Numer. Math. 36, 53–62 (1980)

Numerische Mathematik

An Implementation of Vincent's Theorem

Alkiviadis G. Akritas

University of Kansas, Department of Computer Science, Lawrence, Kansas 66045, USA

Summary. A new method is presented for the isolation of the real roots of a polynomial equation; it is based on Vincent's forgotten theorem of 1836 and has been implemented using exact (infinite precision) integer arithmetic algorithms. A theoretical analysis of the computing time of this method is given along with some empirical results.

Subject Classifications: AMS(MOS): 26A78, 68A15; CR: 5.7.

1. Introduction

Isolation of the real roots of a polynomial equation (or, simply, isolation) is the process of finding real, disjoint intervals such that each contains exactly one real root and every real root is contained in some interval. This process, in itself, is of great interest and use in several areas; as an example we mention Tarski's decision method [16].

In this paper we shall restrict our attention to those isolation methods which have been implemented in computer algebra systems using exact (infinite precision) integer arithmetic algorithms. (For a survey of computer algebra systems see [14].) When we use exact integer arithmetic then, in analyzing an algorithm, the "cost" of an operation on two integers depends not only on the operation itself but on the length (number of bits) of the operants as well. If A is an integer, we define $L(A)$, its length, by

$$L(A) = \begin{cases} 1 & A = 0 \\ \lfloor \log_b |A| \rfloor + 1 & A \neq 0 \end{cases}$$

where b indicates the base of the number system in which the operant A is represented when an operation is being performed. If P is a polynomial with integer coefficients, $|P|_\infty$ represents the maximum coefficient in absolute value.

0029-599X/80/0036/0053/$02.00

The isolation problem was first solved by Sturm (1829), whose main proposition depends on a theorem by Fourier [4]. Sturm's method was shown to be [10]

$$O(n^{13} L(|P|_\infty)^3).$$

In 1975 the method by differentiation, which is based on Rolle's theorem, was proven to be [9]

$$O(n^{10} + n^7 L(|P|_\infty)^3).$$

That same year, in Uspensky's *Theory of Equations* [17], the author of this paper discovered Vincent's forgotten theorem of 1836 [1, 6, 18], which is based on Budan's proposition (1807). (In almost all texts, Fourier's theorem is stated under the name Budan; the statement of Budan's theorem can be found in [18, 4].)

Without careful study of Vincent's theorem, and completely modifying it, Collins and Akritas developed a new method which is an improvement over the previous two; it was shown to be [8]

$$O(n^6 L(|P|_\infty)^2).$$

Subsequently, Vincent's theorem was carefully studied in the Ph.D. thesis of the author [1]; it was proven that, without any modifications, this theorem can be also used in order to solve the isolation problem. Actually, two methods result, Vincent's and ours, corresponding to the two different ways of computing certain quantities a_i. It has been shown [1, 2], that Vincent's method behaves exponentially, whereas, our method has the polynomial computing time bound

$$O(n^5 L(|P|_\infty)^3);$$

this is the best bound achieved thus far using exact integer arithmetic. Moreover, it should be noted that our method is the only one with polynomial computing time bound which isolates the real roots using continued fractions. In what follows we will study this new method in detail.

Historical Remark. In the articles by Collins and Akritas [8] and Collins ([15] pp 35-68) Vincent's exponential method has been erroneously attributed to Uspensky; this was probably due to Uspensky's claim (in the preface of his book [17]) that he himself invented the method. However, as was pointed out ([1] pp 85-86), what can be considered a contribution on Uspensky's part is only the fact that he used the Ruffini-Horner method in order to perform transformations of the form $x = a_i + y$; Vincent on the contrary used Taylor's expansion theorem.

2. Vincent's Theorem

In this section we present the required mathematical background.

