

GENERALIZATIONS OF THE CAUCHY AND FUJIWARA BOUNDS FOR PRODUCTS OF ZEROS OF A POLYNOMIAL*

RAJESH PEREIRA[†] AND MOHAMMAD ALI VALI[‡]

Abstract. The Cauchy bound is one of the best known upper bounds for the modulus of the zeros of a polynomial. The Fujiwara bound is another useful upper bound for the modulus of the zeros of a polynomial. In this paper, compound matrices are used to derive a generalization of both the Cauchy bound and the Fujiwara bound. This generalization yields upper bounds for the modulus of the product of m zeros of the polynomial.

Key words. Zeros of polynomials, Inequalities, Cauchy bound, Companion matrix, Compound matrix.

AMS subject classifications. 15A18, 26D05, 26C10, 30C15.

1. Introduction. We follow [5, Chapter 8] in our introduction to the Cauchy bound.

DEFINITION 1.1. Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be an n -th degree monic polynomial with complex coefficients. Then the *Cauchy bound* of p , denoted as $\rho(p)$ is the unique positive root of the polynomial $z^n - \sum_{k=0}^{n-1} |a_k|z^k$.

THEOREM 1.2. [1] *Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be an n -th degree polynomial with complex coefficients. Then all zeros of p have modulus less than or equal to $\rho(p)$.*

We note that among monic polynomials, the Cauchy bound is a monotone increasing function of the moduli of the coefficients.

LEMMA 1.3. *Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ and $q(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1z + b_0$ be two n -th degree polynomials with complex coefficients. If $|a_k| \leq |b_k|$ for all $0 \leq k \leq n-1$, then $\rho(p) \leq \rho(q)$.*

*Received by the editors on June 19, 2016. Accepted for publication on August 19, 2016. Handling Editor: Michael Tsatsomeros.

[†]Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario N1G 2W1 (pereirar@uoguelph.ca). Supported by NSERC grant 400500.

[‡]Department of Applied Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran (mohamadali_35@yahoo.com).

Proof. Let $p_*(z) = z^n - |a_{n-1}|z^{n-1} - \dots - |a_1|z - |a_0|$ and $q_*(z) = z^n - |b_{n-1}|z^{n-1} - \dots - |b_1|z - |b_0|$, then $p_*(z) \geq q_*(z)$ for all positive z . Let $t = \rho(p)$. Then $q_*(t) \leq p_*(t) = 0$. Since $q_*(z) > 0$ when z is a sufficiently large positive number, $\rho(p) \leq \rho(q)$. \square

We wish to find a generalization of the Cauchy bound which gives an upper bound for all products of k zeros of p . We begin by introducing companion matrices and use them to give a proof of Theorem 1.2. This proof, while short and simple, is not shorter than the usual proof in the literature. The main advantage of this proof is that it can be generalized to give Cauchy-type bounds for products of zeros of p .

For any matrix denoted by a single capital letter, we follow the common convention in using the corresponding lower-case letter together with subscripts to denote the elements of that matrix. So a_{ij} denotes the entry in the i -th row and j -th column of A . We use $\text{diag}(s_1, s_2, \dots, s_n)$ to denote the diagonal matrix whose i -th row, i -th column entry is s_i . We remind the reader of the definition of the l^∞ vector norm which is $|v|_\infty = \max_{1 \leq k \leq n} |v_k|$, the induced operator norm is $\|A\|_\infty = \sup_{v \neq 0} \frac{|Av|_\infty}{|v|_\infty}$ which has the convenient operator norm $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$. It is clear that if λ is an eigenvalue of A , we must have $|\lambda| \leq \|A\|_\infty$.

2. Companion matrices and the higher Cauchy bounds. Let $p(z) = z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0$ be a monic polynomial. Then the matrix

$$M_p = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{bmatrix},$$

is the companion matrix with characteristic polynomial p .

