ZAMM · Z. angew. Math. Mech. 68 (1988) 8, 331 – 337

ALEFELD, G.; POTRA, F. A.

On Two Higher Order Enclosing Methods of J. W. Schmidt

Dedicated to HARRO HEUSER, Karlsruhe, on the occasion of his 60th birthday.

Wir betrachten Modifikationen zweier Iterationsverfahren von J. W. Schmidt, die unter geeigneten Bedingungen monotone Einschließungen für die Lösung einer nichtlinearen Gleichung liefern. Das erste Verfahren hat die Konvergenzordnung 3, und im skalaren Fall erfordert es zwei Funktions- und eine Ableitungsberechnung pro Iterationsschritt, während das zweite die Konvergenzordnung $1 + \sqrt{2}$ hat und nur zwei Funktionsberechnungen pro Schritt erfordert.

We consider modifications of two iterative procedures of J. W. Schmidt which, under appropriate conditions provide monotone enclosures for the solution of a nonlinear equation. The order of convergence of the first method is 3 and in the scalar case it requires two function- and one derivative-evaluation per iteration step, while the second one has the convergence order equal to $1 + \sqrt{2}$ and it requires only two function-evaluations per step.

Рассмотрим модификации двух метода итерации введенных Ё. В. Шмидта. При подходящих условиях эти методы обеспечивают монотонные вложения для решения нелинейного уравнения. Первый метод имеет порядок сходимости от 3, а в скалярном случае требуются два вычисления функции и одно вычисление производной на шаг итерации. Второй метод имеет порядок сходимости от $1 + \sqrt{2}$ и требует только два вычисления функции на шаг.

0. Introduction

Suppose that the real function $f\colon \mathbb{R} \to \mathbb{R}$ is convex and strictly increasing on an interval $[a, b]$ for which $f(a) \leqq 0 \leqq f(b)$. If f is twice continuously differentiable on $[a, b]$ then it is well known from J. B. FOURIER's work from 1818 (see [1, p. 248]) that the Newton-Fourier iterative procedure

Set $y_0 = a$, $z_0 = b$; for $n = 0, 1, \ldots$ compute

$$z_{n+1} = z_n - f'(z_n)^{-1} f(z_n), \qquad y_{n+1} = y_n - f'(z_n)^{-1} f(y_n) \tag{1}$$

produces two sequences $\{y_n\}_{n=0}^{\infty}$, $\{z_n\}_{n=0}^{\infty}$ that are monotonically convergent from below and from above to the unique zero x^* of f on $[a, b]$ i.e.:

$$y_n \leqq y_{n+1} \leqq x^* \leqq z_{n+1} \leqq z_n, \qquad \lim_{n\to\infty} y_n = \lim_{n\to\infty} z_n = x^*. \tag{2}$$

It can be proved that (see [3, p. 70]):

$$\lim_{n\to\infty} \frac{|z_{n+1} - y_{n+1}|}{|z_n - y_n|^2} = \left|\frac{f''(x^*)}{2f'(x^*)}\right| \tag{3}$$

which shows that the diameters of the intervals enclosing the root are Q-quadratically convergent to zero.

We note that the Newton-Fourier method requires the evaluation of two functions and one derivative per iteration step. J. W. SCHMIDT [5] has shown that with the same amount of work it is possible to produce enclosing sequences with a higher order of convergence. The two iterative procedures considered by J. W. SCHMIDT in the above mentioned paper can be described as follows:

Choose $y_0, y_1 \in [a, b]$ such that $f(y_0) \leqq 0$, $f(y_1) \leqq 0$; for $n = 1, 2, \ldots$ compute

$$z_n = y_n - f'(y_n)^{-1} f(y_n), \qquad y_{n+1} = y_n - \delta f(y_n, z_n)^{-1} f(y_n) \tag{4}$$

or

$$z_n = y_n - \delta f(y_{n-1}, y_n)^{-1} f(y_n), \qquad y_{n+1} = y_n - \delta f(y_n, z_n)^{-1} f(y_n), \tag{5}$$

where $\delta f(x, y)$ denotes the divided difference of f at the points x and y. The sequences produced by either (4) or (5) satisfy (2). Moreover, J. W. SCHMIDT has proved that the sequence $\{y_n\}$ given by (4) is Q-cubically convergent while the sequence $\{y_n\}$ given by (5) has the R-order of convergence $1 + \sqrt{2}$. Remarkably enough the iterative procedure (4) requires two function- and one derivative-evaluations (the same as the Newton-Fourier method), while the iterative procedure (5) requires only two function-evaluations per step (with the exception of the first one).

The major inconvenience connected with the iterative procedures (4) and (5) is the fact that in many cases the point z_1 produced by them may fall outside the interval $[a, b]$ where f is supposed to be increasing and convex and then the convergence may break down. In the above quoted paper of J. W. SCHMIDT the fact that z_1 belongs to $[a, b]$ is taken as a hypothesis of the convergence theorem. The paper [5] contains only results on the order of convergence of the sequence $\{y_n\}$. The order of convergence of the sequence $\{z_n\}$ is discussed in [7]. As with many enclosing methods we would be naturally interested in knowing the order with which the diameters of the enclosing intervals converge to zero.

