ELSEVIER

Journal of Computational and Applied Mathematics 121 (2000) 421–464

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

Interval analysis: theory and applications

Götz Alefeld[a,*], Günter Mayer[b]

[a] *Institut für Angewandte Mathematik, Universität Karlsruhe, D-76128 Karlsruhe, Germany*

[b] *Fachbereich Mathematik, Universität Rostock, D-18051 Rostock, Germany*

Received 13 August 1999

Abstract

We give an overview on applications of interval arithmetic. Among others we discuss verification methods for linear systems of equations, nonlinear systems, the algebraic eigenvalue problem, initial value problems for ODEs and boundary value problems for elliptic PDEs of second order. We also consider the item software in this field and give some historical remarks.

Contents

1. Historical remarks and introduction

First, we try to give a survey on how and where interval analysis was developed. Of course, we cannot give a report which covers all single steps of this development. We simply try to list some

* Corresponding author.
E-mail addresses: goetz.alefeld@math.uni-karlsruhe.de (G. Alefeld), guenter.mayer@mathematik.uni-rostock.de (G. Mayer).

0377-0427/00/$ - see front matter
PII: S0377-0427(00)00342-3

important steps and published papers which have contributed to it. This survey is, of course, strongly influenced by the special experience and taste of the authors.

A famous and very old example of an interval enclosure is given by the method due to Archimedes. He considered inscribed polygons and circumscribing polygons of a circle with radius 1 and obtained an increasing sequence of lower bounds and at the same time a decreasing sequence of upper bounds for the aera of the corresponding disc. Thus stopping this process with a circumscribing and an inscribed polygon, each of n sides, he obtained an interval containing the number π. By choosing n large enough, an interval of arbitrary small width can be found in this way containing π.

One of the first references to interval arithmetic as a tool in numerical computing can already be found in [35, p. 346 ff] (originally published in Russian in 1951) where the rules for the arithmetic of intervals (in the case that both operands contain only positive numbers) are explicitly stated and applied to what is called today interval arithmetic evaluation of rational expressions (see Section 2 of the present paper). For example, the following problem is discussed: What is the range of the expression

$$x = \frac{a+b}{(a-b)c}$$

if the exact values of a, b and c are known to lie in certain given intervals. By plugging in the given intervals the expression for x delivers a superset of the range of x.

According to Moore [64] P.S. Dwyer has discussed matrix computations using interval arithmetic already in his book [29] in 1951.

Probably the most important paper for the development of interval arithmetic has been published by the Japanese scientist Teruo Sunaga [88]. In this publication not only the algebraic rules for the basic operations with intervals can be found but also a systematic investigation of the rules which they fulfill. The general principle of bounding the range of a rational function over an interval by using only the endpoints via interval arithmetic evaluation is already discussed. Furthermore, interval vectors are introduced (as multidimensional intervals) and the corresponding operations are discussed. The idea of computing an improved enclosure for the zero of a real function by what is today called interval Newton method is already presented in Sunaga's paper (Example 9.1). Finally, bounding the value of a definite integral by bounding the remainder term using interval arithmetic tools and computing a pointwise enclosure for the solution of an initial value problem by remainder term enclosing have already been discussed there. Although written in English these results did not find much attention until the first book on interval analysis appeared which was written by Moore [64].

Moore's book was the outgrowth of his Ph.D. thesis [63] and therefore was mainly concentrated on bounding solutions of initial value problems for ordinary differential equations although it contained also a whole bunch of general ideas.

After the appearance of Moore's book groups from different countries started to investigate the theory and application of interval arithmetic systematically. One of the first survey articles following Moore's book was written by Kulisch [49]. Based on this article the book [12] was written which was translated to English in 1983 as [13].

The interplay between algorithms and the realization on digital computers was thoroughfully investigated by U. Kulisch and his group. Already in the 1960s, an ALGOL extension was created and

implemented which had a type for real intervals including provision of the corresponding arithmetic and related operators.

During the last three decades the role of compact intervals as independent objects has continuously increased in numerical analysis when verifying or enclosing solutions of various mathematical problems or when proving that such problems cannot have a solution in a particular given domain. This was possible by viewing intervals as extensions of real or complex numbers, by introducing interval functions and interval arithmetics and by applying appropriate fixed point theorems. In addition thoroughful and sophisticated implementations of these arithmetics on a computer together with – partly new – concepts such as controlled roundings, variable precision, operator overloading or epsilon–inflation made the theory fruitful in practice and effected that in many fields solutions could be automatically verified and (mostly tightly) enclosed by the computer.

In this survey article we report on some interval arithmetic tools. In particular, we present various crucial theorems which form the starting point for efficient interval algorithms. In Section 2 we introduce the basic facts of the 'standard' interval arithmetic: We define the arithmetic operations, list some of its properties and present a first way how the range of a given function can be included. We continue this latter topic in Section 3 where we also 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 4 for selected enclosure methods of nonlinear systems. An interesting feature of such 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. We also demonstrate how for a given nonlinear system a test vector can be constructed which will very likely contain a solution.

In Section 5 we address to systems of linear equations $Ax = b$, where we allow A and b to vary within given matrix and vector bounds, respectively. The ideas of Section 4 are refined and yield to interval enclosures of the corresponding set of solutions. As a particularity we restrict A within its bounds to be a symmetric matrix and provide methods for enclosing the associated smaller symmetric solution set. In both cases we show how the amount of overestimation by an interval vector can be measured without knowing the exact solution set.

Section 6 is devoted to mildly nonlinear topics such as the algebraic eigenvalue problem, the generalized algebraic eigenvalue problem, the singular value problem, and – as an application – a particular class of inverse eigenvalue problems.

In Section 7 we present crucial ideas for verifying and enclosing solutions of initial value problems for ordinary differential equations. For shortness, however, we must confine to the popular class of interval Taylor series methods.

Section 8 contains some remarks concerning selected classes of partial differential equations of the second order. We mainly consider elliptic boundary value problems and present an access which leads to a powerful verification method in this field.

The practical importance of interval analysis depends heavily on its realization on a computer. Combining the existing machine arithmetic with direct roundings it is possible to implement an interval arithmetic in such a way that all interval algorithms keep their – theoretically proved –

properties on existence, uniqueness and enclosure of a solution when they are performed on a computer. Based on such a machine interval arithmetic, software is available which delivers verified solutions and bounds for them in various fields of mathematics. We will shortly consider this topic in Section 9.

In the last 20 years both the algorithmic components of interval arithmetic and their realization on computers (including software packages for different problems) were further developed. Today the understanding of the theory and the use of adapted programming languages are indispensible tools for reliable advanced scientific computing.

2. Definitions, notations and basic facts

Let $[a]=[\underline{a},\bar{a}]$, $b=[\underline{b},\bar{b}]$ be real compact intervals and $\circ$ one of the basic operations 'addition', 'subtraction', 'multiplication' and 'division', respectively, for real numbers, that is $\circ\in\{+,-,\cdot,/\}$. Then we define the corresponding operations for intervals $[a]$ and $[b]$ by

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

where we assume $0\notin[b]$ in case of division.

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

$$[a]+[b]=[\underline{a}+\underline{b},\bar{a}+\bar{b}],$$

$$[a]-[b]=[\underline{a}-\bar{b},\bar{a}-\underline{b}],$$

$$[a]\cdot[b]=[\min\{\underline{ab},\underline{a}\bar{b},\bar{a}\underline{b},\bar{a}\bar{b}\},\max\{\underline{ab},\underline{a}\bar{b},\bar{a}\underline{b},\bar{a}\bar{b}\}].$$

If we define

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

then

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

If $\underline{a}=\bar{a}=a$, i.e., if $[a]$ consists only of the element a, then we identify the real number a with the degenerate interval $[a,a]$ keeping the real notation, i.e., $a\equiv[a,a]$. In this way one recovers at once the real numbers $\mathbb{R}$ and the corresponding real arithmetic when restricting $I(\mathbb{R})$ to the set of degenerate real intervals equipped with the arithmetic defined in (1). Unfortunately, $(I(\mathbb{R}),+,\cdot)$ is neither a field nor a ring. The structures $(I(\mathbb{R}),+)$ and $(I(\mathbb{R})/\{0\},\cdot)$ are commutative semigroups with the neutral elements 0 and 1, respectively, but they are not groups. A nondegenerate interval $[a]$ has no inverse with respect to addition or multiplication. Even the distributive law has to be replaced by the so-called subdistributivity

$$[a]([b]+[c])\subseteq[a][b]+[a][c]. \tag{2}$$

The simple example $[-1,1](1+(-1))=0\subset[-1,1]\cdot1+[-1,1]\cdot(-1)=[-2,2]$ illustrates (2) and shows that $-[-1,1]$ is certainly not the inverse of $[-1,1]$ with respect to $+$. It is worth noticing

that equality holds in (2) in some important particular cases, for instance if $[a]$ is degenerate or if $[b]$ and $[c]$ lie on the same side with respect to 0.

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

$$[a] \subseteq [c], [b] \subseteq [d] \Rightarrow [a] \circ [b] \subseteq [c] \circ [d]. \tag{3}$$

Standard interval functions $\varphi \in F = \{\sin, \cos, \tan, \arctan, \exp, \ln, \text{abs}, \text{sqr}, \text{sqrt}\}$ are defined via their range, i.e.,

$$\varphi([x]) = \{\varphi(x) | x \in [x]\}. \tag{4}$$

Apparently, they are extensions of the corresponding real functions. These real functions are continuous and piecewise monotone on any compact subinterval of their domain of definition. Therefore, the values $\varphi([x])$ can be computed directly from the values at the bounds of $[x]$ and from selected constants such as 0 in the case of the square, or $-1, 1$ in the case of sine and cosine. It is obvious that the standard interval functions are inclusion monotone, i.e., they satisfy

$$[x] \subseteq [y] \Rightarrow \varphi([x]) \subseteq \varphi([y]). \tag{5}$$

Let $f\colon D \subseteq \mathbb{R} \to \mathbb{R}$ be given by a mathematical expression $f(x)$ which is composed by finitely many elementary operations $+, -, \cdot, /$ and standard functions $\varphi \in F$. If one replaces the variable x by an interval $[x] \subseteq D$ and if one can evaluate the resulting interval expression following the rules in (1) and (4) then one gets again an interval. It is denoted by $f([x])$ and is usually called (an) interval arithmetic evaluation of f over $[x]$. For simplicity and without mentioning it separately we assume that $f([x])$ exists whenever it occurs in the paper.

From (3) and (5) the interval arithmetic evaluation turns out to be inclusion monotone, i.e.,

$$[x] \subseteq [y] \Rightarrow f([x]) \subseteq f([y]) \tag{6}$$

holds. In particular, $f([x])$ exists whenever $f([y])$ does for $[y] \supseteq [x]$. From (6) we obtain

$$x \in [x] \Rightarrow f(x) \in f([x]), \tag{7}$$

whence

$$R(f; [x]) \subseteq f([x]). \tag{8}$$

Here $R(f; [x])$ denotes the range of f over $[x]$.

Relation (8) is the fundamental property on which nearly all applications of interval arithmetic are based. It is important to stress what (8) 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.

Example 1. Consider the rational function

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

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

$$R(f; [x]) = [-2, -\tfrac{3}{2}],$$

$$f([x]) = [-3, -1],$$

which confirms (8).

For $x \neq 0$ we can rewrite $f(x)$ as

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

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

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

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 $q(\cdot,\cdot)$ between intervals with which $I(\mathbb{R})$ is a complete metric space:

Let $[a] = [\underline{a}, \bar{a}], [b] = [\underline{b}, \bar{b}]$, then

$$q([a],[b]) = \max\{|\underline{a} - \underline{b}|, |\bar{a} - \bar{b}|\}. \tag{9}$$

Furthermore, we use

$$\check{a} = \tfrac{1}{2}(\underline{a} + \bar{a}),$$

$$d[a] = \bar{a} - \underline{a},$$

$$|[a]| = \max\{|a| \, | a \in [a]\} = \max\{|\underline{a}|, |\bar{a}|\},$$

$$\langle [a] \rangle = \min\{|a| | a \in [a]\} = \begin{cases} 0, & \text{if } 0 \in [a], \\ \min\{|\underline{a}|, |\bar{a}|\} & \text{if } 0 \notin [a] \end{cases} \tag{10}$$

and call $\check{a}$ center, $d[a]$ diameter and $|[a]|$ absolute value of $[a]$.

In order to consider multidimensional problems we introduce $m \times n$ interval matrices $[A] = ([a_{ij}])$ with entries $[a_{ij}]$, $i = 1,\ldots,m$, $j = 1,\ldots,n$, and interval vectors $[x] = ([x_i])$ with n components $[x_i]$, $i = 1,\ldots,n$. We denote the corresponding sets by $I(\mathbb{R}^{m\times n})$ and $I(\mathbb{R}^n)$, respectively. Trivially, $[A]$ coincides with the matrix interval $[\underline{A}, \bar{A}] = \{B \in \mathbb{R}^{m\times n} \mid \underline{A} \leqslant B \leqslant \bar{A}\}$ if $\underline{A} = (\underline{a}_{ij}), \bar{A} = (\bar{a}_{ij}) \in \mathbb{R}^{m\times n}$ and if $A = (a_{ij}) \leqslant B = (b_{ij})$ means $a_{ij} \leqslant b_{ij}$ for all i,j. Since interval vectors can be identified with $n \times 1$ matrices, a similar property holds for them. The null matrix O and the identity matrix I have the usual meaning, e denotes the vector $e = (1,1,\ldots,1)^{\mathrm{T}} \in \mathbb{R}^n$. Operations between interval matrices and between interval vectors are defined in the usual manner. They satisfy an analogue of (6)–(8). For example,

$$\{Ax \mid A \in [A], \ x \in [x]\} \subseteq [A][x] = \left(\sum_{j=1}^{n} [a_{ij}][x_j] \right) \in I(\mathbb{R}^m) \tag{11}$$

if $[A] \in I(\mathbb{R}^{m\times n})$ and $[x] \in I(\mathbb{R}^n)$. It is easily seen that $[A][x]$ is the smallest interval vector which contains the left set in (11), but normally it does not coincide with it. An interval item which encloses some set S as tight as possible is called (interval) hull of S. The above-mentioned operations with two interval operands always yield to the hull of the corresponding underlying sets.

An interval matrix $[A] \in I(\mathbb{R}^{n\times n})$ is called nonsingular if it contains no singular real $n \times n$ matrix.

The Hausdorff distance, the center, the diameter and the absolute value in (9), (10) can be generalized to interval matrices and interval vectors, respectively, by applying them entrywise. Note that the results are real matrices and vectors, respectively, as can be seen, e.g., for

$$q([A],[B]) = (q([a_{ij}],[b_{ij}])) \in \mathbb{R}^{m\times n}$$

if $[A]$, $[B] \in I(\mathbb{R}^{m\times n})$. We also use the comparison matrix $\langle [A] \rangle = (c_{ij}) \in \mathbb{R}^{n\times n}$ which is defined for $[A] \in I(\mathbb{R}^{n\times n})$ by

$$c_{ij} = \begin{cases} \langle [a_{ij}] \rangle & \text{if } i = j, \\ -|[a_{ij}]| & \text{if } i \neq j. \end{cases}$$

By $\operatorname{int}([x])$ we denote the interior of an interval vector $[x]$, by $\rho(A)$ the spectral radius of $A \in \mathbb{R}^{n\times n}$ and by $\|\cdot\|_\infty$ the usual maximum norm for vectors from $\mathbb{R}^n$ or the row sum norm for matrices from $\mathbb{R}^{n\times n}$. In addition, the Euclidean norm $\|\cdot\|_2$ in $\mathbb{R}^n$ will be used. We recall that $A \in \mathbb{R}^{n\times n}$ is an M matrix if $a_{ij} \leqslant 0$ for $i \neq j$ and if A^{-1} exists and is nonnegative, i.e., $A^{-1} \geqslant O$. If each matrix A from a given interval matrix $[A]$ is an M matrix then we call $[A]$ an M matrix, too.

Let each component f_i of $f: D \subseteq \mathbb{R}^m \to \mathbb{R}^n$ be given by an expression $f_i(x)$, $i = 1,\ldots,n$, and let $[x] \subseteq D$. Then the interval arithmetic evaluation $f([x])$ is defined analogously to the one-dimensional case.

In this paper we restrict ourselves to real compact intervals. However, complex intervals of the form $[z] = [a] + i[b]$ ($[a],[b] \in I(\mathbb{R})$) and $[z] = \langle \check{z}, r \rangle$ ($\check{z}, r \in \mathbb{R}$, $r \geqslant 0$) are also used in practice. In the first form $[z]$ is a rectangle in the complex plane, in the second form it means a disc with midpoint $\check{z}$ and radius r. In both cases a complex arithmetic can be defined and complex interval functions can be considered which extend the presented ones. See [3,13] or [73], e.g., for details.

