

Kapitel 2

Numerische Verfahren zur Lösung nichtlinearer Gleichungen

2.1 Aufgabenstellung und Motivation

Ist f eine in einem abgeschlossenen Intervall $I = [a, b]$ stetige und reellwertige Funktion, so heißt eine Zahl $\xi \in I$ eine *Nullstelle der Funktion f* oder eine *Lösung der Gleichung*

$$f(x) = 0, \quad (2.1)$$

falls $f(\xi) = 0$ ist.

Wenn f ein *algebraisches Polynom* der Form

$$f(x) \equiv P_n(x) = \sum_{j=0}^n a_j x^j, \quad a_j \in \mathbb{R}, \quad a_n \neq 0, \quad n \in \mathbb{N}$$

ist, heißt die Gleichung (2.1) *algebraisch*, und die natürliche Zahl n heißt der *Grad* des Polynoms bzw. der algebraischen Gleichung. Jede Gleichung (2.1), die nicht algebraisch ist, heißt *transzendent* (z. B. $\ln x - 1/x = 0$; $x - \sin x = 0$).

In diesem Kapitel werden Verfahren zur Bestimmung einfacher und mehrfacher Nullstellen $\xi \in I$ von f vorgestellt. Dabei wird zwischen den klassischen Iterationsverfahren (allgemeines Iterationsverfahren, Newton-Verfahren, Sekantenverfahren) und den sogenannten Einschlussverfahren (Bisektion, Pegasus-Verfahren, Verfahren von Anderson-Björck, Verfahren von King) unterschieden. Die angegebenen Einschlussverfahren benötigen zwei Startwerte, in denen die Funktion f unterschiedliche Vorzeichen hat. Die Startwerte schließen dann (mindestens) eine Nullstelle ungerader Ordnung von f ein. Dieser Einschluss bleibt im Laufe der Rechnung erhalten. Diese Verfahren sind jedoch grundsätzlich unter Anwendung von Satz 2.3 auch im Falle von Nullstellen gerader Ordnung anwendbar. Einschlussverfahren höherer Konvergenzordnung sind im Allgemeinen den klassischen Iterationsverfahren vorzuziehen (vgl. Abschnitt 2.9). Verfahren zur Bestimmung sämtlicher Nullstellen algebraischer Polynome ohne Kenntnis von Startwerten werden im Kapitel 3 behandelt.

Wo in den Anwendungen treten nichtlineare Gleichungen auf? Die meisten derartigen Probleme führen nicht direkt auf eine transzendente Gleichung oder eine Polynomgleichung, sondern diese treten an irgendeiner Stelle des Lösungsprozesses auf. So stößt man etwa auf Polynomgleichungen bei der Lösung von linearen Differentialgleichungen n -ter Ordnung mit konstanten Koeffizienten, bei der Berechnung der Eigenwerte von Matrizen und bei der Bestimmung der Eigenfrequenzen linearer Schwingungssysteme mit n Freiheitsgraden. Transzendente Gleichungen treten beispielsweise bei der Berechnung der kritischen Drehzahl einer Welle in der Gestalt

$$f(x) = \cos x \cosh x \pm 1 = 0 \quad \text{oder} \quad f(x) = \tan x - \tanh x = 0$$

auf. Bei der Berechnung des Druckverlustes in einer Rohrströmung muss man den Rohrreibungsbeiwert λ für hydraulisch glatte Rohre bei turbulenter Strömung nach dem universellen Prandtlschen Widerstandsgesetz bei vorgegebener Reynoldszahl Re aus der transzendenten Gleichung

$$f(\lambda) = \frac{1}{\sqrt{\lambda}} - 2 \lg(\operatorname{Re} \sqrt{\lambda}) + 0.8 = 0$$

ermitteln. Ebenfalls auf eine transzendente Gleichung stößt man bei der Betrachtung der Ausstrahlung eines vollkommenen „schwarzen Körpers“. Wird ein schwarzer Körper erhitzt, so sendet er elektromagnetische Wellen verschiedener Wellenlänge λ (Wärmestrahlen, Licht usw.) aus.

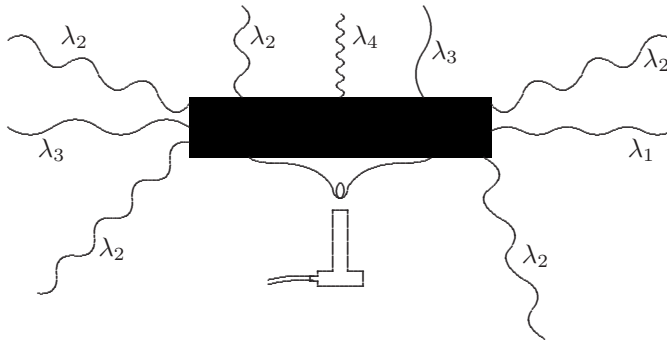


Abb. 2.1. Ausstrahlung eines vollkommenen schwarzen Körpers

Bei einer bestimmten Temperatur T liegt das Maximum der emittierten Strahlung $E(\lambda, T)$ bei λ_{\max} . Mit steigender Temperatur verschiebt sich die Stelle maximaler Emission nach kürzeren Wellenlängen; es gilt (*Wiensches Verschiebungsgesetz*)

$$\lambda_{\max}(T) = \frac{\alpha}{z \cdot T}$$

mit $\alpha = 14.3881 \cdot 10^{-3} \text{ m K}$. Die Konstante z ist die Lösung der transzendenten Gleichung

$$e^{-z} = 1 - \frac{1}{5}z \quad \text{bzw.} \quad f(z) = e^{-z} - 1 + \frac{1}{5}z = 0, \quad z \in \mathbb{R}, \quad z \neq 0.$$

Mit dem so ermittelten $z \approx 4.9651$ lässt sich die Formel für $\lambda_{\max}(T)$ angeben. Aus dieser Formel kann dann z. B. ein Näherungswert für die absolute Temperatur der Sonnenober-

fläche bestimmt werden, wobei für das Sonnenspektrum nach Langley das Maximum der Strahlung bei $\lambda_{\max} = 5 \cdot 10^{-7}$ m liegt.

Weitere Anwendungsbeispiele sind in Abschnitt 2.8 zu finden.

2.2 Definitionen und Sätze über Nullstellen

Im Anschluss werden einige Definitionen und Sätze über Nullstellen angegeben, die in den folgenden Abschnitten benötigt werden. Die Beweise findet man in jedem Werk zur Analysis.

Definition 2.1.

Eine Nullstelle ξ einer Funktion $f \in C[a, b]$ heißt *j-fache Nullstelle* oder Nullstelle der Ordnung j ($j \in \mathbb{N}$), falls f sich auf $[a, b]$ in der Form

$$f(x) = (x - \xi)^j h(x)$$

darstellen lässt mit einer stetigen Funktion h , für die $h(\xi) \neq 0$ ist.

Im Fall $j = 1$ heißt ξ *einfache Nullstelle*, für $j \geq 2$ *mehrfache Nullstelle*.

Ist ξ eine Nullstelle ungerader Ordnung, so hat f in $x = \xi$ einen Vorzeichenwechsel; ist ξ von gerader Ordnung, so berührt der Graph von f die x -Achse, und es gibt keinen Vorzeichenwechsel.

Satz 2.2.

Die Funktion f sei im Intervall I j -mal stetig differenzierbar ($j \in \mathbb{N}$). Dann ist $\xi \in I$ genau dann eine j -fache Nullstelle von f , wenn gilt

$$f(\xi) = f'(\xi) = \dots = f^{(j-1)}(\xi) = 0, \quad f^{(j)}(\xi) \neq 0.$$

Mit der Aussage des folgenden Satzes 2.3 lassen sich Einschlussverfahren auch zur Berechnung von mehrfachen Nullstellen einsetzen, insbesondere auch von Nullstellen gerader Ordnung (also ohne Vorzeichenwechsel).

Satz 2.3.

Ist $\xi \in I$ eine j -fache Nullstelle von f , $j \geq 2$, und ist f $(j+1)$ -mal stetig differenzierbar, so ist ξ eine einfache Nullstelle von g mit

$$g(x) = \frac{f(x)}{f'(x)}.$$

Beweis. Nach Satz 2.2 gelten für $j \geq 2$ $f(\xi) = f'(\xi) = \dots = f^{(j-1)}(\xi) = 0$ und $f^{(j)}(\xi) \neq 0$. Zu zeigen ist, dass für

$$g(x) = \frac{f(x)}{f'(x)} \quad \text{und} \quad g'(x) = 1 - \frac{f(x) f''(x)}{f'^2(x)}$$

nach Satz 2.2 gelten

$$g(\xi) = 0 \quad \text{und} \quad g'(\xi) \neq 0.$$

Für den Nachweis werden die Taylorentwicklungen von f , f' und f'' an der Stelle ξ benötigt (sie finden auch in Abschnitt 2.5.3 Verwendung).

$$f(x) = \frac{(x-\xi)^j}{j!} f^{(j)}(\xi) + O((x-\xi)^{j+1}) = (x-\xi)^j h_0(x) \quad (2.2)$$

$$\text{mit } h_0(x) = \frac{1}{j!} f^{(j)}(\xi) + O(x-\xi) \quad \text{und}$$

$$h_0(\xi) = \frac{1}{j!} f^{(j)}(\xi) \neq 0; \quad (2.3)$$

$$f'(x) = \frac{(x-\xi)^{j-1}}{(j-1)!} f^{(j)}(\xi) + O((x-\xi)^j) = (x-\xi)^{j-1} h_1(x) \quad (2.4)$$

$$\text{mit } h_1(x) = \frac{1}{(j-1)!} f^{(j)}(\xi) + O(x-\xi) \quad \text{und}$$

$$h_1(\xi) = \frac{1}{(j-1)!} f^{(j)}(\xi) \neq 0; \quad (2.5)$$

$$f''(x) = \frac{(x-\xi)^{j-2}}{(j-2)!} f^{(j)}(\xi) + O((x-\xi)^{j-1}) = (x-\xi)^{j-2} h_2(x) \quad (2.6)$$

$$\text{mit } h_2(x) = \frac{1}{(j-2)!} f^{(j)}(\xi) + O(x-\xi) \quad \text{und}$$

$$h_2(\xi) = \frac{1}{(j-2)!} f^{(j)}(\xi) \neq 0. \quad (2.7)$$

Mit (2.2) und (2.4) ergibt sich

$$g(x) = \frac{f(x)}{f'(x)} = (x-\xi) \frac{h_0(x)}{h_1(x)}.$$

Mit (2.3) und (2.5) ist

$$\frac{h_0(\xi)}{h_1(\xi)} = \frac{f^{(j)}(\xi) (j-1)!}{j! f^{(j)}(\xi)} = \frac{1}{j} \neq 0; \quad (2.8)$$

also ist $g(\xi) = 0$.

Weiter ist mit (2.2), (2.4) und (2.6)

$$g'(x) = 1 - \frac{f(x) f''(x)}{f'^2(x)} = 1 - \frac{h_0(x) h_2(x)}{h_1^2(x)}.$$

Für $g'(\xi)$ ergibt sich mit (2.3), (2.5) und (2.7)

$$g'(\xi) = 1 - \frac{h_0(\xi) h_2(\xi)}{h_1^2(\xi)} = 1 - \frac{(j-1)! (j-1)!}{j! (j-2)!} = 1 - \frac{j-1}{j} = \frac{1}{j} \neq 0.$$

Die j -fache Nullstelle ξ von f ist also einfache Nullstelle von g . \square

Satz 2.4. (*Satz von Bolzano, Zwischenwertsatz*)

Sei f in $I = [a, b]$ stetig mit $f(a) \cdot f(b) < 0$. Dann besitzt f in (a, b) mindestens eine Nullstelle ξ ungerader Ordnung.

2.3 Allgemeines Iterationsverfahren

2.3.1 Konstruktionsmethode und Definition

Anstelle der Gleichung $f(x) = 0$ wird eine Gleichung der Form

$$x = \varphi(x) \tag{2.9}$$

betrachtet. Dabei sei

$$\varphi : I \rightarrow \mathbb{R} \quad \text{mit} \quad x \mapsto \varphi(x)$$

eine in einem abgeschlossenen Intervall I stetige Funktion, und $\xi \in I$ heißt eine Lösung von (2.9) bzw. ein Fixpunkt der Abbildung φ , wenn $\varphi(\xi) = \xi$ ist; darum heißt (2.9) auch *Fixpunktgleichung*.

Die Untersuchung von Gleichungen der Form $x = \varphi(x)$ bedeutet keine Beschränkung der Allgemeinheit, denn es gilt der

Hilfssatz 2.5.

Sind f und g stetige Funktionen in einem abgeschlossenen Intervall I und ist $g(x) \neq 0$ für alle $x \in I$, dann besitzen die Gleichungen $f(x) = 0$ und $x = \varphi(x)$ mit

$$\varphi(x) := x - f(x)g(x) \tag{2.10}$$

im Intervall I dieselben Lösungen, d. h. die beiden Gleichungen sind äquivalent.

Beweis. Ist $\xi \in I$ Lösung von $f(x) = 0$, so folgt wegen $f(\xi) = 0$ aus (2.10) $\varphi(\xi) = \xi$. Ist umgekehrt $\xi \in I$ Lösung von $x = \varphi(x)$, so folgt wegen $\xi = \varphi(\xi)$ aus (2.10) $f(\xi)g(\xi) = 0$; wegen $g(\xi) \neq 0$ ist also $f(\xi) = 0$. \square

Jede geeignete Wahl von g liefert eine zu $f(x) = 0$ äquivalente Gleichung $x = \varphi(x)$. Häufig kann eine Gleichung $f(x) = 0$ auf die Form $x = \varphi(x)$ gebracht werden, indem irgendeine Auflösung nach x vorgenommen wird.

Beispiel 2.6.

Gegeben: Die algebraische Gleichung

$$f(x) = x^2 + x - 2 = 0$$

mit den Lösungen $\xi_1 = 1$ und $\xi_2 = -2$.

Gesucht: Zur gegebenen Gleichung äquivalente Gleichungen der Form $x = \varphi(x)$.

Lösung: Durch verschiedenartige Umformung bzw. Auflösung nach x erhält man

$$(I) \quad x = 2 - x^2 = \varphi(x),$$

$$(II) \quad x = \sqrt{2 - x} = \varphi(x), \quad x \leq 2,$$

$$(III) \quad x = \frac{2}{x} - 1 = \varphi(x), \quad x \neq 0.$$

Bei der Angabe eines Intervalls I , in dem die Gleichungen äquivalent zur gegebenen sind, müssen die Einschränkungen für x berücksichtigt werden. \square

Nun sei eine Gleichung der Form $x = \varphi(x)$ mit dem zugehörigen Intervall I gegeben. Dann konstruiert man mit Hilfe eines *Startwertes* $x^{(0)} \in I$ eine Zahlenfolge $\{x^{(\nu)}\}$ nach der Vorschrift

$$x^{(\nu+1)} := \varphi(x^{(\nu)}), \quad \nu = 0, 1, 2, \dots \quad (2.11)$$

Diese Folge lässt sich nur dann sinnvoll konstruieren, wenn für $\nu = 0, 1, 2, \dots$

$$x^{(\nu+1)} = \varphi(x^{(\nu)}) \in I$$

ist, da φ nur für $x \in I$ erklärt ist. Durch φ muss also eine Abbildung des Intervalls I in sich gegeben sein, d. h. der Graph von $y = \varphi(x)$ muss im Quadrat

$$Q = \{(x, y) | x \in I, y \in I\}$$

liegen; in der Abbildung 2.2 ist $I = [a, b]$.

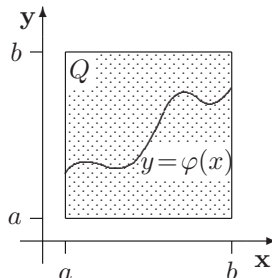


Abb. 2.2.

Wenn die Folge $\{x^{(\nu)}\}$ konvergiert, d. h. wenn die Zahlen $x^{(1)}, x^{(2)}, x^{(3)}, \dots$ gegen ξ streben, somit

$$\lim_{\nu \rightarrow \infty} x^{(\nu)} = \xi$$

ist, dann ist ξ eine Lösung der Gleichung (2.9). Es gilt wegen der Stetigkeit von φ

$$\xi = \lim_{\nu \rightarrow \infty} x^{(\nu)} = \lim_{\nu \rightarrow \infty} x^{(\nu+1)} = \lim_{\nu \rightarrow \infty} \varphi(x^{(\nu)}) = \varphi(\lim_{\nu \rightarrow \infty} x^{(\nu)}) = \varphi(\xi).$$

Ein solches *Verfahren der schrittweisen Annäherung* wird *Iterationsverfahren* genannt. Die Vorschrift (2.11) heißt *Iterationsvorschrift*; sie stellt für jedes feste ν einen *Iterationsschritt* dar. Die Funktion φ wird *Schrittfunktion* genannt. Die Folge $\{x^{(\nu)}\}$ heißt *Iterationsfolge*.

Die Iterationsschritte (2.11) für $\nu = 0(1)N$ bilden zusammen mit dem Startwert $x^{(0)}$ das algorithmische Schema des Iterationsverfahrens:

$$\left\{ \begin{array}{ll} x^{(0)} & = \text{Startwert,} \\ x^{(1)} & = \varphi(x^{(0)}), \\ x^{(2)} & = \varphi(x^{(1)}), \\ \cdot & \\ \cdot & \\ \cdot & \\ x^{(N+1)} & = \varphi(x^{(N)}). \end{array} \right. \quad (2.12)$$

Dabei muss $\varphi(x^{(\nu)}) \in I$ gelten für $\nu = 0(1)N$.

Beispiel 2.7. (Fortsetzung von Beispiel 2.6)

Für die Umformung (I): $x = 2 - x^2$ der Gleichung $x^2 + x - 2 = 0$ lautet die Iterationsvorschrift (2.11)

$$x^{(\nu+1)} = \varphi(x^{(\nu)}) = 2 - (x^{(\nu)})^2, \quad \nu = 0, 1, 2, \dots,$$

und es sei $I = [-50, 0]$. Das algorithmische Schema (2.12) des Iterationsverfahrens lautet mit dem Startwert $x^{(0)} = -3$:

$$\begin{array}{ll} x^{(0)} & = -3 \in I, \\ x^{(1)} & = \varphi(x^{(0)}) = 2 - 3^2 = -7 \in I, \\ x^{(2)} & = \varphi(x^{(1)}) = 2 - 7^2 = -47 \in I, \\ x^{(3)} & = \varphi(x^{(2)}) = 2 - 47^2 = -2207 \notin I. \end{array}$$

Der Verlauf der Rechnung zeigt, dass die so konstruierte Folge $\{x^{(\nu)}\}$ nicht gegen die Lösung $\xi_2 = -2 \in I$ konvergiert.

Mit der Umformung (III): $x = \frac{2}{x} - 1$ der Gleichung $x^2 + x - 2 = 0$ und $I = [-3, -1]$ erhält man dagegen die Iterationsvorschrift

$$x^{(\nu+1)} = \varphi(x^{(\nu)}) = \frac{2}{x^{(\nu)}} - 1, \quad \nu = 0, 1, 2, \dots,$$

und mit dem Startwert $x^{(0)} = -3$ das algorithmische Schema:

$$\begin{aligned}
x^{(0)} &= -3 \in I, \\
x^{(1)} &= \frac{2}{-3} - 1 = -\frac{5}{3} = -1.6666667 \in I, \\
x^{(2)} &= -2.2000000 \in I, \\
x^{(3)} &= -1.9090909 \in I, \\
x^{(4)} &= -2.0476190 \in I.
\end{aligned}$$

Diese vier Iterationsschritte zeigen bereits, dass sich die so konstruierte Folge der Lösung $\xi_2 = -2$ immer mehr nähert, d. h. gegen die gesuchte Lösung konvergiert. \square

Es stellt sich also die Frage nach Bedingungen für die Konvergenz einer Iterationsfolge. Die folgenden Aussagen über die Existenz einer Lösung der Gleichung (2.9) und deren Eindeutigkeit dienen der Beantwortung dieser Frage. Es wird sich zeigen, dass die Bedingungen für die Existenz und Eindeutigkeit auch hinreichend für die Konvergenz der mit (2.11) konstruierten Iterationsfolge sind.

2.3.2 Existenz einer Lösung und Eindeutigkeit der Lösung

Zum Nachweis der Existenz einer Lösung $\xi \in I = [a, b]$ der Gleichung $x = \varphi(x)$ folgen nun lediglich geometrische Überlegungen; zum analytischen Nachweis s. [HENR1972] Bd.1, S. 85/87. Man erhält ξ als Abszisse des Punktes, in dem die Graphen der Gerade $y = x$ und der Funktion $y = \varphi(x)$ sich in dem Quadrat Q schneiden. Es wird also davon ausgegangen, dass für $x \in I$ auch $\varphi(x) \in I$ ist. Wenn $a \neq \varphi(a)$ und $b \neq \varphi(b)$ sind, müssen $\varphi(a) > a$ und $\varphi(b) < b$ sein. Die Punkte $(a, \varphi(a))$ und $(b, \varphi(b))$ liegen also auf verschiedenen Seiten der Diagonale $y = x$ im Quadrat Q . Ist φ stetig, so garantiert der stetige Verlauf des Graphen von $y = \varphi(x)$ in Q die Existenz von mindestens einem Schnittpunkt mit der Gerade $y = x$ (Abb. 2.3(a)). Ist φ dagegen nicht stetig, so existiert nicht notwendig ein solcher Schnittpunkt (Abb. 2.3(b)).

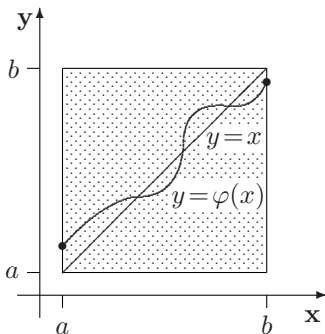


Abb. 2.3. (a) φ stetig

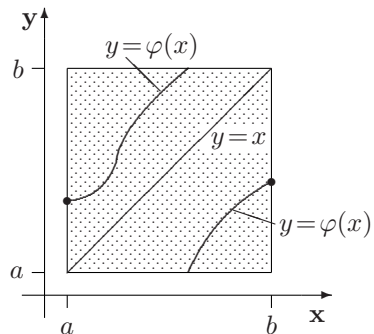


Abb. 2.3. (b) φ nicht stetig

Satz 2.8. (*Existenzsatz*)

Die in dem endlichen, abgeschlossenen Intervall I definierte Funktion $\varphi : I \rightarrow \mathbb{R}$ erfülle die folgenden Bedingungen:

- (i) $\varphi(x) \in I$ für alle $x \in I$,
- (ii) φ ist stetig in I .

Dann besitzt die Gleichung $x = \varphi(x)$ in I mindestens eine Lösung ξ .

Anhand eines Beispiels wird gezeigt, dass die Stetigkeit (ii) für die Existenz einer Lösung $\xi \in I$ allein nicht hinreicht.

Beispiel 2.9.

Gegeben: Die Gleichung $e^x = 0$, $I = [-a, 0]$, $a > 0$, von der bekannt ist, dass sie keine endliche Lösung besitzt.

Gesucht: Eine dazu äquivalente Gleichung der Form $x = \varphi(x)$.

