Proceedings of the Risa Consortium
1997 pp. 83– pp. 99

Using Interval Computation with the Mahler Measure for Zero Determination of Algebraic Numbers

Hiroshi SEKIGAWA*

NTT Communication Science Laboratories

(RECEIVED 1997/5/30 REVISED 1997/9/10)

Abstract. We propose a new zero determination principle for an algebraic number α obtained after performing ring operations among algebraic numbers $\alpha_1, \ldots, \alpha_n$. We assume that each α_i is represented by its minimal polynomial over $\mathbb{Q}$ and its approximate value as an interval that contains only α_i among its conjugates. The principle of zero determination is as follows: by the estimate of the Mahler measure for α and by the interval for α, we can correctly determine whether α is zero or not with a finite precision value of approximation.

We propose two practical usages of the principle. One method computes both intervals and the Mahler measures simultaneously. The other method utilizes a history of computation to compute the Mahler measures only when they are required. Furthermore, we sharpen inequalities on the Mahler measure after ring operations among algebraic numbers.

1. Introduction

Zero determination is important in computational algebra and computational geometry. Here, we consider zero determination of algebraic numbers, especially algebraic numbers obtained after performing ring operations among given algebraic numbers.

When treating an algebraic number α, it is important how it is represented. There are at least three methods of representing α. The following are the representation methods and the associated methods for zero determination.

* E-mail: sekigawa@cslab.kecl.ntt.co.jp

1. [Representation] In terms of a rational coefficient polynomial f and a fixed algebraic number θ, and representing α as $f(\theta)$ [4][13].
   [Zero determination] Refining (if necessary) the interval containing θ [13].
2. [Representation] Using $P \in \mathbb{Z}[x]$, which is square-free and $P(\alpha) = 0$, and an interval containing α but not containing the other roots of P [4].
   [Zero determination] Refining (if necessary) the interval containing α.
3. [Representation] (When α is real) Thom's code: using $P \in \mathbb{Z}[x]$, which is square-free and $P(\alpha) = 0$, and the signs of the ith derivatives of P at α, for $i = 1, \ldots, \deg(P) - 1$ [8].
   [Zero determination] Using a generalized Sturm algorithm for several inequalities [8].

For polynomial root refinement, see [5] (for real root refinement, see [7][6]). If we use either the second or the third representation method, we have to compute a polynomial that has the resulting algebraic number as a root after every ring operation among algebraic numbers. The computational cost of such a polynomial is expensive if algebraic numbers have high degrees. Therefore, the first method is preferable since we can perform ring operations simply. In general, however, it is also expensive to find such a primitive element as θ.

In this article, we propose a principle for zero determination of algebraic numbers that highly depends on numeric computation. The principle uses intervals and the Mahler measure of algebraic numbers instead of primitive elements or integer coefficient polynomials. Johnson proposed a method using interval arithmetic for real algebraic number computation [10], however, the method resorts to exact arithmetic when an interval contains zero.

The motivation of our research is as follows.

In computational algebra, it is dangerous to use an approximation or numerical approach naively. "Reasonably approximate results" cannot be obtained if we *simply* evaluate an original algorithm on approximate inputs. This is because, even if a sequence converges to a given input, the sequence of the outputs for the initial sequence does not necessarily converge to the true output. We will refer to algorithms that have such instability as *unstable algorithms.*

Shirayanagi proposed a method for stabilizing Buchberger's algorithm [21][22]. The method uses interval computation with "zero rewriting," which is the rule of rewriting an

interval into zero if zero lies within the interval. The underlying ideas of this method were generalized by Shirayanagi and Sweedler as a theory of stabilizing algebraic algorithms [23].

Their approach relates to the convergence of the final output; at any step where zero rewriting is performed, irrespective of whether the rewriting is true, the method passes it through toward the output. The output converges to the true output as the precision increases, and at a finite precision value, a reasonable output is arrived at. Along this line, stopping at some step and asking whether a rewritten interval is truly zero has been the motivation.

In Sections 2 and 4, we explain the principle of zero determination and its practical usages. The key inequalities on the Mahler measure are refined in Section 3. Some examples are shown in Section 5. Finally, we summarize our results and describe a future direction.

2. Principle

If the true value of an algebraic number α that is obtained after ring operations among given algebraic numbers is not zero, we can accurately determine that α is not zero solely by interval computation with a sufficiently high precision. If the true value of α is zero, we cannot determine that α is zero by interval computation with any high precision of μ because the resulting interval I_μ will not be a single point zero. However, the width of I_μ approaches 0 as the precision μ increases. Assuming that we can compute a quantity $\gamma > 0$ such that "if $\alpha \neq 0$, then $|\alpha| \geq \gamma$," then, we can determine whether α is zero by comparing the interval I_μ with the above γ. This is the basis of our theory. We use the Mahler measure of α to compute γ.

