The Journal of National Chiao Tung University
Vol. 4, April 1978
pp. 39-44

遞歸多項式零點之分布區域

On the Distribution of Zeros of Some Special Classes of Polynomials

倪維城 Wei-Chen Ni

Institute of Applied Mathematics, N. C. T. U.

(Received March 6, 1978)

Abstract — This paper primarily investigates on the distribution of zeros of polynomials under certain conditions. The idea is to transform such a problem of zeros into a problem of eigenvalues of some related matrix, and apply the Gerschgorin theorem or its generalization to produce upper bounds for the zeros concerned.

I. Introduction

The study of distribution of polynomials' zeros dates from the end of the eighteenth century, when the geometric representation of complex numbers was introduced into mathematics. Since then, many mathematicians have contributed to the growth of this subject. A large amount of articles in this area is, however, centered on the study of the zeros of a polynomial f(z) as functions of different parameters, which are allowed to vary within some prescribed region. The parameters considered are usually the coefficients of f(z) itself, or the coefficients of some related polynomial or polynomials, and so forth. Theory on the relationships between the coefficients of a polynomial and the distribution of its zeros, when applied to special functions [3] widely used in physics, does not provide very satisfactory results. It is the purpose of the present paper to discuss how we can apply the Gerschgorin theorem or its generalization to a particular matrix related to a given polynomial in order to obtain a bound for the moduli of the eigenvalues of the matrix. This in turn indicates how the zeros of the polynomial are distributed over the complex plane. Referring to this idea, we are able to furnish upper bounds for the moduli of zeros of polynomials satisfying certain conditions, including various recurrence relations. The types of recurrence relation considered will, in particular, be satisfied by many special functions, such as the polynomials by the name of Hermite, Lagrange, Laguerre, Legendre and Chebyshev etc.

II. Discussion and Results

In this paper, we shall adopt the following notations and definitions. For the sake of convenience, we set $N=\{1, 2,..,n\}$ and $N_i=\{ j\in N: j\neq i\}$. Let $A=(a_{ij})_{nxn}$ be an nxn matrix over the set of complex numbers $\mathbb{C}$, and $x=(x_1,...,x_n)$ be any vector in R^n with positive components. We define $G_i(x)=\{ z\in C: |z-a_{ii}|\leqslant \frac{1}{x_i}\sum_{j\in N_i}|a_{ij}|x_j\}$ and $G(x)=\bigcup_{i=1}^{n} G_i(x)$. In particular, when $x=(1,1,..,1)$, we use the notations G_i and G to represent $G_i(x)$ and $\bigcup_{i=1}^{n} G_i(x)$ respectively. G_i is called the i^{th} Gerschgorin disk of A and G is called the Gerschgorin configuration of A. Furthermore, we shall denote the set $\{ B=(b_{ij})_{nxn} : b_{ij}\in\mathbb{C};\ b_{ii}=a_{ii},$ for $1\leqslant i\leqslant n;$ and $|b_{ij}|=|a_{ij}|,$ for $1\leqslant i,j\leqslant n\}$ by Ω_A and the set $\{ B=(b_{ij})_{nxn}: b_{ij}\in\mathbb{C}; b_{ii}=a_{ii},$ for $1\leqslant i\leqslant n;$ and $|b_{ij}|\leqslant|a_{ij}|,$ for $1\leqslant i,j\leqslant n\}$ by $\hat{\Omega}_A$. A matrix A is called a Jacobi matrix if $a_{ij}=0$ whenever $|i-j|\geqslant 2$, i.e. if A is tri-diagonal. We shall also use the notation $A(i|j)$ to represent the matrix obtained from A by deleting the i^{th} row and the j^{th} column of A.

As a prelimary result, we state the following well-known theorem:

Theorem 1 (Gerschgorin)

Let $A=(a_{ij})_{nxn}$ be an arbitrary nxn complex matrix. Then, all the eigenvalues λ of A lie in the union of the disks $|z-a_{ii}| \leq \sum_{j \in N_i^J} |a_{ij}|$, $1 \leq i \leq n$. (i.e., all the eigenvalues of A lie in the Gerschgorin configuration of A).

