On the Axiomatization of Interval Arithmetic

Svetoslav Markov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
"Acad. G. Bonchev" st., block 8, 1113 Sofia, Bulgaria
smarkov@bio.bas.bg

Abstract

We investigate some abstract algebraic properties of the system of intervals with respect to the arithmetic operations and the relation inclusion and derive certain practical consequences from these properties. In particular, we discuss the use of improper intervals (in addition to proper ones) and of midpoint-radius presentation of intervals. This work is a theoretical introduction to interval arithmetic involving improper intervals. We especially stress on the existence of special "quasi"-multiplications in interval arithmetic and their role in relevant symbolic computations.

1. Introduction

In this work we discuss several algebraic properties of the system of intervals with the arithmetic operations addition and multiplication and the relation inclusion. Our aim is to point out certain practical advantages of using improper intervals and midpoint-radius presentation of intervals.

Denote by $I\mathbb{R}$ the set of all compact intervals on the real line $\mathbb{R}$ and by $I\mathbb{R}^n$ the set of all n-tuples of intervals. For $A, B \in I\mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$, one defines addition, multiplication by scalars and inclusion, resp., by:

$$A + B = \{\alpha + \beta \mid \alpha \in A, \beta \in B\}, \quad (1)$$
$$\gamma * B = \{\gamma\beta \mid \beta \in B\}, \quad (2)$$
$$A \subseteq B \iff (\alpha \in A \Longrightarrow \alpha \in B), \quad (3)$$

where all operations/relations are understood componentwise. We thus obtain the system $(I\mathbb{R}^n, +, \mathbb{R}, *, \subseteq)$ to be discussed in the sequel.

We shall refer to the definitions of the operations/relations (1)–(3) as *set-theoretic*. These definitions are not suitable for computations with intervals. Our final aim is to derive computationally efficient expressions for these operations based on the intrinsic properties of intervals.

In our study we follow the algebraically natural approach of completing the set $I\mathbb{R}^n$ up to the set $\mathbb{IR}^n$ involving improper intervals. Thus, starting from the above three *basic* operations/relations (1)–(3), we arrive to the system $(\mathbb{IR}^n, +, \mathbb{R}, *, \subseteq)$. In the next Section 2 we briefly recall the above mentioned algebraic construction and the consequent quasivector spaces. It is to be noted that everything said in this section for intervals is also true for the more general case of convex bodies and briefly repeats already published materials, cf. [7], [8]. In Section 3 we concentrate on the system $(\mathbb{IR}^n, +, \mathbb{R}, *, \subseteq)$ obtained by algebraic completion. Using the theoretical foundations given in Section 2 and some specific properties of intervals (distinct from those of general convex bodies) we derive formulae for the operations/relations involved. We show that the familiar midpoint-radius presentation of intervals is a special case of the presentation of elements in a quasivector space. Section 4 is devoted to the system $(\mathbb{IR}, +, \times, \subseteq)$ involving multiplication of one-dimensional intervals. In the Conclusion we discuss various topics like computer implementation of interval arithmetic, symbolic computations, etc.

2. Quasivector spaces

The system $(I\mathbb{R}^n, +)$ is a commutative monoid (semigroup with null) with cancellation law. There is no opposite operator in $(I\mathbb{R}^n, +)$. The operator multiplication by the scalar -1: $\neg A = (-1) * A = \{-\alpha \mid \alpha \in A\},\ A \in I\mathbb{R}$, briefly called *negation* (that may be suspected for opposite), is not an opposite operator, as $A + (\neg A) = 0$ is violated for certain $A \in I\mathbb{R}^n$. Thus $I\mathbb{R}^n$ is not a group; however it can be embedded in a group. The algebraic construction that converts an abelian monoid with cancellation law into a group will be further refered as *embedding construction*. Recall that this approach is used to pass from the monoid of nonnegative reals $(\mathbb{R}^+, +)$ to the set of reals $(\mathbb{R}, +)$. Thus, it is natural instead of the original system $(I\mathbb{R}^n, +, \mathbb{R}, *, \subseteq)$ to consider the extended system $(\mathbb{IR}^n, +, \mathbb{R}, *, \subseteq)$ obtained by the embedding construction.

2.1. The embedding construction

Every abelian monoid $(M,+)$ with cancellation law induces an abelian group $(\mathbb{M},+)$, where $\mathbb{M} = M^2/\sim$ is the *difference (quotient) set* of M consisting of all pairs (A,B) factorized by the congruence relation $\sim$: $(A,B) \sim (C,D)$ iff $A+D = B+C$, for $A,B,C,D \in M$.

Addition in $\mathbb{M}$ is defined by

$$(A,B)+(C,D) = (A+C, B+D). \quad (4)$$

The neutral (null) element of $\mathbb{M}$ is the class (Z,Z), $Z \in M$. Due to the existence of null element in M, we have $(Z,Z) \sim (0,0)$. The opposite element to $(A,B) \in \mathbb{M}$ is $\mathrm{opp}(A,B) = (B,A)$. The mapping $\varphi : M \longrightarrow \mathbb{M}$ defined for $A \in M$ by $\varphi(A) = (A,0) \in \mathbb{M}$ is an *embedding* of monoids. We *embed* M in $\mathbb{M}$ by identifying $A \in M$ with the equivalence class $(A,0) \sim (A+X,X)$, $X \in M$; all elements of $\mathbb{M}$ admitting the form $(A,0)$ are called *proper* and the remaining (new) elements are called *improper*. The set of all proper elements of $\mathbb{M}$ is $\varphi(M) = \{(A,0) \mid A \in M\} \cong M$.

Using the above construction the system $(I\mathbb{R}^n,+)$ is embedded into the group $(\mathbb{IR}^n,+)$ in a unique way.

Multiplication by scalars "$*$" is extended from $\mathbb{R}\times I\mathbb{R}^n$ to $\mathbb{R}\times\mathbb{IR}^n$ by means of

$$\gamma * (A,B) = (\gamma * A, \gamma * B), \quad A,\ B \in I\mathbb{R}^n, \quad \gamma \in \mathbb{R}. \quad (5)$$

In particular, negation is extended by $\neg(A,B) = (-1)*(A,B) = (\neg A, \neg B)$, $\ A,\ B \in I\mathbb{R}^n$.

