EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Methods of operator theory and majorization theory in the geometry of polynomials

av

Jonathan Lundborg

2005 - No 3

Methods of operator theory and majorization theory in the geometry of polynomials

Jonathan Lundborg

Examensarbete i matematik 10 poäng

Handledare: Julius Borcea

2005

Abstract

It was recently noticed that one can gain substantial new insight into the geometry of polynomials by using methods from operator theory and majorization theory. Such methods were used by Pereira and Malamud to prove three long-standing conjectures of de Bruijn-Springer, Schoenberg and Katsoprinakis that go much beyond the classical Gauss-Lucas theorem. In this paper, we present the solutions to these conjectures and several new relationships between the zeros and critical points of arbitrary complex polynomials.

Acknowledgements

I want to thank my tutor Julius Borcea for his great guidance and patience. I am deeply grateful for all the tips and the knowledge he so generously has provided me with. It has been a privilege.

Contents

1 Introduction

The geometry of zeros and critical points of complex polynomials is a classical subject in geometric function theory. There is a vast literature devoted to this topic and its applications (see [1] and references therein.) The well-known Gauss-Lucas theorem says that the critical points of a polynomial lie in the convex hull spanned by its roots. We shall prove three conjectures that give us much more information about relationship between zeros and critical points of an arbitrary polynomial than we already know from the Gauss-Lucas theorem. The conjectures are the de Bruijn-Springer conjecture (1947), Schoenberg's conjecture (1986) and a related conjecture by Katsoprinakis (1997). These long-standing problems have been recently solved by Pereira [2] and Malamud [3] through an ingenious combination of arguments involving operator theory and majorization theory.

In Section 1 we present some general results that are helpful to prove the three conjectures. These preliminary results are regrouped into three subsections. We first review the necessary background on matrix and operator theory in Section 1.1. We will assume that the reader is already familiar with the basic properties of Hilbert spaces and matrix functions. In Section 1.2 we discuss the concept of *differentiator*, first introduced by Davies in 1958 (see [4]). Given an operator that possesses a differentiator we can construct a *compression* of the operator in such a way that the characteristic polynomials of the operator and its compression relate in a similar way that an arbitrary polynomial relates to its derivate. We also define the notion of *trace vector* of an operator and show that the existence of a trace vector implies the existence of a differentiator, and vice versa. We end the subsection by showing that every *normal* operator actually possesses a trace vector and thus a differentiator. To summarize so far, Section 1.2 provides the set-up for studying relations between a polynomial and its derivate via operators and their characteristic polynomials. In Section 1.3 we briefly touch the subject of *majorization*. We shall subsequently see that we can in fact formulate the de Bruijn-Springer and Katsoprinakis conjectures in terms of majorization relations. By making use of the tools presented in section 1, we prove Schoenberg's conjecture, Katsoprinakis conjecture and the de Bruijn-Springer conjecture in section 2, 3 and 4, respectively.

1.1 Some general results in operator theory

Let $\mathcal{H}$ be an n-dimensional Hilbert space, $\mathrm{L}(\mathcal{H})$ be the set of linear operators from $\mathcal{H}$ to $\mathcal{H}$, A be any operator in $\mathrm{L}(\mathcal{H})$ and $e = (e_1, e_2, \ldots, e_n)$ be any basis of $\mathcal{H}$. Each operator in a given basis of $\mathcal{H}$ can be represented by an n by n matrix, so to make a clear distinction between an operator and a matrix, we let $[A]_e$ denote the matrix representation of A in basis $e = (e_1, e_2, \ldots, e_n)$. The (i, j)th element in $[A]_e$ is $e_i^* A e_j = \langle Ae_j, e_i \rangle$. Given two operators A_1 and A_2 we also have the basic property $[A_1 A_2]_e = [A_1]_e [A_2]_e$. For operators we will use the *operator norm* and for matrices the *Euclidian norm* also called the *Frobenius* -or *Hilbert-Schmidt norm.*

Definition 1.1. Define the *operator norm* $\| \cdot \|$ of an operator $A \in \mathrm{L}(\mathcal{H})$ to be

$$\| A \| = \sup_{\substack{\| x \| = 1 \\ x \in \mathcal{H}}} \| Ax \| \,.$$

Definition 1.2. Define the *Euclidian norm* $\| \cdot \|_E$ of an n by n matrix $M = (m_{ij})$ to be

$$\| M \|_E = \Big[\sum_{i=1}^{n} \sum_{j=1}^{n} | m_{ij} |^2 \Big]^{\frac{1}{2}}.$$

We note that the *Euclidian norm* is a *unitarily invariant norm.*[1] Recall that given a matrix $M = (m_{ij})$ one define its Hermitian transpose to be the matrix M^* whose (i, j)th entry is $\overline{m_{ji}}$. M is called *Hermitian* if $M = M^*$ and *normal* if $MM^* = M^*M$. Hence for any operator A, the Euclidian norm of a matrix representation of A is independent of the choice of orthonormal basis in $\mathcal{H}$. This may also be verified by using the following lemma that describes the relation between matrix representations of A in different bases.

Lemma 1.3. *Let $e = (e_1, e_2, \ldots, e_n)$ and $f = (f_1, f_2, \ldots, f_n)$ be two different bases in $\mathcal{H}$ where $f = eQ$ for an n by n matrix Q. Then Q is invertible and $[A]_f = Q^{-1}[A]_e Q$. Furthermore, if $e = (e_1, e_2, \ldots, e_n)$ and $f = (f_1, f_2, \ldots, f_n)$ are orthonormal bases, then $[A]_f = Q^*[A]_e Q$ and Q is a* unitary *matrix where the (i, j)th element of Q is $\langle f_j, e_i \rangle = \overline{\langle e_i, f_j \rangle}$.*

For two given orthonormal bases $e = (e_1, e_2, \ldots, e_n)$ and $f = (f_1, f_2, \ldots, f_n)$ in $\mathcal{H}$ we have according to Lemma 1.3 that $e_i = \sum_{j=1}^{n} \langle e_i, f_j \rangle f_j$ and $f_i = \sum_{j=1}^{n} \langle f_i, e_j \rangle e_j$. By taking the norm of a given base vector, we get $\| e_i \| =$

[1]An n by n matrix U is *unitary* if $U^*U = I = UU^*$. A norm $\| \cdot \|$ on the m by n matrices is *unitarily invariant* if $\| UMV \| = \| M \|$ for all m by n matrices M, m by m unitary matrices U, and n by n unitary matrices V.

