§ 8 Nullstellen-Bestimmung durch Iterationsverfahren

Problemstellung 8.1:

Gegeben sei eine stetige Funktion $f: I \to \mathbb{R}$ auf einem Intervall I. Bestimme ein (oder mehrere oder alle) $\xi \in I$ mit $f(\xi) = 0$.

Ansatz 8.2:

Ausgehend von einem Startwert $x_0 \in I$ berechne weitere Näherungswerte x_n, $n \in \mathbb{N}$, für ξ mit Hilfe einer Iterationsfunktion $\Phi: I \to I$; das heißt, es soll gelten:

$$(8.2) \qquad x_n := \Phi(x_{n-1}) \qquad \text{für } n \in \mathbb{N}.$$

An Φ stellen wir die folgenden

Minimalforderungen:

(I) Für alle $x \in I$ mit $\Phi(x) = x$ ist $f(x) = 0$; das heißt: Jeder Fixpunkt von Φ ist Nullstelle von f.

(II) Φ ist stetig.

Lemma 8.3:

Es seien (I) und (II) erfüllt, und für festes $x_0 \in I$ sei die Folge $(x_n)_{n \geq 0}$ rekursiv durch (8.2) gegeben. Besitzt diese Folge einen Grenzwert $\xi_0 \in I$, so gilt:

$$f(\xi_0) = 0.$$

Beweis:

Aus (II) und (8.2) folgt:

$$\xi_0 = \lim_{n\to\infty} x_n = \lim_{n\to\infty} \Phi(x_{n-1}) = \Phi(\lim_{n\to\infty} x_{n-1}) = \Phi(\xi_0).$$

Aus (I) folgt daher: $f(\xi_0) = 0$. $\square$

Beispiel:

Zur Bestimmung der kleinsten positiven Nullstelle der Cosinus-Funktion (also $\frac{\pi}{2}$) machen wir folgenden Ansatz:

Definiere $f, \Phi: \mathbb{R} \to \mathbb{R}$ durch

$$f(x) := \cos x \quad , \quad \Phi(x) := x + \cos x .$$

Die Minimalforderungen (I) und (II) sind erfüllt; insbesondere gilt für $x \in \mathbb{R}$ folgende Äquivalenz:

$$\Phi(x) = x \Leftrightarrow f(x) = 0 .$$

Mit dem Startwert $x_0 = 1$ erhalten wir mittels (8.2) folgende Näherungswerte für $\frac{\pi}{2}$:

$$x_1 = 1 + \cos 1 \approx 1{,}540302306 \quad ,$$

$$x_2 = x_1 + \cos x_1 \approx 1{,}570791601 \quad ,$$

$$x_3 = x_2 + \cos x_2 \approx 1{,}570796327 \quad ,$$

$$x_n = x_{n-1} + \cos x_{n-1} \approx x_3 \quad \text{für } n \geq 4 .$$

Dabei ist $x_3 = \frac{\pi}{2} + \Theta$ mit $|\Theta| < 10^{-9}$.

Definition 8.5:

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt Cauchy-Folge, falls gilt:

Zu jedem $\varepsilon > 0$ gibt es ein $N = N(\varepsilon) \in \mathbb{N}$, so dass für alle $n, m \in \mathbb{N}$ mit $n, m \geq N$ gilt: $|x_n - x_m| < \varepsilon$.

Bemerkung 8.6:

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ ist genau dann eine Cauchy-Folge, wenn sie konvergiert.

Ist nämlich $x = \lim\limits_{n\to\infty} x_n$, so gibt es zu vorgegebenem $\varepsilon > 0$ ein $N = N(\varepsilon) \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}$ mit $n \geq N$ gilt: $|x_n - x| < \frac{\varepsilon}{2}$.

Dann folgt für alle $n, m \geq N$:

$$|x_n - x_m| \leq |x_n - x| + |x - x_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist jede konvergente Folge eine Cauchy-Folge.

Zum Beweis der Umkehrung siehe Übung 14.41.

Warnung:

Eine Cauchy-Folge mit lauter rationalen Gliedern muss ihren Grenzwert nicht in $\mathbb{Q}$ haben. Beispielsweise ist die irrationale Zahl $\sqrt{2}$ Grenzwert einer Folge von rationalen Zahlen.

