Kapitel 5

Iterationsverfahren für nichtlineare Gleichungen

5.1 Einleitung

In diesem Kapitel betrachten wir das Problem, eine numerische Lösung x^* der Gleichung

$$F(x) = y \tag{5.1}$$

zu finden. Dabei ist $F : \mathcal{G} \subset \mathbb{R}^n \to \mathbb{R}^n$ eine (eventuell) nichtlineare Abbildung und $y \in \mathbb{R}^n$ vorgegeben. (5.1) ist also ein System von n nichtlinearen Gleichungen für die Unbekannten $(x_1, \ldots, x_n)$. Anders als im linearen Fall $F(x) = Ax$, $A \in \mathbb{R}^{n,n}$, wo die Lösung für reguläre A eindeutig durch $x^* = A^{-1}y$ angegeben werden kann, hat man für nichtlineare Gleichungen normalerweise keine allgemeinen Aussagen über Existenz und Eindeutigkeit. Ebenfalls im Gegensatz zum linearen Fall, wo die geläufigsten klassischen Verfahren *direkte Verfahren* sind, die bei Vernachlässigung von Rundungsfehlern nach endlich vielen Schritten die exakte Lösung liefern, ist man bei der Lösung von (5.1) außer in wenigen einfachen Fällen auf *iterative Verfahren* angewiesen. Das bedeutet, daß man eine Lösung x^* als den Grenzwert einer Folge $\{x_k\}$ erhält, die sich typischerweise aus einer Iteration der Form $x_{k+1} = \Phi(x_k, x_{k-1}, \ldots)$ ergibt, d.h. x_{k+1} berechnet sich aus einigen (oder sogar allen) vorher berechneten Folgengliedern.

Nur auf den ersten Blick Spezialfälle von (5.1) sind das *Nullstellenproblem* und das *Fixpunktproblem*: Suche jeweils $x^* \in \mathbb{R}^n$, welches das Problem

$$f(x) = 0 \tag{5.2}$$

bzw.

$$x = g(x) \tag{5.3}$$

löst. (5.1) Kann immer durch Äquivalenzumformungen auf (5.2) und (5.3) gebracht werden, z.B. durch

$$F(x) - y =: f(x) = 0 \Leftrightarrow f(x) + x =: g(x) = x.$$

Welche Form und welche Äquivalenzumformung einen guten Ansatz für einen numerischen Algorithmus bietet, hängt vom konkreten Problem ab.

Beispiel: Wir suchen eine Lösung x^* von $x^2 = 2$ und formen um: Für $x \neq 0$ gilt

$$x^2 = 2 \Leftrightarrow x^2 + x^2 = 2x^2 = 2 + x^2 \Leftrightarrow x = \frac{1}{x} + \frac{x}{2} =: g(x).$$

Das Problem ist jetzt in Fixpunktform. Wir werden zeigen, daß die Folge $\{x_k\}$ definiert durch

$$x_{k+1} = g(x_k), \quad x_0 = a \tag{5.4}$$

für alle $a > 0$ gegen $x^* = \sqrt{2} = 1.414213\ldots$ konvergiert. So erhalten wir z.B. für $a = 5.0$ bei 4-stelliger Rechnung

$$\{x_0, x_1, x_2, x_3, x_4, \ldots\} = \{5.0, 2.7, 1.72, 1.441, 1.415, \ldots\}.$$

x_4 ist schon eine sehr gute Approximation für x^*, das Verfahren konvergiert schnell. Eine andere, naheliegendere Umformung

$$x^2 = 2 \Leftrightarrow x = \frac{2}{x} =: g(x)$$

liefert dagegen bei der Iteration mit $x_0 = a$ die alternierende Folge $\{a, 2/a, a, 2/a, \ldots\}$, welche natürlich nicht (außer für $a = x^*$) gegen eine Lösung x^* konvergiert.

5.2 Fixpunktiteration

Wir haben an obigem Beispiel gesehen, daß geeignete Transformation auf (5.3) ein guter Ansatz zu einer numerischen Lösung sein kann. Der nun folgende klassische Fixpunktsatz von Banach liefert eine gut überprüfbare hinreichende Bedingung dafür, ob eine Iteration der Form (5.4) konvergiert.

Satz 5.1 (Banachscher Fixpunktsatz) *Es sei $\mathcal{V}$ ein Banachraum (d.h. ein vollständiger normierter Vektorraum), $\mathcal{D} \subset \mathcal{V}$ eine abgeschlossene Teilmenge von $\mathcal{V}$ und $g : \mathcal{D} \subset \mathcal{V} \to \mathcal{V}$ eine Abbildung mit $g(\mathcal{D}) \subset \mathcal{D}$ (d.h. $g(x) \in \mathcal{D}$ für alle $x \in \mathcal{D}$). Weiter gebe es ein $q \in [0, 1)$, so daß*

$$\|g(x) - g(y)\| \leq q\|x - y\| \quad \text{für alle } x, y \in \mathcal{D}. \tag{5.5}$$

Dann existiert genau ein $x^* \in \mathcal{D}$ mit $x^* = g(x^*)$, d.h. g besitzt in $\mathcal{D}$ genau einen Fixpunkt. Außerdem konvergiert für jeden Startvektor $x_0 \in \mathcal{D}$ die Folge $\{x_k\}$ definiert durch $x_{k+1} = g(x_k)$ gegen x^*, und es gilt die Fehlerabschätzung

$$\|x^* - x_k\| \leq \frac{q}{1-q}\|x_k - x_{k-1}\| \tag{5.6}$$

für $k \geq 1$.

Beweis: Zunächst gilt für $k \geq 1$:

$$\|x_{k+1} - x_k\| = \|g(x_k) - g(x_{k-1})\| \leq q\|x_k - x_{k-1}\| \leq \cdots \leq q^k \|x_1 - x_0\|.$$

Sei nun $\epsilon > 0$ und $i, j \in \mathbb{N}$. Dann gilt für $i > j$

$$\begin{aligned}
\|x_i - x_j\| &= \|(x_i - x_{i-1}) + (x_{i-1} - x_{i-2}) + \cdots + (x_{j+1} - x_j)\| \\
&= \left\| \sum_{k=i-1}^{j} x_{k+1} - x_k \right\| \leq \sum_{k=i-1}^{j} \|x_{k+1} - x_k\| \leq \sum_{k=i-1}^{j} q^k \|x_1 - x_0\| \\
&= \|x_1 - x_0\| q^{i-1} \sum_{k=0}^{j-i+1} q^k \leq \frac{\|x_1 - x_0\|}{1-q} q^{i-1} \leq \epsilon \quad \text{für } i, j > N,
\end{aligned}$$

wenn $N \in \mathbb{N}$ hinreichend groß gewählt wird. $\{x_k\}$ ist also eine Cauchy-Folge und besitzt demnach einen Grenzwert $x^* \in \mathcal{V}$. Da $\{x_k\} \subset \mathcal{D}$ und $\mathcal{D}$ abgeschlossen ist, gilt auch $x^* \in \mathcal{D}$. Daß x^* tatsächlich ein Fixpunkt von g ist, sieht man wie folgt: (5.5) impliziert offenbar die gleichmäßige Stetigkeit von g auf $\mathcal{D}$. Darum gilt:

$$x^* = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} g(x_k) = g(\lim_{k\to\infty} x_k) = g(x^*).$$

Zur Eindeutigkeit: Falls y^* ein weiterer Fixpunkt von g in $\mathcal{D}$ ist, erhalten wir

$$\|x^* - y^*\| = \|g(x^*) - g(y^*)\| \leq q\|x^* - y^*\|,$$

also $\|x^* - y^*\| = 0$ wegen $q \in [0, 1)$ und damit $x^* = y^*$.

Zur Fehlerabschätzung:

$$\|x^* - x_k\| = \|g(x^*) - g(x_{k-1})\| \leq q\|x^* - x_{k-1}\| \leq q(\|x^* - x_k\| + \|x_k - x_{k-1}\|),$$

also

$$(1-q)\|x^* - x_k\| \leq q\|x_k - x_{k-1}\|.$$

□

Bemerkung:

1. Wenn eine Funktion $g(x)$ die Eigenschaft (5.5) hat, sagt man, daß sie auf $\mathcal{D}$ *kontrahierend* ist. Die Kontraktionseigenschaft hängt von der verwendeten Vektornorm $\|\cdot\|$ ab. Es kann sein, daß eine Funktion kontrahierend ist bzgl. einer Vektornorm, aber nicht kontrahierend bzgl. einer anderen.

