BIT 25 (1985), 256-273

# INTERVAL ITERATION FOR ZEROS OF SYSTEMS OF EQUATIONS

A. NEUMAIER

*Institut für Angewandte Mathematik, Universität Freiburg i. Br., D-7800 Freiburg i. Br., Hermann-Herder Strasse 10, W.-Germany*

**Abstract.**

Easily verifiable existence and convergence conditions are given for a class of interval iteration algorithms for the enclosure of a zero of a system of nonlinear equations. In particular, a quadratically convergent method is obtained which throughout the iteration uses the same interval enclosure of the derivative.

AMS Subject Classifications: 65G10, 65H10.

## 1. Introduction.

Let $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable map such that $F'(x)$ is nonsingular for all $x \in D$, and suppose that $F$ has a zero $x^* \in D$. To determine $x^*$ numerically, a number of good local methods are available, e.g. Newton's method with superlinear (or even quadratic) convergence. Such methods are globally convergent only under very restrictive conditions, like convexity, or they have to be modified suitably, e.g. by a damping strategy. In the latter case, global convergence can be proved under weaker conditions, but the verification of the convergence conditions in concrete cases is difficult or impossible.

In this paper we show that the combination of any good local "predictor" method with certain techniques from interval arithmetic produces a globally convergent strategy under easily verifiable conditions, which is locally at least as fast as the predictor method. Moreover, if the selected initial region contains no zero, this fact is detected after finitely many iterations. Finally, if carried out in rounded interval arithmetic (the usually implemented form), rounding errors are controlled automatically, and the iteration stops after finitely many iterations with an interval vector whose width is roughly proportional to the machine accuracy.

The present method is based on the well-known fact that interval arithmetic may be used to construct, for every estimate $\tilde{x}$ of the zero $x^*$, an interval vector

Received April 1984. Revised August 1984.



bounding the error $\tilde{x} - x^*$. Among the various methods available for this, those which yield a bound of the order of $\|F(\tilde{x})\|$ guarantee a small bound for the error if the estimate $\tilde{x}$ is sufficiently good. This leads us to consider only such enclosure methods where the error bound is given in the form

$$\tilde{x} - x^* \in A^I F(\tilde{x}).$$

This enclosure is valid if $A^I$ is an *inverse* of an interval matrix $A$ containing all $F'(\tilde{x})$ for $\tilde{x}$ in the initial region.

The notion *inverse* is defined here to describe common features of several methods enclosing solution of linear interval equations. The definition is given in terms of *sublinear maps*, a generalization of matrix-vector multiplication introduced in [12]. Properties of sublinear maps in general, and of inverses in particular are presented in Section 2 and 3. Emphasis is given to the problem of establishing whether a given sublinear map or inverse is *regular*; this property, an extension of the same property for matrices, turns out to be crucial for our convergence considerations.

In Section 4 we apply the theory of the previous sections to systems of equations. We give a new proof of a well-known zero existence test. Unlike previous proofs, however, the present proof is based on the inverse function theorem and does not make use of Brouwer's fixpoint theorem. Moreover, some technical inequalities concerning the enclosure

$$x^* \in \tilde{x} - A^I F(\tilde{x})$$

prepare for the subsequent convergence analysis.

In Section 5 we describe a general iteration scheme for finding arbitrarily good interval enclosures $x^l$ for a zero of $F$. In each step, a vector $\tilde{x}^l \in x^l$ is selected and a new enclosure $x^{l+1}$ is constructed by intersecting $x^l$ with $\tilde{x}^l - A^I F(\tilde{x}^l)$. Convergence is shown under the assumption that $A^I$ is regular. This improves and extends results of Alefeld [3]. Then it is discussed how a predictor method can be used to choose a good $\tilde{x}^l$, and for the Newton method as predictor, quadratic convergence of the interval iteration is established. Comments on the cost of the algorithm and on the choice of the inverse $A^I$ conclude the paper. Numerical results are reported in Section 6.

Notation and terminology of the paper follows Neumaier [12]; but for convenience of the reader, some definitions are repeated. In particular, we denote the set of intervals, $n$-dimensional interval vectors, and $n \times n$-interval matrices by $\mathbb{IR}$, $\mathbb{IR}^n$, $\mathbb{IR}^{n \times n}$, respectively, use $\check{x}$, $\check{A}$ for the midpoint and $\varrho(x)$, $\varrho(A)$ for the radius of $x \in \mathbb{IR}^n$, and $A \in \mathbb{IR}^{n \times n}$, respectively, and identify degenerate intervals with the point they contain. The identity matrix is denoted by $I$.

**2. Regularity and continuity of sublinear maps.**

As in [12], we call an interval matrix $A \in \mathbb{IR}^{n\times n}$ *regular* if all $\tilde{A} \in A$ are regular, and by Lemma 2 (ii) of [12], this is equivalent to the property

$$\tilde{x} \in \mathbb{R}^n, 0 \in A\tilde{x} \Rightarrow \tilde{x} = 0.$$

It is useful to extend this concept to sublinear maps, introduced in [12] as a generalization of the maps induced by matrix multiplication. A map $S: \mathbb{IR}^n \to \mathbb{IR}^n$ is called *sublinear* if

(i) $S(x \pm y) \subseteq Sx \pm Sy$,
(ii) $S(x\alpha) \quad = (Sx)\alpha$,
(iii) $x \subseteq y \quad \Rightarrow Sx \subseteq Sy$

for all $x, y \in \mathbb{IR}^n$, $\alpha \in \mathbb{R}$; in particular,

(iv) $S0 \quad = 0$,
(v) $\tilde{x} \in x \quad \Rightarrow S\tilde{x} \in Sx$.

We extend $S$ to matrix arguments by applying it to each column of $A \in \mathbb{IR}^{n\times m}$ separately:

$$SA := (S(Ae^{(1)}), \ldots, S(Ae^{(m)})). \tag{1}$$

The special case where $A$ is the identity matrix appeared already in [12] as the kernel $SI =: \kappa(S)$.

We call a sublinear map $S$ *regular* if

$$\tilde{x} \in \mathbb{R}^n, 0 \in S\tilde{x} \Rightarrow \tilde{x} = 0. \tag{2}$$

A first sufficient condition for regularity is based on the following property of sublinear maps.

