Candidate Sets for Complex Interval Arithmetic

Jürgen Wolff von Gudenberg
Lehrstuhl für Informatik II
Universität Würzburg
D-97074 Würzburg, Germany
email `wolff@informatik.uni-wuerzburg.de`

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
email `vladik@cs.utep.edu`

Abstract

Uncertainty of measuring complex-valued physical quantities can be described by complex sets. These sets can have complicated shapes, so we would like to find a good approximating family of sets. Which approximating family is the best? We reduce the corresponding optimization problem to a geometric one: namely, we prove that, under some reasonable conditions, an optimal family must be shift-, rotation- and scale-invariant. We then use this geometric reduction to conclude that the best approximating low-dimensional families consist of sets with linear or circular boundaries. This result is consistent with the fact that such sets have indeed been successful in computations. It stimulates to study further candidates.

Construction of Optimal Families

A practical problem leading to complex sets. Many physical quantities are complex-valued: wave function in quantum mechanics, complex amplitude and impedance in electrical engineering, etc.

Due to measurement uncertainty, after measuring a value of a physical quantity, we do not get its *exact* value, we only get a *set* of possible values of this quantity. The shapes of these sets can be very complicated, so we would like to approximate them by simpler shapes from an approximating family. Which family should we choose?

In 1-D case, a similar problem has a simple solution: we choose the family of all (real) intervals. This family has many good properties; in particular, it is closed under point-wise arithmetic operations

$$A \odot B = \{a \odot b \,|\, a \in A, b \in B\}$$

such as addition, subtraction, and multiplication, which makes this family perfect for the analysis of how these measurement results get processed in a computer.

Unfortunately, for complex sets, no finite-dimensional family containing real intervals is closed under these operations [Nickel 1980]; moreover, no finite-dimensional family containing real intervals is closed under addition and under multiplication by complex *numbers*. This negative result has a clear geometric meaning, due to the fact that adding a complex number means a shift, and multiplication by a complex number $\rho \cdot \exp(\mathrm{i}\theta)$ means rotation by an angle θ and scaling ρ times. So, Nickel's negative result means it is impossible to have a finite-dimensional family of complex sets which would be closed under addition, invariant under shift, rotation, and scaling, and contain real intervals.

Since we cannot have an approximating family which satisfies all desired properties, we must therefore use families which satisfy only some of them. Several families have been proposed: boxes, polygons, circles, ellipsoids, etc. Some families approximate better, some approximate worse. So, an (informal) problem is: which approximating family is the best?

Of course, the more parameters we allow, the better the approximation. So, the question can be reformulated as follows: for a given number of parameters (i.e., for a given dimension of approximating family), which is the best family? In this paper, we formalize and solve this problem.

Formalizing the problem. All proposed families of sets have analytical (or piece-wise analytical) boundaries, so it is natural to restrict ourselves to such families. By definition, when we say that a piece of a boundary is analytical, we mean that it can be described by an equation $F(x, y) = 0$ for some analytical function

$$F(x, y) = a + bx + cy + dx^2 + exy + fy^2 + \ldots$$

So, in order to describe a family, we must describe the corresponding class of analytical functions $F(x, y)$.

Since we are interested in finite-dimensional families of sets, it is natural to consider finite-dimensional families of functions, i.e., families of the type

$$\{C_1 \cdot F_1(x, y) + \ldots + C_d \cdot F_d(x, y)\},$$

where $F_i(z)$ are given analytical functions, and $C_1, \ldots, C_d$ are arbitrary (real) constants. So, the question is: which of such families is the best?

When we say "the best", we mean that on the set of all such families, there must be a relation $\geq$ describing which family is better or equal in quality. This relation must be transitive (if A is better than B, and B is better than C, then A is better than C). This relation is not necessarily asymmetric, because we can have two approximating families of the same quality. However, we would like to require that this relation be *final* in the sense that it should define a unique *best* family $A_{\rm opt}$ (i.e., the unique family for which $\forall B\,(A_{\rm opt} \geq B)$. Indeed:

- If none of the families is the best, then this criterion is of no use, so there should be *at least one* optimal family.
- If *several* different families are equally best, then we can use this ambiguity to optimize something else: e.g., if we have two families with the same approximating quality, then we choose the one which is easier to compute. As a result, the original criterion was not final: we get a new criterion ($A \geq_{\rm new} B$ if either A gives a better approximation, or if $A \sim_{\rm old} B$ and A is easier to compute), for which the class of optimal families is narrower. We can repeat this procedure until we get a final criterion for which there is only one optimal family.

