Generalization of the Golden–Thompson Inequality $\text{Tr} (e^A e^B) \geq \text{Tr} e^{A+B}$

A. LENARD

Communicated by P. R. HALMOS

1. Introduction. If A and B are two hermitian square matrices with complex entries, of the same size, then

$$(1) \qquad \text{Tr} (e^A e^B) \geq \text{Tr} (e^{A+B})$$

and the equal sign holds only if A and B commute. This fact was discovered independently by Golden [1] and Thompson [2]. (The necessary condition for equality is not explicitly mentioned by these authors but is easily obtained from their proofs.) In this paper we analyze the mathematical situation underlying (1). This leads to a proof which displays a deeper source of the theorem than given in [1] and [2]. In our approach, attention is focused on a large class of functions φ which may replace the trace in (1). This class is characterized by a certain monotonicity property of the functions belonging to it. Our proof of (1) consists of the observation that the trace function has this property. There are many other, equally natural, functions showing this property.

By “matrix” we always mean a finite matrix (n by n , say) with complex entries. The Hermitian adjoint of X is written X^* . A Hermitian matrix with non-negative eigenvalues is called positive. If X is positive, we denote its unique positive square-root by $X^{1/2}$. We write $\text{Spec } X$ for the collection of the eigenvalues of X , each counted with its proper multiplicity. Thus $\text{Spec } X$ is an (unordered) n -tuple of complex numbers. The symbols σ , σ' , etc., will be used to denote such n -tuples. $\sigma \geq 0$ means that all members of σ are non-negative. $\bar{\sigma}$ denotes the n -tuple whose members are the complex conjugates of those of σ .

2. The Partial Order $\succ$. Let X , Y be positive matrices and suppose that unitary matrices U , V exist such that

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

In this case we write $X \succ Y$. Obviously if X and Y are unitarily equivalent, then $X \succ Y$: choose $U = V = 1$.

Proposition 1. *If $X \succ Y$, then $\text{Tr } X \geq \text{Tr } Y$. Equality holds only if X and Y are unitarily equivalent.*

Proof. We use Schwarz's inequality in this form: For any two matrices X, Y

$$(3) \quad |\text{Tr } (X^*Y)|^2 \leq \text{Tr } (X^*X) \text{Tr } (Y^*Y),$$

and equality holds only when X and Y are scalar multiples of each other. Let then $X \succ Y$. We have from (3)

$$\begin{aligned} \text{Tr } Y = \text{Tr } (X^{1/2}UX^{1/2}V) &\leq [\text{Tr } (X^{1/2}U)(X^{1/2}U)^*]^{1/2} \\ &\quad \times [\text{Tr } (X^{1/2}V)^*(X^{1/2}V)]^{1/2} = \text{Tr } X. \end{aligned}$$

Assume now that equality holds. If $X = Y = 0$, there is nothing to prove. If not, we have $\text{Tr } X > 0$. The condition of equality in (3) gives then $U^*X^{1/2} = cX^{1/2}V$ for some complex c , and this implies $X = |c|^2 V^*XV$. So by taking the trace we get $|c| = 1$. Now

$$(4) \quad \text{Spec } Y = \text{Spec } (X^{1/2}UX^{1/2}V) = \bar{c} \text{Spec } (X^{1/2}UU^*X^{1/2}) = \bar{c} \text{Spec } X.$$

But X and Y are positive, and so $c = 1$. This shows that X and Y are unitarily equivalent. Q.E.D.

Note that the relation $\succ$ may be thought of as being defined for pairs of unitary equivalence classes of positive matrices. In other words, if X is unitarily equivalent to X' , Y to Y' , then $X \succ Y$ if and only if $X' \succ Y'$.

The relation $\succ$ is not obviously a partial order, but it may be trivially extended to one as follows. We write $X \succ Y$ for two positive matrices X, Y whenever a finite chain $X \succ T_1, T_1 \succ T_2, \dots, T_s \succ Y$ exists where the T_i s are positive matrices and $\succ$ is used in the previously defined sense. We claim that $\succ$ is a partial order on unitary equivalence classes of matrices. Transitivity and reflexivity are obvious. It must only be shown that $X \succ Y$ and $Y \succ X$ imply the unitary equivalence of X and Y . Indeed, in this case we have $\text{Tr } X = \text{Tr } Y$ from Proposition 1 and if the T_i are as required in the definition of $X \succ Y$, we have $\text{Tr } X \geq \text{Tr } T_1 \geq \dots \geq \text{Tr } T_s \geq \text{Tr } Y = \text{Tr } X$. Thus all signs are $=$, and the conclusion follows again from Proposition 1.