The definition of the Mahler measure is as follows [14]:

Definition 1

For a polynomial $P(x) = \sum_{i=0}^{d} a_i x^i = a_d \prod_{i=1}^{d} (x - \alpha_i) \in \mathbb{C}[x]$ $(a_d \neq 0)$, *the Mahler measure* $M(P)$ *of* P *is defined by the formula*

$$M(P) = |a_d| \prod_{i=1}^{d} \max\{1, |\alpha_i|\}.$$

For an algebraic number α, *the Mahler measure* $M(\alpha)$ *of* α *is defined by the formula* $M(\alpha) = M(P)$, *where* P *is the primitive minimal polynomial of* α *over* $\mathbb{Z}$.

The Mahler measure has the following properties, which are the keys for zero determination.

Proposition 2

Let α and β be algebraic numbers of degrees d and e, respectively.

1. *If $M(\alpha) \leq N$ and $\alpha \neq 0$, then $1/N \leq |\alpha| \leq N$.*
2. $M(\alpha \pm \beta) \leq 2^{de} M(\alpha)^e M(\beta)^d, \qquad M(\alpha\beta) \leq M(\alpha)^e M(\beta)^d.$

Proof. See [3] for instance. ∎

For zero determination, we can use norms of a polynomial, for example,

$$\|P\|_1 = \sum_{i=0}^{d} |a_i|, \qquad \|P\|_2 = \left(\sum_{i=0}^{d} |a_i|^2 \right)^{1/2}, \qquad \|P\|_\infty = \max_{0 \leq i \leq d} \{|a_i|\},$$

for a polynomial $P(x) = \sum_{i=0}^{d} a_i x^i$. Among these norms, the inequalities $\|P\|_\infty \leq \|P\|_2 \leq \|P\|_1$ hold, and the following Landau's inequality holds [12].

$$M(P) \leq \|P\|_2 \tag{1}$$

If we use the norm $\|\cdot\|_\infty$ instead of the Mahler measure, then, for two algebraic numbers α and β, an upper bound for $\|\alpha * \beta\|_\infty$ ($* \in \{+, -, \times\}$) as an expression in $\|\alpha\|_\infty$ and $\|\beta\|_\infty$ becomes more complicated. Moreover, these upper bounds grow in size more rapidly after iteratively performing ring operations among algebraic numbers.

To formulate the principle we need to define the approximate interval sequences. For the complex case, rectangular intervals or circular intervals can be used (see [1], for example).

Definition 3

Let α be an algebraic number and let $\{I_\mu\}_{\mu \in \mathbb{N}}$ be a set of bounded closed intervals. We say that $\{I_\mu\}_{\mu \in \mathbb{N}}$ is an approximate interval sequence for α and that the interval I_μ has a precision of μ if $\alpha \in I_\mu$ for any μ, $I_\mu \supset I_\nu$ for $\mu < \nu$, and $I_\mu \to \{\alpha\}$ as $\mu \to \infty$.

Example 1

Let I_μ be an interval whose endpoints are floating-point numbers to which a real algebraic number α is rounded toward $+\infty$ and $-\infty$ with precision μ. Then, $\{I_\mu\}_\mu$ is an approximate interval sequence for α.

We formulate the principle in terms of the following theorem.

Theorem 4

Let $\alpha_1, \ldots, \alpha_n$ be algebraic numbers and let α be the algebraic number obtained after ring operations among α_i's with a specified order. Let $\{I_{i,\mu}\}_\mu$ be an approximate interval

sequence for α_i, and let I_μ be the resulting interval for α obtained by interval arithmetic among $I_{i,\mu}$'s with the specified order of operations. Let N be a real number such that $M(\alpha) \leq N$.

1. *If there is an integer μ such that $0 \notin I_\mu$, then $\alpha \neq 0$. Furthermore, if α is real, we can determine either $\alpha > 0$ or $\alpha < 0$ from the interval I_μ. Conversely, if $\alpha \neq 0$, then there is a finite precision μ_0 such that $0 \notin I_\mu$ holds for any precision $\mu \geq \mu_0$.*
2. *If there is an integer μ such that $0 \in I_\mu$ and $\max\{|c| \mid c \in I_\mu\} < 1/N$, then $\alpha = 0$. Conversely, if $\alpha = 0$, then $0 \in I_\mu$ holds for any precision μ. Moreover, there is a finite precision μ_0 such that $\max\{|c| \mid c \in I_\mu\} < 1/N$ holds for any precision $\mu \geq \mu_0$.*

Proof. Part (1) follows the definitions of interval arithmetic and approximate interval sequences.

To prove part (2), first suppose that $0 \in I_\mu$ and $\max\{|c| \mid c \in I_\mu\} < 1/N$. Since we have $\alpha \in I_\mu$ and $M(\alpha) \leq N$, by applying Proposition 2 (1), α must be 0. The converse statement is clear from the definitions of interval arithmetic and approximate interval sequences. ∎

3. Improvements of Inequalities

In this section, we will describe some improvements in the inequalities from Proposition 2.

First, we will consider the cases not using values of algebraic numbers.