In what follows we will try to fix these inconveniences by considering the following modifications of J. W. SCHMIDT's iterative procedures:

Set $y_0 = a$, $z_0 = b$; for $n = 0, 1, 2, \ldots$ compute

$$y_{n+1} = y_n - \delta f(y_n, z_n)^{-1} f(y_n), \qquad \bar{z}_{n+1} = y_{n+1} - f'(y_{n+1})^{-1} f(y_n), \qquad z_{n+1} = \inf\{\bar{z}_{n+1}, z_n\} \tag{6}$$

or

$$y_{n+1} = y_n - \delta f(y_n, z_n)^{-1} f(y_n), \qquad \bar{z}_{n+1} = y_{n+1} - \delta f(y_n, y_{n+1})^{-1} f(y_{n+1}), \qquad z_{n+1} = \inf \{\bar{z}_{n+1}, z_n\}. \tag{7}$$

We will study the above iterative procedures in partially ordered Banach spaces, in a framework similar to that considered by J. W. Schmidt [5]. However, for reasons of convenience, we will make some simplifying assumptions. Thus we will assume that all divided differences of f on $[a, b]$ are invertible (which ensures the uniqueness of the root) and we will consider a simpler, but more restrictive, Lipschitz condition on δf which turns out, however, to be satisfied by most examples of interest. (See [4]).

Also, in order to avoid repetition we will prove most of our convergence results for a "more general" iterative procedure of the form

$$\left.\begin{aligned} y_{n+1} &= y_n - \delta f(y_n, z_n)^{-1} f(y_n), \qquad \bar{y}_{n+1} = \alpha_n y_n + (1 - \alpha_n)\, y_{n+1}, \qquad \alpha_n \in [0, 1], \\ \bar{z}_{n+1} &= y_{n+1} - \delta f(\bar{y}_{n+1}, y_{n+1})^{-1} f(y_{n+1}), \qquad z_{n+1} = \inf \{\bar{z}_{n+1}, z_n\} \end{aligned}\right\} \tag{8}$$

which reduces to (6) for $\alpha_n = 0$, $n = 0, 1, \ldots$ and to (7) for $\alpha_n = 1$, $n = 0, 1, \ldots$.

The paper is organized as follows: in section 1 we review some basic definitions concerning partial ordering and divided differences of nonlinear operators; section 2 contains results on the monotonicity and the convergence of the iterative method (8); in section 3 we prove some statements about the order of convergence of the enclosing methods (6) and (7); the last section 4 contains a numerical example.

1. Preliminaries

Let us consider a Banach space B partially ordered by a cone K. This means that K is a closed, convex subset of B which has the property that $x \in K$, $x \neq 0$, implies $\alpha x \in K$ for $\alpha \geqq 0$ and $\alpha x \notin K$ for $\alpha < 0$. The partial ordering in B is then defined by $x \leqq y$ iff $y - x \in K$. The elements of K are called *positive*. If $u \leqq v$, then the set $[u, v] = \{x \in B;\, u \leqq x \leqq v\}$ is called an *interval*.

We assume that the cone K is *normal*, in the sense that there is a constant $\gamma > 0$ such that $0 \leqq x \leqq y$ implies $||x|| \leqq \gamma ||y||$, with γ independent of x and y.

We also assume that the cone is *regular*, which means that every order bounded increasing sequence is convergent in the norm of B.

Finally we assume that the cone K is *minihedral*, which means that each two-element set $\{x, y\}$ has a greatest lower bound $z = \inf \{x, y\}$.

The definitions listed above can be found in [2, p. 133]. The partial ordering of B induces a natural partial ordering in the Banach space $L(B)$ of all bounded linear operators acting on B. Namely if $S, T \in L(B)$ then $S \leqq T$ iff $Sx \leqq Tx$ for all $x \geqq 0$. A linear operator T is called *positive* iff $T \geqq 0$, where 0 denotes the zero operator.

