ON THE POSITIVE SEMIDEFINITE NATURE OF A CERTAIN MATRIX EXPRESSION

EUGENE P. WIGNER AND MUTSUO M. YANASE*

By "positive definite matrices" or, briefly, definite matrices, we mean in this note self-adjoint matrices all the characteristic values of which are positive. Alternatively, they can be defined as matrices, all the hermitian quadratic forms of which are real and positive. It is well known that these two definitions are equivalent and it is the existence of two equivalent definitions which renders the subject particularly interesting. Thus, it follows from the first definition that A^{-1} is positive definite for positive definite A; it follows from the second definition that the sum of two positive definite matrices is also positive definite. Similarly, we shall call a matrix positive semidefinite or, briefly, semidefinite if it is self-adjoint and none of its characteristic values are negative, or if all of its hermitian quadratic forms are non-negative. A matrix which is definite is also semidefinite but a semidefinite matrix is, of course, not necessarily definite.

The concept of a positive definite matrix permits a partial ordering of the self-adjoint matrices. $A > B$ means that $A - B$ is positive definite. The principal and very deep theorem concerning such a partial ordering is due to Löwner (2). It specifies the functions which preserve the partial ordering. It states: If the characteristic values of both A and B are in an interval i, the equation $f(A) > f(B)$ follows from $A > B$ if the anaiytic function f, defined in the real interval i, but extended into the complex plane, is of such nature that the signs of the imaginary part of z and of $f(z)$ are the same. It follows from the last condition immediately that $f(z)$ is real and monotonically increasing in i, but this is only a necessary condition. The analytic extension referred to in Löwner's theorem must not cross the real axis. Thus, as will incidentally be shown also in the present note, $A^{\frac{1}{2}} > B^{\frac{1}{2}}$ if $A > B$ and both A and B are positive definite. Hence, the interval i in this case is $(0, \infty)$ but it may be any interval on the positive real axis. However, $z^{\frac{1}{2}} = r^{\frac{1}{2}}.e^{\frac{1}{2}i\phi}$ if $z = re^{i\phi}$ and $-\pi < \phi < \pi$, the analytic extension not permitting z to cross the negative real axis.

We shall not be concerned here with Löwner's theorem but with the positive semidefinite nature of a set of matrices which occurred in an expression suggested for information content (7). The two theorems of the present note are essentially equivalent and express the fact that the information content of a mixture $\rho = \alpha\rho_1 + (1 - \alpha)\rho_2$ of two ensembles ρ_1 and ρ_2 is a convex function

Received May 21, 1963.

*On leave of absence from Sophia University, Tokyo.

of the mixing parameter α for $0 \leqslant \alpha \leqslant 1$. The expression for the information content in question is $-\text{Trace}[\rho^{\frac{1}{2}}, k]^2$, it measures the amount of information which the ensemble ρ provides as far as quantities are concerned which do not commute with the additive conserved quantity k. The reason for our interest in such an information content is that the measurement of quantities which do not commute with a conserved additive quantity (such as energy) requires, according to the quantum theory of measurement, a macroscopic apparatus, whereas the quantities which commute with all conserved additive quantities can be measured with microscopic means (see **1**, **6**).

We begin with a few lemmas. Lemma 2 follows easily from a theorem of Lyapunov (**4**) and Lemma 3 can also be inferred from (**5**).

LEMMA 1. *If A is self-adjoint, B positive definite, and $AB + BA = 0$, then $A = 0$.*

The last statement is synonymous with the statement that all characteristic values of A are zero. Let us consider, therefore, a characteristic vector ϕ of A and denote its characteristic value by λ. We then calculate

$$(1) \qquad \begin{aligned} 0 &= (\phi, (AB + BA)\phi) = (A\phi, B\phi) + (\phi, BA\phi) \\ &= \lambda(\phi, B\phi) + \lambda(\phi, B\phi). \end{aligned}$$

Since $(\phi, B\phi)$ is positive, $\lambda = 0$ and hence $A = 0$.

