Inequalities and Partial Orders on Matrix Spaces

COLIN J. THOMPSON

Communicated by M. Kac

1. Introduction. In a recent article, Lenard [1] showed that for a large class of real valued functions φ defined on matrices,

$$(1) \quad \varphi(e^A e^B) \geq \varphi(e^{A+B})$$

for Hermitian A and B . This generalized a previous inequality, $\text{Tr}(e^A e^B) \geq \text{Tr}(e^{A+B})$, due to Golden [2] and the author [3].

The class of functions φ constructed by Lenard are monotonic functions with respect to the order $\succ_1$, defined on positive $n \times n$ matrices (*i.e.*, Hermitian with non-negative eigenvalues) X and Y by

$$(2) \quad X \succ_1 Y$$

if there exist unitary matrices U and V such that

$$(3) \quad \text{Spec}(X^{1/2} U X^{1/2} V) = \text{Spec}(Y).$$

$X^{1/2}$ denotes the unique positive square root of X and $\text{Spec}(A)$ denotes the collection of eigenvalues of A , each counted with its proper multiplicity.

Lenard was only able to show that functions of the form $\varphi(A) = |\text{Tr } R(A)|$, where $R(A)$ is a matrix representation of A , are monotonic with respect to $\succ_1$. This, however, led to the interesting result

$$(4) \quad I_s(e^A e^B) \geq I_s(e^{A+B}) \quad s = 1, 2, \dots, n$$

where $I_s(X)$ denotes the s^{th} coefficient in the characteristic polynomial of X and A and B are $n \times n$ Hermitian matrices. Moreover, equality in (4) holds when $s = n(I_n(X) = \text{Det}(X))$.

Lenard did not prove that $\succ_1$ is a partial order, although he did show how it could be trivially extended to one. He also posed the problem of finding necessary and sufficient conditions on $\text{Spec}(Y)$ and $\text{Spec}(X)$ for $X \succ_1 Y$. He was able to solve this problem for 2×2 matrices but not for higher dimensional matrices.

Our first purpose here is to state necessary and sufficient conditions for Lenard's order. As a consequence, it will follow that $\succ_1$ is, in fact, a partial order. It turns out from Theorem 1, stated and proved in the following section, that the order $\succ_1$ is unnecessarily complicated and restrictive: V in (3) may as well be replaced by the identity and one can extend Lenard's concept by defining (below) an order for arbitrary complex $n \times n$ matrices.

In section three, we show that unitarily invariant matrix norms are monotonic with respect to the order $\succ_1$ and in the final section we give a simple direct proof of (4) and some further generalizations. In the appendix, we show that the most obvious generalizations of (1) involving three matrices are in general false.

We will assume throughout that whenever the eigenvalues $a_1, a_2, \dots, a_n$ of an $n \times n$ matrix are real, they are arranged in decreasing order: $a_1 \geq a_2 \geq \dots \geq a_n$. Also, the Hermitian adjoint of a matrix A will be denoted by $A^\dagger$. Finally, we remark that all our results carry over directly to completely continuous operators in Hilbert space.

2. Necessary and sufficient conditions for two partial orders. We begin by defining an order for arbitrary complex $n \times n$ matrices which is slightly different and more general than Lenard's.

Definition. Let X and Y be arbitrary complex $n \times n$ matrices and let $\tilde{X}$ and $\tilde{Y}$ be the unique positive square roots of $XX^\dagger$ and $YY^\dagger$ respectively. We write

$$(5) \quad X \succ Y$$

if there exists a unitary matrix U such that

$$(6) \quad \text{Spec}(\tilde{X}U) = \text{Spec}(\tilde{Y}).$$

It is to be noted that when X and Y are positive (6) reduces to (3) with V the identity (since $\text{Spec}(X^{1/2}UX^{1/2}) = \text{Spec}(XU)$). In other words, for positive X and Y , $X \succ Y$ implies $X \succ_1 Y$.