Lösung: Mit $g(x) = 1$ ergibt sich gemäß (2.10) die Gleichung:

$$x = \varphi(x) = x - e^x.$$

Die Funktion φ ist stetig in $[-a, 0]$. Wegen $\varphi(-a) = -a - 1/e^a < -a$ ist $\varphi(-a) \notin I$ und somit die Bedingung (i) des Satzes 2.8 nicht erfüllt. \square

Die Frage nach der Eindeutigkeit einer Lösung der Gleichung $x = \varphi(x)$ kann man beantworten, wenn φ in I einer sogenannten *Lipschitzbedingung* genügt. Eine Funktion φ heißt *lipschitzbeschränkt*, wenn es eine Konstante L mit $0 \leq L < 1$ gibt, so dass

$$|\varphi(x) - \varphi(x')| \leq L|x - x'| \quad \text{für alle } x, x' \in I \quad (2.13)$$

gilt. L wird *Lipschitzkonstante* genannt, und (2.13) heißt eine *Lipschitzbedingung* für die Funktion φ . Eine differenzierbare Funktion φ ist sicher lipschitzbeschränkt, wenn

$$|\varphi'(x)| \leq L < 1 \quad \text{für alle } x \in I \quad (2.14)$$

gilt; denn nach dem Mittelwertsatz der Differentialrechnung ist

$$\varphi(x) - \varphi(x') = \varphi'(\eta)(x - x'), \quad \eta \in (x, x') \subset I,$$

woraus beim Übergang zu den Beträgen und mit (2.14) folgt

$$|\varphi(x) - \varphi(x')| = |\varphi'(\eta)| |x - x'| \leq L|x - x'|.$$

Abbildungen φ , für die eine Lipschitzbedingung (2.13) bzw. (2.14) gilt, werden auch als *kontrahierende Abbildungen* bezeichnet, weil der Abstand $|\varphi(x) - \varphi(x')|$ der Bilder kleiner ist als der Abstand $|x - x'|$ der Urbilder.

Satz 2.10. (*Eindeutigkeitssatz*)

Die Funktion $\varphi : I \rightarrow \mathbb{R}$ genüge im Intervall I einer Lipschitzbedingung (2.13) oder (2.14). Dann besitzt die Gleichung $x = \varphi(x)$ in I höchstens eine Lösung ξ .

Beweis. Angenommen, es gibt zwei Lösungen ξ_1, ξ_2 im Intervall I , so dass also

$$\xi_1 = \varphi(\xi_1) \quad \text{und} \quad \xi_2 = \varphi(\xi_2) \quad (2.15)$$

gelten. Dann folgt mit (2.13) und (2.15)

$$|\xi_1 - \xi_2| = |\varphi(\xi_1) - \varphi(\xi_2)| \leq L|\xi_1 - \xi_2|$$

und daraus

$$(1 - L)|\xi_1 - \xi_2| \leq 0.$$

Wegen $1 - L > 0$ ist $|\xi_1 - \xi_2| \leq 0$, also kann nur $|\xi_1 - \xi_2| = 0$ sein und damit $\xi_1 = \xi_2$. Die Gleichung $x = \varphi(x)$ besitzt also höchstens eine Lösung $\xi \in I$. \square

Da eine Funktion φ , die in I einer Lipschitzbedingung genügt, überall in I stetig ist (die Umkehrung gilt nicht, d. h. nicht jede stetige Funktion genügt einer Lipschitzbedingung) und da die Stetigkeit von φ zusammen mit der Bedingung (i) des Satzes 2.8 hinreichend ist für die Existenz einer Lösung ξ in I , gilt mit Satz 2.10 weiter der

Satz 2.11. (*Existenz- und Eindeutigkeitssatz*)

Die in dem endlichen, abgeschlossenen Intervall I definierte Funktion $\varphi : I \rightarrow \mathbb{R}$ erfülle die folgenden Bedingungen:

- (i) $\varphi(x) \in I$ für alle $x \in I$.
- (ii) φ ist in I lipschitzbeschränkt, d. h. φ genüge für alle $x, x' \in I$ einer Lipschitzbedingung $|\varphi(x) - \varphi(x')| \leq L|x - x'|$ mit $0 \leq L < 1$ oder, falls φ in I differenzierbar ist, einer Bedingung $|\varphi'(x)| \leq L < 1$.

Dann besitzt die Gleichung $x = \varphi(x)$ in I genau eine Lösung ξ .

Beispiel 2.12. (Fortsetzung von Beispiel 2.9)

Die Funktion $\varphi(x) = x - e^x$ genügt in $I = [-a, 0]$ einer Lipschitzbedingung $|\varphi'(x)| \leq L < 1$, da wegen $0 < e^x \leq 1$ für $x \in I$ mit $\varphi'(x) = 1 - e^x$ gilt $|\varphi'(x)| = |1 - e^x| < 1$.

Falls eine Lösung existieren würde, wäre sie nach Satz 2.10 eindeutig bestimmt. Im vorliegenden Falle existiert aber keine Lösung, da die Bedingung (i) in Satz 2.8 bzw. 2.11 nicht erfüllt ist. \square

2.3.3 Konvergenz eines Iterationsverfahrens

2.3.3.1 Heuristische Betrachtungen

Die Funktion $\varphi : I \rightarrow \mathbb{R}$ sei in I differenzierbar. Dann ergibt sich anschaulich, dass die durch die Iterationsvorschrift $x^{(\nu+1)} = \varphi(x^{(\nu)})$ definierte Folge $\{x^{(\nu)}\}$ konvergiert, wenn φ die Bedingungen des Satzes 2.11 für die Existenz und Eindeutigkeit einer Lösung der Gleichung $x = \varphi(x)$ erfüllt:

- (i) $\varphi(x) \in I$ für alle $x \in I$,
- (ii) $|\varphi'(x)| \leq L < 1$ für alle $x \in I$.

Die Konvergenz ist für $0 \leq \varphi'(x) < 1$ monoton (Abb. 2.4) und für $-1 < \varphi'(x) \leq 0$ alternierend (Abb. 2.5).

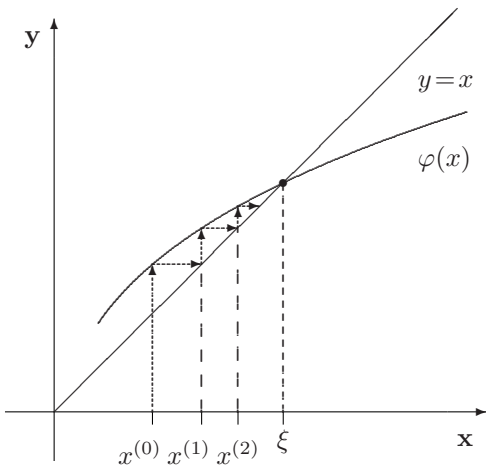


Abb. 2.4. Monotone Konvergenz,
 $0 \leq \varphi'(x) < 1$

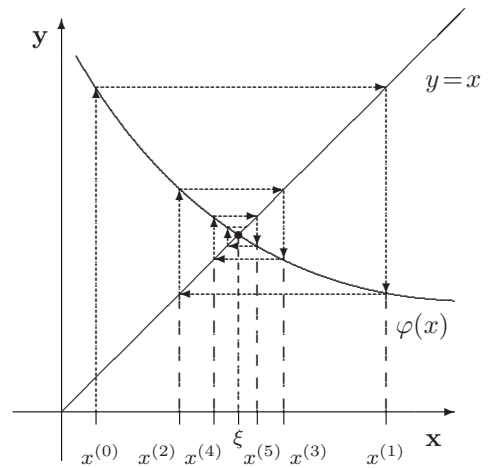


Abb. 2.5. Alternierende Konvergenz,
 $-1 < \varphi'(x) \leq 0$

Je kleiner der Betrag der 1. Ableitung der Schrittfunktion φ ist, desto schneller konvergiert die Iterationsfolge gegen ξ .

Für $|\varphi'(x)| > 1$ divergiert das Verfahren, wie aus den Abbildungen 2.6 und 2.7 zu erkennen ist.

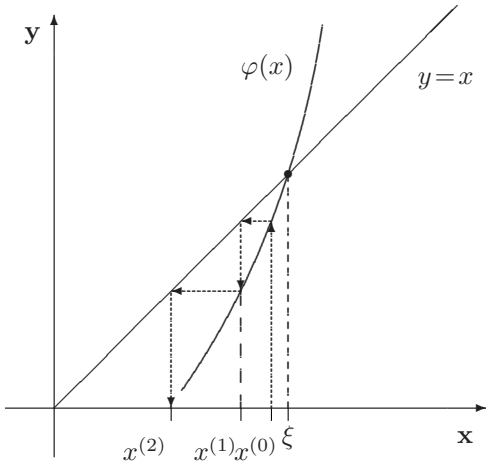


Abb. 2.6. Divergenz, $1 < \varphi'(x)$

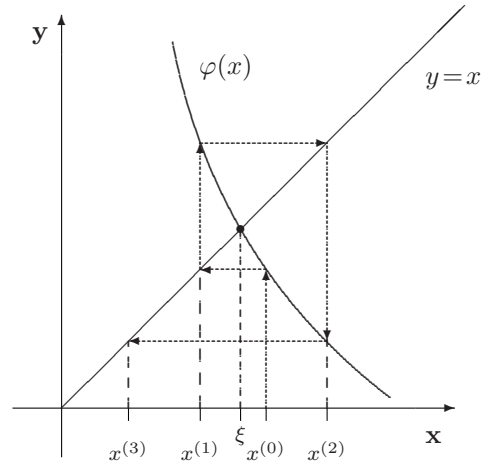


Abb. 2.7. Divergenz, $\varphi'(x) < -1$

Beispiel 2.13. (Fortsetzung von Beispiel 2.6)

Ausgehend von der Umformung (I): $x = 2 - x^2 = \varphi(x)$ der Gleichung $x^2 + x - 2 = 0$ wird $\varphi'(x) = -2x$ untersucht (vgl. Abb. 2.8). Für den Startwert $x^{(0)} = -3$ ist $\varphi'(-3) = 6 > 1$, d. h. die Bedingung $|\varphi'(x)| < 1$ ist bereits für den gewählten Startwert verletzt. Da die Bedingung $|\varphi'(x)| < 1$ nur für $|x| < \frac{1}{2}$ erfüllt werden kann, ist die bei der Umformung (I) entstehende Funktion φ als Schrittfunction für die Iteration zur Bestimmung beider Lösungen 1 und -2 ungeeignet.

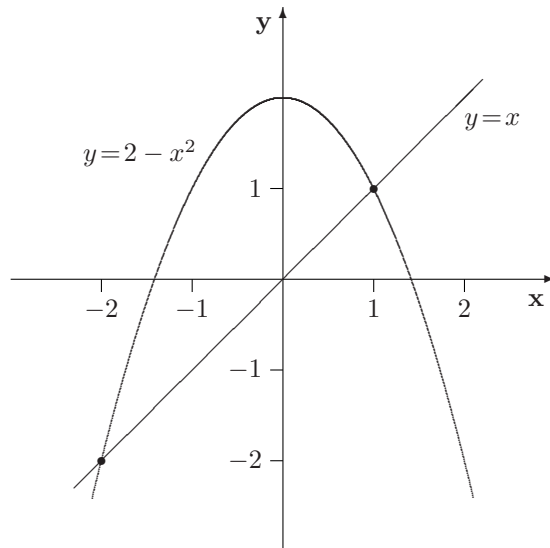


Abb. 2.8. (zu Beispiel 2.13) \square

2.3.3.2 Analytische Betrachtung

Satz 2.14. (*Fixpunktsatz*)

Die Funktion $\varphi : I \rightarrow \mathbb{R}$ erfülle die Voraussetzungen (i) und (ii) des Satzes 2.11. Dann konvergiert die mittels der Iterationsvorschrift

$$x^{(\nu+1)} = \varphi(x^{(\nu)}), \quad \nu = 0, 1, 2, \dots$$

erzeugte Folge $\{x^{(\nu)}\}$ mit einem beliebigen Startwert $x^{(0)} \in I$ gegen den Fixpunkt ξ der Abbildung φ , d. h. es gilt

$$\lim_{\nu \rightarrow \infty} x^{(\nu+1)} = \xi.$$

Beweis. Nach Satz 2.11 hat die Gleichung $x = \varphi(x)$ in I genau eine Lösung ξ , also gilt $\xi = \varphi(\xi)$. Mit $x^{(\nu+1)} = \varphi(x^{(\nu)})$ und der Lipschitzbedingung (2.13) erhält man

$$|x^{(\nu+1)} - \xi| = |\varphi(x^{(\nu)}) - \varphi(\xi)| \stackrel{(LB\,ed)}{\leq} L|x^{(\nu)} - \xi|$$

und in analoger Weise fortfahrend

$$|x^{(\nu+1)} - \xi| \leq L|x^{(\nu)} - \xi| \leq L^2|x^{(\nu-1)} - \xi| \leq \dots \leq L^{\nu+1}|x^{(0)} - \xi|. \quad (2.16)$$

Wegen $0 \leq L < 1$ ist $\lim_{\nu \rightarrow \infty} L^{\nu+1} = 0$, und aus (2.16) folgt

$$\lim_{\nu \rightarrow \infty} |x^{(\nu+1)} - \xi| \leq |x^{(0)} - \xi| \lim_{\nu \rightarrow \infty} L^{\nu+1} = 0, \quad \text{d. h. } \lim_{\nu \rightarrow \infty} x^{(\nu+1)} = \xi.$$

Damit folgt wegen der Stetigkeit von φ

$$\xi = \lim_{\nu \rightarrow \infty} x^{(\nu+1)} = \lim_{\nu \rightarrow \infty} \varphi(x^{(\nu)}) = \varphi\left(\lim_{\nu \rightarrow \infty} x^{(\nu)}\right) = \varphi(\xi). \quad \square$$

Beispiel 2.15. (Fortsetzung von Beispiel 2.6 und 2.7)

Zur Auflösung (III): $x = \frac{2}{x} - 1$ der Gleichung $x^2 + x - 2 = 0$ gehören die Schrittfunktion

$$\varphi(x) = \frac{2}{x} - 1 \quad \text{für } x \neq 0$$

und mit der Fixpunktgleichung $x = \varphi(x)$ die Iterationsvorschrift

$$x^{(\nu+1)} = \varphi(x^{(\nu)}) = \frac{2}{x^{(\nu)}} - 1, \quad \nu = 0, 1, 2, \dots$$

- a) Wegen $\varphi'(x) = -\frac{2}{x^2} < 0$, $x \neq 0$, ist φ streng monoton fallend (Abb. 2.9).

Als Intervall wird $I = [-3, -1]$ gewählt. Wegen $\varphi(-3) = -\frac{5}{3} \in I$ und $\varphi(-1) = -3 \in I$ gilt: φ erfüllt in I die Bedingung (i) des Satzes 2.11.

b) Lipschitzbeschränkung: φ' wird in $I = [-3, -1]$ abgeschätzt:

$$|\varphi'(x)| = \frac{2}{x^2} \leq \frac{2}{\min_{x \in I}(x^2)} = \frac{2}{1} = 2 > 1.$$

Also ist φ in I nicht lipschitzbeschränkt.

Damit $|\varphi'(x)| = 2/x^2 < 1$ ist, muss $x^2 > 2$, also $|x| > \sqrt{2}$ sein. Der rechte Intervallrand muss modifiziert werden, um diese Bedingung zu erfüllen; das neue Intervall sei $I = [-3, -1.5]$.

$|\varphi'(x)|$ nimmt das Maximum am rechten Rand von I an, also gilt

$$|\varphi'(x)| \leq |\varphi'(-1.5)| = 0.\bar{8} < 0.89 = L < 1;$$

in I ist die Voraussetzung (ii) des Satzes 2.11 erfüllt. Damit kann die Lösung $\xi_2 = -2$ iterativ mit der Schrittfunktion aus der Umformung (III) berechnet werden.

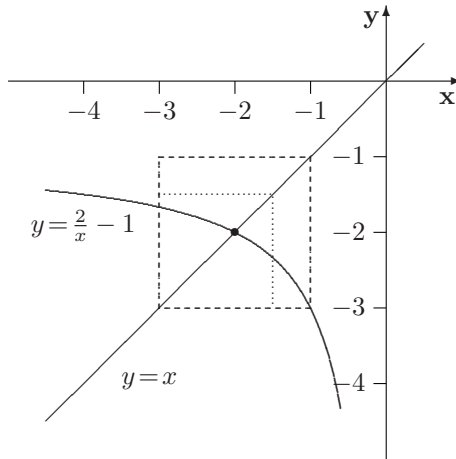


Abb. 2.9. Intervall $[-3, -1]$ wird verkleinert auf $[-3, -1.5]$

Zur Bestimmung der Lösung $\xi_1 = 1$ ist die Umformung (III) ungeeignet. Wählt man nämlich z. B. $I = [0.5, 1.2]$, so gilt dort $|\varphi'(x)| > 1$, so dass die Bedingung (ii) des Satzes 2.11 verletzt ist. \square

2.3.4 Fehlerabschätzungen und Rechnungsfehler

Die nach ν Iterationsschritten erzeugte Näherungslösung $x^{(\nu)}$ unterscheidet sich von der exakten Lösung ξ um den Fehler $\Delta^{(\nu)} := x^{(\nu)} - \xi$ unter der Annahme, dass keine Rechnungsfehler gemacht wurden. Es wird nun für ein festes ν eine Schranke ε für den absoluten Fehler $|\Delta^{(\nu)}|$ gesucht. Ferner interessiert bei vorgegebener Schranke ε die Anzahl ν der Iterationsschritte, die erforderlich ist, damit $|\Delta^{(\nu)}| \leq \varepsilon$ gilt.

(i) Fehlerabschätzungen mit Verwendung der Lipschitzkonstante L

Mit (2.11) und (2.13) folgt

$$\left\{ \begin{array}{lcl} |x^{(\nu+1)} - x^{(\nu)}| & \leq & L|x^{(\nu)} - x^{(\nu-1)}|, \\ |x^{(\nu+2)} - x^{(\nu+1)}| & \leq & L|x^{(\nu+1)} - x^{(\nu)}| \leq L^2|x^{(\nu)} - x^{(\nu-1)}|, \\ \vdots & & \vdots \\ |x^{(\nu+m)} - x^{(\nu+m-1)}| & \leq & L^m|x^{(\nu)} - x^{(\nu-1)}|. \end{array} \right. \quad (2.17)$$

Aus (2.17) erhält man weiter mit der Dreiecksungleichung

$$\begin{aligned} |x^{(\nu+m)} - x^{(\nu)}| &= |x^{(\nu+m)} - x^{(\nu+m-1)} + x^{(\nu+m-1)} - \dots + x^{(\nu+1)} - x^{(\nu)}| \\ &\leq |x^{(\nu+m)} - x^{(\nu+m-1)}| + |x^{(\nu+m-1)} - x^{(\nu+m-2)}| + \dots + |x^{(\nu+1)} - x^{(\nu)}| \\ &\leq (L^m + L^{m-1} + \dots + L) |x^{(\nu)} - x^{(\nu-1)}| \\ &= L \sum_{j=0}^{m-1} L^j |x^{(\nu)} - x^{(\nu-1)}| \\ &= L \frac{1 - L^m}{1 - L} |x^{(\nu)} - x^{(\nu-1)}|. \end{aligned}$$

Wegen $0 < 1 - L^m \leq 1$ folgt daraus

$$|x^{(\nu+m)} - x^{(\nu)}| \leq \frac{L}{1 - L} |x^{(\nu)} - x^{(\nu-1)}|. \quad (2.18)$$

Setzt man in der letzten Zeile von (2.17) $\nu = 1$, $m = \bar{\nu} - 1$ und schreibt anschließend wieder ν statt $\bar{\nu}$, so erhält man $|x^{(\nu)} - x^{(\nu-1)}| \leq L^{\nu-1} |x^{(1)} - x^{(0)}|$. Geht man damit in (2.18) ein, so ergibt sich

$$|x^{(\nu+m)} - x^{(\nu)}| \leq \frac{L}{1 - L} |x^{(\nu)} - x^{(\nu-1)}| \leq \frac{L^\nu}{1 - L} |x^{(1)} - x^{(0)}|. \quad (2.19)$$

Für festes ν und $m \rightarrow \infty$ folgen aus (2.19) mit $\lim_{m \rightarrow \infty} x^{(\nu+m)} = \xi$

1. die *a posteriori*-Fehlerabschätzung

$$|\Delta^{(\nu)}| = |x^{(\nu)} - \xi| \leq \frac{L}{1 - L} |x^{(\nu)} - x^{(\nu-1)}| = \varepsilon_1 \quad (2.20)$$

und

2. die *a priori*-Fehlerabschätzung

$$|\Delta^{(\nu)}| = |x^{(\nu)} - \xi| \leq \frac{L^\nu}{1 - L} |x^{(1)} - x^{(0)}| = \varepsilon_2 \quad (2.21)$$

mit $\varepsilon_1 \leq \varepsilon_2$.

Die a priori-Fehlerabschätzung (2.21) kann bereits nach dem ersten Iterationsschritt vorgenommen werden. Sie dient vor allem dazu, bei vorgegebener Fehlerschranke ε die Anzahl ν der höchstens erforderlichen Iterationsschritte abzuschätzen, denn aus der Forderung

$$|x^{(\nu)} - \xi| \leq \frac{L^\nu}{1-L} |x^{(1)} - x^{(0)}| \stackrel{!}{\leq} \varepsilon$$

ergibt sich mit $L^\nu \leq \frac{\varepsilon(1-L)}{|x^{(1)} - x^{(0)}|} =: K$ und mit $\log L < 0$ für $0 < L < 1$ die Ungleichung

$$\nu \geq \frac{\log K}{\log L}. \quad (2.22)$$

Die a posteriori-Fehlerabschätzung (2.20) kann erst im Verlauf oder nach Abschluss der Rechnung genutzt werden, da sie $x^{(\nu)}$ als bekannt voraussetzt; sie liefert eine bessere Schranke als die a priori-Fehlerabschätzung und wird deshalb vorzugsweise zur Abschätzung des Fehlers verwendet. Um rasche Konvergenz zu erreichen, sollten die Schrittfunction φ und das zugehörige Intervall I so gewählt werden, dass $L < 0.2$ gilt.

Wenn $\frac{L}{1-L} \leq 1$ ist, also $L \leq 1/2$, folgt aus (2.20) $|x^{(\nu)} - \xi| \leq |x^{(\nu)} - x^{(\nu-1)}|$, d. h. der absolute Fehler von $x^{(\nu)}$ ist kleiner (für $L < 1/2$) oder höchstens gleich (für $L = 1/2$) der absoluten Differenz der letzten beiden Näherungen. Jeder Iterationsschritt bringt demnach eine bessere Annäherung der gesuchten Lösung. Für $1/2 < L < 1$ kann jedoch der absolute Fehler von $x^{(\nu)}$ größer sein als $|x^{(\nu)} - x^{(\nu-1)}|$, so dass hier die Iterationsfolge noch nichts über den Fehler aussagen kann.

Im Falle *monotoner Konvergenz* ($0 \leq \varphi'(x) < 1$) folgt aus der Abschätzung $|x^{(\nu)} - \xi| \leq \varepsilon$ (bzw. $x^{(\nu)} - \varepsilon \leq \xi \leq x^{(\nu)} + \varepsilon$) eine schärfere Eingrenzung der Lösung: $x^{(\nu)} \leq \xi \leq x^{(\nu)} + \varepsilon$, falls die $x^{(\nu)}$ von links gegen ξ streben, $x^{(\nu)} - \varepsilon \leq \xi \leq x^{(\nu)}$, falls die $x^{(\nu)}$ von rechts gegen ξ streben.

(ii) Praktikable Fehlerabschätzungen ohne Verwendung der Lipschitzkonstante L

(1) Bei alternierender Konvergenz

Im Falle alternierender Konvergenz ($-1 < \varphi'(x) \leq 0$) schachteln die iterierten Werte die Lösung von links und rechts ein. Aus der anschließenden Skizze lässt sich deshalb leicht der folgende Zusammenhang erklären:

$$\left| x^{(\nu+1)} - \xi \right| \leq \frac{1}{2} \left| x^{(\nu+1)} - x^{(\nu)} \right| = \alpha; \quad (2.23)$$

damit hat man eine Möglichkeit, den Fehler ohne Kenntnis der Lipschitzkonstante abzuschätzen.

(2) Fehlerabschätzung mit dem Satz von Bolzano (Zwischenwertsatz)

Für die Rechenpraxis empfiehlt sich eine in [KIOU1979] angegebene Methode, durch die man im Allgemeinen genauere Schranken erzielen kann als mit der a posteriori-Fehlerabschätzung. Zudem hat sie den großen Vorteil, ohne Lipschitzkonstante auszukommen und sowohl für monotone als auch für alternierende Konvergenz anwendbar zu sein; sie bezieht sich auf die gegebene Funktion f und nicht auf die Schrittfunction φ .

