Computing 42, 69 – 80 (1989)

Computing

A New Class of Interval Methods with Higher Order of Convergence*

G. Alefeld, Karlsruhe, and **F. Potra,** Iowa City

Received October 17, 1988

Abstract — Zusammenfassung

A New Class of Interval Methods with Higher Order of Convergence. In this paper we introduce a new class of interval methods for enclosing a simple root of a nonlinear equation. For each nonnegative integer p we describe an iterative procedure belonging to this class which requires $p+1$ function values and an interval evaluation of the second derivative per step. The order of convergence of the iterative procedure grows exponentially with p. For $p \geq 4$ this order is strictly greater than

$$\left(\frac{1+\sqrt{5}}{2}\right)^{p+2}.$$

Key words: Nonlinear equations, order of convergence.

Eine neue Klasse von Intervall-Methoden mit höherer Konvergenzordnung. In dieser Arbeit wird eine neue Klasse von Intervall-Methoden zur Einschließung einfacher Wurzeln einer nichtlinearen Gleichung vorgestellt. Für eine gegebene ganze Zahl $p \geq 0$ wird ein Iterationsverfahren beschrieben, das $p+1$ Funktionsauswertungen pro Schritt sowie eine Intervallauswertung der zweiten Ableitung benötigt. Die Konvergenzordnung des Iterationsverfahrens wächst exponentiell mit p. Für $p \geq 4$ ist die Ordnung größer als

$$\left(\frac{1+\sqrt{5}}{2}\right)^{p+2}.$$

1. Introduction

Interval arithmetic provides a useful tool for constructing root finders with global convergence and automatic error bounds. Interval arithmetic is nothing else but a natural extension of the usual arithmetic between numbers to an arithmetic between intervals. Let $A=[a_1, a_2]$, $B=[b_1, b_2]$ be two bounded and closed intervals of the real line and let * denote any of the arithmetic operations $+, -, \times, \div$. Then the corresponding operation for intervals is defined as

$$A * B = \{a * b;\ a \in A,\ b \in B\}. \tag{1.1}$$

For the properties of interval arithmetic we refer to Alefeld-Herzberger (1983).

* This paper was written while the second author was visiting the University of Karlsruhe.

The set of all bounded and closed intervals of the real line is denoted by $I(\mathbb{R})$. Any real number x is identified with the trivial interval $[x,x]=\{x\}\in I(\mathbb{R})$. Besides the above introduced arithmetic operations, in $I(\mathbb{R})$ we may also define the intersection

$$A\cap B=\{x;\ x\in A \text{ and } x\in B\}, \tag{1.2}$$

as well as the complementary operation

$$A\vee B=[\min\{a_1,b_1\},\ \max\{a_2,b_2\}]. \tag{1.3}$$

If we consider the set $I(\mathbb{R})$ partially ordered by the inclusion $\subseteq$, then (1.2) and (1.3) are the corresponding lattice operations

$$A\cap B=\inf\{A,B\},\ A\vee B=\sup\{A,B\}.$$

With any interval $A=[a_1,a_2]$ we associate the following three important quantities: the diameter of A

$$d(A)=|a_2-a_1|, \tag{1.4}$$

the absolute value of A

$$|A|=\max\{|a_1|,|a_2|\}, \tag{1.5}$$

the midpoint of A

$$m(A)=(a_1+a_2)/2. \tag{1.6}$$

In the proof of our main result we will frequently use the following relations concerning the diameter and absolute value of an interval.

$$A\subseteq B\Rightarrow d(A)\leq d(B), \tag{1.7}$$

$$d(A\pm B)=d(A)+d(B), \tag{1.8}$$

$$|A+B|\leq|A|+|B|,\ |AB|=|A|\,|B|, \tag{1.9}$$

$$d(AB)\leq|A|\,d(B)+d(A)\,|B|. \tag{1.10}$$

For details see Alefeld-Herzberger (1983).