3. Computing the range of real functions by interval arithmetic tools

Enclosing the range $R(f;[x])$ of a function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ with $[x] \subseteq D$ is an important task in interval analysis. It can be used, e.g., for

- localizing and enclosing global minimizers and global minima of f on $[x]$ if $m = 1$,
- verifying $R(f;[x]) \subseteq [x]$ which is needed in certain fixed point theorems for f if $m = n$,
- enclosing $R(f';[x])$, i.e., the range of the Jacobians of f if $m = n$,
- enclosing $R(f^{(k)};[x])$, i.e., the range of the kth derivative of f which is needed when verifying and enclosing solutions of initial value problems,
- verifying the nonexistence of a zero of f in $[x]$.

According to Section 2 an interval arithmetic evaluation $f([x])$ is automatically an enclosure of $R(f,[x])$. As Example 1 illustrates $f([x])$ may overestimate this range. The following theorem shows how large this overestimation may be.

Theorem 1 (Moore [64]). *Let $f:D\subset\mathbb{R}^n\to\mathbb{R}$ be continuous and let $[x]\subseteq[x]^0\subseteq D$. Then (under mild additional assumptions)*

$$q(R(f;[x]),f([x]))\leqslant\gamma||d[x]||_\infty,\quad \gamma\geqslant 0,$$

$$df([x])\leqslant\delta||d[x]||_\infty,\quad \delta\geqslant 0,$$

where the constants γ and δ depend on $[x]^0$ but not on $[x]$.

Theorem 1 states that if the interval arithmetic evaluation 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?

A first result in this respect shows why the interval arithmetic evaluation of the second expression in Example 1 is optimal:

Theorem 2 (Moore [64]). *Let a continuous function $f:D\subset\mathbb{R}^n\to\mathbb{R}$ be given by an expression $f(x)$ in which each variable x_i, $i=1,\dots,n$, occurs at most once. Then*

$$f([x])=R(f;[x])\quad \textit{for all } [x]\subseteq D.$$

Unfortunately, not many expressions $f(x)$ can be rearranged such that the assumptions of Theorem 2 are fulfilled. In order to propose an alternative we consider first a simple example.

Example 2. Let $f(x)=x-x^2$, $x\in[0,1]=[x]^0$.

It is easy to see that for $0\leqslant r\leqslant\frac{1}{2}$ and $[x]=[\frac{1}{2}-r,\frac{1}{2}+r]$ we have

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

and

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

From this it follows

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

and

$$df([x])\leqslant\delta d[x]\quad \text{with } \delta=2$$

in agreement with Theorem 1.

If we rewrite $f(x)$ as

$$x-x^2=\tfrac{1}{4}-(x-\tfrac{1}{2})(x-\tfrac{1}{2})$$

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

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

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

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

Theorem 3 (The centered form). *Let the function $f:D\subseteq\mathbb{R}^n\to\mathbb{R}$ be represented in the 'centered form'*

$$f(x)=f(z)+h(x)^{\mathrm{T}}(x-z) \tag{12}$$

for some $z\in[x]\subseteq[x]^0\subseteq D$ and $h(x)\in\mathbb{R}^n$. If

$$f([x])=f(z)+h([x])^{\mathrm{T}}([x]-z), \tag{13}$$

then

$$R(f;[x])\subseteq f([x]) \tag{14}$$

and (under some additional assumptions)

$$q(R(f;[x]),f([x]))\leqslant\kappa||d[x]||_\infty^2,\quad \kappa\geqslant 0, \tag{15}$$

where the constant κ depends on $[x]^0$ but not on $[x]$ and z.

Relation (15) is called 'quadratic approximation property' of the centered form. For rational functions it is not difficult to find a centered form, see for example [77].

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. Unfortunately, this is not the case as has been shown recently by Hertling [39]; see also [70].

Nevertheless, in special cases one can use the so-called generalized centered forms to get higher-order approximations of the range; see, e.g., [18]. Another interesting idea which uses a so-called 'remainder form of f' was introduced by Cornelius and Lohner [27].

Finally, we can apply the subdivision principle in order to improve the enclosure of the range. To this end we represent $[x]\in I(\mathbb{R}^n)$ as the union of k^n interval vectors $[x]^l$, $l=1,\dots,k^n$, such that $d[x_i]^l=d[x_i]/k$ for $i=1,\dots,n$ and $l=1,\dots,k^n$. Defining

$$f([x];k)=\bigcup_{l=1}^{k^n} f([x]^l), \tag{16}$$

the following result holds:

Theorem 4. *Let $f:D\subseteq\mathbb{R}^n\to\mathbb{R}$.*

(a) *With the notations and assumptions of Theorem* 1 *and with* (16) *we get*

$$q(R(f;[x]),f([x];k))\leqslant\frac{\hat{\gamma}}{k},$$

where $\hat{\gamma}=\gamma||d[x]^0||_\infty$.

(b) *Let the notations and assumptions of Theorem 3 hold. Then using in* (16) *for* $f([x]^l)$ *the expression* (13) with $z = z^l \in [x]^l$, $l = 1, \ldots, k$, *it follows that*

$$q(R(f;[x]), f([x];k)) \leqslant \frac{\hat{\kappa}}{k^2},$$

where $\hat{\kappa} = \kappa ||d[x]^0||_\infty^2$.

Theorem 4 shows that the range can be enclosed arbitrarily close if k tends to infinity, i.e., if the subdivision of $[x] \subseteq [x]^0$ is sufficiently fine, for details see, e.g., [78].

In passing we note that the principal results presented up to this point provide the basis for enclosing minimizers and minima in global optimization. Necessary refinements for practical algorithms in this respect can be found in, e.g., [36,37,38,42,44] or [79].

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 continuously 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

$$N[x] = m[x] - \frac{f(m[x])}{f'([x])}, \quad m[x] \in [x] \tag{17}$$

and consider the following iteration method:

$$[x]^{k+1} = N[x]^k \cap [x]^k, \quad k = 0, 1, 2, \ldots, \tag{18}$$

which is called interval Newton method.

Properties of operator (17) and method (18) are described in the following result.

Theorem 5. *Under the above assumptions the following holds for* (17) *and* (18):

(a) *If*

$$N[x] \subseteq [x] \subseteq [x]^0, \tag{19}$$

then f *has a zero* $x^* \in [x]$ *which is unique in* $[x]^0$.

(b) *If* f *has a zero* $x^* \in [x]^0$ *then* $\{[x]^k\}_{k=0}^\infty$ *is well defined,* $x^* \in [x]^k$ *and* $\lim_{k\to\infty} [x]^k = x^*$. *If* $df'([x]) \leqslant cd[x]$, $[x] \subseteq [x]^0$, *then* $d[x]^{k+1} \leqslant \gamma (d[x]^k)^2$.

(c) $N[x]^{k_0} \cap [x]^{k_0} = \emptyset$ (= *empty set*) *for some* $k_0 \geqslant 0$ *if and only if* $f(x) \neq 0$ *for all* $x \in [x]^0$.

Theorem 5 delivers two strategies to study zeros in $[x]^0$. By the first it is *proved* that f *has* a unique zero x^* in $[x]^0$. It is based on (a) and can be realized by performing (18) and checking (19) with $[x] = [x]^k$. By the second – based on (c) – it is proved that f *has no* zero x^* in $[x]^0$. While the second strategy is always successful if $[x]^0$ contains no zero of f the first one can fail as the

simple example $f(x)=x^2-4$, $[x]^0=[2,4]$ shows when choosing $m[x]^k > \underline{x}^k$. Here the iterates have the form $[x]^k=[2,a_k]$ with appropriate $a_k > 2$ while $\underline{N[x]^k} < 2$. Hence (19) can never be fulfilled.

In case (b), the diameters are converging quadratically to zero. On the other hand, if method (18) breaks down because of empty intersection after a finite number of steps then from a practical point of view it would be interesting to have qualitative knowledge about the size of k_0 in this case. This will be discussed in the next section in a more general setting.

4. Systems of nonlinear equations

In the present section we consider systems of nonlinear equations in the form

$$f(x)=0 \tag{20}$$

and

$$f(x)=x, \tag{21}$$

respectively, i.e., we look for zeros and for fixed points of f, respectively. (It is well known that problems (20) and (21) are equivalent when choosing f in (21) appropriately.) Using interval arithmetic we want to derive simple criteria which guarantee that a given interval $[x]$ contains at least one zero x^* of f or a corresponding fixed point. We also list conditions for x^* to be unique within $[x]$, and we show how $[x]$ can be improved iteratively to some vector $[x]^*$ which contains x^* and has a smaller diameter.

In the whole section we assume that $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ is at least continuous in D, and often we assume that it is at least once continuously (Fréchet-) differentiable.

We first consider fixed points x^* of f in $[x]\subseteq D$. A simple method for verifying such a point is based on (6)–(8) and Brouwer's fixed point theorem and reads as follows.

Theorem 6. *Let $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ be continuous and let*

$$f([x])\subseteq[x]\subseteq D. \tag{22}$$

Then f has at least one fixed point in $[x]$ and the iteration

$$[x]^0=[x],$$

$$[x]^{k+1}=f([x]^k),\quad k=0,1,\ldots \tag{23}$$

converges to some $[x]^*$ such that

$$[x]^*\subseteq[x]^{k+1}\subseteq[x]^k\subseteq\cdots\subseteq[x]^0=[x]. \tag{24}$$

The limit $[x]^*$ contains all fixed points of f in $[x]$.

We call an interval sequence $\{[x]^k\}_{k=0}^{\infty}$ monotonically decreasing if it fulfills (24).

Theorem 6 says nothing on the uniqueness of $x^*\in[x]$ nor on the width of $[x]^*$. In fact, the simple example $f(x)=-x$, $[x]=[-1,1]$ with $[x]^k=[x]^*=[x]$ shows that $d[x]^*>0$ can occur although $x^*=0$ is the only fixed point of f in $\mathbb{R}$. For P contractions, however, sharper results can be proved

by a direct application of Banach's fixed point theorem. Note that $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ is a P contraction on the set $I([x])$ of all compact intervals contained in $[x]\subseteq D$ if there is a matrix $P\geqslant O\in\mathbb{R}^{n\times n}$ with spectral radius $\rho(P)<1$ and

$$q(f([y]),f([z]))\leqslant Pq([y],[z]) \quad \text{for all } [y],\ [z]\subseteq[x].$$

Trivial examples are linear functions $f(x)=Ax-b$ with $D=\mathbb{R}^n$, $A\in\mathbb{R}^{n\times n}$, $\rho(|A|)<1$, $b\in\mathbb{R}^n$ and $P=|A|$.

Theorem 7. *Let $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ be a P contraction on $I([x])$, $[x]\subseteq D$, and let* (22) *hold. Then f has exactly one fixed point $x^*\in[x]$ and iteration* (23) *converges to x^* for all starting vectors $[x]^0\subseteq[x]$. Moreover, $x^*\in[x]^k$, $k=1,2,\ldots$, if $x^*\in[x]^0$ which holds, in particular, if $[x]^0=[x]$.*

Remark 1. Condition (22) can be omitted in Theorem 7 if f is a P contraction on the whole space $I(\mathbb{R}^n)$ (cf. [13]). For any $[x]^0\in I(\mathbb{R}^n)$ the unique fixed point x^* is then contained in $[-\underline{x}^0-\Delta,\bar{x}^0+\Delta]$, $\Delta=(I-P)^{-1}q([x]^1,[x]^0)$.

Remark 1 is interesting since it is not always an easy task to find an $[x]$ such that (22) holds. There is, however, a method of trial and error which goes back to Rump [81] and which, in practice, mostly ends up with such an $[x]$ in a few steps. The technique is called epsilon inflation and is a quite general interval arithmetic tool. It consists in replacing the current interval iterate by an interval vector which is a proper superset of the iterate and which differs from it by a small parameter ε. This can be done, e.g., in the following way: first compute an approximation $\tilde{x}$ of x^* by applying any appropriate standard method in numerical analysis. Then iterate according to

$$[x]^0=\tilde{x},$$

$$[x]^{k+1}=f([x]^k+d[x]^k\,[-\varepsilon,\varepsilon]+[-\eta,\eta]e),\quad k=0,1,\ldots, \tag{25}$$

where ε, η are some small positive real numbers. If f is a P contraction on $I(\mathbb{R}^n)$ then (25) ends up after finitely many steps with an iterate which fulfills (22). This is stated in our next theorem.

Theorem 8. *Let $f:D=\mathbb{R}^n\to\mathbb{R}^n$ be a P contraction on $I(\mathbb{R}^n)$. With $[x]^0_\varepsilon$ being given, iterate by inflation according to*

$$[x]^{k+1}_\varepsilon=f([x]^k_\varepsilon])+[\delta]^k,\quad k=0,1,\ldots,$$

where $[\delta]^k\in I(\mathbb{R}^n)$ are given vectors which converge to some limit $[\delta]$. If $0\in\operatorname{int}([\delta])$ then there is an integer $k_0=k_0([x]^0_\varepsilon)$ such that

$$f([x]^{k_0}_\varepsilon)\subseteq\operatorname{int}([x]^{k_0}_\varepsilon).$$

In view of (25) we can try to apply Theorem 8 with $[\delta]^k=(df[x]^k_\varepsilon)[-\varepsilon,\varepsilon]+[-\eta,\eta]e$ and $[x]^0_\varepsilon=[x]^0+(d[x]^0)[-\varepsilon,\varepsilon]+[-\eta,\eta]e$. If $[\delta]=\lim_{k\to\infty}[\delta]^k$ exists then $0\in\operatorname{int}([\delta])$ since $0\in[-\eta,\eta]e\subseteq[\delta]^k$ for $k=0,1,\ldots$.

Theorem 8 was originally stated and proved by Rump [83] for linear functions f. It was generalized to P contractions and contractive interval functions in [58,59] where also the case $D\neq\mathbb{R}^n$

is considered and where various examples for epsilon inflations are presented. Unfortunately, Theorem 8 says nothing on the number of steps which are needed to succeed with (22). Therefore, other possibilities become interesting which we are going to present in the second part of this section and in Section 6.

We consider now zeros of a given function f.

A first method is based on a result of C. Miranda (see [62] or Corollary 5.3.8 in [69]) which is equivalent to Brouwer's fixed point theorem. We use it in the following modified interval version.

Theorem 9. *Let* $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *be continuous and let* $[x] \subseteq D$,

$$[l]^i = ([x_1], \ldots, [x_{i-1}], \underline{x}_i, [x_{i+1}], \ldots, [x_n])^{\mathrm{T}},$$

$$[u]^i = ([x_1], \ldots, [x_{i-1}], \bar{x}_i, [x_{i+1}], \ldots, [x_n])^{\mathrm{T}}.$$

If $\overline{f_i([l]^i)} \leqslant 0$, $\underline{f_i([u]^i)} \geqslant 0$ *or* $\underline{f_i([l]^i)} \geqslant 0$, $\overline{f_i([u]^i)} \leqslant 0$ *holds for each* $i = 1, \ldots, n$ *then* f *has at least one zero in* $[x]$.

Combined with subdivisions, lists and exclusion techniques Theorem 9 forms the basis of a simple but efficient verification and enclosure method for zeros of functions $f:D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ even if $m < n$. Curves and surfaces can thus be tightly enclosed and problems in CAGD like ray tracing can be handled. We refer to [31,52,68].

Another method for verifying zeros consists in generalizing the interval Newton method of Section 3 to the multidimensional case. To this end we denote by

$$\mathrm{IGA}([A], [b]),$$

the result of the Gaussian algorithm applied formally to a nonsingular interval matrix $[A] \in I(\mathbb{R}^{n \times n})$ and an interval vector $[b] \in I(\mathbb{R}^n)$, see, for example, [13, Section 15]. Here we assumed that no division by an interval which contains zero occurs in the elimination process. It is easy to see that

$$S = \{x = A^{-1}b \mid A \in [A],\ b \in [b]\} \subseteq \mathrm{IGA}([A], [b]) \tag{26}$$

holds. By

$$\mathrm{IGA}([A])$$

we denote the interval matrix whose ith column is obtained as $\mathrm{IGA}([A], e^i)$ where e^i is the ith unit vector. In other words, $\mathrm{IGA}([A])$ is an enclosure for the inverses of all matrices $A \in [A]$.

Now assume that

$$f:D \subset \mathbb{R}^n \to \mathbb{R}^n \tag{27}$$

is continuously differentiable. If $x, y \in [x] \subseteq D$ then

$$f(x) - f(y) = J(y,x)(x - y), \tag{28}$$

where

$$J(y,x) = \int_0^1 f'(y + t(x - y))\,\mathrm{d}t. \tag{29}$$

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

$$J(y,x)\in f'([x]), \tag{30}$$

where $f'([x])$ denotes the interval arithmetic evaluation of the Jacobian of f. For fixed $y\in[x]$ we obtain from (28) and (30)

$$p(x)=x-J^{-1}(y,x)f(x)=y-J^{-1}(y,x)f(y)\in y-\ \mathrm{IGA}(f'([x]),f(y)). \tag{31}$$

If $x\in[x]$ is a zero of f then (31) implies $x\in y-\ \mathrm{IGA}(f'([x]),f(y))$. This leads to the following definition of the interval Newton operator $N[x]$ which we introduce in analogy to (18): suppose that $m[x]\in[x]$ is a real vector. Then

$$N[x]=m[x]-\ \mathrm{IGA}(f'([x]),f(m[x])). \tag{32}$$

The interval Newton method is defined by

$$[x]^{k+1}=N[x]^k\cap[x]^k,\quad k=0,1,2,\ldots\ . \tag{33}$$

Analogously to Theorem 5 we have the following result.