In what follows we will consider a nonlinear mapping $f\colon D \subset B \to B$ where D is an open convex subset of B. The mapping f is supposed to have a divided difference on D which means that for every pair of elements $u, v \in D$ there is a bounded linear operator $\delta f(u, v) \in L(B)$ such that

$$\delta f(u, v)\,(u - v) = f(u) - f(v), \qquad u, v \in D. \tag{9}$$

We note that, in general, the divided difference is not symmetric (i.e. $\delta f(u, v) \neq \delta f(v, u)$). However, it is easily seen that (9) implies

$$\delta f(v, u)\,(u - v) = f(u) - f(v). \tag{10}$$

We assume that δf is increasing with respect to both arguments i.e.:

$$\delta f(u, v) \leqq \delta f(x, y), \quad \text{whenever} \quad u \leqq x \quad \text{and} \quad v \leqq y. \tag{11}$$

Also we assume that $\delta f(u, v)$ is boundedly invertible for all $u, v \in D$ and that

$$\delta f(u, v)^{-1} \geqq 0, \qquad u, v \in D. \tag{12}$$

Finally we assume that the divided difference is Lipschitz-continuous on D in the sense that there is a constant $\beta > 0$ such that

$$||\delta f(x, y) - \delta f(u, v)|| \leqq \beta(||x - u|| + ||y - v||), \qquad u, v, x, y \in D, \tag{13}$$

where $||\cdot||$ denotes both the norm of B and the operator norm of $L(B)$.

We remark that from (13) it follows that f is Fréchet-differentiable on D and that

$$f'(x) = \delta f(x, x), \qquad x \in D \tag{14}$$

(for a proof see, for example, [4]).

2. Monotone convergence

In the first part of this section we prove a monotone convergence theorem for the iterative method (8) under the general framework considered in the preceding section. The second part of the section contains some refinements of these results for scalar functions.

Theorem 1: *Suppose that the Banach space B and the nonlinear mapping $f\colon D \subset B \to B$ satisfy all the assumptions made in section 1. Assume also that there is an interval $[a, b] \subset D$ which contains a root x^* of the equation $f(x) = 0$ and that the value of f at a is negative i.e.:*

$$a \leqq x^* \leqq b\,, \qquad f(a) \leqq f(x^*) = 0\,. \tag{15}$$

Then the iterative algorithm (8) is well defined for $y_0 = a$, $z_0 = b$, and the sequences $\{y_n\}$, $\{z_n\}$ provide a monotone enclosure of the root x^* satisfying property (2).

Proof: We will prove the following relations:

(i) $f(y_n) \leqq 0$,

(ii) a) $y_n \leqq x^*$, b) $x^* \leqq z_n$,

(iii) a) $y_n \leqq y_{n+1}$, b) $z_{n+1} \leqq z_n$,

(iv) a) $\lim\limits_{n\to\infty} y_n = x^*$, b) $\lim\limits_{n\to\infty} z_n = x^*$.

We will prove (i)—(iii) by induction. For $n = 0$ (i) and (ii) reduce to (15). From (12) and (15) we have

$$y_1 = y_0 - \delta f(y_0, z_0)^{-1} f(y_0) \geqq y_0\,,$$

while from the definition of the greatest lower bound it follows that

$$z_1 = \inf\,\{\bar{z}_1, z_0\} \leqq z_0\,.$$

Thus (i)—(iii) are true for $n = 0$. Let us assume that they hold up to some fixed $n \geqq 0$.

(ii a) Using (8) we can write

$$x^* - y_{n+1} = x^* - y_n + \delta f(y_n, z_n)^{-1} f(y_n) = \delta f(y_n, z_n)^{-1}\,\{f(y_n) - f(x^*) - \delta f(y_n, z_n)\,(y_n - x^*)\} =$$
$$= \delta f(y_n, z_n)^{-1}\,\{\delta f(y_n, x^*) - \delta f(y_n, z_n)\}\,(y_n - x^*)\,.$$

This is positive since $\delta f(y_n, z_n)^{-1} \geqq 0$, $y_n - x^* \leqq 0$, and $\delta f(y_n, x^*) \leqq \delta f(y_n, z_n)$. Hence $y_{n+1} \leqq x^*$.

(ii b) Using (8) we deduce in a similar manner that

$$\bar{z}_{n+1} - x^* = \delta f(\bar{y}_{n+1}, y_{n+1})^{-1}\,\{\delta f(x^*, y_{n+1}) - \delta f(\bar{y}_{n+1}, y_{n+1})\}\,(x^* - y_{n+1})$$

which shows that $x^* \leqq \bar{z}_{n+1}$. By the induction hypothesis we have $x^* \leqq z_n$. Therefore $x^* \leqq \inf\,\{\bar{z}_{n+1}, z_n\} = z_{n+1}$.

(i) Using the first equation in (8) we have

$$f(y_{n+1}) = f(y_{n+1}) - f(y_n) - \delta f(y_n, z_n)\,(y_{n+1} - y_n) = \{\delta f(y_n, y_{n+1}) - \delta f(y_n, z_n)\}\,(y_{n+1} - y_n)\,.$$

Since by the induction hypothesis $y_n \leqq y_{n+1} \leqq x^* \leqq z_{n+1} \leqq z_n$ it follows that $\delta f(y_n, y_{n+1}) \leqq \delta f(y_n, z_n)$, which implies that $f(y_{n+1}) \leqq 0$.