A generalization of Theorem 1 is : [5, thm. 6].

Theorem 2 Let A be an arbitrary nxn matrix. Then, the set of eigenvalues of all $B \in \hat{\Omega}_A$ is $\bigcap_{x>0} G(x)$.

Let p(z) be a monic polynomial function with complex coefficients, i.e. $p(z)=z^n+a_{n-1}z^{n-1}+...+a_o$, where $a_i \in \mathbb{C}_r$ for $0 \leq i \leq n-1$. The matrix

$$C=\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & 0 \\ 0 & \cdots & & 0 & 1 \\ -a_o & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

is known as the companion matrix of p(z). By Theorem 2, the eigenvalues of C are contained in $\bigcap_{x>0} G(x)$. Since the zero of p(z) are simply the eigenvalues of C, we have the following theorem established.

Theorem 3 The set of zeros of a monic polynomial p(z) is a subset of $\bigcap_{x>0} G(x)$.

Theorem 3 actually generalizes many classical results concerning bound for zeros of a polynomial. As an example, [2, thm. 27.2] states below becomes a corallary of this theorem.

Corollary 4 All the zeros of $f(z)=a_n z^n+...+a_o$, where $a_n \neq 0$, lie in the circle $|z| < 1+\max_k\left|\frac{a_k}{a_n}\right|$, k=0,1,2,..n-1.

Proof We first transform f(z) into a polynomial g(z) with leading coefficient equal to 1. An application to C^T, the transpose of the companion matrix C of g(z), yields the desired result.

Next we turn to the study on distribution of zeros of a class of polynomials satisfying some recurrence relation. In particular, many special functions fall into this category. First of all, let us consider polynomials satisfying the recurrence relation

$$b_{i+1,i}f_i(z)+(b_{i,i}-z)f_{i-1}(z)+b_{i-1,i}f_{i-2}(z)+....+b_{1,i}=0. \quad \text{for } 1 \leq i \leq n, \text{ where } f_o=1 \text{ and } f_{-1}=0 \qquad (1)$$

We prove the following two theorems for this type of polynomials.

Theorem 5 Let $A=(a_{ij})_{nxn}$ be an nxn matrix over $\mathbb{C}$ such that $a_{ij}=0$ for i>j-1. The set of zeros of all f_n satisfying (1) for $1 \leq i \leq n$ with $|b_{ij}| \leq |a_{ij}|$ for $1 \leq i,j \leq n$, is precisely $\bigcap_{x>0} G(x)$.

Proof Recurrence relation (1) generates a system of n linear equations as follows:

$$b_{n+1,n}f_n+(b_{n,n}-z)f_{n-1}+b_{n-1,n}f_{n-2}+\cdots\cdots+b_{1,n}=0$$

$$b_{n,n-1}f_{n-1}+(b_{n-1,n-1}-z)f_{n-2}+\cdots\cdots+b_{1,n-1}=0$$

$$\vdots$$

$$b_{3,2}f_2+(b_{2,2}-z)f_1+b_{1,2}=0$$

$$b_{2,1}f_1+(b_{1,1}-z)=0$$

Solving for f_n by means of the well-known Cramers rull and using properties of determinant,

$$f_n=(-1)^n\,\frac{\det\begin{pmatrix} b_{n,n}-z & b_{n-1,n} & \cdots & b_{1,n}\\ b_{n,n-1} & \ddots & & \vdots\\ & \ddots & \ddots & \vdots\\ \bigcirc & & b_{2,1} & b_{1,1}-z\end{pmatrix}}{\det\begin{pmatrix} b_{n+1,n} & b_{n,n}-z & \cdots & b_{2,n}\\ & b_{n,n-1} & \ddots & \vdots\\ & & \ddots & \vdots\\ \bigcirc & & b_{3,2} & b_{2,2}-z\\ & & & b_{2,1}\end{pmatrix}}$$

