PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 131, Number 3, Pages 725–730
S 0002-9939(02)06623-6
Article electronically published on July 25, 2002

THE NUMERICAL RADIUS AND BOUNDS FOR ZEROS OF A POLYNOMIAL

YURI A. ALPIN, MAO-TING CHIEN, AND LINA YEH

(Communicated by David Sharp)

ABSTRACT. Let $p(t)$ be a monic polynomial. We obtain two bounds for zeros of $p(t)$ via the Perron root and the numerical radius of the companion matrix of the polynomial.

Consider the monic polynomial

$$p(t) = t^n + a_{n-1}t^{n-1} + a_{n-2}t^{n-2} + \cdots + a_1 t + a_0. \tag{1}$$

Fujii and Kubo [2] derived from the Buzano inequality a bound for the zeros of $p(z) = 0$:

$$|z| \leq \cos\frac{\pi}{n+1} + \frac{1}{2}\Big(|a_{n-1}| + (|a_{n-1}|^2 + |a_{n-2}|^2 + \cdots + |a_0|^2)^{1/2}\Big). \tag{2}$$

Let $A = (a_{ij}) \in M_n$ be a nonnegative matrix. The famous Perron-Frobenius theorem shows that A has a nonnegative real eigenvalue $\rho(A)$ such that $\rho(A) \geq |\lambda|$, for every eigenvalue λ of A. The eigenvalue $\rho(A)$ is called the Perron root of A. By the directed graph of A, we mean as usual the directed graph with n vertices $1, 2, \ldots, n$ such that there is a directed arc from i to j if $a_{ij} > 0$. Let $s_i = \sum_{j=1}^n a_{ij}$. The mean weight of a directed path σ of the sequential directed arcs $i_1, i_2, \ldots, i_{k+1}$ in the directed graph is the geometric mean $w_\sigma = (s_{i_1}s_{i_2}\cdots s_{i_k})^{1/k}$. Alpin [1] obtained that

$$\rho(A) \leq \max_\sigma w_\sigma,$$

where σ runs over all simple contours in the directed graph of A.

Remark 1. It is known [3, 8.1.18] that the spectral radius of any matrix A is less than or equal to the spectral radius (=Perron root) of the nonnegative matrix, whose (i, j)-element is the absolute value of the (i, j)-element of A.

Received by the editors December 14, 1999 and, in revised form, October 24, 2001.
2000 *Mathematics Subject Classification.* Primary 15A60, 26C10.
Key words and phrases. Perron root, numerical range, numerical radius, companion matrix.
The work of the second author was supported by the National Science Council of the Republic of China.

Relating to the polynomial (1), we consider the companion matrix of the polynomial which is the n-by-n matrix

$$A = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & -a_0 \\ 1 & 0 & \cdots & \cdots & \vdots \\ 0 & 1 & \ddots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \ddots & -a_{n-2} \\ 0 & \cdots & \cdots & 1 & -a_{n-1} \end{pmatrix}. \tag{3}$$

It is known that the characteristic polynomial of the companion matrix A is $p(t)$. Applying Alpin's result and Remark 1 to the companion matrix (3), the following bound is immediate.

Theorem 1. *Let $p(t)$ be the polynomial defined by* (1) *and $p(z) = 0$. Then*

$$|z| \le \max_{1\le k\le n} \Big((1+|a_{n-1}|)(1+|a_{n-2}|)\cdots(1+|a_{n-k}|)\Big)^{1/k}.$$

Remark 2. Let us denote $FK(p)$ as the bound at the right-hand side of (2), $A(p)$ the bound in Theorem 1, and $C(p)$ the maximum absolute value of the coefficients of p. Suppose we can find a sequence of polynomials $p_n(t)$ such that $C(p_n)$ approaches infinity as $n \to \infty$, and $A(p_n)$ is bounded for all n, for example by considering the polynomials

$$p_n(t) = t^n + t + n,\ \ n = 2, 3, \ldots .$$

Then $FK(p_n)$ goes to infinity, and $A(p_n) = 2^{1//2}(1+n)^{1/n}$ converges to 1, $A(p_n)$ is an "infinity sharper" than $FK(p_n)$. Furthermore, we compare the bounds with the well known estimate (cf. [6, p. 65])

$$|z| \le 2\max\{|a_{n-k}|^{1/k} : 1 \le k \le n\}.$$

For the polynomial p_n, this estimate $2n^{1/n}$ is closely comparable to $A(p_n)$.

Let $A \in M_n(\mathbf{C})$. The numerical range of A is the set of complex numbers

$$W(A) = \{x^*Ax : x \in \mathbf{C}^n, |x| = 1\}.$$

By the well known Toeplitz-Hausdorff theorem, the numerical range $W(A)$ is a convex subset of $\mathbf{C}$ containing the eigenvalues of A. The numerical radius $w(A)$ of A is the largest modulus of any point in $W(A)$. (For references on the properties of the numerical range, see, for instance [4, Chapter 1].) From the numerical range standpoint, we investigate the numerical radius of the companion matrix of $p(t)$. The result is applicable to give a bound for the zeros of $p(t)$.

