ELSEVIER

Journal of Computational and Applied Mathematics 60 (1995) 147–169

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

Epsilon-inflation in verification algorithms

Günter Mayer *

Fachbereich Mathematik, Universität Rostock, Universitätsplatz 1, D-18051 Rostock, Germany

Received 23 October 1993; revised 9 May 1994

Abstract

Epsilon-inflation is often used in verification numerics to find an interval vector $[x]_\epsilon$ such that some interval function $[f]$ maps $[x]_\epsilon$ into itself. We recall algorithms which use epsilon-inflation to verify this subset property $[f]([x]_\epsilon) \subseteq [x]_\epsilon$, and we derive criteria which guarantee that already finitely many inflation steps are sufficient to prove it. These criteria require $[f]$ to be a P-contraction. We recall results on P-contractions, and we derive rules to verify them. We also introduce local P-contractions, and we use this concept to guarantee again the subset property above when using epsilon-inflation.

Keywords: Epsilon-inflation; P-contraction; Contraction; Verification algorithms; Self-validating methods; Interval computation; Nonlinear equations; Eigenvalues; Singular values

Dedicated to Prof.Dr. Lothar Berg on the occasion of his 65th birthday

1. Introduction

In recent years, self-validating methods have grown to an important tool in numerical mathematics. Many of these methods use interval arithmetic and are often based on the so called *epsilon-inflation*, a term which has been introduced by Rump [21] (cf. also [6]). Using epsilon-inflation in interval computation means to enlarge a compact interval $[x]$ to some interval $[x]_\epsilon$ which contains $[x]$ in its interior $\operatorname{int}([x]_\epsilon)$, and which differs from $[x]$ in some sense by a real parameter $\epsilon > 0$. Epsilon-inflation for vectors and matrices is defined entrywise where ϵ is allowed to depend on the entries.

The reason why one inflates is illustrated by the following situation. Assume that $f : \mathbb{R} \to \mathbb{R}$ is a continuous function for which a fixed point is to be proved. If there is an interval function $[f] : \mathbb{IR} \to \mathbb{IR}$ on the set $\mathbb{IR}$ of real compact intervals such that

$$f(x) \in [f]([x]), \quad \text{for all } x \in [x] \text{ and any } [x] \in \mathbb{IR}, \tag{1.1}$$

* e-mail: guenter.mayer@mathematik.uni-rostock.de.

0377-0427/95/$09.50 © 1995 Elsevier Science B.V. All rights reserved
SSDI 0377-0427(94)00089-1

then f has a fixed point x^* in $[x]$, provided

$$[f]([x]) \subseteq [x] \quad \text{or} \quad [f]([x]) \subseteq \operatorname{int}([x]) \tag{1.2}$$

can be shown where $\operatorname{int}([x])$ denotes the interior of $[x]$. This follows from Brouwer's fixed-point theorem for f, since (1.1) and (1.2) imply that f maps $[x]$ into itself. To achieve (1.2), one usually starts with an approximation $\tilde{x}$ of x^* which one computes by numerical standard algorithms. One then iterates according to the following pseudo-algorithm in which epsilon-inflation is involved and in which $k_{\max}$ denotes a given positive integer.

Algorithm 1.

(1) $[x] := [\tilde{x}, \tilde{x}]$
(2) **choose** $\epsilon > 0$
    $k := 0$
(3) **repeat**
        $k := k + 1$
        $[y] := [x]_\epsilon$ (ϵ-inflation)
        $[x] := [f]([y])$
    until $[x] \subseteq [y]$ **or** $k = k_{\max}$
(4) **if** $k = k_{\max}$ **then goto** (2) **or stop**

There are essentially two ways to obtain $[x]_\epsilon$.
The first one consists in an *ansatz*

$$[x]_\epsilon := \tilde{x} + [-\epsilon, \epsilon], \tag{1.3}$$

where $\epsilon > 0$ must be computed such that (1.2) holds at once, i.e., one finishes step (3) in Algorithm 1 with $k = 1$.

The second way is a *method of trial and error*. One sets $\epsilon = 0.1$, e.g., defines η to be some small number (the smallest machine number, e.g.), and computes $[x]_\epsilon$ from $[x] = [\underline{x}, \overline{x}] \in \mathbb{IR}$ according to one of the following possibilities:

(i) $[x]_\epsilon = [x] + [-\epsilon, \epsilon]$;

(absolute inflation)

(ii) $$[x]_\epsilon = \begin{cases} [x] + \epsilon[-|\underline{x}|, |\overline{x}|], & \text{if } \underline{x} \cdot \overline{x} \neq 0, \\ [x] + [-\eta, \eta], & \text{if } \underline{x} \cdot \overline{x} = 0; \end{cases}$$

(inflation relatively to the bounds $\underline{x}, \overline{x}$)

(iii) $[x]_\epsilon = [x] + \epsilon[-|\underline{x}|, |\overline{x}|] + [-\eta, \eta]$;

(iv) $$[x]_\epsilon = \begin{cases} [x] + |[x]|[-\epsilon, \epsilon], & \text{if } [x] \neq 0, \\ [x] + [-\eta, \eta], & \text{if } [x] = 0; \end{cases}$$

(inflation relatively to the absolute value $|[x]|$)

(v) $[x]_\epsilon = [x] + |[x]|[-\epsilon, \epsilon] + [-\eta, \eta]$;

(vi) $$[x]_\epsilon = \begin{cases} [x] + d([x])[-\epsilon, \epsilon] = (1+\epsilon)[x] - \epsilon[x], & \text{if } d([x]) := \overline{x} - \underline{x} \neq 0, \\ [x] + [-\eta, \eta], & \text{if } d[x] = 0; \end{cases}$$

(inflation relatively to the diameter $d([x])$)

(vii) $$[x]_\epsilon = [x] + d([x])[-\epsilon, \epsilon] + [-\eta, \eta] = (1+\epsilon)[x] - \epsilon[x] + [-\eta, \eta];$$

(viii) $$\begin{cases} [y] = (1+\epsilon)[x] - \epsilon[x], \\ [x]_\epsilon = [\underline{y}', \overline{y}'], \end{cases}$$

where $\underline{y}'$, $\overline{y'}$ denote the next preceding and the next succeeding machine number of $\underline{y}$ and $\overline{y}$, respectively.

The list is certainly incomplete. The last definition for $[x]_\epsilon$ is implemented in the scientific language PASCAL-XSC (cf. [9]).

Having defined $[x]_\epsilon$ according to one of the eight possibilities above, one runs through step (3) of Algorithm 1 several times hoping that (1.2) will be fulfilled. Often $k_{\max} = 3$, and ϵ is replaced by 5ϵ, if step (2) is reached via step (4) (cf. [7], e.g.) This justifies the name "trial and error". It is obvious that the ideas above can be generalized to vector functions and matrix functions in a straightforward way.

It is a numerical experience that the method of trial and error yields to $[x] \subseteq [y]$ in Algorithm 1 after a few iterations only, provided $\tilde{x}$ approximates x^* sufficiently well and ϵ in $[x]_\epsilon$ is chosen appropriately. Nevertheless it is an open question whether (1.2) will be always obtained theoretically.

For linear systems of equations it is shown in [23] that epsilon-inflation will guarantee (1.2) if $k_{\max}$ in Algorithm 1 is sufficiently large. We will generalize this result to nonlinear systems $f(x) = x$, provided that the interval function $[f]$ in (1.1) is a so-called *P-contraction* (cf. Section 2) on the set $\mathbb{IR}^n$ of interval vectors with n components. We prove a similar result if $[f]$ fulfils the P-contraction property only locally.

We have arranged our paper as follows. In Section 2 we present the notation used later on. In Section 3 we recall classes of examples on which epsilon-inflation is applied. In Section 4 we prove that P-contractions are contractions in the usual sense, if one introduces an appropriate vector norm. In addition, we show how to construct P-contractions. In Section 5 we state and prove results on epsilon-inflation. In Section 6 we illustrate the theory by some numerical examples.

2. Notations

By $\mathbb{R}^n, \mathbb{R}^{n\times n}, \mathbb{IR}, \mathbb{IR}^n, \mathbb{IR}^{n\times n}$ we denote the set of real vectors with n components, the set of real $n \times n$ matrices, the set of intervals, the set of interval vectors with n components and the set of $n \times n$ interval matrices, respectively. By "interval" we always mean a real compact interval. We write interval quantities in brackets with the exception of point quantities (i.e., degenerate interval quantities) which we identify with the element they contain. Examples are the identity matrix I, its ith column $e^{(i)}$ and the vector $e = (1, 1, \ldots, 1)^T$ with n or $n+1$ components (depending on the context). With $[z] \in \mathbb{IR}^n$ we define the subset $I([z]) := \{[x] | [x] \subseteq [z]\}$ of $\mathbb{IR}^n$. We apply the notation $[x] = ([x]_i) = [\underline{x}, \overline{x}] = ([\underline{x}_i, \overline{x}_i]) \in \mathbb{IR}^n$ simultaneously without further reference, and we proceed similarly with the elements of $\mathbb{R}^n$, $\mathbb{R}^{n\times n}$, $\mathbb{IR}$ and $\mathbb{IR}^{n\times n}$. By $\operatorname{int}([a])$, we denote the

topological *interior* of an interval $[a]$, by $\check{a}$ we mean its *midpoint* and by $[a]\cup[b]$ we abbreviate the *convex hull* of the union $[a]\cup[b]$ for $[a],[b]\in\mathbb{IR}$. We define the *absolute value* $|[a]|$ by $|[a]| := \max\{|a|\,|a\in[a]\}$, the *diameter* $d([a])$ by $d([a]) := \overline{a}-\underline{a}$ and the *distance* $q([a],[b])$ by $q([a],[b]) := \max\{|\underline{a}-\underline{b}|,|\overline{a}-\overline{b}|\}$. For interval vectors and interval matrices these terms are applied entrywise. Continuity in $\mathbb{IR}$ and $\mathbb{IR}^n$ is to be understood with respect to q. If $f(x)$ is an expression for some function f, we write $f([x])$ for the interval arithmetic evaluation of this expression (cf. [4]) assuming that $f([x])$ exists. Note, that we distinguish $f([x])$ from the value $[f]([x])$ of the interval function $[f]$ in (1.1). Certainly, $[f]$ can be defined by $[f]([x]) := f([x])$, but it need not necessarily be given in this way.