The notation introduced above is sufficient for understanding the numerical methods to be presented in this paper. However, we strongly advise the reader with no background in interval arithmetic to consult Alefeld-Herzberger (1983). Also for the readers who are interested in applications we recommend the use of PASCAL-SC, a powerful language which supports interval arithmetic (see G. Bohlender et al. (1987)). We mention that the implementation of interval arithmetic on a digital computer is done in such a way that the interval solution resulting from interval arithmetic operations includes all rounding errors which may have occured because of the finite mantissa. Thus interval arithmetic controls the accumulation of the rounding errors, a property which is not shared by the usual floating point arithmetic.

In Section 2 we review some well known results on the interval Newton method and on a class of Newton-like methods which require $p+1$ function values and one interval evaluation of the first derivative per iteration, and which are convergent with Q-order $p+2$. For $p=0$ this reduces to the interval Newton method.

In Section 3 we present a new class of iterative procedures for enclosing roots of nonlinear equations which require $p+1$ function values and one interval evaluation of the second derivative. The corresponding order of convergence grows exponentially with p. For $p \geq 4$ the order is strictly greater than

$$\left(\frac{1+\sqrt{5}}{2}\right)^{p+2}.$$

Section 4 contains some variants of the iterative procedures described in Sections 2 and 3 together with the results of some numerical experiments.

2. The Interval Newton Method

Let us consider the nonlinear equation

$$f(x)=0 \tag{2.1}$$

where f is a continuously differentiable real function of a real variable. We suppose that f is strictly monotone on an interval $X^{(0)} \in I(\mathbb{R})$. Without loss of generality we may assume that f is strictly increasing on $X^{(0)}$. We assume that by using interval arithmetic methods it is possible to compute two positive numbers l_1, l_2 such that

$$0 < l_1 \leq f'(x) \leq l_2, \text{ for all } x \in X^{(0)}. \tag{2.2}$$

Let us denote by L the interval $[l_1, l_2]$. We suppose that the derivative $f'(x) \in \mathbb{R}$, $x \in X^{(0)}$ has an interval extension $f'(X) \in I(\mathbb{R})$, $X \subseteq X^{(0)}$ satisfying the following conditions

$$f'(x) \in f'(X), \text{ for all } x \in X \subseteq X^{(0)}, \tag{2.3}$$

$$f'(X) \subseteq f'(Y), \text{ whenever } X \subseteq Y \subseteq X^{(0)}, \tag{2.4}$$

$$d(f'(X)) \leq c\, d(X), \text{ for all } X \subseteq X^{(0)}, \tag{2.5}$$

where c is a constant independent of X.

We note that condition (2.5) implies the Lipschitz continuity of the point derivative $f'(x)$. With the above notation the interval Newton method can be defined as follows.

Algorithm N_0: for $k=0,1,\ldots$ DO through ES

$$M^{(k)} = f'(X^{(k)}) \cap L, \quad x^{(k)} = m(X^{(k)})$$

ES $\quad X^{(k+1)} = \{x^{(k)} - f(x^{(k)})/M^{(k)}\} \cap X^{(k)}$

In the above algorithm $x^{(k)}$ can be chosen to be any point of the interval $X^{(k)}$. However, the midpoint $x^{(k)} = m(X^{(k)})$ is a natural choice which is also optimal in some sense (see Alefeld-Herzberger (1983) pp. 72–75).

Theorem 1: *Let $f: X^{(0)} \to \mathbb{R}$ be a continuously differentiable function whose derivative has an interval extension satisfying* (2.2)–(2.5). *Suppose that the equation* (2.1) *has a solution $x^* \in X^{(0)}$. Then the sequence of intervals $\{X^{(k)}\}$ given by the Algorithm N_0 satisfies*

$$x^* \in X^{(k)} \textit{ for all } k \geq 0, \tag{2.6}$$

$$X^{(0)} \supseteq X^{(1)} \supseteq \ldots, \lim_{k \to \infty} X^{(k)} = x^*. \tag{2.7}$$

Moreover the convergence is quadratic in the sense that the sequence of diameters $\{d(X^{(k)})\}$ *converges Q-quadratically to zero.* □

We note that, unlike the classical Newton method, the interval Newton procedure can be used as a test for the existence or the nonexistence of a solution of (2.1) in $X^{(0)}$. Thus if (2.2)–(2.5) are satisfied then the following statements hold.

