

A companion matrix approach to the study of zeros and critical points of a polynomial

BY W.S. CHEUNG AND T.W. NG*

*Department of Mathematics,
The University of Hong Kong,
Pokfulam, Hong Kong.*

E-mail address: wshun@graduate.hku.hk, ntw@maths.hku.hk

Abstract

In this paper, we introduce a new type of companion matrices, namely, D -companion matrices. By using these D -companion matrices, we are able to apply matrix theory directly to study the geometrical relation between the zeros and critical points of a polynomial. In fact, this new approach will allow us to prove quite a number of new as well as known results on this topic. For example, we prove some results on the majorization of the critical points of a polynomial by its zeros. In particular, we give a different proof of a recent result of Gerhard Schmeisser on this topic. The same method allows us to prove a higher order Schoenberg-type conjecture proposed by M. G. de Bruin and A. Sharma.

1 Introduction and D -companion matrices

Let p be a non-linear polynomial of one complex variable. A complex number w is a critical point of p if $p'(w) = 0$. Geometry of polynomials is the study of

AMS Classification: Primary, 30C10; Secondary, 15A42.

Key words and phrases: D -companion matrices, polynomials, zeros, critical points, Schoenberg conjecture, De Bruin and Sharma conjecture, majorization, Gerschgorin's disks, ovals of Cassini.

* The research was partially supported by a seed funding grant of HKU and RGC grant HKU 7020/03P.

zeros of polynomials. In this paper, we focus on a class of problems in geometry of polynomials, namely, those problems concern with the geometrical relation between zeros and critical points of a polynomial. Our main goal is to develop a new approach of applying matrix theory to study these problems. Matrix theory has already been applied successfully to the study of the analytic theory of polynomials through the use of companion matrices (see [2], [13], [16] and [21]). An $n \times n$ matrix whose eigenvalues coincide with the zeros of a degree n polynomial p is called a *companion matrix* of p (here we follow the definition of *companion matrix* given in [16, p.265], which is different from the usual definition). For example, if $p(z) = a_n z^n + \cdots + a_1 z + a_0$, then the $n \times n$ matrix below is a companion matrix of p ,

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & -\frac{a_0}{a_n} \\ 1 & \ddots & & \vdots & -\frac{a_1}{a_n} \\ 0 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & -\frac{a_{n-1}}{a_n} \end{pmatrix}.$$

This matrix is called the *Frobenius matrix* of p and we shall denote it by F_p . Other types of companion matrices can be found in [16, Chapter 8].

Companion matrices are very useful in the study of analytic theory of polynomials. For example, by applying Gerschgorin's Theorem in matrix theory (see Theorem 3.1 in Section 3) directly to the matrix F_p , we immediately obtain the result that all the zeros of $p(z) = \sum_{i=0}^n a_i z^i$ are contained in the two disks

$$\left\{ z \in \mathbb{C} : \left| z + \frac{a_{n-1}}{a_n} \right| \leq 1 \right\}$$

and

$$\left\{ z \in \mathbb{C} : |z| \leq \max \left\{ \left| \frac{a_0}{a_n} \right|, 1 + \left| \frac{a_1}{a_n} \right|, \dots, 1 + \left| \frac{a_{n-2}}{a_n} \right| \right\} \right\}.$$

On the other hand, the usual companion matrices are not very useful in the study of problems related to critical points of a polynomial for the following reason. When one tries to study these problems by applying results from matrix theory, it is natural to find an $(n-1) \times (n-1)$ matrix whose eigenvalues

are the critical points of a given degree n polynomial $p(z) = \sum_{i=0}^n a_i z^i$. We shall call such matrix a *derivative companion matrix*. Of course, $F_{p'}$ is a derivative companion matrix of p . However, this matrix is not very useful when one tries to use it to study the relative locations of zeros and critical points of a polynomial as the entries of $F_{p'}$ are expressed directly in terms of the coefficients a_i . In this paper, we shall introduce a new type of derivative companion matrices which is very suitable for the study of problems concerning zeros and critical points of polynomials. By using these derivative companion matrices, we are able to prove quite a number of new and old results in geometry of polynomials systematically. Our approach is based on the following result which provides a unified way to study many problems in the geometry of polynomials through the direct applications of matrix theory. The proof of it is however very elementary.

Theorem 1.1 . Let $p(z) = a_n \prod_{i=1}^n (z - z_i)$ be a polynomial of degree $n \geq 2$. Let

$$D = \begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_{n-1} \end{pmatrix},$$

I and J be the identity matrix of order $n - 1$ and the $(n - 1) \times (n - 1)$ matrix with all entries equal to 1 respectively. Then the set of all eigenvalues of the $(n - 1) \times (n - 1)$ matrix

$$D \left(I - \frac{1}{n} J \right) + \frac{z_n}{n} J$$

is the same as the set of all critical points of the polynomial p .

As we shall see later, Theorem 1.1 opens up the possibilities of applying matrix theory directly to the study of some problems in geometry of polynomials. In view of Theorem 1.1, we shall have the following

Definition. Let $p(z) = a_n \prod_{i=1}^n (z - z_i)$ be a polynomial of degree $n \geq 2$. Let

$$D = \begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_{n-1} \end{pmatrix},$$

I and J be the identity matrix of order $n - 1$ and the

$(n - 1) \times (n - 1)$ matrix with all entries equal to 1 respectively. Then the $(n - 1) \times (n - 1)$ derivative companion matrix,

$$D\left(I - \frac{1}{n}J\right) + \frac{z_n}{n}J$$

is called a *D-companion matrix* of p .

An immediate consequence of Theorem 1.1 and the Gerschgorin disk theorem is the following result about the geometric locations of critical points relative to the zeros of a polynomial.

Theorem 1.2 . *Let p be a polynomial of degree $n \geq 2$ with zeros z_1, \dots, z_n . For each z_k , define the disks*

$$G_i^k = \left\{ z \in \mathbb{C} : \left| z - \frac{n-1}{n}z_i - \frac{1}{n}z_k \right| \leq \frac{n-2}{n}|z_i - z_k| \right\}, \quad i = 1, \dots, n, i \neq k$$

and if $n \geq 3$, the ovals of Cassini

$$C_{ij}^k = \left\{ z \in \mathbb{C} : \left| z - \frac{n-1}{n}z_i - \frac{1}{n}z_k \right| \left| z - \frac{n-1}{n}z_j - \frac{1}{n}z_k \right| \leq \left(\frac{n-2}{n} \right)^2 |z_i - z_k| |z_j - z_k| \right\},$$

for $i, j = 1, \dots, n, i, j \neq k$.

Then the set of all critical points of p is contained in the domain $G^k = \bigcup_{i=1}^n G_i^k$ as well as the domain $C^k = \bigcup_{\substack{i,j=1 \\ i \neq j}}^n C_{ij}^k$.

Remark. Note that $C^k = \bigcup_{\substack{i,j=1 \\ i \neq j}}^n C_{ij}^k$ is actually a subset of $G^k = \bigcup_{i=1}^n G_i^k$.

2 Geometry of polynomials and main results

Theorem 1.2 is about certain relative geometric locations between the zeros and critical points of a polynomial. The first result of this type is the following basic result in geometry of polynomials.

