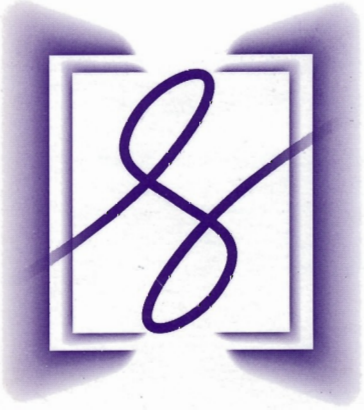
# Reliable Computing

An International Journal Devoted to Reliable Mathematical Computations Based on Finite Representations and Guaranteed Accuracy



Volume 3 No. 2 May 1997
ISSN 1385-3139

Institute of
New Technologies in
Education
St. Petersburg - Moscow
Kluwer Academic Publishers

*Reliable Computing* **3:** 103–135, 1997.



# Algebraic Approach in the "Outer Problem" for Interval Linear Equations

SERGEY P. SHARY*
*Institute of Computational Technologies, 6 Lavrentiev avenue, 630090 Novosibirsk, Russia, e-mail: shary@ict.nsc.ru.*

(Received: 27 July 1996; accepted: 20 December 1996)

**Abstract.** The subject of our work is the classical "outer" problem for the interval linear algebraic system $\mathbf{A}x = \mathbf{b}$ with the square interval matrix $\mathbf{A}$: *find "outer" coordinate-wise estimates of the united solution set* $\Sigma$ *formed by all solutions to the point systems* $Ax = b$ *with* $A \in \mathbf{A}$ *and* $b \in \mathbf{b}$. The purpose of this work is to advance a new *algebraic approach* to the formulated problem, in which it reduces to solving one *noninterval* (point) equation in the Euclidean space of double dimension. We construct a specialized algorithm (subdifferential Newton method) that implements the new approach, then present results of the numerical tests with it. These results demonstrate that the proposed algebraic approach combines unique computational efficency with high quality enclosures of the solution set.

## 1. Introduction

The subject of our work is the classical problem for the interval linear algebraic system

$$\mathbf{A}x = \mathbf{b} \tag{1.1}$$

with the interval $(n \times n)$-matrix $\mathbf{A}$ and interval right-hand side $n$-vector $\mathbf{b}$:

*Find (quick and as sharp as possible) "outer" coordinate-wise estimates of the solution set*

$$\Sigma = \{x \in \mathbb{R}^n \mid (\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b)\} \tag{1.2}$$

*formed by all solutions to the point systems* $Ax = b$ *with* $A \in \mathbf{A}$ *and* $b \in \mathbf{b}$, *that is, evaluate* $\min\{x_k \mid x \in \Sigma\}$ *from below and* $\max\{x_k \mid x \in \Sigma\}$ *from above,* $k = 1, 2, \ldots, n$.

The solution set $\Sigma$ is sometimes called the *united solution set* and denoted by $\Sigma_{\exists\exists}$. That is not superfluous since many different solution sets to interval systems $\mathbf{A}x = \mathbf{b}$ can be defined ($\Sigma_{\forall\exists}$, $\Sigma_{\exists\forall}$, $\Sigma_{\alpha\beta}$ as well as some others). These *generalized solution*

* The results of this paper have been presented as a highlighted talk at the International Conference on Interval Methods and Computer Aided Proofs in Science and Engineering (INTERVAL'96), September 30–October 2, 1996, Würzburg, Germany. The author thanks the organizer and chair of the conference INTERVAL'96 Prof. Dr. Jürgen Wolff von Gudenberg for giving him this opportunity and for kind support.

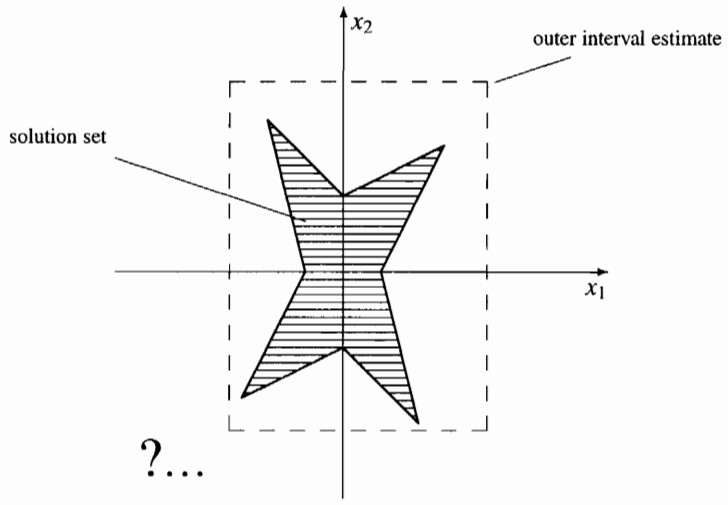



*Figure 1.* "Outer problems" are problems of *outer* (interval) estimation of the solution sets, i.e., system sensitivity problems in interval form.

*sets* to (1.1) naturally come into existence and have interesting and significant applications, for instance, in operations research and decision making [24], [25]. In this work, we shall not consider these generalized solution sets, so, for brevity, we speak simply of the *solution set* instead of the correct expanded term "united solution set" when referring to (1.2).

The main problem to which this entire paper is devoted is often formulated as follows:

> *Find (as quickly and sharply as possible) an interval enclosure of the solution set to the given interval linear equation.* (1.3)

We prefer calling it the *outer problem* for the interval linear equation (1.1) in order to distinguish (1.3) from other possible problem statements for (1.1) (considered, e.g., in [23], [26]). As we mentioned, problem (1.3) is one of the classical interval analysis problems (being in essence an implicit linear sensitivity problem in the interval form), and hundreds of papers and several books have been dedicated to its various aspects, starting in the early 60's. Among the most important theoretical results on this problem, one should mention the recent discoveries concerning the computational complexity of (1.3) and some related problems. (See [11] as well as the extensive references there.) Even the recognition problem, whether the solution set $\Sigma$ is empty or not has been proved to be intractable in general. Further, computing outer coordinatewise estimates for the solution set $\Sigma$ with a prescribed absolute or relative accuracy is intractable too, both in general and in some practically significant particular cases.

Over more than three decades, many good algorithms have been invented to solve the problem (1.3), so that one may think the time has come to finish with the abstract theory and turn to thorough technical improvements of the existing approaches. Nonetheless, in our paper, we are going to throw a fresh light on the subject. We advance one more numerical algorithm for the solution of the "outer problem," its distinctive features being

- high computational efficacy,
- good adaptability to various specific interval linear systems,
- universality—the main algorithm is equally applied to the problem of inner interval estimation of both the solution set (1.2) and the generalized solution sets to (1.1).

The new approach to (1.3) that we are proposing is a further development of the so-called *algebraic approach*, in which the original problem reduces to solving one *noninterval* (point) equation in the Euclidean space of double dimension $\mathbb{R}^{2n}$. The algebraic approach has presented itself as an efficient technique for inner interval estimation of the various solution sets to interval algebraic equations [23]–[25]. It is high time the algebraic approach was extended to outer problems as well.

Below, we construct a specialized algorithm (subdifferential Newton method) that implements the new approach, then we present results of numerical tests with it. These numerical results demonstrate that the proposed algebraic approach combines unique computational efficency with high quality enclosures of the solution set, so that it appears competitive with the popular interval Gaussian method and recent Hansen-Rohn procedure.

Henceforth, we assume familiarity with the basic facts of interval analysis; for an introduction, see, e.g., [1], [14]. Our notation follows mainly that of Kearfott (see, e.g., [16]) who has done a good job modifying and updating the earlier notation of Neumaier [14].

## 2. Fundamentals

PROPOSITION 2.1. *The solution set of the interval system*

$$\mathbf{A}x = \mathbf{b}$$

*coincides with the solution set of the interval system*

$$x = \mathbf{C}x + \mathbf{b} \tag{2.1}$$

*if* $\mathbf{C} = I - \mathbf{A}$.

*Proof.* The solution set of (2.1)

$$= \{x \in \mathbb{R}^n \mid (\exists C \in \mathbf{C})(\exists b \in \mathbf{b})(x = Cx + b)\}$$
$$= \{x \in \mathbb{R}^n \mid \big(\exists (I - C) \in (I - \mathbf{C})\big)(\exists b \in \mathbf{b})\big((I - C)x = b\big)\}$$

$$= \{x \in \mathbb{R}^n \mid (\exists A \in (I - (I - \mathbf{A}))(\exists b \in \mathbf{b})(Ax = b)\}, \quad \text{where } A := I - C$$
$$= \{x \in \mathbb{R}^n \mid (\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b)\}$$
$$= \text{the solution set of } \mathbf{A}x = \mathbf{b},$$

since if $C \in \mathbf{C}$ then $(I - C) \in (I - \mathbf{C})$. □

*Warning.* The equivalent reduction of the interval linear system $\mathbf{A}x = \mathbf{b}$ to the fixed-point form (2.1) described in Proposition 2.1 is not the only possible one. In what follows, we fix this reduction merely for definiteness, while the question will be discussed at length in Section 7.

THEOREM 2.1 [2], [13]. *The iteration*

$$\mathbf{x}^{(k+1)} = \mathbf{C}\mathbf{x}^{(k)} + \mathbf{b}, \quad k \geq 0,$$

*converges to a unique fixed point interval* $\mathbf{x}^*$ *of the equation (2.1) for every starting interval vector* $\mathbf{x}^{(0)}$ *if and only if the spectral radius* $\rho(|\mathbf{C}|)$ *of the matrix* $|\mathbf{C}|$ *made up of the moduli of the elements of* $\mathbf{C}$ *is less than* 1.

THEOREM 2.2 [1], [2]. *Let* $\mathbf{C}$ *be an interval matrix for which* $\rho(|\mathbf{C}|) < 1$. *It then follows that, for the fixed point* $\mathbf{x}^*$ *(that exists and is unique according to Theorem 2.1) of the equation* $\mathbf{x}^* = \mathbf{C}\mathbf{x}^* + \mathbf{b}$, *the relation*

$$\{(I - C)^{-1}b \mid C \in \mathbf{C},\ b \in \mathbf{b}\} \subseteq \mathbf{x}^*$$

*is valid.*

DEFINITION 2.1 [15], [18]. An interval vector is called an *algebraic solution* to an interval equation if substituting this vector into the equation and executing all interval operations according to the rules of interval arithmetic result in the equality.

Thus, taking into account this definition, we can reformulate the above classical results in the following modified form, which will be extensively used in what follows:

THEOREM 2.3. *If the interval matrix* $\mathbf{C} \in \mathbb{IR}^{n \times n}$ *is such that* $\rho(|\mathbf{C}|) < 1$, *then for any* $\mathbf{b}$, *the algebraic solution to the interval equation*

$$\mathbf{x} = \mathbf{C}\mathbf{x} + \mathbf{b} \tag{2.2}$$

*exists, is unique, and gives an interval enclosure of the solution set to this interval linear equation.*

In (2.2) and below, we intentionally designated the unknown variable $\mathbf{x}$ by a boldface letter to emphasize that the required solution is an *interval*. It must obey equation (2.2) in the sense of the *interval arithmetic operations*.

What is the meaning of the reformulation of the well-known results in the form of Theorem 2.3? The point is that the assertion of Theorem 2.3, which is arranged as a pure existence theorem, helps one to recognize the following principal fact: *The way that we prove the theorem need not be connected with the actual way that we find the solutions to the main equation (2.2).*

Of course, the traditional constructive proofs of Theorems 2.1, 2.2 based on Schröder's contraction mapping theorem gave rise to a flow of papers in which various stationary iterative methods for (2.2) have been built. Generally speaking, however, nothing obliges one to confine oneself to using only stationary contractions when developing computational approaches for (2.2), especially since the resulting procedures often prove to be rather slowly converging. The algorithm designer should feel free to exploit other concepts (even symbolic manipulation) in developing algorithms that find the algebraic solution to (2.2). The only underlying principle has to be that the sought-for solution satisfies equation (2.2) in the sense of Definition 2.1. Below, we are just going to advance such more efficient, nonstationary, algorithms.

