Convex Spectral Functions

SHMUEL FRIEDLAND†

*Department of Mathematics and Mathematics Research Center,
University of Wisconsin-Madison, Wisconsin*

In this paper we characterize all convex functionals defined on certain convex sets of hermitian matrices and which depend only on the eigenvalues of matrices. We extend these results to certain classes of non-negative matrices. This is done by formulating some new characterizations for the spectral radius of non-negative matrices, which are of independent interest.

1. INTRODUCTION

Let A be an $n \times n$ matrix with complex entries. We arrange the eigenvalues of A in the following order

$$\text{Re } \lambda_1(A) \geqslant \text{Re } \lambda_2(A) \geqslant \cdots \geqslant \text{Re } \lambda_n(A). \tag{1.1}$$

By H_n we denote the set of all $n \times n$ hermitian matrices. For $A \in H_n$ the classical maximal characterization states

$$\lambda_1(A) = \max_{(x,x)=1} (Ax, x). \tag{1.2}$$

Thus $\lambda_1(A)$ is a convex functional on H_n. Ky Fan extended (1.2) [4]

$$\sum_{i=1}^{k} \lambda_i(A) = \max_{(x_i,x_j)=\delta_{ij}} \sum_{i=1}^{k} (Ax_i, x_i). \tag{1.3}$$

In particular $\sum_{i=1}^{k} \lambda_i(A)$ is a convex functional on H_n. A function

$$\phi : A \to \mathbb{R}(A \subseteq H_n) \tag{1.4}$$

is called a spectral function if

$$\phi(A) = F(\lambda_1(A), \ldots, \lambda_n(A)), \ F : X \to \mathbb{R}, \ X \subseteq \mathbb{R}^n_{\geqslant}. \tag{1.5}$$

Here $\mathbb{R}^n_{\geqslant}$ consists of all vectors $(x_1, ..., x_n)$, $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_n$. In Section 2 of this paper we characterize all F for which ϕ is a convex functional on A. It turns out that F must be convex on X and F Schur's order preserving [13], i.e.

$$F(\alpha) \leqslant F(\beta) \qquad \text{if} \qquad \alpha = (\alpha_1, ..., \alpha_n) < \beta = (\beta_1, ..., \beta_n), \qquad (1.6)$$

$$\sum_{i=1}^{k} \alpha_i \leqslant \sum_{i=1}^{k} \beta_i, \qquad k = 1, ..., n-1, \qquad (1.7)$$

$$\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i. \qquad (1.8)$$

In the rest of the paper we establish similar results for certain classes of non-negative matrices. Let A be an $n \times n$ non-negative matrix. As usual denote by $r(A)$ the spectral radius of A. Contrary to the symmetric case $r(A) = \lambda_1(A)$ is not in general a convex function on non-negative matrices. Consider, for example,

$$A = \begin{pmatrix} 0 & \varepsilon \\ 1 & 0 \end{pmatrix}, \qquad r(A) = \sqrt{\varepsilon}, \qquad \varepsilon \geqslant 0. \qquad (1.9)$$

Recently, Cohen [2] proved that $r(A+D)$ is convex on D_n — the set of $n \times n$ diagonal matrices. By using the general result of Kingman [12] one can show that $\log r(e^D A)$ is convex on $D_n(A \geqslant 0)$—which is similar to Cohen's theorem. In both cases it is virtually impossible to characterize the strict convexity cases. (Some special cases were discussed in [1]) We were able to characterize the strict convexity cases by giving new characterizations of the spectral radius of non-negative matrices. These characterizations extend the Donsker–Varadhan variational formula [3] and provide us with a unified proof of the convexity of the functionals $r(A+D)$ and $\log r(e^D A)$ on D_n. Also if A^{-1} is an M-matrix we get that $r(DA)$ is convex on D_n^+—the subset of non-negative matrices in D_n. This result is stronger than the convexity of $\log r(e^D A)$.

In Section 5 we show how the results of Section 2 can be extended to the non-symmetric case by assuming that A is a totally positive matrix of order $j(TP_j)$. We shall state our results in case that A is a $TP (= TP_n)$ matrix. That is all minors of A (of all orders) are non-negative. In that case we have

$$\lambda_1(e^D A) \geqslant \lambda_2(e^D A) \geqslant \cdots \geqslant \lambda_n(e^D A) \geqslant 0, \qquad D \in D_n. \qquad (1.10)$$

If A is non-singular then the last inequality is strict. Let

$$\phi(D) = F(\log \lambda_1(e^D A), ..., \log \lambda_n(e^D A)). \qquad (1.11)$$

Then ϕ is convex on $A \subset D_n$ if and only if F is convex on X and Schur's order preserving.

We remark that the results in Section 2 hold for symmetric compact

operators in Hilbert space. The results of Section 3–5 can be extended to appropriate integral operators, for example, as it was pointed out in [6].

2. CONVEX FUNCTIONS ON THE SPECTRUM OF HERMITIAN MATRICES

Let A be an $n \times n$ hermitian matrix. We can view A as a self adjoint operator on $\mathbb{C}^n$ endowed with the standard inner product

$$(x, y) = y^* x, x, y \in \mathbb{C}^n. \qquad (2.1)$$

Since the eigenvalues of A are real we arrange them in the decreasing order

$$\lambda_1(A) \geqslant \cdots \geqslant \lambda_n(A). \qquad (2.2)$$

Denote by $\xi_1, ..., \xi_n$ the corresponding set of orthonormal eigen-vectors of A

$$A\xi_i = \lambda_i(A)\xi_i, (\xi_i, \xi_j) = \delta_{ij}, i, j = 1, ..., n. \qquad (2.3)$$

Let H_n denote the set of all $n \times n$ hermitian matrices. Since $\lambda_1(A)$ has the maximal characterization

$$\lambda_1(A) = \max_{(x,x)=1} (Ax, x),$$

$\lambda_1(A)$ is a convex function on H_n. More generally we have [5]

THEOREM 2.1   *Let $\{\alpha_i\}^n$ be a decreasing sequence of real numbers*

$$\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n. \qquad (2.4)$$

Then for any A belonging to H_n

$$\sum_{i=1}^{n} \alpha_i \lambda_i(A) = \max_{(x_i, x_j) = \delta_{ij} l, j = 1, ..., n} \sum_{i=1}^{n} \alpha_i (Ax_i, x_i). \qquad (2.5)$$

Assume that the equality sign holds for some $x_1, ..., x_n$. Let

$$\alpha_1 = \cdots = \alpha_{i_1} > \alpha_{i_1+1} = \cdots = \alpha_{i_2} > \cdots > \alpha_{i_{r-1}+1} = \cdots = \alpha_{i_r} = \alpha_n. \qquad (2.6)$$

Then there exists an orthonormal eigensystem of A such that the following subspaces coincide

$$[\xi_{i_j+1}, ..., \xi_{i_{j+1}}] = [x_{i_j+1}, ..., x_{i_{j+1}}], \qquad j = 0, ..., r-1, \qquad (i_0 = 0). \qquad (2.7)$$

The characterization (2.5) in the case that $\alpha_1 = \cdots = \alpha_i = 1, \alpha_{i+1} = \cdots = \alpha_n = 0$ was established by Fan [4].

