SCIENCES ET TECHNIQUES

EN

PERSPECTIVE

ANNEE 1983–1984

VOLUME IV

Université de Nantes

Equipe de recherche : Sciences, Techniques et Sociétés.

MARS 1984

SCIENCES ET TECHNIQUES EN PERSPECTIVE

Volume IV

Table des matières

A.G. AKRITAS (Lawrence, Kansas)	*Budan's theorem, p. 1-13*
C. GILAIN (Paris)	*La théorie des équations différentielles au début du XIXème siècle, p. 13-20.*
P. BROUZENG (Bordeaux)	*L'oeuvre scientifique de Duhem et l'histoire des sciences, p. 21-26.*
J.L. GARDIES (Nantes)	*Les antécédents scolastiques de la théorie des ensembles, p. 27-50.*
P. BAILHACHE (Nantes)	*Le principe des travaux virtuels vers 1800 , p. 51-64*
N. BRILLOUET (Nantes)	*Le parallélogramme des forces. Son étude fonctionnelle au XVIIIème siècle, p. 65-79.*
A. CHAPERT (Montpellier)	*Malus et l'optique, p. 80-95.*
B. WOJTKOWIAK (Nantes)	*A la découverte de Lavoisier. Son rôle dans la révolution chimique du XVIII ème siècle, p. 95-118.*
J. DHOMBRES (Nantes)	*La méthode fonctionnelle en mathématique histoire conceptuelle (1ère partie) p. 129-171*

Prix 35,00 Frs - Disponible auprès de l'I.R.E.M., Université de Nantes, 2, Chemin de la Houssinière 44072 NANTES CEDEX. Chèque à libeller au nom de Monsieur l'Agent Comptable de l'Université de Nantes.

Budan's Theorem and Its Consequences

By
Alkiviadis G. Akritas
University of Kansas
Department of Computer Science
Lawrence, Kansas 66045

Abstract

The importance of Budan's theorem is underestimated in the existing literature. In this paper it is indicated that it forms the basis of Vincent's forgotten theorem, which in turn leads to a new method for the isolation of the real roots of a polynomial equation. It turns out that if we use exact integer arithmetic algorithms this new method is the fastest existing.

1. Introduction

In the theory of equations it is well known that it was not until the Renaissance, when the Italian mathematicians of the 15th and 16th century Tartaglia, Cardano and Ferrari succeeded in solving by radicals the general equations of the third and fourth degree, that any work of lasting interest was done. In the 17th and 18th century numerous attempts were made in order to solve the general quintic equation but they all failed; as Lagrange put it, this remained "a challenge to the human intellect". Ruffini (1804) was the first to demonstrate the impossibility of solving the quintic equation by a formula composed from the coefficients of the equation by using a finite number of times the operations of addition, multiplication, division and root extraction; he was later followed by Abel who proved (1826) that it is impossible to solve in this way - in general - algebraic equations of degree greater than four.

In the beginning of the 19th century the attention of the mathematicians had already been focused on numerical methods for the solution of general polynomial equations. During this period Fourier conceived the idea to proceed in two steps; that is, first to isolate the roots and then to approximate them to any desired degree of accuracy. In this paper we shall be concerned only with real roots.

Isolation of the real roots of a polynomial equation is the process of finding real intervals such that each contains exactly one real root and every real root is contained in some interval. In order to accomplish this, F.D. Budan and J.B.J. Fourier presented two <u>different</u> - but equivalent - theorems, which enable one to determine the maximum possible number of real roots of an equation within a given interval. Both propositions make use of sign variations which are defined as follows:

<u>Definition 1</u>. We say that a sign variation exists between two numbers c_p and c_q $(p<q)$ of a finite or infinite sequence of real numbers $c_1, c_2, c_3, \ldots$, if c_p and c_q are not zero and have opposite signs, and in case $q \geq p + 2$ (that is c_q does not immediately follow c_p) the numbers $c_{p+1}, \ldots, c_{q-1}$ are all zero.

(Subsequently we will say that a polynomial "has" or "presents" v sign variations and will omit the phrase "in the sequence of its coefficients".)

Fourier's theorem was included in his <u>Analyse des Équations</u> published posthumously by C.L.M.H. Navier [4]. It makes use of the Fourier sequence

(F) $$P(x),\ P^{(1)}(x),\ldots,P^{(m)}(x)$$

where $P^{(i)}(x)$ is the i-th derivative of $P(x)$ with respect to x, and states that the number of real roots of the equation $P(x) = 0$ located in the interval (p,q) can never be more than the number of sign variations lost in passing from the substitution $x=p$ (in Fourier's sequence) to the substitution $x=q$.

