THEOREMS ON NORMAL MATRICES

By H. SCHNEIDER (*Edinburgh*)

[Received 12 September 1951; in revised form 10 December 1951]

1. As a corollary to general considerations on matrices A, B whose commutator $AB-BA$ is nilpotent, Drazin, Dungey, and Gruenberg (**2**) have recently obtained this theorem on normal matrices:† '*Let A and B be normal matrices whose latent roots are α_i, β_i ($i = 1,\ldots, n$). Then A and B commute if and only if the latent roots of every scalar polynomial $f(A, B)$ are $f(\alpha_i, \beta_i)$ ($i = 1,\ldots, n$) for some ordering α_i, β_i.*'

Using a different approach I shall show that for the commutativity of the normal matrices A and B it is sufficient to make apparently weaker assumptions. In these $f(A, B)$ will be considerably specialized and a condition, implied by the theorem quoted, will be imposed on the sum of squared moduli of latent roots only.

Every square matrix A has a 'polar representation' $A = HU$, where H is non-negative definite Hermitian, U is unitary [Wintner and Murnaghan (**7**), Williamson (**6**)]. The matrices H and U commute if and only if A is normal. If A, B, and AB are normal, it is known that the matrices of the polar representation of A and B commute in pairs, except possibly the unitary pair, provided that the latter are properly chosen.‡ This suggested that, if polar representations were considered, a similar theorem might be obtained for normal, but not necessarily commutative, matrices, whose product is also normal.

To avoid repetition in the statement of the theorems I shall assume throughout that A and B are normal $n\times n$ matrices with complex elements and latent roots α_i, β_i ($i = 1,\ldots, n$) respectively. Also that

$$A = HU = UH \quad \text{and} \quad B = KV = VK$$

are the polar representations of A and B, where H and K are non-negative definite Hermitian, U and V unitary.

† A review of several important properties of normal matrices may be found in Drazin (**1**). I shall generally use the results of that paper without giving further references.

‡ Wiegmann (**5**). The unitary polar matrix of a singular matrix is not unique. The proviso is necessary when the rank of either A or B is less than or equal to $n-2$.

Quart. J. Math. Oxford (2), 3 (1952), 241–9.

3695.2.3 R

2. My results rest on a property of matrices due to Schur (4):

Let C be a matrix with latent roots γ_i $(i = 1, \ldots, n)$; then

$$\sum_{i=1}^{n} |\gamma_i|^2 \leqslant \sum_{i,j=1}^{n} |c_{ij}|^2,$$

and

$$\sum_{i=1}^{n} |\gamma_i|^2 = \sum_{i,j=1}^{n} |c_{ij}|^2$$

if and only if C is normal.

I shall first prove a lemma.

LEMMA. *Let C be a normal $n \times n$ matrix, whose latent roots are γ_{i_j} $(j = 1, \ldots, n_i;\ i = 1, \ldots, k)$, where $\sum_{i=1}^{k} n_i = n$. Let C be partitioned so that the k matrices C_{ii} in the diagonal are square. If the sum of squared moduli of latent roots of C_{ii} is $\sum_{j=1}^{n_i} |\gamma_{i_j}|^2$, for all i, then C is the direct sum*

$$C = C_{11} \dot{+} C_{22} \dot{+} \ldots \dot{+} C_{kk}.$$

Proof. Let S_i be the sum of squared moduli of elements of C_{ii} and let

$$S = \sum_{i,j=1}^{n} |c_{ij}|^2.$$

Using Schur's theorem for C_{ii} we have

$$\sum_{j=1}^{n} |\gamma_{i_j}|^2 \leqslant S_i \quad (i = 1, \ldots, k).$$

Since C is normal,

$$S = \sum_{i=1}^{k} \sum_{j=1}^{n_i} |\gamma_{i_j}|^2 \leqslant \sum_{i=1}^{k} S_i.$$

But $S > \sum_{i=1}^{k} S_i$; unless $C_{ij} = 0$ $(i \neq j)$, when $S = \sum_{i=1}^{k} S_i$. It follows that $C_{ij} = 0$ $(i \neq j)$ and therefore that

$$C = C_{11} \dot{+} C_{22} \dot{+} \ldots \dot{+} C_{kk}.$$

It will be convenient to denote the sum of squared moduli of latent roots of a matrix A by $\Lambda(A)$.