(a) equation (2.1) has no solution in $X^{(0)}$ if and only if there is a $k \geq 1$ such that

$$X^{(k)} = \emptyset = \text{the empty set}; \tag{2.8}$$

(b) if there is a $k \geq 0$ such that

$$\{x^{(k)} - f(x^{(k)})/M^{(k)}\} \subseteq X^{(k)} \tag{2.9}$$

then the equation (2.1) has a solution in $X^{(k)}$.

We mention that in case (2.1) has no solution in $X^{(0)}$ then the situation (2.8) will happen rather rapidly (see Alefeld (1988)).

When implementing the interval Newton method on a computer the value $f(x^{(k)})$ has to be evaluated by using an interval extension of the function f. Otherwise, the intersection performed in algorithm N_0 may become empty because of the effect of rounding errors, although (2.1) has a solution in $X^{(0)}$. Therefore, if we consider that the cost of computing $f(x^{(k)})$ and $f'(X^{(k)})$ is about the same, then the efficiency index, in the sense of Ostrowski (1960), of Newton's method is

$$\text{eff}(N_0) = \sqrt{2} = 1.4\ldots. \tag{2.10}$$

We can increase the efficiency index by considering the following class of iterative procedures which are obtained from the interval Newton method by using the same interval derivative for $p+1$ substeps.

Algorithm N_p: for $k=0,1,\ldots$ DO through ES

$X^{(k,0)} = X^{(k)}$, $M^{(k)} = f'(X^{(k)}) \cap L$

for $i=0,1,\ldots,p$ DO through $E2$

$x^{(k,i)} = m(X^{(k,i)})$

$E2$ $\quad X^{(k,i+1)} = \{x^{(k,i)} - f(x^{(k,i)})/M^{(k)}\} \cap X^{(k,i)}$

ES $\quad X^{(k+1)} = X^{(k,p+1)}$

Theorem 2: *Suppose that the hypothesis of Theorem 1 is satisfied. Then the sequence of intervals generated by Algorithm* N_p *satisfies (2.6), (2.7). Moreover, the sequence of diameters* $\{d(X^{(k)})\}$ *converges to zero with Q-order* $p+2$. □

For $p=0$ Theorem 2 reduces to Theorem 1. Algorithm N_p requires $p+1$ function values plus one interval evaluation of the first derivative per iteration step. Its efficiency index is given by

$$\text{eff}(N_p) = \sqrt[p+2]{p+2}. \tag{2.11}$$

It is easily seen that the highest efficiency index is obtained for $p=1$

$$\text{eff}(N_1) = \sqrt[3]{3} = 1.4422496\ldots. \tag{2.12}$$

We also have

$$\mathrm{eff}(N_1) > \mathrm{eff}(N_0) = \mathrm{eff}(N_2) > \mathrm{eff}(N_3) > \mathrm{eff}(N_4) > \dots. \quad (2.13)$$

Of course all this holds under the assumption that the cost of $f(x^{(k)})$ is about the same as the cost of $f'(X^{(k)})$. In general the optimal value of the parameter p is to be established by experimental and theoretical considerations. In the next section we will see that by using $p+1$ function values and one interval evaluation of the *second* derivative we obtain a much higher efficiency index than (2.11).

3. A Class of Secant Type Interval Methods with Higher Convergence Order

Throughout this section we assume that the function f is twice continuously differentiable on $X^{(0)}$ and that its second derivative $f''(x) \in \mathbb{R}, x \in X^{(0)}$ has an interval extension $f''(X) \in I(\mathbb{R}),\ X \subseteq X^{(0)}$ satisfying

$$f''(x) \in f''(X), \text{ for all } x \in X \subseteq X^{(0)}, \quad (3.1)$$

$$f''(X) \subseteq f''(Y), \text{ whenever } X \subseteq Y \subseteq X^{(0)}, \quad (3.2)$$

$$d(f''(X)) \leq c\, d(X), \text{ for all } X \subseteq X^{(0)}, \quad (3.3)$$

where the constant c is independent of X.