$\sum_{j=1}^{n} |\langle e_i, f_j\rangle|^2 = 1$ and $\| f_i\| = \sum_{j=1}^{n} |\langle f_i, e_j\rangle|^2 = 1$. This result is usually known as *Parseval's theorem.*

In our next theorem we will see that given an operator $A \in \mathrm{L}(\mathcal{H})$ we can always find an orthonormal basis $e = (e_1, e_2, \ldots, e_n)$ such that $[A]_e$ becomes an upper triangular matrix. This basis is called a *Schur basis* of A, and $[A]_e$ the *Schur Triangular Form* of A.

Theorem 1.4. *For any operator $A \in L(\mathcal{H})$ we can find an orthonormal basis $e = (e_1, e_2, \ldots, e_n)$ such that Ae_k is a linear combination of $e_1, \ldots, e_k$ where $k = 1, 2, \ldots, n$.*

If $e = (e_1, e_2, \ldots, e_n)$ is a Schur basis of A then the (i,j)th entry in $[A]_e$ is $e_i^* A e_j$ which is 0 whenever $i > j$, and thus $[A]_e$ is upper triangular. An immediate consequence of this theorem is shown in Corollary 1.6, which also follows from *Spectral Theorem for normal matrices.* Let us first define the notions of adjoint operator and normal operator.

Definition 1.5. Let $A \in \mathrm{L}(\mathcal{H})$. There exists a unique operator $A^* \in \mathrm{L}(\mathcal{H})$ such that $\langle Ax, y\rangle = \langle x, A^*y\rangle$ for all $x, y \in \mathcal{H}$. We call A^* the *adjoint* (or *dual*) operator of A. The operator A is called *normal* if $AA^* = A^*A$.

It is easy to see that an operator is normal if and only if its matrix representation in some (and then any) orthonormal basis is normal.

Corollary 1.6. *Let $A \in L(\mathcal{H})$ be a normal operator and let $e = (e_1, e_2, \ldots, e_n)$ be a Schur basis of A. Then $[A]_e$ is a diagonal matrix, and $e_1, \ldots, e_n$ are eigenvectors of $[A]_e$.*

Due to Theorem 1.4 and Corollary 1.6, it is convenient to work with upper triangular matrices or diagonal matrices. For example, we immediately see that an operator is normal if and only if its matrix representation in a Schur basis is diagonal.

By studying the properties of the matrix representations given by an operator, one can in some cases generalize these properties to the operator itself. We list some properties of n by n matrices in the following lemma.

Lemma 1.7. *Let Q be an invertible n by n matrix. For any n by n matrix M let $\tau(M)$ to be the arithmetic mean of the diagonal elements in M. Then the following relations hold:*

1. $\det(M) = \det(Q^{-1}MQ)$.

2. $\det(\lambda I - M) = \det(\lambda I - Q^{-1}MQ)$.

3. $\tau(M) = \tau(Q^{-1}MQ)$.

4. *Let p be any polynomial. Then $p(Q^{-1}MQ) = Q^{-1}p(M)Q$.*

For any operator $A \in \mathrm{L}(\mathcal{H})$, Lemma 1.3 and Lemma 1.7 basically say that the determinant, the characteristic polynomial, the trace and a polynomial of the matrices defined by A are independent of the choice of basis in $\mathcal{H}$. Therefore we define the *determinant* of A to be $\det(A)$, the *characteristic polynomial* of A to be p_A and the *normalized trace* of A to be $\tau(A)$ by fixing a basis $e = (e_1, e_2, \ldots, e_n)$ in $\mathcal{H}$ and setting $\det(A) = \det([A]_e)$, $p_A(\lambda) = p_{[A]_e}(\lambda) = \det(\lambda I - [A]_e)$ and $\tau(A) = \tau([A]_e)$. Let the eigenvalues of A be $\{\lambda_i(A)\}$; then by choosing a Schur Triangular Form of A we can see that $\det(A) = \prod_{i=1}^{n} \lambda_i(A)$ and $\tau(A) = \frac{1}{n}\sum_{i=1}^{n} \lambda_i(A)$.

If $\det(A) \neq 0$, every matrix representation of A is invertible, so we say that A itself is invertible.

We end this subsection with a couple of lemmas which describe some useful properties of normal operators.

Lemma 1.8. *Let $A \in L(\mathcal{H})$ be a normal operator. Then we can express its adjoint operator A^* as a polynomial of A.*

Proof. Choose an orthonormal basis of eigenvectors $\{v_i\}_{i=1}^{n}$ of A corresponding to the eigenvalues $\{\lambda_i\}_{i=1}^{n}$. Let q be a polynomial of degree $n-1$ with complex coefficients $\{a_i\}_{i=1}^{n}$. Denote the distinct eigenvalues of A by $\mu_1, \mu_2, \ldots, \mu_k$ where $k \leq n$. Then $\{\lambda_i\}_{i=1}^{n} = \{\mu_i\}_{i=1}^{k}$ We want to choose the coefficients in q such that $q(A) = A^*$. By Lemma 1.7 it is enough to consider the equality $q([A]_v) = [A^*]_v$, hence we only need to show that there exists a q such that $q(\mu_i) = \overline{\mu_i}$ for $i = 1, 2, \ldots, k$. This system of equations is in matrix form

$$\begin{pmatrix} 1 & \mu_1 & \ldots & \mu_1^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_k & \ldots & \mu_k^{k-1} \end{pmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_{k-1} \end{pmatrix} = \begin{pmatrix} \overline{\mu_1} \\ \vdots \\ \overline{\mu_k} \end{pmatrix}. \tag{1}$$