Gauss-Lucas Theorem : If p is a polynomial of degree n , then all the critical points of p lie inside the closed convex hull of the zeros of p .

The Gauss-Lucas theorem is a result about the general location of all the critical points of a polynomial relative to *all* its zeros. A more refined result was conjectured in 1958 by B. Sendov:

Sendov Conjecture : If $p(z) = a_n \prod_{k=1}^n (z - z_k)$ is a polynomial of degree $n \geq 2$, then each of the disks $|z - z_k| \leq r = \max_i |z_i|$ must contain at least one critical point of p .

The radius $r = \max_i |z_i|$ is best possible as can be seen by considering the polynomial $p(z) = z^n - 1$. This conjecture, also known as ‘‘Ilieff’s Conjecture’’, appeared in 1967 as Problem 4.5 in Hayman’s book, *Research problems in function theory* [7]. Sendov conjecture is still open although attempts to verify this conjecture have led to over 80 papers. To know more about this conjecture, the readers are referred to the survey papers [17] and [20] as well as the two recent books on the analytic theory of polynomials, [16] and [21].

Besides Sendov conjecture, one can also generalize the Gauss-Lucas theorem by studying the majorization of the critical points of a polynomial by its zeros. In fact, very recently Gerhard Schmeisser has succeeded in refining the Gauss-Lucas theorem in this direction (see [18]). In order to state Schmeisser’s result, we shall first introduce the concept of majorization, which is a useful way of comparing the distribution of two sets of real numbers. We shall follow very closely the presentation of majorization given in [18] (see also [14]).

For any vector $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$, we denote by $(x_{[1]}, \dots, x_{[n]})$ a rearrangement of the components of \mathbf{x} such that

$$x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[n]}.$$

Definition. For any two vectors $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ from \mathbb{R}^n , we say that \mathbf{b} weakly majorizes \mathbf{a} , and write this as $\mathbf{a} \prec_w \mathbf{b}$ if

$$\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]} \quad (k = 1, \dots, n).$$

Furthermore, we say that \mathbf{b} (strongly) majorizes \mathbf{a} , and write this as $\mathbf{a} \prec \mathbf{b}$, if, in addition, when $k = n$, we have

$$\sum_{i=1}^n a_{[i]} = \sum_{i=1}^n b_{[i]}.$$

Roughly speaking, $\mathbf{a} \prec \mathbf{b}$ means that the components of \mathbf{a} are less spread out than those of \mathbf{b} .

Very recently, Gerhard Schmeisser has obtained the following result on the majorization of the critical points of a polynomial by its zeros.

Theorem A ([18]). Let p be a polynomial of degree $n \geq 1$ with zeros z_1, \dots, z_n and critical points w_1, \dots, w_{n-1} . Put $w_n = 0$, then

$$(\psi(|w_1|), \dots, \psi(|w_n|)) \prec_w (\psi(|z_1|), \dots, \psi(|z_n|)),$$

for every increasing function $\psi : [0, +\infty) \rightarrow \mathbb{R}$ such that $\psi \circ \exp$ is convex on \mathbb{R} .

Remark. One can take the above ψ to be any of the following functions:

$$\psi(x) \equiv x^p \ (p > 0), \quad \psi(x) \equiv \max\{a, \log x\} \ (a \in \mathbb{R}), \quad \psi(x) \equiv \log x.$$

To see why Theorem A is a refinement of the Gauss-Lucas theorem, let's take $\psi(x) = x$. Then in this case, we have for each $k = 1, \dots, n$,

$$\sum_{i=1}^k |w_{[i]}| \leq \sum_{i=1}^k |z_{[i]}|.$$

In particular it follows from the $k = 1$ case that

$$\max \{|w| : p'(w) = 0\} = |w_{[1]}| \leq |z_{[1]}| = \max \{|z| : p(z) = 0\}.$$

As noticed by Schmeisser in [18], this inequality is actually equivalent to the Gauss-Lucas theorem. First of all, it is clear that the Gauss-Lucas theorem implies that $|w_{[1]}| \leq |z_{[1]}|$. Now suppose we assume that the Gauss-Lucas theorem were false. Then there would exist a polynomial q which has a critical point w lying outside the convex hull \mathcal{H} of the zeros. Clearly, there would exist a circle containing \mathcal{H} in its interior while w lies outside. Let c be the center and r the radius of this circle. Then the moduli of the zeros of $p(z) := q(z + c)$ are bounded by r while p' has a zero of modulus larger than r . This implies that $|w_{[1]}| > |z_{[1]}|$ and we are done.

To prove Theorem A, Schmeisser applies a theorem of de Bruijn and Springer on the zeros of composition-polynomials [5, Theorem 7]. In Section 5, we give a different proof of Theorem A by our matrix theory approach. Besides, we also prove the following new result.

Theorem 2.1 *Let p be a polynomial of degree $n \geq 2$ with zeros z_1, \dots, z_n and critical points w_1, \dots, w_{n-1} . Let p_a be a polynomial of degree $n \geq 2$ with zeros $|z_1|, \dots, |z_n|$. Suppose v_1, \dots, v_{n-1} are the critical points of p_a . If $z_n = 0$, then*

$$(\psi(|w_1|), \dots, \psi(|w_{n-1}|)) \prec_w (\psi(v_1), \dots, \psi(v_{n-1})),$$

for every increasing function $\psi : [0, +\infty) \rightarrow \mathbb{R}$ such that $\psi \circ \exp$ is convex on \mathbb{R} .

Related to Theorem 2.1, we would like to mention the following recent result of R. Pereira which solves a conjecture of Katsoprinakis [10].

Theorem B ([15]). *Let p be a polynomial of degree $n \geq 2$ with zeros z_1, \dots, z_n and critical points w_1, \dots, w_{n-1} . Let p_r be a polynomial of degree $n \geq 2$ with zeros $\operatorname{Re}z_1, \dots, \operatorname{Re}z_n$. Suppose v_1, \dots, v_{n-1} are the critical points of p_r , then*

$$(\psi(\operatorname{Re}w_1), \dots, \psi(\operatorname{Re}w_{n-1})) \prec_w (\psi(v_1), \dots, \psi(v_{n-1})),$$

for every increasing function $\psi : [0, +\infty) \rightarrow \mathbb{R}$ such that $\psi \circ \exp$ is convex on \mathbb{R} .

In view of Pereira's result, one may ask if the condition $z_n = 0$ in Theorem 2.1 is essential. Unfortunately, we are unable to answer this question.

Let's go back to Theorem A. If we take $\psi(x) = x^2$ in Theorem A, then it follows easily from Theorem A that,

$$\sum_{i=1}^{n-1} |w_i|^2 \leq \sum_{i=1}^n |z_i|^2.$$

This inequality is not sharp as we actually have

$$\sum_{i=1}^{n-1} |w_i|^2 \leq \frac{n-1}{n} \sum_{i=1}^n |z_i|^2,$$

which is first proved by de Bruijn and Springer [4] in a more general form. It is certainly possible to further reduce the size of the coefficient of $\sum_{i=1}^n |z_i|^2$ if one is willing to add an extra term to the right hand side of the above inequality of de Bruijn and Springer. The problem is what should be the appropriate term to be added. The following Schoenberg conjecture is related to this problem.

