A Full Function-Based Calculus of Directed and Undirected Intervals: Markov's Interval Arithmetic Revisited

Jürgen Wolff von Gudenberg^{1,2} and Vladik Kreinovich²

¹Lehrstuhl für Informatik II Universität Würzburg Am Hubland D-97074 Würzburg, Germany email wolff@informatik.uni-wuerzburg.de

²Department of Computer Science University of Texas at El Paso 500 W. University El Paso, TX 79968, USA email vladik@cs.utep.edu

Abstract

One of the main objectives of interval computations is to compute the range $f([\underline{x}_1, \overline{x}_1], \dots, [\underline{x}_n, \overline{x}_n])$ of a given continuous function $f(x_1, \dots, x_n)$ over given intervals $[\underline{x}_1, \overline{x}_1], \dots, [\underline{x}_n, \overline{x}_n]$. From this viewpoint, the resulting interval $[\underline{y}, \overline{y}]$ means that we have a function whose range is contained in this interval. In other words, we can associate each interval $[\underline{y}, \overline{y}]$ with the family of all continuous functions f whose range is contained in this interval.

The simplest case is when we consider functions f(x) of one variable $x = x_1$, with $[\underline{x}_1, \overline{x}_1] = [0, 1]$. Then, the range is f([0, 1]), and each interval $[\underline{y}, \overline{y}]$ is associated with the class F of all continuous functions $f: [0, 1] \to R$ for which $f([0, 1]) \subseteq [y, \overline{y}]$.

For every arithmetic operation \odot (+, ·, etc.), we can define, for every two functions f(x) and g(x), the result $f \odot g$ as a function for which, for every x, $(f \odot g)(x) = f(x) \odot g(x)$, and then for every two classes F and G, we can define $F \odot G \stackrel{\text{def}}{=} \{f \odot g \mid f \in F \& g \in G\}$. This definition leads to standard interval arithmetic.

S. Markov showed that if we consider families of monotonic functions, we get Kaucher arithmetic. We propose a general definition of functional intervals, and describe explicit expressions for arithmetic operations with such "intervals".

Keywords: Interval computations; functional interpretation

AMS Subject Classification: 65G20, 65G30, 65G40

1 Introduction: Why Function-Based Calculus of Intervals

One of the main objectives of interval computations is [1, 3, 4, 8]:

Given: intervals $[\underline{x}_1, \overline{x}_1], \dots, [\underline{x}_n, \overline{x}_n]$, and a continuous function $f(x_1, \dots, x_n)$, Compute: the range

$$[\underline{y},\overline{y}] \stackrel{\text{def}}{=} f([\underline{x}_1,\overline{x}_1],\ldots,[\underline{x}_n,\overline{x}_n]) = \{f(x_1,\ldots,x_n) \mid x_i \in [\underline{x}_i,\overline{x}_i]\}$$

of the given function f over the given intervals $[\underline{x}_i, \overline{x}_i]$.

Thus, it seems reasonable to associate each interval $\mathbf{y} = [\underline{y}, \overline{y}]$ with the set of all continuous functions whose range is a subset of \mathbf{Y} .

The problem with this definition is that computing the exact range is difficult; the problem of computing the range is, in general, NP-hard (see, e.g., [5]). As a result, in practice, our main objective is not necessarily to compute the exact range, but rather to compute an enclosure $[\underline{Y}, \overline{Y}] \supseteq [\underline{y}, \overline{y}]$ for the desired range $[\underline{y}, \overline{y}]$ (and to make this enclosure as narrow as possible).

From this practical viewpoint, we would rather associate each interval $\mathbf{y} = [\underline{y}, \overline{y}]$ with the set of all continuous functions whose range is a *subset* of \mathbf{Y} .

2 Describing This Idea in Formal Terms

2.1 Simplest Case

The simplest case of the above problem of interval computations is when we consider functions f(x) of only one variable $x = x_1$, and we take the simplest possible interval $[\underline{x}_1, \overline{x}_1] = [0, 1]$.

