BIT
1998, Vol. 38, No. 4, pp. 619–635

0006-3835/98/3804-0619 $12.00

MODIFICATIONS OF THE INTERVAL-NEWTON-METHOD WITH IMPROVED ASYMPTOTIC EFFICIENCY *

G. E. ALEFELD[1], F. A. POTRA[2] and W. VÖLKER[1]

[1] *Institute für Angewandte Mathematik, Universität Karlsruhe, D-76128 Karlsruhe Germany. email: goetz.alefeld@mathematik.uni-karlsruhe.de*

[2] *Department of Mathematics, University of Iowa, Iowa City, IA 52242 USA. email: potra@math.uiowa.edu*

Dedicated to U. Kulisch, Karlsruhe on the occasion of his 65th birthday

Abstract.

In this paper three new methods are introduced which compute lower and upper bounds of a simple zero of a real function. The lower and upper bounds are converging to this zero. Compared with the well-known Interval-Newton-Method, which has the same properties and asymptotic efficiency 1.414... our optimal method has asymptotic efficiency 1.839.... The new methods have been extensively tested on a large set of test examples.

AMS subject classification: 65G10, 65H05.

Key words: Nonlinear algebraic or transcendental equations, single equations, Interval-Newton-method, modifications.

1 Introduction.

The idea of the Interval-Newton-Method was first discussed in [12]. In [6] a careful investigation of its properties was presented. For a recent discussion see, for example, [4].

Given a real differentiable function f, which has for any interval $[x]$ an interval arithmetic evaluation $f'([x])$ of its derivative, the method reads as follows:

$$\left.\begin{array}{l} \text{Set } [x_0] := [x] \\ \text{Choose } x_n \in [x_n] \\ [x_{n+1}] = \left\{ x_n - \dfrac{f(x_n)}{f'([x_n])} \right\} \cap [x_n] \end{array}\right\} n = 0, 1, 2, \ldots.$$

The method computes a sequence $\{[x_n]\}$ of intervals with the following properties: Assume that there exists a zero x^* of f in $[x_0]$. Provided $0 \notin f'([x_0])$ the method is convergent to x^*, all iterates contain x^*, and under natural additional

*Received July 1997. Communicated by Kaj Madsen.

assumptions on f the R-order of convergence is at least two. See [4]. For more details on interval arithmetic see [1].

If the method is performed on a computer the function value $f(x_n)$ has to be computed by taking into account all rounding errors in order that the inclusion of the zero is guaranteed. This can be achieved by computing $f(x_n)$ in the interval arithmetic sense. Therefore in practice $f(x_n)$ and $f'([x_n])$ together may be considered as two "function values" each of which needs approximately the same amount of work.

The effectivity index (or the asymptotic efficiency) of an iterative method in the sense of Ostrowski is defined as $\sqrt[p]{q}$ where q is the order of convergence and p is the total number of function values per each step. If a fixed absolute error ϵ and two iterative methods are given then asymptotically the method with the higher effectivity index needs less work to reach the given precision. For a discussion of the effectivity index see [10].

From the preceding discussion it follows that the Interval-Newton-Method has the effectivity index at least $\sqrt{2} = 1.414\ldots$. This result is independent of how $x_n \in [x_n]$ is chosen in each step. Usually one chooses x_n to be the center of $[x_n]$. In this paper we show that by choosing $x_n \in [x_n]$ appropriately, the Interval-Newton-Method can be modified in such a manner that the effectivity index is increased. We introduce three methods which have the same convergence and inclusion properties. The first method has its asymptotic efficiency also equal to $\sqrt{2}$. However, the second one has asymptotic efficiency $(1+\sqrt{5})/2 = 1.618\ldots$ and for the third method we get the value $1.839\ldots$. Extensive numerical tests confirm these values. The results of this paper have already been presented without proofs and with fewer numerical examples in [3].

2 The modification.

Before we describe the new methods we introduce some notations. Let $f : D \subseteq \mathbf{R} \to \mathbf{R}$ be a differentiable mapping and assume that for the interval $[x_0] \subseteq D$ an interval arithmetic evaluation of its derivative exists and does not contain zero:

$$0 \notin \Delta = f'([x_0]).$$

Let

$$\alpha = \min_{\delta\in\Delta} |\delta|, \quad \beta = \max_{\delta\in\Delta} |\delta|$$

and

$$\sigma = \operatorname{sign}(\Delta)$$

where

$$\operatorname{sign}(\Delta) = \begin{cases} 1 & \text{if } \delta > 0 \text{ for all } \delta \in \Delta \\ -1 & \text{if } \delta < 0 \text{ for all } \delta \in \Delta. \end{cases}$$

If $[x] = [\underline{x}, \overline{x}]$ is a given interval and $t \in \mathbf{R}$ then we define

$$P_{[x]}(t) := \begin{cases} t & \text{if } t \in [x] \\ \underline{x} & \text{if } t < \underline{x} \\ \overline{x} & \text{if } t > \overline{x} \end{cases}$$

and

$$d[x] := \overline{x} - \underline{x}$$

denotes the diameter of $[x]$.

We set

$$P_{n+1}(t) := P_{[x_{n+1}]}(t) \; .$$

For a given interval $[x] = [\underline{x}, \overline{x}]$ let

$$\text{mid}[x] = \tfrac{1}{2}(\underline{x} + \overline{x})$$

and

$$\text{bisect}([x]) = \begin{cases} [\underline{x}, \text{mid}[x]] & \text{if } \sigma \cdot f(\text{mid}[x]) > 0 \\ [\text{mid}[x], \overline{x}] & \text{if } \sigma \cdot f(\text{mid}[x]) < 0 \end{cases} \; .$$

Let $[x_n]$ denote an infinite sequence of intervals $[x_n] = [\underline{x}_n, \overline{x}_n]$ and assume $m_n \in [x_n] \subseteq D$. Then we set $f_n := f(m_n)$. Instead of $d[x_n]$ we also write d_n if it is clear which interval is meant.

