

KLEINE
NATURWISSENSCHAFTLICHE
BIBLIOTHEK

MATHEMATIK

WILENKIN

*Methoden
der schrittweisen Näherung*

KLEINE NATURWISSENSCHAFTLICHE BIBLIOTHEK

Reihe Mathematik, Band 5

**METHODEN
DER SCHRITTWEISEN
NÄHERUNG**

VON N. J. WILENKIN

MIT 34 ABBILDUNGEN



VERLAG MIR, MOSKAU

BSB B. G. TEUBNER VERLAGSGESELLSCHAFT, LEIPZIG



Titel der Originalausgabe :
Наум Яковлевич Виленкин,
Метод последовательных приближений
Verlag Nauka, Moskau 1968
Deutsche Übersetzung und wissenschaftliche Redaktion:
Dozent Dr. R. Hofmann, Leipzig

Gemeinschaftsausgabe des Verlags MIR, Moskau und der
BSB B. G. Teubner Verlagsgesellschaft, Leipzig
© 1974 Verlag MIR, Moskau und BSB B.G. Teubner
Verlagsgesellschaft, Leipzig
1. Auflage
VLN 294-375/66/74 LZV 1089
Satz und Druck: UdSSR
Best.-Nr. 665 723 5

Aus dem Vorwort des Autors

Hauptziel dieses kleinen Buches ist die Darstellung einiger Näherungsverfahren zur Lösung von Gleichungen. Ungeachtet ihres unbestrittenen praktischen Wertes wird ihnen sowohl in der Grund- als auch in der Hochschule nur geringe Aufmerksamkeit geschenkt. Deshalb ist es häufig so, daß ein Student, dem der Grundkurs der höheren Mathematik sonst keine Schwierigkeiten bereitet, sich nicht zu helfen weiß, wenn er einmal eine höhere transzendente Gleichung lösen soll. Mit der Lösung einer Gleichung befassen sich nicht nur Ingenieure, sondern Neuerer der Produktion und Vertreter vieler anderer Fachrichtungen. Auch für Schüler oberer Klassen ist die Kenntnis von Näherungsmethoden nützlich.

Die meisten Näherungsverfahren hängen mit der ersten Ableitung der beteiligten Funktion zusammen, und wir sahen uns deshalb veranlaßt, diesen Begriff hier einzuführen. Wir bedienen uns dabei vor allem anschaulicher geometrischer Überlegungen, damit dieses Büchlein auch für Schüler von der neunten Klasse an zu verstehen ist.

Beim Schreiben dieses Buches wurden Vorlesungen benutzt, die der Autor vor Schülern der 9. und 10. Klassen an der Moskauer Lomonossow-Universität gehalten hat.

Das Material dieser Vorlesungen wurde von S. I. Schwarzburd ... für die außerunterrichtliche Arbeit in einer 9. Klasse verwendet. Der Autor dankt ihm für die Überlassung der von ihm ausgearbeiteten Aufgaben, die im Buch verwendet worden sind...

Für die zweite Auflage ist das Buch überarbeitet worden. Die Iterationsmethode wurde nunmehr auf der Grundlage des Begriffs der kontraktiven Abbildung dargestellt, so daß auch vor Einführung des Ableitungsbegriffs bereits Konvergenzaussagen gemacht werden konnten. Der Teil des Buches, der der Lösung von Gleichungssystemen gewidmet ist, wurde bedeutend erweitert, und schließlich sind die Aufgaben durch Lösungshinweise ergänzt worden.

Vorwort zur deutschen Ausgabe

Naum Jakowlewitsch Wilenkin ist vielen an mathematischen Problemen interessierten Schülern und Lehrern kein Unbekannter mehr. Sicher haben manche das im gleichen Verlag erschienene Buch „Unterhaltsame Mengenlehre“ gelesen und werden mir zustimmen, daß der Autor es versteht, auch schwierigere mathematische Stoffgebiete anschaulich darzustellen und faßlich zu beschreiben.

Nun legen wir wieder ein kleines Buch von ihm vor und hoffen, daß auch diese populär geschriebene Einführung in spezielle Aufgaben und Methoden der praktischen Mathematik viele Freunde unter Schülern, Lehrern und mathematisch interessierten Werktätigen finden möge.

Auch bei uns ist das behandelte Gebiet leider nicht im Lehrplan der Schulen enthalten, obwohl doch erst die zahlenmäßige Bestimmung von Lösungen einer Gleichung oder eines Gleichungssystems ihre praktische Anwendung ermöglicht.

Mit seinem methodisch geschickten Aufbau, seinen vollständig diskutierten Aufgaben und den an vielen Stellen vorhandenen Ansätzen zur vertieften Betrachtung kann das Buch interessierten Schülern von der 9. Klasse an, Absolventen der POS, Abiturienten und vor allem unseren Lehrern und Zirkelleitern zur Lektüre und zum Gebrauch im fakultativen Unterricht der 11. Klasse oder in der außerunterrichtlichen Beschäftigung mit Mathematik empfohlen werden.

Bei der Übersetzung wurde die Terminologie an die in den Schulen der DDR übliche angeglichen.

Hinweise, die der Beseitigung von Fehlern dienlich sind, die sich trotz sorgfältiger Bearbeitung eingeschlichen haben können, werden gern entgegengenommen.

R. Hofmann

Inhalt

1. Einleitung	7
2. Schrittweise Näherung — Sukzessive Approximation . . .	11
3. Achilles und die Schildkröte	14
4. Wie dividiert eigentlich ein Rechenautomat?	17
5. Quadratwurzelziehen durch sukzessive Approximation . .	19
6. Berechnung der Wurzel mit beliebigem Wurzelexponenten	26
7. Die allgemeine Iterationsmethode	28
8. Geometrische Interpretation der Iterationsmethode . . .	31
9. Kontraktive Abbildungen	33
10. Kontraktive Abbildung und Iterationsverfahren	37
11. Die Sekantenmethode (regula falsi)	43
12. Ein verbessertes Sekantenverfahren	48
13. Die Ableitung eines Polynoms	49
14. Das Newtonsche Verfahren für algebraische Gleichungen	52
15. Geometrische Interpretation der Ableitung	54
16. Die geometrische Interpretation des Newtonschen Verfah- rens	57
17. Die Ableitung einer nichtalgebraischen Funktion	59
18. Die Berechnung der Ableitung	60
19. Wie findet man eine Ausgangsnäherung?	63
20. Eine kombinierte Methode zur Lösung von Gleichungen	66
21. Ein Konvergenzkriterium für die allgemeine Iteration . .	68
22. Über die Konvergenzgeschwindigkeit der Iteration . . .	71
23. Die iterative Lösung eines linearen Gleichungssystems	73
24. Die Lösung eines Systems nichtlinearer Gleichungen mit der Methode der sukzessiven Approximation	79
25. Verschiedene Abstandsdefinitionen	82
26. Konvergenzkriterien für den Iterationsprozeß bei linearen Gleichungssystemen	85
27. Sukzessive Approximation in der Geometrie	91
28. Schlußbemerkungen	94
Aufgaben	95
Lösungen	96

1. Einleitung

Während der gesamten Schulzeit begegnen uns im Mathematikunterricht immer wieder Gleichungen oder Systeme von Gleichungen. Sind es zunächst lineare Gleichungen und eventuell lineare Gleichungssysteme, so werden später auch quadratische, biquadratische oder Wurzelgleichungen behandelt, und schließlich lernt man auch, wie trigonometrische oder logarithmische Gleichungen gelöst werden können.

Es ist keineswegs zufällig, wenn den Gleichungen eine solche Aufmerksamkeit geschenkt wird, sondern folgt zwingend aus ihrer Bedeutung in den Anwendungsbereichen der Mathematik. Wo immer Mathematik verwendet wird, ist zur Beantwortung von Problemstellungen die Lösung von Gleichungen oder Gleichungssystemen erforderlich.

In der Schule treten Gleichungen zur Lösung von Problemen praktischer Natur besonders häufig im Physikunterricht auf, und wir wollen uns jetzt eine solche Aufgabe etwas näher ansehen.

Ein Stein fällt in einen Brunnen. Bestimme die Tiefe des Brunnens, wenn das Geräusch des Aufschlages T Sekunden nach dem Loslassen des Steines vom Ohr wahrgenommen wird!

Bezeichnen wir die Brunnentiefe (gemessen in Meter) mit x , so erhalten wir die Gleichung

$$\sqrt{\frac{2x}{g}} + \frac{x}{c} = T$$

zur Bestimmung von x aus der gegebenen Zeit T , der Schallgeschwindigkeit c (gemessen in Meter pro Sekunde) und der Erdbeschleunigung $g = 9,81 \text{ m/s}^2$ ($\sqrt{2x/g}$ ist die Fallzeit des Steines und x/c die Zeit, die der Schall von der Wasseroberfläche bis zum Brunnenrand benötigt). Es handelt sich um eine Wurzelgleichung, die aber in eine quadratische Gleichung überführt werden kann, indem $\sqrt{x} = y$ als neue Variable eingeführt wird

$$\frac{y^2}{c} + \sqrt{\frac{2}{g}} y - T = 0.$$

Diese Gleichung schließlich kann mittels einer bekannten Formel gelöst werden.

Auch zur Lösung geometrischer Aufgaben werden häufig Gleichungen herangezogen. So führt die Aufgabe der stetigen Teilung einer Strecke AB der Länge l auf die quadratische Gleichung

$$x^2 + lx - l^2 = 0.$$

(Eine Strecke AB wird durch einen Punkt C stetig geteilt, wenn für die Abschnitte die Proportion $\overline{AB} : \overline{AC} = \overline{AC} : \overline{CB}$ gilt. In der obigen Gleichung steht x für \overline{AC} .) Will man einen Winkel α in drei gleiche Teile zerlegen ¹⁾, so hat man die etwas kompliziertere Gleichung

$$4x^3 - 3x - \cos \alpha = 0,$$

in der $x = \cos \frac{\alpha}{3}$ gesetzt wird, zu lösen. Eine solche Gleichung — sie heißt kubische Gleichung — wird in der Schule nicht behandelt. Dennoch gibt es auch für sie eine Lösungsformel, wie in der höheren Algebra hergeleitet wird (vgl. Formel (3)).

Schließlich treten aber in der Physik auch Probleme auf, die auf so komplizierte Gleichungen führen, daß nicht nur in der Schule, sondern auch an der Universität keine Lösungsformel dafür gelehrt werden kann.

Nehmen wir beispielsweise eine eiserne Stange (oder, wie man in der Technik sagt, einen Balken) und mauern ihn an beiden Enden fest ein. Schlagen wir nun auf den Balken, dann wird er in Schwingung geraten. Um die Frequenz der Schwingung zu berechnen, muß man, wie in der mathematischen Physik gezeigt wird, die folgende Gleichung lösen

$$e^x + e^{-x} = \frac{2}{\cos x} \quad (1)$$

in der $e = 2,71828\dots$

In der Schule behandelt man keine Lösungsformel für eine Gleichung von diesem Typ, woraus aber nicht geschlossen werden sollte, daß das etwa nur den Rahmen des Schulstoffes übersteigen würde — nein: eine Lösungsformel in dem in der Schule gebräuchlichen Sinne gibt es für die Gleichung

¹⁾ Gemeint ist hier die Teilung mit Zirkel und Lineal, vgl. Lit. Lugowski, H.; H.-J. Weinert: Grundzüge der Algebra III, BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1967.

(1) überhaupt nicht. Wir wollen diese Behauptung noch etwas präzisieren.

Man sagt, daß für eine gegebene Gleichung eine Lösungsformel vorhanden ist, wenn ihre Wurzeln aus den in der Gleichung gegebenen Größen durch Anwendung der Grundrechenarten, durch Wurzelziehen, Logarithmieren, Anwendung der trigonometrischen Funktionen oder ihrer Umkehrfunktionen berechnet werden können.

In diesem Sinne gibt es für die quadratische Gleichung $x^2 + px + q = 0$ die Lösungsformel

$$x = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}, \quad (2)$$

und auch für die kubische Gleichung ¹⁾

$$x^3 + px + q = 0$$

ist eine solche Formel bekannt:

$$\begin{aligned} x = & \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \\ & + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}. \end{aligned} \quad (3)$$

Bei der praktischen Handhabung dieser letzten Formel ergeben sich allerdings eine Reihe von Schwierigkeiten, und die Einführung der komplexen Zahlen erweist sich als notwendig. Auch die Wurzeln einer allgemeinen Gleichung 4. Grades lassen sich noch formelmäßig berechnen. Die entsprechende Formel ist aber bereits so unübersichtlich, daß wir auf ihre Darstellung verzichten wollen.

Für Gleichungen 5. und höheren Grades steht die Sache prinzipiell schlechter. Im Jahre 1826 hat nämlich der norwegische Mathematiker Abel ²⁾ bewiesen, daß es für $n = 5$ keine allgemeine Lösungsformel für die algebraische Gleichung

¹⁾ Jede beliebige kubische Gleichung

$$a_0x^3 + a_1x^2 + a_2x + a_3 = 0$$

kann durch die Transformation $x + \frac{a_1}{3a_0} = y$ auf diese etwas einfachere Form gebracht werden.

²⁾ Die Ideen von Niels H. Abel wurden von dem berühmten französischen Mathematiker Evariste Galois weiterentwickelt, über dessen Leben und Wirken es das sehr lesenswerte Buch „Wen die Götter lieben“ von L. Infeld gibt.

chung

$$a_0x^n + a_1x^{n-1} + \dots + a_n = 0$$

unter alleiniger Verwendung der vier Grundrechenarten und des Radizierens geben kann. Solche Formeln existieren höchstens für eine Reihe von Sonderfällen ¹⁾).

Würden die Mathematiker das Studium der Gleichungen auf deren *exakte* Lösung, d. h. die Aufstellung von Lösungsformeln und deren Anwendung beschränkt haben, so müßte ein Gespräch zwischen einem Mathematiker und einem Ingenieur etwa den folgenden unbefriedigenden Verlauf nehmen: Ingenieur: „Bei der Berechnung eines Bauwerkes bin ich auf die folgende Gleichung geführt worden (er zeigt eine Gleichung vor). Ich muß sie dringend lösen, denn in einem Monat muß das Projekt abgeschlossen sein.“

Mathematiker: „Ich würde Ihnen ja sehr gern helfen, aber für Gleichungen dieses Typs gibt es keine Lösungsformel.“

Ingenieur: „Dann leiten Sie mir eine solche Formel doch bitte her!“

Mathematiker: „Ich werde es nicht einmal versuchen! Längst ist bewiesen, daß es für Gleichungen dieser Gestalt keine Formel geben kann.“

Man kann sich leicht ausmalen, daß die Meinung des Ingenieurs über die Mathematik und ihre Möglichkeiten nach einer solchen Unterhaltung nicht besonders hoch sein würde. Glücklicherweise wird sie aber in dieser Form nie geführt werden müssen!

In der Tat benötigt ein Ingenieur gewöhnlich gar nicht die Lösungsformel, sondern lediglich die eine oder andere Wurzel einer bestimmten Gleichung mit einer gewissen Genauigkeit, wobei es ihm vollkommen gleichgültig ist, ob er sie mittels einer Formel oder auf irgendeinem anderen Wege erhält. Ja selbst eine etwa vorhandene Lösungsformel verwendet er nur dazu, eine Wurzel bis auf eine gegebene Genauigkeit anzunähern. Möge eine solche Formel für den zu berechnenden Wert den Ausdruck $x = 3 + \sqrt{13}$

¹⁾ Wer sich eingehender mit algebraischen Gleichungen befassen möchte, kann eines der folgenden Bücher zur Hand nehmen: Lugowski, H.; H.-J. Weinert: Grundzüge der Algebra, Teil III, BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1967. Obreschkoff, N.: Verteilung und Berechnung der Nullstellen reeller Polynome. VEB Deutscher Verlag der Wissenschaften, Berlin 1963.

ergeben haben, so ist klar, daß sich dieser Wert niemals realisieren läßt. Für den praktischen Zweck müßte $\sqrt{13}$ als Dezimalbruch dargestellt und mit so vielen Stellen verwendet werden, wie es die vorgelegte Ingenieuraufgabe verlangt.

Deshalb wird ein Ingenieur mit seinem Mathematiker vollkommen zufrieden sein, wenn er ihm einen Weg weisen kann, wie er die Wurzel der betrachteten Gleichung mit der gewünschten Genauigkeit gewinnen kann. In der Mathematik ist denn auch eine ganze Reihe von Verfahren ausgearbeitet worden, Gleichungen näherungsweise zu lösen. Der Darstellung einer bestimmten Klasse solcher Verfahren ist dieses kleine Büchlein gewidmet.

2. Schrittweise Näherung — Sukzessive Approximation

Ein großer Teil der Methoden zur angenäherten Berechnung der Wurzeln von Gleichungen basieren auf der Idee der schrittweisen Annäherung, die aber nicht nur zu diesem Zweck, sondern auch bei der Lösung einer ganzen Reihe praktischer Probleme verwendet wird.

So macht beispielsweise jeder Artillerist von dieser Methode Gebrauch. Will er ein bestimmtes Ziel treffen, so richtet er das Rohr des Geschützes in entsprechender Weise und löst den Schuß. Im Falle eines Fehlschusses verändert er den Neigungswinkel und die Richtung des Geschützes

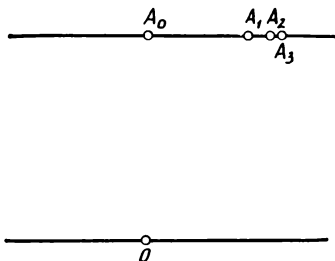


Abb. 1

auf der Grundlage der Beobachtungen des Einschlages und feuert erneut. Nach wenigen Änderungen dieser Art wird er dann mit größter Wahrscheinlichkeit einen Treffer erzielen.

Manchmal ist die Festlegung der Zielrichtung selbst mit schrittweiser Näherung verbunden, wie das Beispiel des Beschlusses eines Flugzeuges mit einem erdgebundenen Fliegerabwehrgeschütz zeigt. Die Kanone möge sich im Punkte O befinden, während das Flugzeug in einer bestimmten Höhe geradeaus fliegt (Abb. 1). Würde man das Geschütz auf den Punkt A_0 richten, in dem sich das Flugzeug im Mo-

ment des Abschusses gerade befindet, so würde sich unweigerlich ein Fehlschuß ergeben, weil sich das Flugzeug während der Flugzeit des Geschosses bereits zum Punkt A_1 bewegt hat. Sind die Geschwindigkeiten von Geschöß und Flugzeug bekannt, so kann dieser Punkt sehr leicht bestimmt werden. Nimmt man ihn als Ziel, so wird man aber wohl immer noch einen Fehlschuß haben. In der Tat ist ja die Flugzeit des Geschosses von O nach A_0 geringer, als die von O nach A_1 , so daß bei seiner Ankunft in A_1 das Flugzeug bereits in A_2 angelangt ist. Auch dieser Punkt kann leicht berechnet und zum Zielpunkt gewählt werden. Jedoch ist auch $\overline{OA_2}$ länger als $\overline{OA_1}$, und das Flugzeug befindet sich in A_3 , wenn das Geschöß A_2 erreicht hat. Nach einigen Schritten dieser Art wird der berechnete Zielpunkt so dicht bei dem wirklichen Ort des Flugzeuges liegen, daß die Granate trifft.

Die Methode der schrittweisen Näherung wird aber auch bei der Lösung vieler weiterer Probleme angewendet. Ein anderes Beispiel soll das illustrieren.

Eine Reihe von Baustellen B_1, B_2, \dots, B_m beziehen ihren Sand von einigen Sandgruben A_1, A_2, \dots, A_n , deren jeweilige Tagesproduktion a_1, a_2, \dots, a_n Tonnen beträgt. Der Tagesverbrauch einer Baustelle B_k werde mit b_k Tonnen angegeben, während die Buchstaben c_{jk} die Unkosten bezeichnen, die beim Transport einer Tonne Kies von der Grube A_j nach der Baustelle B_k entstehen (sie hängen beispielsweise von der Entfernung zwischen A_j und B_k , vom Zustand der Verbindungswege usw. ab). Für die Aufstellung eines Transportplanes fertigen wir uns eine übersichtliche Tabelle an.

Tab. 1

	B_1	B_2		B_m
A_1	x_{11}	x_{12}		x_{1m}
A_2	x_{21}	x_{22}		x_{2m}
\dots	\dots	\dots		\dots
A_n	x_{n1}	x_{n2}		x_{nm}

Hier ist x_{jk} die Bezeichnung für die in Tonnen gemessene Menge Sand, die von A_j nach B_k transportiert wird. Es ist

klar, daß diese Zahlen den folgenden Nebenbedingungen unterworfen sind:

$$x_{j1} + x_{j2} + \dots + x_{jm} \leq a_j$$

(aus der Grube A_j können niemals mehr als a_j Tonnen Kies pro Tag abtransportiert werden) und

$$x_{1k} + x_{2k} + \dots + x_{nk} = b_k$$

(auf der Baustelle B_k werden genau b_k Tonnen Sand am Tag benötigt). Die Gesamtkosten des Transportes belaufen sich bei Realisierung des Planes von Tab. 1 auf

$$\begin{aligned} C = & c_{11}x_{11} + c_{12}x_{12} + \dots + c_{1n}x_{1n} + \\ & + c_{21}x_{21} + c_{22}x_{22} + \dots + c_{2n}x_{2n} + \\ & + c_{m1}x_{m1} + c_{m2}x_{m2} + \dots + c_{mn}x_{mn}. \end{aligned} \quad (4)$$

Man wird bemüht sein, den Plan so zu verändern, daß C möglichst klein wird. Dazu wählt man zunächst einen Orientierungsplan etwa nach folgendem Schema:

Der Grube A_1 wird zunächst die ihr nächstgelegene Baustelle zugeordnet. Fördert sie mehr Sand, als in dieser Baustelle benötigt wird, dann ordnet man ihr eine weitere Baustelle zu und zwar wieder unter den restlichen die nächste. Nach einigen Schritten wird die Produktionskapazität von A_1 erschöpft sein, und wir gehen nunmehr zu A_2 über und ordnen dieser Grube nach dem gleichen Schema eine oder mehrere der noch verbliebenen Baustellen zu, bis auch ihre Kapazität ausgelastet ist und A_3 an die Reihe kommt usw.

Am Ende dieser Überlegungen steht ein Plan, nach dem jede Baustelle durch eine (oder eventuell auch mehrere) Sandgruben ausreichend versorgt wird. Dieser Transportplan wird indessen in der Regel nicht optimal sein, weil gegen Ende des Planungsvorgangs nur noch wenige Baustellen zur Auswahl stehen, die unter Umständen gerade recht weit von den noch verbliebenen Gruben entfernt sind. Deshalb wird der Plan anschließend verändert. Einige der Baustellen, die an Gruben mit niedriger Nummer gebunden waren, werden aus dieser Bindung gelöst und Gruben mit höherer Nummer zugeordnet.

Methoden zur gezielten Veränderung des Transportplanes zum Zwecke der Minimierung der Gesamtkosten werden in

einem Zweig der Mathematik bereitgestellt, der unter dem Namen „Lineare Optimierung“ bekannt ist ¹⁾.

Nach einer Reihe solcher Austauschschritte ist schließlich ein Plan entstanden, in dem die Summe (4) den kleinstmöglichen Betrag besitzt oder ihm so weit nahe kommt, daß weitere Veränderungen keinen wesentlichen Gewinn versprechen.

Auch bei der Aufstellung anderer Pläne (wie etwa der Eisenbahnfahrpläne oder der Stundenpläne einer Schule u. ä.— [Anm. d. deutschen Redaktion]) beginnt man zunächst mit einer groben Näherung, die dann schrittweise verbessert wird, bis ein geeignet gutes Ergebnis erreicht ist.

Die Bearbeitung eines Werkstückes in der Abteilung eines Werkes kann gleichfalls als schrittweise Annäherung an eine vorgegebene Form angesehen werden. Zu Beginn wird eine grobe Näherung — ein Gußstück oder ein Schmiederohling — genommen. Dieser erhält im nächsten Schritt an Drehbänken, auf Bohrwerken und Fräsmaschinen eine dem Sollzustand schon recht nahekommende Gestalt. Feinbearbeitungsgeräte geben ihm schließlich in mehreren Arbeitsgängen die endgültige Form, die mit den Forderungen übereinstimmt.

3. Achilles und die Schildkröte

In der Literatur wird die schrittweise Näherung erstmals von dem griechischen Philosophen Zenon, der um 500 v. u. Z. lebte, dargestellt. Dieser Philosoph versuchte zu beweisen, daß es in der Natur keine Bewegung geben könne, und zwar verlief sein „Beweis“ folgendermaßen:

„Wenn der schnellste Läufer der griechischen Geschichte, Achilles, den Versuch unternehmen würde, eine Schildkröte im Lauf zu überholen, so müßte dieser Versuch unweigerlich scheitern. Ist nämlich der Abstand zwischen ihm und der Schildkröte zum Anfang 1000 Fuß, und legt er in einer Sekunde 10 Fuß zurück, während die Schildkröte in einer Sekunde 1 Fuß vorankommt, so wird er die Distanz von 1000 Fuß in 100 Sekunden bewältigt haben. In dieser Zeit ist aber die Schildkröte ebenfalls nicht faul gewesen und 100

¹⁾ Zu diesem Gegenstand empfehlen wir das Buch BARSOW, A. S.: Was ist lineare Optimierung, das ebenfalls in dieser Reihe erschienen ist.

Fuß weiter gelaufen, und Achilles benötigt weitere 10 Sekunden, um diese Strecke ebenfalls zu durchlaufen. Damit hat aber die Schildkröte Zeit, um weitere 10 Fuß zurückzulegen, und Achilles muß noch 1 Sekunde laufen, um dies ebenfalls zu schaffen. Allerdings ist die Schildkröte inzwischen 1 Fuß weiter, und Achilles bleibt immer hinter ihr zurück, kann sie also jedenfalls nie überholen. Folglich gibt es überhaupt keine Bewegung!“

Es versteht sich, daß diese Schlußfolgerung Zenons nicht mehr als ein geistreiches Paradoxon ist. Bekanntlich ist die Bewegung ja gerade eine Eigenschaft der Materie.

Es gibt ein Gedicht von Alexander Puschkin über Zenon, und darin heißt es dazu :

„Der andere schwieg und begann vor seinen Augen zu gehen. Einen stärkeren Einwand konnte er nicht erheben.

Es priesen alle die tiefsinnige Antwort...“.

Jeder Schüler kann ohne Mühe den Zeitpunkt ausrechnen, an dem Achilles die Schildkröte eingeholt hat. Man braucht dazu lediglich die Gleichung

$$10x - x = 1000 \quad (5)$$

zu lösen, in der x die unbekannte Zeit bezeichnet. Die Lösung ist

$$x = \frac{1000}{9} \text{ s} = 111 \frac{1}{9} \text{ s}.$$

Die Überlegungen Zenons können als originelle Methode zur angenäherten Lösung dieser Gleichung angesehen werden. Addieren wir nämlich in (5) auf beiden Seiten x und dividieren die Gleichung dann durch 10, so erhalten wir die äquivalente Gleichung

$$x = 100 + \frac{x}{10}. \quad (6)$$

Bedenken wir, daß $x/10$ im Vergleich zu x klein ist, dann können wir $x_1 = 100$ als erste Näherung für x ansehen. Diesen Wert setzen wir nun auf der rechten Seite von (6) für x ein und erhalten $x_2 = 100 + 10 = 110$ als genaueren Wert für die gesuchte Lösung x . Setzen wir diese Näherung wiederum rechts ein, dann bekommen wir $x_3 = 111$ und so weiter.

Die Näherungswerte

$$\begin{aligned}x_1 &= 100, \\x_2 &= 110, \\x_3 &= 111, \\x_4 &= 111,1, \\&\dots\end{aligned}$$

sind offenbar gerade die Zahlen, welche in Zenons Überlegungen vorkamen. Je zwei benachbarte Zahlen dieser Folge hängen vermöge der Beziehung

$$x_{n+1} = 100 + \frac{x_n}{10} \quad (7)$$

miteinander zusammen. Diese ermöglicht es, daß wir eine aus der anderen berechnen können (wenn wir von $x_0 = 0$ ausgehen [Anm. d. dtsh. Red.]). Mit wachsendem n nähern sich die Zahlen immer mehr der exakten Lösung $x = 111 \frac{1}{9}$ der Gleichung (5).

Die gerade dargestellte Näherungsmethode für die Lösung einer Gleichung führte deshalb zum Erfolg, weil die Größe $x/10$ klein gegenüber x war. Anderenfalls würden wir Zahlen erhalten, die sich der Lösung nicht nähern. Betrachten wir beispielsweise einen Wettlauf zwischen Achilles und einer leichtfüßigen Gazelle, die 20 Fuß in einer Sekunde zurücklegt. Um die Zeit zu berechnen, die Achilles zum Überholen der Gazelle benötigt, müssen wir jetzt die Gleichung

$$10x - 20x = 1000 \quad (8)$$

lösen und erhalten $x = -100$. Das bedeutet, daß Achilles und die Gazelle vor 100 Sekunden am gleichen Ort gewesen sein könnten, die Gazelle den Läufer inzwischen längst überholt hat und ihm von nun an ständig weiter enteilt.

Versuchen wir die Gleichung (8) nach dem beschriebenen Verfahren zu behandeln, so müssen wir sie zunächst analog zu (5) umformen und bekommen

$$x = 100 + 2x. \quad (9)$$

Wieder wählen wir $x_0 = 0$, setzen ein und erhalten $x_1 = 100$. Setzen wir diesen Wert nun rechts ein, dann erhalten wir die neue Näherung $x_2 = 300$ und so fortfahrend nach-

einander

$$x_0 = 0, x_1 = 100, x_2 = 300, x_3 = 700, \dots$$

Wir sehen, daß die Folge dieser Zahlen keineswegs gegen die exakte Lösung $x = -100$ von (8) strebt.

4. Wie dividiert eigentlich ein Rechenautomat?

Es ist möglich, daß der Leser dieses Buches sich wundert, weshalb wir die Gleichung (5) durch schrittweise Näherung gelöst haben, wenn die Lösung doch ganz genau angegeben werden kann! Aber uns interessierte dabei ja schließlich weniger die Gleichung (5), als vielmehr die Methode der sukzessiven Approximation selbst, die wir nachher auf kompliziertere Gleichungen anwenden wollen.

Im übrigen soll aber auch darauf aufmerksam gemacht werden, daß die Lösung solch einfacher Gleichungen mit der Methode der schrittweisen Näherung im Zusammenhang mit der Entwicklung elektronischer Rechengерäte in den letzten Jahren direkt notwendig geworden ist. Es existiert nämlich eine ganze Reihe von Modellen solcher Automaten, die zunächst nur die drei Grundrechenarten Addition, Subtraktion und Multiplikation ausführen können. Außerdem können sie durch Zahlen der Form 2^n dividieren.

Wie verläuft in solchen Automaten eigentlich die Division durch eine beliebige Zahl?