Together with f and its derivative we will also use its divided differences

$$f[x,y] = \begin{cases} (f(x)-f(y))/(x-y) & \text{if } x \neq y \\ f'(x) & \text{if } x = y, \end{cases} \quad (3.4)$$

$$f[x,y,z] = \begin{cases} (f[x,z]-f[y,z])/(x-y) & \text{if } x \neq y \\ f''(x)/2 & \text{if } x = y. \end{cases} \quad (3.5)$$

As in the previous section we assume that condition (2.2) is satisfied. Then for any nonnegative integer p we can define the following iterative procedure.

Algorithm S_p: for $k=0,1,\dots$ DO through ES

$x^{(k)} = m(X^{(k)})$

if $k=0$ then $Q^{(k)} = L$ & GOTO $E1$

else

$M^{(k)} = \{f[x^{(k)}, x^{(k-1,p)}] + \frac{1}{2} f''(X^{(k-1)})(X^{(k)} - x^{(k-1,p)})\} \cap L$

$Y^{(k)} = \{x^{(k)} - f(x^{(k)})/M^{(k)}\} \cap X^{(k)}$

$Q^{(k)} = \{f[x^{(k)}, x^{(k-1,p)}] + \frac{1}{2} f''(X^{(k-1)})(Y^{(k)} - x^{(k-1,p)})\} \cap L$

$E1 \quad X^{(k,1)} = \{x^{(k)} - f(x^{(k)})/Q^{(k)}\} \cap X^{(k)}$

if $p=0$ then $x^{(k,p)} = x^{(k)}$ & GOTO ES

else

$x^{(k,1)} = m(X^{(k,1)})$

$M^{(k,1)} = \{f[x^{(k)}, x^{(k,1)}] + \frac{1}{2} f''(X^{(k)})(X^{(k,1)} - x^{(k)})\} \cap L$

$Y^{(k,1)} = \{x^{(k,1)} - f(x^{(k,1)})/M^{(k,1)}\} \cap X^{(k,1)}$

$Q^{(k,1)} = \{f[x^{(k)}, x^{(k,1)}] + \frac{1}{2} f''(X^{(k)})(Y^{(k,1)} - x^{(k)})\} \cap L$

$X^{(k,2)} = \{x^{(k,1)} - f(x^{(k,1)})/Q^{(k,1)}\} \cap X^{(k,1)}$

if $p=1$ then GOTO ES

else for $i=2, 3, \ldots, p$ DO through $E2$

$x^{(k,i)} = m(X^{(k,i)})$

$M^{(k,i)} = \{f[x^{(k,i-1)}, x^{(k,i)}] + \frac{1}{2} f''(X^{(k)})(X^{(k,i)} - x^{(k,i-1)})\} \cap L$

$E2 \quad X^{(k,i+1)} = \{x^{(k,i)} - f(x^{(k,i)})/M^{(k,i)}\} \cap X^{(k,i)}$

$ES \quad X^{(k+1)} = X^{(k,p+1)}$

Theorem 3: *Let $f: X^{(0)} \to \mathbb{R}$ be twice continuously differentiable. Suppose its first derivative satisfies condition* (2.2) *and that its second derivative has an interval extension satisfying* (3.1)–(3.3). *Assume also that the equation* (2.1) *has a root $x^* \in X^{(0)}$. Then the sequence of intervals generated by Algorithm S_p satisfies* (2.6), (2.7). *Moreover, the sequence of diameters $\{d(X^{(k)})\}$ converges to zero with R-order w_p defined as*

$$w_p = (f_p + 2f_{p+2} - 1 + \sqrt{12 f_{p+2}^2 + 9 f_p^2 - 20 f_p f_{p+2} - 4 f_{p+2} + 2 f_p + 1})/2 \quad (3.6)$$

where f_j denotes the j-th Fibonacci number, i.e.

$$f_0 = 0, \ f_1 = 1, \ f_{i+1} = f_i + f_{i-1}, \ i = 1, 2, \ldots. \quad (3.7)$$

