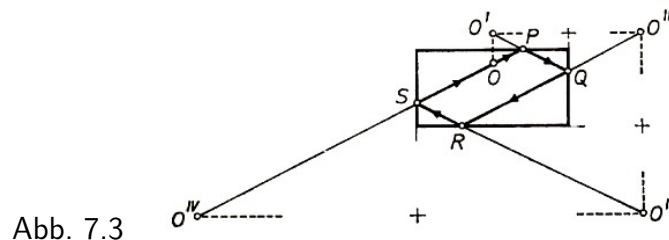


Abb. 7.3

(2.13) a) Man überzeuge sich davon, dass die für (2.12) vorgestellte Lösung auch auf einen beliebigen Punkt P der Ebene als Ausgangspunkt anwendbar ist, aber wieder einen Punkt im Inneren als Lösung der Aufgabe ergibt. Die Lösung ändert sich also nicht.

b) Die Lösung zu (2.12) lässt erkennen, dass sie auch für stumpfwinklige Dreiecke gilt, deren stumpfe Winkel kleiner als 120° sind.

Es sei o. B. d. A. der Winkel bei C mindestens 120° . Dann gilt für alle Punkte $P \neq C$ die Relation $|AP| + |BP| + |CP| > |AC| + |BC|$.

(2.14) Für den Schnittpunkt der Diagonalen wird die Abstandssumme minimal! (Beweis indirekt unter Verwendung der Dreiecksungleichung.)

(2.15) Wir können durch eine Drehung wie zur Lösung der Aufgabe (2.12) (vgl. Abb. 2.6) zu einer vorteilhaften Einsicht kommen.

Da ABC ein gleichseitiges Dreieck ist, gibt es eine Drehung ρ um A mit 60° , bei der $C^\rho = B$ ist. Es sei P ein von A und B verschiedener Punkt. Wegen $|P^\rho P| = |AP|$, $|P^\rho B| = |PC|$ und der Dreiecksungleichung bezüglich P^ρ , P , B gilt $|AP| + |BP| \geq |CP|$. Für $P = A, B$ gilt speziell $|AP| + |BP| = |CP|$.

Also ist das Minimum von $s = |AP| + |BP| - |CP|$ gleich 0.

Gilt⁴⁴ $|AP| + |BP| = |CP|$ und $P \neq A, B$, so müssen - wie man sich näher überlegt - P und C auf verschiedenen Seiten bezüglich g_{AB} liegen, und es muss $P \in \overline{P^\rho B}$ sein. Dann ist $|\angle : APB| = 180^\circ - 60^\circ = 120^\circ$, und demnach liegt P auf dem Umbogen des Dreiecks ABC zwischen A und B .

Umgekehrt ist $s = 0$ für alle Punkte dieses Kreisbogens (einschließlich A und B).

(2.21) Man kann zunächst einmal der Lösung von (2.20) folgen.

⁴⁴Von hier an vgl. mit der Aufgabe 17 12 24 der OJM.

Wir betrachten zuerst alle Wege, die von A nach G über einen Punkt der Kante \overline{BC} verlaufen. Der kürzeste dieser Wege hat die Länge $\sqrt{(a+c)^2 + b^2}$ (vgl. Abb. 2.10a, b).

Der kürzeste aller Wege von A nach G über einen Punkt der Kante \overline{BF} hat die Länge $\sqrt{(a+b)^2 + c^2}$ und der über einen Punkt der Kante \overline{EF} die Länge $\sqrt{(b+c)^2 + a^2}$. Diese Längen lassen sich leicht vergleichen, da sich die Radikanden nur um ein Glied $2xy$ mit $x, y \in \{a, b, c\}$ unterscheiden.

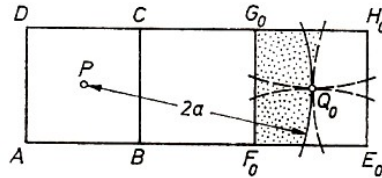


Abb. 7.4

Wegen $a < b < c$ ist demnach die Länge $\sqrt{(a+b)^2 + c^2}$ die kürzeste. Damit gibt es genau zwei Streckenzüge als kürzeste Wege, nämlich \overline{APG} und \overline{AQG} , mit $P \in \overline{BF}$ und $|BP| : |PF| = a : b$ bzw. $Q \in \overline{DH}$ und $|DQ| : |DH| = a : b$.

(2.22) Unter dem Abstand zweier Punkte P und Q auf der Würfeloberfläche ist das Minimum der Weglängen von P nach Q zu verstehen; dies wird hier durch einen Streckenzug realisiert. Eine Verebnung der Würfeloberfläche führt auch hier zur Lösung (vgl. Abb. 7.4).

a) Der Mittelpunkt Q der P gegenüberliegenden Würfelfläche $EFGH$ hat von P offensichtlich den Abstand $2a$, wobei er die Kantenlänge des Würfels ist. Die vier Streckenzüge über die Mittelpunkte der Seitenflächen $ABFE$ bzw. $BCGF$ bzw. $CDHG$ bzw. $DAFH$ sind die kürzesten Wege von P nach Q .

Die Verebnung längs der Kanten \overline{BC} und \overline{FG} (Abb. 7.4) zeigt klar, dass die vier diesbezüglichen "Abstandskreis"-Teile auf der Fläche $EFGH$ mit dem Radius $2a$ diese Seitenfläche überdecken, dass es von P aus also keinen größeren Abstand gibt und alle von Q verschiedenen Punkte einen kleineren Abstand von P besitzen.

b) In entsprechender Weise erkennt man, dass der Mittelpunkt Q von \overline{GH} der von P am weitesten entfernte ist.

(2.24) Die Überlegungen zu (2.23) können entsprechend für den Quader mit ungleich langen Kantenlängen übernommen werden. Stabile Lagen des Gummifadens verlaufen auch hier parallel zu den Quaderkanten oder parallel zu den Diagonalen der Seitenflächen.