Sei $x^{(\nu)}$ eine iterativ bestimmte Näherung für die Nullstelle ξ ungerader Ordnung von f (Abb. 2.10) und gelte für ein vorgegebenes $\varepsilon > 0$

$$f(x^{(\nu)} - \varepsilon) \cdot f(x^{(\nu)} + \varepsilon) < 0, \quad (2.24)$$

so folgt daraus nach dem Zwischenwertsatz von Bolzano (Satz 2.4), dass im Intervall $(x^{(\nu)} - \varepsilon, x^{(\nu)} + \varepsilon)$ eine Nullstelle ξ liegen muss. Damit gilt die Fehlerabschätzung $|x^{(\nu)} - \xi| < \varepsilon$.

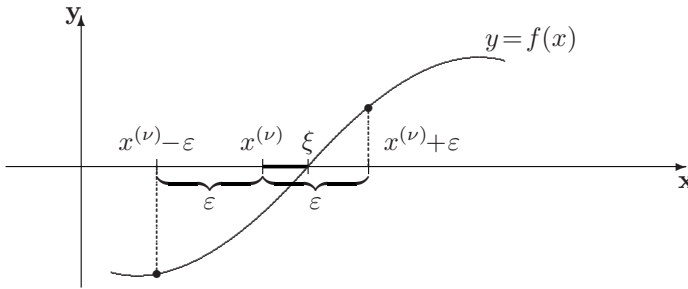


Abb. 2.10. $f(x^{(\nu)} - \varepsilon) \cdot f(x^{(\nu)} + \varepsilon) < 0 \Rightarrow |x^{(\nu)} - \xi| < \varepsilon$

Praktisch geht man bei der Fehlerabschätzung mit Bolzano wie folgt vor: Man setzt zunächst ein ε fest, welches sich über das Abbruchkriterium für die Iteration sinnvoll festlegen lässt (z. B. $\varepsilon = 10^{-k}$, $k \in \mathbb{N}$). Für dieses ε prüft man die Bedingung (2.24), wobei eine Rechnung mit doppelter Genauigkeit zu empfehlen ist. Ist (2.24) erfüllt, so ist ε eine obere Schranke für den absoluten Fehler. Um eine möglichst kleine obere Schranke zu erhalten, führt man die Rechnung noch einmal mit einem kleineren ε durch (z. B. mit $\varepsilon_1 = 10^{-k-1}$). Ist (2.24) auch für ε_1 erfüllt, so ist ε_1 für $|x^{(\nu)} - \xi|$ eine kleinere obere Schranke als ε . Analog fährt man so lange fort, bis sich (2.24) für ein ε_j nicht mehr erfüllen lässt (z. B. $\varepsilon_j = 10^{-k-j}$). Dann ist ε_{j-1} (z. B. $\varepsilon_{j-1} = 10^{-k-j+1}$) die genaueste Fehlerschranke, die man auf diese Weise erhalten hat.

Es wurde bisher vorausgesetzt, dass für diese Art der Fehlerabschätzung eine Nullstelle ξ ungerader Ordnung vorliegen muss. Unter Verwendung des Satzes 2.3 kann man sie aber auch für Nullstellen ξ gerader Ordnung einsetzen, indem man anstelle der Funktion f die Funktion $g = f/f'$ für die Fehlerabschätzung verwendet, da ξ dann eine einfache Nullstelle von g ist. Statt (2.24) ergibt sich hier die analoge Bedingung

$$g(x^{(\nu)} - \varepsilon) \cdot g(x^{(\nu)} + \varepsilon) < 0.$$

Rechnungsfehler

Es sei $\varepsilon^{(\nu)}$ der lokale Rechnungsfehler des ν -ten Iterationsschrittes, der bei der Berechnung von $x^{(\nu)} = \varphi(x^{(\nu-1)})$ entsteht. Gilt $|\varepsilon^{(\nu)}| \leq \varepsilon$ für $\nu = 0, 1, 2, \dots$, so ergibt sich für den akkumulierten Rechnungsfehler des ν -ten Iterationsschrittes

$$|r^{(\nu)}| \leq \frac{\varepsilon}{1-L}, \quad 0 \leq L < 1.$$

Die Fehlerschranke $\varepsilon/(1-L)$ ist also unabhängig von der Anzahl ν der Iterationsschritte; der Algorithmus (2.11) ist somit stabil (vgl. Anmerkung nach Definition 1.36).

Da sich der Gesamtfehler aus dem Verfahrensfehler und dem Rechnungsfehler zusammensetzt, sollten Rechnungsfehler und Verfahrensfehler von etwa gleicher Größenordnung sein. Dann ergibt sich aus (2.21) bei bekanntem L die Beziehung

$$\frac{L^\nu}{1-L} |x^{(1)} - x^{(0)}| \approx \frac{\varepsilon}{1-L};$$

mit $L^\nu \approx \frac{\varepsilon}{|x^{(1)} - x^{(0)}|}$ und $0 < L < 1$ folgt für die Anzahl der höchstens auszuführenden Iterationsschritte

$$\nu \geq \left(\log \frac{\varepsilon}{|x^{(1)} - x^{(0)}|} \right) / \log L.$$

Beispiel 2.16. (Fortsetzung von Beispiel 2.15)

Mit der Iterationsvorschrift $x^{(\nu+1)} = -1 + 2/x^{(\nu)}$ und $I = [-3, -1.5]$ erhält man ausgehend vom Startwert $x^{(0)} = -3$ bei Rechnung mit 14 Stellen die folgenden Werte $x^{(\nu)}$, die mit 8-stelliger Mantisse angegeben sind. Es wird so lange iteriert, bis die Forderung $|x^{(\nu)} - x^{(\nu-1)}| \leq 0.5\text{e-}4$ erfüllt ist.

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ |
|-------|-------------|-----------------------------|
| 0 | -3.0000000 | |
| 1 | -1.6666666 | 1.3333333 |
| 2 | -2.2000000 | 0.5333333 |
| 3 | -1.9090909 | 0.2909091 |
| 4 | -2.0476190 | 0.1385281 |
| 5 | -1.9767442 | 0.0708749 |
| 6 | -2.0117647 | 0.0350205 |
| 7 | -1.9941520 | 0.0176127 |
| 8 | -2.0029326 | 0.0087805 |
| 9 | -1.9985359 | 0.0043967 |
| 10 | -2.0007326 | 0.0021967 |
| 11 | -1.9996338 | 0.0010988 |
| 12 | -2.0001831 | 0.0005493 |
| 13 | -1.9999085 | 0.0002747 |
| 14 | -2.0000458 | 0.0001373 |
| 15 | -1.9999771 | 0.0000687 |
| 16 | -2.0000114 | 0.0000343 |

1. a posteriori-Fehlerabschätzung: Mit $L = 0.89$ (vgl. Beispiel 2.15) folgt

$$|x^{(16)} - \xi| \leq \frac{L}{1-L} |x^{(16)} - x^{(15)}| \leq \frac{0.89}{1-0.89} \cdot 0.35e-4 \leq 0.29e-3$$

2. Die a priori-Fehlerabschätzung hätte man benutzen können, um im Voraus eine Aussage über die maximal notwendige Anzahl ν von Iterationsschritten zu erhalten. Mit $|x^{(\nu)} - \xi| \leq \varepsilon = 0.5e-3$ und $|x^{(1)} - x^{(0)}| = 1.33$ ergeben sich $K = (0.5e-3(1-0.89))/1.33 = 4.135e-5$ und mit (2.22)

$$\nu \geq \frac{\lg K}{\lg L} = \frac{-4.384}{-0.051} = 86.6.$$

Es sind also höchstens $\nu = 87$ Iterationsschritte erforderlich, um die geforderte Genauigkeit zu erreichen. In vielen Fällen wird man aber mit weit weniger Schritten auskommen, denn die a priori-Fehlerabschätzung ist sehr grob. Die a posteriori-Fehlerabschätzung zeigt, dass diese Genauigkeitsforderung bereits nach 16 Iterationsschritten erfüllt ist. Um rasche Konvergenz zu erreichen, sollten die Schrittfunction φ und das zugehörige Intervall I so gewählt werden, dass $L < 0.2$ gilt. Dann sind auch die Fehlerabschätzungen genauer.

3. Fehlerabschätzung bei alternierender Konvergenz:

Wegen $\varphi'(x) = -2/x^2 < 0$ für alle $x \in I$ liegt der Fall der alternierenden Konvergenz vor. Mit $|x^{(16)} - x^{(15)}| = 0.35e-4$ erhält man somit gemäß (2.23)

$$|x^{(16)} - \xi| \leq \frac{1}{2} |x^{(16)} - x^{(15)}| = \frac{0.35}{2} e-4 = 0.175e-4 < 0.18e-4,$$

also eine sogar präzisere Abschätzung als die a posteriori-Fehlerabschätzung, weil $L > \frac{1}{2}$ ist. Für die Praxis ist wichtig, dass diese Fehlerabschätzung ohne Lipschitzkonstante einsetzbar ist.

4. Fehlerabschätzung unter Verwendung des Satzes von Bolzano, also ohne Verwendung einer Lipschitzkonstante. Gegeben ist die Gleichung $f(x) = x^2 + x - 2 = 0$. Untersucht wird der absolute Fehler von $x^{(16)} = -2.0000114$.

1. Wahl: $\varepsilon_1 = 10^{-3}$. Dann folgt aus $f(x^{(16)} - \varepsilon_1) > 0$ und $f(x^{(16)} + \varepsilon_1) < 0$, dass für den absoluten Fehler gilt: $|x^{(16)} - \xi| < 10^{-3}$.
2. Wahl: $\varepsilon_2 = 10^{-4}$. Hier ergibt sich ebenfalls $f(x^{(16)} - \varepsilon_2) > 0$, $f(x^{(16)} + \varepsilon_2) < 0$, also $|x^{(16)} - \xi| < 10^{-4}$.
3. Wahl: $\varepsilon_3 = 10^{-5}$. Man erhält $f(x^{(16)} - \varepsilon_3) > 0$ und $f(x^{(16)} + \varepsilon_3) > 0$, also ist ε_3 zu klein gewählt. Ein weiterer Versuch wird unternommen mit der
4. Wahl: $\varepsilon_4 = 0.5e-4$. Man erhält $f(x^{(16)} - \varepsilon_4) > 0$ und $f(x^{(16)} + \varepsilon_4) < 0$; also gilt für den absoluten Fehler $|x^{(16)} - \xi| < 0.5e-4$.

Tatsächlich ist $|x^{(16)} - \xi| = 0.0000114$. Man sieht, dass diese Fehlerabschätzung nach Bolzano am besten geeignet ist, um den absoluten Fehler möglichst genau anzugeben. Zudem ist nur sie praktikabel, denn meist ist es bedeutend schwieriger als beim vorliegenden Beispiel, eine möglichst kleine Lipschitzkonstante anzugeben. \square

2.3.5 Praktische Durchführung

Bei der Ermittlung der Lösung einer Gleichung $f(x) = 0$ mit Hilfe des allgemeinen Iterationsverfahrens geht man wie folgt vor:

Algorithmus 2.17.

Gesucht ist eine Lösung ξ der Gleichung $f(x) = 0$.

1. Schritt: Festlegung eines Intervalls I , in welchem mindestens eine Nullstelle von f liegt.
2. Schritt: Äquivalente Umformung von $f(x) = 0$ in eine Gleichung der Gestalt $x = \varphi(x)$.
3. Schritt: Prüfung, ob die Funktion φ für alle $x \in I$ die Voraussetzungen des Satzes 2.11 erfüllt.
4. Schritt: Aufstellung der Iterationsvorschrift $x^{(\nu+1)} := \varphi(x^{(\nu)})$, $\nu = 0, 1, 2, \dots$ und Wahl eines beliebigen Startwertes $x^{(0)} \in I$.
5. Schritt: Berechnung der Iterationsfolge $\{x^{(\nu)}\}$, $\nu = 1, 2, \dots$. Die Iteration ist so lange fortzusetzen, bis mit einer Schranke $\delta_1 > 0$ für den relativen Fehler

$$|x^{(\nu+1)} - x^{(\nu)}| \leq \delta_1 |x^{(\nu+1)}| \quad (2.25)$$

oder mit einer Schranke $\delta_2 > 0$ für den absoluten Fehler

$$|x^{(\nu+1)} - x^{(\nu)}| \leq \delta_2 \quad (2.26)$$

und zusätzlich mit $\delta_3 > 0$

$$|f(x^{(\nu+1)})| \leq \delta_3$$

gilt. Dabei ist (2.25) im Allgemeinen der Abbruchbedingung (2.26) vorzuziehen. Beide Bedingungen sind im kombinierten Fehlertest (1.5) enthalten.

6. Schritt: Fehlerabschätzung (s. Abschnitt 2.3.4).

Beispiel 2.18.

Gegeben: $f(x) = \cos x + 1 - \sqrt{x}$, $x \geq 0$.

Gesucht: Eine Lösung ξ der Gleichung $f(x) = 0$ mit dem allgemeinen Iterationsverfahren.

Lösung: (vgl. Algorithmus 2.17)

1. Schritt: Überblick über Lage und Anzahl der Lösungen von $f(x) = 0$. Mit

$$f_1(x) := \cos x + 1, \quad f_2(x) := \sqrt{x}$$

gilt

$$f(x) = f_1(x) - f_2(x) = 0 \iff f_1(x) = f_2(x).$$

Daher liefert die Abszisse eines Schnittpunktes der Graphen von f_1 und f_2 eine Näherung für eine Lösung ξ der Gleichung $f(x) = 0$ (Abbildung 2.11).

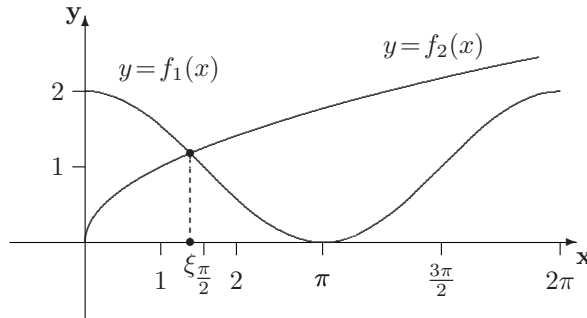


Abb. 2.11.

Im Intervall $[0, \pi]$ sind f_1 monoton fallend und f_2 monoton wachsend. Wegen $f_1(0) = 2 > 0 = f_2(0)$ und $f_1(\pi) = 0 < \sqrt{\pi} = f_2(\pi)$ haben die Graphen von f_1 und f_2 in $[0, \pi]$ genau einen Schnittpunkt.

Wegen $f_1(x) \leq 2$ und $f_2(x) = \sqrt{x} \geq 2$ für $x \geq 4$ gibt es keine weiteren Schnittpunkte.

Aus Abbildung 2.11 liest man ab $\xi \in [1, \frac{\pi}{2}]$. Mit $f(1.3) = 0.127$ und $f(1.5) = -0.154$ ergibt sich $I = [1.3, 1.5]$ als ein engeres Einschlussintervall für die Lösung ξ .

2. Schritt: Äquivalente Auflösung von $f(x) = 0$. Die naheliegende Auflösung

$$x = (\cos x + 1)^2 =: \varphi(x)$$

mit

$$\varphi'(x) = -2(\cos x + 1) \sin x$$

erweist sich wegen $|\varphi'(x)| > 2$ für $x \in I$ als unbrauchbar. Daher wird die Auflösung

$$x = \arccos(\sqrt{x} - 1) =: \varphi(x)$$

gewählt mit

$$\varphi'(x) = \frac{-1}{2\sqrt{x}\sqrt{1-(\sqrt{x}-1)^2}}, \quad x > 0.$$

3. Schritt: Genügt φ für $x \in I$ den Voraussetzungen (i) und (ii) des Satzes 2.11?

zu (i):

Wegen $\varphi'(x) < 0$ für $x > 0$ ist φ streng monoton fallend. Mit $\varphi(1.3) = 1.43 \in I$ und $\varphi(1.5) = 1.34 \in I$ gilt wegen der Monotonie auch $\varphi(x) \in I$ für alle $x \in I = [1.3, 1.5]$. Also ist die Bedingung (i) erfüllt.

zu (ii):

Um eine Lipschitzkonstante L zu ermitteln, wird $\varphi'(x)$ in I grob abgeschätzt:

$$\begin{aligned} |\varphi'(x)| &= \frac{1}{2\sqrt{x} \sqrt{1 - (\sqrt{x} - 1)^2}} \\ &\leq \frac{1}{2 \min_{x \in I} \sqrt{x} \cdot \min_{x \in I} \sqrt{1 - (\sqrt{x} - 1)^2}} \\ &= \frac{1}{2\sqrt{1.3} \sqrt{1 - (\sqrt{1.5} - 1)^2}} = 0.450 = L < \frac{1}{2} < 1. \end{aligned}$$

φ erfüllt also die Voraussetzungen der Sätze 2.11 und 2.14.

4. Schritt: Iterationsvorschrift

$$x^{(\nu+1)} = \varphi(x^{(\nu)}) = \arccos(\sqrt{x^{(\nu)}} - 1), \quad \nu = 0, 1, 2, \dots$$

mit dem Startwert $x^{(0)} = 1.3 \in I$.

5. Schritt: Wegen $\varphi'(x) < 0$ in I liegt alternierende Konvergenz vor, so dass mit der zugehörigen Fehlerabschätzung eine Abbruchbedingung konstruiert werden kann, die eine bestimmte Anzahl genauer Dezimalen garantiert. Für drei genaue Dezimalen erhält man mit (2.23)

$$|x^{(\nu)} - \xi| \leq \frac{1}{2} |x^{(\nu)} - x^{(\nu-1)}| \stackrel{!}{\leq} 0.5\text{e-}3$$

und damit die Abbruchbedingung

$$|x^{(\nu)} - x^{(\nu-1)}| \leq 1\text{e-}3.$$

Iteration:

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ |
|-------|---------------|-----------------------------|
| 0 | 1.3 | . |
| 1 | 1.430 157 740 | . |
| 2 | 1.373 629 308 | . |
| 3 | 1.397 917 137 | . |
| 4 | 1.387 435 119 | |
| 5 | 1.391 950 063 | 0.004 514 944 |
| 6 | 1.390 003 705 | 0.001 946 358 |
| 7 | 1.390 842 462 | 0.000 838 757 |

Bei $\nu = 7$ ist die Abbruchbedingung erfüllt, also ist

$$\xi = 1.391 \approx 1.390\,842\,462 = x^{(7)}.$$

6. Schritt: Fehlerabschätzungen

Die Abbruchbedingung war so gewählt, dass mit der Fehlerabschätzung für alternierende Konvergenz bereits 3 genaue Dezimalen erreicht wurden.

Fehlerabschätzung nach Bolzano: $|x^{(\nu)} - \xi| < \varepsilon$

Annahme: $\varepsilon_1 = 0.5\text{e-}3$

$$\begin{aligned} f(x^{(7)} - \varepsilon_1) &= f(1.390\,342\,462) = 0.000\,348\dots > 0 \\ f(x^{(7)} + \varepsilon_1) &= f(1.391\,342\,462) = -0.001\,059\dots < 0 \\ \Rightarrow |x^{(7)} - \xi| &< 0.5\text{e-}3 \quad \text{d. h. 3 genaue Dezimalen} \end{aligned}$$

Annahme: $\varepsilon_2 = 0.5\text{e-}4$

$$\begin{aligned} f(x^{(7)} - \varepsilon_2) &= f(1.390\,792\,462) = -0.000\,285\dots < 0 \\ f(x^{(7)} + \varepsilon_2) &= f(1.390\,892\,462) = -0.000\,426\dots < 0 \\ \Rightarrow \varepsilon_2 &\text{ zu klein! } \square \end{aligned}$$

2.4 Konvergenzordnung eines Iterationsverfahrens

Bei Iterationsverfahren können die Anzahl der erforderlichen Iterationsschritte und auch der Rechenaufwand im Allgemeinen nicht im Voraus ermittelt werden. Die Konvergenzordnung kann aber als Maßstab für den erforderlichen Rechenaufwand eines Verfahrens dienen.

Definition 2.19. (*Konvergenzordnung*)

Die Iterationsfolge $\{x^{(\nu)}\}$ konvergiert von mindestens p -ter Ordnung gegen ξ , wenn eine Konstante M , $0 \leq M < \infty$ existiert, so dass für $p \in \mathbb{R}$, $p \geq 1$, gilt

$$\lim_{\nu \rightarrow \infty} \frac{|x^{(\nu+1)} - \xi|}{|x^{(\nu)} - \xi|^p} = M. \quad (2.27)$$

Das Iterationsverfahren $x^{(\nu+1)} = \varphi(x^{(\nu)})$ heißt dann ein Verfahren von mindestens p -ter Ordnung; es besitzt genau die Ordnung p , wenn $M \neq 0$ ist.

Durch (2.27) wird also ausgedrückt, dass der Fehler der $(\nu+1)$ -ten Näherung ungefähr gleich M -mal der p -ten Potenz des Fehlers der ν -ten Näherung ist. Die Konvergenzgeschwindigkeit wächst mit der Konvergenzordnung. Bei $p = 1$ spricht man von *linearer Konvergenz*, bei $p = 2$ von *quadratischer Konvergenz* und allgemein bei $p > 1$ von *superlinearer Konvergenz*. Es gilt der

Satz 2.20.

Die Schrittfunktion φ sei für $x \in I$ p -mal stetig differenzierbar. Gilt dann mit $\lim_{\nu \rightarrow \infty} x^{(\nu)} = \xi$

$$\varphi(\xi) = \xi, \varphi'(\xi) = \varphi''(\xi) = \dots = \varphi^{(p-1)}(\xi) = 0, \varphi^{(p)}(\xi) \neq 0,$$

so ist $x^{(\nu+1)} = \varphi(x^{(\nu)})$ ein Iterationsverfahren der Ordnung p mit

$$M = \frac{1}{p!} |\varphi^{(p)}(\xi)| \leq \frac{1}{p!} \max_{x \in I} |\varphi^{(p)}(x)| \leq M_1.$$

Im Fall $p = 1$ gilt zusätzlich $M = |\varphi'(\xi)| < 1$.

Beweis. Die Taylorentwicklung der Schrittfunktion φ an der Stelle ξ lautet

$$\varphi(x^{(\nu)}) = \varphi(\xi) + (x^{(\nu)} - \xi)\varphi'(\xi) + \frac{1}{2}(x^{(\nu)} - \xi)^2\varphi''(\xi) + \dots$$

und mit $x^{(\nu+1)} = \varphi(x^{(\nu)})$ und $\xi = \varphi(\xi)$ ergibt sich

$$x^{(\nu+1)} - \xi = \varphi(x^{(\nu)}) - \varphi(\xi) = (x^{(\nu)} - \xi)\varphi'(\xi) + \frac{1}{2}(x^{(\nu)} - \xi)^2\varphi''(\xi) + \dots \quad (2.28)$$

Wegen $\varphi(\xi) = \xi, \varphi'(\xi) = \dots = \varphi^{(p-1)}(\xi) = 0, \varphi^{(p)}(\xi) \neq 0$ erhält (2.28) die Form

$$x^{(\nu+1)} - \xi = \frac{(x^{(\nu)} - \xi)^p}{p!} \varphi^{(p)}(\xi) + O\left((x^{(\nu)} - \xi)^{p+1}\right)$$

bzw.

$$\frac{x^{(\nu+1)} - \xi}{(x^{(\nu)} - \xi)^p} = \frac{1}{p!} \varphi^{(p)}(\xi) + O\left((x^{(\nu)} - \xi)\right).$$

Durch Grenzübergang folgt weiter

$$\lim_{\nu \rightarrow \infty} \left| \frac{x^{(\nu+1)} - \xi}{(x^{(\nu)} - \xi)^p} \right| = \frac{1}{p!} |\varphi^{(p)}(\xi)| = M$$

mit

$$M \leq \frac{1}{p!} \max_{x \in I} |\varphi^{(p)}(x)| \leq M_1.$$

Im Fall $p = 1$ ist das Iterationsverfahren $x^{(\nu+1)} = \varphi(x^{(\nu)})$ mit $\varphi'(\xi) \neq 0$ ein Verfahren erster Ordnung, und wegen der Lipschitzbedingung (2.14) ist $M = |\varphi'(\xi)| \leq L < 1$. \square

Es gilt außerdem der folgende in [COLL1968], S.231 bewiesene Satz, der zur Konstruktion von Iterationsverfahren beliebig hoher Konvergenzordnung eingesetzt werden kann.

