Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Inequalities and Bounds for the Zeros of Polynomials Using Perron-Frobenius and Gerschgorin Theories

Mohammed A. Hasan
Department of Electrical & Computer Engineering
University of Minnesota Duluth
Email:mhasan@d.umn.edu

Abstract

In this paper, disks containing some or all zeros of a complex polynomial or eigenvalues of a complex matrix are developed. These disks are based on extensions of Cauchy classical bounds, Perron-Frobenius theory of positive matrices, and Gerschgorin theory. As a special case, given a real polynomial with real maximum or minimum zero, intervals containing the extreme zeros are developed. Moreover, methods for computing or refining these intervals are derived. Additionally, a closed form singular value decomposition of a characteristic polynomial was derived and utilized to compute new bounds for the zeros of polynomials. Finally, bounds which are based on zero transformation are given.

1 Introduction

The problem of finding regions that contain some or all eigenvalues of matrices or zeros of polynomials has a long history. Cauchy [1] gave an easily-calculated circular bound for complex coefficient polynomial zeros. Most early research on this problem was based on Cauchy's work. For matrix eigenvalues, Gerschgorin theorem [2] is very powerful tool for improving existing bounds or computing new ones. It can be used to find disks whose union contains all eigenvalues of a complex matrix. Such results may be applied to analyze the stability or the relative stability of discrete-time systems. They may also be applied to determine the eigenvalue distribution in certain disks for continuous-time systems, by shifting the origin of the s-plane.

In recent years, numerous papers and comprehensive books have been published, for finding circular bounds of polynomial zeros which have real or complex coefficients; see for example [2]-[3], and the references therein. An account of several developments on this topic can be found in the comprehensive book by Marden [4]. Extensions of Cauchy's result are given in [5] by introducing polynomial transformations and circular bounds are derived by minimizing Cauchy's bound for a transformed polynomial. In [6], sharper circular bounds have been provided by using quadratic polynomial transformations together with the Cauchy's circular bound. Several other bounds are explored in [7]-[10].

In this paper, we use the Perron-Frobenius Theory and Gerschgorin results to find new bounds for the zeros of polynomials. Additionally, disks of varying centers and radii which contain all or some zeros are developed. The key idea here is to determine polynomials for the rth powers of the zeros of the original polynomial, then using Cauchy and Gerschgorin bounds to derive new bounds. Additionally, circular bounds for interior and extremum zeros are developed. For polynomials whose zeros are all real, these bounds can be refined using Newton's method. We will also developed bounds that are based on the singular value decomposition of the characteristic polynomial. These bounds demonstrated improvement over known results.

2 Gerschgorin-Type Bounds

The Gerschgorin theorem finds a region in the complex plane that contains all the eigenvalues of a complex square matrix. Let $A = [a_{ij}] \in \mathcal{C}^{n\times n}$ and set $A^+ = [|a_{ij}|]$. Assume that R_i and C_i are defines as $R_i = \sum_{\substack{k=1\\k\neq i}}^{n} |a_{ik}|$ and $C_i = \sum_{\substack{k=1\\k\neq i}}^{n} |a_{ik}|$, then each eigenvalue of A is in at least one of the disks

$$|z - a_{ii}| \leq R_i. \tag{1a}$$

Similarly, each eigenvalue of A is in at least one of the disks

$$|z - a_{ii}| \leq C_i. \tag{1b}$$

It is known from the Perron-Frobenius theory on positive matrices that A^+ has a real, positive, and simple eigenvalue of maximum modulus r. The relation between the eigenvalues of A and A^+ are given in the following result which is stated as a Lemma in [11].

Theorem 1 (Wielandt)[11]. *Let A and A^+ be as defined above. Let λ be an eigenvalue of A. If A^+ is irreducible, then $|\lambda| \leq r$.*

The number r can be computed by solving the minimax problem [12]

$$r = \min_{x_i\geq 0} \max_{i=1}^{n} \frac{(A^+x)_i}{x_i}, \tag{2}$$

where $x = [\, x_1 \quad x_2 \quad \cdots \quad x_n \,]$.

To apply this result to polynomials, let

$$P(z) = z^n + \sum_{k=1}^{n} a_k z^{n-k}, \tag{3}$$

be a monic plynomial of degree n with complex coefficients and assume that $a_n \neq 0$. Let

$$C = \begin{bmatrix} 0 & \rho & 0 & 0 & \cdots & 0 \\ 0 & 0 & \rho & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{-a_n}{\rho^{n-1}} & \frac{-a_{n-1}}{\rho^{n-2}} & \frac{-a_{n-2}}{\rho^{n-3}} & \cdots & \frac{-a_2}{\rho} & -a_1 \end{bmatrix}, \tag{4a}$$

be a companion matrix for $P(z)$, where $\rho \neq 0$. It can be shown that the eigenvalues of A are the zeros of $P(z)$ regardless of the value of $\rho \neq 0$. Now C^+ is given as

