

STAN-CS-73-345

THE MINIMUM ROOT SEPARATION OF A
POLYNOMIAL

BY

GEORGE E. COLLINS
ELLIS HOROWITZ

SUPPORTED BY

THE NATIONAL SCIENCE FOUNDATION
AND
ADVANCED RESEARCH PROJECTS AGENCY

ARPA ORDER NO. 457

APRIL 1973

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY



APRIL 1973

COMPUTER SCIENCE DEPARTMENT
REPORT NO. STAN-CS-73-345

THE MINIMUM ROOT SEPARATION OF A POLYNOMIAL

by

George E. Collins (1)

Ellis Horowitz (2)

ABSTRACT: The minimum root separation of a complex polynomial A is defined as the minimum of the distances between distinct roots of A . For polynomials with Gaussian integer coefficients and no multiple roots, three lower bounds are derived for the root separation. In each case the bound is a function of the degree, n , of A and the sum, d , of the absolute values of the coefficients of A . The notion of a semi-norm for a commutative ring is defined, and it is shown how any semi-norm can be extended to polynomial rings and matrix rings, obtaining a very general analogue of Hadamard's determinant theorem.

- (1) Computer Science Department, Stanford University, and Computer Sciences Department, University of Wisconsin-Madison.
- (2) Computer Science Department, Cornell University.

This research is supported by Grants GJ-30125X and GJ-33169 from the National Science Foundation, the Wisconsin Alumni Research Foundation, and (in part) by the Advanced Research Projects Agency of the Office of the Secretary of Defense (SD-183).

The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency, NSF, or the U.S. Government.

Reproduced in the USA. Available from the National Technical Information Service, Springfield, Virginia 22151.

THE MINIMUM ROOT SEPARATION OF A POLYNOMIAL

1. Introduction. Let $A(x)$ be a polynomial of degree $n > 0$ with complex coefficients a_i and complex roots α_j , so that

$$A(x) = \sum_{i=0}^n a_i x^i = a_n \prod_{j=1}^n (x - \alpha_j). \quad (1)$$

We define $\text{sep}(A)$, the minimum root separation of A , by

$$\text{sep}(A) = \min_{\alpha_j \neq \alpha_k} |\alpha_j - \alpha_k|, \quad (2)$$

with the convention that $\text{sep}(A) = \infty$ in case A has only one distinct root.

The computing time required by any algorithm to isolate the zeros of A depends inversely on $\text{sep}(A)$. Hence we are interested in easily computable functions $f(a_0, \dots, a_n)$ of the coefficients such that

$$0 < f(a_0, \dots, a_n) \leq \text{sep}(A). \quad (3)$$

Heindel, [3], in analyzing the computing time of an algorithm based on Sturm's theorem for isolating the real zeros of any polynomial with integer coefficients, used a weak lower bound for $\text{sep}(A)$ due to Collins. Pinkert, [9], presents an analogous algorithm for isolating all zeros, real and complex, of any polynomial with Gaussian integer coefficients. His algorithm is based on Sturm's theorem and the Routh-Hurwitz theorem and uses a stronger lower bound for $\text{sep}(A)$ obtained more recently by Collins. Horowitz, using another simpler approach, has recently obtained a third lower bound, intermediate in strength, but just slightly weaker than the

stronger bound of Collins. In the following, these three bounds are all derived, with the hope of stimulating further research on the problem.

If $A(x)$ has rational complex coefficients, we can easily compute another polynomial, having the same roots, with Gaussian integer coefficients. Further, if $A(x)$ has Gaussian integer coefficients, we can easily compute another polynomial $A^*(x)$ with Gaussian integer coefficients, having the same roots as $A(x)$ and having only simple roots, namely

$$A^*(x) = A(x) / \gcd(A(x), A'(x)), \quad (4)$$

where $A'(x)$ is the derivative of $A(x)$ and "gcd" denotes the greatest common divisor. Hence in the following A is assumed to have Gaussian integer coefficients and no multiple roots.

Also, the three lower bounds to be obtained will all be of the form

$$0 < g(n, d) \leq \text{sep}(A), \quad (5)$$

where $n = \deg(A)$, the degree of A , and $d = \nu(A)$, where ν is some "semi-norm".