For details on the interval arithmetic we refer to [4,5,12,16–18]. Here, we only mention the formulae

$$\left.\begin{aligned} &q([x],[y]) \leqslant q([x],[z]) + q([z],[y]),\\ &q([x]+[z],[y]+[z]) = q([x],[y]),\\ &q([w]+[x],[y]+[z]) \leqslant q([w],[y]) + q([x],[z]),\\ &q([A][x],[A][y]) \leqslant |[A]|q([x],[y]), \end{aligned}\right\} \quad \begin{aligned} &\text{for } [w],[x],[y],[z]\in\mathbb{IR}^n,\\ &[A]\in\mathbb{IR}^{n\times n}. \end{aligned} \tag{2.1}$$

We write $\rho(A)$ for the spectral radius of a matrix $A\in\mathbb{R}^{n\times n}$, and we denote by $A\geqslant 0$ a *nonnegative* $n\times n$ matrix, i.e., $a_{ij}\geqslant 0$ for $i,j=1,\ldots,n$. We call $x\in\mathbb{R}^n$ *positive* writing $x>0$ if $x_i>0$, $i=1,\ldots,n$.

As in [4], we define $[f]:\mathbb{IR}^n\to\mathbb{IR}^n$ to be a *P-contraction* if there is a nonnegative matrix $P\in\mathbb{R}^{n\times n}$ with $\rho(P)<1$ such that

$$q([f]([x]),[f]([y])) \leqslant Pq([x],[y]), \tag{2.2}$$

for all $[x],[y]\in\mathbb{IR}^n$.

If $[f]$ fulfils (2.2) only for all $[x],[y]\subseteq[z]$ with some given $[z]\in\mathbb{IR}^n$, we call $[f]$ a *P-contraction on* $[z]$. Note that this definition is a natural extension of the P-contraction in [20], which is defined for real functions $f:[z]\to\mathbb{R}^n$ as follows.

Definition 2.1 (*Ortega and Rheinboldt* [20]). Let $f:[z]\in\mathbb{IR}^n\to\mathbb{R}^n$ fulfil

$$|f(x)-f(y)| \leqslant P|x-y|, \quad \text{for all } x,y\in[z],$$

with $0\leqslant P\in\mathbb{R}^{n\times n}$ and $\rho(P)<1$. Then f is called a *P-contraction on* $[z]$.

Remember that $q([x,x],[y,y]) = |x-y|$ for point intervals $[x,x]$ and $[y,y]$.

Similarly, we define $[f]:\mathbb{IR}^n\to\mathbb{IR}^n$ to be a *contraction* (*with respect to some vector norm* $\|\cdot\|$) if there is a real constant $\alpha\in(0,1)$ such that

$$\|q([f]([x]),[f]([y]))\| \leqslant \alpha\|q([x],[y])\| \tag{2.3}$$

holds for all $[x],[y]\in\mathbb{IR}^n$.

If $[f]$ fulfils (2.3) only for the interval vector subsets of some given $[z]\in\mathbb{IR}^n$, we call $[f]$ a *contraction* on $[z]$.

Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. Then we term $\|\cdot\|$ *monotone* if $|x| \leqslant |y|$ implies $\|x\| \leqslant \|y\|$. A particular monotone norm is the maximum norm, which we will denote, as usual, by $\|\cdot\|_\infty$. Applied to matrices this symbol means the row sum norm.

3. Examples where epsilon-inflation applies

We want to present now examples in verification numerics where epsilon-inflation is used to prove the self-mapping property of functions in fixed-point theorems which we select to be applied. To this end we assume that some function $g : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is sufficiently smooth (continuous at least) and that its domain D is appropriate for our considerations. Furthermore we suppose that the interval arithmetic evaluation $g([x])$ exists wherever it is needed. With g we construct some function f the values $f(x)$ of which we enclose by some interval function $[f]$ (mostly the interval arithmetic evaluation $[f]([x]) := f([x])$).

Fulfilling

$$[f]([x]) \subseteq [x] \tag{3.1}$$

or

$$[f]([x]) \subseteq \operatorname{int}([x]) \tag{3.2}$$

for some interval vector $[x]$ (derived by inflation, e.g.) will mean

$$f(x^*) = x^*, \tag{3.3}$$

and, with the exception of the subsequent Example 3.1(a),

$$g(x^*) = 0, \tag{3.4}$$

for some vector $x^* \in [x]$. Often (3.2) even guarantees uniqueness of x^* when f, g are restricted to $[x]$.

In Examples 3.1 and 3.2 Brouwer's fixed-point theorem is used to prove (3.3), provided (3.1) or (3.2) are fulfilled.

Example 3.1 (*Systems of linear and nonlinear equations*).

(a) $f(x) := g(x), [f]([x]) := f([x])$.

The properties (3.1) and (3.2) both imply $f(x^*) = g(x^*) = x^*$ for some $x^* \in [x]$.

(b) $$f(x) := \tilde{x} - Cg(\tilde{x}) + \left(I - C\int_0^1 g'(\tilde{x} + t(x - \tilde{x}))\,\mathrm{d}t\right)(x - \tilde{x}), \quad C \in \mathbb{R}^{n\times n},$$

$$[f]([x]) := \tilde{x} - Cg(\tilde{x}) + \left(I - Cg'(\tilde{x} \cup [x])\right)([x] - \tilde{x}).$$

Property (3.1) implies (3.3) and also (3.4) provided C is nonsingular. Property (3.2) implies (3.3) and (3.4) with $x^* \in [x]$ being unique; in addition, it *proves* C to be nonsingular. Iterating with $[f]$ yields the well-known Krawczyk method.

(c) $$f(x) := \tilde{x} - \left(\int_0^1 g'(\tilde{x} + t(x - \tilde{x}))\,\mathrm{d}t\right)^{-1} g(\tilde{x}), \quad [f]([x]) := \tilde{x} - \mathrm{IGA}(g'([x]), g(\tilde{x})),$$

where we assume that the inverse exists and that the interval Gaussian algorithm (cf. [4]), producing the vector IGA$(\cdot,\cdot)$, is feasible. Here, (3.1) implies (3.3) and (3.4) as well as the uniqueness of x^*.

(d) Choosing $g(x) := Ax - b$ in (a) and (b) results in well-known enclosure methods for linear systems of equations. An ansatz of the form (1.3) for (b) yields

$$\epsilon \geqslant \frac{\|C(b - A\tilde{x})\|_\infty}{1 - \|I - CA\|_\infty}$$

as a sufficient criterion for (3.1) provided

$$\|I - CA\|_\infty < 1 \tag{3.5}$$

holds. Note that (3.5) implies the nonsingularity of A and C.

Example 3.2 (*Generalized eigenvalue problem, eigenvalue problem*; *cf.* [1,3,13,14,21,22]). Let A, $B \in \mathbb{R}^{n\times n}$, $\tilde{x} \in \mathbb{R}^n$, $\tilde{\lambda} \in \mathbb{R}$, $\alpha \in \mathbb{R}\backslash\{0\}$, $s \in \{1,\ldots,n\}$ fixed.

(a) $$g(x,\lambda) := \begin{pmatrix} Ax - \lambda Bx \\ (e^{(s)})^{\mathrm{T}} x - \alpha \end{pmatrix},$$

i.e., $g : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$,

$$f(x,\lambda) := \begin{pmatrix} \tilde{x} \\ \tilde{\lambda} \end{pmatrix} - Cg(\tilde{x},\tilde{\lambda}) + \left\{ I - C \begin{pmatrix} A - \tilde{\lambda}B & -Bx \\ (e^{(s)})^{\mathrm{T}} & 0 \end{pmatrix} \right\} \begin{pmatrix} x - \tilde{x} \\ \lambda - \tilde{\lambda} \end{pmatrix},$$

$C \in \mathbb{R}^{(n+1)\times(n+1)}$, $[f]([x],[\lambda]) := f([x],[\lambda])$.

Replace $[x]$ in (3.1) and (3.2) by $([x],[\lambda])$, and use (x^*,λ^*) in (3.3,, (3.4) instead of x^*. Then (3.1) implies (3.3), and also (3.4) provided C is nonsingular. In this case, $([x],[\lambda])$ encloses some eigenpair (x^*,λ^*) of the generalized eigenvalue problem $Ax = \lambda Bx$ with the normalization $x_s^* = \alpha$. Property (3.2) again implies uniqueness of (x^*,λ^*) with respect to $([x],[\lambda])$.

An ansatz like (1.3) yields $\epsilon \in [\beta_1,\beta_2]$ with

$$\beta_{1,2} := \frac{(1-\sigma) \pm \sqrt{(1-\sigma)^2 - 4\rho\tau}}{2\tau},$$

$$\rho := \|Cg(\tilde{x},\tilde{\lambda})\|_\infty, \qquad \sigma := \left\| I - C \begin{pmatrix} A - \tilde{\lambda}B & -B\tilde{x} \\ (e^{(s)})^{\mathrm{T}} & 0 \end{pmatrix} \right\|_\infty, \qquad \tau := \left\| C \begin{pmatrix} B \\ 0 \end{pmatrix} \right\|_\infty.$$

This can be seen analogously to [1].