$$C^+ = \begin{bmatrix} 0 & \rho & 0 & 0 & \cdots & 0 \\ 0 & 0 & \rho & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{|a_n|}{\rho^{n-1}} & \frac{|a_{n-1}|}{\rho^{n-2}} & \frac{|a_{n-2}|}{\rho^{n-3}} & \cdots & \frac{|a_2|}{\rho} & |a_1| \end{bmatrix}, \qquad (4b)$$

where it is assumed that $\rho > 0$ for convenience. The characteristic polynomial for C^+ is

$$P_a(z) = z^n - \sum_{k=1}^{n} |a_k| z^{n-k}, \qquad (4c)$$

The assumption $a_n \neq 0$ implies that C^+ is irreducible and hence the following:

Theorem 2 [12]. *The polynomial $P_a(z)$ of (4c) has a real, positive, simple zero r which is not exceeded by the modulus of any other zero of $P(z)$. If precisely K zeros of $P(z)$ have modulus r, then each of these satisfies*

$$z^K - r^K = 0.$$

The number r is precisely given by $r = \min_{x_i \geq 0} \max_{i=1}^{n} \frac{(C^+x)_i}{x_i}$.

Replacing C^+ by its transpose in (4b) and carrying out the multiplication indicated we deduce the next result.

Corollary 3. *The number r in Theorem 2 is given as*

$$r = \min_{\substack{x_i \geq 0 \\ \rho > 0}} \max_{i=0}^{n} \{\frac{x_n}{x_{i+1}} \frac{a_{n-i}}{\rho^{n-i-1}} + \frac{x_i \rho}{x_{i+1}}\}. \qquad (5)$$

with the understanding that $x_0 = 0$. Hence if z is a zero of $P(z)$ of (3), and $x_1, \cdots, x_n$ are arbitrarily chosen positive numbers, we have $|z| \leq r$.

3 A Few Known Bounds

In the following result we list a few known bounds including Cauchy's result.

Theorem 4. *The zeros of $P(z)$ satisfy $|z| \leq M$ if*

1. *$M = \max\{1 + |a_1|, \cdots, 1 + |a_{n-1}|, |a_n|\}$*
2. *$M = 2\max\{|a_k|^{\frac{1}{k}}\}$*
3. *Assume that $(|a_1|, \cdots, |a_{n-1}| \neq 0)$, then $M = \max\{2|a_1|, \frac{|a_2|}{|a_1|} + |a_1|, \cdots, \frac{|a_{n-1}|}{|a_{n-2}|} + |a_1|, \frac{|a_n|}{|a_{n-1}|}\}$.*

Many other bounds can be found in [4].

3.1 Miscellaneous Bounds

Different choices of the x_is in (5) lead to different bounds. Taking $x_i = 1 (i = 1, \cdots, n)$ and $\rho = 1$, we get

$$|z| \leq \max \left\{ 1 + |a_1|, 1 + |a_2|, \cdots, 1 + |a_{n-1}|, |a_n| \right\}. \qquad (6)$$

due to Cauchy [2].

With $x_i = |a_i| x_1$, $(i = 1, \cdots, n)$ and $\rho = 1$ there results

$$|z| \leq \max \left\{ \frac{|a_n|}{|a_{n-1}|}, 2\frac{|a_{n-1}|}{|a_{n-2}|}, \cdots, 2\frac{|a_2|}{|a_1|} \right\} \qquad (7a)$$

due to Kojima [13]. Next, with $x_i = \rho^i (i = 1, \cdots, n; \rho > 0)$, and $\rho = 1$

$$|z| \leq \max\{|a_1| + \rho, \frac{|a_2|}{\rho} + \rho, \cdots, \frac{|a_n|}{\rho^{n-1}}\}. \qquad (7b)$$

The exact minimum of the right hand expression is very difficult to find. This minimum is a solution of the following minimization problem: For each positive integer m let $f_m(\rho)$ be defined as

$$f_m(\rho) = \{\sum_{k=1}^{n-1} (\frac{|a_k|}{\rho^{k-1}} + \rho)^m + (\frac{|a_n|}{\rho^{n-1}})^m\}^{\frac{1}{m}}, \qquad (8a)$$

and let

$$R = \min_{\rho}\{\lim_{m \to \infty} f_m(\rho)\}. \qquad (8b)$$

Then (7b) can be written as

$$|z| \leq R. \qquad (8c)$$

If ρ is chosen so as to make each term of (7b) minimum, we find

$$|z| \leq \max_{2 \leq k \leq n} \{2|a_1|, \frac{k}{(k-1)^{\frac{k-1}{k}}} |a_k|^{\frac{1}{k}}\}. \qquad (9a)$$

This bound seems to be new. It can be shown that $\frac{k}{(k-1)^{\frac{k-1}{k}}} \leq 2$, for $k \geq 2$. Thus

$$|z| \leq 2 \max_{1 \leq k \leq n} \{|a_k|^{\frac{1}{k}}\}. \qquad (9b)$$

due to Fujiwara [14]. Clearly, better bounds may results from better choices of ρ.