(iii) From the above results it follows immediately that

$$y_{n+2} = y_{n+1} - \delta f(y_{n+1}, z_{n+1})^{-1} f(y_{n+1}) \geqq y_{n+1}\,,$$

while from the definition of the infimum we have

$$z_{n+2} = \inf\,\{\bar{z}_{n+2}, z_{n+1}\} \leqq z_{n+1}\,.$$

According to the induction principle (i)—(iii) hold for all $n \geqq 0$. From the regularity of the cone K there exist y^* and z^* for which $a \leqq y^* \leqq x^* \leqq z^* \leqq b$ hold such that

$$\lim_{n\to\infty} y_n = y^*\,, \qquad \lim_{n\to\infty} z_n = z^*\,.$$

From (11) it follows that $\delta f(y_n, z_n) \leqq \delta f(b, b) = B$ and then by using (iii a) we obtain

$$0 \geqq f(y_n) = \delta f(y_n, z_n)\,(y_n - y_{n+1}) \geqq B(y_n - y_{n+1})\,.$$

Because B is continuous and $\{y_n\}$ is convergent we have $\lim\limits_{n\to\infty} B(y_n - y_{n+1}) = 0$ which implies $0 \geqq f(y^*) \geqq 0$. Hence $f(y^*) = 0$. From the definition of the divided difference we have

$$0 = f(y^*) - f(x^*) = \delta f(y^*, x^*)\,(y^* - x^*)\,,$$

so that from the invertibility of $\delta f(y^*, x^*)$ it follows that $y^* = x^*$. In order to prove $z^* = x^*$ let us note that according to (11) we have

$$A = \delta f(a, a) \leqq \delta f(\bar{y}_n, y_n) \leqq \delta f(b, b) = B\,,$$

wherefrom by virtue of (12) it follows that

$$B^{-1} \leqq \delta f(\bar{y}_n, y_n)^{-1} \leqq A^{-1}\,.$$

Because $f(y_n) \leqq 0$ we can write

$$y_n - B^{-1} f(y_n) \leqq \bar{z}_n = y_n - \delta f(\bar{y}_n, y_n)^{-1} f(y_n) \leqq y_n - A^{-1} f(y_n)\,.$$

The first and the last term above converge to x^*. Therefore, using the normality of the cone, we have $\lim_{n\to\infty} \bar{z}_n = x^*$. Finally from $x^* \leqq z_n \leqq \bar{z}_n$ we deduce that $\lim_{n\to\infty} z_n = x^*$. □

Remark 1: By inspecting the proof of the above theorem we realize that the Lipschitz condition (13) has not been used. In proving $f(y^*) = 0$ we have only used the "continuity from the left" of f, which follows from (10) and (12). Indeed for any $h \geqq 0$ we have

$$Ah \leqq f(x) - f(x-h) = \delta f(x-h, x)\, h \leqq Bh\,,$$

so that for any increasing sequence $\{x_n\}$ which converges to x we have $\lim_{n\to\infty} f(x_n) = f(x)$. However if (13) is not satisfied then f may not be Fréchet-differentiable at all points of D. At such points $\delta f(x, x)$ has to be defined in terms of the "partial derivatives from the left and from the right". For example for scalar convex functions we can take

$$\delta f(x, x) = \tfrac{1}{2}\,\{f'(x-0) + f'(x+0)\}\,. \;\square$$

Remark 2: Note that we have not assumed $f(b) \geqq 0$. It turns out that if in the hypothesis of Theorem 1, we replace condition (15) by the condition

$$a \leqq b\,, \qquad f(a) \leqq 0 \leqq f(b)\,, \tag{16}$$

then the existence of a point $x^* \in [a, b]$ such that $f(x^*) = 0$ follows from a result of Schmidt and Leonhardt [6]. Thus (16) is more restrictive than (15). In fact it is easy to produce examples satisfying the hypothesis of Theorem 1 and for which $f(b) \geqq 0$ does not hold. However if the domain D is "large enough" we can always find a b with $f(b) \geqq 0$. Indeed we can prove that if all the assumptions in the hypothesis of Theorem 1, but the existence of the root $x^* \in [a, b]$, are satisfied and if

$$J = \{x \in B;\, x \geqq a\} \subset D\,,$$

then there is a $c \in J$ such that $f(c) \geqq 0$.

To see this we first observe that according to (10) and (11) we have

$$f(x) - f(a) = \delta f(a, x)\,(x - a) \geqq A(x - a)\,, \qquad x \in J\,,$$

where, as before, $A = \delta f(a, a)$. If we denote $c = a - A^{-1} f(a)$ then we have clearly $c \in J$ and therefore $f(c) \geqq f(a) + A(c - a) = 0$. □

The next result shows that if there is an iteration step for which $z_k = \bar{z}_k$ then the same is true for all subsequent steps.

