Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Journal of Computational and Applied Mathematics 200 (2007) 226–230

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Short Communication

On multiple roots in Descartes' Rule and their distance to roots of higher derivatives

Arno Eigenwillig*

Max-Planck-Institut für Informatik, 66123 Saarbrücken, Germany

Received 11 August 2005; received in revised form 19 December 2005

Abstract

If an open interval I contains a k-fold root α of a real polynomial f, then, after transforming I to $(0, \infty)$, Descartes' Rule of Signs counts exactly k roots of f in I, provided I is such that Descartes' Rule counts no roots of the kth derivative of f. We give a simple proof using the Bernstein basis.

The above condition on I holds if its width does not exceed the minimum distance σ from α to any complex root of the kth derivative. We relate σ to the minimum distance s from α to any other complex root of f using Szegő's composition theorem. For integer polynomials, $\log(1/\sigma)$ obeys the same asymptotic worst-case bound as $\log(1/s)$.

MSC: Primary: 26C10; 30C15; 65H05

Keywords: Descartes' Rule of Signs; Descartes–Jacobi Rule; Bernstein basis; Root isolation; Root separation

1. Introduction

Let $f(x) = \sum_{i=0}^{n} f_i x^i$ be a polynomial of degree n with real coefficients. *Descartes' Rule of Signs* states that the number $v = \mathrm{var}(f_0, \ldots, f_n)$ of sign variations in the coefficient sequence of f exceeds the number p of positive real roots, counted with multiplicities, by an even non-negative integer. See [7, Theorem 2] for a proof with careful historic references. Jacobi [5, IV] made the "little observation" that this statement on the interval $(0, \infty)$ can be extended to any open interval (l, r) by first composing f with the Möbius transform $T(x) = (lx + r)/(x + 1)$ that takes $(0, \infty)$ to (l, r) and then inspecting the coefficients of $g(x) = (x + 1)^n f(T(x))$. We call this the *Descartes test* for the number of roots of f in (l, r). As this test counts roots with multiplicities, it cannot distinguish, say, two simple roots from one double root. However, if multiplicities are known in advance, the Descartes test remains useful even in the presence of multiple roots.

Consider the *Bernstein basis* $B_0^n, B_1^n, \ldots, B_n^n$ defined by

$$B_i^n(x) = B_i^n[l, r](x) = \binom{n}{i} \frac{(x - l)^i (r - x)^{n-i}}{(r - l)^n}. \tag{1}$$

* Tel.: +49 681 9325 129; fax: +49 681 9325 199.

E-mail address: arno@mpi-inf.mpg.de.

0377-0427/$ - see front matter © 2006 Published by Elsevier B.V.

doi:10.1016/j.cam.2005.12.016

The coefficients of $f(x) = \sum_{i=0}^{n} b_i B_i^n(x)$ in the Bernstein basis and the coefficients of $g(x) = \sum_{i=0}^{n} g_i x^i$ in the monomial basis agree up to order and positive constants; that is, $g_i = \binom{n}{i} b_{n-i}$ for $0 \leqslant i \leqslant n$. Hence the Descartes test for f on (l, r) can equivalently be formulated to count the number of sign variations in the sequence $(b_0, \ldots, b_n)$ of Bernstein coefficients. This relation was already known to Pólya and Schoenberg [11, p. 322; 14, Section 1]. It was first applied in a root isolation algorithm by Lane and Riesenfeld [8].

Descartes' Rule of Signs has an immediate geometric interpretation in terms of Bézier curves [4,12]. The graph of f is a Bézier curve with control points $(\mathbf{b}_i)_{i=0}^{n}$ where $\mathbf{b}_i = (i/n, b_i)$. The Descartes test counts how many times the control polygon $\overline{\mathbf{b}_0\mathbf{b}_1} \cup \overline{\mathbf{b}_1\mathbf{b}_2} \cup \cdots \cup \overline{\mathbf{b}_{n-1}\mathbf{b}_n}$ crosses the x-axis. Repeated de Casteljau subdivision is a corner-cutting process on the control polygon, which, in the limit, converges to the graph of f. The number of intersections with the x-axis can never grow and only drop by an even number during corner-cutting.

Let the open interval I contain a simple root α of f. If the Descartes test counts $v = 1$ in I, this implies that α is the unique root of f in I. However, the converse implication does not hold in general. The use of the Descartes test in algorithms for isolating the real roots of square-free polynomials [2,6,13,9] has motivated research on conditions sufficient for the Descartes test to count $v = 1$. A particularly general sufficient condition was given by Ostrowski [10] but has been overlooked until recently [7].