Proposition 1. *Let $S: \mathbb{IR}^n \to \mathbb{IR}^n$ be a sublinear map, and suppose that $A \in \mathbb{IR}^{n\times m}$. Then*

$$S(A\tilde{x}) \subseteq (SA)\tilde{x} \quad \text{for } \tilde{x} \in \mathbb{R}^m, \tag{3}$$

$$S(A\tilde{B}) \subseteq (SA)\tilde{B} \quad \text{for } \tilde{B} \in \mathbb{R}^{m\times p}. \tag{4}$$

Proof. Note first that $(SA)e^{(i)} = S(Ae^{(i)})$ by (1). Now we have

$$A\tilde{x} = A\left(\sum e^{(i)}\tilde{x}_i\right) \subseteq \sum A(e^{(i)}\tilde{x}_i) = \sum (Ae^{(i)})\tilde{x}_i$$



by subadditivity and homogeneity of matrix multiplication, so that by the axioms for a sublinear map,

$$S(A\tilde{x}) \subseteq \sum S((Ae^{(i)})\tilde{x}_i) = \sum (S(Ae^{(i)}))\tilde{x}_i$$
$$= \sum (SA)e^{(i)}\tilde{x}_i = (SA)\tilde{x}.$$

This implies (3), and (4) follows by applying (3) to each column of $\tilde{B}$. ■

Proposition 2. *A sublinear map $S: \mathbb{IR}^n \to \mathbb{IR}^n$ is regular if there is a matrix $\tilde{A} \in \mathbb{R}^{n\times n}$ such that $S\tilde{A}$ is a regular interval matrix.*

Proof. Suppose that $S\tilde{A}$ is regular and $0 \in S\tilde{x}$. If $\tilde{A}\tilde{z} = 0$ for some $\tilde{z} \in \mathbb{R}^n$ then $0 = S(\tilde{A}\tilde{z}) \subseteq (S\tilde{A})\tilde{z}$ by (3), whence $\tilde{z} = 0$ by (2).

Hence $\tilde{A}$ is regular, and $\tilde{y} := \tilde{A}^{-1}\tilde{x}$ exists. Now $0 \in S\tilde{x} = S(\tilde{A}\tilde{y}) \subseteq (S\tilde{A})\tilde{y}$ by (3) whence $\tilde{y} = 0$ by (2). Hence $\tilde{x} = \tilde{A}\tilde{y} = 0$ so that $S$ is regular. ■

Easily recognizable regular interval matrices are $H$-matrices. A matrix $A \in \mathbb{IR}^{n\times n}$ is called an *$H$-matrix* if there is a vector $u \in \mathbb{R}^n$ with positive entries such that the comparison matrix

$$\langle A \rangle := (A'_{ik}), \text{ where } A'_{ik} := \begin{cases} \inf\{|\tilde{a}| \mid \tilde{a} \in A_{ii}\} & \text{if } i = k, \\ -|A_{ik}| & \text{otherwise,} \end{cases} \tag{5}$$

satisfies the relation $\langle A \rangle u > 0$. In particular, the all-one vector $u = (1, \ldots, 1)^T$ works for diagonally dominant matrices. By Alefeld [1], or [12], Lemma 7 (ii), all $H$-matrices are regular, whence we have:

Corollary 1. *A sublinear map $S: \mathbb{IR}^n \to \mathbb{IR}^n$ is regular if there is a matrix $\tilde{A} \in \mathbb{IR}^{n\times n}$ such that $S\tilde{A}$ is an $H$-matrix.* ■

We end this section by showing that every sublinear map is continuous with respect to the vector valued *distance* $q(x, y) := |\check{x} - \check{y}| + |\varrho(x) - \varrho(y)|$ ([4], [12]). As a preparation we need the following characterization of the distance.

Lemma 1. $q(x, y) = \inf\{r \in \mathbb{R}^n \mid r \geqq 0, x \subseteq y + [-r, r], y \subseteq x + [-r, r]\}$.

Proof. Put $p := \varrho(x) - \varrho(y)$. We have $x \subseteq y + [-r, r] =: z$ iff $|\check{x} - \check{z}| + \varrho(x) \leqq \varrho(z)$ iff $|\check{x} - \check{y}| + \varrho(x) \leqq \varrho(y) + r$ (since $\check{z} = \check{y}$, $\varrho(z) = \varrho(y) + r$) iff $p \leqq r - |\check{x} - \check{y}|$. Similarly $y \subseteq x + [-r, r]$ iff $-p \leqq r - |\check{x} - \check{y}|$. Hence both relations hold iff $|p| \leqq r - |\check{x} - \check{y}|$, or equivalently, iff $r \geqq |\check{x} - \check{y}| + |p| = q(x, y)$. ■

Proposition 3. *If $S: \mathbb{IR}^n \to \mathbb{IR}^n$ is sublinear then*

$$q(Sx, Sy) \leqq |S| q(x, y) \text{ for all } x, y \in \mathbb{IR}^n. \tag{6}$$

*In particular, every sublinear map is continuous.*

Proof. Put $r := q(x, y)$. By Lemma 1, $x \subseteq y+[-r, r]$ whence $Sx \subseteq \subseteq Sy+S[-r, r] \subseteq Sy+[-|S|r, |S|r]$ by the axioms for a sublinear map and rule (R7) of [12]. Similarly, $Sy \subseteq Sx+[-|S|r, |S|r]$ whence, by Lemma 1 again, $q(Sx, Sy) \leqq |S|r = |S|q(x, y)$. This implies continuity of $S$. ■

## 3. Regularity of inverses.

We say that the map $A^I := \mathbb{IR}^n \to \mathbb{IR}^n$ is an *inverse* of the regular matrix $A \in \mathbb{IR}^{n \times n}$ if $A^I$ is sublinear and satisfies

(7) $$\tilde{A} \in A, \tilde{x} \in \mathbb{R}^n \Rightarrow \tilde{A}^{-1}\tilde{x} \in A^I\tilde{x}.$$

Particular inverses discussed in [12] are:

1. The *hull inverse* $A^H$, defined for all regular matrices $A \in \mathbb{IR}^{n \times n}$ by

(8) $$A^H x := \square\{\tilde{A}^{-1}\tilde{x} | \tilde{A} \in A, \tilde{x} \in x\},$$

where $\square S := [\inf S, \sup S]$ denotes the smallest interval vector containing a bounded set $S$.

2. The *Gauss inverse* $A^G$ for a class of matrices containing all $H$-matrices (Alefeld [1]) and certain classes of tridiagonal and Hessenberg matrices (Reichmann [16], [17]), where $A^G x$ is defined as the result of Gauss elimination applied to the coefficient matrix $A$ and righthand side $x$.

3. The *fixpoint inverse* $A^F$ for $H$-matrices $A$, where $A^F x$ is defined as the limit of interval Jacobi or Gauss-Seidel iteration applied to the coefficient matrix $A$ and righthand side $x$.