In the following δ_n will denote either $f'(m_n)$ or some approximations to this value. More precisely, in our algorithms δ_n will be defined by one of the following three formulae:

$$\delta_n = f'(m_n), \tag{2.1}$$

$$\delta_n := f[m_n, m_{n-1}] := \frac{f(m_n) - f(m_{n-1})}{m_n - m_{n-1}}, \tag{2.2}$$

and

$$\delta_n = \begin{cases} \hat{\delta}_n & \text{if } \sigma\hat{\delta}_n \geq \dfrac{\alpha}{2} \\ \sigma\dfrac{\alpha}{2} & \text{otherwise} \end{cases} \tag{2.3}$$

where

$$\hat{\delta}_n = f[m_n, m_{n-1}] + f[m_n, m_{n-2}] - f[m_{n-1}, m_{n-2}]. \tag{2.4}$$

In (2.2) the derivative $f'(m_n)$ is approximated by the usual difference quotient. The approximation $\hat{\delta}_n$ of $f'(m_n)$ was considered in a more general setting in [8]. In (2.3) we consider different cases in order that δ_n has the same sign as σ which cannot be guaranteed by (2.4) alone in general. In the algorithm which now follows δ_n denotes one of the expressions in (2.1), (2.2) or (2.3).

ALGORITHM:
Given $[x_0]$
Choose $m_0 \in [x_0]$
For $n = 0, 1, \ldots$ until some stopping criteria is fulfilled do
begin (A1)

$$[y_n] = \left(m_n - \frac{f_n}{\Delta}\right) \cap [x_n]$$

$$q_n = m_n - \frac{f_n}{\delta_n}$$

(A2) **If** $d[y_n] \leq \frac{1}{4} d[x_{n-1}]$

then
begin (A2.a)
$[x_{n+1}] = [y_n]$
$m_{n+1} = P_{n+1}(q_n)$
$f_{n+1} = f(m_{n+1})$
end
else
begin (A2.b)
$[x_{n+1}] = \text{bisect}[y_n]$
$r_n = \text{mid}[y_n]$
$\hat{f}_n = f(r_n)$
$\bar{q}_n = P_{n+1}(q_n)$
$\bar{f}_n = f(\bar{q}_n)$

$$m_{n+1} = \begin{cases} r_n & \text{if } |\hat{f}_n| < |\bar{f}_n| \\ \bar{q}_n & \text{if } |\hat{f}_n| \geq |\bar{f}_n| \end{cases}$$

$$f_{n+1} = \begin{cases} \hat{f}_n & \text{if } |\hat{f}_n| < |\bar{f}_n| \\ \bar{f}_n & \text{if } |\hat{f}_n| \geq |\bar{f}_n| \end{cases}$$

end
end

In the cases (2.1) and (2.2) one needs besides the given $[x_0]$ another interval $[x_{-1}]$ in order that the algorithm can be started. Similarly, in case (2.4) one needs two additional intervals $[x_{-1}]$ and $[x_{-2}]$. For simplicity we could choose $[x_0] = [x_{-1}]$ and $[x_0] = [x_{-1}] = [x_{-2}]$, respectively. If one of the denominators in (2.2) or (2.4) becomes zero then the difference quotient should be replaced by the corresponding derivative.

3 Properties of the new methods.

In this section we first prove some general properties of the algorithm which are independent of the choice of δ_n.

THEOREM 3.1. *Assume that the real function $f : D \subseteq \mathbf{R} \to \mathbf{R}$ is continuously differentiable and its derivative has an interval arithmetic evaluation $f'([x_0])$ with $0 \notin f'([x_0])$ for a given $[x_0] \subseteq D$. Assume that $f(x^*) = 0$ for some $x^* \in [x_0]$. Then*

$$\text{a)} \quad x^* \in [x_{n+1}] \subseteq [x_n],$$

$$\text{b)} \quad \lim_{n\to\infty} [x_n] = x^*,$$

$$\text{c)} \quad d[x_{n+1}] \leq \frac{\beta - \alpha}{\alpha^2} |f(m_n)|. \tag{3.1}$$

PROOF. a) From the Mean Value Theorem we have

$$f(m_n) = f(m_n) - f(x^*) = f'(\xi_n)(m_n - x^*), \quad \xi_n \in [x_n]$$

and therefore, since $x^* \in [x_n]$,

$$x^* = m_n - \frac{f(m_n)}{f'(\xi_n)} \in m_n - \frac{f(m_n)}{\Delta} \subseteq \left\{m_n - \frac{f(m_n)}{\Delta}\right\} \cap [x_n] = [y_n] \subseteq [x_n].$$

The assertion follows by considering cases (A2.a) and (A2.b) of the algorithm.

b) By considering again cases (A2.a) and (A2.b) one proves by mathematical induction that the diameters of the sequence $\{[x_n]\}_{n=0}^{\infty}$ satisfy the relation

$$d_n \leq \left(\frac{1}{2}\right)^n d_0, \quad n \geq 0. \tag{3.2}$$

Together with $x^* \in [x_n]$ the assertion follows.

c) From the algorithm it follows immediately that $d[x_{n+1}] \leq d[y_n]$. Since

$$d\left(\frac{1}{\Delta}\right) = \frac{\beta - \alpha}{\alpha\beta} \leq \frac{\beta - \alpha}{\alpha^2}$$

and

$$d\left(\frac{f(m_n)}{\Delta}\right) = |f(m_n)| \cdot d\left(\frac{1}{\Delta}\right),$$

the assertion follows from

$$d[y_n] \leq d\left(m_n - \frac{f(m_n)}{\Delta}\right) = d\left(\frac{f(m_n)}{\Delta}\right) \leq |f(m_n)| \cdot \frac{\beta - \alpha}{\alpha^2}.$$

□

REMARK 3.1. From inequality (3.2) it follows that the algorithm obtains $d[x_n] < \epsilon$ in at most as many steps as the bisection method.

In what follows we will show that if we choose δ_n as in (2.1), (2.2), (2.3), then for n sufficiently large m_{n+1} is always computed as in (A2.a). As a result of this the convergence of the algorithm is superlinear (of R-order at least 2, $(1+\sqrt{5})/2$, and $1.839\ldots$, respectively).

In order to simplify proofs we denote by f an extension of the given function to the whole real line such that $f \in C^1(\mathbf{R})$ and $f'(x) \in \Delta$ for all $x \in \mathbf{R}$. This can be achieved by extending f by the tangent lines at the endpoints of $[x_0]$.