Based on Fourier's proposition, Sturm (1829) presented an improved theorem, the application of which yields the exact number of real roots of a square-free polynomial equation within a real interval. Thus the real root isolation problem was solved [7]. Since 1830 Sturm's method is the only one widely known and used. Quite recently it was implemented in a computer algebra system using exact integer arithmetic and its computing time was thoroughly analyzed. It was shown [10] that if $P(x) = 0$ is a univariate, polynomial equation of degree $n > 0$, with integer coefficients and without any multiple roots, then Sturm's method is

$$0(n^{13}L(|P|_\infty)^3),$$

where $L(|P|_\infty)$ is the length, in bits, of the maximum coefficient in absolute value.

Budan's theorem, on the other hand, appeared in 1807 and its statement seems to be ignored by most of the existing literature [8],[11],[12],[13],[14]; moreover, due to the fact that it is equivalent to Fourier's theorem, the latter is frequently called Budan-Fourier or Fourier-Budan, or even Budan!! [9],[16]. Recent work by the author of this article, however, showed that Budan's theorem by itself merits special attention because it constitutes the basis of Vincent's forgotten theorem of 1836 [4]. The latter in turn is the foundation of two methods for the isolation of the real roots of a polynomial equation [2]; the first method is due to Vincent [15] (1836), whereas the second has only recently been developed by the author, [1],[2]. We have been able to show [1], [2] that for the polynomial equation $P(x) = 0$ mentioned above, our new method is

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

This is the best bound obtained using exact integer arithmetic.

In what follows we first indicate how Budan's theorem forms the basis of Vincent's forgotten theorem and then we discuss the applications of the latter. Tables are also presented comparing the methods by Sturm and Vincent with ours.

2. Budan's Theorem and Vincent's Forgotten Theorem

As we mentioned in the introduction, recent studies showed that Budan's theorem is of extreme importance; its statement, however, is ignored by most of the existing literature [8], [9], [11],[12],[13],[14],[16]. To our knowledge, Vincent's paper of 1836 [15] is the only reference where one can find Budan's theorem in its original form. It is rendered as follows:

Theorem 1 (Budan 1807). If in an equation in x, $P(x) = 0$, we make two transformations, $x = p + x'$ and $x = q + x''$, where p and q are real numbers such that $p < q$, then

(i) The transformed equation in $x' = x-p$ cannot have fewer sign variations than the transformed equation in $x'' = x-q$,

(ii) the number of real roots of the equation $P(x) = 0$, located between p and q can never be more than the number of sign variations lost in passing from the transformed equation in $x' = x-p$ to the transformed equation in $x'' = x-q$,

(iii) when the first number is less than the second, their difference is always an even number.

A proof of this theorem can be found in the literature. We easily see that Theorem 1 gives us an upper bound on the number of the real roots, which the equation $P(x) = 0$ has inside the interval (p,q). Moreover, it should be noted that only transformations of the form $x = r + x'$ are used. The equivalence of the theorems by Budan and Fourier is easily seen with the help of Taylor's expansion: If we replace x by α in Fourier's sequence (F), then the resulting $m + 1$ numbers are proportional to the corresponding coefficients of the equation $P(x + \alpha) = 0$. Up to now the importance of Budan's theorem was not made clear; however, as we pointed out, it constitutes the basis of the following:

Theorem 2 (Vincent 1836 [15]). If, in a polynomial equation with rational coefficients and without multiple roots, one makes successive transformations of the form

$$x = a_1 + \frac{1}{x'},\ x' = a_2 + \frac{1}{x''},\ x'' = a_3 + \frac{1}{x'''},\ldots\ldots$$

where each $a_1, a_2, a_3, \ldots$ is any positive integer, then the resulting, transformed equation has either zero or one sign variation. In the latter, the equation has a single positive root represented by the continued fraction

$$a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}$$

whereas, in the former case there is no root.

The proof of this theorem can be found in Vincent's original paper and in [3]. Theorem 2 was hinted by Fourier and Vincent indicates his surprise that the former did not try to go further and prove the proposition that was the main subject of Vincent's article. He states however the belief that such a proof may exist in other manuscripts which were not published because of the untimely death of Navier. Vincent's theorem was so totally forgotten that it is not mentioned even in such a capital work as the Enzyclopaedie der mathematischen Wissenschaften. As far as we have been able to determine it is not mentioned by any author with the exception of Uspensky [14] and Obreschkoff [12]. The author of this paper discovered Vincent's theorem while reviewing methods for the isolation of the real roots of equations as presented by Uspensky, [4].