Let $A = (a_{ij}) \in M_n$, and let

$$R_i(A) = \sum_{j=1, j\ne i}^{n} |a_{ij}| \text{ and } g_i(A) = (R_i(A) + R_i(A^*))/2, \quad 1 \le i \le n.$$

The Geršgorin disc thorem (cf. [3, 6.1.1]) shows that the spectrum $\sigma(A)$ of A is contained in Geršgorin discs:

$$\sigma(A) \subset \bigcup_{i=1}^{n} \{z : |z - a_{ii}| \le R_i(A)\}. \tag{4}$$

Brauer [3, 6.4.7] improved the Geršgorin theorem (4), and obtaining a smaller inclusion region, the so-called ovals of Cassini:

$$\sigma(A) \subset \bigcup_{i,j=1, i\neq j}^{n} \{z : |z - a_{ii}|\ |z - a_{jj}| \leq R_i(A)R_j(A)\}. \tag{5}$$

Johnson [5] used Geršgorin disc theorem (4) and obtained a Geršgorin inclusion region for the numerical range of a matrix that

$$W(A) \subset \text{ conv } \Big\{ \bigcup_{i=1}^{n} \{z : |z - a_{ii}| \leq g_i(A)\} \Big\},$$

where "conv" stands for the convex hull of a set. It seems that we may use the inclusion (5) and make the following guess that

$$W(A) \subset \text{ conv } \Big\{ \bigcup_{i,j=1, i\neq j}^{n} \{z : |z - a_{ii}|\ |z - a_{jj}| \leq g_i(A)g_j(A)\} \Big\}. \tag{6}$$

However, the matrix $A = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ gives a counterexample: the inclusion (6) is false. In the following, we are able to extend Johnson's result [5] to a containment region which is the convex hull of the same type of Cassini ovals.

Theorem 2. *Let $A = (a_{ij}) \in M_n$. Then*

$$W(A) \subset G(A),$$

where

$$G(A) = \text{conv } \Big\{ \bigcup_{i,j=1, i\neq j}^{n} \{z : |z - a_{ii}|\ |z - a_{jj}| \leq \big(g_i^2(A)g_j^2(A) + |a_{ii} - a_{jj}|^2\big)^{1/2}\} \Big\}.$$

Proof. Since $W(\beta A) = \beta W(A)$ and $G(\beta A) = \beta G(A)$ for every complex number β, we may assume without loss of generality that the Frobenius norm

$$\|A\|_2 = \Big(\sum_{i,j=1}^{n} |a_{ij}|^2 \Big)^{1/2} = 1.$$

We claim first that if $G(A)$ is contained in the right half plane, then $W(A)$ is also contained in the right half plane. Observe that $G(A)$ enjoys two numerical range properties:

$$G(A + \lambda I) = G(A) + \lambda \text{ and } G(e^{i\theta}A) = e^{i\theta}G(A).$$

Consider a typical curve in the region $G(A)$:

$$|z - a_{ii}|\ |z - a_{jj}| = \alpha_{ij}, \tag{7}$$

where $\alpha_{ij} = \big(g_i^2(A)g_j^2(A) + |a_{ii} - a_{jj}|^2\big)^{1/2}$. Assume $a_{ii} = p_1 + \sqrt{-1}p_2$ and $a_{jj} = q_1 + \sqrt{-1}q_2$, where $p_1, p_2, q_1, q_2 \in \mathbf{R}$. The equation of rectangular coordinates of (7) becomes

$$\Big((x - p_1)^2 + (y - p_2)^2\Big)\Big((x - q_1)^2 + (y - q_2)^2\Big) = \alpha_{ij}^2. \tag{8}$$

Suppose $p_1 \leq q_1$. The case $q_1 \leq p_1$ can be treated in a similar way. It is obvious that a_{ii} belongs to the region bounded by the curve (8) which is contained in $G(A)$. By the assumption that $G(A)$ lies in the right half plane, we obtain that $p_1 > 0$, and the intersections of the curve (8) with the the horizontal line $y = p_2$ are contained in the right half plane. Therefore, the roots of the function

$$f(x) = (x - p_1)^2\Big((x - q_1)^2 + (p_2 - q_2)^2\Big) - \alpha_{ij}^2$$

are positive. We compute that

$$f'(x) = 2(x - p_1)\Big((x - q_1)^2 + (p_2 - q_2)^2\Big) + 2(x - q_1)(x - p_1)^2.$$