The main result in this section is

Theorem 1. If X and Y are complex $n \times n$ matrices and $x_1 \geq x_2 \geq \dots \geq x_n \geq 0$ and $y_1 \geq y_2 \geq \dots \geq y_n \geq 0$ are the eigenvalues of $\tilde{X} = (XX^\dagger)^{1/2}$ and $\tilde{Y} = (YY^\dagger)^{1/2}$ respectively, then there exist unitary matrices U and V such that

$$(7) \quad \text{Spec}(\tilde{X}^{1/2}U\tilde{X}^{1/2}V) = \text{Spec}(\tilde{Y})$$

if and only if

$$x_1x_2 \cdots x_k \geq y_1y_2 \cdots y_k \quad k = 1, 2, \dots, n-1$$

$$(8) \quad \text{and}$$

$$x_1x_2 \cdots x_n = y_1y_2 \cdots y_n.$$

Moreover, the same conclusion holds if (7) is replaced by (6).

As an immediate consequence of the theorem, we deduce that the orders $\succ_1$ and $\succ$ are partial orders, the former for positive matrices and the latter for arbitrary complex matrices. (Equality in both cases means the unitary equivalence of $\tilde{X}$ and $\tilde{Y}$, or in general, since an arbitrary matrix X can be written in the form $X = \tilde{X}U_x$ with U_x unitary, that there exist unitary matrices U and V such that $X = UYV$.)

We remark in passing that condition (8) in the theorem may be reexpressed in terms of the maximum eigenvalues $x^{(k)}$ and $y^{(k)}$ of the skew symmetric tensor products of $\tilde{X}$ taken with itself k times and $\tilde{Y}$ taken with itself k times respectively. This follows from the well-known results $x^{(k)} = x_1x_2 \cdots x_k$ and $y^{(k)} = y_1y_2 \cdots y_k$ (see, for example Ref. 4).

We now prove the theorem for the case of strictly positive $\tilde{X}$ and $\tilde{Y}$ (i.e., $x_n > 0$ and $y_n > 0$). The positive case requires only trivial modifications and is left to the reader.

To prove the necessary part of the theorem we need two lemmas, due to Weyl [4] and Horn [5] respectively.

Lemma 1. *If $y_1, y_2, \dots, y_n$ are the eigenvalues of an $n \times n$ matrix Y and $z_1, z_2, \dots, z_n$ are the eigenvalues of $(YY^\dagger)^{1/2}$ with $|y_1|, |y_2|, \dots, |y_n|$ and $z_1, z_2, \dots, z_n$ arranged in decreasing order, then*

$$(9) \quad z_1z_2 \cdots z_k \geq |y_1| |y_2| \cdots |y_k|, \quad k = 1, 2, \dots, n.$$

Lemma 2. *If $a_i, b_i, c_i, i = 1, 2, \dots, n$ are the eigenvalues of $AA^\dagger, BB^\dagger$, and $(AB)(AB)^\dagger$ respectively arranged in decreasing order, then*

$$(10) \quad a_1a_2 \cdots a_k b_1b_2 \cdots b_k \geq c_1c_2 \cdots c_k \quad k = 1, 2, \dots, n.$$

Choosing $A = \tilde{X}^{1/2}U$ and $B = \tilde{X}^{1/2}V$ in Lemma 2 and noting that $AA^\dagger = BB^\dagger = \tilde{X}$, we obtain

$$(11) \quad x_1x_2 \cdots x_k \geq z_1z_2 \cdots z_k \quad k = 1, 2, \dots, n$$

where $z_1, z_2, \dots, z_n$ are the eigenvalues of

$$[(\tilde{X}^{1/2}U\tilde{X}^{1/2}V)(\tilde{X}^{1/2}U\tilde{X}^{1/2}V)^\dagger]^{1/2}.$$

Now if (7) is assumed, then $\text{Det}(\tilde{X}) = \text{Det}(\tilde{Y})$ and (8) follows from (11) and Lemma 1. If, on the other hand (6) is assumed, (8) follows immediately from Lemma 1 by taking $Y = \tilde{X}U$.