We are now ready to state our principal result.

Theorem. *If A, B are Hermitian matrices, an arbitrary neighborhood of e^{A+B} contains X such that $X \prec e^{A/2}e^Be^{A/2}$.*

An immediate deduction from the theorem is the following fact. If φ is a real-valued continuous function on matrices which is monotonic with respect to $\succ$, i.e., $\varphi(X) \geq \varphi(Y)$ whenever $X \succ Y$, then for Hermitian A, B

$$(5) \quad \varphi(e^{A/2}e^Be^{A/2}) \geq \varphi(e^{A+B}).$$

Indeed, we may choose a sequence $X_k \rightarrow e^{A+B}$ as $k \rightarrow \infty$ and such that for all k $X_k \prec e^{A/2} e^B e^{A/2}$. Then

$$\varphi(e^{A/2} e^B e^{A/2}) \geq \varphi(X_k)$$

and (5) follows by passage to the limit.

An example of a continuous real function φ which is monotonic with respect to $\succ$ is $\varphi(X) = |\text{Tr } X|$ (cf. Proposition 1). We shall see other examples later. (Sec. 6.)

We note that if φ is monotonic with respect to $\succ$, then for positive X , $\varphi(X)$ is determined by the unitary equivalence class of X , i.e., by $\sigma = \text{Spec } X$. In this sense we may write $\varphi = \varphi(\sigma)$ for $\sigma \geq 0$.

The proof of the theorem depends on two preliminary propositions of independent interest. We discuss them in the next sections.

3. The function ψ .

Proposition 2. *Let φ be a bounded, real valued function defined on positive matrices and monotonic with respect to $\succ$. Then there exists a bounded real-valued function ψ , defined on ordered pairs of matrices, possessing the following properties*

- (I) $\psi(X, X) = \varphi(X)$ for positive X ,
- (II) $\psi(XY, ZW) = \psi(YX, WZ)$,
- (III) $\psi(XX^*, YY^*) \geq \psi(XY^*, YX^*)$.

Proof. Let φ be given as stated. We write $\varphi(X) = \varphi(\text{Spec } X) = \varphi(\sigma)$, which is defined for all $\sigma \geq 0$. The monotonicity of φ may be stated as follows:

$$(6) \quad \varphi(\sigma') \geq \varphi(\sigma),$$

whenever positive X, Y and unitary U, V exist such that

$$(7) \quad \begin{cases} \text{Spec } X = \sigma', \\ \text{Spec } Y = \sigma, \\ \text{Spec } (X^{1/2} U X^{1/2} V) = \sigma. \end{cases}$$

We shall define a real valued function $f = f(\sigma, \sigma')$ satisfying the condition

$$(8) \quad f(\sigma, \sigma) = \varphi(\sigma) \quad (\text{if } \sigma \geq 0)$$

and in terms of it we set

$$(9) \quad \psi(X, Y) = f(\text{Spec } X, \text{Spec } Y).$$

Note that (8) and (9) guarantee property (I). Moreover, $\text{Spec } (XY) = \text{Spec } (YX)$ for any pair of matrices; hence property (II) also holds. It remains now only to define f so that property (III) shall hold.

Every matrix X may be written in the polar form $X = P^{1/2} U$ where $P = XX^*$ is positive and U unitary. Hence (III) may be reformulated

$$\psi(P, Q) \geq \psi(P^{1/2} U Q^{1/2}, Q^{1/2} U^* P^{1/2}),$$

when P, Q are positive, U unitary. In terms of the f which is related to ψ by (9), this requirement reads:

$$(10) \quad f(\sigma', \sigma'') \geq f(\sigma, \bar{\sigma})$$

whenever positive P, Q and unitary U exist satisfying

$$(11) \quad \begin{cases} \text{Spec } P = \sigma' \\ \text{Spec } Q = \sigma'' \\ \text{Spec } (P^{1/2} U Q^{1/2}) = \sigma. \end{cases}$$

