INTERVAL ARITHMETIC TOOLS FOR RANGE APPROXIMATION AND INCLUSION OF ZEROS

G. ALEFELD
Institut für Angewandte Mathematik
Universität Karlsruhe
76128 Karlsruhe
e-mail: goetz.alefeld@math.uni-karlsruhe.de

1. Introduction

In this paper we start in section 2 with an introduction to the basic facts of interval arithmetic: We introduce the arithmetic operations, explain how the range of a given function can be included and discuss the problem of overestimation of the range. Finally we demonstrate how range inclusion (of the first derivative of a given function) can be used to compute zeros by a so-called enclosure method.

An enclosure method usually starts with an interval vector which contains a solution 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. This will be discussed in section 3 for so-called Newton-like enclosure methods. An interesting feature of inclusion methods is that they can also be used to prove that there exists no solution in an interval vector. It will be shown that this proof needs only few steps if the test vector has already a small enough diameter. In the last section we demonstrate how for a given nonlinear system a test vector can be constructed which will very likely contain a solution.

A very important point is, of course, the fact that all these ideas can be performed in a safe way (especially with respect to rounding errors) on a computer. We can not go into any details in this paper and refer instead to the survey paper [14] by U. Kulisch and W. Miranker.

H. Bulgak and C. Zenger (eds.), Error Control and Adaptivity in Scientific Computing, 1–21.

2. On Computing the Range of Real Functions by Interval Arithmetic Tools

Let $[a] = [a_1, a_2]$, $b = [b_1, b_2]$ be real intervals and $*$ one of the basic operations 'addition', 'subtraction', 'multiplication' and 'division', respectively, that is $* \in \{+, -, \times, /\}$. Then we define the corresponding operations for intervals $[a]$ and $[b]$ by

$$[a] * [b] = \{a * b \quad | \quad a \in [a], b \in [b]\}, \tag{1}$$

where we assume $0 \notin [b]$ in case of division.
It is easy to prove that the set $\mathbf{I}(\mathbb{R})$ of real intervals is closed with respect to these operations. What is even more important is the fact that $[a] * [b]$ can be represented by using only the bounds of $[a]$ and $[b]$.
The following rules hold:

$$\begin{aligned}
[a] + [b] &= [a_1 + b_1,\ a_2 + b_2]\,, \\
[a] - [b] &= [a_1 - b_2,\ a_2 - b_1]\,, \\
[a] \times [b] &= [\min\{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}, \max\{a_1b_1, a_1b_2, a_2b_1, a_2b_2\}].
\end{aligned}$$

If we define

$$\frac{1}{[b]} = \left\{\frac{1}{b} \quad | \quad b \in [b]\right\} \qquad \text{if } 0 \notin [b]\,,$$

then

$$[a]/[b] = [a] \times \frac{1}{[b]}\,.$$

From (1) it follows immediately that the introduced operations for intervals are *inclusion monotone* in the following sense:

$$[a] \subseteq [c]\,,\, [b] \subseteq [d] \Rightarrow [a] * [b] \subseteq [c] * [d]\,. \tag{2}$$

If we have given a rational function (that is a polynomial or a quotient of two polynomials), and a fixed interval $[x]$, and if we take an $x \in [x]$ then, applying inclusion monotonicity repeatedly, we obtain

$$x \in [x] \Rightarrow f(x) \in f([x])\,. \tag{3}$$

Here $f([x])$ denotes the interval which one obtains by replacing the real variable x by the interval $[x]$ and evaluating this expression following the rules of interval arithmetic. From (3) it follows that

$$R(f; [x]) \subseteq f([x]) \tag{4}$$

where $R(f; [x])$ denotes the range of f over $[x]$. $f([x])$ is usually called (an) *interval arithmetic evaluation* of f over $[x]$.
(4) is the fundamental property on which nearly all applications of interval arithmetic are based. It is important to stress what (4) really is delivering: Without

any further assumptions is it possible to compute lower and upper bounds for the range over an interval by using only the bounds of the given interval.
The concept of interval arithmetic evaluation can by generalized to more general functions without principal difficulties.

EXAMPLE 1.

Consider the rational function

$$f(x) = \frac{x}{1-x} \quad , \qquad x \neq 1$$

and the interval $[x] = [2,3]$. It is easy to see that

$$\begin{aligned} R(f;[x]) &= \left[-2, -\frac{3}{2}\right], \\ f([x]) &= [-3,-1], \end{aligned}$$

which confirms (4).
For $x \neq 0$ we can rewrite $f(x)$ as

$$f(x) = \frac{1}{\frac{1}{x} - 1}, \quad x \neq 0\,, x \neq 1$$

and replacing x by the interval $[2,3]$ we get

$$\frac{1}{\frac{1}{[2,3]} - 1} = \left[-2, -\frac{3}{2}\right] = R(f;[x]).$$

□

From this example it is clear that the quality of the interval arithmetic evaluation as an enclosure of the range of f over an interval $[x]$ is strongly dependent on how the expression for $f(x)$ is written. In order to measure this quality we introduce the so-called Hausdorff *distance* between intervals:
Let $[a] = [a_1, a_2]$, $[b] = [b_1, b_2]$, then

$$q([a],[b]) = \max\{|a_1 - b_1|, |a_2 - b_2|\}.$$

Furthermore we use

$$d[a] = a_2 - a_1$$

and call $d[a]$ *diameter* of $[a]$.