By applying Gerschgorin Theorem to the matrix C, we obtain the following.

Theorem 5. *Suppose $P(z) = 0$ and $\rho > 0$. Then $|z| \leq M(\rho)$, where*

$$M(\rho) = \max\{\rho, |a_1| + |a_2|\rho^{-1} + \cdots + |a_n|\rho^{1-n}\};$$

and hence $|z| \leq M$ if $\rho = M$ satisfies

$$\rho \geq |a_1| + |a_2|\rho^{-1} + \cdots + |a_n|\rho^{1-n}, \qquad (10)$$

and $\rho > 0$.

Note that the equality in (10) is equivalent to $P_a(\rho) = 0$. In the next section, a method for computing the smallest M is given.

4 Zeros' Upper Bound Estimation

From the Perron-Frobenius theory, one deduces that the polynomial $P(z)$ of (3) has a zero of maximum modulus that is positive, real, and simple. The Newton method can be used to solve this equation, however, there is no guarantees that the method converges to the desired solution. The next theorem shows that under mild assumptions the method converges to the largest zero.

Theorem 6. *Suppose $P(z)$ is a real monic polynomial of degree n. Let $\{z_1, z_2, \cdots, z_n\}$ be the zeros of $P(z)$ such that z_n is real and $z_n > Re(z_k), (k = 1, \cdots, n-1)$. For every real number z such that and $z > z_n$, the following inequality holds:*

$$z - \frac{nP(z)}{P'(z)} \leq z_n < z - \frac{P(z)}{P'(z)} < z. \tag{10a}$$

Similarly, if z_1 is real and $z_1 < Re(z_k), (k = 2, \cdots, n)$, then for every real number $z < z_1$ the following inequality holds:

$$z < z - \frac{P(z)}{P'(z)} < z_1 < z - \frac{nP(z)}{P'(z)} < z_n. \tag{10b}$$

Here $Re(z)$ and $Im(z)$ denote the real and imaginary parts of z.

Proof: Assume that each zero of P is simple. Let $z_n > Re(z_k), (k = 1, \cdots, n-1)$ and assume that $z > z_n$, then

$$\frac{P'(z)}{P(z)} = \sum_{k=1}^{n} \frac{1}{z - z_k} = \frac{1}{z - z_m} + \sum_{k=1}^{n-1} \frac{1}{z - z_k}.$$

If one zero is complex, its conjugate is also a zero. Thus assume that $z_1 = \beta_1 + j\gamma_1$ and $z_2 = \beta_1 - j\gamma_1$, where β_1 and γ_1 are real, then

$$\frac{1}{z - z_1} + \frac{1}{z - z_2} = \frac{2(z - \beta_1)}{(z - z_1)(z - z_1^*)}.$$

This means that $\frac{1}{z-z_1} + \frac{1}{z-z_2} > 0$ for $z > \beta_1$, and

$$\begin{aligned} &(z - z_n)\left(\frac{1}{z - z_1} + \frac{1}{z - z_2}\right) \leq \frac{2(z - \beta_1)(z - z_n)}{(z - z_1)(z - z_1^*)} \\ &\leq \frac{2(z - \beta_1)^2}{(z - \beta_1)^2 + z_1 z_1^*} \leq 2 \end{aligned} \quad . \tag{10c}$$

Therefore,

$$z - z_n - \frac{P(z)}{P'(z)} = z - z_n - \frac{1}{\sum_{k=1}^{n} \frac{1}{z - z_k}} = \frac{\sum_{k=1}^{n-1} \frac{z - z_n}{z - z_k}}{\sum_{k=1}^{n} \frac{1}{z - z_k}} > 0 \tag{10d}$$

Since $\frac{P'(z)}{P(z)} > 0$ for $z \geq z_n$, there results

$$z > z - \frac{P(z)}{P'(z)} > z_n$$

Finally, from (10c) and (10d) it follows that

$$z - \frac{nP(z)}{P'(z)} < z_n < z$$

In similar argument, one can prove analogous inequalities involving z_1.

Q.E.D.

If all zeros of $P(z)$ are complex, then a bound on the real part of all zeros can be derived as shown in the next result.

Proposition 7. *Let $P(z)$ be an even degree real polynomial the zeros of which are all complex. Let $\{a_k = Re(z_k)\}_{k=1}^{m}$ where $m = \frac{n}{2}$ be the set of the real parts of the zeros of P. Assume that $a_1 \leq a_2 \leq \cdots \leq a_m$ and let $z > a_n$, then*

$$z - \frac{nP(z)}{P'(z)} < a_m < z, \tag{11}$$

where n is the degree of $P(z)$.