THEOREM 1. *The matrices A and B commute if and only if the sum of squared moduli of latent roots of every scalar polynomial $f(A)B$ is*

$$\sum_{i=1}^{n} |f(\alpha_i)\beta_i|^2$$

for some ordering α_i, β_i $(i = 1, \ldots, n)$.

Proof. By a well-known theorem of Frobenius there is an ordering for which the latent roots of every $f(A)B$ are $f(\alpha_i)\beta_i$, provided that A and B commute. This ordering obviously satisfies the theorem.

Now suppose there is an ordering for which

$$\Lambda\{f(A)B\} = \sum_{i=1}^{n} |f(\alpha_i)\beta_i|^2,$$

for every $f(A)B$.

Since A is normal, there is a unitary matrix Φ such that

$$\bar{\Phi}' A \Phi = D = \mathrm{diag}[\alpha_1, \alpha_2, \ldots, \alpha_n],$$

where equal α_i are arranged consecutively. Thus

$$D = \alpha_1' I_1 \dotplus \alpha_2' I_2 \dotplus \ldots \dotplus \alpha_k' I_k,$$

where $\alpha_i' \neq \alpha_j'$ when $i \neq j$.

Suppose that there are n_i latent roots of A equal to α_i and denote these now by α_{ij} $(j = 1, \ldots, n_i)$. Denote the corresponding roots of B in the enunciation of the theorem by β_{ij}. Let $\bar{\Phi}' B \Phi = C$, and let C be partitioned conformably with D.

We may construct the polynomials $f_i(x)$ $(i = 1, \ldots, k)$ for which

$$f_i(\alpha_i') = 1, \qquad f_i(\alpha_j') = 0, \quad \text{when } j \neq i.$$

Then $f_i(D)C = \bar{\Phi}' f_i(A) B \Phi$, and so, by hypothesis and by the definition of $f_i(x)$, we have

$$\Lambda\{f_i(D)C\} = \Lambda\{f_i(A)B\} = \sum_{h=1}^{k} |f_i(\alpha_h')|^2 \Big(\sum_{j=1}^{n_h} |\beta_{hj}|^2 \Big) = \sum_{j=1}^{n_i} |\beta_{ij}|^2. \tag{1}$$

Also $$f_i(D) = O_1 \dotplus \ldots \dotplus O_{i-1} \dotplus I_i \dotplus O_{i+1} \dotplus \ldots \dotplus O_k,$$

where the O_i are null-matrices, and the I_i are unit-matrices, of order n_i. Hence, if $G_i \equiv f_i(D)C$ is partitioned conformably with D, then $G_{i,kj} = 0$, when $k \neq i$, while $G_{i,ii} = C_{ii}$. It follows that the latent roots of $f_i(D)C$ are those of C_{ii} together with zeros. By (1),

$$\Lambda(C_{ii}) = \Lambda\{f_i(D)C\} = \sum_{j=1}^{n_i} |\beta_{ij}|^2.$$

But the latent roots of $C = \bar{\Phi}' B \Phi$ are β_{ij} $(j = 1, \ldots, n_i;\ i = 1, \ldots, k)$. The lemma applies and we deduce that

$$C = C_{11} \dotplus C_{22} \dotplus \ldots \dotplus C_{kk}.$$

This is sufficient for D and C to commute. Transforming by Φ, we see that A and B commute.

