TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 357, Number 10, Pages 4043–4064
S 0002-9947(04)03649-9
Article electronically published on September 23, 2004

INVERSE SPECTRAL PROBLEM FOR NORMAL MATRICES AND THE GAUSS-LUCAS THEOREM

S. M. MALAMUD

Abstract. We establish an analog of the Cauchy-Poincare interlacing theorem for normal matrices in terms of majorization, and we provide a solution to the corresponding inverse spectral problem. Using this solution we generalize and extend the Gauss–Lucas theorem and prove the old conjecture of de Bruijn-Springer on the location of the roots of a complex polynomial and its derivative and an analog of Rolle's theorem, conjectured by Schoenberg.

1. Introduction

Let $A = A^*$ be a Hermitian $n \times n$ matrix and A_{n-1} its principal $(n-1)\times(n-1)$ submatrix obtained by deleting the last row and column. Let $\sigma(A) = (\lambda_1 \le \ldots \le \lambda_n)$ and $\sigma(A_{n-1}) = (\mu_1 \le \ldots \le \mu_n)$ be the ordered lists of (real) eigenvalues of A and A_{n-1}, respectively. By the Cauchy–Poincare interlacing theorem, the sequences interlace each other, that is,

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \ldots \le \lambda_{n-1} \le \mu_{n-1} \le \lambda_n. \tag{1.1}$$

It is known (see [18], [26]) that the converse is also true, that is, for any two sequences $(\lambda_j)_1^n$ and $(\mu_j)_1^{n-1}$ of real numbers satisfying (1.1), there exists a (non-unique) $n\times n$ Hermitian matrix A such that $\sigma(A) = \{\lambda_j\}_1^n$ and $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$. We say that such a matrix A solves the inverse spectral problem for the sequences $(\lambda_j)_1^n$ and $(\mu_j)_1^{n-1}$. A theorem of Hochstadt [17] asserts that there exists a unique Jacobi, that is, tridiagonal Hermitian matrix A solving the inverse spectral problem for these sequences.

In the present paper we generalize both the Cauchy-Poincare interlacing theorem and the Hochstadt theorem to the case of normal matrices. Since the eigenvalues of a normal matrix are not real, the statements of these generalizations must be different from the Hermitian case. Moreover, the principal submatrix A_{n-1} of a normal matrix A is only normal in trivial cases (see [12] or Lemma 4.14 below).

The main result of the first part of the paper is Theorem 3.11. It provides necessary and sufficient geometric conditions for two sequences $\{\lambda_j\}_1^n$ and $\{\mu_j\}_1^{n-1}$ to be the spectra of a normal matrix A and of its submatrix A_{n-1} respectively. In order to formulate these geometric conditions we introduce in Section 2 several notions of majorization for sequences of vectors from $\mathbb{R}^n$, which are natural generalizations of the classical notions for the case $n = 1$.

Received by the editors July 6, 2003 and, in revised form, November 7, 2003.
2000 *Mathematics Subject Classification.* Primary 15A29; Secondary 30C15, 30C10.
Key words and phrases. Zeros of polynomials, normal matrices, inverse spectral problem, majorization.

©2004 American Mathematical Society

Note that the sufficient conditions for the tuples $\{\lambda_j\}_1^n$ and $\{\mu_j\}_1^{n-1}$ of complex numbers to be the spectra of a normal matrix A and its principal submatrix A_{n-1} respectively can be expressed analytically without the notion of majorization and read

$$c_k := \mathrm{Res}_{\lambda_k}\left(\frac{\prod_{j=1}^{n-1}(\lambda-\mu_j)}{\prod_{j=1}^{n}(\lambda-\lambda_j)}\right) \geq 0, \qquad k \in \{1,\dots,n\}. \tag{1.2}$$

The second part of our paper concerns the location of the roots of a polynomial and its derivative and in particular the Gauss-Lucas theorem [33], stating that the roots $\{\mu_k\}_1^{n-1}$ of the derivative p' of any complex polynomial $p(\in \mathbb{C}[z])$ of degree n lie in the convex hull of the roots $\{\lambda_j\}_1^n$ of p. The following proposition provides a "bridge" between the first and second parts of the paper.

Proposition 1.1. *Let $p(z)$ be a polynomial of degree n with zeros $\{\lambda_j\}_1^n$ and $\{\mu_j\}_1^{n-1}$ the zeros of its derivative. Then there exists a (nonunique) normal matrix $A \in M_n(\mathbb{C})$ such that $\sigma(A) = \{\lambda_j\}_1^n$ and $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$.*

Proposition 1.1 immediately follows from (1.2) since in this case $c_k = 1,\ k \in \{1,\dots,n\}$.

Combining Proposition 1.1 and Theorem 3.11 below, we immediately arrive at the following result (Proposition 4.3):

$$\{\mu_j\}_1^{n-1} \prec_{uds} \{\lambda_j\}_1^n.$$

This is a generalization of the Gauss-Lucas theorem. Proposition 1.1 allows us to apply linear algebra to study the location of the zeros of a polynomial and its derivative. For example, using exterior algebra we obtain a generalization of the Gauss-Lucas theorem for the products of roots (see Proposition 4.3 and Corollary 4.4). In particular we show that $\mu_{k_1}\mu_{k_2} \in \mathrm{conv}\{\lambda_{j_1}\lambda_{j_2} : 1 \leq j_1 < j_2 \leq n\}$ for any pair $\{k_1, k_2\}$ with $1 \leq k_1 \neq k_2 \leq n-1$.

Note that while the second topic has attracted a lot of attention during the last two decades (see [2], [4], [7], [9], [29], [30], [36], [39]), our approach seems to be new and promising. In particular in the framework of this approach we get simple (and short) solutions to the old problems of de Bruijn-Springer [8] and Schoenberg [38].

Let us briefly describe these problems.

In 1947 de Bruijn and Springer [8] conjectured that the following inequality holds for any convex function $f : \mathbb{C} \to \mathbb{R}$:

$$\frac{1}{n-1}\sum_{j=1}^{n-1} f(\mu_j) \leq \frac{1}{n}\sum_{j=1}^{n} f(\lambda_j).$$

They succeeded in proving this inequality for a class of convex functions. It is not difficult to see that this class actually coincides with the class $CVS(\mathbb{C})$, defined in (2.10) below.

We provide a complete proof of this conjecture by showing that a doubly stochastic matrix S in the above-mentioned representation $\mu = S\lambda$ can be chosen in such a way that all entries in the last row equal $1/n$. Moreover, we prove (see

Theorem 4.10) that the following inequality is valid for any $k \in \{1, \dots, n-1\}$

$$\begin{aligned} &\frac{1}{\binom{n-1}{k}} \sum_{1 \le i_i < \dots < i_k \le n-1} f\left(\prod_{j=1}^{k} (\mu_{i_j} - \alpha)\right) \\ &\qquad \le \frac{1}{\binom{n}{k}} \sum_{1 \le i_1 < \dots < i_k \le n} f\left(\prod_{j=1}^{k} (\lambda_{i_j} - \alpha)\right). \end{aligned} \tag{1.3}$$

In 1986 Schoenberg [38] (see also [9]) conjectured that if $\sum_{j=1}^{n} \lambda_j = 0$, then

$$n \sum_{j=1}^{n-1} |\mu_j|^2 \le (n-2) \sum_{j=1}^{n} |\lambda_j|^2 \tag{1.4}$$

and the equality holds if and only if all numbers λ_j lie on the same line. We prove this conjecture below (Theorem 4.15).

Let us briefly sketch the contents of the paper.

In Section 2 we introduce two new notions of majorization for sequences of vectors with non-equal numbers of entries and establish some simple properties of those. We also study the connections between these two notions of majorization and show that they are not equivalent even in the two-dimensional case. This fact provides an answer to a question of Marshall and Olkin [27], p. 433.

In Section 3 we establish an analog of the Cauchy-Poincare interlacing theorem for normal matrices (Theorem 3.11). As a corollary we prove an analogous statement for "convex" combinations of normal matrices with non-commuting weights (Corollary 3.15). We also prove an analog of the Hochstadt theorem for normal matrices.

In Section 4 we apply the results of Section 3 to generalize and extend the Gauss–Lucas theorem and prove the de Bruijn-Springer and Schoenberg conjectures. Finally, we prove an analog of the de Bruijn-Springer conjecture for zeros of the Mason-Shapiro polynomials (see Theorem 4.17).

A preliminary version of this paper has already been published as a preprint [23]. The main results of the paper have been announced without proofs in [24].

2. Majorization

Notation. As usual, $\mathbb{C}^{m\times n}$ ($\mathbb{R}^{m\times n}$) stands for the set of all $m \times n$ matrices with complex (real) entries; $M_n(\mathbb{R}) := \mathbb{R}^{n\times n}$, $M_n(\mathbb{C}) := \mathbb{C}^{n\times n}$.

We denote by $C := A \circ B$ the Schur (elementwise) product of two $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ defined by $(c_{ij}) := (a_{ij} b_{ij})$.

We set $\mathrm{col}(a_1, \dots, a_n) = (a_1, \dots, a_n)^t$ as the vector column.

The convex hull of a set $X \subset \mathbb{R}^m$, i.e., the smallest convex set containing X, is denoted by $\mathrm{conv}(X)$.

The set of extreme points of a convex set X is referred to as $\mathrm{Ext}X$.

The abbreviation $\mathrm{CV}(Y)$ stands for the set of all convex functions on a convex set Y.

2.1. Two definitions of majorization.

We start with several known definitions (see [16], [26], [27]).

Definition 2.1. (a) A matrix $A \in M_n(\mathbb{R})$ is called doubly stochastic if all its entries are nonnegative and the sum of the elements in each row and each column equals one.

