STUDY OF THE PERMANENT CONJECTURE AND SOME OF ITS GENERALIZATIONS.II

BY

O. S. ROTHAUS[1]

ABSTRACT. In this paper we investigate in a more systematic manner some of the topics initiated in part I of the paper with same title [5]. More specifically, we study in greater detail the properties of the function $E(y)$ defined in [5] attached to convex polytopes, whose properties in the special case of the space of doubly stochastic matrices are connected with the permanent conjecture. Some close links with Perron-Frobenius theory are developed, and we obtain as a by-product of our study what is, I believe, a new expression for the maximum eigenvalue of a nonnegative matrix, which leads to some new estimates of the same. A final section of the paper investigates some purely algebraic properties of $E(y)$, and we obtain some very interesting information connecting a doubly stochastic matrix and its transversals.

In order to keep this paper as self-contained as possible, facts used here drawn from part I are stated with as much explicit detail as possible.

We write $A \geqslant 0$ for a matrix A of any size whatsoever whose entries are nonnegative; $A > 0$ if the entries are positive. If A and B have the same dimensions, by A^B we denote $\prod_{i,j} A_{ij}^{B_{ij}}$, $0^0 = 1$. Their Hadamard product (entry by entry multiplication) is denoted $A * B$. Otherwise, multiplication of matrices, in distinction to the usage in part I, [5], of this paper, is ordinary matricial multiplication.

Let α be a $k \times k$ diagonal matrix with positive diagonal entries α_1, $\alpha_2, \ldots, \alpha_k$. We are going to specialize the results of [5, p. 75] to the case where K there described is V_α, V_α being the convex polytope of nonnegative matrices whose ith row sum and ith column sum is α_i for all i, sitting in the affine space $\mathbf{R}^{k^2}$. As a determining set [5, p. 75] for V_α, we take the coordinate functions themselves.

Also as in [5, p. 83], we attach to V_α the group (under Hadamard multiplication) G_α of positive matrices g such that $g^x = 1$, $\forall x \in V_\alpha$. It is easy to see that $g \in G_\alpha$ if and only if

$$g_{ij} = \lambda_i \cdot \mu_j \quad \text{with} \quad \prod \lambda_i^{\alpha_i} = \prod \mu_i^{\alpha_i} = 1, \lambda_i, \mu_i > 0.$$

Put $d = \Sigma \alpha_i$. As in [5, p. 86], we associate to V_α a function $E_\alpha(y)$, defined for all nonnegative $k \times k$ matrices y by:

$$E_\alpha(y) = \frac{1}{d} \min_{g \in G_\alpha} \sum_{i,j} g_{ij} y_{ij} = \frac{1}{d} \min_{\lambda, \mu > 0} \sum_{i,j} y_{ij} \lambda_i \mu_j,$$

where the second minimum is constrained by $\Pi \lambda_i^{\alpha_i} = \Pi \mu_i^{\alpha_i} = 1$. The following properties of E_α are known from [5]:

(i) E_α is continuous, homogeneous of degree one, and $E_\alpha(x + y) \geqslant E_\alpha(x) + E_\alpha(y)$.

(ii) $E_\alpha(x) = 1$ if $x \in V_\alpha$.

(iii) If u and v are diagonal matrices, with positive diagonals, $E_\alpha(uyv) = u^{\alpha/d} v^{\alpha/d} E_\alpha(y)$.

(iv) $E_\alpha(y) = \max_{x \in V_\alpha} [y^x / x^x]^{1/d}$.

(v) If y has the same support as some element of V_α then y may be written $y = E_\alpha(y) u x v$, where $x \in V_\alpha$ and u and v are nonnegative diagonal matrices such that $u^\alpha = v^\alpha = 1$.

(The support of a matrix is simply the set of positions with nonzero entry. (v) above is Theorem 16 of [5, p. 85].)

For any $A \geqslant 0$, define $\lambda(A)$ to be the maximum real eigenvalue of A.

THEOREM 1. $E_\alpha(A) = \min \lambda(\alpha^{-1} A D)$, where the minimum is taken over diagonal matrices $D \geqslant 0$ with $D^\alpha = 1$.