2. Satz 5.1 ist in sehr allgemeiner Form formuliert: $\mathcal{V}$ kann ein beliebiger vollständiger normierter Vektorraum sein, nicht notwendig endlichdimensional. Damit ist Satz 5.1 auch z.B. auf Funktionenräume wie $C[a, b]$ anwendbar, wie wir später sehen werden. Wir werden aber zunächst nur den Fall $\mathcal{V} = \mathbb{R}^n$, speziell $\mathcal{V} = \mathbb{R}$ betrachten.

3. Die Folgen-Indizes k von $\{x_k\}$ sind nicht zu verwechseln mit den Vektorkomponenten von $x \in \mathbb{R}^n$! Weil einzelne Vektorkomponenten in diesem Kapitel jedoch kaum eine Rolle spielen, erscheint diese nicht ganz konsequente Bezeichnung vertretbar.

Beispiel: Wir wollen die Iterationsfunktion $g(x) = x^{-1} + x/2$ zur Bestimmung von $\sqrt{2}$ mit Satz 5.1 untersuchen. In diesem Fall ist $\mathcal{V} = \mathbb{R}$, und wir setzen $\mathcal{D} = [1, \infty)$. $\mathcal{D}$ ist offensichtlich abgeschlossen. Wegen $(x-1)^2 = x^2 - 2x + 1 > -1 \Leftrightarrow x^2 + 2 > 2x$ ist

$$g(x) = \frac{x^2+2}{2x} \geq 1 \quad \text{für alle } x \in \mathbb{R},$$

also $g(\mathcal{D}) \subset \mathcal{D}$. Außerdem ist für $x, y \in \mathcal{D}$ auch $xy \in \mathcal{D}$ und somit

$$-\frac{1}{2} \leq \frac{1}{2} - \frac{1}{xy} = \frac{g(x) - g(y)}{x - y} \leq \frac{1}{2},$$

also

$$|g(x) - g(y)| \leq \frac{1}{2}|x - y|.$$

$g(x)$ erfüllt damit auf $\mathcal{D}$ die Voraussetzungen von Satz 5.1, hat also in $[1, \infty)$ genau einen Fixpunkt x^*, und die Folge (5.4) konvergiert für alle $x_0 \in [1, \infty)$ gegen x^*. Für $x \in (0, 1)$ ist ebenfalls $g(x) \in \mathcal{D}$, so daß insgesamt (5.4) für alle $x_0 \in (0, \infty)$ gegen x^* konvergiert. Man beachte, daß der Satz nicht direkt auf $(0, \infty)$ anwendbar ist, da $(0, \infty)$ nicht abgeschlossen ist und $g(x)$ nicht auf ganz $(0, \infty)$ kontrahierend ist.

Der direkte Beweis der Kontraktionseigenschaft, wie er in obigem Beispiel geführt wurde, ist in vielen Fällen mühselig. Der folgende Satz liefert ein hinreichendes Kriterium dafür, ob eine Funktion kontrahierend ist.

Satz 5.2 *Es sei $\mathcal{D} \subset \mathbb{R}^n$ abgeschlossen und konvex (d.h. mit $x, y \in \mathcal{D}$ liegen auch alle Punkte auf der Verbindungsstrecke zwischen x und y in $\mathcal{D}$), $g : \mathcal{D} \to \mathbb{R}^n$ stetig differenzierbar und*

$$q := \sup_{x \in \mathcal{D}} \|g'(x)\| < 1. \tag{5.7}$$

Dann ist $g(x)$ auf $\mathcal{D}$ kontrahierend mit der Kontraktionskonstanten q.

Beweis: Wir beweisen nur die Aussage für den Fall $n = 1$. In diesem Fall ist der Beweis einfach, da der Mittelwertsatz gilt: $\mathcal{D}$ ist ein abgeschlossenes Intervall, und für alle $x, y \in \mathcal{D}$ gibt es ein $\xi \in \mathcal{D}$ mit $g(x) - g(y) = g'(\xi)(x - y)$. Daraus folgt

$$|g(x) - g(y)| = |g'(\xi)(x - y)| \leq \sup_{\xi \in \mathcal{D}} |g'(\xi)| \cdot |x - y|.$$

$\square$

Bemerkung: Für $n > 1$ ist $g'(x) \in \mathbb{R}^{n,n}$ die Jacobimatrix von g im Punkt $x \in \mathbb{R}^n$. Die Norm in (5.7) ist also eine Matrixnorm, gemeint ist hier die von der verwendeten Vektornorm induzierte Matrixnorm.

Mit Satz (5.2) erhält man unmittelbar, daß $g(x) = 1/x + x/2$ auf $[1, \infty)$ kontrahierend ist, denn es gilt

$$-\frac{1}{2} \leq g'(x) = \frac{1}{2} - \frac{1}{x^2} \leq \frac{1}{2}$$

für alle $x \geq 1$.

Sehr nützlich ist auch folgender Satz:

Satz 5.3 (Lokaler Konvergenzsatz) *Es sei $\mathcal{G} \subset \mathbb{R}^n$ offen, $g : \mathcal{G} \to \mathbb{R}^n$ stetig differenzierbar und $g(x^*) = x^*$ mit $\|g'(x^*)\| < 1$ für ein $x^* \in \mathcal{G}$. Dann gibt es ein $r > 0$, sodaß für alle $x_0 \in \mathcal{D} := \overline{\mathcal{B}_r(x^*)}$ (abg. Kugel um x^* mit Radius r) die Iteration $x_{k+1} = g(x_k)$ gegen x^* konvergiert.*

Beweis: Es gibt ein $r > 0$ und ein $q \in [0, 1)$ mit

$$q = \sup_{x \in \overline{\mathcal{B}_r(x^*)}} \|g'(x)\|.$$

g ist damit nach Satz 5.2 auf $\mathcal{D} = \overline{\mathcal{B}_r(x^*)}$ kontrahierend mit Konstante q. Für $x \in \mathcal{D}$ gilt dann

$$\|g(x) - x^*\| = \|g(x) - g(x^*)\| \leq q\|x - x^*\| \leq r,$$

also $g(x) \in \mathcal{D}$. Damit ist gezeigt worden: $g(\mathcal{D}) \subset \mathcal{D}$. Satz 5.1 liefert jetzt für jeden Startvektor $x_0 \in \mathcal{D}$ Konvergenz der Iteration $x_{k+1} = g(x_k)$ gegen $x^* \in \mathcal{D}$. $\square$

5.3 Einige klassische Iterationsverfahren in $\mathbb{R}$

Das Newton-Verfahren

Das Newton-Verfahren in $\mathbb{R}$ ist wohl das bekannteste Verfahren zur Bestimmung einer Nullstelle einer differenzierbaren Funktion. Die Iterationsvorschrift ist

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} =: g(x_k), \tag{5.8}$$

vorausgesetzt, $f'(x_k) \neq 0$. Ist x^* eine Nullstelle von $f(x)$, so erhalten wir mit

$$g'(x) = 1 - \frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2} \tag{5.9}$$

sofort $g'(x^*) = 0$, falls $f'(x^*) \neq 0$. In diesem Fall garantiert Satz 5.3 die lokale Konvergenz des Newtonverfahrens, d.h. für Startwerte x_0, die schon nahe genug bei der Nullstelle x^* liegen. Daß man i.A. keine *globale* Konvergenz, d.h. Konvergenz für beliebige Startwerte erwarten kann, zeigen folgende

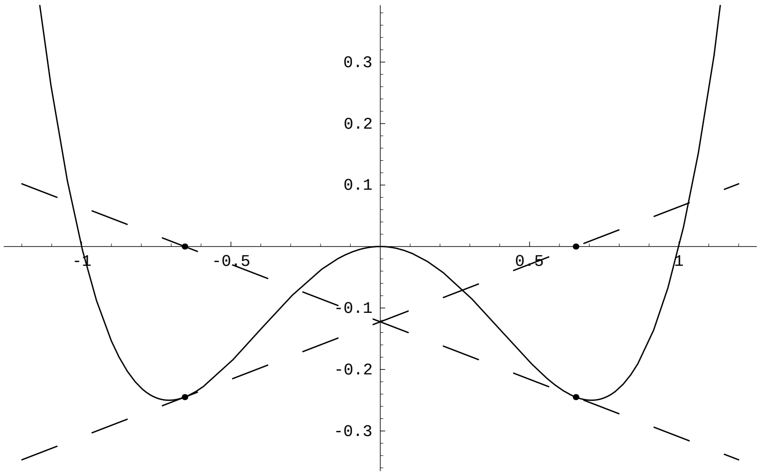

Abbildung 5.1: Beispiel für Nichtkonvergenz des Newtonverfahrens: Die Werte $\pm\sqrt{3/7}$ werden durch das Newtonverfahren wechselseitig aufeinander abgebildet

Beispiele

1. Beim Newtonverfahren zur Bestimmung einer Nullstelle von $f(x) = x^2(x^2-1)$ erhält man mit dem Startwert $x_0 = \sqrt{3/7} \simeq 0.6547$ die alternierende Folge $\{x_0, -x_0, x_0, -x_0, \ldots\}$, siehe dazu Abb. 5.1.