2.2 Resulting Definition

In this case, the range is f([0,1]), and each interval $[\underline{a}, \overline{a}]$ is associated with the class of all continuous functions $f:[0,1]\to R$ for which $f([0,1])\subseteq [\underline{a}, \overline{a}]$. In other words, we arrive at the following definition:

$$[\underline{a}, \overline{a}]_F \stackrel{\text{def}}{=} \{ f \mid f \in C \& f([0, 1]) \subseteq [\underline{a}, \overline{a}] \}, \tag{1}$$

where C denotes the class of all continuous functions.

2.3 Operations on Intervals

Standard operations of interval arithmetic can be naturally interpreted in this functional form. Indeed, for every arithmetic operation \odot (+, -, ·, etc.), we can define,

- for every two functions f(x) and g(x), the result $f \odot g$ as a function for which, for every x, $(f \odot g)(x) = f(x) \odot g(x)$ (i.e., a component-wise operation); and then
- \bullet for every two classes F and G, we can define

$$F \odot G = \{ f \odot g \mid f \in F \& g \in G \}. \tag{2}$$

How are these operations related to standard operations on intervals? It turns out that this definition leads to standard interval arithmetic, for example:

Theorem 1.
$$[\underline{a}, \overline{a}]_F + [\underline{b}, \overline{b}]_F = [\underline{a} + \underline{b}, \overline{a} + \overline{b}]_F$$
.

Comments.

- For readers' convenience, the proofs of all the statements are placed in a separate Proofs section.
- Similar results can be proven for other operations of interval arithmetic (e.g., for subtraction, the proof is basically the same as for addition). So, from the viewpoint of interval computations, we can simply identify each interval $[\underline{a}, \overline{a}]$ with the corresponding family of functions $[\underline{a}, \overline{a}]_F$.

3 Beyond Standard Intervals

We have already mentioned that sometimes, we succeed in getting the *exact* range, but most often, we get an enclosure. It is desirable to separate these two cases, i.e., to consider, in addition to the formula (1), the class of functions for which the range is exactly equal to $[\underline{a}, \overline{a}]$. We will denote this class by

$$\underline{[\underline{a}, \overline{a}]} \stackrel{\text{def}}{=} \{ f \mid f \in C \& f([0, 1]) = [\underline{a}, \overline{a}] \}.$$
(3)

An important case when we can easily compute the exact range is when a function f(x) is monotonic:

- If we know that the function f(x) is non-decreasing (denoted $f \uparrow$), then we can compute the range as f([0,1]) = [f(0), f(1)].
- If we know that the function f(x) is non-increasing (denoted $f \downarrow$), then we can compute its range as f([0,1]) = [f(1), f(0)].

It is therefore reasonable to consider the corresponding sets of monotonic functions:

$$[\underline{a}, \overline{a}]^{\to} \stackrel{\text{def}}{=} \{ f \uparrow \mid f \in C \& f([0, 1]) \subseteq [\underline{a}, \overline{a}] \}; \tag{4}$$

$$[\underline{a}, \overline{a}] \stackrel{\leftarrow}{=} \{ f \downarrow \mid f \in C \& f([0, 1]) \subseteq [\underline{a}, \overline{a}] \}; \tag{5}$$

$$\underline{[\underline{a}, \overline{a}]} \stackrel{\text{def}}{\to} \{ f \uparrow \mid f \in C \& f([0, 1]) = [\underline{a}, \overline{a}] \};$$
(6)

$$[\underline{a}, \overline{a}] \stackrel{\leftarrow}{=} \{ f \downarrow \mid f \in C \& f([0, 1]) = [\underline{a}, \overline{a}] \}.$$
 (7)

Crudely speaking, the family $[\underline{a}, \overline{a}]^{\rightarrow}$ describes not only the *interval* itself, but also the *direction* in which a function f(x) covers this interval. These families can therefore be naturally associated with *directed intervals* as defined by S. Markov (see, e.g., [6, 7]).

4 Beyond Standard Operations

As we have mentioned, if we know that the function is monotonic, then this information drastically simplifies the problem of computing the range of this function. Sometimes, in addition to (or instead of) knowing that the original functions are monotonic, we also know that some intermediate computational results are monotonic. To capture such situations, in addition to the operations (2), it makes sense to consider new "directed" operations in which additionally require that the functions are monotonic.