Satz 2.21.

Sind $x^{(\nu+1)} = \varphi_1(x^{(\nu)})$ und $x^{(\nu+1)} = \varphi_2(x^{(\nu)})$ zwei Iterationsverfahren der Konvergenzordnungen p_1 bzw. p_2 , so ist

$$x^{(\nu+1)} = \varphi_1(\varphi_2(x^{(\nu)}))$$

ein Iterationsverfahren, das mindestens die Konvergenzordnung $p_1 \cdot p_2$ besitzt.

Ist beispielsweise $x^{(\nu+1)} = \varphi(x^{(\nu)})$ ein Iterationsverfahren der Konvergenzordnung $p > 1$, so erhält man durch die Schrittfunction $\varphi_s(x)$ mit

$$\begin{aligned} \varphi_1(x) &= \varphi(x), \\ \varphi_s(x) &= \varphi(\varphi_{s-1}(x)) \quad \text{für } s = 2, 3, \dots \end{aligned}$$

ein Iterationsverfahren der Konvergenzordnung p^s .

2.5 Newtonsche Verfahren

2.5.1 Das Newtonsche Verfahren für einfache Nullstellen

Die Funktion f sei im Intervall $[a, b]$ stetig differenzierbar und besitze in (a, b) eine *einfache Nullstelle* ξ ; also sind $f(\xi) = 0$ und $f'(\xi) \neq 0$. Ferner sei $f'(x) \neq 0$ für alle $x \in [a, b]$.

Zum Newtonschen Iterationsverfahren führt eine einfache geometrische Überlegung.

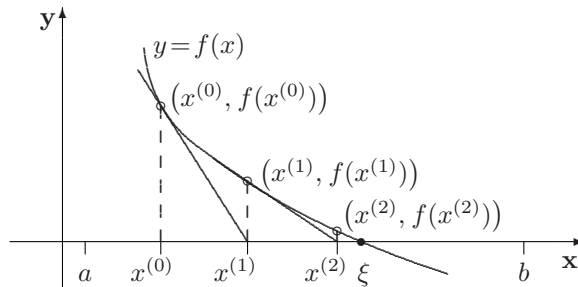


Abb. 2.12. Newtonsches Iterationsverfahren

Mit $x^{(0)} \in [a, b]$ wird im Punkt $(x^{(0)}, f(x^{(0)}))$ die Tangente des Graphen von f erzeugt und dann deren Schnittpunkt $(x^{(1)}, 0)$ mit der x -Achse bestimmt. Aus der Gleichung dieser Tangente

$$y = f'(x^{(0)})(x - x^{(0)}) + f(x^{(0)})$$

folgt mit $x = x^{(1)}$ und $y = 0$

$$x^{(1)} = x^{(0)} - \frac{f(x^{(0)})}{f'(x^{(0)})};$$

$x^{(1)}$ wird als eine gegenüber $x^{(0)}$ verbesserte Näherung für die Nullstelle ξ angesehen. Mit $x^{(1)}$ wird das Verfahren in derselben Weise fortgesetzt. Damit ergibt sich die nächste Näherung

$$x^{(2)} = x^{(1)} - \frac{f(x^{(1)})}{f'(x^{(1)})}.$$

Für dieses geometrisch plausible Verfahren werden im folgenden Satz hinreichende Bedingungen für die Konvergenz der Folge $x^{(0)}, x^{(1)}, x^{(2)}, \dots$ gegen die Nullstelle ξ angegeben.

Satz 2.22.

Die Funktion f sei im Intervall $[a, b]$ zweimal stetig differenzierbar und besitze in (a, b) eine einfache Nullstelle ξ ; also sind $f(\xi) = 0$ und $f'(\xi) \neq 0$.

Dann gibt es ein Intervall $I = (\xi - r, \xi + r) \subset [a, b]$, $r > 0$, so dass die Iterationsfolge

$$x^{(\nu+1)} = x^{(\nu)} - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}, \quad \nu = 0, 1, 2, \dots, \quad (2.29)$$

für das Verfahren von Newton mit jedem Startwert $x^{(0)} \in I$ gegen die Nullstelle ξ konvergiert, und zwar von mindestens zweiter Ordnung.

Beweis. Vergleicht man (2.29) mit (2.4) und (2.2), so ergibt sich für die Schrittfunktion φ des Newton-Verfahrens

$$\varphi(x) = x - \frac{f(x)}{f'(x)}. \quad (2.30)$$

(2.3) zeigt, dass hier $g(x) = 1/f'(x)$ gesetzt wurde.

Um den Fixpunktsatz 2.14 auf die Schrittfunktion (2.30) anwenden zu können, muss ein Intervall I angegeben werden, so dass φ für alle $x \in I$ den Voraussetzungen (i) und (ii) des Satzes 2.11 genügt.

Da f' stetig ist, gibt es wegen $f'(\xi) \neq 0$ eine Umgebung der Nullstelle ξ , in der $f'(x) \neq 0$ ist. Diese Umgebung sei $I_s = (\xi - s, \xi + s) \subset [a, b]$, $s > 0$. Somit ist die Schrittfunktion φ (vgl. (2.30)) stetig im Intervall I_s .

Die Ableitung der Schrittfunktion (2.30) ist

$$\varphi'(x) = \frac{f(x)f''(x)}{f'^2(x)}. \quad (2.31)$$

Auch φ' ist stetig in I_s .

Mit $f(\xi) = 0$ und $f'(\xi) \neq 0$ ist

$$\varphi'(\xi) = 0. \quad (2.32)$$

Wegen der Stetigkeit von φ' in I_s gibt es eine Umgebung von ξ , in der $|\varphi'(x)| \leq L$ ist mit $0 < L < 1$. Diese Umgebung sei das Intervall

$$I = (\xi - r, \xi + r) \subseteq I_s \subset [a, b], \quad 0 < r \leq s. \quad (2.33)$$

Aus

$$|\varphi'(x)| \leq L \quad \text{für alle } x \in I \quad (2.34)$$

mit $0 < L < 1$ folgt nach (2.14), dass die Schrittfunktion φ lipschitzbeschränkt ist, und damit gilt (ii).

Um (i) nachzuweisen, muss gezeigt werden, dass mit $x \in I$, also $|x - \xi| < r$, auch $\varphi(x) \in I$ ist, d. h. $|\varphi(x) - \xi| < r$.

Wegen $\varphi(\xi) = \xi$ (vgl. (2.30)) ist mit (ii)

$$|\varphi(x) - \xi| = |\varphi(x) - \varphi(\xi)| \leq L |x - \xi| < |x - \xi| < r.$$

Damit gilt auch (i).

Nach dem Fixpunktsatz 2.14 konvergiert also die Iterationsfolge (2.29) für jeden Startwert $x^{(0)} \in I$ gegen die Nullstelle ξ .

Wegen (2.32) konvergiert das Newtonsche Verfahren (2.29) nach Satz 2.20 von mindestens zweiter Ordnung, d. h. mindestens quadratisch. \square

Aus dem Beweis des Satzes 2.22 geht hervor, dass das Newtonsche Verfahren nur für solche Startwerte $x^{(0)}$ konvergiert, die genügend nahe bei der Nullstelle ξ liegen.

Wegen der speziellen Schrittfunktion (2.30) kann eine gegenüber der a posteriori-Fehlerabschätzung (2.20) in Abschnitt 2.3.4 verbesserte Fehlerabschätzung für das Newtonsche Verfahren angegeben werden.

Satz 2.23.

Die Newtonsche Iterationsfolge (2.29) besitze gemäß Satz 2.22 das Konvergenzintervall I . Dann gilt unter Verwendung der a posteriori-Fehlerabschätzung (2.20) mit

$$\frac{1}{2} \frac{\max_{x \in I} |f''(x)|}{\min_{x \in I} |f'(x)|} \leq M_1$$

die Fehlerabschätzung für $\nu = 1, 2, \dots$

$$|x^{(\nu+1)} - \xi| \leq M_1 \left(\frac{L}{1-L} \right)^2 |x^{(\nu)} - x^{(\nu-1)}|^2 \quad (2.35)$$

und allgemeiner mit $m = 1, 2, \dots$

$$|x^{(\nu+m)} - \xi| \leq \frac{1}{M_1} \left(M_1 \frac{L}{1-L} |x^{(\nu)} - x^{(\nu-1)}| \right)^{2^m}. \quad (2.36)$$

Beweis. Nach der Taylorschen Formel gilt

$$0 = f(\xi) = f(x^{(\nu)}) + f'(x^{(\nu)})(\xi - x^{(\nu)}) + \frac{1}{2} f''(\tilde{x}^{(\nu)})(\xi - x^{(\nu)})^2, \quad (2.37)$$

wobei $\tilde{x}^{(\nu)}$ zwischen $x^{(\nu)}$ und ξ liegt. Mit Division von (2.37) durch $f'(x^{(\nu)})$ (für $x^{(\nu)} \in I$ ist $f'(x^{(\nu)}) \neq 0$) erhält man

$$-\frac{f(x^{(\nu)})}{f'(x^{(\nu)})} - \xi + x^{(\nu)} = \frac{1}{2} \frac{f''(\tilde{x}^{(\nu)})}{f'(x^{(\nu)})} (\xi - x^{(\nu)})^2$$

und mit (2.29)

$$x^{(\nu+1)} - \xi = \frac{1}{2} \frac{f''(\tilde{x}^{(\nu)})}{f'(x^{(\nu)})} (x^{(\nu)} - \xi)^2. \quad (2.38)$$

Beim Übergang zu den Beträgen folgt aus (2.38)

$$\begin{aligned} |x^{(\nu+1)} - \xi| &\leq \frac{1}{2} \frac{\max_{x \in I} |f''(x)|}{\min_{x \in I} |f'(x)|} |x^{(\nu)} - \xi|^2 \\ &\leq M_1 |x^{(\nu)} - \xi|^2 = \frac{1}{M_1} (M_1 |x^{(\nu)} - \xi|)^2. \end{aligned} \quad (2.39)$$

Wird (2.20)

$$|x^{(\nu)} - \xi| \leq \frac{L}{1-L} |x^{(\nu)} - x^{(\nu-1)}| \quad (2.40)$$

auf der rechten Seite der Ungleichung (2.39) eingesetzt, so ergibt sich die Fehlerabschätzung (2.35).

Die gegenüber (2.39) allgemeinere Ungleichung für $m = 1, 2, \dots$

$$|x^{(\nu+m)} - \xi| \leq \frac{1}{M_1} (M_1 |x^{(\nu)} - \xi|)^{2^m} \quad (2.41)$$

gilt nach (2.39) für $m = 1$. Mit der Annahme, (2.41) gelte für m , wird (2.41) für $m + 1$ bewiesen (vollständige Induktion). Aus (2.41) für $m + 1$ folgt mit (2.39)

$$\begin{aligned} |x^{(\nu+1+m)} - \xi| &\leq \frac{1}{M_1} (M_1 |x^{(\nu+1)} - \xi|)^{2^m} \\ &\leq \frac{1}{M_1} \left(M_1 \frac{1}{M_1} (M_1 |x^{(\nu)} - \xi|)^2 \right)^{2^m} \\ &= \frac{1}{M_1} (M_1 |x^{(\nu)} - \xi|)^{2^{m+1}}. \end{aligned}$$

Wenn auf der rechten Seite der Ungleichung (2.41) $|x^{(\nu)} - \xi|$ mittels (2.40) ersetzt wird, entsteht die Fehlerabschätzung (2.36). \square

Mit (2.35) kann mit zwei Näherungen $x^{(\nu-1)}$ und $x^{(\nu)}$ der absolute Fehler $|x^{(\nu+1)} - \xi|$ der nächsten (noch nicht berechneten) Näherung $x^{(\nu+1)}$ im Voraus abgeschätzt werden.

Eine Lipschitzkonstante L erhält man durch Abschätzung von (2.31) auf dem Intervall I :

$$|\varphi'(x)| = |f(x)f''(x)/f'^2(x)| \leq L < 1.$$

Mit einer Lipschitzkonstante $L \leq 0.5$ ist $L/(1-L) \leq 1$ und (2.40) lautet

$$|x^{(\nu)} - \xi| \leq |x^{(\nu)} - x^{(\nu-1)}|.$$

Gilt für die ν -te Näherung

$$|x^{(\nu)} - x^{(\nu-1)}| \leq \varepsilon,$$

dann ist auch

$$|x^{(\nu)} - \xi| \leq \varepsilon.$$

Der absolute Fehler von $x^{(\nu)}$ ist dann höchstens gleich ε . So erhält man nach Vorgabe von $\varepsilon > 0$ ein geeignetes Abbruchkriterium.

Beispiel 2.24.

Gegeben: Die Gleichung $f(x) = x^2 - a = 0$, $a > 0$.

Gesucht: Die Lösung $x = \sqrt{a}$ mit dem Newton-Verfahren.

Lösung: Wegen $f'(x) = 2x$ ist die Schrittfunktion (2.30)

$$\varphi(x) = x - \frac{x^2 - a}{2x} = \frac{1}{2} \left(x + \frac{a}{x} \right),$$

und die Iterationsvorschrift (2.29) lautet

$$x^{(\nu+1)} = \frac{1}{2} \left(x^{(\nu)} + \frac{a}{x^{(\nu)}} \right), \quad \nu = 0, 1, 2, \dots \quad (2.42)$$

Ferner ist

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

Als numerisches Beispiel wird $a = 5$ gewählt. Wegen $2 < \sqrt{5} < 3$ sei $I = [2, 3]$. Dann ist für alle $x \in I$

$$|\varphi'(x)| = \frac{1}{2} \left| 1 - \frac{5}{x^2} \right| \leq \frac{1}{2} \left| 1 - \frac{5}{9} \right| \leq 0.23 = L < \frac{1}{2}.$$

Die Folge (2.42) mit $a = 5$ konvergiert also mit jedem Startwert $x^{(0)} \in I = [2, 3]$.

Mit der Lipschitzkonstante $L = 0.23$ ist $L/(1-L) = 0.2987$, und die a posteriori-Fehlerabschätzung (2.40) lautet

$$|x^{(\nu)} - \xi| \leq 0.2987 |x^{(\nu)} - x^{(\nu-1)}|, \quad \nu = 0, 1, 2, \dots \quad (2.43)$$

Mit dem Abbruchkriterium

$$|x^{(\nu)} - x^{(\nu-1)}| \leq 0.5 \cdot 10^{-7}$$

erhält man $\xi = \sqrt{5}$ mit 7 Dezimalen.

Für die Fehlerabschätzung (2.35) werden

$$\min_{x \in I} |f'(x)| = \min_{x \in I} |2x| = 4 \quad \text{und} \quad \max_{x \in I} |f''(x)| = 2$$

benötigt. Damit ist $M_1 = \frac{2}{2.4} = 0.25$. Die Fehlerabschätzung (2.35) lautet

$$|x^{(\nu+1)} - \xi| \leq 0.0223 |x^{(\nu)} - x^{(\nu-1)}|^2, \quad \nu = 0, 1, 2, \dots \tag{2.44}$$

Aus der nachfolgenden Tabelle folgt wegen $|x^{(5)} - x^{(4)}| = 0$ für die Nullstelle $\xi = x^{(5)} = 2.236\,067\,98 \approx \sqrt{5}$.

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ | $ x^{(\nu)} - \sqrt{5} $ | (2.43) | (2.44) |
|-------|-------------|-----------------------------|--------------------------|------------|-----------------------|
| 0 | 3.00000000 | | 0.76393202 | | |
| 1 | 2.33333333 | 0.66666667 | 0.09726536 | 0.19913333 | |
| 2 | 2.23809524 | 0.09523810 | 0.00202726 | 0.02844762 | 0.00991352 |
| 3 | 2.23606890 | 0.00202634 | 0.00000092 | 0.00060527 | 0.00020232 |
| 4 | 2.23606798 | 0.00000092 | 0.00000000 | 0.00000027 | 0.00000009 |
| 5 | 2.23606798 | 0.00000000 | 0.00000000 | 0.00000000 | $1.88 \cdot 10^{-14}$ |

Die mit den Fehlerabschätzungen (2.43) und (2.44) ermittelten Schranken für den absoluten Fehler zeigen im Vergleich mit dem exakten Fehler $|x^{(\nu)} - \sqrt{5}|$, dass die Abschätzung (2.44) kleinere Schranken liefert. \square

Für die einfache Funktion f im vorangehenden Beispiel konnten L und M_1 unschwer bestimmt werden. Im Allgemeinen ist die Ermittlung von L und M_1 zu einem gewählten Intervall I mit einem erheblich größeren Aufwand verbunden.

Bemerkung. Da für jeden Iterationsschritt mit dem Newton-Verfahren ein Funktionswert und ein Ableitungswert zu berechnen sind, eignet sich dieses Verfahren in erster Linie für die Berechnung einzelner Nullstellen eines Polynoms, weil die benötigten Werte sich leicht mit dem Horner-Schema (siehe Kapitel 3) ermitteln lassen, weniger dagegen für die Lösung beliebiger transzendenter Gleichungen. In allen Fällen konvergiert das Newton-Verfahren nur für solche Startwerte, die genügend nahe bei der Nullstelle liegen.

Zur Berechnung von Nullstellen beliebiger transzendenter Funktionen eignet sich besser ein Einschlussverfahren. Siehe dazu Abschnitt 2.9 Entscheidungshilfen.

2.5.2 Gedämpftes Newton-Verfahren

Analog zum gedämpften Newton-Verfahren für nichtlineare Systeme (Abschnitt 6.3.1, Algorithmus 6.9) lässt sich das gedämpfte Newton-Verfahren für Einzelgleichungen angeben. Man führt für $\nu = 0, 1, 2, \dots$ folgende Schritte durch:

- (i) Berechne $\Delta x^{(\nu+1)} := x^{(\nu+1)} - x^{(\nu)} = -f(x^{(\nu)})/f'(x^{(\nu)})$ (Newtonschritt (2.29)).
- (ii) Berechne für $i = 0, 1, \dots$

$$x_i^{(\nu+1)} := x^{(\nu)} + \frac{1}{2^i} \Delta x^{(\nu+1)}.$$

Wenn $|f(x_i^{(\nu+1)})| < |f(x^{(\nu)})|$ gilt, wird $x^{(\nu+1)} := x_i^{(\nu+1)}$ gesetzt. Andernfalls wird der Schritt mit dem nächsten i wiederholt.

Wenn mit einem vorgegebenen $i_{\max} \in \mathbb{N}$ für alle $i \leq i_{\max}$ die obige Bedingung nicht erfüllt ist, wird (mit $i = 0$)

$$x^{(\nu+1)} := x^{(\nu)} + \Delta x^{(\nu+1)}$$

gesetzt.

2.5.3 Das Newtonsche Verfahren für mehrfache Nullstellen. Das modifizierte Newtonsche Verfahren

Die Funktion f sei im Intervall $[a, b]$ genügend oft stetig differenzierbar und besitze in (a, b) eine Nullstelle ξ der Vielfachheit j , $j \geq 2$. Nach Satz 2.2 sind also

$$f(\xi) = f'(\xi) = \dots = f^{(j-1)}(\xi) = 0, \quad f^{(j)}(\xi) \neq 0.$$

Für die Ableitung (2.31) der Schrittfunktion (2.30) des Newtonschen Verfahrens für einfache Nullstellen ergibt sich mit (2.2), (2.4) und (2.6)

$$\varphi'(x) = \frac{f(x) f''(x)}{f'^2(x)} = \frac{h_0(x) h_2(x)}{h_1^2(x)}. \quad (2.45)$$

Mit (2.3), (2.5) und (2.7) ist für $j \geq 2$

$$\varphi'(\xi) = \frac{h_0(\xi) h_2(\xi)}{h_1^2(\xi)} = \frac{((j-1)!)^2}{j! (j-2)!} = \frac{j-1}{j} = 1 - \frac{1}{j} \geq \frac{1}{2}. \quad (2.46)$$

Das Newtonsche Verfahren, das für einfache Nullstellen mindestens quadratisch konvergiert (Satz 2.22), konvergiert nach Satz 2.20 in der Umgebung einer mehrfachen Nullstelle nur linear.

Mit der Schrittfunktion

$$\psi(x) = x - j \frac{f(x)}{f'(x)} \quad (2.47)$$

kann die quadratische Konvergenz auch in der Umgebung einer mehrfachen Nullstelle ξ ($j \geq 2$) beibehalten werden. Mit (2.2) und (2.4) lautet (2.47)

$$\psi(x) = x - j(x - \xi) \frac{h_0(x)}{h_1(x)},$$

und es ist

$$\psi(\xi) = \xi.$$

Für die Ableitung der Schrittfunktion (2.47) ergibt sich mit (2.45)

$$\psi'(x) = 1 - j \left(1 - \frac{f(x) f''(x)}{f'^2(x)} \right) = 1 - j \left(1 - \frac{h_0(x) h_2(x)}{h_1^2(x)} \right).$$

Mit (2.46) ist

$$\psi'(\xi) = 1 - j \left(1 - \left(1 - \frac{1}{j} \right) \right) = 1 - j \frac{1}{j} = 0,$$

und daraus folgt nach Satz 2.20 die mindestens quadratische Konvergenz mit der Schrittfunktion (2.47).

Wegen $\psi'(\xi) = 0$ und wegen der Stetigkeit von ψ' in einer Umgebung von ξ gilt: Es gibt ein Intervall $I = (\xi - r, \xi + r) \subset [a, b]$, $r > 0$, so dass für alle $x \in I$ $|\psi'(x)| \leq L < 1$ ist. Somit ist ψ in I lipschitzbeschränkt. Ferner ist wegen $\psi(\xi) = \xi$ für alle $x \in I$

$$|\psi(x) - \xi| = |\psi(x) - \psi(\xi)| \leq L|x - \xi| \leq |x - \xi| < r,$$

also $\psi(x) \in I$. Somit genügt die Schrittfunktion (2.47) den Voraussetzungen (i) und (ii) des Fixpunktsatzes 2.14. Damit folgt der

Satz 2.25. (*Newtonsches Verfahren für mehrfache Nullstellen*)

Die Funktion f sei im Intervall $[a, b]$ genügend oft stetig differenzierbar und besitze in (a, b) eine Nullstelle ξ der Vielfachheit $j \geq 2$. Dann gibt es ein Intervall $I = (\xi - r, \xi + r)$, $r > 0$, $I \subset [a, b]$, so dass die Iterationsfolge

$$x^{(\nu+1)} = x^{(\nu)} - j \frac{f(x^{(\nu)})}{f'(x^{(\nu)})}, \quad \nu = 0, 1, 2, \dots,$$

mit jedem Startwert $x^{(0)} \in I$ von mindestens zweiter Ordnung gegen die j -fache Nullstelle ξ konvergiert.

Die Anwendung des Newtonschen Verfahrens für mehrfache Nullstellen setzt die Kenntnis der Vielfachheit j der gesuchten Nullstelle ξ voraus; j wird allerdings nur in speziellen Fällen bekannt sein.