Definition 5

For a polynomial $P(x) = \sum_{i=0}^{d} a_i x^i = a_d \prod_{i=1}^{d}(x - \alpha_i) \in \mathbb{C}[x]$ $(a_d \neq 0)$, the measures $M_i(P)$ $(i = 0, 1)$ of P are defined by the formulae

$$M_0(P) = |a_d|, \qquad M_1(P) = \prod_{i=1}^{d} \max\{1, |\alpha_i|\}.$$

For an algebraic number α, the measures $M_i(\alpha)$ $(i = 0, 1)$ of α are defined by the formulae $M_i(\alpha) = M_i(P)$, where P is the primitive minimal polynomial of α over $\mathbb{Z}$.

The measures M_i's have the following properties:

Proposition 6

Let α and β be algebraic numbers of degrees d and e, respectively.

1. $M_i(\alpha\beta) \leq M_i(\alpha)^e M_i(\beta)^d \qquad (i = 0,\ 1), \qquad M_0(\alpha \pm \beta) \leq M_0(\alpha)^e M_0(\beta)^d.$

2. *Let f be the degree of $\alpha \pm \beta$. Then, $M_1(\alpha \pm \beta) \leq C$, where C is the product of the f largest numbers among the following de numbers:*

$$M_1(\alpha) + M_1(\beta), \underbrace{M_1(\alpha)+1, \ldots, M_1(\alpha)+1}_{e-1}, \underbrace{M_1(\beta)+1, \ldots, M_1(\beta)+1}_{d-1}, \underbrace{2, \ldots, 2}_{(d-1)(e-1)}.$$

Proof. See Appendix A. ∎

Example 2

Let f be the degree of the algebraic number obtained after a ring operation among two algebraic numbers of degrees d and e, respectively. In general, $f = de$, however, in some case $f < de$. For example, let x_i and y_i $(i = 1, 2, 3)$ be real algebraic numbers of degrees d_i and e_i, respectively. Consider the following determinant:

$$\begin{vmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{vmatrix} = (x_2 - x_1)(y_3 - y_1) - (x_3 - x_1)(y_2 - y_1) \tag{2}$$

Put $\alpha = (x_2 - x_1)(y_3 - y_1)$ and $\beta = (x_3 - x_1)(y_2 - y_1)$. Then, the degrees of α and β are at most $d_1 d_2 e_1 e_3$ and at most $d_1 d_3 e_1 e_2$, respectively. The degree of $\alpha - \beta$ is at most $d_1 d_2 d_3 e_1 e_2 e_3$, which is not greater than $d_1 d_2 e_1 e_3 \cdot d_1 d_3 e_1 e_2 = d_1^2 d_2 d_3 e_1^2 e_2 e_3$.

Let $P_i = (x_i, y_i)$, where x_i and y_i are as described above, be three points in $\mathbb{R}^2$. Then, the sign of the above determinant is plus, minus or zero, depending on whether P_1 is to the left of, to the right of or on the directed line P_2 to P_3.

For an algebraic integer α, we have $M(\alpha) = M_1(\alpha)$. Thus we can obtain the next corollary.

Corollary 7

Let α and β be algebraic integers of degrees d and e, respectively. Then,

$$M(\alpha \pm \beta) \leq 2^{(d-1)(e-1)} (M(\alpha) + M(\beta))(M(\alpha)+1)^{e-1} (M(\beta)+1)^{d-1}.$$

Remark 1

The above inequality is a refinement of the standard inequality in Proposition 2. That is, the following inequality holds:

$$2^{(d-1)(e-1)} (M(\alpha) + M(\beta))(M(\alpha)+1)^{e-1} (M(\beta)+1)^{d-1} \leq 2^{de} M(\alpha)^e M(\beta)^d$$

The equality holds if and only if $M(\alpha) = M(\beta) = 1$. Note that for a nonzero algebraic integer γ, Kronecker proved that $M(\gamma) = 1$ holds if and only if γ is a root of unity [11].

If we know values of algebraic numbers, we can sharpen the inequalities in Propositions 6.

Proposition 8

Let α and β be algebraic numbers of degrees d and e, respectively. To simplify any expressions, we suppose that $d \leq e$ and we thus write

$$a = \min\{1, |\alpha|\}, \qquad b = \min\{1, |\beta|\}, \qquad A = \frac{M_1(\alpha)}{\max\{1, |\alpha|\}}, \qquad B = \frac{M_1(\beta)}{\max\{1, |\beta|\}}.$$

1. $M_1(\alpha\beta) \leq \max\{1, aM_1(\beta)\} \cdot \max\{1, bM_1(\alpha)\} \cdot M_1(\alpha)^{e-1} M_1(\beta)^{d-1}$.
2. $M_1(\alpha \pm \beta) \leq C \cdot \max\{1, |\alpha \pm \beta|\}$, *where C is defined as described below:*
   - *When $d = e = 1$, then $C = 1$.*
   - *When $d = 1$ and $e > 1$, then $C = (|\alpha| + B)(|\alpha| + 1)^{\max\{0, f-2\}}$.*
   - *When $d > 1$ and $e > 1$, then C is equal to the product of the $f - 1$ largest numbers among the following $de - 1$ numbers, where f is the degree of $\alpha \pm \beta$.*

$$A + B, A + |\beta|, B + |\alpha|, \underbrace{A + 1, \ldots, A + 1}_{e-2}, \underbrace{B + 1, \ldots, B + 1}_{d-2},$$

$$\underbrace{|\alpha| + 1, \ldots, |\alpha| + 1}_{e-2}, \underbrace{|\beta| + 1, \ldots, |\beta| + 1}_{d-2}, \underbrace{2, \ldots, 2}_{(d-2)(e-2)}.$$