Theorem 10. *Let $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ be continuously differentiable and assume that* $\mathrm{IGA}(f'([x]^0))$ *exists for some interval vector $[x]^0\subseteq D$: (This is identical to assuming that the Gaussian algorithm is feasible for $f'([x]^0)$. In particular, $f'([x]^0)$ is nonsingular in this case.)*

(a) *If*

$$N[x]\subseteq[x]$$

for some $[x]\subseteq[x]^0$ then f has a zero x^* in $[x]$ which is unique even in $[x]^0$.

Assume that

$$\rho(A)<1,\quad \text{where } A=|I-\mathrm{IGA}(f'([x]^0))f'([x]^0)|. \tag{34}$$

(b) *If f has a zero x^* in $[x]^0$ then the sequence $\{[x]^k\}_{k=0}^{\infty}$ defined by (33) is well defined, $x^*\in[x]^k$ and $\lim_{k\to\infty}[x]^k=x^*$. In particular, $\{[x]^k\}_{k=0}^{\infty}$ is monotonically decreasing and x^* is unique in $[x]^0$.*

Moreover, if

$$df'([x])_{ij}\leqslant\alpha||d[x]||_\infty,\quad \alpha\geqslant 0,\quad 1\leqslant i,j\leqslant n \tag{35}$$

for all $[x]\subseteq[x]^0$ then

$$||d[x]^{k+1}||_\infty\leqslant\gamma||d[x]^k||_\infty^2,\quad \gamma\geqslant 0. \tag{36}$$

(c) *$N[x]^{k_0}\cap[x]^{k_0}=\emptyset$ for some $k_0\geqslant 0$ if and only if $f(x)\neq 0$ for all $x\in[x]^0$.*

The proof of (a) can be quickly done by applying Brouwer's fixed point theorem to p of (31) The results of (b) and (c) can be found in [9].

Note that in contrast to the onedimensional case we need condition (34) in cases (b) and (c).

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 $A=O$ in the limit case $d[x]^0=0$. Schwandt

[86] has discussed a simple example in the case $\rho(A)\geqslant 1$ which shows that for a certain interval vector (33) is feasible, $x^*\in[x]^k$, but $\lim_{k\to\infty}[x]^k\neq x^*$.

In case (a) of the preceding theorem we have by (36) quadratic convergence of the diameters of the enclosing intervals to the zero vector. This is the same favorable behavior 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 (33) 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:

Let $N[x]=[\underline{n},\bar{n}]$ for the interval Newton operator (32). It is easy to prove that

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

if and only if for at least one component i_0 either

$$(\bar{n}-\underline{x})_{i_0}<0 \tag{37}$$

or

$$(\bar{x}-\underline{n})_{i_0}<0 \tag{38}$$

holds. Furthermore, it can be shown that

$$\bar{x}-\underline{n}\leqslant \mathrm{O}(||d[x]||_\infty^2)e+A^2 f(\bar{x}) \tag{39}$$

and

$$\bar{n}-\underline{x}\leqslant \mathrm{O}(||d[x]||_\infty^2)e-A^1 f(\underline{x}) \tag{40}$$

provided (35) holds. Here A^1 and A^2 are two real matrices contained in $\mathrm{IGA}(f'([x]^0))$. 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

$$(A^1 f(\underline{x}))_{i_0}\neq 0 \tag{41}$$

and

$$\mathrm{sign}(A^1 f(\underline{x}))_{i_0}=\mathrm{sign}(A^2 f(\bar{x}))_{i_0}. \tag{42}$$

Assume now that $\mathrm{sign}(A^1 f(\underline{x}))_{i_0}=1$. Then for sufficiently small diameter $d[x]$ we have $(\bar{n}-\underline{x})_{i_0}<0$ by (40) and by (37) the intersection becomes empty. If $\mathrm{sign}(A^1 f(\underline{x}))_{i_0}=-1$ then by (39) we obtain $(\bar{x}-\underline{n})_{i_0}<0$ for sufficiently small $d[x]$ and by (38) 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 $\mathrm{O}(||d[x]||_\infty^2)$ in (39) and (40) 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 behavior by a simple one-dimensional 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 $[x]^0=[-5,6]$. Unfortunately, (18) cannot 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 (18) can be performed for each of these intervals. If such a subinterval contains a zero then (a) of Theorem 5 holds, otherwise (b) is true. Table 1 contains the intervals which were obtained by applying the above-mentioned modification of the 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).

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 approximately 2.75, which is not small, and after only 3 steps the intersection becomes empty. The intervals with the numbers $n = 1, 2, 3, 6, 8$ 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 [8].

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($f'([x]^0)$, $f(m[x]^0)$) can in general only be computed if $d[x]^0$ is sufficiently small. For this reason Krawczyk [48] had the idea to introduce a mapping which today is called the Krawczyk operator:

Assume again that a mapping (27) with the corresponding properties is given. Then analogously to (32) we consider the so-called Krawczyk operator

$$K[x] = m[x] - Cf(m[x]) + (I - Cf'([x]))([x] - m[x]), \tag{43}$$

Table 1
The modified interval Newton method applied to f from Example 3

n	
1	$[-0.356 \cdot 10^1, \ -0.293 \cdot 10^1]$
2	$[-0.141 \cdot 10^1, \ -0.870 \cdot 10^0]$
3	$[-0.977 \cdot 10^0, \ 0.499 \cdot 10^0]$
4	$[0.501 \cdot 10^0, \ 0.633 \cdot 10^0]$
5	$[0.140 \cdot 10^1, \ 0.185 \cdot 10^1]$
6	$[0.188 \cdot 10^1, \ 0.212 \cdot 10^1]$
7	$[0.265 \cdot 10^1, \ 0.269 \cdot 10^1]$
8	$[0.297 \cdot 10^1, \ 0.325 \cdot 10^1]$
9	$[0.327 \cdot 10^1, \ 0.600 \cdot 10^1]$

Table 2
The interval Newton method applied to f from Example 3

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

where C is a nonsingular real matrix and where $m[x] \in [x]$. For fixed C we define the so-called Krawczyk method by

$$[x]^{k+1} = K[x]^k \cap [x]^k, \quad k = 0, 1, 2, \ldots . \tag{44}$$

For this method an analogous result holds as was formulated for the interval Newton method in Theorem 10:

Theorem 11. *Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable and assume that the interval arithmetic evaluation $f'([x]^0)$ of the Jacobian exists for some interval vector $[x]^0 \subseteq D^0$.*

(a) *If*

$$K[x] \subseteq [x] \tag{45}$$

for some $[x] \subseteq [x]^0$ then f has a zero x^* in $[x]$.

If (45) is slightly sharpened to

$$(K[x])_i \subset [x_i] \subseteq [x_i]^0 \quad \text{for } i = 1, \ldots, n, \tag{46}$$

then $\rho(|I - Cf'([x])|) < 1$ holds, $f'([x])$ is nonsingular and x^* is unique in $[x]$.

Let $m[x]$ be the center of $[x]$ and assume that

$$\rho(B) < 1 \quad \text{where } B = |I - Cf'([x]^0)|. \tag{47}$$

(b) *If f has a zero x^* in $[x]^0$ then the sequence $\{[x]^k\}_{k=0}^{\infty}$ defined by (44) is well defined, $x^* \in [x]^k$ and $\lim_{k\to\infty}[x]^k = x^*$. In particular, $\{[x]^k\}_{k=0}^{\infty}$ is monotonically decreasing and x^* is unique in $[x]^0$. Moreover, if $C = C_k$ varies with k such that it is the inverse of some matrix from $f'([x]^k)$, and if*

$$df'([x])_{ij} \leqslant \alpha ||d[x]||_\infty, \quad \alpha \geqslant 0, \quad 1 \leqslant i, j \leqslant n \tag{48}$$

for all $[x] \subseteq [x]^0$ then

$$||d[x]^{k+1}||_\infty \leqslant \gamma ||d[x]^k||_\infty^2, \quad \gamma \geqslant 0. \tag{49}$$

(c) *$K[x]^{k_0} \cap [x]^{k_0} = \emptyset$ for some $k_0 \geqslant 0$ if and only if $f(x) \neq 0$ for all $x \in [x]^0$.*

Proof. (a) Consider for the nonsingular matrix C in $K[x]$ the continuous mapping

$$g : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$$

defined by

$$g(x) = x - Cf(x).$$

It follows, using (28) and the assumption,

$$\begin{aligned} g(x) &= x - Cf(x) \\ &= x - C(f(x) - f(m[x])) - Cf(m[x]) \\ &= m[x] + (x - m[x]) - CJ(m[x], x)(x - m[x]) - Cf(m[x]) \\ &\in m[x] - Cf(m[x]) + (I - Cf'([x]))([x] - m[x]) \\ &= K[x] \subseteq [x], \quad x \in [x]. \end{aligned}$$

By Brouwer's fixed point theorem g has a fixed point $x^* \in [x]$. This fixed point is a zero of f.

If (45) is replaced by (46) then $|I - Cf'([x])|d[x] \leqslant dK[x] < d[x]$. Therefore,

$$\max_{1 \leqslant i \leqslant n} \frac{\sum_{j=1}^{n} |I - Cf'([x])|_{ij} d[x_j]}{d[x_i]} < 1$$

which is equivalent to

$$||\hat{D}^{-1}|I - Cf'([x])|\hat{D}||_\infty < 1.$$

Here, $\hat{D}$ is the diagonal matrix with $\hat{d}_{ii} = d[x_i]$, $i = 1, \ldots, n$. Therefore,

$$\rho(|I - Cf'([x])|) = \rho(\hat{D}^{-1}|I - Cf'([x])|\hat{D}) \leqslant ||\hat{D}^{-1}|I - Cf'([x])|\hat{D}||_\infty < 1.$$

If $f'([x])$ contained a singular matrix A then $I - CA$ would have the eigenvalue 1 and we would get the contradiction

$$1 \leqslant \rho(I - CA) \leqslant \rho(|I - CA|) \leqslant \rho(|I - Cf'([x])|) < 1. \tag{50}$$

Therefore, $f'([x])$ is nonsingular. If f had two zeros x^*, $y^* \in [x]$ then (28) and (30) would imply $x^* = y^*$.

(b) By (28) we have

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

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

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

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 $x^* \in [x]^k$, $k \geqslant 0$.

For the diameters of the sequence $\{[x]^k\}_{k=0}^{\infty}$ we have $d[x]^{k+1} \leqslant dK[x]^k \leqslant Bd[x]^k$, where the last inequality holds because we assumed that $m[x]^k$ is the center of $[x]^k$. Since $\rho(B) < 1$ we have $\lim_{k \to \infty} d[x]^k = 0$, and from $x^* \in [x]^k$ it follows $\lim_{k \to \infty} [x]^k = x^*$. In particular, x^* is unique within $[x]^0$.

Analogously to (a) assumption (47) implies that $f'([x^0])$ is nonsingular. Since it is compact and since the inverse of a matrix $M \in \mathbb{R}^{n \times n}$ depends continuously on the entries of M the set $\{|M^{-1}| \mid M \in f'([x]^0)\}$ is bounded by some matrix $\hat{C}$. The quadratic convergence behavior (49) follows now from

$$\begin{aligned} d[x]^{k+1} &\leqslant |I - C_k f'([x]^k)|d[x]^k \\ &\leqslant |C_k||C_k^{-1} - f'([x]^k)|d[x]^k \\ &\leqslant \hat{C}|f'([x]^k) - f'([x]^k)|d[x]^k \\ &= \hat{C}\, df'([x]^k)d[x]^k \end{aligned}$$

by using (48).

(c) Assume now that $K[x]^{k_0} \cap [x]^{k_0} = \emptyset$ for some $k_0 \geqslant 0$. 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 $x^* \in [x]^k$, $k \geqslant 0$.

If on the other hand $f(x) \neq 0$ and $K[x]^k \cap [x]^k \neq \emptyset$ then $\{[x]^k\}$ is well defined. Because of $\rho(B) < 1$ 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 (44)

$$\hat{x} = K\hat{x} \cap \hat{x} = K\hat{x} = \hat{x} - Cf(\hat{x}).$$

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 11. □

Remark 2. (a) When we defined the Krawczyk operator in (43) we required C to be nonsingular. We need not know this in advance if (45) or (47) holds since either of these two conditions implies the nonsingularity by an analogous argument as in the proof for (a).

(b) It is easy to see that in case (a) of the preceding theorem all the zeros x^* of f in $[x]$ are even in $K[x]$.

(c) If $m[x]$ is not the center of $[x]$ but still an element of it the assertions in (b), (c) remain true if (47) is replaced by $\rho(B) < \frac{1}{2}$.

(d) Assertion (47) certainly holds if (34) is true with $C \in \mathrm{IGA}(f'([x]^0))$.

In case (c) of the Theorem 11, 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 under the following assumptions:

(i) $f'([x]^0)$ is nonsingular,

(ii) (48) holds,

(iii) $C = C_k$ varies with k such that it is the inverse of some matrix from $f'([x]^k)$.

Note that these assumptions certainly hold if the assumptions for (49) are fulfilled.

As for the interval Newton operator we write $K[x] = [\underline{k}, \bar{k}]$. Now $K[x] \cap [x] = \emptyset$ if and only if

$$(\bar{x} - \underline{k})_{i_0} < 0 \tag{51}$$

or

$$(\bar{k} - \underline{x})_{i_0} < 0 \tag{52}$$

for at least one $i_0 \in \{1, 2, \ldots, n\}$. (Compare with (37) and (38).)

We first prove that for $K[x]$ defined by (43) we have the vector inequalities

$$\bar{x} - \underline{k} \leqslant \mathrm{O}(||d[x]||_\infty^2)e + Cf(\bar{x}) \tag{53}$$

and

$$\bar{k} - \underline{x} \leqslant \mathrm{O}(||d[x]||_\infty^2)e - Cf(\underline{x}), \tag{54}$$

where again $e = (1, 1, \ldots, 1)^{\mathrm{T}} \in \mathbb{R}^n$.

We prove (54). For $[x] \subseteq [x]^0$ let $f'([x]) = [\underline{F}', \bar{F}']$ and set $C = \hat{M}^{-1}$ with some matrix $\hat{M} \in f'([x])$. An easy computation shows that

$$I - Cf'([x]) = C[\hat{M} - \bar{F}', \hat{M} - \underline{F}'] \subseteq |C|[\underline{F}' - \bar{F}', \bar{F}' - \underline{F}'] \subseteq [-1, 1]\hat{C}\, df'([x]),$$

where $\hat{C}$ is any upper bound for the set $\{|M^{-1}| \mid M \in f'([x]^0)\}$. Therefore

$$K[x] \subseteq m[x] - Cf(m[x]) + [-1,1]\hat{C}\, df'([x]) \cdot |[x] - m[x]|.$$

Hence,

$$\begin{aligned}\bar{k} - \underline{x} &\leqslant m[x] - \underline{x} - Cf(m[x]) + \hat{C}\, df'([x])d[x] \\ &\leqslant \tfrac{1}{2}d[x] - Cf(m[x]) + \mathrm{O}(||d[x]||_\infty^2)e,\end{aligned}$$

where we have used (48) and $m[x] \in [x]$.

Choosing $x = m[x]$, $y = \underline{x}$ in (28) we obtain

$$f(m[x]) - f(\underline{x}) = J(\underline{x}, m[x])(m[x] - \underline{x}).$$

It follows that

$$\begin{aligned}\bar{k} - \underline{x} &\leqslant \tfrac{1}{2}d[x] - Cf(\underline{x}) - \tfrac{1}{2}CJ(\underline{x}, m[x])d[x] + \mathrm{O}(||d[x]||_\infty^2)e \\ &= \tfrac{1}{2}(I - CJ(\underline{x}, m[x]))d[x] - Cf(\underline{x}) + \mathrm{O}(||d[x]||_\infty^2)e.\end{aligned}$$

Since

$$I - CJ(\underline{x}, m[x]) = C(C^{-1} - J(\underline{x}, m[x])) \in \hat{C}(f'([x]) - f'([x])) = \hat{C}df'([x]),$$

the assertion follows by applying (48).

The second inequality can be shown in the same manner, hence (53) and (54) 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

$$(Cf(\underline{x}))_{i_0} \neq 0 \tag{55}$$

and

$$\operatorname{sign}(Cf(\bar{x}))_{i_0} = \operatorname{sign}(Cf(\underline{x}))_{i_0}. \tag{56}$$

This can be seen as follows: Since $\underline{x} \in [x]$ we have $f(\underline{x}) \neq 0$ and since C is nonsingular it follows that $Cf(\underline{x}) \neq 0$ and therefore $(Cf(\underline{x}))_{i_0} \neq 0$ for at least one $i_0 \in \{1, 2, \ldots, n\}$ which proves (55). Using again (28) with $x = \bar{x}$, $y = \underline{x}$ we get

$$f(\bar{x}) - f(\underline{x}) = J(\underline{x}, \bar{x})(\bar{x} - \underline{x}).$$

It follows

$$Cf(\bar{x}) = Cf(\underline{x}) + CJ(\underline{x}, \bar{x})(\bar{x} - \underline{x}).$$

Since the second term on the right-hand side approaches zero if $d[x] \to 0$ we have (56) for sufficiently small diameter $d[x]$.