Then $f'(x) \le 0$ for $x \le p_1$, and thus f is decreasing on $(-\infty, 0)$. Now since f has no negative roots and $f(x)$ takes positive values for negative sufficient large numbers x, it follows that $f(0) > 0$. The condition $f(0) > 0$ is equivalent to the inequality

$$p_1^2 q_1^2 + p_1^2(p_2 - q_2)^2 > g_i^2(A) g_j^2(A) + |a_{ii} - a_{jj}|^2. \tag{9}$$

From (9) and the assumption that $\|A\|_2 = 1$, we have

$$p_1^2 q_1^2 > g_i^2(A) g_j^2(A) + (p_1 - q_1)^2 + (1 - p_1^2)(p_2 - q_2)^2 \ge g_i^2(A) g_j^2(A),$$

and thus

$$p_1 q_1 > g_i(A) g_j(A). \tag{10}$$

The condition (10) gives a criterion for no nonnegative solutions of the following equation:

$$|x - p_1|\ |x - q_1| = g_i(A) g_j(A).$$

Hence the region

$$\{z : |z - p_1|\ |z - q_1| \le g_i(A) g_j(A)\} \tag{11}$$

is contained in the right half plane.

Let A_H and A_K be the Hermitian parts of A such that $A = A_H + iA_K$. Then

$$W(A) \subset W(A_H) + iW(A_K).$$

By Brauer's result (5),

$$\sigma(A_H) \subset \bigcup_{i,j=1, i\ne j}^{n} \{z : |z - \mathbf{Re}\ a_{ii}|\ |z - \mathbf{Re}\ a_{jj}| \le g_i(A) g_j(A)\}. \tag{12}$$

By (11) and (12), $\sigma(A_H)$ is contained in the right half plane. Since $\mathbf{Re}W(A) = W(A_H)$ is the convex hull of the $\sigma(A_H)$, it follows that $W(A)$ is contained in the right half plane, and this proves the claim.

Now suppose $\alpha \in W(A)$. Then $0 \in W(A - \alpha I)$. If $0 \notin G(A - \alpha I)$, then there would exist θ such that $G(e^{i\theta}(A - \alpha I)) = e^{i\theta} G(A - \alpha I)$ is contained in the right half plane. By the claim, $W(e^{i\theta}(A - \alpha I))$ is contained in the right half plane. But then $0 \notin e^{i\theta} W(A - \alpha I) = W(e^{i\theta}(A - \alpha I))$, a contradiction. Thus $0 \in G(A - \alpha I) = G(A) - \alpha$, and therefore $\alpha \in G(A)$. □

As a consequence of Theorem 2, we locate a bound for the zeros of $p(t)$ defined in (1).

Theorem 3. *Let $p(t)$ be the polynomial* (1) *and $p(z) = 0$, and A be the matrix defined by* (3)*. Let $\alpha = \max_{1\le i<n} g_i(A)$. Then*

$$|z| \le |a_{n-1}|/2 + \Big(|a_{n-1}|^2/4 + (\alpha^2 g_n^2(A) + |a_{n-1}|^2)^{1/2}\Big)^{1/2}.$$

Proof. The possible Cassini ovals in Theorem 2 are

$$|z| \le (g_i(A)g_j(A))^{1/2},\ 1 \le i \ne j < n, \tag{13}$$

and

$$|z|\,|z + a_{n-1}| \le (g_i^2(A)g_n^2(A) + |a_{n-1}^2|)^{1/2},\ 1 \le i < n. \tag{14}$$

From (13), we have $|z| \le \alpha$. From (14), we have

$$|z| \le |a_{n-1}|/2 + \Big(|a_{n-1}|^2/4 + (\alpha^2 g_n^2(A) + |a_{n-1}|^2)^{1/2}\Big)^{1/2}.$$

The inclusion follows from the facts that

$$\alpha \le |a_{n-1}|/2 + \Big(|a_{n-1}|^2/4 + (\alpha^2 g_n^2(A) + |a_{n-1}|^2)^{1/2}\Big)^{1/2}$$

and $\sigma(A) \subset W(A)$. □

The quantities $g_i(A)$ appearing in Theorem 3 are easy to evaluate explicitly in terms of the coefficients of $p(z)$:

$$g_1(A) = (1 + |a_0|)/2, g_2(A) = (2 + |a_1|)/2, \dots , g_{n-1}(A) = (2 + |a_{n-2}|)/2$$

and

$$g_n(A) = (1 + S)/2 \ \text{ where } \ S = \sum_{k=0}^{n-2} |a_k|.$$

Thus $\alpha = \max_i g_i(A)$ is either $(1+S)/2$ or $\beta = \max_{1\le k\le n-2}(2+|a_k|)/2$, depending on which is greater. Theorem 3 may then be stated as follows: any root z of p must satisfy

$$|z| \le |a_{n-1}|/2 + \Big(|a_{n-1}|^2/4 + (\alpha^2((1 + S)/2)^2 + |a_{n-1}|^2)^{1/2}\Big)^{1/2}, \tag{15}$$

where $\alpha = \max\{\beta, (1 + S)/2\}$.