## 3. Interval Arithmetics

It is common knowledge that the main instrument of interval analysis is the so-called *interval arithmetic*, that is, an algebraic system $\langle \mathbb{IR}, +, -, \cdot, / \rangle$, where $\mathbb{IR}$ is the set of all real intervals $[\underline{x}, \overline{x}]$, $\underline{x} \le \overline{x}$, while the binary operations—addition, subtraction, multiplication and division—are defined according to the following fundamental principle:

$$\mathbf{x} \star \mathbf{y} = \{x \star y \mid x \in \mathbf{x},\ y \in \mathbf{y}\} \tag{3.1}$$

for all intervals $\mathbf{x}, \mathbf{y}$ such that $(x \star y)$, $\star \in \{+, -, \cdot, /\}$ makes sense for any $x \in \mathbf{x}$, $y \in \mathbf{y}$ [1], [14].

Unfortunately, the algebraic properties of the interval arithmetic $\mathbb{IR}$ are bad: Most of its elements—all intervals with nonzero width—do not have inverses with respect to the above operations. As a consequence, first, the basic equations

$$\mathbf{a} + x = \mathbf{b}, \quad \mathbf{a} \cdot x = \mathbf{b}$$

and their like may not have solutions in general. Second, the manipulation technique in $\mathbb{IR}$ is quite poor. We cannot even rearrange terms from one side of the equation to the other. Clearly, these features do not favor finding the sought-for algebraic solution to equation (2.2). Besides, $\mathbb{IR}$ is not a lattice [4] with respect to the natural inclusion ordering. The first of the operations

$$\mathbf{x} \wedge \mathbf{y} = [\max\{\underline{\mathbf{x}}, \underline{\mathbf{y}}\}, \min\{\overline{\mathbf{x}}, \overline{\mathbf{y}}\}], \quad \text{— taking minimum (or } \textit{join}),$$

$$\mathbf{x} \vee \mathbf{y} = [\min\{\underline{\mathbf{x}}, \underline{\mathbf{y}}\}, \max\{\overline{\mathbf{x}}, \overline{\mathbf{y}}\}], \quad \text{— taking maximum (or } \textit{meet}),$$

is not always applicable in the classical interval arithmetic.*

* If $\mathbf{x}, \mathbf{y}$ are ordinary one-dimensional intervals intersecting with each other, then $\mathbf{x} \wedge \mathbf{y}$ and $\mathbf{x} \vee \mathbf{y}$ are simply $\mathbf{x} \cap \mathbf{y}$ and $\mathbf{x} \cup \mathbf{y}$ respectively. But in general this is not the case.

Under the circumstances, it is natural to adopt the following plan of action suggested by Shary [21]: We ought to enlarge $\mathbb{IR}$ to a wider algebraic system (or, put differently, to embed $\mathbb{IR}$ into a wider algebraic system) that would have better algebraic properties, that would be richer in manipulation technique, with more powerful analytical tools. Then we seek the algebraic solution in the new wider algebraic system rather than in the common interval arithmetic $\mathbb{IR}$. We can hope that the new problem is easier, but if the interval vector so obtained proves to lie in $\mathbb{IR}$, then it will be a solution to the original problem.

How can one perform the desired enlargement of the classical interval arithmetic? Abstract algebra has come to our rescue in this situation. From the general viewpoint, the arithmetic $\mathbb{IR}$ is a semigroup both with respect to addition and multiplication.* As is well known (see, e.g., [5]), a commutative semigroup with the so-called "cancellation law" can be extended to (or, what is equivalent, can be embedded into) a group, that is, a richer algebraic system, in which every element has an inverse. In our case, the interval arithmetic is exactly the commutative semigroup with the cancellation law, but the multiplicative semigroup satisfying the cancellation law is formed only by the intervals that do not contain zero.

Luckily, all technical constructions that are necessary to build up the extension of interval arithmetic according to the above strategy were already done by Kaucher at the end of the 1970's. In the works [9], [10], Kaucher elaborated an algebraic system called "extended interval arithmetic $\mathbb{IIR}$" that includes the common interval arithmetic $\mathbb{IR}$ as a subset and meets all our requirements. The complete description of $\mathbb{IIR}$ may be found in [7], [10].

Elements of $\mathbb{IIR}$ are pairs of real numbers $[\underline{x}, \overline{x}]$, that are not necessarily related by the condition $\underline{x} \leq \overline{x}$. Therefore, $\mathbb{IIR}$ is obtained by adjoining *improper* intervals $[\underline{x}, \overline{x}]$, $\underline{x} > \overline{x}$, to the set $\mathbb{IR} = \{[\underline{x}, \overline{x}] \mid \underline{x}, \overline{x} \in \mathbb{R}, \underline{x} \leq \overline{x}\}$ of *proper* intervals as well as real numbers identified with the corresponding degenerate intervals. Elements of the Kaucher extended interval arithmetic and other objects formed of these elements shall be denoted by boldface letters, like common intervals.

Proper and improper intervals, the two "halves" of $\mathbb{IIR}$, change places as the result of the *dualization* mapping $\mathrm{dual} : \mathbb{IIR} \to \mathbb{IIR}$, such that

$$\mathrm{dual}\,\mathbf{x} = [\overline{\mathbf{x}}, \underline{\mathbf{x}}].$$

As in classical interval arithmetic, we can define inclusion as

$$\mathbf{x} \subseteq \mathbf{y} \iff \underline{\mathbf{x}} \geq \underline{\mathbf{y}} \quad \& \quad \overline{\mathbf{x}} \leq \overline{\mathbf{y}}. \tag{3.2}$$

This definition makes Kaucher arithmetic $\mathbb{IIR}$ a lattice [4] with respect to the inclusion order relation, in contrast to $\mathbb{IR}$.

Addition and multiplication by real numbers are defined on $\mathbb{IIR}$ by

$$\mathbf{x} + \mathbf{y} := [\underline{\mathbf{x}} + \underline{\mathbf{y}}, \overline{\mathbf{x}} + \overline{\mathbf{y}}],$$

* Strictly speaking, the arithmetic $\mathbb{IR}$ is even a commutative *monoid* [5] with respect to addition and multiplication, that is, a semigroup with a neutral element, but this fact is not so important for our consideration.

$$\lambda \cdot \mathbf{x} := \begin{cases} [\lambda \underline{x}, \lambda \overline{x}], & \text{if } \lambda \in \mathbb{R}^+, \\ [\lambda \overline{x}, \lambda \underline{x}], & \text{otherwise.} \end{cases}$$

Each element $\mathbf{x}$ of $\mathbb{IR}$ has thus a unique additive inverse, denoted "opp $\mathbf{x}$", and

$$\mathbf{x} + \operatorname{opp} \mathbf{x} = 0 \quad \Rightarrow \quad \operatorname{opp}[\underline{x}, \overline{x}] = [-\underline{x}, -\overline{x}].$$

It follows directly from the definition that, with respect to addition, $\mathbb{IR}$ is a commutative group, which is isomorphic to the additive group of the standard linear space $\mathbb{R}^2$. Sometimes, we denote for brevity the inverse operation for addition, i.e., the inner (algebraic) difference of $\mathbb{IR}$, by "$\ominus$", so that

$$\mathbf{x} \ominus \mathbf{y} := \mathbf{x} + \operatorname{opp} \mathbf{y}.$$

The following distributivity properties of addition with respect to the lattice operations sometimes prove useful:

$$\mathbf{x} + (\mathbf{y} \vee \mathbf{z}) = (\mathbf{x} + \mathbf{y}) \vee (\mathbf{x} + \mathbf{z}), \tag{3.3}$$

$$\mathbf{x} + (\mathbf{y} \wedge \mathbf{z}) = (\mathbf{x} + \mathbf{y}) \wedge (\mathbf{x} + \mathbf{z}). \tag{3.4}$$

The nicest fact about Kaucher arithmetic is that the following representation that generalizes formula (3.1) holds:

$$\mathbf{x} \star \mathbf{y} = \bigcup\!\!\!\!\!\!\bigwedge^{\mathbf{x}}_{x \in \operatorname{pro} \mathbf{x}} \bigcup\!\!\!\!\!\!\bigwedge^{\mathbf{y}}_{y \in \operatorname{pro} \mathbf{y}} (x \star y), \tag{3.5}$$

where

$$\bigcup\!\!\!\!\!\!\bigwedge^{\mathbf{x}} := \begin{cases} \bigvee, & \text{if } \mathbf{x} \text{ is proper,} \\ \bigwedge, & \text{otherwise,} \end{cases} \qquad \text{— conditional lattice operation,}$$

$$\operatorname{pro} \mathbf{x} := \begin{cases} \mathbf{x}, & \text{if } \mathbf{x} \text{ is proper,} \\ \operatorname{dual} \mathbf{x}, & \text{otherwise,} \end{cases} \qquad \text{— proper projection of the interval.}$$

This representation expresses the connection between the interval operation $\mathbf{x} \star \mathbf{y}$ and the results of the point operations $x \star y$ for $x \in \operatorname{pro} \mathbf{x}$ and $y \in \operatorname{pro} \mathbf{y}$. Multiplication in Kaucher arithmetic is also defined on the basis of this representation (see, e.g., [7]). Below it will suffice to have the above general property, so that we even shall not write out the explicit formulas (they can be found, e.g., in [10], [23], [24]). The extended interval multiplication turns out to be commutative and associative [7], [9], [10]. But the multiplicative group of $\mathbb{IR}$ is formed only by intervals $[\underline{x}, \overline{x}]$ with $\underline{x}\overline{x} > 0$, since the "cancellation law" does not hold on any wider subset of $\mathbb{IR}$.

Extended arithmetic subtraction and division are defined:

$$\mathbf{x} - \mathbf{y} = \mathbf{x} + (-1) \cdot \mathbf{y},$$

$$\mathbf{x} / \mathbf{y} = \mathbf{x} \cdot [1 / \overline{y}, 1 / \underline{y}] \qquad \text{for } 0 \notin \operatorname{pro} \mathbf{y}.$$

Finally, all extended interval operations are *inclusion monotone*, similar to their classical predecessors:

$$\mathbf{x} \subseteq \mathbf{x}', \mathbf{y} \subseteq \mathbf{y}' \quad \Rightarrow \quad \mathbf{x} * \mathbf{y} \subseteq \mathbf{x}' * \mathbf{y}', \quad * \in \{+, -, \cdot, /\}.$$

The interrelation between multiplication and addition is:

$$\text{if } \mathbf{x} \text{ is proper,} \quad \mathbf{x}\cdot(\mathbf{y}+\mathbf{z}) \subseteq \mathbf{x}\cdot\mathbf{y}+\mathbf{x}\cdot\mathbf{z} \tag{3.6}$$

— subdistributivity,

$$\text{if } \mathbf{x} \text{ is improper,} \quad \mathbf{x}\cdot(\mathbf{y}+\mathbf{z}) \supseteq \mathbf{x}\cdot\mathbf{y}+\mathbf{x}\cdot\mathbf{z} \tag{3.7}$$

— superdistributivity.

These inclusions turn into exact equalities, in particular, if $\mathbf{x}$ is thin, that is, $\mathbf{x} = x \in \mathbb{R}$.

In extended interval arithmetic, the operations with vectors and matrices are defined similarly to those in $\mathbb{IR}$. The sum (difference) of two interval matrices of identical size is an interval matrix of that same size formed by elementwise sums (differences) of operands. If $\mathbf{X} = (\mathbf{x}_{ij}) \in \mathbb{IR}^{m\times l}$ and $\mathbf{Y} = (\mathbf{y}_{ij}) \in \mathbb{IR}^{l\times n}$, then the product of the matrices $\mathbf{X}$, $\mathbf{Y}$ is a matrix $\mathbf{Z} = (\mathbf{z}_{ij}) \in \mathbb{R}^{m\times n}$ such that

$$\mathbf{z}_{ij} = \sum_{k=1}^{l} \mathbf{x}_{ik}\,\mathbf{y}_{kj}.$$

The inclusion ordering on the set of interval vectors and matrices is a direct product [4] of the inclusion orders on the separate components $\mathbb{IR}$. Hence, we set

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \vee \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \vee \mathbf{y}_1 \\ \mathbf{x}_2 \vee \mathbf{y}_2 \\ \vdots \\ \mathbf{x}_n \vee \mathbf{y}_n \end{pmatrix}$$

and

$$\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_n \end{pmatrix} \wedge \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1 \wedge \mathbf{y}_1 \\ \mathbf{x}_2 \wedge \mathbf{y}_2 \\ \vdots \\ \mathbf{x}_n \wedge \mathbf{y}_n \end{pmatrix}.$$

Finally, the topology on the extended interval space $\mathbb{IR}^n$ is defined in the standard way, that is, by the metric

$$\operatorname{dist}(\mathbf{x},\mathbf{y}) := \max\{\|\underline{\mathbf{x}} - \underline{\mathbf{y}}\|, \|\overline{\mathbf{x}} - \overline{\mathbf{y}}\|\}, \quad \mathbf{x},\mathbf{y} \in \mathbb{IR}^n,$$

where $\|\cdot\|$ is a monotonic vector norm on $\mathbb{R}^n$. For $\mathbb{IR}^n$, this metric coincides with Hausdorff distance between interval vectors. All the extended arithmetical operations, the matrix-vector operations in $\mathbb{IR}^n$ as well as the operations $\vee$, $\wedge$, "dual", and "opp" are continuous in the above metric (see [10]).