The k by k matrix $V = (v_{ij}) = (\mu_i^{j-1})$ in (1) is usually known as the *Vandermonde matrix*. Its determinant is $\prod_{i>j\geq 1}^{k}(\mu_i - \mu_j) \neq 0$, since the μ_i's are (pairwise) distinct and thus (1) has a unique solution. □

Let $A \in \mathrm{L}(\mathcal{H})$ be a normal operator with eigenvalues $\{\lambda_i\}_{i=1}^{n}$. Let further $\{v_i\}_{i=1}^{n}$ be an orthonormal basis of eigenvectors of A such that $Av_i = \lambda v_i$. The *Spectral Decomposition* of A is given by

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^*, \qquad 1 \leq i \leq n,$$

and we call $v_i v_i^*$, $1 \leq i \leq n$, the *eigenprojections* of A.

Lemma 1.9. *Any eigenprojection of a normal operator $A \in L(\mathcal{H})$ can be expressed as a polynomial of A.*

Proof. The proof is much similar to the one given for Lemma 1.8 and is therefore omitted. $\square$

An operator $A \in \mathrm{L}(\mathcal{H})$ is Hermitian if $A^* = A$. As usual, we denote by $I \in \mathrm{L}(\mathcal{H})$ the identity operator, i.e., $Iv = v$ for all $v \in \mathcal{H}$. The following lemma may be found in [9].

Lemma 1.10. *The eigenvalues of a normal operator $A \in L(\mathcal{H})$ are collinear*[2] *in the complex plane if and only if A is of the form $A = aH + bI$, for some complex numbers a and b where H is Hermitian and I is the identity operator.*

1.2 Differentiators and compressions

Definition 1.11. Let $\mathcal{H}$ be an n-dimensional Hilbert space, $A \in \mathrm{L}(\mathcal{H})$, ϑ be a unit vector in $\mathcal{H}$ and P be the orthogonal projection onto $\vartheta^{\perp}$. Then we say that $B = PAP_{|P\mathcal{H}}$ is the *compression* of A from $P\mathcal{H}$ to $P\mathcal{H}$.

Example 1.1. Let $A \in L(\mathbb{C}^3)$, let e_1, e_2 and e_3 be the standard basis in $\mathbb{C}^3$ and suppose that

$$[A]_{(e_1,e_2,e_3)} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{22} & a_{33} \end{pmatrix}.$$

Let P be a projection onto $\mathrm{span}\{e_1, e_2\}$; then the associated compression $B = PAP_{|P\mathcal{H}}$ of A in basis (e_1, e_2) is

$$[B]_{(e_1,e_2)} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

In general, if $e = (e_1, e_2, \ldots, e_n)$ is an orthonormal basis in $\mathcal{H}$ and P is a projection onto $e_n^{\perp}$, the matrix $[B]_{(e_1,\ldots,e_{n-1})}$ will be the upper-left hand $n-1$ by $n-1$ principal submatrix of $[A]_e$. Recall that the determinant of every matrix defined by these operators are the same in all bases. Therefore by making use of Cramer's theorem we get the following useful lemma.

Lemma 1.12 (Adjugate relation). *Let A be an invertible operator on an n-dimensional Hilbert space $\mathcal{H}$. Let $e = (e_1, e_2, \ldots, e_n)$ be any basis in $\mathcal{H}$, let*

[2] That is, they lie on a straight line in the complex plane.

P be the projection onto $e_n^\perp$ and let $B = PAP_{|P\mathcal{H}}$. Then the (n,n)th element of $[A]_e^{-1}$ is

$$\frac{\det(B)}{\det(A)} = e_n^* A^{-1} e_n.$$

It was first noticed in [4] that certain relations between p_A and p_B resemble the relations between a polynomial and its derivates. For example, Gauss-Lucas theorem shows that every critical point lies in the convex hull of the roots of a polynomial. When A is normal, one can show that every eigenvalue of B lies in the convex hull of the eigenvalues of A. We therefore study the conditions on P that force the relation $p_B = p_A'/n$.

Definition 1.13. Let $\mathcal{H}$ be an n-dimensional Hilbert space, $A \in \mathrm{L}(\mathcal{H})$, and P a projection from $\mathcal{H}$ onto a subspace of $\mathcal{H}$ having co-dimension one and set $B = PAP_{|P\mathcal{H}}$. Then we shall say that P is a *differentiator* of A if

$$p_B(\lambda) = \frac{1}{n}\frac{d}{d\lambda} p_A(\lambda).$$

Example 1.2. Let $A \in \mathbb{C}^3$, and e_1, e_2, e_3 be the standard basis of $\mathbb{C}^3$. Let P be the projection onto span$\{e_1, e_2\}$ and set $B = PAP_{|P\mathcal{H}}$. Suppose that

$$[A]_{(e_1,e_2,e_3)} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then

$$[B]_{(e_1,e_2)} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

and $p_B(\lambda) = \lambda^2 = p_A'(\lambda)/3$ so P is a differentiator of A.

Now Lemma 1.12 can be used to give a new characterization of differentiators.

Theorem 1.14. *Let $\mathcal{H}$ be a finite dimensional Hilbert space, $A \in L(\mathcal{H})$ and ϑ be a unit vector in $\mathcal{H}$. Let P denote the projection onto $\vartheta^\perp$. Then the following are equivalent.*

(1) P is a differentiator of A.