Let us examine the inequalities (10). The argument on the left is always in $\mathcal{S}_+ = \{(\sigma', \sigma'') : \sigma', \sigma'' \geq 0\}$, and that on the right is always in $\mathcal{S}_* = \{(\sigma, \bar{\sigma})\}$. The intersection of these sets is $\mathcal{S}_+ \cap \mathcal{S}_* = \{(\sigma, \sigma) : \sigma \geq 0\}$. f is defined over it by (8), and one easily checks that for the cases where the arguments on both sides of (10) are in $\mathcal{S}_+ \cap \mathcal{S}_*$ the inequality is satisfied in virtue of (6) and (7). We define f to be piecewise constant over the three sets: $\mathcal{S}_+ \setminus \mathcal{S}_*$, $\mathcal{S}_* \setminus \mathcal{S}_+$, and the complement of $\mathcal{S}_+ \cup \mathcal{S}_*$. Namely, for $(\sigma, \sigma'') \in \mathcal{S}_+ \setminus \mathcal{S}_*$ set $f(\sigma', \sigma'') = \sup f(\sigma, \sigma)$, the supremum being over all $\sigma \geq 0$. Then for $(\sigma, \bar{\sigma}) \in \mathcal{S}_* \setminus \mathcal{S}_+$ set $f(\sigma, \bar{\sigma}) = \inf f(\sigma', \sigma'')$, the infimum being over $(\sigma', \sigma'') \in \mathcal{S}_+$. Finally, for all (σ', σ'') not in $\mathcal{S}_+ \cup \mathcal{S}_*$, set $f(\sigma', \sigma'') = 0$. This defines f completely so that all inequalities (10) are satisfied. Hence ψ has property (III), and the proof of Proposition 2 is complete.

4. An algebraic identity. Next we derive a fact that depends only on the multiplicative structure of matrices and the $*$ -operation. Therefore, we formulate it in the appropriate abstract setting.

We have in mind a system $(\mathfrak{S}, \psi)$ consisting of a $*$ -semigroup $\mathfrak{S}$, and a real valued function ψ on $\mathfrak{S} \times \mathfrak{S}$, possessing cyclic invariance in each variable. This means the following. $\mathfrak{S}$ is a set (with elements $x, y, \dots$) equipped with an associative multiplication and a mapping $x \rightarrow x^*$ into itself which is order inverting with respect to multiplication: $(xy)^* = y^*x^*$. Cyclic invariance of ψ means $\psi(xy, zw) = \psi(yx, wz)$. An element x of $\mathfrak{S}$ is said to be Hermitian if $x = x^*$. For the sake of brevity, we let

$$\omega(x, y) = \psi(xx^*, yy^*) - \psi(xy^*, yx^*).$$

We are interested in a class of monomials formed from two arbitrary elements x, y of $\mathfrak{S}$. They are indexed by finite sequences of $+$ and $-$ signs. We use the letters $\alpha, \beta, \dots$ to denote such sequences, and adopt the convention that $\alpha\beta$ stands for the sequence formed from α followed by β . α is called even (odd) if the number of $-$ signs in it is even (odd). $\delta_\alpha = 1$ for even α , $\delta_\alpha = -1$ for odd α . $|\alpha|$ denotes the number of signs in α . The special symbol φ denotes the empty sequence.

The monomial associated with α is written $m_\alpha(x, y)$. Its definition is inductive on $|\alpha|$,

$$(12) \quad \begin{cases} m_\varphi(x, y) = x, \\ m_{\alpha+}(x, y) = m_\alpha(x, y)m_\alpha(x, y) \\ m_{\alpha-}(x, y) = m_\alpha(x, y)m_\alpha(y, x). \end{cases}$$

For instance, $m_{+-}(x, y) = x^2y^4x^2$. The degree of $m_\alpha(x, y)$ is $2^{|\alpha|}$, and there are 2^p different m_α with $|\alpha| = p$. One verifies easily by induction on $|\alpha|$ that

$$(13) \quad m_\alpha(x, y)^* = \begin{cases} m_\alpha(x^*, y^*) & (\alpha \text{ even}) \\ m_\alpha(y^*, x^*) & (\alpha \text{ odd}). \end{cases}$$