LEMMA 3.2. *If $f'(x) \in \Delta$, $0 \notin \Delta$, $x \in \mathbf{R}$, $x^* \in [x_n] \subseteq [x_0]$, $f(x^*) = 0$, then*

$$|f(P_n(x))| \leq |f(x)|, \quad x \in \mathbf{R}. \tag{3.3}$$

PROOF. If $x \in [x_n] = [\underline{x}, \overline{x}]$ then $P_n(x) = x$ and (3.3) holds. Since $0 \notin \Delta$, assume without loss of generality that $f'(x) > 0$. Then f is increasing and we have $f(\underline{x}) \leq f(x^*) = 0 \leq f(\overline{x})$. Hence, if $x < \underline{x}$ it follows that $f(x) < f(\underline{x})$ and $f(P_n(x)) = f(\underline{x}) \geq f(x)$ or

$$|f(P_n(x))| \leq |f(x)|. \tag{3.4}$$

Similarly, if $x > \overline{x}$, it follows that

$$0 \le f(P_n(x)) \le f(x). \tag{3.5}$$

Inequalities (3.4) and (3.5) prove the assertion in the case $f'(x) > 0$. The case $f'(x) < 0$ can be treated analogously. □

LEMMA 3.3. *If $f'(x) \in \Delta$, $0 \notin \Delta$, $x \in \mathbb{R}$, $x^* \in [x_0]$, $f(x^*) = 0$, then*

$$|x - x^*| \;\le\; \frac{|f(x)|}{\alpha} \tag{3.6}$$

and

$$|f(x)| \;\le\; \beta|x - x^*| \tag{3.7}$$

for all $x \in \mathbb{R}$.

PROOF. From the Mean Value Theorem it follows that

$$|x - x^*| = \left|\frac{f(x) - f(x^*)}{f'(\xi)}\right| = \left|\frac{f(x)}{f'(\xi)}\right| \le \frac{|f(x)|}{\alpha}$$

and

$$|f(x)| = |f'(\xi)||x - x^*| \le \beta|x - x^*|.$$

□

THEOREM 3.4. *Under the hypothesis of the conditions of Theorem 3.1 assume that the mapping f is twice continuously differentiable on the interval $[x_0]$ and that $\gamma := \max_{x\in[x_0]} |f''(x)|$. Assume further that we have extended f as described before Lemma 3.2. If we choose δ_n in the algorithm as in (2.1), that is, $\delta_n = f'(m_n)$, $n = 0, 1, 2, \ldots$, then*

$$|f(m_{n+1})| \;\le\; \frac{\beta\gamma}{2\alpha^3}|f(m_n)|^2 \tag{3.8}$$

and

$$d[x_{n+1}] \;\le\; \frac{\beta(\beta-\alpha)\gamma}{2\alpha^3}(d[x_{n-1}])^2. \tag{3.9}$$

PROOF. We have

$$q_n = m_n - \frac{f(m_n)}{f'(m_n)}$$

and therefore, using Taylor's Theorem,

$$\begin{aligned} |q_n - x^*| &= \left|\frac{1}{f'(m_n)}\left(f(m_n) - f(x^*) - f'(m_n)(m_n - x^*)\right)\right| \\ &= \left|\frac{1}{f'(m_n)}\right| \cdot \frac{1}{2}|f''(\xi)|\,|x^* - m_n|^2 \\ &\le \frac{1}{2}\frac{\gamma}{\alpha}|m_n - x^*|^2. \end{aligned}$$

By considering the definition of m_{n+1} in the algorithm and by applying Lemmas 3.2 and 3.3 we get

$$|f(m_{n+1})| \leq |f(q_n)| \leq \beta |q_n - x^*| \leq \frac{1}{2}\frac{\beta\gamma}{\alpha}|m_n - x^*|^2 \leq \frac{1}{2}\frac{\beta\gamma}{\alpha^3}|f(m_n)|^2,$$

which proves (3.8).

Similarly, by considering the definition of m_n and by using equation (3.7) from Lemma 3.3, we obtain

$$|f(m_n)| \leq |f(q_{n-1})| \leq \beta |q_{n-1} - x^*| \leq \frac{1}{2}\frac{\beta\gamma}{\alpha}|m_{n-1} - x^*|^2.$$

Together with (3.1) from Theorem 3.1 we get

$$d[x_{n+1}] \leq \frac{\beta - \alpha}{\alpha^2}|f(m_n)| \leq \frac{1}{2}\frac{\beta - \alpha}{\alpha^3}\beta\gamma|m_{n-1} - x^*|^2 \leq \frac{1}{2}\frac{\beta - \alpha}{\alpha^3}\beta\gamma(d[x_{n-1}])^2,$$

since m_{n-1}, $x^* \in [x_{n-1}]$. Hence, (3.9) is also proved. □

COROLLARY 3.5. *Let us consider the algorithm with choice* (2.1). *Then the following statements hold:*

a) *There exists an* $N > 0$*, such that*

$$[x_{n+1}] = [y_n], \quad m_{n+1} = P_{n+1}(q_n), \quad n \geq N.$$

b) *The sequence* $\{f(m_n)\}$ *is at least Q-quadratically (and hence at least R-quadratically) convergent.*

c) *The sequence* $\{d[x_n]\}$ *is at least R-quadratically convergent.*

PROOF. a) Since $d[x_n] \to 0$ monotonically, there is an $N > 0$ such that $\beta(\beta - \alpha)\gamma d[x_{n-1}]/(2\alpha^3) \leq 1/8$ for all $n \geq N$. Hence, by (3.9), the inequality $d[x_{n+1}] \leq d[x_{n-1}]/8$ holds for all $n \geq N$. Assume now that $d[y_n] > d[x_{n-1}]/4$ for some $n \geq N$. Then by (A2.b) $d[x_{n+1}] = d[y_n]/2$. However, this contradicts

$$d[x_{n+1}] \leq \tfrac{1}{8}d[x_{n-1}] < \tfrac{1}{8}4d[y_n] = \tfrac{1}{2}d[y_n].$$