Nach Satz 2.3 ist eine j -fache Nullstelle ξ ($j \geq 2$) einer Funktion f eine einfache Nullstelle der Funktion g mit $g(x) = f(x)/f'(x)$. Insofern kann man die Nullstelle von g mit dem Newtonschen Verfahren für einfache Nullstellen ermitteln. Dieses Verfahren beschreibt der Satz

Satz 2.26. (*Modifiziertes Newtonsches Verfahren für mehrfache Nullstellen*)

Die Funktion f sei im Intervall $[a, b]$ genügend oft stetig differenzierbar und besitze in (a, b) eine Nullstelle ξ der Vielfachheit $j \geq 2$. Dann gibt es ein Intervall $I = (\xi - r, \xi + r)$, $r > 0$, $I \subset [a, b]$, so dass mit jedem Startwert $x^{(0)} \in I$ das Verfahren mit der Iterationsvorschrift

$$\begin{aligned} x^{(\nu+1)} &= x^{(\nu)} - J(x^{(\nu)}) \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} \quad \text{mit} \\ J(x^{(\nu)}) &= \frac{1}{1 - \frac{f(x^{(\nu)})f''(x^{(\nu)})}{f'^2(x^{(\nu)})}}, \quad \nu = 0, 1, 2, \dots, \end{aligned}$$

quadratisch konvergiert. Es gelten zugleich

$$\lim_{\nu \rightarrow \infty} x^{(\nu)} = \xi \quad \text{und} \quad \lim_{\nu \rightarrow \infty} J(x^{(\nu)}) = j.$$

Beweis. Die Schrittfunktion des Newtonschen Verfahrens, angewendet auf die Funktion g mit

$$\begin{aligned} g(x) &= \frac{f(x)}{f'(x)}, \\ g'(x) &= 1 - \frac{f(x)f''(x)}{f'^2(x)} \quad \text{und} \\ J(x) &= \frac{1}{g'(x)} \end{aligned}$$

lautet

$$\varphi(x) = x - \frac{g(x)}{g'(x)} = x - J(x) \frac{f(x)}{f'(x)}.$$

Mit den Taylorentwicklungen (2.2), (2.4) und (2.6) an der Stelle ξ sind

$$\begin{aligned} \varphi(x) &= x - J(x)(x - \xi) \frac{h_0(x)}{h_1(x)} \quad \text{und} \\ J(x) &= \frac{1}{1 - \frac{h_0(x)h_2(x)}{h_1^2(x)}}. \end{aligned} \tag{2.48}$$

(1) Mit (2.46) erhält man

$$J(\xi) = \frac{1}{1 - (1 - \frac{1}{j})} = j. \tag{2.49}$$

- (2) Zur Bestimmung der Konvergenzordnung wird $\varphi'(\xi)$ berechnet. Dabei werden (2.49) und (2.8) verwendet. Mit (2.48) und $\varphi(\xi) = \xi$ ist

$$\begin{aligned}\varphi'(\xi) &= \lim_{x \rightarrow \xi} \frac{\varphi(x) - \varphi(\xi)}{x - \xi} \\ &= \lim_{x \rightarrow \xi} \frac{1}{x - \xi} \left(x - J(x) (x - \xi) \frac{h_0(x)}{h_1(x)} - \xi \right) \\ &= \lim_{x \rightarrow \xi} \left(1 - J(x) \frac{h_0(x)}{h_1(x)} \right) = 1 - j \frac{1}{j} = 0.\end{aligned}$$

Nach Satz 2.20 hat die Schrittfunktion φ die Konvergenzordnung 2.

- (3) Ähnlich wie beim Satz 2.25 für (2.47) kann gezeigt werden, dass die Schrittfunktion φ in einem Intervall $I = (\xi - r, \xi + r)$, $r > 0$, den Voraussetzungen des Fixpunktsatzes 2.14 genügt. \square

Beispiel 2.27.

Gegeben: Die Funktion $f : f(x) = 1 - \sin x$. f besitzt bei $\xi = \frac{\pi}{2}$ eine doppelte Nullstelle.

Gesucht: Die Nullstelle ξ mit dem Newtonschen Verfahren für einfache Nullstellen, mit dem Newtonschen Verfahren für mehrfache Nullstellen und mit dem modifizierten Newtonschen Verfahren. Es wird mit 15-stelliger Mantisse gerechnet und die Rechnung abgebrochen, wenn die Bedingung $|x^{(\nu)} - x^{(\nu-1)}| \leq 0.5e-14$ erfüllt ist. Als Startwert wird $x^{(0)} = 2$ gewählt.

1. Newtonsches Verfahren für einfache Nullstellen mit der Iterationsvorschrift (2.29):

$$x^{(\nu+1)} = x^{(\nu)} - \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} = x^{(\nu)} - \frac{1 - \sin(x^{(\nu)})}{-\cos(x^{(\nu)})}$$

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ |
|-------|-------------------|-----------------------------|
| 0 | 2.000000000000000 | |
| 1 | 1.78204190153914 | 0.21795809846086 |
| 2 | 1.67602457140144 | 0.10601733013770 |
| 3 | 1.62336184567011 | 0.05266272573132 |
| 4 | 1.59707303266146 | 0.02628881300865 |
| 5 | 1.58393392371128 | 0.01313910895018 |
| 6 | 1.57736503077218 | 0.00656889293910 |
| 7 | 1.57408066697409 | 0.00328436379809 |
| 8 | 1.57243849540833 | 0.00164217156576 |
| 9 | 1.57161741091709 | 0.00082108449124 |
| 10 | 1.57120686883293 | 0.00041054208416 |
| 11 | 1.57100159781103 | 0.00020527102190 |
| 12 | 1.57089896230260 | 0.00010263550843 |
| 13 | 1.57084764454871 | 0.00005131775390 |

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ |
|-------|------------------|-----------------------------|
| 14 | 1.57082198567180 | 0.00002565887691 |
| 15 | 1.57080915623335 | 0.00001282943845 |
| 16 | 1.57080274151412 | 0.00000641471923 |
| 17 | 1.57079953415451 | 0.00000320735961 |
| 18 | 1.57079793047470 | 0.00000160367980 |
| 19 | 1.57079712863480 | 0.00000080183990 |
| 20 | 1.57079672771482 | 0.00000040091998 |
| 21 | 1.57079652725488 | 0.00000020045994 |
| 22 | 1.57079642702485 | 0.00000010023003 |
| 23 | 1.57079637691009 | 0.00000005011476 |
| 24 | 1.57079635185228 | 0.00000002505781 |
| 25 | 1.57079633932381 | 0.00000001252847 |
| 26 | 1.57079633305860 | 0.00000000626521 |
| 27 | 1.57079632992561 | 0.00000000313298 |
| 28 | 1.57079632836721 | 0.00000000155840 |
| 29 | 1.57079632757422 | 0.00000000079299 |
| 30 | 1.57079632715686 | 0.00000000041736 |
| 31 | 1.57079632700710 | 0.00000000014977 |
| 32 | 1.57079632700710 | 0.00000000000000 |

Die exakte Lösung $\xi = \frac{\pi}{2}$ lautet auf 14 Dezimalen gerundet $\frac{\pi}{2} = 1.57079632679489$. Vergleicht man sie mit dem Näherungswert $x^{(32)}$, so gilt für den absoluten Fehler $|x^{(32)} - \xi| = |x^{(32)} - \frac{\pi}{2}| \leq 0.21\text{e-}9$; obwohl also $|x^{(32)} - x^{(31)}| \leq 0.5\text{e-}14$ ist, erhält man einen Näherungswert $x^{(32)}$, der nur auf 9 Dezimalen genau ist. Die Ursache dafür liegt darin, dass der Ausdruck f/f' in der Iterationsvorschrift für $x^{(\nu)} \rightarrow \xi$ unbestimmt ($\frac{0}{0}$) wird. Wird der Zähler in der mitgeführten Stellenzahl früher Null als der Nenner, so bleibt die Iteration dort stehen; dieser Fall liegt hier vor; andernfalls entfernen sich die iterierten Werte wieder von der Lösung. Wenn der Nenner Null wird, muss das Verfahren abgebrochen werden.

2. Newtonsches Verfahren für doppelte Nullstellen (Satz 2.25 mit Vielfachheit $j = 2$):

$$x^{(\nu+1)} = x^{(\nu)} - 2 \frac{f(x^{(\nu)})}{f'(x^{(\nu)})} = x^{(\nu)} - 2 \frac{1 - \sin(x^{(\nu)})}{-\cos(x^{(\nu)})}$$

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ |
|-------|------------------|-----------------------------|
| 0 | 2.00000000000000 | |
| 1 | 1.56408380307828 | 0.43591619692172 |
| 2 | 1.57079635199940 | 0.00671254892113 |
| 3 | 1.57079632679621 | 0.00000002520319 |
| 4 | 1.57079632679621 | 0.00000000000000 |

Hier ist die Abfrage $|x^{(\nu)} - x^{(\nu-1)}| \leq 0.5\text{e-}14$ bereits nach vier Iterationsschritten erfüllt, da das Verfahren quadratisch konvergiert. Aber auch hier gilt für den absoluten Fehler des Näherungswertes $x^{(4)}$ $|x^{(4)} - \frac{\pi}{2}| \leq 0.13\text{e-}11$; die Ursache dafür ist dieselbe wie im ersten Falle.

3. Modifiziertes Newtonsches Verfahren für mehrfache Nullstellen (Satz 2.26): Rechnung mit doppelter Stellenzahl, Rundung auf einfache.

| ν | $x^{(\nu)}$ | $ x^{(\nu)} - x^{(\nu-1)} $ | $J(x^{(\nu)})$ |
|-------|------------------|-----------------------------|-----------------|
| 0 | 2.00000000000000 | | 1.9092974268257 |
| 1 | 1.5838531634529 | 0.4161468365471 | 1.9999147607192 |
| 2 | 1.5707966977821 | 0.0130564656707 | 2.0000003799090 |
| 3 | 1.5707963267948 | 0.0000003709874 | 1.0000000000000 |
| 4 | 1.5707963267948 | 0.0000000000000 | |

Für den absoluten Fehler des Näherungswertes $x^{(4)}$ gilt $|x^{(4)} - \frac{\pi}{2}| \leq 0.25\text{e-}12$, obwohl für die iterierten Werte die Abfrage $|x^{(4)} - x^{(3)}| \leq 0.5\text{e-}14$ erfüllt ist. Die Ursache liegt wie oben darin, dass $\sin(x^{(3)})$ in den mitgeführten Stellen exakt 1 ist, d. h. $1 - \sin(x^{(3)})$ verschwindet, aber $\cos(x^{(3)}) \neq 0$ ist, so dass $x^{(4)} = x^{(3)}$ gilt. Man sieht hier außerdem, dass $J(x^{(2)})$ der tatsächlichen Vielfachheit $j = 2$ schon sehr nahe ist, $J(x^{(3)})$ jedoch gleich 1 wird. Dies liegt daran, dass $f(x^{(3)}) = 1 - \sin(x^{(3)})$ in den mitgeführten Stellen verschwindet, $f'(x^{(3)}) \neq 0$ ist und damit $f f''/f'^2$ verschwindet. Würde hier f'^2 vor f Null, so würde $J(x^{(\nu)})$ von j in beliebiger Weise abwandern.

Wegen $f(x) = 1 - \sin x$ kann in diesem Fall $J(x^{(\nu)})$ vereinfacht dargestellt werden (Satz 2.26):

$$J(x^{(\nu)}) = \frac{1}{1 - \frac{(1 - \sin(x^{(\nu)})) \sin(x^{(\nu)})}{\cos^2(x^{(\nu)})}} = 1 + \sin(x^{(\nu)}).$$

Damit ergibt sich $J(x^{(3)}) = 1 + 1 = 2$. \square

Empfehlung zum modifizierten Newton-Verfahren

Wegen seiner geringen Effizienz (vgl. Abschnitt 2.9) sollte man mit dem modifizierten Newton-Verfahren nur so lange iterieren, bis die Vielfachheit der Nullstelle geklärt ist, dann aber mit dem Verfahren von Newton für mehrfache Nullstellen weiterrechnen. Die Vielfachheit ist dann geklärt, wenn sich von einem gewissen ν an entweder $J(x^{(\nu)}) = 1$ ergibt oder

$$|J(x^{(\nu)}) - J(x^{(\nu-1)})| > |J(x^{(\nu-1)}) - J(x^{(\nu-2)})|$$

gilt, d. h. $J(x^{(\nu)})$ sich wieder von der Vielfachheit j entfernt. In beiden Fällen ist die zu $J(x^{(\nu-1)})$ nächste ganze Zahl die gesuchte Vielfachheit j . Dieses Verhalten der $J(x^{(\nu)})$ ist bedingt durch die beschränkte Stellenzahl der Maschinenzahlen und den für $x^{(\nu)} \rightarrow \xi$ unbestimmten Ausdruck $f f''/f'^2$ im Nenner von $J(x^{(\nu)})$. Wird nämlich für ein $x^{(\nu)}$ wegen der beschränkten Stellenzahl $f(x^{(\nu)})$ identisch Null, während $f'(x^{(\nu)})$ noch verschieden von Null ist, so erhält man $J(x^{(\nu)}) = 1$, obwohl eine mehrfache Nullstelle vorliegt.

Generelle Empfehlung. Da eine doppelte Nullstelle ξ von f nach Satz 2.3 eine einfache Nullstelle ξ von $g(x) = f(x)/f'(x)$ ist, kann auch zur Berechnung einer doppelten Nullstelle ein Einschlussverfahren verwendet werden. Für Beispiel 2.27 würde das bedeuten, dass man anstelle der doppelten Nullstelle von $f(x) = 1 - \sin x$ die einfache Nullstelle der Funktion

$$g(x) = \frac{f(x)}{f'(x)} = \frac{1 - \sin x}{-\cos x}$$

mit einem Einschlussverfahren ermittelt, das mit größerer Effizienz arbeitet als das Newton-Verfahren (siehe Abschnitt 2.9). Die oben erwähnten numerischen Probleme wegen $g(\xi) = \frac{0}{0}$ treten natürlich auch hier auf.

Anmerkungen zu mehrfachen Nullstellen. Bei einer nichtlinearen Gleichung $f(x) = 0$, die im Rahmen eines technischen Problems gelöst werden muss, werden die im Funktionsterm auftretenden Daten im Allgemeinen mit Fehlern behaftet sein. Darum ist nicht zu erwarten, dass die Funktion f eine mehrfache Nullstelle hat.

Beispielsweise hat die quadratische Funktion $f(x) = x^2 + px + q$ genau dann eine doppelte Nullstelle, wenn exakt $p^2 - 4q = 0$ ist. Wenn p und q mit Messfehlern behaftet sind, wird $|p^2 - 4q| = \varepsilon > 0$ sein. Für $p^2 - 4q > 0$ gibt es dann zwei benachbarte einfache Nullstellen und für $p^2 - 4q < 0$ keine.

Eine doppelte Nullstelle ξ von f mit $f(\xi) = f'(\xi) = 0$, $f''(\xi) \neq 0$ ist eine Stelle, an der die Funktion f ein lokales Minimum oder Maximum besitzt. Wegen $f''(\xi) \neq 0$ ist ξ einfache Nullstelle von f' und kann mit einem Einschlussverfahren berechnet werden. Wenn dann $|f(\xi)| \leq \varepsilon$ ist, kann ξ als doppelte Nullstelle von f akzeptiert werden.

2.6 Das Sekantenverfahren

Das Sekantenverfahren wird hier behandelt, weil es in fast jeder Numerik-Vorlesung als Standardverfahren vorkommt. Für die praktische Anwendung ist es allerdings nicht zu empfehlen, man sollte dem Sekantenverfahren die Einschlussverfahren mit höherer Effizienz unbedingt vorziehen.

2.6.1 Das Sekantenverfahren für einfache Nullstellen

Die Funktion f sei in $I = [a, b]$ stetig und besitze in (a, b) eine einfache Nullstelle ξ . Zur näherungsweisen Bestimmung von ξ mit Hilfe des Verfahrens von Newton ist die Berechnung der Ableitung f' von f erforderlich, so dass die Differenzierbarkeit von f vorausgesetzt werden muss. Das Sekantenverfahren ist ein Iterationsverfahren, das ohne Ableitungen arbeitet und zwei Startwerte $x^{(0)}$, $x^{(1)}$ erfordert (siehe auch Abschnitt 2.7 Einschlussverfahren).

Durch die folgende Überlegung gelangt man zu der Iterationsvorschrift für das Verfahren (Abb. 2.13). Es seien $x^{(0)}, x^{(1)} \in I$. Durch die Punkte $(x^{(0)}, f(x^{(0)}))$ und $(x^{(1)}, f(x^{(1)}))$ legt man die Sekante des Funktionsgraphen und schneidet sie mit der x -Achse. Für die Abszisse $x^{(2)}$ des Schnittpunktes findet man

$$x^{(2)} = x^{(1)} - \frac{x^{(1)} - x^{(0)}}{f(x^{(1)}) - f(x^{(0)})} f(x^{(1)}).$$

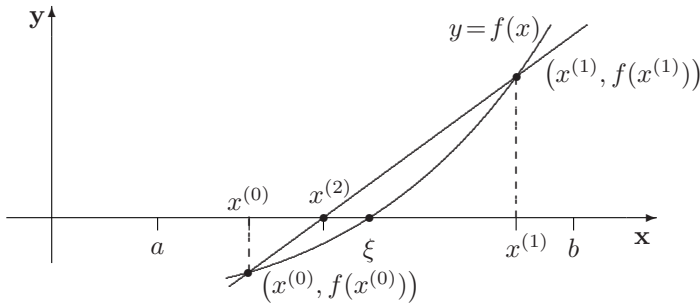


Abb. 2.13. Sekantenverfahren

Die Iterationsvorschrift für das Sekantenverfahren lautet

$$\begin{cases} x^{(\nu+1)} = x^{(\nu)} - \frac{x^{(\nu)} - x^{(\nu-1)}}{f(x^{(\nu)}) - f(x^{(\nu-1)})} f(x^{(\nu)}) & \text{wenn } f(x^{(\nu)}) \neq f(x^{(\nu-1)}), \\ x^{(\nu+1)} = x^{(\nu)} - \frac{x^{(\nu)} - x^{(\nu-1)}}{f(x^{(\nu)}) - 2 \cdot f(x^{(\nu-1)})} f(x^{(\nu)}) & \text{sonst,} \\ \nu = 1, 2, \dots \end{cases} \quad (2.50)$$

Falls $f(x^{(\nu+1)}) = 0$ ist, wird das Verfahren mit $\xi = x^{(\nu+1)}$ abgebrochen. Wesentlich für die Konvergenz des Verfahrens ist, dass die Startwerte $x^{(0)}, x^{(1)}$ hinreichend nahe an der Nullstelle ξ liegen. Es gilt der folgende *Konvergenzsatz*:

Satz 2.28.

Falls die Funktion f für alle $x \in (a, b)$ zweimal stetig differenzierbar ist und mit zwei positiven Zahlen m, M den Bedingungen

$$|f'(x)| \geq m, \quad |f''(x)| \leq M, \quad x \in (a, b),$$

genügt, gibt es immer eine Umgebung $I_r = [\xi - r, \xi + r] \subset (a, b)$, $r > 0$, so dass ξ in I_r die einzige Nullstelle von f ist und das Verfahren für jedes Paar von Startwerten $x^{(0)}, x^{(1)} \in I_r$, $x^{(0)} \neq x^{(1)}$, gegen die gesuchte Nullstelle ξ konvergiert.

Da die Überprüfung der Voraussetzungen dieses Satzes meist nicht praktikabel ist, bleibt nur die Empfehlung, die Startwerte $x^{(0)}$ und $x^{(1)}$ möglichst nahe bei der Nullstelle ξ zu wählen.

Wenn das Startintervall den Bedingungen des Satzes 2.28 nicht genügt, kann beim Sekantenverfahren die Konvergenz von der Bezeichnung der Startwerte $x^{(0)}$ und $x^{(1)}$ abhängen.

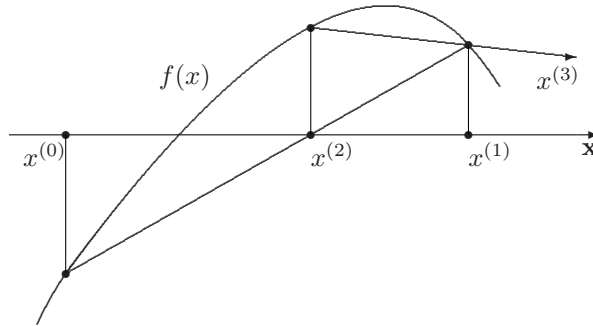


Abb. 2.14. Startwerte, für die das Sekantenverfahren divergiert

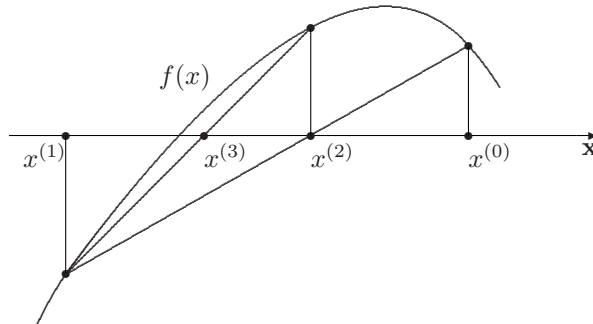


Abb. 2.15. Startwerte, für die das Sekantenverfahren konvergiert

Beispiel 2.29.

Gegeben: Die Funktion $f(x) = \frac{1}{8}x^2 - x + \frac{3}{2}$ und die Startwerte $x^{(0)} = 1$, $x^{(1)} = 5$, die wegen $f(1) = 0.625$, $f(5) = -0.375$ eine Nullstelle von f einschließen.

Gesucht: Die zwischen $x^{(0)}$ und $x^{(1)}$ liegende Nullstelle mit dem Sekantenverfahren.

Lösung: Mit den gegebenen Startwerten konvergiert das Sekantenverfahren gegen die Nullstelle 6, die nicht im Einschlussintervall $[1, 5]$ liegt. Mit den Startwerten $x^{(0)} = 5$, $x^{(1)} = 1$ liefert das Sekantenverfahren die Nullstelle 2.

Ein Einschlussverfahren (siehe Abschnitt 2.7) ermittelt in beiden Fällen die Nullstelle 2. \square

Beispiel 2.30.

Gegeben: Die Funktion $f(x) = \ln x - \sqrt{x} + 1.5$, die für $x > 0$ reell und stetig ist, und die Startwerte $x^{(0)} = 0.2$, $x^{(1)} = 2$, die wegen $f(0.2) \approx -0.557$, $f(2) \approx 0.779$ eine Nullstelle von f einschließen.

Gesucht: Die im Intervall $[0.2, 2]$ liegende Nullstelle mit dem Sekantenverfahren.

Lösung: Mit den gegebenen Startwerten erzeugt das Sekantenverfahren

$$\begin{aligned}x^{(2)} &= 0.950213 \in [0.2, 2] \quad \text{und} \\x^{(3)} &= -0.682864 \notin [0.2, 2].\end{aligned}$$

Also ist $f(x^{(3)})$ nicht reell, und das Sekantenverfahren versagt.

Mit den Startwerten $x^{(0)} = 2$, $x^{(1)} = 0.2$ konvergiert das Sekantenverfahren gegen die Nullstelle $\xi = 0.429\,817\,028$.

Ein Einschlussverfahren (siehe Abschnitt 2.7) erzeugt dagegen Näherungen $x^{(\nu)}$, die stets im Einschlussintervall $[x^{(0)}, x^{(1)}]$ liegen. \square

Die Konvergenzordnung des Sekantenverfahrens mit der Iterationsvorschrift (2.50) ist $p = (1 + \sqrt{5})/2 \approx 1.62$. Prinzipiell kann die Vorschrift (2.50) auch zur näherungsweisen Berechnung mehrfacher Nullstellen verwendet werden, dann geht jedoch die hohe Konvergenzordnung verloren. Das *modifizierte Sekantenverfahren* (Abschnitt 2.6.2) besitzt auch bei mehrfachen Nullstellen die Konvergenzordnung $p \approx 1.62$. Zur Effizienz der Verfahren siehe Abschnitt 2.9.

2.6.2 Das modifizierte Sekantenverfahren für mehrfache Nullstellen

Ist ξ eine Nullstelle der Vielfachheit j , $j \geq 2$, von f und ist $|f^{(j+1)}(x)|$ in der Umgebung von ξ beschränkt, so ist ξ eine einfache Nullstelle der Funktion h

$$h(x) = \frac{f^2(x)}{f(x + f(x)) - f(x)}, \quad (2.51)$$

und $|h''(x)|$ ist in der Umgebung von ξ beschränkt. Verwendet man in der Iterationsvorschrift (2.50) statt f die durch (2.51) definierte Funktion h , so konvergiert dieses modifizierte Sekantenverfahren ebenfalls von der Ordnung $p = (1 + \sqrt{5})/2$ gegen die mehrfache Nullstelle ξ von f (Beweis s. [KIOU1979]). Die Effizienz ist wegen $H = 2$ allerdings nur $E = p^{1/2} = 1.272$ (siehe Abschnitt 2.9).