Proposition 3. For any α , any non-negative integer p , and any Hermitian $x, y \in \mathfrak{S}$, the following identity holds

$$(14) \quad \delta_\alpha \sum_{|\beta|=p} \omega(m_{\alpha\beta}(x, y), m_{\alpha\beta}(y, x)) \\ = \psi([m_{\alpha+}(x, y)]^{2^p}, [m_{\alpha+}(y, x)]^{2^p}) - \psi([m_{\alpha-}(x, y)]^{2^p}, [m_{\alpha-}(y, x)]^{2^p}).$$

Proof. Let x, y be Hermitian, α even. Then

$$\begin{aligned} \omega(m_\alpha(x, y), m_\alpha(y, x)) &= \psi(m_\alpha(x, y)m_\alpha(x, y)^*, m_\alpha(y, x)m_\alpha(y, x)^*) \\ &\quad - \psi(m_\alpha(x, y)m_\alpha(y, x)^*, m_\alpha(y, x)m_\alpha(x, y)^*) = \psi(m_{\alpha+}(x, y), m_{\alpha+}(y, x)) \\ &\quad - \psi(m_{\alpha-}(x, y), m_{\alpha-}(y, x)). \end{aligned}$$

On the other hand, for odd α ,

$$\omega(m_\alpha(x, y), m_\alpha(y, x)) = \psi(m_{\alpha-}(x, y), m_{\alpha-}(y, x)) - \psi(m_{\alpha+}(x, y), m_{\alpha+}(y, x)),$$

i.e., the signs of the two terms are interchanged. Therefore the left hand side of (14) may be rewritten as

$$(15) \quad \sum_{|\beta|=p+1} \delta_\beta \psi(m_{\alpha\beta}(x, y), m_{\alpha\beta}(y, x)).$$

Consider now two terms in (15) corresponding to $\beta = \gamma + - + + \cdots +$ and $\beta = \gamma - - + + \cdots +$ (γ the same, number of $+$ signs at the end k , say). In the first case

$$m_{\alpha\beta}(x, y) = (a^2b^2)^{2^k},$$

where $a = m_{\alpha\gamma}(x, y)$ and $b = m_{\alpha\gamma}(y, x)$, and in the second case

$$m_{\alpha\beta}(x, y) = (ab^2a)^{2^k}.$$

Clearly, these two monomials are cyclically equivalent, and so from the cyclic invariance of ψ it follows that the two terms in (15) are equal apart for sign, since δ_β is opposite for them. We conclude that all terms of the sum which can be paired this way cancel. These include all with the exception of two: $\beta = + + + \cdots +$ and $\beta = - + + \cdots +$. They are the two terms on the right hand side of (14). Q.E.D.

5. Proof of the theorem. We apply the identity (14) to the case at hand: $\mathfrak{S}$ is the set of matrices and $*$ the usual Hermitian adjoint. We suppose that $\psi = \psi(X, Y)$ is a real valued function having properties (II) and (III) of Proposition 2. Then

$$\omega(X, Y) \geq 0$$

and so (14) turns into an inequality. For odd α ,

$$\psi([m_{\alpha-}(X, Y)]^{2^p}, [m_{\alpha-}(Y, X)]^{2^p}) \geq \psi([m_{\alpha+}(X, Y)]^{2^p}, [m_{\alpha+}(Y, X)]^{2^p})$$

(for even α the inequality is opposite). In particular, when $\alpha = ++ \cdots +-$ with $|\alpha| = k + 1$, this reads

$$\psi([X^{2^k} Y^{2^{k+1}} X^{2^k}]^{2^p}, [Y^{2^k} X^{2^{k+1}} Y^{2^k}]^{2^p}) \geq \psi([X^{2^k} Y^{2^k}]^{2^{p+1}}, [Y^{2^k} X^{2^k}]^{2^{p+1}}).$$

Let now $\varphi(X) = \psi(X, X)$, and make use of cyclic invariance. Then

$$\varphi([X^{2^{k+1}} Y^{2^{k+1}}]^{2^p}) \geq \varphi([X^{2^k} Y^{2^k}]^{2^{p+1}}).$$