## 4. Immersion into Linear Space

The result of the preceding sections is that we have reduced the original "outer problem" (1.3) to the problem of finding an algebraic solution of the interval equation

$$\mathbf{C}x \ominus x + \mathbf{b} = 0.$$

Essentially, this is the traditional mathematical problem of the solution of an equation, and much of classical numerical analysis deals with it. The peculiarity of the situation is that the main set $\mathbb{IR}^n$ on which the equation is considered is not a linear space at all: lack of distributivity violates the axiom requiring

$$(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$$

for all vectors $\mathbf{x} \in \mathbb{IR}^n$ and any scalars $\lambda, \mu \in \mathbb{R}$. So most of the existing computational approaches are not directly applicable to our problem.

Furthermore, remaining in $\mathbb{IR}^n$, we cannot carry out comprehensive theoretical analysis of the situation and understand some phenomena. For example, the point matrix

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \tag{4.1}$$

has nonzero determinant, but multiplying by this matrix in $\mathbb{IR}^n$ may nullify even a nonzero vector:

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} [-1, 1] \\ [1, -1] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

What is the reason? One can hardly reveal that from inside the interval space, which is essentially nonlinear. There is an urgent need to transfer our considerations into a linear space, which we denote by $U$ for generality.

Mathematically, we have two different spaces, the interval space $\mathbb{IR}^n$ and the linear space $U$, with totally different structures on them; how can we jump from the first one to the second? We are going to do that in the following way, which somewhat resembles a change of variables. First, we should construct a map $\iota : \mathbb{IR}^n \to U$, that must be a bijection (one-to-one map) to uniquely restore an interval preimage by its image in $U$, and conversely. Then, it is fairly simple to realize that each bijection $\iota : \mathbb{IR}^n \to U$ also induces the bijection from the set of all mappings over $\mathbb{IR}^n$ to the set of all mappings over $U$: each $\varphi : \mathbb{IR}^n \to \mathbb{IR}^n$ is matched to the unique *induced mapping*

$$\iota \circ \varphi \circ \iota^{-1} : U \to U, \tag{4.2}$$

where "$\circ$" stands for composition of mappings. Overall, the situation is described by the commutative diagram (see Figure 2).

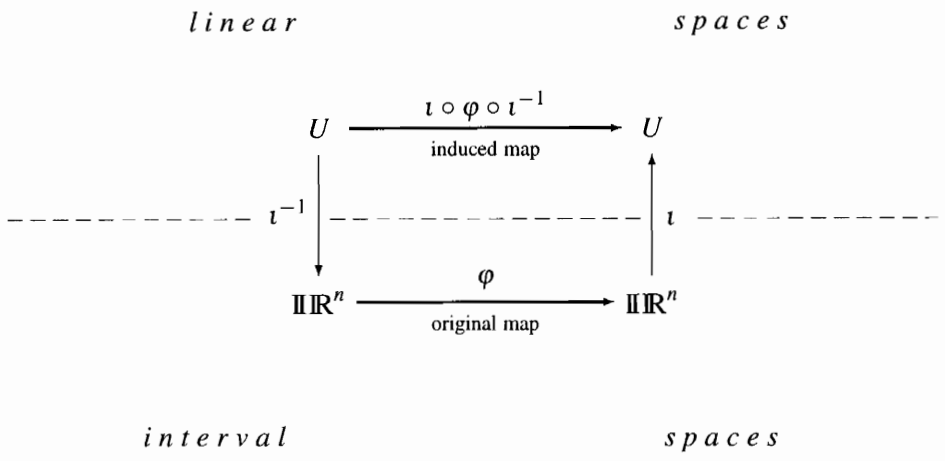



*Figure 2.* How an immersion generates induced mappings.

The fact is, the properties of $\varphi$ and $\iota \circ \varphi \circ \iota^{-1}$ are closely related. Moreover, if $\iota$ is chosen such that $\iota(0) = 0$, then we may change the problem of solution of the equation in $\mathbb{IR}^n$ to the problem of solution of the equation in the linear space $U$, a situation familiar to modern numerical analysts.

Formally, we can organize the above construction as an embedding of the interval space $\mathbb{IR}^n$ into the linear space $U$. The major question about the construction of the embedding is to provide a reasonable compromise between its simplicity and convenient form of the induced mappings (4.2). We adopt the following

DEFINITION 4.1 [23]. For a linear space $U$, a one-to-one mapping $\iota : \mathbb{IR}^n \to U$ is said to be an *immersion* of $\mathbb{IR}^n$ into $U$ provided that it satisfies the following properties:

(i) $\iota$ is an isomorphism of the additive groups $\mathbb{IR}^n$ and $U$,

(ii) $\iota$ is a homeomorphism of the topological spaces $\mathbb{IR}^n$ and $U$.

In other words, we require that the immersion keep the additive algebraic structure and the topological structure of the interval space $\mathbb{IR}^n$.

This definition immediately implies that

$$\iota(0_{\mathbb{IR}^n}) = 0_U,$$
$$\iota(\operatorname{opp} \mathbf{x}) = -\iota(\mathbf{x}), \quad \mathbf{x} \in \mathbb{IR}^n.$$

In addition, the inverse mapping $\iota^{-1} : U \to \mathbb{IR}^n$ also satisfies conditions (i)–(ii) from the definition of $\iota$, and

$$\iota^{-1}(0_U) = 0_{\mathbb{IR}^n},$$

$$\iota^{-1}(-x) = \operatorname{opp} \iota^{-1}(x), \quad x \in U.$$

Invoking, for example, dimension considerations, one can show that $U$ is uniquely determined by Definition 4.1: *$U$ must be* $\mathbb{R}^{2n}$. Further, all the embeddings satisfying Definition 4.1 are easily proved to be equivalent to within a nonsingular linear transformation of the space $\mathbb{R}^{2n}$ [23], so that we can choose a specific immersion from convenience considerations. As in our previous work [23], we shall henceforth use a specific immersion, which is quite rightfully named the *standard immersion*.

DEFINITION 4.2. The immersion $\sigma : \mathbb{IR}^n \to \mathbb{R}^{2n}$ that acts as follows

$$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \quad \mapsto \quad (-\underline{\mathbf{x}}_1, -\underline{\mathbf{x}}_2, \ldots, -\underline{\mathbf{x}}_n, \overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \ldots, \overline{\mathbf{x}}_n), \tag{4.3}$$

i.e., when the left endpoints of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ taken with the opposite signs are identified with the 1st, 2nd, ..., $n$-th components of a real $2n$-vector, while the right endpoints of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ are identified with the $(n+1)$-th, ..., $2n$-th components of the real $2n$-vector, respectively, will be called the *standard immersion* of the interval space $\mathbb{IR}^n$ into $\mathbb{R}^{2n}$.

Each immersion $\iota : \mathbb{IR}^n \to \mathbb{R}^{2n}$ naturally generates an *induced partial order* "$\sqsubseteq$" on the linear space $\mathbb{R}^{2n}$, which is an image of the inclusion order on $\mathbb{IR}^n$ under the immersion $\iota$. Specifically, for $x, y \in \mathbb{R}^{2n}$ one can say that "$x$ does not exceed $y$" and write "$x \sqsubseteq y$" if and only if $\iota^{-1}(x) \subseteq \iota^{-1}(y)$ in $\mathbb{IR}^n$. Since for any $x, y, u, v \in \mathbb{R}^{2n}$ there holds

$$x \sqsubseteq y, \ \alpha \in \mathbb{R}^+ \quad \Rightarrow \quad \alpha x \sqsubseteq \alpha y,$$
$$x \sqsubseteq y, \ u \sqsubseteq v \quad \Rightarrow \quad x + u \sqsubseteq y + v,$$

then the partial order "$\sqsubseteq$" conforms to the linear structure of $\mathbb{R}^{2n}$. Therefore, it is set by some *positive cone* $K_{\sqsubseteq} = \{x \in \mathbb{R}^{2n} \mid x \geq 0\}$ [17], so that

$$x \sqsubseteq y \quad \Longleftrightarrow \quad y - x \in K_{\sqsubseteq}.$$

Clearly, the concrete formulas for "$\sqsubseteq$" depends on the form of immersion, but for the standard immersion (4.3) they look especially simple. It is not hard to see that then

$$x \sqsubseteq y \quad \text{if and only if} \quad x \leq y \quad \text{in the componentwise sense,} \tag{4.4}$$

that is, if $x_i \leq y_i$, $i = 1, 2, \ldots, 2n$. The positive cone under the standard immersion is correspondingly

$$K_{\leq} = \{x \in \mathbb{R}^{2n} \mid x_i \geq 0, \ i = 1, 2, \ldots, 2n\}. \tag{4.5}$$

Thus the induced partial order on the space $\mathbb{R}^{2n}$ coincides with the common componentwise ordering! This is the main justification of the form (4.3) for the standard immersion that we have chosen. The above is a sufficiently compelling argument

for us to treat only the standard immersion of the form (4.3) and corresponding componentwise ordering (4.4) on $\mathbb{R}^{2n}$ in the remaining theoretical development of this paper. However, in practice, other immersions may also be useful.

COROLLARY 4.1. *Directly from the definition of an induced order on* $\mathbb{R}^{2n}$ *and in view of (4.4) we get*

$$\sigma\Big(\bigvee_{\gamma\in\Gamma}\mathbf{x}_\gamma\Big)=\sigma\Big(\max_{\gamma\in\Gamma}{}_{\subseteq}\,\mathbf{x}_\gamma\Big)=\max_{\gamma\in\Gamma}{}_{\leq}\,\sigma(\mathbf{x}_\gamma) \tag{4.6}$$

*for any family* $\{\mathbf{x}_\gamma \mid \gamma\in\Gamma\}$ *of intervals. Thus, the immersion transfers maxima in* $\mathbb{IR}^n$ *to maxima in* $\mathbb{R}^{2n}$ *(and minima to minima).*

DEFINITION 4.3 [4]. For a number $x\in\mathbb{R}$, the quantities

$$\begin{aligned} x^+ &:= \max(x,0),\\ x^- &:= \max(-x,0) \end{aligned}$$

will called the *positive part* and the *negative part* of $x$ respectively.

PROPOSITION 4.1. *If* $\varphi:\mathbb{IR}^n\to\mathbb{IR}^n$ *is an operator of multiplication by a point matrix, that is,*

$$\varphi(\mathbf{x})=Q\mathbf{x}$$

*for some* $Q\in\mathbb{R}^{n\times n}$, $Q=(q_{ij})$, *then the induced mapping* $\iota\circ\varphi\circ\iota^{-1}$ *is a linear transformation of the space* $\mathbb{R}^{2n}$.

*For the standard immersion* $\sigma$, *the matrix of the induced linear transformation* $\sigma\circ\varphi\circ\sigma^{-1}$ *is the following block* $(2n\times 2n)$*-matrix*

$$\left(\begin{array}{c|c} Q^+ & Q^- \\ \hline Q^- & Q^+ \end{array}\right), \tag{4.7}$$

*where the matrices* $Q^+=(q_{ij}^+)$ *and* $Q^-=(q_{ij}^-)$ *are the positive and negative parts of* $Q$ *respectively.*

*Proof.* The first statement immediately follows from the distributivity relation

$$q\cdot(\mathbf{x}+\mathbf{y})=q\cdot\mathbf{x}+q\cdot\mathbf{y}$$

that is valid for thin $q$. The second one is a consequence of the multiplication rule

$$q\cdot[\underline{x},\overline{x}]=\begin{cases}[q\,\underline{x},q\,\overline{x}], & \text{if } q\in\mathbb{R}^+,\\ [q\,\overline{x},q\,\underline{x}], & \text{otherwise}\end{cases}$$

and formula (4.3). □

The block $(2n \times 2n)$-matrix from Proposition 4.1 is so significant in our theory that we shall use a special designation for it.