2.7 Einschlussverfahren

Eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ mit $f(a) \cdot f(b) < 0$ besitzt nach dem Zwischenwertsatz im Intervall (a, b) mindestens eine Nullstelle; ein solches Intervall $[a, b]$ nennt man ein Einschlussintervall. Im Folgenden sei das Intervall $[a, b]$ so gewählt, dass es genau eine Nullstelle ξ der Funktion f einschließt.

Bei den in diesem Abschnitt behandelten Einschlussverfahren wird ein gegebenes Einschlussintervall in zwei Teilintervalle zerlegt; eines von diesen ist wieder ein Einschlussintervall. So können fortlaufend kleinere Einschlussintervalle erzeugt werden (Abschnitt 2.7.1).

Das einfachste dieser Verfahren ist das Bisektionsverfahren (Abschnitt 2.7.2), das stets, aber nur linear, gegen die Nullstelle konvergiert.

Bei der Regula falsi (Abschnitt 2.7.3), die ebenfalls nur linear konvergiert, kann es vorkommen, dass die erzeugten Grenzen der Einschlussintervalle sich der Nullstelle nur von einer Seite her nähern (Abbildung 2.17).

Deshalb wurden einige Einschlussverfahren (Abschnitte 2.7.4 bis 2.7.6) entwickelt, die dieses Verhalten vermeiden und zudem eine höhere Konvergenzordnung besitzen. Allerdings wird diese höhere Konvergenzordnung erst in einer hinreichend kleinen Umgebung der Nullstelle wirksam.

Daher ist es zweckmäßig, ein gegebenes Einschlussintervall zunächst mit dem Bisektionsverfahren zu verkleinern und dann ein Verfahren höherer Ordnung einzusetzen (siehe dazu das Beispiel 2.41).

Falls eine genügend oft stetig differenzierbare Funktion f im Intervall $[a, b]$ eine Nullstelle ξ gerader Ordnung besitzt, ist diese wegen Satz 2.3 eine einfache Nullstelle der Funktion $g(x) = f(x)/f'(x)$ und kann daher mit einem Einschlussverfahren, angewandt auf g , ermittelt werden. Siehe dazu auch die Anmerkungen zu mehrfachen Nullstellen am Ende des Abschnitts 2.5.

2.7.1 Das Prinzip der Einschlussverfahren

Ausgehend von einem Einschlussintervall $[a, b]$ werden $x^{(1)} = a$, $f_1 = f(a)$, $x^{(2)} = b$, $f_2 = f(b)$ gesetzt; damit ist $f_1 \cdot f_2 < 0$. Dann wird eine Zahl

$$x^{(3)} = x^{(2)} + q(x^{(1)} - x^{(2)}), \quad 0 < q < 1, \quad (2.52)$$

erzeugt, die zwischen $x^{(1)}$ und $x^{(2)}$ liegt. Die Verfahren unterscheiden sich in der Wahl von q . Mit $x^{(3)}$ wird der Funktionswert $f_3 = f(x^{(3)})$ berechnet. Wenn $f_3 = 0$ ist, ist $\xi = x^{(3)}$ Nullstelle von f .

Wenn $f_3 \neq 0$ ist, hat f entweder zwischen $x^{(2)}$ und $x^{(3)}$ oder zwischen $x^{(1)}$ und $x^{(3)}$ einen Vorzeichenwechsel. Eines der beiden durch $x^{(3)}$ erzeugten Teilintervalle ist also wieder ein Einschlussintervall, mit dem das Verfahren fortgesetzt werden kann. Das neue Einschlussintervall wird wie folgt ermittelt.

Wenn $f_2 \cdot f_3 < 0$ ist, liegt ξ zwischen $x^{(2)}$ und $x^{(3)}$, und die Intervallgrenzen und Funktionswerte werden umbenannt:

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

Wenn $f_2 \cdot f_3 > 0$ ist und somit $f_1 \cdot f_3 < 0$ ist, liegt ξ zwischen $x^{(1)}$ und $x^{(3)}$. Dann werden gesetzt:

$$x^{(2)} := x^{(3)}, \quad f_2 := f_3.$$

Nun gilt wieder $f_1 \cdot f_2 < 0$, und mit $x^{(1)}$ und $x^{(2)}$ kann das Verfahren fortgesetzt werden. Dabei ist $x^{(2)}$ die zuletzt berechnete Intervallgrenze.

Da ξ zwischen $x^{(1)}$ und $x^{(2)}$ liegt, gelten für die absoluten Fehler dieser Intervallgrenzen

$$|\xi - x^{(1)}| < |x^{(2)} - x^{(1)}|, \quad |\xi - x^{(2)}| < |x^{(2)} - x^{(1)}|.$$

Wenn mit einer positiven Schranke AbsErr für den absoluten Fehler für die Länge des Einschlussintervalls

$$|x^{(2)} - x^{(1)}| \leq \text{AbsErr} \quad (2.53)$$

gilt, sind die absoluten Fehler der beiden Intervallgrenzen kleiner als AbsErr.

Wenn mit einer positiven Schranke RelErr für den relativen Fehler mit der zuletzt berechneten Intervallgrenze $x^{(2)}$ ($\neq 0$)

$$|x^{(2)} - x^{(1)}| \leq |x^{(2)}| \text{RelErr} \quad (2.54)$$

ist, gilt für den relativen Fehler dieser Intervallgrenze

$$\frac{|\xi - x^{(2)}|}{|x^{(2)}|} < \frac{|x^{(2)} - x^{(1)}|}{|x^{(2)}|} \leq \text{RelErr}.$$

Ein Einschlussverfahren kann also außer mit $f_3 = 0$ auch abgebrochen werden, wenn eine der Abfragen (2.53), (2.54) erfüllt ist. Beide können kombiniert werden zu

$$|x^{(2)} - x^{(1)}| \leq |x^{(2)}| \text{RelErr} + \text{AbsErr}.$$

Mit AbsErr > 0 und RelErr $= 0$ entsteht (2.53), und mit AbsErr $= 0$ und RelErr > 0 ergibt sich (2.54).

Von den beiden Intervallgrenzen $x^{(1)}$, $x^{(2)}$ kann man diejenige als die beste Näherung für die Nullstelle ξ wählen, für die der Funktionswert dem Betrage nach kleiner ist.

Die Fehlerschranken AbsErr und RelErr müssen größer als die Maschinengenauigkeit ϱ sein (etwa 2ϱ bis 3ϱ).

Bemerkung. Für die praktische Anwendung ist es sinnvoll, die Abbruchbedingung (2.54) für den relativen Fehler zu verwenden, weil so eine Aussage über die Anzahl der gültigen Ziffern einer Näherungszahl gemacht werden kann (vgl. Satz 1.24). Der absolute Fehler macht eine Aussage über die Anzahl der gültigen Dezimalen (vgl. Definition 1.17).

Das heißt, die Abfrage

$$|x^{(2)} - \xi| \leq \text{AbsErr} = 0.5 \cdot 10^{-k}$$

liefert $x^{(2)}$ mit k genauen Dezimalen, die Abfrage

$$\frac{|x^{(2)} - \xi|}{|x^{(2)}|} \leq \text{RelErr} = 5 \cdot 10^{-m}$$

liefert $x^{(2)}$ mit m genauen Ziffern, beginnend mit der ersten von 0 verschiedenen Ziffer von $x^{(2)}$.

2.7.2 Das Bisektionsverfahren

Bei diesem einfachsten Einschlussverfahren wird in (2.52) $q = 0.5$ gesetzt. Dann wird wegen

$$|x^{(3)} - x^{(2)}| = 0.5 |x^{(1)} - x^{(2)}|$$

die Länge des Einschlussintervalls halbiert. Mit fortgesetzter Intervallhalbierung konvergiert das Bisektionsverfahren linear gegen die Nullstelle.

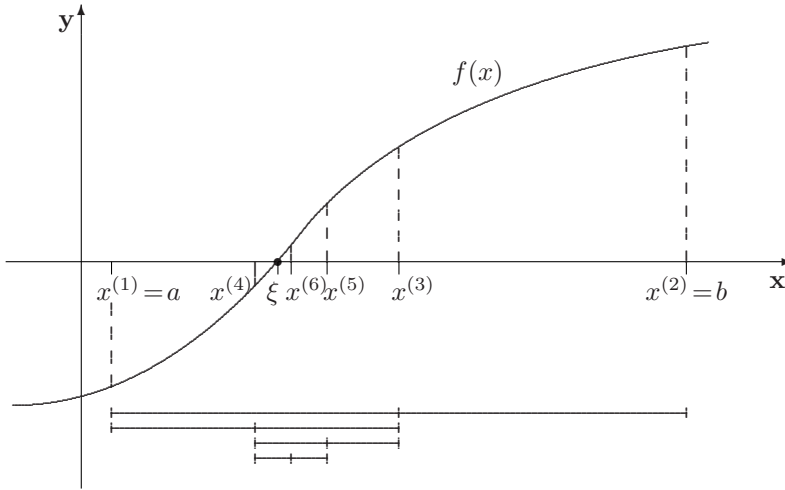


Abb. 2.16. Das Bisektionsverfahren

Algorithmus 2.31. (Bisektionsverfahren)

Gegeben: (i) $f \in C[a, b]$ mit $f(a) \cdot f(b) < 0$,
(ii) Schranken AbsErr und RelErr für den absoluten bzw. den relativen Fehler mit AbsErr > 0 und RelErr = 0
oder mit AbsErr = 0 und RelErr > 0 .

Gesucht: Eine Zahl $\xi \in (a, b)$, für die $f(\xi) = 0$ ist, oder ein Einschlussintervall $[x^{(1)}, x^{(2)}]$ bzw. $[x^{(2)}, x^{(1)}]$ für ξ mit
 $|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{RelErr} + \text{AbsErr}$.

Vorbereitung: $x^{(1)} := a, \quad x^{(2)} := b$
 $f_1 := f(x^{(1)}), \quad f_2 := f(x^{(2)})$.

Pro Iterationsschritt wird wie folgt vorgegangen:

1. Halbierung des Einschlussintervalls durch Ermittlung von

$$x^{(3)} := x^{(2)} + 0.5(x^{(1)} - x^{(2)}).$$

2. *Berechnung des neuen Funktionswertes* $f_3 := f(x^{(3)})$.

Falls $f_3 = 0$ ist, wird die Iteration mit $\xi := x^{(3)}$ abgebrochen, andernfalls geht es mit 3. weiter.

3. *Festlegung des neuen Einschlussintervalls:*

Falls $f_2 \cdot f_3 < 0$ ist, liegt ξ zwischen $x^{(2)}$ und $x^{(3)}$, und es wird gesetzt

$$x^{(1)} := x^{(2)}, \quad x^{(2)} := x^{(3)}, \quad f_1 := f_2, \quad f_2 := f_3;$$

falls $f_2 \cdot f_3 > 0$ ist, liegt ξ zwischen $x^{(1)}$ und $x^{(3)}$, und es wird gesetzt

$$x^{(2)} := x^{(3)}, \quad f_2 := f_3.$$

In beiden Fällen liegt jetzt ξ zwischen $x^{(1)}$ und $x^{(2)}$, und $x^{(2)}$ ist der zuletzt berechnete Näherungswert.

4. *Prüfung der Abbruchbedingung:*

Falls

$$|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{ RelErr} + \text{AbsErr}$$

ist, erfolgt Abbruch. Dann wird gesetzt

$$\xi := x^{(2)}, \quad \text{falls } |f_2| \leq |f_1| \text{ ist,} \quad \text{und sonst} \quad \xi := x^{(1)}.$$

Andernfalls wird die Iteration mit 1. fortgesetzt.

Bemerkung 2.32. Beim Bisektionsverfahren kann die Anzahl n der Intervallhalbierungen, die erforderlich ist, um eine vorgegebene Anzahl k gültiger Dezimalen zu erhalten, vorab bestimmt werden. Für n und k ($k \in \mathbb{N}$) gelte also

$$\frac{|b-a|}{2^n} \leq \frac{1}{2} e^{-k}.$$

Dabei sei $0.5 \cdot 10^{-k} < |b-a|$. Dann folgen

$$\begin{aligned} 2^n &\geq 2 |b-a| 10^k, \\ n \lg 2 &\geq \lg 2 + \lg |b-a| + k, \\ n &\geq 1 + \frac{k + \lg |b-a|}{\lg 2}. \end{aligned} \tag{2.55}$$

Beispiel 2.33.

Gegeben: Die Funktion $f(x) = \sin x + 1 - 1/x$ für $x \in [a, b]$ mit $a = 0.6$, $b = 0.7$ und die Schranke $\text{AbsErr} = 0.5e-6$ für den absoluten Fehler.

Gesucht: Eine Nullstelle ξ der Funktion f im Intervall $(0.6, 0.7)$ mit Hilfe des Bisektionsverfahrens auf 6 Dezimalen genau.

Lösung: Wegen $f(0.6) \approx -0.102$ und $f(0.7) \approx 0.216$ liegt mindestens eine Nullstelle in $(0.6, 0.7)$. Mit $k = 6$, $a = 0.6$ und $b = 0.7$ erhält man nach der Formel (2.55)

die Anzahl der erforderlichen Intervallhalbierungen:

$$n \geq 1 + \frac{6 + \lg 0.1}{\lg 2} = 17.6096.$$

Nach 18 Intervallhalbierungen ist die Länge des letzten Einschlussintervalls $\leq 0.5e-6$.

| ν | $x^{(3)}$ | $f(x^{(3)})$ | $x^{(1)}$ | $f(x^{(1)})$ | $x^{(2)}$ | $f(x^{(2)})$ | $ x^{(2)} - x^{(1)} $ |
|----------|-----------|--------------|-----------|--------------|-----------|--------------|---------------------------|
| | | | 0.6 | $f < 0$ | 0.7 | $f > 0$ | 0.1 |
| 1 | 0.65 | +0.0667249 | 0.6 | $f < 0$ | 0.65 | $f > 0$ | 0.05 |
| 2 | 0.625 | -0.0149027 | 0.65 | $f > 0$ | 0.625 | $f < 0$ | 0.025 |
| 3 | 0.6375 | +0.0265609 | 0.625 | $f < 0$ | 0.6375 | $f > 0$ | 0.0125 |
| \vdots | | | | | | | |
| 16 | 0.6294479 | +0.0000048 | 0.6294464 | $f < 0$ | 0.6294479 | $f > 0$ | 0.0000015 |
| 17 | 0.6294472 | +0.0000023 | 0.6294464 | $f < 0$ | 0.6294472 | $f > 0$ | 0.0000008 |
| 18 | 0.6294468 | +0.0000010 | 0.6294464 | -2.43e-7 | 0.6294468 | 1.03e-6 | 0.0000004 < ε |

In dieser Tabelle steht jeder Schritt in einer Zeile. Die neue Näherung ist $x^{(3)}$. Nach Umspeichern ergeben sich rechts die Grenzen $x^{(1)}$ und $x^{(2)}$ des neuen Einschlussintervalls. In der letzten Spalte kann man die Abbruchbedingung prüfen.

Wegen $|f(x^{(1)})| < |f(x^{(2)})|$ für $\nu = 18$ wird gewählt

$$\xi = 0.629446 \approx x^{(1)}. \quad \square$$

2.7.3 Die Regula falsi

Dieses Verfahren verwendet die Sekante, die die Punkte $(x^{(1)}, f_1)$ und $(x^{(2)}, f_2)$ verbindet, und bestimmt deren Schnittpunkt $(x^{(3)}, 0)$ mit der x -Achse (Sekantenschritt). Mit der Gleichung

$$y = f_2 + \frac{f_1 - f_2}{x^{(1)} - x^{(2)}} (x - x^{(2)})$$

dieser Sekante ergibt sich für $y = 0$

$$x^{(3)} = x^{(2)} + \frac{f_2}{f_2 - f_1} (x^{(1)} - x^{(2)}).$$

Hier ist also mit (2.52)

$$q = \frac{f_2}{f_2 - f_1},$$

und wegen $f_1 \cdot f_2 < 0$ ist $0 < q < 1$.

Algorithmus 2.34. (*Regula falsi*)

Gegeben: (i) $f \in C[a, b]$ mit $f(a) \cdot f(b) < 0$,
(ii) Schranken AbsErr und RelErr für den absoluten bzw. den relativen Fehler mit AbsErr > 0 und RelErr = 0
oder mit AbsErr = 0 und RelErr > 0 .

Gesucht: Eine Zahl $\xi \in (a, b)$, für die $f(\xi) = 0$ ist, oder ein Einschlussintervall $[x^{(1)}, x^{(2)}]$ bzw. $[x^{(2)}, x^{(1)}]$ für ξ mit
 $|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{RelErr} + \text{AbsErr}$.

Vorbereitung: $x^{(1)} := a, \quad x^{(2)} := b$
 $f_1 := f(x^{(1)}), f_2 := f(x^{(2)})$.

Pro Iterationsschritt wird wie folgt vorgegangen:

1. *Bestimmung der neuen Intervallgrenze:*

Berechne

$$\text{tol} := |x^{(2)}| \text{RelErr} + \text{AbsErr}$$

und

$$\Delta x := \frac{f_2}{f_2 - f_1} (x^{(1)} - x^{(2)}).$$

Wenn $|\Delta x| \leq \text{tol}$ ist, setze

$$\Delta x := 0.9 \cdot \text{tol} \cdot \text{sgn}(x^{(1)} - x^{(2)}).$$

Berechne $x^{(3)} := x^{(2)} + \Delta x$.

2. *Berechnung des neuen Funktionswertes* $f_3 := f(x^{(3)})$.

Falls $f_3 = 0$ ist, wird die Iteration mit $\xi := x^{(3)}$ abgebrochen, andernfalls geht es mit 3. weiter.

3. *Festlegung des neuen Einschlussintervalls:*

Falls $f_2 \cdot f_3 < 0$ ist, liegt ξ zwischen $x^{(2)}$ und $x^{(3)}$, und es wird gesetzt

$$x^{(1)} := x^{(2)}, \quad x^{(2)} := x^{(3)}, \quad f_1 := f_2, \quad f_2 := f_3;$$

falls $f_2 \cdot f_3 > 0$ ist, liegt ξ zwischen $x^{(1)}$ und $x^{(3)}$, und es wird gesetzt

$$x^{(2)} := x^{(3)}, \quad f_2 := f_3.$$

In beiden Fällen liegt jetzt ξ zwischen $x^{(1)}$ und $x^{(2)}$, und $x^{(2)}$ ist der zuletzt berechnete Näherungswert.

4. *Prüfung der Abbruchbedingung:*

Falls

$$|x^{(1)} - x^{(2)}| \leq \text{tol}$$

ist, erfolgt Abbruch. Dann wird gesetzt

$$\xi := x^{(2)}, \quad \text{falls } |f_2| \leq |f_1| \text{ ist,} \quad \text{und sonst} \quad \xi := x^{(1)}.$$

Andernfalls, also mit $|x^{(2)} - x^{(1)}| > \text{tol}$, wird die Iteration mit 1. fortgesetzt.

Beispiel 2.35. (vgl. Beispiel 2.33)

Gegeben: Die Funktion $f(x) = \sin x + 1 - 1/x$ für $x \in [a, b]$ mit $a = 0.6$, $b = 0.7$ sowie die Schranke $\text{RelErr} = 5\text{e-}7$ für den relativen Fehler ($\text{AbsErr} = 0$).

Gesucht: Die Nullstelle ξ der Funktion f im Intervall $(0.6, 0.7)$ mit der Regula falsi mit 7 gültigen Ziffern.

Lösung:

| ν | $x^{(3)}$ | $f(x^{(3)})$ | $x^{(1)}$ | $f(x^{(1)})$ | $x^{(2)}$ | $f(x^{(2)})$ | $ x^{(2)} - x^{(1)} / x^{(2)} $ |
|-------|------------|--------------|------------|--------------|------------|--------------|---------------------------------|
| | | | 0.6 | < 0 | 0.7 | > 0 | |
| 1 | 0.63211636 | 9e-3 | 0.6 | < 0 | 0.63211636 | > 0 | 5e-2 |
| 2 | 0.62954848 | 3e-4 | 0.6 | < 0 | 0.62954848 | > 0 | 4.69e-2 |
| 3 | 0.62945038 | 1e-5 | 0.6 | < 0 | 0.62945038 | > 0 | 4.68e-2 |
| 4 | 0.62944663 | 5e-7 | 0.6 | < 0 | 0.62944663 | > 0 | 4.68e-2 |
| 5 | 0.62944635 | -4.5e-7 | 0.62944663 | 5e-7 | 0.62944635 | -4.5e-7 | 4.5e-7 |

$\Rightarrow \xi \approx x^{(2)} = \underbrace{0.6294464}_{7 \text{ genaue Ziffern}} \quad \square$

2.7.4 Das Pegasus-Verfahren

Das Pegasus-Verfahren ist ein gegenüber der Regula falsi verbessertes Verfahren, das in einer hinreichend kleinen Umgebung einer einfachen Nullstelle die Konvergenzordnung $p = 1.642$ besitzt. Falls nach einem Sekantenschritt die älteste Intervallgrenze $x^{(1)}$ nicht ersetzt werden kann, wird der Funktionswert f_1 mittels $g \cdot f_1$, $0 < g < 1$, modifiziert. Damit verbessert sich die Chance, die Grenze $x^{(1)}$ nach dem nächsten Schritt ersetzen zu können. Siehe dazu die geometrische Interpretation und Abbildung 2.18.

Basierend auf den Originalarbeiten [DOWE1971] und [DOWE1972] wurde hier der folgende Algorithmus entwickelt.

Algorithmus 2.36. (*Pegasus-Verfahren*)

Gegeben: (i) $f \in C[a, b]$ mit $f(a) \cdot f(b) < 0$,
(ii) Schranken AbsErr und RelErr für den absoluten bzw. den relativen Fehler mit AbsErr > 0 und RelErr $= 0$
oder mit AbsErr $= 0$ und RelErr > 0 .

Gesucht: Eine Zahl $\xi \in (a, b)$, für die $f(\xi) = 0$ ist, oder ein Einschlussintervall $[x^{(1)}, x^{(2)}]$ bzw. $[x^{(2)}, x^{(1)}]$ für ξ mit
 $|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{RelErr} + \text{AbsErr}$.

Vorbereitung: $x^{(1)} := a, \quad x^{(2)} := b$
 $f_1 := f(x^{(1)}), \quad f_2 := f(x^{(2)})$.

Pro Iterationsschritt ist wie folgt vorzugehen:

1. *Bestimmung der neuen Intervallgrenze:*

Berechne $\text{tol} := |x^{(2)}| \text{RelErr} + \text{AbsErr}$
und $\Delta x := \frac{f_2}{f_2 - f_1} (x^{(1)} - x^{(2)})$.

Wenn $|\Delta x| \leq \text{tol}$ ist, setze

$$\Delta x := 0.9 \cdot \text{tol} \cdot \text{sgn}(x^{(1)} - x^{(2)}).$$

Berechne $x^{(3)} := x^{(2)} + \Delta x$.

2. *Berechnung des neuen Funktionswertes $f_3 := f(x^{(3)})$.*

Falls $f_3 = 0$ ist, wird die Iteration mit $\xi := x^{(3)}$ abgebrochen, andernfalls geht es mit 3. weiter.

3. *Festlegung des neuen Einschlussintervalls:*

Falls $f_2 \cdot f_3 < 0$ ist, wenn also die Nullstelle ξ zwischen $x^{(2)}$ und $x^{(3)}$ liegt, wird gesetzt

$$x^{(1)} := x^{(2)}, \quad x^{(2)} := x^{(3)}, \quad f_1 := f_2, \quad f_2 := f_3.$$

Falls $f_2 \cdot f_3 > 0$ ist, wenn also die Nullstelle ξ zwischen $x^{(1)}$ und $x^{(3)}$ liegt, wird die Stelle $x^{(1)}$ beibehalten, jedoch mit einem mittels

$$f_1 := g f_1, \quad g = \frac{f_2}{f_2 + f_3}, \quad 0 < g < 1$$

abgeänderten Funktionswert (modifizierter Schritt). Dann wird gesetzt

$$x^{(2)} := x^{(3)}, \quad f_2 := f_3.$$

ξ liegt jetzt zwischen $x^{(1)}$ und $x^{(2)}$, und $x^{(2)}$ ist die zuletzt berechnete Näherung.