How large is the overestimation of $R(f;[x])$ by $f([x])$?
This question is answered by the following

THEOREM 1. (MOORE [17])

Let there be given a continuous function $f : D \subset \mathbb{R}$ and assume that the interval arithmetic evaluation exists for all $[x] \subseteq [x]^0 \subseteq D$. Then

$$\begin{aligned} q(R(f;[x]), f([x])) &\leq \gamma \cdot d[x]\,, \quad \gamma \geq 0\,, \\ d\,f([x]) &\leq \delta \cdot d[x]\,, \quad \delta \geq 0\,. \end{aligned}$$

□

We do not discuss here which functions are allowed in order that the interval arithmetic evaluation exists. The theorem states that if it exists then the Hausdorff distance between $R(f;[x])$ and $f([x])$ goes linearly to zero with the diameter $d[x]$. Similarly the diameter of the interval arithmetic evaluation goes linearly to zero if $d[x]$ is approaching zero.
On the other hand we have seen in the second part of Example 1 that $f([x])$ may be dependent on the expression which is used for computing $f([x])$. Therefore the following question is natural:
Is it possible to rearrange the variables of the given function expression in such a manner that the interval arithmetic evaluation gives higher than linear order of convergence to the range of values?
We consider first the simple example

EXAMPLE 2.

Let $f(x) = x - x^2\,, \quad x \in [0,1] = [x]^0$.
It is easy to see that for $0 \leq r \leq \frac{1}{2}$ and $[x] = \left[\frac{1}{2} - r, \frac{1}{2} + r\right]$ we have

$$R(f;[x]) = \left[\frac{1}{4} - r^2, \frac{1}{4}\right]$$

and

$$f([x]) = \left[\frac{1}{4} - 2r - r^2, \frac{1}{4} + 2r - r^2\right]\,.$$

From this it follows

$$q(R(f;[x]), (f([x])) \leq \gamma \cdot d[x] \text{ with } \gamma = 1\,,$$

and

$$d\,f([x]) \leq \delta \cdot d[x] \text{ with } \delta = 2$$

in agreement with Theorem 1.
If we rewrite $f(x)$ as

$$x - x^2 = \frac{1}{4} - \left(x - \frac{1}{2}\right)\left(x - \frac{1}{2}\right)$$

and plug in the interval $[x] = \left[\frac{1}{2} - r, \frac{1}{2} + r\right]$ on the right hand side then we get the interval $\left[\frac{1}{4} - r^2, \frac{1}{4} + r^2\right]$ which, of course, includes $R(f;[x])$ again, and

$$q\left(R(f;[x]), \left[\frac{1}{4} - r^2, \frac{1}{4} + r^2\right]\right) = r^2 = \frac{1}{4}(d[x])^2 .$$

Hence the distance between $R(f;[x])$ and the enclosure interval $\left[\frac{1}{4} - r^2, \frac{1}{4} + r^2\right]$ goes quadratically to zero with the diameter of $[x]$. □

The preceding example is an illustration for the following general result.

THEOREM 2 (*The centered form*)

Let the function $f : \mathbb{R} \to \mathbb{R}$ *be represented in the 'centered form'*

$$f(x) = f(z) + (x - z) \cdot h(x) \tag{5}$$

for some $z \in [x] \subseteq D$. *If we define*

$$f([x]) := f(z) + ([x] - z) \cdot h([x])$$

then

$$a)\ R(f;[x]) \subseteq f([x])$$

and

$$b)\ q(R(f;[x]), f([x])) \le \kappa \cdot (d[x])^2 , \quad \kappa \ge 0 .$$ □

b) is called '*quadratic approximation property*' of the centered form. For rational functions it is not difficult to find a centered form. See [21], for example.
After having introduced the centered form it is natural to ask if there are forms which deliver higher than quadratic order of approximation of the range. Unfortunatly this is not the case as has been shown recently by P. Hertling [11]. See also [18].
Nevertheless in special cases one can use so-called generalized centered forms to get higher order approximations of the range. See [8], e.g. . Another interesting idea which uses a so-called 'remainder form of f' was introduced by Cornelius and Lohner [10].
In passing we note that the principal results presented up to this point also hold for functions of several variables.

As a simple example for the demonstration how the ideas of interval arithmetic can be applied we consider the following problem:
Let there be given a differentiable function $f : D \subset \mathbb{R} \to \mathbb{R}$ and an interval $[x]^0 \subseteq D$ for which the interval arithmetic evaluation of the derivative exists and does not contain zero: $0 \notin f'([x]^0)$. We want to check whether there exists a zero x^* in $[x]^0$, and if it exists we want to compute it by producing a sequence

of intervals containing x^* with the property that the lower and upper bounds are converging to x^*. (Of course, checking the existence is easy in this case by evaluating the function at the endpoints of $[x]^0$. However, the idea following works also for systems of equations. This will be shown in the next section).

For $[x] \subseteq [x]^0$ we introduce the so-called *Interval-Newton-Operator*

$$N[x] = m[x] - \frac{f(m[x])}{f'([x])}, \quad m[x] \in [x], \tag{6}$$

and consider the following iteration method

$$[x]^{k+1} = N[x]^k \cap [x]^k, \quad k = 0, 1, 2, \ldots, \tag{7}$$

which is called *Interval-Newton-Method.*
The properties of the method are described in the following result.

THEOREM 3.

Under the above assumptions the following hold for (7) :

a) $x^* \in [x]^0, f(x^*) = 0 \Longrightarrow \left\{[x]^k\right\}_{k=0}^{\infty}$ *is well defined,* $x^* \in [x]^k$, $\lim_{k\to\infty} [x]^k = x^*$.

If $d\, f'([x]) \leq c \cdot d[x]$, $[x] \subseteq [x]^0$, *then* $d[x]^{k+1} \leq \gamma \left(d[x]^k\right)^2$.

b) $N[x]^{k_0} \cap [x]^{k_0} = \emptyset$ *(= empty set) for some* $k_0 \geq 0$ *iff* $f(x) \neq 0$, $x \in [x]^0$. □

Hence, in case a), the diameters are converging quadratically to zero. On the other hand, if the method (7) breaks down because of empty intersection after a finite number of steps then it is *proved* that there exists no zero of f in $[x]^0$. From a practical point of view it would be interesting to have qualitiative knowledge about the size of k_0 in this case. This will be discussed in the next section in a more general setting.