LEMMA 2. *If B and $AB + BA = C$ are positive definite, A is also positive definite. If C is semidefinite, so is A.*

We first prove that A is self-adjoint. The adjoint of the equation

$$(2) \qquad AB + BA = C$$

is

$$(2a) \qquad BA^\dagger + A^\dagger B = C,$$

the dagger denoting hermitian adjoint. Subtracting (2a) from (2), one obtains

$$(3) \qquad (A - A^\dagger)B + B(A - A^\dagger) = 0.$$

Hence, Lemma 1 applied to B and the self-adjoint $i(A - A^\dagger)$, shows that $i(A - A^\dagger) = 0$, that is $A = A^\dagger$ is self-adjoint.

Let us consider, therefore, a characteristic vector ψ of A and denote its characteristic value by a. Since $AB + BA$ is positive definite,

$$(4) \qquad \begin{aligned} 0 &< (\psi, (AB + BA)\psi) = (\psi, AB\psi) + (\psi, BA\psi) \\ &= (A\psi, B\psi) + a(\psi, B\psi) = 2a(\psi, B\psi) \end{aligned}$$

so that a is positive since $(\psi, B\psi)$ is also positive.

If $AB + BA$ is semidefinite, A is also semidefinite. On the contrary, $AB + BA$ need not be definite if A and B are definite. In order to see this, let us consider first a positive, semidefinite B. This can be given the form

$$B = \begin{Vmatrix} B_{11} & 0 \\ 0 & 0 \end{Vmatrix}$$

in which B is composed of four submatrices and B_{11} is definite. Hence, if we write similarly

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

we obtain

$$AB + BA = \begin{Vmatrix} B_{11}A_{11} + A_{11}B_{11} & B_{11}A_{12} \\ A_{21}B_{11} & 0 \end{Vmatrix}.$$

Unless $A_{12} = A_{21} = 0$, this is surely indefinite, i.e., some of its characteristic values are negative. Now, B can be made positive by an arbitrarily small change. If this is small enough, some of the characteristic values of $AB + BA$ will still be negative.

LEMMA 3. *Again let B be definite. Then, for a given C there is one and only one A such that $AB + BA = C$. If C is self-adjoint, or definite, or semidefinite, so is A.*

In order to prove the existence of A, it seems best to use the principal axes of B as co-ordinate system, i.e., to assume $B_{ik} = b_i \delta_{ik}$ with $b_i > 0$. Then $AB + BA = C$ reads

(5) $$A_{ik} b_k + b_i A_{ik} = C_{ik}$$

or

(6) $$A_{ik} = C_{ik}/(b_i + b_k).$$

This proves the uniqueness of A and shows that it is self-adjoint if C is self-adjoint. The fact that A is definite and semidefinite for definite and semidefinite C is now a consequence of Lemma 2.

The preceding lemmas suggest defining a fraction of matrices C/B which satisfies the equation

(7) $$\frac{C}{B} B + B \frac{C}{B} = 2C.$$

One sees from (6) that C/B is uniquely defined if no characteristic value of B is zero and if no two characteristic values are oppositely equal. However, throughout the present paper, the denominators of fractions will be definite. According to (7), Jordan's quasi-product of B and C/B is just C so that the operation of division implied by (7) is the inverse of Jordan's quasi-multiplication **(3)**.

The fractions of matrices satisfy a few identities and we shall mention those which will be needed in the rest of the present note.

LEMMA 4. *If $C/B = A$, then $C/A = B$. Furthermore*

$$\frac{f(B)C\,g(B)}{B} = f(B)\,\frac{C}{B}\,g(B) \tag{8}$$

for arbitrary functions f and g of B. Finally, for arbitrary C_1, C_2

$$2\frac{C_1}{B}C_2 - 2C_1\frac{C_2}{B} = \left[\frac{C_1}{B}\frac{C_2}{B}, B\right]. \tag{9}$$

The first part of the lemma is a consequence of the fact that C/B and B enter symmetrically into the defining equation (7). Equation (8) follows from (7) by multiplying it on the left by $f(B)$ and on the right by $g(B)$ and noting that B commutes with $f(B)$ and $g(B)$. Hence

$$f(B)\frac{C}{B}g(B)\cdot B + B\cdot f(B)\frac{C}{B}g(B) = 2f(B)Cg(B).$$

This is equivalent to (8). In order to prove (9), we transform the commutator on the right side

$$\begin{aligned}\left[\frac{C_1}{B}\frac{C_2}{B}, B\right] &= \frac{C_1}{B}\cdot\frac{C_2}{B}B - B\frac{C_1}{B}\cdot\frac{C_2}{B} \\ &= \frac{C_1}{B}\left(2C_2 - B\frac{C_2}{B}\right) - \left(2C_1 - \frac{C_1}{B}B\right)\frac{C_2}{B} \\ &= 2\frac{C_1}{B}C_2 - 2C_1\frac{C_2}{B}.\end{aligned} \tag{9a}$$