(2) $\vartheta^*(\lambda I - A)^{-1}\vartheta = \tau((\lambda I - A)^{-1})$ for all $\lambda > \| A\|$.

(3) $\vartheta^* A^i \vartheta = \tau(A^i)$ for every nonnegative integer i.

(4) $\vartheta^* p(A)\vartheta = \tau(p(A))$ for every polynomial p.

Proof. It is well known that $\lambda I - A$ is invertible when $\lambda > \| A\|$. We use the adjugate relation (Lemma 1.12).
(1)$\Rightarrow$(2) Suppose that P is a differentiator; then

$$\begin{aligned}\vartheta^*(\lambda I - A)^{-1}\vartheta =&\frac{p_B(\lambda)}{p_A(\lambda)} = \frac{1}{n}\frac{p'_A(\lambda)}{p_A(\lambda)} = \\ &\frac{1}{n}\sum_{i=1}^{n}(\lambda - \lambda_i(A))^{-1} = \tau\big((\lambda - \lambda_i(A))^{-1}\big).\end{aligned}$$

(2)$\Rightarrow$(1) Suppose that $\vartheta^*(\lambda I - A)^{-1}\vartheta = \tau((\lambda I - A)^{-1})$ for all $\lambda > \| A\|$. Then

$$\begin{aligned}\frac{p_B(\lambda)}{p_A(\lambda)} = \vartheta^*(\lambda I - A)^{-1}\vartheta =&\tau((\lambda I - A)^{-1}) = \\ &\frac{1}{n}\sum_{i=1}^{n}(\lambda - \lambda_i(A))^{-1} = \frac{1}{n}\frac{p'_A(\lambda)}{p_A(\lambda)}.\end{aligned}$$

The equivalence of (2) and (3) follows from the Neumann series (i.e., $(I - A/\lambda)^{-1} = \sum_{i=0}^{\infty}(A/\lambda)^i$ for all $\lambda > \| A\|$) by comparing the coefficients of λ. The equivalence of (3) and (4) is obvious. □

In light of the previous theorem we make the following definition.

Definition 1.15. Let $A \in \mathrm{L}(\mathcal{H})$ and $\vartheta \in \mathcal{H}$. Then we say that ϑ is a trace vector of A if $\vartheta^* p(A)\vartheta = \tau(p(A))$ for all polynomials p.

From Theorem 1.14 we see that there is a $1-1$ correspondence between a differentiator and a trace vector. They relate with each other by the formula $P + \vartheta\,\vartheta^* = I$, and by putting $p = 1$ in the above definition we see that all trace vectors must be of unit length.

We next give an explicit construction of trace vectors (which implies the existence of a differentiators) of normal operators.

Corollary 1.16. *Let $\mathcal{H}$ be an n-dimensional Hilbert space, $A \in L(\mathcal{H})$ be a normal operator and let $e = (e_1, e_2, \ldots, e_n)$ be an orthonormal basis of eigenvectors in $\mathcal{H}$. Set $\vartheta = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} e_i$; then ϑ is a trace-vector of A.*

Proof. Consider the eigenprojection $e_n e_n^*$. By Lemma 1.9, $e_n e_n^*$ is a polynomial of A and therefore ϑ is a trace-vector of A iff $\vartheta^* e_n e_n^* \vartheta = \tau(e_n e_n^*)$. We have that $\vartheta^* e_n e_n^* \vartheta = |\langle \vartheta, e_n\rangle|^2 = \big|\big\langle \frac{1}{\sqrt{n}}\sum_{i=1}^{n} e_i, e_n\big\rangle\big|^2 = 1/n = \tau(e_n e_n^*)$ and thus the proof is complete. □

One has also been able to prove the existence of a trace vector for nonnormal operators, but we will not be needing this result in this paper. Instead we refer to [2] for the interested reader.

Next we give an example of normal operators that share the same differentiator.

Example 1.3. Let $A \in \mathrm{L}(\mathcal{H})$ be a normal operator and let P be a differentiator of A. Then P is also a differentiator of the following operators:

1. The adjoint A^* of A,
2. $\frac{1}{2}(A + A^*)$ and $\frac{1}{2i}(A - A^*)$,
3. $AA^* = A^*A$,
4. Any eigenprojection of A.

These properties follow from Lemma 1.9, Lemma 1.8, Theorem 1.14 and Definition 1.13.

1.3 Majorization

Majorization quantifies the intuitive notion that the components of an n-vector x are less spread out than the components of another such vector y. This is done by means of n inequalities. Hardy, Littlewood and Pólya showed that these inequalities can be expressed as an equality in terms of so called *doubly stochastic matrices.* In turn this led to another characterization of majorization involving arbitrary convex functions. Further studies in this topic have resulted in other characterizations of majorization, as well as generalizations. Specifically Sherman's theorem describes an inequality between two sets of vectors in $\mathbb{R}^m$ which resembles the characterization of the majorization for real numbers by Hardy, Littlewood and Pólya. This theorem will for example allow us to study a type of majorization relation between two sets of complex numbers, not necessarily of the same size. Let us begin with the definition of majorization.

Definition 1.17. Let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ be two n-tuples arranged in descending order. Then we say that $(a_1, \ldots, a_n)$ is majorized by $(b_1, \ldots, b_n)$ if $\sum_{i=1}^{k} a_i \leq \sum_{i=1}^{k} b_i$ for $k = 1, 2, \ldots, n-1$ and $\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i$, and write $(a_1, \ldots, a_n) \prec (b_1, \ldots, b_n)$.

As we already noted, the fact that $(a_1, \ldots, a_n)$ is majorized by $(b_1, \ldots, b_n)$ means roughly that the n-tuple $(a_1, \ldots, a_n)$ is less spread out than $(b_1, \ldots, b_n)$.