From the definition

$$E_\alpha(A) = \frac{1}{d} \min_{\lambda, \mu > 0} \frac{\Sigma A_{ij} \lambda_i \mu_j}{\Pi_i \lambda_i^{\alpha_i/d} \mu_i^{\alpha_i/d}}.$$

Suppose firstly $A > 0$. Let $Au = \lambda(A)u$, $u > 0$, and put $\lambda_i = \alpha_i / u_i$. Then $E_\alpha(A) \leqslant \lambda(A) / \alpha^{\alpha/d}$. Since $E_\alpha(AD) = D^{\alpha/d} E_\alpha(A)$ for a diagonal matrix D, we have $\alpha^{\alpha/d} E_\alpha(A) = E_\alpha(\alpha A) \leqslant \lambda(A)$, or $E_\alpha(A) \leqslant \lambda(\alpha^{-1} A)$, from which we obtain $E_\alpha(A) \leqslant \min_{D^\alpha = 1} \lambda(\alpha^{-1} A D)$, if $A > 0$. Since E_α and λ are continuous functions, the above inequality continues to hold for any $A \geqslant 0$.

Since $A > 0$, we may write $A = \delta D_1 B D_2$, with $B \in V_\alpha$, $\delta = E_\alpha(A)$, and D_1 and D_2 nonnegative diagonal matrices with $D_1^\alpha = D_2^\alpha = 1$. Also $\lambda(\alpha^{-1} B) = 1$, since $\alpha^{-1} B$ has all row sums equal one. Hence

$$\min_{D^\alpha = 1} \lambda(\alpha^{-1} A D) \leqslant \lambda(\delta \alpha^{-1} D_1 B D_2 D_1^{-1} D_2^{-1}) = \lambda(\delta \alpha^{-1} B) = \delta.$$

So the theorem is true if $A > 0$.

Now let A be arbitrary nonnegative, and put $F(A) = \min_{D^\alpha = 1} \lambda(\alpha^{-1} A D)$. If A, B and $A - B$ are nonnegative then it is known that $\lambda(A) \geqslant \lambda(B)$; thus also $F(A) \geqslant F(B)$. Let A_ν be a positive sequence approaching A monotonically from above. Then since E_α is continuous, $\lim E_\alpha(A_\nu) = E_\alpha(A) =$

$\lim F(A_\nu)$. But $F(A_\nu) \geqslant F(A)$, so $E_\alpha(A) \geqslant F(A)$. Since always $E_\alpha(A) \leqslant F(A)$, the theorem is proved.

For our next result, we have first to define the convex set V, which shall consist of all nonnegative $k \times k$ matrices whose ith row sum is the same as the ith column sum for all i, and whose entry sum is one. If $x \in V$, we denote by $\gamma(x)$ the vector of column sums of x. $\underline{V}$ denotes the subset of V for which no row sum is zero.

THEOREM 2. *Let $A \geqslant 0$. Then*

$$\lambda(A) = \max_{x \in V} \left[\frac{\gamma(x)^{\gamma(x)} A^x}{x^x} \right].$$

From Theorem 1, we have $\lambda(A) \geqslant \alpha^{\alpha/d} E_\alpha(A) = E_\alpha(\alpha A)$. First let $A > 0$, and $Au = \lambda(A)u$, $u > 0$; $vA = \lambda(A)v$, $v > 0$. Put $\beta_i = u_i v_i$, and U be the diagonal matrix with diagonal entries u_i. Then $E_\beta(\beta A) = E_\beta(\beta U^{-1} A U) = \lambda(A)$, since $U^{-1} A U$ is $\lambda(A)$ times an element of V_α. Thus for $A > 0$, $\lambda(A) = \max_\alpha \alpha^{\alpha/d} E_\alpha(A)$.