4. *Prüfung der Abbruchbedingung:*

Falls $|x^{(1)} - x^{(2)}| \leq \text{tol}$

ist, erfolgt Abbruch. Dann wird gesetzt

$$\xi := x^{(2)}, \quad \text{falls } |f_2| \leq |f_1| \text{ ist,} \quad \text{und sonst} \quad \xi := x^{(1)}.$$

Andernfalls, also mit $|x^{(2)} - x^{(1)}| > \text{tol}$, wird die Iteration mit 1. fortgesetzt.

Für den modifizierten Schritt wird nachfolgend eine geometrische Interpretation angegeben.

Geometrische Interpretation für den modifizierten Schritt

Die Konstruktion von

$$f_1^* := f_1 \cdot \frac{f_2}{f_2 + f_3} \quad \text{ergibt sich wegen} \quad \frac{f_1^*}{f_1} = \frac{f_2}{f_2 + f_3}$$

geometrisch wie folgt (Abbildung 2.18): Die Verbindungsgeraden der Punkte $(x^{(2)}, f_2 + f_3)$ und $(x^{(1)}, f_1)$ sowie $(x^{(2)}, f_2)$ und $(x^{(1)}, f_1^*)$ schneiden sich in einem Punkt S der x -Achse. Mit dem Punkt $(x^{(1)}, f_1^*)$ verbessert sich die Chance, das Einschlussintervall auf der Seite von $x^{(1)}$ zu verkürzen. Das zeigt die nächste Näherung $x^{(4)}$ in Abbildung 2.18.

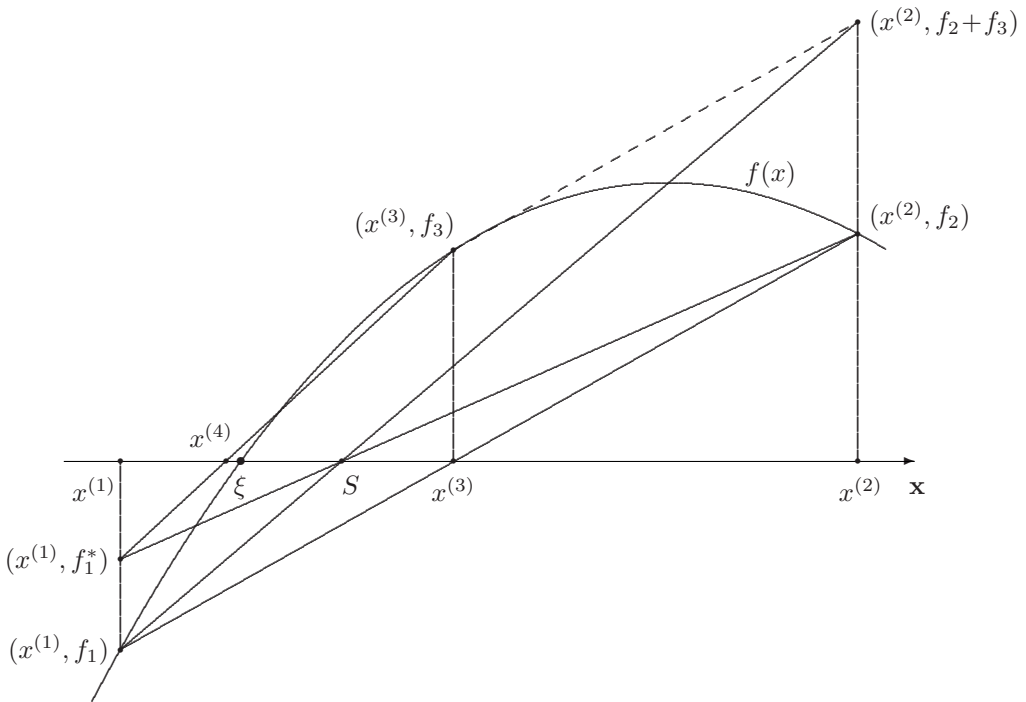


Abb. 2.18. Zum Pegasus-Verfahren, modifizierter Schritt

Beispiel 2.37. (vgl. Beispiel 2.35)

Gegeben: Die Funktion $f(x) = \sin x + 1 - 1/x$ und die Schranke $\text{RelErr} = 5 \cdot 10^{-7}$ für den relativen Fehler ($\text{AbsErr} = 0$).

Gesucht: Die Nullstelle ξ von f im Intervall $(0.6, 0.7)$ mit Hilfe des Pegasus-Verfahrens.

Lösung:

| ν | $x^{(3)}$ | $f(x^{(3)})$ | $x^{(1)}$ | $f(x^{(1)})$ | $x^{(2)}$ | $f(x^{(2)})$ | $ x^{(2)} - x^{(1)} / x^{(2)} $ |
|-------|-----------|--------------|-----------|--------------|-----------|--------------|---------------------------------|
| | | | 0.6 | $f < 0$ | 0.7 | $f > 0$ | |
| 1 | 0.6321164 | +8.9e-3 | 0.6 | $f < 0$ | 0.6321164 | $f > 0$ | 5.08e-2 |
| 2 | 0.6294517 | +1.7e-5 | 0.6 | $f < 0$ | 0.6294517 | $f > 0$ | 4.68e-2 |
| 3 | 0.6294465 | -5.6e-8 | 0.6294517 | $f > 0$ | 0.6294465 | $f < 0$ | 8.26e-6 |
| 4 | 0.6294468 | +8.9e-7 | 0.6294465 | -6e-8 | 0.6294468 | 9e-7 | 4.45e-7 |

Zur Darstellung dieser Tabelle vgl. Beispiel 2.33. Wegen $|x^{(2)} - x^{(1)}|/|x^{(2)}| < 5 \cdot 10^{-7}$ für $\nu = 4$ ist die Nullstelle auf 7 Ziffern genau, also $\xi \approx x^{(1)} = 0.629\,446\,5$. \square

2.7.5 Das Verfahren von Anderson-Björck

Das Verfahren von Anderson-Björck arbeitet ähnlich wie das Pegasus-Verfahren; lediglich im 3. Schritt des Algorithmus (bei der Festlegung des neuen Einschlussintervalls) wird die Modifikation des Funktionswertes f_1 an der Stelle $x^{(1)}$ auf andere Weise vorgenommen. Die Konvergenzordnung des Verfahrens in der Umgebung einer einfachen Nullstelle liegt zwischen 1.682 und 1.710.

Algorithmus 2.38. (*Verfahren von Anderson-Björck*)

Gegeben: (i) $f \in C[a, b]$ mit $f(a) \cdot f(b) < 0$,
(ii) Schranken AbsErr und RelErr für den absoluten bzw. den relativen Fehler mit AbsErr > 0 und RelErr = 0
oder mit AbsErr = 0 und RelErr > 0 .

Gesucht: Eine Zahl $\xi \in (a, b)$, für die $f(\xi) = 0$ ist, oder ein Einschlussintervall $[x^{(1)}, x^{(2)}]$ bzw. $[x^{(2)}, x^{(1)}]$ für ξ mit
 $|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{RelErr} + \text{AbsErr}$.

Vorbereitung: $x^{(1)} := a, \quad x^{(2)} := b$
 $f_1 := f(x^{(1)}), \quad f_2 := f(x^{(2)})$.

Pro Iterationsschritt ist wie folgt vorzugehen:

1. *Bestimmung der neuen Intervallgrenze:*

Berechne $\text{tol} := |x^{(2)}| \text{RelErr} + \text{AbsErr}$
und $\Delta x := \frac{f_2}{f_2 - f_1} (x^{(1)} - x^{(2)})$.

Wenn $|\Delta x| \leq \text{tol}$ ist, setze
 $\Delta x := 0.9 \cdot \text{tol} \cdot \text{sgn}(x^{(1)} - x^{(2)})$.

Berechne $x^{(3)} := x^{(2)} + \Delta x$.

2. *Berechnung des neuen Funktionswertes* $f_3 := f(x^{(3)})$.

Falls $f_3 = 0$ ist, wird die Iteration mit $\xi := x^{(3)}$ abgebrochen, andernfalls geht es mit 3. weiter.

3. *Festlegung des neuen Einschlussintervalls:*

Falls $f_2 \cdot f_3 < 0$ gilt, wenn also die Nullstelle ξ zwischen $x^{(2)}$ und $x^{(3)}$ liegt, wird gesetzt

$$x^{(1)} := x^{(2)}, \quad x^{(2)} := x^{(3)}, \quad f_1 := f_2, \quad f_2 := f_3.$$

Falls $f_2 \cdot f_3 > 0$ gilt, wenn also die Nullstelle ξ zwischen $x^{(1)}$ und $x^{(3)}$ liegt, wird die Stelle $x^{(1)}$ beibehalten und ihr ein abgeänderter Funktionswert

$$f_1 := g f_1 \quad \text{mit} \quad g = 1 - f_3/f_2, \quad g < 1$$

oder, falls $g \leq 0$ ist, mit $g = 0.5$ zugeordnet (modifizierter Schritt). Dann wird gesetzt

$$x^{(2)} := x^{(3)}, \quad f_2 := f_3.$$

Jetzt liegt ξ zwischen $x^{(1)}$ und $x^{(2)}$, und $x^{(2)}$ ist die zuletzt berechnete Näherung.

4. *Prüfung der Abbruchbedingung:*

Falls

$$|x^{(1)} - x^{(2)}| \leq \text{tol}$$

ist, erfolgt Abbruch. Dann wird gesetzt

$$\xi := x^{(2)}, \quad \text{falls } |f_2| \leq |f_1| \text{ ist,} \quad \text{und sonst} \quad \xi := x^{(1)}.$$

Andernfalls, also mit $|x^{(2)} - x^{(1)}| > \text{tol}$, wird die Iteration mit 1. fortgesetzt.

Die Originalarbeit zu dem Verfahren von Anderson-Björck ist [ANDE1973]. Auf dieser Basis wurden hier zusätzlich der Algorithmus formuliert und die geometrische Interpretation entwickelt.

Geometrische Interpretation für den modifizierten Schritt

Die Ersetzung von f_1 durch $g f_1$ im 3. Schritt des Algorithmus lässt sich wie folgt geometrisch interpretieren (Abbildung 2.19): Durch die drei Punkte

$$P_1 = (x^{(1)}, f_1), \quad P_2 = (x^{(2)}, f_2), \quad P_3 = (x^{(3)}, f_3)$$

wird die interpolierende quadratische Parabel gelegt, und im (mittleren) Punkt P_3 wird die Parabeltangente konstruiert. Falls diese Tangente die x -Achse zwischen $x^{(3)}$ und $x^{(1)}$ schneidet, wird dieser Schnittpunkt $x^{(4)}$ als die nächste Näherung für die gesuchte Nullstelle genommen.

Die Parabeltangente kann wie folgt konstruiert werden. H_1 sei der Schnittpunkt der Gerade $P_2 P_3$ mit der Gerade $x = x^{(1)}$ und H_2 der Schnittpunkt der Gerade $P_1 P_3$ mit der Gerade $x = x^{(2)}$. Die Parabeltangente durch P_3 ist parallel zur Verbindungsgerade $H_1 H_2$.

Im Algorithmus wird die Tangente durch den Punkt P_3 und ihren Schnittpunkt $(x^{(1)}, f_1^*)$ mit der Gerade $x = x^{(1)}$ festgelegt.

Nach Anderson-Björck ergibt sich

$$f_1^* = g f_1 \quad \text{mit} \quad g = \frac{s_{23}}{s_{12}}, \quad s_{23} = \frac{f_2 - f_3}{x^{(2)} - x^{(3)}}.$$

Wenn $g > 0$ ist, liegen die Punkte $P_1 = (x^{(1)}, f_1)$ und $(x^{(1)}, f_1^*)$ auf derselben Seite der x -Achse, und der Schnittpunkt $x^{(4)}$ der Parabeltangente mit der x -Achse liegt zwischen $x^{(3)}$ und $x^{(1)}$.

Wegen des vorhergehenden Sekantenschritts sind die Punkte $(x^{(1)}, f_1)$, $(x^{(2)}, f_2)$ und $(x^{(3)}, 0)$ kollinear, und daher gilt

$$s_{12} = \frac{f_1 - f_2}{x^{(1)} - x^{(2)}} = \frac{f_2}{x^{(2)} - x^{(3)}}.$$

Damit ergibt sich für g wesentlich einfacher

$$g = \frac{s_{23}}{s_{12}} = \frac{f_2 - f_3}{x^{(2)} - x^{(3)}} \frac{x^{(1)} - x^{(2)}}{f_1 - f_2} = \frac{f_2 - f_3}{x^{(2)} - x^{(3)}} \frac{x^{(2)} - x^{(3)}}{f_2} = \frac{f_2 - f_3}{f_2} = 1 - \frac{f_3}{f_2}.$$

Wegen $f_3 \cdot f_2 > 0$ ist $g < 1$. Falls $g \leq 0$ ist, schneidet die Tangente die x -Achse nicht zwischen $x^{(3)}$ und $x^{(1)}$. In diesem Fall wird (wie beim Illinois-Verfahren) $g = 0.5$ gesetzt, also $f_1^* = 0.5 f_1$.

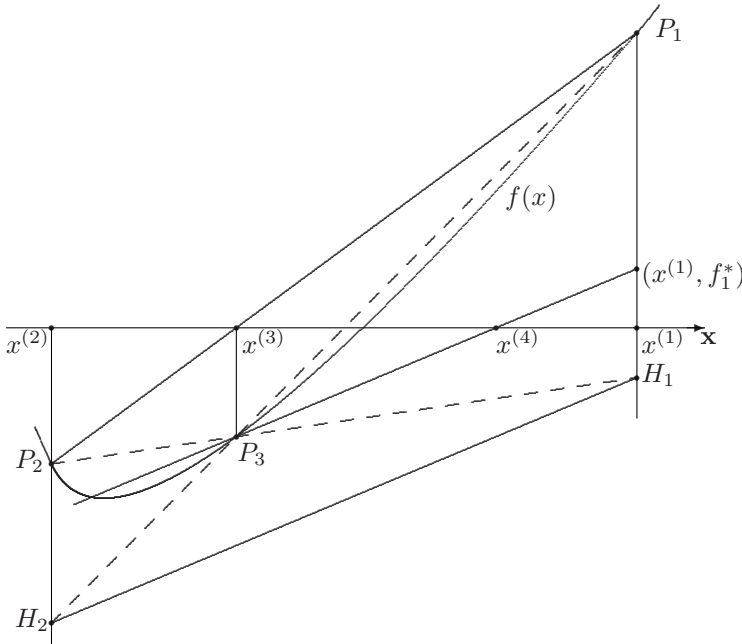


Abb. 2.19. Zum Verfahren von Anderson-Björck. Mit den Geraden P_2P_3 und P_1P_3 ergeben sich die Punkte H_1 und H_2 . Die zu H_1H_2 parallele Gerade durch P_3 ist die Parabeltangente in diesem Punkt.

Beispiel 2.39.

Gegeben: Die Funktion $f(x) = \sin x + 1 - 1/x$ und die Schranke $\text{RelErr} = 5 \cdot 10^{-7}$ für den relativen Fehler ($\text{AbsErr} = 0$).

Gesucht: Die Nullstelle ξ von f im Intervall $(0.6, 0.7)$ mit Hilfe des Verfahrens von Anderson-Björck.

Lösung:

| ν | $x^{(3)}$ | $f(x^{(3)})$ | $x^{(1)}$ | $f(x^{(1)})$ | $x^{(2)}$ | $f(x^{(2)})$ | $ x^{(2)} - x^{(1)} / x^{(2)} $ |
|-------|------------|--------------|------------|--------------|------------|--------------|---------------------------------|
| | | | 0.6 | $f < 0$ | 0.7 | $f > 0$ | |
| 1 | 0.63211636 | +8.9e-3 | 0.6 | $f < 0$ | 0.63211636 | $f > 0$ | 5.08e-2 |
| 2 | 0.62944753 | +3.5e-6 | 0.6 | $f < 0$ | 0.62944753 | $f > 0$ | 4.68e-2 |
| 3 | 0.62944648 | -1.2e-8 | 0.62944753 | $f > 0$ | 0.62944648 | $f < 0$ | 1.67e-6 |
| 4 | 0.62944676 | +9.3e-7 | 0.62944648 | -1.2e-8 | 0.62944676 | 9.3e-7 | 4.45e-7 |

Zur Darstellung dieser Tabelle vgl. Beispiel 2.33. Wegen $|x^{(2)} - x^{(1)}|/|x^{(2)}| < 5 \cdot 10^{-7}$ für $\nu = 4$ ist die Nullstelle auf 7 Ziffern genau, also $\xi \approx x^{(1)} = 0.629\,446\,5$. \square

2.7.6 Die Verfahren von King und Anderson-Björck-King.
Das Illinois-Verfahren

Das Verfahren von King (s. Originalarbeit [KING1973]) unterscheidet sich vom Pegasus-Verfahren nur dadurch, dass nie zwei Sekantenschritte nacheinander ausgeführt werden, sondern auf jeden Sekantenschritt ein modifizierter Schritt folgt. Es besitzt eine etwas höhere Konvergenzordnung (siehe Abschnitt 2.9).

Das Verfahren von Anderson-Björck-King verläuft ganz analog; es arbeitet nach dem Verfahren von Anderson-Björck mit der Zusatzbedingung von King, dass nie zwei Sekantenschritte nacheinander erfolgen dürfen. Hier geschieht der modifizierte Schritt nach der Anderson-Björck-Methode, beim Verfahren von King nach der Pegasus-Methode.

Gemeinsam ist den Einschlussverfahren Pegasus, Anderson-Björck und damit auch King der Sekantenschritt. Lediglich der modifizierte Schritt wird unterschiedlich realisiert; bei Pegasus wird $g = f_2/(f_2 + f_3)$ gesetzt, bei Anderson-Björck $g = 1 - f_3/f_2$ und, falls $g \leq 0$ ist, wird $g = 0.5$ gesetzt. Das Illinois-Verfahren verwendet stets $g = 0.5$.

2.7.7 Ein kombiniertes Einschlussverfahren

In dem folgenden Algorithmus 2.40 wird das gegebene Einschlussintervall $[a, b]$ zunächst mit dem Bisektionsverfahren so lange verkleinert, bis seine Länge nicht größer ist als eine vorgegebene Länge LB . Erst dann kommt ein Einschlussverfahren höherer Konvergenzordnung, das Pegasus-Verfahren, das Verfahren von Anderson-Björck oder auch das Illinois-Verfahren, zum Einsatz.

Die Formulierung der Algorithmen 2.36 und 2.38 ist günstig für die Herleitung und die geometrische Interpretation der Verfahren. Der folgende Algorithmus führt dagegen einen eventuell erforderlichen modifizierten Schritt erst nach einer nicht erfüllten Abbruchbedingung aus; daher weichen die Formeln für g von denen in Abschnitt 2.7.4 und 2.7.5 ab.

Der Algorithmus 2.40 kann in etwas erweiterter Form auch mit den Verfahren von King und Anderson-Björck-King arbeiten. Zugunsten einer einfachen und übersichtlichen Darstellung wird auf diese Verfahren verzichtet (siehe jedoch Beispiel 2.41).

Algorithmus 2.40.

Gegeben: $f \in C[a, b]$ mit $f(a) \cdot f(b) < 0$, Schranken AbsErr und RelErr für den absoluten bzw. den relativen Fehler mit AbsErr > 0 und RelErr $= 0$ oder mit AbsErr $= 0$ und RelErr > 0 , die Länge $LB > 0$ für die Entscheidung zwischen Bisektions- und Sekantenschritt, die maximal zulässige Anzahl $nfmax$ der Funktionsauswertungen.

Gesucht: Eine Zahl $\xi \in (a, b)$, für die $f(\xi) = 0$ ist, oder ein Einschlussintervall $[x^{(1)}, x^{(2)}]$ bzw. $[x^{(2)}, x^{(1)}]$ für ξ mit

$$|x^{(1)} - x^{(2)}| \leq |x^{(2)}| \text{ RelErr} + \text{AbsErr}.$$

Vorbereitung: $x^{(1)} := a$, $x^{(2)} := b$, $f_1 := f(x^{(1)})$, $f_2 := f(x^{(2)})$,
 $nf := 2$, $v := x^{(1)} - x^{(2)}$.

1. Wenn $nf \geq nfmax$ ist, konnte mit $nfmax$ Funktionsauswertungen eine Nullstelle von f nicht ermittelt werden. Die Iteration wird dann abgebrochen.
2. Berechnung der aktuellen Fehlertoleranz: $\text{tol} := |x^{(2)}| \text{ RelErr} + \text{AbsErr}$.
3. Verfahrensschritt wählen und nächste Näherung berechnen:
 Wenn $|v| > LB$ ist, wird ein Bisektionsschritt ausgeführt mit

$$\Delta x := v \cdot 0.5, \quad \text{Bis} := 1,$$

andernfalls ein Sekantenschritt mit

$$\Delta x := v \cdot f_2 / (f_2 - f_1), \quad \text{Bis} := 0.$$

Falls $|\Delta x| \leq \text{tol}$ ist, wird

$$\Delta x := 0.9 \cdot \text{sgn}(v) \cdot \text{tol}$$

gesetzt. Die nächste Näherung ist

$$x^{(3)} := x^{(2)} + \Delta x.$$

4. Berechnung des Funktionswertes

$$f_3 := f(x^{(3)}), \quad nf := nf + 1.$$

Falls $f_3 = 0$ ist, wird die Iteration mit $\xi := x^{(3)}$ abgebrochen.

5. Ermittlung des neuen Einschlussintervalls:

Berechne $e := \operatorname{sgn}(f_2) \cdot f_3$. Wenn $e < 0$ ist, werden $x^{(1)}$ mit $x^{(2)}$ und f_1 mit f_2 getauscht. Dann werden $x^{(2)} := x^{(3)}$ gesetzt, f_2 mit f_3 getauscht und $v := x^{(1)} - x^{(2)}$ berechnet. Jetzt liegt ξ zwischen $x^{(1)}$ und $x^{(2)}$, und $x^{(2)}$ ist die zuletzt berechnete Näherung.

6. Prüfung der Abbruchbedingung: Wenn $|v| \leq \text{tol}$ ist, wird die Iteration abgebrochen mit $\xi := x^{(2)}$, falls $|f_2| \leq |f_1|$ ist, andernfalls mit $\xi := x^{(1)}$.

Wenn $|v| > \text{tol}$ ist, weiter mit 7.

7. Vorbereitung des nächsten Iterationsschrittes. Im Fall $e > 0$ wird der Funktionswert f_1 mittels $f_1 := g \cdot f_1$, $0 < g < 1$, wie folgt modifiziert:

a) Pegasus-Verfahren mit $g = f_3/(f_3 + f_2)$

b) Verfahren von Anderson-Björck:

- nach einem Bisektionsschritt ($Bis = 1$) mit $g = f_3/(f_3 + f_2)$
- nach einem Sekantenschritt ($Bis = 0$) mit $g = 1 - f_2/f_3$
oder, falls $g \leq 0$ ist, mit $g = 0.5$

c) Illinois-Verfahren mit $g = 0.5$

Mit dem ermittelten g wird $f_1 := g \cdot f_1$ gesetzt.

Die Iteration wird mit 1. fortgesetzt.

Bemerkungen.

- Für das kombinierte Verfahren eignet sich eine Länge LB mit $0 < LB \leq 0.2$, beispielsweise $LB = 0.15$. Mit $LB > |b - a|$ werden nur Sekantenschritte, mit $LB = 0$ nur Bisektionsschritte ausgeführt. Siehe dazu auch Beispiel 2.41.
- Zu 1. Die Vorgabe von $nfmax$ (etwa $nfmax = 100$) schützt vor einer Endlosschleife, wenn AbsErr oder RelErr zu klein oder wenn LB zu groß gewählt worden ist.
- Zu 5. Wenn $e < 0$ ist, liegt ξ zwischen $x^{(2)}$ und $x^{(3)}$, andernfalls zwischen $x^{(1)}$ und $x^{(3)}$.
- Zu 7. Wegen der Vertauschung von f_2 mit f_3 in 5. sehen die Formeln für g anders aus als in den Abschnitten 2.7.4 und 2.7.5. Ein Anderson-Björck-Schritt mit $g = 1 - f_2/f_3$ setzt voraus, dass vorher ein Sekantenschritt ($Bis = 0$) ausgeführt wurde (siehe Abschnitt 2.7.5). Deshalb wird er nach einem Bisektionsschritt ($Bis = 1$) durch einen Pegasusschritt ersetzt.