Will man eine Zahl b durch eine Zahl a dividieren, so läuft das auf die Lösung der linearen Gleichung $ax = b$ hinaus. Weil die Maschine multiplizieren und durch Potenzen von 2 dividieren kann, können wir annehmen, daß $1/2 \leq a < 1$ gilt (weil man in jedem anderen Fall durch Multiplikation beider Seiten der Gleichung oder Division durch 2^n diesen Fall erreichen kann).

Nun bringen wir die Gleichung $ax = b$ auf die Form

$$x = (1 - a)x + b. \quad (10)$$

Als erste Näherung für die gesuchte Lösung x wählen wir $x_1 = b$ und bezeichnen den Fehler mit α_1 , so daß

$$x_1 + \alpha_1 = \frac{b}{a}$$

gilt. Nunmehr ergibt sich aus (10)

$$x_1 + \alpha_1 = (1 - a)(x_1 + \alpha_1) + b = (1 - a)x_1 + b + (1 - a)\alpha_1. \quad (11)$$

Wegen $1/2 \leq a < 1$ ist $0 < 1 - a \leq 1/2$, und damit ist der Koeffizient $(1 - a)$ genügend klein, so daß wir auf der rechten Seite von (11) den Summanden $(1 - a) \alpha_1$ weglassen können, weil er von zweiter Ordnung mit α_1 klein wird. Wir bekommen

$$x_1 + \alpha_1 \approx (1 - a) x_1 + b.$$

Die Zahl

$$x_2 = (1 - a) x_1 + b$$

nehmen wir nun als zweite Näherung. Es bezeichne α_2 den Fehler von x_2 gegenüber x , d. h. $x_2 + \alpha_2 = \frac{b}{a}$. Dann erhalten wir aus (10)

$$x_2 + \alpha_2 = (1 - a) x_2 + b + (1 - a) \alpha_2.$$

Lassen wir auf der rechten Seite erneut den von zweiter Ordnung mit α_2 kleinen Term $(1 - a) \alpha_2$ weg, dann erhalten wir die Näherungsgleichung

$$x_2 + \alpha_2 \approx (1 - a) x_2 + b,$$

also können wir

$$x_3 = (1 - a) x_2 + b$$

als dritte Näherung ansehen. Ganz analog sehen wir ein, daß

$$x_4 = (1 - a) x_3 + b$$

die vierte Näherung ist und so weiter. Die Zahlen $x_1, x_2, \dots, x_n, \dots$, die nacheinander nach der Formel

$$x_{n+1} = (1 - a) x_n + b \tag{12}$$

zu berechnen sind, nähern sich mit wachsendem n dem Wert b/a beliebig genau.

In der Formel kommen aber nur die Grundrechenarten Addition, Subtraktion und Multiplikation vor, d. h. die Rechnungen sind mit der betrachteten Maschine ausführbar.

Die soeben dargestellte Methode der Division beruht wesentlich auf der Formel für die Summe der unendlichen geometrischen Reihe.

In der Tat können wir $\frac{b}{a} = \frac{b}{1 - (1 - a)}$ schreiben, woraus sich nach der erwähnten Formel

$$b/a = b + b(1 - a) + b(1 - a)^2 + \dots + b(1 - a)^{n-1} + \dots \tag{13}$$

ergibt. Bezeichnen wir die Summe der ersten n Summanden

jeweils mit x_n

$$x_n = b + b(1-a) + \dots + b(1-a)^{n-1},$$

dann gilt offenbar

$$\begin{aligned} x_{n+1} &= b + b(1-a) + \dots + b(1-a)^n = \\ &= b + (1-a)[b + b(1-a) + \dots + b(1-a)^{n-1}] = \\ &= b + (1-a)x_n. \end{aligned}$$

Diese Formel ist aber mit (12) identisch. Folglich entspricht der Ersetzung von b/a durch x_n die der Summe der unendlichen Reihe in (13) durch die ersten n Glieder. Mit wachsendem n nähert sich der Wert dieser Summe beständig der Zahl b/a . (Die Glieder der Reihe (13) werden mit wachsendem n immer kleiner, weil wir $0 < (1-a) \leq 1/2$ vorausgesetzt hatten.)

5. Quadratwurzelziehen durch sukzessive Approximation

In der Schule wird die Berechnung der Quadratwurzel nicht behandelt. Vielmehr lernt man schon frühzeitig den Gebrauch von Quadrattafeln zu ihrer Berechnung.

Wir wollen nun eine Methode kennenlernen, mit der man die Quadratwurzel einer beliebigen positiven Zahl leicht berechnen kann. Sie war bereits im alten Babylonien lange vor unserer Zeitrechnung bekannt, und auch der Mathematiker und Philosoph Heron von Alexandrien hat sich ihrer bedient. Später war sie in Vergessenheit geraten, bis sie sich als beste Methode zur Berechnung der Quadratwurzel auf Rechenautomaten erwies.

Wir wollen einmal die Wurzel aus 28 nach diesem Verfahren berechnen und gehen zunächst von einer groben Näherung, sagen wir $x_1 = 5$, aus. Wieder bezeichnen wir den Fehler mit α_1 , setzen also $\sqrt{28} = 5 + \alpha_1$. Zur Bestimmung von α_1 quadrieren wir diese Gleichung und erhalten zunächst $28 = 25 + 10\alpha_1 + \alpha_1^2$ und daraus

$$\alpha_1^2 + 10\alpha_1 - 3 = 0. \quad (14)$$

Für α_1 hat sich also eine quadratische Gleichung ergeben. Versuchen wir dieselbe zu lösen, so ergibt sich nach der schon früher benutzten Lösungsformel $\alpha_1 = -5 \pm \sqrt{28}$, d. h., die exakte Lösung von (14) verlangt die Berechnung von $\sqrt{28}$. Es könnte also den Anschein haben, als befänden

wir uns in einem „Teufelskreis“: Um $\sqrt{28}$ zu berechnen, brauchen wir a_1 . Wollen wir a_1 bestimmen, dann benötigen wir wiederum $\sqrt{28}$!

Einen Ausweg aus dieser Situation zeigt uns die folgende Überlegung: Die Abweichung a_1^2 des Näherungswertes $x_1 = 5$ von $x = \sqrt{28}$ ist nicht groß, sie ist jedenfalls kleiner als eins. Dann ist aber a_1^2 noch kleiner! Also versuchen wir einen Näherungswert von a_1 zu berechnen, indem wir in (14) die recht kleine Zahl a_1^2 weglassen. Wir erhalten dann die Näherungsgleichung $10a_1 - 3 \approx 0$, aus der sich $a_1 \approx 0,3$ ergibt.

Wegen $\sqrt{28} = 5 + a_1$ wird $x_2 = 5 + 0,3 = 5,3$ die zweite Näherung für $\sqrt{28}$.

Um eine noch genauere Näherung für $\sqrt{28}$ zu erhalten, wiederholen wir den gesamten Gedankengang (mit x_2 anstelle von x_1 [Anm. d. deutschen Red.]). Wir bezeichnen mit a_2 den Fehler von x_2 , d. h., wir setzen $x_2 = \sqrt{28} - a_2$, vernachlässigen a_2^2 und bekommen $\sqrt{28} \approx x_2^2 + 2x_2a_2$, woraus $a_2 = \frac{28 - x_2^2}{2x_2}$ folgt. Das bedeutet, daß

$$x_3 = x_2 + \frac{28 - x_2^2}{2x_2} = \frac{28 + x_2^2}{2x_2}$$

eine dritte Näherung ist. Aus $x_2 = 5,3$ ergibt sich nach dieser Formel $x_3 = 5,2915\dots$. Genauso bekommen wir mit $x_3 = 5,2915$ die folgende Näherung

$$x_4 = \frac{28 + x_3^2}{2x_3} = 5,2915\dots,$$

und ganz allgemein liefert

$$x_{n+1} = \frac{28 + x_n^2}{2x_n} \quad (15)$$

bei bekanntem x_n die nächste Näherung x_{n+1} .

Dabei wird in jedem Schritt des Prozesses eine bessere Näherung für $\sqrt{28}$ erzeugt, und wir brechen die Rechnung ab, wenn die Differenz zwischen x_{n+1} und x_n innerhalb der mitgeführten Stellenzahl null geworden ist. Verlangen wir etwa, $\sqrt{28}$ mit der Genauigkeit 0,001 zu berechnen, dann genügen nach dem soeben Gesagten bereits vier Näherungen, und wir erhalten $\sqrt{28} = 5,2915$, denn es war ja in der Tat $x_3 = 5,2915\dots$ und $x_4 = 5,2915\dots$.

Nach dem behandelten Verfahren kann mühelos die Quadratwurzel aus jeder beliebigen positiven Zahl gezogen werden. Es sei $a > 0$, dann werden die sukzessiven Näherungen nach

$$x_{n+1} = \frac{a + x_n^2}{2x_n} \quad (16)$$

mit einer geeigneten Anfangsnäherung x_0 gewonnen.

Man kann Formel (16) auch auf einem etwas anderen Wege herleiten.

Nehmen wir an, daß wir bereits eine Näherung x_n für \sqrt{a} haben. \sqrt{a} ist wegen $\sqrt{a} = \sqrt{x_n \cdot \frac{a}{x_n}}$ das geometrische Mittel der beiden positiven Zahlen x_n und a/x_n , also wird das arithmetische Mittel dieser Zahlen eine Näherung für \sqrt{a} sein, d. h.

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right) = \frac{x_n^2 + a}{2x_n},$$

und das ist genau die Formel (16).

Wir können also sagen:

Die oben dargestellte Methode zur Berechnung der Quadratwurzel besteht darin, daß in jedem Schritt der Näherung das geometrische Mittel von x_n und a/x_n durch ihr arithmetisches ersetzt wird.

Jetzt wollen wir uns mit der Frage beschäftigen, ob dieser Iterationsprozeß stets zum Ziele führt oder nicht. Wir wollen also herausbekommen, ob sich die Sache stets so wie bei Achilles und der Schildkröte verhält, oder ob auch der Fall eintreten kann, über den wir beim Wettlauf des griechischen Helden mit der Gazelle diskutiert hatten. (Im ersten Fall spricht der Mathematiker von einem *konvergenten*, im zweiten Fall von einem *bestimmt divergenten* Prozeß.) Wir werden beweisen, daß es beim Wurzelziehen niemals Komplikationen gibt und der Prozeß immer konvergiert.

Dazu vergleichen wir die Fehler $\alpha_n = \sqrt{a} - x_n$ und $\alpha_{n+1} = \sqrt{a} - x_{n+1}$ zweier aufeinanderfolgender Näherungen. Nach Formel (16) können wir α_{n+1} folgendermaßen schreiben

$$\alpha_{n+1} = \sqrt{a} - x_{n+1} = \sqrt{a} - \frac{x_n^2 + a}{2x_n} = -\frac{x_n^2 - 2x_n \sqrt{a} + a}{2x_n}.$$

Nun ist aber

$$x_n^2 - 2x_n \sqrt{a} + a = (x_n - \sqrt{a})^2 = \alpha_n^2$$

und demzufolge

$$\alpha_{n+1} = -\frac{\alpha_n^2}{2x_n}. \quad (17)$$

Wir betrachten nur positive Näherungen x_n für \sqrt{a} , und deshalb folgt aus (17), daß alle Fehler $\alpha_2, \alpha_3, \dots, \alpha_n, \dots$ negativ sind. Mit anderen Worten sind alle Näherungen x_n von der zweiten an ganz sicher größer als \sqrt{a} ¹⁾, während x_1 zunächst sowohl größer als auch kleiner sein kann.

Nun beweist man mit Hilfe der Formel (17) aber leicht, daß der Betrag des Fehlers jeder Näherung x_n in jedem Schritt mindestens um die Hälfte kleiner als im vorigen ist. Man kann nämlich (17) auch so schreiben

$$\alpha_{n+1} = -\frac{\alpha_n}{2x_n} \alpha_n = \frac{x_n - \sqrt{a}}{2x_n} \alpha_n = \left(\frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right) \alpha_n,$$

und demnach ist

$$|\alpha_{n+1}| = \left| \frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right| |\alpha_n| \quad (18)$$

richtig. Wegen $0 < x_n$ ist

$$\frac{1}{2} - \frac{\sqrt{a}}{2x_n} < \frac{1}{2},$$

und wegen der oben abgeleiteten Aussage $x_n > \sqrt{a}$ für $n \geq 2$ gilt

$$\frac{1}{2} - \frac{\sqrt{a}}{2x_n} > 0.$$

Damit erhalten wir schließlich

$$\left| \frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right| < \frac{1}{2}. \quad (19)$$

Aus (18) und (19) zusammen leiten wir aber sofort

$$|\alpha_{n+1}| < \frac{1}{2} |\alpha_n|$$

¹⁾ Das liegt daran, daß das arithmetische Mittel zweier positiver Zahlen stets größer als das geometrische ausfällt ($\sqrt{ab} \leq \frac{1}{2}(a+b)$, falls $a, b \geq 0$ [Anm. d. dtsh. Red.]).

her, und das ist gerade unsere Behauptung: In jedem Schritt der Näherung wird der Absolutbetrag des Fehlers wenigstens um die Hälfte kleiner.

Folglich ist der Betrag des Fehlers im 2. Schritt mindestens viermal so klein wie am Anfang, beträgt im 3. Schritt nur noch $1/8$ des Anfangsfehlers usw. Mit wachsendem n wird der Fehler ersichtlich beliebig klein, er konvergiert gegen null! Das besagt aber wegen $\alpha_n = \sqrt{a} - x_n$, daß x_n mit wachsendem n dem gesuchten Wert \sqrt{a} beliebig nahe kommt (*die Folge $\{x_n\}$ konvergiert für $n \rightarrow \infty$ gegen \sqrt{a}* [Anm. d. dtsh. Red.]).

Jetzt wollen wir noch herausfinden, welchen Einfluß die Wahl der Anfangsnäherung x_1 auf den Verlauf der schrittweisen Näherung hat. Zuerst stellen wir fest, daß x_1 auf das endgültige Ergebnis überhaupt keinen Einfluß hat.

Wir hatten ja schon gesehen, daß die Folge $\alpha_2, \alpha_3, \dots$ $\dots, \alpha_n \dots$ für $n \rightarrow \infty$ ganz unabhängig von x_1 gegen null geht. Das besagt jedoch, daß bei vorgegebener Stellenzahl der Rechnung immer der gleiche Näherungswert für \sqrt{a} herauskommt, wie immer auch x_1 gewählt war. Ja selbst dann, wenn wir x_1 sehr groß wählen, gelangen wir schließlich zum richtigen Resultat. Nach 10 Schritten ist der Betrag des Fehlers mindestens um den Faktor $1/1000$ kleiner geworden (denn $2^{10} = 1024 \approx 1000$, also $1/2^{10} = 1/1024 < 1/1000$), nach 40 Schritten aber bereits auf den billionsten Teil abgesunken ($2^{40} \approx 10^{12}$!). Würden wir also zur Berechnung von $\sqrt{2}$ beispielsweise von $x_1 = 10^6$ ausgehen, dann wäre $\alpha_1 \approx 10^6$ und $|\alpha_{40}| < 10^{-6}$. Selbst bei einem Startwert von 1 Million hätten wir schließlich einen Wert für $\sqrt{2}$ erhalten, der weniger als ein Millionstel vom wahren Wert abweicht.

Die eben angestellten Betrachtungen lehren andererseits auch, daß die Wahl von x_1 einen großen Einfluß auf die Geschwindigkeit der Konvergenz besitzt. Wird die Anfangsnäherung unglücklich gewählt, dann muß man möglicherweise recht lange warten, bis die Differenz zwischen x_{n+1} und x_n genügend klein geworden ist und die gewünschte Genauigkeit anzeigt. Eine gut gewählte Näherung wiederum kann die Sache sehr beschleunigen. Häufig hat man deshalb folgendes Vorgehen: die Anfangsnäherung x_1 entnimmt man

einer Quadrattafel und berechnet mit

$$x_2 = \frac{a + x_1^2}{2x_1} \quad (20)$$

in nur einem Schritt eine hinreichend genaue Lösung. Ein solches Vorgehen ist deshalb besonders wirkungsvoll, weil die Konvergenzgeschwindigkeit der Folge $\{x_n\}$ zunimmt, wenn x_n sich dem Wert \sqrt{a} nähert. In der Tat hatten wir ja bei der Herleitung der Formel $|\alpha_{n+1}| < \frac{1}{2} |\alpha_n|$ in (18)

den Faktor $\left| \frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right|$ durch $\frac{1}{2}$ ersetzt. Wenn aber x_n nahe bei \sqrt{a} liegt, dann ist die Differenz wesentlich kleiner als $\frac{1}{2}$ und $|\alpha_{n+1}| = \left| \frac{1}{2} - \frac{\sqrt{a}}{2x_n} \right| |\alpha_n|$ viel kleiner als $\frac{1}{2} \cdot |\alpha_n|$.

Wir wollen diese Überlegungen noch präzisieren. Dazu betrachten wir neben dem absoluten Fehler $|\alpha_n| = |\sqrt{a} - x_n|$ auch den relativen Fehler β_n von x_n , d. h. das Verhältnis des absoluten Fehlers $|\alpha_n|$ und des exakten Wertes \sqrt{a} :

$$\beta_n = \frac{|\alpha_n|}{\sqrt{a}} = \left| 1 - \frac{x_n}{\sqrt{a}} \right|.$$

Aus (17) erhalten wir dann für β_{n+1} die Formel

$$\beta_{n+1} = \frac{|\alpha_{n+1}|}{\sqrt{a}} = \frac{|\alpha_n^2|}{2x_n \sqrt{a}}.$$

Wegen $x_n > \sqrt{a}$ folgt daraus

$$\beta_{n+1} < \frac{|\alpha_n|^2}{2(\sqrt{a})^2} = \frac{1}{2} \beta_n^2.$$

Für die relativen Fehler zweier aufeinanderfolgender Näherungen gilt demzufolge

$$\beta_{n+1} < \frac{\beta_n^2}{2}. \quad (21)$$

Ist beispielsweise der relative Fehler von x_n gleich 0,01, dann ist derjenige von x_{n+1} nicht größer als 0,00005 und der von x_{n+2} sogar sicher kleiner als 0,000 000 000 13. Wir sehen, daß die Genauigkeit der Näherung immer schneller und schneller zunimmt. Es kann gezeigt werden, daß sich

die Zahl der richtigen Stellen von Schritt zu Schritt verdoppelt, wenn wir einmal nahe genug an \sqrt{a} herangekommen sind. *Beispiel*: Zu berechnen ist $\sqrt{238}$ mit einer Genauigkeit von $\pm 0,00001$.

Diskussion: Der Zahlentafel im „Taschenbuch der Mathematik“ von Bronstein und Semendjajew ¹⁾ entnehmen wir $\sqrt{238} = 15,43$ und wählen diese Zahl als x_1 . Dann berechnen wir nach (20)

$$x_2 = \frac{15,43^2 + 238}{30,86} = 15,42725 \dots$$

Nun schätzen wir die Genauigkeit von x_2 ab. Da der Fehler von x_1 keinesfalls größer als 0,01 ist, können wir mit $\alpha_1 = 0,01$ rechnen und bekommen

$$\beta_1 = \frac{0,01}{15,43} < 0,001.$$

Aber dann ist wegen (21)

$$\beta_2 < \frac{0,001^2}{2} = 0,000\,000\,5,$$

und das besagt, daß der absolute Fehler von x_2 nicht größer als $15,43 \cdot 0,000\,000\,5 < 0,000\,01$ sein kann. Wir sind damit sicher, daß $\sqrt{238} = 15,42725$ in allen 7 Stellen mit dem wahren Wert von $\sqrt{238}$ übereinstimmt.

Würden wir den Wert auf 14 Stellen benötigen, dann wäre bereits die dritte Näherung ausreichend.

Zum Schluß dieses Kapitels gehen wir noch auf eine Besonderheit der Methode der schrittweisen Näherung ein. Die oben betrachtete Unabhängigkeit des Ergebnisses vom Anfangswert führt nämlich dazu, daß man sich auch einmal verrechnen kann und dennoch das richtige Ergebnis bekommt. Die Rechnung dauert dann höchstens etwas länger. Machen wir beispielsweise im n -ten Schritt einen Fehler, dann erhalten wir eine Zahl y_n anstelle der richtigen Näherung x_n . Diese Zahl betrachten wir nun einfach als erste Näherung einer neuen Folge, von der wir sicher sind, daß auch sie gegen \sqrt{a} konvergiert. Diese günstige Eigenschaft der Methode kann man insofern direkt ausnutzen, als man die Rechnung zunächst mit geringerer Stellenzahl beginnt

¹⁾ BRONSTEIN, I. N.; SEMENDJAJEW, K. A.: Taschenbuch der Mathematik. BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1973.

und erst mit zunehmender Konvergenz auf die verlangte Stellenzahl geht, womit unter Umständen Rechenzeit eingespart werden kann.

6. Berechnung der Wurzel mit beliebigem Wurzelexponenten

Wir werden jetzt sehen, daß man die oben dargestellte Methode der Quadratwurzelberechnung auch auf die Bestimmung einer Wurzel mit beliebigem natürlichem Wurzelexponenten übertragen kann. Dazu benötigen wir die folgende Formel ¹⁾:

$$(x + a)^k = x^k + kx^{k-1}a + \dots, \quad (22)$$

in der die Punkte Summanden andeuten, welche die höheren Potenzen a^2, a^3, \dots usw. enthalten.

Wir beweisen diese Formel. Aus der Schule ist bekannt, daß

$$(x + a)^2 = x^2 + 2xa + a^2,$$

$$(x + a)^3 = x^3 + 3x^2a + 3xa^2 + a^3.$$

Diese Gleichungen können wir auch so schreiben:

$$(x + a)^2 = x^2 + 2xa + \dots, \quad (23)$$

$$(x + a)^3 = x^3 + 3x^2a + \dots \quad (24)$$

Formel (22) ist also für $k = 2$ und $k = 3$ richtig.

Nun multiplizieren wir die beiden Seiten von (24) mit $(x + a)$:

$$(x + a)^4 = (x^3 + 3x^2a + \dots)(x + a).$$

Multiplizieren wir hier die rechts stehenden Klammern aus, so ergibt sich genau ein Term — nämlich x^4 — der a nicht enthält, während zwei Terme \sqrt{a} in der ersten Potenz enthalten, nämlich $3ax^3$ und ax^3 .²⁾ Was die übrigen Glieder angeht, so enthalten sie a wenigstens in der zweiten Potenz, und wir können deshalb schreiben

$$(x + a)^4 = x^4 + 3x^3a + x^3a + \dots = x^4 + 4x^3a + \dots, \quad (25)$$

wenn wie früher die drei Punkte für Glieder von mindestens zweiter Potenz in a stehen. Wir haben damit gezeigt, daß die Formel (20) auch für $k = 4$ richtig ist. Nun läßt sich aber aus

¹⁾ Formel (22) folgt sofort aus dem binomischen Lehrsatz, dessen allgemeine Kenntnis wir aber hier nicht voraussetzen wollen.

(25) ganz genauso

$$(x + a)^5 = x^5 + 5x^4a + \dots \quad (26)$$

ableiten, und man sieht, daß sich auf diesem Wege die Formel (22) für jeden positiven ganzzahligen Exponenten k beweisen läßt ¹⁾.

Jetzt wenden wir uns wieder der Berechnung der Wurzel mit beliebigem Wurzelexponenten k zu und nehmen an, daß uns eine erste Näherung x_1 für die k -te Wurzel aus a bekannt ist. Bezeichnen wir mit α_1 wieder den Fehler von x_1 , so bedeutet das vereinbarungsgemäß $x_1 + \alpha_1 = \sqrt[k]{a}$ und $(x_1 + \alpha_1)^k = a$, nach Formel (22) aber

$$x_1^k + kx_1^{k-1}\alpha_1 + \dots = a,$$

wo die drei Punkte Glieder mit α_1^2, α_1^3 , usw. andeuten.

Liegt die gewählte Ausgangsnäherung hinreichend nahe bei $\sqrt[k]{a}$, dann ist ihr Fehler α_1 klein, und man kann die Glieder höherer Ordnung in α_1 vernachlässigen. So erhält man die Näherungsgleichung

$$x_1^k + kx_1^{k-1}\alpha_1 \approx a,$$

aus der

$$\alpha_1 \approx \frac{a - x_1^k}{kx_1^{k-1}}$$

abgeleitet wird. Demzufolge können wir

$$x_2 = x_1 + \frac{a - x_1^k}{kx_1^{k-1}} = \frac{a + (k-1)x_1^k}{kx_1^{k-1}}$$

als zweite Näherung für $\sqrt[k]{a}$ verwenden. Nach der gleichen Überlegung gewinnen wir

$$x_3 = \frac{a + (k-1)x_2^k}{kx_2^{k-1}}$$

als dritte Näherung und erkennen ganz allgemein das Bildungsgesetz der entsprechenden Folge von Näherungswerten

$$x_{n+1} = \frac{a + (k-1)x_n^k}{kx_n^{k-1}}. \quad (27)$$

¹⁾ Der exakte Beweis verläuft nach dem Beweisschema der vollständigen Induktion [Anm. d. dtsh. Red.].

Wie bei der Quadratwurzeliteration kann auch hier gezeigt werden, daß der dargestellte Prozeß bei beliebig gewählter Ausgangsnäherung x_1 konvergiert (sofern x_1 eine positive Zahl ist), was, wie erinnerlich sein wird, gerade bedeutet, daß die Folge der Zahlen $x_2, x_3, \dots, x_n, \dots$ für beliebig gewähltes positives x_1 dem Wert $\sqrt[k]{a}$ beliebig nahe kommt.

Den Iterationsprozeß setzt man so lange fort, bis sich die Zahlen x_n und x_{n+1} innerhalb der verlangten Stellenzahl nicht mehr unterscheiden.

Beispiel: Berechne $\sqrt[3]{970}$ mit einer Genauigkeit von 0,001.

Diskussion: Für $k = 3$ ist (27) durch

$$x_{n+1} = \frac{a + 2x_n^3}{3x_n^2} \quad (28)$$

zu ersetzen. Bei uns ist $a = 970$, und wir nehmen $x_1 = 10$ an. Dann folgt aus (28)

$$x_2 = \frac{970 + 2 \cdot 10^3}{3 \cdot 10^2} = \frac{2970}{300} = 9,900,$$

$$x_3 = \frac{970 + 2 \cdot 9,9^3}{3 \cdot 9,9^2} = \frac{2940,60}{294,03} = 9,899.$$

Wir sehen, daß innerhalb der verlangten Genauigkeit x_2 mit x_3 übereinstimmt. Deshalb gilt $\sqrt[3]{970} = 9,899$ bis auf einen Fehler von $\pm 0,001$.

7. Die allgemeine Iterationsmethode

Alle oben ausgearbeiteten Beispiele sind Spezialfälle einer allgemeinen Lösungsmethode für Gleichungen, die als *Iterationsmethode* oder *sukzessive Approximation* bezeichnet wird. Das Wesen dieses wichtigen Hilfsmittels besteht in folgendem.

Will man die Gleichung $f(x) = 0$ lösen, so formt man sie zunächst einmal so um, daß sie die Gestalt

$$x = \varphi(x) \quad (29)$$

bekommt. Dann verschafft man sich eine Anfangsnäherung x_1 und setzt sie auf der rechten Seite von (29) für die Variable x ein. Den berechneten Wert $x_2 = \varphi(x_1)$ nimmt man als zweite Näherung und verfährt dann wie mit x_1 und so

weiter. Allgemein berechnet man also die nächste Näherung x_{n+1} aus der bereits bekannten nach der Vorschrift

$$x_{n+1} = \varphi(x_n).$$

Ergibt sich nach einigen Schritten innerhalb der mitgeführten Stellenzahl Übereinstimmung zwischen x_n und x_{n+1} , so ist $x_n = \varphi(x_n)$ in eben diesem Sinne eine Lösung der Gleichung (29).

Als wir beispielsweise die Aufgabe mit Achilles und der Schildkröte lösten, hatten wir der Gleichung

$$10x - x - 1000 = 0$$

zunächst die Form

$$x = 100 + \frac{x}{10}$$

gegeben und dann die einzelnen Näherungen nach der Formel

$$x_{n+1} = 100 + \frac{x_n}{10}$$

berechnet.

In der Aufgabe über die Division auf elektronischen Rechenanlagen waren wir von $ax = b$ in der neuen Gestalt

$$x = (1 - a)x + b$$

ausgegangen und hatten die Näherungsfolge rekursiv nach

$$x_{n+1} = (1 - a)x_n + b$$

errechnet.

Schließlich hatten wir zur Berechnung der k -ten Wurzel die Gleichung $x^k = a$ umgeformt zu

$$x = \frac{a + (k-1)x^k}{kx^{k-1}}$$

und danach die Näherungsfolge mit

$$x_{n+1} = \frac{a + (k-1)x_n^k}{kx_n^{k-1}}$$

definiert.

Jetzt stellen wir ein Beispiel für eine kompliziertere Gleichung vor, die ebenfalls iterativ gelöst wird.

Beispiel: Zu lösen sei die Gleichung

$$10x - 1 - \cos x = 0 \tag{30}$$

mit einem Fehler von maximal $\pm 0,001$.

Diskussion: Zuerst geben wir Gleichung (30) die Gestalt

$$x = \frac{1}{10} (1 + \cos x). \quad (31)$$

Dann wählen wir einen Anfangswert, sagen wir $x_1 = 0$, und setzen ihn in die rechte Seite ein. Den Wert $x_2 = \frac{1}{10} (1 + \cos 0) = 0,2$ nehmen wir als zweite Näherung der gesuchten Wurzel und erhalten bei abermaligem Einsetzen in (31)

$$x_3 = \frac{1}{10} (1 + \cos 0,2) \approx 0,198.$$

Weiter errechnen wir

$$x_4 = \frac{1}{10} (1 + \cos 0,198) \approx 0,198$$

und sehen, daß beide Werte mit der geforderten Genauigkeit übereinstimmen. Demzufolge ist wegen

$$x_4 = \frac{1}{10} (1 + \cos x_3)$$

die Zahl $x_3 = 0,198$ Wurzel der Gleichung (31).

Im Zusammenhang mit der beschriebenen allgemeinen Iterationsmethode entstehen einige wesentliche Fragen:

(1) Konvergiert die nach der Iterationsvorschrift berechnete Folge x_1, \dots, x_n, \dots in jedem Falle gegen eine Zahl ξ ?

(2) Wenn dies der Fall ist, wenn also $\lim_{n \rightarrow \infty} x_n = \xi$ gilt, ist dann ξ auch Wurzel der Gleichung $x = \varphi(x)$?

(3) Wie schnell konvergiert die Folge x_1, \dots, x_n, \dots gegen eine Lösung der Gleichung $x = \varphi(x)$?

Am einfachsten fällt die Antwort auf Frage (2) aus, wenn wir voraussetzen, daß $\varphi(x)$ eine stetige Funktion ¹⁾ von x ist.

