Topics in Validated Computations
J. Herzberger (Editor)
1994 Elsevier Science B.V.

Inclusion methods for systems of nonlinear equations – the interval Newton method and modifications

G. Alefeld

Institut für Angewandte Mathematik, Universität Karlsruhe, 76128 Karlsruhe, Germany

1. INTRODUCTION

In this paper we give a survey of methods which can be used for including solutions of a nonlinear system of equations. These methods are called inclusion methods or enclosure methods. An inclusion method usually starts with an interval vector which contains a solution of a given system and improves this inclusion iteratively. The question which has to be discussed is under what conditions is the sequence of including interval vectors convergent to the solution. More often an including interval vector is not known and one tries to compute an interval vector containing a solution by some operator which forms the basis of an inclusion method. In other words, we prove the existence of a solution. Both concepts are discussed and illustrated in this article. An interesting feature of inclusion methods is that they can also be used to prove that there exists no solution in an interval vector.

Our methods are based on interval arithmetic tools. As is well known from the literature there exists a great variety of such inclusion methods (see [2] and [13], e.g.). Since the purpose of this article consists in discussing the main principles of inclusion methods we limit ourselves to only a few methods. Enclosure methods relying on other ideas are not considered.

The paper is organized as follows: In chapter 2.1 we repeat the well known results for the one–dimensional interval Newton method. Chapter 2.2 contains a series of properties of the Gaussian algorithm applied to linear systems with interval data. These results are used in chapter 2.3 where the interval Newton operator is introduced. In chapter 2.4 convergence and divergence statements for the so-called interval Newton method which is based on the interval Newton operator are investigated. Chapter 2.5 contains results about the speed of convergence and divergence, respectively. In chapter 3 the Krawczyk operator is introduced and in the final chapter 4 it is shown how test intervals can be efficiently constructed.

2. THE INTERVAL NEWTON METHOD

2.1. The One–Dimensional Case

In order to motivate the results of the later chapters we first consider a single equation in one unknown. Assume that

$f : [x] \subset D \subset \mathrm{R} \to \mathrm{R}$

and that the derivative of f has an interval arithmetic evaluation $f'([x]^0)$ for some $[x]^0 \subseteq [x]$. Assume that f has a zero x^* in $[x]^0$. If $m[x]^k$ is an arbitrary element contained in $[x]^k$ (usually one chooses $m[x]^k$ as the center of $[x]^k$) then the method

$$[x]^{k+1} = \{m[x]^k - \frac{f(m[x]^k)}{f'([x]^k)}\} \cap [x]^k \tag{1}$$

is well defined, provided $0 \notin f'([x]^0)$, and it holds that $x^* \in [x]^k$ and $\lim_{k\to\infty}[x]^k = x^*$. Moreover the order of convergence is (under some additional conditions) at least two. Method (1) is called the *interval Newton method.* For more details see [2], for example.

2.2. Preliminaries

For interval matrices $[A]$ and $[B]$ and a real vector c it holds in general that

$$[A]([B]c) \subseteq ([A][B])c\ . \tag{2}$$

This was proved in [17], p. 15. If c is equal to one of the unit–vectors the equality–sign holds. This is the content of the next lemma.

Lemma 1 *Let e^i be the i–th unit–vector. Then*

$$[A]([B]e^i) = ([A][B])e^i$$

for arbitrary interval matrices $[A]$ and $[B]$.

Proof. Denote the columns of the matrix $[B]$ by $[b]^i$, $1 \le i \le n$. Then it holds that

$$[B]e^i = ([b]^1, [b]^2, \ldots, [b]^n)e^i = [b]^i$$

and therefore

$$[A]([B]e^i) = [A][b]^i.$$

On the other hand we have

$$([A][B])e^i = ([A][b]^1, \ldots, [A][b]^n)e^i = [A][b]^i$$

and therefore the assertion follows. □

Assume now that we have given an n by n interval matrix $[A] = ([a]_{ij})$ and an interval vector $[b] = ([b]_i)$ with n components. By applying the formulas of the Gaussian algorithm we compute an interval vector $[x] = ([x]_i)$ for which the relation

$$\{x = A^{-1}b \mid A \in [A],\ b \in [b]\} \subseteq [x]$$

holds. See [2], Section 15 or [17], p. 20 ff, for example. If we set $[a]^1_{ij} := [a]_{ij}$, $1 \le i, j \le n$, and $[b]^1_i := [b]_i$, $1 \le i \le n$, then the formulas are as follows:

for $k = 1(1)(n-1)$ **do**
begin
  for $i = (k+1)(1)n$ **do**
  begin
    for $j = (k+1)(1)n$ **do**
    $[a]_{ij}^{k+1} := [a]_{ij}^{k} - [a]_{kj}^{k} \dfrac{[a]_{ik}^{k}}{[a]_{kk}^{k}}$
    $[b]_{i}^{k+1} := [b]_{i}^{k} - [b]_{k}^{k} \dfrac{[a]_{ik}^{k}}{[a]_{kk}^{k}}$
  end;
  for $l = 1(1)k$ **do**
  begin
    for $j = l(1)n$ **do**
    $[a]_{lj}^{k+1} := [a]_{lj}^{k}$
    $[b]_{l}^{k+1} := [b]_{l}^{k}$
  end;
end;
$[x]_n = [b]_n^n / [a]_{nn}^n$
for $i = (n-1)(-1)1$ **do**
$[x]_i = ([b]_i^n - \sum_{j=i+1}^{n} [a]_{ij}^n [x]_j)/[a]_{ii}^n.$

We have assumed that no division by an interval which contains zero occurs. In this case we say that the feasibility of the Gaussian algorithm is guaranteed. The feasibility is not dependent on the right hand side vector $[b]$.
If we define the interval matrices

$$[C]^k := \begin{bmatrix} 1 & & & & & \\ & \ddots & & & O & \\ & & 1 & & & \\ & & -\dfrac{[a]_{k+1,k}^k}{[a]_{kk}^k} & 1 & & \\ & O & \vdots & O & \ddots & \\ & & -\dfrac{[a]_{nk}^k}{[a]_{kk}^k} & & & 1 \end{bmatrix}, \ 1 \le k \le n-1,$$

$$[D]^k := \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & O & & \\ & & 1 & & & & \\ & & & \dfrac{1}{[a]_{kk}^n} & & & \\ & & & & 1 & & \\ & O & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}, \ 1 \le k \le n,$$

$$[T]^k := \begin{bmatrix} 1 & & & & & O \\ & \ddots & & & & \\ & & 1 & -[a]^n_{k,k+1} & \cdots & -[a]^n_{kn} \\ & & & 1 & O & \\ & O & & & \ddots & \\ & & & & & 1 \end{bmatrix}, \ 1 \le k \le n-1,$$

then we have for the interval vector computed by the Gaussian algorithm

$$[x] = [D]^1([T]^1([D]^2([T]^2(\ldots([D]^{n-1}([T]^{n-1}([D]^n[\tilde{b}])\ldots) \tag{3}$$

where

$$[\tilde{b}] = [C]^{n-1}([C]^{n-2}(\ldots([C]^2([C]^1[b])\ldots).$$

The interval vector $[x]$ computed in this manner is usually also denoted by $[x] = \mathrm{IGA}([A],[b])$ ('Interval Gauss algorithm of $[A]$ and $[b]$').