In particular

$$\phi(A) = \sum_{i=1}^{n} \alpha_i \lambda_i(A) \qquad (2.8)$$

is a convex functional on H_n if (2.4) is satisfied. That is

$$\phi(cA + (1-c)B) \leqslant c\phi(A) + (1-c)\phi(B), A, B \in H_n, 0 \leqslant c \leqslant 1. \qquad (2.9)$$

We now are ready to state the problem which we solve in this section. A function

$$\phi: A \to \mathbb{R}, \quad A \subset H_n \tag{2.10}$$

is called a spectral function if

$$\phi(A) = F(\lambda_1(A), ..., \lambda_n(A)). \tag{2.11}$$

That is ϕ is defined on the spectrum of A. Our problem is to characterize all convex spectral functions on A. Thus, we shall assume that A is convex. More restrictions on A will be specified later. To answer this problem we introduce some notation and definitions. Let $\alpha = (\alpha_1, ..., \alpha_n)$ and $\beta = (\beta_1, ..., \beta_n)$ be two vectors satisfying (2.4). According to [8, Sec. 2.18] α is majorized by β, which is denoted by $\alpha < \beta$, if

$$\sum_{i=1}^{k} a_i \leqslant \sum_{i=1}^{k} \beta_i, \quad k = 1, ..., n-1, \tag{2.12}$$

$$\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i. \tag{2.13}$$

Denote

$$\lambda(A) = (\lambda_1(A), ..., \lambda_n(A)). \tag{2.14}$$

From Theorem 2.1 we obtain

LEMMA 2.1   Let $A, B \in H_n$. Then

$$\lambda(A+B) < \lambda(A) + \lambda(B). \tag{2.15}$$

Moreover,

$$\lambda(A+B) = \lambda(A) + \lambda(B) \tag{2.16}$$

if and only A and B have a common eigenvector system

$$A\xi_i = \lambda_i(A)\xi_i, B\xi_i = \lambda_i(B)\xi_i, (\xi_i, \xi_j) = \delta_{ij}, i, j = 1, ..., n. \tag{2.17}$$

Proof   Let

$$(A+B)\xi_i = \lambda_i(A+B)\xi_i, (\xi_i, \xi_j) = \delta_{ij}, i, j = 1, ..., n. \tag{2.18}$$

So for any $\alpha = (\alpha_1, ..., \alpha_n)$ which satisfies (2.4) we get

$$\sum_{i=1}^{n} \alpha_i \lambda_i(A+B) = \sum_{i=1}^{n} \alpha_i((A+B)\xi_i, \xi_i) \leqslant \sum_{i=1}^{n} \alpha_i \lambda_i(A) + \sum_{i=1}^{n} \alpha_i \lambda_i(B). \tag{2.19}$$

This establishes (2.15). Suppose that (2.16) holds. Then we must have

$$\sum_{i=1}^{n} \alpha_i \lambda_i(A) = \sum_{i=1}^{n} \alpha_i(A\xi_i, \xi_i), \sum_{i=1}^{n} \alpha_i \lambda_i(B) = \sum_{i=1}^{n} \alpha_i(B\xi_i, \xi_i). \tag{2.20}$$

for any $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$. Choose $\alpha_i = n - i$. Then the equalities (2.7) imply (2.17). This conclusion is in fact is stated in Theorem 3.1 in [5].

Let $\mathbb{R}^n_{\geqslant}$ be as defined in the Introduction. Clearly

$$\lambda: H_n \to \mathbb{R}^n_{\geqslant}(\lambda(A) = (\lambda_1(A), ..., \lambda_n(A))). \tag{2.21}$$

Let

$$\lambda(A) = X. \tag{2.22}$$

Thus the function F in terms of which ϕ is constructed satisfies $F: X \to \mathbb{R}$.

Let D_n be the set of all $n \times n$ real diagonal matrices and D_n^1 the set of all diagonal matrices

$$D(\alpha) = \text{diag}\{\alpha_1, ..., \alpha_n\}, \alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n. \tag{2.23}$$

Given $X \subseteq \mathbb{R}^n_{\geqslant}$ we require that A should be of the form

$$A = \lambda^{-1}(X). \tag{2.24}$$

In what follows we shall assume that X is of the following type.

DEFINITION 2.1   Let $X \subseteq \mathbb{R}^n_{\geqslant}$. The set X is called strongly convex if X is convex and

$$\text{if} \quad \beta \in X, \alpha < \beta, \quad \text{then} \quad \alpha \in X. \tag{2.25}$$

Indeed, suppose that $\beta \in X$, then $D(\beta) \in A$. Thus the convexity of A implies that X is convex. Let $D(\beta) \in A$ and P be a permutation matrix. Denote by P^T the transpose of P. So

$$PD(\beta)P^T \in A. \tag{2.26}$$

The classical result of [8, sec. 2.19] implies that if $\alpha < \beta$ then

$$D(\alpha) = \sum_{i=1}^{k} a_i P_i D(\beta) P_i^T, \quad a_i > 0, \quad i = 1, ..., k, \quad \sum_{i=1}^{k} a_i = 1, \tag{2.27}$$

for some permutation matrices $P_1, ..., P_k$. Thus $D(\alpha) \in A$ and X satisfies (2.25).

THEOREM 2.2   Let X be a strongly convex set in $\mathbb{R}^n_{\geqslant}$ which contains at least one point α,

$$\alpha_1 > \alpha_2 > \cdots > \alpha_n. \tag{2.28}$$

Let $F: X \to \mathbb{R}$. Assume that $F \in C^{(1)}(X)$. Consider a spectral function $\phi: A \to \mathbb{R}(A \subset H_n)$ where ϕ and A are given by (2.11) and (2.24) accordingly. Then ϕ is convex on A if and only if F is convex on X and

$$\frac{\partial F}{\partial x_1}(\alpha) \geqslant \frac{\partial F}{\partial x_2}(\alpha) \geqslant \cdots \geqslant \frac{\partial F}{\partial x_n}(\alpha) \tag{2.29}$$

for any $\alpha \in X$. Moreover, ϕ is strictly convex on A, i.e.

$$\phi(cA + (1-c)B) < c\phi(A) + (1-c)\phi(B), A \neq B, 0 < c < 1, \tag{2.30}$$

if and only if F is strictly convex on X and

$$\frac{\partial F}{\partial x_i}(\alpha) > \frac{\partial F}{\partial x_{i+1}}(\alpha) \quad \text{if} \quad \alpha_i > \alpha_{i+1}. \tag{2.31}$$

To prove the theorem we need the following theorem of Ostrowski [13] (Theorems VII and VIII).