The eigenvalues of the companion matrix are the zeros of the polynomial $p(z)$. Now consider the polynomial $q_t(z) = t^{-n}p(tz)$ for some fixed $t > 0$. The companion matrix of q_t is the following:

$$M_{q_t} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{a_0}{t^n} & -\frac{a_1}{t^{n-1}} & -\frac{a_2}{t^{n-2}} & \cdots & -\frac{a_{n-1}}{t} \end{bmatrix}.$$

We note that the smallest positive value of t for which $\|M_{q_t}\|_\infty = 1$ must satisfy the equation $\sum_{k=0}^{n-1} \frac{|a_k|}{t^{n-k}} = 1$. Multiplying both sides by t^n , we see that this is the

unique positive solution to $t^n - \sum_{k=0}^{n-1} |a_k| t^k$ which is $\rho(p)$. All the zeros of $q_{\rho(p)}(z)$ must have absolute value less than or equal to $\|M_{q_{\rho(p)}}\| = 1$. Since every zero of p is $\rho(p)$ times a zero of $q_{\rho(p)}$, we obtain another proof of Theorem 1.2.

As mentioned earlier, the proof is useful because it can be generalized to give new results. The generalization uses compound matrices in place of ordinary matrices. We begin by defining them and exploring a few of their properties. A more complete exposition of compound matrices can be found in [3, Section 19F]. We begin by introducing some notation. For any $m, n \in \mathbb{N}$ with $m \leq n$, we define $\mathcal{Q}_{m,n}$ to be the set of all m -tuples of integers $\alpha = \{\alpha_k\}_{k=1}^m$ satisfying $1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_{m-1} < \alpha_m \leq n$. The m -tuples in $\mathcal{Q}_{m,n}$ are ordered lexicographically. If $\alpha, \beta \in \mathcal{Q}_{m,n}$, then $A[\alpha|\beta]$ is the $m \times m$ matrix whose (i, j) -th entry is the (α_i, β_j) -th entry of A .

DEFINITION 2.1. Let A be an $n \times n$ matrix and m be an integer between one and n . Then the m -th compound matrix of A , $C_m(A)$ is an $\binom{n}{m} \times \binom{n}{m}$ matrix whose rows and columns are indexed by the elements of $\mathcal{Q}_{m,n}$ and whose (α, β) entry is the determinant of $A[\alpha|\beta]$ for all $\alpha, \beta \in \mathcal{Q}_{m,n}$.

We note that $C_1(A) = A$ and $C_n(A) = \det(A)$. The most important property of the compound mapping is that it is multiplicative.

LEMMA 2.2. [3, Theorem 19.F.2] Let A and B be $n \times n$ matrices and let $1 \leq k \leq n$, then $C_k(AB) = C_k(A)C_k(B)$.

This property is equivalent to the Binet-Cauchy theorem. The main use of compound matrices are their spectral properties which follow from the previous lemma together with the Jordan Canonical Form.

COROLLARY 2.3. [3, Theorem 19.F.2c] Let A be an $n \times n$ matrix with eigenvalues $\{\lambda_k\}_{k=1}^n$. Then $C_m(A)$ has eigenvalues $\{\prod_{k=1}^m \lambda_{\alpha_k}\}_{\alpha \in \mathcal{Q}_{m,n}}$.

The compounds of companion matrices can be used to study products of roots of polynomials. An excellent example of this can be seen in [4] where there is an extensive description of the compounds of companion matrices as well as some applications. We now examine the absolute values of the entries of the compounds of a companion matrix.

DEFINITION 2.4. Let $\alpha, \beta \in \mathcal{Q}_{m,n}$. β is said to be a forward shift of α if whenever $\alpha_j < n$ for some $1 \leq j \leq m$, there exists $l : 1 \leq l \leq m$ such that $\beta_l = 1 + \alpha_j$.

If $\alpha_m < n$, α has only one forward shift. As an example $\{1, 3, 4\} \in \mathcal{Q}_{3,5}$ has the unique forward shift $\{2, 4, 5\}$. The $\alpha_m = n$ case is more interesting; there are

many forward shifts. To explore this case in more detail, we introduce the following definition.

DEFINITION 2.5. If $\alpha_m = n$ and β is a forward shift of α , we call the integer s the *excluded value* of (α, β) if $s + 1 = \beta_k$ for some k but $s \neq \alpha_j$ for all $j : 1 \leq j \leq m$.