The interval matrix $\mathrm{IGA}([A])$ is *defined* as the product of the interval matrices occuring in the definition of $\mathrm{IGA}([A],[b])$:

$$\mathrm{IGA}([A]) = [D]^1([T]^1(\ldots([D]^{n-1}([T]^{n-1}([D]^n([C]^{n-1}(\ldots([C]^1)\ldots).$$

Note that it is not possible to omit the paranthesis in the expressions for $\mathrm{IGA}([A],[b])$ and $\mathrm{IGA}([A])$, respectively. $\mathrm{IGA}([A])$ has the following property.

Lemma 2 *Let e^i denote the i-th unit vector. Then $IGA([A],e^i) = IGA([A]) \cdot e^i$, $1 \le i \le n$.*

Proof. Starting with the representation of $\mathrm{IGA}([A],e^i)$ and applying repeatedly Lemma 1 we get the assertion. □

The last lemma states that the i-th column of the matrix $\mathrm{IGA}([A])$ is equal to the interval vector which is obtained if one applies the Gaussian algorithm to the interval matrix $[A]$ and the right hand side e^i. In order to compute $\mathrm{IGA}([A])$ it is therefore not necessary to know the matrices appearing on the right hand side in the definition of $\mathrm{IGA}([A])$ explicitly. $\mathrm{IGA}([A])$ can be computed by 'formally inverting ' the interval matrix $[A]$ by applying the Gaussian algorithm.

Finally we need the following result.

Lemma 3 *For an interval matrix $[A]$ and a point vector b it always holds that*

$$IGA([A],b) \subseteq IGA([A]) \cdot b.$$

Proof. Apply (2) repeatedly to the representation (3) of $\mathrm{IGA}([A],b)$. □

2.3. The n–Dimensional Case

Let there be given a mapping

$$f : [x] \subset D \subseteq \mathbf{R}^n \to \mathbf{R}^n$$

and assume that the partial derivatives of f exist in D and are continuous. If $y \in [x]$ is a fixed chosen point then

$$f(x) - f(y) = J(x)(x - y), \; x \in D \tag{4}$$

where the matrix $J(x)$ is defined by

$$J(x) = \int_0^1 f'(y + t(x - y))\, dt. \tag{5}$$

(See [14], p. 71, (10)). Note that J is a continuous mapping of x for fixed y. Since $t \in [0, 1]$ we have that $y + t(x - y) \in [x]$ and therefore that

$$J(x) \in f'([x]) \tag{6}$$

where $f'([x])$ denotes an interval arithmetic evaluation of the Jacobian of f, provided there exists one.

Theorem 1 *Let $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ be a continuously differentiable mapping and assume that an interval arithmetic evaluation $f'([x])$ of the Jacobian exists for some interval vector $[x] \subset D$. Suppose that the Gaussian algorithm is feasible for $f'([x])$ and take for the right-hand side $f(y)$ where $y \in [x]$ is fixed. Then it holds:*

1) If f has a (necessarily unique) zero x^* in $[x]$ then

$$x^* \in y - IGA(f'([x]), f(y)).$$

2) If

$$(y - IGA(f'([x]), f(y))) \cap [x] = \emptyset \tag{7}$$

then f has no zero in $[x]$.

3) If

$$y - IGA(f'([x]), f(y)) \subseteq [x] \tag{8}$$

then f has a unique zero x^* in $[x]$.

Proof. 1) Assume that x^* and x^{**} are two zeros of f in $[x]$. Then by (4)

$$0 = f(x^*) - f(x^{**}) = J(x^*)(x^* - x^{**})$$

where $J(x^*)$ is defined analogously to (5). Since IGA$(f'([x]))$ exists it follows that $f'([x])$ contains no singular matrix and since $J(x^*) \in f'([x])$ it follows that $J(x^*)$ is nonsingular. Therefore $x^* = x^{**}$. By (4) again we have that

$$f(x^*) - f(y) = -f(y) = J(x^*)(x^* - y).$$

Since $J(x^*) \in f'([x])$ it is nonsingular and therefore it follows that

$$\begin{aligned} x^* &= y - J(x^*)^{-1} f(y) \\ &\in y - \mathrm{IGA}(f'([x]), f(y)) \end{aligned}$$

since

$$J(x^*)^{-1} f(y) \in \mathrm{IGA}(f'([x]), f(y))$$

by the inclusion monotonicity of interval arithmetic.
2) The proof follows because of the preceding part.
3) Since the Gaussian algorithm can be carried out on the matrix $f'([x])$ it follows that all the point matrices from $f'([x])$ are nonsingular. In particular $J(x)$ is nonsingular. We consider the mapping

$$p : [x] \subset D \subseteq \mathbf{R}^n \to \mathbf{R}^n$$

where

$$p(x) = x - J(x)^{-1} f(x)$$

and where for some fixed $y \in [x]$ the matrix $J(x)$ is defined by (4) and (5).
It follows

$$\begin{aligned} p(x) &= x - J(x)^{-1} f(y) + J(x)^{-1}(f(y) - f(x)) \\ &= y - J(x)^{-1} f(y) \\ &\in y - \mathrm{IGA}(f'([x]), f(y)) \subseteq [x]. \end{aligned}$$

Hence the continuous mapping p maps the nonempty convex and compact set $[x]$ into itself. Therefore, by the Brouwer fixed point theorem it has a fixed point x^* in $[x]$ from which it follows that f has a solution x^* in $[x]$. By the same ideas as in part 1) it follows that x^* is unique. □

The preceding theorem has a series of implications. If one has an interval vector for which (7) holds then $[x]$ is an *exclusion set* for the zeros of f. If on the other hand (8) holds then $[x]$ and because of 1) also $y - \mathrm{IGA}(f'([x]), f(y)) \subseteq [x]$ is an *inclusion set* for a zero of f and $[x] \backslash (y - \mathrm{IGA}(f'([x]), f(y)))$ is an *exclusion set* for the zeros of f.