Actually, (8) will be used only for definite B and $f(B) = g(B) = B^{\frac{1}{2}}$, whence

$$\frac{B^{\frac{1}{2}}CB^{\frac{1}{2}}}{B} = B^{\frac{1}{2}}\frac{C}{B}B^{\frac{1}{2}}, \tag{8b}$$

where $B^{\frac{1}{2}}$ is to be understood as the positive definite square root of B. Equation (9) will be used only in the form

$$\text{Trace}\left(\frac{C_1}{B}C_2 - C_1\frac{C_2}{B}\right) = 0. \tag{9b}$$

This is an immediate consequence of (9) because the trace of every commutator vanishes

$$\text{Trace}[A, B] = \text{Trace } AB - \text{Trace } BA = \text{Trace } AB - \text{Trace } AB,$$

the trace of a product of two factors being independent of the order of the factors.

LEMMA 5. *If A and B are positive definite (semidefinite), so is $A \times B$, where the cross denotes the direct product.*

This is obvious since the characteristic values of $A \times B$ are the products $a_i b_k$ of the characteristic values of A and B.

LEMMA 6. *If A and B are definite and $A^2 > B^2$* (i.e., $A^2 - B^2$ is also positive definite), *then $A > B$.*

In fact,

$$A^2 - B^2 = \tfrac{1}{2}(A + B)(A - B) + \tfrac{1}{2}(A - B)(A + B). \tag{10}$$

Since $A + B$ is definite, the positive nature of $A - B$ is a consequence of Lemma 2. It follows also from Löwner's theorem. If $A^2 - B^2$ is semidefinite, A and B still being definite, $A - B$ is semidefinite.

On the contrary, if A, B, and $A - B$ are definite, $A^2 - B^2$ need not be positive definite.

LEMMA 7. *If $A > A'$, $B > B'$, and $A + A'$ as well as $B + B'$ are definite, then $A \times B > A' \times B'$.*

This follows from the identity

$$A \times B - A' \times B' = \tfrac{1}{2}(A - A') \times (B + B') + \tfrac{1}{2}(A + A') \times (B - B'). \tag{11}$$

The right side is the sum of two matrices, both of which are definite, because of Lemma 5. Hence, $A \times B - A' \times B'$ is also positive, $A \times B > A' \times B'$.

The preceding lemmas were obtained in order to prove the following theorem.

THEOREM 1. *If S is positive definite, N self-adjoint, $ST + TS = N^2$, then*

$$Q = S \times \bar{T} + T \times \bar{S} - N \times \bar{N} \tag{12}$$

is positive semidefinite.

The bar in (12) denotes complex conjugation, the cross the Kronecker (direct) product. It follows, first, from Lemma 2 and the semidefinite nature of N^2, that $T = \frac{1}{2}N^2/S$ is also semidefinite so that S and T enter our theorem almost interchangeably. The rows and columns of Q can suitably be denoted by double indices, such as $\kappa\kappa'$, $\lambda\lambda'$, ..., the first of which denotes a row or column of the first factor in the Kronecker product, the second a row or column of the second factor. Hence, (12) in terms of its matrix elements reads

$$Q_{\kappa\kappa',\lambda\lambda'} = S_{\kappa\lambda} \bar{T}_{\kappa'\lambda'} + T_{\kappa\lambda} \bar{S}_{\kappa'\lambda'} - N_{\kappa\lambda} \bar{N}_{\kappa'\lambda'}. \tag{12a}$$