Proof: Assume that $z > a_m$, then

$$\frac{P'(z)}{P(z)} = \sum_{k=1}^{n} \frac{1}{z - z_k} \leq \sum_{k=1}^{m} \frac{2}{z - a_k} \leq \frac{n}{z - a_m}.$$

Hence $z - \frac{nP(z)}{P'(z)} < a_m$. In similar argument, one can show that for $z < a_1$, it follows that $z < a_1 < z - \frac{nP(z)}{P'(z)} < a_m$.

Q.E.D.

Corollary 8. *Suppose that $P(z)$ is a real polynomial and suppose that $P(z)$ has a zero of maximum modulus that is positive, real, and simple. Let z_0 be an upper bound for the moduli of all zeros of P, then*

$$z_1 = z_0 - \frac{P(z_0)}{P'(z_0)} \tag{12}$$

is a smaller upper bound for the moduli of all zeros of p. i.e, if $z_0 = 1 + \max\{|a_k|\}$ is the Cauchy bound, then the Newton method converges to the maximum zero of $P(z)$.

Proof: Let $\{z_1, z_2, \cdots, z_n\}$ be the zeros of $P_a(z)$ and let $z_n = r$ be the zero of maximum modulus of $P_a(z)$. This zero is real positive and simple, and $r > \{Re(z_k)\}_{k=1}^{n-1}$. Hence if z_0 is an upper bound (e.g. the bounds given in Theorem 4 and Section 3.1) then

$$z_0 - \frac{nP(z_0)}{P'(z_0)} \leq z_n < z - \frac{P(z_0)}{P'(z_0)} < z_0. \tag{10a}$$

The maximum zero of $P_a(z)$ of (4c) can be computed using the a modified version of the power method as follows: Let

$$C = \begin{bmatrix} 0 & \rho & 0 & 0 & \cdots & 0 \\ 0 & 0 & \rho & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{|a_n|}{\rho^{n-1}} & \frac{|a_{n-1}|}{\rho^{n-2}} & \frac{|a_{n-2}|}{\rho^{n-3}} & \cdots & \frac{|a_2|}{\rho} & |a_1| \end{bmatrix}$$

and let c be the Cauchy bound defined as $c = 1 + \max\{|a_k|\}_{k=1}^{n}$. The following iteration gives an improved bound for the zeros of polynomial:

$$\begin{aligned} c_1 &= c \\ &\text{for } k = 1, 2, \cdots, K, \\ \mathbf{x}_k &= [1 \quad c_k \quad \cdots \quad c_k^{n-1}] \\ \mathbf{x}_{k+1} &= C\mathbf{x}_k \\ c_{k+1} &= \frac{\mathbf{x}_{k+1}^T C \mathbf{x}_{k+1}}{\mathbf{x}_{k+1}^T \mathbf{x}_{k+1}}. \end{aligned}$$

5 Some Properties of the Companion Matrix

The companion matrix C defined in (4a) has the following properties:

Proposition 9.

(a) *For each complex number z let the vector $q(z) = [1 \quad z^2 \quad \cdots \quad z^{n-1}]^T$. If z_i is a zero of $P(z)$, then the vector $q(z_i) = [1 \quad z_i^2 \quad \cdots \quad z_i^{n-1}]^T$ is a right eigenvector of C.*

(b) *The left eigenvector of C is given as $y(z_i)$ where $y(z) = [\,z^{n-1}+\sum_{k=1}^{n-2}a_k z^{n-k} \quad z^{n-2}+\sum_{k=1}^{n-3}a_k z^{n-k} \quad \cdots \quad 1\,]^T$.*

(c) *It can be verified that $x(z_i)^T y(z_i) = P'(z_i)$*

(d) $x(z_i)^T C^T y(z_i) = z_i P'(z_i) - P(z_i)$

From the above observation, one can see that the two sided Rayleigh iteration is equivalent to the Newton method. Specifically, for every complex number z, there holds

$$\frac{x(z)^T C^T y(z)}{x(z)^T y(z)} = z - \frac{P(z)}{P'(z)}. \qquad (11)$$

In the following algorithm, an implementation of the power method in conjunction with Rayleigh iteration [15] is given:

Algorithm 1

$c_0 = c$

for $k = 1, 2, \cdots, K$,

$\mathbf{x}_k = [\,1 \quad c_k \quad \cdots \quad c_k^{n-1}\,]$

$\mathbf{y}_k = [\,c_k^{n-1}+\sum_{k=1}^{n-2}a_k c_i^{n-k} \quad c_i^{n-2}+\sum_{k=1}^{n-3}a_k c_i^{n-k} \quad \cdots \quad 1\,]^T$

$\mathbf{x}_{k+1} = C\mathbf{x}_k$

$\mathbf{y}_{k+1} = C^T\mathbf{y}_k$

$c_{k+1} = \dfrac{\mathbf{y}_{k+1}^T C\mathbf{x}_{k+1}}{\mathbf{y}_{k+1}^T\mathbf{x}_{k+1}}$

To insure convergence to the desired bound, the initial value of c must be chosen as in Theorem 4, for example.