3. Enclosing Solutions of Nonlinear Systems by Newton-like Methods

At the end of the last section we introduced the so-called Interval-Newton-Method for a single equation. In order that we can introduce this and similar methods for systems of simultaneous equations we have to discuss some basic facts about interval matrices and linear equations with intervals as coefficients. For a more general discussion of this subject we refer to [6], especially chapter 10.
An interval matrix is an array with intervals as elements. Operations between interval matrices are defined in the usual manner.

If $[A] = ([a_{ij}])$ and $[B] = ([b_{ij}])$ are interval matrices and if $c = (c_i)$ is a real vector then

$$[A]([B]c) \subseteq ([A][B])c\,. \tag{8}$$

This was proved in [22], p. 15. If, however, c is equal to one of the unit vectors e^i then

$$[A]([B]e^i) = ([A][B])e^i\,. \tag{9}$$

Assume now that we have given an n by n interval matrix $[A] = ([a_{ij}])$ which contains no singular matrix and an interval vector $[b] = ([b_i])$. By applying formally the formulas of the Gaussian algorithm we compute an interval vector $[x] = ([x_i])$ for which the relation

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

holds. See [6], section 15, for example. Here we assumed that no division by an interval which contains zero occurs in the elimination process. Some sufficient conditions for this are contained in [6]. See also [15]. It is an open question to find necessary and sufficient conditions for the feasibility of the Gaussian elimination process in the case of an interval matrix.
Subsequently we denote by

$$IGA([A],[b])$$

the result of the Gaussian algorithm applied to the interval matrix $[A]$ and the right hand side $[b]$, whereas

$$IGA([A])$$

is the interval matrix whose i-th column is obtained as $IGA([A],e^i)$ where e^i denotes the i-th unit vector again. In other words: $IGA([A])$ is an enclosure for the inverses of all matrices $A \in [A]$.

Now, let there be given a mapping

$$f : [x] \subset D \subset \mathbb{R}^n \to \mathbb{R}^n \tag{10}$$

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

$$f(x) - f(y) = J(x) \cdot (x - y)\,, \quad x \in D\,, \tag{11}$$

where

$$J(x) = \int_0^1 f'(y + t(x-y))dt\,, \quad x \in [x]\,. \tag{12}$$

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

$$J(x) \in f'([x]) \tag{13}$$

where $f'([x])$ denotes the interval arithmetic evaluation of the Jacobian of f.

In analogy to (6) we introduce the Interval-Newton-Operator $N[x]$. Suppose that $m[x] \in [x]$ is a real vector. Then

$$N[x] = m[x] - IGA(f'([x]), f(m[x])) \,. \qquad (14)$$

The Interval-Newton-Method is defined by

$$[x]^{k+1} = N[x]^k \cap [x]^k \,, \quad k = 0, 1, 2, \ldots \quad . \qquad (15)$$

Analogously to Theorem 3 we have the following result.

THEOREM 4.

Let there be given an interval vector $[x]^0$ *and a continuously differentiable mapping* $f : [x]^0 \subset D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *and assume that the interval arithmetic evaluation* $f'([x]^0)$ *of the Jacobian exists. Assume that* $IGA(f'([x]))$ *exists (which is identical to assuming that the Gaussian algorithm is feasible for* $f'([x]^0)$*). Assume that* $\rho(A) < 1$ *(*ρ *denotes the spectral radius of the matrix* A*) where*

$$A = |I - IGA(f'([x]^0)) \cdot f'([x]^0)| \,. \qquad (16)$$

(I *denotes the identity. The absolute value of an interval matrix, say* $[H] = ([h_{ij}])$*, is defined elementwise by* $|[H]| = (|h_{ij}|)$*. For a single interval* $[h] = [h_1, h_2]$ *we define* $|[h]| = max\{|h_1|, |h_2|\}$*).*

a) If f *has a (necessarily unique) zero* x^* *in* $[x]^0$ *then the sequence* $\{[x]^k\}_{k=0}^{\infty}$ *defined by (15) is well defined,* $x^* \in [x]^k$ *and* $\lim_{k\to\infty} [x]^k = x^*$*.*
Moreover, if

$$d\, f'([x])_{ij} \le c \cdot \|d[x]\| \,, \quad c \ge 0 \,, \quad 1 \le i, j \le n \,, \qquad (17)$$

for $[x] \subseteq [x]^0$ *(where* $d[x]$ *is a real vector which is obtained by forming componentwise the diameter) then*

$$\|d[x]^{k+1}\| \le \gamma \|d[x]^k\|^2 \,, \quad \gamma \ge 0 \,. \qquad (18)$$

b) $N[x]^{k_0} \cap [x]^{k_0} = \emptyset$ *(= empty set) for some* $k_0 \ge 0$ *iff* $f(x) \ne 0$, $x \in [x]^0$.

□

Note that in contrast to the onedimensional case we need the condition (16). Because of continuity reasons this condition always holds if the diameter $d[x]^0$ of the given interval vector ('starting interval') is componentwise small enough (and if $f'([x]^0)$ contains no singular matrix) since because of Theorem 1 we have $A = 0$ in the limit case $d[x]^0 = 0$. H. Schwandt [22] has discussed a simple example in the case $\rho(A) \ge 1$ which shows that for a certain interval vector (15) is feasible, $x^* \in [x]^k$, but $\lim_{k\to\infty} [x]^k \ne x^*$.