Theorem 2.1. *Let $P(x) = 0$ be a polynomial equation of degree $n > 1$, with rational coefficients and without multiple roots, and let $\Delta > 0$ be the smallest distance*

between any two of its roots. Let m be the smallest index such that

$$F_{m-1}\frac{\Delta}{2}>1 \quad \text{and} \quad F_{m-1}F_m\Delta>1+\frac{1}{\varepsilon_n} \tag{2.1}$$

where F_k is the kth *member of the Fibonacci sequence* $1,1,2,3,5,8,13,\ldots$ *and*

$$\varepsilon_n=\left(1+\frac{1}{n}\right)^{\frac{1}{n-1}}-1.$$

Then the transformation

$$x=a_1+\cfrac{1}{a_2+\ddots+\cfrac{1}{a_m+\cfrac{1}{y}}}, \tag{2.2}$$

(which is equivalent to the series of successive transformations of the form $x=a_i+\frac{1}{y}, i=1,2,\ldots,m$*) with arbitrary, positive, integral elements* $a_1,a_2,\ldots,a_m$*, transforms the equation* $P(x)=0$ *into the equation* $\tilde{P}(y)=0$*, which has not more than one sign variation in the sequence of its coefficients.*

The proof of the above theorem is very long, and it is omitted since it can be found in the literature [3, 17 pp 298–304]. The original form of Theorem 2.1 (that is, without specifying the quantity m) is due to Vincent alone [18, 6, 1] and appeared in 1836; Uspensky [17] extended it in a somewhat erroneous manner, which was corrected in [3].

Definition 2.1. Two real-valued functions f and g, defined on D, are said to be *codominant*, $f\sim g$, in case there exist two positive real numbers c_1, c_2 and $\hat{x}$ in D such that $|f(x)|\leqq c_1|g(x)|$ and $|f(x)|\geqq c_2|g(x)|$ both hold for all x in D, $x>\hat{x}$.

Corollary 2.1. *Under the assumptions of Theorem* 2.1 *we have*

$$m=O(nL(|P|_\infty)+nL(n)).$$

Proof. By definition m is the smallest index such that inequalities (2.1) hold simultaneously. Clearly, one of these inequalities (and possibly both) will not be satisfied if we reduce m by one; suppose that the first one fails, so that

$$F_{m-2}\frac{\Delta}{2}\leqq 1. \tag{2.3}$$

Applying the relation $F_k=\phi^k/\sqrt{5}$, where $\phi=1.618\ldots$, inequality (2.3) yields

$$\phi^{m-2}\leqq 2\sqrt{5}\left(\frac{1}{\Delta}\right)$$

from which we deduce that

$$m\leqq 2+\log_\phi 2+\tfrac{1}{2}\log_\phi 5-\log_\phi \Delta. \tag{2.4}$$

However, from [12] we have

$$\Delta \geqq 3n^{-\frac{(n+2)}{2}} |P|_1^{-(n-1)}, \tag{2.5}$$

where $|P|_1$ is the sum of the absolute values of the coefficients of P. The corollary is now proven if we combine (2.4) and (2.5), taking also into consideration the fact that $L(|P|_1) \sim L(|P|_\infty)$. (The same result is obtained if we assume that the second of the inequalities (2.1) fails.) //

Remark 2.1. In most cases of interest $L(n) \leqq L(|P|_\infty)$ so we can safely conclude that

$$m = O(nL(|P|_\infty)). \tag{2.6}$$

Theorem 2.1 can be used in order to isolate the real roots of a polynomial equation; from its statement we know that a transformation of the form (2.2), with arbitrary, positive integer elements $a_1, a_2, \ldots, a_m$ transforms $P(x)=0$ into an equation $\tilde{P}(y)=0$, which has at most one sign variation. This transformation can be also written as