Schoenberg conjecture: Let z_1, \dots, z_n be the zeros of a polynomial p of degree $n \geq 2$ and w_1, \dots, w_{n-1} be the critical points of p . Then

$$\sum_{i=1}^{n-1} |w_i|^2 \leq \frac{1}{n^2} \left| \sum_{i=1}^n z_i \right|^2 + \frac{n-2}{n} \sum_{i=1}^n |z_i|^2,$$

where equality holds if and only if all z_i lie on a straight line.

This conjecture was first posed by Schoenberg [19] in 1986 and was solved very recently by R. Pereira in [15] and S.M. Malamud ([11] and [12]) independently. Although their approaches are different from each other, the underlying ideas of their proofs are similar. Besides, Schoenberg conjecture, Malamud's approach also allows him to obtain a remarkable generalization of the de Bruijn-Springer conjecture. On the other hand, in Pereira's proof, he uses the concept of differentiators of finite dimensional operators which was first introduced by Chandler Davis [6] in 1959. Our work is inspired by Pereira's approach even though we do not use the concept of differentiators.

Schoenberg conjecture asserts that it is possible to bound $\sum_{i=1}^{n-1} |w_i|^2$ by a suitable combination of the terms $\left| \sum_{i=1}^n z_i \right|^2$ and $\sum_{i=1}^n |z_i|^2$. In [3], M. G. de Bruin and A. Sharma proposed the following higher order Schoenberg-type conjecture.

De Bruin and Sharma's conjecture: Let z_1, \dots, z_n be the zeros of a polynomial p of degree $n \geq 2$ and w_1, \dots, w_{n-1} be the critical points of p . If $\sum_{i=1}^n z_i = 0$, Then

$$\sum_{i=1}^{n-1} |w_i|^4 \leq \frac{n-4}{n} \sum_{i=1}^n |z_i|^4 + \frac{2}{n^2} \left(\sum_{i=1}^n |z_i|^2 \right)^2,$$

where equality holds if and only if all z_i lie on a straight line passing through the origin of the complex plane.

This conjecture has been verified for some classes of polynomials (see [3] and [10]) but as far as we know it had remained open. In Section 6, we solve this conjecture and prove the following

Theorem 2.2 . *The conjecture of de Bruin and Sharma is true.*

Our proof is based on the following general result.

Theorem 2.3 . Let z_1, \dots, z_n be the zeros of a polynomial p of degree $n \geq 2$ and w_1, \dots, w_{n-1} be the critical points of p . Suppose D, I and J are the same as those defined in Theorem 1.1. Then for any positive integer k , we have

$$\sum_{i=1}^{n-1} |w_i|^{2k} \leq \text{tr} \left(\left[D \left(I - \frac{1}{n} J \right) + \frac{z_n}{n} J \right]^k \left[\overline{D} \left(I - \frac{1}{n} J \right) + \frac{\overline{z_n}}{n} J \right]^k \right),$$

where $\text{tr}(A)$ and \overline{A} denote the trace and complex conjugate of a square matrix A respectively.

Remark. If we put $k = 1$ in the above inequality, it is not difficult to check that $\text{tr} \left(\left[D \left(I - \frac{1}{n} J \right) + \frac{z_n}{n} J \right] \left[\overline{D} \left(I - \frac{1}{n} J \right) + \frac{\overline{z_n}}{n} J \right] \right)$ is equal to

$$\frac{1}{n^2} \left| \sum_{i=1}^n z_i \right|^2 + \frac{n-2}{n} \sum_{i=1}^n |z_i|^2,$$

and hence we get back the Schoenberg inequality. Therefore, this gives an alternative proof of the Schoenberg conjecture.

We shall give the proof of Theorem 1.1 and 1.2 in Section 3. In Section 4, we recall some basic results from matrix theory which we shall repeatedly use. We then prove Theorem A and Theorem 2.1 in Section 5. Finally we prove Theorem 2.3 and solve the de Bruin and Sharma's conjecture in Section 6.

3 Proof of Theorem 1.1 and 1.2

Proof of Theorem 1.1. Note that $\frac{p'(z)}{p(z)} = \sum_{i=1}^n \frac{1}{z - z_i}$. If w is a critical point of p and w is not equal to any of the zeros z_i , then we have

$$0 = \frac{p'(w)}{p(w)} = \sum_{i=1}^n \frac{1}{w - z_i}.$$

Therefore, $\sum_{i=1}^{n-1} \frac{1}{z_i - w} = \frac{1}{w - z_n}$ and

$$\sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - w} = \sum_{i=1}^{n-1} \frac{z_i - w + w - z_n}{z_i - w} = n - 1 + \frac{w - z_n}{w - z_n} = n.$$

On the other hand, it is clear from the above discussion that if λ is a complex number such that $\lambda \neq z_i$ for all $1 \leq i \leq n-1$ and $\sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - \lambda} = n$, then $\lambda \neq z_n$ and λ is a critical point of the polynomial $p = a_n \prod_{i=1}^n (z - z_i)$.

Now suppose $\lambda \in \mathbb{C}$ is an eigenvalue of the $(n-1) \times (n-1)$ matrix $D(I - \frac{1}{n}J) + \frac{z_n}{n}J = D - \frac{1}{n}(D - z_n I)J$. We would like to show that λ is a critical point of p . Suppose $(v_1, \dots, v_{n-1})^T$ is an eigenvector associated with the eigenvalue λ . Then

$$\left[\begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_{n-1} \end{pmatrix} - \frac{1}{n} \begin{pmatrix} z_1 - z_n & & 0 \\ & \ddots & \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \right] \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \end{pmatrix}$$

is equal to the vector

$$\lambda \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \end{pmatrix}.$$

Hence,

$$\begin{pmatrix} z_1 - \lambda & & 0 \\ & \ddots & \\ 0 & & z_{n-1} - \lambda \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_{n-1} \end{pmatrix} = \frac{1}{n} \begin{pmatrix} z_1 - z_n & & 0 \\ & \ddots & \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} \sum_{i=1}^{n-1} v_i \\ \vdots \\ \sum_{i=1}^{n-1} v_i \end{pmatrix},$$

$$\begin{pmatrix} (z_1 - \lambda)v_1 \\ \vdots \\ (z_{n-1} - \lambda)v_{n-1} \end{pmatrix} = \begin{pmatrix} \left(\frac{1}{n} \sum_{i=1}^{n-1} v_i\right)(z_1 - z_n) \\ \vdots \\ \left(\frac{1}{n} \sum_{i=1}^{n-1} v_i\right)(z_{n-1} - z_n) \end{pmatrix}.$$

Note that since $(v_1, \dots, v_{n-1})^T$ is an eigenvector, at least one of its components is non-zero. Let's first consider the case that $\sum_{i=1}^{n-1} v_i = 0$. Then at least two of v_1, \dots, v_{n-1} are non-zero and we may assume them to be v_1 and v_2 . Since for $1 \leq k \leq n-1$,

$$(z_k - \lambda)v_k = \left(\frac{1}{n} \sum_{i=1}^{n-1} v_i\right)(z_k - z_n),$$

we have $z_1 - \lambda = z_2 - \lambda = 0$ and hence λ equals to the critical point $z_1 = z_2$. Now consider the case $\sum_{i=1}^{n-1} v_i \neq 0$. If $\lambda = z_k$ for some $1 \leq k \leq n-1$, then $z_k = z_n$ and therefore $\lambda = z_k = z_n$ is a critical point. It remains to consider the case that $\lambda \neq z_i$ for all $1 \leq i \leq n-1$. Note that in this case we have for $1 \leq i \leq n-1$,

$$\frac{v_i}{\sum_{i=1}^{n-1} v_i} = \frac{z_i - z_n}{n(z_i - \lambda)}.$$

Summing up all i from 1 to $n-1$, we have

$$n = \sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - \lambda},$$

and therefore λ is a critical point of the polynomial p .