In case a) of the preceding theorem we have by (18) quadratic convergence of the diameters of the enclosing intervals to the zero vector. This is the same favorable behaviour as it is well known for the usual Newton-Method. If there is no solution x^* of $f(x) = 0$ in $[x]^0$ this can be detected by applying (18) until the intersection becomes empty for some k_0. From a practical point of view it is important that k_0 is not big in general. Under natural conditions it can really be proved that k_0 is small if the diameter of $[x]^0$ is small:

Note that a given interval vector $[x] = ([x_i])$ can be represented by two real vectors x^1 and x^2 which have as its components the lower and upper bounds of $[x_i]$, respectively. Similarly we also write $[x] = ([x_i]) = [x^1, x^2]$ and $N[x] = [n^1, n^2]$ for the Interval-Newton-Operator (14). Now it is easy to prove that

$$N[x] \cap [x] = \emptyset$$

iff for at least one component i_0 either

$$\left(n^2 - x^1\right)_{i_0} < 0 \tag{19}$$

or

$$\left(x^2 - n^1\right)_{i_0} < 0\,. \tag{20}$$

Furthermore it can be shown that

$$x^2 - n^1 \le \mathcal{O}\left(||d[x]||^2\right) + A^2 f(x^2) \tag{21}$$

and

$$n^2 - x^1 \le \mathcal{O}\left(||d[x]||^2\right) - A^1 f(x^1) \tag{22}$$

provided (17) holds. Here A^1 and A^2 are two real matrices contained in $IGA(f'([x]^0))$ and $\mathcal{O}\left(||d[x]||^2\right)$ denotes a real vector whose components all have the order $\mathcal{O}\left(||d[x]||^2\right)$. Furthermore if $f(x) \neq 0\,,\ x \in [x]$, then for sufficiently small diameter $d[x]$ there is at least one $i_0 \in \{1, 2, \ldots, n\}$ such that

$$\left(A^1 \cdot f(x^1)\right)_{i_0} \neq 0 \tag{23}$$

and

$$\operatorname{sign}\left(A^1 \cdot f(x^1)\right)_{i_0} = \operatorname{sign}\left(A^2 \cdot f(x^2)\right)_{i_0}\,. \tag{24}$$

Assume now that $\operatorname{sign}\left(A^1 \cdot f(x^1)\right)_{i_0} = 1$. Then for sufficiently small diameter $d[x]$ we have $\left(n^2 - x^1\right)_{i_0} < 0$ by (22) and by (19) the intersection becomes empty. If $\operatorname{sign}\left(A^1 \cdot f(x^1)\right)_{i_0} = -1$ then by (21) we obtain $x^2 - n^1 < 0$ for sufficiently small $d[x]$ and by (20) the intersection becomes again empty.

If $N[x]^{k_0} \cap [x]^{k_0} = \emptyset$ for some k_0 then the Interval-Newton-Method breaks down and we speak of *divergence* of this method. Because of the terms $\mathcal{O}\left(||d[x]||^2\right)$ in (21) and (22) we can say that in the case $f(x) \neq 0\,,\ x \in [x]^0$, the Interval-Newton-Method is *quadratically divergent*.

We demonstrate this behaviour by a simple onedimensional example.

EXAMPLE 3.

Consider the polynomial

$$f(x) = x^5 + x^4 - 11x^3 - 3x^2 + 18x$$

which has only simple real zeros contained in the interval $[x]^0 = [-5, 6]$. Hence (7) can not be performed since $0 \in f'([x]^0)$. Using a modification of the Interval-Newton-Method described already in [3] one can compute disjoint subintervals of $[x]^0$ for which the interval arithmetic evaluation does not contain zero. Hence (7) can be performed for each of these intervals. If such a subinterval contains a zero then a) of Theorem 3 holds, otherwise b) is true. The following table contains the intervals which were obtained by applying the generalized Interval-Newton-Method until $0 \notin f'([x])$ for all computed subintervals of $[x]^0$ (for simplicity we only give three digits in the mantissa).

TABLE 1.

n	
1	$[-0.356 \times 10^1; \quad -0.293 \times 10^1]$
2	$[-0.141 \times 10^1; \quad -0.870]$
3	$[-0.977; \quad 0.499]$
4	$[0.501; \quad 0.633]$
5	$[0.140 \times 10^1; \quad 0.185 \times 10^1]$
6	$[0.188 \times 10^1; \quad 0.212 \times 10^1]$
7	$[0.265 \times 10^1; \quad 0.269 \times 10^1]$
8	$[0.297 \times 10^1; \quad 0.325 \times 10^1]$
9	$[0.327 \times 10^1; \quad 0.600 \times 10^1]$

TABLE 2.

n	1	2	3	4*	5*	6	7*	8	9*
	5	6	9	1	2	6	1	5	3

The subintervals which do not contain a zero of f are marked by a star in Table 2. The number in the second line exhibits the number of steps until the intersection becomes empty. For $n = 9$ we have a diameter of approx. 2.75, which is not small, and after only 3 steps the intersection becomes empty. The intervals $n = 1, 2, 3, 6, 8$ each contain a zero of f. In the second line the number of steps are given which have to be performed until the lower and upper bound can be no longer improved on the computer. These numbers confirm the quadratic convergence of the diameters of the enclosing intervals. (For $n = 3$ the enclosed zero is $x^* = 0$ and we are in the underflow range). □

For more details concerning the speed of divergence see [4].

The Interval-Newton-Method has the big disadvantage that even if the interval arithmetic evaluation $f'([x]^0)$ of the Jacobian contains no singular matrix its feasibility is not guaranteed, $IGA\left(f'([x]^0), f(m[x]^0)\right)$ can in general only be computed if $d[x]^0$ is sufficiently small. For this reason Krawczyk [12] had the idea to introduce a mapping which today is called the *Krawczyk-Operator:*
Assume again that a mapping (10) with the corresponding properties is given. Then analogously to (14) we consider the so-called *Krawczyk-Operator*

$$K[x] = m[x] - C \cdot f(m[x]) + (I - C \cdot f'([x]))\,([x] - m[x]) \tag{25}$$

where C is a nonsingular real matrix. If we choose $C = m\left(f'([x])\right)^{-1}$ (= the inverse of the center of the interval arithmetic evaluation of the Jacobian) and $m[x]$ as the center of $[x]$ then for the so-called *Krawczyk-Method*

$$[x]^{k+1} = K[x]^k \cap [x]^k\,, \quad k = 0, 1, 2, \ldots \tag{26}$$

the same result as formulated for the Interval-Newton-Method in Theorem 4 holds.