Proposition 1: *Under the hypothesis of Theorem 1 assume that $z_k = \bar{z}_k$ for some $k \geqq 1$. Then $z_n = \bar{z}_n$ for all $n \geqq k$.*

Proof: Obviously, it is sufficient to prove that $z_n = \bar{z}_n$ implies $\bar{z}_{n+1} \leqq z_n$. From the definition of the divided difference we have

$$y_n - y_{n+1} = \delta f(y_n, y_{n+1})^{-1}\,\{f(y_n) - f(y_{n+1})\}\,.$$

Therefore by using (8) and the hypothesis $z_n = \bar{z}_n$ we get

$$\begin{aligned} z_n - \bar{z}_{n+1} &= y_n - y_{n+1} - \delta f(\bar{y}_n, y_n)^{-1} f(y_n) + \delta f(\bar{y}_{n+1}, y_{n+1})^{-1} f(y_{n+1}) = \\ &= \{\delta f(y_n, y_{n+1})^{-1} - \delta f(\bar{y}_n, y_n)^{-1}\}\, f(y_n) + \{\delta f(\bar{y}_{n+1}, y_{n+1})^{-1} - \delta f(y_n, y_{n+1})^{-1}\}\, f(y_{n+1})\,. \end{aligned}$$

The last term above is clearly positive, because $f(y_{n+1}) \leqq 0$ and $\delta f(\bar{y}_{n+1}, y_{n+1}) \geqq \delta f(y_n, y_{n+1})$.

Also $\delta f(\bar{y}_n, y_n) \leqq \delta f(y_n, y_n) \leqq \delta f(y_n, y_{n+1})$ which together with $f(y_n) \leqq 0$ shows that the other term is positive, too. Hence $\bar{z}_{n+1} \leqq z_n$. □

In the remainder of this section we give some more precise results for scalar functions. Namely we will show that in this case $z_n = \bar{z}_n$ for all indices n that are sufficiently large. Also we will prove that $z_n = \bar{z}_n$ for $n = 1, 2, \ldots$ in case f' is concave.

Proposition 2: *Assume that in the hypothesis of Theorem 1 we have $B = \mathbb{R}$ (the set of all real numbers) endowed with the natural ordering and topology. Then there is a positive integer N such that $z_n = \bar{z}_n$ for all $n \geqq N$.*

Proof: According to Proposition 1 it is sufficient to prove that $\bar{z}_{n+1} \leqq z_n$ for some $n \geqq 0$. Using (8) and the definition of the divided difference we have

$$\begin{aligned} \bar{z}_{n+1} - z_n &= y_{n+1} - z_n - \delta f(\bar{y}_{n+1}, y_{n+1})^{-1} f(y_{n+1}) = \\ &= \delta f(\bar{y}_{n+1}, y_{n+1})^{-1}\,\{\delta f(\bar{y}_{n+1}, y_{n+1})\,(y_{n+1} - z_n) - f(y_{n+1})\} = \\ &= \delta f(\bar{y}_{n+1}, y_{n+1})^{-1}\,\{[\delta f(\bar{y}_{n+1}, y_{n+1}) - \delta f(y_{n+1}, z_n)]\,(y_{n+1} - z_n) - f(z_n)\} = \\ &= \delta f(\bar{y}_{n+1}, y_{n+1})^{-1}\,\{\delta f(\bar{y}_{n+1}, y_{n+1}) - \delta f(y_{n+1}, z_n) + \delta f(y_n, z_n)\}\,(y_{n+1} - z_n)\,. \end{aligned} \tag{17}$$

In deducing the last equality we have used the fact that the first equation in (8) can be rewritten as

$$y_{n+1} = z_n - \delta f(y_n, z_n)^{-1} f(z_n)\,.$$

Let us denote

$$d_n = \delta f(\bar{y}_{n+1}, y_{n+1}) - \delta f(y_{n+1}, z_n) + \delta f(y_n, z_n)\,. \tag{18}$$

Under our assumptions we have $\lim_{n\to\infty} d_n = f'(x^*) > 0$, so that there is an $n \geqq 0$ such that $d_n > 0$. But then we have from (17) that $\bar{z}_{n+1} - z_n \leqq 0$.

Proposition 3: *Under the hypothesis of Proposition 2 assume that f' is concave on $[a, b]$. Then $z_n = \bar{z}_n$ for* $n = 1, 2, \ldots$.

Proof: Let us denote by $\delta^2 f$ and $\delta^3 f$, resp., the second and the third, resp., divided difference of f. Under our assumptions we have

$$\delta f(u, v) \geqq 0, \qquad \delta^2 f(u, v, w) \geqq 0, \qquad \delta^3 f(u, v, w, x) \leqq 0, \qquad u, v, w, x \in [a, b].$$

The number d_n defined in (18) can be written as

$$\begin{aligned} d_n &= \delta f(\bar{y}_{n+1}, y_{n+1}) + \delta^2 f(y_n, y_{n+1}, z_n)(y_n - y_{n+1}) = \\ &= \{\delta^2 f(y_n, y_{n+1}, z_n) - \delta^2 f(y_n, \bar{y}_{n+1}, y_{n+1})\}(y_n - y_{n+1})\, \delta f(y_n, \bar{y}_{n+1}) = \\ &= \delta^3 f(y_n, \bar{y}_{n+1}, y_{n+1}, z_n)(z_n - \bar{y}_{n+1})(y_n - y_{n+1}) + \delta f(y_n, \bar{y}_{n+1}). \end{aligned}$$

Because $\bar{y}_{n+1} \leqq z_n$, $y_n \leqq y_{n+1}$ it follows that $d_n \geqq 0$. Hence, according to (17) $\bar{z}_{n+1} \leqq z_n$. The proof is complete.