In the sequel we shall use lower case roman letters to denote the elements of $\mathbb{IR}^n$, writing e. g. $a = (A_1,A_2), A_1,A_2 \in I\mathbb{R}^n$. For example, negation is written: $\neg a = (-1)*a$; below $a \neg b$ means $a + (\neg b)$.

Inclusion "$\subseteq$" is extended in $\mathbb{IR}$ by means of

$$(A,B) \subseteq (C,D) \Longleftrightarrow A+D \subseteq B+C, \quad (6)$$

wherein $A,B,C,D \in I\mathbb{R}^n$ and inclusion of interval n-tuples is meant component-wise. As is immediately seen, under this extension the practically important properties $a \subseteq b \Longleftrightarrow a+c \subseteq b+c$ for $c \in \mathbb{IR}$ and $a \subseteq b \Longleftrightarrow \gamma*a \subseteq \gamma*b$ for $\gamma \in \mathbb{R}$ are preserved. The system $(\mathbb{IR}^n,+,\mathbb{R},*,\subseteq)$ involving improper intervals is now completely defined.

2.2. Quasivector space: definition

The interval system $(\mathbb{IR}^n,+,\mathbb{R},*)$ is a *quasi-vector space* in the sense of the following definition [8]:

Definition. A *quasi-vector space (over* $\mathbb{R}$*)*, denoted $(\mathbb{Q},+,\mathbb{R},*)$, is an abelian group $(\mathbb{Q},+)$ with a mapping (multiplication by scalars) "$*$": $\mathbb{R}\times\mathbb{Q} \longrightarrow \mathbb{Q}$, such that for $a,b,c \in \mathbb{Q}$, $\ \alpha,\beta,\gamma \in \mathbb{R}$:

$$\gamma*(a+b) = \gamma*a+\gamma*b, \quad (7)$$

$$\alpha*(\beta*c) = (\alpha\beta)*c, \quad (8)$$

$$1*a = a, \quad (9)$$

$$(\alpha+\beta)*c = \alpha*c+\beta*c, \quad \text{if } \alpha\beta \ge 0. \quad (10)$$

The only difference between a vector (linear) space and a quasivector space is contained in assumption (10), where the relation $(\alpha+\beta)*c = \alpha*c+\beta*c$ is required to hold just for $\alpha\beta \ge 0$, whereas in a vector space the same relation is assumed to hold for all scalars $\alpha,\beta \in \mathbb{R}$ (known as second distributive law). Thus a vector space is a special quasivector space.

Conjugate elements. From $\mathrm{opp}(a)+a = 0$ we obtain $\neg\mathrm{opp}(a) \neg a = 0$, that is $\neg\mathrm{opp}(a) = \mathrm{opp}(\neg a)$. The element $\neg\mathrm{opp}(a) = \mathrm{opp}(\neg a)$ is further denoted by a_- and the corresponding operator is called *dualization* or *conjugation*. We say that a_- is the conjugate (or dual) of a. In the sequel we shall express the opposite element symbolically as: $\mathrm{opp}(a) = \neg a_-$, minding that $a+(\neg a_-) = 0$ (to be briefly written as $a \neg a_- = 0$). Using conjugate elements the quasistributive law (10) can be written in the form

$$(\alpha+\beta)*c_{\sigma(\alpha+\beta)} = \alpha*c_{\sigma(\alpha)}+\beta*c_{\sigma(\beta)}, \quad (11)$$

wherein σ is the sign functional

$$\sigma(\alpha) = \begin{cases} -, & \alpha < 0; \\ +, & \alpha \ge 0, \quad \alpha \in \mathbb{R}, \end{cases}$$

and the convention $a_+ = a$ has been made (for a proof see [8]). Expression (11) is valid for all values of α,β (not only for equally signed α,β) which allows efficient symbolic calculations.

2.3. A decomposition theorem

An element y with the property $y \neg y = 0$ (equivalently $y = y_-$) is called *linear* or *distributive*; an element z such that $\neg z = z$ (equivalently $z + z_- = 0$) is called *centred* or *0-symmetric*.

Theorem (Decomposition theorem). $(\mathbb{Q},+,\mathbb{R},*)$ is a quasivector space. For every $x \in \mathbb{Q}$ there exist unique $y,z \in \mathbb{Q}$ such that: i) $x = y+z$; ii) $y \neg y = 0$; iii) $\neg z = z$; iv) $y = z \Longrightarrow y = z = 0$.

The proof, see [8], is based on the fact that any $x \in \mathbb{Q}$ can be written in the form:

$$x = y+z = (1/2)*(x+x_-)+(1/2)*(x \neg x). \quad (12)$$

Note that the first summand in (12) is linear, whereas the second one is centred. The subset of all linear elements of $\mathbb{Q}$ is denoted $\mathbb{Q}' = \{x \in \mathbb{Q} \mid x \neg x = 0\}$ and the subset of all centred elements of $\mathbb{Q}$ is denoted $\mathbb{Q}'' = \{x \in \mathbb{Q} \mid x = \neg x\}$.

Note that negation coincides with opposite in $\mathbb{Q}'$, and negation coincides with identity in $\mathbb{Q}''$.

Corollary 1. Every quasivector space $\mathbb{Q}$ is a direct sum of $\mathbb{Q}' = \{x \in \mathbb{Q} \mid x \neg x = 0\}$ and $\mathbb{Q}'' = \{x \in \mathbb{Q} \mid x = \neg x\}$, symbolically $\mathbb{Q} = \mathbb{Q}' \bigoplus \mathbb{Q}''$.

Clearly, $(\mathbb{Q}', +, \mathbb{R}, *)$ with $\mathbb{Q}' = \{x \in \mathbb{Q} \mid x \neg x = 0\}$ is a linear space. Indeed, conjugation in $\mathbb{Q}'$ coincides with identity. Substituting $x = x_-$ in (11), we obtain that the familiar second distributive law $(\alpha + \beta) * c = \alpha * c + \beta * c$, holds for all real α, β as required in the definition of a linear space.