PROOF.

a) By (11) we have

$$f(x^*) - f(m[x]) = J(m[x])(x^* - m[x])$$

and since $f(x^*) = 0$ it follows

$$\begin{aligned} x^* &= m[x] - C \cdot f(m[x]) + (I - C \cdot J(m[x]))(x^* - m[x]) \\ &\in m[x] - C \cdot f(m[x]) + (I - C \cdot f'([x]))([x] - m[x]) \\ &= K[x]\,. \end{aligned}$$

Hence if $x^* \in [x]^0$ then $x^* \in K[x]^0$ and therefore $x^* \in K[x]^0 \cap [x]^0 = [x]^1$.
Mathematical induction proves $x^* \in [x]^k\,, \quad k \geq 0$.
For the diameters of the sequence $\{[x]^k\}_{k=0}^{\infty}$ we have

$$\begin{aligned} d[x]^{k+1} &\leq d\,K[x]^k \\ &\leq |I - C_k \cdot f'([x]^k)|\, d[x]^k \\ &\leq |I - IGA(f'([x]^k)) \cdot f'([x]^k)|\, d[x]^k \\ &\leq A \cdot d[x]^k \end{aligned}$$

where A is defined by (16). Because of $\rho(A) < 1$ we have $\lim\limits_{k\to\infty} d[x]^k = 0$ and since $x^* \in [x]^k$ it follows $\lim\limits_{k\to\infty} [x]^k = x^*$. The proof for the quadratic convergence behaviour (18) follows from

$$\begin{aligned}
d[x]^{k+1} &\leq |I - C_k \cdot f'([x]^k)|\, d[x]^k \\
&\leq |C_k| \cdot |C_k^{-1} - f'([x]^k)|\, d[x]^k \\
&\leq |IGA(f'([x]^0))| \cdot |f'([x]^k) - f'([x]^k)|\, d[x]^k \\
&= |IGA(f'([x]^0))| \cdot d\, f'([x]^k) \cdot d[x]^k
\end{aligned}$$

by using (17).

b) Assume now that $K[x]^{k_0} \cap [x]^{k_0} = \emptyset$ for some $k_0 \geq 0$. Then $f(x) \neq 0$ for $x \in [x]^0$ since if $f(x^*) = 0$ for some $x^* \in [x]^0$ then Krawczyk's method is well defined and $x^* \in [x]^k$, $k \geq 0$.
If on the other hand $f(x) \neq 0$ and $N[x]^k \cap [x]^k \neq \emptyset$ then $\{[x]^k\}$ is well defined. Because of $\rho(A) < 1$ we have $d[x]^k \to 0$ and since we have a nested sequence it follows $\lim_{k\to\infty} [x]^k = \hat{x} \in \mathbb{R}^n$. Since the Krawczyk-Operator is continuous and since the same holds for forming intersections we obtain by passing to infinity in (26)

$$\begin{aligned}
\hat{x} &= K\hat{x} \cap \hat{x} = K\hat{x} \\
&= \hat{x} - f'(\hat{x})^{-1} f(\hat{x}) .
\end{aligned}$$

From this it follows that $f(\hat{x}) = 0$ in contrast to the assumption that $f(x) \neq 0$ for $x \in [x]^0$.
This completes the proof of Theorem 4 for the Krawczyk-Method. □

In case b) of the Theorem 4, that is if $K[x]^{k_0} \cap [x]^{k_0} = \emptyset$ for some k_0, we speak again of divergence (of the Krawczyk-Method). Similar as for the Interval-Newton-Method k_0 is small if the diameter of $[x]^0$ is small. This will be demonstrated subsequently.

As for the Interval-Newton-Operator we can represent $K[x]$ using two real vectors k^1 and k^2 and we write $K[x] = [k^1, k^2]$. Now $K[x] \cap [x] = \emptyset$ iff

$$(x^2 - k^1)_{i_0} < 0 \tag{27}$$

or

$$(k^2 - x^1)_{i_0} < 0 \tag{28}$$

for at least one $i_0 \in \{1, 2, \ldots, n\}$. (Compare with (19) and (20)).
We first prove that for $K[x]$ defined by (25) we have the vector inequalities

$$x^2 - k^1 \leq \mathcal{O}\left(||d[x]||^2\right) + C \cdot f(x^2) \tag{29}$$

and

$$k^2 - x^1 \leq \mathcal{O}\left(||d[x]||^2\right) - C \cdot f(x^1) \tag{30}$$

where $[x] = [x^1, x^2]$ and $\mathcal{O}\left(||d[x]||^2\right)$ denotes a real vector with components all of order $\mathcal{O}\left(||d[x]||^2\right)$.
We prove (29). Let $f'([x]) = [f_1', f_2']$ where f_1', f_2' are real matrices. If $\frac{1}{2}(f_1' + f_2')$ is nonsingular then we set

$$C := \frac{1}{2}(f_1' + f_2')^{-1} .$$

An easy computation shows that

$$I - C \cdot f'([x]) = \left(\frac{f_1' + f_2'}{2}\right)^{-1} \left[\frac{f_1' - f_2'}{2}, \frac{f_2' - f_1'}{2}\right]$$

and therefore

$$K[x] = m[x] - C \cdot f(m[x]) + \left(\frac{f_1' + f_2'}{2}\right)^{-1} \left[\frac{f_1' - f_2'}{2}, \frac{f_2' - f_1'}{2}\right] \cdot \frac{d[x]}{2} .$$