When $\alpha_m = n$, each forward shift of α has a different excluded value. Since any element other than n which is not in α may be an excluded value of a forward shift, there are exactly $n - m + 1$ forward shifts of α when $\alpha_m = n$. As an example $\{1, 2, 5\} \in \mathcal{Q}_{3,5}$ has the three forward shifts: $\{1, 2, 3\}$, $\{2, 3, 4\}$, and $\{2, 3, 5\}$ which have excluded values of zero, three and four respectively. These concepts will allow us to prove the following lemma on the absolute values of the entries of the compounds of the companion matrices.

LEMMA 2.6. Let M_p be the companion matrix of $p = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ and let $\alpha, \beta \in \mathcal{Q}_{m,n}$. Then the (α, β) entry of $C_m(M_p)$ will be zero unless β is a forward shift of α . If β is a forward shift of α and $\alpha_m < n$, then the (α, β) entry of $C_m(M_p)$ will have absolute value one. If β is a forward shift of α and $\alpha_m = n$, then absolute value of the (α, β) entry of $C_m(M_p)$ will be $|a_s|$, where s is the excluded value of (α, β) .

Proof. We now note that if β is not a forward shift of α , then $M_p[\alpha|\beta]$ has at least one row which consists entirely of zeros; so in this case, the (α, β) entry of $C_m(M_p)$ will be zero. For the remainder of the proof, we consider the case where β is a forward shift of α , then the first $m - 1$ rows of $M_p[\alpha|\beta]$ each consist of $m - 1$ zero entries and a single one entry. Each of these one entries are in a different column. If $\alpha_m < n$, the final row will also consist of $m - 1$ zeros and a single one which will be in a different column from all of the other ones in the matrix. In this case, $M_p[\alpha|\beta]$ will be a permutation matrix and the (α, β) entry of $C_m(M_p)$ will have absolute value one. Finally if $\alpha_m = n$, the final row will also consist of $m - 1$ zeros and a single $-a_s$ where s is the excluded value. Since the $-a_s$ entry is in a different column from the ones in the rows above, the absolute value of the (α, β) entry of $C_m(M_p)$ is equal to $|a_s|$. \square

We can now use this lemma to calculate the sums of the absolute entries in each row. Note that for any row α , we need only sum over the entries whose column indices are forward shifts of α since these are the only entries in row α that can be nonzero.

COROLLARY 2.7. Let M_p be the companion matrix of $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ and let $\alpha \in \mathcal{Q}_{m,n}$. Let r_α be the sum of the absolute values of the entries in the α row of $C_m(M_p)$. Then $r_\alpha = 1$ if $\alpha_m < n$, otherwise $r_\alpha = \sum |a_s|$ where the s

is summed over all whole numbers between zero and $n - 1$ which are not equal to α_k for some k . (Note that in this case, the $s = 0$ term is always in the summation).

DEFINITION 2.8. Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be a polynomial. Then a polynomial $q(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1z + b_0$ is called an m -partial polynomial of p if there exists $S \subseteq \{1, 2, \dots, n-1\}$ of cardinality $m-1$ such that $b_k = 0$ if $k \in S$ and $b_k = a_k$ if $k \in \{0, 1, 2, \dots, n-1\} \setminus S$.

In other words, an m -partial polynomial of an n -th degree polynomial $p(z)$ is any n -th degree polynomial that can be formed from $p(z)$ by changing $m-1$ of its non-constant terms of the polynomial to zero. There are $\binom{n-1}{m-1}$ m -partial polynomials of an n -th degree polynomial; these are all distinct if the original polynomial has no zero terms. The sole first partial polynomial of p is p itself. We can now define the higher Cauchy bound. For any polynomial p , let $S_m(p)$ be the set of all m -partial polynomials of p .

EXAMPLE 2.9. Let $p(z) = z^4 + az^3 + bz^2 + cz + d$. Then $S_1(p) = \{p\}$, $S_2(p) = \{z^4 + bz^2 + cz + d, z^4 + az^3 + cz + d, z^4 + az^3 + bz^2 + d\}$, $S_3(p) = \{z^4 + az^3 + d, z^4 + bz^2 + d, z^4 + cz + d\}$ and $S_4(p) = \{z^4 + d\}$.