In 1929, Hardy, Littlewood and Pólya published the following characterization of majorization [5]. First define a *doubly stochastic matrix* to be a matrix with all real positive entries whose columns and rows sum to 1.

Theorem 1.18. *Let I be any interval in $\mathbb{R}$ and let $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ be two n-tuples of real numbers arranged in descending order. Then the following are equivalent.*

(1) $(a_1, \ldots, a_n) \prec (b_1, \ldots, b_n)$.

(2) There exists a doubly stochastic n by n matrix S such that $a_j = \sum_{i=1}^{n} s_{ij} b_i$ for all $j = 1, 2, \ldots, n$.

(3) $\sum_{i=1}^{n} \phi(a_i) \leq \sum_{i=1}^{n} \phi(b_i)$ *for all convex functions $\phi : I \to \mathbb{R}$.*

For vectors in $\mathbb{R}^m$, one defines (so-called multivariate) majorization in the following way (see [6], chapter 15.)

Definition 1.19. Let A and B be $m \times n$ real matrices. Then we say that A is majorized by B if $A = BS$ for some doubly stochastic matrix S, and write $A \prec B$.

In 1950's Sherman took Definition 1.19 one step further (see [8]) and gave a generalized characterization of multivariate majorization.

Theorem 1.20. *Let A and B be $m \times r$ and $m \times s$ real matrices respectively. Denote a_i^C as the ith column in A for $i = 1, \ldots, r$ and b_i^C the ith column in B for $i = 1, \ldots, s$. Then the following are equivalent.*

1.
$$\frac{1}{r} \sum_{i=1}^{r} \phi(a_i^C) \leq \frac{1}{s} \sum_{i=1}^{s} \phi(b_i^C)$$
for all convex functions $\phi : \mathbb{R}^m \to \mathbb{R}$.

2. *There exists a real $s \times r$ matrix $S = (s_{ij})$ that satisfies the following conditions:*
$$A = BS; \quad s_{ij} \geq 0 \text{ for } 1 \leq i \leq s,\ 1 \leq j \leq r;$$
$$\sum_{i=1}^{s} s_{ij} = 1 \text{ for } 1 \leq j \leq r; \ \sum_{j=1}^{r} s_{ij} = \frac{r}{s} \text{ for } 1 \leq i \leq s.$$

This motivates the following definition.

Definition 1.21. Let S be an $s \times r$ matrix with positive real entries. We say that S is *doubly rectangular stochastic* if $\sum_{i=1}^{s} s_{ij} = 1$ for $1 \leq j \leq r$ and $\sum_{j=1}^{r} s_{ij} = \frac{r}{s}$ for $1 \leq i \leq s$.

2 Schoenberg's conjecture

In this section we give a first application of the theory of differentiators. Namely, we prove Schoenberg's 1986 conjecture [8] on the zeros and critical points of arbitrary complex polynomials. Any monic polynomial can be considered as a characteristic polynomial of a normal operator. This can be realized by for example constructing a diagonal matrix whose diagonal elements are the roots of the polynomial. Schoenberg conjectured that the following holds in the special case when $G = 0$.

Conjecture 2.1 (Schoenberg's conjecture). *Let $p(z)$ be an nth degree polynomial. Let $z_1, z_2, \ldots, z_n$ be the roots of $p(z)$ and let $w_1, w_2, \ldots, w_{n-1}$ be the roots of $p'(z)$. Let $G = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n-1}\sum_{i=1}^{n} w_i$. Then*

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

with equality iff the roots of $p(z)$ are collinear in the complex plane.

Later De Bruin et al. [15, Section 3] and independently Katsoprinakis [10] showed that the case $G = 0$ considered by Schoenberg is equivalent to the more general conjecture stated above.

We see that the left -and right-hand side of the above inequality resemble the Euclidian norm of a matrix. Therefore we will investigate the Euclidian norm of matrices defined by an operator and one of its compressions. Let $\mathcal{H}$ be an n-dimensional Hilbert space, $A \in \mathrm{L}(\mathcal{H})$ be a normal operator with eigenvalues $z_1, z_2, \ldots, z_n$. Choose a basis of eigenvectors $v = (v_1, v_2, \ldots, v_n)$ so that $[A]_v$ becomes a diagonal matrix where $z_1, z_2, \ldots, z_n$ are diagonal elements. Let $\vartheta = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} v_i$ be a trace-vector of A and P its associated differentiator. Let $B = PAP_{|P\mathcal{H}}$ be a compression of A, choose an orthonormal basis $u = (u_1, u_2, \ldots, u_{n-1})$ in $P\mathcal{H}$ and let $\widehat{u} = (u_1, u_2, \ldots, u_{n-1}, \vartheta)$. Then

$$[A]_{\widehat{u}} = \begin{pmatrix} [B]_u & C \\ D^* & \tau(A) \end{pmatrix} \qquad \text{where } C \text{ and } D \text{ are } (n-1)\times 1 \text{ matrices.} \quad (2)$$

Indeed, according to Lemma 1.3, the (n,n)th element of $[A]_{\widehat{u}}$ is

$$\sum_{j=1}^{n} |\, \langle \vartheta, v_j \rangle|^2\, z_j = \frac{1}{n}\sum_{j=1}^{n} z_j = \tau(A).$$

By Example 1.3, P is a also a differentiator of AA^* and A^*A, so we can decompose these operators in the same way as A. We get

$$[AA^*]_{\widehat{u}} = \begin{pmatrix} [B]_u[B]_u^* + CC^* & * \\ * & \|D\|_E + |\tau(A)| \end{pmatrix} \tag{3}$$

and

$$[A^*A]_{\widehat{u}} = \begin{pmatrix} [B]_u^*[B]_u + DD^* & * \\ * & \|C\|_E + |\tau(A)| \end{pmatrix}. \tag{4}$$