For $B = I$, one gets, of course, enclosures for an eigenpair of A.

(b) $g(x,\lambda) := Ax - \lambda Bx$, $f(y) := -Cg(\tilde{x},\tilde{\lambda}) + (I - C\hat{A})y + CBy_s\hat{y}$, with the following notation:

$\tilde{x}_s \neq 0$, for some fixed $s \in \{1,\ldots,n\}$,

$y := (x_1 - \tilde{x}_1, \ldots, x_{s-1} - \tilde{x}_{s-1}, \lambda - \tilde{\lambda}, x_{s+1} - \tilde{x}_{s+1}, \ldots, x_n - \tilde{x}_n)^{\mathrm{T}}$,

$\hat{y} := (y_1, \ldots, y_{s-1}, 0, y_{s+1}, \ldots, y_n)^{\mathrm{T}}$,

$(A - \tilde{\lambda}B)_i$ is the ith column of $A - \tilde{\lambda}B$,

$$\hat{A} := ((A - \tilde{\lambda}B)_1, \ldots, (A - \tilde{\lambda}B)_{s-1}, -B\tilde{x}, (A - \tilde{\lambda}B)_{s+1}, \ldots, (A - \tilde{\lambda}B)_n),$$

$$C \in \mathbb{R}^{n\times n}, \qquad [f]([y]) := f([y]).$$

The assertions of (a) hold analogously with $\alpha = \tilde{x}_s$ and $(x^*, \lambda^*) \in (\tilde{x} + [\hat{y}], \tilde{\lambda} + [y]_s)$. Again an ansatz is possible resulting in similar conditions as in (a); cf. [1] for details.

Example 3.3 (*Singular value problem*; *cf.* [2]). Let $A \in \mathbb{R}^{m\times n}$ have full rank and let $\tilde{x} \in \mathbb{R}^n$, $\tilde{y} \in \mathbb{R}^m$, $\tilde{\sigma}, \tilde{\tau} \in \mathbb{R}$.

$$g(x, y, \sigma, \tau) := \begin{pmatrix} Ax - \sigma y \\ A^{\mathrm{T}}y - \tau x \\ x^{\mathrm{T}}x - 1 \\ y^{\mathrm{T}}y - 1 \end{pmatrix}, \quad \text{i.e., } g : \mathbb{R}^{m+n+2} \to \mathbb{R}^{m+n+2},$$

$$f(z) := (\tilde{x}^{\mathrm{T}}, \tilde{y}^{\mathrm{T}}, \tilde{\sigma}, \tilde{\tau})^{\mathrm{T}} - Cg(\tilde{x}, \tilde{y}, \tilde{\sigma}, \tilde{\tau}) + (I - C\hat{A})(z - \tilde{z}) + C\hat{z},$$

with the following notation: I_k is the $k \times k$ identity matrix, $C \in \mathbb{R}^{(m+n+2)\times(m+n+2)}$,

$$z := \begin{pmatrix} x - \tilde{x} \\ y - \tilde{y} \\ \sigma - \tilde{\sigma} \\ \tau - \tilde{\tau} \end{pmatrix} \in \mathbb{R}^{m+n+2}, \qquad \hat{z} := \begin{pmatrix} (\sigma - \tilde{\sigma})(y - \tilde{y}) \\ (\tau - \tilde{\tau})(x - \tilde{x}) \\ -(x - \tilde{x})^{\mathrm{T}}(x - \tilde{x}) \\ -(y - \tilde{y})^{\mathrm{T}}(y - \tilde{y}) \end{pmatrix} \in \mathbb{R}^{m+n+2},$$

$$\hat{A} := \begin{pmatrix} A & -\sigma I_m & -\tilde{y} & 0 \\ -\tau I_n & A^{\mathrm{T}} & 0 & -\tilde{x} \\ 2\tilde{x}^{\mathrm{T}} & 0 & 0 & 0 \\ 0 & 2\tilde{y}^{\mathrm{T}} & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(m+n+2)\times(m+n+2)}.$$

Again, (3.1) and (3.2) imply (3.3) if $[x]$ and x^* are replaced by

$$[z] = \begin{pmatrix} [x] \\ [y] \\ [\sigma] \\ [\tau] \end{pmatrix} \quad \text{and} \quad z^* = \begin{pmatrix} x^* - \tilde{x} \\ y^* - \tilde{y} \\ \sigma^* - \tilde{\sigma} \\ \tau^* - \tilde{\tau} \end{pmatrix},$$

respectively. Property (3.1) also implies $g(z^*) = 0$ provided C is nonsingular. Property (3.2) even guarantees this nonsingularity and, in addition, uniqueness of z^* within $[z]$.

The vector $\tilde{x} + [x]$ encloses x^* which is a right singular vector of A, associated with the right singular value $\sigma^* \in \tilde{\sigma} + [\sigma]$ and normalized by $(x^*)^{\mathrm{T}}x^* = 1$. Analogously, $\tilde{y} + [y]$ encloses the normalized left singular vector y^* which corresponds to the left singular value $\tau^* \in \tilde{\tau} + [\tau]$. As in Example 3.3 an ansatz is possible based on the general criterion

$$[a] \subseteq [b] \Leftrightarrow |\check{a} - \check{b}| \leqslant \tfrac{1}{2}(d([b]) - d([a]))$$

for intervals $[a], [b]$. Cf. [2] for details.

We now turn over to initial-value problems

$$\begin{cases} x'(t) = k(t, x(t)), & t > t_0, \ k \text{ sufficiently smooth,} \\ x(t_0) = x^{(0)}, \end{cases} \tag{3.6}$$

which we will consider in their equivalent integral form

$$x(t) = g(x) := x^{(0)} + \int_{t_0}^{t} k(s, x(s))\,\mathrm{d}s. \tag{3.7}$$

Here Banach's fixed-point theorem can be applied to guarantee the existence and the uniqueness of a solution x^* of (3.6) and (3.7), at least in some nondegenerate interval $[t_0, \bar{t}]$.

Example 3.4 (*Initial-value problem*; *cf.* [16]). $f(x) := g(x)$, choose $s_1 > t_0$, inflate $[x]^{(0)} := x^{(0)}$ to $[x]^{(1)} := [x]_\epsilon^{(0)}$ and adjust $t = t_1 \leqslant s_1$ as large as possible such that

$$[f]([x]^{(1)}) := [x]^{(0)} + [t_0, t_1]k([t_0, s_1], [x]^{(1)}) \subseteq [x]^{(1)}.$$

Thus $f(x(t)) \in [x]^{(1)}$, $t_0 \leqslant t \leqslant t_1$, for all continuous functions $x : [t_0, t_1] \to \mathbb{R}$.

Inflate now $[x]^{(1)}$ to $[x]^{(2)} := [x]_\epsilon^{(1)}$, then choose $s_2 > t_1$ and $t_2 \in (t_1, s_2]$ as large as possible such that

$$[x]^{(1)} + [t_1, t_2]k([t_1, s_2], [x]^{(2)}) \subseteq [x]^{(2)}.$$

Repeating this process up to $t = t_n$ produces an interval function

$$[x](t) := [x]^{(i)}, \quad t_i \leqslant t \leqslant t_{i+1}, \quad i = 0, 1, \ldots, n-1,$$

which contains a solution x^* of (3.6) and (3.7); x^* is unique with respect to $[x](t), t_0 \leqslant t \leqslant t_n$.

4. P-contractions

In this section we derive results on P-contractions which are partly known.

Theorem 4.1 (cf. [4]). *Let $[f] : \mathbb{IR}^n \to \mathbb{IR}^n$ be a P-contraction. Then $[f]$ is continuous. Furthermore there is a monotone norm $\|\cdot\|_P$ in $\mathbb{R}^n$ such that $[f]$ is a contraction in the usual sense when defining the metric in $\mathbb{IR}^n$ by $\|q([x], [y])\|_P$. Each sequence of iterates $[x]^{k+1} := [f]([x]^k), k = 0, 1, \ldots,$ converges to the same limit $[x]^*$, which is the unique fixed point of $[f]$.*

If

$$[f](x) \in \mathbb{R}^n \tag{4.1}$$

holds for all $x \in \mathbb{R}^n$, then $[x]^*$ is a point vector.

If some function $f : \mathbb{R}^n \to \mathbb{R}^n$ satisfies the inclusion property

$$f(x) \in [f]([x]) \tag{4.2}$$

for all $x \in [x]$ and any $[x] \in \mathbb{IR}^n$, then $[x]^*$ contains all fixed points of f. If f is continuous, it has at least one fixed point in $[x]^*$.