We remark that not every test for roots in an interval I has this property of yielding the exact count if the width of I is small enough. The closely related Budan–Fourier test for roots in $(l, r]$ computes $v' := \operatorname{var}(f(l), f'(l), \ldots, f^{(n)}(l)) - \operatorname{var}(f(r), f'(r), \ldots, f^{(n)}(r))$, which is also known to exceed the number of roots in $(l, r]$, counted with multiplicities, by a non-negative even integer [1, Theorem 2.36]. But consider the example $f(x) = x^3 + x$: Substitution into the sequence $(f(x), \ldots, f'''(x))$ yields the sign patterns $(-, +, -, +)$ for $x < 0$ and $(+, +, +, +)$ for $x > 0$. Hence $v' = 3$ for any interval $(l, r]$ containing the simple root 0 in its interior.

Our results: Let the open interval I contain a k-fold root α of f whose multiplicity $k \geqslant 1$ is known. We present sufficient conditions for the Descartes test to count $v = k$ sign variations for f in I and thus to certify uniqueness of the root α in I. Using the Bernstein basis, we can prove very easily in Section 2 that the Descartes test counts $v = k$ for f whenever it counts $v^{(k)} = 0$ for the kth derivative $f^{(k)}$. This condition is met if the width of I does not exceed the minimum distance σ between α and any root of $f^{(k)}$. In Section 3, we relate σ to the minimum distance s between α and any other complex root of f. To do so, we use Szegő's composition theorem in a way that generalizes an approach of Dimitrov [3] from the first to higher derivatives of f, and we obtain a lower bound on the distance of α to roots of $f^{(r)}$ for any $r \leqslant k$. For integer polynomials with τ-bit coefficients, the resulting bound on $\log(1/\sigma)$ has the same worst-case asymptotics as $\log(1/s)$, namely $O(n\tau + n \log n)$.

2. A partial converse via differentiation

From now on, we assume w.l.o.g. $[l, r] = [0, 1]$. Differentiation of the ith Bernstein basis polynomial $B_i^n(x) = \binom{n}{i} x^i (1-x)^{n-i}$ yields $nB_{i-1}^{n-1}(x) - nB_i^{n-1}(x)$, where $B_{-1}^{n-1} = B_n^{n-1} = 0$ by convention. Hence the derivative of $f(x) = \sum_{i=0}^{n} b_i B_i^n(x)$ is $f'(x) = \sum_{i=0}^{n} n(b_{i+1} - b_i) B_i^{n-1}(x)$. The coefficient vector $(c_i)_{i=0}^{n-1}$ of $1/n \cdot f'(x)$ is therefore given by the following difference scheme.

$$\begin{array}{ccccc} b_0 & b_1 & b_2 \cdots b_{n-1} & & b_n \\ c_0 = -b_0 + b_1 & c_1 = -b_1 + b_2 & \cdots & c_{n-1} = -b_{n-1} + b_n & \end{array} \tag{2}$$

The following lemma can be regarded as a piecewise linear analogue of Rolle's Theorem.

Lemma 1. *The numbers of sign variations in* (2) *satisfy* $\operatorname{var}(c_0, \ldots, c_{n-1}) \geqslant \operatorname{var}(b_0, \ldots, b_n) - 1$.

Proof. Each sign variation in $(b_0, \ldots, b_n)$ is an index pair $0 \leqslant i < j \leqslant n$ such that $b_i b_j < 0$ and $b_{i+1} = \cdots = b_{j-1} = 0$. Let there be exactly v such pairs $(i_1, j_1), \ldots, (i_v, j_v)$ with indices $i_1 < j_1 \leqslant i_2 < j_2 \leqslant \cdots \leqslant i_v < j_v$. Sign variations are either "positive to negative" ($b_i > 0$) or "negative to positive" ($b_i < 0$). Obviously, these types alternate. If $b_{i_\ell} > 0$, then $b_{i_\ell+1} \leqslant 0$ and thus $c_{i_\ell} = -b_{i_\ell} + b_{i_\ell+1} < 0$. Similarly, if $b_{i_\ell} < 0$ then $c_{i_\ell} > 0$. Hence the sequence $(c_0, \ldots, c_{n-1})$ contains an alternating subsequence $\operatorname{sgn}(c_{i_1}) \neq \operatorname{sgn}(c_{i_2}) \neq \cdots \neq \operatorname{sgn}(c_{i_v})$, demonstrating that $(c_0, \ldots, c_{n-1})$ has at least $v - 1$ sign variations. □

Theorem 2. *Consider a polynomial f and an interval I. Let the Descartes test in I count v sign variations for f and* $v^{(k)}$ *sign variations for* $f^{(k)}$. *Then* $v \leqslant v^{(k)} + k$.