The dependence of Vincent's theorem on the theorem by Budan is seen if each of the transformations of the form $x = a_i + \frac{1}{y}$ (which make up the continued fraction) is expressed by the equivalent pair of transformations $x = a_i + y'$, $y' = \frac{1}{y}$. Obviously, Theorem 2 does not provide a bound on the number of transformations of the form $x = a_i + \frac{1}{y}$, which have to be performed in order to obtain the polynomial with at most one sign variation in the sequence of its coefficients. Uspensky extended Vincent's theorem in order to give an answer to the above question, ([14], pp. 298-304). His treatment though contains certain errors, in the statement and proof, which were corrected in [3]. The following is a corrected version of the extended theorem (see also [5]):

Theorem 3 ([3]). 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

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

where F_k is the $k^{\underline{th}}$ member of the Fibonacci sequence

$$1,1,2,3,5,8,13,21,\ldots$$

and

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

Then the transformation

$$x = a_1 + \cfrac{1}{a_2 + \cfrac{}{\ddots + \cfrac{1}{a_m + \cfrac{1}{\xi}}}} \tag{3}$$

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

3. Applications of Vincent's Extended Theorem

Theorem 3 can be used in order to isolate the real roots of a polynomial equation. The fact that it holds only for equations without multiple roots does not restrict the generality, because in the opposite case all we have to do is to express $P(x)$ in the form $P = \prod_{i=1}^{e} S_i^i$, where each of the S_i's has only single roots ([14] pp.65-69). Each of these single roots is of multiplicity i for the polynomial $P(x)$ and thus we see that Theorem 3 can be applied on the S_i's. So, in the rest of this discussion it is assumed that $P(x) = 0$ is without multiple roots; moreover, we assume that its coefficients are integers.

From the statement of Theorem 3 we know that a transformation of the form (3), with arbitrary, positive integer elements $a_1, a_2, \ldots, a_m$ transforms $P(x) = 0$ into an equation $\tilde{P}(\xi) = 0$, which has at most one sign variation; this transformation can be also written as

$$x = \frac{P_m\xi + P_{m-1}}{Q_m\xi + Q_{m-1}}, \tag{4}$$

where $\frac{P_k}{Q_k}$ is the k^{th} convergent to the continued fraction $a_1 + \frac{1}{a_2} + \ldots$. Since the elements $a_1, a_2, \ldots, a_m$ are arbitrary there is obviously an infinite number of transformations of the form (3). However, with the help of Budan's theorem 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 $\tilde{P}(\xi) = 0$ is one of these equations; then from the Cardano-Descartes rule of signs we know that it has one root in the interval $(0,\infty)$. If $\hat{\xi}$ was this positive root, then the corresponding root $\hat{x}$ of $P(x) = 0$ could be easily obtained from (4). We only know though that $\hat{\xi}$ lies in the interval $(0,\infty)$; therefore, substituting ξ in (4) once by 0 and once by ∞ we obtain for the positive root $\hat{x}$ its isolating interval, whose unordered endpoints are $\frac{P_{m-1}}{Q_{m-1}}$ and $\frac{P_m}{Q_m}$. 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. Thus we see that we have a procedure for isolating all the real roots of $P(x) = 0$.

The calculation of the quantities $a_1, a_2, \ldots, a_m$ - for the transformations of the form (3) 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; the difference between these two methods can be thought of as being analogous to the difference between the integrals of Riemann and Lebesgue (think of the addition).

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 with an exponential behavior; that is, for big values of the a_i's this method will take a long time (even years in a computer) in order to isolate the real roots of an equation. Therefore, Vincent's method is of little practical importance. Examples of this approach can be found in Vincent's paper [15], and in Uspensky's book ([14] pp. 129-137). The reader should notice that in the preface of his book, Uspensky claims that he himself invented this method. A simple comparison with Vincent's paper though makes clear that what can be considered a contribution on Uspensky's part is only the fact that he used the Ruffini-Horner method [6] in order to perform the transformations $x \leftarrow x + 1$, whereas Vincent used Taylor's expansion theorem. Moreover, Uspensky seems to ignore Budan's theorem and, while computing a particular a_i he performs, after each transormation $x \leftarrow x + 1$ the unnecessary transformation $x \leftarrow \frac{1}{x+1}$, something which Vincent avoids.