We denote the set of doubly stochastic matrices by Ω_n ($\subset M_n(\mathbb{R})$).

(b) A matrix $A \in M_n(\mathbb{R})$ is called unitarily-stochastic (orthostochastic) if there exists a unitary (orthogonal) matrix $U \in M_n(\mathbb{C})(U \in M_n(\mathbb{R}))$ such that A is the Schur product of U and $\bar{U} : A = U \circ \bar{U}$. The set of all unitarily-stochastic matrices is denoted by Ω_n^u. Each unitarily-stochastic $n \times n$ matrix is doubly stochastic, i.e. $\Omega_n^u \subset \Omega_n$. The converse is not true: not every doubly stochastic matrix is unitarily-stochastic (see [26]).

Remark 2.2. The set of all doubly stochastic matrices is convex and contains all permutation matrices. A theorem of Birkhoff states that the set $\mathrm{Ext}(\Omega_n)$ coincides with the set of permutation matrices and thus, by the Krein-Mil'man theorem, Ω_n is the convex hull of all permutation matrices (see [16]).

We recall the notion of majorization ([16], [26], [27]) for sequences of real numbers.

Definition 2.3. Let there be given two real sequences $\alpha := (\alpha_k)_1^m$ and $\beta := (\beta_k)_1^m$. Let also $\hat{\alpha}$ and $\hat{\beta}$ be these sequences, reordered to be decreasing. We write $\beta \prec\prec \alpha$ if

$$\hat{\beta}_1 + \ldots + \hat{\beta}_j \le \hat{\alpha}_1 + \ldots + \hat{\alpha}_j, \qquad j \in \{1, \ldots, m\}. \tag{2.1}$$

If in addition

$$\sum_{k=1}^m \alpha_k = \sum_{k=1}^m \beta_k, \tag{2.2}$$

then the sequence β is said to be majorized by α which is denoted by $\beta \prec \alpha$.

The following famous theorem is due to Weyl, Birkhoff and Hardy-Littlewood-Polya (see [15], [16], [26], [27]):

Theorem 2.4. *Let $\alpha, \beta \in \mathbb{R}^m$ be two real sequences. Then the following five conditions are equivalent:*

(1) $\beta \prec \alpha$.

(2) *The following inclusion holds:*

$$\begin{aligned}\beta &\in \mathrm{conv}(\{A\alpha : \ A \textit{ is a permutation matrix}\}) \\ &= \mathrm{conv}(\{(\alpha_{i_1}, \ldots, \alpha_{i_m}) : \ \{i_1, \ldots, i_m\} \textit{ is a permutation of the set } \{1, \ldots, m\}\}).\end{aligned}$$

(3) *There exists a doubly stochastic matrix $S \in M_m(\mathbb{R})$ such that $\beta = S\alpha$. In fact, the matrix S can be chosen to be orthostochastic.*

(4) *The inequality*

$$\sum_{i=1}^m f(\beta_i) \le \sum_{i=1}^m f(\alpha_i) \tag{2.3}$$

is valid for any convex function f on $\mathbb{R}$.

(5) *The inequality*

$$f(\beta_1, \ldots, \beta_m) \le f(\alpha_1, \ldots, \alpha_m) \tag{2.4}$$

is valid for any convex function f on $\mathbb{R}^m$, invariant under the action of the permutation group.

Remark 2.5. Note that (2.2) and (2.1) imply

$$\hat{\beta}_{j+1} + \ldots + \hat{\beta}_m \geq \hat{\alpha}_{j+1} + \ldots + \hat{\alpha}_m, \qquad j \in \{1, \ldots, m-1\}. \tag{2.5}$$

Since the convex hull of a set of real numbers is the segment between the minimal and the maximal one, we can rewrite the condition of majorization as

$$\begin{aligned} \operatorname{conv}(\beta_{i_1} + \ldots + \beta_{i_j} :\ 1 \leq i_1 < \ldots < i_j \leq m) \\ \subset \operatorname{conv}(\alpha_{i_1} + \ldots + \alpha_{i_j} :\ 1 \leq i_1 < \ldots < i_j \leq m). \end{aligned} \tag{2.6}$$

This observation motivates the following definition.

Definition 2.6. Let $x = \{x_k\}_{k=1}^l$ and $y = \{y_k\}_{k=1}^m$ be two tuples of vectors in $\mathbb{R}^n$ and $l \leq m$. Suppose that the following conditions are fulfilled:

$$\begin{aligned} &\operatorname{conv}(x_{i_1},\ i_1 \in \{1, \ldots, l\}) \subset \operatorname{conv}(y_{i_1} :\ i_1 \in \{1, \ldots, m\}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ &\operatorname{conv}(x_{i_1} + \ldots + x_{i_k} :\ 1 \leq i_1 < \ldots < i_k \leq l) \\ &\subset \operatorname{conv}(y_{i_1} + \ldots + y_{i_k} :\ 1 \leq i_1 < \ldots < i_k \leq m) \\ &\ldots\ldots\ldots\ldots\ldots \\ &x_1 + \ldots + x_l \in \operatorname{conv}(y_{i_1} + \ldots + y_{i_l} :\ 1 \leq i_1 < \ldots < i_l \leq m). \end{aligned} \tag{2.7}$$

Then we say that the sequence $\{y_k\}$ majorizes the sequence $\{x_k\}$ and write $x \prec y$;

Remark 2.7. If $l = m$, the last condition in (2.7) turns into

$$x_1 + \ldots + x_m = y_1 + \ldots + y_m.$$

Theorem 2.4, (3) suggests another analog of the notion of majorization in the multi-dimensional case.

Definition 2.8. Let $x := (x_k)_{k=1}^l$ and $y := (y_k)_{k=1}^m$, $l \leq m$ be two sequences of vectors $x_k, y_k \in \mathbb{R}^n$. We say that x is doubly stochastically majorized by y and write $x \prec_{ds} y$ if there exist vectors $x_{l+1}, \ldots, x_m \in \mathbb{R}^n$ and a doubly stochastic matrix $S \in \Omega_m$ such that $\widetilde{x} := (x_k)_1^m = (I_n \otimes S)y$. If S can be chosen to be unitarily-stochastic, we write $x \prec_{uds} y$.

Remark 2.9. Note that in the case $n = 1$ and $l = m$ Definitions 2.6 and 2.8 are equivalent to the above Definition 2.3 of majorization for real sequences due to Remark 2.5 and Theorem 2.4, (3).

It is easy to see that the following proposition is valid.

Proposition 2.10. *If $x = \{x_k\} \prec_{ds} y = \{y_k\}$, then $x \prec y$.*

One might think that a complete analog of Theorem 2.4 could be valid in the multi-dimensional case, that is, the partial orders $\prec$ and $\prec_{ds}$ could be equivalent. The following example shows that this is not the case.

Example 2.11. Let $n = 2$ and $m = 4$. Set

$$\begin{aligned} x &= \{x_1, \ldots, x_4\} = \{(12, 12), (12, 12), (5, 3), (3, 5)\}, \\ y &= \{y_1, \ldots, y_4\} = \{(8, 16), (16, 8), (0, 0), (8, 8)\}. \end{aligned} \tag{2.8}$$

It is easy to check by hand that $x \prec y$. At the same time it is easy to see that the vector $x_1 = (12, 12)$ can be uniquely expressed as a convex combination of y_k's:

$x_1 = 1/2(y_1 + y_2)$. Suppose now that there exists a doubly stochastic matrix S, such that $x = Sy$. Then S has the form

$$S = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & s_{33} & s_{34} \\ 0 & 0 & s_{43} & s_{44} \end{pmatrix},$$

which is impossible since x_3, x_4 do not belong to the convex hull of y_3, y_4.

Remark 2.12. Definition 2.8 for tuples with equal numbers of elements (i.e. for $l = m$) is contained in [27]. Marshall and Olkin use the sign $\prec$ to denote this partial order. On p. 433 of [27] they suggest another notion of majorization for tuples of vectors with equal numbers of elements. It is not difficult to see that this notion is equivalent to the one introduced in our Definition 2.6 for $l = m$. Marshall and Olkin mention that the relation between $\prec$ and $\prec_{ds}$ is not clear. Example 2.11 provides an answer to this question.

We note the following simple

Proposition 2.13. *Let $l = m$.*

(a) *Let $\{y_k\}_{k=1}^m$ be such that $\operatorname{conv}(\{y_k\}_{k=1}^m)$ is affine isomorphic to the standard simplex*

$$\Sigma_{m-1} := \{(t_1, \dots, t_m) \in \mathbb{R}^m : \ t_k \geq 0, \ \sum_k t_k = 1\}.$$

Then $x = \{x_k\} \prec y = \{y_k\}$ if and only if $x \prec_{ds} y$.

(b) *For $m = 3$ the orders $\prec$ and $\prec_{ds}$ are equivalent.*

Proof. (a) Under assumption (a) each x_k can be uniquely represented as a convex combination of y_k:

$$x_k = \sum_{i=1}^m s_{ki} y_i, \qquad \sum_{i=1}^m s_{ki} = 1, \qquad k \in \{1, \dots, m\}.$$

By definition, $x \prec y$ yields

$$(1/m) \sum_k y_k = (1/m) \sum_k x_k = (1/m) \sum_{i=1}^m \left(\sum_{k=1}^m s_{ki} \right) y_i.$$

Defining

$$\beta_i = 1/m \left(\sum_{k=1}^m s_{ki} \right)$$

we get $\beta_i = 1/m$ for all i since the expression via the extreme points is unique. Therefore the matrix $S = (s_{ki})_{k,i=1}^m$ is doubly stochastic.