In general, since $\tilde{x} \in x$ implies $A^I\tilde{x} \subseteq A^I x$, it is obvious from (7) and (8) that

$$A^H x \subseteq A^I x$$

for every inverse $A^I$ of $A$ and every $x \in \mathbb{IR}^n$. Thus $A^I x$ is an interval enclosure of the solutions of the formal linear interval equation $Ay = x$, interpreted as the set of equations $\tilde{A}\tilde{y} = \tilde{x}$, $(\tilde{A} \in A, \tilde{x} \in x)$.

For applications to nonlinear systems of equations (see Section 4), it is important to know when an inverse is regular. A simple criterion is the application of Corollary 1 with $S = A^I$ and $\tilde{A} = \check{A}$; then $I = \check{A}^{-1}\check{A} \in A^I\check{A}$, and for small radii of $A$ the matrix $A^I\check{A}$ can be expected to be close to $I$, hence diagonally dominant, or at least an $H$-matrix. But this criterion requires the application of $A^I$ to all $n$ columns of $\check{A}$ which may be unpractical if $n$ is large. Hence it is of interest to have simpler tests of regularity. In this section we give such a test for the fixpoint inverse. It is based on the following auxiliary result.



Lemma 2. *If $a, b \in \mathbb{IR}$ then*

(9) $$\varrho(ab) \leqq |a|\varrho(b)+\varrho(a)|\bar{b}|,$$

(10) $$\varrho(ab) = |a|\varrho(b) \quad \text{if} \quad 0 \notin \operatorname{int}(a), 0 \in b,$$

(11) $$\varrho(b/a) = \varrho(b)/\langle a\rangle \quad \text{if} \quad 0 \notin a, 0 \in b.$$

*Here $int(a) = a \backslash \{\underline{a}, \bar{a}\}$, and $\langle a\rangle = \inf\{|\tilde{a}| \; |\tilde{a} \in a\}$ is a special case of* (5).

Proof. (9) is formula (B13) of [12], and (10) is part of Theorem 10 in Chapter 2 of Alefeld and Herzberger [4]. Finally suppose $0 \notin a$, $0 \in b$. If $a > 0$ then $\langle a\rangle = \underline{a}$ and $b/a = [\underline{b}/\underline{a}, \bar{b}/\underline{a}]$ whence $\varrho(b/a) = \frac{1}{2}(\bar{b}/\underline{a} - \underline{b}/\underline{a}) = \frac{1}{2}(\bar{b}-\underline{b})/\underline{a} = \varrho(b)/\langle a\rangle$. The case $a < 0$ is treated in the same way. ■

Proposition 4. *Let $A \in \mathbb{IR}^{n \times n}$ be an $H$-matrix, and define the matrix $\varrho^*(A) \in \mathbb{R}^{n \times n}$ by*

(12) $$\varrho^*(A)_{ik} := \begin{cases} \varrho(A_{ik}) & \text{if } i \neq k, 0 \in int(A_{ik}), \\ 0 & \text{otherwise.} \end{cases}$$

*If there is a vector $u \in \mathbb{R}^n$ with positive entries such that*

(13) $$\langle A\rangle u > \varrho^*(A)u$$

*then the fixpoint inverse $A^F$ is regular.*

Proof. Let $\tilde{x} \in \mathbb{R}^n$ satisfy $0 \in A^F\tilde{x}$, and put $y := A^F\tilde{x}$. Then $0 \in y$, and by [12], Section 7, $y$ satisfies the equations

(14) $$y_i = \Big(\tilde{x}_i - \sum_{k \neq i} A_{ik}y_k\Big)\Big/ A_{ii} \quad (i = 1, \ldots, n).$$

Since $A$ is an $H$-matrix, $0 \notin A_{ii}$, and since $0 \in y_i$, the expression in the brackets contains zero. Hence by repeated application of Lemma 2,

$$\begin{aligned} \varrho(y_i) &= \varrho(\tilde{x}_i - \sum_{k \neq i} A_{ik}y_k)/\langle A_{ii}\rangle \\ &= \Big(\sum_{k \neq i} \varrho(A_{ik}y_k)\Big)\Big/ \langle A_{ii}\rangle \\ &\leqq \Big(\sum_{k \neq i} (|A_{ik}|\varrho(y_k)+\varrho^*(A)_{ik}|\bar{y}_k|)\Big)\Big/ \langle A_{ii}\rangle, \end{aligned}$$

so that by (5)

$$(\langle A\rangle\varrho(y))_i = \langle A_{ii}\rangle\varrho(y_i) - \sum_{k\neq i}|A_{ik}|\varrho(y_k)$$
$$\leqq \sum_{k\neq i}\varrho^*(A)_{ik}|\check{y}_k| = (\varrho^*(A)|\check{y}|)_i.$$

Therefore $\langle A\rangle\varrho(y) \leqq \varrho^*(A)|\check{y}|$. But $|\check{y}| \leqq \varrho(y)$ since $0\in y$, and since $\langle A\rangle^{-1} \geqq 0$ (Lemma 11 of [12]), we have

(15) $$|\check{y}| \leqq \varrho(y) \leqq \langle A\rangle^{-1}\varrho^*(A)|\check{y}|.$$

On the other hand, (13) implies

(16) $$\langle A\rangle^{-1}\varrho^*(A)u < u.$$

Now if $\check{y} \neq 0$ then with

(17) $$\alpha := \max\{|\check{y}_i|/u_i \mid i = 1,\dots,n\} > 0$$

we have $|\check{y}| \leqq \alpha u$ so that by (15) and (16), $|\check{y}| \leqq \langle A\rangle^{-1}\varrho^*(A)\alpha u < \alpha u$, contradicting (17). Therefore $\check{y} = 0$, whence $\varrho(y) = 0$ by (15), so that $y = 0$, and $\tilde{x} = 0$ by (14). Hence (2) holds for $S = A^F$ and $A^F$ is regular. ■

Corollary 2. *If $A \in \mathbb{IR}^{n\times n}$ is an H-matrix such that $0 \notin \operatorname{int}(A_{ik})$ for $i,k = 1,\dots,n$ then $A^F$ is regular. In particular, $A^F$ is regular if A is an M-matrix (i.e. an H-matrix with $A_{ii} \geqq 0$, $A_{ik} \geqq 0$ for $i \neq k$).*

Proof. In this case $\varrho^*(A) = 0$, so that (13) holds for some $u > 0$ by definition of an $H$-matrix. ■

Corollary 3. *Under the assumptions of Proposition 4 or Corollary 2, the hull inverse $A^H$ is regular.*