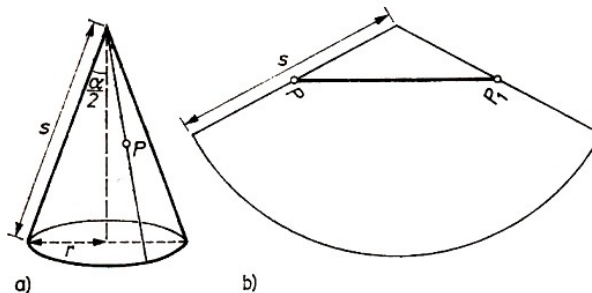


Abb. 7.5

(2.25) Wir schlitzen den Kegelmantel längs der Mantellinie durch P auf und können dann verebnen (Abb. 7.5 b). Jeder Weg von P nach P ist dann nach der Verebnung ein gleichlanger Weg in einem Kreissektor von P nach P_1 (Abb. 7.5 b). Der kürzeste Weg in der Verebnung ist die Verbindungsstrecke $\overline{PP_1}$. Da der Weg nicht über die Spitze S des Kegels geführt werden darf, existiert das Minimum genau dann, wenn $|\angle PSP_1| < 180^\circ$ ist.

Dies ist äquivalent mit $2\pi r < \pi s$, also mit $\sin \frac{\alpha}{2} = \frac{r}{s} < \frac{1}{2}$, d. h. mit $\alpha < 60^\circ$, wobei α der Öffnungswinkel des Kegels ist, r der Radius des Grundkreises und s die Länge einer Mantellinie (Abb. 7.5).

Klebt man den Kreissektor mit der Strecke $\overline{PP_1}$ wieder zu einem Kegelmantel zusammen, so erhält man den räumlichen Verlauf des minimalen Weges. (Häufig ist man von diesem Ergebnis überrascht!)

(2.26) Man "spiegelt" Q am Äquator, d. h., man wählt denjenigen Punkt Q' auf der südlichen Halbkugel, der mit Q auf dem gleichen Meridian liegt und den gleichen Abstand zum Äquator hat.

Ist Q' nicht diametral gegenüberliegend zu P , dann gibt es genau einen Großkreis durch P und Q' , und einer der Bögen von P nach Q' ist der kürzere. Dessen Schnittpunkt R mit dem Äquator ergibt die Lösung. - Sind dagegen P und Q' diametral gegenüberliegend, so leistet nun offenbar jeder Punkt R des Äquators das Gewünschte.

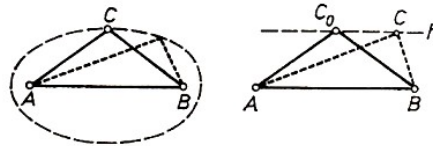


Abb. 7.6 a)

b)

(2.28) Zur Lösung benutzen wir elementare Eigenschaften der Ellipse, auf die wir bereits im Zusammenhang mit der Lösung zur Aufgabe (2.1) zu sprechen kamen. Auf Grund der Voraussetzungen ist die Summe der Seitenlänge $a = |BC|$ und $b = |AC|$ konstant, und demnach liegt C auf einer Ellipse mit den Brennpunkten A und B .

Den größten Abstand von der Symmetrieachse g_{AB} haben diejenigen Ellipsenpunkte, die auf der Mittelsenkrechten von \overline{AB} liegen (Abb. 7.6a). Wegen $F_\Delta = \frac{c \cdot hc}{2}$ und konstantem C liegt hier ein Dreieck mit dem größten Flächeninhalt vor. Andere Methoden zur Lösung stehen später zur Verfügung, beispielsweise:

Nach Voraussetzung gibt es ein gleichschenkliges Dreieck ABC_0 , mit $|AB| = c$ und $|AC_0| = |BC_0| = \frac{a-c}{2}$. Für jeden Punkt C , der bezüglich g_{AB} auf der gleichen Seite wie C_0 liegt, gilt $A(ABC) = A(ABC_0)$ genau dann, wenn C auf der Parallelen h zu g_{AB} durch C_0 liegt.

Nach der Lösung zu (2.1) ist für jeden Punkt $C \in h$, der von C_0 verschieden ist, der Streckenzug \overline{ACB} länger als der Streckenzug $\overline{AC_0B}$ (Abb. 7.6 b), d. h., unter allen Dreiecken mit gleichem Flächeninhalt und gleicher Seite \overline{AB} ist das gleichschenklige Dreieck dasjenige mit dem kleinsten Umfang. Mit Hilfe des Satzes (2.32) folgt daraus die Behauptung.

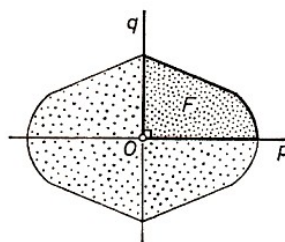


Abb. 7.7

(2.31) Da $\angle p, q$ ein rechter Winkel ist, lässt sich die durch Schnitt der Winkelfläche gegebene Figur F an den Trägergeraden der Schenkel p und q so spiegeln, dass eine bezüglich des Scheitels O von $\angle p, q$ zentralsymmetrische Figur F_1 entsteht (Abb. 7.7).

Diese Figur hat einen konstanten Umfang, nämlich das Vierfache der Länge der Schnittkurve, und wir können in Hinblick auf die Aufgabenstellung o. B. d. A. davon ausgehen, dass F_1 konvex ist.

Nach dem Satz (2.30) hat F_1 und damit F maximalen Flächeninhalt, wenn F_1 eine Kreisfläche, also wenn der Schnitt um O kreisförmig ist.

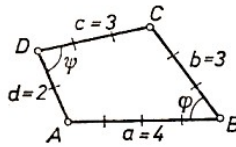


Abb. 7.8

(2.31') φ sei der von den Seiten mit den Längen $a = 4$ und $b = 3$ und ψ der von den Seiten mit den Längen $c = 3$ und $d = 2$ eingeschlossene Winkel (Abb. 7.8). A sei der Flächeninhalt des Vierecks.