If, in addition to (4.1), *the property* (4.2) *holds, then f is a P-contraction in the sense of* [20]. *It has a unique fixed point which can be identified with $[x]^*$.*

Proof. The continuity of $[f]$ follows directly from the definition of a P-contraction. By the continuity of the spectral radius [19] and by the Theorem of Perron and Frobenius [24], there is a $\delta > 0$ and a positive vector x such that

$$\rho := \rho(P + \delta ee^{\mathrm{T}}) < 1 \quad \text{and} \quad (P + \delta ee^{\mathrm{T}})x = \rho x, \tag{4.3}$$

for the matrix P in (2.2). Let $D_x := \operatorname{diag}(x_1, \ldots, x_n) \in \mathbb{R}^{n\times n}$ be the diagonal matrix with x_i as ith diagonal element for $i = 1, \ldots, n$. Define the monotone vector norm $\|\cdot\|_P := \|D_x^{-1} y\|_\infty$ for $y \in \mathbb{R}^n$. Then for the corresponding matrix norm we get

$$\|P\|_P \leqslant \|D_x^{-1}(P + \delta ee^{\mathrm{T}})D_x\|_\infty = \rho < 1;$$

hence by (2.2),

$$\|q(f([x]), f([y]))\|_P \leqslant \rho \cdot \|q([x], [y])\|_P.$$

This proves $[f]$ to be a contraction. Since $(\mathbb{IR}^n, \|q(\cdot,\cdot)\|_P)$ is a complete metric space, the existence and the uniqueness of $[x]^*$ follow from Banach's fixed-point theorem. The remaining part of the theorem follows from Brouwer's fixed-point theorem, or is trivial. □

Theorem 4.1 states, in particular, that any P-contraction is a contraction in the usual sense. The question arises quite naturally whether the properties *P-contraction* and *contraction with respect to some norm* are equivalent. For $n = 1$ this is certainly the case. For $n > 1$, however, the following example, which I owe to Martin Koeber, answers this question negatively even for functions $f : \mathbb{R}^n \to \mathbb{R}^n$ which are P-contractions in the sense of [20].

Example 4.2 (*Koeber* [10]). Let

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2, \\ x \mapsto \frac{1}{2}\|x\|_\infty \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \end{cases}$$

Then f is a contraction with respect to $\|\cdot\|_\infty$, but not a P-contraction.

Proof. Obviously,

$$\|f(x) - f(y)\|_\infty = \tfrac{1}{2}\big|\|x\|_\infty - \|y\|_\infty\big| \leqslant \tfrac{1}{2}\|x - y\|_\infty,$$

for all $x, y \in \mathbb{R}^n$; hence f is a contraction with respect to $\|\cdot\|_\infty$.

Assume now that f is a P-contraction. Then,

$$|f(x)| = |f(x) - f(0)| = q(f(x), f(0)) \leqslant P|x|,$$

for all $x \in \mathbb{R}^n$. Choosing $x = (1, 0)^{\mathrm{T}}$ and $x = (0, 1)^{\mathrm{T}}$, respectively, implies

$$\frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \leqslant \begin{pmatrix} p_{11} \\ p_{21} \end{pmatrix} \quad \text{and} \quad \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \leqslant \begin{pmatrix} p_{12} \\ p_{22} \end{pmatrix},$$

respectively. Thus,

$$P \geqslant \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{hence } \rho(P) \geqslant \rho\left(\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\right) = 1,$$

which contradicts the definition of a P-contraction. □

Example 4.3. Let $[f]: \mathbb{IR}^n \to \mathbb{IR}^n$ be defined by

$$[f]([x]) := [A][x] + [b], \tag{4.4}$$

where $[A] \in \mathbb{IR}^{n\times n}$ satisfies $\rho(|[A]|) < 1$. Then, by (2.1), $[f]$ is a P-contraction.

The following lemma is useful to verify P-contractions for more complicated functions than (4.4).

Lemma 4.4. *Let* $\underline{f}, \overline{f}: \mathbb{R} \to \mathbb{R}$ *satisfy*

$$|\underline{f}(x) - \underline{f}(y)| \leqslant \alpha|x - y|, \qquad |\overline{f}(x) - \overline{f}(y)| \leqslant \alpha|x - y|, \tag{4.5}$$

for some fixed $\alpha > 0$ *and for all* $x, y \in \mathbb{R}$. *Assume that* $[f]: \mathbb{IR} \to \mathbb{IR}$ *is defined by*

$$[f]([x]) := [\underline{f}(x_1), \overline{f}(x_2)], \tag{4.6}$$

where x_1, x_2 *are some elements from* $[x]$ *with*

$$\underline{f}(x_1) = \min\{\underline{f}(x) | x \in [x]\} \leqslant \overline{f}(x_2) = \max\{\overline{f}(x) | x \in [x]\}.$$

Then

$$q([f]([x]), [f]([y])) \leqslant \alpha q([x], [y]) \tag{4.7}$$

holds.

Proof. Let $[x], [y] \in \mathbb{IR}$ and let x_i, y_i, $i = 1, 2$, be the elements from $[x], [y]$ which define $[f]([x])$ and $[f]([y])$, respectively. Without loss of generality, we assume $\underline{f}(y_1) \geqslant \underline{f}(x_1)$. Then by (4.5) we obtain

$$|\underline{f}(y_1) - \underline{f}(x_1)| = \underline{f}(y_1) - \underline{f}(x_1) \leqslant \begin{cases} \underline{f}(\overline{y}) - \underline{f}(x_1) \leqslant \alpha|\overline{y} - x_1|, \\ \overline{f}(\underline{y}) - \underline{f}(x_1) \leqslant \alpha|\underline{y} - x_1|. \end{cases}$$

If $x_1 > \overline{y}$, the upper inequality yields

$$|\underline{f}(y_1) - \underline{f}(x_1)| \leqslant \alpha|\overline{y} - \overline{x}| \leqslant \alpha q([x], [y]).$$

If $x_1 < \underline{y}$, then the lower inequality yields

$$|\underline{f}(y_1) - \underline{f}(x_1)| \leqslant \alpha|\underline{y} - \underline{x}| \leqslant \alpha q([x], [y]).$$

If $\underline{y} \leqslant x_1 \leqslant \overline{y}$, then $\underline{f}(y_1) \leqslant \underline{f}(x_1)$, hence $\underline{f}(y_1) = \underline{f}(x_1)$, which trivially implies

$$|\underline{f}(y_1) - \underline{f}(x_1)| \leqslant \alpha q([x], [y]).$$

Similarly one shows

$$|\overline{f}(y_2) - \overline{f}(x_2)| \leqslant \alpha q([x], [y]),$$

which proves (4.7). □

Example 4.5. The following functions are P-contractions:

(i) $[f]([x]) = \frac{1}{2}\arctan([x])$;

(ii) $[f]([x]) = e^{-[x]^2}$, where $[x]^2 := \{x^2 | x \in [x]\}$;

(iii) $[f]([x]) = \beta \sin([x])$, where $|\beta| < 1$;

(iv) $[f]([x]) = \gamma \cos([x])$, where $|\gamma| < 1$.

Proof. (i) Apply Lemma 4.4 with $\underline{f}(x) = \overline{f}(x) = \frac{1}{2}\arctan x$, $x_1 = \underline{x}$, $x_2 = \overline{x}$, and

$$\alpha = \max\{(\tfrac{1}{2}\arctan x)' | x \in \mathbb{R}\} = \max\left\{\frac{1}{2(1+x^2)} | x \in \mathbb{R}\right\} = \tfrac{1}{2} < 1.$$

(ii) Apply Lemma 4.4 with $\underline{f}(x) = \overline{f}(x) = e^{-x^2}$,

$$x_1 = \begin{cases} \overline{x}, & \text{if } |\overline{x}| \geqslant |\underline{x}|, \\ \underline{x}, & \text{if } |\overline{x}| < |\underline{x}|, \end{cases} \qquad x_2 = \begin{cases} 0, & \text{if } 0 \in [x], \\ \underline{x}, & \text{if } 0 < \underline{x}, \\ \overline{x}, & \text{if } 0 > \overline{x}, \end{cases}$$

and

$$\alpha = \max\{2xe^{-x^2} | x \in \mathbb{R}^n\} = \frac{2}{\sqrt{2}}e^{-1/2} = \sqrt{\frac{2}{e}} < 1.$$

(iii) Apply Lemma 4.4 with

$$\underline{f}(x) = \overline{f}(x) = \beta \sin x, \qquad x_1 = \min\{\beta \sin x | x \in [x]\}, \qquad x_2 = \max\{\beta \sin x | x \in [x]\}$$

and $\alpha = \max\{\beta \cos x | x \in \mathbb{R}^n\} = |\beta| < 1$.

(iv) is proved analogously to (iii). □

The interval arithmetic evaluation of a real function $f : \mathbb{R} \to \mathbb{R}$ is not necessarily a P-contraction, if f is a contraction. This can be seen by the example $f(x) = x - x$. Here, $f([x]) = [x] - [x] = d([x])[-1, 1]$. Choose $[x]$ symmetric with respect to 0, i.e., $-\underline{x} = \overline{x}$, and let $[y] = 0$. Then,

$$q(f([x]), f([y])) = q(d([x]) \cdot [-1, 1], 0) = d([x]) = 2|[x]| = 2q([x], 0) = 2q([x], [y]),$$

although

$$q(f(x), f(y)) = q(0, 0) = 0 \leqslant \alpha q(x, y),$$

for all $x, y \in \mathbb{R}$ and any fixed $\alpha \in [0, 1)$.

Our next lemma helps to construct P-contractions in $\mathbb{IR}^n$.

Lemma 4.6. *Let $[f] : \mathbb{IR}^n \to \mathbb{IR}^n$ be a P-contraction and let $[g] : \mathbb{IR}^n \to \mathbb{IR}^n$ be a Q-contraction. Then the following assertions hold:*

(a) *$[f] \pm [g]$ is a $(P + Q)$-contraction, provided that $\rho(P + Q) < 1$;*

(b) *$[f]([g])$ is a PQ-contraction, provided that $\rho(PQ) < 1$;*

(c) *Let $n = 1$; then $[f] \cdot [g]$ is an $(\alpha Q + \beta P)$-contraction, provided that $|[f]([x])| \leqslant \alpha$, $|[g]([x])| \leqslant \beta$ for all $[x] \in \mathbb{IR}$, and $|\alpha Q + \beta P| < 1$.*

Proof. (a) By (2.1) we get

$$q(([f]\pm[g])([x]),([f]\pm[g])([y]))$$
$$\leqslant q([f]([x]),[f]([y]))+q([g]([x]),[g]([y]))\leqslant (P+Q)q([x],[y]),$$

which proves the assertion.