We now consider more generally the case that we have given an interval vector $[x]$ which contains a zero x^* of the mapping

$$f : [x] \subset D \subseteq \mathbf{R}^n \to \mathbf{R}^n.$$

f is assumed to be continuously differentiable in D. Furthermore we assume that the interval arithmetic evaluation $f'([x])$ of the Jacobian $f'(x)$ exists and that the Gaussian algorithm is feasible for $f'([x])$. In the preceding theorem it was shown that x^* is unique in $[x]$ under these conditions. Assume now that $m[x]$ is an arbitrary chosen real vector contained in $[x]$. Then the so-called *interval Newton operator* $N([x])$ is defined by

$$N([x]) = m[x] - \mathrm{IGA}(f'([x]), f(m[x]))$$

and the *interval Newton method* (IN) is defined by setting $[x]^0 := [x]$ and

$$\left.\begin{cases} m[x]^k & \in & [x]^k \\ N([x]^k) & = & m[x]^k - \mathrm{IGA}(f'([x]^k), f(m[x]^k)) \\ [x]^{k+1} & = & N([x]^k) \cap [x]^k \\ & k = 0,1,2,\dots \,. & \end{cases}\right\} \text{(IN)}$$

Usually one chooses $m[x]^k$ as the midpoint (center) of $[x]^k$ if there is no specific information about the location of x^* in $[x]^k$. However we don't assume this choice.

Since $x^* \in [x]^0$ we have by 1) of the preceding theorem that $x^* \in N([x]^0)$ and therefore that $x^* \in [x]^1$. Since $[x]^1 \subseteq [x]^0$ it follows that $f'([x]^1)$ exists and $f'([x]^1) \subseteq f'([x]^0)$. By the inclusion monotonicity of interval arithmetic we conclude that the Gaussian algorithm for $f'([x]^1)$ is feasible and therefore $N([x]^1)$ and hence $[x]^2$ is well defined. By mathematical induction we can proof that (IN) computes a sequence $\{[x]^k\}_{k=0}^{\infty}$ of interval vectors with

$$x^* \in [x]^k, \; k \geq 0$$

and

$$[x]^0 \supseteq [x]^1 \supseteq \dots \supseteq [x]^k \supseteq [x]^{k+1} \supseteq \dots \,.$$

From these two properties we can conclude that

$$\lim_{k\to\infty} [x]^k = [x]^*$$

and

$$x^* \in [x]^*.$$

The following example, introduced by H. Schwandt in [17] shows that in contrast to the one–dimensional case repeated in chapter 2.1 it does not always hold that $w([x]^*) = 0$. In other words: The sequence $\{[x]^k\}$ may have a limit which is a proper interval.

Example 1 *Let* $x = \begin{pmatrix} u \\ v \end{pmatrix}$ *and* $f(x) = \begin{pmatrix} -u^2 + v^2 - 1 \\ u^2 - v \end{pmatrix}$.

The vector $x^* = \begin{pmatrix} \sqrt{\frac{1+\sqrt{5}}{2}} \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}$ *is the unique solution of the system* $f(x) = 0$ *in the interval vector*

$$[x]^0 = \frac{1}{10} \begin{pmatrix} [11, 19] \\ [11, 19] \end{pmatrix}.$$

Choosing $m[x]^0$ *as the center of* $[x]^0$ *we get*

$$N([x]^0) = \begin{pmatrix} [-\frac{3}{88}, \frac{90771}{12584}] \\ \\ [\frac{7}{8}, \frac{5801}{1144}] \end{pmatrix}$$

and therefore $[x]^1 = [x]^0$. *If we choose* $m[x]^k$ *to be the center of* $[x]^k$ *for all* k *then we have*

$$[x]^0 = [x]^1 = \ldots = [x]^k = [x]^{k+1} = \ldots$$

and

$$\lim_{k\to\infty} [x]^k = [x]^0 \neq x^*. \qquad \square$$

2.4. Convergence and Divergence Statements

In the following theorem we state and prove some convergence and divergence results for the interval Newton method (IN). The definition of these concepts will become clear from the formulation of the next theorem.

Theorem 2 *Let there be given an interval vector* $[x]^0$ *and a continuously differentiable mapping*

$$f : [x]^0 \subset D \subseteq \mathbf{R}^n \to \mathbf{R}^n$$

and assume that an interval arithmetic evaluation $f'([x]^0)$ *of the Jacobian exists. Assume that* $IGA(f'([x]^0))$ *exists (which is identical to assuming that the Gaussian algorithm is feasible for* $f'([x]^0)$ *and an arbitrary right hand side. See the remarks following Lemma 2).*
(a) Suppose that $\rho(A) < 1$ *(* ρ *denotes the spectral radius of a real matrix) where*

$$A = |I - IGA(f'([x]^0)) \cdot f'([x]^0)|, \tag{9}$$

or that $\rho(B) < 1$ *where*

$$B = w(IGA(f'([x]^0))) \cdot |f'([x]^0)|. \tag{10}$$

If f *has a (necessarily unique) zero* x^* *in* $[x]^0$ *then the sequence* $\{[x]^k\}_{k=0}^{\infty}$ *computed by (IN) is well defined and it holds that* $x^* \in [x]^k$ *and* $\lim_{k\to\infty}[x]^k = x^*$.
If one chooses $m[x]^k$ *as the center of* $[x]^k$ *then the condition* $\rho(B) < 1$ *can be replaced by* $\rho(B) < 2$.

(b) If $\rho(A) < 1$ *or* $\rho(B) < 1$ *(where* A *and* B *are defined by (9) and(10), respectively) and if* f *has no zero in* $[x]^0$ *then there is a* $k_0 \geq 0$ *such that* $N([x]^{k_0}) \cap [x]^{k_0} = \emptyset$ *(empty set), that is (IN) will break down after a finite number of steps because of empty intersection.*

Proof. (a) From the discussion in Chapter 2.3 we know that (IN) is well defined if IGA$(f'([x]^0))$ exists and if f has a zero x^* in $[x]^0$. We now assume that $\rho(A) < 1$ holds. Note that analogously to (4), (5) and (6) we have

$$f(m[x]^k) = f(m[x]^k) - f(x^*) = J(m[x]^k)(m[x]^k - x^*) \tag{11}$$

and

$$J(m[x]^k) \in f'([x]^k). \tag{12}$$

Besides of this we use (2), Lemma 3, and the fact that for a real vector c and interval matrices $[X]$ and $[Y]$ the equation $([X]+[Y])c = [X]c + [Y]c$ is valid. Then we get

$$\begin{aligned}
N([x]^k) - x^* &= m[x]^k - x^* - \mathrm{IGA}(f'([x]^k), f(m[x]^k)) \\
&\subseteq m[x]^k - x^* - \mathrm{IGA}(f'([x]^k)) \cdot f(m[x]^k) \\
&= m[x]^k - x^* - \mathrm{IGA}(f'([x]^k)) \cdot \{J(m[x]^k)(m[x]^k - x^*)\} \\
&\subseteq m[x]^k - x^* - \mathrm{IGA}(f'([x]^k))\{f'([x]^k)(m[x]^k - x^*)\} \\
&\subseteq m[x]^k - x^* - \{\mathrm{IGA}(f'([x]^k)) \cdot f'([x]^k)\}(m[x]^k - x^*) \\
&= (I - \mathrm{IGA}(f'([x]^k)) \cdot f'([x]^k))(m[x]^k - x^*) .
\end{aligned}$$