2. Das Newtonverfahren läßt sich analog auch in $\mathbb{C}$ durchführen. Will man z.B. eine der 3 komplexen Nullstellen von $f(z) = z^3 - 1$ bestimmen, so lautet die Newtoniteration
$$z_{k+1} = z_k - \frac{z_k^3 - 1}{3z_k^2}.$$
Man kann sich nun fragen, wie die Menge der Startwerte $z_0 \in \mathbb{C}$ aussieht, für die das Verfahren z.B. gegen die reelle Nullstelle $z^* = 1$ konvergiert. Es stellt sich heraus, daß diese Menge ein Fraktal ist, siehe Abb. 5.2.

Die Herleitung des Newtonverfahrens erfolgt gewöhnlich durch graphische Veranschaulichung, indem man an einer Stelle x_0 mit $f'(x_0) \neq 0$ in der Nähe einer Nullstelle x^* den Funktionsgraphen von f durch seine Tangente ersetzt und als Näherung x_1 den Schnittpunkt der Tangente mit der x-Achse nimmt. Das bedeutet gerade, daß man f durch Abbrechen der Taylorreihe um x_0 nach dem 1. Glied approximiert:
$$0 = f(x^*) \simeq f(x_0) + f'(x_0)(x^* - x_0) \Leftrightarrow x^* \simeq x_0 - \frac{f(x_0)}{f'(x_0)} =: g(x_0),$$
was die Wahl $x_1 = g(x_0)$ nahelegt. Man kann auch die Taylorreihe erst nach dem 2. Glied abbrechen und erhält im Falle $f''(x_0) \neq 0$
$$\begin{aligned} 0 = f(x^*) &\simeq f(x_0) + f'(x_0)(x^* - x_0) + \frac{f''(x_0)}{2}(x^* - x_0)^2 \\ \Leftrightarrow x^* &\simeq x_0 - \frac{f'(x_0) \pm \sqrt{f'(x_0)^2 - 2f(x_0)f''(x_0)}}{f''(x_0)} =: g_\pm(x_0), \end{aligned}$$

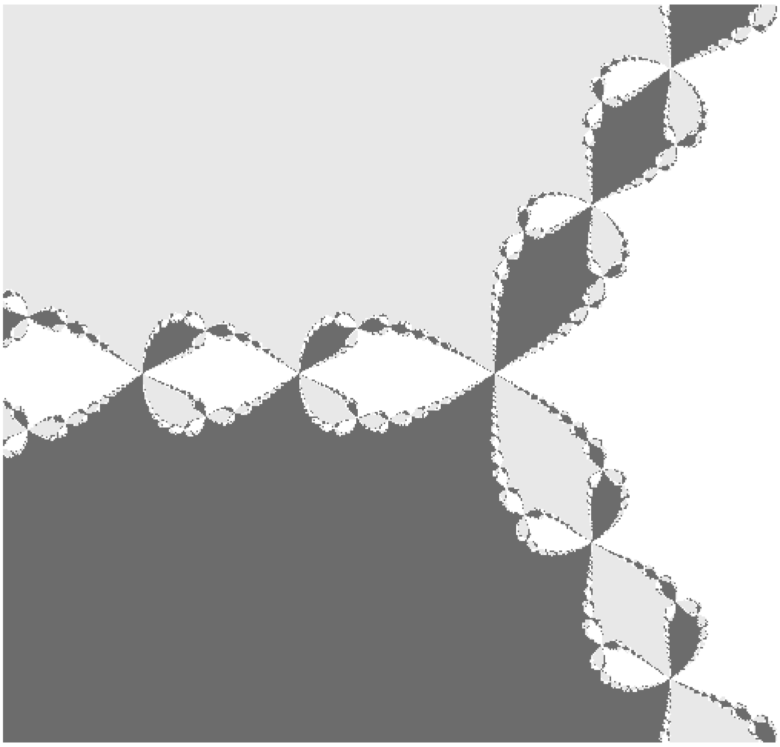

Abbildung 5.2: Der Einzugsbereich des komplexen Newtonverfahrens zur Bestimmung einer Wurzel von $z^3 = 1$. Der Bereich von Startwerten, für den das Verfahren gegen 1 konvergiert, ist weiß, die Bereiche der übrigen beiden Wurzeln hell- bzw. dunkelgrau. Plotbereich: $[-2, 2] \times [-1.5, 1.5]$.

was wiederum die Wahl $x_1 = g_-(x_0)$ oder $x_1 = g_+(x_0)$ nahelegt. Das Vorzeichen ergibt sich daraus, welcher von beiden Werten die Gleichung $f(x) = 0$ besser erfüllt. Dies Verfahren ist das sogenannte *Newton-Verfahren 2. Grades*. Beide Newtonverfahren sind in Abb. 5.3 dargestellt.

Das Sekantenverfahren

Das Sekantenverfahren kann als eine Approximation des Newtonverfahrens (5.8) betrachtet werden. Anstatt eine Tangente an den Graphen von f in einem Punkt $f(x_0)$ zu legen, legt man eine Sekante durch zwei Punkte $f(x_0)$ und $f(x_1)$ und nimmt als neuen Wert x_2 den Schnittpunkt der Sekante mit der x-Achse, siehe Abb. 5.4. Dann macht man dasselbe mit x_1 und x_2 usw. Das Verfahren bietet sich z.B. dann an, wenn die Ableitung von f schwierig auszuwerten ist. Die Formel zur Berechnung von x_{k+1} bei gegebenem x_k und x_{k-1} lautet

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}. \tag{5.10}$$

Man beachte die formale Ähnlichkeit zum Newtonverfahren (5.8). Für nahe beieinanderliegende Werte x_{k-1}, x_k ist der Bruch auf der rechten Seite von (5.10) eine gute Approximation zu $f'(x_k)^{-1}$.

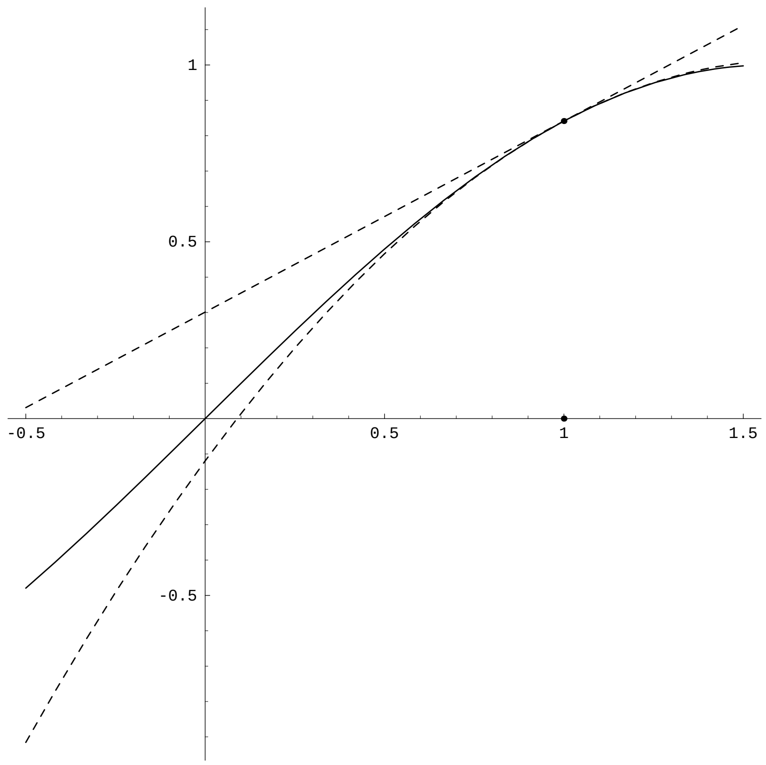

Abbildung 5.3: Die Funktion $f(x) = \sin(x)$ (durchgehende Linie) mit ihren Approximationen 1. und 2. Ordnung (gestrichelte Linien) im Punkt $x_0 = 1$.