Therefore, for $n \geq N$ we have $d[y_n] \leq d[x_{n-1}]/4$ and hence $[x_{n+1}] = [y_n]$ and $m_{n+1} = P_{n+1}(q_n)$ by (A2.a).

b) follows from (3.8) and the fact that the R-order is not smaller than the Q-order (see [7, 9.3.2]).

c) follows from the easily proved fact that if $\{r_n\}$ and $\{s_n\}$ are two real zero sequences such that

$$0 \leq s_{n+1} \leq c \cdot r_n, \quad n \geq 0,$$

then the R-order of the sequence $\{s_n\}$ is not smaller than the R-order of the sequence $\{r_n\}$. From b) we know that the sequence $\{f(m_n)\}$ has at least R-order two. By inequality (3.1) from Theorem 3.1 we conclude that the R-order of the sequence $\{d[x_{n+1}]\}$ is at least 2. □

REMARK 3.2. From part a) of the preceding corollary it follows that the algorithm with choice (2.1) needs asymptotically 2 function values per step since always (A2.a) is performed. Hence its asymptotic efficiency is at least $\sqrt{2} = 1.41\ldots$.

Besides a good asymptotic behaviour our algorithm has also good global convergence properties in the sense of the following corollary.

COROLLARY 3.6. *The cost of obtaining $d[x_n] < \epsilon$ with our algorithm with choice* (2.1) *is at most three times larger than the corresponding cost when using the bisection method.*

PROOF. The bisection method needs one function value per step. If (A2.a) is performed one needs $\delta_n = f'(m_n)$ and $f(m_{n+1})$. In the case (A2.b) one computes $\delta_n = f'(m_n)$, $f(r_n)$ and $f(\overline{q}_n)$. □

THEOREM 3.7. *Assume that the assumptions of Theorem 3.4 hold. If we choose δ_n in the algorithm as in* (2.2), *that is,*

$$\delta_n = \frac{f_n - f_{n-1}}{m_n - m_{n-1}} = f[m_n, m_{n-1}], \quad n = 0, 1, 2, \ldots,$$

then

$$|f(m_{n+1})| \leq \frac{\beta\gamma}{2\alpha^3}|f(m_n)| \cdot |f(m_{n-1})| \tag{3.10}$$

and

$$d[x_{n+1}] \leq \frac{\beta(\beta-\alpha)\gamma}{2\alpha^3} d[x_{n-1}] \cdot d[x_{n-2}]. \tag{3.11}$$

PROOF. We have

$$q_n = m_n - f[m_n, m_{n-1}]^{-1} f(m_n)$$

and therefore

$$\begin{aligned} |q_n - x^*| &= |f[m_n, m_{n-1}]^{-1}(f(m_n) - f(x^*) - f[m_n, m_{n-1}](m_n - x^*))| \\ &= |f[m_n, m_{n-1}]^{-1} f[m_n, m_{n-1}, x^*](m_n - x^*)(m_{n-1} - x^*)| \\ &\leq \frac{\gamma}{2\alpha}|m_n - x^*||m_{n-1} - x^*|, \end{aligned}$$

where $f[m_n, m_{n-1}, x^*]$ denotes the second divided difference and where we have used

$$|f[m_n, m_{n-1}, x^*]| \leq \tfrac{1}{2} \max_{x \in [x_0]} |f''(x)| = \tfrac{1}{2}\gamma.$$

From Lemma 3.2 it follows that $|f(m_{n+1})| \leq |f(q_n)|$, and by applying Lemma 3.3 and (3.1) we get (3.10) and (3.11). □

COROLLARY 3.8. *Let us consider the algorithm with choice* (2.2). *Then*

a) *There is an $N > 0$ such that*

$$[x_{n+1}] = [y_n], \quad m_{n+1} = P_{n+1}(q_n), \quad n \geq N.$$

b) *The R-order of the sequences $\{f(m_n)\}$ and $\{d[x_n]\}$ are both at least equal to $(1+\sqrt{5})/2 = 1.618\ldots$.*

PROOF. a) This follows from (3.11) as in the corresponding statement of Corollary 3.5.

b) The R-order of the sequence $\{f(m_n)\}$ is at least equal to $(1+\sqrt{5})/2$ according to Sec. **9.2.9** in [7]. The remaining is obtained as in part c) of Corollary 3.5. □

REMARK 3.3. From part a) of the preceding corollary it follows that the algorithm with choice (2.2) needs asymptotically one function value per step since always (A2.a) is performed. Hence its asymptotic efficiency is at least equal to $(1+\sqrt{5})/2$.

COROLLARY 3.9. *The cost of obtaining $d[x_n] < \epsilon$ with our algorithm with choice* (2.2) *is at most two times larger than the corresponding cost when using the bisection method.*

THEOREM 3.10. *Under the hypothesis of Theorem* 3.7 *assume that f is three times continuously differentiable and that*

$$\mu = \max_{x\in[x_0]} |f'''(x)|.$$

If we choose δ_n in the algorithm as in (2.3) *then*

$$|f_{n+1}| \leq \nu|f_n|\cdot|f_{n-1}|\cdot|f_{n-2}|, \quad n=2,3,\ldots, \tag{3.12}$$

and

$$d[x_{n+1}] \leq \tilde{\nu} d[x_{n-1}]d[x_{n-2}]d[x_{n-3}], \quad n=3,4,\ldots, \tag{3.13}$$

where $\nu, \tilde{\nu}$ are constants not depending on n.