Using (53), (54) together with (55) and (56) 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 (41), (42) together with (39) and (40). By the same motivation as for the interval Newton method we denote this behavior as 'quadratic divergence' of the Krawczyk method.

Part (a) of the two preceding theorems 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 errorbounds 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

$$x^{k+1}=x^k-f'(x^k)^{-1}f(x^k),\quad k=0,1,\dots\ . \tag{57}$$

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 estimation in conjunction with the quadratic convergence property (36) which we have also proved (under mild additional assumptions) for the Krawczyk method will lead us to a test interval which can be computed using only iterates of Newton's method.

Theorem 12 (See Ortega and Rheinboldt, [71, Theorem 12.6.2]). *Assume that $f:D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ is differentiable in the ball $\{x\,|\,||x-x^0||_\infty\leqslant r\}$ and that*

$$||f'(x)-f'(y)||_\infty\leqslant L||x-y||_\infty$$

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

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

and assume that

$$h_0=B_0\eta_0L\leqslant\frac{1}{2},\quad r_0=\frac{1-\sqrt{1-2h_0}}{h_0}\eta_0\leqslant r.$$

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

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

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

$$||x^*-x^k||_\infty\leqslant\frac{1}{2^{k-1}}(2h_0)^{2^k-1}\eta_0,\quad k\geqslant0 \tag{58}$$

holds.

Since $h_0\leqslant\frac{1}{2}$, the error estimate (58) (for $k=0,1$ and the ∞-norm) leads to

$$||x^*-x^0||_\infty\leqslant2\eta_0=2||x^1-x^0||_\infty,$$

$$||x^*-x^1||_\infty\leqslant2h_0\eta_0\leqslant\eta_0=||x^1-x^0||_\infty.$$

This suggests a simple construction of an interval vector containing the solution x^*. 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 $\{x\,|\,||x-x^1||_\infty\leqslant\eta_0\}$ as starting vector for Newton's method. Because of (36) and since $x^*\in K[x]$ it is reasonable to assume that

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

for

$$[x]=\{x\,|\,||x-x^1||_\infty\leqslant\eta_0\}. \tag{59}$$

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–Kantorovich 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 (59) with x^{k+1} instead of x^1. 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

$$[x] = \{x \in \mathbb{R}^n \mid ||x^{k+1} - x||_\infty \leqslant \eta_k\},$$

$$\eta_k = ||x^{k+1} - x^k||_\infty \tag{60}$$

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

$$\frac{||d[y]||_\infty}{||x^{k+1}||_\infty} \leqslant \text{eps} \tag{61}$$

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

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

Since $||x^*||_\infty$ differs only slightly from $||x^{k+1}||_\infty$ if x^{k+1} is near x^*, condition (61) guarantees that the relative error with which any $y \in [y]$ approximates x^* is close to machine precision. Using (35) it can be shown that

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

and

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

where $\tilde{L} = \max\{\hat{L}, L\}$, and since $||d[x]||_\infty = 2\eta_k$ the inequality (61) holds if

$$4\frac{||f'(x^k)^{-1}||_\infty \tilde{L}\eta_k^2}{||x^{k+1}||_\infty} \leqslant \text{eps} \tag{62}$$

is true.

From Newton's method we have

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

and by 3.2.12 in [71] it follows that

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

Replacing the inequality sign by equality in this relation and eliminating $||f'(x^k)^{-1}||_\infty \tilde{L}$ in (62) we get the following stopping criterion for Newton's method:

$$\frac{8\eta_k^3}{||x^{k+1}||_\infty \eta_{k-1}^2} \leqslant \text{eps}. \tag{63}$$

Of course, this is not a mathematical proof that if (63) 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 [11] the test based on the stopping criterion (63) works extremely well in practice.

Some of the ideas of this section have been generalized to nonsmooth mappings by Chen [24].

Nonlinear interval systems, i.e., systems of nonlinear equations with parameter-dependent input data, have been considered, e.g., in [58].

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 [5], see also [75].

5. Systems of linear equations

Given $[A] \in \mathrm{I}(\mathbb{R}^{n\times n})$, $[b] \in \mathrm{I}(\mathbb{R}^n)$ we want to characterize and to enclose the solution set

$$S = \{x \in \mathbb{R}^n \mid Ax = b,\ A \in [A],\ b \in [b]\} \tag{64}$$

and the symmetric solution set

$$S_{\text{sym}} = \{x \in \mathbb{R}^n \mid Ax = b,\ A = A^{\mathrm{T}} \in [A] = [A]^{\mathrm{T}},\ b \in [b]\}. \tag{65}$$

These sets occur when dealing with systems of linear equations whose input data are afflicted with tolerances (cf., e.g. [13,69] or [84]). This is the case when data $\check{A} \in \mathbb{R}^{n\times n}$, $\check{b} \in \mathbb{R}^n$ are perturbed by errors caused, e.g., by measurements or by a conversion from decimal to binary digits on a computer. Assume that these errors are known to be bounded by some quantities $\Delta A \in \mathbb{R}^{n\times n}$ and $\Delta b \in \mathbb{R}^n$ with nonnegative entries. Then it seems reasonable to accept a vector $\tilde{x}$ as the 'correct' solution of $\check{A}x = \check{b}$ if it is in fact the solution of a perturbed system $\tilde{A}x = \tilde{b}$ with

$$\tilde{A} \in [A] = [\check{A} - \Delta A, \check{A} + \Delta A], \qquad \tilde{b} \in [b] = [\check{b} - \Delta b, \check{b} + \Delta b].$$

The characterization of all such $\tilde{x}$ led Oettli and Prager [72] to statements (a) and (b) of the following theorem.

Theorem 13. *For* $[A] \in \mathrm{I}(\mathbb{R}^{n\times n})$, $[b] \in \mathrm{I}(\mathbb{R}^n)$ *the following properties are equivalent*:

(a) $x \in S$;

(b) $|\check{A}x - \check{b}| \leqslant \frac{1}{2}(d([A])|x| + d([b]))$;

(c) $[A]x \cap [b] \neq \emptyset$;

(d)

$$\left\{ \begin{array}{l} \underline{b}_i - \sum_{j=1}^{n} a_{ij}^{+} x_j \leqslant 0 \\ -\bar{b}_i + \sum_{j=1}^{n} a_{ij}^{-} x_j \leqslant 0 \end{array} \right\}, \quad i = 1, \ldots, n,$$

where a_{ij}^{-} *and* a_{ij}^{+} *are determined by the equality*

$$[\underline{a}_{ij}, \bar{a}_{ij}] = \begin{cases} [a_{ij}^{-}, a_{ij}^{+}] & \text{if } x_j \geqslant 0, \\ [a_{ij}^{+}, a_{ij}^{-}] & \text{if } x_j < 0. \end{cases}$$

The inequality in (b) relates the midpoint residual to the diameters of $[A]$ and $[b]$, (c) is a short interval version of (b) due to Beeck [22] and (d) characterizes S in each orthant as intersection of finitely many half spaces. This last property shows, in particular, that S cannot easily be described. Therefore, one often encloses S by an interval vector $[x]$. According to (26) such a vector can be computed, e.g., by the Gaussian algorithm performed with the interval data as in Section 4. It is an open question to find necessary and sufficient conditions for the feasibility of the Gaussian elimination process if $[A]$ contains nondegenerate entries. For instance, IGA($[A], [b]$) exists if $\langle [A] \rangle$ is an M matrix as was shown in [4]. Other sufficient conditions can be found in [13,55,60]. See also the references there.

Iterative methods can also be used for enclosing S. Two simple ones are the interval Jacobi method

$$[x_i]^{k+1} = \left([b_i] - \sum_{\substack{j=1 \\ j \neq i}}^{n} [a_{ij}][x_j]^k \right) \Big/ [a_{ii}], \quad i = 1, \ldots, n \tag{66}$$

and the interval Gauss–Seidel method

$$[x_i]^{k+1} = \left([b_i] - \sum_{j=1}^{i-1} [a_{ij}][x_j]^{k+1} - \sum_{j=i+1}^{n} [a_{ij}][x_j]^k \right) \Big/ [a_{ii}], \quad i = 1, \ldots, n \tag{67}$$

with $0 \notin [a_{ii}]$ for $i = 1, \ldots, n$. They can be modified by intersecting the right-hand sides of (66) and (67) with $[x_i]^k$ before assigning it to $[x_i]^{k+1}$.

Denote by $[D]$, $-[L]$ and $-[U]$, respectively, the diagonal part, the strictly lower triangular part and the strictly upper triangular part of $[A]$, respectively. Then $[A] = [D] - [L] - [U]$, and the unmodified methods can be written in the form

$$[x]^{k+1} = f([x]^k) \quad \text{with } f([x]) = \mathrm{IGA}([M], [N][x] + [b]), \tag{68}$$

where $[A] = [M] - [N]$ and where we assume that IGA($[M]$) exists. For $[M] = [D]$ we recover the Jacobi method (66) and for $[M] = [D] - [L]$ the Gauss–Seidel method (67). The following result holds for these two cases and for a slight generalization concerning the shape of $[M]$:

Theorem 14. *Let* $[A]=[M]-[N]\in \mathrm{I}(\mathbb{R}^{n\times n})$, $[b]\in \mathrm{I}(\mathbb{R}^n)$ *with* $[M]$ *being a nonsingular lower triangular interval matrix*:

(a) *Iteration* (68) *is equivalent to the iteration*

$$[x_i]^{k+1}=\left([b_i]-\sum_{j=1}^{i-1}[m_{ij}][x_j]^{k+1}+\sum_{j=1}^{n}[n_{ij}][x_j]^k\right)\Big/[m_{ii}],\quad i=1,\dots,n. \tag{69}$$

(b) *Iteration* (68) *is convergent to some limit* $[x]^*\in \mathrm{I}(\mathbb{R}^n)$ (*i.e., each sequence* $\{[x]^k\}_{k=0}^{\infty}$ *of iterates defined by* (68) *is convergent to* $[x]^*$) *if and only if* $\rho(\langle[M]\rangle^{-1}|[N]|)<1$.
In this case $S\subseteq[x]^*$.

(c) *If* $[A]$ *and* $[M]$ *are* M *matrices and if* $\underline{N}\geqslant O$ *then* $\rho(\langle[M]\rangle^{-1}|[N]|)=\rho(\underline{M}^{-1}\bar{N})<1$ *and* $[x]^*$ *from* (*b*) *is the hull of* S.

(d) *Let* $[x]\in \mathrm{I}(\mathbb{R}^n)$. *If* $f([x])$ *from* (68) *satisfies* $(f([x]))_i\subset[x_i]$ *for* $i=1,\dots,n$, *then* $\rho(\langle[M]\rangle^{-1}|[N]|)<1$.

Proof. (a) follows by induction with respect to i taking into account that for lower triangular matrices the ith elimination step of the Gaussian algorithm changes only the ith column of $[A]$.

(b) Let $P=\langle[M]\rangle^{-1}|[N]|$. Since $[M]$ is triangular, $\langle[M]\rangle$ is an M matrix, hence $P\geqslant O$.

'$\Rightarrow$': From (69) we get

$$d[x_i]^{k+1}\geqslant\left(\sum_{j=1}^{i-1}|m_{ij}|d[x_j]^{k+1}+\sum_{j=1}^{n}|n_{ij}|d[x_j]^k\right)\Big/\langle[m_{ii}]\rangle,\quad i=1,\dots,n, \tag{70}$$

which is equivalent to $\langle[M]\rangle d[x]^{k+1}\geqslant|[N]|d[x]^k$. From this, $d[x]^{k+1}\geqslant Pd[x]^k$, and, by induction, $d[x]^k\geqslant P^k d[x]^0$ follow. Choose $[x]^0$ such that $d[x]^0$ is a Perron vector for P with $d[x_{i_0}]^*<d[x_{i_0}]^0$ for some index i_0. If $\rho(P)\geqslant 1$ then

$$d[x_{i_0}]^k\geqslant\rho(P)^k d[x_{i_0}]^0\geqslant d[x_{i_0}]^0>d[x_{i_0}]^*$$

and $k\to\infty$ yields to a contradiction.

'$\Leftarrow$': Let $f([x])=\mathrm{IGA}([M],[N][x]+[b])$. From (69) we get

$$q(f([x]),f([y]))_i\leqslant\frac{1}{\langle[m_{ii}]\rangle}\left(\sum_{j=1}^{i-1}|[m_{ij}]|q(f([x_j]),f([y_j]))+\sum_{j=1}^{n}|[n_{ij}]|q([x_j],[y_j])\right),$$
$$i=1,\dots,n,$$

whence $\langle[M]\rangle q(f([x]),f([y]))\leqslant|[N]|q([x],[y])$ and $q(f([x]),f([y]))\leqslant Pq([x],[y])$. Hence f is a P contraction, and Theorem 7 together with Remark 1 proves the convergence.

Let now (68) be convergent for all $[x]^0$ and choose $\tilde{x}\in S$. There are $\tilde{A}\in[A]$, $\tilde{b}\in[b]$, $\tilde{M}\in[M]$, $\tilde{N}\in[N]$ such that $\tilde{A}\tilde{x}=\tilde{b}$, $\tilde{A}=\tilde{M}-\tilde{N}$ and $\tilde{x}=\tilde{M}^{-1}(\tilde{N}\tilde{x}+\tilde{b})$. Then $\tilde{x}\in\mathrm{IGA}([M],[N]\tilde{x}+[b])$. Start (68) with $[x]^0=\tilde{x}$. Then $\tilde{x}\in[x]^k$ for $k=0,1,\dots$, hence $\tilde{x}\in[x]^*$. This proves $S\subseteq[x]^*$.

(c) The assumptions imply that $\underline{A}=\underline{M}-\bar{N}$ is a regular splitting of $\underline{A}$ and that $\underline{A}^{-1}\geqslant O$. Therefore, 2.4.17 in [71] guarantees $\rho(\langle[M]\rangle^{-1}|[N]|)=\rho(\underline{M}^{-1}\bar{N})<1$.

In order to prove the hull property let $[x]^*$ be the limit of (68), define

$$m_{ij}^*=\begin{cases}\underline{m}_{ij} & \text{if } \underline{x}_j^*\leqslant 0,\\ \bar{m}_{ij} & \text{if } \underline{x}_j^*>0,\end{cases}\qquad n_{ij}^*=\begin{cases}\bar{n}_{ij} & \text{if } \underline{x}_j^*\leqslant 0,\\ \underline{n}_{ij} & \text{if } \underline{x}_j^*>0\end{cases}$$

and let $A^* = M^* - N^*$. Then $A^* \in [A]$, and from (69) with $k \to \infty$ we get $A^*\underline{x}^* = \underline{b}$, hence $\underline{x}^* \in S$. Analogously one can show that $\bar{x}^* \in S$.

(d) Replace $[x_j]^k$ by $[x_j]$ and $[x_i]^{k+1}$ by $f([x])_i$ in (70). Together with the assumption this yields to $Pd[x] \leqslant d\,f([x]) < d[x]$, and analogously to the proof of Theorem 11(a) we get $\rho(P) < 1$. □

For the Richardson splitting $[A] = I - (I - [A])$ parts of Theorem 14 were already stated and proved in [61]. Most of its present form can be found in [69, Chapters 4.4 and 4.5].

We now apply the Krawczyk operator (43) to the function $Ax - b$ and replace $A \in \mathbb{R}^{n\times n}$, $b \in \mathbb{R}^n$ by $[A] \in I(\mathbb{R}^{n\times n})$, $[b] \in I(\mathbb{R}^n)$. Then we get the modified Krawczyk operator

$$K_{\mathrm{mod}}[x] = m[x] + C([b] - [A]m[x]) + (I - C[A])([x] - m[x]) \tag{71}$$

with some nonsingular matrix $C \in \mathbb{R}^{n\times n}$ and any vector $m[x]$ from $\mathbb{R}^n$. For $K_{\mathrm{mod}}[x]$ and for the iteration

$$[x]^{k+1} = K_{\mathrm{mod}}[x]^k \cap [x]^k \tag{72}$$

with fixed C the following analogue of Theorem 11 holds.

Theorem 15. *Let* $[A] \in I(\mathbb{R}^{n\times n})$, $[b] \in \mathrm{I}(\mathbb{R}^n)$:

(a) *If*

$$\rho(|I - C[A]|) < 1, \tag{73}$$

then $[A]$ *is nonsingular, i.e., each linear system* $Ax = b$ *with* $A \in [A]$ *and* $b \in [b]$ *is uniquely solvable. If, in addition,* $S \subseteq [x]^0$ *then the sequence* $\{[x]^k\}_{k=0}^{\infty}$ *defined by* (72) *is well defined,* $S \subseteq [x]^k$ *and* $\lim_{k\to\infty} [x]^k = [x]^* \supseteq S$. *In particular,* $\{[x]^k\}_{k=0}^{\infty}$ *is monotonically decreasing.*

(b) *If*

$$K_{\mathrm{mod}}[x] \subseteq [x] \tag{74}$$

for some $[x] \in I(\mathbb{R}^n)$ *then each linear system* $Ax = b$ *with* $A \in [A]$ *and* $b \in [b]$ *has a solution* $x^* \in [x]$.

