

The Minimum Root Separation of a Polynomial*

By George E. Collins and Ellis Horowitz

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 seminorm for a commutative ring is defined, and it is shown how any seminorm can be extended to polynomial rings and matrix rings, obtaining a very general analogue of Hadamard's determinant theorem.

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

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

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

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

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

The computing times required by known algorithms for isolating the zeros of A depend inversely on $\text{sep}(A)$. Hence, we are interested in easily computable functions $f(a_0, \dots, a_n)$ of the coefficients such that

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

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.

Received April 16, 1973.

AMS (MOS) subject classifications (1970). Primary 12D10.

Key words and phrases. Polynomial zeros, root separation, Hadamard's theorem, seminorms.

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

Copyright © 1974, American Mathematical Society

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

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

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

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

where $n = \deg(A)$, the degree of A , and $d = \nu(A)$, where ν is some "seminorm". In the next section, we introduce the notion of a seminorm for a ring and then derive some lemmas which will be used in deriving the root separation theorems.

2. Seminorms and Resultants. If \mathcal{R} is any commutative ring, a seminorm for \mathcal{R} is any function ν from \mathcal{R} into the nonnegative real numbers satisfying the following three conditions for all $a, b \in \mathcal{R}$:

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

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

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

These conditions imply also

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

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

A *norm* for \mathcal{R} is a seminorm for \mathcal{R} such that

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

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

Any seminorm ν on a commutative ring \mathcal{R} can be extended to a seminorm on the polynomial ring $\mathcal{R}[x]$ by the definition

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

By induction on r , repeated application of (8) extends ν to a seminorm 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 seminorms 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

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

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 seminorm for the ring of all n by n square matrices over \mathcal{R} .

By combining the seminorm 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 seminorm for \mathcal{R} , M an n by n matrix over \mathcal{R} . Then*

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

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,

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

By (6) and (11),

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

The i th row of $M'_{1,j}$ is a subrow of M_{i+1} , so

$$(13) \quad \nu(\det(M'_{1,j})) \leq \prod_{i=2}^{n+1} \nu(M_i)$$

by the induction hypothesis. By (12) and (13),

$$(14) \quad \nu(\det(M)) \leq \left\{ \prod_{i=2}^{n+1} \nu(M_i) \right\} \sum_{j=1}^{n+1} \nu(M_{1,j}).$$

Since $\sum_{i=1}^{n+1} \nu(M_{1,i}) = \nu(M_1)$, this completes the induction. \square

A corollary of Theorem 1, needed in Section 3, will now be obtained by consideration of certain submatrices of the Sylvester matrix of two polynomials, A and B , over \mathcal{R} . Let $m = \deg(A)$, $n = \deg(B)$. The Sylvester matrix of A and B is the $m + n$ by $m + n$ matrix S whose successive rows are the coefficients of the polynomials $x^{n-1}A(x), \dots, xA(x), A(x), x^{m-1}B(x), \dots, xB(x), B(x)$. Diagrammatically, if

$$A(x) = \sum_{i=0}^m a_i x^i \quad \text{and} \quad B(x) = \sum_{i=0}^n b_i x^i,$$

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

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

from which (19) is immediate. \square

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

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

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

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

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

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

Dividing by λ^2 ,

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

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,

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

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

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

from which (22) is immediate. \square

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

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

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

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

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

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

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

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

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

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

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

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

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

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

by Theorem 2. Since $\text{deg}(V) \leq n - 1$ and $|\alpha_1| < d$,

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

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

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

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

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 well-known 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*

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

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

$$\begin{aligned} A^*(x) &= a_n x^n \prod_{i=1}^n (x^{-1} - \alpha_i) = a_n \prod_{i=1}^n (1 - \alpha_i x) \\ &= \left\{ a_n \prod_{i=1}^n (-\alpha_i) \right\} \left\{ \prod_{i=1}^n (x - \alpha_i^{-1}) \right\} = a_n (a_0/a_n) \prod_{i=1}^n (x - \alpha_i^{-1}) = a_0 \prod_{i=1}^n (x - \alpha_i^{-1}). \end{aligned}$$

Hence, $A^*(\alpha^{-1}) = 0$ and, from [4, Exercise 4.6.2.20], we have

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

from which (37) is immediate. \square

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

$$(39) \quad \text{sep}(A) > \frac{1}{2} (e^{1/2} n^{3/2} d)^{-n},$$

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,

$$(40) \quad A(x + y) = a_n \prod_{i=1}^n (y - (\alpha_i - x)).$$

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],

$$(41) \quad B(x) = a_n^{2n} \prod_{1 \leq i, j \leq n} (x - (\alpha_i - \alpha_j)).$$

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

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

is a polynomial of degree $n(n - 1)$ with $\bar{B}(0) \neq 0$. Also, (42) can be written in the form

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

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,

$$(44) \quad A(x + y) = \sum_{i=0}^n \{ A^{(i)}(y)/i! \} x^i,$$

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

$$(45) \quad A^*(x, y) = \{ A(x + y) - A(y) \} / x = \sum_{i=1}^n \{ A^{(i)}(y)/i! \} x^{i-1}.$$