Repetition of this inequality yields

$$\varphi(X^{2^k} Y^{2^k}) \geq \varphi([XY]^{2^k}).$$

Put $X = e^{2^{-k}A}$, $Y = e^{2^{-k}B}$ where A, B are Hermitian. Then taking cyclic invariance of φ into account,

$$(16) \quad \varphi(e^{A/2} e^B e^{A/2}) \geq \varphi(X_k),$$

where X_k is the positive matrix cyclically equivalent to $(e^{2^{-k}A} e^{2^{-k}B})^{2^k}$, obtained from it by shifting a factor $e^{2^{-(k+1)}A}$ from the front to the end of the product. We note that

$$\lim_{k \rightarrow \infty} X_k = e^{A+B}$$

(the "Trotter formula"), so that any neighborhood of e^{A+B} contains all X_k for sufficiently large k .

It remains to show that $X_k \prec e^{A/2} e^B e^{A/2}$. To this end, let φ be defined by

$$\varphi(X) = \begin{cases} 0 & \text{for } X \prec e^{A/2} e^B e^{A/2} \\ 1 & \text{otherwise.} \end{cases}$$

Then φ is bounded and monotonic with respect to $\prec$, and so by Proposition 2 we have the corresponding function ψ . Then, as we have seen, (16) holds. But this means $\varphi(X_k) = 0$, i.e., $X_k \prec e^{A/2} e^B e^{A/2}$. This completes the proof of the theorem.

6. Continuous $*$ -representations. A large class of continuous $\succ$ -monotone functions may be obtained by using the concept of a continuous $*$ -representation. Let $\mathfrak{M}_n$ denote the set of matrices of size n by n and $\mathfrak{M}_\nu$ those of some other size ν by ν . We call a continuous mapping R of $\mathfrak{M}_n$ into $\mathfrak{M}_\nu$ a continuous

-representation if $R(XY) = R(X)R(Y)$, $R(X^) = R(X)^*$, and $R(1) = 1$. It is obvious that the properties of being Hermitian, unitary, non-singular and positive are preserved by R , and $R(X^{-1}) = R(X)^{-1}$. It is less obvious but true that if $\text{Spec } X = \text{Spec } Y$ then also $\text{Spec } R(X) = \text{Spec } R(Y)$. Indeed if X has distinct eigenvalues, then $Y = SXS^{-1}$ for some non-singular S ; then $R(Y) = R(S)R(X)R(S)^{-1}$ and the conclusion follows. If X has repeated eigenvalues, we consider two sequences $X_k \rightarrow X$, $Y_k \rightarrow Y$ such that $\text{Spec } X_k = \text{Spec } Y_k$ and contains no repeated eigenvalues. Then $\text{Spec } R(X) = \lim \text{Spec } R(X_k) = \lim \text{Spec } R(Y_k) = \text{Spec } R(Y)$.

These remarks make it clear that the relation $\succ$ is preserved by R . Therefore the function $\varphi(X) = |\text{Tr } R(X)|$ is continuous and monotonic with respect to $\succ$. Thus we have the following generalization of the Golden-Thompson inequality.

Proposition 4. *If R is a continuous *-representation, then for Hermitian A and B*

$$(17) \quad \text{Tr } R(e^A e^B) \geq \text{Tr } R(e^{A+B}).$$

An interesting continuous *-representation is constructed as follows. Let s be an integer, $1 \leq s \leq n$. Let the symbols (i) , (j) , etc., stand for ordered s -tuples of strictly increasing integers between 1 and n , e.g., $(i) = (i_1, i_2, \dots, i_s)$ where $1 \leq i_1 < i_2 < \dots < i_s \leq n$. The $\nu = \binom{n}{s}$ different values of (i) are used to label the rows and columns of a matrix $R(X)$ as follows

$$(18) \quad R(X)_{(i)(j)} = \sum_{\pi} \delta_{\pi} X_{i_1 i_{\pi(1)}} X_{i_2 i_{\pi(2)}} \cdots X_{i_s i_{\pi(s)}} = \text{Det}_{1 \leq a, b \leq s} (X_{i_a i_b})$$