Note that for any scalar ρ, $E_\alpha(\alpha A) = E_{\rho\alpha}(\rho \alpha A)$. So it is also true for $A > 0$ that $\lambda(A) = \max_{\alpha, \Sigma \alpha_i = 1} \alpha^\alpha E_\alpha(A)$. We also know that $E_\alpha(A) = \max_{x \in V_\alpha}[A^x/x^x]^{1/d}$. So we may conclude for $A > 0$:

$$\lambda(A) = \max_{\alpha_i > 0, \Sigma \alpha_i = 1} \max_{x \in V_\alpha} \left[\alpha^\alpha \frac{A^x}{x^x} \right] = \max_{x \in \underline{V}} \left[\gamma(x)^{\gamma(x)} \frac{A^x}{x^x} \right].$$

Since $\gamma(x)^{\gamma(x)} A^x / x^x$ is a continuous function of x on set V if $A > 0$, we have

$$\lambda(A) = \max_{x \in V} \left[\gamma(x)^{\gamma(x)} \frac{A^x}{x^x} \right].$$

Now let A be arbitrary nonnegative, and A_ν a positive sequence approaching A monotonically from above. Put

$$F(A) = \max_{x \in V} \left[\gamma(x)^{\gamma(x)} \frac{A^x}{x^x} \right].$$

If A, B, and $A - B$ are nonnegative, clearly $F(A) \geqslant F(B)$. Since λ is continuous, we have

$$\lambda(A) = \lim_{\nu \to \infty} (A_\nu) = \lim_{\nu \to \infty} F(A_\nu) \geqslant F(A).$$

But also, since V is compact,

$$\lambda(A_\nu) = \left[\gamma(x_\nu)^{\gamma(x_\nu)} (A_\nu^{x_\nu} / x_\nu^{x_\nu}) \right]$$

for some $x_\nu \in V$. We may suppose that the sequence x_ν has a limit u, and then it is easy to see, since $A_\nu > 0$, that $\lim A_\nu^{x_\nu} = A^u$. Hence

$$\lambda(A) = \gamma(u)^{\gamma(u)} \frac{A^u}{u^u} \leqslant \max_{x \in V} \left[\gamma(x)^{\gamma(x)} \frac{A^x}{x^x} \right] = F(A)$$

and the theorem is proved.

The last result suggests a strong connection between spectral properties of A and entropy. That such indeed is the case is known from some results in [1], [3], where a special case of the above result for A a zero-one matrix is proved. Theorem 2 also leads to a variety of inequalities for $\lambda(A)$, of which we give one example.

Note that

$$\frac{\gamma(x)^{\gamma(x)}}{x^x} = \prod_{i,j} \left[\frac{\gamma_i(x)}{x_{ij}} \right]^{x_{ij}} \leqslant \sum_{i,j} x_{ij} \frac{\gamma_i(x)}{x_{ij}} = k.$$

Hence $\lambda(A) \leqslant k \max_{x \in V} A^x$, and it is clear that A^x takes its max at an extreme point of V.

Now we describe the extreme points of V. Let P be the collection of permutations on k numbers whose cycle decomposition consists of a single cycle of length $r > 1$ for some r, and $k - r$ one cycles. To each $\pi \in P$ associate the $k \times k$ matrix π' which consists of $1/r$ times the permutation matrix for π with the diagonal entries all suppressed. It is easy to see that the collection Σ' of matrices so obtained are extreme points of V. Adjoin to Σ', to obtain Σ, all diagonal matrices having a single one on the diagonal, otherwise zero. Then I claim Σ is the set of extreme points of V. For let x be any element of V. In attempting to write x as a convex combination of points of Σ, I may suppose all diagonal entries of x are already zero. Pick an off diagonal nonzero element, $x_{j_1 j_2}$. Then there must be a nonzero off diagonal element in j_2th row, say $x_{j_2 j_3}$, then nonzero off diagonal element in j_3th row, say $x_{j_3 j_4}$. Continue in this way till one returns at rth stage to a row already visited, which we may suppose, without loss of generality, to be the j_1th. Now we have an r ($r > 1$) cycle $(j_1, j_2, \dots, j_r)$, associated permutation matrix π, associated extreme point π', and taking away from x a suitable multiple of π', we can increase the number of zeros in x and conclude the proof by induction.

Finally, since $\gamma(x)^{\gamma(x)}/x^x$ is always one for the extreme points of V, we have

COROLLARY 3.

$$\max_{\sigma \in \Sigma} A^\sigma \leqslant \lambda(A) \leqslant k \max_{\sigma \in \Sigma} A^\sigma.$$

(The elements of Σ may loosely be described as the "weighted loops on k letters".)