Proof. Apply Lemma 1 repeatedly. □

Theorem 3. *Let f have a root α of multiplicity k in the open interval I. Consider the roots of* $f^{(k)}$ *in the complex plane. If none of them is situated inside the circle C with diameter I, then the Descartes test in I counts* $v^{(k)} = 0$ *sign variations for* $f^{(k)}$ *and* $v = k$ *sign variations for f.*

Proof. The first claim is the well-known One-Circle Theorem [7, Theorem 22] applied to $f^{(k)}$. Regarding the second claim, we have $v \geqslant k$ by Descartes' Rule and $v \leqslant k$ by Theorem 2. □

Corollary 4. *With notation as above, let σ denote the minimum distance between α and any root of* $f^{(k)}$. *Clearly,* $\sigma > 0$. *If the width of I is σ or less, then* $v = k$.

Proof. Apply Theorem 3: any two points inside C have distance less than σ, so no root of $f^{(k)}$ is among them. □

Remark 5. If α is the unique root of f in I, the possible values of v are $k, k+2, k+4, \ldots$, hence $v \leqslant k+1$ suffices to guarantee $v = k$. In this case, we can therefore relax the condition $v^{(k)} = 0$ to $v^{(k)} \leqslant 1$ or replace it by $v^{(k+1)} = 0$.

For the case $k=1$, how does Theorem 3 compare to the Two-Circle Theorem [7, Theorem 21] arising from Ostrowski's condition [10]? For polynomials with only real roots, the Two-Circle Theorem reduces to the optimal statement, namely that the Descartes test yields 1 if and only if there is exactly one real root in I, whereas Theorem 3 fails in the presence of roots of f'. However, there are also examples in which the Two-Circle Theorem fails but Theorem 3 works, e. g., $f(x) = 400\,000\,000\,000\,x^3 - 1\,199\,600\,000\,000\,x^2 + 1\,323\,455\,609\,000\,x - 523\,731\,353\,391$.

3. The distance between complex roots of f and $f^{(r)}$

The results of the preceding section motivate this question about complex polynomials and their complex roots: Given a k-fold root α of f and a number $r \leqslant k$, how close can a root $\xi \neq \alpha$ of $f^{(r)}$ be? The relative position of roots of f and $f^{(r)}$ is invariant under translations, so we assume w.l.o.g. that $\alpha = 0$. In this situation,

$$f(x) = \sum_{i=k}^{n} f_i x^i = x^k g(x) \quad \text{with } g(x) = \sum_{i=0}^{n-k} g_i x^i,$$

$$f^{(r)}(x) = \sum_{i=k-r}^{n-r} \frac{(i+r)!}{i!} f_{i+r} x^i = \frac{n!}{(n-r)!} x^{k-r} h(x) \quad \text{with } h(x) = \sum_{i=0}^{n-k} h_i x^i.$$

The polynomials g and h have exactly those roots of f and $f^{(r)}$, resp., different from α. The coefficients of h satisfy

$$h_i = \frac{(n-r)!}{n!} \frac{(i+k)!}{(i+k-r)!} f_{i+k} = \left(\prod_{d=0}^{r-1} \frac{k+i-d}{n-d} \right) g_i. \tag{3}$$

Following Dimitrov [3], we use Szegő's composition theorem to track how multiplying the coefficients in (3) changes the roots of g into those of h, necessitating this definition: A *closed circular region* in the plane is the closure of the interior of a circle, the closure of the exterior of a circle, or a closed half-plane (its boundary line being a circle of infinite radius).

Theorem 6 (*Szegő's composition theorem [15, Satz 2; 3, Theorem 3]*). *Let* $A(x) = \sum_{i=0}^{n} a_i \binom{n}{i} x^i$, $B(x) = \sum_{i=0}^{n} b_i \binom{n}{i} x^i$, *and* $C(x) = \sum_{i=0}^{n} a_i b_i \binom{n}{i} x^i$. *Let K be a closed circular region in the complex plane containing all roots of A. If ξ is a root of C, there is* $w \in K$ *and a root β of B such that* $\xi = -w\beta$.