Proof. See Appendix B. ∎

4. Methods

In this section, we discuss practical usages of the principle. For each input algebraic number α, we assume that we know a polynomial $P \in \mathbb{Z}[x]$ that has α as a root and we know an interval that contains only α among the roots of P. It is not necessary that P is minimal, however, it is desirable since small upper bounds of the Mahler measures can be established.

First, we describe how to compute the Mahler measures for given algebraic numbers. For special types of algebraic numbers, we know the exact values of the Mahler measures. For example, let $a = m/n$ be a rational number represented in an irreducible fraction and let α be a pth ($p \geq 1$) root of a. If the degree of α is exactly p, then we have

$$M(\alpha) = \max\{|m|, |n|\}, \qquad M_0(\alpha) = |n|, \qquad M_1(\alpha) = \max\left\{1, \left|\frac{m}{n}\right|\right\}.$$

In general, to compute upper bounds of the Mahler measure, we may use Landau's inequality (Section 2. (1)) or its refinements [15][3][16], without computing the roots of a polynomial.

Next, we explain collaboration between interval computation and the Mahler measure computation. When we find an interval containing zero, we need to compute the Mahler measure corresponding to it. We have proposed two methods to collaborate between them.

4.1. Intervals with the Mahler Measure

The first method computes both intervals and the Mahler measures simultaneously. We have previously proposed this method [19][20].

Let $\alpha_1, \ldots, \alpha_n$ be given algebraic numbers and let α be an algebraic number obtained after performing ring operations among α_i's. We fix the order of ring operations. For each α_i, consider a triplet $(I_{i,\mu}, d_i, A_i)$, where $\{I_{i,\mu}\}_\mu$ is an approximate interval sequence for α_i, d_i is the degree of α_i and $A_i \geq M(\alpha_i)$. We call the triplet $(I_{i,\mu}, d_i, A_i)$ an interval with the Mahler measure for α_i. The arithmetic between them is defined as follows:

$$(I_{i,\mu}, d_i, A_i) * (I_{j,\mu}, d_j, A_j) = (I_{i,\mu} * I_{j,\mu}, e, B), \qquad * \in \{+, -, \times\},$$

where $I_{i,\mu} * I_{j,\mu}$ follows the traditional interval arithmetic, e is an upper bound of the degree of $\alpha_i * \alpha_j$, and $B \geq M(\alpha_i * \alpha_j)$. Then, by performing the arithmetic among these triplets, we can obtain the interval containing α and an upper bound of $M(\alpha)$. If we use the improved inequalities in Section 3, we use the pair $(A_{0,i}, A_{1,i})$ instead of A_i, where $A_{k,i} \geq M_k(\alpha_i)$.

Assume that we have a program, e.g., a program for constructing convex hulls, using exact rational arithmetic. Then, we can construct a new program with the above program as the main routine as follows:

1. We write a module that controls the precision of the interval computations.
2. We rewrite each operation among rational numbers into the corresponding operation among intervals with the Mahler measure.
3. We prepare an additional return value *UNDECIDED* other than usual *TRUE* and *FALSE* for the predicates. We rewrite each predicate of zero determination as follows: If an algebraic number cannot be determined whether it is zero, then *UNDECIDED* is returned.

If *UNDECIDED* is returned then the control module raises the precision and initiates the main routine to compute again for the same input. The new program will stop in a finite number of steps as described in Theorem 4.

4.2. Lazy Method

We can easily make a package of intervals with the Mahler measure. Moreover, it is easy to apply the package to real programs because we can use it without changing the main structure of original programs. However, there are some disadvantages. That is, needless computations, for example, computations of the Mahler measure for an algebraic number that can be determined as nonzero by intervals alone, are carried out.

Therefore, a method that delays computations until they are really needed is desirable. Such a method would resemble the so-called lazy rational arithmetic library (see [2], for instance). We have to store the computation history so that for an algebraic number that cannot be determined as being zero or not by intervals alone, only those computations concerned with the number have to be reiterated. For this purpose, we can use an extension of intervals. Each interval consists of a traditional interval and a symbolic definition. A symbolic definition is either an input algebraic number or an unevaluated expression that represents a ring operation of two other symbolic definitions. We can also create a package of the lazy method and apply it to real programs without changing the main structure of the original programs.

However, it should be noted that in general, the memory required to store such a computation history is extremely large. We can reduce the requirements of some programs by removing any part of computation history that becomes obsolete. However, in such a case, the method is strongly dependent on each program and as a result either the structure of the original program or the package needs to be modified accordingly.