It is reasonable to require that the relation $A \geq B$ should not change if we add or multiply all elements of A and B by a complex number; in geometric terms, the relation $A \geq B$ should be shift-, rotation- and scale-invariant.

Now, we are ready for the formal definitions.

Definition 1. *Let $d > 0$ be an integer. By a d-dimensional family, we mean a family A of all functions of the type $\{C_1 \cdot F_1(x,y) + \ldots + C_d \cdot F_d(x,y)\}$, where $F_i(z)$ are given analytical functions, and $C_1, \ldots, C_d$ are arbitrary (real) constants. We say that a set is defined by this family A if its border consists of pieces described by equations $F(x,y) = 0$, with $F \in A$.*

Definition 2. *By an optimality criterion, we mean a transitive relation $\geq$ on the set of all d-dimensional families. We say that a criterion is final if there exists one and only one optimal family, i.e., a family $A_{\rm opt}$ for which $\forall B\,(A_{\rm opt} \geq B)$. We say that a criterion $\geq$ is shift- (corr., rotation- and scale-invariant) if for every two families A and B, $A \geq B$ implies $TA \geq TB$, where TA is a shift (rotation, scaling) of the family A.*

Proposition. ($d \leq 4$) *Let $\geq$ be a final optimality criterion which is shift-, rotation-, and scale-invariant, and let $A_{\rm opt}$ be the corresponding optimal family. Then, the border of every set defined by this family $A_{\rm opt}$ consists of straight line intervals and circular arcs.*

Comment. This result is in good accordance with numerical experiments, according to which such sets indeed provide a good approximation (see, e.g., [Alefeld et al. 1974], [Klatte et al. 1980], [Lerch et al. 1999]).

Proof. This proof is similar to the ones from [Nguyen et al. 1997].

1. Let us first show that the optimal family $A_{\rm opt}$ is itself shift-, rotation-, and scale-invariant.

Indeed, let T be an arbitrary shift, rotation, or scaling. Since $A_{\rm opt}$ is optimal, for every other family B, we have $A_{\rm opt} \geq T^{-1}B$ (where T^{-1} means the inverse transformation). Since the optimality criterion $\geq$ is invariant, we conclude that $TA_{\rm opt} \geq T(T^{-1}B) = B$. Since this is true for every family B, the family $TA_{\rm opt}$ is also optimal. But since our criterion is final, there is only one optimal family and therefore, $TA_{\rm opt} = A_{\rm opt}$. In other words, the optimal family is indeed invariant.

2. Let us now show that all functions from $A_{\rm opt}$ are polynomials.

Indeed, every function $F \in A_{\rm opt}$ is analytical, i.e., can be represented as a Taylor series (sum of monomials). Let us combine together monomials $cx^a y^b$ of the same degree $a+b$; then we get

$$F(z) = F_0(z) + F_1(z) + \ldots + F_k(z) + \ldots,$$

where $F_k(z)$ is the sum of all monomials of degree k. Let us show, by induction over k, that for every k, the function $F_k(z)$ also belongs to $A_{\rm opt}$.

Let us first prove that $F_0(z) \in A_{\rm opt}$. Since the family $A_{\rm opt}$ is scale-invariant, we conclude that for every $\lambda > 0$, the function $F(\lambda z)$ also belongs to $A_{\rm opt}$. For each term $F_k(z)$, we have $F_k(\lambda z) = \lambda^k F_k(z)$, so

$$F(\lambda z) = F_0(z) + \lambda F_1(z) + \ldots \in A_{\rm opt}.$$

When $\lambda \to 0$, we get $F(\lambda z) \to F_0(z)$. The family $A_{\rm opt}$ is finite-dimensional hence closed; so, the limit $F_0(z)$ also belongs to $A_{\rm opt}$. The induction base is proven.