Hence

$$\begin{aligned} k^2 - x^1 &= m[x] - x^1 - C \cdot f(m[x]) + \left(\frac{f_1' + f_2'}{2}\right)^{-1} \cdot \frac{f_2' - f_1'}{2} \cdot \frac{d[x]}{2} \\ &= \frac{x^2 - x^1}{2} - C \cdot f(m[x]) + \mathcal{O}\left(||d[x]||^2\right) \end{aligned}$$

where we have used (17).
Choosing $y := x^1$ in (11) we have

$$f(x) - f(x^1) = J(x) \cdot (x - x^1)$$

where now

$$J(x) = \int_0^1 f'(x^1 + t(x - x^1)dt , \quad x \in [x] . \tag{31}$$

For $x = m[x]$ we therefore have

$$f(m[x]) - f(x^1) = J(m[x]) \cdot (m[x] - x^1)$$

where $J(m[x])$ is defined by (31). It follows that

$$\begin{aligned} k^2 - x^1 &= \frac{1}{2}d[x] - C \cdot f(x^1) - C \cdot J(m[x]) \cdot \frac{d[x]}{2} + \mathcal{O}\left(||d[x]||^2\right) \\ &= \frac{1}{2}(I - C \cdot J(m[x])) - C \cdot f(x^1) + \mathcal{O}\left(||d[x]||^2\right) . \end{aligned}$$

Since

$$\begin{aligned} I - C \cdot J(m[x]) &= C\left(C^{-1} - J(m[x])\right) \\ &\in IGA(f'([x])) \cdot (f'([x]) - f'([x])) \end{aligned}$$

the assertion follows by applying (17).
The second inequality can be shown in the same manner, hence (29) and (30) are proved.

If $f(x) \neq 0$, $x \in [x]$ and $d[x]$ is sufficiently small, then there exists an $i_0 \in \{1, 2, \ldots, n\}$ such that

$$(C \cdot f(x^1))_{i_0} \neq 0 \tag{32}$$

and

$$\text{sign}(C \cdot f(x^2))_{i_0} = \text{sign}(C \cdot f(x^1))_{i_0}\,. \tag{33}$$

This can be seen as follows: Since $x^1 \in [x]$ we have $f(x^1) \neq 0$ and since C is nonsingular it follows that $C \cdot f(x^1) \neq 0$ and therefore $(C \cdot f(x^1))_{i_0} \neq 0$ for at least one $i_0 \in \{1, 2, \ldots, n\}$ which proves (32). Using again (11) with $y = x^1$ we have

$$f(x) - f(x^1) = J(x) \cdot (x - x^1)\,, \quad x \in [x]$$

where

$$J(x) = \int_0^1 f'(x^1 + t(x - x^1))dt\,, \quad x \in [x]\,. \tag{34}$$

By choosing $x = x^2$ we have

$$f(x^2) - f(x^1) = J(x^2) \cdot (x^2 - x^1)$$

where $J(x^2)$ is defined by (34). It follows

$$C \cdot f(x^2) = C \cdot f(x^1) + C \cdot J(x^2) \cdot (x^2 - x^1)\,.$$

Since the second term on the right hand side approaches zero if $d[x] \to 0$ we have (33) for sufficiently small diameter $d[x]$.
Using (29), (30) together with (32) and (33) we can now show that for sufficiently small diameters of $[x]$ the intersection $K[x] \cap [x]$ becomes empty. See the analogous conclusions for the Interval-Newton-Method using (23), (24) together with (21) and (22). By the same motivation as for the Interval-Newton-Method we denote this behaviour as 'quadratic divergence' of the Krawczyk-Method.

It is important that either using the Interval-Newton-Operator or the Krawczyk-Operator one can also *prove* the *existence of a solution* of $f(x) = 0$ in a given interval vector. We formulate this fact as a theorem.

THEOREM 5.

Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable mapping and assume that the interval arithmetic evaluation $f'([x])$ of the Jacobian exists for some interval vector $[x] \subset D$.
a) *Suppose that the Gaussian algorithm is feasible for $f'([x])$ and assume*

$$y - IGA\,(f'([x]), f(y)) \subseteq [x]$$

where $y \in [x]$ is fixed. Then f has a unique zero x^* in $[x]$.
b) *Suppose that C is a nonsingular matrix and assume*

$$y - C \cdot f(y) + (I - C \cdot f'([x])) \cdot ([x] - y) \subseteq [x]$$

for some fixed $y \in [x]$. Then f has a zero x^* in $[x]$.

PROOF.

a) Since the Gaussian algorithm is feasible it follows that $f'([x])$ contains no singular matrices.
For fixed $y \in [x]$ we consider the equation (11) and $J(x)$ defined by (12). $J(x)$ is nonsingular because of (13).
Now consider the mapping

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

where

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

and $y \in [x]$ is fixed.

It follows, using the assumption,

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

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 in $[x]$. The uniqueness follows from the fact that $f'([x])$ contains no singular matrices.

b) Consider for the nonsingular matrix C the continuous mapping

$$q : [x] \subseteq D \subseteq \mathbb{R}^n \to \mathbb{R}^n$$

defined by

$$q(x) = x - C \cdot f(x), \quad x \in \mathbb{R}.$$

It follows, using the assumption,

$$\begin{aligned} q(x) &= x - C \cdot f(x) \\ &= x - C \cdot (f(x) - f(y)) - C \cdot f(y) \\ &= y + (x - y) - C \cdot J(x) \cdot (x - y) - C \cdot f(y) \\ &\in y - C \cdot f(y) + (I - C \cdot f'([x])) \cdot ([x] - y) \\ &\subseteq [x], \quad x \in [x]. \end{aligned}$$

By the same reasoning as before it follows $f(x^*) = 0$ for some $x^* \in [x]$. □

REMARK

It is easy to show that in case a) of the preceding theorem the unique zero of x^* is even in $y - IGA(f'([x]), f(y))$ and in case b) all zeros x^* of f in $[x]$ are even in $y - C \cdot f(y) + (I - C \cdot f'([x])) \cdot ([x] - y)$.