Now we make use of the following theorem attributed to Minkowski.

THEOREM 4. *Let A be a $k \times k$ matrix whose off diagonal elements are nonpositive. The following are equivalent*:

(i) $\exists$ *vector $h > 0$ such that $Ah > 0$.*

(ii) *All eigenvalues of A have positive real part.*

(iii) A^{-1} *exists and $A^{-1} \geqslant 0$.*

The set of matrices satisfying the hypothesis and any one of the conclusions of the last theorem will be denoted by $\mathfrak{M}$. If $A \geqslant 0$, and D is a diagonal matrix such that $D - A \in \mathfrak{M}$, we shall simply write $D > A$.

Since we cannot find a convenient proof of the theorem in the literature, we now sketch a simple demonstration.

We begin by assuming (i) holds. Write A in the form $A = \rho I - C$, $C \geqslant 0$, $\rho > 0$. From an observation of Wielandt, we know that for any $h > 0$,

$$\lambda(C) \leqslant \max_i \frac{(Ch)i}{h_i}.$$

Hence $\lambda(C) < \rho$. But then it is clear that (ii) holds.

If (ii) holds, then clearly $\lambda(C) < \rho$, and it is known [2], that

$$A^{-1} = (\rho I - C)^{-1}$$

exists and is nonnegative, so that (iii) is true.

While if (iii) holds, A^{-1} must have a positive entry in every row. So put $h = A^{-1}k$ for any $k > 0$, to obtain (i).

LEMMA 5. *Let*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

be nonnegative $k \times k$ matrix, where A_{11} and A_{22} are $r \times r$ and $s \times s$ respectively, $r + s = k$. Let

$$D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}$$

be a similarly grouped diagonal matrix. Then the necessary and sufficient conditions that $D > A$ are:

(i) $D_1 > A_{11}$,

(ii) $D_2 > A_{22} + A_{21}(D_1 - A_{11})^{-1}A_{12}$.

As for the necessity, suppose $D > A$. Then $\exists$ positive column vector $\begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$, h_1, r long, h_2 s long, such that

$$\begin{bmatrix} D_1 - A_{11} & -A_{12} \\ -A_{21} & D_2 - A_{22} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} (D_1 - A_{11})h_1 - A_{12}h_2 \\ -A_{21}h_1 + (D_2 - A_{22})h_2 \end{bmatrix} > 0,$$

implying $(D_1 - A_{11})h_1 > A_{12}h_2 \geqslant 0$, implying $D_1 > A_{11}$. Analogously $D_2 > A_{22}$. Now $(D_1 - A_{11})h_1 > A_{12}h_2$, which implies $h_1 > (D_1 - A_{11})^{-1}A_{12}h_2$. Also $(D_2 - A_{22})h_2 > A_{21}h_1$, so $(D_2 - A_{22})h_2 > A_{21}(D_1 - A_{11})^{-1}A_{12}h_2$, which implies $D_2 > A_{22} + A_{21}(D_1 - A_{11})^{-1}A_{12}$.

For the sufficiency, suppose (i) and (ii) are satisfied. There exists $h_2 > 0$ so that

$$D_2 h_2 > A_{22}h_2 + A_{21}(D_1 - A_{11})^{-1}A_{12}h_2.$$

Pick $\delta > 0$ so that $(D_1 - A_{11})\delta > 0$, $D_2 h_2 > A_{22}h_2 + A_{21}(D_1 - A_{11})^{-1}A_{12}h_2 + A_{21}\delta$, and put $h_1 = (D_1 - A_{11})^{-1}A_{12}h_2 + \delta$. Now it is immediate that

$$\begin{bmatrix} D_1 - A_{11} & -A_{12} \\ -A_{21} & D_2 - A_{22} \end{bmatrix}\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} > 0.$$

COROLLARY 6. *The following are equivalent*:
(i) $D_1 > A_{11}$ *and* $D_2 > A_{22} + A_{21}(D_1 - A_{11})^{-1}A_{12}$.
(ii) $D_2 > A_{22}$ *and* $D_1 > A_{11} + A_{12}(D_2 - A_{22})^{-1}A_{21}$.