Similarly, the vectors to which Q can be applied will have double indices. If ψ is such a vector with components $\psi_{\lambda\lambda'}$, these can be considered to form a matrix $M(\psi)$. This will be called the matrix of the vector ψ. One then easily verifies that the matrix of $(A \times \bar{B})\psi$ is

$$M((A \times \bar{B})\psi) = AM(\psi)B^\dagger. \tag{13}$$

Hence, (12) gives

$$M(Q\psi) = SM(\psi)T + TM(\psi)S - NM(\psi)N. \tag{13a}$$

The daggers are omitted because S, T, and N are all self-adjoint. The quadratic form $(\psi, Q\psi)$ becomes, if one sets, for brevity, $M(\psi) = P$ but substitutes $T = \frac{1}{2}N^2/S$,

$$(\psi, Q\psi) = \text{Trace}\left(P^\dagger SP\frac{N^2}{2S} + P^\dagger\frac{N^2}{2S}PS - P^\dagger NPN\right). \tag{14}$$

This is, evidently, a hermitian quadratic form of the elements of P and will become a similar form of the components of N if we replace one of the factors N in every term by $N^\dagger$. This is permissible since N is self-adjoint. We define, therefore, the bihermitian form

$$q_S(N, P) = (\psi, Q\psi) = \text{Trace}\left(P^\dagger SP\frac{N^\dagger N}{2S} + P^\dagger\frac{NN^\dagger}{2S}PS - \frac{1}{2}P^\dagger N^\dagger PN - \frac{1}{2}P^\dagger NPN^\dagger\right). \tag{14a}$$

This is defined no matter whether N is self-adjoint or not. The index S on q is to remind us that the coefficients of the bihermitian form depend on S; this is assumed to be definite and hence self-adjoint.

We shall now show that q_S is invariant under a transformation

$$q_S(N, P) = q_S(S^{\frac{1}{2}}PS^{\frac{1}{2}}, S^{-\frac{1}{2}}NS^{-\frac{1}{2}}). \tag{15}$$

This equation plays a decisive role in the proof of Theorem 1. Since S is definite, its square roots are hermitian, and $S^{\frac{1}{2}}$ will denote the definite square root of S. Similarly $S^{-\frac{1}{2}} = (S^{\frac{1}{2}})^{-1}$ will be definite. We shall show that the first term of $q_S(S^{\frac{1}{2}}PS^{\frac{1}{2}}, S^{-\frac{1}{2}}NS^{-\frac{1}{2}})$ is equal to the first term of $q_S(N, P)$. The same applies to the second term but the third term of $q_S(S^{\frac{1}{2}}PS^{\frac{1}{2}}, S^{-\frac{1}{2}}NS^{-\frac{1}{2}})$ is equal to the fourth term of $q_S(N, P)$, and conversely. However, we shall show only the invariance of the first term under the transformation (15).

The first term of (14a), with N replaced by $S^{\frac{1}{2}}PS^{\frac{1}{2}}$ and P replaced by $S^{-\frac{1}{2}}NS^{-\frac{1}{2}}$, is

$$\text{Trace}\left(S^{-\frac{1}{2}}N^\dagger S^{-\frac{1}{2}}SS^{-\frac{1}{2}}NS^{-\frac{1}{2}}\frac{S^{\frac{1}{2}}P^\dagger S^{\frac{1}{2}}S^{\frac{1}{2}}PS^{\frac{1}{2}}}{2S}\right)$$

and, on account of (8b), this is equal to

$$\text{Trace}\left(S^{-\frac{1}{2}}N^\dagger NS^{-\frac{1}{2}}\cdot S^{\frac{1}{2}}\frac{P^\dagger SP}{2S}S^{\frac{1}{2}}\right) = \text{Trace}\left(N^\dagger N\frac{P^\dagger SP}{2S}\right).$$

Further, because of (9b), this is equal to

$$\text{Trace}\left(\frac{N^\dagger N}{2S}P^\dagger SP\right) = \text{Trace}\left(P^\dagger SP\frac{N^\dagger N}{2S}\right),$$

i.e., equal to the first term of $q_S(N, P)$. The other relations, which were mentioned before, can be proved in the same way so that (15) is established.

We shall next prove the following lemma.

LEMMA 8. *The characteristic vectors of Q can be assumed to have such a form that the corresponding matrix C is self-adjoint.*

In fact, it follows from $Q\psi = q\psi$ with $M(\psi) = C$ and (13*a*) that

$$SCT + TCS - NCN = qC. \tag{16}$$

The hermitian adjoint of (16) is, since q is real,

$$TC^{\dagger}S + SC^{\dagger}T - NC^{\dagger}N = qC^{\dagger}. \tag{16a}$$

Taking the sum and difference of (16) and (16*a*), one sees that $C + C^{\dagger}$ and $i(C - C^{\dagger})$ also correspond to characteristic vectors of Q, and their characteristic value is q. If $C \neq 0$, at least one of these does not vanish and this, or the two of them, can replace C.