In the next section we introduce the notion of a semi-norm for a ring and then derive some lemmas which will be used in deriving the root separation theorems.

2. Semi-Norms and Resultants. If \mathcal{R} is any commutative ring, a semi-norm for \mathcal{R} is any function v from \mathcal{R} into the non-negative real numbers satisfying the following three conditions for all $a, b \in \mathcal{R}$:

$$v(a)=0 \text{ if and only if } a=0, \quad (6a)$$

$$v(a-b) \leq v(a)+v(b), \quad (6b)$$

$$v(ab) \leq v(a)v(b). \quad (6c)$$

These conditions imply also

$$v(-a)=v(a), \quad (6d)$$

$$v(a+b) \leq v(a)+v(b). \quad (6e)$$

A norm for \mathcal{R} is a semi-norm for \mathcal{R} such that

$$v(ab)=v(a)v(b). \quad (7)$$

For the ring G of the Gaussian integers a familiar norm is $v(a+bi)=|a+bi|=(a^2+b^2)^{1/2}$. A semi-norm for G which is not a norm is $v(a+bi)=|a+bi|_1=|a|+|b|$.

Any semi-norm v on a commutative ring \mathcal{R} can be extended to a semi-norm on the polynomial ring $\mathcal{R}[x]$ by the definition

$$v\left(\sum_{i=0}^n a_i x^i\right) = \sum_{i=0}^n v(a_i). \quad (8)$$

By induction on r , repeated application of (8) extends v to a semi-norm on $\mathcal{R}[x_1, \dots, x_r]$, which is easily seen to be independent of the order in which the indeterminates x_i are adjoined.

As a special case, (8) defines $|A|$ and $|A|_1$ for any Gaussian polynomial $A(x_1, \dots, x_r) \in G[x_1, \dots, x_r]$ as extensions of the semi-norms for G defined above. For integral polynomials $A(x_1, \dots, x_r)$ with rational integer coefficients, the norm $|A|_1$ has been used extensively for the analysis of

algebraic algorithms. See, for example, [1], [2], [7] and [8]. Its extension to Gaussian polynomials, however, is new.

If M is an arbitrary matrix (or vector) over \mathcal{R} , we define

$$\nu(M) = \sum_i \sum_j \nu(M_{ij}), \quad (9)$$

where the summation extends over all entries of M . It is easy to verify that the conditions (6a)-(6c) hold for matrices over \mathcal{R} whenever the operations are defined. In particular, this extends ν to a semi-norm for the ring of all n by n square matrices over \mathcal{R} .

By combining the semi-norm extensions for polynomials and matrices, we obtain the following general analogue of Hadamard's determinant theorem ([6], p. 208).

Theorem 1. Let \mathcal{R} be a commutative ring, ν a semi-norm for \mathcal{R} , M an n by n matrix over \mathcal{R} . Then

$$\nu(\det(M)) \leq \prod_{i=1}^n \nu(M_i) \quad (10)$$

where M_i is the i^{th} row of M and $\det(M)$ is the determinant of M .

Proof. By induction on n , the case $n=1$ being trivial. We denote by $M_{i,j}$ the element of M in the i^{th} row and j^{th} column of M , by $M'_{i,j}$ the submatrix of M obtained by deletion of the i^{th} row and j^{th} column. By Laplace expansion,

$$\det(M) = \sum_{j=1}^n \sum_{l=1}^n (-1)^{j+l} M_{lj} \det(M'_{lj}). \quad (11)$$