Dann ergibt der Flächensatz $2A = 12 \sin \varphi + 6 \sin \psi$. Berechnet man die Diagonalenlänge $|AC|$ mittels a, b und φ und andererseits mittels c, d und ψ nach dem Kosinussatz, so erhält man $25 - 24 \cos \varphi = 13 - 12 \cos \psi$. Nach elementaren Vereinfachungen und dem Quadrieren erhält man

$$\frac{A^2}{9} = 4 \sin^2 \varphi + 4 \sin \varphi \sin \psi + \sin^2 \psi$$

und

$$1 = 4 \cos^2 \varphi - 4 \cos \varphi \cos \psi + \cos^2 \psi$$

Die Addition dieser beiden Gleichungen ergibt bei Verwendung von Formeln der Trigonometrie

$$\frac{A^2}{9} = 4(1 - \cos(\varphi + \psi))$$

Damit wird $\frac{A^2}{9}$ maximal für $\cos(\varphi + \psi) = -1$, d. h. für $\varphi + \psi = 180^\circ$, und es gilt $A_{\max} = 6\sqrt{2}$. Das Viereck ist Sehnenviereck.

(2.35) Zunächst folgt aus (2.28), dass unter allen Dreiecken mit gleichem Umfang die gleichseitigen Dreiecke den größten Flächeninhalt besitzen. Nach (2.32) folgt daraus dann die duale Aussage, d. h. die Behauptung.

(2.38) Die Lösung lässt sich entsprechend (2.37) für einen spitzen oder stumpfen Winkel gestalten. - Es gibt wiederum genau eine Tangente g' an den Kreis, die beim Schnitt mit den Schenkeln des Winkels ein gleichschenkliges Dreieck $A'B'C'$ (mit $|CA'| = |CB'|$) liefert. Der Kreisberührungspunkt sei E ; er ist der Mittelpunkt von $\overline{A'B'}$.

Jede andere Gerade durch E , die ebenfalls die Schenkel schneidet, ergibt ein Dreieck mit einem größeren Flächeninhalt. (Der Beweis kann entsprechend zu (2.36) geführt werden!) Und dies gilt nun erst recht für jede von g' verschiedene Tangente an den Kreis.

(2.42) Angeregt durch die Lösung von (2.40) mit Hilfe des Spezialfalls (2.39), beschäftigen wir uns zunächst mit folgender Frage:

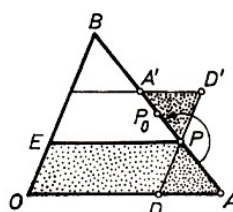


Abb. 7.9

Es sei OAB ein Dreieck. Für welche Punkte $P \in \overline{AB}$ hat das Parallelogramm $ODPE$ mit $D \in \overline{OA}$ und $E \in \overline{BO}$ maximalen Flächeninhalt?

Ist P der Mittelpunkt P_0 von \overline{AB} , dann gilt offenbar $A(ODPE) = \frac{1}{2}A(OAB)$. Ist $P \neq P_0$ und o. B. d. A. ein Punkt zwischen A und P_0 (Abb. 7.9), dann geht bei der Spiegelung an P der Punkt A in einen Punkt A' zwischen P und B über. Das Parallelogramm mit den aufeinanderfolgenden Ecken O, D, D' hat nun - wie man an Hand dieser Spiegelung sofort einsieht - einen kleineren Flächeninhalt als das Dreieck OAB , aber den doppelten des Parallelogramms $ODPE$, d. h., es ist $A(ODPE) < \frac{1}{2}A(OAB)$.

Demnach liegt das Maximum für $P = P_0$ vor.

Mit diesem Hilfssatz ergibt sich nun leicht, dass die Mittelpunkte der Vierecksseiten das gesuchte Maximum bilden. O ist dann der Schnittpunkt der Diagonalen.

(2.43) a) Ist das Dreieck nicht gleichseitig, etwa $|BC| \neq |CA|$, dann kann bei festgehaltener Seite \overline{AB} zu einem von C verschiedenen Kreispunkt D mit $|BD| = |DA|$ und $A(ABD) > A(ABC)$, also zu einem flächengrößeren gleichschenkligen Dreieck übergegangen werden.

Dies leistet derjenige Schnittpunkt D der Mittelsenkrechte von \overline{AB} mit dem Kreis, der mit C auf der gleichen Seite bezüglich g_{AB} liegt.

b) beweist man in analoger Weise.

(2.44) a.) Ist das Tangendendreieck ABC nicht gleichseitig, etwa $|AB| \neq |AC|$, so gibt es nach (2.38) ein Tangendendreieck $AB'C'$ mit $A(AB'C') < A(ABC)$.

b) beweist man nun entsprechend.

(2.45) Zu jeder noch so kleinen Flächeninhaltsgröße können bei festem Punkt P im Innern von \overline{AB} stets so nahe an C liegende Punkte Q und R aus dem Innern von \overline{BC} bzw. \overline{CA} gewählt werden, dass $A(PQR)$ kleiner als die vorgegebene Größe ist.

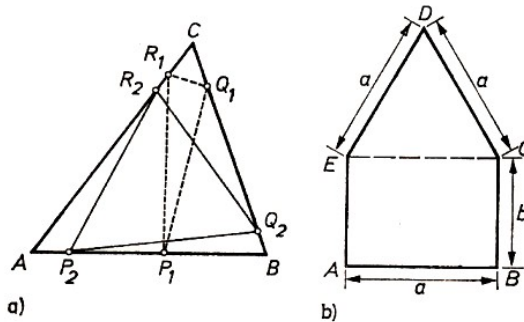


Abb. 7.10

$A(PQR) = 0$ ist jedoch nicht möglich. Andererseits ist offenbar stets $A(PQR) < A(ABC)$, aber zu jeder Flächeninhaltsgröße $F < A(ABC)$ gibt es stets P, Q, R vorgegebener Lage mit $A(PQR) > F$; die Punkte P, Q, R können ja beliebig nahe an A, B bzw. C gewählt werden (Abb. 7.10a).