5. Examples

We carried out experiments in the Risa/Asir system [17] on an HP9000/735 computer. For the experiments, in Risa/Asir we implemented big floating-point numbers with base 10 and rounding toward $+\infty$ and $-\infty$ using big integers, and an arbitrary-precision interval arithmetic package using big floating-points. In the following examples, we take the approximate interval sequence for each real algebraic number as described in Example 1.

Table 1. Intervals and estimates for the Mahler measure in Example 3.

algebraic number	interval		upper bound of the Mahler measure
	precision of 10	precision of 11	
α_1	[1.414213562, 1.414213563]	[1.4142135623, 1.4142135624]	2
α_2	[1.732050807, 1.732050808]	[1.7320508075, 1.7320508076]	3
α_3	[2.449489742, 2.449489743]	[2.4494897427, 2.4494897428]	6
β	[2.449489741, 2.449489745]	[2.4494897425, 2.4494897429]	36
γ	$[-0.2 \times 10^{-8}, 0.3 \times 10^{-8}]$	$[-0.3 \times 10^{-9}, 0.2 \times 10^{-9}]$	429981696

5.1. A Simple Example

We present a simple example illustrating how the principle is used for zero determination.

Example 3

Let $\alpha_1 = \sqrt{2}$, $\alpha_2 = \sqrt{3}$ *and* $\alpha_3 = \sqrt{6}$. *Let* $\beta = \alpha_1 \cdot \alpha_2$ *and* $\gamma = \beta - \alpha_3$. *Determine whether* γ *is equal to zero or not.*

The intervals and the estimates for the algebraic numbers are described in Table 1. First, we set the precision to 10. Since the interval for γ contains 0, to apply Theorem 4, we estimate $M(\gamma)$ using the original inequalities. At a precision of 10, however, we cannot determine γ as zero because

$$0.3 \times 10^{-8} > \frac{1}{429981696} = 0.2325\ldots \times 10^{-8}.$$

Next, we set the precision to 11. At this precision we can determine γ as zero because

$$0.3 \times 10^{-9} < \frac{1}{429981696} = 0.2325\ldots \times 10^{-8}.$$

In this example, $M(\gamma)$ is equal to $M(P)$, where $P = x^4(x^2 - 24)^2$. We will show in Table 2 how the improvements in the inequalities can result in estimates for $M(\gamma)$. To compute upper bounds for $M(\gamma)$ using the inequalities in Proposition 8, we use floating-points with a precision of 10 and rounding toward $+\infty$. The final result E is rounded to the integer closest to and no less than E.

5.2. Application to Graham's Algorithm

We applied the principle to construct convex hulls. Hereafter, we assume that S is a finite set of points in $\mathbb{R}^2$ so that the convex hull of S is a convex polygon. Therefore, as

Table 2. Improvements in inequalities.

inequality	estimate for $M(\gamma)$
original	429981696
Proposition 6	4264176
Proposition 8	203926
exact value	576

the convex hull for a given set S of points, we give an ordered list of vertices of the convex hull. Among several well-known algorithms for constructing two-dimensional convex hulls (see [18] for instance), we chose Graham's algorithm in the experiments. A basic routine in Graham's algorithm is to determine the geometric relation of three points as described in Example 2.

Example 4

S is a set of 1000 points. Each point has coordinates of the form $(\sqrt{X}, \sqrt{Y})$, where X and Y are randomly generated integers satisfying the following conditions:

$$0 \leq X \leq 100, \qquad 0 \leq Y \leq 100, \qquad X + Y \leq 100, \qquad Y \leq 3X.$$

The input points and the convex hull (a polygon with 21 vertices) are described in Figure 1.

To compute the determinant described in Example 2, we used the order of operations as described in the right side of the equation (2) in Example 2. We conducted experiments on the following three methods:

1. Interval arithmetic with the Mahler measure using the original inequalities (Proposition 2).
2. Interval arithmetic with the Mahler measure using the improved inequalities (Proposition 8).
3. Lazy method.

In all three of the methods, for computing the values of algebraic numbers, initially we set the precision to 10. In Methods 1 and 2, we iteratively doubled the precision until the convex hull was obtained. In Method 3, we iteratively doubled the precision until the sign of each number was determined and after that, we reset the precision to 10. To estimate

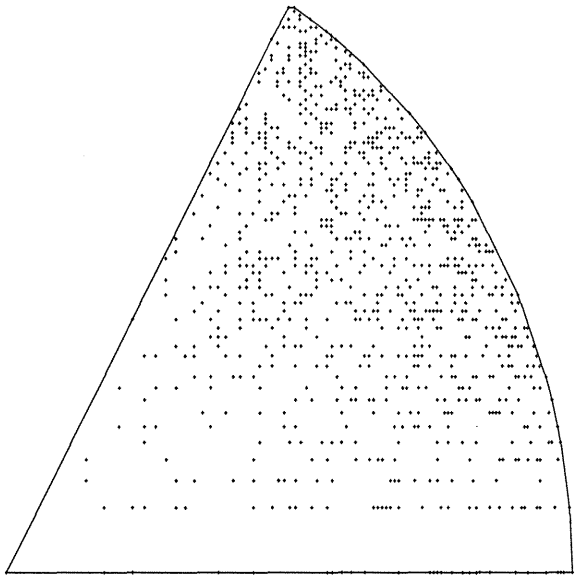

Fig. 1. Input points and the convex hull for Example 4.