Definition 8.7:

Es seien $a, b \in \mathbb{R}$ mit $a < b$. Eine Abbildung $g: [a,b] \to [a,b]$ heißt <u>kontrahierend</u>, falls ein $L \in \mathbb{R}$ mit $0 \leq L < 1$ existiert, so dass für alle $x, y \in [a,b]$ gilt:

(8.7) $\quad |g(x) - g(y)| \leq L \cdot |x - y|.$

Bemerkung 8.8:

Jede kontrahierende Abbildung $g: [a,b] \to [a,b]$ ist stetig.
Umgekehrt ist jedoch etwa die Abbildung $f: [0,1] \to [0,1]$, gegeben durch $f(x) := x^2$ zwar stetig, aber nicht kontrahierend.

Satz 8.9, Der Banachsche Fixpunktsatz:

Seien $a, b \in \mathbb{R}$ mit $a < b$, und die Abbildung $g: [a,b] \to [a,b]$ sei kontrahierend, sei also $L \in [0,1)$ so, dass (8.7) für alle $x, y \in [a,b]$ gilt.
Dann besitzt g genau einen Fixpunkt ξ_g.
Genauer gilt:
Sei $x_0 \in [a,b]$ beliebig, und definiere die Folge $(x_n)_{n \geq 0}$ rekursiv durch $x_{n+1} := g(x_n)$ für $n \geq 0$.
Dann gilt für alle $n \geq 1$:

(8.9 a) $\quad |x_n - \xi_g| \leq \dfrac{L^n}{1-L} \cdot |x_1 - x_0|.$

Insbesondere folgt:

(8.9 b) $\quad \lim\limits_{n \to \infty} x_n = \xi_g.$

Beweis:

Wir zeigen zunächst, dass g höchstens einen Fixpunkt besitzt.

Seien also $\xi_1, \xi_2 \in [a,b]$ mit $g(\xi_1) = \xi_1$ und $g(\xi_2) = \xi_2$.

Zu zeigen ist: $\xi_1 = \xi_2$.

(8.7) liefert:

$$|\xi_1 - \xi_2| = |g(\xi_1) - g(\xi_2)| \le L \cdot |\xi_1 - \xi_2|.$$

Wegen $L < 1$ ist das nur möglich, wenn $|\xi_1 - \xi_2| = 0$ - und damit $\xi_1 = \xi_2$ ist.

Zum Nachweis der Existenz sei nun die Folge $(x_n)_{n\ge 0}$ wie im Satz.

Wir zeigen nun:

(8.9 c) $\quad |x_n - x_{n-1}| \le L^{n-1} \cdot |x_1 - x_0|$ für alle $n \in \mathbb{N}$,

(8.9 d) $\quad |x_{n+k} - x_n| \le \dfrac{L^n}{1-L} \cdot |x_1 - x_0|$ für alle $n, k \in \mathbb{N}$.

(8.9 c) folgt sofort durch Induktion:

(8.9 c) ist trivial für $n = 1$.

Im Induktionsschritt folgt mittels (8.7):

$$|x_{n+1} - x_n| = |g(x_n) - g(x_{n-1})| \le L \cdot |x_n - x_{n-1}| \le L^n \cdot |x_1 - x_0|.$$

Nun folgt weiter für alle $n, k \in \mathbb{N}$ - mittels der Formel für die endliche geometrische Reihe:

$$|x_{n+k} - x_n| = \Big|\sum_{j=0}^{k-1} (x_{n+j+1} - x_{n+j})\Big|$$

$$\le \sum_{j=0}^{k-1} |x_{n+j+1} - x_{n+j}| \le \sum_{j=0}^{k-1} L^{n+j} \cdot |x_1 - x_0|$$

$$= (L^n + L^{n+1} + \dots + L^{n+k-1}) \cdot |x_1 - x_0|$$

$$= \frac{L^n - L^{n+k}}{1-L} \cdot |x_1 - x_0|$$

$$\leq \frac{L^n}{1-L} \cdot |x_1 - x_0| .$$

Damit ist auch (8.9d) bewiesen.

Sei nun $\varepsilon > 0$. Dann gibt es ein $N = N(\varepsilon) \in \mathbb{N}$ mit:

$$\frac{L^N}{1-L} \cdot |x_1 - x_0| < \varepsilon .$$

(8.9d) liefert nun auch, dass für alle $n, m \in \mathbb{N}$ mit $m > n \geq N$ (und $k := m-n$) gilt:

$$|x_m - x_n| \leq \frac{L^n}{1-L} \cdot |x_1 - x_0| < \varepsilon .$$

Das bedeutet: Die Folge $(x_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge in $[a, b]$ - und besitzt daher nach Bemerkung 8.6 einen Grenzwert $\xi \in [a, b]$.