THEOREM 2.3   *Let X and F satisfy the assumptions of Theorem 2.2. Then F satisfies (2.29) if and only if*

$$F(\alpha) \leqslant F(\beta) \quad if \quad \alpha < \beta. \tag{2.32}$$

Moreover

$$F(\alpha) < F(\beta) \quad if \quad \alpha < \beta \quad and \quad \alpha \neq \beta \tag{2.33}$$

if and only if the condition (2.31) holds.

Proof of Theorem 2.2.   Assume first that F is convex on X. So if $\lambda(A)$, $\lambda(B) \in X$ then

$$F\left(\frac{\lambda(A) + \lambda(B)}{2}\right) \leqslant \tfrac{1}{2}(F(\lambda(A)) + F(\lambda(B))). \tag{2.34}$$

According to Theorem 2.3, the assumption (2.29) implies

$$F\left(\frac{\lambda(A + B)}{2}\right) \leqslant F\left(\frac{\lambda(A) + \lambda(B)}{2}\right) \tag{2.35}$$

by the virtue of (2.15). This shows that ϕ is convex on A. Assume furthermore that F is strictly convex on X. So if $\lambda(A) \neq \lambda(B)$ the inequality sign holds in (2.34). This implies (2.30). Suppose that $\lambda(A) = \lambda(B)$ but $A \neq B$. According to Lemma 2.1 $\lambda(A + B) \neq (\lambda(A) + \lambda(B))$. So the additional assumption (2.31) yields the inequality sign in (2.35) according to Theorem 2.3. This manifests that ϕ is strictly convex on A. Assume now that ϕ is convex on A. In particular ϕ is convex on $D_n^1 \cap$ A. This immediately implies that F is convex on X. Furthermore if ϕ is strictly convex then F is strictly convex. Let $\beta \in X$. So $D(\beta) \in$ A. Assume that $\alpha < \beta$. Then $D(\alpha) \in A$. The classical result of [8, sec. 2.19] states that

$$G\beta = \alpha, \tag{2.36}$$

where G is some doubly stochastic matrix. The Birkhoff theorem implies

$$G = \sum_{i=1}^{k} a_i P_i, \, a_i > 0, \, \sum_{i=1}^{k} a_i = 1 \tag{2.37}$$

and P_i is a permutation matrix. Therefore (2.27) holds. So the convexity of ϕ implies

$$\phi(D(\alpha)) \leqslant \sum_{i=1}^{k} a_i \phi(P_i D(\beta) P_i^T) = \phi(D(\beta)), \tag{2.38}$$

which is equivalent to (2.32). Now (2.29) follows from Theorem 2.3. Assume furthermore that ϕ is strictly convex. Then we must have (2.33) which implies (2.31) according to Theorem 2.3. The proof of the theorem is concluded.

Suppose

$$A \subset H_m, \quad m > n. \tag{2.39}$$

When we can define $\phi : A \to \mathbb{R}$ by (2.11). That is ϕ does not depend on $\lambda_{n+1}(A), ..., \lambda_m(A)$, i.e. $\partial F/\partial x_i = 0$ for $i > n$. In that case Theorem 2.2 reads:

COROLLARY 2.1   *Let the assumptions of Theorem 2.2 hold except that we have (2.39). Then ϕ is convex on A if and only if F is convex on X, the inequalities (2.29) hold and in addition*

$$\frac{\partial F}{\partial x_n}(\alpha) \geqslant 0, \quad \alpha \in X. \tag{2.40}$$

3. SOME CHARACTERIZATIONS OF THE SPECTRAL RADIUS

Let A be an $n \times n$ non-negative matrix such that there exists two positive vectors u, v satisfying

$$Au = r(A)u, \, A^T v = r(A)v, \, u^T = (u_1, ..., u_n) > 0, \, v^T = (v_1, ..., v_n) > 0. \tag{3.1}$$

Assume the normalization

$$\sum_{i=1}^{n} u_i v_i = 1. \tag{3.2}$$

Let P_n be the set of probability vectors

$$P_n = \left\{ \alpha \, | \, \alpha = (\alpha_1, ..., \alpha_n), \alpha_i \geqslant 0, \sum_{i=1}^{n} \alpha_i = 1 \right\}. \tag{3.3}$$

In [6, Sec. 3] it was shown

THEOREM 3.1   *Let A be an $n \times n$ non-negative irreducible matrix having positive entries on the diagonal (or fully indecomposable, see Remark 3.3 in [6]). Then for any $\alpha \in P_n$, with positive entries $(\alpha_i > 0)$, the function $f(x) = \sum_{i=1}^{n} \alpha_i \log (Ax)_i/x_i$ has a unique critical point $\xi = (\xi_1, ..., \xi_n)$ in the interior point of $P_n(\xi_i > 0)$ which must satisfy*

$$\min_{x>0} \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i} = \sum_{i=1}^{n} \alpha_i \log \frac{(A\xi)_i}{\xi_i}. \tag{3.4}$$

Thus, if α is chosen to be

$$\alpha = (u_1 v_1, ..., u_n v_n), \tag{3.5}$$

where u and v satisfy (3.1)–(3.2) then

$$\sum_{i=1}^{n} u_i v_i \log \frac{(Ax)_i}{x_i} \geqslant \sum_{i=1}^{n} u_i v_i \log \frac{(Au)_i}{u_i} = \log r(A), \tag{3.6}$$

21

since $x = u$ is a minimal point of $f(x)$.

From Theorem 3.1 we get

THEOREM 3.2   *Let A be an $n \times n$ non-negative matrix such that $r(A) > 0$. Then*

$$\sup_{\alpha \in P_n} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i} = \log r(A). \tag{3.7}$$

Suppose that there exists a positive vector u satisfying (3.1). Assume that

$$\inf_{x > 0} \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i} = \log r(A). \tag{3.8}$$

Then the vector v

$$v^T = (\alpha_1/u_1, \ldots, \alpha_n/u_n) \tag{3.9}$$

fulfills (3.1). In particular if A is irreducible then α is unique and given by (3.5).