DEFINITION 4.4. For a given $(n \times n)$-matrix $Q$, we set

$$Q^{\sigma} := \left(\begin{array}{c|c} Q^{+} & Q^{-} \\ \hline Q^{-} & Q^{+} \end{array}\right) \tag{4.8}$$

and call the matrix $Q^{\sigma}$ the *concomitant matrix* to $Q$.

COROLLARY 4.2. *It follows directly from the definition of the induced map that for any* $x \in \mathbb{R}^{2n}$ *there holds*

$$\sigma(Q\sigma^{-1}(x)) = Q^{\sigma}x. \tag{4.9}$$

The other important feature is that the concomitant matrices $Q^{\sigma} \in \mathbb{R}^{2n\times 2n}$ are always *non-negative*: Such matrices must represent "$\leq$"-isotone operators on $\mathbb{R}^{2n}$ that correspond to inclusion-isotone multiplication on $Q$ in $\mathbb{IR}^{n}$.

We have already noted that the nonsingularity of the point matrix $Q$ in the sense of classical linear algebra does not necessarily mean that the corresponding operator of multiplication by $Q$ in $\mathbb{IR}^{n}$ gives a nonzero result. That is, such a matrix may not generate an invertible operator on $\mathbb{IR}^{n}$. The matrix (4.1) is a counterexample. To distinguish such cases, we give the following

DEFINITION 4.5 [23]. We say that the matrix $Q \in \mathbb{R}^{n\times n}$ is *ι-nonsingular*, if

$$Q\mathbf{x} = 0 \quad \Longleftrightarrow \quad \mathbf{x} = 0 \in \mathbb{IR}^{n}.$$

Otherwise, we call the matrix $Q$ *ι-singular*.

Obviously, if a matrix is singular in the common sense, then, *a fortiori*, it is $\iota$-singular. As a corollary of Proposition 4.1 we get

PROPOSITION 4.2. *The point matrix* $Q \in \mathbb{R}^{n\times n}$ *is ι-nonsingular if and only if its concomitant matrix* $Q^{\sigma} \in \mathbb{R}^{2n\times 2n}$ *is nonsingular in the common sense, i.e., its determinant in nonzero.*

For example, the identity matrix

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

is $\iota$-nonsingular, while the matrix (4.1) is $\iota$-singular. All nonnegative nonsingular matrices are $\iota$-nonsingular.

COROLLARY 4.3. *The operator of multiplication by a point matrix in* $\mathbb{IIR}^n$

$$\varphi : \mathbb{IIR}^n \to \mathbb{IIR}^n, \quad \varphi(\mathbf{x}) = Q\mathbf{x} \quad \textit{for} \quad Q \in \mathbb{R}^{n\times n},$$

*has an inverse operator* $\varphi^{-1} : \mathbb{IIR}^n \to \mathbb{IIR}^n$ *if and only if the matrix* $Q$ *is* $\iota$*-nonsingular, in which case* $\varphi^{-1}$ *acts as follows*

$$\varphi^{-1}(\mathbf{x}) = \sigma^{-1}\left((Q^{\sigma})^{-1} \cdot \sigma(\mathbf{x})\right) \tag{4.10}$$

*(cf. (4.9)).*

*Warning.* Even though the explicit formula (4.10) exists, the operator which is inverse to the operator of multiplication by a point $(n \times n)$-matrix $Q$ in $\mathbb{IIR}^n$ cannot be generally expressed through multiplication by a matrix in $\mathbb{IIR}^n$ (in particular, the inverse operator is not multiplication by the matrix $Q^{-1}$).

Finally, it is worth giving a methodological commentary on the account of this section. We have singled the map $\mathbb{IIR}^n \to \mathbb{R}^{2n}$ out and studied it as a separate notion. Why? Some people would be satisfied by merely identifying the interval's endpoints with the components of the vectors of $\mathbb{R}^{2n}$, thus not introducing unnecessary abstractions.

The point is that the "mere identification" of the interval's endpoints with the components of the vectors of $\mathbb{R}^{2n}$ is nothing but a disguised trick to introduce an immersion (embedding) $\mathbb{IIR}^n \to \mathbb{R}^{2n}$. I am convinced that one had better do that honestly and in an explicit way. Besides, we have at least two reasons to treat the immersion as we did:

1) the map $\mathbb{IIR}^n \to \mathbb{R}^{2n}$ may not be uniquely defined, i.e., there is not just one such immersion that is naturally determined for all possible situations;
2) we can gain meaningful benefits from this non-uniqueness, i.e., if a map $\iota_1 : \mathbb{IIR}^n \to \mathbb{R}^{2n}$ is the most suitable for one situation, then another map $\iota_2 : \mathbb{IIR}^n \to \mathbb{R}^{2n}$ can turn out better for another.

Both these points are actually present in the case considered. The map $\sigma : \mathbb{IIR}^n \to \mathbb{R}^{2n}$, which is referred to as the "standard immersion" is appropriate mainly in theory. When implementing the practical algorithms described in this paper, the author used the simplest immersion

$$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \quad \mapsto \quad (\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \ldots, \underline{\mathbf{x}}_n, \overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \ldots, \overline{\mathbf{x}}_n),$$

which is more convenient when writing a computer code, etc. Moreover, it is not hard to imagine some situations in which immersions different from the above will prove helpful. For example, Lyashko [12] utilizes the immersion

$$(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n) \quad \mapsto \quad (\underline{\mathbf{x}}_1, \overline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \overline{\mathbf{x}}_2, \ldots, \underline{\mathbf{x}}_n, \overline{\mathbf{x}}_n)$$

to vividly represent some results on the asymptotic convergence factor of matrix iterative methods.

## 5. Investigating the Equation

As the result of the embedding, investigation of the mappings $\mathbb{IIR}^n \to \mathbb{IIR}^n$ reduces to investigation of the mappings $\mathbb{R}^{2n} \to \mathbb{R}^{2n}$ of the usual Euclidean space, but we can state even more. As follows from Definitions 4.1, 4.2, for the immersion $\sigma : \mathbb{IIR}^n \to \mathbb{R}^{2n}$

$$\sigma(\mathbf{x}) \neq 0 \text{ of } \mathbb{R}^{2n} \iff \mathbf{x} \neq 0 \text{ of } \mathbb{IIR}^n,$$

while

$$\sigma(0 \text{ of } \mathbb{IIR}^n) = 0 \text{ of } \mathbb{R}^{2n}.$$

We can thus turn the original problem that concerns us, i.e., that of finding zeros of the mapping

$$\psi(\mathbf{x}) = \mathbf{C}\mathbf{x} \ominus \mathbf{x} + \mathbf{b}, \tag{5.1}$$

into the problem of solution of the *induced equation*

$$\Psi(x) = 0 \tag{5.2}$$

in $\mathbb{R}^{2n}$ such that $\Psi = \sigma \circ \psi \circ \sigma^{-1} : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, i.e.,

$$\begin{aligned} \Psi(x) &= \sigma\big(\mathbf{C}\sigma^{-1}(x) \ominus \sigma^{-1}(x) + \mathbf{b}\big) \\ &= \sigma\big(\mathbf{C}\sigma^{-1}(x)\big) - x + \sigma(\mathbf{b}). \end{aligned} \tag{5.3}$$

Overall, the original equation $\psi(\mathbf{x}) = 0$ has a solution $\mathbf{x}^* \in \mathbb{IIR}^n$ if and only if the induced equation $\Psi(x) = 0$ has a solution $x^* \in \mathbb{R}^{2n}$, and there holds

$$\mathbf{x}^* = \sigma^{-1}(x^*).$$

Also, we do not need to prove separate existence and uniqueness results for the solutions of the induced equation (5.2)–(5.3). Those are Theorems 2.1–2.3 for the original equation (2.2).

PROPOSITION 5.1. *The induced mapping* $\Psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ *defined by (5.3) is continuous.*

*Proof.* The original mapping $\psi : \mathbb{IIR}^n \to \mathbb{IIR}^n$ defined by (5.1) is continuous, since interval arithmetic operations in $\mathbb{IIR}$ are continuous. The immersion $\sigma$ as well as its inverse $\sigma^{-1}$ are also continuous. □

Well, what about differentiability, smoothness, etc., of $\Psi$? Unfortunately, we cannot claim any of these properties globally. But we have something even more attractive instead.

To move forward, we recall the following definition.

DEFINITION 5.1. Let $\mathbb{R}^q$ be an ordered linear space with the partial order "$\preccurlyeq$". A mapping $F : \mathbb{R}^p \to \mathbb{R}^q$ is called *order convex* with respect to "$\preccurlyeq$" if

$$F\big(\lambda y + (1-\lambda)z\big) \preccurlyeq \lambda F(y) + (1-\lambda)F(z)$$

for any $y, z \in \mathbb{R}^p$ and $\lambda \in (0,1)$ (see, e.g., [3], [17], [19]).

Being near relatives to linear and affine mappings, convex functions and operators are known to possess many remarkable properties.

PROPOSITION 5.2. *The induced mapping $\Psi(x)$ defined by (5.3), of which we seek the zeros, is order convex with respect to the componentwise partial ordering "$\leq$" of $\mathbb{R}^{2n}$.*

*Proof.* We have for $\lambda \in (0,1)$ and any $\mathbf{u}, \mathbf{v} \in \mathbb{IR}^n$, in view of the subdistributivity (3.6),

$$\mathbf{C}\big(\lambda\mathbf{u} + (1-\lambda)\mathbf{v}\big) \subseteq \lambda\mathbf{C}\mathbf{u} + (1-\lambda)\mathbf{C}\mathbf{v}.$$

Therefore, in $\mathbb{R}^{2n}$,

$$\sigma\Big(\mathbf{C}\big(\lambda\mathbf{u} + (1-\lambda)\mathbf{v}\big)\Big) \leq \lambda\sigma(\mathbf{C}\mathbf{u}) + (1-\lambda)\sigma(\mathbf{C}\mathbf{v}). \tag{5.4}$$

One can conclude from the above that for $y, z \in \mathbb{R}^{2n}$ such that $y = \sigma(\mathbf{u})$, $z = \sigma(\mathbf{v})$ the following chain of inequalities is valid:

$$\begin{aligned}
&\Psi\big(\lambda y + (1-\lambda)z\big) \\
&\qquad = \sigma\Big(\mathbf{C}\sigma^{-1}\big(\lambda y + (1-\lambda)z\big)\Big) - \big(\lambda y + (1-\lambda)z\big) + \sigma(\mathbf{b}) \\
&\qquad = \sigma\Big(\mathbf{C}\big(\lambda\mathbf{u} + (1-\lambda)\mathbf{v}\big)\Big) - \big(\lambda y + (1-\lambda)z\big) + \sigma(\mathbf{b}) \\
&\qquad \leq \lambda\sigma(\mathbf{C}\mathbf{u}) + (1-\lambda)\sigma(\mathbf{C}\mathbf{v}) - \big(\lambda y + (1-\lambda)z\big) + \sigma(\mathbf{b}) \quad \text{by (5.4)} \\
&\qquad = \lambda\big(\sigma(\mathbf{C}\mathbf{u}) - y + \sigma(\mathbf{b})\big) + (1-\lambda)\big(\sigma(\mathbf{C}\mathbf{v}) - z + \sigma(\mathbf{b})\big) \\
&\qquad = \lambda\Psi(y) + (1-\lambda)\Psi(z). \qquad \square
\end{aligned}$$

DEFINITION 5.2. Let $\mathbb{R}^q$ be an ordered linear space with the partial order "$\preccurlyeq$". The *subdifferential* of the mapping $F : \mathbb{R}^p \to \mathbb{R}^q$ at the point $x \in \mathbb{R}^p$ is the set $\partial_{\preccurlyeq}F(x)$ of all linear operators $D : \mathbb{R}^p \to \mathbb{R}^q$ such that

$$D(y-x) \preccurlyeq F(y-x) \tag{5.5}$$

for any $y \in \mathbb{R}^p$. Members of the set $\partial_{\preccurlyeq}F(x)$—the linear operators satisfying (5.5)—are called *subgradients* of the mapping $F$ at the point $x$, while the mapping $F$ itself is said to be *subdifferentiable* at $x$ if its subdifferential is nonempty at that point (see [3], [19]).

