Computing

Archiv für elektronisches Rechnen
Archives for Electronic Computing

Reprint from Vol 13 Fasc 1 1974

G. Alefeld

Stets konvergente Verfahren höherer Ordnung zur Berechnung von reellen Nullstellen

Springer-Verlag Wien New York

Computing 13, 55—65 (1974)

Stets konvergente Verfahren höherer Ordnung zur Berechnung von reellen Nullstellen

Von

G. Alefeld, Karlsruhe

Eingegangen am 26. November 1973

Zusammenfassung — Abstract

Stets konvergente Verfahren höherer Ordnung zur Berechnung von reellen Nullstellen. In dieser Arbeit geben wir *stets* konvergente Verfahren höherer Ordnung an zur Berechnung einer in einem Intervall $[a, b]$ gelegenen Nullstelle einer streng monoton wachsenden oder fallenden reellen Funktion, die dort stetige Ableitungen genügend hoher Ordnung besitzt. Das verwendete Konstruktionsprinzip stellt eine geeignete Verallgemeinerung des Ehrmannschen Vorgehens [5] unter Verwendung intervallarithmetischer Hilfsmittel dar, wodurch die Konvergenz stets erzwungen werden kann.

Always Convergent Higher Order Methods for the Computation of a Real Zero. In this paper we give always convergent higher order methods for the computation of a real zero of a real function which has derivatives of sufficiently high order. The principle of constructing these methods consists in a generalization of that used by Ehrmann [5]. By making appropriate use of interval arithmetic we always can assure convergence.

1. Einleitung

Gegeben sei ein reelles Intervall $X_0=[a, b]$ und eine Abbildung

$$f: X_0 \to \mathbb{R},$$

die in X_0 stetige Ableitungen bis zur $(p+1)$-ten Ordnung einschließlich besitzt. Es sei [1]

$$f(a)\, f(b)<0$$

und

$$f'(x) \neq 0, \quad x \in X_0.$$

Gesucht ist ein Iterationsverfahren, welches

(I) unter Verwendung der Werte von $f(x), f'(x), \ldots, f^{(p)}(x)$, $x \in X_0$, durchführbar ist,

---

[1] Daß die Bedingung $f(a)\, f(b)<0$ erfüllt ist, wird im folgenden ohne nochmalige Erwähnung stets vorausgesetzt.

(II) mindestens die Konvergenzordnung $p+1$ besitzt und

(III) stets gegen die eindeutige Nullstelle $x^* \in X_0$ von $f(x)=0$ konvergiert[2].

Ein allgemeines Prinzip, mit welchem man Iterationsverfahren aufstellen kann, welche den Forderungen (I) und (II) genügen, (III) jedoch nur lokal, d. h. für „genügend nahe" an x^* liegende Startwerte erfüllt ist, besteht in der Taylorentwicklung der Umkehrfunktion von f (siehe [2], [6]). Ein vor allem für die rechnerische Durchführung günstiges Prinzip zur Erzeugung von Iterationsverfahren, welche (I) und (II) erfüllen, hat Ehrmann in [5] angegeben (siehe auch [7]). Auch dabei ist (III) im allgemeinen nur lokal erfüllt.

Im folgenden geben wir Verfahren an, für welche auch (III) erfüllt ist. Darüber hinaus liefern diese Verfahren in jedem Iterationsschnitt verbesserte obere bzw. untere Schranken für die Nullstelle x^*, womit man automatisch in jedem Iterationsschnitt eine Fehlerabschätzung besitzt. Die Folge der oberen bzw. unteren Schranken konvergiert gegen x^*.

Benötigt werden dazu abgeschlossene und beschränkte Intervalle F_i, $i=1\,(1)\,p+1$, für die

$$\begin{aligned} f^{(i)}(x) &\in F_i, \quad i=1\,(1)\,p+1 \quad (x \in X_0) \\ 0 &\notin F_1 \end{aligned}$$

gilt. Die Beschaffung dieser Intervalle bereitet prinzipiell keine Schwierigkeiten, da man sie sehr einfach vor Durchführung der Iteration durch intervallarithmetische Auswertung der Ableitungen von f mit X_0 als Argument erhalten kann. Ist einer dieser Intervallausdrücke nicht definiert (z. B. wegen Division durch ein Nullintervall) oder ist $0 \in f'(X_0)$, so kann man dies aufgrund der Stetigkeit der intervallmäßigen Auswertung durch Halbierung des Intervalls X_0 vermeiden.