To prove the sufficient part of the theorem, we need the following result.

Lemma 3. *If*

$$X = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}$$

and $x_1 \geq \alpha_1 x_1 \geq x_2 > 0$, then there exist unitary matrices U and V such that

$$(12) \quad V X U = \begin{pmatrix} \alpha_1 x_1 & 0 \\ z_1 & \alpha_1^{-1} x_2 \end{pmatrix}.$$

Obviously $\text{Spec } (V X U) = \{\alpha_1 x_1, \alpha_1^{-1} x_2\}$.

Proof. Define

$$(13) \quad U = \begin{pmatrix} c' & s' \\ -s' & c' \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

where $c' = \cos \theta'$, $s' = \sin \theta'$, $c = \cos \theta$, and $s = \sin \theta$. An elementary calculation shows that (12) is satisfied for the choices

$$(14) \quad \begin{aligned} c' &= (\alpha_1^2 x_1^2 - x_2^2)^{1/2} & (\alpha_1^2 x_1^2 - \alpha_1^2 x_2^2)^{-1/2} \\ c &= \alpha_1 c' \end{aligned}$$

with angles θ and θ' chosen such that $s \geq 0$ and $s' \geq 0$. For the above choices, z_1 in (12) is given by

$$(15) \quad z_1 = s c' x_1 - s' c x_2.$$

We are now in a position to complete the proof of the theorem by induction.

Since it is clearly no restriction on generality to take $\tilde{X}$ diagonal, we will do so henceforth.

Consider first the case $n = 2$, i.e., from (8)

$$(16) \quad \begin{aligned} x_1 &\geq y_1 \\ x_1 x_2 &= y_1 y_2. \end{aligned}$$

It follows that there exists $0 < \alpha_1 \leq 1$ such that $y_1 = \alpha_1 x_1$ and $y_2 = \alpha_1^{-1} x_2$. Moreover, from (16), $y_1 = \alpha_1 x_1 \geq y_2 \geq x_2$, so the conditions of Lemma 3 are satisfied. From Lemma 3, there then exist unitary matrices U' and V' such that

$$(17) \quad \text{Spec } (\tilde{Y}) = \text{Spec } (V' \tilde{X} U') = \text{Spec } (\tilde{X} U' V').$$

Since $U = U' V'$ is unitary, (6) is satisfied and also (7) is satisfied with V the identity.

Let us now assume that for $(n-1) \times (n-1)$ matrices X and Y satisfying (8), there exists a unitary matrix U_{n-1} such that $\text{Spec } (\tilde{Y}) = \text{Spec } (\tilde{X} U_{n-1})$.

For n -dimensional matrices, it follows from (8) that $y_n \geq x_n$ and hence that there exists a $2 \leq k \leq n$ such that

$$(18) \quad x_{k-1} \geq y_1 \geq x_k.$$

There then exists an $0 < \alpha_1 \leq 1$ such that $y_1 = \alpha_1 x_1 \geq x_k$. With $\tilde{X}$ diagonal, we can choose from Lemma 3 unitary matrices

$$(19) \quad U_k = \begin{bmatrix} c' & 0 & \cdots & 0 & s' & 0 & \cdots & 0 \\ 0 & 1 & & & & & & \\ \vdots & & \ddots & & & & & \\ 0 & & & 1 & & & & \\ -s' & & & & c' & & & \\ 0 & & & & & 1 & & \\ \vdots & & & & & & \ddots & \\ 0 & & & & & & & 1 \end{bmatrix}$$

and

$$(20) \quad V_k = \begin{bmatrix} c & 0 & \cdots & 0 & -s & 0 & \cdots & 0 \\ 0 & 1 & & & & & & \\ \vdots & & \ddots & & & & & \\ 0 & & & 1 & & & & \\ s & & & & c & & & \\ 0 & & & & & 1 & & \\ \vdots & & & & & & \ddots & \\ 0 & & & & & & & 1 \end{bmatrix}$$