Proof. If $\tilde{x} \in \mathbb{R}^n$ satisfies $0 \in A^H x$ then $0 \in A^F x$ by (8) so that $\tilde{x} = 0$. ■

We end this section with two examples. As our first example we consider

$$A = \begin{pmatrix} 3 & [-2,2] & 0 \\ 0 & 3 & [-2,2] \\ [-2,2] & 0 & 3 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix}.$$

$A$ is a diagonally dominant $H$-matrix, and since

$$\begin{pmatrix} 3 & 2 & 0 \\ 0 & 3 & -1.5 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 2 \\ -1.5 & 0 & 3 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 & -1.5 & 0 \\ 0 & 3 & 0 \\ 2 & 0 & 3 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix},$$



we have $$\begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \in A^H\tilde{x}, \quad \begin{pmatrix} [0,3] \\ [0,3] \\ [0,3] \end{pmatrix} \subseteq A^H\tilde{x}.$$

Therefore $0 \in A^H\tilde{x}$ so that $A^H$ is not regular. Therefore no inverse $A^I$ can be regular. In particular, this example shows the necessity of a condition like (13) in Proposition 4.

As our second example we consider

$$A = \begin{pmatrix} 1 & -2 \\ [-5,-1] & 11 \end{pmatrix}, \quad \tilde{x} = \begin{pmatrix} 2 \\ -9 \end{pmatrix}.$$

$A$ is an $M$-matrix and has the triangular factorization $(L, R)$ with

$$L = \begin{pmatrix} 1 & 0 \\ [-5,-1] & 1 \end{pmatrix}, \quad R = \begin{pmatrix} 1 & -2 \\ 0 & [1,9] \end{pmatrix}.$$

Therefore $A^G\tilde{x} = \begin{pmatrix} [-16,4] \\ [-7,1] \end{pmatrix} \ni 0$, so that $A^G$ is not regular. In particular, Corollary 2 and hence Proposition 4 become false when $A^F$ is replaced by $A^G$.

## 4. Enclosing a zero of nonlinear systems of equations.

In this section we consider the equation $F(x^*) = 0$ for continuously differentiable maps $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. We call an $x^* \in D$ with $F(x^*) = 0$ a *zero* of $F$. The set of all interval vectors $x \in \mathbb{IR}^n$ contained in $D$ will be denoted by $\mathbb{I}D$.

Theorem 1. *Let $F: D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable function, and suppose that $x \in \mathbb{I}D$, $A \in \mathbb{IR}^{n\times n}$ are such that*

(18) $$F'(\tilde{x}) \in A \quad \text{for all } \tilde{x} \in x.$$

*Suppose, moreover, that A is regular and $A^I$ is an inverse of A. Then the following assertions hold.*

(i) *If $x^* \in x$ is a zero of F then*

(19) $$x^* \in \tilde{x} - A^I F(\tilde{x}) \quad \text{for all } \tilde{x} \in x.$$

(ii) *There is at most one $x^* \in x$ with $F(x^*) = 0$.*

(iii) *If, for some $\tilde{x} \in \operatorname{int}(x)$,*

(20) $$\tilde{x} - A^I F(\tilde{x}) \subseteq x$$

*then there is a unique $x^* \in x$ with $F(x^*) = 0$.*

PROOF. Fix $\tilde{x}, \tilde{y} \in x$. Since interval vectors are convex sets, $\tilde{y}+t(\tilde{x}-\tilde{y}) \in x$ for all $t \in [0, 1]$. By (18),

$$\tilde{A}(t) := F'(\tilde{y}+t(\tilde{x}-\tilde{y})) \in A,$$

and, with the integral defined componentwise,

$$\tilde{A} := \int_0^1 \tilde{A}(t)dt \in A$$

since $A$ is convex. In particular, $\tilde{A}$ is regular. We show that

$$\tilde{y} = \tilde{x} - \tilde{A}^{-1}(F(\tilde{x}) - F(\tilde{y})). \tag{21}$$

Indeed, the map $\varphi : [0, 1] \to \mathbb{R}^n$ defined by $\varphi(t) := F(\tilde{y}+t(\tilde{x}-\tilde{y}))$ is continuously differentiable with derivative

$$\varphi'(t) = F'(\tilde{y}+t(\tilde{x}-\tilde{y}))(\tilde{x}-\tilde{y}) = \tilde{A}(t)(\tilde{x}-\tilde{y}),$$

whence $\tilde{A}(\tilde{x}-\tilde{y}) = \int_0^1 \tilde{A}(t)(\tilde{x}-\tilde{y})dt = \int_0^1 \varphi'(t)dt$

$$= \varphi(1) - \varphi(0) = F(\tilde{x}) - F(\tilde{y}).$$

Hence $\tilde{x}-\tilde{y} = \tilde{A}^{-1}(F(\tilde{x})-F(\tilde{y}))$ which implies (21). Now suppose that

$$F(\tilde{y}) = \tau F(\tilde{x}). \tag{22}$$

Then by (21) and (7),

$$\tilde{y} = \tilde{x} - (1-\tau)\tilde{A}^{-1}F(\tilde{x}) \in \tilde{x} - (1-\tau)A^I F(\tilde{x}). \tag{23}$$

In particular, if $\tilde{y} = x^*$ is a zero of $F$ then (22) holds with $\tau = 0$ and (23) simplifies to (19). Hence (i) holds. Moreover, if $\tilde{x}$ is another zero of $F$ then (19) implies $x^* \in \tilde{x}$, so that there is at most one zero of $F$ in $x$, and (ii) holds as well.

To prove (iii), suppose that $\tilde{x} \in \operatorname{int}(x)$ and (20) holds. Then $A^I F(\tilde{x}) = \tilde{x} - (\tilde{x} - A^I F(\tilde{x})) \subseteq \tilde{x} - x$, and (23) implies $\tilde{y} \in \tilde{x} - (1-\tau)(\tilde{x}-x) = \tau\tilde{x} + (1-\tau)x$. In particular,

$$\tilde{y} \in \operatorname{int}(x) \text{ if (22) holds with } \tau \in [0, 1]. \tag{24}$$

By the inverse function theorem there is a nonempty interval $(\tau, 1]$ such that for all $t \in (\tau, 1]$, the equation $F(\tilde{z}) = tF(\tilde{x})$ has a solution $\tilde{z} = \psi(t) \in x$. Let $\tau$ be the smallest nonnegative number with this property. Since $\psi(t) \in x$ and $x$ is compact there is an accumulation point $\tilde{y} \in x$ of $\psi(t)$ for $t \to \tau$, and since $F$ is continuous, $\tilde{y}$ satisfies (22). If $\tau > 0$ then $\tilde{y} \in \operatorname{int}(x)$ by (24); therefore the inverse function theorem again asserts the solvability of $F(\tilde{z}) = tF(\tilde{x})$ by $\tilde{z} \in x$ for $t$ now in a neighbourhood of $\tau$, in contradiction to the construction of $\tau$. Hence $\tau = 0$, and $x^* := \tilde{y}$ is a zero of $F$ in $x$. Uniqueness follows from part (ii). ■