(2.49) Für das Fünfeck $ABCDE$ gilt (Abb. 7.10b)

$$A = ab + \frac{1}{4}a^2\sqrt{3} \quad \text{und} \quad U = 3a + 2b \quad (2)$$

Aus, der Existenz eines Fünfecks folgt, dass die aus (1) und (2) sich ergebende Gleichung

$$a^2 - \frac{2U}{6 - \sqrt{3}}a + \frac{4A}{6 - \sqrt{3}} = 0$$

in a eine Lösung besitzt; und daraus folgt über die Diskriminante

$$U \geq 2\sqrt{A(6 - \sqrt{3})} \quad (3)$$

Für $U = 2\sqrt{A(6 - \sqrt{3})}$ hat die obige quadratische Gleichung in a nur die Lösung $a = \frac{U}{6 - \sqrt{3}}$, und aus dieser Gleichung folgt

$$a = 2\sqrt{\frac{A}{6 - \sqrt{3}}} \quad \text{und} \quad b = (3 - \sqrt{3})\sqrt{\frac{A}{6 - \sqrt{3}}} \quad (4,5)$$

Gibt es also ein Fünfeck der gesuchten Art, bei dem sogar die Gleichheit in (3) gilt, so bestehen (4) und (5).

Umgekehrt lässt sich nun leicht zeigen, dass ein Fünfeck $ABCDE$ mit (4) und (5) den Flächeninhalt A und minimalen Umfang besitzt.

(2.50) Insgesamt gibt es zwölf verschiedene (d. h. inkongruente) Pentominos. Bestimmt man diese, so lässt sich die Aufgabe leicht durch "Auszählen" ermitteln. Man erhält ein recht merkwürdiges Ergebnis: elf der zwölf Pentominos haben trotz verschiedener Formen den gleichen Umfang 12.

Eines der Pentominos hat den Umfang 10. Damit ist das Maximum 12, das Minimum 10 (und "etwas anderes" gibt es nicht!).

(3.4) Die Funktion f sei auf $[a, b]$ streng monoton wachsend. Dann gilt nach (3.3): Wenn $a \leq x_1 < x_2 \leq b$, dann $f(x_1) < f(x_2) \leq f(b)$. Nach (3.2) bedeutet dies, dass $f(b)$ das Maximum ist.

Analoge Betrachtungen führen zum Minimum $f(a)$. (Für eine streng monoton fallende Funktion "kehren sich die Relationszeichen um".)

Für ein offenes Intervall $a < x < b$ führt man die Annahme eines Extremums an der Stelle x_0 etwa wie in (1.2) zum Widerspruch.

(3.5) Wegen $x \neq 0$ und $\frac{1}{x} - 1 \geq 0$ ergibt sich der größtmögliche Definitionsbereich $\{x : 0 < x \leq 1\}$. Aus $x_1 < x_2$ folgt $\frac{1}{x_1} > \frac{1}{x_2}$, $\frac{1}{x_1} - 1 > \frac{1}{x_2} - 1$ und schließlich

$$\sqrt{\frac{1}{x_1} - 1} > \sqrt{\frac{1}{x_2} - 1}$$

d. h., $f(x)$ ist streng monoton fallend. Aus (3.4) folgt damit die Existenz eines Minimums $f(1) = 0$ an der Stelle $x_0 = 1$ und die Nichtexistenz eines Maximums.

2. Lösung. Da ein Wurzelwert stets nichtnegativ ist, kann sein Minimum nicht kleiner als 0 sein. Der Wert 0 wird aber im vorgegebenen Fall für $x_0 = 1$ angenommen. Angenommen, die Funktion f hätte an der Stelle x_1 mit $1 > x_1 > 0$ ein Maximum M , dann gilt $0 < x_1^2 < x_1 < 1$ mit

$$f(x_1^2) = \sqrt{\frac{1}{x_1^2} - 1} = \sqrt{\frac{1}{x_1} - 1} \cdot \sqrt{\frac{1}{x_1} + 1} > M$$

im Widerspruch zur Annahme.

(3.8) a) Wir bezeichnen $|AP| = |RQ|$ mit x . Dann ist für $A(APQR) = x(|AC| - |RC|)$ mit $x \geq 0$ das Maximum gesucht. Nach dem Strahlensatz (vgl. Abb. 2.15a) gilt

$$x : |AB| = |RC| : |AC| \quad \text{bzw.} \quad |RC| = x \cdot \frac{|AC|}{|AB|}$$

Damit erhalten wir

$$A(APQR) = x \cdot |AC| - x^2 \cdot \frac{|AC|}{|AB|}$$

Nach (3.6), (3.6') hat $A(APQR)$ genau an der Stelle $x_0 = \frac{|AB|}{2} > 0$ ein Maximum, d. h., P ist Mittelpunkt von AB , und wegen des Strahlensatzes sind auch R bzw. Q Mittelpunkte von AC bzw. BC .

b) Wir benutzen die Gleichungen (1) und (2). Da in der dualen Aufgabe U als konstant anzunehmen und für A das Maximum zu ermitteln ist, lösen wir (1) nach $2b$ auf und setzen in (2) ein! Es ergibt sich $A = Ur - \left(2 + \frac{3\pi}{2}\right)r^2$ mit $r > 0$.

Nach (3.6), (3.6') erhält man ein Maximum für A genau für

$$r_0 = \frac{U}{3\pi + 4}$$

Das zugehörige b_0 ergibt sich aus $2b_0 = U - 2r_0 - \pi r_0$ zu

$$b_0 = \frac{\pi + 1}{3\pi + 4} \cdot U$$

(3.11) Nach dem Satz des Pythagoras gilt $x^2 + y^2 = d^2$ bzw. $y^2 = d^2 - x^2$. Die Tragkraft T lässt sich als Funktion $T(x) = x(d^2 - x^2)$ der Variablen x erfassen. Für diese Funktion dritten Grades haben die Koeffizienten a bzw. c die Werte -1 bzw. d^2 , die anderen Koeffizienten sind 0.