Im folgenden Abschnitt 2 werden zunächst die wichtigsten Hilfsmittel zusammengestellt. Im Abschnitt 3 werden die bekannten Eigenschaften der von Moore [12] untersuchten stets konvergenten Modifikation des Newton-Verfahrens aufgeführt und in Abschnitt 4 mit einem ähnlich aufgebauten Verfahren verglichen. Die Herleitung dieses Verfahrens ist ähnlich dem Ehrmannschen Vorgehen und zeigt zugleich das allgemeine Prinzip, welches zur Konstruktion stets konvergenter Verfahren $(p+1)$-ter Ordnung verwendet werden kann. Die Angabe dieser Verfahren erfolgt im nächsten Abschnitt.

2. Bezeichnungen und Hilfsmittel

In der Menge $I(\mathbb{R})$ der kompakten (nichtleeren) Intervalle über der reellen Zahlengeraden lassen sich arithmetische Verknüpfungen erklären. Sind

$$X=[x_1, x_2], \ Y=[y_1, y_2] \in I(\mathbb{R}),$$

[2] Mit x^* bezeichnen wir im folgenden stets die eindeutige Nullstelle von $f(x)=0$ im Intervall $[a, b]$.

so gilt

$$X * Y = \{x * y \mid x \in X,\ y \in Y\} \in I(\mathbb{R}),\ * \in \{+, -, \cdot, /\}.$$

Die Division durch ein Intervall Y mit $0 \in Y$ wird ausgeschlossen. Diese Verknüpfungen genügen der Teilmengeneigenschaft:

$$X_i \subseteq Y_i,\ i = 1, 2 \Rightarrow X_1 * X_2 \subseteq Y_1 * Y_2,\ \ * \in \{+, -, \cdot, /\}.$$

Speziell gilt

$$x_i \in X_i,\ i = 1, 2 \Rightarrow x_1 * x_2 \in X_1 * X_2,\ \ * \in \{+, -, \cdot, /\}.$$

Ist $X \in I(\mathbb{R})$, so bezeichnen wir mit $I(X)$ die Menge aller (nichtleeren) kompakten Intervalle, welche in X enthalten sind. Durch

$$q(X, Y) = \max\{|x_1 - y_1|, |x_2 - y_2|\},\quad X, Y \in I(\mathbb{R}),$$

wird $(I(\mathbb{R}), q)$ ein vollständiger metrischer Raum.

$$|X| := q(X, 0),\quad X \in I(\mathbb{R}),$$

heißt Betrag von X.

$$d(X) := x_2 - x_1,\quad X = [x_1, x_2] \in I(\mathbb{R}),$$

heißt Durchmesser von X.

Unter anderem gelten für $X, Y \in I(\mathbb{R})$ die nachfolgend verwendeten Eigenschaften

$$\begin{aligned}
&d(a) = 0, && a \in \mathbb{R},\\
&d(aX) = |a|\, d(X), && a \in \mathbb{R},\\
&d(X^n) \le n\, |X|^{n-1}\, d(X), && n = 1, 2, \ldots, (X^n := \underbrace{X \cdot X \cdot \ldots \cdot X}_{n\text{-mal}}),\\
&X \subseteq Y \Rightarrow d(X) \le d(Y), &&\\
&d(X) = |X - X|, &&\\
&\tfrac{1}{2} d(X) \le q(X, x) \le d(X), && x \in X.
\end{aligned}$$

(Die letzte Ungleichungskette bewirkt, daß für eine Folge $\{X_k\}_{k=0}^{\infty}$ mit $x^* \in X_k$, $k = 0, 1, 2, \ldots$, aus $\lim_{k \to \infty} d(X_k) = 0$ die Konvergenz $\lim_{k \to \infty} X_k = x^*$ folgt.