$$=a\det(B'-Iz)$$

where

$$B'=\begin{pmatrix} b_{n,n} & b_{n-1,n} & \cdots & b_{1,n}\\ b_{n,n-1} & \ddots & & \vdots\\ \bigcirc & \ddots & b_{2,1} & b_{1,1}\end{pmatrix}$$

and a is a constant since the matrix in the denominator is upper triangular. Hence the zeros of f_n are precisely the eigenvalues of B'. Since B' is similar to the matrix

$$B=\begin{pmatrix} b_{1,1} & \cdots & \cdots & b_{1,n}\\ b_{2,1} & \ddots & & \vdots\\ \bigcirc & \ddots & b_{n,n-1} & b_{n,n}\end{pmatrix}$$

in $\hat{\Omega}_A$ they possess the same eigenvalues. Therefore the zeros of f_n are contained in $\bigcap_{x>0} G(x)$. On the other hand, for $z\in\bigcap_{x>0} G(x)$, there exists a matrix B in $\hat{\Omega}_A$ of which z is an eigenvalue. Reversing the arguments above and letting $b_{n+1,n}$ be an arbitary nonzero complex number, we have that z is a zero of f_n satisfying a recurrence relation of the form (1).

Theorem 6 The set of zeros of $\{f_i\}_{i=1}^{n}$ which satisfies the recurrence relation

$$a_i\, f_i(z)+(b_i-z)\, f_{i-1}(z)+c_i f_{i-2}(z)=0 \tag{2}$$

is contained in $G=\bigcup_{i=1}^{n} G_i$

Proof Recurrence relation (2) generates the following system of linear equations:

$$\begin{aligned} &a_nf_n+(b_n-z)f_{n-1}+c_nf_{n-2}=0\\ &\quad a_{n-1}f_{n-1}+(b_{n-1}-z)f_{n-2}+c_{n-1}f_{n-3}=0\\ &\qquad\vdots\\ &\qquad\quad a_2f_2+(b_2-z)f_1+c_2=0\\ &\qquad\quad\quad a_1f_1+(b_1-z)=0\end{aligned}$$

Again, applying Cramer's rule and properties of determinant,

$$f_n = k\ \det(L - zI)$$

where k is a constant and

$$L = \begin{pmatrix} b_n & c_n & & O \\ a_{n-1} & \ddots & \ddots & \\ & \ddots & \ddots & c_2 \\ O & & a_1 & b_1 \end{pmatrix}.$$

Hence the zeros of f_n are just the eigenvalues of L, all of which lie in G by theorem 1. Similarly, the zeros of f_{n-1} are the eigenvalues of $L(1 \mid 1)$, which we clearly contained in G also. The proof is completed by a repetition of the same argument for the zeros of f_i, $1 \leqslant i \leqslant n-2$.

It is worth noting that Theorem 6 differs from Theorem 5 in that Theorem 5 only look at the zeros of polynomials of degree n satisfying recurrence relation (1), while Theorem 6 takes into account all zeros of the family of polynomials $\{f_i\}_{i=1}^{n}$ satisfying a special case of (1).

Corollary 7 The largest zero of $\{f_i\}_{i=1}^{n}$ in modulus is bounded by $\max_i \{ |a_i| + |b_i| + |c_i| \}$.

Proof This is an easy consequence of the proof of Theorem 6 and the triangle inequality.