It is observed from numerical simulations that the number of iterations needed is relatively small ($K \le 4$).

6 Bounds via Singular Value Decomposition

Let x be an eigenvector of A corresponding to the eigenvalue λ, then $Ax = \lambda x$, and hence

$$||\lambda x|| = ||Ax|| \le ||A||||x||.$$

This implies that

$$|\lambda| \le ||A||, \qquad (12)$$

i.e., the moduli of all eigenvalues of A is dominated by the norm of the matrix A. If we consider the Euclidean norm, then $||A||$ is the largest singular value of A which is also the square root of the largest eigenvalue of AA^T or A^TA. The following result reveals the singular value decomposition of the companion matrix C.

Theorem 10. *The matrix $C \in \mathcal{C}^{n\times n}$ ($n \ge 3$) of (4a) has three distinct singular values, $\sigma_1 < \sigma_2 < \sigma_3$ where $\sigma_2 = \rho$ has multiplicity $n-2$. The square of the two extreme singular values σ_1^2 and σ_3^3 are the positive solutions to the quartic equation*

$$\sigma^4 - (\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}})\sigma^2 + \frac{a_n^2}{\rho^{2n-2}} = 0. \qquad (13a)$$

Proof: Consider CC^T and C^TC given by

$$CC^T = \begin{bmatrix} \rho^2 & 0 & 0 & \cdots & -a_n \\ 0 & \rho^2 & 0 & \cdots & -a_{n-1} \\ \vdots & \vdots & \vdots & \vdots & \\ -a_{n-1} & -a_{n-2} & \cdots & \cdots & a_1^2+a_2^2+\cdots+a_n^2 \end{bmatrix}, \qquad (13b)$$

and

$$C^TC = \begin{bmatrix} \frac{a_n^2}{\rho^{2n-2}} & \frac{a_n a_{n-1}}{\rho^{2n-3}} & \cdots & \frac{a_n a_1}{\rho^{n-1}} \\ \frac{a_{n-1}a_n}{\rho^{2n-3}} & \frac{a_{n-1}^2}{\rho^{2n-4}}+\rho^2 & \cdots & \frac{a_{n-1}a_1}{\rho^{n-2}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{a_1 a_n}{\rho^{n-1}} & \frac{a_1 a_{n-1}}{\rho^{n-2}} & \cdots & a_1^2+\rho^2 \end{bmatrix}. \qquad (13c)$$

The matrix CC^T can be expressed as

$$CC^T = \rho^2 I + \begin{bmatrix} O & b \\ b^T & \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}} - \rho^2 \end{bmatrix} = \begin{bmatrix} \rho^2 I & b \\ b^T & \frac{a_n^2}{\rho^{2n-2}} + b^Tb \end{bmatrix}, \qquad (14a)$$

where $b^T = \begin{bmatrix} \frac{a_{n-1}^2}{\rho^{2n-4}} & \cdots & a_1\rho \end{bmatrix}$.

Clearly, for every vector x such that $b^Tx = 0$, the vector $\begin{bmatrix} x \\ 0 \end{bmatrix}$ is an eigenvector corresponding to eigenvalue ρ^2. Since the dimension of the null space of b is $n-2$, then the eigenvalue ρ^2 is of multiplicity $n-2$. Since eigenvectors of CC^T are orthogonal, the other two eigenvectors are of the form $\begin{bmatrix} b \\ \alpha \end{bmatrix}$, for some $\alpha \in \mathcal{C}$. This implies that

$$\begin{aligned} &\rho + \rho^2 = \lambda, \\ &b^Tb + \rho(\frac{b^Tb}{\rho^2} + \frac{a_n^2}{\rho^{2n-2}}) = \rho\lambda. \end{aligned} \qquad (14b)$$

Solving these two equations for ρ and λ yields the quadratic equation

$$\lambda^2 - (\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}})\lambda + \frac{a_n^2}{\rho^{2n-4}} = 0. \qquad (14c)$$

The solutions of this equation are:

$$\lambda_\pm = \frac{\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}} \pm \sqrt{(\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}})^2 - 4\frac{a_n^2}{\rho^{2n-4}}}}{2}. \qquad (14d)$$

Note that both λ_+ and λ_- are positive and $\lambda_+ > \lambda_-$. Hence λ_+ is the maximum eigenvalue of CC^T and λ_- is the smallest eigenvalue of CC^T. Consequently, the maximum and minimum singular values of C are

$$\sigma_\pm = \sqrt{\frac{\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}} \pm \sqrt{(\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}})^2 - 4\frac{a_n^2}{\rho^{2n-4}}}}{2}}, \qquad (15)$$

i.e., $\sigma_1 = \sigma_-$ and $\sigma_3 = \sigma_+ = ||A||$. These singular values are function of ρ. Since

$$|z| \le \min_\rho(||A(\rho)||),$$

for every $\rho \ne 0$, one may choose different values of ρ to obtain different bounds for the zeros of $P(z)$.