Now

$$\begin{aligned}
\|A\|_E^2 &= \|B\|_E^2 + \|C\|_E^2 + \|D\|_E^2 + |\tau(A)|^2 \\
&= \|B\|_E^2 + \big(\|C\|_E^2 + |\tau(A)|^2\big) + \big(\|D\|_E^2 + |\tau(A)|^2\big) - |\tau(A)|^2 \\
&= \|B\|_E^2 + \vartheta^*[AA^*]_{\widehat{u}}\vartheta + \vartheta^*[A^*A]_{\widehat{u}}\vartheta - |\tau(A)|^2 \\
&= \Big[\text{since } AA^* = A^*A \text{ and } [A]_{\widehat{u}}\vartheta = \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i v_i\Big] \\
&= \|B\|_E^2 + \frac{2}{n}\|A\|_E - |\tau(A)|^2.
\end{aligned}$$

It follows that

$$\|B\|_E^2 = |\tau(A)|^2 + \frac{n-2}{n}\|A\|_E^2. \tag{5}$$

This result and the next theorem, a classical inequality by Schur actually proves the inequality part of Schoenberg's conjecture. We use the notation $\lambda_i(A)$ to denote the ith eigenvalue of an operator A.

Theorem 2.2. *Let A be an operator on an n-dimensional Hilbert space and let $Re\,A = \frac{1}{2}(A + A^*)$ and $Im\,A = \frac{1}{2i}(A - A^*)$ Then $\sum_{i=1}^n |\lambda_i(A)|^2 \le \|A\|_E^2$ and $\sum_{i=1}^n |\lambda_i(Re\,A)|^2 \le \|Re\,A\|_E^2$ and $\sum_{i=1}^n |\lambda_i(Im\,A)|^2 \le \|Im\,A\|_E^2$ with equality in any one of the above relations implying the equality in all three and occurring iff A is normal.*

Let p be an nth degree polynomial whose roots are $z_1, \ldots, z_n$ and critical points are $w_1, \ldots, w_{n-1}$ and let $G = \frac{1}{n}\sum_{i=1}^n z_i = \frac{1}{n-1}\sum_{i=1}^{n-1} w_i$. Let A be a normal operator whose characteristic polynomial is $p(z)$, let P be a differentiator of A and let $B = PAP_{|P\mathcal{H}}$. Then by Theorem 2.2 and (5)

$$\sum_{i=1}^{n-1} |w_i|^2 \le \|B\|_E^2 = |\tau(A)|^2 + \frac{n-2}{n}\|A\|_E^2 = |G|^2 + \frac{n-2}{n}\sum_{i=1}^n |z_i|^2.$$

Equality holds iff B is normal. The next proposition (which is also a special case of [9, Theorem 2]) will complete the proof of Conjecture 2.1.

Proposition 2.3. *Let $A \in L(\mathcal{H})$ be normal. Let P be a differentiator and let $B = PAP_{|P\mathcal{H}}$. Then B is normal if and only if all the eigenvalues of A are collinear in the complex plane.*

Proof. Suppose that the eigenvalues of A lie on a straight line in the complex plane. According to Lemma 1.10 we may write A of the form $A = aH + bI$ where H is Hermitian, I is the identity and a, b are complex numbers. The compression $B = PAP_{|P\mathcal{H}} = aPHP_{|P\mathcal{H}} + bI_{\,|P\mathcal{H}}$ where $PHP = PH^*P$ and $P = P^*$ implies that $PHP_{|P\mathcal{H}}$ is Hermitian. Hence B is normal.

To prove the other direction, consider the case where $\tau(A) = 0$. Let $u = (u_1, u_2, \ldots, u_{n-1})$ be an orthonormal basis of eigenvectors of B, $\hat{u} = (u_1, u_2, \ldots, u_{n-1}, u_n)$ be an orthonormal basis in $\mathcal{H}$, and let C and D be as in (2). Since $AA^* = A^*A$ and $BB^* = B^*B$, (3) and (4) implies that $CC^* = DD^*$. The $q = 1$ case of [13, Theorem 3.1] (or the $l = 1$ case of [14, Lemma 2]) states that $CC^* = DD^*$ if and only if $C = \omega D$ for some complex number ω of modulus one. Let $S = A - \omega A^*$ and $T = B - \omega B^* = PTP_{|P\mathcal{H}}$. Both S and T are normal and

$$[S]_{\hat{u}} = \begin{pmatrix} [T]_u & \mathbf{0} \\ \mathbf{0}^* & 0 \end{pmatrix}.$$

Therefore $p_S(\lambda) = \lambda p_T(\lambda)$ and $p_T(\lambda) = p'_S(\lambda)/n$. Hence, $p_T(\lambda) = \lambda^n$ and $S = 0$. Thus

$$A = \omega A^* \text{ and } A = \frac{\omega}{\omega + 1}(A + A^*),$$

so A is a complex multiple of an Hermitian operator. Therefore the eigenvalues are collinear in the complex plane by Lemma 1.10. For the case when $\tau(A) \neq 0$ we know that $\hat{A} = A - \tau(A)I$ is a complex multiple of an Hermitian operator, which give us the relation $A = \hat{A} + \tau(A)I$. □

By using Schur's inequality for the real and imaginary parts of eigenvalues and following the same argument as above, we can also express a Schoenberg-like inequality for the real -and imaginary parts of the roots and critical points of a given polynomial.