REMARK. This theorem is a slight extension of well-known results. In the special case where $A^I$ is multiplication with a matrix containing $A^{-1} = \square\{\tilde{A}^{-1} | \tilde{A} \in A\}$, Theorem 1 goes back to Moore [11] for part (i), and to Kahan [8], Nickel [14] for parts (ii) and (iii). The case where $A^I = A^G$ is treated in Alefeld [2]. A novelty of the present proof is that it replaces the use of Brouwer's fixpoint theorem (needed for older proofs) by the more elementary inverse function theorem. Note that $A$ is easily computed as an interval extension of an arithmetic expression for the derivative $F'(x)$; see [2], [11].

The zero-enclosing vector $\tilde{x} - A^I F(\tilde{x})$ has some useful properties which will be exploited in the next section.

PROPOSITION 5. *Under the assumption of Theorem* 1, *suppose that* $\tilde{x} \in x$ *and*

$$x' := \tilde{x} - A^I F(\tilde{x}). \tag{25}$$

*If $F$ has a zero $x^* \in x$ then*

$$\varrho(x') \leqq \varrho(A^I A)|\tilde{x} - x^*|, \tag{26}$$

$$|x' - x^*| \leqq |I - A^I A||\tilde{x} - x^*|. \tag{27}$$

*Moreover, if $A^I\tilde{A}$ is an H-matrix then*

$$\tilde{x} \notin x' \text{ unless } \tilde{x} \text{ is a zero of } F. \tag{28}$$

PROOF. From (21) with $\tilde{y} = x^*$ we obtain $F(\tilde{x}) = \tilde{A}(\tilde{x} - x^*)$ with $\tilde{A} \in A$ so that by Proposition 1 and sublinearity,

$$A^I F(\tilde{x}) = A^I(\tilde{A}(\tilde{x}-x^*)) \subseteq (A^I\tilde{A})(\tilde{x}-x^*) \subseteq (A^I A)(\tilde{x}-x^*).$$

Hence (25) implies

$$\varrho(x') = \varrho(A^I F(\tilde{x})) \leqq \varrho((A^I A)(\tilde{x}-x^*)) \leqq \varrho(A^I A)|\tilde{x}-x^*|$$

by rule (B14) of [12], which gives (26). Similarly,

$$|x' - x^*| = |\tilde{x} - x^* - A^I F(\tilde{x})| \leqq |\tilde{x} - x^* - (A^I A)(\tilde{x}-x^*)|$$

$$= |(I - A^I A)(\tilde{x}-x^*)| \leqq |I - A^I A||\tilde{x}-x^*|$$

by (B21) and (B4) of [12], which gives (27). Finally, if $A^I \check{A}$ is an $H$-matrix then $A^I$ is regular by Proposition 2. Hence if $\tilde{x} \in x'$ then $0 \in A^I F(\tilde{x})$ by (25) and $F(\tilde{x}) = 0$ by regularity of $A^I$. By part (ii) of Theorem 1, $\tilde{x} = x^*$, which proves (28). ■

REMARK. The condition "$A^I \check{A}$ is an $H$-matrix" in Proposition 5 can be replaced by any other condition which implies regularity of $A^I$.

Proposition 5 is related to a recent paper of Alefeld [3]. In a discussion of some interval methods for the solution of nonlinear equations (cf. Section 5 below), Alefeld uses the Gauss inverse $A^G$ of the matrix $A = F'(x^0)$, an interval evaluation of the derivative, and relates with it two matrices

$$P := |I - \kappa(A^G)A|, \qquad Q := 2\varrho(\kappa(A^G))|A|.$$

(In [3], $A^G x$ is written as $IGA(A, x)$, and the kernel $\kappa(A^G)$ is written as $IGA(A)$). The proof of the main result of [3] then depends essentially on the fact that the vector $x' := \tilde{x} - A^G F(\tilde{x})$ satisfies

$$|x' - x^*| \leq P|\tilde{x} - x^*|, \qquad \varrho(x') \leq \tfrac{1}{2}Q|\tilde{x} - x^*|,$$

inequalities closely related to (26) and (27). Moreover, Lemma 5 of [3] contains a proof of (28) under the assumption that the spectral radius of $P$ is less than one.

## 5. Convergence theorems for interval iteration.

With the terminology of the last section we consider here the question how to find arbitrarily good interval enclosures $x^l$ $(l = 0, 1, 2, \ldots)$ for a zero $x^*$ of $F$ in an initial interval vector $x$ satisfying (18) with regular $A$.

THEOREM 2. *Under the assumptions of Theorem 1, suppose that $x^l$ $(l = 0, 1, 2, \ldots)$ is a sequence with $x^0 = x$, $x^l \in \mathbb{I}D \cup \{\phi\}$ $(l = 1, 2, \ldots)$ such that $x^{l+1} = \phi$ if $x^l = \phi$, and otherwise*

(29) $$x^{l+1} = x^l \cap (\tilde{x}^l - A^I F(\tilde{x}^l))$$

*for suitable $\tilde{x}^l \in x^l$. Then*

(30) $$x^{l+1} \subseteq x^l \subseteq \ldots \subseteq x^0 = x,$$

*and the following assertions hold:*

(i) *If $F$ has a (unique) zero $x^*$ in $x$ then $x^* \in x^l$ for all $l \geq 0$.*

(ii) *If $x^l = \phi$ for some $l > 0$ then $x$ contains no zero of $F$.*

*Moreover, if $A^I$ is regular then:*

(iii) *If $F$ has a (unique) zero $x^*$ in $x$ then $\lim_{l \to \infty} x^l = x^*$.*

(iv) *If $x$ contains no zero of $F$ then $x^l = \phi$ for some $l > 0$.*

PROOF. (30) is obvious, and by Theorem 1 (ii), $F$ has at most one zero $x^*$ in $x$. If such a zero is in $x^l$, $x^* \in x^l$, then by Theorem 1 (i) – with $x^l$, $\tilde{x}^l$ in place of $x$, $\tilde{x}$ – we have $x^* \in \tilde{x}^l - A^I F(\tilde{x}^l)$ so that $x^* \in x^{l+1}$ by (29). Induction therefore implies (i). If $x^l = \phi$ for some $l > 0$ then the conclusion of (i) cannot be valid; hence in this case $F$ has no zero in $x$, and (ii) holds.

