A NECESSARY AND SUFFICIENT CONDITION FOR SIMULTANEOUS DIAGONALIZATION OF TWO HERMITIAN MATRICES AND ITS APPLICATION

by YIK-HOI AU-YEUNG†

(Received 4 February, 1969)

1. Introduction and statement of the theorems. We denote by $F$ the field $R$ of real numbers, the field $C$ of complex numbers, or the skew field $H$ of real quaternions, and by $F^n$ an $n$-dimensional left vector space over $F$. If $A$ is a matrix with elements in $F$, we denote by $A^*$ its conjugate transpose. In all three cases of $F$, an $n \times n$ matrix $A$ is said to be hermitian if $A = A^*$, and we say that two $n \times n$ hermitian matrices $A$ and $B$ with elements in $F$ can be diagonalized simultaneously if there exists a non singular matrix $U$ with elements in $F$ such that $UAU^*$ and $UBU^*$ are diagonal matrices. We shall regard a vector $u \in F^n$ as a $1 \times n$ matrix and identify a $1 \times 1$ matrix with its single element, and we shall denote by diag $\{A_1, \ldots , A_m\}$ a diagonal block matrix with the square matrices $A_1, \ldots , A_m$ lying on its diagonal.

Let $A = \text{diag} \{A_1, \ldots , A_m\}$ and $B = \text{diag} \{B_1, \ldots , B_m\}$ be any two hermitian block matrices such that, for each $k = 1, \ldots , m$, $A_k$ and $B_k$ are of the same size. Then it is obvious that, if each pair $A_k$ and $B_k$ can be diagonalized simultaneously, so also can the pair $A$ and $B$. Whether the converse is true or not is not at all obvious. In this note the author gives a simple proof of the converse (Theorem 2) by first proving the following theorem on a necessary and sufficient condition for simultaneous diagonalization of two hermitian matrices.

**Theorem 1.** Let $A$ and $B$ be two $n \times n$ hermitian matrices with elements in $F$. Then $A$ and $B$ can be diagonalized simultaneously if and only if there exists a basis $\{\mathbf{u}_1, \ldots , \mathbf{u}_n\}$ of $F^n$ such that, for each $i = 1, \ldots , n$, the two vectors $u_iA$ and $u_iB$ are linearly dependent over $R$.

**Theorem 2.** Let $A = \text{diag} \{A_1, \ldots , A_m\}$ and $B = \text{diag} \{B_1, \ldots , B_m\}$ be two hermitian diagonal block matrices with elements in $F$ such that, for each $k = 1, \ldots , m$, $A_k$ and $B_k$ are of the same size. If $A$ and $B$ can be diagonalized simultaneously, then so also can the pair $A_k$ and $B_k$ for each $k$.

A theorem similar to Theorem 1, on the simultaneous diagonalization of two nondegenerate symmetric bilinear forms over a field of characteristic not equal to 2, has been established by M. J. Wonenburger [3, Theorem 1, p. 617].

2. Proof of Theorem 1. Suppose that $A$ and $B$ can be diagonalized simultaneously. Then there exists a basis $\{\mathbf{u}_1, \ldots , \mathbf{u}_n\}$ such that $u_iA_j^* = u_jB_i^* = 0$ for all $i \neq j$ $(i, j = 1, \ldots , n)$. Now, for each fixed $i$, if $u_iA_i^* = u_iB_i^* = 0$, then $u_iA = 0 = u_iB$, while if $u_iA_i^*$ and $u_iB_i^*$ are not both zero, then $(u_iB_i^*)u_iA - (u_iA_i^*)u_iB = 0$. Hence in both cases $u_iA$ and $u_iB$ are linearly dependent over $R$.