Betrachten wir nämlich die definierende Relation $x_{n+1} = \varphi(x_n)$ der Folge und setzen deren Konvergenz gegen eine Zahl ξ voraus, dann strebt für wachsendes n deren linke Seite gegen ξ , die rechte aber gegen $\varphi(\xi)$. Demnach haben wir aber auch $\xi = \varphi(\xi)$, d. h., ξ ist eine Wurzel der behandelten Gleichung $x = \varphi(x)$.

Die erste Frage müssen wir in der allgemeinen Form, in der sie hier gestellt worden ist, verneinen. Wir geben dazu

¹⁾ Über Stetigkeit einer Funktion s. GÜNTHER/BEYER/GOTTWALD/WÜNSCH: Grundkurs Analysis I. BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1972.

ein Gegenbeispiel an und betrachten die Gleichung

$$x = 10^x - 2.$$

Wählen wir $x_1 = 1$, so bekommen wir nacheinander

$$x_2 = 8, \quad x_3 = 10^8 - 2, \quad x_4 = \dots, \dots$$

Die x_n wachsen also mit n über alle Grenzen, es kann deshalb keinen endlichen Grenzwert geben. Andererseits konvergiert der Iterationsprozeß jedoch für die durch Umformung gewonnene äquivalente Gleichung

$$x = \lg(x + 2)$$

sehr rasch und liefert schon nach drei Schritten das Ergebnis $\xi = 2,38$.

Wir schließen daraus, daß die Frage (1) in der vorliegenden Form nicht sachgemäß ist und etwa so lauten sollte:

(1') Wie muß die Funktion $\varphi(x)$ beschaffen sein, damit die Iterationsfolge $x_1, x_2, \dots, x_n, \dots$ konvergiert?

Zur Klärung dieser Frage soll uns eine geometrische Veranschaulichung der Iteration verhelpen.

8. Geometrische Interpretation der Iterationsmethode

Offenbar handelt es sich bei der Bestimmung einer Wurzel ξ der Gleichung $x = \varphi(x)$ darum, die Abszisse des

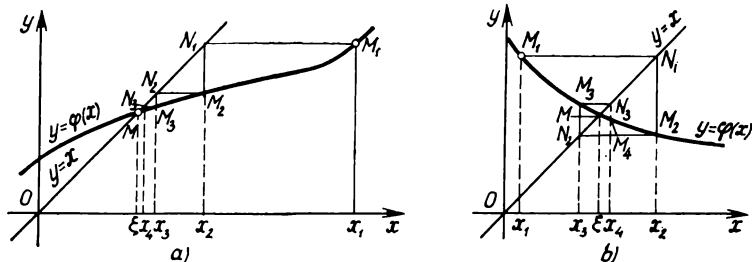


Abb. 2

Schnittpunktes M der Kurve $y = \varphi(x)$ mit der Geraden $y = x$ zu finden. Jetzt schauen wir uns einmal die Abbildung 2 an.

Wenn x_1 der Anfangswert ist, dann liegt der Punkt M_1 mit den Koordinaten $(x_1, \varphi(x_1))$ auf der Kurve $y = \varphi(x)$. Die horizontale Gerade durch diesen Punkt schneidet die Gerade

$y = x$ in einem Punkt $N_1(\varphi(x_1), \varphi(x_1)) = N_1(x_2, x_2)$, wobei wir $x_2 = \varphi(x_1)$ verwendet haben. Die Senkrechte durch N_1 schneidet die Kurve in einem Punkt $M_2(x_2, \varphi(x_2))$. Wiederholen wir den Vorgang, dann bekommen wir $N_2(\varphi(x_2), \varphi(x_2)) = N_2(x_3, x_3)$ und $M_3(x_3, \varphi(x_3))$ mit $x_3 = \varphi(x_2)$ usw.

Konvergiert der Iterationsprozeß, so nähern sich die Punkte M_1, M_2, \dots , aber auch N_1, N_2, \dots ständig dem gesuchten Schnittpunkt M .

Der geometrische Gehalt der Iterationsmethode besteht also darin, daß wir uns dem gesuchten Schnittpunkt von

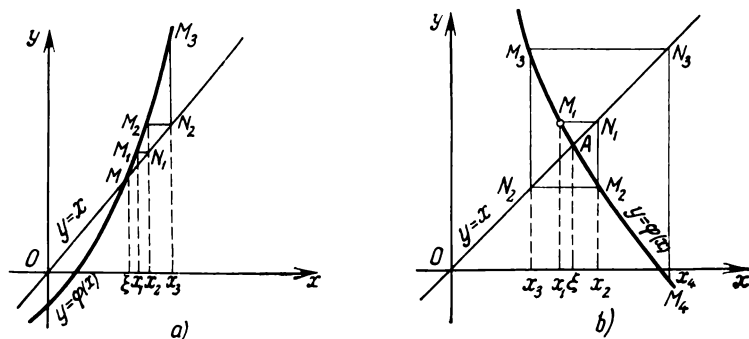


Abb. 3

Kurve und Gerade längs eines Polygonzuges nähern, dessen Ecken abwechselnd auf der Kurve und auf der Geraden liegen und dessen Seiten dabei abwechselnd horizontal und vertikal verlaufen.

Liegen Kurve und Gerade so zueinander, wie es Abb. 2a zeigt, dann entsteht ein Treppenvolygon. Bei den in Abb. 2b angedeuteten Verhältnissen entsteht dagegen ein spiralgiges Polygon.

Der angedeutete Iterationsprozeß kann natürlich auch divergieren, wie es im Falle von Achilles und der Gazelle war, d. h. zu einem Polygon führen, dessen „Treppenstufen“ oder „Spiralwindungen“ mit jedem Schritt größer werden und wo die Punktfolgen M_i bzw. N_i von M wegstreben (Abb. 3).

Der wesentliche Unterschied zwischen den Abbildungen 2 und 3 besteht in folgendem. Wir zeichnen durch M , dem Schnittpunkt der Geraden $y = x$ mit der Kurve $y = \varphi(x)$, die zur Geraden $y = x$ senkrechte Gerade. Durch diese

beiden Geraden wird die Ebene in vier Quadranten zerlegt, und die Methode konvergiert genau dann, wenn die Kurve $y = \varphi(x)$ zwischen M und M_1 ganz im Inneren des linken und rechten Quadranten verläuft. Sie divergiert dagegen, sobald die Kurve dort im oberen oder unteren Viertel liegt.

Wollen wir diese Regel anwenden, so müssen wir zunächst den Graphen der Funktion $y = \varphi(x)$ zeichnen, was aber nicht immer mit der nötigen Genauigkeit möglich ist. Wir benötigen somit ein anderes Kriterium für die Konvergenz, eine Regel, die aufgrund von Rechnungen die Frage zu klären erlaubt und keinen Bezug auf die Geometrie mehr nimmt. Ein solches Kriterium wird in Kap. 10. bereitgestellt werden. Zunächst aber müssen wir den zentralen Begriff der Kontraktivität einer Abbildung einführen und erläutern.

9. Kontraktive Abbildungen

Wir betrachten eine auf dem Intervall $[a, b]$ definierte Funktion $y = \varphi(x)$, durch die jedem Punkt x_0 dieses Intervalls ein Punkt $y_0 = \varphi(x_0)$ auf der Ordinatenachse zugeordnet wird. Will man diesen Punkt auf zeichnerischem Wege bestimmen, dann legt man durch x_0 die vertikale Gerade und bestimmt ihren Schnittpunkt mit dem Graphen von $y = \varphi(x)$. Durch diesen Schnittpunkt legt man die horizontale Gerade, welche die Ordinatenachse in y_0 schneidet (Abb. 4). Die Funktion $y = \varphi(x)$ vermittelt somit eine

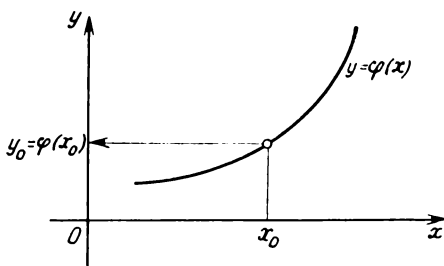


Abb. 4

eindeutige Abbildung des Intervalls $[a, b]$ in die Menge der Punkte der Ordinatenachse. Die Menge aller Bildpunkte von Punkten aus $[a, b]$ wird *Bild* dieses Intervalls genannt. So ist etwa $[4, 25]$ das Bild von $[2, 5]$ unter der Abbildung

$y = x^2$, während bei der gleichen Abbildung das Intervall $[-1, 6]$ das Bild $[0, 36]$ hat. (Man zeichne die grafische Darstellung der Funktion $y = x^2$!)

Es ist bewiesen worden, daß das Bild eines Intervalls $[a, b]$ stets wieder ein Intervall der Ordinatenachse ist, wenn die Funktion $\varphi(x)$ stetig ist. Darüberhinaus bemerkt man sofort, daß das Bildintervall von $[a, b]$ gleich $[\varphi(a), \varphi(b)]$ ist, wenn $y = \varphi(x)$ in $[a, b]$ zusätzlich monoton wachsend ist, während eine monoton fallende Funktion $y = \varphi(x)$ das Bild $[\varphi(b), \varphi(a)]$ erzeugt (Abb. 5).

Wir können aber auch leicht eine durch eine Funktion erzeugte Abbildung des Intervalls $[a, b]$ auf ein Intervall der Abszissenachse angeben. Dazu brauchen wir nur die Ordinatenachse anschließend um -90° zu drehen und mit der Abszissenachse zur Deckung zu bringen, dann kommt das Bild auf die Abszissenachse zu liegen, und wir haben unser Ziel erreicht. Die so definierte Abbildung wollen wir durch $x \rightarrow \varphi(x)$ symbolisieren. Ist $\varphi(x)$ stetig, dann ist das Bild eines Intervalls $[a, b]$ ein anderes Intervall $[a_1, b_1]$ der x -Achse.

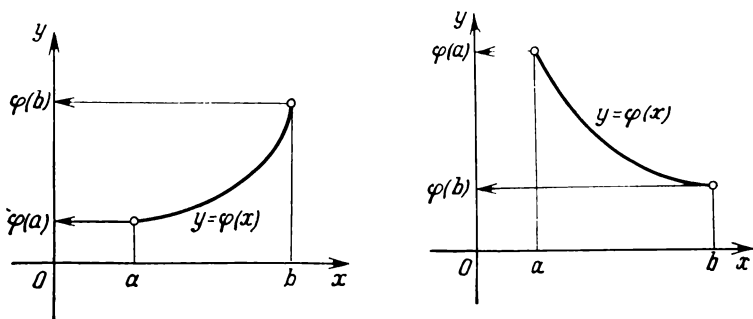


Abb. 5

Es kann nun vorkommen, daß das Bildintervall $[a_1, b_1]$ ein Teil seines Urbildes $[a, b]$ ist. Als Beispiel nennen wir die Abbildung $x \rightarrow 1 + \sqrt{x}$, die das Intervall $[0, 4]$ auf das Teilintervall $[1, 3]$ abbildet. In einem solchen Fall wollen wir sagen, daß die Funktion $\varphi(x)$ das Intervall *in sich* abbildet.

Es sei das Bild $[a_1, b_1]$ ein Teilintervall von $[a, b]$, dann hat offensichtlich jedes Teilintervall von $[a, b]$ ein Bildintervall, das ganz in $[a_1, b_1]$ gelegen ist. Demzufolge geht

$[a_1, b_1]$ selbst bei der Abbildung $\varphi(x)$ in ein Teilintervall $[a_2, b_2]$ über, dieses wiederum wird auf ein Intervall $[a_3, b_3]$ abgebildet usw. Im Endergebnis haben wir eine Folge von Intervallen $[a, b], [a_1, b_1], [a_2, b_2], \dots$ erhalten, von denen jedes im vorigen enthalten ist und als Bild des vorhergehenden Intervalls durch die Abbildung $x \rightarrow \varphi(x)$ entsteht (Symbolisch:

$$[a_{n+1}, b_{n+1}] = \varphi([a_n, b_n]).$$

Zum Beispiel bildet $x \rightarrow 1 - (x + 2)^{-1}$ das Intervall $[0, 4]$ auf das Teilintervall $[1/2, 5/6]$ ab. Dieses wiederum geht in $[3/5, 11/17]$ über und so fort. Jedes folgende Intervall ist ein Teil des vorhergehenden.

Dann sind genau zwei Fälle möglich: entweder gibt es ein Intervall $[c, d]$, das jedem Intervall der Folge angehört, oder es existiert genau ein Punkt ξ , der allen Intervallen gemeinsam ist. In diesem letzten Fall sagt man, daß sich die Intervallfolge auf den Punkt ξ zusammenzieht.

Nachfolgend werden die Bedingungen für dieses Zusammenziehen der Intervallfolge näher untersucht, und dazu wollen wir zunächst den wichtigen Begriff der kontraktiven Abbildung einführen.

Man nennt eine Abbildung $x \rightarrow \varphi(x)$, die ein Intervall $[a, b]$ in sich abbildet, *kontraktiv* oder *kontrahierend*, wenn sie den Abstand zwischen zwei beliebigen Punkten mindestens um das m -fache verkleinert, wobei $m > 1$ sein muß. Da der Abstand durch $|x_1 - x_2|$ gegeben wird, kann diese Forderung genauer formuliert werden:

Es gibt eine Zahl q mit $0 < q < 1$ und für je zwei Punkte $x_1, x_2 \in [a, b]$ gilt

$$|\varphi(x_1) - \varphi(x_2)| \leq q |x_1 - x_2| \quad (32)$$

(Hier ist also $q = 1/m$).

Bei einer kontraktiven Abbildung verkürzen sich demnach die Längen aller in $[a, b]$ enthaltenen Teilintervalle $[c, d]$ mindestens um den Faktor $m = 1/q$ (sofern die Abbildung stetig ist [Anm. d. dtsh. Red.]). Ist $[c_1, d_1]$ ein Teil von $[c, d]$, dann sind c_1 und d_1 Bildpunkte gewisser Punkte x_1, x_2 aus $[c, d]$:

$$c_1 = \varphi(x_1), \quad d_1 = \varphi(x_2)$$

und deshalb

$$|d_1 - c_1| = |\varphi(x_2) - \varphi(x_1)| \leq q |x_2 - x_1|.$$

Es ist $|x_2 - x_1| < d - c$, weil x_2, x_1 Punkte aus $[c, d]$ sind, und folglich haben wir mit $|d_1 - c_1| \leq q |d - c|$ unsere Behauptung erhalten.

Wir sind jetzt in der Lage, eine Bedingung dafür anzugeben, daß sich die oben betrachtete Intervallfolge $[a, b], [a_1, b_1], \dots$, die durch fortgesetzte Anwendung der Abbildung $\varphi(x)$ auf $[a, b]$ erzeugt wird, auf einen Punkt ξ zusammenzieht.

Bildet die Funktion $x \rightarrow \varphi(x)$ das Intervall $[a, b]$ auf ein Teilintervall $[a_1, b_1] \subseteq [a, b]$ ab, und ist sie kontraktiv, dann zieht sich die Folge der Intervalle $[a_n, b_n]$ (mit

$$[a_{n+1}, b_{n+1}] = \varphi([a_n, b_n])$$

$$[a_0, b_0] = [a, b] \text{ [Anm. d. dtsh. Red.]}$$

auf einen Punkt ξ des Intervalls $[a, b]$ zusammen.

Beweis: Wegen der Kontraktivität gilt für beliebiges n

$$|b_n - a_n| \leq q |b_{n-1} - a_{n-1}|, \text{ aber auch}$$

$$|b_{n-1} - a_{n-1}| \leq q |b_{n-2} - a_{n-2}| \text{ und folglich}$$

$$|b_n - a_n| \leq q^2 |b_{n-2} - a_{n-2}|.$$

Indem wir in dieser Form fortfahren, erhalten wir

$$|b_n - a_n| \leq q^n |b - a|.$$

Da $0 < q < 1$ vorausgesetzt ist, konvergiert die Folge $q^1, q^2, q^3, \dots, q^n, \dots$ monoton gegen null, woraus folgt, daß die Länge $|b_n - a_n|$ des Intervalls $[a_n, b_n]$ ebenfalls beliebig klein wird, wenn n gegen unendlich strebt. Dann kann es aber kein Intervall $[c, d]$ geben, das allen Intervallen der Folge gemeinsam ist, und diese Folge zieht sich auf einen Punkt zusammen.

Zum Abschluß betrachten wir eine Funktion $\varphi(x)$, für welche die Ungleichung (32) für zwei beliebige reelle Zahlen x_1 und x_2 erfüllt ist. Sie ist auf der ganzen reellen Achse kontraktiv. Wir können leicht zeigen, daß es auch dann ein Intervall gibt, das durch die Abbildung auf einen Punkt zusammengezogen wird. Da die Bedingung der Kontraktivität immer erfüllt ist, genügt es, wenn wir ein Intervall angeben, das durch $x \rightarrow \varphi(x)$ in sich abgebildet wird. Wir wählen dazu eine Zahl a und berechnen $b = \varphi(a)$. Ferner sei $q_1 < 1$ eine Zahl mit $q_1 > q$. Bilden wir jetzt $R = |b - a| / (1 - q_1)$, dann hat das Intervall $[a - R, a + R]$ ganz sicher die verlangte Eigenschaft, denn für

ein beliebiges $x \in [a - R, a + R]$ haben wir $|x - a| < R$ und wegen (32)

$$|\varphi(x) - b| = |\varphi(x) - \varphi(a)| \leq q|x - a| \leq q \cdot R.$$

Dann ist weiter

$$\begin{aligned} |\varphi(x) - a| &= |\varphi(x) - b + b - a| \leq |\varphi(x) - b| + |b - a| \\ &\leq qR + |b - a| = qR + (1 - q_1)R = \\ &= (1 + q - q_1)R < R. \end{aligned}$$

Das bedeutet, daß der Punkt $\varphi(x)$ ebenfalls im betrachteten Intervall liegt, und das war gerade unsere Behauptung.

10. Kontraktive Abbildung und Iterationsverfahren

Jetzt kehren wir zum Iterationsverfahren zurück, das ja gerade zur Lösung von Gleichungen der Form $x = \varphi(x)$ verwendet wird. Wenn ξ Wurzel dieser Gleichung ist, dann bleibt der Punkt ξ bei der Abbildung $x \rightarrow \varphi(x)$ an seinem Platz, denn es gilt ja $\xi = \varphi(\xi)$. Das heißt aber, daß *die Aufgabe, eine Gleichung zu lösen, gleichwertig damit ist, einen sogenannten Fixpunkt der Abbildung $x \rightarrow \varphi(x)$ zu finden.*

Ist die Abbildung $\varphi(x)$ auf dem Intervall $[a, b]$ kontraktiv, dann gibt es dort immer einen Fixpunkt. Das sieht man folgendermaßen ein:

Eine über $[a, b]$ kontraktive Abbildung $\varphi(x)$ erzeugt eine Intervallfolge $[a_1, b_1], [a_2, b_2], \dots, [a_n, b_n], \dots$ aus $[a, b]$, die sich auf einen Punkt ξ zusammenzieht, der eindeutig bestimmt ist und in allen Intervallen $[a_n, b_n]$ liegt. Dieser Punkt ist Fixpunkt der Abbildung $\varphi(x)$.

In der Tat bildet $\varphi(x)$ jedes Intervall $[a_n, b_n]$ in ein Teilintervall $[a_{n+1}, b_{n+1}]$ ab. Demnach liegt das Bild $\varphi(c)$ eines beliebigen Punktes c aus $[a_n, b_n]$ in diesem Teilintervall $[a_{n+1}, b_{n+1}]$ und insbesondere in $[a_n, b_n]$ selbst. Weil nun ξ *allen* Intervallen $[a_n, b_n]$ angehört, muß das auch auf sein Bild $\varphi(\xi)$ zutreffen. Es gibt aber nur einen Punkt, der in allen Intervallen liegt. Folglich ist $\varphi(\xi) = \xi$, d. h., ξ ist Fixpunkt von $\varphi(x)$.

Ebenso leicht zeigt man, daß es in diesem Falle immer nur einen Fixpunkt in $[a, b]$ geben kann. Wäre nämlich η ein weiterer Fixpunkt im Intervall, dann müßte gelten

$$|\xi - \eta| = |\varphi(\xi) - \varphi(\eta)| \leq q|\xi - \eta|.$$

Da aber $0 < q < 1$ ist, kann das nur richtig sein, wenn

$$|\xi - \eta| = 0, \text{ d. h. } \xi = \eta \text{ ist.}$$

Jetzt können wir endgültig eine *hinreichende Bedingung* für die Konvergenz des Iterationsverfahrens angeben.

Die Funktion $\varphi(x)$ möge eine kontraktive Abbildung des Intervalls $[a, b]$ bewirken. Dann konvergiert die nach der Vorschrift $x_{n+1} = \varphi(x_n)$ mit einem beliebigen Anfangswert x_0 aus diesem Intervall gebildete Folge $x_0, x_1, \dots, x_n, \dots$ gegen die einzige in diesem Intervall gelegene Wurzel ξ der Gleichung $x = \varphi(x)$.

Beweis: Wenn $[a_n, b_n], n = 1, 2, \dots$, die Intervallfolge ist, die durch sukzessive Anwendung der Abbildung $x \rightarrow \varphi(x)$ auf das Intervall $[a, b]$ entsteht, dann liegt x_0 in $[a, b]$, x_1 in $[a_1, b_1]$ und allgemein x_n in $[a_n, b_n]$. Die Folge muß somit gegen ξ konvergieren, denn die Länge $|b_n - a_n|$ der Intervalle konvergiert gegen null.

Wir heben nochmals hervor, daß der Beweis gleichzeitig zeigt, daß x_0 in $[a, b]$ ganz beliebig gewählt werden kann.

Jetzt wenden wir uns der dritten der aufgeworfenen Probleme zu und wollen Aussagen über die *Konvergenzgeschwindigkeit* der Folge x_0, x_1, \dots, x_n , gegen ξ herleiten.

Wegen $\xi = \varphi(\xi)$ gilt

$$|\varphi(c) - \xi| = |\varphi(c) - \varphi\xi| \leq q|c - \xi| \quad (33)$$

für einen beliebigen Punkt c aus $[a, b]$.

Wir wenden (33) jetzt auf die Punkte x_n ($n = 0, 1, 2, \dots$) an und erhalten wegen $x_{n+1} = \varphi(x_n)$

$$|x_n - \xi| = |\varphi(x_{n-1}) - \xi| \leq q|x_{n-1} - \xi| \text{ und durch}$$

fortgesetzte Anwendung dieser Abschätzung schließlich

$$|x_n - \xi| \leq q|x_{n-1} - \xi| \leq q^2|x_{n-2} - \xi| \leq \dots \leq q^n|x_0 - \xi|.$$

Daraus lesen wir ab, daß die Folge $|x_n - \xi|$ mit wachsendem n mindestens mit der Geschwindigkeit einer geometrischen Folge mit dem Quotienten $q < 1$ gegen null konvergiert.

Wir geben nun einige Beispiele für die soeben bewiesene Aussage.

Beispiel 1: Kann die Iterationsmethode auf die Lösung der Gleichung $x = 1/(4 + x^2)$ angewendet werden? (34)

Diskussion: In (34) ist $\varphi(x) = 1/(4+x^2)$. Für beliebige x_1, x_2 haben wir

$$\begin{aligned} |\varphi(x_2) - \varphi(x_1)| &= \left| \frac{1}{4+x_2^2} - \frac{1}{4+x_1^2} \right| = \\ &= \frac{|x_1^2 - x_2^2|}{(4+x_2^2)(4+x_1^2)} = \frac{|x_1+x_2|}{|4+x_1^2||4+x_2^2|} |x_2 - x_1|. \end{aligned}$$

Ferner gilt $|x| = \frac{1}{2} \sqrt{4x^2} \leq \frac{4+x^2}{4}$ wegen der Ungleichung zwischen geometrischem und arithmetischem Mittel. Also gilt

$$\begin{aligned} |x_1+x_2| &\leq |x_1| + |x_2| \leq \frac{(4+x_1^2) + (4+x_2^2)}{4} = \\ &= 2 + \frac{x_1^2+x_2^2}{4} \leq 2 + \frac{x_1^2+x_2^2}{2} + \frac{x_1^2x_2^2}{8} = \\ &= \frac{1}{8} (4+x_1^2)(4+x_2^2). \end{aligned}$$

Wir haben damit gezeigt, daß für beliebige reelle Zahlen x_1 und x_2 stets die Ungleichung

$$\frac{|x_1+x_2|}{(4+x_1^2)(4+x_2^2)} \leq \frac{1}{8}$$

richtig ist, und daraus folgt sofort

$$|\varphi(x_2) - \varphi(x_1)| \leq \frac{1}{8} |x_2 - x_1|.$$

Das beweist aber, daß die Abbildung $x \rightarrow \varphi(x)$ auf dem gesamten Zahlenstrahl kontraktiv ist.

Wir wissen bereits, daß es in diesem Falle sicher ein Intervall gibt, das bei dieser Abbildung in sich übergehen wird. Um dasselbe zu finden, setzen wir $a = 0$ und berechnen $b = \varphi(a)$, d. h. $b = 1/4 = \varphi(0)$. Wir sahen oben, daß $q = 1/8$ ist und wählen deshalb $q_1 = 1/4$, woraus sich $R = \frac{|b-a|}{1-q_1} = \frac{1}{3}$ ergibt. Das Intervall $[-1/3, 1/3]$ hat demnach

die gewünschte Selbstabbildungseigenschaft, so daß es in diesem Intervall genau eine Wurzel der Gleichung (34) gibt, die wir nun durch Iteration bestimmen wollen.

Als beliebige Ausgangsnäherung aus dem Intervall nehmen wir etwa $x_0 = 0$ und rechnen nacheinander

aus:

$$x_1 = \frac{1}{4} = 0,25,$$

$$x_2 = \frac{1}{4 + 0,25^2} = \frac{1}{4,0625} = 0,2461,$$

$$x_3 = \frac{1}{4 + 0,2461^2} = \frac{1}{4,0605} = 0,2463,$$

$$x_4 = \frac{1}{4 + 0,2463^2} = \frac{1}{4,0605} = 0,2463.$$

Die Werte x_3 und x_4 stimmen bis auf vier Stellen überein. Folglich ist die im Intervall $[-1/3, 1/3]$ gelegene Wurzel von (34) gleich $0,2463 \pm 0,001$.

Das ist aber überhaupt die einzige Wurzel dieser Gleichung, weil $\varphi(x)$ auf der gesamten Zahlengeraden kontraktiv ist.

Beispiel 2: Untersuchen Sie die Anwendbarkeit der Iterationsmethode auf die Gleichung

$$x = 1 + \sqrt[3]{x}$$

im Intervall $[-1, 8]!$

Diskussion: Hier ist $\varphi(x) = 1 + \sqrt[3]{x}$, und wegen $\varphi(-1) = 0$ und $\varphi(8) = 3$ wird das Intervall $[-1, 8]$ in sich abgebildet. Kontraktiv ist die Abbildung indessen dort nicht, denn für $x_1 = -0,008$ und $x_2 = 0,008$ wird

$$\begin{aligned} |\varphi(x_2) - \varphi(x_1)| &= \\ &= |\sqrt[3]{0,008} - \sqrt[3]{-0,008}| = \\ &= 0,4 > |x_2 - x_1|. \end{aligned}$$

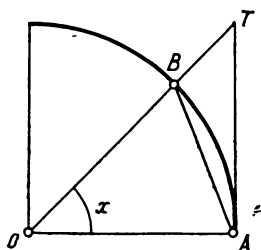


Abb. 6

Beim Beweis der Kontraktivität haben wir in *Beispiel 1* die Ungleichung $\sqrt{ab} \leq (a+b)/2$ angewendet. Jetzt wollen wir noch einige weitere Ungleichungen einführen, die bei analogen Beweisen immer wieder nützlich sind.

Zunächst beweisen wir die Gültigkeit der Ungleichung

$$\sin x < x < \tan x \quad (35)$$

für alle $x > 0$.

Dazu bemerken wir, daß in Abb. 6 die Fläche S_{OAB} des Sektors OAB mit dem Winkel x größer als die Dreiecksfläche $S_{\Delta OAB}$ und

kleiner als die Fläche $S_{\Delta OAT}$ des Dreiecks ΔOAT ist :

$$S_{\Delta OAB} < S_{OAB} < S_{\Delta OAT}.$$

Es gilt aber auch

$$S_{OAB} = \frac{R^2 \sin x}{2}, \quad S_{\Delta OAT} = \frac{R^2 \tan x}{2},$$

wo R der Kreisradius ist. Die Fläche des Sektors selbst ist gleich $R^2 x/2$ (den Winkel x messen wir im Bogenmaß). Demnach gilt

$$\frac{R^2 \sin x}{2} < \frac{R^2 x}{2} < \frac{R^2 \tan x}{2},$$

woraus durch Multiplikation mit $\frac{2}{R^2} > 0$ die Behauptung folgt.

Aus der Ungleichung (35) lesen wir auch ab, daß für $0 < x < 1$ stets $x < \arcsin x$ und für $0 < x$ stets $x > \arctan x$ gilt.

Ferner notieren wir noch die Ungleichungen

$$e^x > 1 + x \quad \text{für } x > 0 \text{ und} \\ \ln(1 + x) < x \quad \text{für } 0 < x < 1,$$

deren Beweis etwas schwieriger ist.

Beispiel 3: Klären Sie die Frage der Konvergenz des Iterationsverfahrens für die Gleichung

$$x = 1 + \frac{1}{2} \arctan x. \quad (36)$$

Diskussion: Diese Gleichung kann wegen $1 + \frac{1}{2} \arctan x > 0$ für alle x nur positive Wurzeln haben. Es ist hier nun

$$\varphi(x) = 1 + \frac{1}{2} \arctan x \quad (37)$$

und deshalb

$$\begin{aligned} |\varphi(x_2) - \varphi(x_1)| &= \left| \left(1 + \frac{1}{2} \arctan x_2 \right) - \right. \\ &\left. - \left(1 + \frac{1}{2} \arctan x_1 \right) \right| = \frac{1}{2} |\arctan x_2 - \arctan x_1|. \end{aligned}$$

Für $x_1 \geq 0, x_2 \geq 0$ gilt stets

$$\arctan x_2 - \arctan x_1 = \arctan \frac{x_2 - x_1}{1 + x_1 x_2}$$

und demnach auch

$$\begin{aligned} |\varphi(x_2) - \varphi(x_1)| &= \frac{1}{2} \left| \arctan \frac{x_2 - x_1}{1 + x_1 x_2} \right| < \\ &< \frac{1}{2} \left| \frac{x_2 - x_1}{1 + x_1 x_2} \right| < \frac{1}{2} |x_2 - x_1|. \end{aligned}$$

Also ist die Abbildung (37) auf der Halbachse $[0, \infty]$ kontraktiv und bildet das Intervall $[0, \sqrt[3]{3}]$ auf das Teilintervall $\left[1, 1 + \frac{\pi}{6}\right]$ ab. Deshalb gibt es genau eine Wurzel der Gleichung (36) in diesem Intervall. Zur Berechnung dieser Wurzel setzen wir nun $x_1 = 1$ und bekommen nacheinander

$$x_2 = 1 + \frac{1}{2} \arctan 1 = 1 + \frac{\pi}{8} \simeq 1,39,$$

$$x_3 = 1 + \frac{1}{2} \arctan 1,39 = 1,474,$$

$$x_4 = 1 + \frac{1}{2} \arctan 1,474 = 1,487,$$

$$x_5 = 1 + \frac{1}{2} \arctan 1,487 = 1,489,$$

$$x_6 = 1 + \frac{1}{2} \arctan 1,489 = 1,490,$$

$$x_7 = 1 + \frac{1}{2} \arctan 1,490 = 1,490.$$

Innerhalb der mitgeführten Stellenzahl stimmen x_6 und x_7 überein, und die gesuchte Wurzel ist bis auf einen Fehler von $\pm 0,001$ gleich 1,490. Die Abbildung $x \rightarrow \varphi(x)$ ist auf der gesamten Halbachse kontraktiv, woraus folgt, daß (36) keine weiteren Wurzeln haben kann.