Außerdem sind die beiden Aussagen

$$\left.\begin{aligned}
&d(X_{k+1}) \le \gamma\, d(X_k)^p,\\
&q(X_{k+1}, x^*) \le \tilde{\gamma}\, q(X_k, x^*)^p
\end{aligned}\right\}\quad (p \ge 1)$$

äquivalent, d. h. falls die Folge der Durchmesser mindestens von der Ordnung $p \ge 1$ gegen Null konvergiert, so konvergiert auch die Folge der Abstände mindestens mit der Ordnung p gegen Null und umgekehrt.)

Mit

$$m(X) = \tfrac{1}{2}(x_1 + x_2)$$

bezeichnen wir den Mittelpunkt des Intervalls $X = [x_1, x_2]$.

3. Das Verfahren von Moore [12]

Ist f einmal stetig differenzierbar in $X_0=[a,b]$ und $f'(x)\in F_1$ für $x\in X_0$, $F_1\in I(\mathbb{R})$, $0\notin F_1$, und ist die intervallarithmetische Auswertung von f' an der Stelle X_0 definiert, so liefert das Verfahren

$$\begin{cases} x_k\in X_k \\ X_{k+1}=\left\{x_k-\dfrac{f(x_k)}{F_1\cap f'(X_k)}\right\}\cap X_k \end{cases}$$

$$k=0,1,2,\ldots,$$

eine Folge von Intervallen $X_k\in I(\mathbb{R})$ mit den Eigenschaften[1,2]

(a) $x^*\in X_k$, $k=0,1,2,\ldots$,

(b) $\lim\limits_{k\to\infty} X_k=x^*$,

(c) $d(X_{k+1})\le c\,d(X_k)^2$,
falls $f'(X)$ für $X\in I(X_0)$ einer Bedingung $d(f'(X))\le\gamma\,d(X)$ genügt.

Diese Aussagen gelten unabhängig von der Wahl von $x_k\in X_k$. Moore wählt $x_k=m(X_k)$, was in gewissem Sinne optimal ist ([8]). In der von Moore angegebenen Form des obigen Verfahrens fehlt die Bildung von $F_1\cap f'(X_k)$ und es wird $0\notin f'(X_0)$ vorausgesetzt. Da $f'(X_0)$ den Wertebereich von $f'(x)$, $x\in X$, i. a. überschätzt, kann eine anders berechnete Einschließung F_1 mit $0\notin F_1$ den Start des Verfahrens sichern ([9]).

Das Mooresche Verfahren benötigt in jedem Iterationsschnitt den Wert von f an der Stelle x_k sowie die intervallarithmetische Auswertung der Ableitung von f mit dem Intervall X_k als Argument.

4. Ein quadratisch konvergentes Verfahren ohne intervallarithmetische Auswertung der Ableitung

Es sei f zweimal stetig differenzierbar in X_0 und mit $F_1, F_2\in I(\mathbb{R})$ gelte

$$f'(x)\in F_1,\quad f''(x)\in F_2 \quad \text{für } x\in X_0$$
$$0\notin F_1.$$

Für $x_0=m(X_0)$ folgt dann aus dem Mittelwertsatz

$$f(x_0)-f(x^*)=f'(\xi_1)(x_0-x^*),$$

(ξ_1 zwischen x_0 und x^*), und der Teilmengeneigenschaft

$$x^*=x_0-\frac{f(x_0)}{f'(\xi_1)}\in x_0-\frac{f(x_0)}{F_1},$$

$$x^*\in Y_1:=\left\{x_0-\frac{f(x_0)}{F_1}\right\}\cap X_0,$$

$$x^*-x_0\in Y_1-x_0.$$

Aus der Taylorformel

$$0=f(x^*)=f(x_0)+f'(x_0)(x^*-x_0)+\tfrac{1}{2}f''(\xi_2)(x^*-x_0)^2,$$

(ξ_2 zwischen x_0 und x^*), folgt mit der Teilmengeneigenschaft

$$x^*=x_0-\frac{1}{f'(x_0)}[f(x_0)+\tfrac{1}{2}f''(\xi_2)(x^*-x_0)^2]$$

$$\in\left\{x_0-\frac{1}{f'(x_0)}[f(x_0)+\tfrac{1}{2}F_2(Y_1-x_0)^2]\right\}\cap Y_1=:X_1.$$