Since $m[x]^k \in [x]^k$ and $|I - \mathrm{IGA}(f'([x]^k)) \cdot f'([x]^k)| \leq A$ it follows that

$$|N([x]^k) - x^*| \leq A|[x]^k - x^*|$$

or, equivalently, that

$$q(N([x]^k), x^*) \leq A\ q([x]^k, x^*).$$

Since $x^* \in N([x]^k) \cap [x]^k = [x]^{k+1} \subseteq N([x]^k)$ it also holds that

$$q([x]^{k+1}, x^*) \leq q(N([x]^k), x^*) \leq A\ q([x]^k, x^*)$$

and therefore

$$q([x]^{k+1}, x^*) \leq q(N([x]^k), x^*) \leq A^{k+1} q([x]^0, x^*).$$

Because of $\rho(A) < 1$ the assertion $\lim_{k\to\infty}[x]^k = x^*$ holds.
We now proof (a) under the assumption $\rho(B) < 1$. Applying Lemma 3 we get for $k \geq 0$

$$\begin{aligned}
N([x]^k) &= m[x]^k - \mathrm{IGA}(f'([x]^k), f(m[x]^k)) \\
&\subseteq m[x]^k - \mathrm{IGA}(f'([x]^k)) \cdot f(m[x]^k)
\end{aligned}$$

and for the width of $N([x]^k)$

$$\begin{aligned}
w(N([x]^k)) &\leq w(\mathrm{IGA}(f'([x]^k))) \cdot |f(m[x]^k)| \\
&\leq w(\mathrm{IGA}(f'([x]^0))) \cdot |J(m[x]^k)(m[x]^k - x^*)| \\
&\leq w(\mathrm{IGA}(f'([x]^0))) \cdot |f'([x]^0)| \cdot w([x]^k),
\end{aligned}$$

where we have used (4) and (5), the fact that $[x]^k \subseteq [x]^0$ and the inequality $|m[x]^k - x^*| \leq w([x]^k)$.
Since $[x]^{k+1} = N([x]^k) \cap [x]^k$ we have $w([x]^{k+1}) \leq w(N[x]^k)$ and therefore

$$w([x]^{k+1}) \leq B\ w([x]^k)$$

and

$$w([x]^{k+1}) \leq B^{k+1} w([x]^0),$$

from which the assertion $\lim_{k\to\infty}[x]^k = x^*$ follows. If $m[x]^k$ is the center of $[x]^k$ then $|m[x]^k - x^*| \leq \frac{1}{2}w[x]^k$ and the proof can be completed also in the case $\rho(B) < 2$.

(b) We now prove (b) under the assumption $\rho(A) < 1$. We assume that for all $k \geq 0$ the intersection $N([x]^k) \cap [x]^k$ is not empty. Then (IN) is well defined and it holds that

$$[x]^0 \supseteq [x]^1 \supseteq \ldots \supseteq [x]^k \supseteq [x]^{k+1} \supseteq \ldots$$

from which it follows that the sequence is converging to an interval vector $[x]^*$. We now consider the sequence $\{m[x]^k\}_{k=0}^{\infty}$. This sequence is contained in the compact set $[x]^0$. By the Bolzano–Weierstrass theorem we conclude that there exists a convergent sub-sequence $\{m[x]^{k_i}\}_{i=0}^{\infty}$. Suppose that $\lim_{i\to\infty} m[x]^{k_i} = x^*$. Since $m[x]^{k_i} \in [x]^{k_i}$ it also holds that $x^* \in [x]^*$. Using the continuity of the functions and operations involved in the method (IN) we get from

$$\begin{aligned} N([x]^{k_i}) &= m([x]^{k_i}) - \mathrm{IGA}(f'([x]^{k_i}), f(m[x]^{k_i})) \\ [x]^{k_i+1} &= N([x]^{k_i}) \cap [x]^{k_i} \end{aligned}$$

the pair of equations

$$\begin{aligned} [u]^* &= x^* - \mathrm{IGA}(f'([x]^*), f(x^*)) \\ [x]^* &= [u]^* \cap [x]^* \end{aligned}$$

where $[u]^* = \lim_{i\to\infty} N([x]^{k_i}) = N([x]^*)$.
From the second equation it follows that $[x]^* \subseteq [u]^*$ and therefore that $x^* \in [u]^*$. Since $x^* \in [x]^*$. Therefore we get from the first equation

$$x^* \in x^* - \mathrm{IGA}(f'([x]^*), f(x^*))$$

or

$$0 \in \mathrm{IGA}(f'([x]^*), f(x^*))$$

and by Lemma 3 and the inclusion monotonicity

$$\begin{aligned} 0 &\in \mathrm{IGA}(f'([x]^*)) \cdot f(x^*) \\ &\subseteq \mathrm{IGA}(f'([x]^0) \cdot f(x^*). \end{aligned}$$

Since for an interval matrix $[X]$ and a real vector c we have $[X]c = \{Xc | X \in [X]\}$ it follows that there exists a matrix $X \in \mathrm{IGA}(f'([x]^0)$ such that $Xf(x^*) = 0$. If X is nonsingular then we have the contradiction $f(x^*) = 0$. The nonsingularity of X can be seen in the following manner: If $Y \in f'([x]^0)$ then

$$|I - XY| \leq |I - \mathrm{IGA}(f'([x]^0)) \cdot f'([x]^0)| = A.$$

By the Perron–Frobenius theory it follows that

$$\rho(I - XY) \leq \rho(A) < 1.$$

Therefore

$$(I - (I - XY))^{-1} = (XY)^{-1}$$

exists, that is X is nonsingular.
In order to complete the proof, we have to show that under the assumption $\rho(B) < 1$

there is no singular matrix contained in $\mathrm{IGA}(f'([x]^0))$. This can be seen as follows (see [9], where also relations between $\rho(A)$ and $\rho(B)$ are discussed): Let $Y \in f'([x]^0)$ and $X \in \mathrm{IGA}(f'([x]^0))$. Since $\mathrm{IGA}(f'([x]^0))$ exists it follows that Y is nonsingular.
Therefore

$$\begin{aligned} |I - XY| &\leq |Y^{-1} - X||Y| \\ &\leq w(\mathrm{IGA}(f'([x]^0))) \cdot |f'([x]^0)| \\ &= B. \end{aligned}$$

Again we have $\rho(I - XY) \leq \rho(B) < 1$ by the Perron–Frobenius theory and hence the nonsingularity of X. This completes the proof. □

How easy or difficult are the conditions $\rho(A) < 1$ and/or $\rho(B) < 1$ to fulfill? The answer is that these conditions always hold if the components of the width $w([x]^0)$ of $[x]^0$ are all sufficiently small. In this case the product $\mathrm{IGA}(f'([x]^0)) \cdot f'([x]^0)$ differs only slightly from the identity matrix and A is therefore a small nonnegative matrix. Analogously is $\mathrm{IGA}(f'([x]^0))$ close to a real matrix if $w([x]^0)$ is small, hence $w(\mathrm{IGA}(f'([x]^0)))$ is a small nonnegative matrix and therefore also B.