To characterize $\mathbb{Q}'' = \{x \in \mathbb{Q} \mid x = \neg x\}$, note that $\mathbb{Q}''$ is a quasivector space of centred elements, to be briefly called *centred quasivector space*. As $x = \neg x$ is equivalent to $x + x_- = 0$, conjugation coincides with the opposite operator in $\mathbb{Q}''$. The centred quasivector space $\mathbb{Q}''$ can be converted into a linear space by re-defining multiplication by scalars as follows:

$$\alpha \cdot c = \alpha * c_{\sigma(\alpha)}, \quad c \in \mathbb{Q}''. \tag{13}$$

Now $(\mathbb{Q}'', +, \mathbb{R}, \cdot)$ is a linear space. Indeed, substituting (13) in (11), we obtain that the second distributive law $(\alpha + \beta) \cdot c = \alpha \cdot c + \beta \cdot c$ is valid for all real α, β; furthermore, properties (7)–(9) remain valid: $\gamma \cdot (a + b) = \gamma \cdot a + \gamma \cdot b, \alpha \cdot (\beta \cdot c) = (\alpha\beta) \cdot c, 1 \cdot a = a$.

For a better distinction between the two multiplications by scalars, we call "$*$" *quasi-vector* multiplication by scalars and "$\cdot$" — *linear* multiplication by scalars. Note that the linear multiplication by scalars "$\cdot$" defined by (13) coincides in $\mathbb{Q}'$ with the quasi-vector one, as in $\mathbb{Q}'$ relation (13) becomes $\alpha \cdot c = \alpha * c$ (due to $c_- = c$). Therefore

Corollary 2. The space $(\mathbb{Q}, +, \mathbb{R}, \cdot) = (\mathbb{Q}', +, \mathbb{R}, \cdot) \bigoplus (\mathbb{Q}'', +, \mathbb{R}, \cdot)$, with "$\cdot$" defined by (13), is a linear space.

Formula (13) gives an expression in $\mathbb{Q}''$ for the linear multiplication by scalars in terms of the quasi-vector one. Conversely, in $\mathbb{Q}''$ the quasi-vector multiplication by scalars is expressed by the linear one via

$$\alpha * c = |\alpha| \cdot c, \quad c \in \mathbb{Q}''. \tag{14}$$

Naturally, the quasi-vector multiplication by scalars is inclusion isotone, that is $a \leq b \Longrightarrow \gamma * a \leq \gamma * b$, $\gamma \in \mathbb{R}$, as $a \leq b \Longrightarrow |\gamma| a \leq |\gamma| b$ (which is not true for the linear multiplication by scalar).

2.4. Presentation of elements

The spaces involved in Corollary 2 are vector spaces. If we want to efficiently compute within these spaces we should assume them to be finite dimensional so that their elements have finite presentation. Thus we shall assume that the two vector spaces $(\mathbb{Q}', +, \mathbb{R}, \cdot)$, $(\mathbb{Q}'', +, \mathbb{R}, \cdot)$ are m-, resp. n-dimensional and therefore they are isomorphic to $\mathbb{R}^m$, resp. $\mathbb{R}^n$. Hence any $a \in \mathbb{Q}$ is a direct sum of the form $a = (a'; a'')$, $a' \in \mathbb{R}^m$, $a'' \in \mathbb{R}^n$. As elements of $\mathbb{Q}$ the elements of $\mathbb{Q}'$ are of the form $(a'; 0)$ and the elements of $\mathbb{Q}''$ — of the form $(0; a'')$, so that:

$$a = (a'; a'') = (a'; 0) + (0; a''). \tag{15}$$

Because of the presence of the special operation "$*$" and its different meaning in the two spaces, we shall continue to make a distinction between the spaces $(\mathbb{Q}', +, \mathbb{R}, \cdot)$, $(\mathbb{Q}'', +, \mathbb{R}, \cdot)$, calling the first one linear, and the second one centred (quasivector); also the elements of $\mathbb{Q}'$ will be called linear, and the elements of Q'' — centred. Let us recall that "$*$" and "$\cdot$" coincide in $\mathbb{Q}'$, but are distinct in Q'', so $\mathbb{Q}''$ will be always considered together with the two multiplications: $(\mathbb{Q}'', +, \mathbb{R}, \cdot, *)$.

Applying the Decomposition theorem we have that any finite dimensional quasivector space $\mathbb{Q}$ is a direct sum of two spaces — the linear space $(\mathbb{R}^m, +, \mathbb{R}, \cdot)$ and the centred quasivector space $(\mathbb{R}^n, +, \mathbb{R}, \cdot, *)$, symbolically: $(\mathbb{Q}, +, \mathbb{R}, \cdot, *) = (\mathbb{R}^m, +, \mathbb{R}, \cdot) \bigoplus (\mathbb{R}^n, +, \mathbb{R}, \cdot, *)$. Thus we can write down the operations addition and multiplication by scalars for intervals $a = (a'; a''),\ b = (b'; b'') \in \mathbb{Q}$ in the form:

$$a + b \ = \ (a'; a'') + (b'; b'') = (a' + b'; a'' + b''), \tag{16}$$
$$\gamma * b \ = \ \gamma * (b'; b'') = (\gamma * b'; \gamma * b'') = (\gamma b'; |\gamma| b''), \tag{17}$$

wherein $a', b' \in \mathbb{R}^m$, $a'', b'' \in \mathbb{R}^n$, $\gamma \in \mathbb{R}$ and in the last expression relation (14) is applied.