2.7.8 Das Zeroin-Verfahren

Der als Zeroin-Verfahren bezeichnete programmierte Algorithmus entspricht dem bei [DEKK1969], [FORS1977], [BREN1971] angegebenen Verfahren. Es ist eine geschickte Kombination des Bisektionsverfahrens ($p = 1$), des Sekantenverfahrens ($p = 1.618$) sowie der inversen quadratischen Interpolation ($p = 2$). Es ist ein Einschlussverfahren, bei dem ständig mit drei iterierten Näherungswerten für die gesuchte Nullstelle operiert wird. Aufgrund geometrischer Überlegungen wird jeweils das für die momentane Situation sinnvollste Verfahren eingesetzt. Eine ausführliche Beschreibung findet man bei [QUAR2002], Abschnitt 6.2.3.

Beispiel 2.41.

Für einen Vergleich verschiedener Einschlussverfahren werden üblicherweise die Nullstellen von gewissen Testfunktionen berechnet und die dafür benötigten Funktionsauswertungen gezählt.

Für 12 solcher Funktionen (als eine Auswahl aus einem umfangreicheren Test) sind die Ergebnisse in der nachfolgenden Tabelle zusammengestellt. Verglichen werden das Bisektionsverfahren, das Zeroin-Verfahren, das Illinois- und das Pegasus-Verfahren sowie die Verfahren von Anderson-Björck, King und Anderson-Björck-King.

Für die letzten fünf Verfahren wird der Algorithmus 2.40 (mit einer Erweiterung für die letzten beiden Verfahren) angewendet. Sie werden jeweils einmal ohne eine vorangehende Verkleinerung des Startintervalls $[a, b]$ ($LB > |b - a|$) und einmal mit einer solchen Verkleinerung mittels Bisektion ($LB < |b - a|$) eingesetzt. Mit $LB = 0$ führt der Algorithmus das Bisektionsverfahren aus.

Damit beim Test alle Verfahren mit einem Intervall der Länge LB beginnen, haben die Startintervalle die Längen $4 \cdot LB$, $8 \cdot LB$ oder $16 \cdot LB$, so dass nach 2, 3 bzw. 4 Bisektionsschritten die Länge LB erreicht ist. Für die Tabelle ist $LB = 0.15$ gewählt.

Die Abbruchbedingungen bei ν Funktionsauswertungen sind $f(x^{(\nu)}) = 0$ oder $|x^{(\nu)} - x^{(\nu-1)}| \leq |x^{(\nu)}| \cdot \text{RelErr}$ mit $\text{RelErr} = 2e-11$. Die angegebenen Nullstellen haben 10 gültige Ziffern.

Funktionen und Startintervalle:

1. $f(x) = x^2 \left(\frac{x^2}{3} + \sqrt{2} \sin x \right) - \frac{\sqrt{3}}{18}$, $[0, 1.2]$, $\xi = 0.399\,422\,2917$
2. $f(x) = 11x^{11} - 1$, $[0.4, 1.6]$, $\xi = 0.804\,133\,0975$
3. $f(x) = 35x^{35} - 1$, $[-0.5, 1.9]$, $\xi = 0.903\,407\,6632$
4. $f(x) = 2(xe^{-9} - e^{-9x}) + 1$, $[-0.5, 0.7]$, $\xi = 0.077\,014\,241\,35$
5. $f(x) = x^2 - (1 - x)^9$, $[-1.4, 1]$, $\xi = 0.259\,204\,4937$
6. $f(x) = (x - 1)e^{-9x} + x^9$, $[-0.8, 1.6]$, $\xi = 0.536\,741\,6626$

7. $f(x) = x^2 + \sin\left(\frac{x}{9}\right) - \frac{1}{4}, \quad [-0.5, 1.9], \quad \xi = 0.447\,541\,7621$

8. $f(x) = \frac{1}{8}\left(9 - \frac{1}{x}\right), \quad [0.001, 1.201], \quad \xi = \frac{1}{9} = 0.111\,111\,1111$

9. $f(x) = \tan x - x - 0.046\,3025, \quad [-0.9, 1.5], \quad \xi = 0.500\,000\,0340$

10. $f(x) = x^2 + x \sin(x\sqrt{75}) - 0.2, \quad [0.4, 1], \quad \xi = 0.679\,808\,9215$

11. $f(x) = x^9 + 0.0001, \quad [-1.2, 0], \quad \xi = -0.359\,381\,3664$

12. $f(x) = \ln x + \frac{x^2}{2 \cdot e} - 2 \frac{x}{\sqrt{e}} + 1, \quad [1, 3.4],$

Die Verfahren liefern bei Abbruch mit $f(x^{(\nu)}) = 0$ unterschiedliche Werte für die Nullstelle ξ , die alle im Intervall $[1.6483, 1.6492]$ liegen. Die exakte Nullstelle ist $\xi = \sqrt{e} = 1.648\,721\,270\,70 \dots$

Die Tabelle zeigt, dass bei der Mehrzahl der Funktionen eine vorbereitende Intervallverkleinerung zweckmäßig oder sogar unerlässlich ist. X bedeutet, dass die Nullstelle mit 100 Funktionsauswertungen nicht gefunden wurde.

Anzahl der Funktionsauswertungen

| Verfahren/Funktion | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|--|----|----|----|----|----|----|----|----|----|----|----|----|
| Bisektion | 40 | 39 | 40 | 41 | 41 | 40 | 41 | 42 | 40 | 38 | 40 | 12 |
| Illinois ohne Bisektion | 14 | 21 | 63 | 18 | 22 | 22 | 14 | 18 | 19 | 13 | 28 | 25 |
| Illinois mit Bisektion | 12 | 13 | 19 | 14 | 14 | 14 | 13 | 15 | 13 | 12 | 15 | 21 |
| Pegasus ohne Bisektion | 12 | 19 | 63 | 17 | 20 | 28 | 11 | 18 | 18 | 12 | 26 | 36 |
| Pegasus mit Bisektion | 11 | 12 | 16 | 12 | 12 | 11 | 11 | 16 | 11 | 10 | 14 | 29 |
| King ohne Bisektion | 11 | 17 | 60 | 16 | 19 | 28 | 12 | 18 | 17 | 10 | 25 | 38 |
| King mit Bisektion | 11 | 12 | 15 | 12 | 12 | 11 | 11 | 15 | 11 | 10 | 13 | 29 |
| Anderson-Björck ohne Bisektion | 14 | 74 | X | 29 | 11 | 22 | 14 | 7 | 19 | 12 | X | 30 |
| Anderson-Björck mit Bisektion | 10 | 11 | 16 | 11 | 11 | 11 | 12 | 11 | 12 | 10 | 14 | 24 |
| Anderson-Björck-King ohne Bisektion | 16 | 67 | X | 26 | 11 | 22 | 13 | 7 | 18 | 11 | X | 29 |
| Anderson-Björck-King mit Bisektion | 10 | 11 | 15 | 11 | 11 | 11 | 11 | 10 | 11 | 10 | 13 | 24 |
| Zeroin | 12 | 14 | 17 | 10 | 11 | 11 | 13 | 13 | 15 | 12 | 14 | 28 |

□

2.8 Anwendungsbeispiele

Beispiel 2.42. (Rohrleitungsbeispiel)

Es geht um die Ermittlung des Rohrdurchmessers d zur Optimierung der in Abb. 2.20 skizzierten Druckrohrleitung einer Freistrahlturbine. Die Düse am Rohrleitungsende ist in Abb. 2.20 rechts noch einmal getrennt skizziert.

Eingabeparameter:

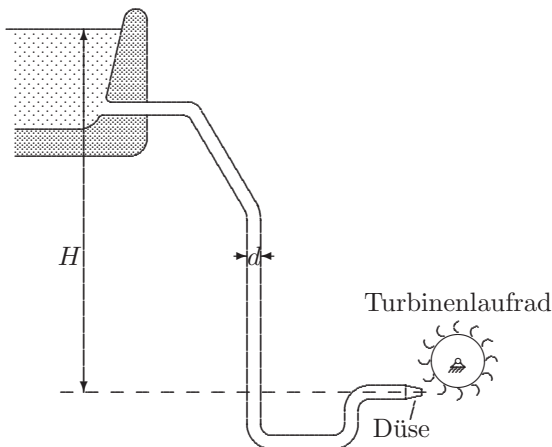
| | | |
|------------|---|--|
| α_0 | : | Geschätztes Flächenverhältnis Düse/Rohr, $\alpha \in [0.1, 0.25]$ |
| \dot{V} | : | Volumenstrom in $\frac{\text{m}^3}{\text{s}}$ |
| H | : | Gesamtgefälle in m |
| L | : | Gesamtlänge der Rohrleitung in m |
| ζ_D | : | Widerstandsbeiwert der Düse bezogen auf die Strahlgeschwindigkeit |
| ζ_R | : | Summe aller Formwiderstände der Rohrleitung bezogen auf die Strömungsgeschwindigkeit im Rohr |
| g | : | Gravitationskonstante in $\frac{\text{m}}{\text{s}^2}$ |
| ϱ | : | Dichte des Wassers in $\frac{\text{kg}}{\text{m}^3}$ |
| λ | : | Rohrreibungsbeiwert |

Zu den Eingabeparametern gehören noch Größen aus den Abbruchkriterien für die Iteration, z. B.

| | | |
|---------------|---|--|
| ε | : | $ d^{(\nu)} - d^{(\nu-1)} \leq \varepsilon$, $\varepsilon > 0$, ν = Iterationsindex |
| N_{\max} | : | Maximale Anzahl der Iterationen |

Ausgabeparameter:

| | | |
|-----|---|-----------------------|
| d | : | Rohrdurchmesser in m. |
|-----|---|-----------------------|



Düse am Rohrleitungsende:

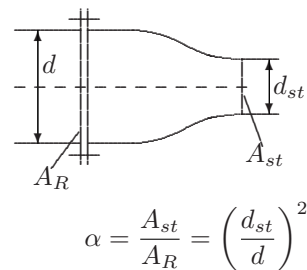


Abb. 2.20.

Vorgehensweise:

- (a) Zunächst berechnet man das optimale Flächenverhältnis der Düse α_{opt} aus der Forderung, dass die Strahlleistung P_{st} maximal wird.
- (b) Für dieses optimierte Flächenverhältnis α_{opt} ist dann derjenige Rohrdurchmesser d zu bestimmen, der für den vorgegebenen Volumenstrom erforderlich ist.

Die *Energiegleichung* (erweiterte *Bernoulli-Gleichung*) lautet

$$gH = \frac{v_{st}^2}{2} + \left(\lambda \frac{L}{d} + \zeta_R\right) \frac{v_R^2}{2} + \zeta_D \frac{v_{st}^2}{2}. \quad (2.56)$$

Aus der *Kontinuitätsgleichung*

$$\dot{V} = A_{st} v_{st} = A_R v_R \quad (2.57)$$

folgt mit $\alpha = A_{st}/A_R$ für die Strömungsgeschwindigkeit im Rohr

$$v_R = \alpha v_{st}.$$

Damit kann aus (2.56) die *Strahlgeschwindigkeit* v_{st} berechnet werden:

$$v_{st} = \sqrt{\frac{2gH}{1 + \zeta_D + \alpha^2(\lambda \frac{L}{d} + \zeta_R)}}.$$

Die *Strahlleistung* ist

$$\begin{aligned} P_{st} &= \frac{1}{2} \rho \dot{V} v_{st}^2 = \frac{1}{2} \rho A_R v_R v_{st}^2 = \frac{1}{2} \rho A_R \alpha v_{st}^3 \\ &= \frac{1}{2} \rho A_R \alpha \left[\frac{2gH}{1 + \zeta_D + \alpha^2(\lambda \frac{L}{d} + \zeta_R)} \right]^{3/2}. \end{aligned}$$

Zu (a): Mit der notwendigen Bedingung $\frac{\partial P_{st}}{\partial \alpha} = 0$ für eine maximale Strahlleistung ergibt sich das optimale Flächenverhältnis

$$\alpha_{opt} = \sqrt{\frac{1 + \zeta_D}{2(\lambda \frac{L}{d} + \zeta_R)}}.$$

(Wegen $\frac{\partial^2 P_{st}}{\partial \alpha^2}(\alpha_{opt}) < 0$ liegt tatsächlich ein Maximum vor.)

Setzt man α_{opt} für α in die Formel für die Strahlgeschwindigkeit v_{st} ein, so ergibt sich die optimale Strahlgeschwindigkeit

$$v_{st}(\alpha_{opt}) = \sqrt{\frac{4gH}{3(1 + \zeta_D)}}.$$

Damit folgt für die optimale Strömungsgeschwindigkeit im Rohr

$$v_R(\alpha_{opt}) = \alpha_{opt} \cdot v_{st}(\alpha_{opt}) = \sqrt{\frac{2gH}{3(\lambda \frac{L}{d} + \zeta_R)}}.$$

Zu (b): Mit der Querschnittsfläche des Rohres

$$A_R = \pi \frac{d^2}{4}$$

ist nach (2.57) der optimale Volumenstrom

$$\dot{V} = \pi \frac{d^2}{4} v_R(\alpha_{opt}) = \pi \frac{d^2}{4} \sqrt{\frac{2gH}{3(\lambda \frac{L}{d} + \zeta_R)}},$$

und daraus folgt für den optimalen Rohrdurchmesser d die nichtlineare Gleichung

$$d = \sqrt{\frac{4\dot{V}}{\pi} \sqrt{\frac{3}{2gH}} \sqrt{\lambda \frac{L}{d} + \zeta_R}} = \varphi(d). \quad (2.58)$$

Zweimaliges Quadrieren liefert

$$f(d) = d^4 - \frac{24\dot{V}^2}{\pi^2 g H} \left(\lambda \frac{L}{d} + \zeta_R \right) = 0. \quad (2.59)$$

Zahlenbeispiel.

$\dot{V} = 6.2 \text{ m}^3/\text{s}$, $H = 1130 \text{ m}$, $L = 1300 \text{ m}$, $\zeta_D = 0.04$, $\zeta_R = 2.5$, $g = 9.81 \text{ m/s}^2$,
 $\varrho = 10^3 \text{ kg/m}^3$, $\lambda = 0.02$.

Schranke für den absoluten Fehler $\varepsilon = 0.5 \cdot 10^{-6}$.

Mit den Startwerten $d^{(1)} = 0.7$ und $d^{(2)} = 0.8$ sind $f(0.7) < 0$ und $f(0.8) > 0$; also ist $0.7 < d < 0.8$. Für die Lösung der Gleichung (2.59)

$$d^4 - 0.008432327 \left(\frac{26}{d} + 2.5 \right) = 0$$

liefert das Pegasus-Verfahren mit 5 Iterationsschritten

$$d = 0.748551.$$

Somit ergibt sich für den gesuchten optimalen Rohrdurchmesser bei Rundung auf 3-stellige Mantisse $d \approx 0.749 \text{ m}$.

Mit dem Quadrat der optimalen Strahlgeschwindigkeit

$$v_{st}^2(\alpha_{opt}) = \frac{4gH}{3(1 + \zeta_D)} = \frac{4 \cdot 9.81 \cdot 1130}{3 \cdot 1.04} \frac{\text{m}^2}{\text{s}^2} = 14211.9 \text{ m}^2/\text{s}^2$$

erhält man die maximale Strahlleistung

$$P_{st} = \frac{1}{2} \varrho \dot{V} v_{st}^2(\alpha_{opt}) = 44057 \text{ kW}.$$

Das allgemeine Iterationsverfahren, angewendet auf die Gleichung (2.58)

$$d = \sqrt{0.091\,827\,70 \sqrt{\frac{26}{d}} + 2.5} = \varphi(d)$$

benötigt mit dem Startwert $d^{(0)} = 0.7$ und mit derselben Fehlerschranke 10 Iterationsschritte. Es konvergiert wesentlich schlechter als das Pegasus-Verfahren. \square

Beispiel 2.43. (*Turbinenbeispiel*)

Eine Turbine arbeitet zwischen zwei Wasserspeichern, deren Spiegel die Höhendifferenz h besitzen. Stromabwärts von der Turbine wird das Wasser über einen Diffusor, dessen Endquerschnitt die Fläche A hat, in das untere Becken eingeleitet. An der Turbinenwelle wird die Leistung P angenommen. Frage: Welcher Volumenstrom \dot{V} fließt durch die Turbine, wenn von den Verlusten nur der Austrittsverlust am Diffusor berücksichtigt wird?

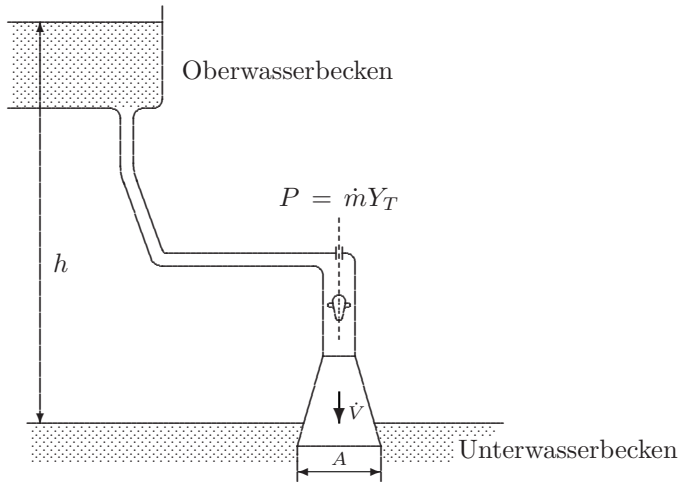


Abb. 2.21.

Ausgehend von der erweiterten Bernoulli-Gleichung erhält man

$$g \cdot h - Y_T = \frac{1}{2} \left(\frac{\dot{V}}{A} \right)^2 ;$$

dabei sind Y_T die Stutzenarbeit der Turbine und $0.5(\dot{V}/A)^2$ der Austrittsverlust am Diffusor. Durch Auflösung nach Y_T und wegen $Y_T = P/(\varrho \dot{V})$ ergibt sich die Beziehung

$$g \cdot h - \frac{1}{2} \left(\frac{\dot{V}}{A} \right)^2 = \frac{P}{\varrho \dot{V}} . \quad (2.60)$$

Durch einfache Umstellung folgt aus (2.60) die kubische Gleichung für \dot{V}

$$f(\dot{V}) := \dot{V}^3 - 2ghA^2\dot{V} + \frac{2PA^2}{\varrho} = 0 .$$

Die praktisch interessierende Lösung dieser Gleichung $f(\dot{V}) = 0$ kann z. B. mit dem Newton-Verfahren ermittelt werden. Ein Startwert $\dot{V}^{(0)}$ für die Iteration ergibt sich aus (2.60) bei Vernachlässigung des Austrittsverlustes und anschließender Auflösung nach \dot{V} zu

$$\dot{V}^{(0)} = \frac{P}{\varrho \cdot h \cdot g}$$

mit $\varrho = 10^3 \text{ kg/m}^3$, $g = 9.81 \text{ m/s}^2$. Man kann aber auch mit jedem Einschussverfahren arbeiten.

Zur Lösung der kubischen Gleichung $f(\dot{V}) = 0$ könnte aber ebenso ein Verfahren zur Bestimmung sämtlicher Lösungen einer algebraischen Gleichung ohne Kenntnis von Startwerten (z. B. das Verfahren von Muller) verwendet werden (s. Abschnitt 3.3); hier ist dann aus den drei Lösungen die für den Anwendungsfall sinnvolle Lösung auszuwählen (siehe dazu das Beispiel 3.11). \square

2.9 Effizienz der Verfahren und Entscheidungshilfen

Mit Hilfe des *Effizienzindex* E von Traub (s. [TRAU1984], App. C) lassen sich die Iterationsverfahren gut vergleichen. Seien H die *Hornerzahl* (Anzahl der erforderlichen Funktionsauswertungen pro Iterationsschritt) und p die Konvergenzordnung eines Iterationsverfahrens, so ist

$$E := p^{1/H}$$

der Effizienzindex.

Die folgende Tabelle gibt eine Übersicht über die Konvergenzordnung und den Effizienzindex bei Verfahren zur Berechnung einfacher und mehrfacher Nullstellen; *je größer E , desto effizienter ist das Verfahren in der Umgebung der Nullstelle.*

| Verfahren | Konvergenzordnung p | Hornerzahl H | Effizienz E | Einschlussverfahren |
|--|-----------------------|----------------|---------------|---------------------|
| Newton, einfache Nullstelle | 2 | 2 | 1.414 | nein |
| Newton, mehrfache Nullstelle | 2 | 2 | 1.414 | nein |
| Newton, modifiziert | 2 | 3 | 1.260 | nein |
| Bisektion | 1 | 1 | 1 | ja |
| Regula falsi | 1 | 1 | 1 | ja |
| Sekantenverfahren, einfache Nullstelle | 1.618 | 1 | 1.618 | nein |

| Verfahren | Konvergenz- ordnung p | Horner- zahl H | Effizienz E | Einschluss- verfahren |
|--------------------------|----------------------------|---------------------|------------------|--------------------------|
| Illinois | 1.442 | 1 | 1.442 | ja |
| Pegasus | 1.642 | 1 | 1.642 | ja |
| Anderson-Björck | 1.682...1.710 | 1 | 1.682...1.710 | ja |
| King | 1.710...1.732 | 1 | 1.710...1.732 | ja |
| Anderson- Björck-King | 1.710...1.732 | 1 | 1.710...1.732 | ja |

Es zeigt sich, dass man (nicht zuletzt wegen der sicheren Konvergenz) am effektivsten mit dem Pegasus-Verfahren, dem Verfahren von Anderson-Björck oder den Verfahren von King und Anderson-Björck-King arbeitet. Das Bisektionsverfahren wird benutzt, um ein Startintervall mit Einschluss zu verkleinern, bevor eines der genannten Verfahren eingesetzt wird.

Will man einzelne *Polynom*-Nullstellen berechnen, so lässt sich auch das Newton-Verfahren zusammen mit dem Hornerschema zur Berechnung der Funktions- und Ableitungswerte effektiv einsetzen. Für transzendente Gleichungen sind weder das Sekantenverfahren noch das Newton-Verfahren zu empfehlen.

Die Ermittlung der Lösung einer Gleichung $f(x) = 0$ mit dem Algorithmus 2.17 für das allgemeine Iterationsverfahren erfordert umfangreiche Vorbereitungen. Der größte Aufwand steckt im 3. Schritt, in dem zu prüfen ist, ob bei der zu $f(x) = 0$ äquivalenten Gleichung $\varphi(x) = x$ die Funktion φ den Voraussetzungen des Existenz- und Eindeutigkeitssatzes 2.11 genügt. Dagegen kann schon nach dem 1. Schritt – Festlegung eines Intervalls I , in dem mindestens eine Nullstelle von f liegt – ein Einschlussverfahren zur Bestimmung dieser Nullstelle eingesetzt werden. Darum hat das allgemeine Iterationsverfahren für die praktische Ermittlung von Nullstellen keine Bedeutung.

Für den Fall mehrfacher Nullstellen sind die „Anmerkungen zu mehrfachen Nullstellen“ in Abschnitt 2.5.3 zu beachten.

Ergänzende Literatur zu Kapitel 2

[BARW2007], 7.1-3; [DAHM2008], 5.5; [HAMM1994], 8.; [HERM2001], Kap.4; [MAES1988], 5.; [PREU2001], Kap.2; [SCHW1997], 5.; [STOE1989], 5.; [STOE2002A], 5.9; [WERN1993], II §§1-5.

Numerik-Algorithmen

Verfahren, Beispiele, Anwendungen

Engeln-Müllges, G.; Niederdrenk, K.; Wodicka, R.

2011, XXII, 756 S. 200 Abb. Mit Online-Extras.,

Hardcover

ISBN: 978-3-642-13472-2