4. Verification of Solutions of Nonlinear Systems

The result of the last theorem in the preceding section can be used in a systematic manner for verifying the existence of a solution of a nonlinear system in an interval vector. Besides of the existence of a solution also componentwise error-bounds are delivered by such an interval vector. We are now going to discuss how such an interval vector can be constructed.
For a nonlinear mapping $f : D \subset \mathbb{R}^n \to \mathbb{R}^n$ we consider Newton's method

$$x^{k+1} = x^k - f'(x^k)^{-1} f(x^k), \quad k = 0, 1, \ldots \quad . \tag{35}$$

The Newton-Kantorowich 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 estimation in conjunction with the quadratic convergence property (18) which we have also proved for the Krawczyk-Method will lead us to a test interval which can be computed using only iterates of Newton's method.

THEOREM 6. (See [19], Theorem 12.6.2)

Assume that $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is differentiable in the ball $\{x \mid ||x - x^0|| \leq r\}$ and that

$$||f'(x) - f'(y)|| \leq L||x - y||$$

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

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

and assume that

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

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

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

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

$$\|x^* - x^k\| \leq \frac{1}{2^{k-1}}(2h_0)^{2^k-1}\eta_0\,,\, k \geq 0 \tag{36}$$

holds. □

Since $h_0 \leq \frac{1}{2}$, the error estimate (36) (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 x^*. The situation is illustrated in Figure 1.

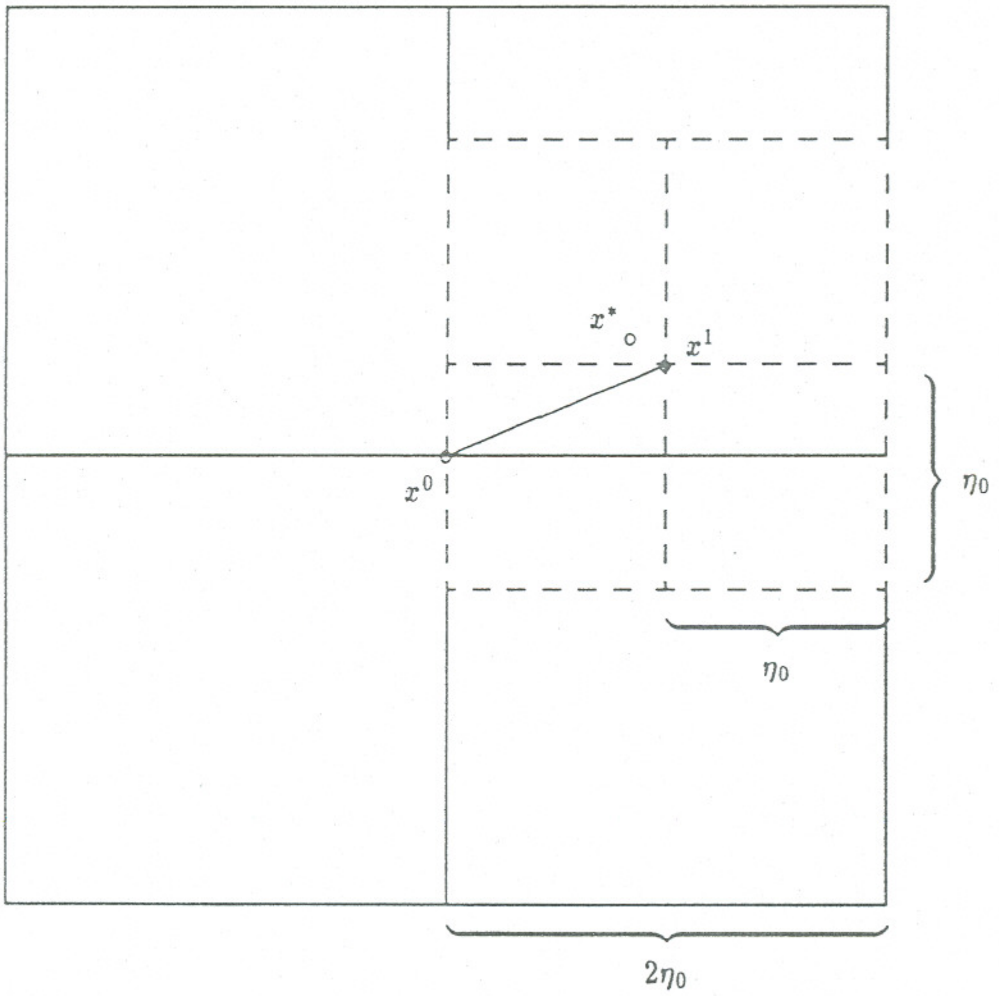

Figure 1: Error estimates (36) for $k = 1$ and the ∞–norm

If x^0 is close enough to the solution x^* then x^1 is much closer to x^* than x^0 since Newton's method is quadratically convergent. The same holds if we choose any vector ($\neq x^*$) from the ball $\{x \mid ||x - x^1||_\infty \leq \eta_0\}$ as starting vector for Newton's method. Because of (18) and since $x^* \in K[x]$ it is reasonable to assume that

$$K[x] = x^1 - f'(x^0)^{-1} \cdot f(x^1) + (I - f'(x^0)^{-1} \cdot f'([x])) \cdot ([x] - x^1) \subseteq [x]$$

for

$$[x] = \{x \mid ||x - x^1||_\infty \leq \eta_0\}. \tag{37}$$

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

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 (37) 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^*. For ease of notation we set