Die fortwährende Wiederholung dieses Schlusses liefert das folgende Iterationsverfahren:

$$\left\{\begin{aligned} x_k&=m(X_k)\\ Y_{k+1}&=\left\{x_k-\frac{f(x_k)}{F_1}\right\}\cap X_k\\ X_{k+1}&=\left\{x_k-\frac{1}{f'(x_k)}[f(x_k)+\tfrac{1}{2}F_2(Y_{k+1}-x_k)^2]\right\}\cap Y_{k+1}\end{aligned}\right. \tag{1}$$

$k=0,1,2,\dots.$

Pro Iterationsschnitt werden bei diesem Verfahren (abgesehen von einigen zusätzlichen arithmetischen Operationen) der Funktionswert und die Ableitung an der Stelle x_k benötigt. Da die intervallarithmetische Auswertung der Ableitung für ein echtes Intervall im allgemeinen mit größerem Aufwand als die Berechnung der Ableitung an der Stelle x_k verbunden ist, kommen wir hier mit geringerem Aufwand als beim Mooreschen Verfahren aus. Ansonsten besitzt das angegebene Verfahren die gleichen Eigenschaften. Dies ist der Inhalt des folgenden Satzes.

Satz 1:

Für das Verfahren (1) *gilt*[2]

(*a*) $x^*\in X_k,\quad k=0,1,2,\dots,$

(*b*) $\lim\limits_{k\to\infty} X_k=x^*,$

(*c*) $d(X_{k+1})\le c_1\,d(X_k)^2.$

Beweis:

Die Behauptung (a) zeigt man durch vollständige Induktion, indem man die bei der Herleitung des Verfahrens (1) durchgeführten Überlegungen mit X_k als Ausgangsintervall durchführt.

Aus der Beziehung

$$Y_{k+1}=\left\{x_k-\frac{f(x_k)}{F_1}\right\}\cap X_k$$

folgt, daß (außer für den Fall $x_k = x^*$, womit dann (b) gilt) die Beziehung $x_k \notin Y_{k+1}$ besteht. Daraus folgt $d(Y_{k+1}) < \frac{1}{2} d(X_k)$. Wegen der Durchschnittsbildung folgt dann

$$d(X_{k+1}) < d(Y_{k+1}) < \tfrac{1}{2} d(X_k), \quad k = 0, 1, 2, \ldots,$$

womit

$$\lim_{k \to \infty} d(X_k) = 0$$

gilt. Da nach (a) $x^* \in X_k$, $k = 0, 1, 2, \ldots$, gilt, folgt $\lim_{k \to \infty} X_k = x^*$. Damit ist (b) gezeigt.

Unter Verwendung der in Abschnitt 2 angegebenen Durchmesserregeln ergibt sich aus der Iterationsvorschrift (1)

$$\begin{aligned} d(X_{k+1}) &\le d\left(x_k - \frac{1}{f'(x_k)}\left[f(x_k) + \tfrac{1}{2} F_2 (Y_{k+1} - x_k)^2\right]\right) \\ &\le \tfrac{1}{2} d\left(\frac{F_2}{F_1}(X_k - X_k)^2\right) \\ &\le \tfrac{1}{2} d\left(\frac{F_2}{F_1}\left[-d(X_k)^2, d(X_k)^2\right]\right) \\ &= \left|\frac{F_2}{F_1}\right| d(X_k)^2 . \end{aligned}$$

Damit ist (c) mit $c_1 = \left|\frac{F_2}{F_1}\right|$ gezeigt und der Satz ist bewiesen.

Bemerkung 1:

Die Aussagen von Satz 1 bleiben bestehen, wenn man anstelle von $x_k = m(X_k)$ im Verfahren (1) einen beliebigen Punkt $x_k \in X_k$ wählt. Der Beweis erfordert nur geringe Abänderungen.