2.5. Speed of Convergence and Divergence

In the next theorems we will present results concerning the speed of convergence and divergence, respectively.

Theorem 3 *Let there be given an interval vector $[x]^0$ and a continuously differentiable mapping*

$$f : [x]^0 \subset D \subseteq \mathbf{R}^n \to \mathbf{R}^n$$

and assume that an interval arithmetic evaluation $f'([x]^0)$ of the Jacobian exists. Furthermore assume that for all elements of the Jacobian and some norm an inequality of the form

$$w(f'([x])_{ij}) \leq c\|w([x])\|, \ c \geq 0, \tag{13}$$

holds for all $[x] \subseteq [x]^0$. If $[x]^0$ contains a zero x^* of f and if (IN) is convergent to x^* then

$$\|w([x]^{k+1})\| \leq \gamma\|w([x]^k)\|^2, \ \gamma \geq 0,$$

that is the sequence of widths is quadratically convergent to zero (and therefore also the sequence of distances $q([x]^k, x^*)$ between $[x]^k$ and x^*. See [2], Appendix A).

Proof. As in [2], Chapter 19, Lemma 6, it can be shown that (13) implies for some norm

$$\|w(\mathrm{IGA}(f'([x])))\| \leq \kappa\|w([x])\|, \ \kappa \geq 0, \tag{14}$$

for all $[x] \subseteq [x]^0$. Because of (4), (5) and (6) we have

$$\begin{aligned} w([x]^{k+1}) &\leq w(N([x]^k)) \\ &= w(m[x]^k - \mathrm{IGA}(f'([x]^k), f(m[x]^k))) \\ &\leq w(\mathrm{IGA}(f'([x]^k))) \cdot |f(m[x]^k)| \\ &= w(\mathrm{IGA}(f'([x]^k))) \cdot |J(m[x]^k)(m[x]^k - x^*)| \\ &\leq w(\mathrm{IGA}(f'([x]^k))) \cdot |f'([x]^0)| \cdot w([x]^k). \end{aligned}$$

Using a monotone vector norm and the equivalence of all vector norms and matrix norms, respectively, we get

$$\|w([x])^{k+1}\| \leq \gamma \|w([x]^k)\|^2$$

where

$$\gamma = \delta \cdot \kappa \cdot \| \, |f'([x]^0)| \, \|$$

and where δ is a consequence of the equivalence of all norms. □

We remark that in practice the condition (13) is not very strong. It holds, for example, for all functions with continuous partial derivatives of second order.

Quadratic convergence behaviour is a property which an iterative method usually only exhibits asymptotically. Therefore the following behaviour is of interest.

Lemma 4 *Suppose that the assumptions of Theorem 2 concerning f and $f'([x]^0)$ are true, that f has a zero x^* in $[x]^0$ and that $\rho(A) < 1$. Then $m[x]^k \notin [x]^{k+1}$ if $m[x]^k \neq x^*$.*

Proof. Suppose that $m[x]^k \in N([x]^k)$. Using the relation

$$N([x]^k) - x^* \subseteq (I - \mathrm{IGA}(f'([x]^k)) \cdot f'([x]^k))(m[x]^k - x^*)$$

which was derived in part (a) of the proof of Theorem 2 we get, because of

$$m[x]^k - x^* \in N([x]^k) - x^*,$$

the inequalities

$$\begin{aligned} |m[x]^k - x^*| &\leq |I - \mathrm{IGA}(f'([x]^k)) \cdot f'([x]^k)| \cdot |m[x]^k - x^*| \\ &\leq A|m[x]^k - x^*|. \end{aligned}$$

Since $\rho(A) < 1$ we get the contradiction $m[x]^k = x^*$. □

From Lemma 4 it follows that if one chooses $m[x]^k$ as the center of $[x]^k$ and if $m[x]^k \neq x^*$ then at least one of the components of $[x]^{k+1}$ has its width smaller than half of the width of the corresponding component of $[x]^k$. The next lemma shows that a similar result holds if there is no solution in $[x]^0$.

Lemma 5 *Suppose that the assumptions of Theorem 2 concerning f and $f'([x]^0)$ are true, that f has* no *zero x^* in $[x]^0$ and that $\rho(A) < 1$ or $\rho(B) < 1$. Then $m[x]^k \notin [x]^{k+1}$ (provided $[x]^{k+1}$ is defined at all).*

Proof. Assume that $m[x]^k \in [x]^{k+1}$. Then $m[x]^k \in N[x]^k$ and therefore

$$0 \in \mathrm{IGA}(f'[x]^k) \cdot f(m[x]^k).$$

Similarly as in the proof of (b) in Theorem 2 this leads to the contradiction that $m[x]^k$ is a zero of f. □

From the preceding Lemma 5 it is not clear how many steps it will take until the intersection in (IN) becomes empty if there is no zero of f in $[x]^0$. We now will show that only a few steps are necessary if the width $w([x]^0)$ is small enough. Because of Lemma 3 we have for any $[x] \subseteq [x]^0$

$$\begin{aligned} N([x]) &= m[x] - \mathrm{IGA}(f'([x]), f(m[x])) \\ &\subseteq m[x] - \mathrm{IGA}(f'([x])) \cdot f(m[x]) =: \tilde{N}([x]). \end{aligned}$$

We denote the vector of lower bounds and upper bounds of the components of $N([x])$ by n^1 and n^2, respectively. Analogously we define $\tilde{n}^1$ and $\tilde{n}^2$. Then $N([x]) \subseteq \tilde{N}([x])$ is equivalent to

$$\tilde{n}^1 \leq n^1 \leq n^2 \leq \tilde{n}^2 \tag{15}$$

where the partial ordering is defined componentwise. As in the proof of Theorem 2 we use the fact that for an interval matrix $[X]$ and a real vector c we have $[X]c = \{Xc | X \in [X]\}$. Therefore from the definition of $\tilde{N}([x])$ we obtain

$$\left\{ \begin{aligned} \tilde{n}^1 &= m[x] - A^2 f(m[x]) \\ \tilde{n}^2 &= m[x] - A^1 f(m[x]) \end{aligned} \right. \tag{16}$$

where A^1 and A^2 are real matrices contained in $\mathrm{IGA}(f'([x]))$. If the real vectors x^1 and x^2 are defined analogously to n^1 and n^2, respectively, then we get by using (15) and (16)

$$\begin{aligned} x^2 - n^1 &\leq x^2 - \tilde{n}^1 \\ &= x^2 - m[x] + A^2 f(m[x]) \\ &= x^2 - m[x] + A^2\{f(x^2) + J(m[x])(m[x] - x^2)\} \\ &= (I - A^2 J(m[x]))(x^2 - m[x]) + A^2 f(x^2). \end{aligned}$$