(8.9d) liefert weiter durch Grenzübergang für $k \to \infty$:

$$|\xi - x_n| \leq \frac{L^n}{1-L} \cdot |x_1 - x_0| \quad \text{für alle } n \in \mathbb{N}.$$

Damit folgen (8.9a) und (8.9b) für $\xi_0 := \xi$.

Schließlich folgt aus der Stetigkeit von g:

$$g(\xi_0) = g(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} g(x_n) = \lim_{n\to\infty} x_{n+1} = \xi_0 .$$

ξ_0 ist also Fixpunkt von g - und aufgrund der eingangs gezeigten Eindeutigkeit unabhängig von dem speziell gewählten Anfangswert $x_0 \in [a, b]$.

$\square$

Definition 8.10, Das Newton-Verfahren:

Sei $I \subseteq \mathbb{R}$ ein Intervall, und $f: I \to \mathbb{R}$ sei stetig differenzierbar mit $f'(x) \neq 0$ für alle $x \in I$.
Definiere $\Phi: I \to \mathbb{R}$ durch

$$\Phi(x) := x - \frac{f(x)}{f'(x)} . \tag{8.10}$$

Ist $\Phi(I) \subseteq I$, so heißt das durch die Iterationsfunktion Φ gegebene Iterationsverfahren das Newton-Verfahren zur Bestimmung einer Nullstelle ξ von f.

Bemerkungen 8.11:

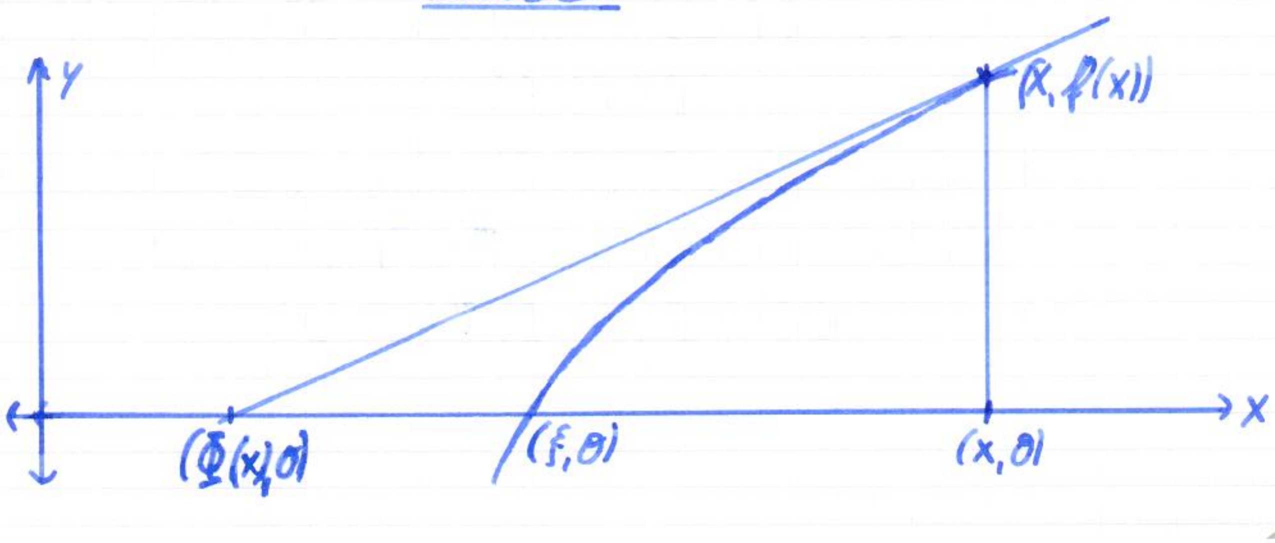
i) Seien f, Φ wie in Definition 8.10 - mit $\Phi(I) \subseteq I$.
Dann erfüllt Φ auch die geforderten Minimal=forderungen (I) und (II).

ii) (8.10) besagt auch:

$$\frac{f(x)}{x - \Phi(x)} = f'(x) . \tag{8.11}$$

Das bedeutet:
Der Punkt $(\Phi(x), 0)$ ist der Schnittpunkt der x-Achse mit der Tangente an f durch den Punkt $(x, f(x))$.