Let us now suppose that we have already proven that for all $k < s$, $F_k(z) \in A_{\rm opt}$. Let us prove that $F_s(z) \in A_{\rm opt}$. For that, let us take

$$G(z) = F(z) - F_1(z) - \ldots - F_{s-1}(z).$$

We already know that $F_1, \ldots, F_{s-1} \in A_{\rm opt}$; so, since $A_{\rm opt}$ is a linear space, we conclude that

$$G(z) = F_s(z) + F_{s+1}(z) + \ldots \in A_{\rm opt}.$$

The family $A_{\rm opt}$ is scale-invariant, so, for every $\lambda > 0$, the function

$$G(\lambda z) = \lambda^s F_s(z) + \lambda^{s+1} F_{s+1}(z) + \ldots$$

also belongs to $A_{\rm opt}$. Since $A_{\rm opt}$ is a linear space, the function

$$H_\lambda(z) = \lambda^{-s} G(\lambda z) =$$
$$F_s(z) + \lambda F_{s+1}(z) + \lambda^2 F_{s+2}(z) + \ldots$$

also belongs to $A_{\rm opt}$.

When $\lambda \to 0$, we get $H_\lambda(z) \to F_s(z)$. The family $A_{\rm opt}$ is finite-dimensional hence closed; so, the limit $F_s(z)$ also belongs to $A_{\rm opt}$. The induction is proven.

Now, monomials of different degree are linearly independent; therefore, if we have infinitely many non-zero terms $F_k(z)$, we would have infinitely many linearly independent functions in a finite-dimensional family $A_{\rm opt}$ – a contradiction. Thus, only finitely many monomials $F_k(z)$ are different from 0, and so, $F(z)$ is a sum of finitely many monomials, i.e., a polynomial.

3. Let us prove that if a function $F(x,y)$ belongs to $A_{\rm opt}$, then its partial derivatives $F_{,x}(x,y)$ and $F_{,y}(x,y)$ also belong to $A_{\rm opt}$.

Indeed, since the family $A_{\rm opt}$ is shift-invariant, for every $h>0$, we get $F(x+h,y)\in A_{\rm opt}$. Since this family is a linear space, we conclude that a linear combination $h^{-1}(F(x+h,y)-F(x,y))$ of two functions from $A_{\rm opt}$ also belongs to $A_{\rm opt}$. Since the family $A_{\rm opt}$ is finite-dimensional, it is closed and therefore, the limit $F_{,x}(x,y)$ of such linear combinations also belongs to $A_{\rm opt}$. (For $F_{,y}$, the proof is similar).

4. Due to Parts 2 and 3 of this proof, if any polynomial from $A_{\rm opt}$ has a non-zero part F_k of degree $k>0$, then it also has a non-zero part ($(F_k)_{,x}$ or $(F_k)_{,y}$) of degree $k-1$. Similarly, it has non-zero parts of degrees $k-2,\ldots,1,0$.

So, in all cases, $A_{\rm opt}$ contains a non-zero constant and a non-zero linear function $F_1(x,y)=bx+cy$. We can now use the fact that the family $A_{\rm opt}$ is rotation-invariant; let T be a rotation which transforms (b,c) into the x-axis, then we conclude that

$$F_1(Tz)=b'x\in A_{\rm opt},$$

and hence $x\in A_{\rm opt}$. Similarly, $y\in A_{\rm opt}$. So, the family $A_{\rm opt}$ contains at least 3 linearly independent functions: a non-zero constant, x, and y.

If $d=3$, then the 3-D family $A_{\rm opt}$ cannot contain anything else, and all the pieces of borders $F(x,y)=0$ of all the sets defined by this family are straight lines.

If $d=4$, then we cannot have any cubic or higher order terms in $A_{\rm opt}$, because then, due to Part 3, we would have both this cubic part *and* a (linearly independent) quadratic part, and the total dimension of $A_{\rm opt}$ would be at least $3+2=5$. So, all functions from $A_{\rm opt}$ are quadratic. Since $\dim(A_{\rm opt})=4$, and the dimension of 0- and 1-D parts is 3, the dimension of possible parts of second degree is 1. Since $A_{\rm opt}$ is rotation-invariant, the quadratic part $dx^2+exy+fy^2$ must be also rotation-invariant (else, we would have two linearly independent quadratic terms in $A_{\rm opt}$: the original expression and its rotated version). Thus, this quadratic part must be proportional to x^2+y^2.

Hence, every function $F\in A_{\rm opt}$ has the form

$$F(x,y)=a+bx+cy+d(x^2+y^2),$$

and therefore, all the pieces of borders $F(x,y)=0$ of all the sets defined by this family are either straight lines or circular arcs. The proposition is proven.

Properties of Sub-Families

In this section we recall that the usually used implementations of complex interval arithmetic, e.g. axis parallel rectangles, circles, and circular sectors are subfamilies of $A_{\rm opt}$ and collect their properties.