Bemerkung 2:

Die obigen Ausführungen, wonach Verfahren (1) mit geringerem Aufwand als das Mooresche Verfahren auskommt, gelten theoretisch, d. h. bei exakter Rechnung. Um beim Rechnen mit endlicher Stellenzahl auf einer Rechenanlage die Einschließung der Nullstelle zu sichern, kann man etwa alle Operationen maschinenintervallarithmetisch ausführen. Dies erfordert beim Mooreschen Verfahren neben der Berechnung von $f'(X_k)$ in der Maschinenintervallarithmetik ([1], [3]) auch die Berechnung eines Maschinenintervalls, welches $f(x_k)$ enthält. Ein solches Intervall kann man durch Auswertung von $f([x_k, x_k])$ in der Maschinenintervallarithmetik gewinnen. Beim Verfahren (1) wird — neben der Berechnung von $f([x_k, x_k])$ in der Maschinenintervallarithmetik — anstelle von $f'(X_k)$ die Berechnung von $f'([x_k, x_k])$ in der Maschinenintervallarithmetik benötigt. Die Berechnung dieser beiden Intervalle erfordert im allgemeinen den gleichen Aufwand. Unter diesen Gesichtspunkten erfordern das Mooresche Verfahren und Verfahren (1) — abgesehen von einigen zusätzlichen arithmetischen Operationen, welche höchstens die erreichbare Genauigkeit beeinflussen können — den gleichen Aufwand. Der Vorteil des Mooreschen Verfahrens besteht darin, daß (zu seiner Herleitung und Durchführung) nicht die Existenz der zweiten Ableitung von f benötigt wird. Demgegenüber lassen sich die bei der Herleitung von Verfahren (1) verwendeten Überlegungen unmittelbar zur Gewinnung von Verfahren mit höherer Konvergenzordnung heranziehen. Dies wird im nächsten Abschnitt durchgeführt.

Bemerkung 3:

In [11], Abschnitt 9, Satz 3, wird unter der Voraussetzung, daß f zweimal stetig differenzierbar ist, das folgende Verfahren angegeben:

$$\begin{cases} x_k \in X_k \\ X_{k+1} = \left\{ x_k - \dfrac{1}{f'(x_k)} \left[f(x_k) + \tfrac{1}{2} f''(X_k)(X_k - x_k)^2 \right] \right\} \cap X_k, \end{cases}$$

$$k = 0, 1, 2, \ldots .$$

Es gilt $x^* \in X_k$, $k = 0, 1, 2, \ldots$. Bedingungen für die Konvergenz $\lim_{k \to \infty} X_k = x^*$ sind nicht angegeben. Im Falle $\lim_{k \to \infty} X_k = x^*$ konvergiert die Folge der Durchmesser quadratisch gegen Null. Gegenüber Verfahren (1) wird zusätzlich in jedem Schritt die zweite Ableitung für das Intervall X_k ausgewertet [3]. Bei der praktischen Durchführung dieses Verfahrens, d. h. falls die Rundungsfehler bei der Berechnung von $f(x_k)$ und $f'(x_k)$ durch Auswertung in der Maschinenintervallarithmetik miterfaßt werden sollen, erfordert es ein Drittel Aufwand mehr. Abgesehen von der ungeklärten Konvergenzfrage erscheint es daher wenig attraktiv.

5. Verfahren höherer Ordnung

Es sei f $(p+1)$-mal stetig differenzierbar in $X_0 = [a, b]$ und es gelte für $F_i \in I(\mathbb{R})$, $i = 1\,(1)\,p+1$,

$$f^{(i)}(x) \in F_i, \quad x \in X_0, \quad i = 1\,(1)\,p+1,$$

$$0 \notin F_1 .$$

Wir betrachten das folgende Iterationsverfahren

[3] Dies verkleinert den Konvergenzfaktor, liefert jedoch keine Erhöhung der Konvergenzordnung. (Dasselbe gilt für Verfahren (1), falls man das konstante Intervall F_2 in jedem Iterationsschritt durch $f''(X_k)$ ersetzt.)