On the other hand, we may more directly estimate the roots based on Brauer's containment of $\sigma(A)$. Brauer's theorem (5) applied to the companion matrix A tells us that any root z of p satisfies one of the inequalities

$$|z|^2 \le R_i(A)R_j(A) \quad (i < j < n)$$

or

$$|z(z + a_{n-1})| \le R_n(A)R_j(A) \quad (j < n).$$

Now $R_1(A) = |a_o|, R_j(A) = 1 + |a_{j-1}|, 1 < j < n$, and $R_n(A) = 1$. Let M and m denote the largest and second largest numbers among

$$|a_0|, 1 + |a_1|, \dots , 1 + |a_{n-2}|.$$

Clearly a root z must satisfy either $|z| \le (Mm)^{1/2}$ or $|z|(|z| - |a_{n-1}|) \le M$, i.e.,

$$|z| \le \max\{(Mm)^{1/2}, (|a_{n-1}| + (|a_{n-1}|^2 + 4M)^{1/2})/2\}. \tag{16}$$

Applying Brauer's theorem to the adjoint of the comapnion matrix A gives us corresponding inequality

$$|z| \le \max\{1, (|a_{n-1}| + (|a_{n-1}|^2 + 4S)^{1/2})/2\}. \tag{17}$$

In fact, it turns out that (16) and (17), taken together, are stronger than the inequality (15) of Theorem 3, unless S is quite small. The right-hand side of (15) is no less than

$$Q = |a_{n-1}|/2 + \Big(|a_{n-1}|^2/4 + (((1+S)/2)^4 + |a_{n-1}|^2)^{1/2}\Big)^{1/2}$$

and

$$(|a_{n-1}| + (|a_{n-1}|^2 + 4S)^{1/2})/2 \le Q.$$

Thus (15) cannot be stronger than (17) unless the right-hand side of (15) is smaller than 1. Then (15) is also stronger than (16) since $M, m \ge 1$ when $n \ge 4$. But if the right-hand side of (15) is less than 1, then we have $\alpha(1+S)/2 < 1$ with $\alpha = (2 + |a_k|)/2 > (1+S)/2$, putting strong restrictions on the size of S.

Remark 3. We give a comparison among the numerical bounds obtained by Fujii-Kubo (2), Theorem 1, Theorem 3, and (17). Consider the polynomial $p(t) = t^5 + 2t^4 + 1$. The numerical bounds obtained by Fujii-Kubo (2), Theorem 1, Theorem 3, and (17) are approximated to 2.984, 3.000, 2.750 and 2.414, respectively. (17) is the best bound among them.

On the other hand, consider the polynomial $p(t) = t^5 + 0.01t^4 + 0.01t + 0.01$ with small coefficients. The numerical bounds obtained by Fujii-Kubo (2), Theorem 1, Theorem 3, and (17) are $0.719, 1.010, 0.716$, and 1.000 respectively. In this case, Theorem 3 is the best.

Acknowledgements

The authors are grateful to the referee for his/her helpful suggestions, especially for providing the estimates (16) and (17) based on Brauer's theorem. The referee also pointed out that the new bound for the polynomial roots is usually stronger than Theorem 3, unless the coefficients of the polynomial are small.

References

[1] Y. A. Alpin, Bounds for the Perron root of a nonnnegative matrix involving the properties of its graph, Math. Notes 58(1995), 1121-1123. MR **97a:**15032

[2] M. Fujii and F. Kubo, Buzano's inequality and bounds for roots of algebraic equations, Proc. Amer. Math. Soc. 117(1993), 359-361. MR **93d:**47014

[3] R. A. Horn and C. R. Johnson, *Matrix analysis*, Camb. Univ. Press, New York, 1990. MR **91i:**15001

[4] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Camb. Univ. Press, New York, 1991. MR **92e:**15003

[5] C. R. Johnson, A Geršgorin inclusion set for the field of values of a finite matrix, Proc. Amer. Math. Soc. 41(1973), 57-60. MR **47:**6738

[6] M. Marden, *Geometry of polynomials*, Math. Surveys and Monographs, No. 3, Amer. Math. Soc., 1989. MR **37:**1562 (review of 2nd edition)

Department of Mathematics and Mechanics, Kazan State University, Kazan, Russia, 420008

E-mail address: `Yuri.Alpin@ksu.ras.ru`

Department of Mathematics, Soochow University, Taipei, Taiwan 11102

E-mail address: `mtchien@math.scu.edu.tw`

Department of Mathematics, Soochow University, Taipei, Taiwan 11102

E-mail address: `yehlina@math.scu.edu.tw`