In general, checking whether the subdifferential exists and, in case of existence, computing the subdifferential are not easy tasks. Fortunately, the situation is crucially simplified in the finite-dimensional space: it is known (see, e.g., [3], [19]) that convex continuous functions are always subdifferentiable in the interior of their domain.

The mapping under consideration $\Psi : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$, being order convex with respect to the common componentwise $\leq$-order on $\mathbb{R}^{2n}$, is equivalent to the functionals $\Psi_i : \mathbb{R}^{2n} \to \mathbb{R}$—coordinate components of $\Psi$—being convex for $i = 1, 2, \ldots, 2n$. In addition, the components $\Psi_i(x)$ are everywhere continuous. Therefore, none of the subdifferentials $\Psi_i(x)$ is empty at any point $x \in \mathbb{R}^{2n}$, so that for each $i = 1, 2, \ldots, n$ there exists a $2n$-vector $d_{(i)}$, such that

$$\Psi_i(x + v) - \Psi_i(x) \geq d_{(i)}^{\top} v \quad \text{for all} \quad v \in \mathbb{R}^{2n}.$$

Then, constructing a $(2n \times 2n)$-matrix $D = (d_{(1)}, d_{(2)}, \ldots, d_{(2n)})^{\top}$, we may conclude that at any $x \in \mathbb{R}^{2n}$ the set of matrices $D \in \mathbb{R}^{2n \times 2n}$ that satisfy the inequality

$$\Psi(x + v) - \Psi(x) \geq Dv$$

for any $v \in \mathbb{R}^{2n}$ is nonempty. That is, we have proved

PROPOSITION 5.3. *The $\leq$-order subdifferential $\partial_{\leq}\Psi(x)$ of the order convex map $\Psi$ is always nonempty, i.e., the induced mapping $\Psi$ defined by (5.3) is everywhere subdifferentiable.*

For brevity, we shall denote the subdifferential simply by $\partial\Psi(x)$, since no other order on $\mathbb{R}^{2n}$ is considered below.

To give explicit formulas for the subdifferential $\partial\Psi(x)$ and to derive estimates for it which will be used later, we need to know more about the mapping $\Psi$ under study.

PROPOSITION 5.4. *For the mapping $\Psi$ defined by (5.3), the following representation holds:*

$$\Psi(x) = \max_{C \in \mathbf{C}} C^{\sigma} x - x + \sigma(\mathbf{b}). \tag{5.6}$$

*Proof.* First, for any proper $(n \times n)$-matrix $\mathbf{C}$ and arbitrary interval $n$-vector $\mathbf{v}$, there holds

$$\mathbf{C} \cdot \mathbf{v} = \bigvee_{C \in \mathbf{C}} C \cdot \mathbf{v}.$$

Indeed, if $\mathbf{C} \cdot \mathbf{v} = \big((\mathbf{C} \cdot \mathbf{v})_1, (\mathbf{C} \cdot \mathbf{v})_2, \ldots, (\mathbf{C} \cdot \mathbf{v})_n\big)^{\top}$, then, using (3.5) and distributivity of the operation "$\vee$" with respect to addition (3.3), we get for any $i = 1, 2, \ldots, n$

$$(\mathbf{C}\cdot\mathbf{v})_i = \sum_{j=1}^{n} \mathbf{c}_{ij}\mathbf{v}_j = \sum_{j=1}^{n} \bigvee_{c_{ij}\in\mathbf{c}_{ij}} c_{ij}\mathbf{v}_j = \bigvee_{c_{i1}\in\mathbf{c}_{i1}} \bigvee_{c_{i2}\in\mathbf{c}_{i2}} \cdots \bigvee_{c_{in}\in\mathbf{c}_{in}} \sum_{j=1}^{n} c_{ij}\mathbf{v}_j$$

$$= \bigvee_{C\in\mathbf{C}} \sum_{j=1}^{n} c_{ij}\mathbf{v}_j = \bigvee_{C\in\mathbf{C}} (C\cdot\mathbf{v})_i.$$

Hence, in view of (4.6) and (4.9), we conclude

$$\sigma\big(\mathbf{C}\sigma^{-1}(x)\big) = \sigma\Big(\bigvee_{C\in\mathbf{C}} C\sigma^{-1}(x)\Big) = \max_{C\in\mathbf{C}} \le \sigma\big(C\sigma^{-1}(x)\big) = \max_{C\in\mathbf{C}} \le C^{\sigma}x,$$

so that overall the formula (5.6) follows. □

DEFINITION 5.3. For a function $f : \mathbb{R}^p \to \mathbb{R}$, the *epigraph* is the set

$$\operatorname{epi} f := \{(x,t) \mid x\in\mathbb{R}^p,\ t\in\mathbb{R},\ f(x)\le t\}$$

(see, e.g., [3], [19]).

A *polyhedral convex set* in $\mathbb{R}^p$ is a set that can be represented as intersection of a finite number of closed half-spaces of $\mathbb{R}^p$, i.e., as the set of solutions to a finite system of linear inequalities

$$h_{(i)}^{\top}x \le \xi_i, \quad i = 1,2,\ldots,m, \quad h_{(i)}^{\top}\in\mathbb{R}^p, \quad \xi_i\in\mathbb{R}.$$

A *polyhedral convex function* is a (convex) function of which the epigraph is a polyhedral set (see [19]).

We stress that polyhedral functions, in their turn, are the simplest amongst the convex functions: they are *locally affine* almost everywhere because their graphs are composed of pieces of *hyperplanes*.

PROPOSITION 5.5. *All the components $\Psi_i(x)$, $i = 1,2,\ldots,2n$, of the map $\Psi$ defined by (5.3) are convex polyhedral functions.*

*Proof.* Utilizing the result of Proposition 5.4, we arrive at

$$\Psi_i(x) = \Big(\max_{C\in\mathbf{C}} C^{\sigma}x\Big)_i - x_i + \big(\sigma(\mathbf{b})\big)_i \tag{5.7}$$

for any $i = 1,2,\ldots,2n$. In (5.7), due to the special form (4.7) of the matrix $C^{\sigma}$, the maximum may be attained only at *endpoints* of the interval entries $\mathbf{c}_{ij}$, $j = 1,2,\ldots,n$, or at some *zeros* among them if the corresponding $\mathbf{c}_{ij} \ni 0$. In either case, one can equivalently replace the interval $\mathbf{c}_{ij}$ by a finite number of points (two or three) to obtain the maximum in (5.7). With these considerations, we state that

$$\Psi_i(x) = \Big(\max_{C\in\operatorname{Vert}\mathbf{C}} C^{\sigma}x\Big)_i - x_i + \big(\sigma(\mathbf{b})\big)_i, \tag{5.8}$$

where $\operatorname{Vert}\mathbf{C}$ is the extended set of vertex matrices of $\mathbf{C}$ defined as

$$(\operatorname{Vert}\mathbf{C})_{ij} := \begin{cases} \{\underline{\mathbf{c}}_{ij}, \overline{\mathbf{c}}_{ij}\}, & \text{if } 0\notin\mathbf{c}_{ij}, \\ \{\underline{\mathbf{c}}_{ij}, 0, \overline{\mathbf{c}}_{ij}\}, & \text{otherwise.} \end{cases}$$

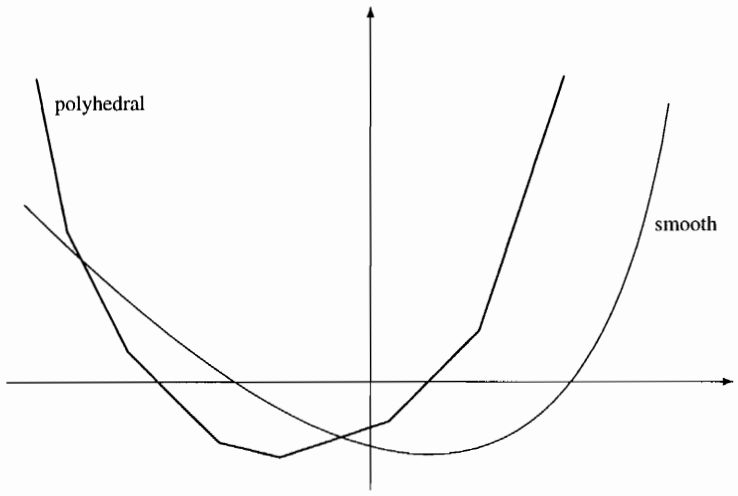



*Figure 3.* A graph of a smooth convex function as compared with a graph of a polyhedral convex function: the latter is made up of pieces of hyperplanes.

Since Vert **C** is finite, the proposition follows. □

DEFINITION 5.4. For a function $f : \mathbb{R}^p \to \mathbb{R}$, the *one-sided derivative at a point x with respect to a direction* $y \in \mathbb{R}^p$ (or, briefly, *directional derivative*) is defined as the limit

$$\frac{\partial f(x)}{\partial y} = \lim_{\alpha \searrow 0} \frac{f(x + \alpha y) - f(x)}{\alpha},$$

provided it exists (see, e.g., [3], [19]).

We will especially need the following particular cases of the above definition

$$\frac{\partial \Psi_i(x)}{\partial x_j^-} := \lim_{\alpha \searrow 0} \frac{\Psi_i(x_1, \ldots, x_{j-1}, x_j - \alpha, x_{j+1}, \ldots, x_{2n}) - \Psi_i(x_1, \ldots, x_{2n})}{\alpha}$$

and

$$\frac{\partial \Psi_i(x)}{\partial x_j^+} := \lim_{\alpha \searrow 0} \frac{\Psi_i(x_1, \ldots, x_{j-1}, x_j + \alpha, x_{j+1}, \ldots, x_{2n}) - \Psi_i(x_1, \ldots, x_{2n})}{\alpha},$$

one-sided partial derivatives of the component $\Psi_i$ at the point $x$, from the left and from the right, with respect to the $j$-th coordinate direction.

DEFINITION 5.5. The *support function* of a convex set $W \subseteq \mathbb{R}^p$ is the function

$$\delta_W(x) := \sup \{x^\top w \mid w \in W\}.$$

(see, e.g., [3], [19]).

THEOREM 5.1 ([19], Section 23). *Let $f : \mathbb{R}^p \to \mathbb{R}$ be a polyhedral convex function which is finite at a point $x$. Then $f$ is subdifferentiable at the point $x$ and $\partial f(x)$ is a polyhedral convex set. In addition, the directional derivative of $f$ at $x$, as a function of the direction, is a support function of the subdifferential $\partial f(x)$.*

Finally, we are able to write out the explicit form of the subdifferential $\partial\Psi(x)$ under investigation.

PROPOSITION 5.6. *The subdifferential $\partial\Psi(x)$ of the mapping $\Psi$ defined by (5.3) is the following (proper) interval $(2n \times 2n)$-matrix:*

$$\begin{pmatrix} \left[\dfrac{\partial\Psi_1(x)}{\partial x_1^-}, \dfrac{\partial\Psi_1(x)}{\partial x_1^+}\right] & \cdots & \left[\dfrac{\partial\Psi_1(x)}{\partial x_{2n}^-}, \dfrac{\partial\Psi_1(x)}{\partial x_{2n}^+}\right] \\ \vdots & \ddots & \vdots \\ \left[\dfrac{\partial\Psi_{2n}(x)}{\partial x_1^-}, \dfrac{\partial\Psi_{2n}(x)}{\partial x_1^+}\right] & \cdots & \left[\dfrac{\partial\Psi_{2n}(x)}{\partial x_{2n}^-}, \dfrac{\partial\Psi_{2n}(x)}{\partial x_{2n}^+}\right] \end{pmatrix}. \tag{5.9}$$

*Proof.* The natural componentwise partial order on $\mathbb{R}^{2n}$ is the direct product of the orders "$\leq$" on $\mathbb{R}$. Therefore, the order subdifferential is the direct product of the common subdifferentials of the separate components $\Psi_i : \mathbb{R}^{2n} \to \mathbb{R}$, and for each one of them we can avail ourselves of Theorem 5.1, as follows.