Now, we claim that for each critical point w of p , w is an eigenvalue of the $(n-1) \times (n-1)$ matrix $D(I - \frac{1}{n}J) + \frac{z_n}{n}J = D - \frac{1}{n}(D - z_n I)J$.

We first consider the case that w is not equal to any of the zeros z_i . In this case, $\sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - w} = n$ and we shall show that $\left(\frac{z_1 - z_n}{n(z_1 - w)}, \dots, \frac{z_{n-1} - z_n}{n(z_{n-1} - w)} \right)^T$ is actually an eigenvector associated with the eigenvalue w of $D - \frac{1}{n}(D - z_n I)J$.

In fact,

$$\begin{aligned} & \left[\begin{pmatrix} z_1 & & \\ & \ddots & 0 \\ 0 & & z_{n-1} \end{pmatrix} - \frac{1}{n} \begin{pmatrix} z_1 - z_n & & \\ & \ddots & 0 \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \right] \begin{pmatrix} \frac{z_1 - z_n}{n(z_1 - w)} \\ \vdots \\ \frac{z_{n-1} - z_n}{n(z_{n-1} - w)} \end{pmatrix} \\ &= \begin{pmatrix} \frac{z_1(z_1 - z_n)}{n(z_1 - w)} \\ \vdots \\ \frac{z_{n-1}(z_{n-1} - z_n)}{n(z_{n-1} - w)} \end{pmatrix} - \frac{1}{n} \begin{pmatrix} z_1 - z_n & & \\ & \ddots & 0 \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} \frac{1}{n} \sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - w} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^{n-1} \frac{z_i - z_n}{z_i - w} \end{pmatrix} \\ &= \begin{pmatrix} \frac{z_1(z_1 - z_n)}{n(z_1 - w)} \\ \vdots \\ \frac{z_{n-1}(z_{n-1} - z_n)}{n(z_{n-1} - w)} \end{pmatrix} - \frac{1}{n} \begin{pmatrix} z_1 - z_n & & \\ & \ddots & 0 \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} \frac{z_1(z_1 - z_n)}{n(z_1 - w)} \\ \vdots \\ \frac{z_{n-1}(z_{n-1} - z_n)}{n(z_{n-1} - w)} \end{pmatrix} - \begin{pmatrix} \frac{z_1 - z_n}{n} \\ \vdots \\ \frac{z_{n-1} - z_n}{n} \end{pmatrix} = w \begin{pmatrix} \frac{z_1 - z_n}{n(z_1 - w)} \\ \vdots \\ \frac{z_{n-1} - z_n}{n(z_1 - w)} \end{pmatrix}.$$

Now suppose the critical point w is equal to one of the zeros, say z_i . This implies that $w = z_i = z_j$ for some $j \neq i$. If $j = n$, we claim that $w = z_n$ is an eigenvalue of $D - \frac{1}{n}(D - z_n I)J$ and $(1, \dots, 1)^T$ is a corresponding eigenvector.

In fact,

$$\begin{aligned} & \left[\begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_{n-1} \end{pmatrix} - \frac{1}{n} \begin{pmatrix} z_1 - z_n & & 0 \\ & \ddots & \\ 0 & & z_{n-1} - z_n \end{pmatrix} \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \right] \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} z_1 \\ \vdots \\ z_{n-1} \end{pmatrix} - \begin{pmatrix} z_1 - z_n \\ \vdots \\ z_{n-1} - z_n \end{pmatrix} \\ &= \begin{pmatrix} z_n \\ \vdots \\ z_n \end{pmatrix} = z_n \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = w \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}. \end{aligned}$$

Finally, it remains to consider the case that $w = z_i = z_j$ for some $1 \leq j \leq n - 1$. Let $\mathbf{v} = (v_1, \dots, v_{n-1})^T$ be the column vector with $v_i = 1$, $v_j = -1$ and $v_k = 0$ otherwise. Then $J\mathbf{v}$ equals to the zero vector. Hence, $(D - \frac{1}{n}(D - z_n I)J)\mathbf{v} = D\mathbf{v} = w\mathbf{v}$ and we are done. ■

Proof of Theorem 1.2. The proof of Theorem 1.2 is a direct application of Theorem 1.1 and the following well-known results in matrix theory. To state these results, for any square matrix $A = (a_{ij})$ of order $n \geq 2$, we shall use the following notation:

$$R_i(A) = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i = 1, \dots, n.$$

Theorem 3.1 (Gerschgorin's theorem) ([8, p.344]). The eigenvalues of any square matrix $A = (a_{ij})$ of order $n \geq 2$, lie in the union $G = \bigcup_{i=1}^n G_i$ of the Gerschgorin disks

$$G_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i(A)\}, \quad i = 1, \dots, n.$$

Theorem 3.2 (Brauer's theorem) ([8, p.380]). All the eigenvalues of a square matrix $A = (a_{ij})$ of order $n \geq 2$, are contained in the union $C = \bigcup_{\substack{i,j=1 \\ i \neq j}}^n C_{ij}$ of the ovals of Cassini

$$C_{ij} = \{z \in \mathbb{C} : |z - a_{ii}| |z - a_{jj}| \leq R_i(A)R_j(A)\}, \quad i \neq j, \quad i, j = 1, \dots, n.$$

Now, let $A = D - \frac{1}{n}DJ + \frac{z_n}{n}J$. Then A is an $(n-1) \times (n-1)$ matrix with entry $a_{ii} = \frac{n-1}{n}z_i + \frac{1}{n}z_n$ and $a_{ij} = \frac{1}{n}(z_n - z_i)$ for all $j \neq i$. Then $R_i(A) = \frac{n-2}{n}|z_n - z_i|$ and the results follow from Gerschgorin's theorem and Brauer's theorem and the fact that z_n can be any zero of p . \blacksquare

Remark. Other than Gerschgorin's theorem and Brauer's theorem, one can find many other eigenvalue inclusion theorems in the recent book of R.S. Varga [22]. These eigenvalue inclusion theorems can also be used to obtain results similar to Theorem 1.2.

4 Reviews on matrix theory

Before we continue, it would be helpful to fix some notations and recall a few facts from matrix theory. The readers are referred to [1], [8] and [9] for other basic results in matrix theory.