By (6) and (11),

$$\nu(\det(M)) \leq \sum_{j=1}^n \sum_{l=1}^n \nu(M_{lj}) \nu(\det(M'_{lj})). \quad (12)$$

with semi-norm v . Let $m = \deg(A) > 0$, $n = \deg(B) > 0$, $c = \text{res}(A, B)$. Then

$$v(c) \leq v(A)^n v(B)^m. \quad (16)$$

Also, there exist polynomials U and V over \mathcal{R} such that $AU + BV = c$, $\deg(U) < n$, $\deg(V) < m$,

$$v(U) \leq v(A)^{n-1} v(B)^m, \quad (17)$$

and

$$v(V) \leq v(A)^n v(B)^{m-1}. \quad (18)$$

Proof. If S_i is the i^{th} row of S then $v(S_i) = v(A)$ for $1 \leq i \leq n$ and $v(S_i) = v(B)$ for $n+1 \leq i \leq m+n$, and (16) follows from Theorem 1. Now consider the matrix S^* which is obtained by adding to the last column of S x^{m+n-i} times the i^{th} column of S , for $1 \leq i < m+n$. $\det(S^*) = \det(S) = c$ and the last column of S^* contains the entries $x^{n-1}A(x), \dots, xA(x), A(x), x^{m-1}B(x), \dots, xB(x), B(x)$. Applying the Laplace determinant expansion theorem to the last column of S^* we obtain $AU + BV = c$ with $\deg(U) \leq n-1$ and $\deg(V) \leq m-1$, where the coefficients of U and V are the cofactors of the last column of S^* . Each coefficient of U is the determinant of a matrix obtained from S by deleting one row of coefficients of A and the last column, and so Theorem 1 yields (17), and similarly (18) holds. \square

3. Root Separation Bounds. For each of the first two root separation bounds we will use the following upper bound on the roots of a polynomial.

Theorem 3. Let A be any non-zero Gaussian polynomial, and let α be a root of A . Then

$$|\alpha| < |A| / |a_n| \quad (19)$$

where $a_n = \text{ldcf}(A)$.

Proof. $|A| \geq |a_n|$, so **(19)** holds for $|\alpha| < 1$. Let $A(x) = \sum_{i=0}^n a_i x^i$ and assume $|\alpha| \geq 1$. Then $a_n \alpha^n = \sum_{i=0}^{n-1} a_i \alpha^i$, so

$$|a_n| \cdot |\alpha|^n \leq \sum_{i=0}^{n-1} |a_i| \cdot |\alpha|^i. \tag{20}$$

Dividing (20) by $|\alpha|^{n-1}$,

$$|a_n| \cdot |\alpha| \leq \sum_{i=0}^{n-1} |a_i| \cdot |\alpha|^{i-n+1} \leq \sum_{i=0}^{n-1} |a_i| < |A|, \tag{21}$$

from which **(19)** is immediate. ■

Theorem 4. (Collins, 1970) Let A be a Gaussian polynomial of degree $n \geq 2$ with only simple roots, and $d = |A|$. Then

$$\text{sep}(A) > (2d)^{-n(n-1)/2}. \tag{22}$$

Proof. Let $\alpha_1, \dots, \alpha_n$ be the zeros of A and $\lambda = \text{sep}(A)$. We may choose notation so that $\lambda = |\alpha_1 - \alpha_2|$. Let D be the discriminant of A , so that

$$D = a_n^{2n-2} \prod_{j < k} (\alpha_j - \alpha_k)^2, \tag{23}$$

and ([10], Section 28), D is a Gaussian integer. Since the α_j are distinct, $D \neq 0$ and hence $|D| \geq 1$. Combining this with (23), we have

$$1 \leq |a_n|^{2n-2} \prod_{j < k} |\alpha_j - \alpha_k|^2. \tag{24}$$

Dividing by λ^2 ,

$$\lambda^{-2} \leq |a_n|^{2n-2} \prod_{\substack{j < k \\ (j,k) \neq (1,2)}} |\alpha_j - \alpha_k|^2. \tag{25}$$