(b) By the assumptions we get

$$q([f]([g]([x])),[f]([g]([y])))\leqslant Pq([g]([x]),[g]([y]))\leqslant PQq([x],[y]);$$

hence the assertion follows.

(c) Using (2.1) again, we obtain

$$\begin{aligned} q(([f]\cdot[g])([x]),([f]\cdot[g])([y])) &\leqslant q(([f]\cdot[g])([x]),[f]([x])\cdot[g]([y])) \\ &\quad + q([f]([x])\cdot[g]([y]),([f]\cdot[g])([y])) \\ &\leqslant \alpha Qq([x],[y])+Pq([x],[y])\beta \\ &= (\alpha Q+\beta P)q([x],[y]), \end{aligned}$$

which proves the lemma. □

Applying this lemma to the functions of Example 4.5 shows immediately the P-contraction property of the function $[f]:\mathbb{IR}^2\to\mathbb{IR}^2$, defined by

$$[f]([x]_1,[x]_2):=\begin{pmatrix} \frac{1}{2}\arctan(\frac{1}{2}\sin([x]_1))+\frac{1}{2}e^{-[x]_2^2} \\ \frac{1}{4}\sin([x]_1)-\frac{1}{4}\sin([x]_2) \end{pmatrix}. \tag{4.8}$$

Unfortunately, such simple functions as

$$[f]:\begin{cases} \mathbb{IR}\to\mathbb{IR}, \\ [x]\mapsto [x]\cdot[x] \end{cases}$$

are not P-contractions. This can be seen for $[x]\in\mathbb{IR}$ and $[y]=0$ from

$$\begin{aligned} q([f]([x]),[f]([y])) &= q([x]\cdot[x],0)=|[x]\cdot[x]| \\ &= |[x]|^2=|[x]|q([x],0)=|[x]|q([x],[y]), \end{aligned}$$

which follows from elementary rules of $q(\cdot,\cdot)$ and $|\cdot|$ (cf. [4,18], e.g.). For any $[z]\in\mathbb{IR}$ with $|[z]|<\frac{1}{2}$, however, $[f]$ is a P-contraction on $[z]$. This follows for $[x],[y]\subseteq[z]$ from

$$\begin{aligned} q(f([x]),f([y])) &\leqslant q([x]\cdot[x],[x]\cdot[y])+q([x]\cdot[y],[y]\cdot[y]) \\ &\leqslant |[x]|q([x],[y])+|[y]|q([x],[y])\leqslant 2|[z]|q([x],[y]). \end{aligned}$$

Therefore, we state two local forms of Theorem 4.1.

Theorem 4.7. *For a fixed vector* $[z]\in\mathbb{IR}^n$ *let* $[f]:I([z])\to I([z])$ *be a P-contraction on* $[z]$. *Then the assertions of Theorem* 4.1 *hold analogously with* $[z]$ *replacing* $\mathbb{R}^n$ *and* $I([z])$ *replacing* $\mathbb{IR}^n$.

Proof. The theorem is proved analogously to Theorem 4.1. □

Note that the definition of $[f]$ in Theorem 4.7 implies, in particular, $[f]([z]) \subseteq [z]$. Therefore, each continuous function $f : [z] \to \mathbb{IR}^n$ has at least one fixed point in $[z]$, provided

$f(x) \in [f]([x])$ holds for all $x \in [x]$ and any $[x] \in I([z])$;

no P-contraction property is needed for this observation which, again, is based on Brouwer's fixed-point theorem. If $[f]$ is inclusion monotone, i.e.,

$[x] \subseteq [y] \subseteq [z]$ implies $[f]([x]) \subseteq [f]([y])$,

then the fixed points of $[f]$ are contained in each of the iterates

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

when starting this process with $[x]^0 = [z]$.

Theorem 4.8. *For a fixed vector $[z] \in \mathbb{IR}^n$ and an $n \times n$ matrix $P \geqslant 0$ with $\rho(P) < 1$ let $[z]_P \in \mathbb{IR}^n$ be a vector which satisfies*

$$[z]_P \supseteq [z] + (I - P)^{-1} d([z])[-1, 1]. \tag{4.9}$$

Let $[f] : I([z]_P) \to \mathbb{IR}^n$ be a P-contraction on $[z]_P$ with the contraction matrix P from (4.9). Choose

$$[x]^0 \subseteq [z] \tag{4.10}$$

and assume

$$[x]^1 := [f]([x]^0) \subseteq [z]. \tag{4.11}$$

Then the iterates

$$[x]^{k+1} := [f]([x]^k) \tag{4.12}$$

are defined for $k = 0, 1, \ldots$ and converge to some vector $[x]^* \subseteq [z]_P$ which is independent of $[x]^0$ as long as (4.10) and (4.11) are fulfilled.

Proof. Assume that $[x]^i \subseteq [z]_P$ holds for $i = 0, 1, 2, \ldots, k$. This is certainly true for $k = 0$ and $k = 1$. Then we get

$$\begin{aligned}
q([x]^{k+1}, [x]^0) &\leqslant \sum_{i=0}^{k} q([x]^{i+1}, [x]^i) \\
&= \sum_{i=1}^{k} q([f]([x]^i), [f]([x]^{i-1})) + q([x]^1, [x]^0) \\
&\leqslant \sum_{i=1}^{k} P q([x]^i, [x]^{i-1}) + q([x]^1, [x]^0) \\
&\leqslant \cdots \leqslant \left(\sum_{i=0}^{k} P^i \right) q([x]^1, [x]^0) \leqslant \left(\sum_{i=0}^{\infty} P^i \right) q([x]^1, [x]^0) \\
&= (I - P)^{-1} q([x]^1, [x]^0) \leqslant (I - P)^{-1} d([z]),
\end{aligned} \tag{4.13}$$

where the last inequality follows from (4.10), (4.11) and from the definition of $q(\cdot,\cdot)$. Hence,

$$[x]^{k+1} \subseteq [x]^0 + (I-P)^{-1}d([z])[-1,1] \subseteq [z]_P, \tag{4.14}$$

and $[x]^k$ exists for all nonnegative integers k. Since

$$\begin{aligned} q([x]^{k+m},[x]^k) &= q([f]([x]^{k-1+m}),[f]([x]^{k-1})) \\ &\leqslant Pq([x]^{k-1+m},[x]^{k-1}) \leqslant \cdots \leqslant P^k q([x]^m,[x]^0) \\ &\leqslant P^k(I-P)^{-1}d([z]), \quad \text{for all } m=0,1,\ldots, \end{aligned}$$

and since $\rho(P)<1$ by assumption, the sequences $\{\underline{x}^k\}, \{\overline{x}^k\}$ are Cauchy sequences in $\mathbb{R}^n$, hence they converge to limits $\underline{x}^*$ and $\overline{x}^*$, respectively, with $\underline{x}^* \leqslant \overline{x}^*$. Therefore, $\lim_{k\to\infty}[x]^k = [\underline{x}^*, \overline{x}^*] =: [x]^*$ with $[x]^* \subseteq [z]_P$ by (4.14). To show the uniqueness, let $[y]^k$ be constructed analogously to (4.12) and assume that (4.10) and (4.11) hold for $[y]^0$ and $[y]^1$. Let $[y]^*$ be the limit of $\{[y]^k\}$. Since P-contractions are continuous functions, we get

$$q([x]^*,[y]^*) = q([f]([x]^*),[f]([y]^*)) \leqslant Pq([x]^*,[y]^*) \leqslant \cdots \leqslant P^k q([x]^*,[y]^*),$$

$k=0,1,\ldots$, which implies $q([x]^*,[y]^*)=0$ and $[x]^*=[y]^*$. □

We point out that Theorem 4.8 suffers from the dependency of $[f]$ and $[z]_P$ from the same matrix P. How it can be applied is shown by the following steps for which we assume $[f]([x]) := f([x])$ with f being some function for which a fixed point is looked for.

(1) Choose a "sufficiently good" approximation $\tilde{x}$ of x^*, say, by some noninterval method.

(2) Let $[x]^0 := [\tilde{x}, \tilde{x}]$.

(3) Set $[z] := \tilde{x} + \delta[-1,1]e$ where $\delta := \|\tilde{x} - f(\tilde{x})\|_\infty$, or let $[z] := \tilde{x} + \delta[-1,1]$ where $\delta := |\tilde{x} - f(\tilde{x})|$. (Both possibilities imply $[x]^1 := [f]([x]^0) = [f(\tilde{x}), f(\tilde{x})] = \tilde{x} + [f(\tilde{x})-\tilde{x}, f(\tilde{x})-\tilde{x}] \subseteq [z]$.)

(4) Check whether $[f]$ is an αP-contraction on $[z]$ where $\alpha \in (0,1)$ is sufficiently small, say $\alpha = 0.1$, and where $0 \leqslant P \in \mathbb{IR}^{n\times n}$ with $\rho(P)<1$.

(5) Choose $[z]_P := [z] + (I-P)^{-1}d([z])[-1,1] \supseteq [z]$, and check whether $[f]$ is a P-contraction on $[z]_P$.

If $\tilde{x}$ approximates x^* sufficiently well, then δ is small, hence $d([z])$ will be small. Assuming $[f]$ to be an αP-contraction as required in (4) lets $[f]$ expect to be a P-contraction on $[z]_P$ since $[z]_P$ differs "only slightly" from $[z]$. This is heuristics, of course. That it can work is illustrated by the following example which deals with enclosures of eigenpairs and which modifies the steps (1)–(5) slightly.