LEMMA 9. $C = 1$ *is the matrix of a characteristic vector of Q; its characteristic value is* 0.

This follows immediately from (16) and $ST + TS = N^2$. The characteristic vector itself is

$$\psi_{\kappa\lambda}{}^0 = \delta_{\kappa\lambda}. \tag{17}$$

Since all characteristic vectors which belong to another characteristic value of Q are orthogonal to this one, we have the following lemma.

LEMMA 10. *The trace of each matrix C of the characteristic vector of a non-zero characteristic value of Q vanishes; no such matrix can be definite or semidefinite.*

The last conclusion is obvious since the sum of the characteristic values of C is zero and they cannot all vanish. Hence, at least one of them must be negative.

If a quadratic form such as (14*a*) can become negative for any N and P, it will be negative also if these are normalized. We can confine our attention, therefore, to the value of $q_S(N, P)$ for normalized N and P

$$\sum_{\kappa\lambda} |N_{\kappa\lambda}|^2 = \text{Trace}\, NN^{\dagger} = 1, \tag{18a}$$

$$\sum_{\kappa\lambda} |\psi_{\kappa\lambda}|^2 = \text{Trace}\, PP^{\dagger} = 1. \tag{18b}$$

We next note that $q_S(N, P)$ is defined in a closed bounded domain of its arguments $N_{\kappa\lambda}$, $P_{\kappa\lambda}$ if these are restricted by (18*a*), (18*b*). Hence, it will assume its minimum value for some values of these arguments and our task is just to prove that it leads to a contradiction to assume that this minimum is negative. We can even further stipulate that both N and P are self-adjoint—N is naturally so, and the expression $q_S(N, P) = (\psi, Q\psi)$ will assume its minimum value for given N for a ψ which is a characteristic vector of Q. Lemma 8 shows, however, that the matrix of any characteristic vector of Q can be assumed to be self-adjoint. We note incidentally that if N and P are self-adjoint, the variables of q_S on the right side of (15) are also self-adjoint.

The form $q_S(N, P)$ is surely not increased if N is replaced by the positive definite or semidefinite square root N_p of its square. Such a substitution does not affect (18a) because this merely fixes the sum of the squares of the characteristic values of N and the transition from N to N_p merely changes the sign of the negative characteristic values of N. Further, the T of (12) is completely determined by S and $N^2 = N_p^2$ so that the first two terms of Q are not affected by the change. It follows, however, from Lemma 7 that $N_p \times N_p > N \times N$ since $N_p > N$, as one can see best in the diagonal form of these matrices. It follows that $q_S(N, P)$ assumes its minimum for a definite or semidefinite N. It may assume the same value, of course, for other N as well.

It now follows from (15) that $q_S(S^{\frac{1}{2}}PS^{\frac{1}{2}}, S^{-\frac{1}{2}}NS^{-\frac{1}{2}})$ also assumes its minimum for some semidefinite N. Writing $N' = S^{\frac{1}{2}}PS^{\frac{1}{2}}$ and $P' = S^{-\frac{1}{2}}NS^{-\frac{1}{2}}$, the latter becomes semidefinite and we find that $q_S(N', P')$ assumes its minimum value for a semidefinite P' if the normalization of N' and P' now is, instead of (18a) and (18b),

(19a) $$\text{Trace}(S^{-\frac{1}{2}}N'S^{-\frac{1}{2}})^2 = 1,$$

(19b) $$\text{Trace}(S^{\frac{1}{2}}P'S^{\frac{1}{2}})^2 = 1.$$

It will be demonstrated now that if the minimum of $q_S(N', P')$ is negative, this minimum cannot be assumed for a definite or semidefinite P', the normalization being given by (19). In fact, the P' which makes $q_S(N', P')$ a minimum cannot be semidefinite if the minimum is negative, no matter what N' is. This would be an immediate consequence of Lemma 10 if the normalization were given by Trace $P'P'^{\dagger} = 1$, the analogue of (18b). In this case P' would be the matrix which corresponds to a characteristic vector with a non-zero characteristic value of a Q. Lemma 10 shows, however, that such a matrix cannot be semidefinite. Since the normalization of P' is actually given by

(20) $$\sum_{\kappa\lambda} |P'_{\kappa\lambda}|^2 s_\kappa s_\lambda = 1,$$

we shall have to prove that the P' which makes $q_S(N', P')$ a negative minimum cannot be semidefinite with the normalization (20) either. For (20), S was assumed to be diagonal, and its diagonal elements were denoted by s. These are, by the assumption of Theorem 1, positive.