Proof: We wirst prove (2.6). This is clearly satisfied for $k=0$. Suppose it is satisfied for a certain $k \geq 0$. To prove that (2.6) holds with k replaced by $k+1$ we use the identity

$$x^* = x - f(x)/f[x, x^*].$$

Then for any point u we can write

$$x^* = x - f(x)/\{f[x, u] + f[x, u, x^*](x^* - u)\}.$$

From the mean value theorem for second order divided differences it follows that

$$x^* \in x - f(x)/\{f[x, u] + \tfrac{1}{2} f''(U)(V - u)\},$$

for any intervals U, V satisfying

$$x, u, x^* \in U \subseteq X^{(0)}, \quad x^* \in V \subseteq X^{(0)}.$$

By using this fact we obtain successively

$$x^* \in Y^{(k)},\ x^* \in X^{(k,1)},\ x^* \in X^{(k,i+1)},\ i=1,\dots,p.$$

For $i=p$ this reduces to

$$x^* \in X^{(k,p+1)} = X^{(k+1)}$$

which proves that (2.6) holds for any $k \geq 0$.

The inclusions from (2.7) are obvious. In fact, we have

$$X^{(k)} \supseteq Y^{(k)} \supseteq X^{(k,1)} \supseteq Y^{(k,1)} \supseteq X^{(k,2)} \supseteq \dots \supseteq X^{(k,p+1)} = X^{(k+1)}.$$

Because $x^{(k)}$ is the midpoint of $X^{(k)}$ and by (2.2) it follows that

$$d(X^{(k+1)}) \leq d(Y^{(k)}) \leq d\left(x^{(k)} - f(x^{(k)})/L\right) \leq \tfrac{1}{2} d(X^{(k)}).$$

According to (2.6) this implies that $\lim_{k\to\infty} X^{(k)} = x^*$.

Let us now establish the order of convergence of the sequence $\{d(X^{(k)})\}$. It will be convenient to denote by c a generic constant which depends on f and p but is independent of k. Because the value of this constant does not matter we convene to write $c+c=c$, $c \times c = c$ etc. as long as we perform a finite number of operations (the number of operations may depend on p but not on k). In our majorizations we will often use the following obvious facts

$$|f(x)| \leq l_2\, d(X) \text{ whenever } x, x^* \in X \subseteq X^{(0)}, \tag{3.8}$$

$$d(1/S) \leq d(S)/l_1^2 \text{ whenever } S \subseteq L. \tag{3.9}$$

Let us fix an arbitrary $k \geq 1$ and denote

$$X^{(k,0)} = X^{(k)},\ \xi_{k,i} = d(X^{(k,i)}),\ i=0,1,\dots,p+1;$$

$$\xi_k = d(X^{(k)}),\ \eta_k = d(X^{(k-1,p)}),\ \zeta_k = d(X^{(k-1)}).$$

Clearly,

$$\xi_k = \xi_{k,0} = \xi_{k-1,p+1},\ \eta_k = \xi_{k-1,p},\ \zeta_k = \xi_{k-1}.$$

From the definition of $M^{(k)}$ we have

$$d(M^{(k)}) \leq \tfrac{1}{2} d\left(f''(X^{(k-1)})\right) |X^{(k)} - x^{(k-1,p)}| + \tfrac{1}{2} |f''(X^{(k-1)})|\, d(X^{(k)} - x^{(k-1,p)}).$$

Because

$$X^{(k)} \subseteq X^{(k-1,p)},\ x^{(k-1,p)} \in X^{(k-1,p)},\ d\left(f''(X^{(k-1)})\right) \leq c\, d(X^{(k-1)}),$$

$$|f''(X^{(k-1)})| \leq |f''(X^{(0)})| = c$$

we can write

$$d(M^{(k)}) \leq c(\xi_{k-1}\eta_k + \xi_k) = c(\eta_k \zeta_k + \xi_k).$$

By using (3.8) and (3.9) we obtain

$$d(Y^{(k)}) \leq c\, d(M^{(k)})\, d(X^{(k)}) \leq c(\eta_k \zeta_k + \xi_k)\, \xi_k.$$