Let $A = (a_{ij})$ be an $n \times n$ matrix. The n eigenvalues of A will be written as $\lambda_i(A)$ ($1 \leq i \leq n$). If $f(z)$ is a polynomial, then the eigenvalues of the matrix $f(A)$ are precisely $f(\lambda_i(A))$.

Let $\bar{A} = (\bar{a}_{ij})$ be the complex conjugate of A , then \bar{A}^T is the conjugate transpose of A and it will be denoted by A^* . A matrix A is called Hermitian if $A = A^*$, unitary if $AA^* = I = A^*A$ and normal if $AA^* = A^*A$. Let $\text{tr}(A)$

be the trace of A , then the Schur inequality [16, p.56] says that

$$\sum_{i=1}^n |\lambda_i(A)|^2 \leq \text{tr}(A^*A),$$

and equality holds if and only if A is normal.

A Hermitian matrix A is said to be positive semidefinite if $\mathbf{x}^*A\mathbf{x} \geq 0$ for all column vector \mathbf{x} in \mathbb{C}^n and we shall then write $A \geq 0$. It is known that a Hermitian matrix is positive semidefinite if and only if all its eigenvalues are nonnegative [8, p.402]. Moreover, it is easy to see that if A is positive semidefinite, then so is S^*AS for any $n \times n$ matrix S .

Note that if A is positive semidefinite, then there exists a unique positive semidefinite matrix \sqrt{A} such that $\sqrt{A}\sqrt{A} = A$. Moreover the matrix A^*A is always positive semidefinite and its unique positive semidefinite $\sqrt{A^*A}$ will be denoted by $|A|$. The n eigenvalues of $|A|$ counted with multiplicities are called the singular values of A , denoted by $\sigma_i(A)$ ($1 \leq i \leq n$). We shall always enumerate the singular values and the modulus of the eigenvalues of A in decreasing order, i.e.

$$\sigma_1(A) \geq \cdots \geq \sigma_n(A) \quad \text{and} \quad |\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|.$$

We say A is similar to B (denoted by $A \sim B$) if there exists an invertible matrix S such that $A = S^{-1}BS$. A is unitarily similar to B if such S is an unitary matrix. Note that A and B share the same set of eigenvalues (singular values) if A and B are similar (unitarily similar). Finally, it is well-known that $MN \sim NM$ for any two square matrices M and N .

Throughout the remaining part of this paper, I is the identity matrix of order $n - 1$ and J is the $(n - 1) \times (n - 1)$ matrix with all entries equal to 1. We shall always use D to denote the $(n - 1) \times (n - 1)$ diagonal matrix below

$$\begin{pmatrix} z_1 & & 0 \\ & \ddots & \\ 0 & & z_{n-1} \end{pmatrix}.$$

In this case, $|D|$ is simply be the following $(n-1) \times (n-1)$ diagonal matrix

$$\begin{pmatrix} |z_1| & & 0 \\ & \ddots & \\ 0 & & |z_{n-1}| \end{pmatrix}.$$

5 Proof of Theorem A and Theorem 2.1

Proof of Theorem A. Let $\alpha = \frac{\sqrt{n}+1}{\sqrt{n}(n-1)}$. We claim that the set of eigenvalues of the matrix $A_1 = (I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ is exactly the same as the set of critical points of the polynomial p . This follows directly from the following lemma and Theorem 1.1.

Lemma 5.1 *Let $\alpha = \frac{\sqrt{n}+1}{\sqrt{n}(n-1)}$. Then the symmetric matrix $(I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ and $D(I - \frac{1}{n}J) + \frac{z_n}{n}J$ have the same set of eigenvalues.*

Proof of Lemma 5.1. It suffices to show that $(I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ and $D(I - \frac{1}{n}J) + \frac{z_n}{n}J$ are similar. First of all, notice that for $\alpha = \frac{\sqrt{n}+1}{\sqrt{n}(n-1)}$, we have $(I - \alpha J)(I - \alpha J) = I - \frac{1}{n}J$. Moreover, one can also check easily that $(I - \alpha J)^{-1} = I - \frac{\alpha}{(n-1)\alpha-1}J = I - \frac{\sqrt{n}+1}{n-1}J$. Hence both $(I - \alpha J)$ and $(I - \alpha J)^{-1}$ are degree one polynomials in J and therefore they commute with J . This implies that

$$\begin{aligned} & (I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J \\ &= (I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}(I - \alpha J)J(I - \alpha J)^{-1} \\ &= (I - \alpha J)\left(D(I - \alpha J) + \frac{z_n}{n}J(I - \alpha J)^{-1}\right) \\ &\sim \left(D(I - \alpha J) + \frac{z_n}{n}J(I - \alpha J)^{-1}\right)(I - \alpha J) \\ &= D(I - \alpha J)^2 + \frac{z_n}{n}J \\ &= D(I - \frac{1}{n}J) + \frac{z_n}{n}J. \end{aligned}$$

Therefore, $(I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ is similar to $D(I - \frac{1}{n}J) + \frac{z_n}{n}J$. ■

By Lemma 5.1 and Theorem 1.1, the set of eigenvalues of the matrix $(I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ is exactly equal to $\{w_1, \dots, w_{n-1}\}$, the set of critical points of the polynomial p .

Let $\mathbf{v} \in \mathbb{C}^{n-1}$ be a column vector with all entries equal to one and E be the following $n \times n$ Hermitian matrix

$$\begin{bmatrix} I - \alpha J & \frac{1}{\sqrt{n}}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T & \frac{1}{\sqrt{n}} \end{bmatrix}.$$

Then it is easy to verify that $E^*E = I_n$ and therefore $E^{-1} = E^* = E$. Hence,

$\tilde{D} = \begin{bmatrix} D & O \\ O & z_n \end{bmatrix}$ is unitarily similar to

$$\begin{aligned} & \begin{bmatrix} I - \alpha J & \frac{1}{\sqrt{n}}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T & \frac{1}{\sqrt{n}} \end{bmatrix} \begin{bmatrix} D & O \\ O & z_n \end{bmatrix} \begin{bmatrix} I - \alpha J & \frac{1}{\sqrt{n}}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T & \frac{1}{\sqrt{n}} \end{bmatrix} \\ &= \begin{bmatrix} (I - \alpha J)D & \frac{z_n}{\sqrt{n}}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T D & \frac{z_n}{\sqrt{n}} \end{bmatrix} \begin{bmatrix} I - \alpha J & \frac{1}{\sqrt{n}}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T & \frac{1}{\sqrt{n}} \end{bmatrix} \\ &= \begin{bmatrix} (I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J & \frac{1}{\sqrt{n}}(I - \alpha J)D\mathbf{v} + \frac{z_n}{n}\mathbf{v} \\ \frac{1}{\sqrt{n}}\mathbf{v}^T D(I - \alpha J) + \frac{z_n}{n}\mathbf{v}^T & \frac{1}{n}\mathbf{v}^T D\mathbf{v} + \frac{z_n}{n} \end{bmatrix}. \end{aligned}$$

We shall denote the last matrix by A . Since A is unitarily similar to \tilde{D} , the set of the singular values of A is equal to that of \tilde{D} , which is simply $\{|z_1|, \dots, |z_n|\}$.