For the functions $\Psi_i(x)$, the existence of the one-sided derivative with respect to any coordinate direction in no way affects the existence of the derivative with respect to the other variables. Besides, the matrices of the form (5.9) composed of the one-sided derivatives obviously belong to the subdifferential $\partial\Psi(x)$. Hence, $\partial\Psi(x)$ is a direct product of the "partial" subdifferentials, that is, an interval matrix whose elements are $[\partial\Psi_i(x) / \partial x_j^-,\ \partial\Psi_i(x) / \partial x_j^+]$. □

COROLLARY 5.1. *Assuming differentiability of $\Psi$ at $x$,*

$$\frac{\partial\Psi_i(x)}{\partial x_j^-} = \frac{\partial\Psi_i(x)}{\partial x_j^+} = \textit{the common partial derivative } \frac{\partial\Psi_i(x)}{\partial x_j},$$

*so that the structure of the matrix (5.9) is simplified. Then, the subdifferential $\partial\Psi(x)$ consists of the only element, namely, of the Jacobi matrix*

$$\begin{pmatrix} \dfrac{\partial\Psi_1(x)}{\partial x_1} & \cdots & \dfrac{\partial\Psi_1(x)}{\partial x_{2n}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial\Psi_{2n}(x)}{\partial x_1} & \cdots & \dfrac{\partial\Psi_{2n}(x)}{\partial x_{2n}} \end{pmatrix}.$$

DEFINITION 5.6. The *positive part* $\mathbf{x}^+$ and the *negative part* $\mathbf{x}^-$ of a proper interval $\mathbf{x}$ are the following intervals

$$\mathbf{x}^+ := \{x^+ \mid x \in \mathbf{x}\} = \{\max(x, 0) \mid x \in \mathbf{x}\},$$

$$\mathbf{x}^- := \{x^- \mid x \in \mathbf{x}\} = \{\max(-x, 0) \mid x \in \mathbf{x}\}.$$

For example,

$$[-1,2]^+ = [0,2], \quad [-1,2]^- = [0,1],$$
$$[1,2]^+ = [1,2], \quad [1,2]^- = [0,0].$$

One can consider taking positive and negative parts of an interval as interval extensions of the functions $(\cdot)^+$ and $(\cdot)^-$ introduced in Definition 4.3. As usual, the above operations will be applied to interval vectors and matrices in the componentwise manner.

PROPOSITION 5.7. *For the subdifferential $\partial\Psi(x)$ of the mapping $\Psi$ defined by (5.3), the following estimate is valid:*

$$\partial\Psi(x) \subseteq \begin{pmatrix} \mathbf{C}^+ & \mathbf{C}^- \\ \mathbf{C}^- & \mathbf{C}^+ \end{pmatrix} - I. \tag{5.10}$$

*Proof.* Below, it will be convenient to write $\pm$ instead of either of the signs $+$ and $-$. By virtue of the representation (5.8),

$$\begin{aligned} \frac{\partial\Psi_i(x)}{\partial x_j^\pm} &= \frac{\partial}{\partial x_j^\pm}\left(\left(\max_{C \in \operatorname{Vert}\mathbf{C}} C^\sigma x\right)_i - x_i + (\sigma(\mathbf{b}))_i\right) \\ &= \frac{\partial}{\partial x_j^\pm}\left(\max_{C \in \operatorname{Vert}\mathbf{C}} C^\sigma x\right)_i - \delta_{ij}, \end{aligned} \tag{5.11}$$

where $\delta_{ij}$ is Kronecker symbol:

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

We can avail ourselves of the differentiation rule for a maximum-type function (see, e.g., [6], Section III.2):

$$\frac{\partial}{\partial x_j^\pm}\left(\max_{C \in \operatorname{Vert}\mathbf{C}} \leq C^\sigma x\right)_i = \begin{array}{l} ij\text{-th element of that matrix } C^\sigma \text{ at which} \\ \text{the considered } \max\limits_{C \in \operatorname{Vert}\mathbf{C}} \leq C^\sigma x \text{ is attained.} \end{array} \tag{5.12}$$

Overall, combining (4.7), (5.11) and (5.12), we get the general form of the directional derivative matrix:

$$\left(\frac{\partial\Psi(x)}{\partial x_{ij}^\pm}\right) = \begin{pmatrix} (C')^+ & (C')^- \\ (C'')^- & (C'')^+ \end{pmatrix} - I,$$

where $C', C'' \in \mathbb{R}^{n\times n}$, $C', C'' \in \operatorname{Vert}\mathbf{C}$. So, in view of Proposition 5.6, we arrive at the required inclusion (5.10). □

## 6. Algorithm

To solve the equations (5.2)–(5.3) in the enveloping space $\mathbb{R}^{2n}$, we propose the following iterative algorithm.

ALGORITHM I. Subdifferential Newton method with a special starting approximation

As the starting vector $x^{(0)}$ take the solution of the "midpoint" system

$$\left(I - (\operatorname{mid} \mathbf{C})^{\sigma}\right)x = \sigma(\mathbf{b}).$$

If the $k$-th approximation $x^{(k)}$, $k = 0, 1, \ldots$, has already been found, then compute any $D^{(k)} \in \partial\Psi(x^{(k)})$ and put

$$x^{(k+1)} := x^{(k)} - \tau(D^{(k)})^{-1}\left(\Psi(x^{(k)})\right)$$

for some constant $\tau \in (0, 1]$.

Here, $\tau$ is a damping factor, which we incline to recommend to be equal to or close to 1. Our computational experience shows that then, as a rule, Algorithm I gives an *exact* solution to the problem in a small finite number of iterations, which usually does not exceed the dimension of the system. In this kind of method, the damping factor was originally introduced to prevent divergence. However, whether taking $\tau < 1$ really improves convergence of the subdifferential Newton method is not quite clear to me yet. Anyway, in practice it appears that the smaller $\tau$, the slower Algorithm I works.

It is worth noting that the above Algorithm I is exactly the same we used in [23] to find *inner* interval estimates of the solution sets to interval linear systems: we applied Algorithm I to the computation of algebraic solutions of the equation (dual $\mathbf{A}$)$x = \mathbf{b}$ there.

Complete investigation of the subdifferential Newton method is beyond the scope of the present work. Below, we shall prove, based on the standard technique (see, e.g., [17]), a local convergence theorem, which is identical to the similar result from [23]. It amounts to the following:

THEOREM 6.1. *If the proper interval matrix* $\mathbf{C}$ *is sufficiently narrow and all point* $(2n \times 2n)$*-matrices* $S$ *that satisfy*

$$S \in \begin{pmatrix} \mathbf{C}^{+} & \mathbf{C}^{-} \\ \mathbf{C}^{-} & \mathbf{C}^{+} \end{pmatrix} - I$$

*are nonsingular, then Algorithm I converges to* $\sigma(\mathbf{x}^{*})$*, where* $\mathbf{x}^{*}$ *is an algebraic solution of the main system (2.2).*

*Proof.* Let us specify what is meant by a "sufficiently narrow" interval matrix $\mathbf{A}$. We shall require that

the convex hull of the set

$$\bigcup \left\{ S^{-1}K_{\leq} \;\middle|\; S \in \mathbb{R}^{2n\times 2n},\ S \in \begin{pmatrix} \mathbf{C}^{+} & \mathbf{C}^{-} \\ \mathbf{C}^{-} & \mathbf{C}^{+} \end{pmatrix} - I \right\}, \tag{6.1}$$

where $S^{-1}K_{\leq}$ denotes the preimage of the positive cone (4.5) under the linear transformation $S$, itself is a cone $K_{\trianglelefteq}$ in $\mathbb{R}^{2n}$.

This is not an arbitrary condition. If the matrix $\mathbf{C}$ is thin, that is, $\mathbf{C} = C$, then $(C^{\sigma})^{-1}K_{\leq}$ is actually a cone, being the image of the cone in a linear transformation. If the matrices $S', S'' \in \mathbb{R}^{2n\times 2n}$ are "sufficiently close," then the cones $(S')^{-1}K_{\leq}$ and $(S'')^{-1}K_{\leq}$ are close too, and their convex hull is still a cone. The condition (6.1) thus indeed reflects "narrowness" of the interval matrix $\mathbf{C}$ in some sense.

When the set (6.1) is a cone, it defines a partial ordering "$\trianglelefteq$" of $\mathbb{R}^{2n}$:

$$x \trianglelefteq y \quad \Longleftrightarrow \quad x - y \in K_{\trianglelefteq}.$$

The main idea of our proof is to demonstrate that the sequence of approximations generated by Algorithm I is *decreasing* as well as *bounded* from below with respect to this specially constructed order "$\trianglelefteq$".

First, the special choice of the starting vector $x^{(0)}$ implies

$$\begin{aligned} \Psi(x^{(0)}) &= \sigma\big(\mathbf{C}\sigma^{-1}(x^{(0)})\big) - x^{(0)} + \sigma(\mathbf{b}) \\ &\geq \sigma\big((\operatorname{mid}\mathbf{C})\sigma^{-1}(x^{(0)})\big) - x^{(0)} + \sigma(\mathbf{b}) \quad \text{by inclusion monotonicity} \\ &= (\operatorname{mid}\mathbf{C})^{\sigma}x^{(0)} - x^{(0)} + \sigma(\mathbf{b}) \quad \text{by property (4.9)} \\ &= \big((\operatorname{mid}\mathbf{C})^{\sigma} - I\big)x^{(0)} + \sigma(\mathbf{b}) = 0. \end{aligned}$$

So

$$\Psi(x^{(0)}) \geq 0.$$

Next, directly from the definition of subdifferential,

$$\Psi(x^{(k+1)}) \geq \Psi(x^{(k)}) + D^{(k)}(x^{(k+1)} - x^{(k)})$$

for $D^{(k)} \in \partial\Psi(x^{(k)})$ and any $k = 0, 1, 2, \ldots$, while by virtue of Algorithm I

$$D^{(k)}(x^{(k+1)} - x^{(k)}) = -\tau\Psi(x^{(k)}). \tag{6.2}$$

Hence, for $0 < \tau \leq 1$ we get by induction

$$\Psi(x^{(k)}) \geq 0, \quad k = 1, 2, \ldots \tag{6.3}$$

What is behind the inequality (6.3)? Recall the representation (5.6): for each $k$, we have

$$\Psi(x^{(k)}) = \max_{C\in\mathbf{C}} C^{\sigma}x^{(k)} - x^{(k)} + \sigma(\mathbf{b}) \geq 0,$$

which implies

$$S^{(k)}x^{(k)} + \sigma(\mathbf{b}) \in K_{\leq}$$

for some

$$S^{(k)} \in \begin{pmatrix} \mathbf{C}^+ & \mathbf{C}^- \\ \mathbf{C}^- & \mathbf{C}^+ \end{pmatrix} - I$$

at which $(\max C^{\sigma}x^{(k)} - x^{(k)})$ is attained.

The latter is equivalent, inasmuch as $S^{(k)}$ must be nonsingular, to

$$x^{(k)} + (S^{(k)})^{-1}\sigma(\mathbf{b}) \in (S^{(k)})^{-1}K_{\leq} \subseteq K_{\trianglelefteq},$$

while this inclusion simply means

$$x^{(k)} \trianglerighteq -(S^{(k)})^{-1}\sigma(\mathbf{b}).$$

If we set

$$\beta = \min_{\trianglelefteq}\left\{-S^{-1}\sigma(\mathbf{b}) \;\middle|\; S \in \mathbb{R}^{2n\times 2n},\ S \in \begin{pmatrix} \mathbf{C}^+ & \mathbf{C}^- \\ \mathbf{C}^- & \mathbf{C}^+ \end{pmatrix} - I\right\},$$

then each $x^{(k)} \trianglerighteq \beta$, i.e., the sequence $\{x^{(k)}\}$ turns out to be $\trianglelefteq$-bounded from below.

The other important point: the sequence $\{x^{(k)}\}$ produced by our algorithm is monotonically decreasing with respect to the order "$\trianglelefteq$", that is

$$x^{(k)} \trianglerighteq x^{(k+1)} \tag{6.4}$$

for all $k = 0, 1, 2, \ldots$ Indeed, combining (6.2) and (6.3) one obtains

$$D^{(k)}(x^{(k+1)} - x^{(k)}) \leq 0.$$

As we could see, the inclusion (5.10)

$$D^{(k)} \in \begin{pmatrix} \mathbf{C}^+ & \mathbf{C}^- \\ \mathbf{C}^- & \mathbf{C}^+ \end{pmatrix} - I$$

holds for each $D^{(k)}$. Thereby, in view of (6.1), (6.4) follows.