3. The order of convergence

In this section we will prove that the diameters of the enclosing intervals provided by (6) tend to zero Q-cubically, while those produced by (7) have the R-order of convergence $1 + \sqrt{2}$. In the onedimensional case we will prove some results similar to (3). In particular it will follow that if $f''(x^*) \neq 0$ then the diameters of the enclosing intervals provided by (7) converge to zero with Q-order $1 + \sqrt{2}$.

Theorem 2: *Under the hypothesis of Theorem 1 there is a constant $\mu > 0$ such that*

$$\|z_{n+1} - y_{n+1}\| \leqq \mu \|z_n - y_n\|^2 \|z_n - \bar{y}_n\|, \qquad n = 1, 2, \ldots.$$

Proof: From (8) and (10) it follows that

$$\begin{aligned} 0 \leqq z_{n+1} - y_{n+1} \leqq \bar{z}_{n+1} - y_{n+1} &= -\delta f(\bar{y}_{n+1}, y_{n+1})^{-1} f(y_{n+1}) = \\ &= -\delta f(y_{n+1}, y_{n+1})^{-1}\{f(y_{n+1}) - f(y_n) - \delta f(y_n, z_n)(y_{n+1} - y_n)\} = \\ &= -\delta f(\bar{y}_{n+1}, y_{n+1})^{-1}\{\delta f(y_n, y_{n+1}) - \delta f(y_n, z_n)\}(y_{n+1} - y_n). \end{aligned} \tag{19}$$

Using the fact that $0 \leqq y_{n+1} - y_n \leqq z_n - y_n$ and

$$\delta f(\bar{y}_{n+1}, y_{n+1})^{-1} \leqq A^{-1}, \qquad A = \delta f(a, a), \qquad \delta f(y_n, y_{n+1}) \leqq \delta f(y_n, z_n)$$

we obtain

$$0 \leqq z_{n+1} - y_{n+1} \leqq A^{-1}\{\delta f(y_n, z_n) - \delta f(y_n, y_{n+1})\}(z_n - y_n)$$

so that from (13) and the normality of the cone we have

$$\|z_{n+1} - y_{n+1}\| \leqq \mu_1 \|z_n - y_{n+1}\| \, \|z_n - y_n\|, \tag{20}$$

where $\mu_1 = \beta\gamma\|A^{-1}\|$. The first equation in (8) can be rewritten as

$$y_{n+1} = z_n - \delta f(y_n, z_n)^{-1} f(z_n). \tag{21}$$

Using (11), (12) and the fact that $\bar{y}_n < y_n \leqq y_{n+1} \leqq z_n$ it follows that

$$\begin{aligned} \delta f(\bar{y}_n, y_n)^{-1} f(z_n) - \delta f(y_n, z_n)^{-1} f(z_n) &= \delta f(\bar{y}_n, y_n)^{-1}\{\delta f(y_n, z_n) - \delta f(\bar{y}_n, y_n)\}\, \delta f(y_n, z_n)^{-1} f(z_n) = \\ &= \delta f(\bar{y}_n, y_n)^{-1}\{\delta f(y_n, z_n) - \delta f(\bar{y}_n, y_n)\}(z_n - y_{n+1}) \geqq 0. \end{aligned}$$

We deduce that

$$0 \leqq z_n - y_{n+1} = \delta f(y_n, z_n)^{-1} f(z_n) \leqq \delta f(\bar{y}_n, y_n)^{-1} f(z_n).$$

On the other hand from the third equation in (8) and the fact that $y_n \leqq z_n \leqq \bar{z}_n$ we have

$$0 = -(\bar{z}_n - y_n) - \delta f(\bar{y}_n, y_n)^{-1} f(y_n) \leqq -(z_n - y_n) - \delta f(\bar{y}_n, y_n)^{-1} f(y_n).$$

The above inequalities imply that

$$\begin{aligned} 0 \leqq z_n - y_{n+1} &\leqq \delta f(\bar{y}_n, y_n)^{-1}\{f(z_n) - f(y_n) - \delta f(\bar{y}_n, y_n)(z_n - y_n)\} = \\ &= \delta f(\bar{y}_n, y_n)^{-1}\{\delta f(z_n, y_n) - \delta f(\bar{y}_n, y_n)\}(z_n - y_n) \leqq A^{-1}\{\delta f(z_n, y_n) - \delta f(\bar{y}_n, y_n)\}(z_n - y_n). \end{aligned}$$