Notice that $A_1 = (I - \alpha J)D(I - \alpha J) + \frac{z_n}{n}J$ is the upper-left hand $n - 1$ by $n - 1$ principal submatrix of A . In order to compare the singular values of A_1 with that of A , we shall need the following interlacing theorem for singular values of complex matrices which follows from Corollary 3.1.3 in [9].

Theorem 5.2 . *Let A be an $n \times n$ complex matrix and $\sigma_1(A) \geq \dots \geq \sigma_n(A)$ be the ordered singular values of A . Let A_1 be the upper-left hand $n - 1$ by $n - 1$ principal submatrix of A and $\sigma_1(A_1) \geq \dots \geq \sigma_{n-1}(A_1)$ be the ordered*

singular values of A_1 . Then for each $1 \leq k \leq n$, we have

$$\sigma_k(A) \geq \sigma_k(A_1) \geq \sigma_{k+2}(A).$$

Remark. Note that the above submatrix A_1 corresponds to the matrix A_2 in Corollary 3.1.3 of [9] because A_1 is obtained by deleting a total of two "lines" (one row and one column) from A .

We also need the following result on the majorization of the eigenvalues of a matrix by its singular values.

Theorem 5.3 ([9, p.166]). Let A be an $n \times n$ complex matrix and $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A) \geq 0$ be the ordered singular values of A . Let $\{\lambda_1(A), \dots, \lambda_n(A)\}$ be the set of eigenvalues of A ordered so that $|\lambda_1(A)| \geq \cdots \geq |\lambda_n(A)|$. Then, we have for each $k = 1, \dots, n$,

$$\sum_{i=1}^k \psi(|\lambda_i(A)|) \leq \sum_{i=1}^k \psi(\sigma_i(A)),$$

for every increasing function $\psi : [0, +\infty) \rightarrow \mathbb{R}$ such that $\psi \circ \exp$ is convex on \mathbb{R} .

Now without loss of generality assume that $|w_1| \geq \cdots \geq |w_{n-1}| \geq |w_n| = 0$. Apply Theorem 5.2 and Theorem 5.3 and use the fact that $\{\lambda_1(A_1), \dots, \lambda_{n-1}(A_1)\} = \{w_1, \dots, w_{n-1}\}$, we have for each $k = 1, \dots, n-1$,

$$\sum_{i=1}^k \psi(|w_i|) = \sum_{i=1}^k \psi(|\lambda_i(A_1)|) \leq \sum_{i=1}^k \psi(|\sigma_i(A_1)|) \leq \sum_{i=1}^k \psi(|\sigma_i(A)|) = \sum_{i=1}^k \psi(|z_i|).$$

The $k = n$ case follows from the $k = n-1$ case and the fact that $|z_n| \geq |w_n| = 0$ and ψ is an increasing function on $[0, +\infty)$. This completes the proof of Theorem A. ■

Proof of Theorem 2.1. Without loss of generality we shall assume that

$$|w_1| \geq \cdots \geq |w_{n-1}| \quad \text{and} \quad |v_1| \geq \cdots \geq |v_{n-1}|.$$

Now fix a branch of \sqrt{z} and define

$$D^{\frac{1}{2}} = \begin{pmatrix} \sqrt{z_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{z_{n-1}} \end{pmatrix}.$$

Let $A = D^{\frac{1}{2}}(I - \frac{1}{n}J)D^{\frac{1}{2}}$. Then $A = D^{\frac{1}{2}}(I - \frac{1}{n}J)D^{\frac{1}{2}} \sim D^{\frac{1}{2}}D^{\frac{1}{2}}(I - \frac{1}{n}J) = D(I - \frac{1}{n}J)$. It follows from Theorem 1.1 that for each $1 \leq i \leq n-1$, $\lambda_i(A) = w_i$. Now we would like to find the singular values of A , which are the same as the eigenvalues of $|A|$. We claim that $|A|$ is equal to $\overline{D}^{\frac{1}{2}}(I - \frac{1}{n}J)D^{\frac{1}{2}}$. First of all, the eigenvalues of $(I - \frac{1}{n}J)$ are equal to 1 and $\frac{1}{n}$, hence $(I - \frac{1}{n}J)$ is positive semidefinite. It follows easily that $\overline{D}^{\frac{1}{2}}(I - \frac{1}{n}J)D^{\frac{1}{2}} = (D^{\frac{1}{2}})^*(I - \frac{1}{n}J)D^{\frac{1}{2}}$ is also positive semidefinite. Now it remains to check that

$$\begin{aligned} & \overline{D}^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)D^{\frac{1}{2}}\overline{D}^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)D^{\frac{1}{2}} \\ &= \overline{D}^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)|D|\left(I - \frac{1}{n}J\right)D^{\frac{1}{2}} \\ &= \overline{D}^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)\overline{D}^{\frac{1}{2}}D^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)D^{\frac{1}{2}} \\ &= A^*A. \end{aligned}$$

Since D is a diagonal matrix, it is possible to choose a unitary matrix S of the form

$$\begin{pmatrix} e^{i\theta_1} & & 0 \\ & \ddots & \\ 0 & & e^{i\theta_{n-1}} \end{pmatrix}$$

such that $S^{-1}\overline{D}^{\frac{1}{2}} = |D|^{\frac{1}{2}} = D^{\frac{1}{2}}S$. Therefore $|A| = \overline{D}^{\frac{1}{2}}(I - \frac{1}{n}J)D^{\frac{1}{2}}$ is unitary similar to $|D|^{\frac{1}{2}}(I - \frac{1}{n}J)|D|^{\frac{1}{2}}$. Hence, we have

$$\sigma_i(A) = \lambda_i(|A|) = \lambda_i\left(|D|^{\frac{1}{2}}\left(I - \frac{1}{n}J\right)|D|^{\frac{1}{2}}\right).$$

Note that $|D|^{\frac{1}{2}}(I - \frac{1}{n}J)|D|^{\frac{1}{2}}$ is similar to $|D|(I - \frac{1}{n}J)$ and by Theorem 1.1, the eigenvalues of $|D|(I - \frac{1}{n}J)$ are the critical points of the polynomial p_a . Therefore, we have for each $1 \leq i \leq n-1$, $\sigma_i(A) = v_i$.

Apply Theorem 5.3 and use the fact that $\lambda_i(A) = w_i$, we have for each $k = 1, \dots, n-1$,

$$\sum_{i=1}^k \psi(|w_i|) = \sum_{i=1}^k \psi(|\lambda_i(A)|) \leq \sum_{i=1}^k \psi(\sigma_i(A)) \leq \sum_{i=1}^k \psi(v_i),$$

and we are done. ■

6 A proof of de Bruin and Sharma's conjecture

We prove Theorem 2.3 first, our proof of de Bruin and Sharma's conjecture will then follow easily from this theorem.

Proof of Theorem 2.3. We shall need the following simple lemma in our proof.