Intervallschachtelung

Die Idee bei der Intervallschachtelung ist sehr einfach. Angenommen, wir haben ein Intervall $[a_k, b_k]$ mit $f(a_k) < 0$ und $f(b_k) > 0$. (Der umgekehrte Fall $f(a_k) > 0$ und $f(b_k) < 0$ ist analog zu behandeln). Wenn $f(x)$ stetig ist, muß nach dem Zwischenwertsatz f in $[a_k, b_k]$ mindestens eine Nullstelle besitzen. Wir setzen dann $c_k = (a_k + b_k)/2$ (der Intervallmittelpunkt). Falls $c_k = 0$, sind wir fertig. Ansonsten nehmen wir als neues Intervall

$$[a_{k+1}, b_{k+1}] = \begin{cases} [a_k, c_k] : & f(c_k) > 0 \\ [c_k, b_k] : & f(c_k) < 0 \end{cases}. \tag{5.11}$$

Offensichtlich hat f an den Intervallenden a_{k+1}, b_{k+1} ebenfalls entgegengesetzte Vorzeichen, sodaß $[a_{k+1}, b_{k+1}]$ eine Nullstelle von f enthalten muß. Insgesamt erhält man eine Folge von kompakten Intervallen $\{[a_k, b_k]\}$, deren Länge sich in jedem Schritt halbiert. Der Durchschnitt $\cap_k [a_k, b_k]$ aller Intervalle enthält genau einen Punkt, nämlich eine gesuchte Nullstelle.

5.4 Die Konvergenzordnung eines Iterationsverfahrens

Die Konvergenzordnung eines Verfahrens kann als ein Maß betrachtet werden, wie schnell das Verfahren gegen eine Lösung konvergiert.

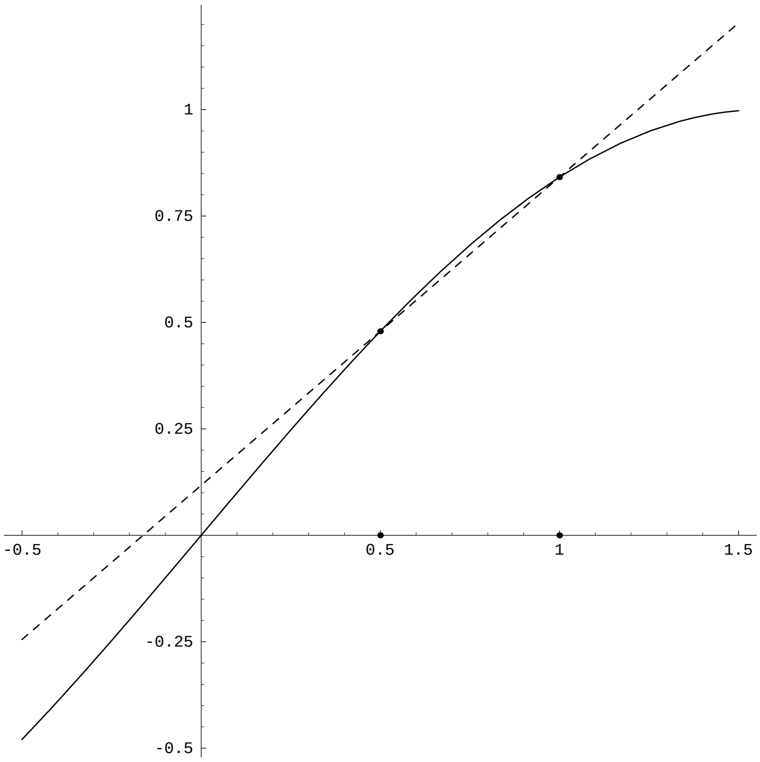

Abbildung 5.4: Das Sekantenverfahren für die Funktion $f(x) = \sin(x)$ zu den Startwerten $x_0 = 1$, $x_1 = 1/2$.

Definition 5.1 (Konvergenzordnung) *Es sei* $\{x_k\} \subset \mathcal{V}$ *eine Folge in einem normierten Vektorraum mit* $\lim_{k\to\infty} x_k = x^*$ *für ein* $x^* \in \mathcal{V}$. *Es gebe positive reelle Zahlen* p, C *und* $N \in \mathbb{N}$ *mit*

$$\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|^p \tag{5.12}$$

für alle $k \geq N$. *Dann heißt* p *die* (lokale) Konvergenzordnung *der Folge* $\{x_k\}$ *bzw. eines Iterationsverfahrens* $x_{k+1} = g(x_k)$, *welches die Folge* $\{x_k\}$ *definiert.*

Bemerkung: Die Konvergenzordnung sagt im wesentlichen, mit welcher Potenz der absolute Fehler $\epsilon_k := x_k - x^*$ sich in jedem Iterationsschritt verringert. (5.12) ist gleichbedeutend damit, daß $\|\epsilon_{k+1}\| = O(\|\epsilon_k\|^p)$. Für den Fall $p = 1$ und $C < 1$ erhält man z.B., daß der Fehler sich in jedem Schritt um den Faktor C verringert. In diesem Fall spricht man von *linearer Konvergenz*, im Fall $p = 2$ von *quadratischer Konvergenz*.

Für die wichtige Klasse der Fixpunktiterationen läßt sich folgendes zeigen:

Satz 5.4 (Konvergenzordnung von Fixpunktiterationen) *Es sei* $\mathcal{G} \subset \mathbb{R}$ *ein offenes Intervall,* $p \in \mathbb{N}$, $g : \mathcal{G} \to \mathbb{R}$ p *mal stetig differenzierbar und* $x^* \in \mathcal{G}$ *mit*

$$g(x^*) = x^*, \quad g^{(i)}(x^*) = 0 \quad \textit{für } i = 1, \ldots, p-1.$$

Dann ist das Iterationsverfahren $x_{k+1} = g(x_k)$ zur Bestimmung von x^* mindestens von der Konvergenzordnung p.

Beweis: Wir stellen $g(x_k)$ um $x = x^*$ mit der Formel von Taylor dar:

$$\begin{aligned} x_{k+1} = g(x_k) &= g(x^*) + g'(x^*)(x_k - x^*) + \cdots + \frac{g^{(p-1)}(x^*)}{(p-1)!}(x_k - x^*)^{p-1} + \frac{g^{(p)}(\xi)}{p!}(x_k - x^*)^p \\ &= x^* + \frac{g^{(p)}(\xi)}{p!}(x_k - x^*)^p \end{aligned}$$

nach Voraussetzung, wobei ξ zwischen x_k und x^* liegt. Hieraus folgt

$$|x_{k+1} - x^*| = O(|x_k - x^*|^p).$$

□

Als Anwendung hiervon erhalten wir:

Satz 5.5 (Konvergenzordnung des Newtonverfahrens) *Es sei $f : \mathbb{R} \to \mathbb{R}$ in einer offenen Umgebung einer Nullstelle x^* zweimal stetig differenzierbar. Das Newtonverfahren (5.8) zur Bestimmung von x^* konvergiert (lokal) quadratisch, falls $f'(x^*) \neq 0$. Für $f'(x^*) = 0$, aber $f''(x^*) \neq 0$ konvergiert es noch linear.*

Beweis: Für $f'(x^*) \neq 0$ folgt $g'(x^*) = 0$ unmittelbar aus (5.9). Die Voraussetzungen für Satz 5.4 sind also für $p = 2$ erfüllt.
Für $f'(x^*) = 0$, aber $f''(x^*) \neq 0$ ist der Beweis etwas aufwendiger: Für x_k in der Nähe der Nullstelle x^* gilt

$$\begin{aligned} f(x_k) &= f(x^*) + f'(x^*)(x_k - x^*) + \frac{f''(\xi_1)}{2}(x_k - x^*)^2, \\ f'(x_k) &= f'(x^*) + f''(\xi_2)(x_k - x^*), \end{aligned}$$

wobei ξ_1 und ξ_2 zwischen x_k und x^* liegen. Wegen $f''(x^*) \neq 0$ können wir k so groß wählen, daß x_k derart nahe bei x^* liegt, daß $f''(x)$ an den Stellen ξ_1 und ξ_2 von Null verschieden ist und daß

$$\frac{1}{2} < \frac{f''(\xi_1)}{f''(\xi_2)} \leq \frac{3}{2}$$

gilt. Daraus erhalten wir

$$x_{k+1} - x^* = x_k - x^* - \frac{f(x_k)}{f'(x_k)} = x_k - x^* - \frac{f''(\xi_1)}{2f''(\xi_2)}(x_k - x^*) = \left(1 - \frac{f''(\xi_1)}{2f''(\xi_2)}\right)(x_k - x^*)$$

und mit dem bisher gezeigten

$$|x_{k+1} - x^*| \leq \frac{3}{2}|x_k - x^*|.$$

□

Bemerkung:

1. Dieses Ergebnis bestätigt die Erfahrung, daß mehrfache Nullstellen numerisch schwieriger zu bestimmen sind als einfache Nullstellen.