A suboptimal way to minimize $||A(\rho)||$ is by minimizing the trace of CC^T, i.e.,

$$\min_\rho(trace(CC^T)) = \min_\rho\{(n-1)\rho^2 + \sum_{k=1}^{n}\frac{a_k^2}{\rho^{2k-2}}\}. \qquad (16a)$$

This amounts to solving the equation

$$\frac{\partial trace(CC^T))}{\partial \rho} = 2(n-1)\rho - \sum_{k=1}^{n} \frac{(2k-2)a_k^2}{\rho^{2k-3}} = 0. \quad (16b)$$

The matrix C^TC can be analyzed analogously. Let $b = \begin{bmatrix} \frac{a_{n-1}}{\rho^{n-2}} & \cdots & a_1 \end{bmatrix}^T$, then

$$C^TC = \begin{bmatrix} \frac{a_n^2}{\rho^{2n-2}} \\ b \end{bmatrix} \begin{bmatrix} \frac{a_n^2}{\rho^{2n-2}} & b^T \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & \rho^2 I \end{bmatrix}.$$

As in the previous analysis, it can be shown ρ^2 is an eigenvalue of multiplicity $n-2$. The corresponding eigenvectors are $\begin{bmatrix} 0 \\ x \end{bmatrix}$, where $x^Tb = 0$. Note that the null space of b is $(n-2)$-dimensional. The other two eigenvectors are of the form $\begin{bmatrix} \gamma \\ b \end{bmatrix}$ for some complex number γ. The corresponding eigenvalues are solutions of the quartic equation (13a).

Remark: Sharper bounds may be obtained via the singular value decomposition of C^r for some positive integer r. It can be shown that

$$||C(\rho)|| > \sqrt[r]{||C(\rho)^r||},\ r = 2, 3, \cdots. \quad (18)$$

Therefore, if $P(z) = 0$, then

$$|z| \le \sqrt[r]{||C(\rho)^r||},\ r = 2, 3, \cdots. \quad (19)$$

In the next section, extension to this remark will be given.

7 Bounds via Zero Transformations

Given the polynomial $P(z)$ in (3), define a new set of polynomials so that for any positive integer $r > 1$, the polynomial $P(z)$ can be written as

$$P(z) = \sum_{k=0}^{r-1} z^k P_k(z^r), \quad (20)$$

where each P_k is a polynomial of degree $\le n$. Let w be a primitive rth root of 1, $w^r = 1$, and generate the following polynomial:

$$\begin{aligned} Q_r(z) &= \Pi_{k=0}^{r-1} \left(\sum_{k=0}^{r-1} w^{kl} z^k P_k(z^r) \right) \\ &= \Pi_{k=0}^{r-1} P(w^k z). \end{aligned} \quad (21a)$$

The polynomial $Q_r(z)$ is of degree nr. It can be verified that the right hand side of (21a) is in fact a polynomial in z^r and that the coefficients of Q_r are real if the coefficients of P are real. Let $q_r(u)$ be a polynomial of degree n obtained from the change of variable $u = z^r$, i.e.,

$$q_r(u) = Q_r(z)_{u=z^r}. \quad (21b)$$

The goal is to use $q_r(u)$ for computing new bounds for the zeros of P.

When $r = 2$, the above procedure reduces to expressing P as a sum of even and odd degree polynomials

$$P(z) = P_0(z^2) + zP_1(z^2).$$

If $P(z) = 0$, then $P_0(z)^2 - z^2P_1^2(z) = 0$. Thus if $Q_2(z) = 0$ then $Q_2(-z) = 0$. Using this observation, the following bound can be obtained.

Theorem 11. *All zeros of P lie in the disk*

$$|z| < \sqrt{A_2} \quad (22)$$

with $A_2 = 1 + \max\{|b_k|\}_{k=1}^n$, where b_k are the coefficients of the polynomial $q_2(u) = P_2(z^2)|_{u=z^2}$ where $q_2(u) = Q_2(z) = P_0(z)^2 - z^2P_1(z)^2$.

Proof: from the observation that $P(z) = P_0(z^2) + zP_1(z^2)$, it follows that

$$P(z)P(-z) = P_0(z^2)^2 - z^2P_1(z^2)^2.$$

This implies that $q_2(u) = P(z)P(-z)|_{u=z^2}$ is a polynomial in u. The conclusion follows directly from applying Theorem 1 to $q_2(u)$. Q.E.D.