DEFINITION 2.10. Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be a polynomial and let $1 \leq m \leq n$. We define the m -th Cauchy bound as follows: $\rho_m(p) = \max\{\rho(f) : f \in S_m(p)\}$.

We note that $\rho_1(p) = \rho(p)$, and it follows from Lemma 1.3 that $|a_0|^{\frac{1}{n}} = \rho_n(p) \leq \rho_{n-1}(p) \leq \cdots \leq \rho_2(p) \leq \rho_1(p)$.

We can now prove our main theorem of the paper. Note that the $m = 1$ case is the ordinary Cauchy bound (i.e., Theorem 1.2).

THEOREM 2.11. Let $p(z)$ be a complex polynomial of degree n and let z_1, z_2, \dots, z_n be the n not necessarily distinct zeros of $p(z)$ listed in descending order of modulus. Then $(\prod_{k=1}^m |z_k|)^{\frac{1}{m}} \leq \rho_m(p)$.

Proof. We note that for any positive t , $t^{-m} \prod_{k=1}^m |z_k|$ is an eigenvalue of $C_m(M_{q_t})$. If we choose t to be the smallest positive number for which $\|C_m(M_{q_t})\|_{\infty} \leq 1$, then $\prod_{k=1}^m |z_k| \leq t^m$. By Corollary 2.7, $\|C_m(M_{q_t})\|_{\infty} \leq 1$ if and only if $\sum_{s \in \alpha^c} |a_s| t^{n-s} \leq 1$ for all $\alpha \in \mathcal{Q}_{m,n}$. (Note that here and in what follows, by α^c , we mean $\{0, 1, \dots, n-2, n-1\} \setminus \alpha$). This means that $\|C_m(M_{q_t})\|_{\infty} \leq 1$ if and only if $t^n - \sum_{s \in \alpha^c} |a_s| t^s \geq 0$ for all $\alpha \in \mathcal{Q}_{m,n}$ which will hold if $t \geq \rho_m(p)$. Hence, $(\prod_{k=1}^m |z_k|)^{\frac{1}{m}} \leq \rho_m(p)$. \square

We may get equality in the above theorem. An example of this occurs when $p(z) = z^n - a$, then the only m -partial polynomial of p for any m is p itself. Hence, $\rho_m = |a|^{\frac{1}{n}}$ for all m . In general, since $\rho_n(p) \leq \rho_{n-1}(p) \leq \cdots \leq \rho_2(p) \leq \rho_1(p)$ with strict inequalities if p has no zero coefficients, the bounds on the products of the zeros

given by Theorem 2.11 will be better than those given by the original Cauchy bounds in general. As an example of this, consider the polynomial $p(x) = x^3 - x^2 - 4x + 4$ which has as zeros $(z_1, z_2, z_3) = (2, -2, 1)$. The original Cauchy bound $\rho_1(p)$ is the positive zero of $x^3 - x^2 - 4x - 4$ which is approximately 2.8751 which gives us the inequality $\sqrt{|z_1 z_2|} \leq 2.8751$. We can obtain a better estimate by finding the higher Cauchy bound $\rho_2(p)$ which is the positive zero of $x^3 - 4x - 4$ or approximately 2.3830; hence, we get $\sqrt{|z_1 z_2|} \leq 2.3830$ which cuts the error by more than half.

3. A generalization of Fujiwara's bound. Theorem 2.11 can be used to derive a generalization of the following important root bound of Fujiwara.

THEOREM 3.1. [2] *Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ be a polynomial all of with complex coefficients. If z is any zero of p , then*

$$|z| \leq K_1 \max_{1 \leq j \leq n} |a_j|^{\frac{1}{n-j}},$$

where K_1 is the unique positive zero of the polynomial $q(z) = z^n - \sum_{k=0}^{n-1} z^k$. (K_1 is also the unique positive zero of the polynomial $(z-1)q(z) = z^{n+1} - 2z^n + 1$).

It will be useful to calculate the higher Cauchy bounds of the polynomial $q(z) = z^n - \sum_{k=0}^{n-1} z^k$.