Suppose now that $A^I$ is regular. If $x^l \neq \phi$ for all $l > 0$ then the (bounded) sequence $\tilde{x}^l$, $l = 0, 1, 2, \ldots$ contains a convergent subsequence $\tilde{x}^{l_i}$ with limit $\bar{x}$, say. By (30) and (29), $\bar{x} \in x^{l+1} \subseteq x^l - A^I F(\tilde{x}^l)$ whence by continuity of $F$ and of the sublinear map $A^I$ (Proposition 3) we have $\bar{x} \in \bar{x} - A^I F(\bar{x})$. Therefore $0 \in A^I F(\bar{x})$, and since $A^I$ is regular, $F(\bar{x}) = 0$. Hence $x^* := \bar{x}$ is a zero of $F$ in $x^0 = x$. At this stage, (iv) follows. Moreover, if $i \to \infty$ then $\tilde{x}^{l_i} - A^I F(\tilde{x}^{l_i}) \to \bar{x} - A^I F(\bar{x}) = \bar{x} = x^*$ whence $x^{l_i+1} \to x^*$ by (29) and $\lim_{l \to \infty} x^l = x^*$ by (30). Therefore (iii) holds. ■

An interesting and important feature of the preceding theorems is the freedom in the choice of $\tilde{x}^l \in x^l$. We should like to exploit this by choosing $\tilde{x}^l$ in such a way that the new enclosure $x^{l+1}$ has a small radius.

Assuming the existence of a zero $x^* \in x$ we have by (29) and Proposition 5 the relation

(31) $$\varrho(x^{l+1}) \leq \varrho(\tilde{x}^l - A^I F(\tilde{x}^l)) \leq \varrho(A^I A)|\tilde{x}^l - x^*|.$$

Therefore, if $\tilde{x}^l$ is chosen as a good approximation of $x^*$, the new radius will be small.

In order to predict a good approximation $\tilde{x}^l$ of $x^*$, a considerable number of local methods are available; many of them are described in the book by Ortega and Rheinboldt [15]. For the sake of definiteness we shall use as "predictor" one step of Newton's method from the previous approximation $x^{l-1}$:

(32) $$\tilde{u}^l := \tilde{x}^{l-1} - F'(\tilde{x}^{l-1})^{-1} F(\tilde{x}^{l-1}).$$

Since it may happen that $\tilde{u}^l$ falls outside the most recent enclosure $x^l$ we move $\tilde{u}^l$ into $x^l$ by means of the *cut-off operator* $\kappa: \mathbb{R}^n \times \mathbb{IR}^n \to \mathbb{R}^n$ defined by

(33) $$\kappa(\tilde{u}, x)_i := \begin{cases} \underline{x}_i & \text{if } \tilde{u}_i < \underline{x}_i, \\ \bar{x}_i & \text{if } \tilde{u}_i > \bar{x}_i, \\ \tilde{u}_i & \text{otherwise} \end{cases}$$

(cf. Cornelius and Alefeld [6]). For all $\tilde{u} \in \mathbb{R}^n, x \in \mathbb{IR}^n$, the cut-off operator satisfies

(34) $$\kappa(\tilde{u}, x) \in x,$$

$$\kappa(\tilde u, x) = \tilde u \qquad \text{if } \tilde u \in x, \tag{35}$$

$$|\kappa(\tilde u, x) - x^*| \leqq |\tilde u - x^*| \qquad \text{if } x^* \in x. \tag{36}$$

Since, in our application, the zero $x^*$ is in $x^l$, the corrected vector

$$\tilde x^l := \kappa(\tilde u^l, x^l)$$

belongs to $x^l$, agrees with the Newton approximation if $\tilde u^l \in x^l$, and is otherwise a better approximation for $x^*$. Regularity of $A^I$ and local quadratic convergence of the predictor method (Newton's iteration) now force global quadratic convergence of the interval iteration:

THEOREM 3. *Under the assumptions of Theorem 1, suppose that $F$ has a zero $x^* \in x$ and $A^I$ is regular. Then the three sequences $x^l$, $\tilde u^l$, $\tilde x^l$ ($l = 0, 1, 2, \ldots$) with*

$$x^0 := x, \tilde x^0 := \tilde u^0 := \check x,$$

*and for $l = 0, 1, 2, \ldots$*

$$x^{l+1} := x^l \cap (\tilde x^l - A^I F(\tilde x^l)),$$
$$\tilde u^{l+1} := \tilde x^l - F'(\tilde x^l)^{-1} F(\tilde x^l),$$
$$\tilde x^{l+1} := \kappa(\tilde u^{l+1}, x^{l+1}),$$

*are well-defined and converge to $x^*$ superlinearly. Moreover, if $F'(x)$ is Lipschitz continuous then there are constant $\gamma_1$ and $\gamma_2$ such that*

$$\|\tilde x^{l+1} - x^*\|_\infty \leqq \gamma_1 \|\tilde x^l - x^*\|_\infty^2 \tag{37}$$

$$\|\varrho(x^{l+1})\|_\infty \leqq \gamma_2 \|\tilde x^l - x^*\|_\infty^2. \tag{38}$$

*In particular, $\tilde x^l \to x^*$ and $\varrho(x^l) \to 0$ with $R$-convergence order $\geqq 2$.*

PROOF. Global convergence follows from (34) and Theorem 2, and (37) follows from (36) and $\|\tilde u^{l+1} - x^*\| \leqq \gamma_1 \|\tilde x^l - x^*\|^2$, the well-known ($Q$-)quadratic convergence property of Newton's iteration. (38) follows from (31). Finally, the remark on the convergence order is an immediate consequence of (37), (38), and the discussion of $R$-order in Ortega and Rheinboldt [15]. ■

REMARKS. 1. If, instead of Newton's method a different predictor is chosen such that the relation

$$\|\tilde u^{l+1} - x^*\| \leqq \gamma_1 \|\tilde x^l - x^*\|^{c_0} \|\tilde x^{l-1} - x^*\|^{c_1} \ldots \|\tilde x^{l-s} - x^*\|^{c_s}$$

is satisfied then similar arguments show that the $R$-order of $\tilde x^l$ and $\varrho(x^l)$ is at least the positive solution $r$ of

$$r^{s+1} = c_0 r^s + c_1 r^{s-1} + \ldots + c_s.$$

In particular, many local methods can be "globalized" by an interval iteration in the spirit of Theorem 3, without loss of convergence speed.