It should be mentioned that the idea provided in the proof of Theorem 11 is the basis of Graeffe method [3]. The coefficients b_k were explicitly computed in [9]. If $r = 3$ we obtain the following bound.

Theorem 12. *All zeros of P lie in the disk*

$$|z| < (A_3)^{\frac{1}{3}} \quad (23)$$

with $A_3 = 1 + \max\{|b_k|\}_{k=1}^n$, where b_k are the coefficients of the polynomial $q_3(u) = Q_3(z)|_{u=z^3}$ where

$$Q_3(z) = P_0(z)^3 + z^3P_1(z)^3 + z^6P_2(z)^3 - 3z^3P_0(z)P_1(z)P_2(z).$$

Here P_0, P_1, P_2 are as described in (20) with $r = 3$.

Proof: The polynomial $P(z)$ can be expressed as

$$P(z) = P_0(z^3) + zP_1(z^3) + z^2P_2(z^3).$$

Let $w \ne 1$ such that $w^3 = 1$ and define

$$\begin{aligned} Q_3(z) = &\left(P_0(z^3) + zP_1(z^3) + z^2P_2(z^3)\right) \times \\ &\left(P_0(z^3) + wzP_1(z^3) + w^2z^2P_2(z^3)\right) \times \\ &\left(P_0(z^3) + w^2zP_1(z^3) + wz^2P_2(z^3)\right). \end{aligned}$$

It can be shown that $Q_3(z)$ is a function of z^3 and that all zeros of P are zeros of Q. In fact if z_k is a zero of P then z_k, wz_k, w^2z_k are zeros of Q.

Hence any bound on the zeros of the polynomial $q_3(u) = Q_3(z)_{u=z^3}$ is the cube of a bound of the zeros of P. Q.E.D.

The next result generalizes Theorem 11 and 12.

Theorem 13. *All zeros of P lie in the disk*

$$|z| < (A_r)^{\frac{1}{n}} \quad (24)$$

with $A_r = 1 + \max\{|b_k|\}_{k=1}^n$, where b_k are the coefficients of the polynomial $q_r(u) = Q_r(z)|_{u=z^r}$, and Q_r is as defined in (21).

In Eq.(21) if

$$q_r(u) = Q_r(z)|_{u=z^r} = b_0u^n + b_1u^{n-1} + b_2u^{n-2} + \cdots + b_n,$$

then a closed form of the coefficients b_k can be obtained for different values of r. It can be shown that the coefficients of the polynomial $q_r(z)$ are given by the formula:

$$b_m = \sum_{i_1+i_2+\cdots+i_r=mr} \Pi_{k=1}^r a_{i_k} w^{(k-1)(n-i_k)},\ m = 1, \cdots, n \quad (25)$$

where w is a primitive rth root of 1. In the above formula, which is derived directly using (21a), we are only considering the non-negative integer solutions of the equation $i_1 + i_2 + \cdots + i_r = mr$.

8 Bounds Using Matrix Powers and Gerschgorin Disks

As stated earlier the eigenvalues of the matrix

$$C = \begin{bmatrix} 0 & \rho & 0 & 0 & \cdots & 0 \\ 0 & 0 & \rho & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{-a_n}{\rho^{n-1}} & \frac{-a_{n-1}}{\rho^{n-2}} & \frac{-a_{n-2}}{\rho^{n-3}} & \cdots & \frac{-a_2}{\rho} & -a_1 \end{bmatrix},$$

are the zeros of the polynomial $P(z)$ for $\rho \neq 0$. Now

$$C^2 = \begin{bmatrix} 0 & 0 & \rho & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{-a_n}{\rho^{n-2}} & \frac{-a_{n-1}}{\rho^{n-3}} & \frac{-a_{n-2}}{\rho^{n-4}} & \cdots & -a_1\rho \\ \frac{a_1 a_n}{\rho^{n-1}} & \frac{a_1 a_{n-1} - a_n}{\rho^{n-2}} & \frac{a_1 a_{n-2} - a_{n-1}}{\rho^{n-3}} & \cdots & a_1^2 - a_2 \end{bmatrix}.$$

The matrices C and C^2 have $(n-2)$ rows in common. For every positive integer r it is the case that C^r and C^{r+1} have (n-2) common rows. Thus a recursive formula for computing powers of C can be developed. For simplicity, consider the matrix C^2. Applying Gerschgorin theorem, the zeros of P are contained in the three disks:

$$\begin{aligned} |z^2| &\leq |\rho|, \\ |z^2 + a_2| &\leq \sum_{\substack{k=1 \\ k\neq 2}}^{n} \left|\frac{a_k}{\rho^{k-2}}\right|, \\ |z^2 + a_2 - a_1^2| &\leq \sum_{\substack{k=1 \\ k\neq 2}}^{n} \left|\frac{a_1 a_{k-1} - a_k}{\rho^{k-2}}\right|. \end{aligned} \tag{26a}$$