In a similar manner we deduce that

$$d(Q^{(k)}) \le c\big(\zeta_k \eta_k + d(Y^{(k)})\big) \le c\big(\eta_k \zeta_k + (\eta_k \zeta_k + \xi_k)\xi_k\big) \le c\,\eta_k \zeta_k,$$
$$\xi_{k,1} = d(X^{(k,1)}) \le c\, d(Q^{(k)})\, d(X^{(k)}) \le c\, \xi_k \eta_k \zeta_k,$$
$$d(M^{(k,1)}) \le c\big(\xi_k^2 + d(X^{(k,1)})\big) \le c(\xi_k^2 + \xi_k \eta_k \zeta_k),$$
$$d(Y^{(k,1)}) \le c\, d(M^{(k,1)})\, d(X^{(k,1)}) \le c(\xi_k^2 + \xi_k \eta_k \zeta_k)\, \xi_k \eta_k \zeta_k,$$
$$d(Q^{(k,1)}) \le c\big(\xi_k^2 + d(Y^{(k,1)})\big) \le c\, \xi_k^2,$$
$$\xi_{k,2} = d(X^{(k,2)}) \le c\, d(Q^{(k,1)})\, d(X^{(k,1)}) \le c\, \xi_k^3 \eta_k \zeta_k.$$

We will prove that

$$\xi_{k,j} \le c\, \xi_k^{2f_{j+1}-1}\, \eta_k^{f_j} \zeta_k^{f_j}, \; j = 0, 1, \ldots, p+1. \tag{3.10}$$

This has been already verified for $j = 0, 1, 2$.

Suppose that (3.10) holds for $j = 0, 1, \ldots, i$ with $2 \le i \le p$. From the last equations in algorithm S_p we deduce that

$$d(M^{(k,i)}) \le c\big(d(X^{(k)})\, d(X^{(k,i-1)}) + d(X^{(k,i)})\big) \le$$
$$\le c(\xi_k^{2f_i} \eta_k^{f_{i-1}} \zeta_k^{f_{i-1}} + \xi_k^{2f_{i+1}-1} \eta_k^{f_i} \zeta_k^{f_i}) \le c\, \xi_k^{2f_i} \eta_k^{f_{i-1}} \zeta_k^{f_{i-1}},$$
$$\xi_{k,i+1} = d(X^{(k,i+1)}) \le c\, d(M^{(k,i)})\, d(X^{(k,i)}) \le$$
$$\le c\, \xi_k^{2f_i} \eta_k^{f_{i-1}} \zeta_k^{f_{i-1}}\, \xi_k^{2f_{i+1}-1}\, \eta_k^{f_i} \zeta_k^{f_i} = c\, \xi_k^{2f_{i+2}-1}\, \eta_k^{f_{i+1}} \zeta_k^{f_{i+1}}.$$

This ends the proof of (3.10). By writing (3.10) for $j = p$, $j = p+1$ and using the fact that

$$\xi_{k,p} = \eta_{k+1}, \; \xi_{k,p+1} = \xi_{k+1}, \; \zeta_{k+1} = \xi_k$$

we obtain

$$\xi_{k+1} \le c\, \xi_k^{2f_{p+2}-1}\, \eta_k^{f_{p+1}} \zeta_k^{f_{p+1}},$$
$$\eta_{k+1} \le c\, \xi_k^{2f_{p+1}-1}\, \eta_k^{f_p} \zeta_k^{f_p},$$
$$\zeta_{k+1} = \xi_k.$$

According to a result of J. W. Schmidt (1981) on the R-order of convergence of coupled sequences it follows that the R-order of convergence of the sequence $\{\xi_k\}$ is equal to the largest positive eigenvalue of the matrix

$$\begin{bmatrix} 2f_{p+2}-1 & f_{p+1} & f_{p+1} \\ 2f_{p+1}-1 & f_p & f_p \\ 1 & 0 & 0 \end{bmatrix}$$

It is easily verified that this is exactly w_p. The proof of the theorem is complete. □