with non-zero off-diagonal entries ($\pm s'$ and $\pm s$) in the k^{th} rows and k^{th} columns, such that

$$(21) \quad V_k \tilde{X} U_k = \left[\begin{array}{c|ccc} \alpha_1 x_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ \vdots & & & \\ z_1 & & \tilde{X}' & \\ \vdots & & & \\ \vdots & & & \\ 0 & & & \end{array} \right]$$

where z_1 is in the k^{th} row and

$$(22) \quad \tilde{X}' = \begin{bmatrix} x_2 & & & & & \\ & x_3 & & & 0 & \\ & & \ddots & & & \\ & & & \alpha_1^{-1} x_k & & \\ 0 & & & & \ddots & \\ & & & & & x_n \end{bmatrix}$$

is a diagonal $(n - 1) \times (n - 1)$ matrix. Now since $x_1 \geq x_2 \geq \cdots \geq x_n$ and $y_1 \geq y_2 \geq \cdots \geq y_n$ it follows from (18) that $x_i \geq y_i$, $i = 2, 3, \cdots, k - 1$ and hence that

$$(23) \quad x_2 x_3 \cdots x_\ell \geq y_2 y_3 \cdots y_\ell \quad \ell = 2, 3, \cdots, k - 1.$$

Also, since $y_1 = \alpha_1 x_1$ it follows from (8) that

$$x_2 x_3 \cdots \alpha_1^{-1} x_k \cdots x_\ell \geq y_2 y_3 \cdots y_\ell \quad \ell = k, \cdots, n - 1$$

and

$$(24) \quad x_2 x_3 \cdots \alpha_1^{-1} x_k \cdots x_n = y_2 y_3 \cdots y_n.$$

Defining

$$(25) \quad x'_i = \begin{cases} x_i, & i \neq k \\ \alpha_1^{-1} x_k, & i = k, \end{cases} \quad i = 2, 3, \cdots, n$$

(23) and (24) imply that

$$(26) \quad x'_2 x'_3 \cdots x'_\ell \geq y_2 y_3 \cdots y_\ell \quad \ell = 2, 3, \cdots, n - 1$$

and

$$x'_2 x'_3 \cdots x'_n = y_2 y_3 \cdots y_n,$$

i.e., that the eigenvalues of the matrix $\tilde{X}'$, Eq. (22), satisfy condition (8). By our inductive assumption, there then exists a unitary U_{n-1} such that

$$(27) \quad \text{Spec}(\tilde{X}' U_{n-1}) = \{y_2, y_3, \cdots, y_n\}$$

and hence from (21)

$$(28) \quad \text{Spec}(\tilde{Y}) = \text{Spec}(V_k \tilde{X} U_k U_n) = \text{Spec}(\tilde{X} U)$$

where $U = U_k U_n V_k$ and

$$(29) \quad U_n = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ \vdots & & & \\ 0 & & & \end{array} \begin{array}{c} \\ \\ U_{n-1} \\ \\ \end{array} \right).$$

The proof of the theorem is then complete.

3. Gauge functions and unitarily invariant norms. In general, a real valued function $\|A\|$ on the space of complex $n \times n$ matrices is a *norm* if

- (a) $\|A\| \geq 0$ and $= 0$ if and only if $A = 0$,
- (b) $\|cA\| = |c| \|A\|$ for any complex number c ,
- (c) $\|A + B\| \leq \|A\| + \|B\|$.

If, in addition,

(d) $\|AU\| = \|A\|$ for all unitary U , we say that the norm is *unitarily invariant*.