Auch wenn die Gleichung $x = \varphi(x)$ nicht mit Iteration gelöst werden kann, läßt sich durch Umformung häufig eine Gleichung herleiten, auf die die Methode anwendbar ist. Wir betrachten als Beispiel hierfür die Gleichung

$$x = x^3 - 2. \quad (38)$$

Wegen $\varphi(1) = -1 < 1$, $\varphi(2) = 6 > 2$ besitzt sie im Intervall $[1, 2]$ sicher eine Wurzel. Die Abbildung $x^3 - 2$ ist dort aber nicht kontraktiv, sie bildet nicht einmal das Intervall in sich ab.

Jetzt formen wir (38) folgendermaßen um:

$$x = \sqrt[3]{x+2},$$

und nun ist $\psi(x) = \sqrt[3]{x+2}$ und deshalb

$$\begin{aligned} |\psi(x_2) - \psi(x_1)| &= \left| \sqrt[3]{x_2+2} - \sqrt[3]{x_1+2} \right| = \\ &= \left| \frac{x_2 - x_1}{\sqrt[3]{(x_2+2)^2} + \sqrt[3]{(x_1+2)(x_2+2)} + \sqrt[3]{(x_1+2)^2}} \right|. \end{aligned}$$

Im betrachteten Intervall haben wir $x_1 \geq 1$, $x_2 \geq 1$, und deshalb gilt dort

$$|\psi(x_2) - \psi(x_1)| \leq \frac{1}{3\sqrt[3]{9}} |x_2 - x_1|,$$

womit die Kontraktivität von $\psi(x)$ im Intervall $[1, 2]$ gezeigt ist.

Wir wählen $x_1 = 1$ und führen die Iterationsmethode durch, wobei wir

$$x_2 = 3 = 1,442,$$

$$x_3 = 3,442 = 1,510,$$

$$x_4 = 3,510 = 1,520,$$

$$x_5 = 3,521 = 1,521,$$

$$x_6 = 3,521 = 1,521$$

als Folge der Näherungen ausrechnen. Wie wir sehen, stimmt $x_5 = 1,521$ bis auf einen Fehler von $\pm 0,001$ mit der einzigen im Intervall $[1, 2]$ vorhandenen Wurzel von (38) überein. Andere reelle Wurzeln hat diese Gleichung aber überhaupt nicht.

Die bisher gewonnene Charakterisierung der Konvergenz der Iterationsmethode ist für die Anwendung nicht besonders geeignet, weil sie die Gewinnung genügend scharfer Abschätzungen voraussetzt. Wir werden später im Kapitel 21 eine andere Form der Konvergenzbedingung gewinnen, deren Verwendung den Konvergenzbeweis der Methode beträchtlich vereinfachen kann.

11. Die Sekantenmethode (regula falsi)

Es gibt eine Reihe von Näherungsverfahren zur Lösung von Gleichungen, die schon lange im Gebrauch sind und eine wesentlich speziellere Gestalt haben. Viele von ihnen lassen sich relativ leicht als Spezialfälle des bisher betrachteten allgemeinen Iterationsverfahrens interpretieren. Wir werden uns in diesem Sinne jetzt zunächst mit dem sogenannten Sekanten-Näherungsverfahren — der „regula falsi“, wie es auch vielfach genannt wird — beschäftigen.

Wir gehen wieder aus von einer Gleichung $f(x) = 0$ und bemerken zunächst, daß die Aufgabe, eine solche Gleichung zu lösen, gleichwertig ist mit der Bestimmung des Schnittpunktes des Graphen von $f(x)$ mit der Abszissenachse.

Wir setzen voraus, daß die Funktion $y = f(x)$ stetig ist

und in den beiden Punkten $x = a$ und $x = b$ Werte mit entgegengesetztem Vorzeichen annimmt. Dann besitzt $y = f(x)$ im Intervall $[a, b]$ mindestens eine Nullstelle.

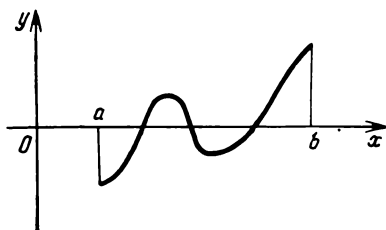


Abb. 7

Die grafische Darstellung der Funktion schneidet die Abszissenachse in diesem Intervall wenigstens einmal. Es kann aber natürlich auch mehrere Nullstellen geben (Abb. 7).

Nur wenn die Funktion $y = f(x)$ in $[a, b]$ außerdem noch monoton ist, schneidet ihre grafische Darstellung die x -Achse in genau einem Punkt ξ . Um denselben annähernd zu finden, ersetzen wir den Kurvenbogen $y = f(x)$ in $[a, b]$ durch die Sekante MN und bestimmen deren Schnittpunkt T mit der Abszissenachse (Abb. 8).

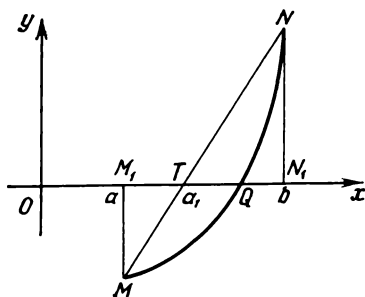


Abb. 8

Dazu betrachten wir die beiden ähnlichen Dreiecke MM_1T und NN_1T und folgern aus der Ähnlichkeit die Gleichung

$$\frac{\overline{M_1T}}{\overline{MM_1}} = \frac{\overline{TN_1}}{\overline{NN_1}}.$$

Aus der Zeichnung entnehmen wir $\overline{M_1T} = a_1 - a$, $\overline{TN_1} = b - a_1$, $\overline{MM_1} = -f(a)$ und $\overline{NN_1} = f(b)$, wenn a_1 die Abszisse von T ist. Also gilt

$$\frac{a_1 - a}{-f(a)} = \frac{b - a_1}{f(b)}$$

oder

$$a_1 = \frac{af(b) - bf(a)}{f(b) - f(a)}.$$

Die letzte Gleichung können wir auch so aufschreiben:

$$a_1 = b - f(b) \frac{b-a}{f(b) - f(a)} \text{ bzw.} \quad (39)$$

$$a_1 = a - f(a) \frac{b-a}{f(b) - f(a)}. \quad (40)$$

(Beweisen Sie, daß beide Formeln äquivalent sind!)

Die Zahl a_1 ist dann offensichtlich ein Näherungswert für die zwischen a und b gelegenen Wurzeln der Gleichung $f(x) = 0$.

Weil $f(a)$ und $f(b)$ verschiedene Vorzeichen haben, muß entweder das Vorzeichen von $f(a)$ oder das von $f(b)$ vom Vorzeichen der Zahl $f(a_1)$ verschieden sein (sofern nicht a_1 zufällig bereits eine Nullstelle ist [Anm. d. dtsh. Red.]). Wenn die Funktion beispielsweise in a und a_1 verschiedenes Vorzeichen hat, dann wendet man Formel (39) auf dieses Intervall $[a, a_1]$ an und bekommt die folgende Näherung

$$a_2 = a_1 - f(a_1) \frac{a_1 - a}{f(a_1) - f(a)} \quad (41)$$

für die gesuchte Wurzel. Haben dagegen $f(a_1)$ und $f(b)$ unterschiedliche Vorzeichen, so ist Formel (40) anzuwenden, und es ergibt sich

$$a_2 = a_1 - f(a_1) \frac{b - a_1}{f(b) - f(a_1)}. \quad (42)$$

Hat man a_2 ermittelt, so wird Formel (39) auf $[a, a_2]$ oder (40) auf $[a_2, b]$ angewendet, womit man a_3 als dritte Näherung erhält. Ganz allgemein berechnen wir die $(n + 1)$ -te Näherung aus der n -ten nach der Formel (43) oder (44):

$$a_{n+1} = a_n - f(a_n) \frac{a_n - a}{f(a_n) - f(a)}, \quad (43)$$

$$a_{n+1} = a_n - f(a_n) \frac{b - a_n}{f(b) - f(a_n)}. \quad (44)$$

Wir haben zwei Formeln erhalten und werden jetzt überlegen, unter welchen Voraussetzungen die eine oder andere sachgemäß ist.

Die Kurve sei *von unten gesehen konvex*, dann ist derjenige Intervallendpunkt M oder N zu nehmen, in dem die

Funktion *positiv* ist. Ist die Kurve dagegen *von unten gesehen konkav*, dann muß gerade der Endpunkt *negativer* Ordinate genommen werden. Die vier möglichen Fälle sind in Abb. 9

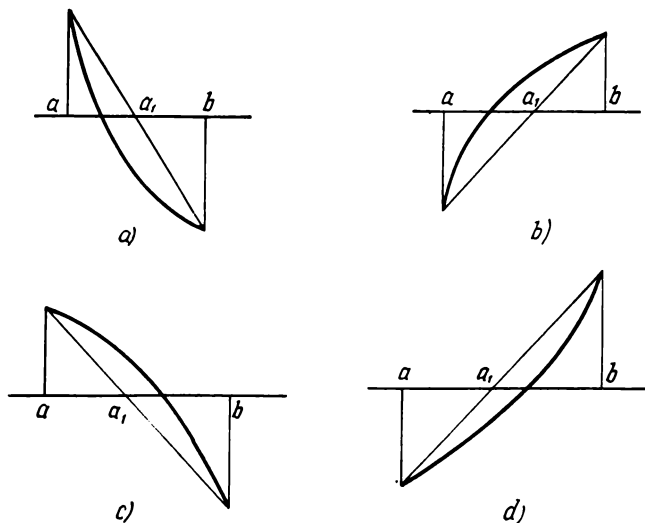


Abb. 9

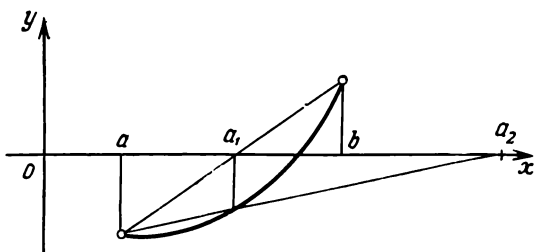


Abb. 10

dargestellt, und man liest die geometrisch anschauliche Aussage ab:

Die Funktion $f(x)$ sei auf dem Intervall stetig und monoton, habe im ganzen Intervall gleichbleibende Krümmungsrichtung und in den Intervallendpunkten entgegengesetzte Vorzeichen, dann liefert die Sekantenmethode bei richtiger Wahl der Formel in jedem Schritt eine Punktfolge, die gegen die Lösung der Gleichung $f(x) = 0$ konvergiert.

Wird die falsche Formel genommen, dann führt das dazu, daß der Punkt a_2 außerhalb des Intervalls $[a, b]$ liegt, wie es Abbildung 10 zeigt.

Wir wollen jetzt sehen, inwiefern die Sekantenmethode ein Spezialfall des allgemeinen Iterationsverfahrens ist,

Wenn $f(a) \neq 0$ ist, dann ist die Gleichung $f(x) = 0$ äquivalent zur Gleichung

$$x = x - f(x) \frac{x-a}{f(x)-f(a)}. \quad (45)$$

Beweis:

a) Es sei $f(\xi) = 0$, dann ist

$$\xi = \xi - f(\xi) \frac{\xi-a}{f(\xi)-f(a)} \quad (46)$$

richtig und (45) offenbar erfüllt.

b) Wenn andererseits $\xi \neq a$ eine Lösung von (46) ist, so gilt $f(\xi) = 0$.

$$(0 = -f(\xi) \frac{\xi-a}{f(\xi)-f(a)} \text{ und wegen } \xi \neq a \rightarrow f\xi = 0,$$

was wir beweisen wollten [Anm. d. dtsh. Red.]

Gleichung (45) ist aber genau von der Gestalt $x = \varphi(x)$ mit

$$\varphi(x) = x - f(x) \frac{x-a}{f(x)-f(a)} = \frac{af(x)-xf(a)}{f(x)-f(a)}.$$

Wir wählen $x_0 = b$ und führen die Iterationsmethode durch. Dabei erhalten wir genau dieselbe Zahlenfolge $a_1, a_2, \dots, \dots, a_n \dots$ wie mit dem Sekantenverfahren:

$$a_{n+1} = a_n - f(a_n) \frac{a_n-a}{f(a_n)-f(a)}.$$

Beispiel: Wir lösen die Gleichung

$$x^3 + 3x - 1 = 0 \quad (47)$$

nach dem Sekantenverfahren.

Diskussion: Hier ist $f(x) = x^3 + 3x - 1 = 0$ und $f(0) = -1, f(1) = 3$, so daß (47) im Intervall $[0, 1]$ mindestens eine Lösung haben muß. Der Graph der Funktion ist im Intervall von unten gesehen konvex, wie eine Skizze zeigt, und wir müssen deshalb mit der Formel (39) arbeiten:

$$x_1 = b - f(b) \frac{b-a}{f(b)-f(a)} = 1 - 3 \cdot \frac{1-0}{3-(-1)} = 0,25.$$

Die gleiche Formel liefert auch die zweite und alle weiteren Näherungen

$$x_2 = b - f(b) \frac{b - x_1}{f(b) - f(x_1)} = 1 - 3 \frac{1 - 0,25}{3 + 0,23} = 0,31,$$

$$x_3 = 1 - 3 \cdot \frac{1 - 0,31}{3 + 0,040} = 0,319,$$

$$x_4 = 1 - 3 \cdot \frac{1 - 0,319}{3 + 0,010} = 0,322,$$

$$x_5 = 1 - 3 \cdot \frac{1 - 0,322}{3 + 0,0006} = 0,322.$$

Die im Intervall $[0, 1]$ gelegene Wurzel ist also $0,322 \pm 0,001$.

12. Ein verbessertes Sekantenverfahren

Wenn die *regula falsi* konvergiert, dann stimmt die Konvergenzgeschwindigkeit im allgemeinen mit der des Iterationsverfahrens überein — der Fehler verkleinert sich wie die Glieder einer geometrischen Folge.

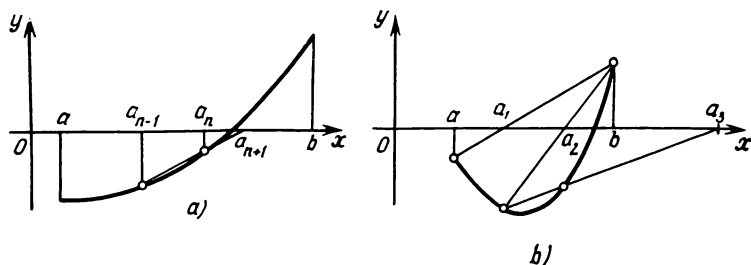


Abb. 11

Es gibt aber eine verbesserte Variante des eben beschriebenen Verfahrens, die durch eine wesentlich größere Konvergenzgeschwindigkeit gekennzeichnet ist.

Bei der gewöhnlichen *regula falsi* wird in jedem Schritt einer der beiden Intervallendpunkte von $[a, b]$ mit dem gerade berechneten Näherungswert kombiniert. Stattdessen kann man aber zwei aufeinanderfolgende Näherungen benutzen, denn sie liegen gewiß näher an der gesuchten Lösung als die Intervallendpunkte. Die entsprechende Formel ist (vgl. Abb. 11a)

$$a_{n+1} = a_n - f(a_n) \frac{a_n - a_{n-1}}{f(a_n) - f(a_{n-1})}. \quad (48)$$

Dabei berechnet sich a_1 zunächst nach Formel (39) und a_2 je nach Vorzeichen von $f(a)$, $f(b)$ und $f(a_1)$ nach (41) oder (42): Ist $f(a) < 0$, $f(b) > 0$, so wird für $f(a_1) < 0$ Formel (42) und für $f(a_1) > 0$ (41) genommen.

Liegt der nach Formel (48) berechnete Wert a_3 zufällig außerhalb des Intervalls $[a, b]$, dann muß man im nächsten Schritt das ihm benachbarte Intervallende an seiner Stelle verwenden (Abb. 11b).

Die Konvergenzgeschwindigkeit der verbesserten Sekantenmethode ist, wie man zeigen kann, wesentlich größer als die der gewöhnlichen *regula falsi*. Bezeichnet nämlich ξ wieder die Wurzel der Gleichung $f(x) = 0$, dann ist

$$|a_{n+1} - \xi| < C |a_n - \xi|^t, \quad (49)$$

wobei

$$t = \frac{1 + \sqrt{5}}{2} \approx 1,618 \text{ ist.}$$

Zur Illustration lösen wir die Gleichung

$$x^3 + 3x - 1 = 0$$

von vorhin mit der verbesserten Methode. Die Anfangsnäherung $a_1 = 0,25$ und $a_2 = 0,31$ übernehmen wir von der einfachen Sekantenmethode. Die nächste Näherung ergibt sich dann nach (48) zu

$$\begin{aligned} a_3 &= a_2 - f(a_2) \frac{a_2 - a_1}{f(a_2) - f(a_1)} = \\ &= 0,31 + 0,040 \cdot \frac{0,31 - 0,25}{-0,040 + 0,234} = 0,3223. \end{aligned}$$

Wir haben $f(0,3223) = 0,0004$, und demnach ist $x = 0,3223$ bis auf mindestens 3 Stellen nach dem Komma mit der gesuchten Wurzel identisch.

13. Die Ableitung eines Polynoms

Bei der Lösung der Gleichung $f(x) = 0$ mit der Methode der Iteration hängt sehr viel davon ab, ob man eine geeignete Umformung auf die Gestalt $x = \varphi(x)$ findet. In sehr vielen Fällen erweist sich die von Newton entwickelte allgemeine Methode als eine der besten. Sie gründet sich auf den allgemeinen Begriff der *Ableitung* einer Funktion.

In diesem Paragraphen wollen wir zunächst herausarbeiten, was man unter der Ableitung eines Polynoms zu ver-

stehen hat. Das wird uns dann in die Lage versetzen, das Newtonsche Verfahren auf die Lösung algebraischer Gleichungen, d. h. Gleichungen der Form

$$a_0x^h + a_1x^{h-1} + \dots + a_h = 0 \quad (50)$$

anzuwenden.

Es sei $f(x) = a_0x^h + \dots + a_h$ ein beliebiges Polynom. Wir betrachten das Polynom $f(x + \alpha)$, d. h. den Ausdruck

$$f(x + \alpha) = a_0(x + \alpha)^h + a_1(x + \alpha)^{h-1} + \dots + a_h. \quad (51)$$

Lösen wir hier die Klammern auf, dann erhalten wir einige Terme, in denen α nicht vorkommt, während in anderen die erste, zweite, dritte bis k -te Potenz von α vorkommt. Ordnen wir die Terme nach Potenzen von α , so ergibt sich

$$f(x + \alpha) = f_0(x) + f_1(x)\alpha + \dots + f_h(x)\alpha^h. \quad (52)$$

(Die höchste Potenz von α ist die k -te, denn $f(x)$ hat den Grad k .) Man überlegt sich leicht, daß die Glieder dieser Folge ebenfalls Polynome in x sind.

Beispiel:

$$f(x) = 2x^3 - 3x^2 + 6x - 1.$$

Diskussion: Hier ist

$$\begin{aligned} f(x + \alpha) &= 2(x + \alpha)^3 - 3(x + \alpha)^2 + 6(x + \alpha) - 1 = \\ &= 2(x^3 + 3x^2\alpha + 3x\alpha^2 + \alpha^3) - 3(x^2 + 2x\alpha + \alpha^2) \\ &+ 6(x + \alpha) - 1 = (2x^3 - 3x^2 + 6x - 1) \\ &+ (6x^2 - 6x + 6)\alpha + (6x - 3)\alpha^2 + 2\alpha^3 \end{aligned}$$

und folglich

$$\begin{aligned} f_0(x) &= 2x^3 - 3x^2 + 6x - 1, \\ f_1(x) &= 6x^2 - 6x + 6, \\ f_2(x) &= 6x - 3, \\ f_3(x) &= 2. \end{aligned}$$

Wir sehen, daß $f_0(x)$ und $f(x)$ identisch sind. Das ist nicht zufällig. Setzen wir nämlich in (52) $\alpha = 0$, dann erhalten wir $f(x) = f_0(x)$.

Verweilen wir jetzt ein wenig beim nächsten Glied $f_1(x)\alpha$. Der Koeffizient von α in (52), d. h. das Polynom $f_1(x)$, heißt *Ableitung des Polynoms $f(x)$* . Die Ableitung des

Polynoms $2x^3 - 3x^2 + 6x - 1$ ist zum Beispiel gleich $6x^2 - 6x + 6$. Man schreibt gewöhnlich $f'(x)$ für die erste Ableitung von $f(x)$.

Die Ableitung $f'(x)$ des Polynoms $f(x)$ ist der Koeffizient von a in der Entwicklung des Polynoms $f(x+a)$ nach Potenzen von a .

Formel (52) beginnt also mit der neuen Bezeichnung so:

$$f(x+a) = f(x) + f'(x)a + \dots \quad (53)$$

Die drei Punkte stehen wieder, wie schon früher, für Glieder mit a^2, a^3, \dots, a^k .

In unserem Beispiel sieht das so aus:

$$\begin{aligned} 2(x+a)^3 - 3(x+a)^2 + 6(x+a) - 1 &= \\ &= 2x^3 - 3x^2 + 6x - 1 + (6x^2 - 6x + 6)a + \dots \end{aligned}$$

Nachdem wir den Begriff der Ableitung eines Polynoms definiert haben, suchen wir nach einer einfacheren Möglichkeit der Berechnung. Dazu betrachten wir ein Polynom

$$f(x+a) = a_0(x+a)^k + a_1(x+a)^{k-1} + \dots + a_{k-1}(x+a) + a_k$$

und lösen alle vorkommenden Klammern unter Benutzung der *binomischen Formel* $(x+a)^m = x^m + mx^{m-1}a +$ (vgl. Kap. 6) auf

$$\begin{aligned} f(x+a) &= a_0(x^k + kx^{k-1}a + \dots) + \\ &\quad + a_1[x^{k-1} + (k-1)x^{k-2}a + \dots] + \dots \\ &\quad \dots + a_{k-1}(x+a) + a_k = \\ &= a_0x^k + a_1x^{k-1} + \dots + a_k + a[ka_0x^{k-1} + \\ &\quad + (k-1)a_1x^{k-2} + \dots + a_{k-1}] + \dots \end{aligned}$$

Ein Vergleich der letzten Formel mit (53) zeigt uns die Richtigkeit der folgenden Aussage:

Die Ableitung eines Polynoms

$$f(x) = a_0x^k + a_1x^{k-1} + \dots + a_{k-1}x + a_k \quad (54)$$

hat die Gestalt

$$f'(x) = ka_0x^{k-1} + (k-1)a_1x^{k-2} + \dots + a_{k-1}. \quad (55)$$

Beispielsweise ist die Ableitung von $f(x) = 6x^7 + 8x^3 - 3x^2 - 1$ gleich $f'(x) = 42x^6 + 24x^2 - 6x$.

14. Das Newtonsche Verfahren für algebraische Gleichungen

Wir betrachten in diesem Abschnitt sogenannte *algebraische Gleichungen*

$$a_0x^k + a_1x^{k-1} + \dots + a_k = 0. \quad (56)$$

Wir gehen wie früher von einem irgendwie beschafften *Anfangswert* x_1 einer Wurzel von (56) aus und suchen diesen iterativ zu verbessern. Wie schon in früheren Abschnitten bezeichnen wir den Fehler von x_1 mit α_1 , so daß $\xi = x_1 + \alpha_1$ eine gesuchte Wurzel darstellt. Dann gilt also die Gleichheit

$$a_0(x + \alpha_1)^k + a_1(x_1 + \alpha_1)^{k-1} + \dots + a_k = 0 \quad (57)$$

oder kurz

$$f(x_1 + \alpha_1) = 0.$$

Nach Formel (53) ist $f(x_1 + \alpha_1) = f(x_1) + \alpha_1 f'(x_1) + \dots$ und folglich muß für α_1 die Gleichung

$$f(x_1 + \alpha_1) = f(x_1) + \alpha_1 f'(x_1) + \dots = 0 \quad (58)$$

erfüllt sein, wenn die Punkte Glieder mit $\alpha_1^2, \alpha_1^3, \dots, \alpha_1^k$ darstellen.

War die Ausgangsnäherung x_1 bereits hinreichend gut, dann ist ihr Fehler klein. Die Glieder mit höheren Potenzen können also in (58) keinen sehr großen Beitrag leisten, und α_1 muß annähernd der Gleichung

$$f(x_1) + \alpha_1 f'(x_1) = 0 \quad (59)$$

genügen, woraus wir

$$\alpha_1 \approx -\frac{f(x_1)}{f'(x_1)} \quad (60)$$

folgern. Wir können demnach annehmen, daß

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)} \quad (61)$$

unter hinreichend allgemeinen Bedingungen eine verbesserte Näherung ist. Danach kann die erhaltene Näherung erneut verbessert werden. Die dritte Näherung erhält man durch folgende Formel

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

Allgemein ergibt sich, wenn die n -te Näherung bekannt ist, die folgende durch

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (62)$$

oder explizit geschrieben

$$x_{n+1} = x_n - \frac{a_0 x_n^k + a_1 x_n^{k-1} + \dots + a_{k-1} x_n + a_k}{k a_0 x_n^{k-1} + (k-1) a_1 x_n^{k-2} + \dots + a_{k-1}}. \quad (63)$$

Unser Näherungsprozeß ist innerhalb der mitgeführten Stellenzahl beendet, wenn zwei Näherungen x_{n+1} und x_n innerhalb dieses Bereichs übereinstimmen.

Die eben dargestellte Näherungsmethode geht auf den großen englischen Gelehrten Sir Isaac Newton zurück.

Das Newtonsche Verfahren hängt eng mit der allgemeinen Iterationsmethode zusammen. Besitzen nämlich die Funktionen $y = f(x)$ und $y = f'(x)$ keine gemeinsame Nullstelle, dann ist die Gleichung $f(x) = 0$ der Gleichung

$$x = x - \frac{f(x)}{f'(x)} \quad (64)$$

äquivalent. Wenden wir hierauf die Iterationsmethode an, dann bekommen wir eine Zahlenfolge $x_1, x_2, x_3, \dots, x_n, \dots$ nach der Formel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (65)$$

die auch im Newtonschen Verfahren verwendet wird.

Beispiel: Wir wollen die schon mehrfach betrachtete Gleichung $x^3 - 3x - 5 = 0$ bis auf $\pm 0,001$ mit dem Newtonschen Verfahren lösen.

Diskussion: Als erste Näherung verwenden wir $x_1 = 3$. Die Ableitung von $f(x)$ ist $f'(x) = 3x^2 - 3$, so daß (62) in diesem konkreten Fall

$$x_{n+1} = x_n - \frac{x_n^3 - 3x_n - 5}{3x_n^2 - 3}$$

lautet. Danach ist

$$x_2 = 3 - \frac{27 - 9 - 5}{27 - 3} = 3 - \frac{13}{24} = 2,46,$$

$$x_3 = 2,46 - \frac{14,89 - 7,38 - 5}{18,16 - 3} = 2,46 - 0,165 = 2,295,$$

$$x_4 = 2,295 - \frac{12,088 - 6,885 - 5}{15,801 - 3} = 2,295 - 0,016 = 2,279,$$

$$x_5 = 2,279 - \frac{11,837 - 6,807 - 5}{15,582 - 3} = 2,279,$$

und wir sehen, daß bis auf einen Fehler von $\pm 0,001$

$$x_4 = x_5.$$

Deshalb ist die Wurzel der Gleichung $x^3 - 3x - 5 = 0$ gleich 2,279 mit einer Genauigkeit von 0,001.

Die in Kap. 6 beschriebene Methode der Quadratwurzeliteration ist ebenfalls ein Spezialfall des Newton-Verfahrens. Bedeutet doch die Berechnung von $\sqrt[k]{a}$ nichts anderes als Lösung der Gleichung $x^k - a = 0$.

Die Ableitung des Polynoms $x^k - a = f(x)$ ist $f'(x) = kx^{k-1}$, und deshalb erhält (62) die spezielle Form

$$x_{n+1} = x_n - \frac{x_n^k - a}{kx_n^{k-1}} = \frac{a + (k-1)x_n^k}{kx_n^{k-1}}.$$

Diese Formel ist aber mit der früher gegebenen Iterationsvorschrift vollkommen identisch. Wir wollen noch auf den folgenden entscheidenden Unterschied zwischen der Lösung von $x^k - a = 0$ und der allgemeinen algebraischen Gleichung (56) aufmerksam machen:

Für die Lösung der Gleichung $x^k - a = 0$ war die Wahl der Anfangsnäherung x_1 ganz willkürlich, und es ergab sich nach endlich vielen Schritten eine hinreichend gute Näherung von $\sqrt[k]{a}$. Ganz anders liegen die Dinge aber bei Gleichung (56).

Jetzt führt der eine Anfangswert zu dieser, ein anderer aber zu jener Wurzel, und einige Anfangswerte führen überhaupt nicht zu einer Lösung. Die nach (62) berechnete Folge $x_1, x_2, \dots, x_n, \dots$ konvergiert gegen keinen endlichen Wert, sie divergiert also.

15. Geometrische Interpretation der Ableitung

Das Newtonsche Verfahren können wir bis jetzt ausschließlich auf algebraische Gleichungen anwenden. Um es auf beliebige Gleichungen zu verallgemeinern, müssen wir den Begriff der Ableitung so erweitern, daß er auch für beliebige Funktionen seinen Sinn behält. Um zu einer solchen Verallgemeinerung zu kommen, wollen wir uns den geometrischen Inhalt des Begriffes klarmachen.