Formally, if \odot be an arithmetic operation, and F and G are function families, then, in addition to the family $F \odot G$ as defined by the formula (2), we can also define the following families:

$$F \odot^{\to} G \stackrel{\text{def}}{=} \{h \uparrow \mid h \in F \odot G\}; \tag{8}$$

$$F \odot^{\leftarrow} G \stackrel{\text{def}}{=} \{h \downarrow \mid h \in F \odot G\}. \tag{9}$$

It is also reasonable to consider the case when the resulting function h is monotonic, but we do not whether it is non-increasing or non-decreasing, i.e., to consider the family

$$F \odot^{\leftrightarrow} G \stackrel{\text{def}}{=} \{h \uparrow \lor h \downarrow \mid h \in F \odot G\} = (F \odot^{\rightarrow} G) \cup (F \odot^{\leftarrow} G). \tag{10}$$

These new definitions are in line with "directed" operations between directed intervals – as defined by S. Markov.

5 Technical Comment: What Is the Functional Analogue of Interval Hull?

5.1 Problem

For original interval functional classes (corresponding to undirected intervals), simple arithmetic operations lead result in similar interval classes; however, for more complex interval classes and more complex operations, we may get very complex classes if we apply the above definitions.

For example, if we apply the operation $+^{\leftrightarrow}$ to two intervals [0,1] and [0,1], the resulting set consists of all monotonic (non-increasing or non-decreasing) functions with a range in [0,1]; this is not one of the originally defined functional intervals.

Since our objective is to design a calculus of (directed and undirected) *intervals*, we should thus take, as a result of the corresponding interval operation, not the resulting function class itself, but rather the smallest interval class which contains this result.

5.2 Similar Problem

A similar problem occurs in interval arithmetic, where in some problems, the solution set is not an interval. For example, the set of possible solutions to the interval equation $x^2 \in [1, 4]$ is not an interval, it is a union $X = [-2, -1] \cup [1, 2]$ of two intervals.

5.3 Usual Solution

When the solution set is not an interval, and we want to stay within interval arithmetic, we take the narrowest interval enclosure instead of the original set X. In other words, instead of the set X, we take its *interval hull* [X] that is defined as the intersection of all the intervals that contain the set X:

$$[X] \stackrel{\text{def}}{=} \bigcap \{\text{all intervals } \supseteq X\}.$$

The possibility to define the interval hull as the intersection comes from the fact that the intersection of an arbitrary family of intervals is also an interval (unless, of course, it is empty set).

For example, for the above solution set $X = [-2, -1] \cup [1, 2]$, the interval hull is equal to [X] = [-2, 2].

5.4 Natural Idea

It is natural to define, for functional interval classes, an interval hull of a given set as the intersection of all interval classes which contain this set.

5.5 Minor Problem

We are not yet ready to implement this idea because, for this idea to work, we must make sure that the intersection of interval classes is always an interval class, and this is not always the case: the intersection of the family $[\underline{a}, \overline{a}]^{\rightarrow}$ of all non-decreasing functions and the family $[\underline{a}, \overline{a}]^{\leftarrow}$ of all non-increasing functions is the set of all constant functions, which is neither a directed nor an undirected interval:

$$[\underline{a},\overline{a}]^{\,\rightarrow}\cap[\underline{a},\overline{a}]^{\,\leftarrow}=\{f\,|\,f\in C\,\&\,f([0,1])\subseteq[\underline{a},\overline{a}]\,\&\,f=\mathrm{const}\}.$$

It is therefore necessary to add these classes to our list of interval classes, i.e., define

$$[\underline{a}, \overline{a}]^{=} \stackrel{\text{def}}{=} \{ f \mid f([0, 1]) \subseteq [\underline{a}, \overline{a}] \& f = \text{const} \}; \tag{11}$$

$$[\underline{a}, \overline{a}] \stackrel{\text{def}}{=} \{ f \mid f([0, 1]) = [\underline{a}, \overline{a}] \& f = \text{const} \}.$$
 (12)

Let \mathcal{F} denote the class of all families of the type (1), (3), (4)–(7), and (11)–(12). Then, it is easy to prove that the resulting class is closed under intersection:

Theorem 2. The class \mathcal{F} is closed under intersection.