The polynomial $C(x)$ is called the *composition* of $A(x)$ and $B(x)$. According to (3), $h(x)$ is the composition of $g(x)$ and

$$T_r(x) = \sum_{i=0}^{n-k} \left(\prod_{d=0}^{r-1} \frac{k+i-d}{n-d} \right) \binom{n-k}{i} x^i. \tag{4}$$

It remains to find a suitable circular region K enclosing the roots of T_r.

Lemma 7. *All roots of T_r are real and contained in the interval* $[-1, -T_r(0)]$.

Proof. We start with a minor generalization of Lemma 1 of Dimitrov [3]: the polynomial

$$U_d(x) := \sum_{i=0}^{n-k} \frac{k+i-d}{n-d} \binom{n-k}{i} x^i = \left(x + \frac{k-d}{n-d} \right) (x+1)^{n-k-1} \tag{5}$$

has the roots $-(k-d)/(n-d)$ and -1; both are elements of $[-1, 0]$. To verify (5), observe that the coefficient of x^i on the right is

$$\frac{k-d}{n-d} \binom{n-k-1}{i} + \binom{n-k-1}{i-1} = \frac{(k-d)(n-k-i)+(n-d)i}{n-d} \frac{(n-k-1)!}{i!(n-k-i)!} = \frac{k+i-d}{n-d} \binom{n-k}{i}.$$

We need a second auxiliary result, namely a remark by Schur [15, p. 37] on Theorem 6: If the set K is a half-plane containing 0, and if all roots β of B are elements of $[-1, 0]$, then K contains not only w but also $\xi = (-\beta) \cdot w$. By intersecting families of such half-planes, it follows that any convex set containing 0 as well as all roots of A also contains all roots of C.

T_r is the result of composing U_0 with $U_1, U_2, \ldots, U_{r-1}$. By Schur's remark, the enclosure of all roots of U_0 in the convex set $[-1, 0]$ is preserved by these compositions. Note that $T_{r,0} := T_r(0) = \prod_{d=0}^{r-1} (k-d)/(n-d) > 0$ is the product of all roots of T_r, up to sign. As none of them has magnitude larger than one, $T_{r,0}$ is a lower bound on the magnitude of each of them, proving the claim. □

Theorem 8. *Let α be a k-fold root of f, and let $s > 0$ be the minimum distance of α to any other root β of f. Let $r \leqslant k$. If $\xi \neq \alpha$ is a root of $f^{(r)}$, then $|\alpha - \xi| \geqslant sT_{r,0}$ where $T_{r,0} = \prod_{d=0}^{r-1} (k-d)/(n-d)$.*

Proof. W.l.o.g., $\alpha = 0$. Let K be the closed disc with diameter $[-1, -T_{r,0}]$. It contains all roots of T_r, so Theorem 6 implies for the composition $h(x)$ of $T_r(x)$ and $g(x)$ that its root ξ has the form $\xi = -w\beta$ with $w \in K$ and $g(\beta) = 0$. One has $|w| \geqslant T_{r,0}$ and $|\beta| \geqslant s$, so that $|\xi| \geqslant sT_{r,0}$. □

For $k = r = 1$, this bound is tight: $B_n^1(x)$ has a maximum at $1/n$, cf. [3]. The following corollary combines Theorem 8 with Corollary 4.

Corollary 9. *Let I be an open interval enclosing a k-fold root α of the real polynomial f. Let s and $T_{r,0}$ be as above. If the width of I is $sT_{r,0}$ or less, then the Descartes test for f in I counts exactly k sign variations.*

How good is this result? For $k = 1$, we have $T_{1,0} = 1/n$, so Corollary 9 requires I to have width $|I| \leqslant 1/n \cdot s$ to guarantee that the Descartes test counts $v = 1$. The Two-Circle Theorem after Ostrowski [7, Theorem 21] already applies if $|I| \leqslant \sqrt{3}/2 \cdot s$. However, Ostrowski's result is inapplicable for $k > 1$. Furthermore, the relevant quantity in the analysis of root isolation by recursive bisection is $\log(1/|I|)$, and regarding that, we attain the same asymptotics: suppose that f has integer coefficients of at most τ bits (that is, with magnitude less than 2^τ) in the *monomial* basis. Its root separation $\mathrm{sep}(f)$, that is the minimum distance between any two *different* roots, satisfies $\log(1/\mathrm{sep}(f)) = O(n\tau + n \log n)$, see [1, Proposition 10.21]. Using the estimate $T_{r,0} > n^{-n}$, Theorem 8 implies that the minimum distance σ between any k-fold root α of f and the nearest root of $f^{(k)}$ satisfies the same asymptotic bound: $\log(1/\sigma) \leqslant \log(1/\mathrm{sep}(f)) + n \log n = O(n\tau + n \log n)$. From this point of view, Corollary 9 improves upon the state of the art by removing the restriction on α to be a simple root.