PROOF. We have

$$\begin{aligned}
\hat{\delta}_n &= f[m_n,m_{n-1}] + f[m_n,m_{n-2}] - f[m_{n-1},m_{n-2}] \\
&= f[m_n,x^*] + f[m_n,m_{n-1}] - f[m_n,x^*] + f[m_n,m_{n-2}] - f[m_{n-1}m_{n-2}] \\
&= f[m_n,x^*] + f[m_{n-1},m_n,x^*](m_{n-1}-x^*) \\
&\quad + f[m_n,m_{n-2},m_{n-1}](m_n - x^* + x^* - m_{n-1}) \\
&= f[m_n,x^*] - f[m_n,m_{n-1},m_{n-2}](x^*-m_n) \\
&\quad + f[m_{n-1},m_n,x^*](m_{n-1}-x^*) - f[m_n,m_{n-1}m_{n-2}](m_{n-1}-x^*) \\
&= f[m_n,x^*] - f[m_n,m_{n-1},m_{n-2}](x^*-m_n) \\
&\quad + f[m_n,m_{n-1},m_{n-2},x^*](x^*-m_{n-2})(m_{n-1}-x^*),
\end{aligned}$$

where $f[m_n,m_{n-1},m_{n-2},x^*]$ denotes the third divided difference. Since for $n\to\infty$ the second and the third term in the last equation tend to zero, whereas the first term approaches $f'(x^*)$, it follows that there is an $N_1 > 0$ such that $\sigma\hat{\delta}_n \geq \frac{\alpha}{2}$ for all $n \geq N_1$. Since then $\delta_n = \hat{\delta}_n$ by the definition of δ_n (see (2.3)), it follows that $|\delta_n^{-1}| \leq \frac{2}{\alpha}$ for $n \geq N_1$. Therefore

$$\begin{aligned}
|q_n - x^*| &= |\delta_n^{-1}(f(m_n) - f(x^*) - \delta_n(m_n - x^*))| \\
&\leq \frac{2}{\alpha}|f[m_n,x^*] - \delta_n||m_n - x^*| \\
&\leq \frac{2}{\alpha}|m_n - x^*|\left\{\frac{\gamma}{2}|m_n - x^*| + \frac{\mu}{6}|m_{n-1}-x^*||m_{n-2}-x^*|\right\},
\end{aligned}$$

where we have used the preceding representation of $\hat{\delta}_n = \delta_n$. Using (3.6) we can further bound the right hand side to get

$$|q_n - x^*| \le \frac{2}{\alpha^2}|f_n| \left\{ \frac{\gamma}{2\alpha}|f_n| + \frac{\mu}{6\alpha^2}|f_{n-1}||f_{n-2}| \right\}.$$

Using this inequality, Lemma 3.2 and (3.7) we arrive at

$$\begin{aligned} |f_{n+1}| &\le |f(q_n)| \le \beta|q_n - x^*| \le \frac{2\beta}{\alpha^2}|f_n| \left\{ \frac{\gamma}{2\alpha}|f_n| + \frac{\mu}{6\alpha^2}|f_{n-1}||f_{n-2}| \right\} \\ &\le \gamma_1|f_n|\{|f_n| + |f_{n-1}||f_{n-2}|\} \quad \text{for } n > N_1, \end{aligned} \tag{3.14}$$

where $\gamma_1 = 2\beta \max\{\gamma/2\alpha, \mu/(6\alpha^2)\}/\alpha^2$.

Since $|f_n| \to 0$, it is clear from the last inequality that there exists an $N_2 > N_1$ such that

$$|f_{n+1}| < |f_n| < 1$$

for $n \ge N_2$. Using (3.14) once more it follows that

$$\begin{aligned} |f_{n+1}| &\le \gamma_1|f_n||f_n| + \gamma_1|f_n||f_{n-1}||f_{n-2}| \\ &\le \gamma_1|f_n||f_{n-1}| + \gamma_1|f_n||f_{n-1}| = 2\gamma_1|f_n||f_{n-1}| \end{aligned}$$

for $n > N_2 + 2$, or

$$|f_n| \le 2\gamma_1|f_{n-1}||f_{n-2}| \quad \text{for } n > N_2 + 3.$$

Therefore, according to (3.14) we have

$$|f_{n+1}| \le \gamma_1|f_n|\{2\gamma_1|f_{n-1}||f_{n-2}| + |f_{n-1}||f_{n-2}|\} = \gamma_1(2\gamma_1 + 1)|f_n||f_{n-1}||f_{n-2}|$$

for all $n > N_2 + 3$. By choosing $\nu > \gamma_1(2\gamma_1 + 1)$ sufficiently large we may now assume that (3.12) holds for all $n = 2, 3, \ldots$. From statement c) of Theorem 3.1 and (3.12) it follows that

$$d[x_{n+1}] \le \frac{\beta - \alpha}{\alpha^2}|f(m_n)| \le \frac{\beta - \alpha}{\alpha^2}\nu|f(m_{n-1})||f(m_{n-2})||f(m_{n-3})|$$

for $n \ge 3$. Using (3.7) we obtain

$$|f(m_{n-1})| \le \beta|m_{n-1} - x^*| \le \beta d[x_{n-1}].$$

Similarly we have

$$\begin{aligned} |f(m_{n-2})| &\le \beta d[x_{n-2}], \\ |f(m_{n-3})| &\le \beta d[x_{n-3}]. \end{aligned}$$

Therefore from the preceding inequality we get

$$d[x_{n+1}] \le \beta^3 \frac{\beta - \alpha}{\alpha^2}\nu d[x_{n-1}]d[x_{n-2}]d[x_{n-3}]$$

for $n \ge 3$. This is (3.13) with $\tilde{\nu} = \beta^3(\beta - \alpha)\nu/\alpha^2$. □

COROLLARY 3.11. *Let us consider the algorithm with choice* (2.3). *Then*

Table 4.1: Examples.

Ex.	function $f(x)$	$[a, b]$	parameter
1	$\sin x - \frac{x}{2}$	$[\frac{\pi}{2}, \pi]$	
2	$-2\sum_{i=1}^{20} \frac{(2i-5)^2}{(x-i^2)^3}$	$[a_n, b_n]$ $a_n = n^2 + 10^{-9}$ $b_n = (n+1)^2 - 10^{-9}$	$n = 1, 5, 10$
3	axe^{bx}	$[-9, 31]$	$a = -40, b = -1$ $a = -100, b = -2$ $a = -200, b = -3$
4	$2xe^{-n} - 2e^{-nx} + 1$	$[0, 1]$	$n = 1, 5, 20, 100$
5	$(1 + (1-n)^2)x - (1-nx)^2$	$[0, 1]$	$n = 5, 10, 20$
6	$x^2 - (1-x)^n$	$[0, 1]$	$n = 2, 5, 10, 15, 20$
7	$(1 + (1-n)^4)x - (1-nx)^4$	$[0, 1]$	$n = 1, 2, 4, 5, 8,$ $15, 20$
8	$e^{-nx}(x-1) + x^n$	$[0, 1]$	$n = 1, 5, 10, 15, 20$
9	$\frac{nx-1}{(n-1)x}$	$[0.01, 1]$	$n = 2, 5, 15, 20$
10	$x^{1/n} - n^{1/n}$	$[1, 100]$	$n = 2, 3, 4, 5, 7, 10,$ $15, 20, 25, 30, 33$

a) *There is an $N > 0$ such that*

$$[x_{n+1}] = [y_n], \quad m_n = P_{n+1}(q_n), \quad n \geq N.$$

b) *The R-orders of the sequences $\{f(m_n)\}$ and $\{d[x_n]\}$ are both at least equal to the unique positive root t^* of the polynomial*

$$p(t) = t^3 - t^2 - t - 1.$$

We have $t^* \approx 1.839\ldots$.