(b) $m = 3$. If the vectors $y_2 - y_1$ and $y_3 - y_1$ are not co-linear, then (a) applies. If they are, the problem is reduced to the case $n = 1$, contained in Theorem 2.4. $\square$

Remark 2.14. It is clear that if $\{y_k\}$ are linearly independent, then they satisfy the assumptions of (a). It is interesting to note that in the case (a) the first condition $x_k \in \operatorname{conv}(\{y_k\}_{k=1}^m)$ together with the last one $\sum x_k = \sum y_k$ are already enough to imply the existence of a doubly stochastic matrix.

2.2. **The set of extreme points of the set Maj(y).** Let $n = 1$ and $x = (x_k)_1^m,\ y = (y_k)_1^m \in R^m$. It is known [25] that in this case the set of extreme points of the set $\mathrm{Maj}(y) := \{x :\ x \in \mathbb{R}^m,\ x \prec y\}$ is

$$\mathrm{Ext}(\mathrm{Maj}(y)) = \{Py :\ P \text{ is a permutation matrix}\}. \tag{2.9}$$

The following statement easily follows from the scalar case.

Proposition 2.15. *Let $x = \{x_k\}_1^m,\ y = \{y_k\}_1^m$ be two tuples of vectors in $\mathbb{R}^n$ and $\mathrm{Maj}(y) := \{x :\ x \prec y\}$. Then*

$$\mathrm{Ext}(\mathrm{Maj}(y)) \supset \{(I_n \otimes P)y :\ P \textit{ is a permutation matrix}\}.$$

The following questions naturally arise in this context:

Questions. (1) Find some additional geometric conditions, which together with (2.7) imply $x \prec_{ds} y$.

(2) What are the extreme points of the set $\mathrm{Maj}(y)$?

(3) Under which conditions on the sequence $y = \{y_k\}$ do the sets $\{x :\ x \prec y\}$ and $\{x :\ x \prec_{ds} y\}$ coincide?

Let $CVS(\mathbb{R}^n)$ be the closed cone (in the pointwise convergence topology in $CV(\mathbb{R}^n)$), generated by the set of convex functions

$$\{f(\langle x, y\rangle) : f \in \mathrm{CV}(\mathbb{R}),\ y \in \mathbb{R}^n\}. \tag{2.10}$$

The class $CVS(\mathbb{R}^n)$ arises naturally in the following proposition, which is a partial generalization of Theorem 2.4.

Proposition 2.16. *Let $x := \{x_j\}_1^l$ be a k-tuple and $y := \{y_k,\}_1^m$ an m-tuple of vectors from $\mathbb{R}^n$ and $l \le m$. The following conditions are equivalent:*

(1) $x \prec y$;

(2) *for all vectors $h \in \mathbb{R}^n$,*

$$(\langle x_1, h\rangle, \ldots, \langle x_l, h\rangle) \prec (\langle y_1, h\rangle, \ldots, \langle y_m, h\rangle);$$

(3) *the inequality*

$$\sum_{i=1}^{l} f(x_i) \le \sum_{i=1}^{m} f(y_i) \tag{2.11}$$

holds for any nonnegative $f \in CVS(\mathbb{R}^n)$.

Proof. (1)$\Longleftrightarrow$(2). By the separating hyperplane theorem, a vector $x \in \mathbb{R}^n$ lies in a convex set $Y \subset \mathbb{R}^n$ iff its projection onto any line through the origin belongs to the projection of Y onto the same line. Thus

$$x_{i_1} + \ldots + x_{i_k} \in \mathrm{conv}(\{y_{j_1} + \ldots + y_{j_k}\})$$

iff

$$\langle x_{i_1}, h\rangle + \ldots + \langle x_{i_k}, h\rangle \in \mathrm{conv}(\{\langle y_{j_1}, h\rangle + \ldots + \langle y_{j_k}, h\rangle\})$$

for any $h \in \mathbb{R}^n$. By Remark 2.5, this is equivalent to (2).

(2)$\Longleftrightarrow$ (3). By a theorem of Fisher and Holbrook [13], $\{\langle x_k, h\rangle\}_1^l \prec \{\langle y_k, h\rangle\}_1^m$ if and only if

$$\sum_{k=1}^{l} f(\langle x_k, h\rangle) \le \sum_{k=1}^{m} f(\langle y_k, h\rangle)$$

for any nonnegative function $f \in \mathrm{CV}(\mathbb{R})$ (see Remark 2.18). Thus, (2.11) implies (2). The implication (2) $\Rightarrow$ (3) follows from Proposition 2.19 below. $\square$

Let us also mention the following beautiful theorem due to Sherman [37], which is another partial generalization of the Hardy-Littlewood-Polya theorem.

Theorem 2.17 ([37]). *Let $x := \{x_i\}_1^l$ be an l-tuple and $y := \{y_j\}_1^m$ an m-tuple of vectors in $\mathbb{R}^n$. Let also $\{\mu_i\}_1^l$ and $\{\nu_j\}_1^m$ be two collections of positive weights. Then the following are equivalent:*

(1) *there exists a matrix $S = (s_{ij}) \in \mathbb{R}^{l\times m}$ with nonnegative entries such that $\{x_i\}_{i=1}^l = (I \otimes S)\{y_j\}_{j=1}^m$ and*

$$\sum_{j=1}^m s_{ij} = 1, \quad i \in \{1,\dots,l\}, \quad \text{and} \quad \sum_{i=1}^l s_{ij}\mu_i = \nu_j, \quad j \in \{1,\dots,m\}; \tag{2.12}$$

(2) *the weights $\{\mu_i\}_1^l$ and $\{\nu_j\}_1^m$ satisfy $\sum_{i=1}^l \mu_i = \sum_{j=1}^m \nu_j$ and the inequality*

$$\sum_{i=1}^l \mu_i f(x_i) \le \sum_{j=1}^m \nu_j f(y_j) \tag{2.13}$$

holds for any $f \in \mathrm{CV}(\mathbb{R}^n)$.

Remark 2.18. This theorem is very nontrivial and the proof differs in an essential way from the scalar case.

There is an analog of this theorem for general measures on locally-convex spaces due to Cartier-Fell-Meyer. It plays an important role in convex analysis (see [32], [6]).

Fisher and Holbrook [13] have obtained a generalization of this theorem to the case when $\sum_{i=1}^l \mu_i \le \sum_{j=1}^m \nu_j$. Namely, they show that there exists a stochastic (by rows) matrix S, such that $\sum_{i=1}^l s_{ij}\mu_i \le \nu_j$ for all $j \in \{1,\dots,m\}$ if and only if inequality (2.13) holds for any nonnegative convex function f.

Combining Proposition 2.16, Theorem 2.17 (with $l = m$, $\mu_i = \nu_j = 1$) and Example 2.11 we get that $CVS(\mathbb{R}^n) \ne \mathrm{CV}(\mathbb{R}^n)$.[1]

To finish the section, we mention the following elegant proposition due to F. Petrov.

Proposition 2.19. *Let $\{x_i\}_1^l$ and $\{y_j\}_1^m$ be two tuples of vectors in $\mathbb{R}^n$ and $l < m$. The following three conditions are equivalent:*

(1) $\{x_i\}_1^l \prec \{y_j\}_1^m$;

(2) *for any $k = \{1,\dots,l\}$ and any k-tuple of indices $1 < i_1 < \ldots < i_k \le l$ there exist numbers $\beta_j \in [0,1]$ with $\sum_{j=1}^m \beta_j = k$ such that*

$$x_{i_1} + \ldots + x_{i_k} = \sum_{j=1}^m \beta_j y_j; \tag{2.14}$$

(3) *there exist vectors $x_{l+1},\dots,x_m \in \mathbb{R}^n$ such that $\{x_i\}_1^m \prec \{y_j\}_1^m$.*

Proof. (i) Let us show that (1) and (2) are equivalent. The implication (1) $\Rightarrow$ (2) easily follows from Definition 2.6.

Conversely, let (2) be fulfilled. By Proposition 2.16 (2), it suffices to treat the one-dimensional case $n = 1$. Assume without loss of generality that $y_1 \ge y_2 \ge \ldots \ge y_m$. Take any k-tuple $(i_j)_{j=1}^k$ with $1 \le i_1 < \ldots < i_k \le l$ and assume that (2.14) holds.

[1]Using our Example 2.11, F. V. Petrov constructed a simple explicit example of a function $f \in \mathrm{CV}(\mathbb{R}^n) \setminus CVS(\mathbb{R}^n)$.

The following inequality easily follows from the monotonicity of y_j's:

$$\sum_{j=1}^{k}(1-\beta_j)y_j \geq y_k\sum_{j=1}^{k}(1-\beta_j) = y_k\sum_{j=k+1}^{m}\beta_j \geq \sum_{j=k+1}^{m}\beta_j y_j. \tag{2.15}$$

Thus

$$\sum_{j=1}^{k} y_j \geq \sum_{j=1}^{m}\beta_j y_j = x_{i_1}+\ldots x_{i_k}$$

and similarly $\sum_{j=m-k+1}^{m} y_j \leq \sum_{j=1}^{m}\beta_j y_j = x_{i_1}+\ldots+x_{i_k}$. The implication (2) $\Rightarrow$ (1) follows by Remark 2.5.

(ii) The implication (3) $\Rightarrow$ (1) is obvious. Let us prove the implication (1) $\Rightarrow$ (3). Take

$$x_{l+1} := \frac{\sum_{j=1}^{m} y_j - \sum_{i=1}^{l} x_i}{m-l}.$$

Let us show that $\{x_i\}_1^{l+1} \prec \{y_j\}_1^m$.