LEMMA 3.2. *Let $q(z) = z^n - \sum_{k=0}^{n-1} z^k$ and let $1 \leq m \leq n$. Then $\rho_m(q)$ is the unique positive zero of $q_m(z) = z^n - \sum_{k=0}^{n-m} z^k$.*

Proof. Let p_1 and p_2 be two polynomials in $S_m(q)$ which differ in at most two terms; hence, there exists a polynomial $r(z)$ such that $p_1(z) = r(z) + z^a$ and $p_2(z) = r(z) + z^b$. We note that if $a > b$, then $p_1(z) - p_2(z) = z^a - z^b > 1$ if $z > 1$. It follows that $\rho(p_1) < \rho(p_2)$. It follows that the polynomial in $S_m(q)$ with the smallest degree nonzero nonleading coefficients will have the largest Cauchy bound; this polynomial is q_m from which this result follows. \square

We also note that we can generalize Lemma 1.3 to the higher Cauchy bounds.

LEMMA 3.3. *Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ and $q(z) = z^n + b_{n-1}z^{n-1} + \cdots + b_1z + b_0$ be two n -th degree polynomials with complex coefficients. If $|a_k| \leq |b_k|$ for all $0 \leq k \leq n-1$, then $\rho_m(p) \leq \rho_m(q)$ for all m such that $1 \leq m \leq n$.*

The result follows from applying Lemma 1.3 to the corresponding partial polynomials of p and q .

We are now ready to prove our generalization of the Fujiwara bound,

THEOREM 3.4. *Let $p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0$ and let z_1, z_2, \dots, z_n be the n not necessarily distinct zeros of $p(z)$ listed in descending order of modulus*

and let m be an integer less than or equal to n . Then

$$\prod_{j=1}^m |z_k|^{\frac{1}{m}} \leq K_m \max_{0 \leq j \leq n-1} |a_j|^{\frac{1}{n-j}},$$

where K_m is the unique positive zero of the polynomial $q_m(z) = z^n - \sum_{k=0}^{n-m} z^k$. (K_m is also the unique positive zero of the polynomial $(z-1)q_m(z) = z^{n+1} - z^n - z^{n-m+1} + 1$).

Proof. Let $t = \max_{0 \leq j \leq n-1} |a_j|^{\frac{1}{n-j}}$. Let $f(z) = t^{-n}p(tz)$, then $f(z)$ is a monic polynomial all of whose coefficients have modulus less than or equal to one. Let $q(z) = z^n - \sum_{k=0}^{n-1} z^k$. It follows from Lemma 3.3 that $\rho_m(f) \leq \rho_m(q) = \rho(q_m) = K_m$ for all m such that $1 \leq m \leq n$. Since the zeros of $f(z)$ listed in descending order of modulus are $\frac{z_1}{t}, \frac{z_2}{t}, \dots, \frac{z_n}{t}$, it follows from Theorem 2.11 that $\prod_{j=1}^m |z_k|^{\frac{1}{m}} \leq tK_m = K_m \max_{0 \leq j \leq n-1} |a_j|^{\frac{1}{n-j}}$. \square

We note that the $m = 1$ special case of this result is exactly Fujiwara's bound.

Acknowledgment. The first author would like to acknowledge the support of NSERC discovery grant number 400550. The second author would like to thank the Department of Mathematics and Statistics of the University of Guelph for its hospitality during his stay there. The authors would like to thank the referee for many suggestions which improved this paper.

REFERENCES

- [1] A.-L. Cauchy. Sur la résolution des équations numériques et sur la théorie de l'élimination. *Oeuvres complètes: Series 2, Vol. 9.*, pp. 87–161. First published in 1891. Cambridge University Press, 2009.
- [2] M. Fujiwara. Über die obere Schranke des absoluten Betrages der Wurzeln einer algebraischen Gleichung. *Tôhoku. Math. J.*, 10:167–171, 1916.
- [3] A.W. Marshall, I. Olkin, and B.C. Arnold. *Inequalities: Theory of Majorization and Its Applications*, second edition. Springer, New York, 2011.
- [4] O. Moussa. An inequality about the largest roots of a polynomial. *Applicable Algebra in Engineering, Communication and Computing*, 8:497–504, 1997.
- [5] Q.I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials*. London Mathematical Society Monographs, Clarendon Press, Oxford, 2002.