If (74) *is slightly sharpened to*

$$(K_{\mathrm{mod}}[x])_i \subset [x_i] \quad \textit{for } i = 1, \ldots, n, \tag{75}$$

then $\rho(|I - C[A]|) < 1$, *i.e., the properties in* (a) *hold with* $S \subset [x]$.

(c) *If*

$$||\,|I - C[A]|\,||_\infty < 1, \tag{76}$$

then the properties in (a) *hold. In addition,*

$$S \subseteq [\tilde{x}] = [\tilde{x} - \alpha e, \tilde{x} + \alpha e], \tag{77}$$

where

$$\alpha = \frac{||\,|C([b] - [A]\tilde{x})|\,||_\infty}{1 - ||\,|I - C[A]|\,||_\infty}.$$

Therefore, the second part of (a) *holds for any* $[x]^0 \supseteq [\tilde{x}]$.

Proof. (a) Can be proved via an analog of (50) and by using the representation

$$x^* = m[x] + C(b - Am[x]) + (I - CA)(x^* - m[x]) \in K_{\mathrm{mod}}[x] \tag{78}$$

for $x^* = A^{-1}b,\ A \in [A],\ b \in [b]$.

(b) Is proved analogously to part (a) of Theorem 11.

(c) Since the assertion implies $\rho(|I - C[A]|) < 1$ all properties of (a) hold. Let $x^* \in S$. Then there are $A \in [A],\ b \in [b]$ such that $Ax^* = b$. Hence

$$||x^* - \tilde{x}||_\infty = ||A^{-1}(b - A\tilde{x})||_\infty \leqslant ||\{I - (I - CA)\}^{-1}||_\infty ||C(b - A\tilde{x})||_\infty \leqslant \alpha,$$

where we used the Neumann series for the last inequality. □

Remark 3. (a) As in Remark 2 it is not necessary to know whether C is nonsingular if (73), (75) or (76) hold. Either of these assumptions guarantees the nonsingularity of C.

(b) If (74) or (75) holds then $S \subseteq K_{\mathrm{mod}}[x]$.

(c) If $[A]$ and $[b]$ are degenerate, i.e., $[A] \equiv A$, $[b] \equiv b$ then the assumption $\rho(|I - CA|) < 1$ in Theorem 15 implies

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

where $Ax^* = b$.

Remark 3(b) leads to the question how good the enclosures are which one gets as iterates obtained by (72). The following result is due to Rump [82] and answers this question if (75) holds. To this end we define S_i as the projection of S to the ith coordinate axis, i.e.,

$$S_i = \{x_i \,|\, x \in S\} \subseteq \mathbb{R}. \tag{79}$$

For nonsingular $[A]$ Cramer's rule shows that x_i depends continuously on $A \in [A]$ and $b \in [b]$. Since $[A]$ and $[b]$ are connected and compact, the sets S_i are compact intervals.

Theorem 16. *Let* $[A] \in I(\mathbb{R}^{n \times n})$, $[b] \in I(\mathbb{R}^n)$, S_i *as in* (79). *Compute* $K_{\mathrm{mod}}[x]$ *from* (71) *with any* $m[x] = \tilde{x} \in \mathbb{R}^n$ *and any nonsingular* $C \in \mathbb{R}^{n \times n}$, *and let*

$$[z] = C([b] - [A]\tilde{x}), \quad [\delta] = (I - C[A])([x] - \tilde{x}).$$

If $(K_{\mathrm{mod}}[x])_i \subset [x_i]$ *for* $i = 1, \ldots, n$ *then*

$$\tilde{x}_i + \underline{z}_i + \underline{\delta}_i \leqslant \min S_i \leqslant \tilde{x}_i + \underline{z}_i + \overline{\delta}_i, \tag{80}$$

$$\tilde{x}_i + \overline{z}_i + \underline{\delta}_i \leqslant \max S_i \leqslant \tilde{x}_i + \overline{z}_i + \overline{\delta}_i, \tag{81}$$

i.e., $d[\delta]$ *is a measure for the overestimation of* S *by* $K_{\mathrm{mod}}[x]$.

Proof. The left inequality of (80) and the right inequality of (81) follow directly from Remark 3(b). In order to prove the two remaining inequalities note that the interval $[z_i]$ is the interval arithmetic evaluation of the function $f : \mathbb{R}^{n^2+n} \to \mathbb{R}$ which is defined by $f(A, b) = (C(b - A\tilde{x}))_i$. In $f(A, b)$ each variable occurs only once. Therefore, Theorem 2 implies

$$f([A], [b]) = R(f; [A], [b]), \tag{82}$$

i.e., there are some $A^* \in [A]$, $b^* \in [b]$ such that $\underline{z}_i = f(A^*, b^*)$. From (78) for $x^* = (A^*)^{-1}b^* \in S$ and with $\delta^* = (I - CA^*)(x^* - \tilde{x})$ we get

$$\min S_i \leqslant x_i^* = \tilde{x}_i + \underline{z}_i + \delta_i^* \leqslant \tilde{x}_i + \underline{z}_i + \bar{\delta}_i,$$

which shows the right inequality of (80). The left inequality of (81) is proved analogously. □

Remark 4. Let (75) holds with C being the inverse of the center of $[A]$ and let $\tilde{x}$ be a good approximation of some element of S. Assume that $d[A]$, $d[b]$ are small and that (75) holds for some $[x]$ with $m[x] = \tilde{x} \in [x]$. Then $d[z] = |C|(d[b] + d[A]\tilde{x})$ can be expected to be small and from

$$[\delta] = |C|[-\tfrac{1}{2}d[A], \tfrac{1}{2}d[A]]([x] - \tilde{x}) = |C|[-\tfrac{1}{2}d[A], \tfrac{1}{2}d[A]]|[x] - \tilde{x}|,$$

we get $d[\delta] \leqslant |C|d[A]d[x]$. Hence if $d[x]$ is also small (which can be expected if some $A \in [A]$ is not ill-conditioned) then $d[\delta]$ is quadratically small, i.e., $d[\delta] \ll d[z]$. This indicates a small overestimation of S by $K_{\mathrm{mod}}[x]$.

If, in fact, at least $d[\delta] \leqslant d[z]$ holds then $\underline{z} + \bar{\delta} \leqslant \bar{z} + \underline{\delta}$ and $[x]^{\mathrm{int}} = [\underline{x}^{\mathrm{int}}, \bar{x}^{\mathrm{int}}] = \tilde{x} + [\underline{z} + \bar{\delta}, \bar{z} + \underline{\delta}]$ is an interval vector which satisfies $\min S_i \leqslant \underline{x}_i^{\mathrm{int}} \leqslant \bar{x}_i^{\mathrm{int}} \leqslant \max S_i$ for $i = 1, \ldots, n$. Such a vector is called an inner enclosure of S by Rump [84]. If an inner enclosure of S is known one can estimate the quality of an enclosure (in the set-theoretical sense) of S in a straightforward way. Inner enclosures and related topics are considered for instance in [84,87].

Now we address to the symmetric solution set S_{sym} from (65), i.e., we are interested in linear systems $Ax = b$ with symmetric matrices $A \in [A] \in I(\mathbb{R}^{n\times n})$. For simplicity, we assume

$$[A] = [A]^{\mathrm{T}}. \tag{83}$$

Otherwise the subsequent results hold for the largest interval matrix which is contained in $[A]$ and which has property (83).

Trivially, S_{sym} is a subset of S. Its shape is even more complicated than that of S: Curved boundaries can occur as the following theorem indicates.

Theorem 17. *Let S_{sym} be defined for a given nonsingular interval matrix $[A] = [A]^{\mathrm{T}} \in I(\mathbb{R}^{n\times n})$ and a given interval vector $[b] \in I(\mathbb{R}^n)$. Then for any closed orthant $O \subseteq \mathbb{R}^n$ the set $S_{\mathrm{sym}} \cap O$ can be represented as the intersection of finitely many closed sets whose boundaries are quadrics or hyperplanes. These sets can be described by inequalities which result, e.g., from a Fourier–Motzkin elimination process.*

The proof of this theorem can be found in [15], corresponding properties on classes of matrices with more general dependencies in [16,17]. For the Fourier–Motzkin elimination see, for instance, [85].

We want to enclose S_{sym} by an interval vector. Trivially, each method for enclosing S delivers such a vector. But the symmetric solution set often contains much less elements than S. Therefore, it is useful to look for methods which enclose S_{sym} but not necessarily S. Such a method is the interval Cholesky method which is defined by applying formally the formulas of the Cholesky method to the interval data $[A] = [A]^{\mathrm{T}}$ and $[b]$. It produces an interval vector which we denote by $\mathrm{ICh}([A], [b])$. In the algorithm the squares and the square roots are defined via (4). We assume that no division

by an interval occurs which contains zero. If $\langle [A] \rangle$ is an M matrix with $\underline{a}_{ii} > 0$ for $i = 1, \ldots, n$ then $\mathrm{ICh}([A], [b])$ exists. This was shown in [19] where the interval version of the Cholesky method was introduced and studied in detail. See also [21].

Another method to enclose S_{sym} was considered by Jansson in [41]. He starts with a modification of $K_{\mathrm{mod}}[x]$ from (71): Let

$$K^{\mathrm{sym}}_{\mathrm{mod}}[x] = m[x] + [z]^{\mathrm{sym}} + (I - C[A])([x] - m[x]), \tag{84}$$

where $[z]^{\mathrm{sym}} = ([z_i]^{\mathrm{sym}}) \in I(\mathbb{R}^n)$ is defined by

$$[z_i]^{\mathrm{sym}} = \sum_{j=1}^{n} c_{ij}([b_j] - [a_{jj}](m[x])_j) - \sum_{j=1}^{n} \sum_{l=1}^{j-1} (c_{ij}(m[x])_l + c_{il}(m[x])_j)[a]_{jl}.$$

Iterate analogously to (72) with $K^{\mathrm{sym}}_{\mathrm{mod}}[x]$ replacing $K_{\mathrm{mod}}[x]$. Since by the same reasoning as above

$$[z_i]^{\mathrm{sym}} = \{(C(b - Am[x]))_i \mid A = A^{\mathrm{T}} \in [A],\ b \in [b]\},$$

Theorems 15 and 16 hold with S, $[z]$ being replaced by S_{sym}, $[z]^{\mathrm{sym}}$.

6. The algebraic eigenvalue problem and related topics

In this section we look for intervals $[\lambda] \in I(\mathbb{R})$ and interval vectors $[x] \in I(\mathbb{R}^n)$ such that $[\lambda]$ contains an eigenvalue $\lambda^* \in \mathbb{R}$ and $[x]$ contains an associated eigenvector $x^* \in \mathbb{R}^n \setminus \{0\}$ for a given matrix $A \in \mathbb{R}^{n \times n}$. We restrict ourselves only to real eigenpairs. Complex ones have also been studied; cf. [56,57], e.g., for an overview.

We start with the mild nonlinear equation

$$f(x, \lambda) = \begin{pmatrix} Ax - \lambda x \\ x_{i_0} - \alpha \end{pmatrix} = 0, \tag{85}$$

where i_0 is a fixed index from $\{1, \ldots, n\}$ and $\alpha \neq 0$ is a constant. It is obvious that (x^*, λ^*) is a solution of (85) if and only if (x^*, λ^*) is an eigenpair of A with the normalization $x^*_{i_0} = \alpha$ of the eigenvector x^*. Expanding f into a Taylor series at an approximation $(\tilde{x}, \tilde{\lambda})$ of (x^*, λ^*) yields to

$$f(x, \lambda) = f(\tilde{x}, \tilde{\lambda}) + \begin{pmatrix} A - \tilde{\lambda} I_n & -\tilde{x} \\ (\mathrm{e}^{(i_0)})^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} - \begin{pmatrix} \Delta\lambda\, \Delta x \\ 0 \end{pmatrix}, \tag{86}$$

where $\Delta x = x - \tilde{x}$, $\Delta \lambda = \lambda - \tilde{\lambda}$, I_k is the $k \times k$ identity matrix and $\mathrm{e}^{(i_0)}$ is the i_0th column of I_n. Multiplying (86) by a preconditioning matrix $-C \in \mathbb{R}^{(n+1)\times(n+1)}$ and adding $((\Delta x)^{\mathrm{T}}, \Delta\lambda)^{\mathrm{T}}$ on both sides results in the fixed point equation

$$\begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix} = g(\Delta x, \Delta\lambda) = -Cf(\tilde{x}, \tilde{\lambda}) + \left\{ I_{n+1} - C \begin{pmatrix} A - \tilde{\lambda} I_n & -\tilde{x} - \Delta x \\ (\mathrm{e}^{(i_0)})^{\mathrm{T}} & 0 \end{pmatrix} \right\} \begin{pmatrix} \Delta x \\ \Delta \lambda \end{pmatrix}, \tag{87}$$

for the error $(\Delta x, \Delta\lambda) = (\Delta x^*, \Delta\lambda^*) = (x^* - \tilde{x}, \lambda^* - \tilde{\lambda})$ of an eigenpair (x^*, λ^*). The following theorem is due to Rump [81].

Theorem 18. *Let* $A \in \mathbb{R}^{n\times n}$, $\tilde{\lambda} \in \mathbb{R}$, $\tilde{x} \in \mathbb{R}^n$, $C \in \mathbb{R}^{(n+1)\times(n+1)}$, *and define* g *by* (87). *Let* $\tilde{x}$ *be normalized by* $\tilde{x}_{i_0} = \alpha \neq 0$. *If* g *fulfills the inclusion*

$$g([\Delta x], [\Delta\lambda]) \subseteq \operatorname{int}([\Delta x]^{\mathrm{T}}, [\Delta\lambda])^{\mathrm{T}} \tag{88}$$

then the following assertions hold:

(a) C *is nonsingular.*

(b) *There exists exactly one eigenvector* $x^* \in \tilde{x} + [\Delta x]$ *of* A *which is normalized by* $x^*_{i_0} = \alpha$.

(c) *There exists exactly one eigenvalue* $\lambda^* \in \tilde{\lambda} + [\Delta\lambda]$ *of* A.

(d) $Ax^* = \lambda^* x^*$ *with* x^* *from* (b) *and* λ^* *from* (c).

(e) *The eigenvalue* λ^* *from* (d) *is geometric simple.*

(f) *If* $(\tilde{x}, \tilde{\lambda})$ *is a sufficiently good approximation of the eigenpair* (x^*, λ^*) *from* (d) *then it can be guaranteed that* λ^* *is algebraic simple.*

(g) *If one starts the iteration*

$$\begin{pmatrix} [\Delta x]^{k+1} \\ [\Delta\lambda]^{k+1} \end{pmatrix} = g([\Delta x]^k, [\Delta\lambda]^k), \quad k = 0, 1, \ldots, \tag{89}$$

with

$$([\Delta x]^0, [\Delta\lambda]^0) = ([\Delta x], [\Delta\lambda])$$

from (88) *then the iterates converge satisfying*

$$([\Delta x]^{k+1}, [\Delta\lambda]^{k+1}) \subseteq ([\Delta x]^k, [\Delta\lambda]^k), \quad k = 0, 1, \ldots$$

and

$$(x^*, \lambda^*) \in (\tilde{x}, \tilde{\lambda}) + ([\Delta x]^k, [\Delta\lambda]^k), \quad k = 0, 1, \ldots$$

for the eigenpair (x^*, λ^*) *from* (d).

Interval quantities $[x]$, $[\lambda]$ with (88) can be found, e.g., via ε-inflation; cf. [58] or [59]. Another way was indicated in [6] by the following theorem.

Theorem 19. *With the notations of Theorem* 18 *define*

$$\rho = \left\| C \begin{pmatrix} A\tilde{x} - \tilde{\lambda}\tilde{x} \\ 0 \end{pmatrix} \right\|_\infty, \quad \sigma = \left\| I_{n+1} - C \begin{pmatrix} A - \tilde{\lambda} I_n & -\tilde{x} \\ (e^{(i_0)})^{\mathrm{T}} & 0 \end{pmatrix} \right\|_\infty, \quad \tau = \|C\|_\infty \tag{90}$$

and assume

$$\sigma < 1, \quad \Delta = (1-\sigma)^2 - 4\rho\tau \geqslant 0. \tag{91}$$

Then the numbers

$$\beta^- = (1 - \sigma - \sqrt{\Delta})/(2\tau) = \frac{2\rho}{1 - \sigma + \sqrt{\Delta}},$$

$$\beta^+ = (1 - \sigma + \sqrt{\Delta})/(2\tau)$$

are nonnegative, and the condition (88) *of Theorem* 18 *is fulfilled for* $([\Delta x]^{\mathrm{T}}, [\Delta\lambda])^{\mathrm{T}} = [-\beta, \beta]e \in I(\mathbb{R})^{(n+1)\times(n+1)}$ *with arbitrary* $\beta \in (\beta^-, \beta^+)$. *In particular, all the assertions of that theorem hold.*

If β is restricted to $[\beta^-,(\beta^- + \beta^+)/2)$ then the iterates of (89) *converge to the error* $\binom{\Delta x^*}{\Delta\lambda^*}$.