Instead of normalizing P' by (20), we can use an unnormalized P' but use for q_S the expression

(21) $$q_S(N', P') = \frac{(\psi', Q'\psi')}{\sum |\psi'_{\kappa\lambda}|^2 s_\kappa s_\lambda}.$$

Q' is the Q in which N is replaced by N', the ψ' is the vector to which P' corresponds, i.e., $\psi'_{\kappa\lambda} = P'_{\kappa\lambda}$. In order to calculate $q_S(N', P')$, we decompose ψ' into two parts; one of these is a multiple $x\,\psi^0$ of the characteristic vector ψ^0 of (17), the other, ψ^1, a linear combination of the other characteristic vectors of Q'.

(22) $$\psi' = \psi^1 + x\psi^0 \quad \text{or} \quad \psi'_{\kappa\lambda} = \psi^1_{\kappa\lambda} + x\delta_{\kappa\lambda}.$$

Since $Q'\psi^0 = 0$, the expression

$$(\psi', Q'\psi') = (\psi^1 + x\psi^0, Q'(\psi^1 + x\psi^0)) = (\psi^1 + x\psi^0, Q'\psi^1)$$
$$= (\psi^1, Q'\psi^1) + \bar{x}(Q'\psi^0, \psi^1) = (\psi^1, Q'\psi^1)$$

is independent of x. If it is negative, q_S will assume its smallest value for the x which makes the denominator of (21) smallest. Hence, we calculate

$$\text{(24)} \quad \sum_{\kappa\lambda} |\psi'_{\kappa\lambda}|^2 s_\kappa s_\lambda = \sum_{\kappa\lambda} [|\psi^1_{\kappa\lambda}|^2 + \delta_{\kappa\lambda}(\bar{x}\psi^1_{\kappa\lambda} + x\bar{\psi}^1_{\kappa\lambda}) + |x^2|\delta_{\kappa\lambda}]s_\kappa s_\lambda$$
$$= \sum_{\kappa\lambda} |\psi^1_{\kappa\lambda}|^2 s_\kappa s_\lambda + \sum_{\kappa} ((x + \bar{x})\psi^1_{\kappa\kappa} + |x^2|)s_\kappa^2.$$

This assumes its minimum for

$$\text{(24a)} \qquad x = -\frac{\sum_\kappa \psi^1_{\kappa\kappa} s_\kappa^2}{\sum_\kappa s_\kappa^2}.$$

Hence, the matrix element $P'_{\mu\mu}$ becomes

$$\text{(25)} \qquad P'_{\mu\mu} = \psi'_{\mu\mu} = \psi^1_{\mu\mu} + x = \frac{\sum_\kappa (\psi^1_{\mu\mu} - \psi^1_{\kappa\kappa})s_\kappa^2}{\sum_\kappa s_\kappa^2}.$$

For the μ for which $\psi^1{}_{\mu\mu}$ is the smallest of all, $\psi'{}_{\kappa\kappa}$, this is surely negative so that P' cannot be semidefinite. This argument is valid unless all $\psi^1{}_{\kappa\kappa}$ are equal. In this case, however, all $P'_{\mu\mu} = 0$ and P' can be semidefinite only if all its elements are zero—a case which is incompatible with the normalization (20). We see, therefore, that the assumption $q_S(N', P') = q_S(N, P) < 0$ leads to a contradiction so that our theorem is established.

It was mentioned that S and T play almost equivalent roles in Theorem 1. The difference is that whereas S was assumed to be definite, only the semidefinite nature of T follows from $ST + TS = N^2$. We state, therefore, the following theorem.