2. In finite precision arithmetic, the attainable limit accuracy is mainly determined by the accuracy with which the "residual" $F(\tilde x)$ can be enclosed. Rounded interval arithmetic (as discussed e.g. in Moore [11] or Alefeld and Herzberger [4]) automatically provides an enclosure or $F(\tilde x)$ if $F$ is given by an arithmetic expression and is evaluated at the degenerate interval $\tilde x = [\tilde x, \tilde x]$. The radius of the calculated interval vector is then roughly proportional to the relative accuracy of single machine operations (see Theorem 5 in Chapter 4 of [4]); therefore the same will hold for the limiting zero-enclosing vector.

3. The cost of the interval algorithm exceeds the cost of the predictor by

(i) the interval evaluation of $F'(x)$ to compute $A$,
(ii) the construction of an inverse $A^I$,
(iii) the evaluation of $A^I F(\tilde x^l)$.

The cost of (i) is a small multiple of the cost for the ordinary evaluation of $F'(\tilde x)$. The cost of (ii) and (iii) depends on the inverse, but is generally $O(n^3)$ for (ii) and $O(n^2)$ for (iii). In particular, the extra $O(n^3)$-effort has to be made only once.

4. In the one-dimensional case, $A^I x := x/A$ defines a regular inverse $A^I$ of $A$. The contents of this section may therefore be regarded as an extension of the one-dimensional iteration considered in Neumaier [13].

For $A^I = A^G$, the iteration (29) is considered in Alefeld [3] as the "method (SN)". His main result concerning (SN) is that (iii), (iv) of Theorem 2 hold if the spectral radius of the matrix $|I - B|$, where

$$B := \kappa(A^G) A \in \mathbb{IR}^{n \times n},$$

is less than one (his other sufficient condition, that $Q := 2\varrho(B)$ has spectral radius less than one is weaker since $I = \check A^{-1} \check A \in A^G A \subseteq \kappa(A^G) A = B$ so that $|I - B| = |I - \check B| + \varrho(B) \leqq 2\varrho(B) = Q$). Now if $|I - B|$ has spectral radius $< 1$ then there is a vector $u > 0$ such that $|I - B| u < u$ (rule (P7) of [12]); hence $\langle B \rangle u = \langle I - (I - B) \rangle u \geqq (\langle I \rangle - |I - B|) u = u - |I - B| u > 0$, so that $B$ is an $H$-matrix. In particular, since

$$A^G \check A \subseteq A^G A \subseteq \kappa(A^G) A = B,$$

$A^G$ is regular. Therefore, Theorem 2 strengthens the convergence condition in

[3] for method (SN). It is not difficult to see that this also implies a corresponding stronger convergence condition for method (N) in [3].

A disadvantage of the Gauss inverse $A^G$ is its restricted applicability to matrices for which $A^G$ is defined. To increase its range it is sensible to precondition the matrix $A$ by an approximation $C \in \mathbb{R}^{n \times n}$ to the midpoint inverse $\check{A}^{-1}$; see [7], [12]. This corresponds to the use of the iteration (29) with the map $A^I$ defined by

$$A^I x := (CA)^G (Cx); \tag{39}$$

sublinearity and the inverse property of $A^I$ follow immediately from the corresponding properties of the Gauss inverse.

Another iteration of the type (29) is defined for $H$-matrices $A$ through the inverse $A^K$ defined by

$$A^K x := x + \langle A \rangle^{-1} |(I-A)x| [-1, 1]. \tag{40}$$

Sublinearity is obvious, and the inverse property is shown as follows. If $\tilde{A}\tilde{z} = \tilde{x}$, $\tilde{A} \in A$, $\tilde{x} \in x$ then $\tilde{z} - \tilde{x} = \tilde{A}^{-1}\tilde{x} - \tilde{x} = \tilde{A}^{-1}(I-\tilde{A})\tilde{x} \in A^{-1}(I-A)x$ whence by [12], Lemma 11, we have $|\tilde{z}-\tilde{x}| \leqq |A^{-1}(I-A)x| \leqq \langle A \rangle^{-1} |(I-A)x|$ so that $\tilde{z} \in \tilde{x} + \langle A \rangle^{-1} |(I-A)x| [-1, 1] \subseteq A^K x$.

A sufficient condition for the regularity of $A^K$ is the existence of a vector $u > 0$ such that

$$\langle A \rangle u > |I - A| u; \tag{41}$$

indeed then the nonnegative matrix $R := \langle A \rangle^{-1} |I-A|$ satisfies $Ru < u$ so that the spectral radius of $R$ is less than one (rule (P7) of [12]). Therefore, $\kappa(A^K) = A^K I = I + R[-1, 1]$ is regular and hence $A^K$ is regular.

As before with $A^G$, it is sensible to use $A^K$ in conjunction with preconditioning; hence (29) is used with the inverse $A^I$ defined by

$$A^I x := (CA)^K Cx = Cx + \langle CA \rangle^{-1} |(I - CA)Cx| [-1, 1] \tag{42}$$

which is defined if $CA$ is an $H$-matrix, and regular if a vector $u > 0$ exists with

$$\langle CA \rangle u > |I - CA| u. \tag{43}$$

In the particular case where $C = \check{A}^{-1}$, the matrix $CA$ has midpoint $I$ and radius $|C|\varrho(A)$ whence $\langle CA \rangle = I - |C|\varrho(A)$ and $I - CA = |C|\varrho(A)[-1, 1]$; hence (42) can be written as

$$A^I x = Cx + (I - |C|\varrho(A))^{-1} |C|\varrho(A)|x| [-1, 1], \quad C = \check{A}^{-1}, \tag{44}$$



defined if there is a vector $u > 0$ such that $|C|\varrho(A)u < u$, and regular if there is a vector $u > 0$ such that $|C|\varrho(A)u < \frac{1}{2}u$. The iteration (29) with the inverse (44) appears already in the literature [9], [10].

## 6. Numerical results.

We illustrate our method with several examples. The algorithm described in Theorem 3 is used with the Gauss inverse without preconditioning in Example 1 and 2, and the preconditioned Gauss inverse in Example 3. The results were obtained on a SPERRY 1100/82 (mantissa length 27 binary digits) with the INTERVAL FORTRAN package of Yohe [19].

EXAMPLE 1. The map $F: \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 + 9x_1 + x_2 - 36 \\ x_1 + x_2^2 + 10x_2 - 3 \end{pmatrix}$$

has a unique zero $x^* = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ in the interval $x^{(0)} = \begin{pmatrix} [-4, 4] \\ [-4, 4] \end{pmatrix}$.