4. Conclusion

We have given a new partial converse of Descartes' Rule of Signs for k-fold roots based on roots of the kth derivative. Using the Bernstein basis, the proof has been rather simple. To make the result quantitative, we have extended an approach of Dimitrov and used Szegő's composition theorem to bound the minimum distance between a kth root and a root of the rth derivative. This second result, which may also be of independent interest, has allowed us to extend the asymptotic bound on the depth of recursive bisection needed to obtain isolating intervals certified by the Descartes test from simple to multiple real roots of integer polynomials.

Acknowledgements

I thank Kurt Mehlhorn and Raimund Seidel for pointing me towards the problem, Michael Kerber and Nicola Wolpert for discussions about it.

References

[1] S. Basu, R. Pollack, M.-F. Roy, Algorithms in Real Algebraic Geometry, Springer, Berlin, 2003, revised electronic version at ⟨http://www.math.gatech.edu/∼saugata/bpr-posted1.html⟩.
[2] G.E. Collins, A.G. Akritas, Polynomial real root isolation using Descartes' Rule of Signs, in: R.D. Jenks (Ed.), Proceedings of the 1976 ACM Symposium on Symbolic and Algebraic Computation, ACM Press, New York, 1976, pp. 272–275.
[3] D.K. Dimitrov, A refinement of the Gauss–Lucas theorem, Proc. Amer. Math. Soc. 126 (1998) 2065–2070.
[4] G.E. Farin, Curves and Surfaces for Computer-Aided Geometric Design, fourth ed., Academic Press, New York, 1997.
[5] C.G.J. Jacobi, Observatiunculae ad theoriam aequationum pertinentes, J. Reine Angewandt. Math. 13 (1835) 340–352, available from ⟨http://gdz.sub.uni-goettingen.de⟩.
[6] W. Krandick, Isolierung reeller Nullstellen von Polynomen, in: J. Herzberger (Ed.), Wissenschaftliches Rechnen, Akademie-Verlag, Berlin, 1995, pp. 105–154.
[7] W. Krandick, K. Mehlhorn, New bounds for the Descartes method, J. Symbolic Comput. 41 (2006) 49–66.
[8] J.M. Lane, R.F. Riesenfeld, Bounds on a polynomial, BIT 21 (1981) 112–117.
[9] B. Mourrain, F. Rouillier, M.-F. Roy, Bernstein's basis and real root isolation, Rapport de recherche 5149, INRIA-Rocquencourt, ⟨http://www.inria.fr/rrrt/rr-5149.html⟩ (March 2004).
[10] A. M. Ostrowski, Note on Vincent's theorem, Annals of Mathematics, Second Series 52 (1950) 702–707, reprinted in: Alexander Ostrowski, Collected Mathematical Papers, vol. 1, Birkhäuser, Basel, 1983, pp. 728–733.
[11] G. Pólya, I.J. Schoenberg, Remarks on de la Vallée Poussin means and convex conformal maps of the circle, Pacific J. Math. 8 (1958) 295–334.
[12] H. Prautzsch, W. Boehm, M. Paluszny, Bézier and B-Spline Techniques, Springer, Berlin, 2002.
[13] F. Rouillier, P. Zimmermann, Efficient isolation of [a] polynomial's real roots, J. Comput. Appl. Math. 162 (2004) 33–50.
[14] I.J. Schoenberg, On variation diminishing approximation methods, in: R.E. Langer (Ed.), On Numerical Approximation, University of Wisconsin Press, Madison, 1959, pp. 249–274.
[15] G. Szegő, Bemerkungen zu einem Satz von J. H. Grace über die Wurzeln algebraischer Gleichungen, Math. Z. 13 (1922) 28–55, available from ⟨http://gdz.sub.uni-goettingen.de⟩.