Example 4.9. Let $(\tilde{x}, \tilde{\lambda})$ be an approximation of some eigenpair (x^*, λ^*) of a given matrix $A \in \mathbb{R}^{n\times n}$ with λ^* being an algebraic simple eigenvalue. Assume x^*, $\tilde{x}$ to be normalized by $x_n^* = \tilde{x}_n = 1$. We will use the notation in the Example 3.2 with $B = I$, $\alpha = 1$ and $s = n$. We choose

$$[x]^0 := \tilde{z} := \begin{pmatrix} \tilde{x} \\ \tilde{\lambda} \end{pmatrix}, \qquad \delta := \|Cg(\tilde{x},\tilde{\lambda})\|_\infty, \qquad [z] := \tilde{z} + \delta[-1,1]e \in \mathbb{R}^{n+1}.$$

Let P be a nonnegative $(n+1)\times(n+1)$ matrix to be determined later on such that

$$\|P\|_\infty \leqslant \tfrac{1}{2} \tag{4.15}$$

holds, the bound $\frac{1}{2}$ being arbitrary within the interval $(0,1)$. Then

$$[z]_P := [z] + 2d([z])[-1,1] = \tilde{z} + 5\delta[-1,1]e$$

fulfils

$$[z]_P \supseteq [z] + (I-P)^{-1}d([z])[-1,1],$$

since

$$\|(I-P)^{-1}\|_\infty = \left\|\sum_{k=0}^{\infty} P^k\right\|_\infty \leqslant \sum_{k=0}^{\infty} \|P\|_\infty^k = (1-\|P\|_\infty)^{-1} \leqslant 2.$$

Thus (4.9) of Theorem 4.8 holds, and the assumptions (4.10) and (4.11) of this theorem can also be seen at once. We will show that $[f]$ is a P-contraction on $[z]_P$ with the same P as above, provided δ is small enough, i.e., $(\tilde{x}, \tilde{\lambda})$ approximates (x^*, λ^*) sufficiently well. To this end let

$$[v] = \begin{pmatrix} [v]' \\ [v]_{n+1} \end{pmatrix} \subseteq [z]_P - \tilde{z}, \qquad [w] = \begin{pmatrix} [w]' \\ [w]_{n+1} \end{pmatrix} \subseteq [z]_P - \tilde{z},$$

and choose

$$C = \begin{pmatrix} A - \tilde{\lambda} I & -\tilde{x} \\ (e^{(n)})^{\mathrm{T}} & 0 \end{pmatrix}^{-1}.$$

It is known (cf. [15]) that the inverse of

$$\begin{pmatrix} A - \lambda I & -x \\ (e^{(n)})^{\mathrm{T}} & 0 \end{pmatrix}$$

exists at least in some neighbourhood of $(x, \lambda) = (x^*, \lambda^*)$. With (2.1) and

$$\begin{pmatrix} A - \tilde{\lambda} I & -x \\ (e^{(n)})^{\mathrm{T}} & 0 \end{pmatrix} = C^{-1} + \begin{pmatrix} O & -x + \tilde{x} \\ O & 0 \end{pmatrix},$$

we get

$$\begin{aligned} &q([f](\tilde{z} + [v]), [f](\tilde{z} + [w])) \\ &\quad = q\left(\left\{C\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right\}[v], \left\{C\begin{pmatrix} O & -[w]' \\ O & 0 \end{pmatrix}\right\}[w]\right) \\ &\quad \leqslant q\left(\left\{C\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right\}[v], \left\{C\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right\}[w]\right) \end{aligned}$$

$$
\begin{aligned}
&+ q\left(\left\{C\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right\}[w], \left\{C\begin{pmatrix} O & -[w]' \\ O & 0 \end{pmatrix}\right\}[w]\right) \\
&\leqslant \left|C\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right| q([v],[w]) + |C| q\left(\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}, \begin{pmatrix} O & -[w]' \\ O & 0 \end{pmatrix}\right)|[w]| \\
&\leqslant |C|\left\{\left|\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right| q([v],[w]) + q\left(\begin{pmatrix} -[v]' \\ 0 \end{pmatrix}, \begin{pmatrix} -[w]' \\ 0 \end{pmatrix}\right)|[w]_{n+1}|\right\} \\
&\leqslant |C|\left\{\left|\begin{pmatrix} O & -[v]' \\ O & 0 \end{pmatrix}\right| + |[w]_{n+1}| I\right\} q([v],[w]).
\end{aligned}
$$

With

$$
P := 5\delta|C| \begin{pmatrix} 1 & & & 1 \\ & \ddots & O & \vdots \\ & & \ddots & \vdots \\ & O & & \vdots \\ & & & 1 \end{pmatrix} = 5\delta|C|(I + (e - e^{(n+1)})(e^{(n+1)})^{\mathrm{T}}),
$$

we thus get

$$
q([f](\tilde{z} + [v]), [f](\tilde{z} + [w])) \leqslant Pq([v],[w]) = Pq(\tilde{z} + [v], \tilde{z} + [w]).
$$

For sufficiently small δ, i.e., for sufficiently good approximations $(\tilde{x}, \tilde{\lambda})$, the condition $\|P\|_\infty \leqslant \frac{1}{2}$ is certainly fulfilled, so that Theorem 4.8 can be applied. Note, that C depends on $(\tilde{x}, \tilde{\lambda})$, but it is bounded if $(\tilde{x}, \tilde{\lambda})$ lies in some fixed neighbourhood of (x^*, λ^*).

If it is known that $[f]$ is a P-contraction on some interval vector $[z]_P$, one often can *construct* a vector $[z]$ such that

$$
[z]_P = [z] + (I - P)^{-1} d([z])[-1, 1]
$$

holds. How this can be done is stated in our next theorem.

Theorem 4.10. *Let* $[f]$ *be a P-contraction on* $[z]_P \in \mathbb{IR}^n$. *Then*

$$
[z]_P = [z] + (I - P)^{-1} d([z])[-1, 1] \tag{4.16}
$$

implies

$$
[z] = [\underline{z}_P + (3I - P)^{-1} d([z]_P), \overline{z}_P - (3I - P)^{-1} d([z]_P)], \tag{4.17}
$$

provided that $\underline{z} \leqslant \overline{z}$.

Proof. Since (4.16) implies

$$
\underline{z}_P = \underline{z} - (I - P)^{-1}(\overline{z} - \underline{z}),
$$

we get

$$
\overline{z} = \underline{z} + (I - P)(\underline{z} - \underline{z}_P). \tag{4.18}
$$

By (4.16) and (4.18), it follows

$$\overline{z}_P = \overline{z} + (I-P)^{-1}(\overline{z} - \underline{z}) = \underline{z} + (I-P)(\underline{z} - \underline{z}_P) + \underline{z} - \underline{z}_P;$$

hence,

$$d([z]_P) = (3I - P)(\underline{z} - \underline{z}_P)$$

and

$$\underline{z} = \underline{z}_P + (3I-P)^{-1} d([z]_P). \tag{4.19}$$

Substituting (4.19) in (4.18) yields

$$\begin{aligned}\overline{z} &= \underline{z}_P + (3I-P)^{-1}d([z]_P) + (I-P)(3I-P)^{-1}d([z]_P)\\ &= \overline{z}_P - d([z]_P) + (3I-P)^{-1}(2I-P)d([z]_P)\\ &= \overline{z}_P - (3I-P)^{-1}\{3I - P - (2I-P)\}d([z]_P)\\ &= \overline{z}_P - (3I-P)^{-1}d([z]_P).\end{aligned}$$

Note that $(3I-P)^{-1} = \frac{1}{3}(I-\frac{1}{3}P)^{-1}$ exists by Neumann's series, since $\rho(\frac{1}{3}P) = \frac{1}{3}\rho(P) < \frac{1}{3} < 1$. □

The inequality $\underline{z} \leqslant \overline{z}$ in Theorem 4.10 is equivalent to

$$2(3I-P)^{-1}d([z]_P) \leqslant d([z]_P).$$

Therefore, in the case $n = 1$, we get

$$\underline{z} \leqslant \overline{z} \iff d([z]_P) \leqslant \tfrac{1}{2}(3I-P)d([z]_P),$$

where $3I - P > 2$ because of $\rho(P) < 1$. This means that $\underline{z} \leqslant \overline{z}$ is always true for $n = 1$.

For $n > 1$, this inequality can fail, as is shown by the example

$$P := \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{pmatrix}, \quad d([z]_P) = \begin{pmatrix} 0.001 \\ 1 \end{pmatrix},$$

which yields

$$\rho(P) = \tfrac{3}{4} < 1$$

and

$$\begin{aligned}2(3I-P)^{-1}d([z]_P) &= \frac{2}{3}\begin{pmatrix} \frac{5}{6} & -\frac{1}{12} \\ -\frac{1}{12} & \frac{5}{6} \end{pmatrix}^{-1}\begin{pmatrix} 0.001 \\ 1 \end{pmatrix}\\ &= \frac{8}{99}\begin{pmatrix} 10 & 1 \\ 1 & 10 \end{pmatrix}\begin{pmatrix} 0.001 \\ 1 \end{pmatrix} = \frac{8}{99}\begin{pmatrix} 1.01 \\ 10.001 \end{pmatrix} \not\leqslant d([z]_P).\end{aligned}$$

5. Success in epsilon-inflation

If the ansatz method described in Section 1 works, the conditions in (1.2) are trivially fulfilled since they are used to compute ϵ. In the method which we called "trial and error" it is a priori unknown whether (1.2) can be achieved when $k_{\max}$ in Algorithm 1 is sufficiently large. For P-contractions $[f]$, however, success is guaranteed provided that the inflated vectors $[x]_\epsilon^k$ of Algorithm 1 can be written in the form

$$[x]_\epsilon^{k+1} = [f]([x]_\epsilon^k) + [\delta]^k,$$

with $[\delta]^k$ tending to some limit $[\delta]$ which contains 0 in its interior. For inflations like (vii) in Section 1 these assumptions seem to be reasonable at least for sequences $\{[x]_\epsilon^k\}$ which converge to some limit $[x]_\epsilon^*$ since then, for $k \to \infty$, and with $[x]^{k+1} := [f]([x]_\epsilon^k)$ the interval vectors

$$[\delta]^k = d([x]^{k+1})[-\epsilon, \epsilon] + [-\eta, \eta] \text{ tend to } [\delta] := d([x]_\epsilon^*)[-\epsilon, \epsilon] + [-\eta, \eta],$$

which apparently fulfils $0 \in \operatorname{int}([\delta])$.