We have thus proved

$$x^{(k)} \trianglerighteq x^{(k+1)} \trianglerighteq \beta. \tag{6.5}$$

In general, the interrelation between order and topology in a partially ordered linear topological space may be quite complicated, but things luckily get better in the finite-dimensional situation: a sequence that is monotonic and bounded with respect to a partial order (conforming to the linear structure) always has a

topological limit as well. So, we can conclude from (6.5) that there exists a limit $x^*$ of the sequence $\{x^{(k)}\}$ produced by Algorithm I. We can find it by solving the fixed-point equation

$$x^* = x^* - \tau(D^*)^{-1}\big(\Psi(x^*)\big)$$

with some $D^* \in \partial\Psi(x^*)$ which must be nonsingular. Therefore, $\Psi(x^*) = 0$. □

## 7. Modifications and Further Development

What is the position of our algebraic approach amongst other approaches to the solution of the "outer problem" (1.3)? The algebraic approach inherits good enclosure properties from the interval iterative schemes, while its low computational complexity is nearly equal to that of the finite algorithms. Overall, the algebraic approach, appropriately modified and updated, may serve as a tool for quick computation of reasonably good enclosures of the solution set to quite general interval linear systems. Our immediate purpose is to briefly outline some possible ways of doing this.

PROPOSITION 7.1. *The solution set of the interval system*

$$\mathbf{A}x = \mathbf{b}$$

*coincides with the solution set of the interval system*

$$\mathbf{G}x = (\text{dual } \mathbf{G} - \mathbf{A})x + \mathbf{b} \tag{7.1}$$

*for any proper interval matrix* $\mathbf{G}$ *of the same size as* $\mathbf{A}$ *and such that* $\operatorname{rad}\mathbf{G} \leq \operatorname{rad}\mathbf{A}$.

*Proof.* Notice, as a remark to the formulation of Proposition 7.1, that the interval matrix (dual $\mathbf{G} - \mathbf{A}$) is proper if $\operatorname{rad}\mathbf{G} \leq \operatorname{rad}\mathbf{A}$.

We have

$$\begin{aligned}
&\text{the solution set of (7.1)}\\
&= \{x \in \mathbb{R}^n \mid (\exists G \in \mathbf{G})\big(\exists H \in (\text{dual } \mathbf{G} - \mathbf{A})\big)(\exists b \in \mathbf{b})(Gx = Hx + b)\}\\
&= \{x \in \mathbb{R}^n \mid (\exists G \in \mathbf{G})\big(\exists H \in (\text{dual } \mathbf{G} - \mathbf{A})\big)(\exists b \in \mathbf{b})\big((G - H)x = b\big)\}\\
&= \{x \in \mathbb{R}^n \mid \big(\exists A \in \mathbf{G} - (\text{dual } \mathbf{G} - \mathbf{A})\big)(\exists b \in \mathbf{b})(Ax = b)\}\\
&= \{x \in \mathbb{R}^n \mid (\exists A \in \mathbf{A})(\exists b \in \mathbf{b})(Ax = b)\}\\
&= \text{the solution set of } \mathbf{A}x = \mathbf{b},
\end{aligned}$$

since if $G \in \mathbf{G}$ and $H \in (\text{dual } \mathbf{G} - \mathbf{A})$, then $(G - H) \in \mathbf{G} - (\text{dual } \mathbf{G} - \mathbf{A}) = \mathbf{A}$. □

Now, the main idea of our algebraic approach can be applied to the equation (7.1). Let us designate by $\Gamma : \mathbb{IR}^n \to \mathbb{IR}^n$ the map with the action

$$\Gamma(\mathbf{x}) = \mathbf{G}\mathbf{x}, \tag{7.2}$$

that is, multiplication by the matrix $\mathbf{G}$. Then, if the inverse $\Gamma^{-1} : \mathbb{IR}^n \to \mathbb{IR}^n$ exists and the mapping

$$\mathbf{x} \mapsto \Gamma^{-1}\big((\mathrm{dual}\,\mathbf{G} - \mathbf{A})\mathbf{x} + \mathbf{b}\big)$$

is Lipschitz continuous, where the Lipschitz operator is a contraction of $\mathbb{IR}^n$, then the algebraic solution of the equation

$$\mathbf{G}\mathbf{x} = (\mathrm{dual}\,\mathbf{G} - \mathbf{A})\mathbf{x} + \mathbf{b} \tag{7.3}$$

is a solution to the "outer problem" (1.3). The corresponding general result has been proved in somewhat other terms as early as by Alefeld and Herzberger [1] (Chapter 11, Corollary 6). Next, to compute the sought-for algebraic solution, we can make use of equivalent transformation (in Kaucher arithmetic), embedding into $\mathbb{IR}^{2n}$, as well as various numerics, other than traditional stationary contractions.

Before proceeding further, we emphasize that it makes good sense to restrict ourselves to splitting only *thin* (point) matrices $\mathbf{G}$ from $\mathbf{A}$ in (7.3), that is, when $\mathbf{G} = G$. The reason behind that is twofold and quite important:

- To solve the main equation (7.3), we intend to avail ourselves of the subdifferential Newton method, which requires the induced map $\Psi = \sigma \circ \psi \circ \sigma^{-1}$ matching

  $$\psi(\mathbf{x}) = (\mathrm{dual}\,\mathbf{G} - \mathbf{A})\mathbf{x} \ominus \mathbf{G}\mathbf{x} + \mathbf{b}$$

  to be order convex. Tracing the chain of transformations from the proof of Proposition 5.2, one can easily check that the desired convexity holds only if the matrix $\mathbf{G}$ is thin.
- For thin $\mathbf{G}$, invertibility of the map (7.2), which is decisive for the solvability of the implicit relationship (7.3), can be investigated comprehensively and easy implemented numerically. As was revealed in Section 4, the mapping (7.2) with $\mathbf{G} = G$ is invertible if and only if the matrix $G$ is $\imath$-nonsingular, while the inverse mapping $\Gamma^{-1}$ is determined by the formula (4.10).

Thus, let $\mathbf{G} = G$ be a $\imath$-nonsingular point matrix to make sure that the inverse function to (7.2) actually exists. Now then, the appropriate question is: what are advantages of the modified form (7.3) as compared with the original basic equation (2.2)?

Under the circumstances, the operator $\Gamma^{-1}$, being an inverse to a linear map, must be a linear map too, so that the modified form (7.3) is equivalent to

$$\mathbf{x} = \Gamma^{-1}\big((G - \mathbf{A})\mathbf{x}\big) + \Gamma^{-1}(\mathbf{b}).$$

Both $\Gamma$ and $\Gamma^{-1}$ were identity operators in the initially treated case (2.2). But now, through appropriate choice of $G$, we can try to make the composition of $\Gamma^{-1}$ and multiplication by $(G - \mathbf{A})$ a "greater contraction" than the multiplication by

$(I - \mathbf{A})$ alone. For one thing, if the algebraic approach is not applicable to the simplest equation

$$\mathbf{x} = \mathbf{C}\mathbf{x} + \mathbf{b}, \tag{2.2}$$

then we can hope that, after a suitable transformation, the resulting equation (7.3) will have a contracting Lipschitz operator. For another, even if $\rho(|\mathbf{C}|) < 1$ and the algebraic approach is applicable to the equation (2.2), then the width of its fixed point is shown (e.g., in [2]) to crucially depend on $\|\mathbf{C}\|$. Smaller $\|\mathbf{C}\|$ shall produce better enclosures for the solution set, all other factors being equal.

How can we optimize the choice of the matrix $G$? This is the subject of our future work.

## 8. Computational Tests

In this section, we summarize numerical experiments carried out with an algorithm (subdifferential Newton method) that realizes the algebraic approach to the outer problem for interval linear equations on a PC/AT 486. We take the term "algebraic approach" to mean the simplest version of the modification described in Section 7, with the splitted matrix $G$ defined as follows.

DEFINITION 8.1. By the *deviation* of a proper interval $\mathbf{x}$, we mean the quantity

$$\operatorname{dev} \mathbf{x} := \begin{cases} \underline{\mathbf{x}}, & \text{if } |\underline{\mathbf{x}}| \geq |\overline{\mathbf{x}}|, \\ \overline{\mathbf{x}}, & \text{otherwise,} \end{cases}$$

that is, the most distant (from 0) point of the interval $\mathbf{x}$.

We put

$$G := \operatorname{dev} \operatorname{diag} \mathbf{A} = \begin{pmatrix} \operatorname{dev} \mathbf{a}_{11} & & 0 \\ & \ddots & \\ 0 & & \operatorname{dev} \mathbf{a}_{nn} \end{pmatrix},$$

that is, a diagonal matrix made up of the deviations of the diagonal elements of $\mathbf{A}$. Then the map which is inverse to the multiplication by the matrix $G$ is simply multiplication by the matrix $G^{-1}$, and we have to solve

$$\mathbf{x} = G^{-1}(G - \mathbf{A})\mathbf{x} + G^{-1}\mathbf{b}$$

instead of (7.3). We experimentally exhibit that even such simplest choice of $G$ leads to sufficiently good results. In addition, both the applicability domain of the approach extends and the quality of interval enclosure improves in comparison to the nonmodified form (2.2).

In all the following examples we set the damping factor $\tau = 1$ to show the computational efficacy of the algebraic approach. Checking out whether $\rho(|G^{-1}(G -$

$\mathbf{A})|) < 1$, which is so essential for feasibility of the algebraic approach, was changed to checking the stronger condition $\| |G^{-1}(G - \mathbf{A})| \| < 1$ by virtue of the well-known inequality between spectral radius and matrix norms. The algorithm was implemented using Turbo C in standard double precision floating point arithmetic. The answers below are presented to three digits.

The natural rivals of the algebraic approach are the famous interval Gauss algorithm (see, e.g., [1], [14]) and recent Hansen-Rohn's procedure [8], [20]. Each of the above methods has its own applicability scope, and for each one there exists a special class(es) of interval linear equations on which the method gives the best enclosures, with the least overestimation. To make a definite conclusion on comparative qualities of a method, much routine work is to be done. A few examples are hardly convincing that one method is actually better than the others. Still, we have tried to consider, for completeness, various test examples within the present paper, including such as the following artificial

EXAMPLE 8.0 from [8].

$$\begin{pmatrix} [0.7,1.3] & [-0.3,0.3] & [-0.3,0.3] \\ [-0.3,0.3] & [0.7,1.3] & [-0.3,0.3] \\ [-0.3,0.3] & [-0.3,0.3] & [0.7,1.3] \end{pmatrix} x = \begin{pmatrix} [-14,-7] \\ [9,12] \\ [-3,3] \end{pmatrix}.$$

Here, the interval Gaussian elimination gives the interval

$$\begin{pmatrix} [-101,71] \\ [-62.25,99] \\ [-90,90] \end{pmatrix},$$

Hansen's approach gives

$$\begin{pmatrix} [-101,17] \\ [-15,99] \\ [-90,90] \end{pmatrix},$$

while the algebraic approach converges after two iterations to

$$\begin{pmatrix} [-101,71] \\ [-69,99] \\ [-90,90] \end{pmatrix}.$$

No wonder the Hansen's approach turns out to be best: the midpoint matrix of the system is the unit matrix.

EXAMPLE 8.1. Let us turn to a more realistic interval linear system from [16]

$$\begin{pmatrix} [15,17] & [-3,3.01] & [-3,3.01] & [-3,3.01] \\ [-3,3.01] & [15,17] & [-3,2.99] & [-3,2.99] \\ [-3,2.99] & [-3,2.99] & [15,17] & [-3,3.01] \\ [-3,3.01] & [-3,3.01] & [-3,2.99] & [15,17] \end{pmatrix} x = \begin{pmatrix} [-6,-2] \\ [4,5] \\ [-2,4] \\ [8,10] \end{pmatrix}.$$

By interval Gauss algorithm, we obtain the box

$$\begin{pmatrix} [-1.03, 0.495] \\ [-0.347, 0.974] \\ [-0.770, 0.917] \\ [0.150, 1.25] \end{pmatrix},$$

and, using Hansen's technique of [8], we obtain the box

$$\begin{pmatrix} [-1.03, 0.363] \\ [-0.223, 0.975] \\ [-0.752, 0.919] \\ [0.149, 1.25] \end{pmatrix}.$$

By our algebraic approach we get

$$\begin{pmatrix} [-1.03, 0.495] \\ [-0.372, 0.974] \\ [-0.785, 0.917] \\ [-0.05, 1.25] \end{pmatrix},$$

which is not so bad: again, the midpoint matrix of the system is nearly diagonal. That favors good quality of the results obtained by Hansen-Rohn's procedure.