As with the interval Newton method we note that in the numerical implementation of algorithm S_p one has to use interval extensions for the function values $f(x^{(k,i)})$ because otherwise the intervals $X^{(k,i)}$ may become empty although the equation (2.1) has a solution in $X^{(0)}$. Under the assumption that the cost of an interval evaluation of the second derivative is about the same as a function evaluation, the efficiency index of the algorithm S_p is given by

$$\mathrm{eff}(S_p) = \sqrt[p+2]{w_p}. \tag{3.11}$$

In Table 1 we give the values of w_p and

$$\sqrt[p+2]{w_p} \text{ for } p=0,1,2,\ldots,10.$$

Table 1

p	w_p	$\sqrt[p+2]{w_p}$
0	2.00000000000 $E+00$	1.41421356237 $E+00$
1	3.73205080757 $E+00$	1.55113351807 $E+00$
2	6.46410161514 $E+00$	1.59450925267 $E+00$
3	1.10000000000 $E+01$	1.61539426620 $E+00$
4	1.82736184955 $E+01$	1.62294608383 $E+00$
5	3.00996688705 $E+01$	1.62638403519 $E+00$
6	4.92032386541 $E+01$	1.62741835990 $E+00$
7	8.01372644808 $E+01$	1.62756060099 $E+00$
8	1.30176682947 $E+02$	1.62724640258 $E+00$
9	2.11151549918 $E+02$	1.62677223759 $E+00$
10	3.42166585524 $E+02$	1.62624684244 $E+00$

It can be proved that

$$w_p > \left(\frac{1+\sqrt{5}}{2}\right)^{p+2} \quad \text{for } p \geq 4, \tag{3.12}$$

$$\lim_{p\to\infty} \sqrt[p+2]{w_p} = \frac{1+\sqrt{5}}{2}, \tag{3.13}$$

$$\max_{p\geq 0} \sqrt[p+2]{w_p} = \sqrt[9]{w_7} = 1.62756\ldots. \tag{3.14}$$

Thus, under the assumption that the cost of $f(x)$ is about the same as the cost of $f''(X)$ we have maximum efficiency for $p=7$. Of course this result has to be understood only in an asymptotic sense. In practice the optimal value of the parameter p is to be determined through numerical experiments. Nevertheless, we remark that the convergence order of the algorithm S_p grows exponentially with p while that of the algorithm N_p grows only linearly with p.

4. Practical Procedures and Some Numerical Results

In this section we give some modifications of the algorithms N_p and S_p. With the same number of function evaluations per step, but with a slight increase in the number of arithmetic operations we obtain enclosing intervals which are in most cases contained in the corresponding intervals given by N_p and S_p. Although this will not change the order of convergence of the respective procedures we will obtain in many cases a faster reduction of the diameters of the enclosing intervals, especially during the first iterations. For a motivation of such modifications see Alefeld-Herzberger (1983, pp. 76–78).