Für (***) gilt speziell

$$-x^3 + d^2x - h = -x^3 + (r + 2s)x^2 - s(s + 2r)x + rs^2$$

Mittels Koeffizientenvergleich ergibt sich $r = -2s$, $d^2 = 3s^2$ und $h = 2s^3$. Dabei ist $x_0 = s$ die Extremalstelle und h der Extremalwert: $s = \frac{\sqrt{3}}{3}d$, $h = T_{\max} = \frac{2}{9}\sqrt{3}d^3$.

(3.13) Wir gehen wieder von den Formeln (1) und (2) aus. Wir schreiben (1) in der Gestalt $U = (2 + \pi)r + 2b$.

Soll (3.12) Verwendung finden, so müsste sich b als $b = \frac{c}{r} + er$ darstellen lassen; das ist aber mittels (2) möglich! (2) ist äquivalent zu (2') $2b = \frac{A}{r} + \frac{\pi}{2}r$.

Damit folgt

$$U = \left(2 + \frac{3}{2}\pi\right)r + \frac{A}{r} \quad \text{bzw.} \quad \frac{2U}{4 + 3\pi} = r + \frac{d}{r}$$

mit $d = \frac{2A}{4+3\pi}$, d.h., $\frac{2U}{4+3\pi}$ bzw. U wird genau für $r = \frac{2A}{4+3\pi}$ minimal. Setzt man diesen Wert in (2') für T ein, so erhält man den zugehörigen Wert von b .

(3.17) a) Zu (3.14): Für den äußeren Widerstand $x = \sqrt{d} = \sqrt{R_i^2} = R_i$ (da $R_i > 0$) wird $\frac{e^2}{N}$ minimal und damit N maximal.

b) Zu (3.15): Eine Lieferung umfasst $\sqrt{6400} = 80$ Stück. Es müssen demnach in Abständen von einem halben Monat die Lieferungen erfolgen.

c) Zu (3.16): $f(x)$ wird minimal für $x = 1$, d. h., es gilt $|EF| = 1$ und $|FG| = u$. Dann ist aber das Viereck $EFGH$ in Abb. 3.7b ein Drachenviereck, d. h. $\alpha = \delta$ und $\beta = \gamma$. Aus Abb. 3.7a ist ersichtlich, dass der Strahlendurchgang symmetrisch erfolgen muss.

(4.1) Für alle reellen Zahlen x ist $x^2 \geq 0$ und $x^2 + 1 \geq 1 > 0$:

Demnach ist stets $f(x) \geq 0$ und $f(x) = 0 \Leftrightarrow x^2 = 0 \Leftrightarrow x = 0$. Die Funktion hat also für $x = 0$ das Minimum $f(x) = 0$. Außerdem gilt (*) $(u - 1)^2 \geq 0$ bzw. $u^2 + 1 \geq 2u$ und damit $x^4 + 1 \geq 2x^2$, wenn man u durch x^2 ersetzt. Daraus folgt (wegen $x^4 + 1 > 0$) die Ungleichung $\frac{x^2}{x^4 + 1} \leq \frac{1}{2}$.

Da die Gleichheit in (*) genau für $u = 1$ gilt, hat $f(x)$ genau an den Stellen x mit $x^2 = 1$, also bei $x = 1$ und $x = -1$ das Maximum. (Betrachtet man $\frac{1}{f(x)}$, so kann man die Ergebnisse aus Abschnitt 3.3 benutzen.)

(4.9) Es sei M der Mittelpunkt und R der Radius des Umkreises. Mit den üblichen Bezeichnungen gilt $|\angle BMC| = 2\alpha$, $|\angle CMA| = 2\beta$, $|\angle AMB| = 2\gamma$ (Peripheriewinkelsatz !) und damit

$$A_{\Delta} = \frac{1}{2}R^2(\sin 2\alpha + \sin 2\beta + \sin 2\gamma) \quad (\text{Flächensatz})$$

Mit $A_K = \pi R^2$ folgt

$$\frac{1}{p} = \frac{1}{2\pi}(\sin 2\alpha + \sin 2\beta + \sin 2\gamma)$$

und p wird genau dann minimal, wenn $\frac{1}{p}$ maximal wird.

Das Dreieck sei nicht stumpfwinklig. Dann gelten die Ungleichungen $0 \leq 2\alpha, 2\beta, 2\gamma \leq 180^\circ$. In diesem Intervall ist die sin-Funktion konkav, folglich gilt nach (4.5)

$$\sin 2\alpha + \sin 2\beta + \sin 2\gamma = 3 \sin \frac{2\alpha + 2\beta + 2\gamma}{3}$$

d. h. $\frac{1}{p} = \frac{3}{2\pi} \sin 120^\circ = \frac{3\sqrt{3}}{4\pi}$. Das Gleichheitszeichen gilt nach (4.6) nur für das gleichseitige Dreieck mit $p_{\min} = \frac{4\sqrt{3}\pi}{9}$.

Ist das Dreieck ABC stumpfwinklig mit $\gamma > 90^\circ$, so existiert für den gleichen Umkreis - also bei festem A_K - ein spitzwinkliges Dreieck $A'B'C$ mit größerem Flächeninhalt. (A' bzw. B' sind dabei die Bilder von A bzw. B bei Spiegelung an M .) Folglich kann p_{\min} nicht für stumpfwinklige Dreiecke angenommen werden.

(4.12) Die Betrachtungen zur Lösung von (4.11) können entsprechend auf konvexe Oktaeder und konvexe Ikosaeder übertragen werden. Maximalität liegt genau dann vor, wenn der Körper regelmäßig ist. In diesem Fall ist dann $O_{\max} = \frac{\sqrt{3}}{12}q$ für Oktaeder und $O_{\max} = \frac{\sqrt{3}}{30}q$ für Ikosaeder.