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

where

$$\begin{aligned} [x] &= \{x \in \mathbb{R}^n \mid ||x^{k+1} - x||_\infty \leq \eta_k\}, \\ \eta_k &= ||x^{k+1} - x^k||_\infty \end{aligned} \tag{38}$$

for some fixed k. Our goal is to terminate Newton's method as soon as

$$\frac{||d[y]||_\infty}{||x^{k+1}||_\infty} \leq eps \tag{39}$$

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

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

Since $||x^*||_\infty$ differs only slightly from $||x^{k+1}||_\infty$ if x^{k+1} is near x^*, condition (39) guarantees that the relative error with which any $y \in [y]$ approximates x^* is close to machine precision. Using (17) it can be shown that

$$||df'([x])||_\infty \leq \hat{L}||d[x]||_\infty$$

and

$$||d[y]||_\infty \leq ||f'(x^k)^{-1}||_\infty \cdot \tilde{L}||d[x]||_\infty^2$$

where $\tilde{L} = \max\{\hat{L}, L\}$, and since $||d[x]||_\infty = 2\eta_k$ the inequality (39) holds if

$$4\frac{||f'(x^k)^{-1}||_\infty \tilde{L}\eta_k^2}{||x^{k+1}||_\infty} \leq eps \tag{40}$$

is true.
From Newton's method we have

$$x^{k+1} - x^k = f'(x^k)^{-1}\left\{f(x^k) - f(x^{k-1}) - f'(x^k)^{-1}(x^k - x^{k-1})\right\}$$

and by 3.2.12 in [19] it follows that

$$\eta_k \leq \frac{1}{2}||f'(x^k)^{-1}||_\infty \tilde{L}\eta_{k-1}^2 .$$

Replacing the inequality sign by equality in this relation and eliminating $||f'(x^k)^{-1}||_\infty \tilde{L}$ in (40) we get the following stoping criterion for Newton's method:

$$\frac{8\eta_k^3}{||x^{k+1}||_\infty \eta_{k-1}^2} \leq eps . \tag{41}$$

Of course this is not a mathematical proof that if (41) is satisfied then the interval $[y]$ constructed as above will contain x^* and that the vectors in $[y]$ will approximate x^* with a relative error close to eps. However as has been shown in [5] the test based on the stopping criterion (41) works extremely well in practice.

The idea of this section has been generalized to nonsmooth mappings by X. Chen [9].

A very important point is also the fact that for the verification of solutions of nonlinear systems one can often replace the interval arithmetic evaluation of the Jacobian by an interval arithmetic enclosure of the slope-matrix of f. In this connection slopes have first been considered in [1]. See also [20].

Verification techniques have been applied to a series of fundamental problems by rewriting them as nonlinear systems. We mention the eigenvalue problem for matrices, the singular value problem, the generalized eigenvalue problem and the generalized singular value problem, e.g. . See [16] where one can find many references.
Interval arithmetic can also be applied in a systematic manner to bound the solution set of a given problem if the data are already contained in intervals. An interesting question in this context is how the solution set looks like. A couple of recent papers are concerned with the discussion of this question in the case of a linear system with interval entries. See [7], for example.

References

[1] Alefeld, G. (1981) Bounding the Slope of Polynomials and some Applications, Computing 26, 227-237

[2] Alefeld, G. (1994) Inclusion Methods for Systems of Nonlinear Equations - The Interval-Newton-Method and Modifications, in Herzberger, J.(ed.), Topics in Validated Computation, Elsevier

[3] Alefeld, G. (1968) Intervallrechnung über den komplexen Zahlen und einige Anwendungen, Ph. D. Thesis, University Karlsruhe

[4] Alefeld, G. (1991) Über das Divergenzverhalten des Intervall-Newton-Verfahrens, Computing 46, 289-294

[5] Alefeld, G., Gienger, A. and Potra, F. (1994) Efficient Numerical Validation of Solutions of Nonlinear Systems, SIAM J. Numer. Anal., 31(1), 252-260

[6] Alefeld, G. and Herzberger, J. (1983) Introduction to Interval Computations, Academic Press

[7] Alefeld, G., Kreinowich, V. and Mayer, G. (1998) The Shape of the Solution Set of Linear Interval Equations with Dependent Coefficients, Mathematische Nachrichten 192, 23-36

[8] Alefeld, G. and Lohner, R. (1985) On Higher Order Centered Forms, Computing 35, 177-184

[9] Chen, X. (1997) A Verification Method for Solutions of Nonsmooth Equations, Computing 58, 281-294

[10] Cornelius, H. and Lohner, R. (1984) Computing the Range of Values of Real Functions with Accuracy Higher than Second Order, Computing 33, 331-347

[11] Hertling, P. (Jan. 1998) A Lower Bound for Range Enclosure in Interval Arithmetic, Centre for Discrete Mathematics and Theoretical Computer Science Research Report Series, Department of Computer Science, University of Auckland

[12] Krawczyk, R. (1969) Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, Computing 4, 187-201

[13] Krawczyk, R. and Neumaier, A. (1985) Interval Slopes for Rational Functions and Associated Centered Forms, SIAM J. Numer. Anal., 22(3), 604-616

[14] Kulisch, U. and Miranker, W. (1986) The Arithmetic of the Digital Computer: A New Approach, SIAM Review 28(1), 1-40

[15] Mayer, G. (1992) Old and New Aspects for the Interval Gaussian Algorithm, in Kaucher, E., Markov, S. M. and Mayer, G. (eds.), Computer Arithmetic, Scientific Computation and Mathematical Modelling, J. C. Baltzer AG Scientific Publishing Co., Basel

[16] Mayer, G. (1994) Result Verification for Eigenvectors and Eigenvalues, in Herzberger, J. (ed.), Topics in Validated Computations, Elsevier

[17] Moore, R. E. (1966) Interval Analysis, Prentice Hall, Englewood Cliffs, New Jersey

[18] Nguyen, H. T., Kreinowich, Y., Nesterov, V. and Nakumura, M. (1997) On Hardware Support for Interval Computations and for Soft Computing: Theorems, IEEE Transactions on Fuzzy Systems, Vol 5, No 1, 108-127

[19] Ortega, J. M. and Rheinboldt, W. C. (1970) Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York and London

[20] Rall, L. B. (1969) Computational Solution of Nonlinear Operator Equations, Wiley, New York

[21] Ratschek, H. (1980) Centered Forms, SIAM J. Numer. Anal., 17, 656-662

[22] Schwandt, H. (1981) Schnelle fast global konvergente Verfahren für die Fünf-Punkte-Diskretisierung der Poissongleichung mit Dirichletschen Randbedingungen auf Rechteckgebieten, Thesis, Fachbereich Mathematik der TU Berlin