Another type of recurrence relation that we are interested in is of the form

$$f_i(z) + (a_i + b_i z) f_{i-1}(z) + (c_i + d_i z) f_{i-2}(z) = 0 .$$

For the case in which $|b_i| = |d_i|$, distribution of the zeros of f_n was investigated in [4]. In this paper, we shall discuss the case when $|b_i| > |d_i|$. Using a similar argument as in Theorem 5, we obtain that the zeros of f_n are simply the zeros of an equation of the form

$$\det \begin{pmatrix} a_{1,1} - b_1\lambda & a_{1,2} - c_1\lambda & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} - b_2\lambda & \cdots & a_{2,n} \\ & \ddots & \ddots & \vdots \\ & & & a_{n-1,n} - c_{n-1}\lambda \\ O & & a_{n,n-1} & a_{n,n} - b_n\lambda \end{pmatrix} = 0 . \qquad (3)$$

If we let $B = \begin{pmatrix} b_1 & c_1 & & \\ & \ddots & \ddots & \\ & & \ddots & c_{n-1} \\ O & & & b_n \end{pmatrix}$, (3) becomes

$$\det(A - \lambda B) = 0 ,$$

where

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & & \vdots \\ & \ddots & \ddots & \vdots \\ O & & a_{n,n-1} & a_{n,n} \end{pmatrix}$$

This is nothing but a generalized eigenvalue problem, i.e. $Ax = \lambda Bx$. Before we investigate on how the zeros of the polynomials $\{f_i\}_{i=1}^{n}$ are distributed are distributed over the complex plane, we prove a lemma.

Lemma 8 Let $A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & \cdots & a_{1,n} \\ a_{2,1} & \ddots & & & \vdots \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ O & & & a_{n,n-1} & a_{n,n} \end{pmatrix}$ and $B = \begin{pmatrix} b_1 & c_1 & & O \\ & \ddots & \ddots & \\ & & \ddots & c_{n-1} \\ O & & & b_n \end{pmatrix}$

where $|b_i| > |c_i|$. For the generalized eigenvalue problem $Ax = \lambda Bx$, all the generalized eigenvalues λ lie in the disk $|z| \leq \max_i \dfrac{\sum_j |a_{ij}|}{|b_i| - |c_i|}$.

Proof Let λ be a generalized eigenvalue and x be its corresponding generalized eigenvector. Normalizing the vector x so that its largest component in modulus is unity, we have that

$$\sum_{j=1}^{n} a_{ij}x_j = \lambda(b_i x_i + c_i x_{i+1}), \quad 1 \leq i \leq n.$$

Suppose the component of x of largest modulus occurs in the r^{th} position, then $|x_r| = 1$ after normalization and when $i=r$,

$$\left|\sum_{j=1}^{n} a_{rj}x_j\right| = |\lambda(b_r x_r + c_r x_{r+1})|.$$

This implies that

$$\sum_{j=1}^{n} |a_{rj}|\,|x_j| \geq |\lambda|\,(|b_r|\,|x_r| - |c_r|\,|x_{r+1}|)$$

$$\geq |\lambda|\,(|b_r| - |c_r|).$$

Since $|b_r| - |c_r|$ is positive by hypothesis, we thus have

$$|\lambda| \leq \frac{\sum_j |a_{rj}| \cdot |x_j|}{|b_r| - |c_r|} \leq \frac{\sum_j |a_{ij}|}{|b_r| - |c_r|}.$$

That is λ lies in the circle centered at the origin with radius being $\dfrac{\sum_j |a_{rj}|}{|b_r| - |c_r|}$

In general, any generalized eigenvalue λ will be in a circle $|z| \leq \dfrac{\sum_j |a_{rj}|}{|b_r| - |c_r|}$ for some r. Hence all generalized eigenvalues for $Ax = \lambda Bx$ will fall inside the circle $|z| \leq \max_{1 \leq i \leq n} \dfrac{\sum_j |a_{ij}|}{|b_i| - |c_i|}$ which completes the proof.

Theorem 9 All zeros of $\{f_i\}_{i=1}^{n}$ satisfying the recurrence relation of the form

$$f_i(z) + (a_i + b_i z)f_{i-1}(z) + (c_i + d_i z)f_{i-2}(z) = 0$$

lie within the circle $|z| \leq \max_{1 \leq i \leq n} \dfrac{1 + |a_i| + |c_i|}{|b_i| - |d_i|}$.