(4.16) Analog zur Lösung von (4.15) erhält man das Maximum von p für $x_i = \frac{k_i s}{a_i t}$ ($i = 1, \dots, n$) mit $t = k_1 + k_2 + \dots + k_n$;

$$p_{\max} = \left(\frac{s}{t}\right)^k \cdot \left(\frac{k_1}{a_1}\right)^k \cdot \dots \cdot \left(\frac{k_n}{a_n}\right)^k$$

(4.17) Bei (4.10) ist (4.15) mit $k_i = 1$ für alle i anwendbar. In (4.11) wird außerdem die Ungleichung (5) benötigt, ebenso in (4.12). Bei (4.18) ist (4.15) mit $n = 2$, $k_1 = m$ und $k_2 = 1$ anzuwenden, bei (4.14) kann man dann entsprechend (4.15) auf die Funktion $g(x, y, z)$ - aber eben nicht auf $f(x, y, z)$ - anwenden.

(4.18) a) Zu (2.29): $k_1 = k_2 = 1$, $a_1 = 1$, $a_2 = 2$, $s = 64$ m. Mit (4.16) ergibt sich $x_1 = a_{\max} = 32$ m, $x_2 = b_{\max} = 16$ m.

b) Zu (2.39): Ist P ein beliebiger Punkt im Innern der Strecke \overline{AB} , so gilt $|AP| + |PB| = |AB| = s$. Ist $x_1 = |AP|$ und x_2 die zweite Rechtecksseite, so folgt bei Benutzung des Strahlensatzes $x_1 + \frac{|AB|}{|AC|}x_2 = |AB|$.

Mit Bezug auf (4.16) gilt also $a_1 = 1$, $a_2 = \frac{|AB|}{|AC|}$, $k_1 = k_2 = 1$. $x_1 \cdot x_2$ wird maximal für

$$x_1 = \frac{1 \cdot |AB|}{1 \cdot 2} = \frac{1}{2}|AB| \quad \text{und} \quad x_2 = \frac{|AC|}{|AB|} \cdot \frac{|AB|}{2} = \frac{1}{2}|AC|$$

c) Zu (2.48): Nach (1) und (2) gilt $A = x_1 \cdot x_2$ mit $x_1 = r$ und $x_2 = 2b - \frac{\pi}{2}r$ und somit

$$U = \left(2 + \frac{3}{2}\pi\right)x_1 + x_2$$

d. h., in (4.15) gilt $k_1 = k_2 = 1$, $a_1 = 2 + \frac{3}{2}\pi$, $a_2 = 1$, $p = A$, $s = U$ und $t = 2$. Aus (**)
ergibt sich

$$x_1 = \frac{2}{4 + 3\pi} \sqrt{\frac{4 + 3\pi}{2}} A \quad \text{bzw.} \quad x_2 = \sqrt{\frac{4 + 3\pi}{2}} A$$

Aus x_2 und $x_1 = r$ errechnet man das gesuchte b .

(4.19) Als Analogon wählen wir: Man bestimme unter allen Quadern mit gleicher Oberfläche O diejenigen mit dem größten Volumen V .

Sind die Kantenlängen a , b und c , so ist also das Produkt abc unter der Nebenbedingung $s = \frac{O}{2} = ab + bc + ca = \text{const}$ zu maximieren. Wir benutzen folglich (4.16) und setzen $x_1 = ab$, $x_2 = bc$, $x_3 = ca$ und $p = V^2 = x_1x_2x_3$.

Mit dem Resultat von (4.16) ergibt sich sofort, dass nur der Würfel die Problemstellung erfüllt.

(4.20) Mit den üblichen Bezeichnungen gilt $V = \frac{2}{3}\pi r^2 h$ und $O = 2\pi r \sqrt{r^2 + h^2}$. Da O zu minimieren ist, kann man versuchen, (4.15) anzuwenden. Dann müsste aber eine Summe auftreten.

Diese Forderung erfüllt

$$s = \frac{O^2}{4\pi^2} = r^2(r^2 + h^2) = r^4 + r^2h^2$$

(O wird genau dann minimal, wenn s minimal wird.) Wir setzen $x_1 = r^4$ und $x_2 = r^2h^2$. Die Bedingung $V = \text{const}$ ergibt dann $V = \frac{2}{3}\pi x_1^{1/4} x_2^{1/2}$. Um ganzzahlige k_i zu erhalten, potenzieren wir.

Es ist $p = \left(\frac{3V}{2\pi}\right)^4 = x_1x_2^2$. In (4.15) gilt nun $a_1 = a_2 = 1$, $k_1 = 1$, $k_2 = 2$, $t = 3$. Aus (**)
ist

$x_1 = \sqrt[3]{\frac{1}{4} \left(\frac{3V}{2\pi}\right)^4}$ und $x_2 = 2x_1$ ablesbar. Mit der oben eingeführten Bedeutung von x_1 und x_2 ergibt die letzte Gleichung $h = \sqrt{2}r$. Daraus folgt aber schon die Behauptung! (Der Leser fertige sich bitte eine Skizze an.)

(4.21) Wir betrachten zunächst das ebene Problem. a_i ($i = 1, 2, 3$) seien die Längen der Seiten i des Dreiecks, x_i die Abstände eines Punktes P von den Seiten i .

Um eventuell (4.16) anwenden zu können, ist eine Summe $s = a_1x_1 + a_2x_2 + a_3x_3$ gesucht, die für alle (x_1, x_2, x_3) konstant ist. Die zufällige Übereinstimmung in den Bezeichnungen erweist sich inhaltlich als richtig.

Bezeichnen x_i die Abstände und a_i die Seitenlängen, so ist für jeden Punkt P im Inneren des Dreiecks s der doppelte Flächeninhalt des Dreiecks. Nach (4.16) ergibt sich das Maximum für $x_i = \frac{s}{3a_i}$.

Ein zugehöriger Punkt P_{\max} zerlegt also mit den von ihm ausgehenden "Eckpunktsstrahlen" das Dreieck in drei inhaltsgleiche Dreiecke. Da sich die Schwerlinien im Dreieck im Verhältnis 2 : 1 teilen, ist der Schwerpunkt gerade ein solcher Punkt. Da sich aber für jeden anderen Punkt die Ecktransversalen in anderen Verhältnissen teilen, gibt es auch nur diesen einen Punkt.