For example, in the space $(\mathbb{R}^n, +, \mathbb{R}, \cdot, *)$ for $n = 3$, the quasi-multiplication of centred elements by scalars looks as follows: $-2 * (1, 2, 1) = (2, 4, 2)$; $-1 * (-1, 2, -2) = (-1, 2, -2)$. Multiplication by -1 (negation) coincides with identity. To change the signs of the components of a centred element one should take the opposite (or dual) and not negation, e. g. $\mathrm{opp}(-1, 2, -2) = (-1, 2, -2)_- = (1, -2, 2)$. For negation, subtraction, opposite and dual in $\mathbb{Q}$ we have:

$$\neg a \ = \ -1 * (a'; a'') = (-a'; a''), \tag{18}$$
$$a \neg b \ = \ a + (-1) * b = (a' - b'; a'' + b''), \tag{19}$$
$$\mathrm{opp}\,(a) \ = \ \neg a_- = \neg (a'; a'')_- = (-a'; -a''), \tag{20}$$
$$\mathrm{dual}\,(a) \ = \ a_- = (a'; a'')_- = (a'; -a''). \tag{21}$$

3. Presentation of intervals

We see that any element of a finite quasivector space is represented in the form (15) as a sum of a linear and a centred component. This refers to arbitrary finite quasivector spaces, such as quasivector spaces of zonotopes [9]. We next wish to interpret formulae (18)–(15) for the case of interval vectors. To this end we shall consider the presentation of proper intervals.

Intervals (interval vectors) are usually represented by sets of real numbers. Two familiar presentations of intervals (interval vectors) are those by pairs of real numbers (vectors) either in end-point or midpoint-radius form. Up to now we intentionally avoided to represent intervals in any form. We now discuss the presentation of intervals by drawing consequences directly from their algebraic properties.

3.1. The case of proper intervals

Proper intervals as elements of $\mathbb{IR}^n$ are pairs of the form $(A, 0)$, where $A \in I\mathbb{R}^n$. Assume first that $(A, 0)$ is a linear element, that is $(A, 0) \neg (A, 0) = 0$; using (4) this means $A \neg A = 0$. As we know such properties possess exactly the point intervals $A \in I\mathbb{R}^n$, that is vectors from $\mathbb{R}^n$. Assume now that $(A, 0)$ is centred, that is $(A, 0) = \neg (A, 0)$; according to (4) this means $A = \neg A$, that is $A \in I\mathbb{R}^n$ is a centred interval (symmetric with respect to the origin).

Recall that point intervals are represented as $[\alpha, \alpha]$ and proper centred ones — as $[-\beta, \beta]$, where $\beta \geq 0$ is the radius of the centred interval. Any proper interval is written in *end-point form* as $[\alpha - \beta,\ \alpha + \beta],\ \beta \geq 0$.

In *midpoint-radius form* (briefly: MR-form) a proper interval $A \in I\mathbb{R}^n$ is written as $A = (a'; a'')$, $a'' \geq 0$ and is interpreted as the set:

$$A = (a'; a'') = \{\xi \in \mathbb{R}^n \mid |\xi - a'| \leq a''\}, \quad (22)$$

wherein the module of a vector is meant component-wise, i. e. $|(\alpha_1, ..., \alpha_n)| = (|\alpha_1|, ..., |\alpha_n|)$. Using MR-form a point interval is written as $(a'; 0)$ and a proper centred one as $(0; a'')$, $a'' \geq 0$. The component a' is called the *midpoint* and the component a'' — the *radius* or *error (bound)*.

Within the interpretation (22) the component a' is an element of the vector space $\mathbb{R}^n$, whereas the component $a'' \geq 0$ is an element of the monoid $(\mathbb{R}^n)^+$. The relation between the MR-presentation (22) and the end-point presentation is given by $(a'; a'') = [a' - a'', a' + a'']$, resp. $[e^-, e^+] = ((e^+ + e^-)/2;\ (e^+ - e^-)/2]$.

Let us see how proper elements of the form $a = (A, 0)$, $A \in I\mathbb{R}$, are decomposed according to the Decomposition theorem. For the linear and the centred parts (summands) of $a = (A, 0)$ we obtain respectively, cf. (12):

$$(1/2) * (a + a_-) = (1/2) * (A, \neg A), \quad (23)$$
$$(1/2) * (a \neg a) = (1/2) * (A \neg A, 0). \quad (24)$$

It is immediately seen that the centred part (24) of a proper interval is a proper interval. Denoting $A = (a'; a'')$, $a'' \geq 0$, using (19), we have $A \neg A = (a'; a'') + (-a'; a'') = (0; 2a'')$, which gives a value of (24) equal to $(0; a'')$.

Let us check if the linear part (23) of a proper element $(A, 0)$ is a proper element. Clearly, $(A, \neg A)$ is a proper element if there exists $X \in \mathbb{M}$ such that $(A, \neg A) = (X, 0)$, that is $A = X \neg A$. As we know this property is satisfied by intervals, indeed, we have $X = (2a'; 0)$, where a' is the midpoint of A. Hence, for the value of (23) we obtain $1/2 * X = (a'; 0)$. Summarizing we obtain for the proper interval $(a'; a'')$, $a'' \geq 0$: $(a'; a'') = (a', 0) + (0; a'')$, which is exactly of the form (15).

Remarks. 1. Presentation $(a'; a'') = (a', 0) + (0; a'')$ corresponds (for $a'' > 0$) to the symbolic form $a = a' \pm a''$ often used in engineering sciences. 2. Convex bodies with so-called Minkowski operations [7], [14] also form a quasivector space. However, for convex bodies the equation $A = X \neg A$ is not solvable in general and consequently, the linear part of a proper convex body may not be proper. For example, in the case of two-dimensional convex bodies, if A is a (proper) triangle, then such X does not exist and consequently the linear part of A is an improper element.

In the case of proper intervals ($a'' \geq 0$) we have

$$A \neg B \quad = \{\alpha - \beta \mid \alpha \in A, \beta \in B\}, \quad (25)$$

and, in particular, $\neg B = \{-\beta \mid \beta \in B\}$.

MR-presentation of improper intervals. Formula (15) demonstrates the practical meaning of the decomposition theorem: there the linear (point) interval of the form $(a'; 0)$ corresponds to the midpoint of a and the centred interval $(0; a'')$ corresponds to the radius (error bound) of a. A negative radius $a'' < 0$ corresponds to improper interval $(a'; a'')$. In the end-point form $[a' - a'', a' + a'']$ of an improper interval we have that the left end-point is greater than the right one, $a' - a'' \geq a' + a''$.