Let $k \in \{1,\ldots,l\}$. It suffices to consider the tuple $(i_1,\ldots,i_k) = (1,\ldots,k)$. By the equivalence of (1) and (2) established on step (i), there exist $\beta_j \in [0,1]$, $j \in \{1,\ldots,m\}$, such that

$$x_1+\ldots+x_k = \sum_{j=1}^{m}\beta_j y_j, \qquad \sum_{j=1}^{m}\beta_j = k. \tag{2.16}$$

Similarly, there exist $\gamma_1,\ldots,\gamma_m$ with $\gamma_i \in [0,1]$ and satisfying

$$x_{k+1}+\ldots+x_l = \sum_{j=1}^{m}\gamma_j y_j, \qquad \sum_{j=1}^{m}\gamma_j = l-k. \tag{2.17}$$

Combining (2.16) and (2.17) we get

$$\begin{aligned} x_1+\ldots+x_k+x_{l+1} &= \sum_{j=1}^{m}\beta_j y_j + \frac{\sum_{j=1}^{m} y_j - \sum_{i=1}^{l} x_i}{m-l} \\ &= \sum_{j=1}^{m}(\beta_j + 1/(m-l) - \beta_j/(m-l) - \gamma_j/(m-l))y_j =: \sum_{j=1}^{m}\alpha_j y_j. \end{aligned} \tag{2.18}$$

Clearly, $\alpha_j \in [0,1]$ for all $j \in \{1,\ldots,m\}$ and

$$\sum_{j=1}^{m}\alpha_j = k + m/(m-l) - k/(m-l) - (l-k)/(m-l) = k+1.$$

Therefore the $(l+1)$-tuple $\{x_i\}_1^{l+1}$ satisfies (2). One completes the proof by induction. □

3. Inverse spectral problem and interlacing theorem for normal matrices

In this section we use the partial orders $\prec$ and $\prec_{ds}$ for vectors with complex entries. In this case we identify $\mathbb{C}$ with $\mathbb{R}^2$ and hence Definitions 2.6 and 2.8 do not change.

Let $A_{i_1,\ldots,i_k}^{j_1,\ldots,j_k}$ denote the submatrix of A with the rows $i_1,\ldots,i_k$ and columns $j_1,\ldots,j_k$. For the sake of brevity denote $A_{n-1} := A_{1,\ldots,n-1}^{1,\ldots,n-1}$.

3.1. **A preliminary solution to the inverse spectral problem by two spectra.**

Proposition 3.1. *Let $\{\lambda_k\}_1^n$ and $\{\mu_j\}_1^{n-1}$ be two sequences of complex numbers and define*

$$\Delta(\lambda) := \frac{\prod_{j=1}^{n-1}(\lambda-\mu_j)}{\prod_{k=1}^{n}(\lambda-\lambda_k)}. \tag{3.1}$$

Then the following two conditions are equivalent:

(1) *The rational function Δ has only simple poles and the residues in all poles are nonnegative, that is,*

$$\mathrm{Res}_{\lambda_k}(\Delta(\lambda)) \geq 0, \quad k \in \{1,\dots,n\}. \tag{3.2}$$

(2) *There exists a normal matrix A with the spectrum $\sigma(A) = \{\lambda_1,\dots,\lambda_n\}$ such that $\sigma(A_{n-1}) = \{\mu_1,\dots,\mu_{n-1}\}$.*

Proof. Necessity. Let A be a normal matrix with the spectrum $\sigma(A) = \{\lambda_1,\dots,\lambda_n\}$ and such that $\sigma(A_{n-1}) = \{\mu_1,\dots,\mu_{n-1}\}$. Let $e = (0,\dots,0,1)$. Consider the function

$$\Delta(\lambda) := ((\lambda-A)^{-1}e, e) = \sum_{k=1}^{n} \frac{|x_k|^2}{\lambda-\lambda_k} = \frac{\det(\lambda - A_{n-1})}{\det(\lambda - A)}, \tag{3.3}$$

where x_k are the coordinates of e in the orthonormal basis of eigenvectors of A. Clearly, the poles of $\Delta(\lambda)$ are in the spectrum of A and the residues in these poles are equal to $|x_k|^2$ and hence are nonnegative.

Sufficiency. Let (3.2) be fulfilled. Then

$$\Delta(\lambda) = \sum_{k=1}^{n} \frac{|x_k|^2}{\lambda-\lambda_k} \tag{3.4}$$

with some complex numbers $x_k,\ k = 1,\dots,n$, and

$$\sum_{k=1}^{n} |x_k|^2 = \lim_{\lambda\to\infty} \lambda\Delta(\lambda) = 1.$$

Therefore, considering the diagonal matrix $A = \mathrm{diag}\{\lambda_1,\dots,\lambda_n\}$ and writing it in an orthonormal basis with the last basis vector $e_n = (x_1,\dots,x_n)$, we get the required normal matrix. □

Remark 3.2. Note that the poles of the function $\Delta(\lambda)$ (see (3.3)) are simple (i.e. of multiplicity one), since the matrix A is (unitarily-) diagonalizable. Therefore it easily follows from (3.3), as well as from general dimension arguments, that if A has a k-multiple eigenvalue λ_0, then λ_0 is an eigenvalue of A_{m-1} of multiplicity at least $k-1$. Note that A_{n-1} is normal only in some special cases (see Lemma 4.14 below). It may even happen that A_{n-1} is non-diagonalizable.

Remark 3.3. (a) By the definition of a residue, (3.2) can be rewritten as

$$\lim_{\lambda\to\lambda_k} \frac{\prod_{j=1}^{n-1}(\lambda-\mu_j)}{\prod_{1\leq j\leq n, j\neq k}(\lambda-\lambda_j)} \geq 0. \tag{3.5}$$

In the case of real numbers $\lambda_1 \le \ldots \le \lambda_n$ and $\mu_1 \le \ldots \le \mu_{n-1}$ these conditions mean that the numbers $\{\mu_j\}$ and $\{\lambda_k\}$ are interlacing, that is,

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \ldots \le \mu_{n-1} \le \lambda_n. \tag{3.6}$$

This is the Cauchy-Poincare interlacing theorem [26].

(b) The criterion (3.5) ((3.2)) is trivial and provides no information on the location of the complex numbers $\{\lambda_k\}_1^n$ and $\{\mu_j\}_1^{n-1}$. It is even unclear how far μ_j's can lie from λ_k's. Therefore it would be desirable to have a more "geometric" answer which would be in some sense an analog of the Cauchy-Poincare theorem.

Corollary 3.4. *Let two tuples of complex numbers $\{\mu_j\}_1^{n-1}$ and $\{\lambda_j\}_1^n$ satisfy*

$$\Delta(\lambda) := \frac{\prod_{j=1}^{n-1}(\lambda-\mu_j)}{\prod_{j=1}^{n}(\lambda-\lambda_j)} = \sum_{j=1}^{n} \frac{|x_j|^2}{\lambda-\lambda_j} \tag{3.7}$$

with some complex numbers $\{x_j\}_1^n$. Then for any unitary matrix $U = (u_{ij})_{i,j=1}^n$ with the last row $(u_{n1},\ldots,u_{nn}) = (x_1,\ldots,x_n)$, the normal matrix $A := U\,\mathrm{diag}\{\lambda_j\}_1^n U^*$ satisfies the hypothesis of Proposition 3.1 (2).

3.2. **Quasi-Jacobi normal matrices and an analog of the Hochstadt theorem.** It is known and easy to see that any Hermitian matrix (bounded operator) is unitarily equivalent to a Hermitian tridiagonal (Jacobi) matrix.

We shall find an analog of such a form for a normal matrix and use it to obtain an analog of the Hochstadt theorem [17] on the unique recovery of a Jacobi matrix from two spectra.

Proposition 3.5. *Every normal matrix A is unitarily equivalent to a direct sum of normal matrices A_i, $i \in \{1,\ldots,k\}$, satisfying $(A_i)_{r,s} = 0$ for $s \ge r+2$ and $(A_i)_{j,j+1} \ne 0$. The matrix A has simple spectrum iff $k = 1$.*

Proof. The proof is very simple and standard. Clearly, it suffices to treat only the simple spectrum case. Taking any cyclic vector x, we get that $\{A^j x\}_0^{m-1}$ forms a basis in $\mathbb{C}^m$. We get the required orthonormal basis after the Gramm-Schmidt procedure. □

Definition 3.6. A matrix A is called quasi-Jacobi if $(A)_{r,s} = 0$ for all $s \ge r+2$.

Thus, the quasi-Jacobi form is in a sense a normal form for a normal matrix. We can now complement Proposition 3.1 with a uniqueness statement.

Theorem 3.7. *For any two tuples of complex numbers $\{\lambda_j\}_1^n$ and $\{\mu_j\}_1^{n-1}$ satisfying* (3.2) *there exists a unique normal quasi-Jacobi matrix A such that $\sigma(A) = \{\lambda_j\}_1^n$ and $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$.*

Proof. Writing the function $\Delta(\lambda)$ from (3.7) in the form

$$\Delta(\lambda) = \int \frac{d\mu(z)}{\lambda - z}$$

with $d\mu = \sum_{k=1}^n x_k^2 \delta_{\lambda_k}$, we can introduce the orthogonal polynomials with respect to this measure just as in the case of Jacobi matrices (see, e.g. [1] and Gesztesy and Simon [14]) and then, following the same lines as in [14], we get what we need. □

Remark 3.8. The function $\Delta(\lambda)$ is an analog of the Weyl M-function in this case (see [14], [20]).

3.3. **The set of all possible diagonals in the unitary orbit of a normal matrix.** In this subsection we start with an arbitrary normal matrix A and give a (trivial) description of the set of diagonals of its unitary orbit.