Proof   As the left-hand side of (3.7) is a continuous function of A it is enough to prove (3.7) for A positive. Let $u > 0$ be the corresponding eigenvector of A. So

$$\inf_{x > 0} \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i} \leqslant \sum_{i=1}^{n} \alpha_i \log \frac{(Au)_i}{u_i} = \log r(A)$$

for any α such that $\sum_{i=1}^{n} \alpha_i = 1$. Thus

$$\sup_{\alpha \in P_n} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i} \leqslant \log r(A).$$

The above inequality together with (3.6) yields (3.7). Suppose that (3.8) holds. If $u > 0$ satisfies (3.1) then $x = u$ is a minimal point for $f(x) = \sum_{i=1}^{n} \alpha_i \log (Ax)_i/x_i$. So

$$0 = \frac{\partial f}{\partial x_j}\bigg|_{x=u} = \sum_{i=1}^{n} \frac{\alpha_i a_{ij}}{(Ax)_i} - \frac{\alpha_j}{x_j}\bigg|_{x=u} = r(A)^{-1} \sum_{i=1}^{n} \alpha_i u_i^{-1} a_{ij} - \alpha_j u_j^{-1}.$$

This shows that v given by (3.9) is a left eigenvector of A corresponding to $r(A)$. If A is irreducible, then u and v are unique up to a multiple of a positive scalar. Thus α is of the form (3.5) and since $\alpha \in P_n$, α unique. The proof of the theorem is completed.

We now bring an extended version of Theorem 3.2 which includes (3.7) and the Donsker–Varadhan characterization [3] as its special cases.

THEOREM 3.3   *Let $\Psi : \mathbb{R} \to \mathbb{R}$ be a continuous convex function on $\mathbb{R}$. Define $\Phi : \mathbb{R}_+ \to \mathbb{R}$*

$$\Phi(x) = \Psi(\log x). \tag{3.10}$$

Let A be an $n \times n$ non-negative matrix such that $r(A) > 0$. Assume

$$\Psi'(\log r(A)) \geqslant 0. \tag{3.11}$$

Then

$$\sup_{\alpha \in P_n} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \Phi\left(\frac{(Ax)_i}{x_i}\right) = \Phi(r(A)), \tag{3.12}$$

Assume that the inequality sign holds in (3.11) and suppose that there exists a positive vector u satisfying (3.1). If

$$\inf_{x > 0} \sum_{i=1}^{n} \alpha_i \Phi\left(\frac{(Ax)_i}{x_i}\right) = \Phi(r(A)), \tag{3.13}$$

then the vector v (3.9) satisfies (3.1). In particular if A is irreducible then α is unique and given by (3.5).

Proof   Let $t_0 = \log r(A)$, $\Psi'(t_0) = s$. Then the convexity of Ψ implies

$$\Psi(t) \geqslant \Psi(t_0) + (t - t_0)\Psi'(t_0).$$

So

$$\sum_{i=1}^{n} \alpha_i \Phi\left(\frac{(Ax)_i}{x_i}\right) \geqslant \Phi(r(A)) - s \log r(A) + s \sum_{i=1}^{n} \alpha_i \log \frac{(Ax)_i}{x_i}, \quad \alpha \in P_n. \tag{3.14}$$

As $s \geqslant 0$ from Theorem 3.2 and the above inequality we get

$$\sup_{\alpha \in P_n} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \Phi\left(\frac{(Ax)_i}{x_i}\right) \geqslant \Phi(r(A)). \tag{3.15}$$

Since Φ is continuous we may assume that A is positive. By choosing $x = u$ the left-hand side of (3.15) we deduce an opposite inequality of (3.15). This establishes (3.12). In case the $s > 0$ we use the arguments of Theorem 3.2 to analyze the equality (3.13). End of proof.

Letting $\psi(x) = e^x$ in Theorem 3.3 we obtain the Donsker–Varadhan characterization [3].

COROLLARY 3.1   *Let the assumptions of Theorem 3.2 hold. Then*

$$\sup_{\alpha \in P_n} \inf_{x > 0} \sum_{i=1}^{n} \alpha_i \frac{(Ax)_i}{x_i} = r(A). \tag{3.16}$$

Suppose that

$$\inf_{x > 0} \sum_{i=1}^{n} \alpha_i \frac{(Ax)_i}{x_i} = r(A). \tag{3.17}$$

If A has a positive eigenvector u then the conclusions of Theorem 3.2 apply.

Recall the classical characterization due to Wielandt [14]

$$\inf_{x > 0} \max_{1 \leqslant i \leqslant n} \frac{(Ax)_i}{x_i} = r(A) \tag{3.18}$$

for any non-negative A. Assume that Φ is an increasing function of x on $\mathbb{R}_+$.

So

$$\inf_{x>0} \sup_{\alpha \in P_n} \sum_{i=1}^{n} \alpha_i \Phi\left(\frac{(Ax)_i}{x_i}\right) = \inf_{x>0} \Phi\left(\max_{1 \le i \le n} \frac{(Ax)_i}{x_i}\right)$$

$$= \Phi\left(\inf_{x>0} \max_{1 \le i \le n} \frac{(Ax)_i}{x_i}\right) = \Phi(r(A)). \qquad (3.19)$$

Thus if Φ is increasing and satisfies the assumptions of Theorem 3.3 then we can interchange sup with inf in (3.12). The characterization (3.19) is completely equivalent to the Wielandt characterization (3.18) while (3.12) seems to be a deeper characterization.

Let A be a non-negative and non-singular. Assume furthermore that A^{-1} is an M-matrix, i.e. the off-diagonal elements of A^{-1} are non-positive. Following [6] we bring another characterization of $r(A)$.

THEOREM 3.4   *Let A be a non-negative and non-singular matrix such that A^{-1} is an M-matrix. Then*

$$\inf_{\alpha \in P_n} \sup_{x>0} \sum_{i=1}^{n} \alpha_i \frac{x_i}{(Ax)_i} = \frac{1}{r(A)}. \qquad (3.20)$$

Assume that there exists a positive vector u satisfying (3.1) and suppose

$$\sup_{x>0} \sum_{i=1}^{n} \alpha_i \frac{x_i}{(Ax)_i} = \frac{1}{r(A)}. \qquad (3.21)$$

Then v given by (3.9) satisfies (3.1). In particular if A is irreducible then α is unique and given by (3.5).