Let us check how improper one-dimensional elements of the form $a = (0, A)$, $A \in I\mathbb{R}$, are decomposed according to the Decomposition theorem.

For the linear and the centred part (summand) of $a = (0, A)$ we obtain respectively:

$$(1/2) * (a + a_-) = (1/2) * (\neg A, A), \quad (26)$$
$$(1/2) * (a \neg a) = (1/2) * (0, A \neg A). \quad (27)$$

Clearly, the centred part (27) of the improper interval $(0, A)$ is an improper interval. Denoting $A = (a'; a'')$, $a'' \geq 0$, we have $A \neg A = (a'; a'') + (-a'; a'') = (0; 2a'')$, hence the value of (27) is $\mathrm{opp}(0; a'') = (0; -a'')$.

Let us check if the linear part (26) of an improper element $(0, A)$ is an improper element. Clearly, $(\neg A, A)$ is an improper element if there exists $X \in \mathbb{M}$ such that $(\neg A, A) = (0, X)$, that is $A = X \neg A$. We have $X = (2a'; 0)$, where a' is the midpoint of a. Hence, for the value of (26) we obtain $\mathrm{opp}(1/2 * X) = (-a'; 0)$. We conclude that the linear part of any interval is always a (proper) point interval.

Summarizing we obtain for the improper interval $(0, A) = \mathrm{opp}(a'; a'')$, $a'' \geq 0$: $(-a'; -a'') = (-a'; 0) + (0; -a'')$, which is in agreement with (15). Therefore we can consider (15) as a generalization of the MR-presentation of intervals. The difference is that the value of a'' in (15) can be negative. We can thus speak of *negative radii* (or *negative errors*) corresponding to improper intervals.

In the n-dimensional case a'' in (15) belongs to $\mathbb{R}^n$ and thus the components of a'' can have negative values corresponding to improper one-dimensional intervals. We can again speak of n-dimensional *radii* (or *errors*) whose components are not necessarily nonnegative.

Proposition 1. Let $A, B \in I\mathbb{R}$ and $(A, B) \in \mathbb{IR}$ as defined in Section 2.1. If $A = (a'; a''), B = (b'; b'')$, then $(A, B) = (a' - b'; a'' - b'')$.

Proof. We have

$$(A, B) \quad = \quad \begin{cases} (X, 0), & \text{if } \ A = B + X, \\ (0, Y), & \text{if } \ A + Y = B, \end{cases}$$

where $X, Y \in I\mathbb{R}$. Minding that $(X, 0) = X = (a' - b'; a'' - b'')$, $Y = (b' - a'; b'' - a'')$, $(0, Y) = \mathrm{opp}(Y, 0) = (a' - b'; a'' - b'')$ we obtain $(A, B) = (a' - b'; a'' - b'')$.

Improper interval vectors as elements of $\mathbb{IR}^n$ are pairs of the form (A, B), where $A, B \in I\mathbb{R}^n$ and at least one pair (A_i, B_i) is an improper interval. In principle, we can assume that a' and a'' are vectors of different dimensions, as this is often needed in applications. For an example of $m > n$, one sometimes considers two-dimensional points that contain errors only in one of the components, say the ordinate. Alternatively, as an example of $m < n$, one may consider a point in the 3D space, which moves on one of the space axes (or on a curve in one of the coordinate planes) with an error bounded within a 3D box.

Due to the above we can state that formulae (18)–(21) are valid for all intervals (proper and improper).

3.2. Examples

Linear algebraic equations with interval right-hand side. Consider a linear interval $(n \times n)$-system of the form

$$A * x = b, \tag{28}$$

where $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ is a nonsingular matrix of reals and $b \in \mathbb{IR}^n$ is an interval vector. We are interested in vectors $x \in \mathbb{IR}^n$ satisfying (28); such x are called *interval solutions* (other authors call them *algebraic solutions* [1], p. 65, or *formal solutions* [13], [15]).

Denoting $x = (x'; x'') \in I\mathbb{R}^n$ we have $A * x = A * (x'; x'') = (A * x'; A * x'') = (Ax'; |A|x'')$. Thus the interval problem (28) reduces to two linear algebraic problems, one for the midpoints and one for the radii:

$$Ax' \quad = \quad b', \tag{29}$$
$$|A|\, x'' \quad = \quad b''. \tag{30}$$

Assuming that the real matrices A and $|A|$ are nonsingular we obtain for the solution of (29)–(30): $x' = A^{-1}b'$, $x'' = |A|^{-1}b''$. We must assume $|A|^{-1}b'' \geq 0$, so that $x'' \geq 0$. We thus obtain the following corollary:

Corollary. Given system (28), such that $A \in \mathbb{R}^{n \times n}$, $b = (b', b'') \in I\mathbb{R}^n$, assume that the real matrices A and $|A|$ are nonsingular and $|A|^{-1}b'' \geq 0$. Then there exists a unique interval solution x to (28).

Sets of solutions. For the determination of solution sets of algebraic problems with uncertain parameters it is of practical significance [2], [3], [15] to be able to find the set

$$\{x \in \mathbb{R}^n \mid A * x \subseteq b\}, \tag{31}$$

where $A = (A'; A'') \in \mathbb{IR}^{n \times n}$ is an interval matrix (with $A' = (a_{ij}{}') \in \mathbb{R}^{n \times n}$, $A'' = (a''_{ij}) \in \mathbb{R}^{n \times n}$) and $b \in \mathbb{IR}^n$ is an interval vector.

The multiplication in (31) is multiplication by scalars, hence we have the equivalences

$$\begin{aligned} A * x \subseteq b \quad & \Longleftrightarrow \quad (A'; A'') * x \subseteq (b'; b'') \\ & \Longleftrightarrow \quad (A'x; A''|x|) \subseteq (b'; b'') \\ & \Longleftrightarrow \quad |b' - A'x| \leq b'' - A''|x|. \end{aligned} \tag{32}$$

We know that the last equivalence is true for proper intervals. However, in the sequel we show that it is also true for arbitrary intervals, which is the practically important case in the above problem. At any case the solution sets problem (31) is reduced to a familiar system of algebraic inequalities.