In [58] it is shown how (87) can be reduced to an n-dimensional problem which, originally, formed the starting point in [6]. It is also indicated there how (87) has to be modified if the normalization $x^*_{i_0} = \alpha$ is replaced by $||x^*||_2 = 1$.

A second method for enclosing eigenpairs starts with the centered form

$$f(x,\lambda) = f(\tilde{x},\tilde{\lambda}) + \begin{pmatrix} A - \tilde{\lambda} I_n & -\tilde{x} - \Delta x \\ (e^{(i_0)})^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix}.$$

It is obvious that the subdivision principle discussed in Section 3 can be applied to any initial domain $([x]^0,[\lambda]^0)$ chosen by the user. The crucial problem remains to verify that $0 \in f([\hat{x}],[\hat{\lambda}])$ yields to $f(x^*,\lambda^*) = 0$ in a subdomain $([\hat{x}],[\hat{\lambda}]) \subseteq ([x]^0,[\lambda]^0)$.

A third method is due to H. Behnke and F. Goerisch. It assumes A to be symmetric and is based on a complementary variational principle. For details see, e.g., [23, Section 6], and the references there.

Symmetric matrices can also be handled by an access due to Lohner [54]. First A is reduced to nearly diagonal form using Jacobi rotations and a sort of staggered correction. Finally Gershgorin's theorem is applied in order to obtain bounds for the eigenvalues. A theorem due to Wilkinson allows the enclosure of eigenvectors.

There is no problem to generalize the ideas above to the generalized eigenvalue problem $Ax = \lambda Bx$, $x \neq 0$, $B \in \mathbb{R}^{n\times n}$ nonsingular. The analogue of (85) reads

$$f(x,\lambda) = \begin{pmatrix} Ax - \lambda Bx \\ x_{i_0} - \alpha \end{pmatrix} = 0.$$

In a similar way one can treat the singular value problem for a given $m \times n$ matrix A with $m \geqslant n$. Here, we look for orthogonal matrices $U \in \mathbb{R}^{n\times n}$, $V \in \mathbb{R}^{m\times m}$ and for a diagonal matrix $\Sigma = \mathrm{diag}(\sigma_1,\ldots,\sigma_r,\ldots,\sigma_n) \in \mathbb{R}^{m\times n}$ with the singular values $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_r > \sigma_{r+1} = 0 = \cdots = \sigma_n$, $r = \mathrm{rank}(A)$, such that $A = V\Sigma U^{\mathrm{T}}$. One starts with

$$f(u,v,\sigma) = \begin{pmatrix} Au - \sigma v \\ A^{\mathrm{T}}v - \sigma u \\ u^{\mathrm{T}}u - 1 \end{pmatrix} \quad \text{or with} \quad f(u,v,\sigma,\sigma') = \begin{pmatrix} Au - \sigma v \\ A^{\mathrm{T}}v - \sigma' u \\ u^{\mathrm{T}}u - 1 \\ v^{\mathrm{T}}v - 1 \end{pmatrix}.$$

In the first case a zero of f satisfies $v^{\mathrm{T}}v = 1$, in the second one gets $\sigma = \sigma'$. In either of the cases u is a column of U, v a corresponding column of V and σ a singular value of A associated with u and v. For details, additional remarks and references to further methods for verifying and enclosing singular values see [7,57].

We also mention verification methods in [14] for generalized singular values (c^*,s^*) of a given matrix pair (A,B), $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{q\times n}$, which are defined as the zeros of the function $f(c,s) = \det(s^2A^{\mathrm{T}}A - c^2B^{\mathrm{T}}B)$ restricted to $c, s \geqslant 0$, $c^2 + s^2 = 1$. For applications of generalized singular values see [33].

The methods and results of the Sections 4–6 can be combined in order to study the following inverse eigenvalue problem:

Given $n+1$ symmetric matrices $A_i \in \mathbb{R}^{n\times n}$, $i=0,1,\ldots,n$. Find n real numbers c_i^*, $i=1,\ldots,n$, such that the matrix $A(c)=A_0+\sum_{i=1}^n c_iA_i$, $c=(c_i)\in\mathbb{R}^n$, has for $c=c^*=(c_i^*)$ prescribed eigenvalues

$$\lambda_1^* < \lambda_2^* < \cdots < \lambda_n^*. \tag{92}$$

Here one starts with the function $f(c)=\lambda(c)-\lambda^*\in\mathbb{R}^n$, c sufficiently close to c^*, where the components $\lambda_i(c)$ of $\lambda(c)$ are the eigenvalues of $A(c)$ ordered increasingly, and where $\lambda^*=(\lambda_i^*)$ is defined with (92). One can show that the equation for Newton's method reads

$$(x^i(c^k))^TA_j(x^i(c^k))(c^{k+1}-c^k)=-(\lambda(c^k)-\lambda^*); \tag{93}$$

$x^i(c^k)$ are the eigenvectors of $A(c^k)$ associated with the eigenvalues $\lambda_i(c^k)$ and normalized by $x^i(c^k)^Tx^i(c^k)=1$, $\operatorname{sign}(x^i_{i_0}(c^k))=1$ for some fixed index $i_0\in\{1,\ldots,n\}$.

In a first step approximations of $x^i(c^k)$, $\lambda_i(c^k)$ are computed for $i=1,\ldots,n$. With these values Eq. (93) is formed and solved. This is done for $k=0,1,\ldots$ up to some k_0. In a second step the verification process is performed using the interval Newton method and results from Section 6 which are generalized from point matrices to interval matrices. For details see [10,20] or [57].

7. Ordinary differential equations

Many contributions to verification numerics refer to initial value problems for ordinary differential equations

$$y'=f(y), \tag{94}$$

$$y(x_0)=y^0, \tag{95}$$

where we assume that $f{:}D\subseteq\mathbb{R}^n\to\mathbb{R}^n$ is sufficiently smooth and that (94) has a unique solution in some given interval $[x_0,x_0+T]$ for any initial value $y^0\in[y^0]\subseteq D$. For ease of presentation we choose (94) to be autonomous. This is not a severe restriction since any nonautonomous initial value problem can be reduced to an autonomous one by introducing the additional component $y_{n+1}=x$, the additional differential equation $y'_{n+1}=1$ and the additional initial value $y_{n+1}(x_0)=y^0_{n+1}=x_0$. We shall use a grid $x_0<x_1<\cdots<x_k<\cdots<x_K=x_0+T$ with grid points x_k and stepsizes $h_k=x_{k+1}-x_k$ to be determined later on, and we shall consider (94) with initial values $y(x_k)$ from some intermediate interval vectors $[y^k]$. To this end we introduce the set

$$y(x;x_k,[y^k])=\{y(x)\mid y'=f(y),\ y(x_k)\in[y^k]\} \tag{96}$$

of all solutions of (94) with initial values in $[y^k]$. In the sequel, we shall need the following auxiliary result.

Theorem 20. *If* $[\tilde y]+[0,h]f([\hat y])\subseteq[\hat y]$ *for* f *from* (94) *and some* $h>0$, $[\tilde y]\subseteq[\hat y]\subseteq D$, *then* $y(x;\tilde x,[\tilde y])\subseteq[\hat y]$ *for all* $x\in[\tilde x,\tilde x+h]$.

Proof. For fixed $\tilde y^0\in[\tilde y]$ apply Banach's fixed point theorem to the Picard–Lindelöf operator $(Tu)(x)=\tilde y^0+\int_{\tilde x}^x f(u(t))\,\mathrm{d}t$, to the set $U=\{u\mid u\in C^0[\tilde x,\tilde x+h]$ and $u(x)\in[\hat y]$ for $x\in[\tilde x,\tilde x+h]\}$ and to the metric $\|u\|_\alpha=\max_{\tilde x\leqslant x\leqslant\tilde x+h}\{\mathrm{e}^{-\alpha(x-\tilde x)}\parallel u(x)\parallel_\infty\}$ with any $\alpha>\||\partial f([\hat y])/\partial y|\|_\infty$. □

One of the most popular methods for verifying and enclosing solutions of initial value problems is known as interval Taylor series method. It goes back to R.E. Moore and was modified in various ways – cf., for instance, [30,53], and overviews in [26,66,80]. In order to describe this method we assume that we know the grid point $x_k < x_K$ and an enclosure $[y^k]$ of $y(x_k; x_0, [y^0])$. Such an enclosure is given for $k = 0$. The method consists of two major steps:

In the first step a new stepsize $h_k > 0$ and a rough a priori enclosure $[\hat{y}^k]$ is computed such that

$$y(x; x_k, [y^k]) \subseteq [\hat{y}^k] \quad \text{for all } x \in [x_k, x_k + h_k]. \tag{97}$$

To this end let $[\hat{y}^k]$ be any vector which contains $[y^k]$ in its interior and choose $h_k > 0$ so small that $[y^k] + [0, h_k] f([\hat{y}^k]) \subseteq [\hat{y}^k]$. Then (97) is guaranteed by Theorem 20. With h_k we know $x_{k+1} = x_k + h_k$, and from (97) with $x = x_{k+1}$ we see that $[\hat{y}^k]$ is a candidate for $[y^{k+1}]$.

In the second step of the method this candidate is improved in the following way: consider any particular solution y^* of (94) with $y^*(x_k) \in [y^k]$. Using (94) and the Taylor expansion of y^* at x_k we get for a fixed $p \in \mathbb{N}$ and $h = x - x_k$

$$y^*(x) = \psi(h, y^*(x_k)) + r_p(h, y^*) \tag{98}$$

with

$$\psi(h, y) = y + \sum_{j=1}^{p} h^j f^{[j]}(y), \quad f^{[1]} = f, \quad f^{[j]} = \frac{1}{j}(f^{[j-1]})' = \frac{1}{j}\frac{\partial f^{[j-1]}}{\partial y} f \text{ for } j \geqslant 2$$

and with the remainder term $r_p(h, y^*) \in h^{p+1} f^{[p+1]}([\hat{y}^k])$. Throughout this section we assume that the Taylor coefficients $f^{[j]}(y^*(x_k))$ exist. They can be computed recursively by means of automatic differentiation which is described, e.g., in [34] or [76]. Obviously,

$$y(x; x_0, [y^0]) \subseteq y(x; x_k, [y^k]) \subseteq \psi(h, [y^k]) + h^{p+1} f^{[p+1]}([\hat{y}^k]) \quad \text{for } x_k \leqslant x \leqslant x_{k+1}. \tag{99}$$

By virtue of $d\psi(h_k, [y^k]) \geqslant d[y^k]$ the right expression in (99) with $h = h_k$ seems not yet to be suited as a good candidate for $[y^{k+1}]$ since its diameter dominates $d[y^k]$. Therefore, we represent $\psi(h, y)$ as centered form

$$\psi(h, y) = \psi(h, \tilde{y}^k) + \left\{ I + \sum_{j=1}^{p} h^j J(y, \tilde{y}^k; f^{[j]}) \right\} (y - \tilde{y}^k) \tag{100}$$

$$\in \psi(h, \tilde{y}^k) + \left\{ I + \sum_{j=1}^{p} h^j \frac{\partial f^{[j]}([y^k])}{\partial y} \right\} ([y^k] - \tilde{y}^k), \tag{101}$$

where $y, \tilde{y}^k \in [y^k]$ and where $J(y, z; f)$ is defined as $J(y, z)$ in (29) using the third argument as underlying function. With y^* as in (98) and

$$S_k^* = I + \sum_{j=1}^{p} h_k^j J(y^*, \tilde{y}^k; f^{[j]}), \tag{102}$$

$$[S_k] = I + \sum_{j=1}^{p} h_k^j \frac{\partial f^{[j]}([y^k])}{\partial y}, \tag{103}$$

$$[\tilde{y}^{k+1}] = \psi(h_k, \tilde{y}^k) + h_k^{p+1} f^{[p+1]}([\hat{y}^k]) \tag{104}$$

for $k = 0, 1, \ldots, K-1$ we therefore get

$$y^*(x_{k+1}) = \psi(h_k, \tilde{y}^k) + r_p(h_k, y^*) + S_k^*(y^*(x_k) - \tilde{y}^k) \tag{105}$$

$$\in [\tilde{y}^{k+1}] + [S_k]([y^k] - \tilde{y}^k). \tag{106}$$

The partial derivatives in (101) and (103) can again be computed using automatic differentiation or by differentiating the code list of $f^{[j]}$. Formula (105) represents the basis for most variants of the interval Taylor series method as long as they differ in their second step. Obviously,

$$y(x_{k+1}; x_0, [y^0]) \subseteq y(x_{k+1}; x_k, [y^k]) \subseteq [\tilde{y}^{k+1}] + [S_k]([y^k] - \tilde{y}^k), \tag{107}$$

so that the right expression is a candidate for $[y^{k+1}]$, this time with $d[y^{k+1}] \leqslant d[y^k]$ being possible. The successive construction of $[y^{k+1}]$ via (106) is called mean value method. Since $0 \in [S_k]([y^k] - \tilde{y}^k)$, we get $[\tilde{y}^{k+1}] \subseteq [y^{k+1}]$. Therefore, we can assume for the succeeding interval $[x_{k+1}, x_{k+2}]$ that $\tilde{y}^{k+1} \in [y^{k+1}]$ in (100) is chosen from $[\tilde{y}^{k+1}]$ – preferably its midpoint – which justifies our notation.

Unfortunately, $y(x_{k+1}; x_k, [y^k])$ is not necessarily an interval vector. Therefore, $[y^{k+1}]$ can overestimate this set and, consequently, $y(x_{k+1}; x_0, [y^0])$. This phenomenon which occurs at each grid point x_k, $k > 0$, is called wrapping effect. Its existence is an intrinsic feature of interval arithmetic and does not depend on the particular method. Its size, however, is strongly influenced by the choice of the method. In order to reduce this size the original mean value method often has to be modified. If $h_k > 0$ is small and p is large one can expect that the second summand $[S_k]([y^k] - \tilde{y}^k)$ in (106) contributes most to the wrapping effect. It can be influenced by preconditioning with a regular matrix $A_k \in \mathbb{R}^{n \times n}$ which yields to the following variant of the mean value method:

- Choose $\tilde{y}^0 \in [y^0]$ and let $[r^0] = [y^0] - \tilde{y}^0$, $A_0 = I \in \mathbb{R}^{n \times n}$.

For $k = 0, 1, \ldots, K-1$ do the following steps:

- Compute $[S_k]$, $[\tilde{y}^{k+1}]$ as in (103), (104).
- Choose $\tilde{y}^{k+1} \in [\tilde{y}^{k+1}]$.
- Choose $A_{k+1} \in \mathbb{R}^{n \times n}$ (regular) as described below.
- Compute

$$[r^{k+1}] = \{A_{k+1}^{-1}([S_k]A_k)\}[r^k] + A_{k+1}^{-1}([\tilde{y}^{k+1}] - \tilde{y}^{k+1}), \tag{108}$$

$$[y^{k+1}] = [\tilde{y}^{k+1}] + ([S_k]A_k)[r^k]. \tag{109}$$

Before we consider particular choices of matrices A_k we prove an analogue of (107).

Theorem 21. *Let $\tilde{y}^k$, $[\tilde{y}^k]$, $[y^k]$, $[r^k], A_k$ be defined for $k = 0, 1, \ldots, K$ as in the preceding variant of the mean value method and let, formally, $x_{-1} = x_0$, $[y^{-1}] = [y^0]$. Then for $k = 0, 1, \ldots, K$ we get*

$$y(x_k; x_{k-1}, [y^{k-1}]) \subseteq [y^k], \tag{110}$$

$$A_k^{-1}(y^*(x_k) - \tilde{y}^k) \in [r^k] \quad \textit{for any solution } y^* \textit{ of } (94) \textit{ with } y^*(x_{k-1}) \in [y^{k-1}]. \tag{111}$$

Proof. The assertion is true for $k = 0$ by the definition of x_{-1}, $[y^{-1}]$ and by $A_0 = I$. Let it hold for some $k < K$ and let y^* be a solution of (94) with $y^*(x_k) \in [y^k]$. From (105), (111) and (109) we get

$$y^*(x_{k+1}) \in [\tilde{y}^{k+1}] + S_k^*(y^*(x_k) - \tilde{y}^k) = [\tilde{y}^{k+1}] + (S_k^* A_k)\{A_k^{-1}(y^*(x_k) - \tilde{y}^k)\} \tag{112}$$

$$\subseteq [\tilde{y}^{k+1}] + ([S_k]A_k)[r^k] = [y^{k+1}], \tag{113}$$

hence (110) follows for $k+1$. Since (112) implies $y^*(x_{k+1}) - \tilde{y}^{k+1} \in [\tilde{y}^{k+1}] - \tilde{y}^{k+1} + S_k^*(y^*(x_k) - \tilde{y}^k)$ we obtain

$$A_{k+1}^{-1}(y^*(x_{k+1}) - \tilde{y}^{k+1}) \in A_{k+1}^{-1}([\tilde{y}^{k+1}] - \tilde{y}^{k+1}) + (A_{k+1}^{-1} S_k^* A_k)\{A_k^{-1}(y^*(x_k) - \tilde{y}^k)\}$$

$$\subseteq A_{k+1}^{-1}([\tilde{y}^{k+1}] - \tilde{y}^{k+1}) + (A_{k+1}^{-1}[S_k]A_k)[r^k] = [r^{k+1}],$$

where we used (111) and (108). □

An easy induction shows that one can retrieve the mean value method from its variant above if $A_k = I$ for $k = 0, 1, \ldots, K$.