Suppose that α_{i_j}, β_{i_j} and α_{i_j}, β'_{i_j} are orderings of the latent roots of A and B satisfying Frobenius's theorem. The latent roots of $f_i(A)B$ are β_{i_j} or β'_{i_j} ($j = 1,\dots, n_i$), together with $n-n_i$ zeros. Hence in Frobenius's theorem the latent roots of B associated with a set of equal latent roots of A form a unique set. This is not true in the case of our theorem. For suppose that α_{i_j}, β_{i_j} and α_{i_j}, β'_{i_j} satisfy

$$\Lambda\{f(A)B\} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}|f(\alpha_{i_j})\beta_{i_j}|^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}|f(\alpha_{i_j})\beta'_{i_j}|^2,$$

for all polynomials $f(A)B$. Then we must have

$$\sum_{j=1}^{n_i}|\beta_{i_j}|^2 = \sum_{j=1}^{n_i}|\beta'_{i_j}|^2 \quad (i = 1,\dots, k), \tag{2}$$

by putting $f(A)B = f_i(A)B$ ($i = 1,\dots, k$). Conversely, if α_{i_j}, β_{i_j} is an ordering satisfying Theorem 1, and (2) is satisfied for some ordering α_{i_j}, β'_{i_j} of the latent roots of B, then

$$\Lambda\{f(A)B\} = \sum_{i=1}^{k}\sum_{j=1}^{n_i}|f(\alpha_{i_j})\beta_{i_j}|^2 = \sum_{i=1}^{k}|f(\alpha'_i)|^2\Big(\sum_{j=1}^{n_i}|\beta_{i_j}|^2\Big)$$
$$= \sum_{i=1}^{k}|f(\alpha'_i)|^2\Big(\sum_{j=1}^{n_i}|\beta'_{i_j}|^2\Big) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}|f(\alpha_{i_j})\beta'_{i_j}|^2,$$

and therefore the ordering α_{i_j}, β'_{i_j} also satisfies the theorem. I state these results in a corollary.

COROLLARY. *Let the latent roots of A be α_{i_j} ($j = 1,\dots, n_i$; $i = 1,\dots, k$), where $\alpha_{i_j} = \alpha'_i$ and $\alpha'_i \neq \alpha'_j$ when $i \neq j$; and let the latent roots of B be β_{i_j} or β'_{i_j} ($j = 1,\dots, n_i$; $i = 1,\dots, k$). Let A and B commute. Then, if α_{i_j}, β_{i_j} is an ordering satisfying Theorem 1, α_{i_j}, β'_{i_j} is another if and only if*

$$\sum_{j=1}^{n_i}|\beta_{i_j}|^2 = \sum_{j=1}^{n_i}|\beta'_{i_j}|^2 \quad (i = 1,\dots, k).$$

It is easily seen from an example that there may be various groupings of the β_{i_j} into k sets of n_i members each, the sum of squared moduli of the ith set being $\sum_{j=1}^{n_i}|\beta_{i_j}|^2$.

3. I shall next prove a theorem concerning a unitary matrix B. It has no analogue for general normal B.

THEOREM 2. *Let $B = V$ be unitary. Then AV is normal if and only if the sum of squared moduli of latent roots of AV is $\sum_{i=1}^{n}|\alpha_i|^2$.*

Proof. It is well known that multiplication by a unitary matrix does not alter the sum of squared moduli of elements of a matrix.

Let $C = AV$, and let the latent roots of C be γ_i ($i = 1,..., n$). Then

$$S = \sum_{i,j=1}^{n} |c_{ij}|^2 = \sum_{i,j=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{n} |\alpha_i|^2,$$

since A is normal. By Schur's theorem C is normal if and only if

$$\sum_{i=1}^{n} |\gamma_i|^2 = S,$$

and the theorem follows from $S = \sum_{i=1}^{n} |\alpha_i|^2$.

When AV is normal the polar matrices of A and those of AV commute. The latent roots of the unitary matrices V and UV are of unit modulus. Denote the latent roots of H, U, UV, by h_i, u_i, w_i ($i = 1,..., n$) respectively. If $C = AV$, then $C = H.UV$ is a polar representation of C, and H and UV commute. By Frobenius's theorem,

$$h_i u_i = \alpha_i,$$

and $$h_i w_i = \gamma_i;$$

whence $$|\gamma_i| = |\alpha_i|.$$

On using Theorem 2 we obtain the non-trivial part of Corollary 1.