Wir betrachten die grafische Darstellung der Funktion $y = a_0x^k + a_1x^{k-1} + \dots + a_k$ und wählen auf der Kurve zwei benachbarte Punkte M und N aus (Abb. 12). Die Ab-

szissen in M und N seien x und $x + \alpha$, die Ordinaten ergeben sich dann zu

$$f(x) = a_0 x^k + a_1 x^{k-1} + \dots + a_k,$$

beziehungsweise

$$f(x + \alpha) = a_0 (x + \alpha)^k + a_1 (x + \alpha)^{k-1} + \dots + a_k.$$

Durch die beiden Punkte legen wir die Sekante und bestimmen ihre Steigung (den Anstieg) k_s .¹⁾ Der Zeichnung ent-

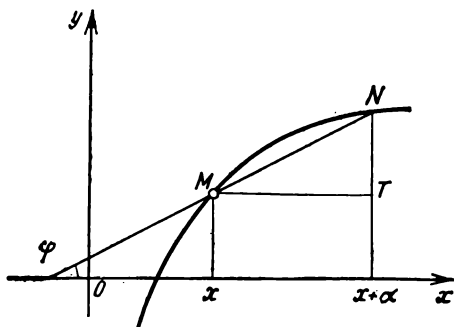


Abb. 12

nehmen wir $\tan \psi = \overline{TN}/\overline{MT}$. Die Streckenlängen \overline{MT} bzw. \overline{TN} sind Koordinatendifferenzen der Endpunkte, d. h.

$$\overline{MT} = (x + \alpha) - x = \alpha,$$

$$\overline{TN} = f(x + \alpha) - f(x),$$

und folglich ist

$$\tan \psi = \frac{\overline{TN}}{\overline{MT}} = \frac{f(x + \alpha) - f(x)}{\alpha}.$$

Nach Formel (53) gilt $f(x + \alpha) = f(x) + \alpha f'(x) + \dots$, wo die Punkte für Glieder höherer Ordnung in α stehen, und somit ist

$$\tan \psi = \frac{\alpha f'(x) + \dots}{\alpha} = f'(x) + \dots, \text{ wo dieses}$$

eine Mal die Punkte Terme mit $\alpha, \alpha^2, \alpha^3, \dots, \alpha^{k-1}$ ersetzen.

¹⁾ Als Anstieg oder Steigung einer Geraden bezeichnen wir die Zahl $k = \tan \psi$, wenn ψ der Winkel ist, den Gerade und positive x-Achse miteinander bilden.

Der Anstieg der Sekante durch M und N ist deshalb durch

$$k_S = \tan \psi = f'(x) + \dots \quad (66)$$

gegeben.

Jetzt wollen wir α verkleinern. Dann wird sich die Sekante irgendwie um den Punkt M drehen und dabei für $\alpha \rightarrow 0$ in die Tangente an die Kurve $y = f(x)$ im Punkte M übergehen. In Abb. 13 ist dieser Vorgang für eine Folge von α -Werten ($\alpha = 1, 1/2, 1/4, 0$) dargestellt.

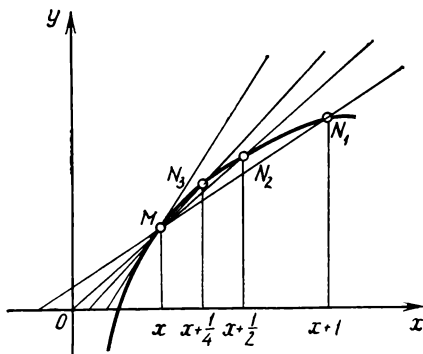


Abb. 13

Für $\alpha = 0$ sind aber alle in Formel (66) durch Punkte ersetzten Glieder gleich null, so daß der Anstieg der Tangente an die Kurve $y = f(x)$ im Punkte $M(x, f(x))$ gleich

$$k_T = f'(x) \quad (67)$$

ist.

Die Ableitung $f'(x)$ des Polynoms $f(x)$ ist gleich dem Anstieg der Tangente an die Bildkurve der Funktion $y = f(x)$ im Punkt mit der Abszisse x .

Beispiel: Bestimme den Winkel, unter welchem die Tangente an die Bildkurve von $f(x) = x^3 - 4x^2 + 5x + 1$ im Punkt mit der Abszisse $x = 2$ die positive x -Achse schneidet!

Diskussion: Wegen $f'(x) = 3x^2 - 8x + 5$ ist $f'(2) = 1$ der Anstieg der Tangente im gegebenen Punkt. Es gilt $\tan \varphi = 1$, woraus $\varphi = 45^\circ$ abzulesen ist.

16. Die geometrische Interpretation des Newtonschen Verfahrens

Jetzt sind wir so weit, daß uns die geometrische Interpretation des Verfahrens von Newton möglich ist. Wir hatten schon früher gesehen, daß die Lösung einer Gleichung

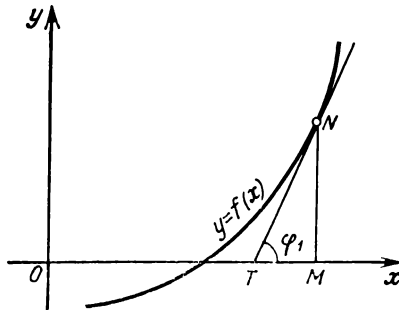


Abb. 14

$f(x) = 0$ gleichbedeutend damit ist, Schnittpunkte des zugehörigen Graphen mit der Abszissenachse zu ermitteln.

Wir wollen annehmen, daß uns die Näherung x_1 einer Wurzel dieser Gleichung bekannt sei. Im Punkte N mit der Abszisse x_1 legen wir die Tangente an die Kurve $f(x)$, die die Abszissenachse im Punkt $T(x_2, 0)$ schneidet, der unter gewissen Bedingungen näher am gesuchten Schnittpunkt liegt als M (Abb. 14).

Die Abszisse x_2 von T können wir leicht berechnen, wenn wir das rechtwinklige Dreieck $\triangle TMN$ betrachten, in dem TM mit $\overline{TM} = x_2 - x_1$ Ankathete und MN mit $\overline{MN} = f(x_1)$ Gegenkathete von φ_1 ist. Deshalb haben wir

$$\tan \varphi_1 = f(x_1)/(x_1 - x_2) \quad (68)$$

und

$$x_2 = x_1 - \frac{f(x_1)}{\tan \varphi_1}. \quad (69)$$

Nun war aber NM Tangente an $y = f(x)$ in N , und demzufolge ist der Anstieg gerade $\tan \varphi_1 = f'(x_1)$, Formel (69) lautet also

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}.$$

Zur Berechnung einer dritten Näherung für die gesuchte Wurzel können wir ganz genauso verfahren, wir brauchen die Tangente lediglich im Punkt $(x_2, f(x_2))$ an die Kurve zu legen, dann ergibt sich die Abszisse ihres Schnittpunktes mit der x -Achse nach der analogen Formel

$$x_3 = x_2 - \frac{f(x_2)}{f'(x_2)}.$$

Ganz allgemein können wir auf diese Weise eine ganze Folge x_1, x_2, x_3, \dots nach der Formel

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{f(x_n)}{\tan \varphi_n} \quad (70)$$

berechnen. Jeweils wird die nächste Näherung x_{n+1} aus der schon bekannten x_n so gewonnen, daß im Punkt mit der Abszisse x_n die Tangente an die Kurve $y = f(x)$ gelegt wird.

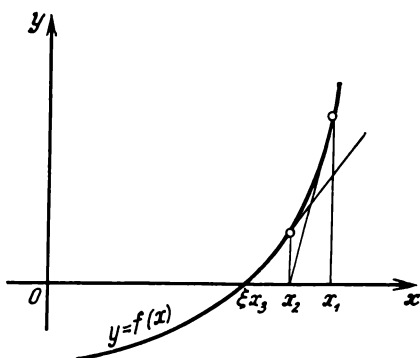


Abb. 15

Die Abszisse des Schnittpunktes dieser Tangente mit der x -Achse liefert x_{n+1} .

Der Winkel φ_n ist der Anstiegswinkel dieser Tangente. Die Formel (70) ist aber mit der Formel (62) identisch. Unser geometrisches Modell gibt genau das Verfahren von Newton wieder.

Wir haben also eine geometrische Interpretation des letztgenannten Näherungsverfahrens gewonnen:

In jedem Schritt wird ein Kurvenbogen von $y = f(x)$ durch eine Tangente an die Kurve ersetzt. Deshalb wird die Methode auch oft „Tangenten-Näherungsverfahren“ genannt. In Abb. 15 ist der Annäherungsvorgang veranschaulicht.

17. Die Ableitung einer nichtalgebraischen Funktion

Die soeben gefundene geometrische Interpretation des Newtonschen Verfahrens erlaubt sofort eine Verallgemeinerung auf andere Gleichungen der Form $f(x) = 0$, wo $f(x)$ auch kein Polynom sein kann (da der Begriff der Tangente und damit der der Ableitung anschaulich unabhängig von der speziellen Gestalt der definierenden Gleichung $y = f(x)$ ist [Anm. d. dtsh. Red.]).

Dazu wählen wir eine Näherungslösung x_1 für eine Wurzel, legen im Punkt mit der Abszisse x_1 die Tangente an die Kurve $y = f(x)$ und bezeichnen ihren Schnittpunkt mit der Abszissenachse mit x_2 . Durch den Kurvenpunkt mit der Abszisse x_2 zeichnen wir erneut die Tangente usw.

Genau wie im Falle der algebraischen Gleichung, wo $f(x)$ ein Polynom ist, wird leicht überprüft, daß

$$x_{n+1} = x_n - \frac{f(x_n)}{\tan \varphi_n} \quad (71)$$

gilt, wo $\tan \varphi_n$ der Anstieg der Tangente im Punkte $(x_n, f(x_n))$ ist.

Diese Gleichung ist für die rechnerische Durchführung nicht geeignet, weil wir den Anstieg $\tan \varphi_n$ nicht direkt berechnen können.

Wir benötigen deshalb zunächst eine Methode zur Berechnung des Anstieges der Tangente an eine durch $y = f(x)$ gegebene Kurve, wobei $f(x)$ kein Polynom zu sein braucht. Um sie zu finden, erinnern wir uns der geometrischen Interpretation der Ableitung eines Polynoms in Kap. 15 und bemerken, daß

$$k_S = \tan \psi = \frac{f(x+a) - f(x)}{a} \quad (72)$$

der Anstieg der Sekante durch die Kurvenpunkte M und N mit den Abszissen x und $x + a$ ist.

Wird $|a|$ verkleinert, dann bewegt sich der Punkt N längs der Kurve in Richtung auf M hin, die Sekante dreht sich um M und geht schließlich für $a \rightarrow 0$ in die Tangente an die Kurve im Punkt M über (vgl. Abb. 12). Also kann man schreiben

$$k_T = \tan \varphi = \lim_{a \rightarrow 0} \frac{f(x+a) - f(x)}{a}. \quad (73)$$

Den rechts stehenden Grenzwert nennen wir, falls er existiert, wiederum *Ableitung der Funktion* $f(x)$ und bezeichnen ihn mit $f'(x)$. Wir haben also die allgemeine Definition der Ableitung gefunden:

$$f'(x) = \lim_{\alpha \rightarrow 0} \frac{f(x+\alpha) - f(x)}{\alpha}. \quad (74)$$

Die Gleichung (73) lautet nun

$$k_T = \tan \varphi = f'(x). \quad (75)$$

Diese Definition ist als Verallgemeinerung des entsprechenden Begriffs so gewählt, daß die geometrische Deutung erhalten bleibt: die Ableitung ist ganz allgemein als Anstieg der Tangente an die Kurve $y = f(x)$ im Punkt M definiert. Eine Konsequenz dieser Tatsache ist beispielsweise, daß der Grenzwert in (74) nicht existieren kann, wenn es im betrachteten Punkt überhaupt keine Kurventangente gibt, wie etwa an einer Spitze. (Anm. d. deutsch. Red.: An dieser Stelle wurde eine Fußnote in den Text eingearbeitet.)

Jetzt können wir der Formel (71) wegen $\tan \varphi_n = f'(x_n)$ endgültig die Form

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad (76)$$

geben, die uns zeigt, daß wir jetzt das Newtonsche Verfahren wegen der Übereinstimmung mit (62) auf alle Gleichungen $f(x) = 0$ (mit differenzierbarem $f(x)$) ausgedehnt haben. (Eine Funktion heißt differenzierbar, wenn sie eine Ableitung besitzt [Anm. d. deutsch. Red.])

18. Die Berechnung der Ableitung

Im vorigen Abschnitt haben wir gesehen, daß zur Berechnung der Ableitung einer Funktion $y = f(x)$ die Bestimmung eines Grenzwertes nötig ist (siehe Formel (74)!).

Das ist aber in den meisten Fällen eine recht schwierige Aufgabe. Es wäre deshalb wünschenswert, daß man für möglichst viele der wichtigsten Funktionen die Ableitung ganz allgemein in geschlossener Form angeben könnte.

Das ist auch in der Tat möglich, und wir stellen nachfolgend eine kleine Tabelle der wichtigsten elementaren

Funktionen und ihrer Ableitungen zusammen:

1. $(a)' = 0$.
2. $(x^k)' = kx^{k-1}$.
3. $(a^x)' = a^x \ln a$.
4. $(\sin ax)' = a \cos ax$.
5. $(\cos ax)' = -a \sin ax$.
6. $(\tan ax)' = \frac{a}{\cos^2 ax}$.
7. $(\cot ax)' = -\frac{a}{\sin^2 ax}$.
8. $(\log_a x)' = \frac{1}{x \ln a}$.
9. $(\arcsin ax)' = \frac{a}{\sqrt{1-a^2x^2}}$.
10. $(\arctan ax)' = \frac{a}{1+a^2x^2}$.

($\ln a$ ist der Logarithmus zur Basis $e = 2,71828\dots$, der natürliche Logarithmus.)

Wir möchten noch besonders darauf hinweisen, daß der Exponent k in Formel 2. keine natürliche Zahl zu sein braucht, sondern beliebig! rational sein darf. So gilt beispielsweise

$$(\sqrt{x})' = (x^{\frac{1}{2}})' = \frac{1}{2} x^{\frac{1}{2}-1} = \frac{1}{2\sqrt{x}},$$

$$\left(\frac{1}{x^2}\right)' = (x^{-2})' = -2x^{-3} = -\frac{2}{x^3}.$$

Die Grundformeln 1.—10. reichen aber zur Berechnung der Ableitungen beliebig zusammengesetzter Funktionen noch nicht aus und müssen durch einige Rechenregeln ergänzt werden, die die Berechnung von Ableitungen solcher Funktionen ermöglichen, die mit Hilfe der Grundrechenarten aus einfachen elementaren Funktionen zusammengesetzt sind. Diese Regeln werden wie auch die Formeln 1.—10. in jedem Anfängerkurs der Differentialrechnung hergeleitet und lauten:

1. Die Ableitung der Summe zweier Funktionen ist gleich der Summe der Ableitungen der einzelnen Funktion:

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

2. Ein konstanter Zahlenfaktor kann vor das Ableitungszeichen gezogen werden:

$$(af(x))' = af'(x).$$

3. Die Ableitung des Produktes zweier Funktionen berechnet sich nach der folgenden „Produktregel“:

$$(f_1(x) \cdot f_2(x))' = f_1'(x) f_2(x) + f_1(x) f_2'(x).$$

4. Für die Ableitung eines Quotienten gilt die „Quotientenregel“

$$\left[\frac{f_1(x)}{f_2(x)} \right]' = \frac{f_1'(x) f_2(x) - f_1(x) f_2'(x)}{[f_2(x)]^2}$$

sofern $f_2(x) \neq 0$ ist — [Anm. d. dtsh. Red.].

Man kann leicht nachprüfen, daß die in Kap. 13 gegebene Ableitungsformel für Polynome aus Formel 2 und den Regeln 1. und 2. gefolgert werden kann.

Wir wollen uns jetzt einige Beispiele ansehen.

Beispiel 1:

$$f(x) = \frac{3x^2 - x + 1}{2x^3 + 5}.$$

Lösung: Anwendung von Regel 4. liefert zunächst

$$f'(x) = \frac{(3x^2 - x + 1)'(2x^3 + 5) - (3x^2 - x + 1)(2x^3 + 5)'}{(2x^3 + 5)^2},$$

woraus unter Anwendung der Regel für die Ableitung eines Polynoms wegen $(3x^2 - x + 1)' = 6x - 1$ und $(2x^3 + 5)' = 6x^2$ nach elementaren Umformungen

$$\begin{aligned} f'(x) &= \frac{(2x^3 + 5)(6x - 1) - (3x^2 - x + 1)6x^2}{(2x^3 + 5)^2} = \\ &= \frac{-6x^4 + 4x^3 + 6x^2 + 30x - 5}{(2x^3 + 5)^2} \end{aligned}$$

gefolgert werden kann.

Beispiel 2: $f(x) = \frac{1}{10} \left(\arcsin 3x - \frac{1}{x^2} \right).$

Lösung: Regel 1. und 2. zusammen mit Formel 2. und 9. führen unmittelbar auf

$$f'(x) = \frac{1}{10} \frac{3}{\sqrt{1-9x^2}} - \frac{1}{10} \left(\frac{-2}{x^3} \right) = \frac{3}{10\sqrt{1-9x^2}} + \frac{1}{5x^3}.$$

Beispiel 3: $f(x) = 10^x \sin 2x.$

Lösung: Hier braucht man die Formeln 3. und 4. und errechnet unter Anwendung der Produktregel 3.

$$\begin{aligned} f'(x) &= (10^x)' \sin 2x + 10^x (\sin 2x)' \\ &= 10^x \sin 2x \cdot \ln 10 + 10^x \cdot 2 \cos 2x \\ &= 10^x (\sin 2x \cdot \ln 10 + 2 \cos 2x). \end{aligned}$$

Auch die Regeln 1.—4. reichen noch nicht zur geschlossenen Darstellung aller vorkommenden Ableitungen aus und müssen noch durch die besonders wichtige Formel für die Ableitung einer mittelbaren Funktion ergänzt werden:

5. Wenn eine Funktion $y = f(x)$ in der Gestalt $y = F(z)$, $z = \varphi(x)$ dargestellt werden kann, dann wird ihre Ableitung durch

$$f'(x) = F'(z) \varphi'(x) \quad (77)$$

mit $z = \varphi(x)$ bestimmt.

Beispiel: Betrachten wir die Funktion $y = \sin(x^3)$, so kann sie leicht auf die oben benutzte Form $y = \sin z$, $z = x^3$, gebracht werden. Die Ableitung der „äußeren Funktion“ $F(z) = \sin z$ ist $F'(z) = \cos z$, während die Ableitung der „inneren Funktion“ $\varphi(x) = x^3$ bekanntlich $\varphi'(x) = 3x^2$ lautet. Wenden wir jetzt die Formel (77) an, so ergibt sich

$$(\sin(x^3))' = F'(z) \varphi'(x) = \cos z \cdot 3x^2 = 3x^2 \cdot \cos(x^3).$$

Eine ausführliche und mathematisch exakte Darstellung des Ableitungsbegriffs kann der Leser beispielsweise in dem Buch: Günther/Beyer/Gottwald/Wünsch: Grundkurs der Analysis, Teil 1. BSB B. G. Teubner Verlagsgesellschaft, Leipzig 1972 finden. Wir wollen uns jetzt wieder unserem numerischen Problem zuwenden.

19. Wie findet man eine Ausgangsnäherung?

Eine naheliegende Möglichkeit für die Bestimmung einer Ausgangsnäherung zur Lösung einer Gleichung $f(x) = 0$ ist die grafische Wurzelbestimmung durch grafische Darstellung der Funktion $y = f(x)$ und Ermittlung der Schnittpunkte des Funktionsbildes mit der Abszissenachse. (In diesem Punkt ist $y = 0$ und deshalb $f(x) = 0$.)

Wenn aber das Zeichnen des Graphen aus irgendeinem Grunde unmöglich oder auch nur unpassend ist (z. B. wenn die Rechnung auf einem Rechenautomaten erfolgen soll!), dann muß man sich nach anderen Möglichkeiten umsehen.

Man kann beispielsweise die Funktionswerte für eine Reihe von Argumentwerten ausrechnen und nach solchen x -Werten suchen, zwischen denen $f(x)$ offensichtlich das Vorzeichen wechselt. Ist nämlich die Funktion stetig, dann hat ihr Graph keine Lücken oder Sprünge, und zwischen x -Werten a und b , mit $f(a) f(b) < 0$ muß notwendigerweise wenigstens eine Stelle ξ mit $f(\xi) = 0$ liegen (Abb. 16).

Wieso dieser Schluß bei unstetigen Funktionen falsch ist, macht Abb. 16 b) anschaulich klar (dann kann sich das Vorzeichen ändern, ohne daß die Kurve dabei die Abszissenachse schneidet).

Wir wollen auch noch darauf hinweisen, daß bei diesem Vorgehen Nullstellen „vergessen“ werden können, weil auch

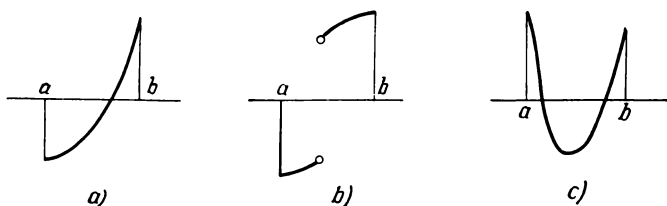


Abb. 16

zwischen Argumentwerten, an denen $f(x)$ gleiches Vorzeichen hat, die Kurve die x -Achse durchaus schneiden oder wenigstens berühren kann und das unter Umständen sogar mehrmals (vgl. Abb 16 c). Wir wollen jetzt nur den Fall von Abb. 16a) im Auge behalten und uns die Frage vorlegen, welche der beiden Werte a oder b wir im Newtonschen Verfahren günstig als Ausgangsnäherung für ξ verwenden sollten. Wir schauen uns dazu die Abbildung 17 an. Abb. 17 a) und 17 b) zeigen uns, daß wir bei Funktionen, deren Graphen von unten gesehen konvex sind, immer den Punkt mit positivem Funktionswert als Ausgangsnäherung zu wählen haben, während der Punkt mit negativem Funktionswert sachgemäß ist, wenn die Kurve von unten konkav ist, wie eine Betrachtung von Abb. 17 c) und d) lehrt.

Diese Regel nützt uns natürlich gar nichts, wenn wir die graphische Darstellung nicht kennen, und deshalb hätten wir gern eine rechnerische Möglichkeit für die Bestätigung der richtigen Auswahl gehabt.

Dazu benötigen wir aber den Begriff der *zweiten Ableitung* einer Funktion $f(x)$.

Unter der *zweiten Ableitung* $f''(x)$ einer Funktion $f(x)$ verstehen wir die *Ableitung der ersten Ableitung* von $f(x)$, d. h. $f''(x) = (f'(x))'$.

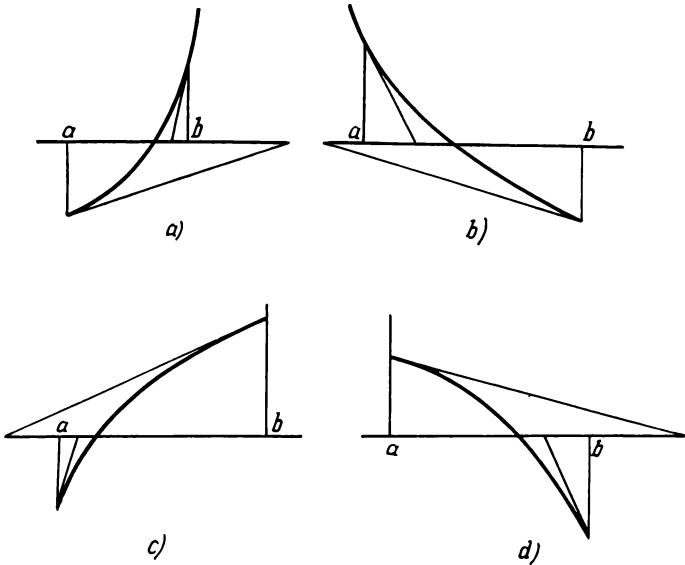


Abb. 17

So ist beispielsweise die erste Ableitung von $f(x) = x^3 - 4x^2 + 3x - 1$ gleich $f'(x) = 3x^2 - 8x + 3$ und die zweite Ableitung gleich

$$f''(x) = 6x - 8.$$

In jedem Einführungskurs in die höhere Mathematik wird bewiesen, daß die zweite Ableitung einer Funktion überall dort positiv ist, wo der Graph eine von unten konvexe Kurve ist, während sie negativ ausfällt, wo die Kurve von unten konkaven Verlauf zeigt. Unter Verwendung dieser Aussage läßt sich die Frage nach der richtigen Wahl der Anfangsnäherung beim Newtonschen Verfahren auch ohne Zeichnung richtig beantworten:

Wenn die Funktion $y = f(x)$ in zwei Punkten $x = a$ und $x = b$ entgegengesetztes Vorzeichen hat, dann wird immer

derjenige von ihnen zum Startwert genommen, an dem die Vorzeichen von Funktion und 2. Ableitung übereinstimmen.

20. Eine kombinierte Methode zur Lösung von Gleichungen

Die Methode von Newton und das Sekantenverfahren werden manchmal auch kombiniert eingesetzt:

Wenn die Kurve $y = f(x)$ von unten konvex ist, dann erhalten wir die Punkte a_1 und x_1 nach den Formeln

$$a_1 = a - f(a) \frac{b-a}{f(b)-f(a)}, \quad (78)$$

$$x_1 = b - \frac{f(b)}{f'(b)}. \quad (79)$$

Ist das Bild von $y = f(x)$ dagegen von unten konkav, so tritt an die Stelle von (79) die Gleichung

$$x_1 = a - \frac{f(a)}{f'(a)}. \quad (80)$$

Den Abbildungen 18 entnehmen wir anschaulich die Aussage, daß die Wurzel ξ der Gleichung $f(x) = 0$ dann zwischen den beiden Punkten x_1 und a_1 liegt. Wenden wir nun

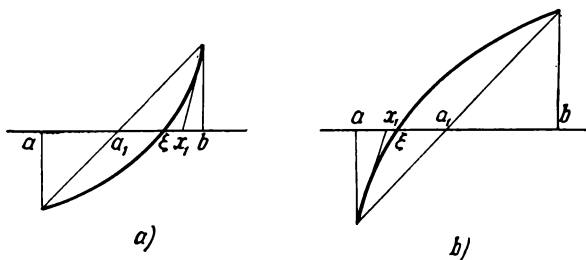


Abb. 18

auf diese beiden Werte erneut die Formeln (78) und (79) (bzw. (80)) an, dann erhalten wir wieder ein entsprechendes Paar a_2, x_2 usw.

So bekommt man eine Folge $a_1, a_2, \dots, a_n, \dots$ und eine zweite Folge $x_1, x_2, \dots, x_n, \dots$, die sich beide von verschiedenen Seiten her der gesuchten Wurzel ξ nähern, und darin besteht gerade der Vorzug dieses Vorgehens, weil man auf diese Weise stets genaue Fehlergrenzen nach oben und unten hat.

Beispiel: Wir lösen die Gleichung $x - \sin x - 0,5 = 0$ nach diesem kombinierten Verfahren mit der Genauigkeit von 0,001.

Lösung: Zunächst tabellieren wir die Funktion

$$f(x) = x - \sin x - 0,5$$

x	-1	0	1	2
$f(x)$	-0,659	-0,5	-0,341	0,591

und stellen fest, daß zwischen $x = 1$ und $x = 2$ wenigstens eine Wurzel liegt. Unter Verwendung der Formeln 2. und 4. aus Kap. 18 erhalten wir für die Ableitung

$$f'(x) = 1 - \cos x.$$

Die Rekursionsformel des Newtonschen Verfahrens bekommt damit folgende Gestalt

$$x_{n+1} = x_n - \frac{x_n - \sin x_n - 0,5}{1 - \cos x_n}. \quad (81)$$

Um zu erfahren, welche der beiden Zahlen wir für x_0 zu nehmen haben, bilden wir die zweite Ableitung der Funktion. Nach Formel 5. von Kap. 18 gilt $f''(x) = \sin x$. Nun wissen wir aber, daß die Funktion $y = \sin x$ im Intervall $[1, 2]$ positiv ist, denn sie ist ja sogar in dem umfassenden Intervall $[0; \pi] = [0; 3,1415\dots]$ nichtnegativ.

Das bedeutet aber, daß die Funktion $f(x)$ selbst in diesem Intervall von unten konkav ist und wir nach der früher erarbeiteten Regel den Wert $x_0 = 2$ als Startwert zu nehmen haben.

Nun rechnen wir nach der Formel (81)

$$x_1 = 2 - \frac{2 - \sin 2 - 0,5}{1 - \cos 2} = 2 - \frac{2 - 0,909 - 0,5}{1 + 0,416} = 1,583$$

und nach Formel (78)

$$a_1 = 1 - (-0,341) \frac{2-1}{0,591 - (-0,341)} = 1,366.$$

Verwenden wir die gleichen Formeln auch weiterhin, dann ergibt sich

$$x_2 = 1,583 - \frac{1,583 - 1,000 - 0,5}{1 + 0,012} = 1,501$$

und

$$a_2 = 1,366 + 0,113 \frac{1,583 - 1,366}{0,083 + 0,113} = 1,491$$

und weiter $x_3 = 1,498$ bzw. $a_3 = 1,498$. Die Wurzel unserer Gleichung ist bis auf einen Fehler von $\pm 0,001$ gleich 1,498.

21. Ein Konvergenzkriterium für die allgemeine Iteration

In diesem Abschnitt wollen wir sehen, wie mit Hilfe des Begriffs der Ableitung ein neues Konvergenzkriterium für

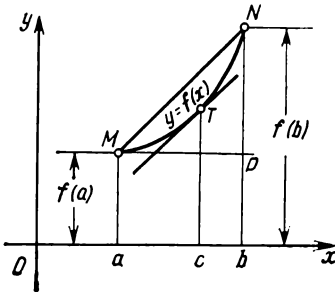


Abb. 19

Iterationen gefunden werden kann. Wir benötigen dafür eine Formel, die „Formel von Lagrange“ genannt wird (Lagrange war ein französischer Mathematiker des 18. Jahrhunderts).— Die Formel ist bei uns als Mittelwertsatz der Differentialrechnung bekannt [Anm. d. dtsh. Red.]. Über dem Intervall $[a, b]$ betrachten wir die Bildkurve einer Funktion $y = f(x)$ (Abb. 19). Bezeichnen wir den Anfangspunkt des Kurvenstücks mit M und seinen Endpunkt mit N , so berechnet sich der Anstieg der Sekante durch diese beiden Punkte nach der Formel

$$k_S = \tan \psi = \overline{PN} / \overline{MP}.$$

Unter Verwendung von $\overline{MP} = b - a$, $\overline{PN} = f(b) - f(a)$, wird daraus

$$k_S = \frac{f(b) - f(a)}{b - a}.$$

Nun suchen wir den Punkt T auf dem Kurvenbogen, der von der Sekante den größtmöglichen Abstand hat, und zeichnen durch T die Tangente. Es zeigt sich, daß diese Tangente parallel zu MN sein muß, da man durch einen Punkt außerhalb einer Geraden nur eine Parallele zeichnen kann und die Parallele durch T die Kurve in keinem Punkt (innerhalb des betrachteten Intervalls) echt schneiden kann, denn

sonst würde es Punkte geben, die weiter als T von MN entfernt sind. Der Anstieg der Tangente in T ist gleich $f'(c)$, wenn c die Abszisse von T bezeichnet, und nach dem oben Gesagten muß deshalb

$$f'(c) = \frac{f(b) - f(a)}{b - a} \quad (82)$$

erfüllt sein. Das ist der wesentliche Inhalt des sogenannten *Mittelwertsatzes der Differentialrechnung*.