Table 3. Computation times and maximal precisions.

	Method 1	Method 2	Method 3
maximal precision	2560	1280	1280
CPU time (sec.)	3.17×10^4	2.71×10^4	304
GC time (sec.)	6.37×10^3	6.90×10^3	157

the Mahler measure, we used big integers in Method 1 and floating-points with a precision of 10 in Methods 2 and 3.

When determining geometric relations among the three points described in Example 2, there can be the decrease of the degrees of the intermediate expressions of algebraic numbers. In Method 3, we took this decrease into account, while we did not in Methods 1 and 2.

The maximal precisions in the computations, CPU times, and garbage collection (GC) times are described in Table 3. Comparing Method 1 with Method 2, the maximal precision and computation time decreased when using Method 2 due to the improvements achieved in inequalities. The computation time could be decreased drastically when using Method 3.

The lazy method is efficient if that part of the computation history that has become obsolete can be removed efficiently. In this example, we assigned indices to input algebraic numbers and changed the arguments from numbers to these indices in each subroutine containing a sign determination so that we achieved efficiency. If we used a package of the lazy method, the memory requirements would become extremely large.

The number of cases where the principle actually contributed to zero determination was nonzero; in Method 3, the number was 389, which was the same as those at the maximal precision in Methods 1 and 2 (at 2560 for Method 1 and at 1280 for Method 2).

6. Conclusion

We have proposed a principle using interval computation and the Mahler measure for zero determination of an algebraic number α obtained after ring operations among algebraic numbers. If α is not zero, we can determine correctly that it is not zero by using interval computation. Otherwise, we can determine correctly that α is zero by using both the Mahler measure and interval computation. We have also proposed two methods using this principle.

One future direction of research is to estimate the precision needed to determine zero correctly before computation. Hiyoshi's work [9] is the first attempt in this direction.

Acknowledgements

The author would like to thank Nobuki Takayama and Kiyoshi Shirayanagi for their valuable comments.

A. Proof of Proposition 6

We prove only part (2) since proofs for part (1) are similar to those for the Mahler measure. It is enough to show the case of addition. Let $\alpha_1 = \alpha$, α_2, ..., α_d and $\beta_1 = \beta$, β_2, ..., β_e be all the conjugates of α and β over $\mathbb{Q}$. Let S be the set of pairs (i, j) such that the set $\{\alpha_i + \beta_j \mid (i, j) \in S\}$ is equal to the set of the all conjugates of $\alpha + \beta$ over $\mathbb{Q}$. Then,

$$M_1(\alpha + \beta) = \prod_{(i,j)\in S} \max\{1, |\alpha_i + \beta_j|\} \leq \prod_{(i,j)\in S} (\max\{1, |\alpha_i|\} + \max\{1, |\beta_j|\}),$$

and since $\#S = f$, part (2) follows the following lemma. ∎

Lemma 9

Let S be a set of s pairs (i,j) $(1 \le i \le m,\ 1 \le j \le n)$ and let $F(x_1,\ldots,x_m,y_1,\ldots,y_n)$ be $\prod_{(i,j)\in S}(x_i+y_j)$. For A, $B \ge 1$, the function F has the maximum on the domain

$$D = \left\{ (x_1,\ldots,x_m,y_1,\ldots,y_n) \in \mathbb{R}^{m+n} \;\middle|\; x_i \ge 1,\ \ y_j \ge 1,\ \prod_{i=1}^{m} x_i = A,\ \prod_{j=1}^{n} y_j = B \right\}.$$

The maximum is not greater than the product of the s largest numbers among the following mn numbers:

$$A+B, \underbrace{A+1,\ldots,A+1}_{n-1}, \underbrace{1+B,\ldots,1+B}_{m-1}, \underbrace{2,\ldots,2}_{(m-1)(n-1)}.$$

Proof. Since the domain D is compact, the continuous function F takes the maximum on D. To estimate the maximum, first we consider the case where $m = 1$. Let $(a_1,b_1,\ldots,b_n) \in D$ be a point where F takes the maximum. We will prove that all b_j's except one are equal to 1. Assume that there are two distinct numbers p and q such that $(1,p)$, $(1,q) \in S$ and b_p, $b_q > 1$. Then, the point $(a_1,b'_1,\ldots,b'_n)$, where $b'_p = b_pb_q$, $b'_q = 1$ and $b'_j = b_j$ if $j \ne p$, q, belongs to D. However, the following inequality shows the contradiction.

$$F(a_1,b'_1,\ldots,b'_n) - F(a_1,b_1,\ldots,b_n) = a_1(b_p-1)(b_q-1) \prod_{\substack{(1,j)\in S \\ j\ne p,q}} (a_1+b_j) > 0$$