$$\left.\begin{aligned}
&x_k = m(X_k)\\
&X_{k+1,0} = \left\{x_k - \frac{f(x_k)}{F_1}\right\} \cap X_k\\
&X_{k+1,i} = \left\{x_k - \frac{1}{f'(x_k)}\left(f(x_k) + \sum_{\nu=2}^{i} \frac{1}{\nu!} f^{(\nu)}(x_k)(X_{k+1,i-1} - x_k)^\nu\right.\right.\\
&\qquad\left.\left. + \frac{1}{(i+1)!} F_{i+1}(X_{k+1,i-1} - x_k)^{i+1}\right)\right\} \cap X_{k+1,i-1}\\
&\quad i = 1\,(1)\,p\\
&X_{k+1} = X_{k+1,p}\\
&\quad k = 0, 1, 2, \ldots.
\end{aligned}\right\} \qquad (2)$$

Für $p=1$ geht (2) in (1) über. Das angegebene Verfahren erfordert pro Schritt die Berechnung der Werte von $f(x_k), f'(x_k), \ldots, f^{(p)}(x_k)$ und besitzt die folgenden Eigenschaften.

Satz 2:

Für das Verfahren (2) *gilt*[2]

(a) $x^* \in X_k, \quad k = 0, 1, 2, \ldots,$

(b) $\lim_{k\to\infty} X_k = x^*,$

(c) $d(X_{k+1}) \le \gamma\, d(X_k)^{p+1}.$

Beweis:

Zu (a): Es sei $x^* \in X_k$ für ein $k \ge 0$, was nach Voraussetzung für $k=0$ richtig ist. Wie bei der Herleitung von Verfahren (1) zeigt man dann mit Hilfe des Mittelwertsatzes

$$x^* \in X_{k+1,0}.$$

Ist $x^* \in X_{k+1,i}$ für ein $i \ge 0$, was für $i=0$ richtig ist, so gilt

$$x^* - x_k \in X_{k+1,i} - x_k.$$

Damit folgt aus der Taylorschen Formel

$$0 = f(x^*) = f(x_k) + f'(x_k)(x^* - x_k) + \ldots + \frac{1}{(i+1)!} f^{(i+1)}(x_k)(x^* - x_k)^{i+1}$$

$$+ \frac{1}{(i+2)!} f^{(i+2)}(\xi_{i+2})(x^* - x_k)^{i+2},$$

(ξ_{i+2} zwischen x_k und x^*), und aufgrund der Teilmengeneigenschaft

$$x^* = x_k - \frac{1}{f'(x_k)}\left[f(x_k) + \sum_{\nu=2}^{i+1} \frac{1}{\nu!} f^{(\nu)}(x_k)(x^* - x_k)^\nu\right.$$

$$+\frac{1}{(i+2)!} f^{(i+2)}(\xi_{i+2})(x^*-x_k)^{i+2}\Bigg]$$

$$\in \left\{x_k - \frac{1}{f'(x_k)}\left[f(x_k) + \sum_{\nu=2}^{i+1} \frac{1}{\nu!} f^{(\nu)}(x_k)(X_{k+1,i}-x_k)^\nu\right.\right.$$

$$\left.\left.+\frac{1}{(i+2)!} F_{i+2}(X_{k+1,i}-x_k)^{i+2}\right]\right\} \cap X_{k+1,i}$$

$$= X_{k+1,i+1}.$$

Damit gilt $x^* \in X_{k+1,i}$, $i=0\,(1)\,p$, und $x^* \in X_{k+1} = X_{k+1,p}$.

Zu (b): Wie beim Beweis von Satz 1 folgt

$$d(X_{k+1,0}) \leq \tfrac{1}{2}\, d(X_k)$$

und aufgrund der Durchschnittsbildung

$$d(X_{k+1,i}) \leq \tfrac{1}{2}\, d(X_k), \quad i=0\,(1)\,p,$$

also insbesondere

$$d(X_{k+1}) = d(X_{k+1,p}) \leq \tfrac{1}{2}\, d(X_k), \quad k=0,1,2,\ldots.$$

Die Konvergenzaussage folgt jetzt wie in Satz 1.