Let M be the Sylvester matrix of $A(y)$ and $A(x + y)$. If we subtract the i th row of M from the $(n + i)$ 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

$$(46) \quad \bar{B}(x) = a_n B^*(x)$$

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

$$(47) \quad A_k^*(x, y) = \sum_{i=1}^{k+1} \{ A^{(i)}(y)/i! \} x^{i-1},$$

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

$$(48) \quad |b_k^*| \leq |B_k^*|.$$

Now, $|A^{(i)}(y)/i!| \leq \binom{n}{i} d$, so, by (47),

$$(49) \quad |A_k^*(x, y)| \leq \sum_{i=1}^{k+1} \binom{n}{i} d \leq en^{k+1} d.$$

By Theorem 2 and (49),

$$(50) \quad |B_k^*| \leq e^n n^{(k+1)n} d^{2n-1}.$$

By (48) and (50), together with $|b_0^*| \geq 1$,

$$(51) \quad |b_0^*/b_{2k}^*|^{1/2k} \geq e^{-n/2} n^{-3n/2} d^{-n+1/2}$$

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

$$(52) \quad |\alpha_i - \alpha_j| > \frac{1}{2}(e^{1/2} n^{3/2} d)^{-n},$$

completing the proof. \square

The computing time of Pinkert's algorithm in [9] for isolating the zeros of a Gaussian polynomial A is dominated (in the sense of [2]) by a polynomial function of $n = \text{deg}(A)$, $\log d$ where $d = |A|$, and $\log \lambda^{-1}$ where $\lambda = \text{sep}(A)$. Specifically, $n^7(\log d \lambda^{-1})(\log nd \lambda^{-1})^2$ is derived by Pinkert as a dominating function. 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

$$(53) \quad \log C_1(n, d)^{-1} \sim n^2 \log d,$$

whereas

$$(54) \quad \log H(n, d)^{-1} \sim \log C_2(n, d)^{-1} \sim n \log nd.$$

Thus, although $C_2(n, d)$ is generally a much sharper bound than $H(n, d)$, the use of either yields $n^{10} (\log nd)^3$ as a dominating function for the computing time of Pinkert's algorithm.

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

$$(55) \quad \text{sep}(A) = |b^2 - 4ac|^{1/2}/|a|.$$

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

$$(56) \quad L(n, d) = \min\{\text{sep}(A) : \text{deg}(A) = n \text{ \& \ } |A| \leq d\}.$$

Then, we have just shown,

$$(57) \quad L(2, d) \sim d^{-1}.$$

It does not seem unreasonable to ask for an explicit relation such as (57) for $L(3, d)$, but we have thus far not succeeded with this apparently simple problem. We know only, by Theorem 7 and some obvious examples, that

$$(58) \quad d^{-3} \leq L(3, d) \leq d^{-1},$$

where " \leq " is the dominance relation.

Computer Science Program
University of Southern California
Los Angeles, California 90007

1. G. E. COLLINS, "Computing time analyses for some arithmetic and algebraic algorithms," *Proc. 1968 Summer Institute on Symbolic Mathematical Computation*, IBM Corp., Cambridge, Mass., 1961, pp. 197-231.
2. G. E. COLLINS, "The calculation of multivariate polynomial resultants," *J. Assoc. Comput. Mach.*, v. 18, 1971, pp. 515-532. MR 45 #7970.
3. L. E. HEINDEL, "Integer arithmetic algorithms for polynomial real zero determination," *J. Assoc. Comput. Mach.*, v. 18, 1971, pp. 533-548. MR 45 #9480.
4. D. E. KNUTH, *The Art of Computer Programming*. Vol. 2: *Seminumerical Algorithms*, Addison-Wesley, Reading, Mass., 1969. MR 44 #3531.
5. R. G. K. LOOS, "A constructive approach to algebraic numbers," *Math. of Comp.* (submitted.)
6. H. MINC & M. MARCUS, *Introduction to Linear Algebra*, Macmillan, New York, 1965. MR 32 #5660.
7. M. T. MCCLELLAN, "The exact solution of systems of linear equations with polynomial coefficients," *J. Assoc. Comput. Mach.*, v. 20, 1973, pp. 563-588.
8. D. R. MUSSER, *Algorithms for Polynomial Factorization*, Univ. of Wisconsin Comp. Sci. Dept. Technical Report No. 134 (Ph.D Thesis), Sept. 1971, 174 pp.
9. J. R. PINKERT, *Algebraic Algorithms for Computing the Complex Zeros of Gaussian Polynomials*, Univ. of Wisconsin Comp. Sci. Dept. Ph.D. Thesis, May 1973, Technical Report No. 188, July 1973.
10. B. L. VAN DER WAERDEN, *Moderne Algebra*. Vol. I, Springer, Berlin, 1930; English transl., Ungar, New York, 1949. MR 10, 587.