(The sum is over the permutations π of $\{1, 2, \dots, s\}$ and $\delta_{\pi} = \pm 1$ according to even or odd π .) One verifies that R is a continuous *-representation (in algebraic terminology, $R(X)$ is the anti-symmetrized tensor product of X with itself s times). In the trivial case $s = 1$, $R(X) = X$; in the other trivial case $s = n$, $R(X) = \text{Det } X$ (i.e., is a 1 by 1 matrix). Let $I_s(X) = \text{Tr } R(X)$. We have then $I_1(X) = \text{Tr } X$, $I_n(X) = \text{Det } X$. We proceed to calculate $I_s(X)$ in general.

$$\begin{aligned} I_s(X) &= \sum_{(i)} \sum_{\pi} \delta_{\pi} X_{i_1 i_{\pi(1)}} X_{i_2 i_{\pi(2)}} \cdots X_{i_s i_{\pi(s)}} \\ &= \frac{1}{s!} \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_s=1}^n \sum_{\pi} \delta_{\pi} X_{i_1 i_{\pi(1)}} X_{i_2 i_{\pi(2)}} \cdots X_{i_s i_{\pi(s)}} \\ &= \frac{1}{s!} \sum_{\pi} \delta_{\pi} (\text{Tr } X)^{h_1(\pi)} (\text{Tr } X^2)^{h_2(\pi)} \cdots (\text{Tr } X^s)^{h_s(\pi)} \end{aligned}$$

where $h_1(\pi)$, $h_2(\pi)$, etc., are the numbers of 1-cycles, 2-cycles, etc., in π . Now

$$\delta_{\pi} = (-1)^{h_2(\pi) + h_4(\pi) + \dots},$$

and the number of permutations π with cycle character $(h_1, h_2, \dots, h_s) = (h)$ is

$$\frac{s!}{1^{h_1} h_1! 2^{h_2} h_2! \cdots s^{h_s} h_s!}.$$

Therefore

$$(19) \quad I_s(X) = \sum_{(h)} \frac{1}{h_1!} (\text{Tr } X)^{h_1} \frac{1}{h_2!} \left(-\frac{\text{Tr } X^2}{2} \right)^{h_2} \cdots$$

where the sum is over systems (h) of non-negative integers subject to $h_1 + 2h_2 + 3h_3 + \cdots = s$. We note that the sum in (19) is meaningful for any positive s even if $s > n$, so it may be used to extend the definition of $I_s(X)$. The final step in the calculation is the introduction of the generating function

$$(20) \quad 1 + \sum_{s=1}^{\infty} \lambda^s I_s(X) = e^{\lambda \text{Tr } X - (\lambda^2/2) \text{Tr } X^2 + (\lambda^3/3) \text{Tr } X^3 - \cdots} \\ = e^{\text{Tr } \log(1 + \lambda X)} = \text{Det}(1 + \lambda X).$$

This is just the characteristic polynomial of X and we see that $I_1(X)$, $I_2(X)$, etc., are its coefficients (the “invariants” of X). We have thus proved the following.

Proposition 5. *The coefficients $I_s(X)$ in the characteristic polynomial of X are monotonic with respect to $\succ$. Hence for Hermitian A, B*

$$(21) \quad I_s(e^A e^B) \geq I_s(e^{A+B}).$$

For $s = n$ (21) is always an equality.

Indeed, $\text{Det}(e^A e^B) = \text{Det}(e^A) \text{Det}(e^B) = e^{\text{Tr } A} e^{\text{Tr } B} = e^{\text{Tr } (A+B)} = \text{Det}(e^{A+B})$.

Other choices of R lead to other interesting inequalities. For instance, $R(X)$ may be defined by a formula analogous to (18) but with the omission of the factor δ_π (“symmetrized tensor product”). Here the subscripts (i) and (j) are non-decreasing s -tuples of positive integers between 1 and n (equivalently: un-ordered s -tuples). Thus $R(X)$ is a ν by ν matrix where $\nu = n^s/s!$. The calculation of its trace is completely along the above pattern. Denoting it by $J_s(X)$, one obtains

$$(22) \quad 1 + \sum_{s=1}^{\infty} \lambda^s J_s(X) = \text{Det}(1 - \lambda X)^{-1} = \text{Det}(1 + \lambda X + \lambda^2 X^2 + \cdots).$$