Either (i) or (ii) is equivalent to

$$\begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix} > \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

the second after similar renumbering of the rows and columns of A.

COROLLARY 7. *Let* a, b *be nonnegative column vectors,* ρ *a positive scalar, and* $M \in \mathfrak{M}$. *Then* $b'M^{-1}a < \rho$ *if and only if* $M - ab'/\rho \in \mathfrak{M}$.

This corollary follows readily from the last on letting $s = 1$, $A_{22} = 0$, $A_{12} = a$, and $A_{21} = b'$.

THEOREM 8. *The set* $L(A) = \{D, D \text{ diagonal} | D > A\}$ *is convex.*

For the proof we first introduce a little notation: If ρ is a scalar, by $A^{\underline{\rho}}$ we denote the matrix whose i, jth entry is A_{ij}^{ρ}.

Now let $D_1 > A$, $D_2 > B$. I claim that $D_1^{\underline{p}}D_2^{\underline{q}} > A^{\underline{p}} * B^{\underline{q}}$ for any $p, q \geqslant 0$, $p + q = 1$. For let h and k be positive vectors such that $(D_1 - A)h > 0$ and $(D_2 - B)k > 0$. Then:

$$\sum_j A_{ij}^p B_{ij}^q h_j^p k_j^q \leqslant \left(\sum_j A_{ij}h_j\right)^p \left(\sum_j B_{ij}b_j\right)^q \leqslant (D_1 h)_i^p (D_2 h)_i^q;$$

and the claim is established.

Since $D_1^{\underline{p}}D_2^{\underline{q}} \leqslant pD_1 + qD_2$, the theorem follows immediately from the special case $A = B$.

COROLLARY 9. *Let* a *and* b *be nonnegative. Then* $\log(b'(D - A)^{-1}a)$ *is a convex function of* D *in* $L(A)$.

Consider the matrix $B = \begin{bmatrix} A & a \\ b' & 0 \end{bmatrix}$. If $D_1 > A$ and $D_2 > A$, then

$$\begin{bmatrix} D_1 & 0 \\ 0 & \lambda_1 \end{bmatrix} > B \quad \text{and} \quad \begin{bmatrix} D_2 & 0 \\ 0 & \lambda_2 \end{bmatrix} > B$$

if and only if $\lambda_1 > b'(D_1 - A)^{-1}a$, and $\lambda_2 > b'(D_2 - A)^{-1}a$. From the proof of the last theorem, it now follows that $\lambda_1^p \lambda_2^q > b'(D_1^p D_2^q - A)^{-1}a$.

Let x be one of the diagonal entries of D. Then with an obvious notation

$$\partial (D - A)^{-1}/\partial x = -(D - A)^{-1} \partial D (D - A)^{-1}/\partial x,$$

so that the entries of $(D - A)^{-1}$ are monotone nonincreasing functions of the entries of D. Hence $b'(D_1^p D_2^q - A)^{-1}a \geqslant b'(pD_1 + qD_2 - A)^{-1}a$, and the corollary follows readily.

THEOREM 10.

$$E_\alpha(A) = \min_{D \in L(\alpha^{-1}A)} D^{\alpha/d} = \frac{1}{\alpha^{\alpha/d}} \min_{D \in L(A)} D^{\alpha/d}.$$

Let I be the $k \times k$ identity matrix. By Perron-Frobenius theory $\lambda(A) = \min_{\rho > 0} \rho$ subject to the contraint $\rho I > A$. Hence $\lambda(\alpha^{-1}AD) = \min\{\rho | \rho I > \alpha^{-1}AD\}$. Now using Theorem 1, we obtain

$$E_\alpha(A) = \min_D \left\{ \lambda(\alpha^{-1}AD) | D^{\alpha/d} = 1 \right\} = \min_{\rho, D} \left\{ \rho | D^{\alpha/d} = 1, \rho I > \alpha^{-1}AD \right\}$$

$$= \min_D \left\{ D^{\alpha/d} | D > \alpha^{-1}A \right\} = \min_{D \in L(\alpha^{-1}A)} D^{\alpha/d}.$$

Since $E_\alpha(\alpha A) = \alpha^{\alpha/d} E_\alpha(A)$, the second equality is clear.