Note that the enclosure of rounding errors can also be incorporated in $[f]$: think of $[f]([x])$ to be the outward rounded interval arithmetic evaluation $f([x])$, so that

$$f([x]) \subseteq [f]([x])$$

holds. The crucial condition, however, is the P-contraction property for $[f]$.

Theorem 5.1. *Let $[f] : \mathbb{IR}^n \to \mathbb{IR}^n$ be a P-contraction satisfying (1.1). Iterate by inflation according to*

$$[x]_\epsilon^{k+1} := [f]([x]_\epsilon^k) + [\delta]^k, \tag{5.1}$$

where $[\delta]^k \in \mathbb{IR}^n$ are given vectors which converge to some limit $[\delta]$. If $[\delta]$ contains 0 in its interior, then there is an integer $k_0 = k_0([x]_\epsilon^0)$ such that

$$[f]([x]_\epsilon^{k_0}) \subseteq \operatorname{int}([x]_\epsilon^{k_0}) \tag{5.2}$$

holds.

Proof. Let $[h]([x]) := [f]([x]) + [\delta]$. Then by (2.1) we get

$$q([h]([x]), [h]([y])) = q([f]([x]), [f]([y])) \leqslant Pq([x], [y]); \tag{5.3}$$

hence $[h]$ is a P-contraction. By Theorem 4.1 it has a unique fixed point $[x]^*$ which satisfies

$$[x]^* = [f]([x]^*) + [\delta]. \tag{5.4}$$

Assume for the moment that

$$\lim_{k\to\infty} [x]_\epsilon^k = [x]^* \tag{5.5}$$

holds for the sequence in (5.1). By the continuity of $[f]$ we have

$$\lim_{k\to\infty} [f]([x]_\epsilon^k) = [f]([x]^*). \tag{5.6}$$

Since $0 \in \operatorname{int}([\delta])$, Eq. (5.4) implies

$$[f]([x]^*) \subseteq \operatorname{int}([x]^*).$$

Together with (5.5) and (5.6) this yields (5.2) for all sufficiently large integers k_0.

We will prove now the assumption (5.5). With (2.1) we obtain

$$\begin{aligned} q([x]_\epsilon^{k+1}, [x]^*) &= q([f]([x]_\epsilon^k) + [\delta]^k, [f]([x]^*) + [\delta]) \\ &\leqslant Pq([x]_\epsilon^k, [x]^*) + q([\delta]^k, [\delta]) \\ &\leqslant P^2 q([x]_\epsilon^{k-1}, [x]^*) + Pq([\delta]^{k-1}, [\delta]) + q([\delta]^k, [\delta]) \\ &\leqslant \cdots \leqslant P^{k+1} q([x]_\epsilon^0, [x]^*) + \sum_{i=0}^{k} P^i q([\delta]^{k-i}, [\delta]). \end{aligned} \tag{5.7}$$

Fix $\theta > 0$. Since $\rho(P) < 1$ there is an integer m such that

$$P^i \leqslant \theta e e^{\mathrm{T}}, \quad \text{for all } i \geqslant m.$$

Since $\lim_{k\to\infty} [\delta]^k = [\delta]$, there are a constant $c > 0$ and an integer $k' > m$ with

$$q([\delta]^i, [\delta]) \leqslant ce, \quad i = 0, 1, \ldots,$$

and

$$q([\delta]^{k-i}, [\delta]) \leqslant \theta e, \quad k \geqslant k', \ i = 0, 1, \ldots, m.$$

For $k \geqslant k'$ we thus get with (5.7)

$$\begin{aligned} q([x]_\epsilon^{k+1}, [x]^*) &\leqslant \theta e e^{\mathrm{T}} q([x]_\epsilon^0, [x]^*) + \sum_{i=0}^{m} P^i \theta e + P^{m+1} \sum_{i=m+1}^{k} P^{i-(m+1)} ce \\ &\leqslant \theta \{ e e^{\mathrm{T}} q([x]_\epsilon^0, [x]^*) + (I - P)^{-1} e + e e^{\mathrm{T}} (I - P)^{-1} ce \}. \end{aligned}$$

Since the expression in braces is independent of θ, m and k, and since θ can be chosen arbitrarily small, (5.5) holds. □

Let $[b] \in \mathbb{IR}^n$, $[A] \in \mathbb{IR}^{n\times n}$, and let $[D]_\epsilon \in \mathbb{IR}^{n\times n}$ be diagonal with $I \in [D]_\epsilon$ and $\rho(|[D]_\epsilon[A]|) < 1$. Define $[f]$ by $[f]([x]) := [D]_\epsilon([b] + [A][x])$. Then Theorem 5.1 is essentially identical with [23, Theorem 5].

In many cases, $[f]$ is only a local P-contraction, i.e., a P-contraction with respect to some vector $[z]$. Then based on Theorem 4.7 one can formulate a local analogon of Theorem 5.1 which we first formulate and prove for $[\delta]^k = [\delta], k = 0, 1, \ldots$.

Theorem 5.2. *Let* $[f] : \mathbb{IR}^n \to \mathbb{IR}^n$ *satisfy* (1.1) *and let* $[\delta], [z] \in \mathbb{IR}^n$ *fulfil the following properties*:

(i) $0 \in \operatorname{int}([\delta])$;

(ii) $[f]$ *is a local P-contraction on*

$$[z]_P \supseteq [z] + (I - P)^{-1} d([z])[-1, 1]. \tag{5.8}$$

If $[x]_\epsilon^0 \subseteq [z]$ *and* $[x]_\epsilon^1 \subseteq [z]$ *hold for the iterates from* (5.1) *with* $[\delta]^k := [\delta]$, *then there is an integer* $k_0 = k_0([x]_\epsilon^0)$ *such that* (5.2) *is true. In particular,* f *from* (1.1) *has a fixed point in* $[x]_\epsilon^{k_0}$.

Proof. Since $[h]([x]) := [f]([x]) + [\delta]$ fulfils (5.3) for all $[x], [y] \subseteq [z]_P$, the function $[h]$ is a local P-contraction on $[z]_P$. By Theorem 4.8 there is a vector $[x]^* \subseteq [z]_P$ which satisfies

$$\lim_{k\to\infty} [x]_\epsilon^k = [x]^* = h([x]^*) = [f]([x]^*) + [\delta]. \tag{5.9}$$

Since $0 \in \operatorname{int}([\delta])$, this yields

$$[f]([x]^*) \subseteq \operatorname{int}([x]^*), \tag{5.10}$$

and the assertion follows from (5.6), (5.10) and from the first equality in (5.9). □

To get full analogy of Theorem 5.1, condition (5.8) has to be replaced by a slightly more complicated one.

Theorem 5.3. *Let* $[f] : \mathbb{IR}^n \to \mathbb{IR}^n$ *satisfy* (1.1), *and let* $[\delta]^k$, $[\delta]$, $[z] \in \mathbb{IR}^n$ *fulfil the following properties*:

(i) $\lim_{k\to\infty} [\delta]^k = [\delta]$;
(ii) $0 \in \operatorname{int}([\delta])$;
(iii) $[f]$ *is a local P-contraction on*

$$[z]_P \supseteq [z] + (I - P)^{-1}(d([z]) + v)[-1, 1],$$

with

$$v := \max\{w, d([z]) + (I - P)^{-1} d([z])\},$$

where

$$w := \left(\max_k \{q([\delta]_i^k, [\delta]_i)\}\right) \in \mathbb{R}^n.$$

If $[x]_\epsilon^0 \subseteq [z]$ *and* $[x]_\epsilon^1 \subseteq [z]$ *hold for the iterates from* (5.1), *then there is an integer* $k_0 = k_0\left([x]_\epsilon^0\right)$ *such that* (5.2) *is true. In particular,* f *from* (1.1) *has a fixed point in* $[x]_\epsilon^{k_0}$.

Proof. As in the proof of Theorem 5.2 there is a vector $[x]^* \subseteq [z]_P$ which satisfies

$$[x]^* = [f]([x]^*) + [\delta].$$

It remains to show

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

then the proof can be finished as in Theorem 5.1. To this end, we prove by induction

$$[x]_\epsilon^k \subseteq [z]_P, \quad k = 0, 1, \ldots. \tag{5.11}$$

For $k = 0$ and $k = 1$ this holds by assumption.