Lemma 6.1 *Let C be an $m \times m$ square matrix. Suppose there exists some positive semidefinite matrix S such that $\bar{C} = S^{-1}C^*S$. Then for any positive integer k ,*

$$\sum_{i=1}^m |\lambda_i(C)|^{2k} \leq \text{tr}(C^k \bar{C}^k).$$

Proof. Let \sqrt{S} be the unique positive semidefinite matrix such that $\sqrt{S}\sqrt{S} = S$. Note that this \sqrt{S} must be Hermitian, i.e. $\sqrt{S} = (\sqrt{S})^*$. From the assumption, we have

$$\begin{aligned} C^k \bar{C}^k &= C^k S^{-1} (C^*)^k S \\ &= C^k (\sqrt{S})^{-1} (\sqrt{S})^{-1} (C^*)^k \sqrt{S} \sqrt{S} \\ &\sim \sqrt{S} C^k \sqrt{S}^{-1} \sqrt{S}^{-1} (C^k)^* \sqrt{S} \\ &= \sqrt{S} C^k \sqrt{S}^{-1} ((\sqrt{S})^*)^{-1} (C^k)^* (\sqrt{S})^* \\ &= \sqrt{S} C^k \sqrt{S}^{-1} (\sqrt{S} C^k \sqrt{S}^{-1})^*. \end{aligned}$$

Therefore,

$$\begin{aligned} \text{tr}(C^k \bar{C}^k) &= \text{tr}(\sqrt{S} C^k \sqrt{S}^{-1} (\sqrt{S} C^k \sqrt{S}^{-1})^*) \\ &\geq \sum_{i=1}^m |\lambda_i(\sqrt{S} C^k \sqrt{S}^{-1})|^2 \\ &= \sum_{i=1}^m |\lambda_i(C^k)|^2 \\ &= \sum_{i=1}^m |\lambda_i(C)|^{2k}. \end{aligned}$$

■

Let $B = D(I - \frac{1}{n}J) + \frac{z_n}{n}J$, then $\bar{B} = \bar{D}(I - \frac{1}{n}J) + \frac{\bar{z}_n}{n}J$ and $B^* =$

$(I - \frac{1}{n}J)\bar{D} + \frac{\bar{z}_n}{n}J$. In order to apply Lemma 6.1, we first notice that

$$\begin{aligned} \left(I - \frac{1}{n}J\right)^{-1} B^* \left(I - \frac{1}{n}J\right) &= (I + J)B^* \left(I - \frac{1}{n}J\right) \\ &= (I + J) \left(\left(I - \frac{1}{n}J\right)\bar{D} + \frac{\bar{z}_n}{n}J \right) \left(I - \frac{1}{n}J\right) \\ &= \bar{D} \left(I - \frac{1}{n}J\right) + \frac{\bar{z}_n}{n}J = \bar{B}. \end{aligned}$$

Now $S = (I - \frac{1}{n}J)$ is positive semidefinite because it is Hermitian and its eigenvalues are nonnegative. By Lemma 6.1, we have

$$\sum_{i=1}^{n-1} |w_i|^{2k} \leq \text{tr} \left(\left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^k \left[\bar{D} \left(I - \frac{1}{n}J \right) + \frac{\bar{z}_n}{n}J \right]^k \right).$$

■

Proof of Theorem 2.2. Take $k = 2$ in Theorem 2.3, then we have

$$\sum_{i=1}^{n-1} |w_i|^4 \leq \text{tr} \left(\left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^2 \left[\bar{D} \left(I - \frac{1}{n}J \right) + \frac{\bar{z}_n}{n}J \right]^2 \right).$$

Using the fact that for a diagonal matrix E , we have $JEJ = \text{tr}(E)J$ as well as the assumption that $\sum_{i=1}^n z_i = 0$ (which is the same as $\text{tr}(D) = -z_n$), one can check easily that

$$\left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^2 = D^2 - \frac{1}{n}D^2J - \frac{1}{n}DJ D + \frac{z_n}{n}JD + \frac{z_n^2}{n}J.$$

Let $A = \left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^2$, then $A\bar{A} = \left[D \left(I - \frac{1}{n}J \right) + \frac{z_n}{n}J \right]^2 \left[\bar{D} \left(I - \frac{1}{n}J \right) + \frac{\bar{z}_n}{n}J \right]^2$ is equal to

$$\left[D^2 - \frac{1}{n}D^2J - \frac{1}{n}DJ D + \frac{z_n}{n}JD + \frac{z_n^2}{n}J \right] \left[\bar{D}^2 - \frac{1}{n}\bar{D}^2J - \frac{1}{n}\bar{D}J\bar{D} + \frac{\bar{z}_n}{n}J\bar{D} + \frac{\bar{z}_n^2}{n}J \right].$$

To compute $\text{tr}(A\bar{A})$, it is useful to notice that $\text{tr}(MN) = \text{tr}(NM)$ and for any diagonal matrix E , we have $JEJ = \text{tr}(E)J$ and $\text{tr}(EJ) = \text{tr}(JE) = \text{tr}(E)$.

Expand $A\bar{A}$ and after simplifications, we find that

$$\begin{aligned}
\operatorname{tr}(A\bar{A}) &= \frac{n-4}{n} \operatorname{tr}(|D|^4) + 2 \operatorname{Re} \left\{ \frac{z_n}{n} \operatorname{tr}(\bar{D}|D|^2) \right\} \\
&+ 2 \operatorname{Re} \left\{ \frac{z_n^2}{n} \operatorname{tr}(\bar{D}^2) \right\} + \frac{1}{n^2} \operatorname{tr}(\bar{D}^2) \operatorname{tr}(D^2) \\
&+ 2 \operatorname{Re} \left\{ \frac{1}{n^2} \operatorname{tr}(D) \operatorname{tr}(\bar{D}|D|^2) \right\} - 2 \operatorname{Re} \left\{ \frac{n-1}{n^2} z_n \operatorname{tr}(\bar{D}|D|^2) \right\} \\
&- 2 \operatorname{Re} \left\{ \frac{n-1}{n^2} z_n^2 \operatorname{tr}(\bar{D}^2) \right\} + \frac{1}{n^2} [\operatorname{tr}(|D|^2)]^2 \\
&- 2 \operatorname{Re} \left\{ \frac{z_n}{n^2} \operatorname{tr}(\bar{D}) \operatorname{tr}(|D|^2) \right\} - 2 \operatorname{Re} \left\{ \frac{z_n^2}{n^2} [\operatorname{tr}(\bar{D})]^2 \right\} \\
&+ \frac{|z_n|^2}{n^2} |\operatorname{tr}(D)|^2 + 2 \operatorname{Re} \left\{ \frac{(n-1)z_n|z_n|^2}{n^2} \operatorname{tr}(\bar{D}) \right\} + \left(\frac{n-1}{n} \right)^2 |z_n|^4.
\end{aligned}$$