Wir wollen ausdrücklich hervorheben, daß die Abszisse c des Punktes T stets zwischen a und b liegt.

In einer etwas anders aufgeschriebenen Form

$$f(b) - f(a) = f'(c) \cdot (b - a) \quad (83)$$

wird die Ordinatendifferenz direkt als $f'(c) = \text{fache Abszissendifferenz}$ dargestellt.

Jetzt wenden wir uns wieder unserem Problem der iterativen Berechnung einer Wurzel der Gleichung $x = \varphi(x)$ zu. Nehmen wir an, daß die Funktion $x \rightarrow \varphi(x)$ das Intervall $[a, b]$ in sich abbildet, wobei im ganzen Intervall gleichmäßig die Ungleichung $|\varphi'(x)| \leq q$ erfüllt ist, wo q eine beliebige reelle Zahl zwischen 0 und 1 bezeichnet. Nun greifen wir zwei beliebige Punkte x_1 und x_2 des Intervalls heraus. Dann liegen auch $\varphi(x_1)$ und $\varphi(x_2)$ im Intervall, und die Formel (82) garantiert die Existenz eines Punktes c zwischen x_1 und x_2 , also jedenfalls aus dem Intervall $[a, b]$, mit

$$\varphi(x_2) - \varphi(x_1) = \varphi'(c) \cdot (x_2 - x_1).$$

Aus dieser Gleichung folgt

$$|\varphi(x_2) - \varphi(x_1)| = |\varphi'(c)| |x_2 - x_1| \leq q |x_2 - x_1|, \quad (84)$$

und die Voraussetzung $q < 1$ erweist sich gerade als hinreichende Bedingung für die *Kontraktivität* der Abbildung $x \rightarrow \varphi(x)$ in $[a, b]$. Wir wissen aber bereits, daß eine solche Abbildung stets einen Fixpunkt im Intervall hat, den man mit einem aus dem Intervall völlig beliebig gewählten Anfangswert x_0 nach der Vorschrift $x_{n+1} = \varphi(x_n)$ durch Iteration beliebig genau annähern kann. Wir formulieren das erhaltene Ergebnis in einem

Satz: Wenn eine Funktion $y = \varphi(x)$ ein Intervall $[a, b]$ in sich abbildet, wobei für alle $x \in [a, b]$ eine Ungleichung $|\varphi'(x)| \leq q < 1$ gilt, dann konvergiert die mit einem be-

liebigen Startwert $x_0 \in [a, b]$ gebildete Iterationsfolge

$$x_{n+1} = \varphi(x_n)$$

gegen die einzige Lösung der Gleichung $x = \varphi(x)$ im betrachteten Intervall.

Grob gesprochen besteht der Gehalt des eben formulierten Satzes darin, daß der Iterationsprozeß für solche Wurzeln ξ der Gleichung $x = \varphi(x)$ konvergiert, für die die Ungleichung $|\varphi'(\xi)| < 1$ erfüllt ist. Man kann davon sprechen, daß diese Punkte den Polygonzug, der geometrisch der Iterationsfolge entspricht (vgl. Kap. 8), „anziehen“, während Punkte, wo $|\varphi'(\xi)| > 1$ ist, denselben von sich „abstoßen“.

Wenn die Ungleichung $|\varphi'(x)| \leq q < 1$ für alle x erfüllt ist, dann konvergiert das Iterationsverfahren bei vollkommener freier Wahl des Anfangswertes x_0 (vgl. Kap. 10).

Beispiel 1: Wir wollen untersuchen, ob zur Lösung der Gleichung $x = \frac{1}{4}(\cos x + \sin x)$ das Iterationsverfahren benutzt werden kann.

Diskussion: Hier ist $\varphi(x) = \frac{1}{4}(\cos x + \sin x)$, daher $\varphi'(x) = \frac{1}{4}(-\sin x + \cos x)$. Bekanntlich ist $|\sin x| \leq 1$ und $|\cos x| \leq 1$ für alle x und daher

$$|\varphi'(x)| = \left| \frac{-\sin x + \cos x}{4} \right| \leq \frac{|\sin x| + |\cos x|}{4} < \frac{1}{2}$$

auf der gesamten reellen Achse. Das Iterationsverfahren konvergiert somit (und zwar für jeden Anfangswert x_0 gegen die einzige Lösung der Gleichung [Anm. d. deutsch. Red.]).

Beispiel 2: Jetzt wollen wir die gleiche Frage für die Gleichung

$$x = 4 - 2^x \tag{85}$$

entscheiden.

Diskussion: Eine Wurzel muß im Intervall $[1, 2]$ liegen, denn die Funktion $y = f(x) = 4 - 2^x - x$ ändert wegen $4 - 2^1 - 1 > 0$ und $4 - 2^2 - 2 < 0$ hier das Vorzeichen. Es ist $\varphi'(x) = -2^x \ln 2$ und $|\varphi'(x)| = 2^x \ln 2$. Für $1 \leq x \leq 2$ ist $2 \leq 2^x \leq 4$ und demnach

$$2 \ln 2 \leq 2^x \ln 2 \leq 4 \ln 2.$$

Aus einer Tabelle der natürlichen Logarithmen (Logarithmen mit der Basis $e = 2,78 \dots$) entnehmen wir $\ln 2 = 0,69 \dots$

und erhalten die Beziehung

$$1,38 \dots \leq 2^x \ln 2 \leq 2,76 \dots,$$

die für alle x aus $[1, 2]$ erfüllt ist. Also divergiert das Iterationsverfahren. Wollen wir die Gleichung (85) dennoch durch Iteration lösen, dann müssen wir sie äquivalent umformen. Zunächst folgt aus (85) $2^x = 4 - x$ und daraus durch Logarithmieren zur Basis 2

$$x = \log_2(4 - x).$$

In dieser Gleichung ist jetzt $\varphi(x) = \log_2(4 - x)$ und $\varphi'(x) = -\frac{1}{(4-x)\ln 2}$, und im Intervall $[1, 2]$ gilt sicher $|\varphi'(x)| < \frac{1}{2 \ln 2} = \frac{1}{1,38} < 1$. Das Iterationsverfahren für diese Form der Gleichung konvergiert!

22. Über die Konvergenzgeschwindigkeit der Iteration

In diesem Abschnitt, dessen Inhalt übrigens später nicht weiter verwendet wird, wollen wir sehen, wie die Ableitung $\varphi'(x)$ auch zur Abschätzung der Konvergenzgeschwindigkeit herangezogen werden kann. Dazu wollen wir die

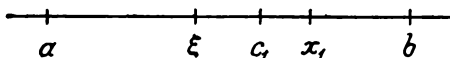


Abb. 20

Änderungsgeschwindigkeit des Fehlers $\alpha_n = \xi - x_n$ der Näherungswerte $x_1, x_2, \dots, x_n, \dots$ studieren. Wir erinnern daran, daß die beiden Gleichungen $\xi = \varphi(\xi)$ und $x_{n+1} = \varphi(x_n)$ gelten, so daß

$$\alpha_{n+1} = \xi - x_{n+1} = \varphi(\xi) - \varphi(x_n)$$

und wegen Formel (82)

$$\varphi(\xi) - \varphi(x_n) = \varphi'(c_n)(\xi - x_n) = \varphi'(c_n)\alpha_n$$

mit Zahlen c_n zwischen x_n und ξ gelten.

Demnach ist

$$\alpha_{n+1} = \varphi'(c_n)\alpha_n. \quad (86)$$

Den Inhalt dieser Formel können wir folgendermaßen umschreiben:

Eine Wurzel ξ der Gleichung $x = \varphi(x)$ möge im Intervall $[a, b]$ liegen. Gilt in diesem Intervall $|\varphi'(x)| \leq q < 1$ und liegt auch der Anfangswert x_1 im Intervall, dann gilt für alle natürlichen Zahlen n die Abschätzung

$$|\alpha_{n+1}| \leq q^n |\alpha_1|. \quad (87)$$

Das sieht man leicht ein, denn nach (86) ist

$$|\alpha_2| = |\varphi'(c_1)| |\alpha_1|$$

$|\alpha_2| \leq q |\alpha_1|$, denn c_1 liegt in $[a, b]$ (Abb. 20), wo $|\varphi'(c_1)| \leq q < 1$ immer erfüllt ist. Dasselbe gilt auch für α_3

$$|\alpha_3| = |\varphi'(c_2)| |\alpha_2| < q |\alpha_2| < q^2 |\alpha_1|$$

und allgemein

$$|\alpha_{n+1}| < q^n |\alpha_1|$$

(Anmerkung: Der exakte Beweis wird durch vollständige Induktion erbracht).

Wegen $0 < q < 1$ konvergiert die geometrische Folge $q, q^2, q^3, \dots, q^n, \dots$ und damit auch die Folge der Fehler α_n mit wachsendem n gegen null. Mit anderen Worten konvergiert die Iterationsfolge $x_1, x_2, \dots, x_n, \dots$ unter den weiter oben gemachten Voraussetzungen gegen ξ , wobei die Differenz $|\xi - x_{n+1}|$ schneller als $|\alpha_1| \cdot q^n$ klein wird.

Mit den gleichen Schlüssen können wir jetzt auch leicht noch einmal die Divergenz des Prozesses für den Fall beweisen, daß $|\varphi'(x)| > 1$ im Intervall $[a, b]$ ist.

Besonders rasch konvergiert das Iterationsverfahren, wenn die Ableitung von $\varphi(x)$ im Punkte ξ gleich null ist. Dann strebt nämlich nicht nur α_n , sondern auch $\varphi'(c_n)$ für $n \rightarrow \infty$ gegen null, so daß wegen

$$|\alpha_{n+1}| = |\varphi'(c_n)| |\alpha_n|$$

sich der Konvergenzvorgang mit wachsendem n ständig beschleunigt.

Wir waren diesem Phänomen schon früher bei der Berechnung der Quadratwurzel begegnet. Wir erinnern uns, daß wir dazu die Gleichung $x^2 = a$ auf die Gestalt $x = \frac{x^2 + a}{2x}$ gebracht hatten. Die Ableitung der rechten Seite

berechnen wir aber so :

$$\begin{aligned}\varphi'(x) &= \frac{(x^2 + a)' \cdot 2x - (x^2 + a) (2x)'}{4x^2} \\ &= \frac{2x \cdot 2x - (x^2 + a) 2}{4x^2} = \frac{x^2 - a}{2x^2}.\end{aligned}$$

(Dabei haben wir Regel 4. aus Kap. 18 und die Formel (55) aus Kap. 13 verwendet.) Folglich wird

$$\varphi'(\sqrt{a}) = \frac{(\sqrt{a})^2 - a}{2(\sqrt{a})^2} = 0.$$

Im Punkte $x = \sqrt{a}$ wird nicht nur $\varphi(x)$, sondern auch die Ableitung gleich null, und das führt zu einer ständigen Beschleunigung der Konvergenz bei Annäherung an diesen Punkt.

Diese Erscheinung der Beschleunigung des Iterationsprozesses während der Annäherung an die Wurzel der Gleichung ist aber auch für das Newtonsche Verfahren ganz allgemein charakteristisch. Wir hatten doch schon herausgearbeitet, daß dieses Verfahren mit der zu lösenden Gleichung $f(x) = 0$ über die Gleichung

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

zusammenhängt, und die sukzessive Approximation hierauf angewendet eine Näherungsfolge für die Lösung jener Gleichung erzeugt. Deshalb haben wir hier $\varphi(x) = x - \frac{f(x)}{f'(x)}$, also

$$\begin{aligned}\varphi'(x) &= 1 - \left[\frac{f(x)}{f'(x)} \right]' = 1 - \frac{f'(x)[f(x)]' - f(x)[f'(x)]'}{[f'(x)]^2} \\ &= 1 - \frac{[f'(x)]^2 - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2},\end{aligned}$$

und weil im Punkte ξ die Gleichung $f(\xi) = 0$ gelten sollte, ist dort auch $\varphi'(\xi) = 0$. Das aber bedeutet, wie wir oben erläutert haben, gerade die Beschleunigung der Konvergenz bei Annäherung an die Wurzel ξ .

23. Die iterative Lösung eines linearen Gleichungssystems

Bisher haben wir unsere ganze Diskussion auf Gleichungen mit einer Unbekannten ausgerichtet. Jetzt wollen wir zur Lösung von Gleichungssystemen übergehen und

beginnen mit dem einfachsten Fall linearer Gleichungen.

Es seien m Gleichungen ersten Grades mit m Unbekannten x_1, x_2, \dots, x_m und reellen Koeffizienten a_{ij} ($i = 1, 2, \dots, m$ gibt die Nummer der Gleichung, $j = 1, 2, \dots, m$ die der Unbekannten an) gegeben:

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m &= b_2, \\ \dots & \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mm}x_m &= b_m. \end{aligned} \right\} \quad (88)$$

Solchen Systemen begegnet man in den verschiedensten Anwendungen, und zwar treten dabei häufig sehr viele Unbekannte und Gleichungen auf. So hat beispielsweise ein Geodät bei der Auswertung der Vermessung eines größeren Teiles der Erdoberfläche manchmal Systeme mit mehreren hundert Gleichungen zu lösen. Im Zusammenhang mit der Konstruktion von Fachwerken hat auch ein Bauingenieur häufig mit solchen Systemen zu tun.

Bei der Lösung mit den herkömmlichen Eliminationsverfahren (etwa dem Gaußschen Algorithmus oder der Methode des Variablen austausches [Anm. d. dtsh. Red.]) entstehen dabei oft beträchtliche Schwierigkeiten. Die Anwendung von Näherungsmethoden ist hier weitaus vorteilhafter.

Wir wollen zunächst an einem Beispiel den Lösungsweg zeigen und zwar gehen wir vom System

$$\begin{aligned} 10x_1 - 2x_2 + x_3 &= 9, \\ x_1 + 5x_2 - x_3 &= 8, \\ 4x_1 + 2x_2 + 8x_3 &= 39 \end{aligned}$$

aus und suchen eine Lösung x_1, x_2, x_3 mit Fehlern von höchstens $\pm 0,001$ pro Komponente.

Wir lösen die erste Gleichung nach x_1 , die zweite nach x_2 und die dritte nach x_3 auf und erhalten

$$\begin{aligned} x_1 &= 0,9 + 0,2x_2 - 0,1x_3, \\ x_2 &= 1,6 - 0,2x_1 + 0,2x_3, \\ x_3 &= 4 - 0,5x_1 - 0,25x_2. \end{aligned} \quad (89)$$

Dann nehmen wir drei beliebige Zahlen als Anfangsnäherungen, also z. B. $x_1^{(0)} = x_2^{(0)} = x_3^{(0)} = 0$. Setzen wir diese Werte auf der rechten Seite von (89) ein, dann erhalten wir

drei neue Zahlen $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}$, in unserem Beispiel

$$x_1^{(1)} = 0,9;$$

$$x_2^{(1)} = 1,6;$$

$$x_3^{(1)} = 4;$$

die wir als erste Näherungen erneut rechts einsetzen:

$$x_1^{(2)} = 0,9 + 0,2 \cdot 1,6 - 0,1 \cdot 4 = 0,82;$$

$$x_2^{(2)} = 1,6 - 0,2 \cdot 0,9 + 0,2 \cdot 4 = 2,22;$$

$$x_3^{(2)} = 4 - 0,5 \cdot 0,9 - 0,25 \cdot 1,6 = 3,15$$

und die zweite Näherung berechnen. Haben wir so fortfahrend schließlich die n -te Näherung $x_1^{(n)}, x_2^{(n)}, x_3^{(n)}$ ausgerechnet, dann wird die folgende Näherung nach dieser Formel gebildet:

$$\begin{aligned} x_1^{(n+1)} &= 0,9 + 0,2x_2^{(n)} - 0,1x_3^{(n)}, \\ x_2^{(n+1)} &= 1,6 - 0,2x_1^{(n)} + 0,2x_3^{(n)}, \\ x_3^{(n+1)} &= 4 - 0,5x_1^{(n)} - 0,25x_2^{(n)}. \end{aligned} \quad (90)$$

Wir haben die jeweiligen Resultate in einer Tabelle zusammengestellt:

Tab. 2

n	1	2	3	4	5	6
$x_1^{(n)}$	0,9	0,82	1,03	1,01	1,00	1,00
$x_2^{(n)}$	1,6	2,22	2,07	2,00	1,99	2,00
$x_3^{(n)}$	4,0	3,15	3,03	2,97	3,00	3,00

Wir sehen, daß innerhalb der vorgegebenen Fehlergrenzen

$$x_1^{(5)} = x_1^{(6)}, \quad x_2^{(5)} = x_2^{(6)}, \quad x_3^{(5)} = x_3^{(6)} \quad (91)$$

gilt. Setzen wir in (90) $n = 5$ und verwenden (91), dann erhalten wir

$$x_1^{(5)} \approx 0,9 + 0,2x_2^{(5)} - 0,1x_3^{(5)},$$

$$x_2^{(5)} \approx 1,6 - 0,2x_1^{(5)} + 0,2x_3^{(5)},$$

$$x_3^{(5)} \approx 4 - 0,5x_1^{(5)} - 0,25x_2^{(5)}$$

(dieses System gilt hier sogar exakt, aber das muß nicht sein). $x_1 = 1,00$, $x_2 = 2,00$ und $x_3 = 3,00$ sind die gesuchten Näherungslösungen.

Das Vorgehen im allgemeinen Fall unterscheidet sich von dem soeben im Beispiel praktizierten überhaupt nicht.

Es sei ein Gleichungssystem (88) gegeben. Dann lösen wir die erste Gleichung nach x_1 , die zweite nach x_2 usw. auf und erhalten das äquivalente System

$$\begin{aligned} x_1 &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2 - \dots - \frac{a_{1m}}{a_{11}} x_m, \\ x_2 &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1 - \dots - \frac{a_{2m}}{a_{22}} x_m, \\ x_m &= \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1 - \frac{a_{m2}}{a_{mm}} x_2 - \dots - \frac{a_{m, m-1}}{a_{mm}} x_{m-1}. \end{aligned} \quad (92)$$

Setzen wir gewisse Anfangsnäherungen $x_1^{(1)}$, $x_2^{(1)}$, \dots , $x_m^{(1)}$ für die Unbekannten x_1 , x_2 , \dots , x_m in die rechte Seite des Systems (92) ein, dann errechnen wir zweite Näherungen $x_1^{(2)}$, \dots , $x_m^{(2)}$ nach

$$\begin{aligned} x_1^{(2)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2^{(1)} - \dots - \frac{a_{1m}}{a_{11}} x_m^{(1)}, \\ x_2^{(2)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(1)} - \dots - \frac{a_{2m}}{a_{22}} x_m^{(1)}, \\ x_m^{(2)} &= \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1^{(1)} - \dots - \frac{a_{m, m-1}}{a_{mm}} x_{m-1}^{(1)} \end{aligned}$$

und $(n + 1)$ -te Näherungen durch die Formeln

$$\begin{aligned} x_1^{(n+1)} &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} x_2^{(n)} - \dots - \frac{a_{1m}}{a_{11}} x_m^{(n)}, \\ x_2^{(n+1)} &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x_1^{(n)} - \dots - \frac{a_{2m}}{a_{22}} x_m^{(n)}, \\ &\vdots \\ x_m^{(n+1)} &= \frac{b_m}{a_{mm}} - \frac{a_{m1}}{a_{mm}} x_1^{(n)} - \dots - \frac{a_{m, m-1}}{a_{mm}} x_{m-1}^{(n)}, \end{aligned} \quad (93)$$

wenn wir annehmen, daß uns die n -te Näherung bekannt ist.

Bevor wir weiter auf diese iterative Methode zur Lösung linearer Gleichungssysteme eingehen, indem wir insbesondere die Konvergenzfrage zu klären versuchen, wollen wir

uns erst einmal mit einem umgekehrten Vorgehen bekannt machen. Es kann nämlich auch vorkommen, daß die Lösung eines linearen Gleichungssystems zur Bestimmung der Grenzwerte eines gewissen Näherungsprozesses notwendig wird. (Wir erinnern an die analoge Fragestellung bei einer Variablen, wie sie uns bei Achilles und der Schildkröte begegnet ist [Anm. d. dtsh. Red.]).

Von drei gleichen Eimern möge genau einer gefüllt sein und 12 l Wasser enthalten. Jetzt gießen wir die Hälfte seines Inhaltes in den zweiten Eimer und davon wieder die Hälfte in den dritten. Anschließend gießen wir die Hälfte des Inhaltes des dritten Eimers in den ersten zurück und beginnen wieder wie oben beschrieben. Den Zyklus durchlaufen wir 20-mal und fragen, wieviel Wasser sich dann in jedem Eimer befindet (auf 0,0001 l genau!).

Es ist klar, daß es sich hier um die sukzessive Approximation einer bestimmten Verteilung des Wassers handelt. Dieser Grenzzustand ist offenbar dadurch charakterisiert, daß er sich nicht mehr ändert, wenn der Zyklus noch einmal durchlaufen wird (von Verlusten sehen wir natürlich ganz und gar ab).

Zu Beginn des Zyklus seien im ersten Eimer x Liter, im zweiten y Liter und im dritten die restlichen $12-x-y$ Liter Wasser. Dann würde sich bei Anwendung des Zyklus die folgende Tabelle ergeben.

	1. Eimer	2. Eimer	3. Eimer
Anfangszustand	x	y	$12 - x - y$
1. Eimer umgefüllt	$x/2$	$\frac{x}{2} + y$	$12 - x - y$
2. Eimer umgefüllt	$x/2$	$\frac{x}{4} + \frac{y}{2}$	$12 - \frac{3}{4}x - \frac{y}{2}$
3. Eimer umgefüllt	$6 + \frac{x}{8} - \frac{y}{4}$	$\frac{x}{4} + \frac{y}{2}$	$6 - \frac{3}{8}x - \frac{y}{4}$

Soll dieser Zyklus den Zustand nicht mehr verändern, dann müssen notwendig folgende Gleichungen erfüllt

sein :

$$x = 6 + x/8 - y/4,$$

$$y = x/4 + y/2.$$

Wenn wir dieses lineare System lösen, dann finden wir $x = 6$ und $y = 3$. Der Grenzzustand ist demnach dann erreicht, wenn im ersten Eimer 6 l, im zweiten 3 l und im dritten gleichfalls 3 l sind.

Wir wollen jetzt feststellen, mit welcher Geschwindigkeit sich dieser Endzustand einstellt. Sind am Anfang im ersten Eimer a l und im zweiten b l Wasser, dann sind es nach dem ersten Zyklus

$$a_1 = 6 + a/8 - b/4 \quad (94)$$

beziehungsweise

$$b_1 = a/4 + b/2 \quad (95)$$

Liter geworden. Nennen wir $a - 6 = \alpha$ und $b - 3 = \beta$, $a_1 - 6 = \alpha_1$ und $b_1 - 3 = \beta_1$, dann folgt aus (94) und (95)

$$\alpha_1 = a_1 - 6 = \frac{a-6}{8} - \frac{b-3}{4} = \frac{\alpha}{8} - \frac{\beta}{4} \text{ und}$$

$$\beta_1 = b_1 - 3 = \frac{a-6}{4} + \frac{b-3}{2} = \frac{\alpha}{4} + \frac{\beta}{2}.$$

Nach dem zweiten Zyklus werden die entsprechenden Fehler so berechnet :

$$\begin{aligned} \alpha_2 &= \frac{\alpha_1}{8} - \frac{\beta_1}{4} = \frac{1}{8} \left(\frac{\alpha}{8} - \frac{\beta}{4} \right) - \frac{1}{4} \left(\frac{\alpha}{4} + \frac{\beta}{2} \right) \\ &= -\frac{3}{64} \alpha - \frac{5}{32} \beta, \end{aligned}$$

$$\begin{aligned} \beta_2 &= \frac{\alpha_1}{4} + \frac{\beta_1}{2} = \frac{1}{4} \left(\frac{\alpha}{8} - \frac{\beta}{4} \right) + \frac{1}{2} \left(\frac{\alpha}{4} + \frac{\beta}{2} \right) \\ &= \frac{5}{32} \alpha - \frac{3}{16} \beta. \end{aligned}$$

Wenn wir annehmen, daß $|\alpha| < \varepsilon$ und $|\beta| < \varepsilon$ war, dann ist

$$|\alpha_2| < \frac{13}{64} \varepsilon \approx 0,2\varepsilon,$$

$$|\beta_2| < \frac{11}{32} \varepsilon \approx 0,34\varepsilon.$$

Wir sehen, daß zweimaliges Durchlaufen des Zyklus die Größen α und β mindestens um den Faktor $1/3$ verkleinern,

nach 20 Zyklen wird der Fehler also mindestens um den Faktor $1/3^{10} \approx 1/70\,000$ kleiner geworden sein.

Diese Überlegungen zeigen uns, daß nach 20 Zyklen mit einer Genauigkeit von $\pm 0,0001$ l im ersten Eimer 6 l, im zweiten 3 und im dritten ebenfalls 3 l enthalten sind.

24. Die Lösung eines Systems nichtlinearer Gleichungen mit der Methode der sukzessiven Approximation

Auch viele nichtlineare Gleichungssysteme können iterativ behandelt werden. Als Beispiel dafür schauen wir uns das System

$$\left. \begin{aligned} x &= 2 + \frac{x^2 + y}{20}, \\ y &= 1 + \frac{x + y^2}{20} \end{aligned} \right\} \quad (96)$$

an. Als Ausgangsnäherung wählen wir $x_0 = 0$, $y_0 = 0$, setzen diese Werte auf der rechten Seite ein und bekommen $x_1 = y_1 = 1$ als weitere Näherung. Erneutes Einsetzen führt auf

$$\begin{aligned} x_2 &= 2 + \frac{2^2 + 1}{20} = 2,25, \\ y_2 &= 1 + \frac{2 + 1^2}{20} = 1,15. \end{aligned}$$

Die Fortsetzung des Iterationsvorganges sieht dann so aus:

$$\begin{aligned} x_3 &= 2 + \frac{2,25^2 + 1,15}{20} = 2,31, \\ y_3 &= 1 + \frac{2,25 + 1,15^2}{20} = 1,18, \\ x_4 &= 2 + \frac{2,31^2 + 1,18}{20} = 2,33, \\ y_4 &= 1 + \frac{2,31 + 1,18^2}{20} = 1,18, \\ x_5 &= 2 + \frac{2,33^2 + 1,18}{20} = 2,33, \\ y_5 &= 1 + \frac{2,33 + 1,18^2}{20} = 1,18. \end{aligned}$$

Offenbar folgt aus den bis auf einen Fehler von maximal 0,01 geltenden Gleichungen $x_4 = x_5 = 2,33$ und $y_4 = y_5 = 1,18$, daß wir die gesuchte Lösung bereits erhalten haben.

So kann man auch dann vorgehen, wenn ein System

$$\begin{aligned}x &= \varphi(x, y), \\ y &= \psi(x, y)\end{aligned}\tag{97}$$

mit zwei beliebigen Funktionen φ und ψ gegeben ist. Wir wählen eine Anfangsnäherung x_0, y_0 , setzen sie rechts ein und verfahren immer weiter nach der Vorschrift

$$\begin{aligned}x_{n+1} &= \varphi(x_n, y_n), \\ y_{n+1} &= \psi(x_n, y_n).\end{aligned}\tag{98}$$

Wir brechen die Iteration ab, wenn innerhalb der betrachteten Stellenzahl $x_{n+1} = x_n$ und $y_{n+1} = y_n$ ist, denn dann gilt ebenfalls innerhalb dieser Grenzen $x = x_n$ und $y = y_n$.

Genauso lösen wir Gleichungssysteme mit drei oder mehr Unbekannten. (Die Frage der Konvergenz wird hier allerdings nicht so relativ einfach zu klären sein, wie bei einer Gleichung mit nur einer Unbekannten [Anm. d. dtsh. Red.])

Wir werden jetzt untersuchen, unter welchen Bedingungen die Konvergenz des Iterationsprozesses bei der Lösung von Systemen garantiert ist. Wir setzen dazu voraus, daß die Funktionen $\varphi(x, y)$ und $\psi(x, y)$ in einem gewissen begrenzten, geschlossenen Gebiet D der (x, y) -Ebene definiert sind. Mit anderen Worten, wir fordern, daß das Gebiet D samt seinen Randpunkten vollkommen im Inneren eines hinreichend groß gewählten Quadrates liegt. Um eine Vorstellung zu haben, kann man bei diesem Gebiet etwa an eine Kreisfläche, ein Polygon, eine Ellipse (sämtlich unter Einschluß der Randlinien) u. ä. denken.

Weiter werden wir annehmen, daß beide Funktionen in D stetig sind. Dann beschreiben sie gemeinsam eine Abbildung des Bereiches D auf einen gewissen anderen Bereich derselben Ebene. Die Koordinaten eines Bildpunktes N_0 ($\varphi(x_0), \psi(x_0)$) des Punktes $M_0(x_0, y_0) \in D$ bekommt man durch Einsetzen von x_0, y_0 , in φ und ψ .

Beispiel:

$$\begin{aligned}\varphi(x, y) &= x^2 + y^2, \\ \psi(x, y) &= 2xy.\end{aligned}$$

Der Bildpunkt von $M_0(1, 3)$ ist offensichtlich $N_0(10, 6)$.

Wir wollen die durch die beiden Funktionen φ und ψ vermittelte Abbildung von D im folgenden mit einem Buch-

staben Φ kennzeichnen und unter $\Phi(M)$ den Bildpunkt von M , unter $\Phi(D)$ das Bildgebiet von D verstehen.

Es sei dieses Bildgebiet in einem gewissen konkreten Fall ein Teilbereich D_1 von D selbst, dann kann man auf D_1 erneut die Abbildung Φ anwenden, wobei man ein Gebiet $D_2 = \Phi(D_1)$ als Bildgebiet erhält, welches wiederum ganz in D_1 enthalten sein muß. So fortfahrend können wir eine

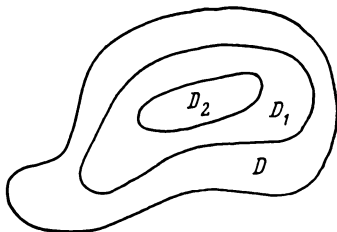


Abb. 21

ganze Folge D, D_1, D_2, \dots von Gebieten konstruieren, von denen jedes im vorangehenden enthalten und gleichzeitig dessen Bild ist (Abb. 21).

Wir nennen eine solche Abbildung Φ wieder kontraktiv, wenn es eine Zahl $q, 0 < q < 1$, gibt, so daß für zwei beliebige Punkte M_1 und M_2 aus dem Bereich D stets die Ungleichung

$$r(\Phi(M_1), \Phi(M_2)) \leq q \cdot r(M_1, M_2)$$

erfüllt ist. Dabei bezeichnen wir mit $r(P, Q)$ den Abstand der beiden Punkte P und Q .