$$x = \frac{P_m y + P_{m-1}}{Q_m y + Q_{m-1}}, \tag{2.7}$$

where P_k/Q_k is the k-th convergent to the continued fraction

$$a_1 + \cfrac{1}{a_2 + \ddots},$$

and as we recall

$$P_{k+1} = a_{k+1} P_k + P_{k-1}$$
$$Q_{k+1} = a_{k+1} Q_k + Q_{k-1}.$$

Since the elements $a_1, a_2, \ldots, a_m$ are arbitrary, there is obviously an infinite number of transformations of the form (2.2). However, with the help of Budan's theorem [4] we can easily determine those that are of interest to us; namely, there is a finite number of them (equal to the number of the positive roots of $P(x)=0$) which lead to an equation with *exactly* one sign variation in the sequence of its coefficients. Suppose that $\tilde{P}(y)=0$ is one of those equations; then from the Cardano-Descartes rule of signs we know that it has one root in the interval $(0, \infty)$. If $\hat{y}$ was this positive root, then the corresponding root $\hat{x}$ of $P(x)=0$ could be easily obtained from (2.7). We only know though that $\hat{y}$ lies in the interval $(0, \infty)$; therefore, substituting y in (2.7) once by 0 and once by ∞ we obtain for the positive root $\hat{x}$ its isolating interval whose unordered endpoints are P_{m-1}/Q_{m-1} and P_m/Q_m. In this fashion we can isolate all the positive roots of $P(x)=0$. If we subsequently replace x by $-x$ in the original equation, the negative roots become positive and hence, they too can be isolated in the way mentioned above.

The calculation of the quantities $a_1, a_2, \ldots, a_m$ - for the transformations of the form (2.2) which lead to an equation with exactly one sign variation -

constitutes the polynomial real root isolation procedure. Two methods actually result, Vincent's and ours, corresponding to the two different ways in which the computation of the a_i's may be performed.

Vincent's method basically consists of computing a particular a_i by a series of unit incrementations; that is, $a_i \leftarrow a_i + 1$, which corresponds to the substitution $x \leftarrow x+1$. This brute force approach results in a method which will behave exponentially when the values of the a_i's are big. Examples of this approach can be found in [18] and in [17].

Our method, on the contrary, consists of immediately computing a particular a_i as the lower bound b on the values of the positive roots of a polynomial; that is, $a_i \leftarrow b$, which corresponds to the substitution $x \leftarrow x+b$ performed on the particular polynomial under consideration. It is obvious that our method is independent of how big the values of the a_i's are. An unsuccessful treatment of the big values of the a_i's can be found in ([17] p 136).

This interpretation of each a_i as a lower root bound is made clear if we consider that our objective is to force one of the positive roots inside the interval $(0, 1)$ and all others inside the interval $(1, \infty)$ or vice versa. The following lemmas are relevant.

Lemma 2.1. *Let $P(x)=0$ be an univariate polynomial equation of degree $n \geqq 2$, with integer coefficients and without multiple roots, which has p real roots in the interval $(0,1)$, $2 \leqq p \leqq n$, and let $\Delta_p > 0$ be the smallest distance between any two of them. Then the inversion $x \leftarrow 1/x$, performed on $P(x)=0$ maps these p roots in the interval $(1, \infty)$, where now the smallest distance between any two of them is $\Delta'_p > \Delta_p$.*

Proof. Let $0 < \alpha_1 < \alpha_2 < \ldots < \alpha_i < \alpha_j < \alpha_k < \alpha_l < \ldots < \alpha_p < 1$ be the p roots of $P(x)=0$ in the interval $(0,1)$ and suppose that

$$\Delta_p = \alpha_j - \alpha_i, \qquad \text{whereas } \Delta'_p = \frac{1}{\alpha_k} - \frac{1}{\alpha_l}.$$

The lemma now follows immediately since

$$\Delta'_p = \frac{\alpha_l - \alpha_k}{\alpha_k \alpha_l} > \alpha_l - \alpha_k \geqq \alpha_j - \alpha_i = \Delta_p. \quad //$$

Lemma 2.2. *Let $P(x)=0$ be an univariate polynomial equation of degree $n \geqq 2$, with integer coefficients and without multiple roots, which has two complex conjugate roots α_1 and α_2 inside the circle with center $(1/2, 0)$ and radius $1/2$, and let $\delta_p = |\alpha_1 - \alpha_2|$. Then the inversion $x \leftarrow 1/x$, performed on $P(x)=0$, maps a_1 and α_2 in the half-plane with real part > 1, where now their distance is $\delta'_p > \delta_p$.*

Proof. Similar to the previous one. //

The lower bound b on the values of the positive roots is computed with Cauchy's rule ([13] pp 50–51); observe that we are computing the upper bound on the values of the positive roots of $P(1/x)=0$.