COROLLARY 1. *The sum of squared moduli of latent roots of AV is $\sum_{i=1}^{n} |\alpha_i|^2$ if and only if the moduli of latent roots of AV are $|\alpha_i|$ ($i = 1,..., n$).*

Next let H be a non-negative definite Hermitian matrix, with latent roots h_i ($i = 1,..., n$). It is known that $h_i \geqslant 0$. By Theorem 2,

$$\sum_{i=1}^{n} h_i^2 = \sum_{i=1}^{n} |\gamma_i|^2$$

if and only if $C = HV$ is normal. But C is normal if and only if $HV = VH$. Corollary 2 [cf. Parker (3)] now follows.

COROLLARY 2. *Let H be a non-negative definite Hermitian, and V be unitary. Then H and V commute if and only if the sum of squared moduli of latent roots of H is equal to the sum of squared moduli of latent roots of HV.*

It will be seen that Theorem 1 is a generalization of this corollary.

4. In Theorems 4 and 5 we shall be considering polynomials $f(A)$ whose polar representation is $f(A) = g(H)W$. The significance of this restriction is brought out by Theorem 3.

THEOREM 3. *The polar representation*† *of* $f(A)$ *is* $f(A) = g(H)W$, *where* $g(H)$ *is non-negative definite Hermitian and* W *is unitary, if and only if*
$$|f(\alpha_i)| = |f(\alpha_j)| = g(|\alpha_i|), \tag{3}$$
when $|\alpha_i| = |\alpha_j|$.

Proof. Let
$$D = \overline{\Phi}' A \Phi = \text{diag}[\alpha_1,..., \alpha_n].$$
Then
$$D = \overline{\Phi}' H \Phi . \overline{\Phi}' U \Phi$$
is a polar representation of D. If
$$R = \text{diag}[|\alpha_1|,..., |\alpha_n|],$$
there is a diagonal matrix $Z = \text{diag}[z_1,..., z_n]$, $|z_i| = 1$, for which $D = RZ$ is also a polar representation of D. The Hermitian polar matrix is unique, and therefore
$$R = \overline{\Phi}' H \Phi.$$
Suppose that
$$f(A) = g(H)W.$$
Then
$$f(D) = \overline{\Phi}' f(A) \Phi = \overline{\Phi}' g(H) \Phi . \overline{\Phi}' W \Phi = g(R) . \overline{\Phi}' W \Phi.$$
Also
$$f(D) = \text{diag}[f(\alpha_1),..., f(\alpha_n)], \tag{4}$$
and
$$g(R) = \text{diag}[g(|\alpha_1|),..., g(|\alpha_n|)], \tag{5}$$
while, by an argument similar to that for D and R above, it follows from
$$f(D) = g(R) . \overline{\Phi}' W \Phi$$
that
$$g(R) = \text{diag}[|f(\alpha_1)|,..., |f(\alpha_n)|]. \tag{6}$$
By comparing (5) and (6), we deduce (3).

Conversely, let us assume (3). We obtain (4) and (5), by the definition of D and R. Then (6) follows by (3). Thus $f(D) = g(R)X = X g(R)$ is a polar representation, where $X = \text{diag}[x_1,..., x_n]$, $|x_i| = 1$. Hence
$$f(A) = \Phi f(D) \overline{\Phi}' = \Phi g(R) X \overline{\Phi}' = g(H) W,$$
for
$$W = \Phi X \overline{\Phi}',$$
and the theorem is proved.

I shall use this theorem due to Wiegmann (5): '*Let* A *and* B *be normal. Then* AB *is normal if and only if* H *and* B, A *and* V, *commute.*' A version of this theorem has already been quoted in the introductory remarks.

THEOREM 4. *Let* A, B *and* AB *be normal. Then, if the Hermitian polar matrices of* $f(A)$, $l(B)$ *are functions of the Hermitian polar matrices of* A, B *respectively,* $f(A)l(B)$ *is normal.*

† Since the unitary matrix W is here left unrestricted we may say '*the* polar representation' meaning '*all* polar representations'.

Proof. Let $f(A) = g(H)W, \qquad l(B) = m(K)X.$

By Wiegmann's theorem, H and B, A and K commute. Hence $g(H)$ and $l(B)$, $f(A)$ and $m(K)$ commute. Using Wiegmann's theorem in the reverse direction, we see that $f(A)l(B)$ is normal.