Zu (c): Es ist $d(X_{k+1,0}) \leq \frac{1}{2}\, d(X_k)$ und wie in Satz 1 folgt

$$d(X_{k+1,1}) \leq c_1\, d(X_k)^2$$

mit der von k unabhängigen Konstanten $c_1 = \left|\frac{F_2}{F_1}\right|$.

Es gelte mit einer von k unabhängigen Konstanten c_i

$$d(X_{k+1,i}) \leq c_i\, d(X_k)^{i+1}$$

für ein $i \geq 1$, was für $i=1$ richtig ist.

Dann folgt aus der Iterationsvorschrift (2) unter Verwendung der im Abschnitt 2 angegebenen Regeln

$$d(X_{k+1,i+1}) \leq d\left(\sum_{\nu=2}^{i+1} \frac{1}{\nu!} \frac{f^{(\nu)}(x_k)}{f'(x_k)}(X_{k+1,i}-x_k)^\nu\right.$$

$$\left.+\frac{1}{(i+2)!} \frac{F_{i+2}}{f'(x_k)}(X_{k+1,i}-x_k)^{i+2}\right)$$

$$= \sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left|\frac{f^{(\nu)}(x_k)}{f'(x_k)}\right| d\left((X_{k+1,i}-x_k)^\nu\right)$$

$$+\frac{1}{(i+2)!}\, d\left(\frac{F_{i+2}}{f'(x_k)}(X_{k+1,i}-x_k)^{i+2}\right)$$

$$\leq \sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left|\frac{F_\nu}{F_1}\right| \cdot \nu \cdot |X_{k+1,i}-x_k|^{\nu-1} \cdot d(X_{k+1,i}-x_k)$$

$$
\begin{aligned}
&+\frac{1}{(i+2)!}\, d\left(\frac{F_{i+2}}{f'(x_k)}\,(X_{k+1,i}-x_k)^{i+2}\right)\\
&\le \sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left(\frac{F_\nu}{F_1}\right) \nu\, |X_k - X_k|^{\nu-1}\, d(X_{k+1,i})\\
&+\frac{1}{(i+2)!}\, d\left(\frac{F_{i+2}}{F_1}\,(X_k-X_k)^{i+2}\right)\\
&\le \sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left|\frac{F_\nu}{F_1}\right| \nu\cdot d(X_k)^{\nu-1}\cdot c_i\cdot d(X_k)^{i+1}\\
&+\frac{1}{(i+2)!}\, d\left(\frac{F_{i+2}}{F_1}\,[-d(X_k)^{i+2}, d(X_k)^{i+2}]\right)\\
&= d(X_k)^{i+2} \quad \sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left|\frac{F_\nu}{F_1}\right|\cdot \nu\cdot c_i\cdot d(X_k)^{\nu-2}\\
&+\frac{2}{(i+2)!}\left|\frac{F_{i+2}}{F_1}\right| d(X_k)^{i+2}\\
&\le \left(\sum_{\nu=2}^{i+1} \frac{1}{\nu!}\left|\frac{F_\nu}{F_1}\right|\cdot \nu\cdot c_i\cdot d(X_0)^{\nu-2}\right.\\
&\left.+\frac{2}{(i+2)!}\left|\frac{F_{i+2}}{F_1}\right|\right) d(X_k)^{i+2}\\
&= c_{i+1}\, d(X_k)^{i+2}
\end{aligned}
$$

mit der Konstanten

$$
c_{i+1}=c_i \sum_{\nu=2}^{i+1} \frac{1}{(\nu-1)!}\left|\frac{F_\nu}{F_1}\right| d(X_0)^{\nu-2}+\frac{2}{(i+2)!}\left|\frac{F_{i+2}}{F_1}\right|,
$$

die unabhängig von k ist.