Es ist nun sehr interessant, daß jetzt genau derselbe Satz wie im Falle einer Variablen bewiesen werden kann:

Wenn Φ eine Abbildung von D in sich ($\Phi(D) \subseteq D$) bewirkt und kontraktiv ist, dann gibt es in D einen einzigen Punkt N , für den $N = \Phi(N)$ ist. Er gehört allen Bildgebieten D_n an, und seine Koordinaten (ξ, η) erfüllen das Gleichungssystem (97). Dieser „Fixpunkt“ $N(\xi, \eta)$ kann durch sukzessive Approximation nach dem Schema

$$M_{n+1}(x_{n+1}, y_{n+1}) = \Phi(M_n(x_n, y_n)),$$

d. h.

$$x_{n+1} = \varphi(x_n, y_n),$$

$$y_{n+1} = \psi(x_n, y_n),$$

gewonnen werden. $M_0(x_0, y_0)$ ist dabei ein beliebiger Punkt aus D . Dann gilt:

$$\lim_{n \rightarrow \infty} x_n = \xi, \quad \lim_{n \rightarrow \infty} y_n = \eta, \quad \text{d. h. } M_n \rightarrow N.$$

25. Verschiedene Abstandsdefinitionen

Wie wir sahen, ist der kontraktive Charakter der Abbildung Φ eine Bedingung dafür, daß der Iterationsprozeß konvergiert. Diese Bedingung ist aber keinesfalls notwendige Voraussetzung für die Konvergenz!

Man kann durchaus Abbildungen angeben, die nicht kontraktiv sind und doch die Iteration gestatten. Ein Beispiel dafür ist die Abbildung

$$\Phi = \begin{cases} \varphi(x, y) = 1 + 2y, \\ \psi(x, y) = 3 + \frac{x}{8}, \end{cases}$$

die nicht kontraktiv ist, weil für $A(8, 0)$ und $B(8, 4)$ der Abstand $r(A, B) = 4$ ist und wegen $\Phi(A) = (1, 4)$ und $\Phi(B) = (9, 4)$

$$r(\Phi(A), \Phi(B)) = 8 > r(A, B)$$

gilt. Dennoch konvergiert die Folge M_0, M_1, M_2, \dots , M_0 beliebig vorgegeben, stets gegen $N\left(9\frac{1}{3}, 4\frac{1}{6}\right)$.

In einigen Fällen gelingt der Konvergenznachweis für den Iterationsprozeß dadurch, daß ein neues Abstandsmaß für zwei Punkte der Ebene eingeführt wird. Es handelt sich dabei durchaus um eine ganz natürliche und anschauliche Sache, denn auch im praktischen Leben kennen wir die Möglichkeit, den Abstand zweier Orte verschieden zu messen. (In der Zeit der Postkutsche wurde ganz natürlich die Entfernung zwischen zwei Orten in Stunden angegeben, wovon die sächsischen Postmeilensäulen beredete Kunde geben.) [Anm. d. dtsh. Red.] Für dieses Entfernungsmaß ist aber nun beispielsweise die gerade Verbindungslinie zwischen A und B keineswegs immer die Linie geringster Entfernung! Ein Beispiel hierfür ist in Abb. 22a gegeben, wo der „Abstand“ zwischen A und B die Summe der Abstände zwischen A und C , C und D bzw. D und B ist, wenn unterwegs eine Brücke zwischen C und D zu überqueren ist.

Ein weiteres Entfernungsmaß ergibt sich so:

In der Ebene sei ein rechtwinkliges Koordinatensystem gegeben, dann werde die Summe der Projektionen der Strecke \overline{AB} auf die beiden Achsen als „Abstand“ $r(A, B)$ genommen (Abb. 22 b). Eine weitere Definition legt die größere der beiden Strecken \overline{AC} oder \overline{CB} als „Abstand“ $r_1(A, B)$ fest, und so kann man sich noch eine ganze Menge weiterer Abstände ausdenken.¹⁾

Gewöhnlich fordert man von einem *Abstandsmaß* $r(A, B)$ für zwei Punkte die folgenden drei Eigenschaften:

- (1) Der Abstand $r(A, B)$ ist für zwei beliebige Punkte A, B als *nichtnegative, reelle Zahl* definiert und genau dann gleich null, wenn $A = B$ ist.

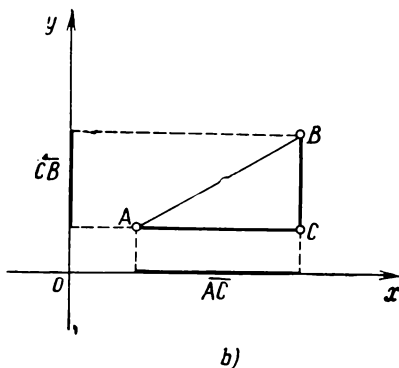
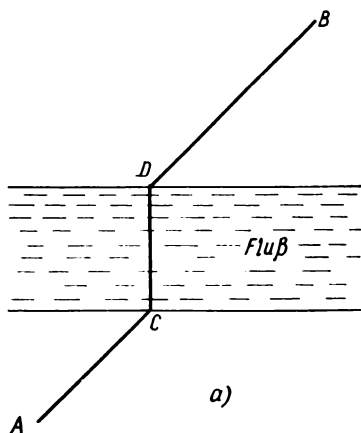


Abb. 22

- (2) Der Abstand ist eine *symmetrische Funktion* der beiden Punkte A und B , d. h. $r(A, B) = r(B, A)$.
 (3) Für drei beliebige Punkte A, B und C ist die „*Dreiecksungleichung*“

$$r(A, B) = r(A, C) + r(C, B)$$

erfüllt.

Ist in irgendeiner Menge von Objekten ein Abstand definiert, der diese Bedingungen erfüllt, dann nennt man die

¹⁾ Zur Erweiterung des Wissens über diesen Gegenstand empfehlen wir das Büchlein Belkner, H.: *Metrische Räume*. BSB. B. G. Teubner Verlagsgesellschaft, Leipzig 1972.

Menge einen metrischen Raum und ihre Elemente Punkte des Raums.

In diesem Sinne können auch Funktionen „Punkte“ eines metrischen Raumes sein: der Abstand zweier auf $[a, b]$ stetiger Funktionen $\varphi(x)$ und $\psi(x)$ ist das Maximum der Funktion $|\varphi(x) - \psi(x)|$ auf diesem Intervall:

$$r(\varphi, \psi) = \max_{a \leq x \leq b} |\varphi(x) - \psi(x)|.$$

Weiter oben hatten wir schon mehrere Möglichkeiten einer Abstandsdefinition in der Ebene kennengelernt, für welche die Bedingungen (1) — (3) erfüllt sind.

Wir möchten hier auf den interessanten Grenzfall hinweisen, wo zwar die Bedingungen (1) und (3), nicht aber (2) erfüllt sind. Das trifft z. B. dann zu, wenn wir den Abstand zweier Punkte A, B im Gebirge durch die Zeit messen, die man von A nach B benötigt. Offenbar wird hier in fast allen Fällen $r(A, B) \neq r(B, A)$ sein, denn man braucht bergauf wesentlich mehr Zeit als talwärts.

Es ist bewiesen worden, daß für die Konvergenz des Iterationsprozesses zur Lösung des Gleichungssystems

$$\begin{aligned} x &= \varphi(x, y), \\ y &= \psi(x, y) \end{aligned} \tag{99}$$

die folgenden Voraussetzungen hinreichend sind:
Die Funktion

$$\Phi = \begin{cases} \varphi(x, y) \\ \psi(x, y) \end{cases}$$

bildet D in sich ab und ist bezüglich irgendeiner Abstandsfunktion $r(A, B)$ kontraktiv, d. h. erfüllt die Bedingung

$$r(\Phi(M_1), \Phi(M_2)) \leq q \cdot r(M_1, M_2)$$

für alle Punkte $M_1, M_2 \in D$ und $0 < q < 1$.

Betrachten wir z. B. die Funktionen $\varphi(x, y) = 1 + 2y$ und $\psi(x, y) = 3 + \frac{x}{8}$. Wie wir sehen, ist die Abbildung Φ kontraktiv mit dem Faktor $q = \frac{1}{2}$, wenn wir den Abstand $r(A, B)$ zwischen den Punkten $A(x_1, y_1)$ und $B(x_2, y_2)$ wie folgt definieren:

$$\begin{aligned} r(A, B) &= \left| \frac{1}{2}(x_2 - x_1) + 2(y_2 - y_1) \right| + \\ &+ \left| \frac{1}{2}(x_2 - x_1) - 2(y_2 - y_1) \right|. \end{aligned}$$

Für unsere beiden Punkte $A(8, 0)$ und $B(8, 4)$ von vorhin beispielsweise wird $r(A, B) = 16$, und für ihre Bilder $\Phi(A)$ und $\Phi(B)$ erhalten wir $r(\Phi(A), \Phi(B)) = 8$. In dieser „Metrik“ ist die Abbildung Φ somit kontraktiv, und die Konvergenz des Iterationsverfahrens zur Lösung des Gleichungssystems

$$\begin{aligned}x &= 1 + 2y \\y &= 3 + x/8\end{aligned}$$

ist gesichert, obwohl die Abbildung Φ bezüglich der gewöhnlichen Abstandsdefinition nicht kontraktiv ist.

26. Konvergenzkriterien für den Iterationsprozeß bei linearen Gleichungssystemen

Wir wenden jetzt den im vorigen Abschnitt formulierten Konvergenzatz auf lineare Gleichungssysteme an.

Indem wir verschiedenartige Abstandsbegriffe einführen, erhalten wir für solche Systeme Konvergenzkriterien, zu deren Überprüfung lediglich die gegebenen Koeffizienten herangezogen werden müssen.

Betrachten wir zunächst Systeme zweier linearer Gleichungen mit zwei Unbekannten:

$$\begin{aligned}a_{11}x + a_{12}y &= b_1, \\a_{21}x + a_{22}y &= b_2.\end{aligned}\tag{100}$$

Wenn wir voraussetzen, daß $a_{11} \neq 0$ und $a_{22} \neq 0$ ist, dann können wir die erste Gleichung nach x , die zweite nach y auflösen und erhalten das folgende neue System

$$\begin{aligned}x &= \frac{b_1}{a_{11}} - \frac{a_{12}}{a_{11}} y, \\y &= \frac{b_2}{a_{22}} - \frac{a_{21}}{a_{22}} x.\end{aligned}$$

Schreiben wir hier abkürzend β_1 für b_1/a_{11} , β_2 für b_2/a_{22} , α_1 für $-a_{12}/a_{11}$ und α_2 für $-a_{21}/a_{22}$, dann geht (100) über in

$$\left. \begin{aligned}x &= \alpha_1 y + \beta_1, \\y &= \alpha_2 x + \beta_2.\end{aligned} \right\}\tag{100'}$$

Die betrachtete Abbildung Φ hat demzufolge jetzt eine ganz einfache Gestalt, nämlich

$$\varphi = \begin{cases} \varphi(x, y) = \alpha_1 y + \beta_1, \\ \psi(x, y) = \alpha_2 x + \beta_2. \end{cases}$$

Welchen Bedingungen müssen nun die Koeffizienten α_1 und α_2 genügen, damit diese Abbildung kontraktiv ist?

Bekanntlich wird der geometrische Abstand $r(A, B)$ zwischen zwei Punkten $A(x_1, y_1)$ und $B(x_2, y_2)$ nach der Formel

$$r(A, B) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

ausgerechnet. Die Transformation führt den Punkt A über in $A_1(\alpha_1 y_1 + \beta_1, \alpha_2 x_1 + \beta_2)$ und den Punkt B in den Punkt $B_1(\alpha_1 y_2 + \beta_1, \alpha_2 x_2 + \beta_2)$. Dann ist der Abstand der beiden Bildpunkte durch

$$\begin{aligned} r(A_1, B_1) &= \sqrt{(\alpha_1 y_2 - \alpha_1 y_1)^2 + (\alpha_2 x_2 - \alpha_2 x_1)^2} \\ &= \sqrt{\alpha_1^2 (y_2 - y_1)^2 + \alpha_2^2 (x_2 - x_1)^2} \end{aligned} \quad (101)$$

gegeben. Bezeichnen wir die größte der beiden positiven Zahlen $|\alpha_1|$ und $|\alpha_2|$ mit q :

$$q = \max(|\alpha_1|, |\alpha_2|),$$

dann folgt aus (101) die Ungleichung

$$r(A_1, B_1) \leq q \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} = qr(A, B).$$

Ist $q < 1$, dann konvergiert das Iterationsverfahren, wie wir wissen, denn nach der letzten Ungleichung ist die Abbildung Φ eine kontraktive Abbildung der ganzen Ebene auf sich.

Wir haben damit die Aussage gewonnen, daß das Iterationsverfahren zur Lösung des Gleichungssystems (100') immer konvergiert, wenn der größere der beiden Koeffizientenbeträge kleiner als eins ist:

$$\max(|\alpha_1|, |\alpha_2|) < 1. \quad (102)$$

Rechnen wir diese Bedingung wieder auf die ursprünglichen Koeffizienten in (100) um, so erhalten wir, da

$$\begin{aligned} \alpha_1 &= -\frac{a_{12}}{a_{11}}, \\ \alpha_2 &= -\frac{a_{21}}{a_{22}}, \end{aligned}$$

den folgenden Konvergenzsatz:

Für die Konvergenz des Iterationsverfahrens zur Lösung des linearen Gleichungssystems

$$\begin{aligned} a_{11}x + a_{12}y &= b_1, \\ a_{21}x + a_{22}y &= b_2 \end{aligned}$$

ist hinreichend, daß

$$\max \left(\left| \frac{a_{12}}{a_{11}} \right|, \left| \frac{a_{21}}{a_{22}} \right| \right) < 1$$

erfüllt ist.

Diese Regel drückt aus, daß die Koeffizienten der *Hauptdiagonalen* betragsmäßig größer sein müssen, als die Koeffizienten außerhalb der Diagonalen in der gleichen Zeile.

Wollen wir beispielsweise das System

$$\begin{aligned} x - 3y &= -11, \\ 6x + y &= 10 \end{aligned}$$

iterativ lösen, dann müssen wir die erste Gleichung nach y und die zweite nach x auflösen:

$$y = \frac{11}{3} + \frac{1}{3}x; \quad x = \frac{5}{3} - \frac{y}{6}.$$

Manchmal ist es nützlich, die Unbekannten eines Systems zunächst durch neue zu ersetzen, damit die Regel erfüllt werden kann. Nehmen wir etwa das System

$$\left. \begin{aligned} 12x + y &= 14, \\ 3x - 2y &= -1, \end{aligned} \right\} \quad (103)$$

für das

$$\max \left(\left| \frac{a_{12}}{a_{11}} \right|, \left| \frac{a_{21}}{a_{22}} \right| \right) = \max \left(\frac{1}{12}, \frac{3}{2} \right) = \frac{3}{2},$$

gilt. Die Konvergenzbedingung ist damit nicht erfüllt! Setzen wir aber $x = \frac{1}{3}z$, dann bekommen wir

$$\left. \begin{aligned} 4z + y &= 14, \\ z - 2y &= -1, \end{aligned} \right\} \quad (104)$$

und hierfür ist

$$\max \left(\left| \frac{a_{12}}{a_{11}} \right|, \left| \frac{a_{21}}{a_{22}} \right| \right) = \max \left(\frac{1}{4}, \frac{1}{2} \right) = \frac{1}{2}.$$

Also können wir das neue Gleichungssystem (104) durch Iteration lösen.

Es versteht sich von selbst, daß niemand ernsthaft daran denkt, ein solch einfaches System mit zwei Unbekannten wie (104) iterativ zu lösen. Der Vorteil der Methode kommt aber sofort zum Tragen, wenn sehr große Systeme zu lö-

So entspricht der Ausdruck (106) dem Abstandsmaß

$$r(A, B) = \max(|x_1 - y_1|, \dots, |x_n - y_n|),$$

der Ausdruck (107) entspricht

$$r(A, B) = \sum_{i=1}^n |x_i - y_i|$$

und (108) schließlich entspricht

$$r(A, B) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2},$$

wobei in jedem Fall die Punkte $A(x_1, \dots, x_n)$ und $B(y_1, \dots, y_n)$ gemeint sind.

Genau wie im zweidimensionalen Fall ist es manchmal nützlich, die Unbekannten x_1, x_2, \dots, x_n durch neue Variable $y_j = p_j x$ ($j = 1, 2, \dots, n$) mit $p_j > 0$ zu ersetzen. Dann lauten die Regeln (106) – (108)

$$1') \quad \max_i \sum_{j=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \frac{p_i}{p_j} < 1, \quad (106')$$

$$2') \quad \max_j \sum_{i=1}^n \left| \frac{a_{ij}}{a_{ii}} \right| \frac{p_i}{p_j} < 1, \quad (107')$$

$$3') \quad \max_k \sum_{i,j=1}^n \binom{(k)}{i,j} \left| \frac{a_{ij}}{a_{ii}} \right| \frac{2p_i^2}{p_j^2} < 1, \quad (108')$$

woraus mit $p_i = |a_{ii}|$ folgt

$$1'') \quad \max_i \sum_{j=1}^n \left| \frac{a_{ij}}{a_{jj}} \right| < 1, \quad (106'')$$

$$2'') \quad \max_j \sum_{i=1}^n \left| \frac{a_{ij}}{a_{jj}} \right| < 1, \quad (107'')$$

$$3'') \quad \max_k \sum_{i,j=1}^n \binom{(k)}{i,j} \left| \frac{a_{ij}}{a_{jj}} \right| < 1, \quad (108'')$$

Um auch hier ein Beispiel vor Augen zu haben, schauen wir uns das System

$$\begin{aligned}x - 0,6y - 0,5z &= -2,6, \\ -0,2x + y - 0,4z &= 3, \\ -0,1x + 0,5y + z &= 3,9\end{aligned}$$

an.

Hier sind weder (106) und (107), noch (106'') und (107'') erfüllt. Ebenso gilt die Ungleichung (109) nicht, denn die Summe der Quadrate der Nichtdiagonalelemente ist gleich 1,07... . Die Gleichung

$$\begin{aligned}\max_k \sum^{(k)} \left| \frac{a_{ij}}{a_{ii}} \right|^2 &= \max (0,6^2 + 0,5^2 + 0,2^2 + 0,4^2; \\ &0,6^2 + 0,5^2 + 0,1^2 + 0,5^2; \quad 0,2^2 + 0,4^2 \\ &+ 0,1^2 + 0,5^2) = 0,87 < 1\end{aligned}$$

zeigt aber, daß dieses System mit der Iterationsmethode lösbar ist.

Zum Schluß wollen wir nochmals hervorheben, daß alle bisher formulierten Bedingungen für die Konvergenz *hinreichende* Bedingungen sind und sich in keinem Falle als notwendige Voraussetzungen erwiesen haben. Andererseits können wir durch die Wahl neuer Abstandsdefinitionen beliebig viele weitere Konvergenzkriterien ableiten.

Wir werden aber die hiermit zusammenhängenden interessanten Fragen in diesem Buch nicht weiter verfolgen.

Im Übrigen gelten für lineare Gleichungssysteme die gleichen Bemerkungen, wie wir sie im Kapitel 5 gemacht hatten.

So hängt das Ergebnis der Iteration nicht von der Wahl der Anfangsnäherung ab, Rechenfehler können somit das Endergebnis niemals verfälschen, sondern nur die Zeit bis zu dessen Gewinnung verlängern.

Es gibt eine Reihe von Varianten des Näherungsverfahrens zur Lösung linearer Gleichungssysteme. So wird beispielsweise bei einigen Verfahren bei der Berechnung von $x_2^{(m+1)}$ außer $x_3^{(m)}, \dots, x_n^{(m)}$ bereits das gerade zuvor neu berechnete $x_1^{(m+1)}$ in die rechte Seite eingesetzt, $x_3^{(m+1)}$ bekommt man unter Verwendung von $x_1^{(m+1)}, x_2^{(m+1)}, x_4^{(m)}, \dots, x_n^{(m)}$ und so fort.

Es würde aber ein eigenes Buch erfordern, wollten wir alle möglichen Näherungsverfahren für lineare Gleichungssysteme aufschreiben.

27. Sukzessive Approximation in der Geometrie

Wir haben ausführlich dargestellt, wie iterative Methoden zur Lösung von Gleichungen und Gleichungssystemen verwendet werden und wollen es nicht versäumen, auf einige geometrische Anwendungen einzugehen.

Es dürfte ziemlich allgemein bekannt sein, daß man die Länge des Kreisumfangs iterativ ermitteln muß. Dazu berechnet man zuerst den Umfang eines dem Kreis einbeschriebenen Quadrates, danach den Umfang eines regulären 8-Ecks, eines 16-Ecks usw. Der Grenzwert dieser Umfänge muß mit dem Kreisumfang übereinstimmen.

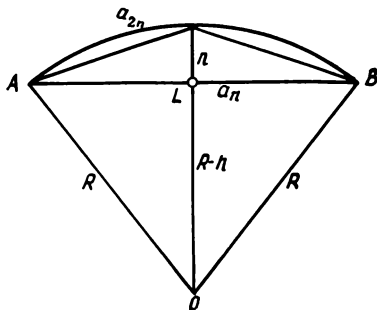


Abb. 23

Dabei berechnet man jeden folgenden Umfang mit Hilfe des vorangegangenen. Das geschieht so:

Wir nennen den Umfang des einbeschriebenen regelmäßigen 2^n -Ecks P_n und eine Seite A_n . Zum Beispiel ist A_2 die Seitenlänge des Quadrates im Kreis und beträgt somit $A_2 = R\sqrt{2}$, woraus $P_2 = 4R\sqrt{2}$ folgt.

Setzen wir voraus, daß uns P_n bereits bekannt ist, dann ist offenbar $A_n = P_n/2^n$. Man kann relativ leicht beweisen, daß sich die Seitenlänge a_{2n} eines dem Kreis vom Radius R einbeschriebenen regelmäßigen $2n$ -Ecks aus der Seitenlänge a_n des entsprechenden n -Ecks nach der Formel

$$a_{2n} = R \sqrt{2 - \sqrt{4 - \frac{a_n^2}{R^2}}} \quad (110)$$

berechnen läßt.

Wir betrachten dazu die Abbildung 23 und finden:

$$\left(\frac{a_n}{2}\right)^2 + (R-h)^2 = R^2 \quad \text{und} \quad a_{2n}^2 = (a_n/2)^2 + h^2.$$

Daraus folgt

$$(a_n/2)^2 + h^2 - 2Rh = 0 \quad \text{und deshalb} \quad a_{2n}^2 = 2Rh$$

und

$$h = R - \sqrt{R^2 - (a_n/2)^2}.$$

Daraus folgt nun aber insgesamt

$$a_{2n}^2 = 2R^2 - 2R \sqrt{R^2 - a_n^2/4}, \quad \text{d. h.}$$

$$a_{2n}^2 = R^2 \left(2 - \sqrt{4 - a_n^2/R^2} \right),$$

$$a_{2n} = R \sqrt{2 - \sqrt{4 - a_n^2/R^2}},$$

was wir beweisen wollten. ¹⁾

Diese Formel besagt aber, daß sich die Seitenlänge A_{n+1} des regulären 2^{n+1} -Ecks aus der des regulären 2^n -Ecks nach der Formel

$$A_{n+1} = R \sqrt{2 - \sqrt{4 - \frac{A_n^2}{R^2}}}$$

bestimmt. Mit $A_n = \frac{P_n}{2^n}$ und $A_{n+1} = \frac{P_{n+1}}{2^{n+1}}$ haben wir

$$P_{n+1} = 2^{n+1}R \sqrt{2 - \sqrt{4 - \frac{P_n^2}{2^{2n}R^2}}}. \quad (110')$$

Die Folge der Zahlen $P_2, P_3, \dots, P_n, \dots$ konvergiert gegen den Kreisumfang, d. h. gegen $2\pi R$. Also kann die Formel (110) als Iterationsformel zur Berechnung von $2\pi R$ angesehen und verwendet werden, womit uns eine Möglichkeit für die Berechnung von π auf beliebig viele Dezimalstellen gegeben ist.

Es gibt auch noch eine weitere Näherungsmethode zur Berechnung dieser wichtigen Zahl, die man „Methode des konstanten Umfangs“ nennt. Hierbei wird ein regelmäßiges 2^n -Eck in ein regelmäßiges 2^{n+1} -Eck mit gleichem Umfang verwandelt.

Wir bezeichnen den Inkreisradius des regelmäßigen 2^n -Ecks mit l_n , den Umkreisradius mit r_n . Den Inkreisradius des regelmäßigen 2^{n+1} -Ecks, das den gleichen Umfang besitzt, bezeichnen wir mit l_{n+1} , den Radius des Umkreises mit r_{n+1} und kommen jetzt zur Konstruktion.

¹⁾ Der Beweis wurde gegenüber dem Original geändert, um ohne Verwendung der Trigonometrie auszukommen [Anm. d. dtsh. Red.].

Es sei AB eine Seite des 2^n -Ecks, d. h. $\overline{AO} = r_n$ und $\overline{AF} = l_n$. Wir halbieren den Bogen AB in C und verbinden C mit A und B , halbieren die Seiten AC und BC in D bzw. E und verbinden beide Punkte. Man bestätigt leicht, daß

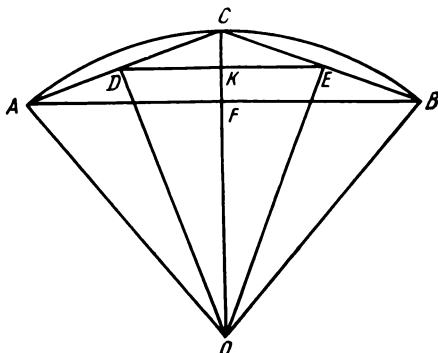


Abb. 24

$\sphericalangle DOE = \frac{1}{2} \sphericalangle AOB$ ist. Damit aber DE Seite eines regulären 2^{n+1} -Ecks mit dem Umkreisradius $r_{n+1} = \overline{OD}$ und dem Inkreisradius $l_{n+1} = \overline{OK}$.

Wegen $\overline{DE} = \frac{1}{2} \overline{AB}$ ist der Umfang dieses 2^{n+1} -Ecks gleich dem des 2^n -Ecks, von dem wir ausgegangen sind. Es läßt sich leicht nachrechnen, daß

$$l_{n+1} = \frac{r_n + l_n}{2}. \quad (111)$$

Ferner gilt im rechtwinkligen Dreieck $\triangle ODC$ der Kathetensatz $\overline{OD}^2 = \overline{OK} \cdot \overline{OC}$, mit anderen Symbolen

$$r_{n+1} = \sqrt{r_n \cdot l_{n+1}}. \quad (112)$$

Beide Formeln zusammen gestatten die Berechnung von r_{n+1} und l_{n+1} aus r_n und l_n .

Mit wachsendem n nähern sich die Zahlen r_n und l_n einem gemeinsamen Grenzwert, während sich der Umfang des Polygons nicht ändert. Deshalb muß dieser Grenzwert mit dem Radius desjenigen Kreises zusammenfallen, der ebenfalls diesen fest vorgegebenen Umfang hat.

Wählen wir das Anfangspolygon so, daß der Umfang gleich zwei ist, dann konvergiert sowohl $\{r_n\}$ als auch $\{l_n\}$ gegen die Zahl $1/\pi$: $\lim_{n \rightarrow \infty} r_n = \frac{1}{\pi}$; $\lim_{n \rightarrow \infty} l_n = \frac{1}{\pi}$.

Beginnen wir etwa mit dem Quadrat der Seitenlänge $1/2$, dann wird $r_2 = \frac{1}{4}\sqrt{2}$, $l_2 = \frac{1}{4}$. Berechnen wir davon ausgehend die Werte r_{n+1} , l_{n+1} , für $n = 2, 3, \dots$ nach den Formeln (111) und (112), so ist

$$\lim_{n \rightarrow \infty} r_n = \lim_{n \rightarrow \infty} l_n = \frac{1}{\pi}$$

Also können wir den Wert $\frac{1}{\pi}$ mit Hilfe dieser Formeln näherungsweise bestimmen, wenn wir die Berechnung so lange fortsetzen, bis sich r_n und l_n bei vorgegebener Genauigkeit nicht mehr unterscheiden. Dann sind diese Werte r_n und l_n Näherungen von $\frac{1}{\pi}$ mit geforderter Genauigkeit.

28. Schlußbemerkungen

Wir haben uns in diesem Büchlein mit einer Reihe von Anwendungen der schrittweisen Näherung vertraut gemacht, so etwa bei der Aufstellung von Plänen, bei der Berechnung der Wurzel einer Zahl, der Lösung von Gleichungen oder schließlich bei der Umfangsberechnung eines Kreises.

Damit ist die Vielfalt der Anwendungsmöglichkeiten jedoch längst nicht ausgeschöpft worden.

Sehr viele praktisch wichtige Aufgaben führen auf Differentialgleichungen (in denen es um die Bestimmung einer unbekanntten Funktion geht), auf Integralgleichungen und auf noch kompliziertere Gleichungen.

Auch bei der näherungsweisen Lösung eines solchen Problems bedient man sich mit großem Erfolg der Methode der sukzessiven Approximation. Es versteht sich wohl von selbst, daß deren Durchführung in diesen Fällen weitaus schwieriger ist, als etwa im Falle der Lösung einer algebraischen Gleichung. Dennoch kann man ohne Übertreibung sagen, daß ohne die Anwendung der Iteration keine der grandiosen physikalisch-technischen Leistungen hätte vollbracht werden können, die gerade in den letzten Jahren so großes Aufsehen erregt haben.

Ob bei der Bahnberechnung der Sputniks, der Berechnung von Atomreaktoren oder der Untersuchung des Baues eines Atoms — Iterationsverfahren waren immer im Spiel.

Die Erläuterung dieser außerhalb der Grenzen der Elementarmathematik liegenden Anwendungen der Methode der schrittweisen Näherung übersteigt allerdings den Rahmen dieser kleinen Schrift.