If $f(A)l(B)$ is normal, it is easily seen that the Hermitian polar matrices of $f(A)$, $l(B)$ may not be functions of H and K.

THEOREM 5. *If $f(A)l(B)$ is normal, the moduli of its latent roots are $|f(\alpha_i)l(\beta_i)|$ $(i = 1,..., n)$, for some ordering α_i, β_i.*

Proof. Let $f(A) = GW$, $l(B) = MX$ be polar representations; $f(A)$ and $l(B)$ are normal. As in Theorem 3, we may prove that the latent roots of G, M are $|f(\alpha_i)|$, $|l(\beta_i)|$ $(i = 1,..., n)$ respectively. Now $f(A)l(B)$ is normal and therefore G and M, W and M, G and X commute, provided that W and X are suitably chosen, by the version of Wiegmann's theorem quoted in the introduction. Hence†

$$f(A)l(B) = GWMX = GMWX = WXGM.$$

By Frobenius's theorem the latent roots of GM are $|f(\alpha_i)l(\beta_i)|$ for some ordering, and the latent roots of $f(A)l(B)$ are $|f(\alpha_i)l(\beta_i)|y_i$, where the y_i $(|y_i| = 1)$ are the latent roots of WX. The theorem follows.

5. The matrix AB is normal when $AB = BA$. Hence, by Theorem 1, if

$$\Lambda\{f(A)B\} = \sum_{i=1}^{n} |f(\alpha_i)\beta_i|^2$$

for every $f(A)$, then AB is certainly normal. In Theorem 6, I shall prove that it is sufficient to assume rather less for the normality of AB.

THEOREM 6. *The matrix AB is normal if and only if, for some ordering α_i, β_i $(i = 1,..., n)$, $\sum_{i=1}^{n} |f(\alpha_i)\beta_i|^2$ equals the sum of squared moduli of latent roots of every scalar polynomial $f(A)B$ for which the Hermitian polar matrix of $f(A)$ is a polynomial in the Hermitian polar matrix of A.*

Proof. Suppose that AB is normal. Then, if $f(A) = g(H)W$ and $B = KV$ are polar representations, we may use Theorem 4 with $l(B) = B$. It follows that $f(A)B$ is normal. By Theorem 5 the moduli of latent roots of $f(A)B$ are thus $|f(\alpha_i)\beta_i|$ $(i = 1,..., n)$, for some ordering α_i, β_i. Hence

$$\Lambda\{f(A)B\} = \sum_{i=1}^{n} |f(\alpha_i)\beta_i|^2. \tag{7}$$

† When $f(A)l(B)$ is normal, then $f(A)l(B) = GMWX$ for all polar matrices W and X. This is easily proved.

Suppose now that there is an ordering α_i, β_i for which (7) is satisfied for every $f(A)$ that has a polar representation $f(A) = g(H)W$. We have already proved that, if

$$\bar{\Phi}'A\Phi = D = \operatorname{diag}[\alpha_1,\ldots,\alpha_n],$$

then
$$\bar{\Phi}'H\Phi = R = \operatorname{diag}[|\alpha_1|,\ldots,|\alpha_n|].$$

On constructing the polynomial $\psi(x)$ satisfying $\psi(\alpha_i) = |\alpha_i|$ $(i = 1,\ldots,n)$, it follows that

$$\psi(D) = R,$$

and
$$\psi(A) = \Phi\psi(D)\bar{\Phi}' = \Phi R\bar{\Phi}' = H.$$

Since H is a polynomial in A, any polynomial in H is also a polynomial in A. Now let

$$f(A) = g(H)W.$$

We may suppose that $g(H) = \phi(A)$, and this implies $g(R) = \phi(D)$. From this equation and from Theorem 3 it follows that

$$\phi(\alpha_i) = g(|\alpha_i|) = |f(\alpha_i)| \quad (i = 1,\ldots,n).$$

Since, of course, $\phi(A) = g(H)$ is its own Hermitian polar matrix,

$$\Lambda\{g(H)B\} = \Lambda\{\phi(A)B\} = \sum_{i=1}^{n} |\phi(\alpha_i)\beta_i|^2 = \sum_{i=1}^{n} |g(|\alpha_i|)\beta_i|^2, \tag{8}$$