Since $\operatorname{tr}(D) = -z_n$, we have

$$\begin{aligned}
\operatorname{tr}(A\bar{A}) &= \frac{n-4}{n} \operatorname{tr}(|D|^4) + \frac{2}{n^2} \operatorname{Re} \left\{ z_n^2 \operatorname{tr}(\bar{D}^2) \right\} \\
&+ \frac{1}{n^2} |\operatorname{tr}(D^2)|^2 + \frac{1}{n^2} [\operatorname{tr}(|D|^2)]^2 \\
&+ \frac{2}{n^2} |z_n|^2 \operatorname{tr}(|D|^2) + \left(\frac{n-4}{n} + \frac{2}{n^2} \right) |z_n|^4 \\
&\leq \frac{n-4}{n} \sum_{i=1}^n |z_i|^4 + \frac{2}{n^2} \{ [\operatorname{tr}(|D|^2)]^2 + 2|z_n|^2 \operatorname{tr}(|D|^2) + |z_n|^4 \} \\
&= \frac{n-4}{n} \sum_{i=1}^n |z_i|^4 + \frac{2}{n^2} \left(\sum_{i=1}^n |z_i|^2 \right)^2.
\end{aligned}$$

The inequality above follows from the simple fact that

$$\operatorname{Re} \left\{ z_n^2 \operatorname{tr}(\bar{D}^2) \right\} \leq |z_n|^2 \operatorname{tr}(|D|^2) \quad \text{and} \quad |\operatorname{tr}(D^2)|^2 \leq [\operatorname{tr}(|D|^2)]^2.$$

It remains to show that equality holds if and only if z_1, \dots, z_n lie on a straight line passing through the origin of the complex plane. Note that when the equality holds, we must have $|\operatorname{tr}(D^2)|^2 = [\operatorname{tr}(|D|^2)]^2$ which is the same as

$$|(z_1, \dots, z_{n-1})(\bar{z}_1, \dots, \bar{z}_{n-1})^*|^2 = \left(\sum_{i=1}^n |z_i|^2 \right) \left(\sum_{i=1}^n |\bar{z}_i|^2 \right).$$

By Cauchy-Schwarz inequality, there exists some complex number λ such that $(z_1, \dots, z_{n-1}) = \lambda(\bar{z}_1, \dots, \bar{z}_{n-1})$. This implies that for those non-zero z_i , their argument are all equal to some fixed angle θ . This is also true for z_n

because $z_n = -\sum_{i=1}^{n-1} z_i$. Therefore, all z_1, \dots, z_n lie on a straight line passing through the origin.

Finally, assume that z_1, \dots, z_n lie on a straight line passing through the origin. This is equivalent to saying that there exists some $0 \leq \theta \leq 2\pi$ such that $z_i = |z_i|e^{i\theta}$ for all $1 \leq i \leq n$. By the Gauss-Lucas theorem, all w_i will also lie on the same straight line and therefore $w_i = |w_i|e^{i\theta}$ for all i . Then it follows easily from the identity,

$$\sum_{i=1}^{n-1} w_i^4 = \frac{n-4}{n} \sum_{i=1}^n z_i^4 + \frac{2}{n^2} \left(\sum_{i=1}^n z_i^2 \right)^2,$$

that the equality holds.

One way to obtain the above identity is applying the *formulae of Newton* (see [16, p.8]). Here we suggest another way to get the above identity. Notice that if $A = D(I - \frac{1}{n}J) + \frac{z_n}{n}J$, then Theorem 1.1 implies that the eigenvalues of A^4 are w_1^4, \dots, w_{n-1}^4 . Hence, we have $\sum_{i=1}^{n-1} w_i^4 = \text{tr}(A^4)$. Expand $\text{tr}(A^4)$ and use the assumption $\text{tr}(D) = -z_n$ to simplify the expression as we have done before, we will have

$$\text{tr}(A^4) = \frac{n-4}{n} \sum_{i=1}^n z_i^4 + \frac{2}{n^2} \left(\sum_{i=1}^n z_i^2 \right)^2,$$

and we are done.

Acknowledgments. The authors would like to thank the referee for the extremely helpful and valuable suggestions. ■

References

- [1] R. Bhatia, Matrix analysis. Graduate Texts in Mathematics, **169**, Springer-Verlag, New York, 1997.
- [2] P. Borwein, T. Erdelyi, Polynomials and polynomial inequalities. Graduate Texts in Mathematics, **161**, Springer-Verlag, New York, 1995.

- [3] M.G. de Bruin, A. Sharma, On a Schoenberg-type conjecture. Continued fractions and geometric function theory (CONFUN) (Trondheim, 1997). *J. Comput. Appl. Math.* **105** (1999), no. 1-2, 221–228.
- [4] N.G. de Bruijn, T.A. Springer, On the zeros of a polynomial and its derivative II, *Indag.Math.* **9** (1947), 264-270.
- [5] N.G. de Bruijn, T.A. Springer, On the zeros of composition-polynomials, *Indag. Math.* **9** (1947), 406-414.
- [6] C. Davis, Eigenvalues of compressions, *Bull. Math. Soc. Math. Phys. RPR* **51** (1959) 3-5.
- [7] W.K. Hayman, *Research problems in function theory*. The Athlone Press University of London, London 1967.
- [8] R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, 1990.
- [9] R.A. Horn and C.R. Johnson. *Topics in matrix analysis*. Cambridge University Press, Cambridge, 1991.
- [10] E.S. Katsoprinakis, On the complex Rolle set of a polynomial. *Computational methods and function theory 1997 (Nicosia)*, 315–326, Ser. Approx. Decompos., 11, World Sci. Publishing, River Edge, NJ, 1999.
- [11] S.M. Malamud, An Analog of the Poincare Separation Theorem for Normal Matrices and the Gauss-Lucas Theorem. *Functional Analysis and Its Applications* **37** (2003), no.3, 232-235.
- [12] S.M. Malamud, Inverse spectral problem for normal matrices and the Gauss-Lucas theorem, *Trans. Amer. Math. Soc.* (to appear).
- [13] M. Marden, *Geometry of polynomials*. Second edition. *Mathematical Surveys*, **3** American Mathematical Society, 1966.

- [14] A.W. Marshall, I. Olkin, Inequalities: theory of majorization and its applications. Mathematics in Science and Engineering, **143**. Academic Press, Inc., New York-London, 1979.
- [15] R. Pereira, Differentiators and the geometry of polynomials. J. Math. Anal. Appl. **285** (2003), no. 1, 336–348.
- [16] Q.I. Rahman and G. Schmeisser, Analytic theory of polynomials. London Mathematical Society Monographs. New Series, 26. The Clarendon Press, Oxford University Press, Oxford, 2002.
- [17] G. Schmeisser, The conjectures of Sendov and Smale. Approximation theory: A volume dedicated to Blagovest Sendov, 353–369, DARBA, Sofia, 2002.
- [18] G. Schmeisser, Majorization of the critical points of a polynomial by its zeros, Computational Methods and Function Theory **3** (2003), 95-103.
- [19] I.J. Schoenberg, A conjectured analogue of Rolle’s theorem for polynomials with real or complex coefficients. Amer. Math. Monthly 93 (1986), no. 1, 8–13.
- [20] Bl. Sendov, Hausdorff geometry of polynomials, East J. Approx. **7** (2001), no. 2, 123–178.
- [21] T. Sheil-Small, Complex polynomials. Cambridge Studies in Advanced Mathematics, 75. Cambridge University Press, Cambridge, 2002.
- [22] R.S. Varga, Geršgorin and his circles. Springer Series in Computational Mathematics, 36. Springer-Verlag, Berlin, 2004.