As an example consider the case when the matrix A consists only of improper intervals, that is $A'' \leq 0$. Then dual $A = (A'; -A'')$ is a proper interval matrix (as $-A'' \geq 0$) and the inclusion $A*x \subseteq b$ is equivalent to dual $A*x \cap b \neq \emptyset$ (such x are called week solutions, cf. [1]). In this case (32) obtains the form: $|b' - A'x| \leq b'' + (-A'')|x|$, which is known as Oettli-Prager characterization, see, e.g. [1]), [10].

3.3. Inclusion

The system $(\mathbb{IR}, \subseteq)$. Interval inclusion (3) as defined for proper intervals $A = (a'; a''), B = (b'; b'') \in I\mathbb{R}$ is expressed in MR-form by: $A \subseteq B \Longleftrightarrow |b' - a'| \leq b'' - a''$. We shall prove that a relation of exactly the same form is true for any intervals (proper or improper).

Proposition 2. Interval inclusion (6) as defined for arbitrary intervals $a = (a'; a''), b = (b'; b'') \in \mathbb{IR}$ is expressed in MR-form by:

$$a \subseteq b \Longleftrightarrow |b' - a'| \le b'' - a''. \tag{33}$$

Proof. As in Section 2.1. define a, b by $a = (X, Y)$, $b = (U, V)$, where $X = (x'; x'')$, $Y = (y'; y'')$, $U = (u'; u'')$, $V = (v'; v'')$ are intervals from $I\mathbb{R}$. According to (6):

$$(X, Y) \subseteq (U, V) \Longleftrightarrow X + V \subseteq Y + U. \tag{34}$$

In MR-presentation (34): reads:

$$\begin{aligned}&(x' - y'; x'' - y'') \subseteq (u' - v'; u'' - v'')\\ &\quad\Longleftrightarrow (x' + v'; x'' + v'') \subseteq (y' + u'; y'' + u'')\\ &\quad\Longleftrightarrow |y' + u' - x' - v'| \le y'' + u'' - x'' - v''\\ &\quad\Longleftrightarrow |u' - v' - (x' - y')| \le u'' - v'' - (x'' - y''),\end{aligned}$$

which according to Proposition 1 is equivalent to (33).

Remark. From (33) we see that the case a proper and b improper is impossible (in this case we have $b'' - a'' < 0$ and (33) cannot be satisfied).

Lattice operations. Consider now the lattice operations $x \vee_{\subseteq} y = \sup_{\subseteq}(x, y)$, $x \wedge_{\subseteq} y = \inf_{\subseteq}(x, y)$ in the case $x, y \in \mathbb{IR}$. For $x \subseteq y$ we have $\sup_{\subseteq}(x, y) = y$ etc., but the general case is more involved. The proof of the next proposition is rather technical and will be omitted.

Proposition 3. If neither $x \subseteq y$ nor $y \subseteq x$ then denoting $\varepsilon = \mathrm{sign}(x' - y')(x'' - y'')$ we have for $x \vee y = \sup_{\subseteq}(x, y)$ and $x \wedge y = \inf_{\subseteq}(x, y)$, respectively, the values:

$$\left(\tfrac{x' + y' + \varepsilon|x'' - y''|}{2}; \ \tfrac{|x' - y'| + x'' + y''}{2}\right),$$

$$\left(\tfrac{x' + y' - \varepsilon|x'' - y''|}{2}; \ \tfrac{-|x' - y'| + x'' + y''}{2}\right).$$

Clearly, the values of $-|x' - y'| + x'' + y''$ may be negative, that is the interval $x \wedge y = \inf_{\subseteq}(x, y)$ may be improper. Hence, the lattice $(\mathbb{IR}, \subseteq)$ is complete, whereas $(I\mathbb{R}, \subseteq)$ is not.

Inclusion isotonicity is preserved in $\mathbb{IR}$ with respect to (16)–(17), i. e.

$$a \subseteq b \iff a + c \subseteq b + c \tag{35}$$

$$a \subseteq b \iff \gamma * a \subseteq \gamma * b. \tag{36}$$

Remark. A generalization of (33) for the n-dimensional case $a, b \in \mathbb{IR}^n$ is straightforward. The same holds true for Proposition 3 regarding the lattice operations.

4. Interval multiplication

This section is devoted to the system $(\mathbb{IR}, +, \times, \subseteq)$ involving multiplication of intervals. Multiplication of one-dimensional proper intervals $A, B \in I\mathbb{R}$ is defined by the set-theoretic expression:

$$A \times B = \{\alpha\beta \mid \alpha \in A, \ \beta \in B\}. \tag{37}$$

As before we are looking for a computationally efficient expression of (37). It has been proved [4] that multiplication (37) is extended uniquely in $\mathbb{IR}$ in such a way that :

$$a \subseteq b \Longrightarrow c \times a \subseteq c \times b, \quad a, b, c \in \mathbb{IR}. \tag{38}$$

Hence, we can assume that the system $(\mathbb{IR}, +, \times, \subseteq)$ is well-defined. To obtain a formula of MR type we first define a functional "relative error" κ as follows.

For $a \in I\mathbb{R}$ any non-centred interval $(a' \neq 0)$ can be written as: $a = (a'; a'') = |a'| * (\sigma(a'); a''/|a'|)$, wherein $\sigma(\alpha) = \mathrm{sign}(\alpha)$.

Examples: $(10; 1) = 10 * (1; 0.1)$; $(-2; 4) = 2 * (-1; 2)$.

Thus we can write: $a = (a'; a'') = |a'| * (\sigma(a'); \kappa(a))$, where $\kappa(a) = a''/|a'|$ is the relative error in a. When multiplication is considered the functional $\kappa(a)$ plays important roles. Note that the condition $\kappa(a) < 1$ means that a does not contain zero (in the sense of (3). The case of proper intervals has been fully discussed in [5]. We next extend the definition of the functional κ in $\mathbb{IR}$.