Für das analoge räumliche Problem läuft lediglich i von 1 bis 4, und die a_i sind Flächeninhalte.

(4.23) Wir wählen entsprechend dem Anliegen unseres Abschnitts eine Lösung durch geometrische Interpretation. Da nur quadratische Terme in den Radikanden auftreten und bei Ersetzung von c durch $-c$ wieder die gegebene Ungleichung entsteht, kann man sich auf nichtnegative a, b, c beschränken.

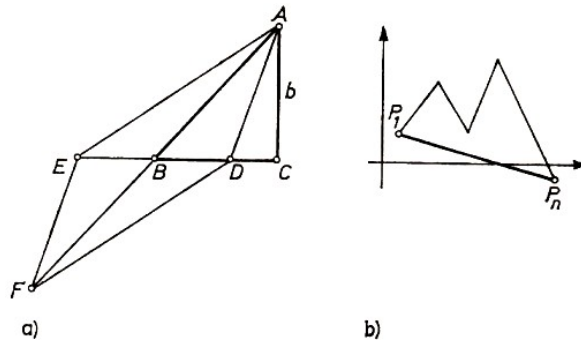


Abb. 7.11

Wir betrachten ein rechtwinkliges Dreieck ABC mit dem rechten Winkel bei C und $|BC| = a$, $|AC| = b$ (vgl. Abb. 7.11a).

Wir tragen auf g_{BC} zwei Punkte D und E so ein, dass $D \neq E$ und $|DB| = |EB| = c$ gilt. Durch F ergänzen wir EAD zu einem Parallelogramm $EADF$. Nach dem Satz von Pythagoras gilt

$$|BA| = \sqrt{a^2 + b^2}, \quad |DA| = \sqrt{(a - c)^2 + b^2}, \quad |EA| = \sqrt{(a + c)^2 + b^2} \quad (*)$$

Da B der Diagonalenmittelpunkt des Parallelogramms $EFDA$ ist, gilt nach der Dreiecksungleichung, angewandt auf das Dreieck AFD , $|DA| + |FD| \geq 2|BA|$. Wegen $(*)$ und $|FD| = |EA|$ folgt daraus die Behauptung.

Im Falle, dass einige der Größen a, b oder c gleich 0 sind, tritt einer der möglichen "Entartungsfälle" ein.

(4.24) (x_i, y_i) sei der Punkt P_i im kartesischen x, y -Koordinatensystem. Dann gilt nach dem Satz des Pythagoras und den Grundlagen der Koordinatengeometrie

$$|P_i P_{i+1}| = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad \text{und} \quad |P_1 P_n| = \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2}$$

Die linke Seite der Ungleichung in (4.24) entspricht damit der Länge eines Polygonzuges P_1, P_2, \dots, P_n , die rechte Seite dagegen der kürzesten Entfernung zwischen den Endpunkten des Polygonzuges (Abb. 7.11 b).

Nach Archimedes' Grundsatz ist damit die Aufgabe gelöst.

(5.4)

$$\begin{aligned} f(x_0 + h) &= (x_0 + h)^3 = x_0^3 + 3x_0^2h + 3x_0h^2 + h^3 \\ f(x_0 + h) - f(x_0) &= 3x_0^2h + 3x_0h^2 + h^3 \\ \frac{f(x_0 + h) - f(x_0)}{h} &= 3x_0^2 + 3x_0h + h^2 \\ f'(x_0) &= \lim_{h \rightarrow 0} (3x_0^2 + 3x_0h + h^2) = 3x_0^2 \end{aligned}$$

(5.5)

$$\begin{aligned} \frac{p(x_0 + h) - p(x_0)}{h} &= \frac{f(x_0 + h) + g(x_0 + h) - f(x_0) - g(x_0)}{h} \\ &= \frac{f(x_0 + h) - f(x_0)}{h} + \frac{g(x_0 + h) - g(x_0)}{h} \end{aligned}$$

Da nach Voraussetzung $f'(x_0)$ und $g'(x_0)$ existieren, existiert auch der Grenzwert $p'(x_0)$ der linken Seite.

Die Gültigkeit von $k'(x_0) = af'(x_0)$ beweist man de facto durch "Ausklammern von a ". (Bei diesen Ableitungen wurden "intuitiv" Rechenregeln für Grenzwerte benutzt.)

(6.3) a) Für $(6; 5)$ hat die Zielfunktion den Wert $f(6; 5) = 12, 2$, für $(8; 3)$ dagegen den größten Wert $f(8; 3) = 12, 6$. Außerdem erfüllt $(8; 3)$ die Nebenbedingungen (2), (3) und (4), wie man leicht nachrechnet, und die Nichtnegativitätsbedingung (5).

b) In Abb. 6.2 ändert sich lediglich der Anstieg der Geraden, die sich aus der Zielfunktion ergeben. Diese Geraden liegen aber jetzt so, dass sie mit dem Eckpunkt $(6, 36; 5, 45)$ auch den Eckpunkt $(8; 3)$ des Zulässigkeitsbereiches gemeinsam haben. Lösung des Optimierungsproblems sind alle (x_1, x_2) mit $6, 36 \leq x_1 \leq 8$ und $x_2 = 15 - 1, 5x_1$.

Für alle diese Paare ist der Wert der Zielfunktion der maximale und beträgt 15.

(6.4) Für $a = 0, 1; 1; 10$ ergeben sich (in der angegebenen Reihenfolge) die maximalen Funktionswerte $2,632; 4; \approx 26,323$.

(6.6) Bei dieser Aufgabe ist zwar keine Nichtnegativitätsbedingung gestellt, die gegebenen Ungleichungen erzeugen aber wieder einen konvexen Zulässigkeitsbereich in der Gestalt einer vierseitigen Pyramide. Die Eckpunkte dieser Pyramide sind $(0, 0, -1)$, $(1, 0, -1)$, $(1, 1, -1)$, $(0, 1, -1)$ und $(\frac{4}{5}, \frac{4}{5}, 10)$.