Using the triangle inequality, we can consider larger regions so that:

$$\begin{aligned} |z^2| &\leq |\rho| \\ |z^2| &\leq \sum_{k=1}^{n} \left|\frac{a_k}{\rho^{k-2}}\right| \\ |z^2| &\leq \sum_{k=1}^{n} \left|\frac{a_1 a_{k-1} - a_k}{\rho^{k-2}}\right|. \end{aligned} \tag{26b}$$

Consequently, all zeros are contained in the disk of radius R, where

$$R = \min_{\rho} \max \left\{ \sqrt{|\rho|}, \sqrt{\sum_{k=1}^{n} \left|\frac{a_k}{\rho^{k-2}}\right|}, \sqrt{\sum_{k=1}^{n} \left|\frac{a_1 a_{k-1} - a_k}{\rho^{k-2}}\right|} \right\}. \tag{26c}$$

This idea can be generalized as follows. Let $r \geq 2$ be a positive integer and assume that $B_r(\rho) = C^r = [b_{ij}^{(r)}(\rho)]$. Then all zeros of P are contained in a disk of radius R, where

$$R = \min_{\rho} \max \left\{ \sqrt[r]{\sum_{k=1}^{n} |b_{ik}^{(r)}(\rho)|} \right\}_{i=1}^{n}. \tag{27a}$$

Analogous result holds if columns of B_r are used, i.e.,

$$R = \min_{\rho} \max \left\{ \sqrt[r]{\sum_{k=1}^{n} |b_{kj}^{(r)}(\rho)|} \right\}_{j=1}^{n}. \tag{27b}$$

9 Conclusion

It is known that bounds that are based on Cauchy's result, stated in Theorem 4, are conservative. The main focus of this paper has been the development of sharper circular bounds for the zeros of real and complex-coefficient polynomials. This is done via improving the classical result of Cauchy using Gerschgorin theorem and transformations of zeros. In addition to exploring methods for computing new bounds, methods for improving existing bounds are also derived. Additionally, disks based on the singular value decomposition of the characteristic polynomial and zero transformation are also developed. Finally, some of the results of this work can be modified so as to obtain lower bounds for the modulus of the smallest zero. Many of the developed bounds are tested on several examples but the results are not reported here due to space limitation.

References

[1] A. L. Cauchy, Exercises de Mathematique in Oeuvres, Vol. 9, 1829.

[2] P. Henrici, Applied and Computational Complex Analysis, Vol. 1, John Wiley & Sons, New York, 1974.

[3] A. S. Householder, The Numerical Treatment of a Single Nonlinear Equation, McGraw-Hill, New York, 1970.

[4] M. Marden, "Geometry of polynomials," Math. Surveys 3. Providence, RI: American Mathematical Society, 1966.

[5] Yeong-Jeu Sun and Jer-Guang Hsieh, "A Note on the Circular Pound of Polynomial Zeros," IEEE Trans. Circuits Syst I, Vol. 43, No. 6, June 1996.

[6] A. Joyal, G. Labelle, and Q. I. Rahman, "On the locations of zeros of polynomials," Canadian Math. Bulletin, Vol. 10, pp. 53-63, 1967.

[7] F. G. Boese and W. J. Luther, "A note on a classical bound for the moduli of all zeros of polynomial," IEEE Trans. Automat. Contr vol. 34, pp. 998-1001, 1989.

[8] E. Zeheb, "On the largest modulus of polynomial zeros," IEEE Trans. Circuits Syst., Vol. 38, pp. 333-337, 1991.

[9] M. S. Zilovic, L. M. Roytman, P. L. Combettes, and M. N. S. Swamy, "A bound for the zeros of polynomials," IEEE Trans. Circuits Syst I, Vol. 39, pp. 476-478, 1992.

[10] B. Dan and N. K. Govil, "On the location of the zeros of polynomial." J. Approx. Theory, Vol. 24, pp. 78-82, 1978.

[11] F. R. Gantmacher, Applications of the theory of matrices, New York, Interscience Publishers, 1959.

[12] H. S. Wilf, Perron-Frobenius theory and the zeros of polynomials, Proc. Amer. Math. Soc., 12 (1961) 247-250.

[13] T. Kojima, On the limits of the roots of an algebraic equation, Tohoku Math. J., (1) 11 (1917) 119-127.

[14] M. Fujiwara, Uber die obere Schranke des absoluten Betrages der Wurzeln einer algebraischen Gleichen, Tohoku Math. J., (1) 10 (1916) 167-171.

[15] B. N. Parlett, The Rayleigh Quotient Iteration and some generalizations for non-normal matrices, Mathematics of Computation 28 (1974), no. 127. pp:679-693.