In general we can describe a family of sets by choosing a specific number of boundary curves of the type:

$$\mathcal{C}: a+bx+cy+d(x^2+y^2)$$

Interesting for practical applications are those which lead to closed bounded sets, are characterized by a few parameters and have efficient implementations of the arithmetic operations. Since no family can be closd for addition and multiplication multiaspect combinations of several representations are favorable, see [Lerch et al. 1999]

Circles

- 1 boundary curve: $d\neq 0$
- 3 real parameters needed: a/d, b/d, c/d
- equivalent representation: 1 point, 1 radius
- properties
  - shift-invariant
  - rotation-invariant
  - scale-invariant
  - addition closed
  - multiplication not closed
  - multiplicative inversion closed
  - intersection not closed

Rectangles

- 4 boundary curves: two vertical lines $d=0, c=0, b\neq 0$ and two horizontal lines $d=0, b=0, c\neq 0$
- 4 real parameters needed: a_1/b_1, a_2/b_2, a_3/c_3, a_4/c_4
- equivalent representation: 2 points
- properties
  - shift-invariant
  - not rotation-invariant
  - scale-invariant
  - addition closed
  - multiplication not closed
  - multiplicative inversion not closed
  - intersection closed

Sectors

- 4 boundary curves: two straight lines $d=0, a=0, c\neq 0$ and two circles $c=0, b=0, d\neq 0$ origin centered
- 4 real parameters needed: b_1/c_1, b_2/c_2, a_3/d_3, a_4/d_4
- equivalent representation: 2 real intervals (radius, angle)
- properties
  - not shift-invariant

- rotation-invariant
- scale-invariant
- addition not closed
- multiplication closed
- multiplicative inversion closed
- intersection not closed

Parallelograms

- 4 boundary curves: 4 straight lines $d = 0, c \neq 0$ with $b_1/c_1 = b_2/c_2$ and $b_3/c_3 = b_4/c_4$
- 6 real parameters needed: a_1, a_2, a_3, a_3, a_4, b_1/c_1, b_3/c_3
- equivalent representation: 3 points (1 corner and 2 edges)
- properties
  - shift-invariant
  - rotation-invariant
  - scale-invariant
  - addition not closed
  - multiplication not closed
  - multiplicative inversion not closed
  - intersection not closed

The operations like addition or membership test are more efficient than for sectors. A generalization to parallelepipeds or zonotopes is used in dynamic systems to avoid the wrapping effect, see [Kühn 1998], [Lohner 1988]

Acknowledgments. This work was supported in part by NASA under cooperative agreement NCC5-209, by NSF grants No. DUE-9750858 and CDA-9522207, by United Space Alliance, grant No. NAS 9-20000 (PWO C0C67713A6), by the 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 grant number F49620-95-1-0518, by the National Security Agency under Grant No. MDA904-98-1-0564, and by Nortel Co.

References

G. Alefeld and J. Herzberger, *Einführung in die Intervallrechnung*, Bibliographisches Institut, Mannheim, Wien, Zürich, 1974.

R. Hammer, M. Hocks, U. Kulisch, D. Ratz, *Numerical toolbox for verified computing. I. Basic numerical problems*, Springer Verlag, Heidelberg, N.Y., 1993.

E. R. Hansen, *Global optimization using interval analysis*, Marcel Dekker, N.Y., 1992.

R. B. Kearfott, *Rigorous global search: continuous problems*, Kluwer, Dordrecht, 1996.

R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

R. Klatte and Ch. Ullrich, "Complex sector arithmetic", *Computing*, 1980, Vol. 24, pp. 139–148.

W. Kühn, " Rigorously Computed Orbits of Dynamical Systems without the Wrapping Effect", Computing 61, 1998, pp. 47–67.

M. Lerch and J. Wolff von Gudenberg, "Multiaspect interval types",in Developments in Reliable Computing, T. Csendes (ed.), Kluwer, p. 365-372, 1999, to appear

R. Lohner, *Einschließung der Lösung gewöhnlicher Anfangs- und Randwertaufgaben und Anwendungen*, PhD thesis, Univ. Karlsruhe, 1988

R. Moore, *Methods and Applications of Interval Analysis*, SIAM, Philadelphia, 1979.

H. T. Nguyen and V. Kreinovich, *Applications of continuous mathematics to computer science*, Kluwer, Dordrecht, 1997.

K. Nickel, "Arithmetic of complex sets", *Computing*, 1980, Vol. 24, pp. 97–105.