Von Neumann [6] proved some time ago that any unitarily invariant norm can be constructed from *symmetric gauge functions*, such functions $\Phi(\alpha_1, \alpha_2, \dots, \alpha_n)$ being real valued and defined on the space of n -tuples $(\alpha_1, \alpha_2, \dots, \alpha_n)$ of real numbers with the properties

- (a) $\Phi(\alpha_1, \alpha_2, \dots, \alpha_n) \geq 0$ and $= 0$ if and only if $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$,
- (b) $\Phi(c\alpha_1, c\alpha_2, \dots, c\alpha_n) = |c| \Phi(\alpha_1, \alpha_2, \dots, \alpha_n)$ for any constant c ,
- (c) $\Phi(\alpha_1 + \alpha'_1, \alpha_2 + \alpha'_2, \dots, \alpha_n + \alpha'_n)$
 $\leq \Phi(\alpha_1, \alpha_2, \dots, \alpha_n) + \Phi(\alpha'_1, \alpha'_2, \dots, \alpha'_n)$,
- (d) $\Phi(\alpha_1, \alpha_2, \dots, \alpha_n) = \Phi(\epsilon_1\alpha_{1'}, \epsilon_2\alpha_{2'}, \dots, \epsilon_n\alpha_{n'})$

where $\epsilon_i = \pm 1$ and $1', 2', \dots, n'$ is any permutation on $1, 2, \dots, n$.

Some simple examples of symmetric gauge functions are

$$(i) \quad \Phi_p(\alpha_1, \alpha_2, \dots, \alpha_n) = \left(\sum_{i=1}^n |\alpha_i|^p \right)^{1/p}, \quad 1 \leq p < \infty$$

and

$$(ii) \quad \Phi_\infty(\alpha_1, \alpha_2, \dots, \alpha_n) = \max_{1 \leq i \leq n} |\alpha_i|.$$

Obviously, if $\alpha_1, \alpha_2, \dots, \alpha_n$ are the eigenvalues of $(AA^\dagger)^{1/2}$, then for any symmetric gauge function Φ

$$(30) \quad \|A\|_\Phi = \Phi(\alpha_1, \alpha_2, \dots, \alpha_n)$$

is unitarily invariant. What von Neumann [6] was able to show was that (30) is a matrix norm and moreover that any unitarily invariant norm can be expressed in this form.

In terms of the partial order $\succ$ defined by (5) and (6), we have the following results.

Theorem 2. *If for arbitrary complex $n \times n$ matrices A and B , $A \succ B$, then $\|A\| \geq \|B\|$ for any unitarily invariant norm $\|\cdot\|$.*

We merely indicate that this result follows from Theorem 1, the following two lemmas due to Pólya [7] and Fan [8] respectively, and von Neumann's theorem [6] mentioned above.

Lemma 4. *If $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ are two non-negative sequences arranged in decreasing order such that*

$$(31) \quad a_1 a_2 \cdots a_k \geq b_1 b_2 \cdots b_k \quad k = 1, 2, \dots, n$$

then

$$(32) \quad \sum_{i=1}^k \omega(a_i) \geq \sum_{i=1}^k \omega(b_i) \quad k = 1, 2, \dots, n$$

for any increasing function $\omega(x)$ of positive argument such that $\omega(e^\xi)$ is a convex function of ξ .

Lemma 5. If $a_1, a_2, \dots, a_n$ and $b_1, b_2, \dots, b_n$ are two non-negative sequences arranged in decreasing order, then for any symmetric gauge function Φ ,

$$(33) \quad \Phi(a_1, a_2, \dots, a_n) \geq \Phi(b_1, b_2, \dots, b_n)$$

if and only if condition (32) with $\omega(x) = x$ is satisfied.

As a consequence of Lenard's analysis, it follows from Theorem 2 that for any unitarily invariant norm $\|\cdot\|$ and Hermitian A and B ,

$$(34) \quad \|e^A e^B\| \geq \|e^{A+B}\|$$

We conclude in the next section with simple direct proofs of (4) and (34) and some further generalizations.

4. Further inequalities. The main result for proving inequalities of the type (1) is the following

Lemma 6. If, for a continuous real valued matrix function $\varphi(X)$

$$(35) \quad \varphi(XY) = \varphi(YX) \text{ for } Y \text{ positive,}$$

and

$$(36) \quad \varphi([XX^\dagger]^s) \geq |\varphi(X^{2s})| \quad s = 1, 2, \dots$$

for an arbitrary complex matrix X , then

$$(37) \quad \varphi(e^A e^B) \geq \varphi(e^{A+B})$$

for Hermitian A and B .