2. Ähnliche Ergebnisse lassen sich auch für die anderen in diesem Kapitel vorgestellten Verfahren herleiten: Das Intervallhalbierungsverfahren konvergiert offensichtlich linear mit $C = 1/2$. Die Konvergenzordnung des Sekantenverfahrens ist $p = (\sqrt{5}+1)/2 \simeq 1.62$. Hierfür ist der Beweis aber schwieriger.

5.5 Nullstellen von Polynomen: Das Verfahren von Laguerre

Das Verfahren von Laguerre ist eine der besten Methoden zur Bestimmung der Nullstellen eines Polynoms $p(x)$. Für den Fall, daß

$$p(x) = (x-\alpha_1)(x-\alpha_2)\dots(x-\alpha_n) = \prod_{i=1}^{n}(x-\alpha_i) \tag{5.13}$$

mit paarweise verschiedenen Nullstellen α_i, werden wir zeigen, daß das Verfahren global (d.h. für jeden Startwert) mit der Ordnung 3 gegen eine Nullstelle α_i konvergiert. Eine exakte Herleitung des Verfahrens erfordert erheblichen Aufwand. Wir beschränken uns auf eine heuristische Herleitung. Dazu bilden wir zunächst den Logarithmus von (5.13) und leiten zweimal ab:

$$\ln|p(x)| = \sum_{i=1}^{n} \ln|x-\alpha_i|, \tag{5.14}$$

$$\frac{d}{dx}\ln|p(x)| = \sum_{i=1}^{n}(x-\alpha_i)^{-1} = \frac{p'(x)}{p(x)} =: a, \tag{5.15}$$

$$-\frac{d^2}{dx^2}\ln|p(x)| = \sum_{i=1}^{n}(x-\alpha_i)^{-2} = \frac{p'(x)^2 - p''(x)p(x)}{p(x)^2} =: b. \tag{5.16}$$

Nehmen wir nun ein x nahe bei der zu bestimmenden ersten Nullstelle α_1 und setzen $a_i := (x-\alpha_i)^{-1}$ für $i = 1,\dots,n$. Wenn die Abstände zwischen allen Nullstellen ungefähr gleich sind, bedeutet dies, daß die Verhältnisse a_1/a_i für $i = 2,\dots,n$ ungefähr dieselbe Größenordnung haben. Wir nehmen also an, daß $a_2 = a_3 = \dots = a_n$ gilt. Aus (5.15) und (5.16) erhalten wir dann das Gleichungssystem

$$a_1 + (n-1)a_2 = a, \quad a_1^2 + (n-1)a_2^2 = b.$$

Wir eliminieren hieraus a_2, indem wir die zweite Gleichung mit $n-1$ multiplizieren und von der ersten Gleichung $(n-1)a_2 = a - a_1$ einsetzen. Das Resultat ist

$$na_1^2 - 2aa_1 + a^2 - (n-1)b = 0,$$

eine quadratische Gleichung für a_1, die formal nach

$$a_1 = \frac{a \pm \sqrt{(n-1)(nb-a^2)}}{n} = \frac{p'(x) \pm \sigma(x)}{np(x)} \tag{5.17}$$

aufgelöst werden kann, wobei wir zur Abkürzung

$$\sigma(x) := \Big\{(n-1)[(n-1)p'(x)^2 - np(x)p''(x)]\Big\}^{1/2}$$

gesetzt haben. Dabei haben wir für a und b die Terme gemäß (5.15) und (5.16) eingesetzt, in denen p, p', p'' vorkommt. Wenn wir a_1 und x kennen, können wir α_1 bestimmen, es gilt $\alpha_1 = x - a_1^{-1}$. Für a_1 in diesem Ausdruck können wir jetzt (5.17) einsetzen. Insgesamt erhalten wir aus (5.15), (5.16) und (5.17) also:

$$\alpha_1 = x - \frac{1}{a_1} = x - \frac{np(x)}{p'(x) \pm \sigma(x)} =: g_\pm(x). \tag{5.18}$$

Für $x \to \alpha_1$ erhalten wir, daß α_1 ein Fixpunkt von $g_\pm$ ist, was die Iteration $x_{k+1} = g_\pm(x_k)$ nahelegt.

Wir werden zeigen, daß das Verfahren tatsächlich wohldefiniert werden kann (d.h. eine sinnvolle Entscheidung g_+ oder g_- ist möglich, alle Wurzeln sind reell, keine Nullen im Nenner treten auf etc.) und global konvergiert. Dazu zunächst ein Hilfssatz:

Satz 5.6 *Ist $p(x)$ von der Form (5.13) mit n einfachen reellen Nullstellen $\alpha_1 < \alpha_2 < \cdots < \alpha_n$ und ist $x \in (\alpha_j, \alpha_{j+1})$, so gilt $\sigma(x) \in \mathbb{R}_+$, $p'(x) \pm \sigma(x) \neq 0$ (d.h. $g_\pm(x) \in \mathbb{R}$) und*

$$\begin{aligned} \alpha_j \le g_+(x) < x < g_-(x) \le \alpha_{j+1}, \quad &\textit{falls } p(x) > 0, \\ \alpha_j \le g_-(x) < x < g_+(x) \le \alpha_{j+1}, \quad &\textit{falls } p(x) < 0. \end{aligned}$$

Beweis: Zunächst leiten wir eine Hilfsungleichung her: Es seien allgemein

$$a_1, \ldots, a_n \in \mathbb{R}, \quad a := \sum_{i=1}^n a_i, \quad b := \sum_{i=1}^n a_i^2, \tag{5.19}$$

dann gilt

$$(a-a_1)^2 = \left(\sum_{i=2}^n a_i\right)^2 = \left(\sum_{i=2}^n a_i \cdot 1\right)^2 \le \left(\sum_{i=2}^n a_i^2\right) \cdot \left(\sum_{i=2}^n 1^2\right) = (n-1)\sum_{i=2}^n a_i^2 = (n-1)(b-a_1^2).$$

Dabei haben wir die Cauchy-Schwarzsche Ungleichung benutzt. Daraus folgt, daß

$$(a-a_1)^2 - (n-1)(b-a_1^2) = na_1^2 - 2a_1 a + a^2 - (n-1)b \le 0.$$

Das Polynom $q(z) := nz^2 - 2za - (n-1)b$ zweiten Grades hat also zwei reelle Nullstellen, denn der höchste Koeffizient ist n, und $q(a_1) \le 0$. Bezeichnen wir diese Nullstellen mit z_1 und z_2, dann liegt a_1 zwischen diesen beiden:

$$z_1 = \frac{a - \sqrt{(n-1)(nb-a^2)}}{n} \le a_1 \le \frac{a + \sqrt{(n-1)(nb-a^2)}}{n} = z_2. \tag{5.20}$$

Da z_1 und z_2 nur von a und b abhängen, kann man (5.20) für alle a_i aus (5.19) herleiten. Insbesondere gilt (5.20) für die speziellen a_i, a und b aus (5.15) und (5.16), da diese von der Form (5.19) sind. Die Ungleichung (5.20) ergibt für diese, falls $p(x) > 0$:

$$z_1 = \frac{p'(x) - \sigma(x)}{np(x)} \le (x - \alpha_i)^{-1} \le \frac{p'(x) + \sigma(x)}{np(x)} = z_2, \quad i = 1, \dots, n. \tag{5.21}$$

Man beachte, daß mit den Wurzeln in (5.20) auch $\sigma(x) \in \mathbb{R}_+$ ist.

Jetzt folgt aus (5.21) und $\alpha_j < x < \alpha_{j+1}$ das Ergebnis: Für $p(x) > 0$ gilt

$$0 < (x - \alpha_j)^{-1} \le \frac{p' + \sigma}{np} \Rightarrow 0 < \frac{np}{p' + \sigma} \le x - \alpha_j \Rightarrow \alpha_j \le x - \frac{np}{p' + \sigma} = g_+ < x,$$

$$\frac{p' - \sigma}{np} \le (x - \alpha_{j+1})^{-1} < 0 \Rightarrow x - \alpha_{j+1} \le \frac{np}{p' - \sigma} < 0 \Rightarrow x < x - \frac{np}{p' - \sigma} = g_- \le \alpha_{j+1}.$$

Für den Fall $p(x) < 0$ kehrt sich die Ungleichung (5.21) um, und man verfährt analog. □

Auswahl des Vorzeichens von $g_\pm$

Die Aussage von 5.6 ist in Abb. 5.5 dargestellt. Wenn z.B. (wie im ersten Teil der Abbildung) $p(x) > 0$ und $p'(x) > 0$ ist, liegt es nahe zu vermuten, daß x näher zur linken Nullstelle α_j als zur rechten Nullstelle α_{j+1} gelegen ist. In diesem Fall würde man dasjenige $g(x)$ als neue Approximation nehmen, welches zwischen α_j und x liegt, also in diesem Fall $g_+(x)$. Für den Fall $p(x) < 0$ und $p'(x) < 0$ (wie im zweiten Teil der Abbildung) würde man $g_-(x)$ nehmen. Die übrigen Fälle würde man analog betrachten. Insgesamt sieht man, daß die Auswahl g_+ oder g_- nicht vom Vorzeichen von $p(x)$, sondern nur vom Vorzeichen von $p'(x)$ abhängt. Daher erscheint folgende Iterationsfunktion sinnvoll:

$$x_{k+1} = \begin{cases} g_+(x): & \text{falls } p'(x) \ge 0 \\ g_-(x): & \text{falls } p'(x) < 0 \end{cases}$$

Wir werden nun zeigen, daß diese Iteration für jeden Startwert x_0 immer gegen eine Nullstelle von $p(x)$ konvergiert.