Our method, on the contrary, is an aesthetically pleasing interpretation of Theorem 3. Basically it consists of immediately computing a particular the 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. (The reader should notice that the absolute lower bound on the roots does not work. Why?) 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 Uspensky ([14] p 136). In this discussion it is assumed that $b = \lfloor \alpha_s \rfloor$ the greatest integer $\le \alpha_s$, where α_s is the smallest positive roots.) Since the substitutions $x \leftarrow x + 1$ and $x \leftarrow x + b$ can be performed in about the same time [6], we easily see that our method results in enormous savings of computing time. We have implemented our method in a computer algebra system using exact integer arithmetic and have been able to show [1], [2] that its computing time bound is

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

Our method is the only one with polynomial computing time bound which isolates the real roots of a polynomial equation using continued fractions, and it turns out that this is the best approach.

In this paper we present two tables showing the observed computing times in seconds. These times were obtained by using the SAC-1 computer algebra system on the IBM S/370 Model 165 computer located at the Triangle Universites Computation Center (North Carolina), where a subroutine CCLOCK is available, which reads the computer clock. Table 1 compares Sturm's method with ours, for randomly generated polynomials of degree 5 through 20. Table 2 compares Vincent's method with ours for polynomials of degree 5, whose roots have been randomly generated. (All the coefficients were integer.)

Degree	Sturm	Our Method
5	2.05	.26
10	33.28	.46
15	156.40	.94
20	524.42	2.36

Table 1

Polynomials with randomly generated coefficients

Interval Containing Roots	Vincent	Our Method
$(0,10^2)$	.45	.16
$(0,10^3)$	1.16	.71
$(0,10^4)$	16.43	2.01
$(0,10^5)$	175.62	4.81

Table 2

Polynomials of degree 5 with randomly generated roots

We believe that these tables along with the preceeding discussion adequately stress the importance of Budan's ignored theorem. It is our hope that in the future, this important theorem will not be confused with the one by Fourier, and that it will find its place in the history of mathematics.

References

1. Akritas, A.G., "Vincent's Theorem in Algebraic Manipulation", Ph.D. Thesis, Operations Research Program, North Carolina State University, Raleigh, N.C., 1978.

2. Akritas, A.G., "A New Method for Polynomial Real Root Isolation", Proceedings of the 16th Annual Southeast Regional ACM Conference, Atlanta, Georgia, April 1978, 39-43 (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, Vol. 19, 278-285, 1978.

4. Akritas, A.G. and S.D. Danielopoulos "On the Forgotten Theorem of Mr. Vincent", Historia Mathematica, Vol.5, 427-435, 1978.

5. Akritas, A.G. and S.D. Danielopoulos, "On a Special Case of the Cardano-Descartes Rule of Signs", submitted for publication.

6. Akritas, A.G. and S.D, Danielopoulos, "On the Complexity of Algorithms for the Translation of Polynomials", Computing, Vol. 24, p. 51-60, 1980.

7. Bôcher, M., "The Published and Unpublished Work of Charles Sturm on Algebraic and Differential Equations", Bulletin of the American Mathematical Society, Vol. 18, 1-18, 1911.

8. Burnside, W.S. and A.W. Panton, "The Theory of Equations", Dover, New York, Second Edition, 1960, Vol. 1.

9. Dickson, L.E. "First Course in the Theory of Equations", J. Wiley & Sons, New York, 1922.

10. Heindel, L.E. ,"Integer Arithmetic Algorithms for Polynomial Real Zero Determination", Journal of the ACM, Vol. 18, 1971, 533-548.

11. Hurwitz. A., "Ueber den Satz von Budan-Fourier", Mathematische Annalen, Vol. 71, 1912, 584-591.

12. Obreschkoff, N., "Verteilung und Berechnung der Nullstellen reeller Polynome", VEB Deutscher Verlag der Wissenschaften, Berlin, 1963.

13. Turnbull, H.W., "Theory of Equations", Oliver and Boyd, Edinburgh and London, Fifth Edition, 1957.

14. Uspensky, J.V., "Theory of Equations", McGraw-Hill, New York, 1948.

15. Vincent, A.J.H., "Sur la Résolution des Équations Numériques", Journal de Mathématiques Pures et Appliquées, Vol. 1, 1836, 341-372.

16. Weisner, L. "Introduction to the Theory of Equations", The Macmillan Co., New York, 1938.