Proof. The proof is almost identical with the proof of Lemma 2, Ref. [3]. Thus, if we take $X = AB$ and $s = 2^{k-1}$ in (36) with A and B positive, we obtain from (35)

$$(38) \quad \varphi([A^2 B^2]^{2^{k-1}}) = \varphi([AB^2 A]^{2^{k-1}}) \geq |\varphi([AB]^{2^k})|.$$

Repeated iteration of (38) then gives

$$(39) \quad \varphi(A^{2^k} B^{2^k}) \geq |\varphi([AB]^{2^k})|.$$

Replacing A and B by $\exp(2^{-k}A)$ and $\exp(2^{-k}B)$ respectively in (39), (37) is obtained from the Trotter formula [9],

$$\lim_{k \rightarrow \infty} [\exp(2^{-k}A) \exp(2^{-k}B)]^{2^k} = \exp(A+B),$$

and the continuity of φ .

In the following, we will only consider functions $\varphi(X)$ of the spectrum of X .

When Y is positive, $\text{Spec}(XY) = \text{Spec}(YX)$ so condition (35) of Lemma 6 will always hold.

The object now is to construct functions φ satisfying conditions (35) and (36) of Lemma 6. For this purpose, we need the following result due to Weyl [4], which follows directly from Lemmas 1 and 4.

Lemma 7. *If $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\kappa_1, \kappa_2, \dots, \kappa_n$ are the eigenvalues of A and $AA^\dagger$ respectively, arranged so that $|\alpha_1| \geq |\alpha_2| \geq \dots \geq |\alpha_n| \geq 0$ and $\kappa_1 \geq \kappa_2 \geq \dots \geq \kappa_n \geq 0$, then for any increasing function $\omega(x)$ of positive argument x such that $\omega(e^\xi)$ is a convex function of ξ ,*

$$(40) \quad \sum_{i=1}^k \omega(\kappa_i) \geq \sum_{i=1}^k \omega(|\alpha_i|^2) \quad k = 1, 2, \dots, n.$$

For convenience, we will call functions $\omega(x)$ satisfying the conditions of Lemma 7 W -functions.

The simplest example of a W -function is $\omega(x) = x^s$, $s = 1, 2, \dots$, in which case (40) implies that

$$(41) \quad \sum_{i=1}^k \kappa_i^s \geq \sum_{i=1}^k |\alpha_i|^{2s} \geq \left| \sum_{i=1}^k \alpha_i^{2s} \right| \quad k = 1, 2, \dots, n.$$

When $k = n$, (41) is equivalent to

$$(42) \quad \text{Tr}(AA^\dagger)^s \geq |\text{Tr} A^{2s}|.$$

The basic inequality $\text{Tr}(e^A e^B) \geq \text{Tr}(e^{A+B})$ follows directly from (42) and Lemma 6 (this, in fact, was how it was originally proved in [3]).

As noted previously, if $a_1 \geq a_2 \geq \dots \geq a_n \geq 0$ are the eigenvalues of a matrix A , then $a_1 a_2 \dots a_k$ is the maximum eigenvalue of the skew symmetric tensor product $A^{(k)}$ of A taken with itself k times. Applying Lemma 1 and Lemma 7 in its simplest form (41) to $A^{(2)}, A^{(3)}, \dots$, then gives

$$(43) \quad \begin{aligned} \sum_{i_1 < i_2} \kappa_{i_1 i_2}^s &\geq \sum_{i_1 < i_2} |\alpha_{i_1}|^{2s} |\alpha_{i_2}|^{2s} \\ \sum_{i_1 < i_2 < i_3} \kappa_{i_1 i_2 i_3}^s &\geq \sum_{i_1 < i_2 < i_3} |\alpha_{i_1}|^{2s} |\alpha_{i_2}|^{2s} |\alpha_{i_3}|^{2s} \quad \text{etc.} \end{aligned}$$