PROOF. a) follows from (3.13) as in the proof of the corresponding statement in Corollary 3.5.

b) The R-order of the sequence $\{f(m_n)\}$ is at least equal to t^* according to [11, 4.2.4]. The remaining part is obtained as in part c) of Corollary 3.5. □

REMARK 3.4. From part a) of the preceding corollary it follows that the algorithm with choice (2.3) needs asymptotically one function value per step since always (A2.a) is performed. Hence its asymptotic efficiency is at least $t^* \approx 1.839\ldots$.

COROLLARY 3.12. *The cost of obtaining $d[x_n] < \epsilon$ with our algorithm with choice (2.3) is at most two times larger than the corresponding cost when using the bisection method.*

4 Numerical examples.

In order to test the global convergence, the methods were tested using the following examples from Table 4.1.

In Table 4.2 we present for a user given tolerance of $tol := 0$ the number of function values used by the Interval-Newton-Method(IVN), case (2.1), case (2.2) and case (2.3). The precise meaning of tol is carefully described in [13]. See also [2] and [5]. For every example we have tested the methods with 15 different starting intervals which were obtained by increasing and decreasing the lower and upper bound of the interval $[a, b]$, respectively, in such a manner that the zero is still contained in the new interval. The precise values can be obtained from the authors on request (in some of the examples from Table 4.1 we could not start with the given interval $[a, b]$ since $0 \in \Delta$).

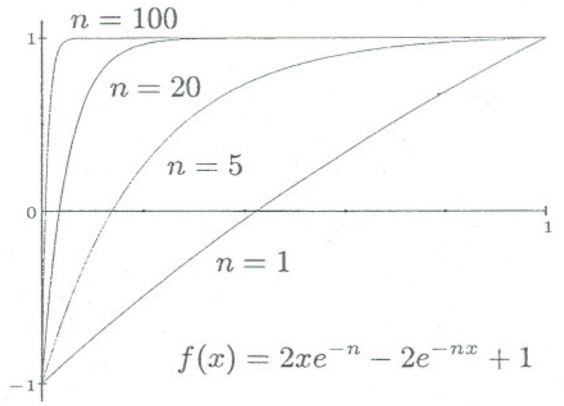

Figure 4.1: Example 4.

The last column contains the function values necessary for Brent's method (see [5]). Brent's method is one of the most popular nonlinear equation solvers. Note, however, that because rounding errors are not taken into account Brent's method does not guarantee enclosure of a zero. Also as shown in [9], there are examples for which even if exact arithmetic is used the diameters of the intervals produced by Brent's method do not converge to zero. Nevertheless, as shown by our numerical examples, Brent's method is very efficient in practice. Finally we mention that Example 4 (see Figure 4.1) is a very hard one for our new methods. The initial enclosure Δ of the derivative has a lower bound very close to zero and the upper bound is slightly bigger than $2n$. This enclosure is never improved during the course of computation. As a result of this, it is very unlikely that the first part of (A2) in the Algorithm (namely (A2.a)) is ever performed with the given floating point system, and therefore the Algorithm behaves approximately like the Bisection-Method. This example is the reason that altogether our optimal method (the Algorithm with δ_n defined by (2.3)) needs slightly more function values than the Interval-Newton-Method. (See $\sum^*$ in the last row of Table 4.2.)

Subsequently we illustrate the contents of Table 4.2 via diagrams. For each example we have added together the total number of function values necessary to fulfill the stopping criteria. In the diagrams the heights of the bars are depicting these values in relation to the Interval-Newton-Method.