Skizze



Satz 8.12:

Sei $f: I \to \mathbb{R}$ zweimal stetig differenzierbar auf einem Intervall $I \subseteq \mathbb{R}$, sei $f'(x) \neq 0$ für alle $x \in I$, und sei $\xi \in I$ eine Nullstelle von f, die kein Randpunkt von I ist. Dann gibt es ein $\delta > 0$ mit $[\xi - \delta, \xi + \delta] \subseteq I$, so dass für Φ wie in (8.10) folgt:

i) Es gibt ein $L \in [0, 1)$, so dass für alle $x \in [\xi - \delta, \xi + \delta]$ gilt:

(8.12 a) $$|\Phi'(x)| \leq L.$$

ii) Es ist $\Phi([\xi - \delta, \xi + \delta]) \subseteq [\xi - \delta, \xi + \delta]$.

iii) Für jeden Startwert $x_0 \in [\xi - \delta, \xi + \delta]$ folgt für die durch (8.2) gegebene Folge $(x_n)_{n \geq 0}$:

(8.12 b) $$\lim_{n \to \infty} x_n = \xi.$$

Beweis:

i) Weil f auf I zweimal stetig differenzierbar ist, ist Φ auf I (mindestens) einmal stetig differenzierbar, und für alle $x \in I$ folgt:

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

Insbesondere ist $\Phi'(\xi) = 0$. Weil Φ' stetig ist, gibt es daher sogar zu vorgegebenem $L \in [0, 1)$ ein $\delta > 0$ mit $[\xi - \delta, \xi + \delta] \subseteq I$, so dass für alle $x \in [\xi - \delta, \xi + \delta]$ gilt:

$$|\Phi'(x)| \leq L.$$

ii) Es seien $x, y \in [\xi - \delta, \xi + \delta]$. Dann gibt es nach dem Mittelwertsatz ein $\eta \in I[x, y]$ mit:

$$\Phi(x) - \Phi(y) = \Phi'(\eta) \cdot (x - y).$$

ii) liefert damit weiter für alle $x, y \in [\xi-\delta, \xi+\delta]$:

(8.12 c) $\qquad |\Phi(x) - \Phi(y)| \leq L \cdot |x-y|.$

Insbesondere folgt wegen $\Phi(\xi) = \xi$:

$$|\Phi(x) - \xi| = |\Phi(x) - \Phi(\xi)| \leq |x-\xi| \quad \text{für alle } x \in [\xi-\delta, \xi+\delta].$$

Das bedeutet: $\quad \Phi([\xi-\delta, \xi+\delta]) \subseteq [\xi-\delta, \xi+\delta].$

iii) Nach (8.12c) ist Φ auf dem abgeschlossenen Intervall $[\xi-\delta, \xi+\delta]$ kontrahierend. Somit folgt aus dem Banachschen Fixpunktsatz, angewendet auf die Restriktion g von Φ auf $[\xi-\delta, \xi+\delta]$:

ξ ist der einzige Fixpunkt von g, und (8.12 b) gilt.

$\square$

Beispiel:

Wir berechnen numerisch die - eindeutig bestimmte - Nullstelle der - streng monoton steigenden - Funktion $f: \mathbb{R} \to \mathbb{R}$, gegeben durch

$$f(x) := x^3 + 3x + 1.$$

Für alle $x \in \mathbb{R}$ ist dann

$$f'(x) = 3x^2 + 3 > 0,$$

und die durch (8.10) gegebene Iterationsfunktion $\Phi: \mathbb{R} \to \mathbb{R}$ ist gegeben durch:

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

Mit dem Startwert $x_0 = 0$ erhalten wir weiter:

$$x_1 = -\frac{1}{3}, \quad x_2 = -\frac{29}{90} = -0{,}3\overline{2},$$

$$x_3 \approx -0{,}322185355,$$

$$x_n \approx x_3 \quad \text{für } n \geq 4.$$

Definition 8.13, Die Sekantenmethode:

Für eine stetige und injektive Funktion $f: I \to \mathbb{R}$ auf einem Intervall $I \subseteq \mathbb{R}$ betrachte folgendes Iterationsverfahren zur Bestimmung einer Nullstelle ξ von f:

Wähle zwei verschiedene Startwerte $x_0, x_1 \in I$.

Sei $n \geq 1$, und seien $x_0, x_1, \ldots, x_n$ bereits bestimmt.

Stopp, falls $x_n \notin I$.

Setze $x_m := x_n$ für alle $m > n$, falls $x_{n-1} = x_n$.