Define

$$U(A) := \{UAU^* : \ U \in M_n(\mathbb{C}), \ U \in \mathrm{SU}(n)\}$$

where $SU(n)$ is the special unitary group. Put

$$\mathrm{diag}(U(A)) := \{x = (x_i)_{i=1}^n \in \mathbb{C}^n : \ x = \mathrm{diag}\, B, \ B = (b_{ij}) \in U(A)\}$$

where $\mathrm{diag}\, B := (b_{11}, \dots, b_{nn})$.

Proposition 3.9. *Let $A \in M_n(\mathbb{C})$ be a normal matrix with the spectrum $\sigma(A) = \{\lambda_j\}_{j=1}^n$. Then*

(a) $\{a_{jj}\}_{j=1}^n \prec_{uds} \{\lambda_j\}_{j=1}^n$ *and*

$$\mathrm{diag}(U(A)) = \{x \in \mathbb{C}^n : \ x \prec_{uds} \{\lambda_j\}_1^n\} = \Omega_n^u((\lambda_j)_1^n);$$

(b) *if A is Hermitian, then the set* $\mathrm{diag}(U(A))$ *is convex;*

(c) *if A is not Hermitian, then the set* $\mathrm{diag}(U(A))$ *is not necessarily convex.*

Proof. The validity of (a) is obvious.

(b) is due to Horn.

(c) Clearly, since permutation matrices are unitary, $\mathrm{diag}(U(A))$ contains the orbit of $(\lambda_j)_1^n$ under the permutation group. Thus, if $\mathrm{diag}(U(A))$ were convex, we would have $\mathrm{diag}(U(A)) = \Omega_n^u((\lambda_j)_1^n) = \Omega_n((\lambda_j)_1^n)$. Take

$$S = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

It is known (see [26]) that S is not unitarily stochastic. Set $(\lambda_1, \lambda_2, \lambda_3) = (1, i, 0)$. We have

$$S \cdot \mathrm{col}(\lambda_1, \lambda_2, \lambda_3) = \mathrm{col}(\alpha_1, \alpha_2, \alpha_3) = \mathrm{col}(1 + i, 1, i).$$

If there existed a unitarily-stochastic matrix O, such that

$$O \cdot \mathrm{col}(\lambda_1, \lambda_2, \lambda_3) = \mathrm{col}(\alpha_1, \alpha_2, \alpha_3),$$

it would immediately yield $O = S$. □

Remark 3.10. The only nontrivial part of Proposition 3.9 is (b). The statement of (b) is due to Horn [18], [26]. It is a special case (for the semi-simple Lie group $\mathrm{SU}(n)$) of the famous Kostant convexity theorem (see [19]).

3.4. **Analog of the Cauchy-Poincare interlacing theorem and a solution to the inverse spectral problem for normal matrices.** Let $X \subset \mathbb{C}^m$. We denote by X' the set of limit points of X.

Define for any vector $(\lambda_j)_1^m \in \mathbb{C}^m$ the vector

$$C_k((\lambda_j)_1^m) := (\lambda_{i_1} \cdots \lambda_{i_k})_{1 \le i_1 < \dots < i_k \le m} \in \mathbb{C}^{\binom{m}{k}}. \tag{3.8}$$

Theorem 3.11. *Let $(\lambda_1, \dots, \lambda_n)$ and $(\mu_1, \dots, \mu_{n-1})$ be two complex vectors. Let $X \subset \mathbb{C}$ be any set such that $X' \supset \{\lambda_1, \dots, \lambda_n\}$.*

Then for the existence of a normal matrix A such that $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ and $\sigma(A_{n-1}) = \{\mu_1, \dots, \mu_{n-1}\}$ it is necessary that the condition

$$C_k((\mu_i - \alpha)_1^{n-1}) \prec_{uds} C_k((\lambda_j - \alpha)_1^n) \tag{3.9}$$

be fulfilled for all complex numbers $\alpha \in \mathbb{C}$ and all $k \in \{1, \dots, n-1\}$ and sufficient that it be fulfilled for $k = n-1$ and all $\alpha \in X$.

Proof. Sufficiency. Suppose first that $\mu_i \neq \lambda_j$ for all $i \in \{1, \dots, n-1\}$ and all $k \in \{1, \dots, n\}$ and that the eigenvalues are simple, i.e. $\lambda_i \neq \lambda_j$ for all $i \neq j$. Take $\alpha = \lambda_j$. Then (3.9) for $k = n-1$ reads

$$\prod_{i=1}^{n-1}(\mu_i - \lambda_j) \in \operatorname{conv}\left(\prod_{1\le i\le n, i\neq l}(\lambda_i - \lambda_j) : \ l \in \{1, \dots, n\}\right) = \operatorname{conv}\left(0; \prod_{1\le i\le n, i\neq j}(\lambda_i - \lambda_j)\right) \tag{3.10}$$

and hence (3.5) is valid. The general case follows by continuity. Proposition 3.1 yields the required assertion.

Necessity. Let A be a normal matrix with the spectrum $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$ and $\sigma(A_{n-1}) = \{\mu_1, \dots, \mu_{n-1}\}$. By the Schur Theorem [18], there exists a unitary matrix $V_1 \in M_{n-1}(\mathbb{C})$ such that the matrix $V_1^* A_{n-1} V_1$ is upper triangular. Therefore, considering the matrix $U_1 := V_1 \oplus 1 \in M_n(\mathbb{C})$ we get the normal matrix $B := U_1^* A U_1$ with $\sigma(B) = \sigma(A)$, but the numbers $\{\mu_i\}_1^{n-1}$ are on the diagonal of B. Proposition 3.9 implies $\{\mu_i\}_1^{n-1} \prec_{uds} \{\lambda_j\}_1^n$.

Take an arbitrary $\alpha \in \mathbb{C}$. Let us consider the exterior power $C_k(B - \alpha I) := \wedge^k(B - \alpha I)$ acting on the space $\wedge^k H := \underbrace{H \wedge \cdots \wedge H}_{k \text{ times}}$. This matrix is normal with the spectrum

$$\sigma(C_k(B - \alpha I)) = \{(\lambda_{j_1} - \alpha) \cdot (\lambda_{j_2} - \alpha) \cdots (\lambda_{j_k} - \alpha)\}_{1 \le j_1 < j_2 < \dots < j_k \le n}.$$

The diagonal elements of $C_k(B - \alpha I)$ are the $k \times k$ principal minors of $B - \alpha I$. Since $B_{n-1} - \alpha I_{n-1}$ is upper triangular, the $k \times k$ principal minors of $B_{n-1} - \alpha I$ equal

$$(\mu_{i_1} - \alpha) \cdot (\mu_{i_2} - \alpha) \cdots (\mu_{i_k} - \alpha), \quad 1 \le i_1 < \dots < i_k \le n-1.$$

Proposition 3.9 (a) implies what was required. □

Remark 3.12. (i) If $\lambda_i \neq \lambda_j$ for all $i \neq j$, then it suffices to take $X = \{\lambda_1, \dots, \lambda_n\}$.

(ii) Theorem 3.11 and Proposition 2.10 together yield that $C_k((\mu_i - \alpha)_1^{n-1}) \prec C_k((\lambda_j - \alpha)_1^n)$ for any $k \in \{1, \dots, n-1\}$. Therefore these vectors satisfy the simple geometric conditions (2.7), clarifying the relative location of the numbers $\{\mu_i\}_1^{n-1}$ and $\{\lambda_j\}_1^n$.

(iii) It is easy to see from the proof that the unitarily-stochastic matrices in (3.9) are independent of α.

Example 3.13. In the case $n = 3$ the orders $\prec$ and $\prec_{ds}$ are equivalent. Therefore in this case conditions (3.9) take a simple form:

$$\mu_1, \mu_2 \in \operatorname{conv}\{\lambda_j\}_1^3, \ \mu_1 + \mu_2 \in \operatorname{conv}\{\lambda_1 + \lambda_2, \lambda_2 + \lambda_3, \lambda_1 + \lambda_3\},$$
$$\mu_1\mu_2 - \alpha(\mu_1 + \mu_2) \in \operatorname{conv}\{\lambda_k\lambda_p - \alpha(\lambda_k + \lambda_p)\}_{1\le k<p\le 3}, \ \alpha \in \mathbb{C}.$$

An immediate consequence of Theorem 3.11 is:

Corollary 3.14. *Let A be a normal matrix with the spectrum $\sigma(A) := \{\lambda_1, \dots, \lambda_n\}$ and P an $m(\leq n)$-dimensional orthoprojection in $\mathbb{C}^n$. Let also $B := PA\lceil P\mathbb{C}^n$ and $\sigma(B) := \{\mu_1, \dots, \mu_m\}$. Then*

$$C_k((\mu_i - \alpha)_1^m) \prec_{uds} C_k((\lambda_j - \alpha)_1^n) \tag{3.11}$$

for all $k \in \{1 \dots, m\}$ and any $\alpha \in \mathbb{C}$.

Corollary 3.15. *Let $\{A_i\}_1^p$ be a p-tuple of normal matrices, $A_i \in M_{n_i}(\mathbb{C})$, $i \in \{1, \dots, p\}$, and $\sigma(A_i) = \{\lambda_k^i\}_{k=1}^{n_i}$ their spectra. Let $S_i \in \mathbb{C}^{m \times n_i}$ for $i \in \{1, \dots, p\}$ and satisfy*

$$\sum_{i=1}^{p} S_i^* S_i = I_m, \tag{3.12}$$

where I_m is the $m \times m$ identity matrix. Consider $B = \sum_{i=1}^p S_i^* A_i S_i$ and let $\sigma(B) = (\mu_1, \dots, \mu_m)$. Set $n = n_1 + \dots + n_p$ and

$$(\lambda_1, \dots, \lambda_n) = (\lambda_1^1, \dots, \lambda_{n_1}^1, \lambda_1^2, \dots, \lambda_{n_2}^2, \dots, \lambda_{n_p}^p). \tag{3.13}$$

Then the tuples $(\lambda_1, \dots, \lambda_n)$ and $(\mu_1, \dots, \mu_m)$ satisfy (3.11) for all $k \in \{1, \dots, m\}$ and any $\alpha \in \mathbb{C}$.