From the preceding discussion, especially concerning Example 4, one could try

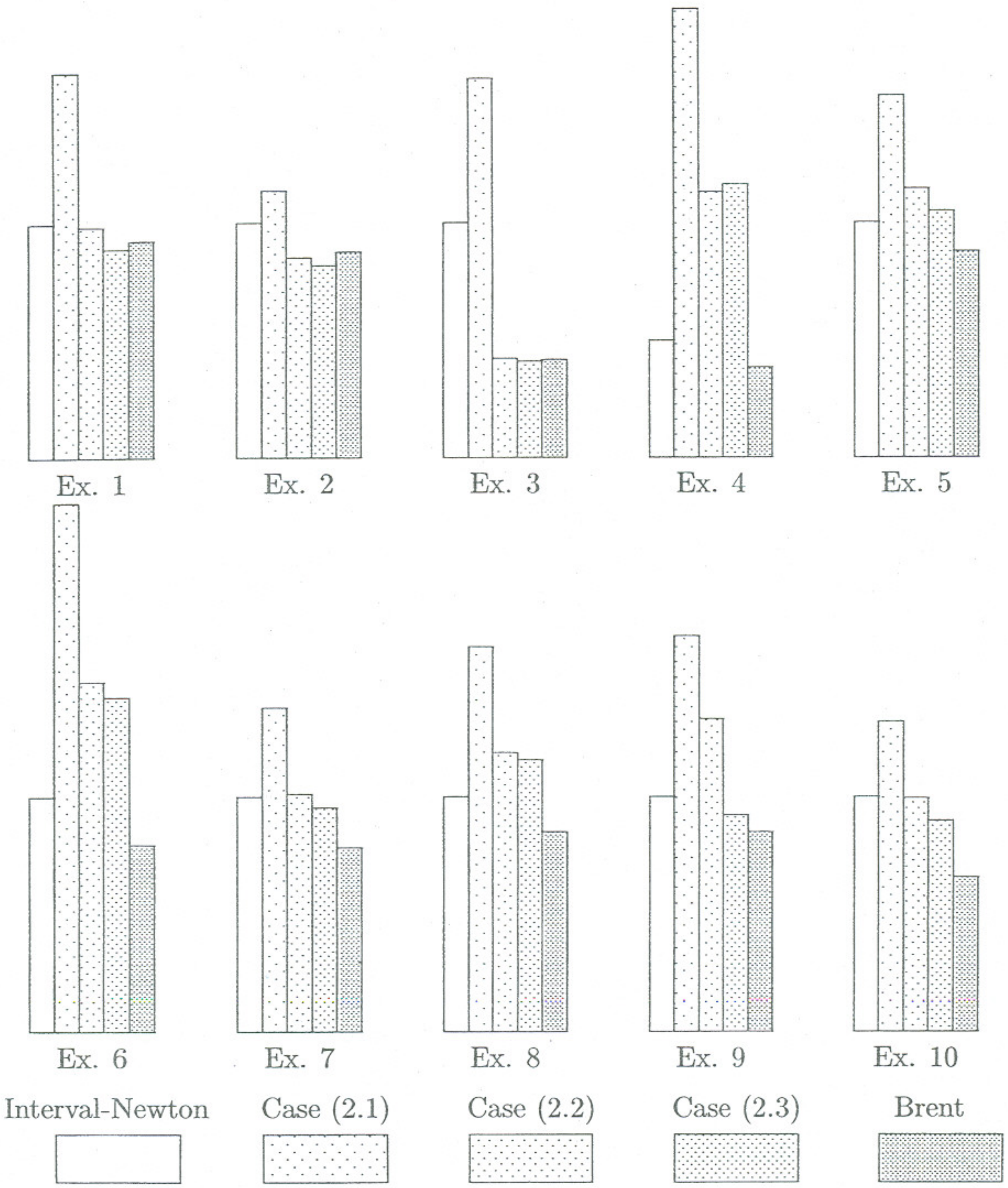

Figure 4.2: Total number of function values necessary to fulfill the stopping criteria for each example for the examples of Table 4.2. The heights of the bars depict these values in relation to the Interval-Newton-Method.

Table 4.2: User given tolerance of $tol := 0$.

Ex.	IVN	(2.1)	(2.2)	(2.3)	Br
1	122	159	120	109	113
2_1	170	201	149	138	144
2_2	156	179	133	128	140
2_3	164	178	135	135	145
3_1	426	627	144	146	164
3_2	252	601	166	169	158
3_3	480	646	179	161	162
4_1	122	159	118	108	105
4_2	170	316	230	216	135
4_3	212	448	498	483	165
4_4	284	2088	934	1027	201
5_1	138	227	165	148	125
5_2	120	186	135	126	101
5_3	124	176	138	127	109
6_1	30	45	45	45	46
6_2	154	346	214	228	127
6_3	176	380	268	219	142
6_4	194	456	296	305	149
6_5	214	504	323	299	146
7_1	134	179	135	119	137
7_2	136	196	144	128	137
7_3	156	229	163	154	112
7_4	144	182	137	129	100

Ex.	IVN	(2.1)	(2.2)	(2.3)	Br
7_5	140	188	139	134	87
7_6	126	168	122	122	88
7_7	120	177	128	128	87
8_1	118	159	117	106	124
8_2	172	211	154	145	126
8_3	210	282	223	208	164
8_4	218	393	297	290	184
8_5	218	490	320	336	198
9_1	118	161	124	81	135
9_2	170	250	191	145	132
9_3	202	428	302	208	165
9_4	200	325	303	202	153
10_1	174	225	167	149	84
10_2	194	254	188	164	129
10_3	194	254	206	187	132
10_4	204	291	220	209	144
10_5	200	277	211	186	139
10_6	192	262	194	177	142
10_7	202	284	206	181	132
10_8	186	261	191	163	132
10_9	198	245	184	172	130
10_{10}	196	247	184	166	125
10_{11}	214	247	188	181	125
$\sum^*$	8444	14787	9528	8887	6120

to improve the practical behavior of our new methods by computing improved values of Δ during the course of computation. For example, one could compute a new enclosure $f'([x_n])$ of f' over $[x_n]$ and use this instead of $f'([x_0])$. Another modification could be to compute the interval arithmetic evaluation of the derivative only after k steps again, where $k > 1$ is a fixed integer. The order of convergence is improved in this way, however, the cost is also increased. We do not discuss here the dependence of the effectivity index on k, since this is related only to the asymptotic behaviour of the method. In the following parts of Tables 4.3–4.5 we report on the corresponding values $\sum^*$ of Table 4.2 for $tol = 10^{-5}, 10^{-10}, 10^{-12}, 10^{-14}, 10^{-16}, 0$. In each table the integer k specifies the number of steps after which a new enclosure of the derivative has been computed. ($k = \infty$ means that the initial value of Δ was fixed. This corresponds to the original version of our new algorithms.) Under each table we have also listed the number of function values which were needed for the Interval-Newton-Method and for Brent's method, respectively. These values are independent of k. It can be seen from the tables that for values of tol between 10^{-5} and 10^{-12} our optimal method (2.3) for $k = \infty$ is superior to the Interval-Newton-Method. For $tol = 10^{-14}$, 10^{-16} and $tol = 0$ the latter needs less function values than

Table 4.3: Updating the derivative after k steps: $tol = 10^{-5}, 10^{-10}$.

k	$tol = 10^{-5}$ (2.1)	(2.2)	(2.3)
1	8915	6872	6617
2	7041	5813	5546
3	6764	5465	5187
4	6713	5369	5095
5	6697	5314	4944
6	6735	5313	4946
7	6740	5300	4953
8	6759	5299	4937
9	6790	5292	4912
10	6821	5284	4909
∞	6983	5280	4913
Interval-Newton:			5028
Brent's Method:			4915

k	$tol = 10^{-10}$ (2.1)	(2.2)	(2.3)
1	9662	8216	7823
2	8514	6980	6493
3	8215	6614	6057
4	8037	6399	5893
5	8066	6350	5802
6	8019	6342	5748
7	8041	6275	5702
8	8028	6264	5683
9	8077	6232	5659
10	8131	6222	5670
∞	8978	6298	5811
Interval-Newton:			6192
Brent's Method:			5710