The $J_s(X)$ are symmetric polynomials in the eigenvalues of X and so, by a fundamental theorem of algebra, are expressible as polynomials in $I_1(X)$, $I_2(X)$, $\cdots$. Explicitly,

$$(23) \quad J_s = \begin{vmatrix} I_1 & I_2 & I_3 & \cdots & I_s \\ 1 & I_1 & I_2 & \cdots & I_{s-1} \\ 0 & 1 & I_1 & \cdots & I_{s-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I_1 \end{vmatrix}$$

where one must set $I_s = 0$ for $s > n$. For later use we note

$$(24) \quad J_2 = I_1^2 - I_2.$$

7. Alternate proof of Proposition 4. It must be pointed out that the result embodied in Proposition 4 is not quite the grand generalization of the Golden-Thompson inequality (1) as it appears to be at first sight. The fact is that it is possible to derive (17) in a completely elementary manner, using only the original Golden-Thompson inequality itself. (This was pointed out to the writer by Professor Oscar E. Lanford, III. His enlightening remarks are gratefully acknowledged.)

Let R be a continuous $*$ -representation (see Section 6 for definition). Then for a fixed X the function $F(t) = R(e^{tX})$ is not merely continuous but also differentiable. Indeed $F(t + t') = F(t)F(t')$, and therefore

$$\frac{1}{h} [F(t + h) - F(t)] \int_{-e}^e F(\tau) d\tau = \frac{1}{h} \int_{-e}^{e+h} F(t + \tau) d\tau - \frac{1}{h} \int_{-e}^{-e+h} F(t + \tau) d\tau.$$

The right hand side here has the limit $F(t + e) - F(t - e)$ as $h \rightarrow 0$. Therefore

$$S(X) = \left[\frac{d}{dt} R(e^{tX}) \right]_{t=0}$$

exists, and one obtains in the standard manner

$$R(e^X) = e^{S(X)}.$$

It is easily verified that the function $S(X)$ has the properties $S(X + Y) = S(X) + S(Y)$ and $S(X^*) = S(X)^*$. Therefore if X is Hermitian so is $S(X)$.

Now let A and B be Hermitian. Then

$$\begin{aligned} \text{Tr } R(e^A e^B) &= \text{Tr } R(e^A) R(e^B) = \text{Tr } e^{S(A)} e^{S(B)} \geq \text{Tr } e^{S(A) + S(B)} \\ &= \text{Tr } e^{S(A+B)} = \text{Tr } R(e^{A+B}) \end{aligned}$$

where the inequality is the application of the Golden-Thompson inequality itself in the representing space $\mathfrak{M}_n$.

This proof of (17) shows that from a certain point of view the original proofs [1], [2] of (1) are preferable to the method of the present paper, especially the proof of Golden [1] who uses no other tool than Schwarz's Inequality in the form (3). Our method has only the virtue of deriving (1) as well as its generalization (17) by a single reasoning whose conclusion is the theorem in Section 2. A true generalization of (1) would be obtained if a continuous function $\varphi = \varphi(X)$ were produced which is monotonic with respect to $\succ$ but is not of the form $|\text{Tr } R(X)|$ (or some monotonic function of it) when restricted to positive matrices. The writer has not been able to find such a function.

8. Alternate characterization of $\succ$. The definition of the partial order $\succ$, while mathematically impeccable, is not particularly enlightening, since it

gives no indication how for given positive matrices X , Y it can be decided whether $X \succ Y$ or not. It is desirable, therefore, to give some easily managed necessary and sufficient condition. This lends interest to the following fact.

Proposition 6. *For 2 by 2 matrices, $X \succ Y$ is equivalent to the conditions $\text{Det } X = \text{Det } Y$, $\text{Tr } X \geq \text{Tr } Y$.*

Thus for 2 by 2 matrices all functions monotonic with respect to $\succ$ are of the form $\varphi(X) = f(\text{Det } X, \text{Tr } X)$ where f is monotonic in its second argument for every fixed value of its first argument.

Proof. In view of Proposition 5 and the definition of $\succ$ (see Section 2) we only have to show that if $\text{Det } X = \text{Det } Y$ and $\text{Tr } X \geq \text{Tr } Y$ for positive matrices X and Y , we can find unitary U and V such that $\text{Spec } Y = \text{Spec } (X^{1/2}UX^{1/2}V)$. Since for 2 by 2 matrices the spectrum is determined by the determinant and the trace, this is equivalent to $\text{Det } Y = \text{Det } X \text{Det } U \text{Det } V$ and $\text{Tr } Y = \text{Tr } (X^{1/2}UX^{1/2}V)$. It is no restriction on generality to consider only diagonal X and Y .