Proof. Consider the normal matrix $A := \bigoplus_{i=1}^p A_i$. Condition (3.12) means that the operator

$$V := \begin{pmatrix} S_1 \\ \vdots \\ S_p \end{pmatrix} : \mathbb{R}^m \to \mathbb{R}^n$$

is an isometry and hence the operator $B = V^* A V$ is unitarily equivalent to $PA\lceil P\mathbb{C}^n$ where P is the projector onto the image of V. Corollary 3.14 completes the proof. □

Remark 3.16. There is another way to implement the trick of going from Corollary 3.14 to Corollary 3.15. See [21], [22].

4. Location of the roots of a polynomial and its derivative

4.1. **Generalization of the Gauss-Lucas theorem.** Recall the Gauss-Lucas theorem.

Theorem 4.1. *The roots of the derivative p' of a polynomial $p \in C[z]$ lie in the convex hull of the roots of p.*

Numerous papers are devoted to different generalizations and improvements of this theorem (see e.g. [2], [4], [10], [30], [36], [39]).

In what follows we denote by $p \in C[z]$ a polynomial of degree n with complex coefficients. Let also $\{\lambda_j\}_1^n$ be the roots of p and $\{\mu_j\}_1^{n-1}$ the roots of its derivative p'. We set additionally $\mu_n := (\sum_1^n \lambda_j)/n$. Then the Gauss-Lucas theorem reads: there exists a stochastic (by rows) matrix S' such that $\mu = S'\lambda$, where $\mu := \{\mu_j\}_1^n$ and $\lambda := \{\lambda_j\}_1^n$. We start with the following proposition, which is already mentioned in the introduction.

Proposition 4.2. *Let $p(z)$ be a polynomial of degree n with zeros $\{\lambda_j\}_1^n$, and $\{\mu_j\}_1^{n-1}$ the zeros of its derivative. Then there exists a (non-unique) normal matrix $A \in M_n(\mathbb{C})$ such that $\sigma(A) = \{\lambda_j\}_1^n$ and $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$.*

Proof. The claim immediately follows from the identity

$$p'(z)/p(z) = \sum_{j=1}^{n} 1/(z - \lambda_j) \tag{4.1}$$

and Corollary 3.4. □

Therefore we can apply Theorem 3.11 and get the following generalization of the Gauss-Lucas theorem.

Proposition 4.3. *Let p ($\in \mathbb{C}[z]$) be a degree n complex polynomial with roots $\{\lambda_j\}_1^n$, and let $\{\mu_k\}_1^{n-1}$ be the roots of its derivative p'. Then the vectors $(\lambda_j)_1^n$ and let $(\mu_k)_1^{n-1}$ satisfy (3.9) for $k \in \{1, \dots, n-1\}$.*

Corollary 4.4. *Let $\lambda := (\lambda_j)_1^n$ and $\mu := (\mu_i)_1^{n-1}$ be as above and $k \in \{1, ..., n-1\}$. Then $C_k((\mu_i)_1^{n-1}) \prec C_k((\lambda_j)_1^n)$ and in particular*

$$\mu_{i_1}\mu_{i_2}\cdots\mu_{i_k} \in \operatorname{conv}\{\lambda_{j_1}\lambda_{j_2}\cdots\lambda_{j_k} :\ 1 \le j_1 < j_2 < ... < j_k \le n\} \tag{4.2}$$

for any k-tuple $\{i_1, \dots, i_k\}$ with $1 \le i_1 < i_2 < ... < i_k \le n-1$.

The Viete formulae immediately yield

$$\frac{1}{C_{n-1}^k} \sum_{1 \le i_1 < \dots < i_k \le n-1} \prod_{j=1}^{k} (\mu_{i_j} - \alpha) = \frac{1}{C_n^k} \sum_{1 \le i_1 < \dots < i_k \le n} \prod_{j=1}^{k} (\lambda_{i_j} - \alpha). \tag{4.3}$$

It turns out that in this case we can get more information on the matrices $S \in \Omega_n^u$, realizing majorization, using the special form of the identity (4.1) (compare with (3.7)).

Definition 4.5. A matrix $S = (s_{ij}) \in \mathbb{R}^{m \times n}$ is put in the class $\Omega(m, n)$ if $s_{ij} \ge 0$ for $i \in \{1, ..., m\},\ j \in \{1, ..., n\}$ and

$$\sum_{j=1}^{n} s_{ij} = 1,\ i \in \{1, \dots, m\}, \quad \text{and} \quad \sum_{i=1}^{m} s_{ij} = \frac{m}{n},\ j \in \{1, ..., n\}. \tag{4.4}$$

It is clear that $\Omega(n, n) = \Omega_n$. The following multiplicative property of the classes $\Omega(m, n)$ is obvious.

Lemma 4.6. $\Omega(m, k) \cdot \Omega(k, n) \subset \Omega(m, n)$.

For every $S \in \Omega_n$ and every $m < n$ we denote by $S(m, n)$ the $(m \times n)$ submatrix of S, lying in the first m rows.

Theorem 4.7. *Let $k \in \{1, \dots, n-1\}$, $a := \binom{n-1}{k}$, $b := \binom{n}{k}$. Then there exists a matrix $S_k = (s_{ijk}) \in \Omega_b^u$ such that $S_k(a, b) \in \Omega(a, b)$ and $C_k((\mu_j - \alpha)_1^{n-1}) = S_k(a, b) C_k((\lambda_j - \alpha)_1^n)$ (see (3.8)).*

In particular there exists an $S_1 \in \Omega_n^u$ such that $s_{nj1} = 1/n,\ j \in \{1, \dots, n\}$ and $(\mu_j)_1^n = S_1(\lambda_j)_1^n$.

Proof. Set $D := \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Corollary 3.4 and (4.1) together yield that for any unitary matrix V with the last row consisting of $1/\sqrt{n}$, the normal matrix $A := VDV^*$ solves the inverse spectral problem for the pair of tuples $\{\lambda_j\}_1^n$ and $\{\mu_j\}_1^{n-1}$, that is, $\sigma(A) = \{\lambda_j\}_1^n$ and $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$. Further let U_1 be the same unitary matrix as in the proof of Theorem 3.11. Then $U := U_1 V (\in M_n(\mathbb{C}))$ is a unitary matrix with the last row consisting of $1/\sqrt{n}$. Moreover, it follows from the proof of Theorem 3.11 that $(\mu_j)_1^n = S_1(\lambda_j)_1^n$, where $S_1 := U \circ \bar{U} (\in \Omega_n^u)$ is a unitarily-stochastic matrix with the last row consisting of $1/n$.

Therefore, passing to the exterior powers as in the proof of Theorem 3.11, we get that the unitarily-stochastic matrix

$$S_k := C_k(U) \circ C_k(\bar{U}), \qquad k \in \{1, \dots, n\}$$

realizes the unitarily-stochastic majorization of the tuples of numbers $C_k((\mu_j)_1^{n-1})$ and $C_k((\lambda_j)_1^n)$. We shall prove that S_k has the additional property

$$\sum_{1 \le i_1 < \dots < i_{k-1} \le n-1} \left| \det U^{1,\dots,k}_{i_1,\dots,i_{k-1},n} \right|^2 = \frac{n-k}{n}, \tag{4.5}$$

By symmetry, the same identity is certainly valid for any other choice of k columns, that is, $S_k(a,b) \in \Omega(a,b)$.

Expanding each minor with respect to the last row we get

$$\det U^{1,\dots,k}_{i_1,\dots,i_{k-1},n} = \frac{1}{\sqrt{n}} \sum_{l=1}^{k} (-1)^{l+1} \det U^{1,\dots,\hat{l},\dots,k}_{i_1,\dots,i_{k-1}} \tag{4.6}$$

and thus

$$|\det U^{1,\dots,k}_{i_1,\dots,i_{k-1},n}|^2 = \frac{1}{n} \sum_{p,q=1}^{k} (-1)^{p+q} \det U^{1,\dots,\hat{p},k}_{i_1,\dots,i_{k-1}} \det \bar{U}^{1,\dots,\hat{q},\dots,k}_{i_1,\dots,i_{k-1}}. \tag{4.7}$$

Consider now the matrix $V := U^{1,\dots,k}_{1,\dots,n-1}$. Then, since U is unitary, we get

$$T := V^*V = \{t_{ij}\}_{i,j=1}^k \quad \text{with } t_{ij} = \begin{cases} \frac{n-1}{n}, & i = j, \\ -\frac{1}{n}, & i \ne j. \end{cases}$$

Therefore

$$C_{k-1}(V)^* \cdot C_{k-1}(V) = C_{k-1}(T).$$

By (4.7), one easily sees that the required sum in (4.5) is a linear combination of the elements of the matrix $C_{k-1}(V)^* \cdot C_{k-1}(V)$. Therefore this sum depends only on the matrix T, is independent of the choice of U (with $u_{nj} = 1/\sqrt{n},\ j \in \{1, \dots, n\}$), and is the same for any other choice of k columns.