Now, we are ready for the final definition:

5.6 Final Definition

For each operation \odot (such as +, + \leftarrow , etc.), we define the new operation $\odot_{\mathcal{F}}$ as follows:

$$F \odot_{\mathcal{F}} G \stackrel{\text{def}}{=} \bigcap \{ H \in \mathcal{F} \mid F \odot G \subseteq H \}. \tag{13}$$

This is the operation that we will be using.

6 Example: Markov's Explanation of Kaucher Arithmetic Reformulated

Following Markov, let us consider intervals of the type $[\underline{a}, \overline{a}] \leftarrow \text{and } [\underline{a}, \overline{a}] \rightarrow$. To get Kaucher's arithmetic, in this section, we will:

- identify a proper interval $[\underline{a}, \overline{a}]$ $(\underline{a} \leq \overline{a})$ with $[\underline{a}, \overline{a}] \stackrel{\rightarrow}{\rightarrow}$, and
- identify an improper interval $[\overline{a}, \underline{a}]$ $(\underline{a} \leq \overline{a})$ with $[\underline{a}, \overline{a}] \leftarrow$.

For example, [0, 1] will mean $[0, 1] \rightarrow$, and [1, 0] will mean $[0, 1] \leftarrow$. As an operation +, we take $+ \leftrightarrow_{\mathcal{F}}$. The following result shows that we do get Kaucher arithmetic (originally proposed in [2]):

Theorem 3. In the above interpretation, for arbitrary real numbers a_1 , b_1 , a_2 , and b_2 , we have

$$[a_1, b_1] + [a_2, b_2] = [a_1 + a_2, b_1 + b_2].$$

In particular, we have [2,3] + [1,0] = [3,3] and [0,1] + [2,-2] = [2,1].

7 What Did We Gain? Some Understanding

We can produce similar results for operations with different interval classes. As a result, we get a uniform description of different interval operations. It is always a good idea to have a uniform explanation for several different things; but, as we will show, on top of that, we gain some deeper understanding of interval operations.

Example: interval arithmetic lacks some algebraic properties of operations with real numbers, e.g., in general, there is no distributivity:

$$A \cdot (B + C) \neq A \cdot B + A \cdot C$$
.

Our functional interpretation enables us to explain this non-distributivity. For example, let us consider the interpretation from the previous section in which each operation \odot is interpreted as \odot^{\leftrightarrow} ; in other words, the functions corresponding to all the intermediate results must be monotonic. Hence:

• In both sides, we compute, for every $x \in [0,1]$, the same value

$$a(x) \cdot (b(x) + c(x)) \equiv a(x) \cdot b(x) + a(x) \cdot c(x)$$
.

- In the left-hand side, the intermediate results are b+c and $a \cdot (b+c)$. This means that in defining this expression, we assume that the functions b+c and $a \cdot (b+c)$ are monotonic.
- In the right-hand side, the intermediate results are $a \cdot b$, $a \cdot c$, and $a \cdot (b+c)$. This means that in defining this expression, we assume that the functions $a \cdot b$, $a \cdot c$, and $a \cdot (b+c)$ are monotonic.

These are two different monotonicity assumptions, so no wonder we sometimes get different results.

8 What Next?

In our definitions, as the domain of all the functions, we used the simplest possible case when we have only one variables, and we the simplest possible interval – the interval [0, 1]. Natural questions:

- What is we use some other interval as a domain? one can easily see that the choice of the basic interval does not matter all the results are valid for any basic interval.
- What if we use a multi-D base? This is an open question.

9 Proofs

Proof of Theorem 1. To prove the equality of the two sets, we must prove that every element from the left set belongs to the right set, and vice versa.

1°. Let us first prove that every element h(x) of the set $[\underline{a}, \overline{a}]_F + [\underline{b}, \overline{b}]_F$ belongs to the set $[\underline{a} + \underline{b}, \overline{a} + \overline{b}]_F$.