Cauchy's Rule. Let $P(x)=x^n+c_{n-1}x^{n-1}+\ldots+c_1x+c_0=0$ be a polynomial equation of degree n with integer coefficients and let λ be the number of its negative coefficients. Then

$$b=\max_{\substack{1\leqq k\leqq n\\ c_{n-k}<0}}|\lambda c_{n-k}|^{1/k} \tag{2.8}$$

is an upper bound on the values of the positive roots of $P(x)=0$.

We have implemented Cauchy's rule using exact integer arithmetic and we have shown that its computing time bound is [5]

$$O(n^2L(|P|_\infty)). \tag{2.9}$$

For the next theorem we need the following: Consider an infinite binary tree in which the root corresponds to an original polynomial equation $P(x)=0$ and each node corresponds to a transformed equation resulting from the original after a series of successive transformations of the form $x=a_i+\frac{1}{y}$. The path from each node to the right descendent corresponds to the substitution $x\leftarrow 1+x$, whereas, the path to the left descendent corresponds to the substitution $x\leftarrow\frac{1}{1+x}$. All the nodes belonging to a specific path, finite or infinite, will be considered as members of disjoint sets which can be of three types. A set of type V_0, V_1 or V_n contains nodes corresponding to polynomials with zero, one or more than one sign variations respectively. Sets of type V_0 or V_1 are called *terminal sets.* In the case of sets belonging to the same path, a set X is said to *precede* a set Y if and only if for all x in X and all y in Y pathlength$(x)<$pathlength(y). In a terminal set, the node having the shortest path from the root of the tree will be called a *terminal node.* With these definitions in mind, the difference between Vincent's method and ours is seen in Fig. 1.

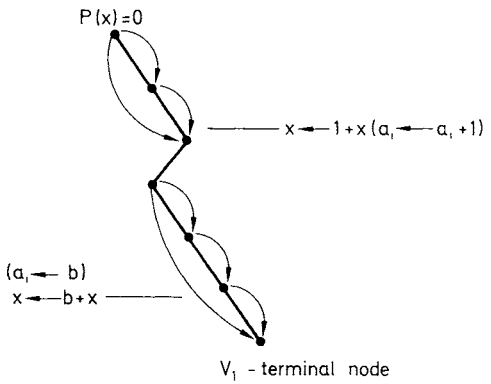

Fig. 1. Geometric interpretation of the two different ways of computing the value of a particular a_i (the length of a branch)

Moreover, in [7] we proved that if P is a polynomial of degree n the transformation $x \leftarrow b+x$, $b \geqq 1$, can be performed in time

$$O(n^3 L(b)^2 + n^2 L(b) L(|P|_\infty)). \tag{2.10}$$

Theorem 2.2. *Let $P(x)=c_n x^n + \ldots + c_1 x + c_0 = 0$ be a polynomial equation of degree $n>1$, with rational coefficients and without any multiple roots, which corresponds to the root of the binary tree. Suppose, moreover, that the transformation*

$$x = a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_l + \cfrac{1}{y}}}$$

with positive, integer elements $a_1, a_2, \ldots, a_l$, $1 \leqq l \leqq m$ (where m is defined in Theorem 2.1) transforms $P(x)=0$ into a new equation corresponding to a V_0 or V_1-terminal node. Then for all k, $1 \leqq k \leqq l$, we have

$$a_k = O(|P|_\infty). \tag{2.11}$$

Proof. For $k=1$ consider the following:

(i) an upper bound on the values of the roots of $P(x)=0$ is given by

$$b = 2 \max_{1 \leqq k \leqq n} \left| \frac{c_{n-k}}{c_n} \right|^{1/k} \tag{2.12}$$

(see for example [11] p 398) and this $b \leqq 2|P|_\infty$;

(ii) provided that $l>1$, a_1 is equal to the integer part of some root of $P(x)$ with positive real part.