If $A_{k+1} \in [S_k]A_k$ then $I \in A_{k+1}^{-1}([S_k]A_k)$, and $(A_{k+1}^{-1}[S_k]A_k)[r^k] \approx [r^k]$ can be expected if A_k is not ill-conditioned (cf. [66, p. 32]). Therefore, the wrapping effect should not lead to large overestimations in this case. Unfortunately, A_k is not always well-conditioned. So, other choices for A_k become important. R. Lohner starts in [53] with $\tilde{A}_{k+1} \in [S_k]A_k$ and performs a QR-decomposition of $\tilde{A}_{k+1}$ (eventually after having permuted the columns of this matrix), i.e., $\tilde{A}_{k+1} = Q_{k+1}R_{k+1}$. Then he chooses $A_{k+1} = Q_{k+1}$ which effects a rotation of the coordinate system. For details cf. [53] or [66].

We also mention variants due to Eijgenraam [30] and Rihm [80] and Lohner's implementation AWA. For further reading we recommend [66] in which an interval Hermite–Obreschkoff method is considered, and [67] in which an enclosure method for the solution of linear ODEs with polynomial coefficients is given.

Based on the preceding ideas boundary value problems can be handled via the well-known shooting method as it was done in [53].

The stability of the Orr–Sommerfeld equation for different parameters was investigated in [51] by enclosure methods.

ODEs are closely related to integral equations. Therefore, it is interesting to ask for verified enclosures of such equations and of definite integrals. Due to space limit, however, we must refer the reader to the literature, for instance to [25,32,43] and to various contributions in [1].

8. Partial differential equations

Like the theory of partial differential equations the verification methods in this field are very heterogeneous. As in many cases in the previous sections they are mostly based on fixed point theorems and on particular function spaces. In order to give a taste of some ideas we outline a method due to Plum [74] which applies for second order elliptic boundary value problems of the form

$$-\Delta u + F(x, u, \nabla u) = 0 \quad \text{in } \Omega, \tag{114}$$

$$B[u] = 0 \qquad \text{on } \partial\Omega, \tag{115}$$

where $\Omega \subseteq \mathbb{R}^n$, $n \in \{2,3\}$, is a bounded domain whose boundary $\partial\Omega$ is at least Lipschitz continuous. The boundary operator B is defined by

$$B[u] = \begin{cases} u & \text{on } \Gamma_0, \\ \dfrac{\partial u}{\partial \nu} = \nu \cdot \nabla u & \text{on } \partial\Omega \setminus \Gamma_0 \end{cases}$$

with $\Gamma_0 \subseteq \partial\Omega$ being closed and with ν denoting the unit outward normal vector. The function F is given by $F:\bar{\Omega} \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ with $|F(x,y,z)| \leqslant C(1 + ||z||_2^2)$ for some $C \geqslant 0$ and all $x \in \bar{\Omega}$, $y \in \mathbb{R}$, $|y| \leqslant \alpha$, $z \in \mathbb{R}^n$. We assume that F and its derivatives $F_y = \partial F/\partial y$, $F_z = (\partial F/\partial z_1, \ldots, \partial F/\partial z_n)^T$, are continuous.

In view of the theory for (114) we assume that for some $\sigma \in \mathbb{R}$ and each $r \in L^2(\Omega)$ (= set of square integrable functions) the boundary value problem $-\Delta u + \sigma u = r$ in Ω is uniquely solvable in $H_B^2 = \mathrm{cl}\{u \in C^2(\bar{\Omega}) \mid B[u] = 0 \text{ on } \partial\Omega\}$ where 'cl' means the closure in the Sobolev space $H^2(\Omega)$.

We start with a function $\omega \in H_B^2(\Omega)$ which can be thought to be an approximation of a solution u^* of (114), (115), although – at the moment – we do not know whether such a solution exists.

We will apply the operator $L: H_B^2(\Omega) \to L^2(\Omega)$ given by

$$L[u] = -\Delta u + b \cdot \nabla u + cu, \quad b = F_z(\cdot, \omega, \nabla\omega),\ c = F_y(\cdot, \omega, \nabla\omega). \tag{116}$$

In order to guarantee the invertibility of L needed later on we assume $\nabla\omega \in (L^\infty(\Omega))^n$ and we have to check numerically that all eigenvalues of L on $H_B^2(\Omega)$ are nonzero. In addition, we suppose that, for some Banach space $X \supseteq H_B^2(\Omega)$ with some norm $||\cdot||_X$:

(a) the function

$$\Phi: \begin{cases} X & \to & L^2(\Omega), \\ u & \mapsto & b \cdot \nabla u + cu - F(\cdot, u, \nabla u) \end{cases} \tag{117}$$

is continuous, bounded on bounded sets, and Fréchet differentiable at ω with $\Phi'(\omega) = 0$,

(b) the imbedding $H_B^2(\Omega) \hookrightarrow X$ is compact.

As fixed point operator we choose the simplified Newton operator

$$Tu = u - \mathscr{F}'(\omega)^{-1}\mathscr{F}(u) \tag{118}$$

with $\mathscr{F}(u) = -\Delta u + F(\cdot, u, \nabla u)$, with the Fréchet derivative $\mathscr{F}'$ of $\mathscr{F}$ and with ω as above. Since $\mathscr{F}'(\omega) = L$ and $-\Delta u = L[u] - b \cdot \nabla u - cu$ we obtain

$$Tu = u - L^{-1}[-\Delta u + F(\cdot, u, \nabla u)] = L^{-1}[b \cdot \nabla u + cu - F(\cdot, u, \nabla u)] = L^{-1}[\Phi(u)]. \tag{119}$$

Due to our assumptions it can be shown that $T: X \to X$ is continuous, compact and Fréchet differentiable at ω with $T'(\omega) = 0$. If we can find some closed, bounded, convex function set $U \subseteq X$ such that

$$TU \subseteq U, \tag{120}$$

then Schauder's fixed point theorem guarantees the existence of some fixed point $u^* \in U$ of T which, by virtue of (119), is a solution of (114), (115). In order to construct U we first apply a shift $u \mapsto v = u - \omega$ which yields to a set $V = U - \omega$ and which emphasizes the approximative

character of ω. Moreover, it follows the lines of centered forms which we exploited successfully already several times. From $u^* = Tu^*$ and $v^* = u^* - \omega \in X$ we get

$$v^* = T\omega - \omega + \{T(\omega + v^*) - T\omega\} = L^{-1}[-\delta[\omega] + \varphi(v^*)] \tag{121}$$

with

$$\delta[\omega] = -\Delta\omega + F(\cdot, \omega, \nabla\omega),$$

$$\varphi(v) = -\{F(\cdot, \omega + v, \nabla\omega + \nabla v) - F(\cdot, \omega, \nabla\omega) - b \cdot \nabla v - cv\}. \tag{122}$$

If we replace (120) by

$$L^{-1}[-\delta[\omega] + \varphi(V)] \subseteq V, \tag{123}$$

then Schauder's fixed point theorem applies again yielding to a fixed point v^* such that $u^* = \omega + v^*$ is a solution of (114), (115). We now construct a closed, bounded, convex set V which satisfies (123). Since $T'(\omega) = 0$ by definition of ω, we have $T(\omega + v) - T(\omega) = T'(\omega)[v] + o(||v||_X) = o(||v||_X)$, hence, by virtue of (121), v^* can be expected to be small if ω is a good approximation of a solution u^* of (114), (115). Therefore, we assume V to be some small ball

$$V = \{v \in X \mid ||v||_X \leqslant \alpha\} \tag{124}$$

with some $\alpha > 0$. In [74] X is suggested to be the space $H^{1,4}(\Omega)$ with the norm

$$||u||_X = \max\{||u||_\infty, \gamma||\nabla u||_4\} \tag{125}$$

and with

$$||u||_p = \left\{\frac{1}{\text{meas}(\Omega)} \int_\Omega |v(x)|^p \, dx\right\}^{1/p} \quad \text{for } p \in \{2, 4\}$$

here and in the remaining part of this section. The constant $\gamma > 0$ is adapted such that

$$||L^{-1}[r]||_X \leqslant K||r||_2 \quad \text{for all } r \in L^2(\Omega) \tag{126}$$

with a computable constant $K > 0$. Due to $\Phi'(\omega) = 0$ we have

$$||\varphi(v)||_2 = ||\Phi(\omega + v) - \Phi(\omega)||_2 = o(||v||_X) \quad \text{for } ||v||_X \to 0.$$

Let $G:[0, \infty) \to [0, \infty)$ be a majorizing monotonically nondecreasing function such that

$$||\varphi(v)||_2 \leqslant G(||v||_X) \quad \text{for all } v \in X \tag{127}$$

and

$$G(t) = o(t) \quad \text{for } t \to +0. \tag{128}$$

Such a function can be found explicitly via an ansatz according to the lines in [74]. The following theorem is then crucial in view of (123).

Theorem 22. *With the notation and the assumptions above let* $||\delta[\omega]||_2 \leqslant \beta$ *for some* $\beta > 0$. *If*

$$\beta \leqslant \frac{\alpha}{K} - G(\alpha), \tag{129}$$

then V *from* (124) *satisfies* (123), *i.e.*, *there exists a solution* $u^* \in H_B^2(\Omega)$ *of* (114), (115) *with* $||u^* - \omega||_X \leqslant \alpha$.

The proof follows immediately from

$$||L^{-1}[-\delta[\omega] + \varphi(v)]||_X \leqslant K(||\delta[\omega]||_2 + ||\varphi(v)||_2) \leqslant K(\beta + G(||v||_X)) \leqslant K(\beta + G(\alpha)) \leqslant \alpha$$

for each $v \in V$. Note that the right-hand side of (129) is positive for small α, hence (129) can be fulfilled if ω is a sufficiently good approximation of u^* which makes the defect $\delta[\omega]$ small. Some care has to be taken when computing the constants for the inequalities. It is here, among others, where interval arithmetic comes into play. For instance, in order to obtain the constant K in (126) and to check the invertibility of L (on $H_B^2(\Omega)$) one has to verify $\lambda_1 > 0$ for the smallest eigenvalue λ_1 of the eigenvalue problem (in weak formulation)

$$u \in H_B^2(\Omega), \quad \langle L[u], L[\psi] \rangle = \lambda \langle u, \psi \rangle \quad \text{for all } \psi \in H_B^2(\Omega)$$

with $\langle \cdot, \cdot \rangle$ denoting the canonical inner product in $L^2(\Omega)$. By means of interval arithmetic one is able to provide verified bounds for λ_1 and K. Details on the method including the computation of the approximation ω via finite elements can be found in [74] and in papers cited there.

While Plum's method can be characterized as an analytic one there are other methods for elliptic differential equations which use intervals in a more direct way. Thus for the Dirichlet problem

$$\begin{aligned} -\Delta u &= f(u) \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

Nakao [65] works with some set U which has the form

$$U = \omega + \sum_{j=1}^{m} [a_j]\phi_j + \{\phi \in S^{\perp} \mid ||\phi||_{H_0^1} \leqslant \alpha\},$$

where $S \subseteq H_0^1(\Omega)$ is a finite-dimensional (finite element) subspace, $S^{\perp}$ is its orthogonal complement in H_0^1, $\{\phi_1, \ldots, \phi_m\}$ forms a basis of S and α is some constant which has to be determined numerically.

We also mention verification methods for hyperbolic equations – cf. for instance [28,47] and the literature there.

The investigation of old and the introduction of new ideas for the enclosure of solutions of differential equations is still a very active part of research.

9. Software for interval arithmetic

Interval arithmetic has been implemented on many platforms and is supported by several programming languages. The extended scientific computation (XSC) languages provide powerful tools necessary for achieving high accuracy and reliability. They provide a large number of predefined numerical data types and operations to deal with uncertain data.

PASCAL-XSC [46] is a general purpose programming language. Compared with PASCAL it provides an extended set of mathematical functions that are available for the types `real`, `complex`, `interval` and `cinterval` (complex interval) and delivers a result of maximum accuracy. Routines

for solving numerical problems have been implemented in PASCAL-XSC. PASCAL-XSC systems are available for personal computers, workstations, mainframes and supercomputers.

Similar remarks hold for the languages C-XSC [45] and FORTRAN-XSC [89].

ACRITH-XSC [40] is an extension of FORTRAN 77. It was developed in a joint project between IBM/Germany and the Institute of Applied Mathematics of the University of Karlsruhe (U. Kulisch). Unfortunately, it can be used only on machines with IBM/370 architecture that operates under the VMCMS operating system. It is a FORTRAN like programing library. Its features are dynamic arrays, subarrays, interval and vector arithmetic and problem solving routines for mathematical problems with verified results.

In the last section of the paper [50] one can find a general discussion of the availability of the necessary arithmetic for automatic result verification in hardware and suitable programming support. A detailed information of latest developments in the group of U. Kulisch can be found under `http://www.uni-karlsruhe.de/~iam`.

Via `http://interval.usl.edu/kearfott` one can get an overview on software written in the Computer Science Department of the University of South Louisiana, Lafayette, under the guidance of R. Baker Kearfott. Here is a short outline of available software:

- INTBIS (FORTRAN 77 code to find all solutions to polynomial systems of equations),
- INTLIB (ACM TOMS Algorithm 737 – A FORTRAN 77 library for interval arithmetic and for rigorous bounds on the ranges of standard functions),
- INTERVAL ARITHMETIC (A FORTRAN 77 module that uses INTLIB to define an interval data type).

Programmer's Runtime Optimized Fast Library (PROFIL) developed at the Technical University of Hamburg–Harburg (S.M. Rump) is a C++ class library which has available usual real operations and the corresponding ones for intervals. Presently, the following data types are supported: `int`, `real`, `interval`, vectors and matrices for these types and complex numbers. For more details see `http://www.ti3.tu-harburg.de/Software/PROFIL.html`.

Recently, Rump announced the availability of an interval arithmetic package for MATLAB, called "INTLAB – A MATLAB library for interval arithmetic routines". Elements (toolboxes) of INTLAB are

- arithmetic operations for real and complex intervals, vectors and matrices over those, including sparse matrices,
- rigorous (real) standard functions,
- automatic differentiation including interval data,
- automatic slopes including interval data,
- multiple precision including interval data,
- rigorous input and output,
- some sample verification routines.

All INTLAB code is written in MATLAB for best portability. There is exactly one exception to that statement, that is one assembly language routine for switching the rounding mode of the processor (provided for some hardware platform).

Major objective of INTLAB is speed and ease of use. The first is achieved by a special concept for arithmetic routines, the second by the operator concept in MATLAB.

INTLAB code is easy to read and to write, almost as a specification. INTLAB is available for WINDOWS and UNIX systems, prerequisite is MATLAB Version 5. For more details and downloading see `http://www.ti3.tu-harburg.de/rump/intlab/`.