Proof   We have available the representation

$$A^{-1} = rI - B, \qquad B \ge 0, \qquad r > r(B) \qquad (3.22)$$

and B is reducible if and only if A is reducible (e.g. [9, chap. 8]). Again, as in the proof of Theorem 3.2 one may assume that B is positive. By letting x to be equal to the positive eigenvector u of A we immediately deduce

$$\inf_{\alpha \in P_n} \sup_{x>0} \sum_{i=1}^{n} \alpha_i \frac{x_i}{(Ax)_i} \ge \frac{1}{r(A)}. \qquad (3.23)$$

Let α be given by (3.5). Obviously for any $x > 0$ and $y = Ax$

$$\sum_{i=1}^{n} u_i v_i \frac{x_i}{(Ax)_i} = \sum_{i=1}^{n} u_i v_i \frac{(A^{-1}y)_i}{y_i} = r - \sum_{i=1}^{n} u_i v_i \frac{(By)_i}{y_i}. \qquad (3.24)$$

From Corollary 3.1 it follows

$$\sum_{i=1}^{n} u_i v_i \frac{(By)_i}{y_i} \ge r(B). \qquad (3.25)$$

So

$$\sum_{i=1}^{n} u_i v_i \frac{x_i}{(Ax)_i} \le r - r(B) = \frac{1}{r(A)} \qquad (3.26)$$

and the equality sign holds if $x = u$. This establishes (3.20). The equality (3.21) is analyzed in the same way as in Theorem 3.2.

Remark 3.1   Theorem 3.4 does not hold for arbitrary non-negative matrices, take for example A to be a permutation matrix $P \ne I$. Therefore Theorem 3.4 is not a special case of Theorem 3.3.

4. CONVEXITY PROPERTIES OF THE SPECTRAL RADIUS

Let A be an $n \times n$ non-negative matrix. Consider the matrix $A + D$, $D \in D_n$. Assume that the eigenvalues of $A + D$ arranged in the order

$$\operatorname{Re} \lambda_1(A) \ge \operatorname{Re} \lambda_2(A) \ge \cdots \ge \operatorname{Re} \lambda_n(A). \qquad (4.1)$$

Let

$$\rho(D) = \lambda_1(A + D). \qquad (4.2)$$

We claim that $\rho(D)$ is real. If D is non-negative this fact is a consequence of the Perron–Frobenius theorem. For an arbitrary D consider $A + D + aI$

$$\lambda_k(A + D + aI) = \lambda_k(A + D) + a, \qquad k = 1, \ldots, n. \qquad (4.3)$$

Thus if a is big enough, $A + D + aI \ge 0$ and (4.3) implies that $\rho(D)$ is real. Moreover by considering the matrix $B = A + D + aI$ and using the Donsker–Varadhan characterization for B we get the following characterization for $\rho(D)$

$$\rho(D) = \sup_{\alpha \in P_n} L_1(D, \alpha). \qquad (4.4)$$

Here $L_1(D, \alpha)$ is a linear functional on D_n

$$L_1(D, \alpha) = \sum_{i=1}^{n} \alpha_i d_i + \inf_{x>0} \sum_{i=1}^{n} \alpha_i \frac{(Ax)_i}{x_i}, \qquad (4.5)$$

$$\alpha = (\alpha_1, \ldots, \alpha_n), \qquad D = \operatorname{diag}\{d_1, \ldots, d_n\}.$$

It is a standard fact that (4.4) and (4.5) imply the convexity of $\rho(D)$ on the set D_n. More precisely we have:

THEOREM 4.1   *Let A be a fixed $n \times n$ non-negative matrix. Assume that $\rho(D)$. $D \in D_n$, is given by (4.2). Then $\rho(D)$ is a convex functional on D_n.*

$$\rho((D_1 + D_2)/2) \le (\rho(D_1) + \rho(D_2))/2. \qquad (4.6)$$

Moreover if A is irreducible then the equality sign holds in (4.6) if and only if

$$D_2 - D_1 = aI \qquad (4.7)$$

for some a.

Proof As we pointed out (4.6) is a consequence of (4.4). So it is enough to analyze the equality case. Let

$$A_1 = A + (D_1 + D_2)/2, \ A_1 u = r_1 u, \ A_1^T v = r_1 v, r_1 = \rho((D_1 + D_2)/2). \quad (4.8)$$

As A is irreducible we may assume that $u, v > 0$ and the normalization (3.2) holds. Let α be given by (3.5). So

$$L_1((D_1 + D_2)/2, \alpha) = \inf_{x > 0} \sum_{i=1}^n \alpha_i \frac{(A_1 x)_i}{x_i} = \sum_{i=1}^n \alpha_i \frac{(A_1 u)_i}{u_i} = r_1. \quad (4.9)$$

Applying the results of Section 3 in [6] for the function

$$f(x, B) = \sum_{i=1}^n \alpha_i \frac{(Bx)_i}{x_i}, \quad (4.10)$$

where $B + bI$ is irreducible matrix, for some positive b, we deduce that $f(x, B)$ has a unique critical point in the interior of P_n which must be the minimum point ($f(x) = +\infty$ on the boundary of P_n). The equality sign in (4.6) implies

$$L_1(D_1, \alpha) = \rho(D_1), L_1(D_2, \alpha) = \rho(D_2). \quad (4.11)$$

That is

$$f(x, A + D_1) \geqslant f(u, A + D_1) = \rho(D_1), f(x, A + D_2) \geqslant f(u, A + D_2) = \rho(D_2). \quad (4.12)$$

The uniqueness of the minimal point of $f(X, B)$ implies

$$(A + D_1)u = \rho(D_1)u, (A + D_2)u = \rho(D_2)u. \quad (4.13)$$

As $u > 0$ (4.7) follows the above equality. The proof of the theorem is completed.

The inequality (4.6) was established by Cohen in [2]. (Consult [1] for partial cases of (4.6).) The equality case for an irreducible A was conjectured in [2].

Let A be a non-negative matrix such that $r(A) > 0$. Clearly, for any $D \in D_n$, $r(e^D A)$ is also positive. Define

$$R(D) = \log r(e^D A). \quad (4.14)$$

According to Theorem 3.2,

$$R(D) = \sup_{\alpha \in P_n} L_2(D, \alpha), \quad (4.15)$$

where

$$L_2(D, \alpha) = \sum_{i=1}^n \alpha_i d_i + \inf_{x > 0} \sum_{i=1}^n \alpha_i \log \frac{(Ax)_i}{x_i}. \quad (4.16)$$

Combining (4.15) and (4.16) and using the uniqueness result stated in Theorem 3.1 as in the proof of Theorem 4.1 we deduce.

THEOREM 4.2   *Let A be a fixed $n \times n$ non-negative matrix having a positive spectral radius. Assume that $R(D)$ is given by (4.14). Then $R(D)$ is a convex*

functional on D_n.

$$R((D_1 + D_2)/2) \leqslant (R(D_1) + R(D_2))/2. \quad (4.17)$$

Moreover if A is irreducible and the diagonal entries of A are positive (or A is fully indecomposable) then the equality sign holds in (4.17) if and only if (4.7) holds for some a.