From (i) and (ii) it follows easily that $a_1 = O(|P|_\infty)$. In order to prove the general case set $P_0(x)=P(x)=0$ and assume that $\tilde{P}_{i+1}(x)=0$ and $P_i(x)=0$ are two equations such that $\tilde{P}_{i+1}(x)=P_i(a_i+x)$. Subsequently apply the previous considerations to equation $P_i(x)=0$, taking also into account the fact that $|P_i|_\infty \sim |\tilde{P}_{i+1}|_\infty$ and $|\tilde{P}_{i+1}|_\infty = |P_{i+1}|_\infty$, where $P_{i+1}(x)=\tilde{P}_{i+1}(1/x)=0$. //

3. Our Method

In the previous section we described how a real root is isolated, once we have obtained the corresponding polynomial with one sign variation in the sequence of its coefficients. In what follows we give a recursive description of the way in which such a polynomial is obtained.

Recursive Description. Let

$$P(x)=0 \tag{3.1}$$

be a polynomial equation with v sign variations in the sequence of its integer coefficients and without multiple roots.

Base. $v=0$ or $v=1$. From the Cardano-Descartes rule of signs we know that $v=0$ implies that (3.1) has no positive roots, whereas, $v=1$ indicates that (3.1) has exactly one positive root, in which case $(0, \infty)$ is its isolating interval; in either case, no transformation of (3.1) is necessary, and the method terminates.

Recursion. $v>1$. In this case (3.1) has to be further investigated. We first compute the lower bound b on the values of the positive roots and then we obtain the translated equation $P_b(x)=P(b+x)=0$, which also has v sign variations provided $P(b)\neq 0$ (if $P(b)=0$, we have found an integer root of the original equation and v is decreased). $P_b(x)=0$ is now transformed by the substitutions $x \leftarrow 1+x$ and $x \leftarrow \frac{1}{1+x}$, and the procedure is applied again twice, once with $P_b\left(\frac{1}{1+x}\right)=0$ in place of (3.1) and once with $P_b(1+x)=0$.

Theorem 3.1. *Let $P(x)=0$ be an univariate polynomial equation of degree $n>0$, with integer coefficients and without multiple roots. If the method described above is applied twice (once for the positive and once for the negative roots), the real roots of $P(x)=0$ will be isolated in time*

$$O(n^5 L(|P|_\infty)^3).$$

Proof. First let us see what happens for the positive roots. Let

$$x=a_1+\cfrac{1}{a_2+\ddots+\cfrac{1}{a_m+\cfrac{1}{y}}} \tag{3.2}$$

be a transformation which transforms the given equation into another one corresponding to a V_0 or V_1-terminal node. In this proof we assume that the set of type V_0 is preceded by a set of type V_n, which implies that we tried to isolate complex conjugate roots; this assumption is quite justified since there is no reason for us to arrive at a set of type V_0 which is preceded by a set of type V_1. Moreover, we assume that $b=\lfloor R \cdot p \cdot (\alpha_s) \rfloor$, where α_s is the root with the smallest positive real part.

In the worst possible case a transformation of the form (3.2) will be performed for each root with positive real part. An inversion $x \leftarrow 1/x$ is performed in time $\sim n$ by simply inverting the order of the coefficients; therefore, we will concern ourselves only with the translations of the form $x \leftarrow a_k+x$. Given (2.11) and (2.10) each substitution $x \leftarrow a_k+x$ will be performed in time

$$O(n^3 L(|P|_\infty)^2). \tag{3.3}$$

However, there will be at most m such substitutions, where m is bounded by (2.6). Therefore, the time needed for each transformation of the form (3.2) is

$$O(n^4 L(|P|_\infty)^3).$$

The theorem now follows from the fact that there are exactly n roots. //

4. Empirical Results and Conclusions

In this section we present two tables showing the observed computing times for the methods of Sturm, Vincent and ours. We compare only these methods because we think that only these can be considered classical; we have shown that they are based on two very old and related theorems [4]. All times are in seconds and were obtained by using the SAC-1 computer algebra system on the IBM S/370 Model 165 computer located at the Triangle Universities Computation Center, where a subroutine is available which reads the computer clock. The polynomials used in the second table are the same as those used in [8] and [9], so that the reader can easily verify that our method is faster by a simple comparison of the ratios of the times of the various methods to the corresponding times obtained with Sturm's method.