COROLLARY 11.

$$E_\alpha(A) = \frac{1}{\alpha^{\alpha/d}} \min_{h > 0} \prod_i \left[\frac{\sum_j A_{ij} h_j}{h_i} \right]^{\alpha_i/d}.$$

For if $D > A$, there exists $h > 0$ so that $Dh > Ah$; i.e., $D_i > \sum_j A_{ij} h_j / h_i$, and the result is obvious.

(An alternative proof to the above may also be obtained by taking the constrained bilinear form whose minimum is $E_\alpha(A)$, and minimizing it with respect to one of the vector variables, holding the other one fixed.)

For the remainder of the paper, we will concentrate on the function $E_I(A)$, which we will denote henceforth by $E(A)$, independently of the dimension k of A, and I will denote the generic identity matrix.

Let A be of dimension k and write

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} is $r \times r$, and A_{22} is $s \times s$, $r + s = k$.

LEMMA 12. $E^k(A) = \min_{D > A_{11}} D^I E^s(A_{22} + A_{21}(D - A_{11})^{-1}A_{12})$.

For according to Lemma 5,

$$\begin{bmatrix} D & 0 \\ 0 & D' \end{bmatrix} > A$$

if and only if $D > A_{11}$ and $D' > A_{22} + A_{21}(D - A_{11})^{-1}A_{12}$. The result is now obvious.

Let A be of dimension k and write

$$A = \begin{bmatrix} B & a \\ b' & c \end{bmatrix},$$

where B is of dimension $k - 1$. Then

COROLLARY 13.

$$E^k(A) = \min_{D > B} D^I\left(c + b'(D - B)^{-1}a\right) = \min_{\rho > 0}(c + \rho)E^{k-1}\left(B + \frac{1}{\rho}ab'\right).$$

Both results are simply special cases of Lemma 12.

The second equality above gives a convenient inductive defintion of $E(A)$. Thus if

$$B = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix}$$

is 2×2, we find easily that $E(B) = \sqrt{b_1 b_4} + \sqrt{b_2 b_3}$. If

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is 3×3, we thus obtain

$$E^3(A) = \min_{\rho > 0}(i + \rho)\left\{ \sqrt{\left(a + \frac{1}{\rho}cg\right)\left(e + \frac{1}{\rho}fh\right)} \right.$$
$$\left. + \sqrt{\left(b + \frac{1}{\rho}ch\right)\left(d + \frac{1}{\rho}fg\right)} \right\}^2,$$

which is, unfortunately, not a very manageable minimization problem.

The last results suggest an inductive proof of the permanent conjecture along the following lines. Let m_k be the minimum value of the permanent for $k \times k$ doubly stochastic matrices. We have proved in [5, Theorem 24, p. 90], that for every $A > 0$, Perm $A \geqslant m_k E^k(A)$. Write

$$A = \begin{bmatrix} B & a \\ b' & c \end{bmatrix}$$

with B of dimension $k - 1$. Then from our last result

$$E^k(A) \leqslant (c + \rho)E^{k-1}(B + ab'/\rho), \quad \text{for any } \rho > 0.$$

If A is doubly stochastic $E^k(A) = 1$, so $(c + \rho)\text{Perm}(B + ab'/\rho) \geqslant m_{k-1}$. The interesting fact to observe is that among the terms occurring in the expansion of $(c + \rho)\text{Perm}(B + ab'/\rho)$, one finds the expansion for $\text{Perm}(A)$, but we have not been able to treat the remaining terms in a systematic enough fashion to get an inductive proof.

(It is probably worth noting that $B + ab'/(1 - c)$ is doubly stochastic $(k - 1) \times (k - 1)$ matrix.)