We distinguish two cases. Let first $\text{Det } X = \text{Det } Y = 0$. We let

$$X = \begin{pmatrix} 0 & 0 \\ 0 & \xi \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 \\ 0 & \eta \end{pmatrix},$$

where in view of $\text{Tr } X \geq \text{Tr } Y$ we have $0 \leq \eta \leq \xi$. Let, for some real θ ,

$$(25) \quad U = V = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.$$

Then

$$X^{1/2}UX^{1/2}V = \begin{pmatrix} 0 & 0 \\ -\xi \cos \theta \sin \theta & \xi \cos^2 \theta \end{pmatrix}.$$

Thus, for appropriate θ , $\text{Tr } (X^{1/2}UX^{1/2}V) = \xi \cos^2 \theta = \eta = \text{Tr } Y$.

In the second case, let $\text{Det } X = \text{Det } Y > 0$. It is no restriction to assume $\text{Det } X = \text{Det } Y = 1$, since multiplying all matrices by a common positive scalar factor leaves the relation $\succ$ unchanged. We put then

$$X = \begin{pmatrix} \xi & 0 \\ 0 & \xi^{-1} \end{pmatrix}, \quad Y = \begin{pmatrix} \eta & 0 \\ 0 & \eta^{-1} \end{pmatrix},$$

where $0 < \xi \leq \eta \leq 1$ now corresponds to $\text{Tr } X \geq \text{Tr } Y$. For U and V we again take (25). A straightforward computation then yields the result that

$$\text{Tr } (X^{1/2}UX^{1/2}V) = (\text{Tr } X) \cos^2 \theta - 2 \sin^2 \theta.$$

Thus for θ satisfying

$$\cos^2 \theta = \frac{2 + \text{Tr } Y}{2 + \text{Tr } X}$$

we have $\text{Tr}(X^{1/2}UX^{1/2}V) = \text{Tr}Y$. This proves Proposition 6.

In view of this result and Proposition 5, a reasonable conjecture might be that the conditions

$$(26) \quad \begin{cases} I_s(X) \geq I_s(Y) & (s = 1, 2, \dots, n-1) \\ I_n(X) = I_n(Y) \end{cases}$$

are sufficient as well as necessary for $X \succ Y$. This is however not the case for $n \geq 3$. Indeed, let X be positive, with strictly positive and distinct eigenvalues. Furthermore, let $\lambda = \eta_1, \eta_2, \dots, \eta_n$ be the solutions of the polynomial equation

$$1 + I_1(X)\lambda + [I_2(X) - \epsilon]\lambda^2 + I_3(X)\lambda^3 + \dots + I_n(X)\lambda^n = 0$$

where $\epsilon > 0$ is so small that all solutions are still > 0 . Now let Y be the diagonal matrix with non-zero entries $\eta_1, \eta_2, \dots, \eta_n$. Then $I_s(X) = I_s(Y)$ for $s = 1, 3, 4, \dots, n$, whereas $I_2(X) - \epsilon = I_2(Y) < I_2(X)$. Thus the conditions (26) are fulfilled, yet $X \succ Y$ is false because this implies $J_2(X) \geq J_2(Y)$, whereas from (24) one gets $J_2(X) = J_2(Y) - \epsilon < J_2(Y)$.

The writer has not succeeded to discover any simple characterization of $\succ$ for $n \geq 3$.

Note added in proof. The desired characterization was found by C. J. Thompson after the present paper was completed. It, together with a number of elegant results related to this subject, appears as the succeeding article in this journal.

REFERENCES

1. S. GOLDEN, *Lower bounds for the Helmholtz function*, Phys. Rev. **137** (1965), B1127-B1128.
2. C. J. THOMPSON, *Inequality with applications in statistical mechanics*, J. Math. Phys. **6** (1965), 1812-1813.

This work was supported by NSF Grant GP-22926.

Indiana University

Date communicated: SEPTEMBER 7, 1971