Let $[x]_\epsilon^0, [x]_\epsilon^1, \ldots, [x]_\epsilon^k \subseteq [z]_P$. Then as in (5.7) we have

$$q([x]_\epsilon^{k+1}, [x]^*) \leqslant P^{k+1} q([x]_\epsilon^0, [x]^*) + \sum_{i=0}^{k} P^i q([\delta]^{k-i}, [\delta]). \tag{5.12}$$

Since $[x]_\epsilon^0 \subseteq [z]$ by assumption, and since $[x]^* \subseteq [z] + (I-P)^{-1} d([z])[-1,1]$ by Theorem 4.8, we get

$$q([x]_\epsilon^0, [x]^*) \leqslant |\underline{z} - \overline{z}| + (I-P)^{-1} d([z]) = d([z]) + (I-P)^{-1} d([z]);$$

hence (5.12) implies

$$\begin{aligned} q([x]_\epsilon^{k+1}, [x]^*) &\leqslant P^{k+1}\{d([z]) + (I-P)^{-1} d([z])\} + \sum_{i=0}^{k} P^i w \\ &\leqslant P^{k+1} v + \sum_{i=0}^{k} P^i v \leqslant \sum_{i=0}^{\infty} P^i v = (I-P)^{-1} v. \end{aligned}$$

Therefore,

$$\begin{aligned} [x]_\epsilon^{k+1} &\subseteq [x]^* + (I-P)^{-1} v[-1,1] \\ &\subseteq [z] + (I-P)^{-1} d([z])[-1,1] + (I-P)^{-1} v[-1,1] \subseteq [z]_P, \end{aligned}$$

and (5.11) is true.

The proof finishes now following the lines of Theorem 5.1. The inclusion (5.11) is needed to justify the first estimate in (5.7) which is now based on the P-contraction property on $[z]_P$. □

6. Numerical results

In our final section we illustrate our theoretical results by some examples. We start by a rather trivial one which fulfils the conditions of Theorem 5.2.

Example 6.1. Let $f(x) := x \cdot x$, $[f]([x]) := [x] \cdot [x]$, $[z] := [-2,2] \cdot 10^{-2}$, $P := \frac{1}{2}$, $[x]_\epsilon := [x] + [\delta]$ with $[\delta] := [-1,1] \cdot 10^{-4}$. Then $0 \in \operatorname{int}([\delta])$,

$$[z]_P := [z] + (I-P)^{-1} d([z])[-1,1] = [-0.1, 0.1],$$

$$\begin{aligned} q([x]\cdot[x], [y]\cdot[y]) &\leqslant q([x]\cdot[x], [x]\cdot[y]) + q([x]\cdot[y], [y]\cdot[y]) \\ &\leqslant (|[x]| + |[y]|) q([x], [y]) \\ &\leqslant 0.2\, q([x], [y]), \quad \text{for all } [x], [y] \subseteq [z]_P; \end{aligned}$$

hence $[f]$ is a local P-contraction on $[z]_P$.

With $[x]^0 := 10^{-2}$ we get

$$[x]_\epsilon^0 = [0.99, 1.01] \cdot 10^{-2} \subseteq [z],$$

$$[x]^1 = [f]([x]_\epsilon^0) = [0.9801, 1.0201] \cdot 10^{-4},$$

$$[x]_\epsilon^1 = [-0.0199, 2.0201] \cdot 10^{-4} \subseteq [z].$$

Thus the assumptions of Theorem 5.2 are fulfilled, whence (5.2) holds for some integer k_0. Indeed,

$$[x]^2 = [f]([x]_\epsilon^1) = [-0.040\,199\,99, 4.080\,804\,01] \cdot 10^{-8} \subseteq \operatorname{int}([x]_\epsilon^1).$$

Example 6.2. Let $A \in \mathbb{R}^{21\times 21}$ be the Wilkinson matrix W_{21}^+ defined by

$$A = W_{21}^+ := \begin{pmatrix} 10 & 1 & & & & & & & & & & & & & & & & & & & \\ 1 & 9 & 1 & & & & & & & & & & & & & & & & & & \\ & 1 & 8 & 1 & & & & & & & & & & & & & & & & & \\ & & \ddots & \ddots & \ddots & & & & & & & & & & & & & & & & \\ & & & & 1 & 0 & 1 & & & & & & & & & & & & & & \\ & & & & & \ddots & \ddots & \ddots & & & & & & & & & & & & & \\ & & & & & & & 1 & 8 & 1 & & & & & & & & & & & \\ & & & & & & & & 1 & 9 & 1 & & & & & & & & & & \\ & & & & & & & & & 1 & 10 & & & & & & & & & & \end{pmatrix}.$$

To verify the largest eigenvalue λ_{21}, apply Algorithm 1.1 with $\tilde{z} := \begin{pmatrix} \tilde{x} \\ \tilde{\lambda} \end{pmatrix}$ replacing $\tilde{x}$. Compute the eigenpair approximations with a software package like EISPACK or LAPACK, and choose $[x]_\epsilon$ from (viii) in Section 1, f and $[f]$ from Example 3.2 with C being an approximate inverse of

$$\begin{pmatrix} A - \tilde{\lambda} I & -\tilde{x} \\ (e^{(n)})^{\mathrm{T}} & 0 \end{pmatrix}.$$

Example 4.9 lets expect that Theorem 5.2 holds, and, in fact, after ten steps of iterations (with ϵ-inflation) one ends up with the verified inclusion

$$[\lambda] = [10.746\,194\,182\,903\,39, 10.746\,194\,182\,903\,40],$$

skipping the corresponding verification of the eigenvector.

How to choose ϵ generally in verification algorithms remains an open question. Some remarks on this problem — primarily based on experience — can be found in [11,21].

For an effective estimate of $(I - P)^{-1}$ which occurs in Theorems 4.8, 5.2 and 5.3, we refer to [8].

References

[1] G. Alefeld, Berechenbare Fehlerschranken für ein Eigenpaar unter Einschluß von Rundungsfehlern bei Verwendung des genauen Skalarprodukts, *Z. Angew. Math. Mech.* **67** (1987) 145–152.

[2] G. Alefeld, Rigorous error bounds for singular values of a matrix using the precise scalar product, in: E. Kaucher, U. Kulisch and Ch. Ullrich, Eds., *Computerarithmetic* (Teubner, Stuttgart, 1987) 9–30.

[3] G. Alefeld, Berechenbare Fehlerschranken für ein Eigenpaar beim verallgemeinerten Eigenwertproblem, *Z. Angew. Math. Mech.* **68** (1988) 181–184.

[4] G. Alefeld and J. Herzberger, *Introduction to Interval Computations* (Academic Press, New York, 1983).

[5] H. Bauch, K.-U. Jahn, D. Oelschlägel, H. Süsse and V. Wiebigke, *Intervallmathematik* (BSB Teubner, Leipzig, 1987).
[6] O. Caprani and K. Madsen, Iterative methods for interval inclusion of fixed points, *BIT* **18** (1978) 42–51.
[7] K. Grüner, Solving the complex algebraic eigenvalue problem with verified high accuracy, in: Ch. Ullrich and J. Wolff von Gudenberg, Eds., *Accurate Numerical Algorithms, A Collection of Research Papers*, Research Reports ESPRIT, Project 1072, DIAMOND, Vol. 1 (Springer, Berlin, 1989) 59–78.
[8] G. Heindl, Computable hypernorm bounds for the error of an approximate solution to a system of nonlinear equations, in: L. Atanassova and J. Herzberger, Eds., *Computer Arithmetic and Enclosure Methods* (North-Holland, Amsterdam, 1992) 451–461.
[9] R. Klatte, U. Kulisch, M. Neaga, D. Ratz and Ch. Ullrich, *PASCAL-XSC, Sprachbeschreibung mit Beispielen* (Springer, Berlin, 1991).
[10] M. Koeber, Private communication, Karlsruhe, 1993.
[11] S. König, On the inflation parameter used in self-validating methods, in: Ch. Ullrich, Ed., *Contributions to Computer Arithmetic and Self-Validating Numerical Methods*, IMACS Ann. Comput. Appl. Math. **7** (Baltzer, Basel, 1990) 127–132.
[12] G. Mayer, Grundbegriffe der Intervallrechnung, in: U. Kulisch, Ed., *Wissenschaftliches Rechnen mit Ergebnisverifikation* (Vieweg, Braunschweig, 1989) 101–117.
[13] G. Mayer, Enclosures for eigenvalues and eigenvectors, in: L. Atanassova and J. Herzberger, Eds., *Computer Arithmetic and Enclosure Methods* (North-Holland, Amsterdam, 1992) 49–67.
[14] G. Mayer, Taylor-Verfahren für das algebraische Eigenwertproblem, *Z. Angew. Math. Mech.* **73** (1993) T921–T924.
[15] G. Mayer, A unified approach to enclosure methods for eigenpairs, *Z. Angew. Math. Mech.* **74** (1994) 115–128.
[16] R.E. Moore, *Interval Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1966).
[17] R.E. Moore, *Methods and Applications of Interval Analysis* (SIAM, Philadelphia, PA, 1979).
[18] A. Neumaier, *Interval Methods for Systems of Equations* (Cambridge Univ. Press, Cambridge, 1990).
[19] J.M. Ortega, *Numerical Analysis: A Second Course*, Classics Appl. Math. **3** (SIAM, Philadelphia, PA, 1990).
[20] J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic Press, New York, 1970).
[21] S.M. Rump, Kleine Fehlerschranken bei Matrixproblemen, Thesis, Univ. Karlsruhe, 1980.
[22] S.M. Rump, Solving algebraic problems with high accuracy, in: U.W. Kulisch and W.L. Miranker, Eds., *A New Approach to Scientific Computation* (Academic Press, New York, 1983) 53–120.
[23] S.M. Rump, On the solution of interval linear systems, *Computing* **47** (1992) 337–353.
[24] R.S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962).