Satz 5.7 (Laguerre) *Das Verfahren $x_{k+1} = g(x_k)$ mit*

$$g(x) = x - \frac{np(x)}{p'(x) + \textbf{\textit{sign}}(p'(x)) \cdot \sigma(x)} \tag{5.22}$$

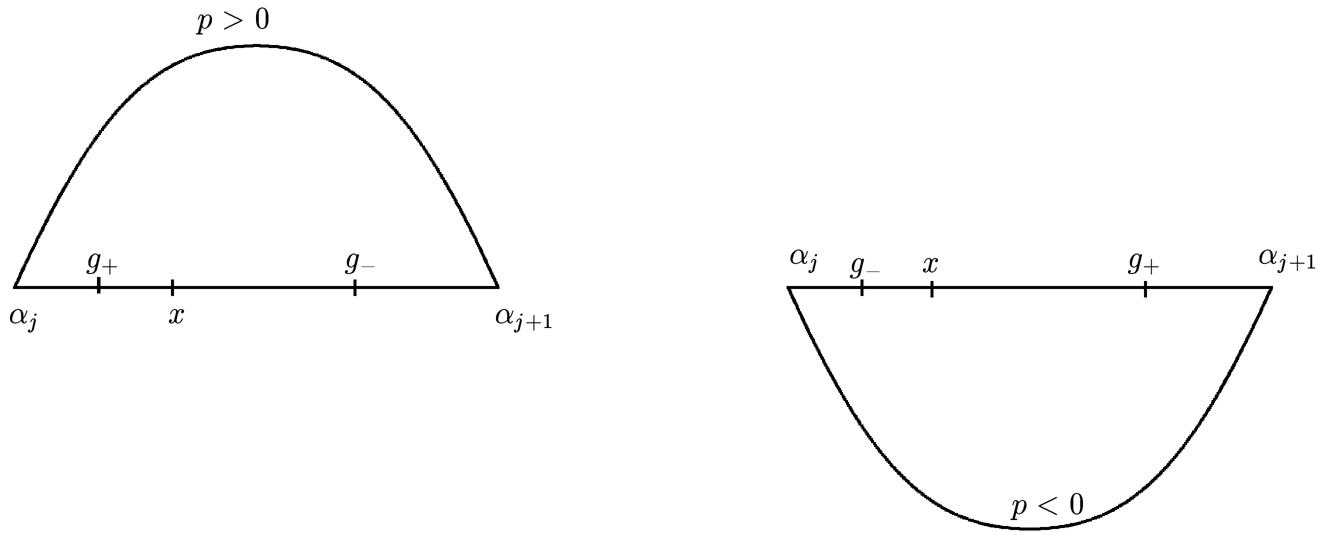

Abbildung 5.5: Schematische Darstellung der Aussage von Satz (5.6)

konvergiert für alle Startwerte x_0 mit der Konvergenzordnung 3 gegen eine Nullstelle von $p(x)$ (sign(0)=1).

Beweis: Zum Beweis der Konvergenz: Man betrachte hierzu Abb. 5.5. Es können im wesentlichen 3 Fälle auftreten: $\alpha_j < x_0 < \alpha_{j+1}$, $x_0 > \alpha_n$ oder $x_0 < \alpha_1$. Den Fall $x_0 = \alpha_j$ brauchen wir nicht zu betrachten.

1. Für den ersten Fall nehmen wir an, daß $p(x_0) > 0$ und $p'(x_0) \geq 0$. Dann ist $\alpha_j < x_1 < x_0$ mit $p(x_1) > 0$ und $p'(x_1) > 0$, also $\alpha_j < x_2 < x_1 < x_0$ usw. Insgesamt sieht man, daß $\{x_k\}$ eine streng monoton fallende Folge ist, die nach unten durch α_j beschränkt ist, also gegen ein $x^* \geq \alpha_j$ konvergiert. Wegen $x^* = \lim x_k = \lim g(x_k)$ muß dieses x^* dann ein Fixpunkt von $g(x)$ sein. Wegen $g(x^*) = x^* \Leftrightarrow p(x^*) = 0$ ist x^* also eine Nullstelle von $p(x)$. Da offensichtlich $x^* < \alpha_{j+1}$, folgt $x^* = \alpha_j$.
   Die übrigen drei Unterfälle werden genauso behandelt.

2. Falls $x_0 > \alpha_n$, sind $p(x)$ und $p'(x)$ entweder beide positiv oder beide negativ. In jedem Fall gilt $\alpha_n < x_1 < x_0$, und der Konvergenzbeweis verläuft analog zu 1).

3. Der Fall $x_0 < \alpha_1$ ist analog zu beweisen.

Zum Beweis der Konvergenzordnung benutzt man Satz 5.4, was auf eine Berechnung der ersten Ableitungen von $g(x)$ hinausläuft (zur Übung). □

Bemerkung: Die Anwendung des Verfahrens ist auch nicht auf den Fall einfacher Nullstellen beschränkt. In vielen Fällen mehrfacher Nullstellen liefert es durchaus brauchbare Ergebnisse. Allerdings ist es in diesen Fällen schwieriger, globale Konvergenz zu beweisen.

```
p2 = 0.0; p1 = 0.0; p0 = c(n);
do k=n-1,0,-1
  p2 = x*p2 + 2*p1
  p1 = x*p1 + p0
  p0 = x*p0 + c(k)
end do
```

Algorithmus 5.1: Auswertung von p und seiner ersten beiden Ableitungen an einer Stelle x. Nach Ausführung enthält p0 $p(x)$, p1 $p'(x)$ und p2 $p''(x)$.

Es gibt auch komplexe Erweiterungen des Verfahrens zur Berechnung von nicht-reellen Nullstellen.

Effiziente Auswertung von Polynomen

Wir nehmen an, daß ein Polynom $p \in \mathcal{P}_n$ in der Form

$$p(x) = \sum_{k=0}^{n} c_k x^k \tag{5.23}$$

gegeben ist (wäre es in der Form (5.13) gegeben, brauchte man die Nullstellen nicht mehr ausrechnen). Für die ersten beiden Ableitungen von $p(x)$ gilt

$$p'(x) = \sum_{k=1}^{n} k c_k x^{k-1} = \sum_{k=0}^{n-1} (k+1) c_{k+1} x^k, \tag{5.24}$$

$$p''(x) = \sum_{k=2}^{n} k(k-1) c_k x^{k-2} = \sum_{k=0}^{n-2} (k+1)(k+2) c_{k+2} x^k. \tag{5.25}$$

Beim Verfahren von Laguerre müssen die ersten beiden Ableitungen wiederholt an verschiedenen Stellen ausgewertet werden. Man überlegt sich leicht, daß z.B. eine Berechnung von $p(x)$ durch die (naheliegende) Zuweisung (hier für $n = 3$)

```
p = c(0) + c(1)*x + c(2)*x*x + c(3)*x*x*x
```

sehr ineffizient ist. Stattdessen verwendet man üblicherweise das sogenannte *Hornerschema*, welches eine Darstellung von $p(x)$ in der Form

$$p(x) = ((c_3 x + c_2)x + c_1)x + c_0$$

(hier wieder für $n = 3$) ausnutzt. Algorithmus 5.1 liefert die Auswertung von p an einer Stelle x, und die Auswertungen von p' und p'' gleich mit:

Satz 5.8 *Bei Vernachlässigung von Rundungsfehlern enthält nach Ausführung von Alg. 5.1 p0 den Wert von $p(x)$, p1 den Wert von $p'(x)$ und p2 den Wert von $p''(x)$.*