by hypothesis. We have proved (8) only for any non-negative Hermitian $g(H)$. From an inspection of the proof of Theorem 1 it is immediately clear that H and B commute if (8) is assumed only for the polynomials $g_i(H)$ defined for H in the same way as $f_i(A)$ is defined for A. These polynomials $f_i(A)$ used in the proof of Theorem 1 are non-negative definite Hermitian, being normal matrices with non-negative latent roots. Hence the corresponding $g_i(H)$ are also non-negative definite Hermitian; (8) holds for them and the commutativity of H and B follows.

If $\psi(A) = H$, we have the polar representations $\psi(A) = HI$ and $A = HU$. Hence, by the hypothesis,

$$\Lambda(HB) = \sum_{i=1}^{n} |\alpha_i\beta_i|^2$$

and
$$\Lambda(AB) = \sum_{i=1}^{n} |\alpha_i\beta_i|^2.$$

We deduce that $\Lambda(AB) = \Lambda(HB)$, where HB is normal, since H and B commute. On using Theorem 2 we see that AB is normal. The theorem is thus proved.

A corollary may be obtained by an argument similar to that which led to the corollary of Theorem 1.

COROLLARY. *Let the latent roots of A be α_{i_j} ($j = 1,\ldots, m_i$; $i = 1,\ldots, r$), where $|\alpha_{i_j}| = |\alpha''_i|$ and $|\alpha''_i| \neq |\alpha''_j|$ when $i \neq j$; and let the latent roots of B be β_{i_j} or β'_{i_j} ($j = 1,\ldots, m_i$; $i = 1,\ldots, r$). Let AB be normal. Then, if α_{i_j}, β_{i_j} is an ordering satisfying Theorem 6, α_{i_j}, β'_{i_j} is another if and only if*

$$\sum_{j=1}^{m_i} |\beta_{i_j}|^2 = \sum_{j=1}^{m_i} |\beta'_{i_j}|^2 \quad (i = 1,\ldots, r).$$

6. We may remark in conclusion, without proof, that the polynomials $f_i(A)$ and $g_i(H)$ are the principal idempotent elements of A and H respectively [cf. Drazin (1)].

In this connexion, as some work in the proof of Theorem 6 indicates, the condition

$$\Lambda\{f_i(A)B\} = \sum_{j=1}^{n_i} |\beta_{i_j}|^2 \quad (i = 1,\ldots, k)$$

is a sufficient guarantee of the commutativity of A and B. Analogously, the normality of AB is ensured by the condition

$$\Lambda\{g_i(H)UB\} = \sum_{j=1}^{m_i} |\beta_{i_j}|^2 \quad (i = 1,\ldots, r),$$

for *some* polar representation $A = HU$, though this requires much fuller amplification.

REFERENCES

1. M. P. Drazin, 'On diagonable and normal matrices', *Quart. J. of Math.* (Oxford) (2), 2 (1951), 189–98.
2. M. P. Drazin, J. W. Dungey, and K. W. Gruenberg, 'Some theorems on commutative matrices', *J. of London Math. Soc.* 26 (1951), 221–8.
3. W. V. Parker, 'The characteristic roots of matrices', *Duke Math. J.* 12 (1945), 519–26.
4. I. Schur, 'Über die charakterischen Wurzeln einer linearen Substitution mit einer Anwendung auf die Theorie der Integralgleichungen', *Math. Annalen,* 66 (1909), 488–510.
5. N. A. Wiegmann, 'Normal products of matrices', *Duke Math. J.* 15 (1948), 633–8.
6. J. Williamson, 'Note on a principal-axis transformation of non-Hermitian matrices', *Bull. American Math. Soc.* 45 (1939), 920–2.
7. A. Wintner and F. D. Murnaghan, 'On a polar representation of non-singular matrices', *Proc. National Academy Sciences, U.S.A.* 17 (1931), 676–8.