EXAMPLE 8.2, an example of Hansen (see, e.g., [8]).

$$\begin{pmatrix} [2, 3] & [0, 1] \\ [1, 2] & [2, 3] \end{pmatrix} x = \begin{pmatrix} [0, 120] \\ [60, 240] \end{pmatrix}.$$

As in the previous case, the algebraic approach with the diagonal deviation splitting converges to the exact solution

$$\begin{pmatrix} [-120, 90] \\ [-60, 240] \end{pmatrix}$$

in 2 iterations. The same answer can be obtained by the interval Gauss algorithm.

The interval obtained is the optimal interval enclosure of the solution set, while applying Hansen's approach [8] we get the bounds on the solution set

$$\begin{pmatrix} \left[-120, \dfrac{1845}{11}\right] \\ \left[-60, \dfrac{2940}{11}\right] \end{pmatrix},$$

which is considerably worse (due to a crude preconditioning that is intrinsic to Hansen-Rohn's procedure).

The next interesting range of Examples 8.3–8.7 with a fixed interval matrix is due to Ning and Kearfott [16] who advocate and further develop the Hansen-Rohn approach.

EXAMPLE 8.3 [16]. Let the interval equation system $\mathbf{A}x = \mathbf{b}$ be given with

$$\mathbf{A} = \begin{pmatrix} [3.7, 4.3] & [-1.5, -0.5] & [0, 0] \\ [-1.5, -0.5] & [3.7, 4.3] & [-1.5, -0.5] \\ [0, 0] & [-1.5, -0.5] & [3.7, 4.3] \end{pmatrix} \tag{8.1}$$

and $\mathbf{b} = ([-14, 14], [-9, 9], [-3, 3])^\top$.

Then, using interval Gaussian elimination, we obtain the box

$$\begin{pmatrix} [-6.38, 6.38] \\ [-6.40, 6.40] \\ [-3.40, 3.40] \end{pmatrix}, \tag{8.2}$$

while using Hansen's technique [8] or Rohn's reformulation of Hansen's technique [20] or Ning-Kearfott's modification [16] gives the same result. The box (8.2) proves to be the exact hull of the solution set of the system under consideration, and the algebraic approach gives this result too, in only 1 iteration.

EXAMPLE 8.4 [16]. Let the interval equation system $\mathbf{A}x = \mathbf{b}$ be given, with $\mathbf{A}$ as in Example 8.3 and $\mathbf{b} = ([-14, 0], [-9, 0], [-3, 0])^\top$.

Using interval Gaussian elimination, we obtain the box

$$\begin{pmatrix} [-6.38, 0] \\ [-6.40, 0] \\ [-3.40, 0] \end{pmatrix}. \tag{8.3}$$

Since the matrix (8.1) is an interval $M$-matrix and the right-hand side components have one sign, this box is the smallest one containing the solution set. Using Hansen's technique of [8] or Rohn's reformulation of [20], we obtain the wider box

$$\begin{pmatrix} [-6.38, 1.12] \\ [-6.40, 1.54] \\ [-3.40, 1.40] \end{pmatrix},$$

and using Ning-Kearfott's modification, we obtain an even wider box

$$\begin{pmatrix} [-6.38, 1.67] \\ [-6.40, 2.77] \\ [-3.40, 2.40] \end{pmatrix}.$$

The algebraic approach gives, to within roundoff errors, the exact hull of the solution set (8.3) after 1 (one) iteration.

EXAMPLE 8.5 [16]. Let the interval equation system $\mathbf{A}x = \mathbf{b}$ be given, with $\mathbf{A}$ as in Example 8.3 and $\mathbf{b} = ([0, 14], [0, 9], [0, 3])^\top$.

Using interval Gauss algorithm, we obtain the box

$$\begin{pmatrix} [0, 6.38] \\ [0, 6.40] \\ [0, 3.40] \end{pmatrix}, \tag{8.4}$$

which is the hull of the solution set to the system considered. Using Hansen's technique of [8], we obtain the wider box

$$\begin{pmatrix} [-1.12, 6.38] \\ [-1.54, 6.40] \\ [-1.40, 3.40] \end{pmatrix},$$

and Ning-Kearfott's modification gives an even worse result

$$\begin{pmatrix} [-1.67, 6.38] \\ [-2.77, 6.40] \\ [-2.40, 3.40] \end{pmatrix}.$$

Again, the box obtained by the algebraic approach after 1 iteration is the optimal box (8.4).

EXAMPLE 8.6 [16]. Let the interval equation system $\mathbf{A}x = \mathbf{b}$ be given, with $\mathbf{A}$ as in Example 8.3 and $\mathbf{b} = ([2, 14], [-9, -3], [-3, 1])^{\top}$.

Using the interval Gauss algorithm, we obtain the box

$$\begin{pmatrix} [-1.09, 4.29] \\ [-4.02, 1.24] \\ [-2.44, 0.773] \end{pmatrix}.$$

Using Hansen's technique of [8], we obtain the wider box

$$\begin{pmatrix} [-0.995, 5.01] \\ [-4.64, 1.52] \\ [-2.69, 1.38] \end{pmatrix},$$

while the smallest interval enclosure of the solution set is

$$\begin{pmatrix} [-0.995, 4.29] \\ [-3.79, 1.24] \\ [-2.35, 0.773] \end{pmatrix}.$$

It is a particular pleasure that, like in the preceding cases, the algebraic approach gives this interval answer after only 1 (one) iteration.

EXAMPLE 8.7 [16]. Let the interval equation system $\mathbf{A}x = \mathbf{b}$ be given, with $\mathbf{A}$ as in Example 8.3 and $\mathbf{b} = ([2, 14], [3, 9], [-3, 1])^{\top}$.

Using the interval Gauss algorithm, we obtain the box

$$\begin{pmatrix} [0.517, 6.25] \\ [0.450, 6.07] \\ [-0.881, 2.73] \end{pmatrix}.$$

Using Hansen's technique of [8], we obtain the wider box

$$\begin{pmatrix} [-0.206, 6.25] \\ [-0.386, 6.07] \\ [-2.01, 2.73] \end{pmatrix},$$

while the smallest interval enclosure of the solution set is

$$\begin{pmatrix} [0.523, 6.25] \\ [0.499, 6.07] \\ [-0.743, 2.73] \end{pmatrix}.$$

And again, the algebraic approach demonstrates its superiority: it gives the above best interval answer after only 2 iterations.

To sum up, for interval linear systems with "sufficiently arbitrary" matrices that are not close to the midpoint diagonal, the algebraic approach seems to work better than Hansen-Rohn-Ning-Kearfott's procedure. Arnold Neumaier observed that, with our choice of a diagonal splitted matrix $\mathbf{G}$ in the fixed-point form (7.1), the algebraic approach is easily shown to work precisely for $H$-matrices [14] and will give good results for sufficiently diagonally dominant matrices $\mathbf{A}$ as in the above examples. What will happen if we do not confine ourselves to diagonal splittings in the reduction to (7.1) is an interesting open question.

## 9. Advertisement

Public domain software that implements the algebraic approach to the outer problem for interval linear systems (subdifferential Newton method) as well as its text in C is available. They can be downloaded from ftp://www-sbras.ict.nsc.ru in the file pub/interval/shary.zip

## References

1. Alefeld, G. and Herzberger, J.: *Introduction to Interval Computations*, Academic Press, New York, 1983.
2. Apostolatos, N. and Kulisch, U.: Grundzüge einer Intervallrechnung für Matrizen und einige Anwendungen, *Electron. Rechenanl.* **10** (1968), pp. 73–83.
3. Aubin, J.-P.: *L'analyse Non Lineáire et ses Motivations Economiques*, Masson, Paris, 1984.
4. Birkhoff, G.: *Lattice Theory*, American Mathematical Society, Providence, R.I., 1967.
5. Clifford, A. H. and Preston, G. B.: *The Algebraic Theory of Semigroups*, American Mathematical Society, Providence, 1964.
6. Dem'yanov, V. F. and Malozemov, V. N.: *Introduction to Minimax*, John Wiley & Sons, New York, 1974.
7. Gardeñes, E. and Trepat, A.: Fundamentals of SIGLA, an Interval Computing System over the Completed Set of Intervals, *Computing* **24** (1980), pp. 161–179.
8. Hansen, E.: Bounding the Solution of Interval Linear Equations, *SIAM J. Numer. Anal.* **29** (1992), pp. 1493–1503.
9. Kaucher, E.: Algebraische Erweiterungen der Intervallrechnung unter Erhaltung Ordnungs- und Verbandsstrukturen, *Computing Suppl.* **1** (1977), pp. 65–79.
10. Kaucher, E.: Interval Analysis in the Extended Interval Space $\mathbb{IR}$, *Computing Suppl.* **2** (1980), pp. 33–49.
11. Kreinovich, V., Lakeyev, A., and Rohn, J.: Computational Complexity of Interval Algebraic Problems: Some Are Feasible and Some Are Computationally Intractable—A Survey, in: Alefeld, G., Frommer, A., and Lang, B. (eds), *Scientific Computing and Validated Numerics*, Akademie Verlag, Berlin, 1996, pp. 293–306.
12. Lyashko, M. A.: *On the Speed of Convergence of the Total Step Iteration Method for Interval SLAE*, Balashov State Pedagogical Institute, Balashov, 1996, deposited in VINITI, No. 430–B96.

13. Mayer, O.: Algebraische und metrische Strukturen in der Intervallrechnung und einige Anwendungen, *Computing* **5** (1970), pp. 144–162.
14. Neumaier, A.: *Interval Methods for Systems of Equations*, Cambridge University Press, Cambridge, 1990.
15. Nickel, K.: Die Auflösbarkeit linearer Kreisscheiben- und Intervall-Gleichungssystemen, *Linear Algebra Appl.* **44** (1982), pp. 19–40.
16. Ning, S. and Kearfott, R. B.: A Comparison of Some Methods for Solving Linear Interval Equations, *SIAM J. Numer. Anal.* **34** (1997), to appear.
17. Ortega, J. M. and Rheinboldt, W. C.: *Iterative Solutions of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
18. Ratschek, H. and Sauer, W.: Linear Interval Equations, *Computing* **28** (1982), pp. 105–115.
19. Rockafellar, R. T.: *Convex Analysis*, Princeton University Press, Princeton, 1970.
20. Rohn, J.: Cheap and Tight Bounds: The Recent Results by E. Hansen Can Be Made More Efficient, *Interval Computations* 4 (1993), pp. 13–21.
21. Shary, S. P.: On One Interval Problem of Linear Algebra, in: *Informacionno-operativnyi material*, Preprint 2 (1987), Computer Center of Siberian Department of the USSR Academy of Sciences, Krasnoyarsk, pp. 45–46 (in Russian).
22. Shary, S. P.: On Optimal Solution of Interval Linear Equations, *SIAM J. Numer. Anal.* **32** (1995), pp. 610–630.
23. Shary, S. P.: Algebraic Approach to the Interval Linear Static Identification, Tolerance and Control Problems, or One More Application of Kaucher Arithmetic, *Reliable Computing* **2** (1) (1996), pp. 3–33.
24. Shary, S. P.: Algebraic Solutions to Interval Linear Equations and Their Applications, in: Alefeld, G. and Herzberger, J. (eds), *Numerical Methods and Error Bounds*, Akademie Verlag, Berlin, 1996, pp. 224–233.
25. Shary, S. P.: A New Approach to the Analysis of Static Systems under Interval Uncertainty, in: Alefeld, G., Frommer, A., and Lang, B. (eds), *Scientific Computing and Validated Numerics*, Akademie Verlag, Berlin, 1996, pp. 118–132.
26. Shokin, Yu. I.: On Interval Problems, Interval Algorithms and Their Computational Complexity, in: Alefeld, G., Frommer, A., and Lang, B. (eds), *Scientific Computing and Validated Numerics*, Akademie Verlag, Berlin, 1996, pp. 314–328.