Beweis: Wir zeigen die Richtigkeit der Behauptung mit Induktion über n. Für $n = 0$ wird die Schleife nicht durchlaufen, und die Behauptung ist offensichtlich richtig. Sei die Behauptung nun für $n-1$ bewiesen. Man überlegt sich leicht, daß Algorithmus 5.1 äquivalent ist zu folgendem Algorithmus:

```
1  p2 = 0.0; p1 = 0.0; p0 = c(n);
2  do k=n-2,0,-1
3    p2 = x*p2 + 2*p1
4    p1 = x*p1 + p0
5    p0 = x*p0 + c(k+1)
6  end do
7  p2 = x*p2 + 2*p1
8  p1 = x*p1 + p0
9  p0 = x*p0 + c(0)
```

(der letzte Schleifendurchlauf für $k = 0$ wurde einfach separat geschrieben). Nach Induktionsvoraussetzung und (5.23)–(5.25) enthalten p0, p1 und p2 in Zeile 7 (also nach Ablauf der do-Schleife) folgende Werte:

$$\mathrm{p0} = \sum_{k=0}^{n-1} c_{k+1}x^k, \quad \mathrm{p1} = \sum_{k=0}^{n-2}(k+1)c_{k+2}x^k, \quad \mathrm{p2} = \sum_{k=0}^{n-3}(k+2)(k+1)c_{k+3}x^k.$$

In den Zeilen 7–9 wird also folgendes berechnet:

$$\begin{aligned}
\mathrm{p2} &= x\sum_{k=0}^{n-3}(k+2)(k+1)c_{k+3}x^k + 2\sum_{k=0}^{n-2}(k+1)c_{k+2}x^k = \sum_{k=0}^{n-2}(k+1)(k+2)c_{k+2}x^k,\\
\mathrm{p1} &= x\sum_{k=0}^{n-2}(k+1)c_{k+2}x^k + \sum_{k=0}^{n-1} c_{k+1}x^k = \sum_{k=0}^{n-1}(k+1)c_{k+1}x^k,\\
\mathrm{p0} &= x\sum_{k=0}^{n-1} c_{k+1}x^k + c(0) = \sum_{k=0}^{n} c_k x^k.
\end{aligned}$$

□

Herauskürzen einer berechneten Nullstelle

Das Laguerre–Verfahren läßt sich wie folgt zur Berechnung *aller* reeller Nullstellen eines Polynoms benutzen: Man wendet das Verfahren auf $p \in \mathcal{P}_n$ an und erhält eine Nullstelle α. Dann kürzt man $x-\alpha$ aus $p(x)$ heraus, d.h. man berechnet $q \in \mathcal{P}_{n-1}$ mit $p(x) = (x-\alpha)q(x)$ und wendet das Verfahren wieder auf q an, usw. Wegen der globalen Konvergenz des Verfahrens erhält man so nach und nach alle reellen Nullstellen. Das Verfahren bricht irgendwann ab, wenn Quadratwurzeln in $g(x)$ nicht mehr reell sind (was gerade bedeutet, daß keine reellen Nullstellen mehr vorhanden sind). Allerdings ist bei dieser Vorgehensweise der Rundungsfehlereinfluß ein großes Problem, weil die durch das

Herauskürzen berechneten Polynome (und damit ihre Nullstellen!) nach und nach immer mehr an Genauigkeit verlieren. Jede ungenaue Berechnung einer Nullstelle impliziert automatisch ungenaue Berechnung aller nachfolgender Nullstellen.

Zur Berechnung der Koeffizienten c'_k von q machen wir den Ansatz

$$p(x) = \sum_{j=0}^{n} c_k x^k = (x-\alpha)q(x) = (x-\alpha)\sum_{k=0}^{n-1} c'_k x^k = \sum_{k=1}^{n} c'_{k-1}x^k - \sum_{k=0}^{n-1} \alpha c'_k x^k.$$

Durch Koeffizientenvergleich erhalten wir:

$$c_0 = -\alpha c'_0, \quad c_1 = c'_0 - \alpha c'_1, \quad \ldots, \quad c_{n-1} = c'_{n-2} - \alpha c'_{n-1}, \quad c_n = c'_{n-1},$$

und man kann die c'_k wie folgt rückwärts berechnen:

$$c'_{n-1} = c_n, \quad c'_{n-2} = c_{n-1} + \alpha c'_{n-1}, \quad \ldots \quad c'_0 = c_1 + \alpha c'_1.$$

5.6 Beispiele für höherdimensionale Probleme

Nullstellen bzw. Fixpunkte von Funktionen, die auf Teilmengen des $\mathbb{R}^n$ für $n > 1$ definiert sind, sind i.A. viel schwieriger zu bestimmen als im Fall $n = 1$. Bei Fixpunktproblemen z.B. liegt die Schwierigkeit meistens schon darin, die Gültigkeit der Voraussetzungen des Banachschen Fixpunktsatzes zu überprüfen.

Fixpunktiteration im $\mathbb{R}^2$

Wir suchen einen Fixpunkt der Funktion $g : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$g(x,y) = \begin{pmatrix} g_1(x,y) \\ g_2(x,y) \end{pmatrix} = \begin{pmatrix} \frac{\cos(x)}{5} - xy^2 + \frac{1}{2} \\ \frac{x}{3} - \frac{\sin(y)}{2} \end{pmatrix}.$$

Zur Anwendung des Banachschen Fixpunktsatzes muß zunächst ein abgeschlossenes $\mathcal{D}$ gefunden werden mit $g(\mathcal{D}) \subset \mathcal{D}$. Das ist weitaus schwieriger als im $\mathbb{R}^1$, weil simultan zwei Gleichungen mit zwei Unbekannten kontrolliert werden müssen. Am einfachsten erscheint die Suche nach einem Rechteck $\mathcal{D}$ mit (x_l, y_l) als linker unterer Ecke und (x_r, y_r) als rechter oberer Ecke. Für gegebene Ecken überprüft man dann, ob $g(x_l, y_l)$ und $g(x_r, y_r)$ ebenfalls in $\mathcal{D}$ liegen. Ist das der Fall, besteht Hoffnung, daß $g(\mathcal{D}) \subset \mathcal{D}$, und man muß genauer

$$x_l \le g_1(x,y) \le x_r, \quad y_l \le g_2(x,y) \le y_r$$

zeigen. Andernfalls muß man neue Ecken suchen. In unserem Beispiel erhält man nach einigem Ausprobieren $(x_l, y_l) = (0.6, 0.1)$ und $(x_r, y_r) = (0.7, 0.2)$ als mögliche Kandidaten: Es gilt $g(x_l, y_l) \simeq (0.66, 0.15) \in \mathcal{D}$ und $g(x_r, y_r) \simeq (0.65, 0.13) \in \mathcal{D}$. Nähere Betrachtungen (Monotonität etc.) zeigen nun

$$\begin{aligned} 0.62 \simeq g_1(x_r, y_r) \le g_1(x,y) \le g_1(x_l, y_l) \simeq 0.66, \\ 0.1007 \simeq g_2(x_l, y_r) \le g_2(x,y) \le g_2(x_r, y_l) \simeq 0.18 \end{aligned}$$

was gleichbedeutend ist mit $g(\mathcal{D}) \subset \mathcal{D}$. Zu Überprüfung der Kontraktionseigenschaft berechnen wir

$$g'(x,y) = \begin{bmatrix} -\frac{\sin(x)}{5} - y^2 & -2\,xy \\ \frac{1}{3} & -\frac{\cos(y)}{2} \end{bmatrix}$$

und die Zeilensummennorm $\|g'(x,y)\|_\infty$, für die gilt

$$\begin{aligned} \|g'(x,y)\|_\infty &= \max\left\{\frac{\sin(x)}{5} + y^2 + 2xy, \frac{1}{3} + \frac{\cos(y)}{2}\right\} \\ &\le \max\left\{\frac{\sin(x_r)}{5} + y_r^2 + 2x_r y_r, \frac{1}{3} + \frac{\cos(y_l)}{2}\right\} \le \max\{0.45, 0.84\} = 0.84 \end{aligned}$$

und damit

$$\sup_{(x,y)\in\mathcal{D}} \|g'(x,y)\|_\infty \le 0.84.$$

Nach Satz 5.2 ist g in $\mathcal{D}$ bzgl. $\|\cdot\|_\infty$ kontrahierend und besitzt demnach genau einen Fixpunkt in $\mathcal{D}$. Beginnend mit $(x_0, y_0) = (0.65, 0.15)$ (dem Mittelpunkt von $\mathcal{D}$) erhalten wir bei der Iteration folgende Werte:

k	x_k	y_k
0	0.65	0.15
1	0.64459175	0.14014485
2	0.64720887	0.14589301
3	0.64577827	0.14257141
4	0.64659984	0.14448882
⋮	⋮	⋮
22	0.64630120	0.14378750
23	0.64630118	0.14378745

Die ersten 6 Stellen ändern sich hiernach nicht mehr. Der Fixpunkt (auf 6 Stellen genau) ist also $(x^*, y^*) \simeq (0.646301, 0.143787)$.