Identity (4.3) implies that these sums must equal $(n-k)/n$. The proof is complete. □

Corollary 4.8. *Let $U (\in \mathbb{C}^{n \times n})$ be a unitary matrix with $u_{nj} = 1/\sqrt{n}$. Let $C_k(U)$ be its k-th exterior power, and $S_k := C_k(U) \circ C_k(\bar{U})$. Then for any $k \in \{1, \dots, n-1\}$ the matrix S_k is unitarily stochastic and satisfies*

$$\sum_{i=1}^{\binom{n-1}{k}} s_{ij} = \frac{\binom{n-1}{k}}{\binom{n}{k}} = \frac{n-k}{n}, \qquad j \in \left\{1, \dots, \binom{n}{k}\right\},$$

that is, $S_k(a,b) \in \Omega(a,b)$.

4.2. **The conjecture of de Bruijn-Springer and its generalization.** In 1947 de Bruijn and Springer [8] conjectured that the inequality

$$\frac{1}{n-1}\sum_{j=1}^{n-1} f(\mu_j) \le \frac{1}{n}\sum_{j=1}^{n} f(\lambda_j) \tag{4.8}$$

holds for any convex function $f : \mathbb{C} \to \mathbb{R}$.

In order to prove this conjecture and its generalization we need the following simple lemma, which is actually contained in Theorem 2.17.

Lemma 4.9. *Let $(x_j)_1^k,\ (y_j)_1^n$ be two sequences of vectors from $\mathbb{R}^m$. Suppose that there exists a matrix $S = (s_{ij}) \in \Omega(k,n)$ such that $(x_j)_1^k = (I_m \otimes S)(y_j)_1^n$. Then the inequality*

$$\frac{1}{k}\sum_{j=1}^{k} f(x_j) \le \frac{1}{n}\sum_{j=1}^{n} f(y_j) \tag{4.9}$$

holds for any function $f \in CV(\mathbb{R}^m)$.

Proof. The proof follows by an application of the Jensen inequality. □

The following theorem contains in particular a positive solution to the conjecture of de Bruijn and Springer [8].

Theorem 4.10. *The following inequality holds for any convex function $f : \mathbb{C} \to \mathbb{R}$ and any $k,\ 1 \le k \le n-1$:*

$$\begin{aligned} &\frac{1}{\binom{n-1}{k}} \sum_{1\le i_i<\ldots<i_k\le n-1} f\left(\prod_{j=1}^{k}(\mu_{i_j}-\alpha)\right) \\ &\qquad \le \frac{1}{\binom{n}{k}} \sum_{1\le i_1<\ldots<i_k\le n} f\left(\prod_{j=1}^{k}(\lambda_{i_j}-\alpha)\right). \end{aligned} \tag{4.10}$$

Proof. The inequality immediately follows by combining Theorem 4.7 with Lemma 4.9. □

Remark 4.11. By Theorem 2.17, Theorem 4.7 and Theorem 4.10 are equivalent.

4.3. **The Schoenberg conjecture.** We will need two lemmas. The first one is known [26], but we present it with a proof for the reader's convenience.

Lemma 4.12. *Any matrix $A = (a_{ij})_{i,j=1}^n \in M_n(\mathbb{C})$ with spectrum $\sigma(A) = \{\lambda_j\}_1^n$ satisfies the inequality*

$$\sum_{j=1}^{n} |\lambda_j|^2 \le \|A\|_2^2 = \sum_{i,j=1}^{n} |a_{ij}|^2 \tag{4.11}$$

and the equality holds if and only if A is normal.

Proof. It is clear that the equality holds for a normal matrix. Conversely, let A satisfy the equality. By the Schur theorem, A is unitarily equivalent to an upper-triangular matrix with λ_j's on the diagonal. Since $\|A\|_2$ is unitarily invariant, this upper-triangular matrix is diagonal, that is, A is normal. □

Lemma 4.13. *Let* $\varepsilon = e^{2\pi i/n}$ *and* $U = n^{-1/2}(\varepsilon^{k(j-1)})_{k,j=1}^n$. *Let also* $\sum_{j=1}^n \lambda_j = 0$. *Define*

$$r(z) := \sum_{j=1}^n \lambda_j z^{j-1}$$

and

$$A := U \operatorname{diag}(\lambda_j)_{j=1}^n U^* = \frac{1}{n}(r(\varepsilon^{k-j}))_{k,j=1}^n =: (a_{ij})_{i,j=1}^n.$$

Then A *is a normal matrix with* $\sigma(A) = \{\lambda_j\}_1^n$ *and* $\sigma(A_{n-1}) = \{\mu_j\}_1^{n-1}$ *and satisfies*

$$n\|A_{n-1}\|_2^2 = (n-2)\|A\|_2^2.$$

Proof. The first statement follows from Corollary 3.4, since the last row of U consists of $\frac{1}{\sqrt{n}}$. It remains to prove the last identity.

It is easy to see that $a_{jj} = 0$ and $a_{nj} = a_{n-k,j-k}$ for $k < j$ and $a_{nj} = a_{k-j,k}$ for $k > j$. Therefore the required identity takes the form

$$\sum_{j=1}^n |\lambda_j|^2 = n \sum_{j=1}^{n-1} |a_{nj}|^2.$$

But

$$n \sum_{j=1}^{n-1} |a_{nj}|^2 = \frac{(n-1)}{n} \sum_{j=1}^n |\lambda_j|^2 + \frac{1}{n} 2 \sum_{i<j} \Re\left(\lambda_i \bar{\lambda}_j c_{ij}\right)$$

with

$$c_{ij} = \sum_{k=1}^{n-1} \varepsilon^{(i-j)(n-k)} = -1$$

for all $1 \le i \ne j \le n$. We have

$$\left|\sum_{j=1}^n \lambda_j\right|^2 = 0 \Longleftrightarrow 2\sum_{i<j} \Re(\lambda_i \bar{\lambda}_j) = -\sum_{j=1}^n |\lambda_j|^2.$$

This completes the proof. □

The next lemma is due to Fan and Pall [12].

Lemma 4.14. *Let* A *be a normal matrix such that its submatrix* A_{n-1} *is also normal and* $A \ne A_{n-1} \oplus a_{nn}$. *Then all eigenvalues of* A *lie on the same line.*

The following inequality has been conjectured by Schoenberg [38] (see also [9]).

Theorem 4.15. *Let* $\sum_{j=1}^n \lambda_j = 0$. *Then*

$$n \sum_{j=1}^{n-1} |\mu_j|^2 \le (n-2) \sum_{j=1}^n |\lambda_j|^2$$

and the equality holds if and only if all numbers λ_j *lie on the same line.*

Proof. Combining Lemma 4.12 and Lemma 4.13, we get

$$n \sum_{j=1}^{n-1} |\mu_j|^2 \le n\|A_{n-1}\|_2^2 = (n-2)\|A\|_2^2 = (n-2) \sum_{j=1}^n |\lambda_j|^2. \tag{4.12}$$

By Lemma 4.12, equality holds in (4.12) if and only if A_{n-1} is normal. By Lemma 4.14, this is possible if and only if all λ_j's lie on the same line. □

4.4. **The Mason-Shapiro polynomials.** In [28] Gisli Mason and Boris Shapiro initiated study of a class of differential operators T_Q defined as follows: let Q be a degree k monic polynomial. Then T_Q is defined via

$$T_Q : \ f \to (Qf)^{(k)}.$$

They have shown that for each m there exists a unique polynomial eigenfunction p_m of T_Q of degree m. Moreover,

$$T_Q p_m = \lambda_{m,k} p_m \tag{4.13}$$

and $\lambda_{m,k}$ depends only on k, m, namely

$$\lambda_{m,k} = (m+1)(m+2)\cdots(m+k).$$

One of the results of [28] is the following interesting analog of the Gauss-Lucas theorem:

Theorem 4.16 ([28])**.** *The zeros of p_m are contained in the convex hull of the zeros of Q for each m.*

Mason and Shapiro have also made a number of beautiful conjectures on the asymptotic distribution of the zeros of p_m's, partly proved in [3].

We strengthen Theorem 4.16 in the following way:

Theorem 4.17. *Let $\{z_j\}_{j=1}^k$ be the zeros of Q and $\{w_i\}_1^m$ the zeros of p_m. Then*

(i) *there exists a matrix $S \in \Omega(m,k)$ such that*

$$\{w_i\}_1^m = S\{z_j\}_1^k; \tag{4.14}$$

(ii) *for any convex function $f : \mathbb{C} \to \mathbb{R}$:*

$$\frac{\sum_{i=1}^m f(w_i)}{m} \le \frac{\sum_{j=1}^k f(z_j)}{k}. \tag{4.15}$$

Proof. Denote by $\{w_i^{(r)}\}_{i=1}^{m+k-r}$ the roots of $(QP_m)^{(r)}$, $r \in \{1, ..., k\}$, and note that due to (4.13) $w_i = w_i^{(k)}$, $i \in \{1, ..., m\}$. By Theorem 4.7, there exists a matrix $S_r \in \Omega(m+k-r, m+k-r+1)$ such that

$$\mathrm{col}(w_1^{(r)}, ..., w_{m+k-r}^{(r)}) = S_r \cdot \mathrm{col}(w_1^{(r-1)}, \ldots, w_{m+k-r+1}^{(r-1)}), \quad r \in \{1, ..., k\}.$$

It follows that

$$\mathrm{col}(w_1, ..., w_m) = \widetilde{S} \cdot \mathrm{col}(w_1, \ldots, w_m, z_1, \ldots, z_k) \tag{4.16}$$

where $\widetilde{S} := S_k \cdot S_{k-1} \cdot ... \cdot S_1$. By Lemma 4.6, $\widetilde{S} \in \Omega(m, m+k)$.

It is possible to construct the matrix $S \in \Omega(m,k)$ satisfying (4.14) by resolving identity (4.16) for the w_i's. We present here an indirect proof, appealing to Theorem 2.17. Namely, (4.16) and Lemma 4.9 together yield

$$\frac{1}{m}\sum_{j=1}^m f(w_i) \le \frac{1}{k+m}\left(\sum_{j=1}^k f(z_j) + \sum_{i=1}^m f(w_i)\right)$$

which is equivalent to (4.15).