The algorithm gives the following result:

| step | predictor | zero approx. | inclusion vector |
|---|---|---|---|
| 1 | – | .0 | (−4., 4.) |
| | | .0 | (−4., .8823530) |
| 2 | 4.0112359 | 4.0000000 | (−4., 3.0617024) |
| | −.1012359 | −.1012359 | (−4., .8823530) |
| 3 | 3.0591171 | same | ( 1.2978356, 3.0069026) |
| | −.0049900 | | (−.0026027, .8710299) |
| 4 | 3.0002325 | same | ( 2.9933211, 3.0000281) |
| | −.0000208 | | (−.0000108, .0034226) |
| 5 | 3.0000000 | same | ( 2.9999981, 3.0000011) |
| | −.0000000 | | (−.0000006, .0000010) |

One observes a cutoff in the first step. As expected, convergence is quadratic; the limit accuracy is attained after 5 steps.

EXAMPLE 2. A simple discretization procedure turns the non-linear boundary value problem $y'' = 4-4y+y^3$, $y(0) = y(1) = 0$ into the following system of $n$ equations, with $h = (n+1)^{-1}$:

$$2y_1 - y_2 + h^2(4-4y_1+y_1^3) = 0,$$
$$2y_i - y_{i+1} - y_{i-1} + h^2(4-4y_i+y_i^3) = 0, \quad (i = 2, \ldots, n-1)$$
$$2y_n - y_{n-1} + h^2(4-4y_n+y_n^3) = 0.$$

The problem was tested for $n = 15$ with starting vectors having all components equal to $[\alpha, \beta]$, for various $\alpha, \beta$. For $[\alpha, \beta] = [0, 100]$ nonexistence of a solution was found after 9 steps. For $[\alpha, \beta] = [-100, 0]$, the (unique) solution was enclosed after 14 steps with a maximal error (radius) of $3 \cdot 10^{-8}$; quadratic convergence was observed beginning with step 9 where the maximal error still was 0.834. To check the efficiency of the nonexistence test we tried $[\alpha, \beta] = [-.786, 0]$ which ruled out a solution after 2 steps, and $[\alpha, \beta] = [-.786, 10]$ which did not rule out a solution although none exists: the iteration did not further improve after 8 steps with a maximal error of $7 \cdot 10^{-4}$. Since the absolutely largest component of the solution was $-.7862003$, only $2 \cdot 10^{-4}$ off the bound, the closeness of the solution to the bound impairs here the nonexistence test over the second interval. This is due to the fact that $A^G$ was not regular.

Generally, if $A$ is regular and $A^I$ is defined, the present algorithm yields good results. However, as the final example shows, there are problems in establishing regularity of $A$ if the derivative is singular near points of $x^0$.

EXAMPLE 3. The map $F: \mathbb{R}^4 \to \mathbb{R}^4$ defined by

$$F_i(x) = (x_1^i + x_2^i + x_3^i + x_4^i)/i - c_i \qquad (i = 1, \ldots, 4)$$

has as derivative a Vandermonde matrix; hence $F'(\tilde{x})$ is regular for all $\tilde{x} \in x^0$ for every $x^0$ whose components are disjoint intervals. Unfortunately, the enclosing matrix $F'(x^0)$ was found regular by the (preconditioned) Gauss algorithm only for a much more restrictive choice of $x^0$. For example, the triangular decomposition of the preconditioned matrix $A^{-1}A$, for $A = F'(x^0)$ with $x^0 = (-7, -2, 2, 7)^T + [-\alpha, \alpha](1, 1, 1, 1)^T$, is defined for $\alpha = 0.7$ but not defined for $\alpha = 0.75$; however, $x^0$ has disjoint components for all $\alpha < 2$. The same phenomenon occurs for other highly nonlinear problems where the derivative is singular on an extended manifold close to $x^0$, and makes preliminary bisection steps necessary.

## REFERENCES


1. G. Alefeld. *Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten*, Computing Suppl. 1 (1977), 15–19.
2. G. Alefeld. *Intervallanalytische Methoden bei nichtlinearen Gleichungen*, in: *Jahrbuch Überblicke Mathematik*, Bibl. Inst., Mannheim 1979, pp. 63–78.





3. G. Alefeld, *On the convergence of some intervalarithmetic modifications of Newton's method*, in: *Scientific Computing* (R. S. Stepleman, ed.), North-Holland, Amsterdam 1983, pp. 223–230.
4. G. Alefeld and J. Herzberger, *Introduction to Interval Computations*. Acad. Press, New York-London 1983.
5. H. Cornelius, *Untersuchungen zu einem intervallarithmetischen Iterationsverfahren mit Anwendungen auf eine Klasse nichtlinearer Gleichungssysteme*. Dissertation, Technische Universität Berlin, 1981.
6. H. Cornelius and G. Alefeld, *A device for the acceleration of convergence of a monotonously enclosing iteration method*, in: *Iterative solution of nonlinear systems of equations*, Springer Lecture Notes in Mathematics 953, 1982, pp. 68–79.
7. E. Hansen and R. Smith, *Interval arithmetic in matrix computations*, part II, SIAM J. Numer. Anal. 4 (1967), 1–9.
8. W. M. Kahan, *A more complete interval arithmetic*. Lecture Notes, University of Michigan, 1968.
9. R. Krawczyk, *Interval iterations for including a set of solutions*, Computing, 32 (1984), 13–31.
10. R. Krawczyk and A. Neumaier, *An improved interval Newton operator*, to appear.
11. R. E. Moore, *Interval Analysis*, Prentice-Hall, Englewood Cliffs, N.Y., 1966.
12. A. Neumaier, *New techniques for the analysis of linear interval equations*, Linear Algebra Appl., 58 (1984), 273–325.
13. A. Neumaier, *An interval version of the secant method*, BIT 24 (1984), 366–372.
14. K. Nickel, *On the Newton method in interval analysis*, MRC Techn. Sum. Rep. 1136, University of Wisconsin, Madison 1971.
15. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. Acad. Press, New York-London 1970.
16. K. Reichmann, *Abbruch beim Intervall-Gauß-Algorithmus*, Computing 22 (1979), 355–361.
17. K. Reichmann, *Ein hinreichendes Kriterium für die Durchführbarkeit des Intervall-Gauss-Algorithmus bei Intervall-Hessenbergmatrizen ohne Pivotsuche*, Z. Angew. Math. Mech. 59 (1979), 373–379.
18. H. Schwetlick, *Numerische Lösung nichtlinearer Gleichungen*. R. Oldenburg, München-Wien 1979.
19. J. M. Yohe, *Software for interval arithmetic: a reasonably portable package*. ACM Trans. Math. Software 5 (1979), 50–63.