Using again (13) and the normality of the cone we get

$$\|z_n - y_{n+1}\| \leqq \mu_1 \|z_n - y_n\| \, \|z_n - \bar{y}_n\|. \tag{22}$$

Finally (20) and (22) imply the inequality stated in our theorem. □

Corollary 1: *Under the hypothesis of Theorem 1, the sequence $\{\|z_n - y_n\|\}$ of the diameters of the enclosing intervals produced by the iterative method (6) converges to zero Q-cubically.*

Corollary 2: *Under the hypothesis of Theorem 1, the sequence $\{\|z_n - y_n\|\}$ of the diameters of the enclosing intervals produced by the iterative method (7) converges to zero with R-order $1 + \sqrt{2}$.*

Proof: Note that in this case

$$\|z_n - \bar{y}_n\| = \|z_n - y_{n-1}\| \leqq \gamma \|z_{n-1} - y_{n-1}\| . \qquad \square$$

Proposition 4: *Under the hypothesis of Proposition* 2 *suppose that* f *is twice continuously differentiable on* $[a, b]$. *Then*

$$\lim_{n\to\infty} \frac{|z_{n+1} - y_{n+1}|}{|z_n - y_n|^2 \, |z_n - \bar{y}_n|} = \left| \frac{f''(x^*)}{2f'(x^*)} \right|^2 .$$

Proof: As in the proof of Proposition 3 we denote by $\delta^2 f$ the second divided difference of f. Using (19) and (8) we can write

$$z_{n+1} - y_{n+1} = -\delta f(\bar{y}_{n+1}, y_{n+1})^{-1} \, \delta^2 f(y_n, y_{n+1}, z_n) \, (y_{n+1} - z_n) \, (y_{n+1} - y_n) =$$
$$= \delta f(\bar{y}_{n+1}, y_{n+1})^{-1} \, \delta^2 f(y_n, y_{n+1}, z_n) \, \delta f(y_n, z_n)^{-1} f(y_n) \, (y_{n+1} - z_n) .$$

If we denote

$$\beta_n = \delta f(\bar{y}_{n+1}, y_{n+1})^{-1} \, \delta^2 f(y_n, y_{n+1}, z_n) \, \delta f(y_n, z_n)^{-2} ,$$

then according to (21) we have

$$z_{n+1} - y_{n+1} = -\beta_n f(y_n) \, f(z_n) . \tag{23}$$

From Proposition 2 it follows that $z_n = \bar{z}_n$. Therefore the third equation in (8) can be rewritten as

$$z_n = \bar{y}_n - \delta f(\bar{y}_n, y_n)^{-1} f(\bar{y}_n) . \tag{24}$$

Consequently

$$f(z_n) = f(z_n) - f(\bar{y}_n) - \delta f(\bar{y}_n, y_n) \, (z_n - \bar{y}_n) =$$
$$= \{\delta f(\bar{y}_n, z_n) - \delta f(\bar{y}_n, y_n)\} \, (z_n - \bar{y}_n) = \delta^2 f(\bar{y}_n, y_n, z_n) \, (z_n - y_n) \, (z_n - \bar{y}_n) .$$

Using again the third equation in (8) and $z_n = \bar{z}_n$ we get

$$f(y_n) = -\delta f(\bar{y}_n, y_n) \, (z_n - y_n) . \tag{25}$$

Substituting these expressions for $f(z_n)$, $f(y_n)$ in (23) gives

$$z_{n+1} - y_{n+1} = \beta_n \delta^2 f(\bar{y}_n, y_n, z_n) \, \delta f(\bar{y}_n, y_n) \, (z_n - y_n)^2 \, (z_n - \bar{y}_n) .$$

For $n \to \infty$ in this relation we obtain the desired result. $\square$

Corollary 3: *Under the hypothesis of Proposition* 4 *suppose that* $f''(x^*) \neq 0$. *Then the sequence* $\{|z_n - y_n|\}$ *of the diameters of the enclosing intervals produced by* (7) *converges to zero with Q-order* $1 + \sqrt{2}$.

Proof: For (7) we have $\bar{y}_n = y_{n-1}$ so that by using (24) and (25) we have

$$z_n - \bar{y}_n = z_n - y_{n-1} = -\delta f(y_{n-1}, y_n)^{-1} f(y_{n-1}) = \delta f(y_{n-1}, y_n)^{-1} \, \delta f(y_{n-2}, y_{n-1}) \, (z_{n-1} - y_{n-1}) .$$

Thus $\lim_{n\to\infty} (|z_n - \bar{y}_n| / |z_{n-1} - y_{n-1}|) = |f'(x^*) / f'(x^*)| = 1$ so that by using Proposition 4 we deduce that

$$\lim_{n\to\infty} \frac{|z_{n+1} - y_{n+1}|}{|z_n - y_n|^2 \, |z_{n-1} - y_{n-1}|} = \left| \frac{f''(x^*)}{2f'(x^*)} \right|^2 > 0 .$$

Our Corollary follows then from a result of [9]. $\square$

For the definition of the Q-order and the R-order of convergence and the relation between these two notations the reader may consult, for example, [8].