Inequality (4.15) allows us to apply Theorem 2.17 with the weights $\mu_i = \frac{1}{m},\ i \in \{1, \dots, m\}$, and $\nu_j = \frac{1}{k},\ j \in \{1, \dots, k\}$. Thus, there exists a matrix $S = (s_{ij}) \in \mathbb{R}^{m \times k}$ with nonnegative entries, satisfying relations (2.12), that is, $S \in \Omega(m, k)$ and (i) is proved. $\square$

Remark 4.18. In the case of real numbers w_j and z_j inequality (4.15) was mentioned without proof by Harold Shapiro [35].

Final remarks and open problems. (1) Let $A \in M_n(\mathbb{C})$ be a normal matrix, $e \in \mathbb{C}^n$ a vector and P the orthoprojection onto the orthogonal complement of e. It is not difficult to construct examples of a nondiagonalizable $A_e := PA\lceil P\mathbb{C}^n$. It would be interesting to investigate the Jordan normal form and unitary invariants of the operator A_e.

(2) The same questions as in (1), but for projections P of arbitrary dimension seem to be even more complicated. We do not even know whether conditions (3.11) in Corollary 3.14 are sufficient. And the Jordan normal form of a submatrix of high codimension may certainly be very complicated (you can get any matrix with norm less or equal to one as a principal submatrix of a unitary matrix).

(3) The location of the roots of higher (even second) order derivatives of a polynomial may be very complicated (see [34]). The following natural question arises in connection with (2):

Question. Let p be a polynomial. Does there exist a normal matrix with characteristic polynomial p, such that its principal submatrix of codimension k has $p^{(k)}$ as its characteristic polynomial.

I conjecture that the answer is negative, but perhaps it still holds under some conditions? (For example, it clearly holds for real roots).

(4) The most famous open problems connected with the Gauss-Lucas theorem are the conjectures of Sendov and Smale (see [4] for very interesting and nontrivial recent theorems in this direction and [36] for a survey on this topic). Can one in any way apply the linear algebraic approach to these conjectures?

Finally, we mention a very interesting recent paper [5] by J. Borcea and B. Shapiro, where the authors also study relations between majorization and the location of the roots of a polynomial and of its derivative, and formulate several interesting conjectures.

Acknowledgements

I would like to express my deep gratitude to Boris Shapiro, M. M. Malamud and F. V. Petrov for numerous comments, remarks and discussions which essentially improved the paper. I am also very grateful to J. Borcea and F. V. Petrov who have informed me about the conjecture of de Bruijn and Springer [8] and the papers [37] and [13].

I am very grateful to Sergey Fomin for all his editorial help and to Professor Oscar E. Lanford III for helping to correct my English.

Note added in proof

When the paper was under review, R. Pereira [31] published a solution to the de Bruijn-Springer and Schoenberg conjectures, based on similar ideas. The author thanks the referee for informing him about it.

References

[1] N. I. Akhiezer, *The classical moment problem,* Oliver and Boyd, Edinburgh, 1965. MR32:1518

[2] A. Aziz and N. A. Rather, *On an inequality of S. Bernstein and the Gauss-Lucas theorem.* Rassias, Themistocles M. (ed.) et al., Analytic and geometric inequalities and applications. Dordrecht: Kluwer Academic Publishers. Math. Appl., Dordr. 478, 29-35 (1999). MR2001e:30006

[3] T. Bergkvist and H. Rullgard, *On polynomial eigenfunctions for a class of differential operators,* Math. Res. Let. v. 9, pp. 153-171, 2002. MR2003d:34192

[4] J. Borcea, *Maximal and inextensible polynomials and the geometry of the spectra of normal operators.* preprint, math.CV/0309233.

[5] J. Borcea and B. Shapiro, *Hyperbolic polynomials and spectral order,* C. R. Math. Acad. Sci. Paris 337 (2003), no. 11, 693–698.

[6] O. Bratteli and D. W. Robinson, Operator Algebras and Quantum Statistical Mechanics I. Springer, 1987. MR88d:46105

[7] T. Craven and G. Csordas, *The Gauss-Lucas theorem and Jensen polynomials.* Trans. Amer. Math. Soc. 278 (1983), no. 1, 415–429. MR85d:30031

[8] N. G. de Bruijn and T. A. Springer, *On the zeros of a polynomial and of its derivative II,* Indagationes Math. 9, 264-270 (1947). MR9:30b

[9] M. G. de Bruin, K. G. Ivanov and A. Sharma, *A conjecture of Schoenberg,* J. Inequal. Appl. 4, No. 3 (1999), 183-213. MR2000k:30007

[10] D. Dimitrov, *A refinement of the Gauss-Lucas theorem,* Proc. Amer. Math. Soc. 126, No.7 (1998), 2065-2070. MR98h:30005

[11] W. F. Donoghue, Monotone Matrix functions and analytic continuation. Springer, 1974. MR58:6279

[12] Ky Fan and G. Pall, *Imbedding conditions for Hermitian and normal matrices,* Canad. J. Math. 9 (1957), 298-304. MR19:6e

[13] P. Fischer and J. A. R. Holbrook, *Balayage defined by the nonnegative convex functions,* Proc. AMS 79 (1980), 445-448. MR81f:46012

[14] F. Gesztesy and B. Simon, *m-functions and inverse spectral analysis for finite and semi-infinite Jacobi matrices,* J. Anal. Math. 73, 267-297 (1997). MR99c:47039

[15] G. H. Hardy, J. E. Littlewood and G. Polya, Inequalites. Cambridge, 1988. MR89d:26016

[16] L. Hörmander, *Notions of convexity.* Birkhäuser, 1994. MR95k:00002

[17] H. Hochstadt, *On the construction of a Jacobi matrix from spectral data,* Lin. Algebra and Appl., 8 (1974), 435-446. MR52:3199

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge Univ. Press, Cambridge, 1993. MR91i:15001

[19] B. Kostant, *On the convexity, the Weyl group and the Iwasawa decomposition,* Annales scientificues de l'Ecole Normale Superieure, 6, 413-455 (1973). MR51:806

[20] M. M. Malamud, *On the formula of generalized resolvents of a nondensely defined Hermitian operator,* Ukr. Matem. Zhurn. vol. 44, N12 (1992), 1658-1688 (in Russian) (translation in Ukr. Math. J. v. 44 (1992), 1522-1547). MR95i:47046

[21] S. M. Malamud, *Operator inequalities, converse to the Jensen inequality,* Mathematical Notes, v. 69, No. 4 (2001), 633-637. MR2002g:47031

[22] S. M. Malamud, *A converse to the Jensen inequality, its matrix extensions and inequalities for minors and eigenvalues,* Linear Algebra and Applications, v. 322, (2001), 19-41. MR2001m:15048

[23] S. M. Malamud, *An inverse spectral problem for normal matrices and a generalization of the Gauss-Lucas theorem,* math.CV/0304158

[24] S. M. Malamud, *Analog of the Poincare separation theorem for normal matrices and the Gauss-Lucas theorem,* Funct. Anal. Appl., v. 37, No 3 (2003), 72-76.

[25] A. S. Markus, *Eigenvalues and singular values of the sum and product of linear operators,* Russian Math. Surveys 19 (1964), 91-120. MR29:6318

[26] M. Marcus and H. Minc, *A survey on matrix theory and matrix inequalities.* Allyn and Bacon, 1964. MR29:112

[27] A. W. Marshall and I. Olkin, *Inequalities: Theory of majorization and its applications.* Acad. Press, 1979. MR81b:00002

[28] G. Mason and B. Shapiro, *A note on polynomial eigenfunctions of a hypergeometric type operator,* Experimental mathematics, 10, 609-618. MR2003e:33038

[29] M. J. Miller, *Maximal polynomials and the Ilieff-Sendov conjecture,* Trans. Amer. Math. Soc. 321 (1990), 285–303. MR90m:30007

[30] P. Pawlowski, *On the zeros of a polynomial and its derivatives,* Trans. Amer. Math. Soc. 350 (1998), no. 11, 4461–4472. MR99a:30008

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

[32] R. R. Phelps, *Lectures on Choquet's Theorem.* Van Nostrand-Reinhold (1966). MR33:1690

[33] G. Polya and G. Szego, *Problems and theorems in analysis, vol. II.* Springer, 1976. MR53:2

[34] B. Shapiro and M. Shapiro, *This strange and misterious Rolle's Theorem,* preprint, math.CA/0302215

[35] Harold S. Shapiro, *Spectral aspects of a class of differential operators,* Operator Theory Adv. Appl., 132, pp. 361-385. Birkhäuser, Basel, 2002. MR2003i:47035

[36] G. Schmeisser, *The conjectures of Sendov and Smale,* Approx. Theory: A volume dedicated to Blagovest Sendov, DARBA, Sofia, 2002, 353-369. MR2004c:30007

[37] S. Sherman, *On a theorem of Hardy, Littlewood, Polya, and Blackwell,* Proc. Nat. Acad. Sci. USA 37 (1951), 826-831. MR13:633g

[38] I.J. Schoenberg, *A conjectured analogue of Rolle's theorem for polynomials with real or complex coefficients,* Amer. Math. Mon. 93, 8-13 (1986). MR87d:30010

[39] D. Tischler, *Critical points and values of complex polynomials,* J. Complexity 5 (1989), pp. 438-456. MR91a:30004

Departement Mathematik, HG G33.1, ETH-Zentrum, Raemistrasse 101, 8092 Zürich, Switzerland

E-mail address: semka@math.ethz.ch