Table 1. Polynomials of degree 5 with randomly generated roots

Roots are in the interval	Vincent	Our method
$(0, 10^2)$	0.45	0.16
$(0, 10^3)$	1.61	0.71
$(0, 10^4)$	16.43	2.01
$(0, 10^5)$	175.62	4.81

Table 1 indicates the exponential nature of Vincent's method. Each of the polynomials was formed by taking the product of 5 linear terms.

Table 2. Polynomials with randomly generated coefficients

Degree	Sturm	Our method
5	2.05	0.26
10	33.28	0.46
15	156.40	0.94
20	524.42	2.36

All the coefficients of the polynomials in Table 2 were nonzero, each 10 decimal digits long and randomly generated.

References

1. Akritas, A.G.: Vincent's theorem in algebraic manipulation, Ph.D. thesis. Operations Research Program, North Carolina State University, Raleigh 1978
2. Akritas, A.G.: A new method for polynomial real root isolation. Proc. of the 16th annual southeast regional ACM conference, Atlanta, Georgia, 39-43 (1978). (This paper received the First Prize in the student paper competition)

3. Akritas, A.G.: A correction on a theorem by Uspensky. Bulletin of the Greek Mathematical Society **19**, 278–285 (1978)
4. Akritas, A.G.: The two different ways of expressing the Budan-Fourier theorem and their consequences. In: 5th volume of lectures given at the General Mathematical Seminar of the University of Patras, 127–146 (1979) (in Greek)
5. Akritas, A.G.: Exact algorithms for the implementation of Cauchy's rule (submitted, 1980)
6. Akritas, A.G., Danielopoulos, S.D.: On the forgotten theorem of Mr. Vincent. Historia Math. **5**, 427–435 (1978)
7. Akritas, A.G., Danielopoulos, S.D.: On the complexity of algorithms for the translation of polynomials. Computing **24**, 51–60 (1980)
8. Collins, G.E., Akritas, A.G.: Polynomial real root isolation using Descartes' rule of signs. Proc. of the 1976 ACM symposium on symbolic and algebraic computation, Yorktown Heights, New York, 272–275 (1976)
9. Collins, G.E., Loos, R.: Polynomial real root isolation by differentiation. Proc. of the 1976 ACM symposium on symbolic and algebraic computation, Yorktown Heights, New York, 15–20 (1976)
10. Heindel, L.E.: Integer arithmetic algorithms for polynomial real zero determination. J. Assoc. Comput. Mach. **18**, 533–548 (1971)
11. Knuth, D.E.: The art of computer programming, Vol. II: Seminumerical Algorithms, Reading: Addison-Wesley 1969
12. Mahler, K.: An inequality for the discriminant of a polynomial. Michigan Mathm. J. **11**, 257–262 (1964)
13. Obreschkoff, N.: Verteilung und Berechnung der Nullstellen reeller Polynome, Berlin: VEB Deutscher Verlag der Wissenschaften 1963
14. S.R. Petricle (ed.) Proceedings of the 2nd Symposium on Symbolic and Algebraic Manipulation. ACM 1971
15. Rice, J.: Mathematical software III. New York and London: Academic Press 1977
16. Tarski, A.: A decision method for elementary algebra and geometry, Berkeley: University of California Press 1951
17. Uspensky, J.V.: Theory of equations, New York: McGraw-Hill 1948
18. Vincent, A.J.H.: Sur la résolution des équations numériques. J. Math. Pures Appl. **1**, 341–372 (1836)

Received December 4, 1979/July 25, 1980