Proof. This theorem follows from Lemma 8 and arguments similar to Theorem 5.

For a general polynomial $f(z) = a_n z^n + a_{n-1}z^{n-1} + \dots + a_o$, when one of the coefficients changes, although zeros of the polynomial changes continuously with it, the amount of change can be large at times. However, in the case when the polynomial satisfies (z), its zero can be estimated by the following theorem.

Theorem 10 If the coefficients a_i, b_i and c_i in the recurrence relation (2) changes to a_i', b_i' and c_i' such that

$|a_i'| < |a_i| + \mathcal{E}_1$, $|b_i'| < |b_1| < |b_i| + \mathcal{E}_2$, $|c_i'| < |c_i| + \mathcal{E}_3$, then the zeros in absolute values of the newly generated polynomials are bounded above by $\max_i \{ |a_i| + |b_i| + |c_i| + \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 \}$.

Proof An application of Corollary 7 to the newly generated family of functions $\{ f_i' \}_{i=1}^{n}$ due to the changes in coefficients, yields the desired result.

It is known [1, p.166] that if L is an nxn real Jacobi matrix and $a_i c_{i-1} > 0$, for i=2,....n, then

(i) All eigenvalues of A are real and simple;

and

(ii) Between any two eigenvalues of A lies exactly one eigenvalue of the principal matrix L_{n-1}=L (n/n) of L.

When the coefficients in the recurrence relaton (2) are real and $a_{i-1}c_i > 0$ for $2 \leq i \leq n$, the above Theorem and similar arguments as in Theorem 5 show that Corollary 7 can be sharpened as:

Theorem 11 If a_i, b_i and c_i are real numbers such that $a_{i-1}c_i > 0$ for $2 \leq i \leq n$, and $\{ f_i \}_{i=1}^{n}$ satisfies.

$$a_i f_i + (b_i \pm z) f_{i-1} + c_i f_{i-2} = 0$$

then

(i) All zeros of f_n are real and simple;

and

(ii) Between any two zeros of f_n lies exactly one zero of f_{n-1}.

Corollary 12 If the coefficients a_i, b_i, and c_i in the recurrence relation (2) are positive real numbers, then the zeros of $\{ f_i \}_{i=1}^{n}$ all lie in the interval [-k, 0] where $k = \max_i \{ |a_i| + |b_i| + |c_i| \}$.

Proof Theorem 11 together with Corollary 7 says that the zeros of $\{ f_i \}_{i=1}^{n}$ all lie in the interval [-k,k] where $k = \max_i \{ |a_i| + |b_i| + |c_i| \}$. Due to the fact that polynomials with positive coefficients have no positive roots, we obtain the desired result.

From the discussion above, we should note that given a polynomial f(z), using different related matrices with eigenvalues being the zeros of the polynomial, we can come up with different bounds for the zeros by means of the Gerschgorin theorem or its generalization. Moreover, some results in this paper are proved under the assumption that x=(1,1,...,1). By making appropriate changes of this positive vector x, are can again obtain various bounds for the zeros concerned. However, the question of how we should choose the related matrix or the vector x in order to produce a best possible bound for the zeros using matrix method is still left open.

References

1. M. Marcus and H. Ming, "A Survey of Matrix Theory and Matrix Inequalities", Boston, Allyn and Bacon, 1964, pp.166-167.
2. M. Marden, "The Geometry of the Zeros of a Polynomial in a Complex Variable", New York, Amer. Math. Soc.,1949, pp.95-108.
3. E. D. Rainville, "Special Functions", New York, The Macmillan,1960, pp.233-245.
4. E. B. Saff and R. S. Varga, "Zero-free parabolic regions for sequences of polynomials", SIAM Jol. of Math. Anal, 7, No.3,344-357 (1976).
5. R. S. Varga, "Minimal Gerschgorin sets", Pacific Jol. of Mathematics, 15, No.2, 719-729 (1965).