We now assume that (13) from Theorem 3 and therefore also (14) holds. $J(m[x])$ is nonsingular since $J(m[x]) \in f'([x])$. Therefore we get

$$\begin{aligned} |(I - A^2 J(m[x]))(x^2 - m[x])| &= |(J(m[x])^{-1} - A^2)\ J(m[x])(x^2 - m[x])| \\ &\leq w(\mathrm{IGA}(f'([x]))) \cdot |f'([x]^0)| \cdot w([x]) \\ &= O(\|w([x])\|^2). \end{aligned}$$

The right hand side denotes a vector whose components are all of the order $O(\|w([x])\|^2)$. The preceding inequality can therefore be written as

$$x^2 - n^1 \leq O(\|w([x])\|^2) + A^2 f(x^2). \tag{17}$$

By similar considerations we get

$$n^2 - x^1 \leq O(\|w([x])\|^2) - A^1 f(x^1). \tag{18}$$

We now show that for sufficiently small width $w([x])$ the matrices A^1 and A^2 are nonsingular: Since $\mathrm{IGA}(f'([x]))$ exists it follows that all $Y \in f'([x])$ are nonsingular. For an arbitrary $X \in \mathrm{IGA}(f'([x]))$ we have

$$|I - XY| \leq |Y^{-1} - X| \cdot |Y| \leq w(\mathrm{IGA}(f'([x]))) \cdot |f'([x])|.$$

Since $w(\mathrm{IGA}(f'([x]))) \to 0$ for $w([x]) \to 0$ it follows that the matrix on the right hand side has spectral radius less than one for sufficiently small $w([x])$. From this it follows $\rho(I - XY) < 1$ by the Perron-Frobenius theory and therefore the nonsingularity of X and hence of A^1 and A^2 for sufficiently small $w([x])$.

Lemma 6 *Assume that under our assumptions $f(x) \neq 0$ for $x \in [x]$. Then, if $w([x])$ is sufficiently small, there is at least one $i_0 \in \{1, 2, \ldots, n\}$ such that*

$$(A^1 f(x^1))_{i_0} \neq 0 \tag{α}$$

and

$$sign(A^1 f(x^1))_{i_0} = sign(A^2 f(x^2))_{i_0} \ . \tag{β}$$

Proof. Since $f(x^1) \neq 0$ and because A^1 is nonsingular for sufficiently small $w([x])$ we have (α). Since $A^1,\ A^2 \in \mathrm{IGA}(f'([x]))$ it follows

$$|A^2 - A^1| \leq w(\mathrm{IGA}(f'([x])))$$

and therefore

$$A^2 = A^1 + T$$

where $\|T\| = O(\|w([x])\|)$. Hence we have

$$A^2 f(x^2) = A^1 f(x^2) + T f(x^2)$$

with $T f(x^2) \to 0$ for $w([x]) \to 0$. By (4) and (5) we have

$$f(x^2) = f(x^1) + J(x^2)(x^2 - x^1).$$

Multiplying this by A^1 and inserting the resulting equation in the preceding one we get

$$A^2 f(x^2) = A^1 f(x^1) + A^1 J(x^2)(x^2 - x^1) + T f(x^2)$$

where $A^1 J(x^2)(x^2 - x^1) \to 0$ and $T f(x^2) \to 0$ for $w([x]) \to 0$. For sufficiently small $w([x])$ there is by (α) an index i_0 such that $(A^1 f(x^2))_{i_0} \neq 0$ and therefore

$$\mathrm{sign}(A^2 f(x^2))_{i_0} = \mathrm{sign}(A^1 f(x^1))_{i_0}$$

which is (β). □

Assume now that $\mathrm{sign}(A^1 f(x^1))_{i_0} = 1$. Then by (18)

$$(n^2 - x^1)_{i_0} \leq (O(\|w([x])\|^2) - A^1 f(x^1))_{i_0} < 0 \tag{19}$$

for sufficiently small $w([x])$. If $\mathrm{sign}(A^1 f(x^1))_{i_0} = -1$ then by the preceding lemma $\mathrm{sign}(A^2 f(x^2))_{i_0} = -1$ and therefore by (17)

$$(x^2 - n^1)_{i_0} \leq (O(\|w([x])\|^2) + A^2 f(x^2))_{i_0} < 0 \tag{20}$$

for sufficiently small $w([x])$.

Now the next lemma can easily be shown.

Lemma 7 *We have $N([x]) \cap [x] = \emptyset$ if for (at least) one $i_0 \in \{1, 2, \ldots, n\}$*

$$(\frac{1}{2} w([x]) + \frac{1}{2} w(N([x])))_{i_0} < |(m[x] - m(N([x])))_{i_0}|.$$

This is eqivalent to

$$(n^2 - x^1)_{i_0} < 0 \quad or \quad (x^2 - n^1)_{i_0} < 0$$

for at least one $i_0 \in \{1, 2, \ldots, n\}$. □

The inequalities (19) and(20) show that because of the term $O(\|w([x])\|^2)$ on the right hand side the intersection will become empty as soon as the width $w([x])$ is small enough. Because of the term $O(\|w[x]\|^2)$ we can speak of *quadratic divergence behaviour*.

3. MODIFICATIONS OF THE INTERVAL NEWTON METHOD

In our assumptions on the interval Newton method we always had to assume that the Gaussian algorithm is feasible with $f'([x]^0)$ (and an arbitrary right hand side). If $f'([x]^0)$ contains no singular matrix this is by continuity arguments always the case if the width of the elements of $f'([x]^0)$ is small enough. There also exist a series of sufficient criteria for the feasibility of the Gaussian algorithm. See [10], for example. On the other hand there exist simple examples of interval matrices which contain no singular matrix, nevertheless the Gaussian algorithm will break down because of division by an interval which contains zero. This is the case even if one takes into account all possible pivoting strategies. R. Krawczyk had the idea to avoid this problem by introducing what today is called the Krawczyk–operator. See [8].

Assume again that $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ is a continuously differentiable mapping and assume that an interval arithmetic evaluation $f'([x])$ exists for some $[x] \subset D$. Furthermore let C be a real nonsingular matrix and $x \in [x]$. Then the *Krawczyk operator* is defined as

$$K([x], x, C) = x - Cf(x) + (I - Cf'([x]))([x] - x).$$

$K([x], x, C)$ is again an interval vector. A complete analogue to Theorem 1 could be formulated. We limit ourselves to the following result.

Theorem 4 *If $K([x], x, C) \subseteq [x]$ then f has a zero x^* in $K([x], x, C)$ (and therefore also in $[x]$).* □

The proof is based on the Brouwer fixed point theorem and can be found in [2], Theorem 10 in Chapter 13 or in [12]. We omit the details.

Analogously as it was done with $N([x])$ we can also construct an iteration method using the Krawczyk operator. There exist a series of possibilities. We limit ourselves to one special case: Let $[x]^0 \subseteq [x]$ be a given interval vector.