† The author wishes to thank Professor Y. C. Wong for his advice during the preparation of this note.
To prove the sufficiency of the condition, suppose that there exists a basis \( X = \{u_1, \ldots, u_n\} \) of \( F^n \) such that, for each \( i = 1, \ldots, n, u_i A \) and \( u_i B \) are linearly dependent over \( R \). Then, for each \( i \), there exist \( \alpha_i, \beta_i \in R \), not both zero, such that
\[
\alpha_i u_i A + \beta_i u_i B = 0.
\]

Now in the set \( X = \{u_1, \ldots, u_n\} \) we define a relation \( \sim \) by setting \( u_i \sim u_j \) if \( \alpha_i \beta_j - \alpha_j \beta_i = 0 \). Obviously this is an equivalence relation. Let

\[
X = X_1 \cup X_2 \cup \ldots \cup X_m
\]

be the partition defined by this relation. Then, for each \( k = 1, \ldots, m \), there exist \( a_k, b_k \in R \), not both zero, such that
\[
a_k u A + b_k u B = 0, \quad \text{for all } u \in X_k; \tag{1}
\]
\[
a_k b_l - a_l b_k \neq 0, \quad \text{for all } k \neq l (k, l = 1, \ldots, m). \tag{2}
\]

From these two properties and \( (u A v^*)_* = v A u^* \), it follows immediately that
\[
u A v^* = u B v^* = 0, \quad \text{for all } u \in X_k \text{ and } v \in X_l \text{ with } k \neq l. \tag{3}
\]

Without loss of generality we may assume that \( u_1, \ldots, u_{n_1} \in X_1, u_{n_1} + 1, \ldots, u_{n_1 + n_2} \in X_2, \ldots, u_{n_1 + n_2 + \ldots + n_{m-1} + 1}, \ldots, u_n \in X_m \). Let \( U \) be the matrix whose elements in \( i \)th row are the components of \( u_i \). Then \( U \) is non singular and, by (3), we have
\[
U A U^* = \text{diag} \{A_1, \ldots, A_m\},
\]
\[
U B U^* = \text{diag} \{B_1, \ldots, B_m\},
\]
where \( A_k \) and \( B_k \) are hermitian matrices of size \( n_k \) and, by (1), we have
\[
a_k u A v + b_k u B v = 0 \quad \text{for all } u, v \in X_k.
\]

Hence
\[
a_k A_k + b_k B_k = 0 \quad \text{for each } k = 1, \ldots, m.
\]

Since any hermitian matrix can be diagonalized (for \( F = R \) or \( C \), this is well-known; for \( F = H \), see [1] or [2]) and \( a_k, b_k \) are not both zero, \( A_k \) and \( B_k \) can be diagonalized simultaneously for each \( k \). Hence \( A \) and \( B \) can be diagonalized simultaneously.

3. Proof of Theorem 2. It suffices to prove the theorem for \( m = 2 \). Let \( A = \text{diag} \{A_1, A_2\} \) and \( B = \text{diag} \{B_1, B_2\} \), where \( A_1 \) and \( B_1 \) are of size \( n_1 \) and \( A_2 \) and \( B_2 \) are of size \( n_2 \), and let \( n = n_1 + n_2 \). If \( A \) and \( B \) can be diagonalized simultaneously, then, by Theorem 1, there exists a basis \( \{u_1, \ldots, u_n\} \) of \( F^n \) such that, for each \( i = 1, \ldots, n \), \( u_i A \) and \( u_i B \) are linearly dependent over \( R \).

Let \( u_i = (x_i, y_i) \), where \( x_i \in F^{n_1} \) and \( y_i \in F^{n_2} \). Then \( (x_i A_1, y_i A_2) \) and \( (x_i B_1, y_i B_2) \) are linearly dependent over \( R \) for each \( i \). Hence \( x_i A_1 \) and \( x_i B_1 \) are linearly dependent over \( R \) for each \( i \). Since \( \{u_1, \ldots, u_n\} \) is a basis of \( F^n \), there exists \( \{x_{i_1}, \ldots, x_{i_{n_1}}\} \) which forms a basis of \( F^{n_1} \). By Theorem 1, \( A_1 \) and \( B_1 \) can be diagonalized simultaneously. Similarly, \( A_2 \) and \( B_2 \) can be diagonalized simultaneously. This completes the proof.
REFERENCES