Here, however, is a proof for $k = 3$, assuming the case $k = 2$ settled. Let

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

be a 3×3 doubly stochastic matrix for which the permanent assumes its minimum. W.l.o.g., we will further assume that i is the largest element in its row; so, in particular, $i \geqslant \frac{1}{3}$. The assumptions A has minimum permanent and $i \neq 0$ guarantee that $ae + bd = \text{Perm}(A)$, [5, Theorem 9, p. 81]. Next,

$$H = \begin{pmatrix} a & b \\ d & e \end{pmatrix} + \frac{1}{1 - i}\begin{pmatrix} eg & ch \\ fg & fh \end{pmatrix}$$

is doubly stochastic, so $\text{Perm } H \geqslant \frac{1}{2}$. One finds readily, however, that

$$\text{Perm } H = (ae + bd) + \frac{1}{1 - i}(afh + ecg + bfg + dch) + \frac{2}{(1 - i)^2} cfgh$$

$$= \text{Perm } A + \frac{1}{1 - i}\{\text{Perm } A - i(ab + de)\} + \frac{2}{(1 - i)^2} cfgh$$

$$= 2\,\text{Perm } A + \frac{2}{(1 - i)^2} cfgh.$$

Since $c + f = 1 - i$, $cf \leqslant (1 - i)^2/4$, analogously for gh. Hence $\text{Perm } H \leqslant 2\text{Perm } A + \frac{1}{8}(1 - i)^2$. Since $i \geqslant \frac{1}{3}$, $\text{Perm } H \leqslant 2\text{Perm } A + \frac{1}{18}$; since $\text{Perm } H \geqslant \frac{1}{2}$, $\text{Perm } A \geqslant \frac{2}{9}$.

There are some simple relationships between the order relations we have introduced above, and the order coming from positive semidefiniteness. Let A and B be symmetric square matrices of the same dimension. We write $A \rhd B$ if $A - B$ is positive definite, $A \unrhd B$ if $A - B$ is positive semidefinite.

LEMMA 14. *Let A and B be nonnegative symmetric matrices, such that $A \unrhd$*

B. Then $L(A) \subset L(B)$. Moreover, $E(A) \geqslant E(B)$.

The requirement $D > A$ is equivalent to $D \rhd A$. But if $D \rhd A$, then $D - B = D - A + A - B \rhd 0$, so $D > B$.

THEOREM 15. *Let X be symmetric, $X \geqslant 0$, and $X \trianglerighteq 0$. Write*

$$X = \begin{bmatrix} A & B \\ B' & C \end{bmatrix}$$

with A $r \times r$, and C $s \times s$, $r + s = k$. Then

$$E^r(A)E^s(C) \geqslant r^r s^s E^k(X)/k^k.$$

We know that $V^I E^r(A + B(V - C)^{-1}B') \geqslant E^k(X)$, for any $V > C$. Pick a real number $p > 1$, and select $V > pC$. Hence $V \rhd pC$, so $V - C \rhd (p - 1)C$. We will now suppose X is positive definite, so that C is also positive definite. The general case can be established by passage to the limit.

Since $V - C \rhd (p - 1)C$, it follows that $(V - C)^{-1} \lhd C^{-1}/(p - 1)$, and hence $B(V - C)^{-1}B' \trianglelefteq BC^{-1}B'/(p - 1)$. Also

$$\begin{bmatrix} I, & -BC^{-1} \end{bmatrix} \begin{bmatrix} A & B \\ B' & C \end{bmatrix} \begin{bmatrix} I \\ -C^{-1} & B' \end{bmatrix} = A - BC^{-1}B' \rhd 0.$$

Thus $A + B(V - C)^{-1}B' \lhd pA/(p - 1)$. Hence

$$p^r V^I E^r(A)/(p - 1)^r \geqslant E^k(X).$$

Since $V > pC$, we may by Theorem 10 select V (in limit if necessary) so that $V^I = E^s(pC)$. Thus

$$E^r(A)E^s(C) \geqslant (p - 1)^r E^k(X)/p^k.$$

The most favorable choice of p is $p = k/s$, which gives stated theorem.

Note that equality holds if X is the matrix with all entries one.