Table 4.4: Updating the derivative after k steps: $tol = 10^{-12}, 10^{-14}$.

k	$tol = 10^{-12}$ (2.1)	(2.2)	(2.3)
1	10115	8520	8105
2	8899	7285	6753
3	8587	6916	6322
4	8417	6667	6113
5	8461	6614	5994
6	8399	6636	5958
7	8406	6564	5890
8	8403	6537	5860
9	8450	6521	5854
10	8517	6499	5859
∞	9753	6799	6258
Interval-Newton:			6578
Brent's Method:			5843

k	$tol = 10^{-14}$ (2.1)	(2.2)	(2.3)
1	10860	9149	8722
2	9619	7832	7263
3	9266	7431	6768
4	9075	7184	6532
5	9083	7089	6366
6	9062	7091	6380
7	9091	7058	6313
8	9169	7013	6290
9	9224	6994	6320
10	9310	7012	6336
∞	11023	7672	7079
Interval-Newton:			7064
Brent's Method:			5947

(2.3) for $k = \infty$.

On the other hand it can be seen that for each given value of tol there is a $k \geq 1$ such that our optimal method (2.3) needs slightly less function values than the Interval-Newton-Method. Therefore an adaptive technique could be used to decide for a fixed n whether a new value of $f'([x_n])$ should be computed or not. We proceed as follows: If (A2.b) is performed then we have three different points contained in $[y_n]$ available, namely $r_n, \overline{q}_n$ and m_n as well as the

Table 4.5: Updating the derivative after k steps: $tol = 10^{-16}, 0$.

k	$tol = 10^{-16}$ (2.1)	(2.2)	(2.3)
1	11445	9535	9112
2	10001	8167	7518
3	9664	7734	7082
4	9480	7483	6789
5	9450	7375	6656
6	9500	7379	6676
7	9631	7355	6657
8	9868	7363	6705
9	10027	7412	6793
10	10178	7487	6869
∞	12496	8746	8134
Interval-Newton:			7362
Brent's Method:			6022

k	$tol = 0$ (2.1)	(2.2)	(2.3)
1	14108	10184	9818
2	12199	8661	8047
3	11718	8164	7515
4	11467	7872	7217
5	11455	7781	7113
6	11549	7753	7106
7	11660	7765	7155
8	11930	7835	7231
9	12105	7908	7342
10	12270	8009	7430
∞	14787	9528	8887
Interval-Newton:			8444
Brent's Method:			6120

Table 4.6: Values for the adaptive technique.

$\sum^*$	IVN	1/2	1/3	1/4	(2.3)
10^{-5}	5028	4838	4837	4844	4913
10^{-10}	6192	5581	5579	5618	5811
10^{-12}	6578	5775	5774	5782	6258
10^{-14}	7064	6224	6231	6242	7079
10^{-16}	7362	6706	6733	6756	8134
0	8444	7321	7338	7350	8887

corresponding function values. Using these three points we compute the slopes $f[r_n, \overline{q}_n]$, $f[r_n, m_n]$ and $f[\overline{q}_n, m_n]$. The minimum and maximum of these three values are called min and max, respectively. If with some constant $c > 0$ we have $\max - \min \leq c \cdot d(f'([x_n]))$, where $f'([x_n])$ is the enclosure of the derivative over $[x_n]$, then we compute a new enclosure over $[x_{n+1}]$. In every case a new enclosure is computed if during the last 5 steps this has not been done. In Table 4.6 we list for different values of *tol* the total number of function values needed for the Interval-Newton-Method for $c = \frac{1}{2}, \frac{1}{3}, \frac{1}{4}$, and the unmodified method (2.3). It can be seen that with the exception of $tol = 10^{-16}$ and $tol = 0$ only a small amount of the total number of function values can be saved. However the adaptive method is always better than the Interval-Newton-Method.

The results have been computed using the programming language PASCAL XSC on a SUN Workstation. The mantissa length is 16 decimal digits.

Acknowledgement.

The authors are greatful to the referees. Their comments helped to improve the paper considerably.

REFERENCES

1. G. Alefeld and J. Herzberger, *Introduction to Interval Computations*, Academic Press, New York, 1983.
2. G. Alefeld, F. A. Potra, and Y. Shi, *Algorithm 748: Enclosing zeros of continuous functions*, ACM Trans. Math. Software, 21:3 (1995), pp. 327–344.
3. G. Alefeld, F. A. Potra, and W. Völker, *Effective improvements of the interval-Newton-method*, in Scientific Computing and Validated Numerics, G. Alefeld, A. Frommer, and B. Lang, eds., Akademie-Verlag, Berlin, 1996, pp. 133–139.
4. G. Alefeld (1994). *Inclusion methods for systems of nonlinear equations—the interval Newton method and modifications*, in Topics in Validated Computations, J. Herzberger, ed., Elsevier, Amsterdam, 1994.
5. R. P. Brent, *Algorithms for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, NJ, 1972.
6. R. E. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1966.
7. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
8. F. A. Potra, *On an iterative algorithm of order* 1.839... *for solving nonlinear operator equations*, Numer. Funct. Anal. Optim., 7:1 (1984–1985), pp. 75–106.
9. F. A. Potra (1995), *A note on Brent's rootfinding method*, in Numerical Methods and Error Bounds, G. Alefeld and J. Herzberger, eds., Akademie-Verlag, Berlin, 1996, pp. 188–197.
10. A. Ralston and P. Rabinowitz, *A First Course in Numerical Analysis*, McGraw-Hill, Tokyo, 1978.
11. H. Schwetlick, *Numerische Lösung nichtlinearer Gleichungen*, Deutscher Verlag der Wissenschaften, Berlin, 1979.
12. T. Sunaga, *Theory of an interval algebra and its applications to numerical analysis*, in RAAG Memoirs II, Gakujutsu, Bunken Fukyu-Kai, 1958.
13. W. Völker, *Effektive Verbesserungen des Intervall-Newton-Verfahrens*, Diplomarbeit, Institut für Angewandte Mathematik, Universität Karlsruhe, 1995.