It follows, as noted by Weyl [4], that the polynomial $Q_s(z) = \prod_{i=1}^n (1 + |\alpha_i|^{2s} z)$ is majorized coefficient for coefficient by $P_s(z) = \prod_{i=1}^n (1 + \kappa_i^s z)$. The coefficients $I_s(X)$ in the characteristic polynomial $\text{Det}(I + zX)$ of X therefore satisfy the conditions of Lemma 6, and hence Lenard's result (4) follows.

As a further example, if in Lemma 7 we take $\omega(x) = \varphi(x^s)$, with $\varphi(|x|) \geq |\varphi(x)|$ and φ a W -function, we obtain

$$(44) \quad \begin{aligned} \sum_{i=1}^k \varphi(\kappa_i^s) &\geq \sum_{i=1}^k \varphi(|\alpha_i|^{2s}) \\ &\geq \sum_{i=1}^k |\varphi(\alpha_i^{2s})| \\ &\geq \left| \sum_{i=1}^k \varphi(\alpha_i^{2s}) \right|, \quad k = 1, 2, \dots, n. \end{aligned}$$

If we define

$$(45) \quad \varphi_k(A) = \sum_{i=1}^k \varphi(a_i) \quad k = 1, 2, \dots, n$$

for an arbitrary complex $n \times n$ matrix A with eigenvalues $a_1, a_2, \dots, a_n$ arranged such that $|a_1| \geq |a_2| \geq \dots \geq |a_n|$, and φ satisfying the conditions stated in the previous paragraph, then (44) implies (36) and hence we obtain

$$(46) \quad \varphi_k(e^A e^B) \geq \varphi_k(e^{A+B}) \quad k = 1, 2, \dots, n$$

for Hermitian $n \times n$ matrices A and B .

Notice that when $k = 1$, (46) gives an inequality between the maximum eigenvalues of $e^A e^B$ and e^{A+B} . Also, when $k = n$, (45) and (46) give, with obvious notation,

$$(47) \quad \text{Tr} [\varphi(e^A e^B)] \geq \text{Tr} [\varphi(e^{A+B})].$$

A final generalization of (1) results immediately from (44) and Lemma 5. Thus, if we define

$$(49) \quad \Phi(\varphi(A)) = \Phi(|\varphi(a_1)|, |\varphi(a_2)|, \dots, |\varphi(a_n)|),$$

where Φ is a symmetric gauge function, $\varphi(x)$ satisfies the conditions required for (44), and $a_1, a_2, \dots, a_n$ are the eigenvalues of an arbitrary complex $n \times n$ matrix A , then Lemma 5 and Lemma 6 give

$$(50) \quad \Phi[\varphi(e^A e^B)] \geq \Phi[\varphi(e^{A+B})]$$

for Hermitian $n \times n$ matrices A and B .

Equation (34) is obviously a special case of (50).

Appendix. We show here that the most obvious generalizations of the basic inequality involving three matrices, namely

$$(A1) \quad |\text{Tr} (e^A e^B e^C)| \geq \text{Tr} (e^{A+B+C})$$

or, in symmetrical form,

$$(A2) \quad \frac{1}{2}[\text{Tr} (e^A e^B e^C) + \text{Tr} (e^A e^C e^B)] \geq \text{Tr} (e^{A+B+C})$$

for Hermitian A, B , and C are, in general, false.

It is to be noted first of all that

$$(A3) \quad \begin{aligned} \frac{1}{2}[\text{Tr} (e^A e^B e^C) + \text{Tr} (e^A e^C e^B)] &= \frac{1}{2}[\text{Tr} (e^A e^B e^C) + \text{Tr} (e^A e^B e^C)^\dagger] \\ &= \Re \text{Tr} (e^A e^B e^C) \\ &\leq |\text{Tr} (e^A e^B e^C)| \end{aligned}$$

where $\Re$ denotes real part and in the second and third steps we have used the elementary results $\text{Tr} (XY) = \text{Tr} (YX)$ and $\text{Tr} (X^\dagger) = [\text{Tr} (X)]^*$ (* denoting complex conjugate). It follows from (3) that if (1) is false then so is (2).