There are $(n^2 - n - 2)/2$ factors $|\alpha_j - \alpha_k|^2$ in (25) and $|\alpha_j - \alpha_k| \leq |\alpha_j| + |\alpha_k| < 2d/|a_n|$ by Theorem 3. Hence,

$$\lambda^{-2} \leq (2d)^{n^2 - n - 2} / |a_n|^{n^2 - 3n}. \tag{26}$$

Now $n^2 - 3n + 2 \geq 0$ and $|a_n| \geq 1$ so

$$\lambda^{-2} \leq (2d)^{n^2-n-2} |a_n|^2 < (2d)^{n^2-n}, \quad (27)$$

from which (22) is immediate. ■

Theorem 5. (Horowitz, 1973) Let A be a Gaussian polynomial of degree $n \geq 2$ with only simple roots, and $d = |A|$. Then

$$\text{sep}(A) \geq (nd)^{-4n+5}. \quad (28)$$

Proof. Let $\alpha_1, \dots, \alpha_n$ be the zeros of A and $\lambda = \text{sep}(A)$. We may suppose that $\lambda = |\alpha_1 - \alpha_2|$. By Theorem 2, there exist Gaussian polynomials U and V such that

$$AU + A'V = c, \quad (29)$$

$\deg(U) \leq n-2$ and $\deg(V) \leq n-1$, where $c = \text{res}(A, A')$. Since $A(x) = a_n \prod_{i=1}^n (x - \alpha_i)$, we have

$$A'(x) = a_n \sum_{j=1}^n \prod_{1 \leq i \leq n, i \neq j} (x - \alpha_i). \quad (30)$$

Evaluating (30) at $x = \alpha_1$, we obtain

$$A'(\alpha_1) = a_n \prod_{i=2}^n (\alpha_1 - \alpha_i). \quad (31)$$

Hence, evaluating (29) at $x = \alpha_1$ and using (31),

$$\left\{ a_n \prod_{i=2}^n (\alpha_1 - \alpha_i) \right\} V(\alpha_1) = c. \quad (32)$$

By [10], Section 28, $c = a_n D$, where D is the discriminant of A , a non-zero Gaussian integer. Hence $V(\alpha_1) \neq 0$ and by (32),

$$\text{sep}(A) = D / V(\alpha_1) \prod_{i=3}^n (\alpha_1 - \alpha_i). \quad (33)$$

$|A'| \leq n|A|$ so

$$|V| \leq n^{n-1} d^{2n-2} \quad (34)$$

by Theorem 2. Since $\deg(V) \leq n-1$ and $|\alpha_i| < d$,

$$|V(\alpha_1)| \leq |V| \cdot d^{n-1} \leq n^{n-1} d^{3n-3}. \quad (35)$$

From (33) and (35), using $|D| \geq 1$ and $|\alpha_1 - \alpha_i| < 2d$,

$$\text{sep}(A) \geq 2^{-n+2} n^{-n+1} d^{-4n+5}. \quad (36)$$

The proof is completed by observing that $n \geq 2$. ■

In order to obtain the third root separation bound, we construct a Gaussian polynomial B^* whose roots are all the differences $\alpha_i - \alpha_j$ with $i \neq j$. The idea of constructing B^* as a resultant was suggested by some current research of R. Loos, [5]. After obtaining upper bounds for the coefficients of B^* , we will apply the following theorem to obtain a lower bound for the roots of B^* , and hence for $\text{sep}(A)$.

Theorem 6. Let $A(x) = \sum_{i=0}^n a_i x^i$ be a complex polynomial of degree $n > 0$, with $a_0 \neq 0$. If α is any root of A , then

$$|\alpha| > \frac{1}{2} \min_{\substack{1 \leq i \leq n \\ a_i \neq 0}} |a_0/a_i|^{1/i}. \quad (37)$$

Proof. Let $A^*(x) = x^n A(x^{-1}) = \sum_{i=0}^n a_{n-i} x^i$. A^* is a polynomial of degree n whose roots are the reciprocals of the roots of A , for $A^*(x) = a_n x^n \prod_{i=1}^n (x^{-1} - \alpha_i) = a_n \prod_{i=1}^n (1 - \alpha_i x) = \{a_n \prod_{i=1}^n (-\alpha_i)\} \{\prod_{i=1}^n (x - \alpha_i^{-1})\} = a_n (a_0/a_n) \prod_{i=1}^n (x - \alpha_i^{-1}) = a_0 \prod_{i=1}^n (x - \alpha_i^{-1})$. Hence $A^*(\alpha^{-1}) = 0$ and from [4], Exercise 4.6.2.20, we have

$$|\alpha^{-1}| < 2 \max_{1 \leq i \leq n} |a_i/a_0|^{1/i}, \quad (38)$$

from which (37) is immediate. ■ (39)

Theorem 7. (Collins, 1973) Let A be a Gaussian polynomial of degree $n \geq 2$ with only simple roots and $d = |A|$. Then

$$\text{sep}(A) > \frac{1}{2} (e^{\frac{1}{2}} n^{3/2} d)^{-n}, \quad (40)$$

where e is the base of the natural logarithm.