Next, we consider the general cases. Let $N(i)$ be $\#\{(i,j) \in S\}$. For simplicity, we may assume that $\{i \mid N(i) > 0\} = \{1,\ldots,p\}$ and $N(1) \ge N(2) \ge \cdots \ge N(p)$. From the case $m = 1$,

$$\prod_{\substack{(i,j)\in S \\ i=k}} (x_i+y_j) \le (x_k+B)(x_k+1)^{N(k)-1}, \qquad k = 1,\ldots,p,$$

hold on the domain D. Therefore,

$$F(x_1,\ldots,x_m,y_1,\ldots,y_n) \le \prod_{i=1}^{p}(x_i+B)(x_i+1)^{N(i)-1} = \prod_{i=1}^{p}(x_i+B) \cdot \prod_{i=1}^{p}(x_i+1)^{N(i)-1}.$$

From similar arguments for when $m = 1$, we obtain

$$\prod_{i=1}^{p}(x_i+B) \cdot \prod_{i=1}^{p}(x_i+1)^{N(i)-1} \le (A+B)(1+B)^{p-1} \cdot (A+1)^{N(1)-1} \cdot 2^c,$$

where $c = \sum_{i=2}^{p}(N(i)-1)$. The statement follows the equality $1+(p-1)+(N(1)-1)+c = s$, and the three inequalities $p-1 \le m-1$, $N(1)-1 \le n-1$ and $c \le (m-1)(n-1)$. ∎

B. Proof of Proposition 8

Let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_d$ and $\beta_1 = \beta, \beta_2, \ldots, \beta_e$ be all the conjugates of α and β over $\mathbb{Q}$.

First, we will prove the case of multiplication. When $|\alpha|, |\beta| \geq 1$, a proof is not required. To prove the other cases, first we will show that the following inequality holds when $|\alpha| < 1$.

$$\prod_{j=1}^{e} \max\{1, |\alpha\beta_j|\} \leq \max\{1, |\alpha| M_1(\beta)\}$$

If $|\alpha\beta_j| \leq 1$ holds for any j then the statement is clear. Otherwise, there is a number k such that $|\beta_k| > |\alpha\beta_k| > 1$. Therefore,

$$\prod_{j=1}^{e} \max\{1, |\alpha\beta_j|\} = |\alpha\beta_k| \prod_{\substack{1 \leq j \leq e \\ j \neq k}} \max\{1, |\alpha\beta_j|\} \leq |\alpha| \prod_{j=1}^{e} \max\{1, |\beta_j|\} = |\alpha| M_1(\beta).$$

Suppose that $|\alpha| < 1$ and $|\beta| \geq 1$ (the case where $|\alpha| \geq 1$ and $|\beta| < 1$ is similar). Then,

$$M_1(\alpha\beta) \leq \prod_{j=1}^{e} \max\{1, |\alpha_1\beta_j|\} \cdot \prod_{i=2}^{d} \prod_{j=1}^{e} \max\{1, |\alpha_i\beta_j|\} \leq \max\{1, aM_1(\beta)\} \cdot M_1(\alpha)^e M_1(\beta)^{d-1}.$$

When $|\alpha| < 1$ and $|\beta| < 1$, by a similar argument to the above case, we have

$$M_1(\alpha\beta) \leq \max\{1, aM_1(\beta)\} \cdot \max\{1, bM_1(\alpha)\} \cdot M_1(\alpha)^{e-1} M_1(\beta)^{d-1}.$$

Next, we will prove the case of addition (the case of subtraction is similar). When $d = e = 1$, the statement is clear. When $d = 1$ and $e > 1$, the statement follows Lemma 9.

For the case $d, e \geq 2$, let S be the set of pairs (i, j) such that the set $\{\alpha_i + \beta_j \mid (i, j) \in S\}$ is equal to the set of the all conjugates of $\alpha + \beta$ over $\mathbb{Q}$. Then we have

$$M_1(\alpha + \beta) \leq \max\{1, |\alpha + \beta|\} \cdot \prod_{\substack{(1,j) \in S \\ j \geq 2}} (|\alpha| + \max\{1, |\beta_j|\} \cdot \prod_{\substack{(i,1) \in S \\ i \geq 2}} (\max\{1, |\alpha_i|\} + |\beta|)$$
$$\times \prod_{\substack{(i,j) \in S \\ i \geq 2 \\ j \geq 2}} (\max\{1, |\alpha_i|\} + \max\{1, |\beta_j|\}).$$