Algorithm MN_p: set $M^{(-1)}=L$

for $k=0,1,\ldots$ DO through ES

$x^{(k)}=m(X^{(k)}),\ x^{(k,0)}=x^{(k)},\ X^{(k,0)}=X^{(k)}$

$Y^{(k)}=\{x^{(k)}-f(x^{(k)})/M^{(k-1)}\}\cap X^{(k)}$

$Z^{(k)}=Y^{(k)}\vee x^{(k)},\ M^{(k)}=f'(Z^{(k)})\cap L$

for $i=0,1,\ldots,p$ DO through $E2$

$E2$ $X^{(k,i+1)}=\{x^{(k,i)}-f(x^{(k,i)})/M^{(k)}\}\cap X^{(k,i)}$

ES $X^{(k+1)}=X^{(k,p+1)}$

Algorithm MS_p: for $k=0,1,\ldots$ DO through ES

$x^{(k)}=m(X^{(k)})$

if $k=0$ then $Q^{(k)}=L$ & GOTO $E1$

else

$M^{(k)}=\{f[x^{(k)},x^{(k-1,p)}]+T^{(k-1)}(X^{(k)}-x^{(k-1,p)})\}\cap L$

$Y^{(k)}=\{x^{(k)}-f(x^{(k)})/M^{(k)}\}\cap X^{(k)}$

$Q^{(k)}=\{f[x^{(k)},x^{(k-1,p)}]+T^{(k-1)}(Y^{(k)}-x^{(k-1,p)})\}\cap L$

$E1$ $X^{(k,1)}=\{x^{(k)}-f(x^{(k)})/Q^{(k)}\}\cap X^{(k)}$

$x^{(k,0)}=x^{(k)}$

$U^{(k)}=X^{(k,1)}\vee x^{(k)},\ T^{(k)}=\frac{1}{2}f''(U^{(k)})$

if $p=0$ then GOTO ES

else

for $i=1,2,\ldots,p$ DO through $E2$

$x^{(k,i)}=m(X^{(k,i)})$

$M^{(k,i)}=\{f[x^{(k,i-1)},x^{(k,i)}]+T^{(k)}(X^{(k,i)}-x^{(k,i-1)})\}\cap L$

$Y^{(k,i)}=\{x^{(k,i)}-f(x^{(k,i)})/M^{(k,i)}\}\cap X^{(k,i)}$

$Q^{(k,i)}=\{f[x^{(k,i-1)},x^{(k,i)}]+T^{(k)}(Y^{(k,i)}-x^{(k,i-1)})\}\cap L$

$E2$ $X^{(k,i+1)}=\{x^{(k,i)}-f(x^{(k,i)})/Q^{(k,i)}\}\cap Y^{(k,i)}$

ES $X^{(k+1)}=X^{(k,p+1)}$

It is easy to prove that Theorem 2 and Theorem 3 remain valid if we replace the algorithms N_p, S_p by the algorithms MN_p, MS_p, respectively. We have applied the algorithms N_p, MN_p, S_p, MS_p to the following example considered in Alefeld-Herzberger (1983, pp. 77–78).

Example: $f(x)=x^2(\frac{1}{3}x^2+\sqrt{2}\sin x)-\sqrt{3}/19$, $X^{(0)}=[0.1, 1]$.

We have stopped the iterative procedure when the diameter of some inclosing interval became less than 10^{-10}. The number of function evaluations ($\# f$), as well as the number of interval evaluations of the first derivative ($\# f'$) or the number of interval evaluations of the second derivative ($\# f''$) are given in Table 2.

Table 2

p	N_p		MN_p		S_p		MS_p	
	$\# f$	$\# f'$	$\# f$	$\# f'$	$\# f$	$\# f''$	$\# f$	$\# f''$
0	6	6	5	5	6	5	5	4
1	7	4	6	3	6	3	5	2
2	8	3	7	3	6	2	5	2
3	8	2	7	2	7	2	6	2
4	9	2	7	2	7	2	6	1
5	9	2	8	2	7	1	6	1
6	10	2	9	2	7	1	6	1
7	10	2	10	2	7	1	6	1

Acknowledgement

The authors would like to thank Mr. Bernd Illg for providing many numerical experiments related to the subject of the present paper.

References

[1] Alefeld, G., Herzberger, J.: Introduction to Interval Computations. New York: Academic Press 1983.

[2] Alefeld, G.: Über die Divergenzgeschwindigkeit des Intervall-Newton-Verfahrens. Submitted for publication (1988).

[3] Bohlender, G., Ullrich, C., Wolff von Guldenberg, J., Rall, L. B.: PASCAL-SC. New York: Academic Press (1987).

[4] Ostrowski, A. M.: Solution of Equations and Systems of Equations. New York: Academic Press 1960.

[5] Schmidt, J. W.: On the *R*-order of coupled sequences. Computing *26,* 333–342 (1981).

G. Alefeld
Institut für Angewandte Mathematik
Universität Karlsruhe
Kaiserstrasse 12
D-7500 Karlsruhe
Federal Republic of Germany

F. Potra
Department of Mathematics
University of Iowa
Iowa City, 11752242
U.S.A.

Printed in Austria

Druck: Paul Gerin, A-1021 Wien