References

[1] E. Adams, U. Kulisch (Eds.), Scientific Computing with Automatic Result Verification, Academic Press, Boston, 1993.
[2] R. Albrecht, G. Alefeld, H.J. Stetter (Eds.), Validation Numerics, Theory and Applications, Springer, Wien, 1993.
[3] G. Alefeld, Intervallrechnung über den komplexen Zahlen und einige Anwendungen, Ph.D. Thesis, Universität Karlsruhe, Karlsruhe, 1968.
[4] G. Alefeld, Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten, Comput. Suppl. 1 (1977) 15–19.
[5] G. Alefeld, Bounding the slope of polynomials and some applications, Computing 26 (1981) 227–237.
[6] G. Alefeld, Berechenbare Fehlerschranken für ein Eigenpaar unter Einschluß von Rundungsfehlern bei Verwendung des genauen Skalarprodukts, Z. Angew. Math. Mech. 67 (1987) 145–152.
[7] G. Alefeld, Rigorous error bounds for singular values of a matrix using the precise scalar product, in: E. Kaucher, U. Kulisch, C. Ullrich (Eds.), Computerarithmetic, B.G. Teubner, Stuttgart, 1987, pp. 9–30.
[8] G. Alefeld, Über das Divergenzverhalten des Intervall-Newton-Verfahrens, Computing 46 (1991) 289–294.
[9] G. Alefeld, Inclusion methods for systems of nonlinear equations – the interval Newton method and modifications, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 7–26.
[10] G. Alefeld, A. Gienger, G. Mayer, Numerical validation for an inverse matrix eigenvalue problem, Computing 53 (1994) 311–322.
[11] G. Alefeld, A. Gienger, F. Potra, Efficient numerical validation of solutions of nonlinear systems, SIAM J. Numer. Anal. 31 (1994) 252–260.
[12] G. Alefeld, J. Herzberger, Einführung in die Intervallrechnung, Bibliographisches Institut, Mannheim, 1974.
[13] G. Alefeld, J. Herzberger, Introduction to Interval Computations, Academic Press, New York, 1983.
[14] G. Alefeld, R. Hoffmann, G. Mayer, Verification algorithms for generalized singular values, Math. Nachr. 208 (1999) 5–29.
[15] G. Alefeld, V. Kreinovich, G. Mayer, On the shape of the symmetric, persymmetric, and skew-symmetric solution set, SIAM J. Matrix Anal. Appl. 18 (1997) 693–705.
[16] G. Alefeld, V. Kreinovich, G. Mayer, The shape of the solution set of linear interval equations with dependent coefficients, Math. Nachr. 192 (1998) 23–36.
[17] G. Alefeld, V. Kreinovich, G. Mayer, On the solution sets of particular classes of linear systems, submitted for publication.
[18] G. Alefeld, R. Lohner, On higher order centered forms, Computing 35 (1985) 177–184.
[19] G. Alefeld, G. Mayer, The Cholesky method for interval data, Linear Algebra Appl. 194 (1993) 161–182.
[20] G. Alefeld, G. Mayer, A computer-aided existence and uniqueness proof for an inverse matrix eigenvalue problem, Int. J. Interval Comput. 1994 (1) (1994) 4–27.
[21] G. Alefeld, G. Mayer, On the symmetric and unsymmetric solution set of interval systems, SIAM J. Matrix Anal. Appl. 16 (1995) 1223–1240.
[22] H. Beeck, Über Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten, Computing 10 (1972) 231–244.
[23] H. Behnke, F. Goerisch, Inclusions for eigenvalues of selfadjoint problems, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 277–322.
[24] X. Chen, A verification method for solutions of nonsmooth equations, Computing 58 (1997) 281–294.
[25] G.F. Corliss, Computing narrow inclusions for definite integrals, MRC Technical Summary Report # 2913, University of Madison, Madison, Wisconsin, February 1986.
[26] G.F. Corliss, Introduction to validated ODE solving, Technical Report No. 416. Marquette University, Milwaukee, Wisconsin, March 1995.

[27] H. Cornelius, R. Lohner, Computing the range of values of real functions with accuracy higher than second order, Computing 33 (1984) 331–347.
[28] H.-J. Dobner, Einschließungsalgorithmen für hyperbolische Differentialgleichungen, Thesis, Universität Karlsruhe, Karlsruhe, 1986.
[29] P.S. Dwyer, Linear Computations, Wiley, New York, 1951.
[30] P. Eijgenraam, The solution of initial value problems using interval arithmetic. Formulation and analysis of an algorithm, Thesis, Mathematisch Centrum, Amsterdam, 1981.
[31] W. Enger, Intervall Ray Tracing – Ein Divide-and-Conquer Verfahren für photorealistische Computergrafik, Thesis, Universität Freiburg, Freiburg, 1990.
[32] A. Gienger, Zur Lösungsverifikation bei Fredholmschen Integralgleichungen zweiter Art, Thesis, Universität Karlsruhe, Karlsruhe, 1997.
[33] G.H. Golub, C.F. van Loan, Matrix Computations, 3rd Edition, John Hopkins, Baltimore, 1995.
[34] A. Griewank, G.F. Corliss (Eds.), Automatic Differentiation of Algorithms, SIAM, Philadelphia, PA, 1992.
[35] H. Grell, K. Maruhn, W. Rinow (Eds.), Enzyklopädie der Elementarmathematik, Band I Arithmetik, Dritte Auflage, VEB Deutscher Verlag der Wissenschaften, Berlin, 1966.
[36] R. Hammer, M. Hocks, U. Kulisch, D. Ratz, Numerical Toolbox for Verified Computing I, Springer, Berlin, 1993.
[37] E. Hansen, An overview of global optimization using interval analysis, in: R.E. Moore (Ed.), Reliability in Computing, The Role of Interval Methods in Scientific Computing, Perspectives in Computing, Vol. 19, Academic Press, Boston, 1988, pp. 289–307.
[38] E. Hansen, Global Optimization Using Interval Analysis, Dekker, New York, 1992.
[39] P. Hertling, 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, January 1998.
[40] IBM High Accuracy Arithmetic-Extended Scientific Computation (ACRITH-XSC), General Information GC 33-646-01, IBM Corp., 1990.
[41] C. Jansson, Rigorous sensitivity analysis for real symmetric matrices with interval data, in: E. Kaucher, S.M. Markov, G. Mayer (Eds.), Computer Arithmetic, Scientific Computation and Mathematical Modelling, IMACS Annals on Computing and Applied Mathematics, Vol. 12, J.C. Baltzer AG, Scientific Publishing, Basel, 1994, pp. 293–316.
[42] C. Jansson, On self-validating methods for optimization problems, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 381–438.
[43] E.W. Kaucher, W.L. Miranker, in: Self-Validating Numerics for Function Space Problems, Computations with Guarantees for Differential and Integral Equations, Academic Press, Orlando, 1984.
[44] R.B. Kearfott, A review of techniques in the verified solution of constrained global optimization problems, in: R.B. Kearfott, V. Kreinovich (Eds.), Applications of Interval Computations, Kluwer, Dordrecht, 1996, pp. 23–59.
[45] R. Klatte, U. Kulisch, C. Lawo, M. Rauch, A. Wiethoff, C-XSC. A C^{++} Class Library for Extended Scientific Computing, Springer, Berlin, 1993.
[46] R. Klatte, U. Kulisch, M. Neaga, D. Ratz, C. Ullrich, PASCAL-XSC, Language Reference with Examples, Springer, Berlin, 1992.
[47] M. Koeber, Lösungseinschließung bei Anfangswertproblemen für quasilineare hyperbolische Differentialgleichungen, Thesis, Universität Karlsruhe, Karlsruhe, 1997.
[48] R. Krawczyk, Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken, Computing 4 (1969) 187–201.
[49] U. Kulisch, Grundzüge der Intervallrechnung, in: Jahrbuch Überblicke Mathematik, Vol. 2, Bibliographisches Institut, Mannheim, 1969.
[50] U. Kulisch, Numerical algorithms with automatic result verification, Lectures in Applied Mathematics, Vol. 32, 1996, pp. 471–502.
[51] J.-R. Lahmann, Eine Methode zur Einschließung von Eigenpaaren nichtselbstadjungierter Eigenwertprobleme und ihre Anwendung auf die Orr-Sommerfeld-Gleichung, Thesis, Universität Karlsruhe, Karlsruhe, 1999.
[52] B. Lang, Lokalisierung und Darstellung von Nullstellenmengen einer Funktion, Diploma Thesis, Universität Karlsruhe, Karlsruhe, 1989.

[53] R. Lohner, Einschließung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen, Thesis, Universität Karlsruhe, Karlsruhe, 1988.
[54] R. Lohner, Enclosing all eigenvalues of symmetric matrices, in: C. Ullrich, J. Wolff von Gudenberg (Eds.), Accurate Numerical Algorithms. A Collection of Research Papers, Research Reports, ESPRIT, Project 1072, Diamond, Vol. 1, Springer, Berlin, 1989, pp. 87–103.
[55] G. Mayer, Old and new aspects for the interval Gaussian algorithm, in: E. Kaucher, S.M. Markov, G. Mayer (Eds.), Computer Arithmetic, Scientific Computation and Mathematical Modelling, IMACS Annals on Computing and Applied Mathematics, Vol. 12, J.C. Baltzer AG, Scientific Publishing, Basel, 1991, pp. 329–349.
[56] G. Mayer, Enclosures for eigenvalues and eigenvectors, in: L. Atanassova, J. Herzberger (Eds.), Computer Arithmetic and Enclosure Methods, Elsevier, Amsterdam, 1992, pp. 49–67.
[57] G. Mayer, Result verification for eigenvectors and eigenvalues, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 209–276.
[58] G. Mayer, Epsilon-inflation in verification algorithms, J. Comput. Appl. Math. 60 (1995) 147–169.
[59] G. Mayer, Epsilon-inflation with contractive interval functions, Appl. Math. 43 (1998) 241–254.
[60] G. Mayer, J. Rohn, On the applicability of the interval Gaussian algorithm, Reliable Comput. 4 (1998) 205–222.
[61] O. Mayer, Über die in der Intervallrechnung auftretenden Räume und einige Anwendungen, Thesis, Universität Karlsuhe, Karlsruhe, 1968.
[62] C. Miranda, Un' osservazione su un teorema di Brouwer, Bol. Un. Mat. Ital. Ser. II 3 (1941) 5–7.
[63] R.E. Moore, Interval Arithmetic and Automatic Error Analysis in Digital Computing, Thesis, Stanford University, October 1962.
[64] R.E. Moore, Interval Analysis, Prentice-Hall, Englewood Cliffs, NJ, 1966.
[65] M.T. Nakao, State of the art for numerical computations with guaranteed accuracy, Math. Japanese 48 (1998) 323–338.
[66] N.S. Nedialkov, Computing rigorous bounds on the solution of an initial value problem for an ordinary differential equation, Thesis, University of Toronto, Toronto, 1999.
[67] M. Neher, An enclosure method for the solution of linear ODEs with polynomial coefficients, Numer. Funct. Anal. Optim. 20 (1999) 779–803.
[68] A. Neumaier, The enclosure of solutions of parameter-dependent systems of equations, in: R.E. Moore (Ed.), Reliability in Computing. The Role of Interval Methods in Scientific Computing, Perspectives in Computing, Vol. 19, Academic Press, Boston, 1988, pp. 269–286.
[69] A. Neumaier, Interval Methods for Systems of Equations, University Press, Cambridge, 1990.
[70] H.T. Nguyen, V. Kreinovich, V. Nesterov, M. Nakumura, On hardware support for interval computations and for soft computing: theorems, IEEE Trans. Fuzzy Systems 5 (1) (1997) 108–127.
[71] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.
[72] W. Oettli, W. Prager, Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides, Numer. Math. 6 (1964) 405–409.
[73] M. Petković, L.D. Petković, Complex Interval Arithmetic and Its Applications, Wiley, New York, 1998.
[74] M. Plum, Inclusion methods for elliptic boundary value problems, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 323–379.
[75] L.B. Rall, Computational Solution of Nonlinear Operator Equations, Wiley, New York, 1969.
[76] L.B. Rall, Automatic Differentiation: Techniques and Applications, Lecture Notes in Computer Science, Vol. 120, Springer, Berlin, 1981.
[77] H. Ratschek, Centered forms, SIAM J. Numer. Anal. 17 (1980) 656–662.
[78] H. Ratschek, J. Rokne, Computer Methods for the Range of Functions, Ellis Horwood, Chichester, 1984.
[79] H. Ratschek, J. Rokne, New Computer Methods for Global Optimization, Ellis Horwood, Chichester, UK, 1988.
[80] R. Rihm, Interval methods for initial value problems in ODE's, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 173–207.
[81] S.M. Rump, Solving algebraic problems with high accuracy, in: U.W. Kulisch, W.L. Miranker (Eds.), A New Approach to Scientific Computation, Academic Press, New York, 1983, pp. 53–120.
[82] S.M. Rump, Rigorous sensitivity analysis for systems of linear and nonlinear equations, Math. Comp. 54 (1990) 721–736.

[83] S.M. Rump, On the solution of interval linear systems, Computing 47 (1992) 337–353.
[84] S.M. Rump, Verification methods for dense and sparse systems of equations, in: J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994, pp. 63–135.
[85] A. Schrijver, Theory of Linear and Integer Programming, Wiley, New York, 1986.
[86] 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, Berlin, 1981.
[87] P.S. Shary, Solving the linear interval tolerance problem, Math. Comput. Simulation 39 (1995) 53–85.
[88] T. Sunaga, Theory of an interval algebra and its application to numerical analysis, RAAG Memoirs 2 (1958) 29–46.
[89] W.V. Walter, FORTRAN-XSC: a portable FORTRAN 90 module library for accurate and reliable scientific computing, in: R. Albrecht, G. Alefeld, H.J. Stetter (Eds.), Validation Numerics, Theory and Applications, Springer, Wien, 1993, pp. 265–285.

Further reading

[1] G. Alefeld, A. Frommer, B. Lang (Eds.), Scientific Computing and Validated Numerics, Mathematical Research, Vol. 90, Akademie Verlag, Berlin, 1996.
[2] G. Alefeld, R.D. Grigorieff (Eds.), Fundamentals of Numerical Computation, Computer-Oriented Numerical Analysis, Computing Supplementum, Vol. 2, Springer, Wien, 1980.
[3] G. Alefeld, J. Herzberger (Eds.), Numerical Methods and Error Bounds, Mathematical Research, Vol. 89, Akademie Verlag, Berlin, 1996.
[4] L. Atanassova, J. Herzberger (Eds.), Computer Arithmetic and Enclosure Methods, Elsevier, Amsterdam, 1992.
[5] H. Bauch, K.-U. Jahn, D. Oelschlägel, H. Süße, V. Wiebigke, Intervallmathematik, Theorie und Anwendungen, Mathematisch-Naturwissenschaftliche Bibliothek, Vol. 72, BSB B.G. Teubner, Leipzig, 1987.
[6] E. Hansen (Ed.), Topics in Interval Analysis, Oxford University Press, Oxford, 1969.
[7] J. Herzberger (Ed.), Topics in Validated Computations, Elsevier, Amsterdam, 1994.
[8] J. Herzberger (Ed.), Wissenschaftliches Rechnen, Eine Einführung in das Scientific Computing, Akademie Verlag, Berlin, 1995.
[9] S.A. Kalmykov, J.I. Shokin, S.C. Yuldashev, Methods of Interval Analysis, Novosibirsk, 1986 (in Russian).
[10] E. Kaucher, U. Kulisch, C. Ullrich (Eds.), Computerarithmetic, B.G. Teubner, Stuttgart, 1987.
[11] E. Kaucher, S.M. Markov, G. Mayer (Eds.), in: Computer Arithmetic, Scientific Computation and Mathematical Modelling, IMACS Annals on Computing and Applied Mathematics, Vol. 12, J.C. Baltzer AG, Scientific Publishing, Basel, 1991.
[12] R.B. Kearfott, V. Kreinovich (Eds.), Applications of Interval Computations, Kluwer, Dordrecht, 1996.
[13] V. Kreinovich, A. Lakeyev, J. Rohn, P. Kahl, Computational Complexity and Feasibility of Data Processing and Interval Computations, Kluwer, Dordrecht, 1998.
[14] U. Kulisch (Ed.), Wissenschaftliches Rechnen mit Ergebnisverifikation, Vieweg, Braunschweig, 1989.
[15] U.W. Kulisch, W.L. Miranker, Computer Arithmetic in Theory and Practice, Academic Press, New York, 1981.
[16] U.W. Kulisch, W.L. Miranker (Eds.), A New Approach to Scientific Computation, Academic Press, New York, 1983.
[17] U. Kulisch, W. Miranker, The arithmetic of the digital computer: a new approach, SIAM Rev. 28 (1986) 1–40.
[18] U. Kulisch, H.J. Stetter (Eds.), Scientific Computation with Automatic Result Verification, Computing Supplementum, Vol. 6, Springer, Wien, 1988.
[19] S.M. Markov (Ed.), Scientific Computation and Mathematical Modelling, DATECS Publishing, Sofia, 1993.
[20] W.L. Miranker, R.A. Toupin (Eds.), Accurate Scientific Computations, Lecture Notes in Computer Science, Vol. 235, Springer, Berlin, 1986.
[21] R.E. Moore, Methods and Applications of Interval Analysis, SIAM, Philadelphia, 1979.
[22] R.E. Moore, Computational Functional Analysis, Ellis Horwood, Chichester, 1985.
[23] R.E. Moore (Ed.), Reliability in Computing. The Role of Interval Methods in Scientific Computing, Perspectives in Computing, Vol. 19, Academic Press, Boston, 1988.
[24] K. Nickel (Ed.), Interval Mathematics, Lecture Notes in Computer Science, Vol. 29, Springer, Berlin, 1975.

[25] K.L.E. Nickel (Ed.), Interval Mathematics 1980, Academic Press, New York, 1980.
[26] K. Nickel (Ed.), Interval Mathematics 1985, Lecture Notes in Computer Science, Vol. 212, Springer, Berlin, 1985.
[27] M. Petković, Iterative Methods for Simultaneous Inclusion of Polynomial Zeros, Lecture Notes in Mathematics, Vol. 1387, Springer, Berlin, 1989.
[28] C. Ullrich (Ed.), Computer Arithmetic and Self-Validating Numerical Methods, Academic Press, Boston, 1990.
[29] C. Ullrich (Ed.), Contributions to Computer Arithmetic and Self-Validating Numerical Methods, IMACS Annals on Computing and Applied Mathematics, Vol. 7, J.C. Baltzer AG, Scientific Publishing, Basel, 1990.
[30] C. Ullrich, J. Wolff von Gudenberg (Eds.), Accurate Numerical Algorithms. A Collection of Research Papers, Research Reports, ESPRIT, Project 1072, Diamond, Vol. 1, Springer, Berlin, 1989.
[31] T. Csendes (Ed.), Developments in Reliable Computing, Kluwer, Dordrecht, 1999.