THEOREM 2. *Let S and T be semidefinite, N self-adjoint, $ST + TS = N^2$; then $S \times \bar{T} + T \times \bar{S} - N \times \bar{N}$ is also semidefinite.*

Note that even though N^2 is naturally semidefinite, since S is only semidefinite, the semidefinite nature of T does not follow from $ST + TS = N^2$ and had to be postulated.

In order to establish Theorem 2, it is simplest to introduce again the characteristic vectors of S as co-ordinate axes. S, T, and N then will be written as supermatrices

$$\text{(26)} \qquad S = \begin{Vmatrix} 0 & 0 \\ 0 & S_{22} \end{Vmatrix}, \qquad T = \begin{Vmatrix} T_{11} & T_{12} \\ T^\dagger_{12} & T_{22} \end{Vmatrix}, \qquad N = \begin{Vmatrix} N_{11} & N_{12} \\ N^\dagger_{12} & N_{22} \end{Vmatrix}.$$

S_{22} is a diagonal matrix with real positive diagonal elements, T_{11}, T_{22}, N_{11}, N_{22} are self-adjoint. The equation $ST + TS = N^2$ shows that the submatrix of

N^2 in the upper left corner vanishes. However, the vanishing of a diagonal element of the square of a self-adjoint matrix entails the vanishing of the corresponding row and column of the matrix itself. Hence, $N_{11} = N_{12} = 0$ and it now follows from $ST + TS = N^2$ that $T_{12} = 0$ also:

$$T = \begin{Vmatrix} T_{11} & 0 \\ 0 & T_{22} \end{Vmatrix}, \qquad N = \begin{Vmatrix} 0 & 0 \\ 0 & N_{22} \end{Vmatrix}. \tag{27}$$

It now follows from the semidefinite nature of T that T_{11} and T_{22} are also semidefinite.

The subdivision of S, T, and N into submatrices corresponds to a decomposition of the underlying vector space into two orthogonal parts and this leads to a decomposition of the product space of this space with itself into four parts. Since the non-diagonal submatrices of S, T, and N all vanish, all four of these parts are invariant under any direct product of the three matrices. All vectors in the first of these spaces are annihilated by Q, i.e., $Q_{11;11} = 0$. In the second space $Q_{12;12} = T_{11} \times \bar{S}_{22}$ and this is the direct product of a semidefinite and a definite matrix and, hence, semidefinite by Lemma 5. The same applies in the third space $Q_{21;21} = S_{22} \times T_{11}$. The effect of Q in the last space is

$$Q_{22;22} = S_{22} \times \bar{T}_{22} + T_{22} \times \bar{S}_{22} - N_{22} \times \bar{N}_{22} \tag{28}$$

and the semidefinite nature of this is a consequence of Theorem 1. This then proves Theorem 2.

We wish to express our appreciation to Drs. S. Bochner and D. Lowdenslager who followed our efforts to prove the theorem of this paper with interest and understanding. One of the authors (M.M.Y.) is greatly indebted to Professor J. R. Oppenheimer and the members of the Institute for Advanced Study for the kind hospitality extended to him.

References

1. H. Araki and M. M. Yanase, Phys. Rev., *120* (1960), 662; also E. P. Wigner, Z. Physik, *133* (1952), 101.
2. K. Löwner, Math. Z., *38* (1934), 177.
3. P. Jordan, Nachr. Gött. Ges. Wiss. (1932), 569; (1933), 209.
4. F. R. Gantmacher, *Applications of the theory of matrices* (New York, 1959), p. 224, or the particularly adapted form in Richard Bellman's *Introduction to matrix analysis* (New York, 1960), p. 244, and O. Taussky, J. Soc. Indust. Appl. Math., *9* (1961), p. 640.
5. One of the writers (E.P.W.) wishes to use this opportunity to point out that Theorem I of his article *On weakly positive matrices* (Can. J. Math., *15* (1963), 313) is partially contained in O. Taussky's problem 4846 published in Am. Math. Monthly, *66* (1959), 429.
6. E. P. Wigner, Article in Physikertagung Wien 1961 (Physik Verlag, Mosbach, Baden, 1962), p. 1.
7. Eugene P. Wigner and Mutsuo M. Yanase, *Information contents of distributions*, Proc. Natl. Acad. Sci., *49* (1963), 910.

Princeton University and
Institute for Advanced Study, Princeton