Theorem 2.4. *Let $p(z)$ be an nth degree polynomial. Let $z_1, \ldots, z_n$ be the roots of $p(z)$, let $w_1, \ldots, w_{n-1}$ be the roots of $p'(z)$, and let*

$$G = \frac{1}{n}\sum_{i=1}^{n} z_i = \frac{1}{n-1}\sum_{i=1}^{n-1} w_i.$$

Then

$$\sum_{i=1}^{n-1} |\,Re\, w_i|^2 \leq \|\,Re\, B\|_E^2 = |\,Re\, G|^2 + \frac{n-2}{n}\sum_{i=1}^{n} |\,Re\, z_i|^2$$

and

$$\sum_{i=1}^{n-1} |\,Im\, w_i|^2 \leq \|\,Im\, B\|_E^2 = |\,Im\, G|^2 + \frac{n-2}{n}\sum_{i=1}^{n} |\,Im\, z_i|^2$$

with equality iff all the roots of $p(z)$ are collinear on the complex plane.

3 Katsoprinakis conjecture

In this section we state and solve a conjecture due to Katsoprinakis [10]. Let us first make the following definition.

Definition 3.1. Let $p(z)$ be a polynomial whose roots are $\{z_i\}_{i=1}^n$ and $p^*(z)$ be the polynomial whose roots are $\{\operatorname{Re} z_i\}_{i=1}^n$. Let $\{w_i\}_{i=1}^{n-1}$ be the critical points of $p(z)$ and let $\{w_i^*\}_{i=1}^{n-1}$ be the critical points of $p^*(z)$. Then we say that $p(z)$ satisfies the *majorization condition* if $(\operatorname{Re} w_1, \dots, \operatorname{Re} w_{n-1}) \prec (w_1^*, \dots, w_{n-1}^*)$.

We note that by Theorem 1.18, a polynomial $p(z)$ also satisfies the majorization condition if $\sum_{i=1}^{n-1} \phi(\operatorname{Re} w_i) \leq \sum_{i=1}^{n-1} \phi(w_i^*)$ for any convex function $\phi : \mathbb{R} \to \mathbb{R}$.

Katsoprinakis conjectured that every polynomial actually satisfies the majorization condition. We give an example to illustrate the definition.

Example 3.1. Let $p(z) = z^4 - 1$, the roots of p are $\{1, -1, i, -i\}$ which have real parts $\{1, -1, 0, 0\}$. Hence $p^*(z) = (z-1)(z+1)z^2 = z^4 - z^2$ which has critical points $\{0, 1/\sqrt{2}, -1/\sqrt{2}\}$. The critical points of $p(z)$ are $\{0, 0, 0\}$, so $p(z)$ satisfies the majorization condition.

By using theory of diffentiators we can solve Katsoprinakis conjecture. First we need the following result by Ky Fan [11] (see also [6, Theorem 9.F.1]) which describes a majorization relation between eigenvalues of an operator and its real part operator.

Lemma 3.2. *Let A be an operator on an n-dimensional Hilbert space and let $\operatorname{Re} A = (1/2)(A + A^*)$. Then we have the majorization relation*

$$(\operatorname{Re} \lambda_1(A), \operatorname{Re} \lambda_2(A), \dots \operatorname{Re} \lambda_n(A)) \prec (\lambda_1(\operatorname{Re} A), \lambda_2(\operatorname{Re} A), \dots, \lambda_n(\operatorname{Re} A)).$$

Theorem 3.3 (Katsoprinakis' conjecture). *Every polynomial satisifes the majorization condition.*

Proof. Let $A \in \mathrm{L}(\mathcal{H})$ be a normal operator with characteristic polynomial p_A and eigenvalues $z_1, \dots, z_n$. Then $\operatorname{Re} A$ is a normal operator with eigenvalues $\operatorname{Re} z_1, \dots, \operatorname{Re} z_n$ and characteristic polynomial p_A^* (which we see for example by choosing an orthonormal basis of eigenvectors of A.) Let the critical points of p_A and p_A^* be $w_1, \dots, w_{n-1}$ and $w_1^*, \dots, w_{n-1}^*$ respectively. The operator $\operatorname{Re} A$ has the same differentiator as A (cf. Example 1.3) and

$$P\Big(\frac{1}{2}(A + A^*)\Big)P_{|P\mathcal{H}} = \frac{1}{2}(B + B^*) = \operatorname{Re} B,$$

so the eigenvalues of $\operatorname{Re} B$ are $w_1^*, \dots, w_{n-1}^*$ Thus by Lemma 3.2 we have $(\operatorname{Re} w_1, \dots, \operatorname{Re} w_{n-1}) \prec (w_1^*, \dots, w_{n-1}^*)$. $\square$

In [10, Proposition 2.g] Katsoprinakis has shown that Theorem 2.4 would follow from Theorem 3.3. Hence this section along with [10, Proposition 2.g] would give us a second proof of Schoenberg's conjecture. He also showed in [10] that Theorem 3.3 would imply a whole family of inequalities between roots and critical points of a polynomial.

4 De Bruijn-Springer conjecture

In the mid 1940's several papers were written about the inequalities of the form

$$\frac{1}{n-1}\sum_{i=1}^{n-1}\phi(w_i) \le \frac{1}{n}\sum_{i=1}^{n}\phi(z_i), \tag{6}$$

where $\phi : \mathbb{C} \to \mathbb{R}$, p is an arbitrary polynomial, $n = \deg(p)$, $\{z_i\}_{i=1}^n$ are the roots of p, and $\{w_i\}_{i=1}^{n-1}$ are the critical points of p.

Such questions were studied by Erdos, Niven, de Bruijn, Springer among others. In particular, de Bruijn and Springer showed that any continuous function ϕ that satisfies (6) for all complex polynomials must be convex. They also proved that (6) is true for all convex functions ϕ and polynomials p with all real zeros as well as for convex functions $\phi : \mathbb{C} \to \mathbb{R}$ of the form $\phi(z) = |\, z|^r$, $r \ge 1$. It was natural to conjecture that (6) actually holds for all convex functions, which de Bruijn and Springer did in [12].