Proof. Let $B(x)$ be the resultant of $A(y)$ and $A(x+y)$. If the coefficients and roots of A are given by (1), then,

$$A(x+y) = a_n \prod_{j=1}^n (y - (\alpha_j - x)). \quad (41)$$

Expressing the resultant $B(x)$ as a symmetric function of the roots of $A(y)$ and $A(x+y)$ by the theorem of van der Waerden ([10], Section 28),

$$B(x) = a_n^{2n} \prod_{1 \leq i, j \leq n} (x - (\alpha_i - \alpha_j)). \quad (42)$$

Since $\alpha_i = \alpha_j$ if and only if $i = j$, $B(x) = x^n \bar{B}(x)$, where

$$\bar{B}(x) = a_n^{2n} \prod_{i \neq j} (x - (\alpha_i - \alpha_j)), \quad (43)$$

is a polynomial of degree $n(n-1)$ with $\bar{B}(0) \neq 0$. Also, (43) can be written in the form

$$\bar{B}(x) = a_n^{2n} \prod_{i < j} (x^2 - (\alpha_i - \alpha_j)^2), \quad (44)$$

so that if $\bar{B}(x) = \sum_{i=0}^{n(n-1)} \bar{b}_i x^i$ then $\bar{b}_i = 0$ for i odd.

Expanding $A(x+y)$ in a Taylor series,

$$A(x+y) = \sum_{i=0}^{\infty} \frac{A^{(i)}(y)}{i!} x^i, \quad (45)$$

where $A^{(i)}$ is the i^{th} derivative of A . Let

$$A^*(x, y) = \{A(x+y) - A(y)\} / x = \sum_{i=1}^{\infty} \frac{A^{(i)}(y)}{i!} x^{i-1}. \quad (46)$$

Let M be the Sylvester matrix of $A(y)$ and $A(x+y)$. If we subtract the i^{th} row of M from the $(n+i)^{\text{th}}$ row and then divide the latter by x , for $1 \leq i \leq n$ we obtain a matrix \bar{M} such that $\det(M) = x^n \det(\bar{M})$. The first column of \bar{M} contains a_n in the first row and zeros elsewhere. Hence $\det(\bar{M}) = a_n \det(M^*)$, where M^* results from \bar{M} upon deletion of its first row and column. But M^* is the Sylvester matrix of $A(y)$ and $A^*(x, y)$, so

$$\bar{B}(x) = a_n B^*(x) \quad (47)$$

where $B^*(x)$ is the resultant of $A(y)$ and $A^*(x, y)$.

We now proceed to obtain bounds for the coefficients of B^* . Let

$$A_k^*(x,y) = \sum_{i=1}^{k+1} \{A^{(i)}(y)/i!\} x^{i-1}, \quad (48)$$

so that A_k^* is the result of deleting from A^* all terms of degree $k+1$ or greater in x . Since A^* and A_k^* are both of degree $n-1$ in y , $B^*(x) \equiv B_k^*(x)$ (modulo x^{k+1}) for $k \geq 0$. Hence the coefficients of x^k in $B^*(x)$ and $B_k^*(x)$ are identical, and if $B^*(x) = \sum_{i=0}^{n(n-1)} b_i^* x^i$ then

$$|b_k^*| \leq |B_k^*|. \quad (49)$$

Now $|A^{(i)}(y)/i!| \leq \binom{n}{i} d$, so, by (48),

$$|A_k^*(x,y)| \leq \sum_{i=1}^{k+1} \binom{n}{i} d \leq n^{k+1} d. \quad (50)$$

By Theorem 2 and (50),

$$|B_k^*| \leq e^n n^{(k+1)n} d^{2n-1}. \quad (51)$$

By (49) and (51), together with $|b_0^*| \geq 1$,

$$|b_0^*/b_{2k}^*| \geq e^{-1/2k} n^{-n/2} d^{-3n/2} n^{n+1/2} \quad (52)$$

for $k \geq 1$. Since $b_i^* = 0$ for i odd, by Theorem 6,

$$|\alpha_i - \alpha_j| > \frac{1}{2} (e^{1/2} n^{3/2} d)^{-n}, \quad (53)$$

completing the proof.)