The inequality (4.17) follows easily from the Kingman's result [12]. Indeed, since each entry of $e^D A$ is log convex on D_n then Kingman's theorem implies that $r(e^D A)$ is also log convex. However, the equality case in (4.17) cannot be analyzed by the methods given in [12].

Assume that $A, B \in H_n$ and furthermore A is positive definite ($(Ax, x) > 0$ for $x \neq 0$). Then BA is similar to $A^{1/2} B A^{1/2}$. This shows that $\lambda_1(BA)$ is a convex functional on H_n for a fixed positive definite A. If in addition A has non-negative entries then $\lambda_1(DA)$ is convex on D_n. This result does not apply in general for non-negative matrices. For example, take A to be a permutation matrix $P \neq I$. However, $\lambda_1(DA)$ is convex on D_n^+—the set of $n \times n$ non-negative diagonal matrices if A^{-1} is an M-matrix.

THEOREM 4.3   *Let A^{-1} be an M-matrix. Then $r(DA)$ is a convex functional on D_n^+.*

$$r\left(\frac{(D_1 + D_2)}{2} A\right) \leqslant \tfrac{1}{2}(r(D_1 A) + r(D_2 A)). \quad (4.18)$$

Moreover if A is irreducible then the equality sign in (4.18) holds if and only if

$$D_2 = a D_1 \quad (4.19)$$

for some positive a provided that D_1 or D_2 have positive diagonal elements.

Proof Using the continuity argument we may assume that in the decomposition (3.22) B is positive (irreducible), i.e. A is positive (irreducible). Thus if all diagonal elements of $D_0 = \text{diag}\{d_1^0, ..., d_n^0\}$ are positive then $D_0 A$ is positive (irreducible). According to the Perron–Frobenius theorem $r(D_0 A)$ is a simple root of the $\det(\lambda I - D_0 A) = 0$. By the implicit function theorem $r(DA)$ is an analytic function of D in the neighbourhood of D_0. Then the convexity of $r(DA)$ would follow if we show that

$$r(DA) \geqslant r(D_0 A) + \sum_{i=1}^n (d_i - d_i^{(0)}) \frac{\partial r(DA)}{\partial d_i}\Big|_{D^0}, \quad (4.20)$$

for any D_0 with positive diagonal elements. Let ξ, η be the eigenvectors corresponding to $D_0 A$ and $A^T D_0$

$$D_0 A \xi = r(D_0 A)\xi, \ A^T D_0 \eta = r(D_0 A)\eta, \ 0 < \xi = (\xi_1, ..., \xi_n), \quad (4.21)$$

$$0 < \eta = (\eta_1, ..., \eta_n), \ \sum_{i=1}^n \xi_i \eta_i = 1.$$

It can be shown that

$$\frac{\partial r(DA)}{\partial d_i}\Big|_{D^0} = \eta^T \frac{\partial D}{\partial d_i} A\xi = r(D_0 A)\frac{\eta_i \xi_i}{d_i^{(0)}} \qquad i = 1, ..., n. \qquad (4.22)$$

This can be done by bringing $D_0 A$ to the Jordan form and using the simplicity of $r(D_0 A)$. See for example [11, II, §5.4]. Thus (4.20) is equivalent to

$$r(DA) \geq r(D_0 A) \sum_{i=1}^{n} \frac{d_i}{d_i^{(0)}} \xi_i \eta_i. \qquad (4.23)$$

This inequality was established in [6]. It follows directly from (3.26). Indeed suppose that D has positive diagonal elements and let

$$DAw = r(DA)w, \quad w = (w_1, ..., w_n) > 0. \qquad (4.24)$$

Then according to (3.26)

$$\frac{1}{r(D_0 A)} \geq \sum_{i=1}^{n} \xi_i \eta_i \frac{w_i}{(D_0 Aw)_i} = \sum_{i=1}^{n} \xi_i \eta_i \frac{d_i}{d_i^{(0)}} \frac{w_i}{(DAw)_i} = \frac{1}{r(DA)} \sum_{i=1}^{n} \xi_i \eta_i \frac{d_i}{d_i^{(0)}},$$

which establishes (4.23) for D with positive diagonal. So (4.18) holds in the interior of D_n^+. The continuity argument implies the validity of (4.18) on D_n^+. Suppose that A is also irreducible. Then B in the decomposition (3.22) is also irreducible, since the inverse of block triangular matrix is also a block triangular one. As in the proof of Theorem 4.1 strict inequality holds in (4.23) unless $D_0 A$ and DA have the same positive eigenvector. So $D = aD_0$ for some $a > 0$. This shows that we have strict inequality in (4.18) unless (4.19) holds provided that D_0 (which is either D_1 or D_2) have positive diagonal. The proof of the theorem is completed.

We conclude this section by pointing out that the convexity of $r(DA)$ on D_n^+ is a stronger result then the convexity of $\log r(e^Q A)$ on D_n. Indeed, let

$$D_0 = e^{Q_0}, \quad Q_0 = \text{diag}\{q_1^{(0)}, ..., q_n^{(0)}\}, \quad q_i^{(0)} = \log d_i^{(0)}, \quad i = 1, ..., n. \qquad (4.25)$$

Suppose that $\log r(e^Q A)$ is convex at $Q = Q_0$. This means

$$\log r(e^Q A) \geq \log r(D_0 A) + r(D_0 A)^{-1} \sum_{i=1}^{n} \frac{\partial r(e^Q A)}{\partial q_i}\Big|_{Q=Q_0} (q_i - q_i^{(0)}), \qquad (4.26)$$

$$Q = \text{diag}\{q_1, ..., q_n\}.$$

As in the proof of Theorem 4.3

$$\frac{\partial r(e^Q A)}{\partial q_i}\Big|_{Q=Q_0} = \eta^T \frac{\partial e^Q}{\partial q_i}\Big|_{Q_0} A\xi = r(D_0 A)\eta_i \xi_i, \quad i = 1, ..., n \qquad (4.27)$$

where η, ξ given by (4.21).