Indeed, let $h \in [\underline{a}, \overline{a}]_F + [\underline{b}, \overline{b}]_F$. By definition of the sum of two function classes, the fact that the function h(x) belongs to the sum of the two function classes means that h(x) can be represented

as a (component-wise) sum h(x) = f(x) + g(x) of two function f(x) and g(x), where $f \in [\underline{a}, \overline{a}]_F$ and $g \in \left| \underline{b}, \overline{b} \right|_{\overline{b}}$.

By definition of these classes, $f \in [\underline{a}, \overline{a}]_F$ means that $f([0,1]) \subseteq [\underline{a}, \overline{a}]$, i.e., that for every $x \in [0,1]$, we have

$$\underline{a} \le f(x) \le \overline{a}$$
.

Similarly, we conclude that for every $x \in [0,1]$, we have

$$\underline{b} \le g(x) \le \overline{b}.$$

Adding the two displayed inequalities, we conclude that for every $x \in [0,1]$, we have

$$\underline{a} + \underline{b} \le h(x) = f(x) + g(x) \le \overline{a} + \overline{b}.$$

In other words, we have $h([0,1]) \subseteq \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right]$. By definition of the class $[.,.]_F$, this means that $h \in \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right]_F$.

 2° . To complete our proof, let us show that every element $h \in \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right]_F$ also belongs to the set $[\underline{a}, \overline{a}]_F + [\underline{b}, \overline{b}]_F$

Indeed, let $h \in \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right]_F$. To prove this statement, we must be able to represent the function h(x) in the form f(x)+g(x), where $f\in [\underline{a},\overline{a}]_F$ and $g\in \left[\underline{b},\overline{b}\right]_F$. For that, we will define an auxiliary function $F: \left[\underline{a} + \underline{b}, \overline{a} + \overline{b}\right] \to \left[\underline{a}, \overline{a}\right]_F$ and then take $f(x) \stackrel{\text{def}}{=} F(h(x))$ and $g(x) \stackrel{\text{def}}{=}$ h(x) - F(h(x)). We want to define the function F(z) in such a way that for every real number z for which this function is defined, we have $\underline{a} \leq F(z) \leq \overline{a}$ (this is true for any mapping to the interval $[\underline{a}, \overline{a}]$) and $\underline{b} \leq z - F(z) \leq \overline{b}$. Then, we will automatically have $f(x) \in [\underline{a}, \overline{b}]$ and $g(x) \in |\underline{b}, \overline{b}|$, hence $f \in [\underline{a}, \overline{a}]_F$ and $g \in [\underline{b}, \overline{b}]_F$. We define this function F(z) as follows:

- $F(z) = z \underline{b}$ for $\underline{a} + \underline{b} < z < \overline{a} + \underline{b}$, and
- $F(z) = \overline{a} \text{ for } \overline{a} + \underline{b} \le z \le \overline{a} + \overline{b}.$

In this case,

- when $\underline{a} + \underline{b} \le z \le \overline{a} + \underline{b}$, we have $z F(z) = \underline{b} \in [\underline{b}, \overline{b}]$;
- when $\overline{a} + \underline{b} \leq z \leq \overline{a} + \overline{b}$, then, by subtracting $\overline{a} = F(z)$ from all three terms in this inequality, we conclude that $\underline{b} \leq z F(z) \leq \overline{b}$.

In both cases, $z - F(z) \in \left[\underline{b}, \overline{b}\right]$, so we get the desired function F(z). The theorem is proven.

9.1 Proof of Theorem 2: Main Idea

In order to prove that the intersection belongs to the class, we can first intersect functional intervals belonging to the same class, and then take the intersection of the results.

First, we prove, for each class, that within this class, the intersection belongs to the same class (if it is non-empty at all); this is straightforward.

Then, by enumerating all possible pairs of classes, we prove that their intersection is also one of the standard functional intervals; this is somewhat tedious but also straightforward.