Aufgaben

Wir geben einige Aufgaben an, mit denen der Leser überprüfen soll, in welchem Grade er die Darstellungen des Buches verstanden hat und diese anzuwenden versteht. (1) Die nachfolgenden Gleichungen sollen durch Iteration gelöst werden. Dabei sind einige von ihnen zuerst in eine entsprechende geeignete Form zu bringen.

$$1. \quad x = \frac{1}{(x+1)^2}.$$

$$2. \quad x = (x+1)^3.$$

$$3. \quad x = 4 + \sqrt[3]{\frac{x-1}{x+1}}.$$

$$4. \quad x = 2 \pm \sqrt[4]{x}.$$

$$5. \quad x = \sqrt[3]{5-x}.$$

$$6. \quad 4-x = \tan x.$$

$$7. \quad x^2 = \sin x.$$

$$8. \quad x^3 = \sin x.$$

$$9. \quad x = \arcsin \frac{x+1}{4}.$$

$$10. \quad x = \cos x.$$

$$11. \quad x = \frac{1}{\cos x}.$$

$$12. \quad x = \frac{1}{10} \sin x.$$

$$13. \quad x = \pm \sqrt{\lg(x+2)}.$$

$$14. \quad x^2 = \ln(x+1).$$

$$15. \quad \ln x = 4 - x^2.$$

$$16. \quad \ln x = 2 - x.$$

$$17. \quad x^2 = e^x + 2.$$

$$18. \quad \lg x = 0,1 x.$$

$$19. \quad \tan x = \lg x.$$

$$20. \quad x = \frac{1}{10} e^{-x}.$$

(2) Es folgen einige Aufgaben, die nach dem Newtonschen Verfahren behandelt werden sollen:

$$21. \quad x^3 - 5x + 1 = 0.$$

$$24. \quad x^5 + 5x + 1 = 0.$$

$$22. \quad x^3 - 9x^2 + 20x - 11 = 0. \quad 25. \quad \sin x + x = 1.$$

$$23. \quad x^3 - 3x^2 - 3x + 11 = 0. \quad 26. \quad x^2 - 10 \lg x - 3 = 0.$$

(3) Wir geben zum Schluß noch drei Gleichungssysteme, die mit einer Genauigkeit von $\pm 0,001$ durch sukzessive Appro-

Iteration zu lösen sind:

$$27. \quad \begin{cases} x = 0,2y - 0,1z + 0,89, \\ y = 0,3x + 0,15z + 1,383, \\ z = 0,25x - 0,4y + 3,677. \end{cases}$$

$$28. \quad \begin{cases} x = \frac{1}{4} \sin(x+y) + 0,336, \\ y = -\frac{1}{4} \sin(x-y) + 0,362. \end{cases}$$

$$29. \quad \begin{cases} x = \sqrt{x+2y} - 0,710, \\ y = \sqrt{y-x} + 1. \end{cases}$$

Lösungen

Anmerkung d. dtsh. Red. zu den Lösungen:

Die Darstellung ist gegenüber dem Original in vielen Fällen gestrafft worden. Dies geschah, damit der Leser bei der Durcharbeitung der Aufgaben auch unter Verwendung der Lösungshinweise nochmals zum intensiven Nachdenken angeregt wird.

1. $\varphi(x) = (x+1)^{-2}$, $\varphi'(x) = -2(x+1)^{-3}$.
 $\varphi(0) = 1 > 0$, $\varphi(1) = 1/4 < 1$, eine Wurzel liegt deshalb sicher in $[0, 1]$. Die Iteration ist aber nicht im ganzen Intervall konvergent wegen $|\varphi'(0)| = 2 > 1$. Wir verkleinern deshalb das Intervall unter Beachtung von $\varphi(0,4) = 1/1,96 > 0,4$. Für $0,4 \leq x \leq 1$ ist $|\varphi'(x)| \leq \frac{2}{1,4^3} < 1$, und die Iteration konvergiert. Mit $x_1 = 0,4$ wird $x_{11} \approx x_{12} \approx 0,4655$, d. h. $x = 0,4655 \pm 0,001$ ist die Lösung.
 2. $\varphi(x) = (x+1)^3$, $\varphi'(x) = 3(x+1)^2$; $\varphi(-2) = -1 > -2$, $\varphi(-3) = -8 < -3$, eine Wurzel liegt in $[-3, -2]$. Iteration ist nicht anwendbar, da $|\varphi'(x)| > 1$ im ganzen Intervall. Nach Umformung auf $x = \sqrt[3]{x} - 1$ ist Iteration möglich, denn für $\psi(x) = \sqrt[3]{x} - 1$ gilt $\psi'(x) = (\sqrt[3]{x^2})^{-1}$, und das ist im ganzen Intervall kleiner als eins. Mit $x_1 = -2$ wird $x_6 \approx x_7 \approx -2,325$ — die Lösung bis auf einen Fehler von $\pm 0,001$.

$$3. \varphi(x) = 4 + \sqrt[3]{\frac{x-1}{x+1}}, \quad \varphi'(x) = \frac{2}{3\sqrt[3]{(x-1)^2(x+1)^4}}.$$

Abb. 25 zeigt, daß die Gerade $y = x$ genau zwei Schnittpunkte mit der Kurve $y = \varphi(x)$ hat, die im Intervall $[-1, 0]$

und $[4,5]$ liegen. In $[4,5]$ haben wir $|\varphi'(x)| \leq \frac{2}{15\sqrt[3]{45}} < 1$. Mit $x_1 = 4$ wird $x_3 \approx x_4 \approx 4,870$, also $x = 4,870 \pm \pm 0,001$. Im anderen Intervall ist die Methode nicht konvergent. Umformung von $(x-4)^3 = \frac{x-1}{x+1}$ auf

$$x = \frac{x-1}{(x-4)^3} - 1$$

führt auf eine geeignete Form.

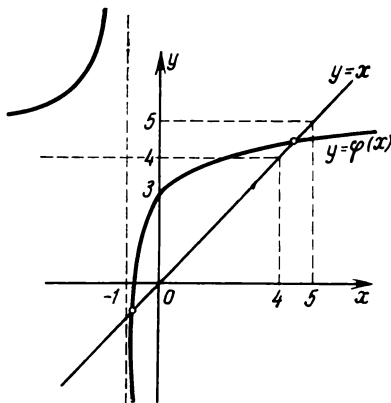


Abb. 25

Es ist $\psi(x) = \frac{x-1}{(x-4)^4} - 1$ und $\psi(x) = \frac{-2x-1}{(x-4)^4}$. Für $-1 \leq x < 0$ erhalten wir $|\psi'(x)| \leq \frac{1}{256} < 1$ und können nun die Methode der schrittweisen Näherung anwenden. Mit $x_1 = 0$ wird $x_2 = x_3 = -0,9840 \pm 0,0001$. Damit haben wir die beiden Wurzeln $x = -0,9840$ und $x = 4,870$ bestimmt.

4. Hier ist $\varphi_1(x) = 2 + \sqrt[4]{x}$ und $\varphi_2(x) = 2 - \sqrt[4]{x}$. Abb. 26 läßt eine Wurzel von $x = \varphi_1(x)$ im Intervall $[3, 4]$ erkennen, und dort ist in der Tat $|\varphi'(x)| = \frac{1}{4\sqrt[4]{x^3}} < 1$. Die

Iteration konvergiert aus diesem Grunde, und man erhält bei $x_1 = 4$ mit der verlangten Genauigkeit $x_3 \approx x_4 \approx \approx 3,353$, d. h. $x = 3,353 \pm 0,001$.

Die zweite Gleichung hat ebenfalls nur eine Lösung (wie aus der Abbildung zu sehen ist), und diese ist $x = 1$, wie man sofort nachrechnet. Damit sind die entsprechenden Lösungen: $x = 3,353$ und $x = 1$.

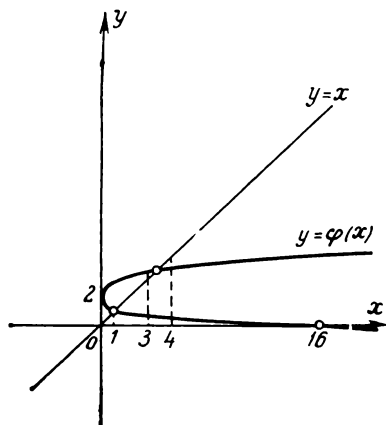


Abb. 26

$$5. \quad \varphi(x) = \sqrt[3]{(5-x)},$$

$$\varphi'(x) = \frac{-1}{3\sqrt[3]{(5-x)^2}}.$$

$\varphi(1) = \sqrt[3]{4} > 1$, $\varphi(2) = \sqrt[3]{3} < 2$. Die somit im Intervall $[1, 2]$ sicher vorhandene Lösung gewinnen wir wegen $|\varphi'(x)| \leq \frac{1}{3\sqrt[3]{9}} < 1$ mit $x_1 = 1$ zu $x_5 = x_6 = x = 1,516 \pm \pm 0,001$.

6. In der vorliegenden Form nicht kontraktiv. Wir formen um: $x = \arctan(4-x) = \varphi(x)$. Man erhält $\varphi(1) \approx 1,25$, $\varphi(2) = \arctan 2 \approx 1,10$. Die somit in $[1, 2]$ sicher vorhandene Wurzel kann wegen $|\varphi'(x)| = \frac{1}{1+(4-x)^2} \leq \frac{1}{2} < 1$ iterativ ermittelt werden. Mit $x_1 = 1$ wird $x = x_3 = x_4 = 1,225 \pm 0,001$ die gesuchte Näherung.

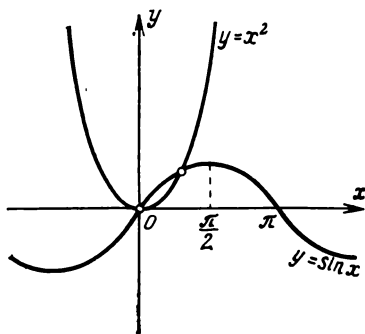


Abb. 27

7. Die gegebene Gleichung hat $x = 0$ als eine Wurzel. Abb. 27 zeigt, daß es noch eine weitere Wurzel gibt, die der Gleichung $x = \sqrt{\sin x}$ genügt. Es ist

$\varphi(x) = \sqrt{\sin x}$ und $\varphi'(x) = \frac{\cos x}{2\sqrt{\sin x}}$. Man überzeugt sich,

daß die Nullstelle im Intervall $\left[\frac{1}{2}, 1\right]$ zu suchen ist, und

dort gilt

$$|\varphi'(x)| \leq \frac{\cos \frac{1}{2}}{2 \sqrt{\sin \frac{1}{2}}} \approx \frac{0,8703}{1,3846} < 1,$$

woraus die Konvergenz des Iterationsverfahrens folgt. $x_1 = 1$ führt auf $x_7 \approx x_8 \approx 0,8768$, also ist $x = 0,8768 \pm \pm 0,0001$ die gesuchte Näherung.

8. Auch diese Gleichung hat $x = 0$ als Wurzel. Nach dem Vorbild von Aufgabe 7 schreiben wir $x = \sqrt[3]{\sin x}$. Mit $x_1 = 1$ erhält man $x_8 \approx x_7 \approx 0,9286$, also $x = 0,9286 \pm \pm 0,0001$. Da die beiden Funktionen $y = x^3$ und $y = \sin x$

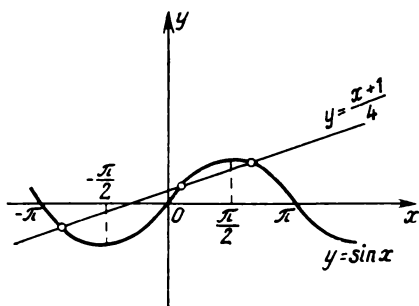


Abb. 28

ungerade sind, erhalten wir noch die Wurzel $x = -0,9286 \pm \pm 0,0001$.

9. In der Form $\frac{1}{4}(x+1) = \sin x$ ergibt sich für die einzige in $[0, \frac{\pi}{2}]$ gelegene Wurzel (Abb. 28) wegen $\varphi(x) = \arcsin \frac{x+1}{4}$ und $\varphi'(x) = \frac{1}{\sqrt{16-(x+1)^2}}$, $|\varphi'(x)| < \frac{1}{\sqrt{3}} < 1$, ein konvergentes Iterationsverfahren, das mit $x_1 = 0$ bei $x_8 = x_9 = 0,3422 \pm 0,0001$ mit der gewünschten Genauigkeit abgebrochen wird.

10. Die Wurzel ist wegen $\cos 0 = 1$ und $\cos 1 > 0$ im Intervall $[0, 1]$ zu suchen. Wegen $|\varphi'(x)| \leq |\sin 1| < 1$ konvergiert die sukzessive Approximation: $x_1 = 1$, $x = 0,7391 \pm 0,0001$.

11. $\cos x = 1/x$. Abb. 29 zeigt, daß die Wurzeln der Gleichung nahe bei den Schnittpunkten des Graphen von $y = \cos x$ mit der Abszissenachse liegen und zwar rechts von ihnen in der Nähe der Punkte $\frac{\pi}{2} + (2k + 1)\pi$ und links davon bei $\frac{\pi}{2} + 2k\pi$. Um die Lösung zu berechnen, die in

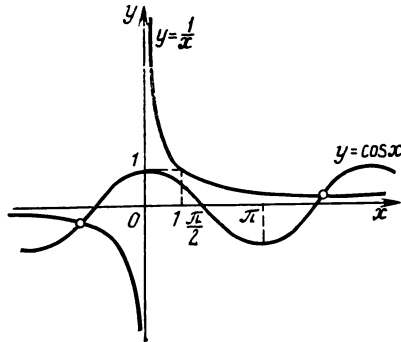


Abb. 29

der Nähe von $x = \frac{\pi}{2} + n\pi$ liegt, setzen wir $x = n\pi + \frac{\pi}{2} + y$, womit die Gleichung eine neue Form annimmt:

$$y + n\pi + \frac{\pi}{2} \frac{1}{\cos\left(y + n\pi + \frac{\pi}{2}\right)} = \frac{(-1)^{n+1}}{\sin y}.$$

Dabei ist $-\pi/2 \leq y \leq \pi/2$, und die Gleichung kann auch so geschrieben werden:

$$y = (-1)^{n+1} \arcsin \frac{1}{y + n\pi + \frac{\pi}{2}}.$$

Jetzt ist

$$\varphi(y) = (-1)^{n+1} \arcsin \frac{1}{y + n\pi + \frac{\pi}{2}}$$

und

$$\varphi'(y) = \frac{(-1)^n}{\sqrt{\left(y + n\pi + \frac{\pi}{2}\right)^2 - 1} \left(y + n\pi + \frac{\pi}{2}\right)}.$$

Nah bei $y = 0$ ist sicher $|\varphi'(y)| \leq q < 1$, und deshalb konvergiert die Iterationsfolge. So ist die Lösung für $n = 1$ mit $y_0 = 0$ zu $y_2 \approx y_3 \approx 0,204$ berechnet. Demnach gilt $y = 0,204 \pm 0,001$, $x = \frac{3}{2} \pi + y = 4,917 \pm 0,001$.

Um die erste negative Wurzel zu bekommen, setzen wir $n = -1$. Wir erhalten die Gleichung

$$y = \arcsin \frac{1}{y - \frac{\pi}{2}}.$$

Für die Ausgangsnäherung $y_0 = 0$ erhalten wir $y_6 \approx y_7 \approx 0,503$, also $y = 0,503 \pm 0,001$ und deshalb $x = -2,074 \pm 0,001$.

Für große Werte von $|n|$ liefert das Iterationsverfahren eine Näherungsformel für y :

$$y = \varphi(y_0) = (-1)^{n+1} \arcsin \frac{1}{n\pi + \frac{\pi}{2}} \approx \frac{(-1)^{n+1} \cdot 2}{(2n+1)\pi}$$

und x :

$$x \approx \frac{\pi}{2} (2n+1) + \frac{(-1)^{n+1} \cdot 2}{(2n+1)\pi}.$$

12. Mit $x_1 = 0$ wird $x = x_3 = x_4 = 1,088$ bis auf einen Fehler vom maximalen Betrag 0,001 die Lösung.

13. Zuerst lösen wir die Gleichung $x = \sqrt{\lg(x+2)}$ und haben $\varphi(x) = \sqrt{\lg(x+2)}$ mit

$$\varphi'(x) = \frac{\lg e}{2(x+2)\sqrt{\lg(x+2)}}.$$

Wegen $\varphi(0) = \sqrt{\lg 2} \stackrel{!}{\approx} 0$, $\varphi(1) = \sqrt{\lg 3} < 1$, ist eine Wurzel in $[0, 1]$ vorhanden. In diesem Intervall gilt aber auch die Ungleichung $|\varphi'(x)| \leq q < 1$. Das Iterationsverfahren kann angewendet werden und liefert mit $x_1 = 1$ die Lösung $x_5 \approx x_6 \approx 0,6507$; $x = 0,6507 \pm 0,0001$.

Jetzt betrachten wir die Gleichung $x = -\sqrt{\lg(x+2)}$. Auch hier geht alles nach dem bekannten Schema, und wir bekommen für das Intervall $[-\frac{1}{2}, 0]$ mit $x_1 = 0$ die

Lösung $x_8 \approx x_9 \approx -0,4397$, also $x = -0,4397 \pm 0,0001$.

14. Die eine Wurzel ist $x = 0$, und das ist auch die einzige Lösung von $x = -\sqrt{\ln(x+1)}$. Deshalb brauchen wir nur

die Gleichung $x = \sqrt{\ln(x+1)}$ zu behandeln. Hier ist $\varphi(1/2) = \sqrt{\ln(3/2)} > 1/2$; $\varphi(1) = \sqrt{\ln 2} < 1$. Folglich haben wir die Wurzel im Intervall $[1/2, 1]$ zu suchen. Dort ist aber wegen

$$\varphi'(x) = \frac{1}{2(x+1)\sqrt{\ln(x+1)}}$$

$|\varphi'(x)| \leq q < 1$. Setzen wir $x_1 = 1$, dann erhalten wir $x_9 \approx x_{10} \approx 0,7469$ und somit $x = 0,7469 \pm 0,0001$.

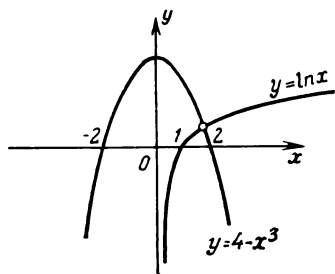


Abb. 30

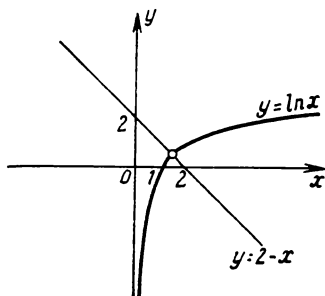


Abb. 31

15. Wir behandeln die äquivalente Form

$$x = \sqrt{4 - \ln x}.$$

Hier ist

$$\varphi(1) = 2, \quad \varphi(2) = \sqrt{4 - \ln 2}, \quad \varphi'(x) = \frac{-1}{2x\sqrt{4 - \ln x}},$$

und man überzeugt sich leicht, daß in $[1, 2]$ wenigstens eine Wurzel liegen muß. Abb. 30 zeigt, daß es keine weiteren Wurzeln gibt ($x_1 = 2, x_4 \approx x_5 \approx 1,841, x = 1,841 \pm 0,001$).

16. Behandlung in der Form $x = 2 - \ln x$, $\varphi(x) = 2 - \ln x$, $\varphi'(x) = -\frac{1}{x}$. Abb. 31 zeigt eine Wurzel in $[1, 2]$ an. Dort ist $|\varphi'(x)| \leq 1$ ($x_1 = 1,5; x_{13} \approx x_{14} \approx 1,557$, d. h. $x = 1,557 \pm 0,001$).

17. Die Gleichung besitzt nur eine reelle Lösung, wie Abb. 32 zeigt. Wir verwenden die äquivalente Form

$$x = -\sqrt{e^x + 2}.$$

Dann ist

$$\varphi(x) = -\sqrt{e^x + 2}, \quad \varphi'(x) = \frac{-e^x}{2\sqrt{e^x + 2}}$$

und

$$\varphi(-1) = -\sqrt{2 + e^{-1}} \approx -1,54;$$

$$\varphi(-2) = -\sqrt{2 + e^{-2}} \approx -1,46.$$

Die Wurzel liegt also im Intervall $[-1, -2]$ und errechnet sich mit $x_1 = -1$ und $x_4 \approx x_5 \approx -1,492$ zu $x = -1,492 \pm \pm 0,001$.

18. Eine Wurzel ist ersichtlich $x = 10$. Eine zweite gewinnen wir bei Anwendung der Iterationsmethode auf die äquivalente Gleichung $x = 10^{0,1x}$. Hier ist $\varphi(x) = 10^{0,1x}$, $\varphi'(x) = 0,1 \cdot 10^{0,1x} \ln 10$. Außerdem ist $\varphi(1) = 10^{0,1} > 1$, $\varphi(2) = 10^{0,2} < 2$. Also liegt eine Lösung im Intervall $[1, 2]$. Für dieses Intervall gilt

$$|\varphi'(x)| \leq 0,1 \cdot 10^{0,2} \ln 10 \approx 0,37 < 1,$$

also konvergiert das Verfahren. Für $x_1 = 2$ erhält man $x_7 \approx x_8 \approx 1,372$, also $x = 1,372 \pm 0,001$.

19. Abbildung 33 zeigt, daß die Gleichung in jedem der Intervalle

$$\left[\frac{\pi}{2} + n\pi, \quad \frac{\pi}{2} + (n+1)\pi \right], \quad n = 0, 1,$$

genau eine Wurzel besitzt, wobei dieselbe stets in der rechten Intervallhälfte liegt. Zur Berechnung der ersten positiven Nullstelle transformieren wir die Gleichung durch $x = \frac{3\pi}{2} - y$ auf die Gleichung

$$\cot y = \lg \left(\frac{3\pi}{2} - y \right),$$

woraus $y = \arccot \left[\lg \left(\frac{3\pi}{2} - y \right) \right]$, wegen $0 < y < \pi$ folgt.

Hier ist $\varphi(y) = \operatorname{arccot} \left[\lg \left(\frac{3}{2} \pi - y \right) \right]$, und es wird $\varphi'(y) =$

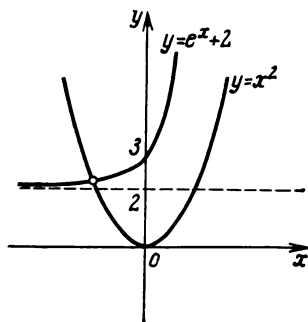


Abb. 32

$= -\lg e \left[\left(1 + \lg^2 \left(\frac{3}{2} \pi - y \right) \right) \left(\frac{3}{2} \pi - y \right) \right]^{-1}$. Demzufolge gilt $\varphi'(y) < 1$ im Intervall $0 < y < \pi$. Mit $y_1 = 0$ errechnen

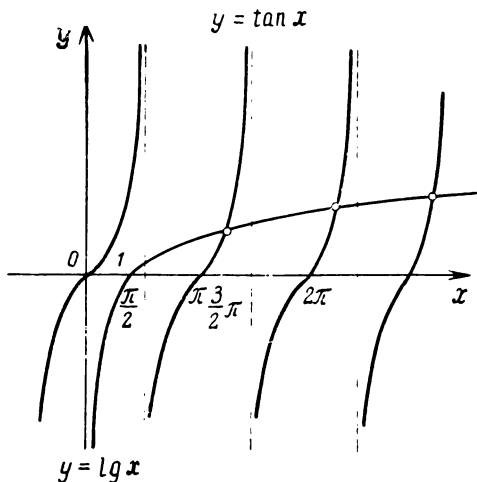


Abb. 33

wir durch sukzessive Approximation $y_4 \approx y_5 \approx 1,059$, also $y = 1,059 \pm 0,001$ und somit $x = 3,654 \pm 0,001$.

Zur Berechnung der zweiten Wurzel müssen wir $x = \frac{5}{2} \pi - y$ setzen. Wir erhalten

$$y = \arctan \left[\lg \left(\frac{5}{2} \pi - y \right) \right].$$

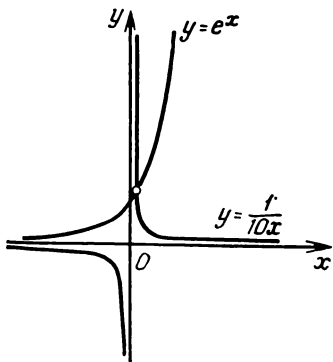


Abb. 34

Mit $y_1 = 0$ wird $y_4 \approx y_5 \approx 0,876$, also $y = 0,876 \pm 0,001$ und $x = 6,984 \pm 0,001$.

20. Die Gleichung besitzt genau eine Lösung, welche zwischen 0 und 1 liegt (Abb. 34). $\varphi(x) = 0,1 e^{-x}$, $\varphi'(x) = -0,1 e^{-x}$. In $[0, 1]$ ist $|\varphi'(x)| \leq 0,1$. Das Verfahren

konvergiert! Mit $x_1 = 0$ erhalten wir $x_4 \approx x_5 \approx 0,091$, d. h. $x = 0,091 \pm 0,001$.

21. $f(x) = x^3 - 5x + 1$, $f'(x) = 3x^2 - 5$, $f''(x) = 6x$. Die Newtonsche Formel lautet dann

$$\beta_{n+1} = \beta_n - \frac{\beta_n^3 - 5\beta_n + 1}{3\beta_n^2 - 5}.$$

Eine Wertetafel

x	-3	-2	-1	0	1	2	3
$f(x)$	-11	3	5	1	-3	-1	13

signalisiert Nullstellen in $[-3, -2]$, $[0, 1]$, $[2, 3]$. Wenden wir uns der Berechnung der negativen Nullstelle zu. Da in diesem Intervall $f''(x) < 0$ gilt, ist der linke Intervallendpunkt $\beta_0 = -3$ als Startpunkt zu nehmen ($f(-3) = -11$), und wir bekommen

$$\beta_1 = -3 - \frac{(-3)^3 - 5(-3) + 1}{3(-3)^2 - 5} = -2,5$$

als erste Näherung und nach Fortsetzung der Rechnung mit $\beta_3 = \beta_4 \approx -2,331$ die Lösung $x = -2,331 \pm 0,001$. Jetzt wenden wir uns dem Intervall $[0, 1]$ zu, in welchem offenbar $f''(x) \geq 0$ ausfällt. Demnach ist hier $\beta_0 = 0$ zu nehmen, und damit errechnen wir

$$\beta_1 = 0 - \frac{0^3 - 5 \cdot 0 + 1}{3 \cdot 0^2 - 5} = 0,2, \quad \beta_3 = \beta_4 = 0,202,$$

$$x = 0,202 \pm 0,001.$$

Zur Bestimmung der in $[2, 3]$ gelegenen Wurzel schließlich wählen wir $\beta_0 = 3$ und bekommen

$$\beta_1 = 3 - \frac{3^3 - 5 \cdot 3 + 1}{3 \cdot 3^2 - 5} \approx 2,409.$$

Die Rechnung ergibt $\beta_4 = \beta_5 \approx 2,128$, also $x = 2,128 \pm \pm 0,001$. Wir erhalten somit drei Wurzeln: $x_1 = -2,331$; $x_2 = 0,202$; $x_3 = 2,128$ ($\pm 0,001$).

Zum Vergleich lösen wir dieselbe Aufgabe mit dem Sekantenverfahren. In $[-3, -2]$ wird

$$\begin{aligned} \alpha_1 &= -3 - f(-3) \frac{-3 - (-2)}{f(-3) - f(-2)} = \\ &= -3 + 11 \frac{(-3)}{-14} \approx -2,214. \end{aligned}$$

Wegen $f''(x) < 0$ in diesem Intervall, ist die Kurve von unten gesehen konvex, und wir berechnen α_2 nach

$$\alpha_2 = -3 - f(-3) \frac{-3 - (-2,214)}{f(-3) - f(-2,214)} \approx -2,293.$$

Weiter bekommen wir

$$\alpha_3 = -2,293 - f(-2,293) \frac{-2,293 + 2,214}{f(-2,293) - f(-2,214)} \approx -2,331,$$

und dieser Wert stimmt innerhalb der mitgeführten Stellenzahl mit unserem früheren Ergebnis überein.

Ebenso löst man die Gleichung für die Intervalle $[0, 1]$ und $[2, 3]$.

22. Hier haben wir

$$f(x) = x^3 - 9x^2 + 20x - 11,$$

$$f'(x) = 3x^2 - 18x + 20,$$

$$f''(x) = 6x - 18 = 6(x - 3).$$

Die Wertetafel ist jetzt

x	0	1	2	3	4	5	6
$f(x)$	-11	1	1	-5	-11	-11	1

Die Lösungen der Gleichung liegen in den Intervallen $[0, 1]$, $[2, 3]$, $[5, 6]$. In $[0, 1]$ wählen wir $\beta_0 = 0$ und errechnen $\beta_4 \approx \beta_5 \approx 0,834$. In $[2, 3]$ wird mit $\beta_0 = 3$, $\beta_2 \approx \beta_3 \approx 2,216$, und in $[5, 6]$ schließlich errechnen wir $\beta_4 \approx \beta_5 \approx 5,249$ unter Verwendung von $\beta_1 = 0$. Die drei Wurzeln der Gleichung sind

$$x_1 = 0,834; x_2 = 2,216; x_3 = 5,249 \text{ (jeweils } \pm 0,001).$$

23. $f(x) = x^3 - 3x^2 - 3x + 11$, $f'(x) = 3x^2 - 6x - 3$,
 $f''(x) = 6x - 6 = 6(x - 1)$. Wertetafel:

x	-2	-1	0	1	2	3
$f(x)$	-3	10	11	6	1	2

Die Gleichung hat eine im Intervall $[-2, -1]$ gelegene Wurzel. Mit $\beta_0 = -2$ bekommen wir bereits im zweiten Schritt die gewünschte Genauigkeit: $\beta_2 \approx \beta_3 = -1,847$, also $x = -1,847 \pm 0,001$.

24. $f(x) = x^5 + 5x + 1$, $f'(x) = 5x^4 + 5$, $f''(x) = 20x^3$. Die Wertetafel

x	-1	0	1
$f(x)$	-5	1	7

zeigt, daß eine Wurzel im Intervall $[-1, 0]$ vorhanden sein muß, und man berechnet sie mit $\beta_0 = -1$ beginnend zu $\beta_3 \approx \beta_4 \approx -0,1999$, d. h. $x = -0,1999 \pm 0,0001$.

25. $f(x) = \sin x + x - 1$. $f'(x) = \cos x + 1$, $f''(x) = -\sin x$. Eine kleine Wertetafel:

x	0	1	2
$f(x)$	-1	0,8115	1,9093

zeigt eine Wurzel zwischen 0 und +1 an, die mit $\beta_0 = 0$ schnell berechnet ist: $\beta_2 \leq \beta_3 \approx 0,5110$, also $x = 0,5110 \pm 0,001$.

26. $f(x) = x^2 - 10 \lg x - 3$. $f'(x) = 2x - \frac{10}{x \ln 10}$,

$f''(x) = 2 + \frac{10}{x^2 \ln 10}$. Auch hier zeigt die Wertetafel

x	0,5	1	2	3
$f(x)$	0,26	-2	-2,01	1,23

daß Lösungen in $[0,5, 1]$ und $[2, 3]$ liegen.

1. Wurzel: $\beta_0 = 0,5$, $\beta_2 \approx \beta_3 \approx 0,535$; $x_1 = 0,535 \pm 0,001$,

2. Wurzel: $\beta_0 = 3$, $\beta_2 \approx \beta_3 \approx 2,705$; $x_2 = 2,705 \pm 0,001$.

27. $x_0 = 0$, $x = 1,021 \pm 0,001$,

$y_0 = 0$, $y = 2,150 \pm 0,001$,

$z_0 = 0$, $z = 3,072 \pm 0,001$.

28. $x_0 = 0$, $x = 0,520 \pm 0,001$,

$y_0 = 0$, $y = 0,310 \pm 0,001$.

29. $x_0 = 0$, $x = 1,000 \pm 0,001$,

$y_0 = 0$, $y = 2,000 \pm 0,001$.