Thus (4.26) is equivalent to

$$r(e^Q A) \geq r(D_0 A) \prod_{i=1}^{n} \left(\frac{e^{q_i}}{d_i^{(0)}}\right)^{\xi_i \eta_i} = r(D_0 A) \prod_{i=1}^{n} \left(\frac{d_i}{d_i^{(0)}}\right)^{\xi_i \eta_i},$$

$$q_i = \log d_i, \quad i = 1, ..., n. \qquad (4.28)$$

Using the relation between the arithmetic and the geometric means from (4.23) we get

$$r(e^Q A) \geq r(D_0 A) \sum_{i=1}^{n} \frac{e^{q_i}}{d_i^{(0)}} \xi_i \eta_i \geq r(D_0 A) \prod_{i=1}^{n} \left(\frac{e^{q_i}}{d_i^{(0)}}\right)^{\eta_i \xi_i}. \qquad (4.29)$$

That is the convexity of $r(DA)$ at $D_0 \in D_n^+$ implies the convexity of $\log r(e^Q A)$ at $Q_0 = \log D_0$. This demonstrates that the convexity of $r(DA)$ on D_n^+ implies the convexity of $\log r(e^D A)$ on D_n. On the other hand if A is a permutation matrix $P \neq I$ then $r(PA)$ is not convex on D_n^+ (for details see [6], Section 3), while $\log r(e^\circ P)$ is convex on D_n.

5. CONVEX FUNCTIONS ON THE SPECTRUM OF TOTALLY POSITIVE MATRICES

A real valued $n \times n$ matrix is called a totally (strictly totally) positive matrix of order j if all minors of A of order less or equal to j are non-negative (positive). We denote these matrices by $TP_j(STP_j)$. For $j = n$ we call these matrices simply by $TP(STP)$. A matrix A is called oscillating if A is TP and some power of A is STP. It is known that a TP matrix if oscillating if and only if

$$a_{ii} > 0, a_{i(i-1)} > 0, a_{i(i+1)} > 0, i = 1, ..., n, A = (a_{ij})_1^n \geq 0. \qquad (5.1)$$

In that case A is totally indecomposable.

If A is TP_j then

$$\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_j(A) \geq |\lambda_k(A)|, \quad k = j+1, ..., n. \qquad (5.2)$$

If A is STP_j then we have strict inequalities in (5.2). See [7] and [10] for proofs of these results and additional properties of these matrices. Let A be TP_j. Define $\phi : A \to \mathbb{R}(A \subset D_n)$ as follows

$$\phi(D) = F(\log \lambda_1(e^D A), ..., \log \lambda_j(e^D A)). \qquad (5.3)$$

As in Section 2 we are looking for necessary and sufficient conditions on F which imply that ϕ is a convex function on $A \subset D_n$ for any A which is TP_j. It turns out that we have an analogous result to Theorem 2.2. To do so we need few notations and definitions. Let $1 \leq j \leq n$. Put $\bar{\alpha} = (\alpha_1, ..., \alpha_j)$ and $\bar{\beta} = (\beta_1, ..., \beta_j)$. We define $\bar{\alpha} \ll \bar{\beta}$ if (2.12) holds for $k = 1, ..., j$. Thus if $\alpha = (\alpha_1, ..., \alpha_n)$, $\beta = (\beta_1, ..., \beta_n)$ and $\alpha < \beta$ then $\bar{\alpha} \ll \bar{\beta}$. Conversely, if $\bar{\alpha} \ll \bar{\beta}$ we can extend $\bar{\alpha}$ to α and $\bar{\beta}$ to β such that $\alpha < \beta$. A set $\bar{X} \subseteq \mathbb{R}_{\geq}^j$ is called a super convex if $\bar{X}$ is convex and

$$\text{if} \quad \bar{\beta} \in \bar{X}, \bar{\alpha} \ll \bar{\beta}, \quad \text{then} \quad \bar{\alpha} \in \bar{X}. \qquad (5.4)$$

Clearly $\bar{X}$ is super convex in $\mathbb{R}^j_{\geqslant}$ if and only if it could be extended to $X \subseteq \mathbb{R}^n_{\geqslant}$ such that X is strongly convex in $\mathbb{R}^n_{\geqslant}$. Using the above arguments and Ostrowski's result (Theorem 2.3) we get

LEMMA 5.1   Let $\bar{X}$ be a super convex set in $\mathbb{R}^j_{\geqslant}$. Let $F : \bar{X} \to \mathbb{R}$. Assume that $F \in C^{(1)}(\bar{X})$. Then

$$F(\bar{\alpha}) \leqslant F(\bar{\beta}) \quad \text{if} \quad \bar{\alpha} \ll \bar{\beta} \tag{5.5}$$

if and only if

$$\frac{\partial F}{\partial x_1}(\bar{\alpha}) \geqslant \frac{\partial F}{\partial x_2}(\bar{\alpha}) \geqslant \cdots \geqslant \frac{\partial F}{\partial x_j}(\bar{\alpha}) \geqslant 0 \tag{5.6}$$

for any $\bar{\alpha} \in \bar{X}$. Moreover strict inequality in (5.5) holds for $\bar{\alpha} \neq \bar{\beta}$ if and only if

$$\frac{\partial F}{\partial x_i}(\bar{\alpha}) > \frac{\partial F}{\partial x_{i+1}}(\bar{\alpha}) \quad \text{if} \quad \alpha_i > \alpha_{i+1}, \quad \frac{\partial F}{\partial x_j}(\bar{\alpha}) > 0 \quad \text{if} \quad \alpha_j > 0. \tag{5.7}$$

Assume that A is TP_j. Denote

$$\lambda^{(j)}(A) = (\lambda_1(A), \ldots, \lambda_j(A)), \quad \log \lambda^{(j)}(A) = (\log \lambda_1(A), \ldots, \log \lambda_j(A)). \tag{5.8}$$

THEOREM 5.1   Let A be an $n \times n$ non-singular TP_j matrix. If $j < n$ then

$$\log \lambda^{(j)}(e^{(D_1 + D_2)/2}A) \ll \tfrac{1}{2}\log \lambda^{(j)}(e^{D_1}A) + \log \lambda^{(j)}(e^{D_2}A)). \tag{5.9}$$

If $j = n$ then

$$\log \lambda(e^{(D_1 + D_2)/2}A) < \tfrac{1}{2}[\log \lambda(e^{D_1}A) + \log \lambda(e^{D_2}A)]. \tag{5.10}$$

If in addition A satisfies (5.1), or more generally A is totally indecomposable, then

$$\log \lambda^{(j)}(e^{(D_1 + D_2)/2}A) = \tfrac{1}{2}[\log \lambda^{(j)}(e^{D_1}A) + \log \lambda^{(j)}(e^{D_2}A)] \tag{5.11}$$

for any $1 \leqslant j \leqslant n$ if and only if (4.7) is satisfied for some a.