A simpler counterexample to (1) is obtained by choosing

$$(A4) \quad A = \epsilon \sigma_1, B = \epsilon \sigma_2, \text{ and } C = \epsilon(\sigma_3 - \sigma_2 - \sigma_1)$$

where $\sigma_1, \sigma_2, \sigma_3$ are the 2×2 Pauli matrices ($i^2 = -1$)

$$(A5) \quad \sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

An elementary calculation shows that

$$(A6) \quad \left(\sum_{i=1}^3 a_i \sigma_i \right)^2 = \left(\sum_{i=1}^3 a_i^2 \right) I, \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

So, by straightforward expansion we obtain

$$(A7) \quad \exp \left(\sum_{i=1}^3 a_i \sigma_i \right) = I \cosh \alpha + \left(\sum_{i=1}^3 a_i \sigma_i \right) \alpha^{-1} \sinh \alpha,$$

where

$$\alpha = \left(\sum_{i=1}^3 a_i^2 \right)^{1/2}.$$

Applying (7) to (1) for the choice (4), the left hand side of (1) is given by

$$(A8) \quad \begin{aligned} & 2 |\cosh^2 \epsilon \cosh \epsilon 3^{1/2} - 2 \times 3^{-1/2} \cosh \epsilon \sinh \epsilon \sinh \epsilon 3^{1/2} + i 3^{-1/2} \sinh^2 \epsilon \sinh \epsilon 3^{1/2}| \\ & = 2 \cosh \epsilon \left\{ 1 - \frac{\epsilon^4}{12} + O(\epsilon^6) \right\}. \end{aligned}$$

The right hand side of (1), on the other hand, is $2 \cosh \epsilon$. So, from (8) the inequality (1) (and hence (2) also) is violated for sufficiently small ϵ .

Acknowledgments. I am grateful to A. Lenard for posing the problem and for several helpful remarks, to F. J. Dyson for comments, and to Dr. Carl Kaysen for his kind hospitality at the Institute for Advanced Study.

REFERENCES

1. A. LENARD, *Theory of the Golden-Thompson inequality*, Indiana Univ. Math. J. 21 (1971), 457-467.
2. S. GOLDEN, *Lower bounds for the Helmholtz function*, Phys. Rev. 137 (1965), B1127-B1128.
3. C. J. THOMPSON, *Inequality with applications in statistical mechanics*, J. Math. Phys. 6 (1965), 1812-1813.
4. H. WEYL, *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Natl. Acad. Sci. U. S. 35 (1949), 408-411.
5. A. HORN, *On the singular values of a product of completely continuous operators*, Proc. Natl. Acad. Sci. U. S. 36 (1950), 374-375.
6. J. von NEUMANN, Tomsk Univ. Rev. 1 (1937) 286; *Collected works*, Vol. IV, 205-218. For further properties of symmetric gauge functions and their relationship to unitarily invariant norms, see R. Schatten, *A theory of cross spaces*, Princeton University Press, 1950, 84-95.

7. G. PÓLYA, *Remark on Weyl's note "Inequalities between the two kinds of eigenvalues of a linear transformation,"* Proc. Natl. Acad. Sci. U. S. **36** (1950), 49.
8. K. FAN, *Maximum properties and inequalities for the eigenvalues of completely continuous operators,* Proc. Natl. Acad. Sci. U. S. **37** (1951), 760-766.
9. H. F. TROTTER, *On the product of semi-groups of operators,* Proc. Amer. Math. Soc. **10** (1959), 545-551.

Supported at IAS by a grant from the Alfred P. Sloan Foundation.

The Institute for Advanced Study
Date communicated: AUGUST 2, 1971