4. A numerical example

We will give only an example in the onedimensional case (a rather "ill-conditioned" one) to illustrate the importance of taking the infimum in (8). We want to find a monotone enclosure for $(1/11)^{1/11} \in [0.1, 1.0]$, by using the iterative procedures (6) and (7). It is easily seen that the function $f(x) = 11x^{11} - 1$ and the points $a = 0.1$, $b = 1.0$ satisfy the hypothesis of Theorem 1. The results obtained by applying (6) and (7) are given in tables 1 and 2.

We note that both methods take the infimum 11 times. After that method (6) needs 6 steps to attain full accuracy (18 digits) while method (7) needs 7 steps. (In the tables are only displayed 12 digits of the mantissa.) This is in accordance with the respective orders of convergence. We also note that without taking the infimum the solution would have been much slower (if not impossible because of the overflow). We have worked out several multidimensional examples (some of which arise from discretizing nonlinear integral equations or nonlinear two-point boundary value problems) that exhibit the same pattern of behaviour. However, taking the infimum in the multidimensional case has a more dramatic effect because in many cases z_{n+1} is sensibly closer to the root than both z_n and $\bar{z}_{n+1}$. We conjecture that there are multidimensional problems for which the infimum must be taken infinitely many times but we have not been able to find such an example so far.

Table 1. Results for method (6)

n	y_n	z_n	$\bar{z}_n$
1	0.181818181809	1	209335.010110
2	0.256198341760	1	6783.69629225
⋮			
10	0.649855750431	1	1.20601852658
11	0.678883880368	1	1.01459626279
12	0.703896078337	0.916682000648	0.916682000648
13	0.744859870517	0.834351242786	0.843351242789
14	0.792482522776	0.805028398613	0.805028398613
15	0.804066504121	0.804133125087	0.804133125087
16	0.804133097492	0.804133097503	0.804133097503
17	0.804133097503	0.804133097503	0.804133097503

Table 2. Results for method (7)

n	y_n	z_n	$\bar{z}_n$
1	0.181818181809	1	1037653.16044
2	0.256198341760	1	22173.4151338
⋮			
10	0.649855750431	1	1.36250590216
11	0.678883880368	1	1.09261792176
12	0.703896078337	0.957216014306	0.95721601430
13	0.733544662495	0.875553625862	0.87555362586
14	0.774865855929	0.820824556317	0.82082455631
15	0.801021492377	0.804746555927	0.80474655592
16	0.804121155031	0.804133330360	0.80413333036
17	0.804133097486	0.804133097503	0.80413309750
18	0.804133097503	0.804133097503	0.80413309750

Acknowledgements

1. This work is the result of a visit of the first author to the University of Iowa and of a visit of the second author to the University of Karlsruhe. We want to thank both Universities for their support.
2. We would like to thank Mr. MARKUS NEHER, University of Karlsruhe, for helping us with many numerical experiments.

References

1 DARBOUX, G.: Sur la méthode d'approximation de Newton. Nouv. Ann. de Math. ,2e sér. 8 (1869), 17–27.
2 KRASNOSELSKII, M. A.; ZABREIKO, P. P.: Geometrical methods of nonlinear analysis. Springer-Verlag, Berlin 1984.
3 OSTROWSKI, A. M.: Solution of equations in Euclidian and Banach spaces. Academic Press, New York 1973.
4 POTRA, F. A.; PTÁK, V.: Nondiscrete induction and iterative processes. Pitman, London 1984.
5 SCHMIDT, J. W.: Eingrenzung von Lösungen nichtlinearer Gleichungen durch Verfahren mit höherer Konvergenzgeschwindigkeit. Computing 8 (1971), 208–215.
6 SCHMIDT, J. W.; LEONHARDT, H.: Eingrenzung von Lösungen mit Hilfe der Regula falsi. Computing 6 (1970), 318–329.
7 SCHMIDT, J. W.: On the R-order of coupled sequences. Computing 26 (1981), 333–342.
8 SCHWETLICK, H.: Numerische Lösung nichtlinearer Gleichungen. Deutscher Verlag der Wissenschaften, Berlin 1979.
9 TAPIA, R. A.; WHITLEY, D.: On the superquadratic convergence of projected Newton for the symmetric eigenvalue problem. Techn. Rep. 87–14, May 1987. Math. Sci. Dept., Rice University, Houston, Texas.

Received October 8, 1987

Addresses: Prof. Dr. G. ALEFELD, Institut für Angewandte Mathematik, Universität Karlsruhe, D-7500 Karlsruhe, F.R.G.; Prof. Dr. F.A. POTRA, Department of Mathematics, The University of Iowa, Iowa City, IA 52242, U.S.A.