Proof   Denote by $C_k(A)$ the kth compound of A. Thus

$$C_k(e^D) = e^{\varphi_k(D)} \tag{5.12}$$

where φ_k is well defined map $\varphi_k : D_n \to D_{\binom{n}{k}}$. It is easy to see using the properties of the compound matrices that φ_k is a linear map. According to Theorem 4.2 $\log r(e^D C_k(A))$ is convex on $D_{\binom{n}{k}}$ for $k = 1, \ldots, j$. Note that the non-singularity of A implies that $r(C_k(A)) > 0$. Thus $\log r(e^{\varphi_k(D)}C_k(A))$ is convex on D_n. Let

$$R_k(D) = \sum_{i=1}^{k} \log \lambda_i(e^D A). \tag{5.13}$$

It is well known that

$$R_k(D) = \log r(C_k(e^D A)). \tag{5.14}$$

Therefore $R_k(D)$ is convex on D_n for $k = 1, \ldots, j$. This equivalent to (5.9) for

---

$j < n$. For $j = n$, $R_n(D)$ is linear on D as

$$R_n(D) = \log \det(e^D A) = \sum_{i=1}^{n} d_i + \log \det(A). \tag{5.15}$$

This verifies (5.10) if A is a TP matrix. Suppose that in addition A is totally indecomposable. According to Theorem 4.2 we have a strict inequality in (4.17) unless (4.7) holds. Thus (5.11) can be satisfied if only (4.7) holds. Trivially (4.7) implies (5.11). The proof of the theorem is completed.

THEOREM 5.2   Let $\bar{X}$ be a super convex set in $\mathbb{R}^j_{\geqslant}$ for $1 \leqslant j \leqslant n$ (a strongly convex set containing a point $\alpha, \alpha_1 > \cdots > \alpha_n$, if $j = n$). Let $F : \bar{X} \to \mathbb{R}$. Assume that $F \in C^{(1)}(\bar{X})$. Let A be a given $n \times n$ non-singular TP_j matrix. Consider a spectral function $\phi : \mathbf{A} \to \mathbb{R}$, given by (5.3), where $\mathbf{A}$ is a convex set in D_n such that

$$\log \lambda^{(j)}(e^D A) \subseteq \bar{X}, \quad D \in \mathbf{A}. \tag{5.16}$$

Then, for all such A, ϕ is convex if and only if F is convex on $\bar{X}$ and satisfies (5.6) in case that $1 \leqslant j \leqslant n$. Moreover, if A is totally indecomposable then ϕ is strictly convex if and only if F is strictly convex and satisfies (5.7). In case $j = n$, ϕ is convex (strictly convex provided that A is totally indecomposable) if and only if F satisfies the assumptions of Theorem 2.2.

Proof   A proof of this theorem can be achieved by modifying in the obvious way the proof of Theorem 2.2. In fact, all the arguments of the proof of Theorem 2.2 carry over if one notices that the identity matrix is TP.

References†

[1] J. E. Cohen, Derivatives of the spectral radius as a function of non-negative matrix elements, Math. Proc. Phil. Soc. 83 (1978), 183–190.

[2] J. E. Cohen, Random evolutions and the spectral radius of a non-negative matrix, Math. Proc. Camb. Phil. Soc. 86 (1979), 345–350.

[3] M. D. Donsker and R. S. Varadhan, On a variational formula for the principal eigenvalue for operators with maximum principle, Proc. Acad. Sci. U.S.A., 72 (1975), 780–783.

[4] K. Fan, On a theorem of Weyl concerning eigenvalues of linear transformations, Proc. Nat. Acad. Sci. U.S.A. 35 (1949), 652–655.

[5] S. Friedland, Inverse eigenvalue problems, Linear Algebra Appl. 17 (1977), 15–51.

[6] S. Friedland and S. Karlin, Some inequalities for the spectral radius of non-negative matrices and applications, Duke Math. J. 42 (1975), 459–490.

[7] F. R. Gantmacher and M. G. Krein, Oscillating Matrices and Kernels and Small Vibrations of Mechanical Systems, Moscow, 1950.

[8] G. H. Hardy, J. E. Littlewood and G. Pólya, Inequalities, 2nd edition, Cambridge, 1952.

[9] S. Karlin, Mathematical Methods and Theory of Games, Programming and Economics, Addison Wesley, 1959.

[10] S. Karlin, Total Positivity, vol. I. Stanford University, 1968.

[11] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.

---

[12] J. F. C. Kingman, A convexity property of positive matrices, *Quart. J. Math. Oxford* **12** (1961), 283–284.

[13] A. Ostrowski, Sur quelques applications des fonctions convexes et concaves au sen de I. Schur, *J. Math. Pures Appl.* **31** (1952), 253–292.

[14] H. Wielandt, Unzerlegbare, nicht negative Matrizen, *Math. Z.* **52** (1950), 642–648.

On the Đoković Conjecture for Matrices of Rank Two

DAVID LONDON

Department of Mathematics, Technion, I.I.T., Haifa, Israel

The Đoković conjecture asserts that

$$(*) \quad p_k(A) \geqslant \frac{(n-k+1)^2}{nk} p_{k-1}(A), \qquad k = 2, \ldots, n.$$

Here A is any doubly stochastic $n \times n$ matrix and $p_k(A)$ is the sum of the permanents of all $k \times k$ submatrices of A.

We prove that if the case $k = l$ of (*) holds for all $l \times l$, $l \leqslant n$, doubly stochastic matrices of rank two, then all the other cases of (*) also hold for the same matrices. In a preceding paper a similar result was obtained for the generalized van der Waerden conjecture. In both cases the proof is based on a representation of $p_k(A)$ by means of a function $(f(x), g(x))$ of a pair of associated polynomials $\{f(x), g(x)\}$. Using this representation we also obtain some results on minimizing matrices corresponding to the van der Waerden and to the Đoković conjectures. Finally we outline some properties of the function $(f(x), g(x))$.

1. INTRODUCTION

Let A be an $n \times n$ matrix, let $p(A)$ be the permanent of A and let $p_k(A)$ be the sum of the permanents of all $\binom{n}{k}^2$ $k \times k$ sub-matrices of A. Denote by Ω_n the set of all $n \times n$ doubly stochastic matrices and by J_n the $n \times n$ matrix all of whose entries are equal to $1/n$.

The generalized van der Waerden conjecture [11] asserts that if $A \in \Omega_n$, then

$$p_k(A) \geqslant p_k(J_n), \qquad k = 1, \ldots, n, \tag{1}$$

with equality, for $k > 1$, if and only if $A = J_n$. The case $k = n$ of (1) is the (ordinary) van der Waerden conjecture [12]

$$p(A) \geqslant p(J_n) = \frac{n!}{n^n}. \tag{2}$$