We set

$$\left.\begin{array}{rcl} x^k & = & m[x]^k \in [x]^k \\ C^k & = & (m(f'([x]^k)))^{-1} \\ K([x]^k, x^k, C^k) & = & x^k - C^k f(x^k) + (I - C^k f'([x]^k))([x]^k - x^k) \\ [x]^{k+1} & = & K([x]^k, x^k, C^k) \cap [x]^k \\ k & = & 0, 1, 2, \ldots \end{array}\right\} \quad (K)$$

(K) is called the *Krawczyk method.*

Usually one chooses for $m[x]^k$ and $m(f'([x]^k))$ the center of $[x]^k$ and $f'([x]^k)$, respectively, but this is not a must. In contrast to (IN) where we have to perform the Gaussian algorithm with an interval matrix we have to invert a real matrix which always can be done if it is nonsingular. For (K) similar results as in Theorem 2 can be formulated and proved. In [3] a Krawczyk–like operator has been considered where only triangular factorizations (and no inversions) of matrices are performed.

Finally we mention that the so-called *Hansen–Sengupta operator* which is a nonlinear version of interval Gauss–Seidel iteration is occasionally prefered in practice (see [13], pp. 177 ff).

4. THE EFFICIENT CONSTRUCTION OF TEST INTERVALS

Both the interval Newton operator $N([x])$ and the Krawczyk operator $K([x], x, C)$ can be used for proving the existence of a zero of a given mapping f in an interval vector. This has been stated in part 3) of Theorem 1 for $N([x])$ and in Theorem 4 for $K([x], x, C)$. The main problem is how one can find an interval vector $[x]$ for which $N([x])$ or $K([x], x, C)$ is contained in $[x]$.

In this chapter we propose the use of the Kantorovich theorem in order to efficiently produce a good test interval that presumably contains a solution. Namely we proceed with Newton's method performed on a computer in normal floating point arithmetic. For a given *eps* equal to the machine precision we devise a stopping criterion and construct a test interval $[x]$ such that $K([x], x, C) \subset [x]$ is very likely to be satisfied. Moreover our method is designed in such a way that the condition

$$\frac{\|y - x^*\|_\infty}{\|x^*\|_\infty} \le \ eps$$

is eventually also satisfied. Here y denotes any point of the interval $K([x], x, C)$. Besides of having an elegant theoretical justification, the resulting algorithm turns out to be very efficient in practice. It gives highly accurate results and in the same time provides a tool for establishing the existence of solutions of certain equations.

We start with the following well known result concerning the Krawczyk operator.

Lemma 8 *Assume that the mapping $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ is continuously differentiable and the Jacobian has an interval arithmetic evaluation $f'([x])$ for all $[x] \subset D$ such that*

$$\|w(f'([x]))\|_\infty \le \hat{L}\|w([x])\|_\infty, \ [x] \subset D, \tag{21}$$

for some $\hat{L} \ge 0$. If $C^{-1} \in f'([x])$ then the following inequality

$$\|w(K([x], x, C))\|_\infty \le \gamma \|w([x])\|_\infty^2 \tag{22}$$

holds with $\gamma = \|C\|_\infty \hat{L}$.

Proof. For the width of $K([x], x, C)$ we get

$$\begin{aligned} w(K([x], x, C)) &= w(x - Cf(x) + (I - Cf'([x]))([x] - x)) \\ &\le w(I - Cf'([x])) \cdot w([x] - x) \\ &= w(C(C^{-1} - f'([x]))) \cdot w[x] \\ &= |C| \cdot w(f'([x])) \cdot w([x]). \end{aligned}$$

Using (21) we obtain (22). □

Consider now Newton's method

$$x^{k+1} = x^k - f'(x^k)^{-1} f(x^k), \ k = 0, 1, 2, \dots, \tag{23}$$

applied to a mapping $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$. The Newton–Kantorovich theorem gives sufficient conditions for the convergence of Newton's method starting at x^0. Furthermore it contains an error estimation. A simple discussion of this estimate in conjunction with Lemma 8 will lead us to a test interval which can be computed using only iterates of Newton's method.

Theorem 5 *(Newton–Kantorovich. See [14], Theorem 12.6.2)*
Assume that $f : D \subseteq \mathbf{R}^n \to \mathbf{R}^n$ is differentiable in the ball $\{x \| \|x - x^0\| \leq r\}$ and that

$$\|f'(x) - f'(y)\| \leq L\|x - y\| \tag{24}$$

for all x, y from this ball. Suppose that $f'(x^0)^{-1}$ exists and that $\|f'(x^0)^{-1}\| \leq B_0$. Let

$$\|x^1 - x^0\| = \|f'(x^0)^{-1} f(x^0)\| \leq \eta_0$$

and assume that

$$h_0 = B_0 \eta_0 L \leq \frac{1}{2}, \; r_0 = \frac{1 - \sqrt{1 - 2h_0}}{h_0} \eta_0 \leq r.$$

Then the Newton iterates (23) are well defined, remain in the ball $\{x \| \|x - x^0\| \leq r_0\}$ and converge to a solution x^* of $f(x) = 0$ which is unique in $D \cap \{x \| \|x - x^0\| \leq r_1\}$ where

$$r_1 = \frac{1 + \sqrt{1 - 2h_0}}{h_0} \eta_0$$

provided $r \geq r_1$. Moreover the error estimate

$$\|x^* - x\| \leq \frac{1}{2^{k-1}} (2h_0)^{2^k - 1} \eta_0, \; k \geq 0, \tag{25}$$

holds. □

Please note that there also exists an affine–invariant form of the Newton–Kantorovich theorem. See [5].

Theorem 5 has been used in [16] to prove the existence of solutions by explicitly computing L (this can be done by interval arithmetic evaluation of the second partial derivatives provided they exist) and the bounds B_0 and η_0.

Since $h_0 \leq \frac{1}{2}$ the error estimate (25) (for $k = 0, 1$ and the ∞–norm) leads to

$$\begin{aligned} \|x^* - x^0\|_\infty &\leq 2\eta_0 &&= 2\|x^1 - x^0\|_\infty, \\ \|x^* - x^1\|_\infty &\leq 2h_0\eta_0 &&\leq \eta_0 = \|x^1 - x^0\|_\infty. \end{aligned}$$

This suggests a simple construction of an interval vector containing the solution. The situation is illustrated in Figure 1.

If x^0 is close enough to the solution x^* then x^1 is much closer than x^0 since Newton's method is quadratically convergent. The same holds if we choose any vector ($\neq x^*$) from the ball $\{x \| \|x - x^1\|_\infty \leq \eta_0\}$ as starting vector for Newton's method. Because of (22) and since $x^* \in K([x], x, C)$ it is reasonable to assume that $K([x], x^1, f'(x^0)^{-1}) \subseteq [x]$ for

$$[x] = \{x \| \|x - x^1\|_\infty \leq \eta_0\}. \tag{26}$$

The important point is that the test interval $[x]$ can be computed without knowing B_0 and L. Of course all the arguments above are based on the assumption that the hypothesis of the Newton–Kantorovich theorem is satisfied, which may not be the case if x^0 is far away from x^*.