9.2 Proof of Theorem 3

We want to prove that for arbitrary real numbers a_1 , b_1 , a_2 , and b_2 , we have

$$[a_1, b_1] + [a_2, b_2] = [a_1 + a_2, b_1 + b_2].$$

To prove this result, let us recall what an interval [a, b] means in our functional interpretation:

- When $a \leq b$, this interval means the set of all non-decreasing functions f(x) for which the range is exactly [a, b]. Since the function is non-decreasing, its smallest value is attained at x = 0 and its largest value is attained at x = 1, hence f(0) = a and f(1) = b.
- When a > b, this interval means the set of all non-increasing functions f(x) for which the range is exactly [a, b]. Since the function is non-increasing, its largest value is attained at x = 0 and its smallest value is attained at x = 1, hence f(0) = a and f(1) = b.

In both cases, f(0) = a and f(1) = b.

In this interpretation, + is interpreted as the set of all monotonic functions h that can be represented as $h = f_1 + f_2$, where $f_1 \in [a_1, b_1]$ and $f_2 \in [a_2, b_2]$. Since $f_1 \in [a_1, b_1]$, we have $f_1(0) = a_1$ and $f_1(1) = b_1$. Similarly, we have $f_2(0) = a_2$ and $f_2(1) = b_2$, hence $h(0) = f_1(0) + f_2(0) = a_1 + a_2$ and $h(1) = f_1(1) + f_2(1) = b_1 + b_2$.

The function h is monotonic, so, we have two possibilities:

- if $a_1 + a_2 \le b_1 + b_2$, this function is non-decreasing hence it belongs to $[a_1 + a_2, b_1 + b_2]^{\rightarrow}$, i.e., to $[a_1 + a_2, b_1 + b_2]$;
- if $a_1 + a_2 > b_1 + b_2$, this function is non-increasing hence it belongs to $[b_1 + b_2, a_1 + a_2] \leftarrow$, i.e., also to $[a_1 + a_2, b_1 + b_2]$.

In both cases, every function from $[a_1, b_1] + [a_2, b_2]$ belongs to the interval class $[a_1 + a_2, b_1 + b_2]$. To complete the proof, it is sufficient to observe that $[a_1 + a_2, b_1 + b_2]$ cannot belong to a smaller class, because it attains both endpoints as values.

Acknowledgments

This work was supported in part by German Science Foundation, by NASA under cooperative agreement NCC5-209 and grant NCC2-1232, by Future Aerospace Science and Technology Program (FAST) Center for Structural Integrity of Aerospace Systems, effort sponsored by the Air Force Office of Scientific Research, Air Force Materiel Command, USAF, under grants numbers F49620-95-1-0518 and F49620-00-1-0365, by NSF grants CDA-9522207, EAR-0112968, EAR-0225670, and 9710940 Mexico/Conacyt, by IEEE/ACM SC2001 and SC2002 Minority Serving Institutions Participation Grants, and by a research grant from Sandia National Laboratories as part of the Department of Energy Accelerated Strategic Computing Initiative (ASCI).

This research was partly done when J.W.v.G. was a Visiting Researcher at the University of Texas at El Paso. The authors are thankful to all the participants of SCAN'02, especially to S. Markov and S. Utyubaev, for valuable discussions.

References

[1] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics, Springer-Verlag, London, 2001.

- [2] E. Kaucher, "Über Eigenschaften und Anwendungsmöglichkeiten der erweiterten Intervallrechnung und des hyperbolische Fastkörpers über R", Computing, Suppl. 1, 1977, pp. 81–94.
- [3] R. B. Kearfott, Rigorous Global Search: Continuous Problems, Kluwer, Dordrecht, 1996.
- [4] R. B. Kearfott and V. Kreinovich (eds.), Applications of Interval Computations, Kluwer, Dordrecht, 1996.
- [5] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, Computational complexity and feasibility of data processing and interval computations, Kluwer, Dordrecht, 1997.
- [6] S. M. Markov, "On Directed Interval Arithmetic and its Applications", *Journal of Universal Computer Science*, 1995, Vol. 1, No. 7, pp. 514–526.
- [7] S. M. Markov, "On the Algebra of Intervals and Convex Bodies", Journal of Universal Computer Science, 1995, Vol. 4, No. 1, pp. 34–47.
- [8] R. E. Moore, Methods and Applications of Interval Analysis, SIAM, Philadelphia, 1979.