Our next and last result may simply be a curiosity, but we believe it indicates a direction worth further study. To explain it, we must recall some further facts from [5, p. 82]. Let x be a $k \times k$ nonnegative matrix, and $P(x)$ its permanent. We have shown that the map h: $h(x) = (1/P(x))\Sigma_\pi x^\pi \pi$, π ranging over permutation matrices, is a bijection of the space of doubly stochastic matrices, with inverse map denoted l. If we agree to call the numbers x^π the transversals of x, then the above implies that a doubly stochastic matrix is a single valued algebraic function of its transversals, since h is a rational map. What we are going to show is that in fact a doubly stochastic matrix is a rational function of its transversals, except possibly on a not identically vanishing polynomial locus.

Let again x be an arbitrary positive matrix. We have also shown in [5, p. 85] that x may be written uniquely $x_{ij} = \delta \lambda_i A_{ij} u_j$, where $\delta = E(x)$, A is doubly stochastic, and $\Pi_i \lambda_i = \Pi_i u_i = 1$. Moreover $\partial E(x)/\partial x_{ij} = 1/k\lambda_i u_j$.

LEMMA 16. *Let $x > 0$. Then $E(x)$ satisfies an irreducible polynomial equation whose coefficients are polynomials in the transversals of x.*

For the proof, let $A > 0$ be doubly stochastic. Clearly, $P(l(A))$ is algebraic over the field generated by the entries of A. Hence, there exist polynomials $Q_v(A)$ in the entries of A such that $\sum_{v=0}^{r} Q_v(A) P^v(l(A)) = 0$. Replace A by $h(A)$ to obtain $\sum_{v=0}^{r} Q_v(h(A)) P^v(A) = 0$. If now x is an arbitrary positive matrix, write $x_{ij} = \delta \lambda_i A_{ij} u_j$, with $\delta = E(x)$. Since $h(A) = h(x)$, we get

$$\sum_{v=0}^{r} Q_v(h(x)) \frac{P^v(x)}{E^{nv}(x)} = 0.$$

Since $Q_v(h(x))$ is a rational function of the transversals of x, the result follows.

THEOREM 17. *Doubly stochastic $A > 0$ is a rational function of its transversals, except possibly on a not identically vanishing polynomial locus.*

Let x be arbitrary positive, and suppose that $\sum_{v=0}^{r} Q_v(x) y^v$ is the irreducible polynomial satisfied by $y = E(x)$, with coefficients polynomials in the transversals of x. On differentiating, we obtain

$$x_{ij} \frac{\partial E(x)}{\partial x_{ij}} = - \frac{\sum_v x_{ij} E^v(x) \partial Q_v(x) / \partial x_{ij}}{\sum_v v Q_v(x) E^{v-1}(x)} = \frac{\sum_v H_v^{ij}(x) E^v(x)}{\sum_v v Q_v(x) E^{v-1}(x)}.$$

The denominator above is not identically zero because we started with an irreducible equation, and it is easy to see that it does not vanish identically on the space of doubly stochastic matrices. For consideration of homogeneity shows that the $Q_v(x)$ are all homogeneous polynomials, and if the denominator vanished for all doubly stochastic matrices, it would vanish identically. It is also obvious that $H_v^{ij}(x) = - x_{ij} \partial Q_v(x) / \partial x_{ij}$ is a polynomial in the transversals of x. Hence, letting $x = A$ a doubly stochastic matrix, we obtain, since $E(A) = 1$,

$$A_{ij} = \frac{\sum_v H_v^{ij}(A)}{\sum_v v Q_v(A)}.$$

We have not been able to find the irreducible polynomial equation for $E(x)$, even for 3×3 matrices, but is clear that it contains a great deal of combinatorial information.

REFERENCES

1. R. Adler and B. Weiss, *Similarity of automorphisms of the torus*, Mem. Amer. Math. Soc. No. 98 (1970). MR **41** #1966.

2. F. R. Gantmaher, *The theory of matrices*, Vol. I, GITTL, Moscow, 1953; English transl., Chelsea, New York, 1959. MR **21** #6372c.

3. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112** (1964), 55–66. MR **28** #4579.

4. G. Letac, *A unified treatment of some theorems on positive matrices*, Proc. Amer. Math. Soc. **43** (1974), 11–17. MR **49** #2804.

5. O. Rothaus, *Study of the permanent conjecture and some of its generalizations*, Israel J. Math. **18** (1974), 75–96. MR **51** #8137.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853