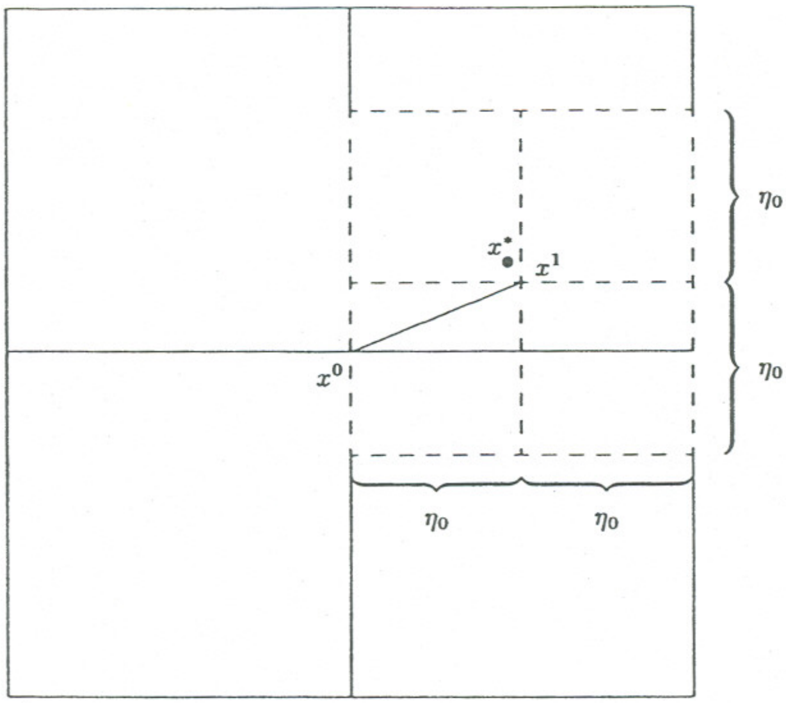

Figure 1. Error estimate (25) for $k = 1$ and the ∞–norm

We try to overcome this difficulty by performing first a certain number of Newton steps until we are close enough to a solution x^* of $f(x) = 0$. Then we compute the interval vector (26) and using the Krawczyk operator we test whether this interval contains a solution. The question of when to terminate the Newton iteration is answered by the following considerations.

Our general assumption is that the Newton iterates are convergent to x^*. We set

$$[y] := K([x], x^{k+1}, f'(x^k)^{-1})$$

where

$$\begin{aligned} [x] &= \{x \in \mathbf{R}^n \,|\, \|x^{k+1} - x\|_\infty \leq \eta_k\} \\ \eta_k &= \|x^{k+1} - x^k\|_\infty \end{aligned}$$

for some fixed k.

Our goal is to terminate Newton's method as soon as

$$\frac{\|w([y])\|_\infty}{\|x^{k+1}\|_\infty} \leq eps \tag{27}$$

holds, where eps is the machine precision of the floating point system. If $x^* \in [x]$ then $x^* \in [y]$ so that for any $y \in [y]$ we have

$$\frac{\|x^* - y\|_\infty}{\|x^*\|_\infty} \leq \frac{\|w([y])\|_\infty}{\|x^*\|_\infty}.$$

Since $\|x^*\|_\infty$ differs only slightly from $\|x^{k+1}\|_\infty$ if x^{k+1} is near x^*, condition (27) guarantees that the relative error with which any $y \in [y]$ approximates x^* is close to machine precision.

A discussion which has been performed in [4] leads to the following result:

As soon as the inequality

$$\frac{8\eta_k^3}{\|x^{k+1}\|_\infty \cdot \eta_{k-1}^2} \leq eps \tag{28}$$

is satisfied Newton's method is stopped. This stopping criterion needs only eps, x^{k-1}, x^k and x^{k+1}. Hence (28) can be checked at each step of Newton's method as soon as three successive iterates have been computed. Extensive numerical testing has shown that the proposed method has very good practical performance (see [4] and [6]).

Finally we note that the so-called ϵ-inflation considered first in [16] is another method for computing test intervals. A more theoretical investigation of this approach can be found in [11].

REFERENCES

1. G. Alefeld, Über das Divergenzverhalten des Intervall-Newton-Verfahrens. Computing 46, 289-294 (1991).
2. G. Alefeld, J. Herzberger, Introduction to Interval Computations. Academic Press (1983).
3. G. Alefeld, L. Platzöder, A Quadratically Convergent Krawczyk-like Algorithm. SIAM J. Numer. Anal. 20, 211-219 (1983).
4. G. Alefeld, A. Gienger, F. A. Potra, Efficient Numerical Validation of Solutions of Nonlinear Systems. Reports on Computational Mathematics Nr. 35. The University of Iowa, Department of Mathematics. October 1992.
5. P. Deuflhard, G. Heindl, Affine invariant convergence theorems for Newton's method and extensions to related methods. SIAM J. Numer. Anal. 16, 1-10 (1979).
6. A. Gienger, Zur Lösungsverifikation bei nichtlinearen Gleichungssystemen. Diplomarbeit, Universität Karlsruhe, 1992 (TEX-File).
7. P. Henrici, Elements of Numerical Analysis. John Wiley & Sons, Inc. New York, London, Sidney (1964).
8. R. Krawczyk, Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. Computing 4, 187-201 (1969).
9. G. Mayer, Some Remarks on Two Interval-Arithmetic Modifications of the Newton Method. Computing 48, 125-128 (1992).
10. G. Mayer, Old and New Aspects for the Interval Gaussian Algorithm. In: Computer Arithmetic, Scientific Computation and Mathematical Modelling. E. Kaucher, S.M. Markov, G. Mayer (editors). J.C. Baltzer AG Scientific Publishing Co. Basel (1992).
11. G. Mayer, Epsilon-Inflation in Verification Algorithms. Submitted for Publication.
12. R. E. Moore, A test for existence of solutions to nonlinear systems. SIAM J. Numer. Anal. 14, 611-615 (1977).
13. A. Neumaier, Interval methods for systems of equations. Cambridge: Cambridge University Press 1990.

14. J. M. Ortega, W. C. Rheinholdt, Iterative Solution of Nonlinear Equations in Several Variables. Academic Press. New York and London (1970).
15. L. B. Rall, A comparison of the existence theorems of Kantorovich and Moore. SIAM J. Numer. Anal. 17 (1980), 148–161.
16. S. M. Rump, Kleine Fehlerschranken bei Matrixproblemen. Thesis. Fakultät für Mathematik. Universität Karlsruhe (1980).
17. H. Schwandt, Schnelle fast global konvergente Verfahren für die Fünf–Punkte–Diskretisierung der Poissongleichung mit Dirichletschen Randbedingungen auf Rechteckgebieten. Thesis. Fachbereich Mathematik der TU Berlin (1981).