Da für jeden Wert $y_0 > 0$ der Zielfunktion die Gleichung $y_0 = x_1^2 + x_2^2 + x_3^2$ einer Kugelfläche mit dem Mittelpunkt $(0, 0, 0)$ und dem Radius $\sqrt{y_0}$ entspricht, ist die Kugel mit dem größten Radius gesucht, auf deren Oberfläche wenigstens noch ein Punkt der Pyramide liegt.

Diese Kugelfläche geht dann durch den am weitesten von $(0, 0, 0)$ entfernten Eckpunkt der Pyramide, das ist der Punkt $(\frac{4}{5}, \frac{4}{5}, 10)$; es folgt $f(\bar{x}_1, \bar{x}_2, \bar{x}_3) = 101, 28$.

(6.8) Für die ersten $m - 2$ Jahre sind alle Mittel im Zweig I, für die letzten beiden Jahre im Zweig II zu investieren.

Literatur

- [1] BOLTJANSKI, W. G., und I. M. JAGLOM, Geometrische Extremwertaufgaben. In: Enzyklopädie der Elementarmathematik, Band V, VEB Deutscher Verlag der Wissenschaften, Berlin 1971, S. 261-338 (Übersetzung aus dem Russischen).
- [2] COURANT, R., und H. ROBBINS, Was ist Mathematik?. Springer-Verlag, Berlin-Göttingen-Heidelberg 1962, Kap. VII.
- [3] DÖRRIE, H., Ein neues elementares Verfahren zur Lösung von Extremwertaufgaben. Zeitschrift für den math. u. naturw. Unterr. aller Schulgattungen 50 (1919), 153-177; Nachdruck in: Der Mathematikunterricht 18 (1972) 5, 23-51, Ernst Klett Verlag Stuttgart.
- [4] KAISER, H., Numerische Mathematik und Rechentechnik II. - MfL Bd. 10, VEB Deutscher Verlag der Wissenschaften, Berlin 1980, Kap. 7.
- [5] Kleine Enzyklopädie Mathematik. VEB Bibliographisches Institut, 10. Aufl. Leipzig 1977
- [6] LANGER, K., Extremwertaufgaben in der Geometrie mit elementarer Lösung. Diplomarbeit, PH Potsdam 1983.
- [7] MÜLLER, M., Zur Geschichte der Extremalaufgaben. Mathematik in der Schule 4 (1966), 463-478, 777-782.
- [8] NAGIBIN, F. F, Extrema. Moskau 1966. (russisch)
- [9] NATANSON, I. P., Einfachste Maxima- und Minima-Aufgaben. 7. Aufl., VEB Deutscher Verlag der Wissenschaften, Berlin 1975 (Übersetzung aus dem Russischen), MSB Nr. 15.
- [10] POLYA, G., Mathematik und plausibles Schliessen, Bd. 1. Birkhäuser Verlag, Basel und Stuttgart 1969, Kap. VII-X.
- [11] QUAISSER, E., Bewegungen in der Ebene und im Raum. VEB Deutscher Verlag der Wissenschaften, Berlin 1983, MSB Nr. 116.
- [12] QUAISSER, E., und H.-J. SPRENGEL, Räumliche Geometrie. VEB Deutscher Verlag der Wissenschaften, Berlin 1981, MSB Nr. 103.
- [13] SPRENGEL, H.-J., und O. WILHELM, Funktionen und Funktionalgleichungen. VEB Deutscher Verlag der Wissenschaften, Berlin 1984, MSB Nr. 114.
- [14] STURM, R., Maxima und Minima in der elementaren Geometrie. B. G. Teubner Verlag, Leipzig und Berlin 1910.
- [15] WENTZEL, J. S., Elemente der dynamischen Optimierung. B. G. Teubner, Verlagsgesellschaft, Leipzig 1966.
- [16] WUSSING, H., Mathematik in der Antike, B. G. Teubner, Verlagsgesellschaft, Leipzig 1962.
- [17] WUSSING, H., Vorlesungen zur Geschichte der Mathematik. MfL Bd. 13, VEB Deutscher Verlag der Wissenschaften, Berlin. 1979.

Erklärung verwendeter Symbole

\overline{AB}	Strecke mit den Endpunkten A und B
$ AB $	Länge der Strecke \overline{AB} (bzw. deren Maßzahl)
g_{AB}	Verbindungsgerade der Punkte A und B
$\overline{ABC\dots YZ}$	Streckenzug ($\overline{AB} \cup \overline{BC} \cup \dots \cup \overline{YZ}$)
AB^+	Halbgerade (mit dem Scheitel A , die B enthält)
$\sphericalangle p, q$	Winkel mit den Schenkeln (Halbgeraden) p und q
$\sphericalangle ABC$	Winkel mit den Schenkeln BA^+ und BC^+
$ \sphericalangle p, q $	Größe des Winkels $\sphericalangle p, q$ (bzw. deren Maßzahl)
$g \parallel h$	Gerade g ist parallel zur Geraden h
$A(F)$	Inhalt der Figur F (bzw. dessen Maßzahl)
$f(x)$	Wert der Funktion f an der Stelle x
\mathbb{R}	Menge der reellen Zahlen
\mathbb{R}_+^*	Menge der positiven reellen Zahlen
\mathbb{N}	Menge der natürlichen Zahlen
\mathbb{N}^*	Menge der positiven natürlichen Zahlen, $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$
(a, b)	offenes Intervall, d. h. Menge aller reellen Zahlen x mit $a < x < b$
$[a, b]$	abgeschlossenes Intervall ($[a, b] = (a, b) \cup \{a, b\}$)
OJM	Olympiaden Junger Mathematiker der DDR
•	Aufgabe, die selbständig zu lösen ist
*	bezeichnet etwas schwierigere Probleme