Andernfalls setze

$$(8.13) \qquad x_{n+1} := x_n - f(x_n) \cdot \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} = \frac{x_{n-1} \cdot f(x_n) - x_n \cdot f(x_{n-1})}{f(x_n) - f(x_{n-1})}.$$

Dieses Rekursionsverfahren heißt die <u>Sekantenmethode</u>.

Bemerkungen 8.14:

i) Ist $x_{n+1} = x_n$ für ein $n \geq 1$ und ist n minimal mit dieser Eigenschaft, so ist $\xi := x_n$ nach (8.13) eine Nullstelle von f.

ii) (8.13) besagt auch:

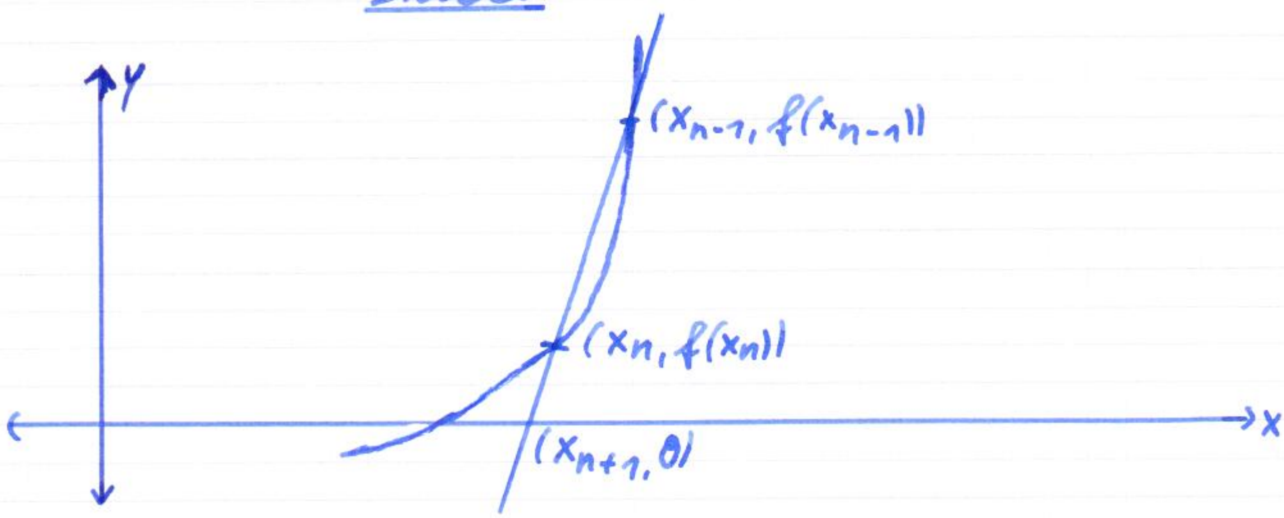
$$(8.14) \qquad \frac{f(x_n)}{x_n - x_{n+1}} = \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}}.$$

Das bedeutet:

Der Punkt $(x_{n+1}, 0)$ ist der Schnittpunkt der x-Achse mit der Sekante zu f durch die Punkte $(x_{n-1}, f(x_{n-1}))$ und $(x_n, f(x_n))$.

8.11

Skizze



iii) Bei der Regula falsi wird die Sekantenmethode dahingehend modifiziert, dass $f(x_{n-1})$ und $f(x_n)$ stets unterschiedliche Vorzeichen aufweisen - solange $f(x_n) \neq 0$ ist.
Das hat - unter anderem - den Vorteil, dass x_{n+1} stets zwischen x_{n-1} und x_n und damit auch in dem gegebenen Intervall I liegt.
Insbesondere sind die Startwerte $x_0, x_1 \in I$ so zu wählen, dass $f(x_0) \cdot f(x_1) \leq 0$ ist.

In Analogie zu Satz 8.12 gilt folgendes - allerdings schwieriger zu beweisendes - Ergebnis:

Satz 8.15:
Sei $f: I \to \mathbb{R}$ zweimal stetig differenzierbar auf einem Intervall $I \subseteq \mathbb{R}$, sei $f'(x) \neq 0$ für alle $x \in I$, und sei $\xi \in I$ eine Nullstelle von f, die kein Randpunkt von I ist. Dann gibt es ein $\delta > 0$, so dass für alle $x_0, x_1 \in [\xi - \delta, \xi + \delta]$ mit $x_0 \neq x_1$ die Sekantenmethode eine Folge $(x_n)_{n \geq 0}$ liefert mit $\lim\limits_{n \to \infty} x_n = \xi$.