Intervals from $\mathbb{IR}$. For $a = (a'; a'') \in \mathbb{IR}$ we define κ by:

$$\kappa(a) = \begin{cases} |a''|/|a'|, & a' \neq 0; \\ \infty, & a' = 0, \end{cases}$$

Condition $\kappa(a) < 1$ means that a does not contain zero in the sense of (6), whereas $\kappa(a) \le 1$ means that a may contain zero only as an end-point.

To write an MR-formula for interval multiplication denote by $E(a)$ condition $\kappa(a) \le 1$ and by $C(a, b)$ condition

$$\kappa(a) > 1 \ge \kappa(b) \ \text{ or } \ (\kappa(a) \ge \kappa(b) > 1, a''b'' \ge 0).$$

Note that $E(b)$ means $\kappa(b) \le 1$ and $C(b, a)$ means

$$\kappa(b) > 1 \ge \kappa(a) \ \text{ or } \ (\kappa(b) \ge \kappa(a) > 1, a''b'' \ge 0).$$

The following MR-expression is equivalent to the respective end-point formula given by E. Kaucher [4]:

$$a \times b = (a'; a'') \times (b'; b'') = \tag{39}$$

$$\begin{cases} (a'b' + \sigma(a')\sigma(b')a''b''; |a'|b'' + |b'|a''), & \text{if } E(a) \& E(b); \\ (b' + \sigma(b')\sigma(a'')b'') * (a'; a''), & \text{if } C(a, b); \\ (a' + \sigma(a')\sigma(b'')a'') * (b'; b''), & \text{if } C(b, a); \\ 0, \quad \text{if } \kappa(a) \ge 1, \ \kappa(b) \ge 1, \ a''b'' \le 0. \end{cases}$$

Remark. The two intermediate cases involve multiplication by scalars; these cases can be viewed as a single case by a, b interchanging places, as in the FORTRAN-like algorithm below.

An algorithm for interval multiplication.

IF $E(a)$ and $E(b)$
THEN $a \times b = (a'b' + \sigma(a')\sigma(b')a''b''; |a'|b'' + |b'|a'')$
ELSE IF $\kappa(a) \geq 1,\ \kappa(b) \geq 1, a''b'' \leq 0$ THEN $a \times b = 0$
ELSE IF $C(a, b)$
THEN $a \times b = (b' + \sigma(b')\sigma(a'')b'') * (a'; a'')$
ELSE $a \times b = (a' + \sigma(a')\sigma(b'')a'') * (b'; b'')$

A compact end-point formula is published in [11]. A tool for visualizing the operation multiplication (39) using the above algorithm can be downloaded from [16]. To run the program the .NET Framework Redistributable Package (dotnetfx.exe) should be installed. The visualization tool offers the possibility to move continuously one of the arguments a, b in the product $a \times b$ by dragging the corresponding point by the mouse. This allows to demonstrate various properties of the product, like its continuity (which is not obvious from the conditional expression (39)), morphism w. r. t. conjugation, etc.

In the simple case of centred intervals we should have $(0; a) \subseteq (0; b) \Longrightarrow (0; c)\times(0; a) \subseteq (0; c)\times(0; b)$ for any $(0; c)$, or, equivalently $a \leq b \Longrightarrow c\times a \leq c\times b$ for all real a, b, c. This property should hold for the case when a and b take possibly negative values (as we know such a property does not hold for the familiar multiplication). This is why it is interesting to find a relation between the "quasi-multiplication" of radii and the familiar multiplication of reals. Substituting $a' = b' = 0$ in (39) we obtain:

$$a\times b = \begin{cases} ab, & \text{if } a \geq 0,\ b \geq 0, \\ -ab, & \text{if } a \leq 0,\ b < 0, \\ 0, & \text{if } a > 0,\ b < 0 \text{ or } a < 0,\ b > 0. \end{cases} \qquad (40)$$

Some examples: $(-2)\times(-3) = -6,\quad 2\times(-3) = 0.$

Apart of being inclusion isotone, we have for the quasi-multiplication: $-(a\times b) = (-a)\times(-b)$, which is also not true for the familiar multiplication.

Clearly the system of the radii $(\mathbb{R}, +, \times)$ is not a ring, but if we define a multiplication by $a \cdot b = \text{sign}(a)\text{sign}(b)(|a|\times|b|)$, then system $(\mathbb{R}, +, \cdot)$ is a ring, cf. [6]. Therefore the system of readiuses can be considered as a ring of reals endowed additionally with a quasi-multiplication (40), i. e. $(\mathbb{R}, +, \cdot, \times)$.

5. Conclusions

This work is a theoretical introduction to interval arithmetic. We show that the rigorous theory naturally involves improper intervals and points out the basic role of the midpoint-radius presentation of intervals. Let us summarize our arguments in favor of using the extended interval systems $(\mathbb{IR}^n, +, \mathbb{R}, *, \subseteq)$, $(\mathbb{IR}, +, \times, \subseteq)$ involving improper intervals ($\mathbb{IR}$-systems). The advantages of using $\mathbb{IR}$-systems instead of the respective $I\mathbb{R}$-systems $(I\mathbb{R}^n, +, \mathbb{R}, *, \subseteq)$, $(I\mathbb{R}, +, \times, \subseteq)$ can be compared with the advantages of using the familiar real vector space $\mathbb{R}^n$ and real field $\mathbb{R}$ instead of the respective systems $(\mathbb{R}^+)^n$, $\mathbb{R}^+$, involving only nonnegative reals. Recall that the system $(\mathbb{IR}^n, +)$ is a group, so an equation of the form $a + x = b$ is always solvable, whereas it is not always solvable in $(I\mathbb{R}, +)$. Moreover, $(I\mathbb{R}^n, \subseteq)$ is not a complete lattice, whereas $(\mathbb{IR}, \subseteq)$ is, so that lattice operations induced by the relation "$\subseteq$" always exist. In particular, the lattice operation "meet" $\inf_{\subseteq}(a, b)$ is well-defined, whereas it is not always available in $(I\mathbb{R}, \subseteq)$ (two proper intervals may not intersect).