Conjecture 4.1. *Let $p(z)$ be an arbitrary complex polynomial with roots $\{z_i\}_{i=1}^n$ and critical points $\{w_i\}_{i=1}^{n-1}$. Then*

$$\frac{1}{n-1}\sum_{i=1}^{n-1}\phi(w_i) \le \frac{1}{n}\sum_{i=1}^{n}\phi(z_i),$$

where $\phi : \mathbb{C} \to \mathbb{R}$ is any convex function.

We shall give a proof of this conjecture by again using the tools in Section 1. We first note that by Sherman's theorem (Theorem 1.20), the de Bruijn-Springer conjecture (Conjecture 4.1) is in fact equivalent to a generalized majorization relation between zeros and critical points of a polynomial.

Theorem 4.2 (De Bruijn-Springer conjecture). *Let p be an arbitrary complex polynomial whose roots are $\{z_i\}_{i=1}^n$ and whose critical points are $\{w_i\}_{i=1}^{n-1}$. Then there exists a doubly rectangular stochastic matrix S such that*

$$(w_1, w_2, \ldots, w_{n-1}) = (z_1, z_2, \ldots, z_n)S.$$

Proof. Let $A \in \mathrm{L}(\mathcal{H})$ be a normal operator whose eigenvalues are $\{z_i\}_{i=1}^n$. Choose an orthonormal basis of eigenvectors $\{v_i\}_{i=1}^n$ of A. Let further $\vartheta = \frac{1}{\sqrt{n}}\sum_{i=1}^n v_i$ be a trace vector of A, P be its associated differentiator and $B = PAP_{|P\mathcal{H}}$. Choose a Schur basis $u = (u_1, u_2, \ldots, u_{n-1})$ that triangulizes B. Then $w_i = u_i^* B u_i = u_i^* A u_i$ for $i = 1, 2, \ldots, n-1$. Recall from Lemma 1.3 that $u_i = \sum_{j=1}^n v_j \langle u_i, v_j\rangle$, so

$$w_i = \sum_{j=1}^n z_j \,|\, \langle u_i, v_j\rangle|^2 .$$

Let $S = (s_{ij})$ denote the $n \times n-1$ matrix where $s_{ij} = |\langle v_i, u_j\rangle|^2$. By Parseval's theorem S is doubly rectangular stochastic since

$$\| u_j \| = \sum_{i=1}^{n} | v_i, u_j |^2 = 1, \qquad 1 \leq j \leq n-1$$

and

$$\begin{aligned}
\| v_i \| &= \sum_{j=1}^{n-1} | \langle v_i, u_j \rangle|^2 + \left| \left\langle v_i, \frac{1}{\sqrt{n}} \sum_{i=1}^{n} v_i \right\rangle \right|^2 \\
&= \sum_{j=1}^{n-1} | \langle v_i, u_j \rangle|^2 + \frac{1}{n} \\
&\Rightarrow \quad \sum_{j=1}^{n-1} | \langle v_i, u_j \rangle|^2 = \frac{n-1}{n}, \quad 1 \leq i \leq n.
\end{aligned}$$

This completes the proof. $\square$

We note that Theorem 4.2 has been generalized in [3] by Malamud where he showed that

$$\frac{1}{\binom{n-1}{k}} \sum_{1 \leq i_i < \ldots < i_k \leq n-1} \phi\left(\prod_{j=1}^{k} w_{i_j} \right) \leq \frac{1}{\binom{n}{k}} \sum_{1 \leq i_i < \ldots < i_k \leq n} \phi\left(\prod_{j=1}^{k} z_{i_j} \right).$$

.

References

[1] Q. I. Rahman, G. Schmeisser, *Analytic theory of polynomials,* London Math. Soc. Monogr. (N.S.) vol. **26**, Oxford Univ. Press, New York, NY, 2002.

[2] R. Pereira, *Differentiators and the geometry of polymomials*, J. Math. Anal. Appl. **285** (2003) 336-348.

[3] S. M. Malamud, *Inverse spectral problem for normal matrices and a generalization of the Gauss-Lucas theorem,* arXiv:math.CV/0304158v1-v3.

[4] C. Davis, *Eigenvalues of compressions*, Bull. Math. Soc. Sci. Math. Phys. RPR **51** (1959) 3-5.

[5] G. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge University Press, 1988.

[6] A. W. Marshall and Ingram Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, 1979.

[7] S. Sherman, *On a theorem of Hardy, Littlewood, Pólya and Blackwell*, Proc. Nat. Acad. Sci. U.S.A. **37** (1951) 826-831.

[8] I. J. Schoenberg, *A conjectured analogue of Rolle's theorem for polynomials with real or complex coefficients*, Amer. Math. Monthly **93** (1986) 8-13.

[9] K. Fan and G. Pall, *Imbedding conditions for Hermitian and normal matrices*, Canad. J. Math. **9** (1957) 298-304.

[10] E. S. Katsoprinakis, *On the complex role set of a polynomial*, in: N. Papamichael, St. Ruschuwyeh and E. B. Saff (Eds.), *Computational Methods and Function Theory*, 1997 (Nicosia), in: *Series in Approximations and Decompositions, Vol. 11*, World Scientific, River Edge, NJ, 1999.

[11] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations (II)*, Proc. Nat. Acad. Sci. U.S.A. **36** (1950) 31-35.

[12] N. G. de Bruijn, *On the zeros of polynomial and its derivate (II)*, Indagationes Mathematicae **9** (1947), 264-270.

[13] R. A. Horn, I. Olkin, *When does $A^*A = B^*B$ and why does one want to know?*, Amer. Math. Monthly **103** (1996) 470-482.

[14] Kh. D. Ikramov, L. Elsner, *On normal matrices with normal principal submatrices*, J. Math. Sci. **89** (1998) 1631-1651.

[15] M. G. de Bruin, K. G. Ivanov, A. Sharma, *A conjecture of Schoenberg*, J.Inequal, Appl. **4** (1999) 183-213.