Bemerkung: In vielen Fällen konvergiert eine Iteration $x_{k+1} = g(x_k)$ nicht gegen einen Fixpunkt, sondern in einem gewissen Sinne gegen eine Menge von Punkten. Diese Mengen (in diesem Zusammenhang spricht man auch von *Attraktoren*) können unter Umständen selbst bei sehr harmlos aussehenden Iterationsfunktionen $g(x)$ höchst kompliziert sein. Ein Beispiel ist die Iteration $(x_{k+1}, y_{k+1}) = g(x_k, y_k)$ mit der sogenannten *Hénon-Abbildung*

$$g(x,y) = \begin{pmatrix} 1 + y - ax^2 \\ bx \end{pmatrix},$$

wobei a, b Parameter sind. Abb. 5.6 zeigt den Attraktor für $a = 1.4$, $b = 0.4$, ein Fraktal.

Das Newtonverfahren im $\mathbb{R}^2$

Das Newtonverfahren für eine differenzierbare Funktion kann man im $\mathbb{R}^n$ für $n > 1$ analog zu $n = 1$ herleiten. Allerdings muß man darauf achten, daß die Ableitung hier eine Matrix

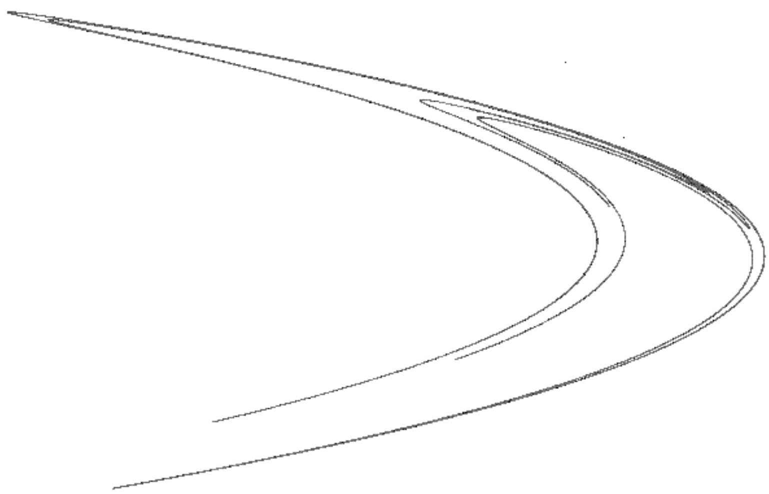

Abbildung 5.6: Der Hénon-Attraktor für $(x, y) \in [-1.4, 1.4] \times [-0.5, 0.5]$, $a = 1.4$, $b = 0.4$.

(die Jacobimatrix) ist. Ist x^* eine Nullstelle von f, so erhalten wir für x_0 in der Nähe von x^*

$$0 = f(x^*) \simeq f(x_0) + f'(x_0)(x^* - x_0) \Leftrightarrow x^* \simeq x_0 - f'(x_0)^{-1} f(x_0) =: g(x_0),$$

und das Newtonverfahren ist $x_{k+1} = g(x_k)$. In Anwendungen formuliert man das Newtonverfahren meistens um wie folgt mit $\delta_{k+1} := x_{k+1} - x_k$:

$$f'(x_k)\delta_{k+1} = -f(x_k), \quad x_{k+1} = x_k + \delta_{k+1},$$

das heißt, man löst ein Gleichungssystem für δ_{k+1} mit Koeffizientenmatrix $f'(x_k)$ und rechter Seite $-f(x_k)$ und erhält x_{k+1}, indem man δ_{k+1} zu x_k addiert. Auf diese Weise vermeidet man die direkte Invertierung von $f'(x_k)$. Man kann zeigen, daß analoge Konvergenzeigenschaften wie im Fall $n = 1$ gelten (insbesondere quadratische Konvergenz, falls $f'(x^*)$ regulär ist). Als numerisches Beispiel betrachten wir das Nullstellenproblem: Suche $(x^*, y^*) \in \mathbb{R}^2$ als eine Nullstelle von

$$f(x, y) = \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \end{pmatrix} = \begin{pmatrix} \frac{\cos(x)}{5} - x(y^2 + 1) + \frac{1}{2} \\ \frac{x}{3} - y - \frac{\sin(y)}{2} \end{pmatrix}$$

Dies ist eine Umformulierung des obigen Fixpunktproblems als Nullstellenproblem. Beginnend wieder mit dem Startvektor $(x_0, y_0) = (0.65, 0.15)$ liefert die Newtoniteration

k	x_k	y_k
0	0.65	0.15
1	0.64632908	0.14379274
2	0.64630119	0.14378747

nach nur 2 Iterationsschritten hat man also schon eine Genauigkeit von 6 Stellen erreicht, was eindrucksvoll die Überlegenheit gegenüber der einfachen Fixpunktiteration demonstriert. Der Nachteil ist natürlich der erhöhte Aufwand, im Wesentlichen Berechnung der Jacobimatrix und Lösen eines linearen Gleichungssystems.

Funktional-Iteration

Als abschließendes Beispiel wollen wir folgendes Problem mit dem Banachschen Fixpunktsatz behandeln: Gegeben ein Intervall $[0, a]$, wobei a noch zu bestimmen ist, gesucht ist eine Funktion $f \in C[0, a]$ mit

$$f(x) = 1 + \int_0^x f(t)\, dt := (Gf)(x), \quad 0 \leq x \leq a. \tag{5.26}$$

Ähnlich wie bei der Fredholmschen Integralgleichung ist dies eine Funktionalgleichung, d.h. eine Lösung f^* ist kein Vektor im $\mathbb{R}^n$, sondern ein Element des Funktionenraumes $C[0, a]$. (5.26) entsteht bei der Integration der Differentialgleichung $f'(x) = f(x)$ zum Anfangswert $f(0) = 1$, welche die eindeutige Lösung $f(x) = e^x$ besitzt. Da (5.26) in Fixpunktform ist, versuchen wir, den Banachschen Fixpunktsatz anzuwenden. Dafür nehmen wir $\mathcal{V} = C[0, a]$ mit der Norm

$$\|f\|_\infty := \max_{x \in [0,a]} |f(x)|.$$

Man kann zeigen, daß $\mathcal{V}$ bezüglich dieser Norm ein vollständiger normierter Vektorraum, also ein Banachraum ist. Wir setzen $\mathcal{D} = \mathcal{V}$ und sehen, daß $G(\mathcal{D}) \subset \mathcal{D}$ (mit anderen Worten: Falls $f(x)$ stetig ist, ist auch $(Gf)(x)$ stetig). Um zu zeigen, daß G in $\mathcal{D}$ genau einen Fixpunkt hat, müssen wir noch zeigen, daß G bezüglich der obigen Norm auf $\mathcal{D}$ kontrahierend ist. Das geht wie folgt: Seien zwei Funktionen $f_1, f_2 \in \mathcal{D}$, dann ist

$$\begin{aligned}
|(Gf_1)(x) - (Gf_2)(x)| &= \left| \int_0^x f_1(t) - f_2(t)\, dt \right| \\
&\leq \int_0^x |f_1(t) - f_2(t)|\, dt \\
&\leq \int_0^a \max_{t \in [0,a]} |f_1(t) - f_2(t)|\, dt \\
&= a\|f_1 - f_2\|_\infty, \quad \text{also auch} \\
\|Gf_1 - Gf_2\|_\infty &= \max_{x \in [0,a]} |(Gf_1)(x) - (Gf_2)(x)| \\
&\leq a\|f_1 - f_2\|_\infty.
\end{aligned}$$

Falls $a < 1$, haben wir also gezeigt, daß G auf $\mathcal{D}$ eine Kontraktion ist und somit nach dem Banachschen Fixpunktsatz einen eindeutigen Fixpunkt $f^* \in \mathcal{D}$ besitzt. Für jedes $f_0 \in \mathcal{D}$ konvergiert außerdem die Folge $f_{k+1} = Gf_k$ bezüglich der ∞-Norm, also gleichmäßig, gegen f^*. Wir berechnen nun einmal die ersten Glieder der Folge, beginnend mit $f_0(x) = 1$ und erhalten:

$$f_1(x) = 1 + \int_0^x 1\, dt = 1 + x, \quad f_2(x) = 1 + \int_0^x f_1(t)\, dt = 1 + x + \frac{x^2}{2}, \quad \text{usw.}$$

Insgesamt ist es nicht schwer zu zeigen, daß $f_k(x)$ gerade die k-te Partialsumme $\sum_{i=0}^k x^i/i!$ der Exponentialreihe ist.