Damit besteht die Beziehung

$$
d(X_{k+1,i}) \le c_i\, d(X_k)^{i+1}
$$

für $i=1\,(1)\,p$ und es gilt

$$
d(X_{k+1})=d(X_{k+1,p})\le \gamma\, d(X_k)^{p+1},
$$

wobei γ von k unabhängig ist. Damit ist der Beweis von Satz 2 abgeschlossen.

Bemerkung 4:

Die Ausführungen von Bemerkung 1 gelten auch für das Verfahren $(p+1)$-ter Ordnung.

Bemerkung 5:

Ist $f(a)\,f(b)>0$, d. h. besitzt f in $X_0=[a,b]$ keine Nullstelle, so tritt nach einer endlichen Anzahl k^* von Schritten der Fall ein, daß einer der in Verfahren (2) zu bildenden Durchschnitte leer ist. Dies folgt durch einen einfachen Widerspruchsbeweis.

6. Schlußbemerkungen

Bei der Anwendung von Verfahren (2) muß man sich für eine bestimmte Ordnung entscheiden. Ohne auf Einzelheiten einzugehen, bemerken wir, daß man unter den in [5] bei entsprechenden Überlegungen gemachten Voraussetzungen auch hier zu dem Ergebnis kommt, daß das Verfahren (2) mit $p=2$, also das Verfahren dritter Ordnung optimal ist in dem Sinne, daß man für $p=2$ mit geringstem Aufwand (asymptotisch) eine vorgegebene Genauigkeit erreichen kann.

Abschließend bemerken wir, daß die in dieser Arbeit durchgeführten Überlegungen zur Gewinnung von stets konvergenten Verfahren höherer Ordnung auch auf Gleichungssysteme übertragen werden können. Darauf soll an anderer Stelle eingegangen werden.

Literatur

[1] Apostolatos, N., und U. Kulisch: Grundlagen einer Maschinenintervallarithmetik. Computing **2**, 89—104 (1967).

[2] Berezin, I., and N. Zhidkov: Computing Methods. Moscow: Fizmatgiz. Transl. by Blunn, O., and A. Booth. Oxford: Pergamon Press. 1965.

[3] Christ, H.: Realisierung einer Maschinenintervallarithmetik mit beliebigen ALGOL-60 Compilern. Elektron. Rechenanlagen **10**, 217—222 (1968).

[4] Collatz, L.: Funktionalanalysis und numerische Mathematik. Berlin-Göttingen-Heidelberg-New York: Springer. 1964.

[5] Ehrmann, H.: Konstruktion und Durchführung von Iterationsverfahren höherer Ordnung. Arch. Rational Mech. Anal. **4**, 65—88 (1959).

[6] Euler, L.: Institutiones Calculi Differentialis, II, Cap. IX. Opera Omnia, Ser. I, Vol. X, 422—455.

[7] Heinrich, H.: Numerische Behandlung nichtlinearer Gleichungen. Überblicke Mathematik 2. Mannheim: Bibliograph. Inst. 1969.

[8] Herzberger, J.: Optimale Verfahren zur Bestimmung einer reellen Nullstelle einer reellen Funktion. Habilitationsschrift, Universität Karlsruhe (1972).

[9] Herzberger, J.: Bemerkungen zu einem Verfahren von R. E. Moore. ZAMM **53**, 356—358 (1973).

[10] Kulisch, U.: Grundzüge der Intervallrechnung. Überblicke Mathematik 2. Mannheim: Bibliograph. Inst. 1969.

[11] Krawczyk, R.: Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. Computing **4**, 187—201 (1969).

[12] Moore, R. E.: Interval Analysis. Englewood Cliffs N. J.: Prentice Hall. 1966.

[13] Schröder, E.: Über unendlich viele Algorithmen zur Auflösung der Gleichungen. Math. Ann. **2**, 317—365 (1870).

[14] Alefeld, G., und J. Herzberger: Einführung in die Intervallrechnung. Mannheim: Bibliographisches Institut. 1974.

Prof. Dr. G. Alefeld
Kaiserstraße 12
Postfach 6380
D-7500 Karlsruhe
Bundesrepublik Deutschland