We write $N_1 = \#\{j \mid (1, j) \in S\}$, $N_2 = \#\{i \mid (i, 1) \in S\}$, $N = \#\{(i, j) \in S \mid i \geq 2, \ j \geq 2\}$, $c_{11} = \min\{1, N_1 - 1\}$, $c_{12} = \max\{0, N_1 - 2\}$, $c_{21} = \min\{1, N_2 - 1\}$ and $c_{22} = \max\{0, N_2 - 2\}$. Then, by similar arguments as to that in the proof of Lemma 9, we can derive

$$\prod_{\substack{(1,j) \in S \\ j \geq 2}} (|\alpha| + \max\{1, |\beta_j|\}) \leq (|\alpha| + B)^{c_{11}} (|\alpha| + 1)^{c_{12}},$$

$$\prod_{\substack{(i,1)\in S \\ i\geq 2}} (\max\{1,|\alpha_i|\}+|\beta|) \leq (A+|\beta|)^{c_{21}}(1+|\beta|)^{c_{22}},$$

$$\prod_{\substack{(i,j)\in S \\ i\geq 2 \\ j\geq 2}} (\max\{1,|\alpha_i|\}+\max\{1,|\beta_j|\}) \leq C',$$

where C' is the product of the N largest numbers among the following $(d-1)(e-1)$ numbers:

$$A+B, \underbrace{A+1,\ldots,A+1}_{e-2}, \underbrace{1+B,\ldots,1+B}_{d-2}, \underbrace{2,\ldots,2}_{(d-2)(e-2)}.$$

The statement follows the equality $c_{11}+c_{12}+c_{21}+c_{22}+N=f-1$, and the four inequalities $c_{11}\leq 1$, $c_{12}\leq e-2$, $c_{21}\leq 1$ and $c_{22}\leq d-2$. ∎

References

[1] Alefeld, G. and Herzberger, J.: *Introduction to Interval Computations*, Academic Press, 1983.

[2] Benouamer, M., Michelucci, D. and Peroche, B.: Error-free boundary evaluation using lazy rational arithmetic: a detailed implementation, *Proc. 2nd Symposium on Solid Modeling and Applications*, 1993, 115–126.

[3] Cerlienco, L., Mignotte, M. and Piras, F.: Computing the measure of a polynomial, *J. Symb. Comput.*, **4**, 1987, 21–33.

[4] Cohen, H.: *A Course in Computational Algebraic Number Theory*, Springer-Verlag, 1993.

[5] Collins, G. E. and Krandick, W.: An efficient algorithm for infallible polynomial complex root isolation, *Proc. ISSAC'92*, 1992, 189–194.

[6] Collins, G. E. and Krandick, W.: A hybrid method for high precision calculation of polynomial real roots, *Proc. ISSAC'93*, 1993, 47–52.

[7] Collins, G. E. and Loos, R.: Real zeros of polynomials, *Computer Algebra Symbolic and Algebraic Computation* (Buchberger, B., Collins, G. E. and Loos, R., eds.), Springer-Verlag, 1983, 83–94.

[8] Coste, M. and Roy, M.-F.: Thom's lemma, the coding of real algebraic numbers and the computation of the topology of semi-algebraic sets, *J. Symb. Comput.*, **5**, 1988, 121–129.

[9] Hiyoshi, H.: Applications of approximate computation to computational algebra and computational geometry, Master thesis, Univ. Tokyo (in Japanese), 1997.

[10] Johnson, J. R.: Real algebraic number computation using interval arithmetic, *Proc. ISSAC'92*, 1992, 195–205.

[11] Kronecker, L.: Zwei Sätze über Gleichungen mit ganzzahligen Coefficienten, *J. für und angew. Math.*, **53**, 1857, 173–175.

[12] Landau, E.: Sur quelques théorèmes de M. Petrovic relatifs aux zéros des fonctions analytiques, *Bull. Soc. Math. France*, **33**, 1905, 251–261.

[13] Loos, R.: Computing in algebraic extensions, *Computer Algebra Symbolic and Algebraic Computation* (Buchberger, B., Collins, G. E. and Loos, R., eds.), Springer-Verlag, 1983, 173–187.

[14] Mahler, K.: An application of Jensen's formulae to polynomials, *Mathematica*, **7**, 1960, 98–100.

[15] Mignotte, M.: An inequality about factors of polynomials, *Math. Comp.*, **28**, 1974, 1153–1157.

[16] Mignotte, M. and Glesser, P.: Landau's inequality via Hadamard's, *J. Symb. Comput.*, **18**, 1994, 379–383.

[17] Noro, M. and Takeshima, T.: Risa/Asir—A computer algebra system, *Proc. ISSAC'92*, 1992, 387–396.

[18] Preparata, F. P. and Shamos, M. I.: *Computational Geometry*, Springer-Verlag, 1985.

[19] Sekigawa, H.: An interval arithmetic with algebraic complexity to determine the signs of algebraic expressions, *Abstracts of MEGA'96*, 1996, 43.

[20] Sekigawa, H.: Using interval arithmetic and polynomial norms to determine signs of algebraic numbers, *Proc. ASCM'96*, 1996, 43–53.

[21] Shirayanagi, K.: An algorithm to compute floating point Gröbner bases, *Mathematical Computation with Maple V: Ideas and Applications* (Lee, T. ed.), Birkhäuser, 1993, 95–106.

[22] Shirayanagi, K.: Floating point Gröbner bases, *Mathematics and Computers in Simulation*, **42**, 1996, 509–528.

[23] Shirayanagi, K. and Sweedler, M.: A theory of stabilizing algebraic algorithms, *Technical Report 95-28, Mathematical Sciences Institute, Cornell University*, 1995.