The computing time of an algorithm, e.g. [9], for isolating the zeros of a Gaussian polynomial A is dominated (in the sense of [2]) by a polynomial function of $n = \deg(A)$, $\log d$ where $d = |A|$, and $-\log \text{sep}(A)$. If " \sim " denotes codominance of functions as in [2] and if $C_1(n,d)$, $H(n,d)$ and $C_2(n,d)$ are the bounds on $\text{sep}(A)$ given by Theorems 4, 5 and 7, then we have

$$-\log C_1(n,d) \sim n^2 \log d, \quad (54)$$

whereas

$$-\log H(n,d) \sim -\log C_2(n,d) \sim n \log nd. \quad (55)$$

In this sense the last two bounds are equivalent.

When $n=2$, $\text{sep}(A)$ can be given explicitly. If $A(x) = ax^2 + bx + c$ has two distinct roots, then

$$\text{sep}(A) = |b^2 - 4ac|^{1/2} / |a|. \quad (56)$$

Also, by Theorem 4, $\text{sep}(A) > 1/2d$. Let $a=k$, $b=2k-1$ and $c=k-1$ with $k \geq 1$.

Then $d = |A| = 4k-2$ and $\text{sep}(A) = 1/k < 4/(4k-2) = 4/d$.

Define

$$L(n,d) = \min\{\text{sep}(A) : \deg(A) = n, |A| \leq d\}. \quad (57)$$

Then, we have just shown,

$$L(2,d) \sim d^{-1}. \quad (58)$$

It does not seem unreasonable to ask for an explicit relation such as (58) for $L(3,d)$, but we have thus far not succeeded with this apparently simple problem. We know only, by Theorem 57 and some obvious examples, that

$$d^{-3} \leq L(3,d) \leq d^{-1}, \quad (59)$$

where " \leq " is the dominance relation.

REFERENCES

1. G. E. Collins, Computing Time Analyses for Some Arithmetic and Algebraic Algorithms, Proc. 1968 Summer Institute on Symbolic Mathematical Computation, pp. 197-231, IBM Corp., Cambridge, Mass., **1961**.
2. G. E. Collins, The Calculation of Multivariate Polynomial Resultants, Jour. A.C.M., Vol. **18**, No. **4** (Oct. 1971), pp. 515-532.
3. L. E. Heindel, Integer Arithmetic Algorithms for Polynomial Real Zero Determination, Jour. A.C.M., Vol. **18**, No. **4** (Oct. 1971), pp. 533-548.
4. D. E. Knuth, The Art of Computer Programming, Vol. 2: Seminumerical Algorithms, Addison-Wesley, Reading, Mass., **1969**.
5. R. Loos, Resultant Algorithms for Exact Operations on Algebraic Numbers, in preparation.
6. M. Marcus and H. Minc, Introduction to Linear Algebra, Macmillan co., New York, **1965**.
7. M. T. McClellan, The Exact Solution of Systems of Linear Equations with Polynomial Coefficients, Jour. A.C.M., to appear.
8. D. R. Musser, Algorithms for Polynomial Factorization, Univ. of Wisconsin Comp. Sci. Dept. Technical Report No. **134** (Ph.D Thesis), Sept. 1971, **174** pages.
9. J. R. Pinkert, Algebraic Algorithms for Computing the Complex Zeros of Gaussian Polynomials, Univ. of Wisconsin Comp. Sci. Dept. Ph.D. Thesis, in preparation.
10. B. L. van der Waerden, Modern Algebra, Vol. I, Ungar Publishing co., New York, **1953**.