The $\mathbb{IR}$-systems incorporate the respective $I\mathbb{R}$-systems, so everything that can be done in an $I\mathbb{R}$-system, can be also performed in the respective $\mathbb{IR}$-system using improper intervals. An user not familiar with improper intervals may not notice that the $\mathbb{IR}$-system handles them — unless certain final results cannot be expressed by proper intervals — then the $\mathbb{IR}$-system may produce improper intervals, whereas the $I\mathbb{R}$-system will issue a message for non-existing solutions.

Computations in the extended $\mathbb{IR}$-systems are in no way more complicated or time consuming than computations with proper intervals. On the contrary, the corresponding formulae may be simpler as nonnegativity tests may be avoided. Formulae (16), (17) hold for intervals from $\mathbb{IR}^n$ and thus, in particular, for proper intervals. This means that (16), (17) provide MR-expressions for the set-theoretic operations addition (1) and multiplication by scalars (2) involving proper intervals. Note that there are no simpler formulae than (16), (17) for the operations addition (1) and multiplication by scalars (2) for proper intervals; we mention this in order to dispel the myth that the arithmetic using improper intervals is more involved than the arithmetic for proper intervals. A close examination of formula (39) leads us to a similar conclusion: this expression is computationally as efficient as the respective expressions for proper intervals.

An axiomatic introduction to interval arithmetic necessarily uses improper intervals and has certain methodological advantages. Such an introduction shows which are the basic concepts. We have shown that the operations/relations (1)–(3) and (37) are in the bases of the whole theory.

The axiomatic approach passing through improper intervals demonstrates the different algebraic nature of the spaces of midpoints and radii. Many authors have pointed out practical arguments regarding the need of a differentiated approach in the presentation of midpoints and radii in the computational practice, cf. [10] (p. 29–30), [12]. [17].

The general idea is that for the computer presentation of the radii we normally need a much smaller number of digits than for the midpoints — maximum 2–3 decimal digits. From computational point of view this makes the MR-form faster than the end-point form and should be preferred as a basic form in a computer implementation; end-point form of I/O data can be easily available as a secondary optional presentation.

As another argument in favor of improper intervals we mention the increasing number of practical problems, such as solution sets of linear and nonlinear problems, functional ranges, etc., and numerical methods that require such intervals, cf. [2], [3], [13], [15], [18].

There is an ongoing discussion on the implementation of symbolic interval arithmetic in computer algebra systems. Hopefully this work offers convincing motivation in favor to an implementation involving improper intervals and MR-presentation of intervals. Expression (11) is an instructive example that shows a possible way of overcoming problems related to symbolic computations with conditional formulas like (10), typical for interval arithmetic. Note that there is no similar formula in the monoid $I\mathbb{R}$, so we have to use a space involving improper intervals to this end.

Note that when *end-point presentation* $[\alpha - \beta, \alpha + \beta]$, $\beta \geq 0$ is used, then the point and centred intervals cannot belong to spaces of different dimensions (as then we shall not know what is, say, $\alpha + \beta$. Clearly, the notation $I\mathbb{R}^n$ is closely related to the end-point presentation. However, as we have seen, the midpoint and radius vectors need not be of same dimension. For example, one often considers intervals that have two-dimensional midpoint component, and one-dimensional radius component (errors in one of the two coordinates only). Different dimensions of the spaces of linear and centred elements, as prescribed by the Decomposition theorem, are possible within the midpoint-radius presentation.

References

[1] Fiedler, M., Nedoma, J., Ramik, J., Rohn, J., Zimmermann, K., Linear Optimization Problems with Inexact Data, Springer, 2006.

[2] Goldsztejn, A., G. Chabert, On the approximation of linear AE-solution sets. In: 12th GAMM IMACS International Symposion on Scientific Computing, Computer Arithmetic and Validated Numerics, Duisburg, Germany (2006), submitted.

[3] Goldsztejn, A., G. Chabert, A Generalized Interval LU-decomposition for the Solution of Interval Linear Systems, submitted to LNCS.

[4] Kaucher, E., Interval Analysis in the Extended Interval Space $\mathbb{IR}$. Computing Suppl. **2** (1980) 33–49.

[5] Kulpa, Z., S. Markov, On the inclusion properties of interval multiplication: A diagrammatic study, BIT 43(4), 741–810, 2003.

[6] Markov, S. On the Algebraic Properties of Errors, PAMM (Proc. Appl. Math. Mech.) 1, 2002, 506-507.

[7] Markov, S., On Quasilinear Spaces of Convex Bodies and Intervals, Journal of Computational and Applied Mathematics 162 (1), 93-112, 2004.

[8] Markov, S., Quasivector Spaces and their Relation to Vector Spaces, EJMC (Electronic Journal on Mathematics of Computation) **2-1** (2005), 1–21. (http://www.cin.ufpe.br/ejmc)

[9] Markov, S., On the Quasivector Space of Zonotopes in the Plane, PAMM (Proc. Appl. Math. Mech.) 5 (1), 2005, 711–712.

[10] Neumaier, A., Interval Methods for Systems of Equations, Cambridge University Press, 1990.

[11] Popova, E., Multiplication distributivity of proper and improper intervals. Reliable computing **7**(2) (2001) 129–140.

[12] Rump, S. M., Fast and Parallel Interval Arithmetic, BIT **39** (1999), 3, 534–554.

[13] Sainz M. A., Gardenyes E., L. Jorba, Formal Solution to Systems of Interval Linear or Non-Linear Equations. Reliable computing, 8(3), 189211, 2002.

[14] Schneider, R. Convex Bodies: The Brunn-Minkowski Theory. Cambridge Univ. Press, 1993.

[15] Shary, S., A new technique in systems analysis under interval uncertainty and ambiguity, Reliable computing **8** (2002) 321–418.

[16] Surilov, R., S. Markov, A visualization software tool for interval operations, freely downloadable from: http://www.math.bas.bg/˜bio/IntervalOperations.exe

[17] Van Emden, M. H., On the Significance of Digits in Interval Notation, Reliable Computing 10, 1, 45–58.

[18] Wolff von Gudenberg, J., V. Kreinovich, A Full Function-Based Calculus of Directed and Undirected Intervals: Markov's Interval Arithmetic Revisited, Numerical Algorithms 37, 2004, 417–428.