A THEOREM ON A MAPPING FROM A SPHERE TO
THE CIRCLE AND THE SIMULTANEOUS DIAGONALIZATION OF TWO HERMITIAN MATRICES

YIK-HOI AU-YEUNG

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 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 unitary if $AA^* = I$, where $I$ is the $n \times n$ identity matrix. An $n \times n$ hermitian matrix $A$ is said to be positive definite if $uAu^* > 0$ for all $u (\neq 0)$ in $F^n$. Here and in what follows we regard $u$ as a $1 \times n$ matrix and identify a $1 \times 1$ matrix with its single element.

The purpose of this note is to prove Theorem 1 on a mapping from a sphere to the circle, and use it to prove Theorem 2 on the simultaneous diagonalization of two hermitian matrices.

Theorem 1. Let $A$ and $B$ be two $n \times n$ hermitian matrices with elements in $F$ such that $(uAu^*)^2 - (uBu^*)^2 > 0$ for all $u (\neq 0)$ in $F^n$, and $S(F^n)$ the unit sphere in $F^n$ (i.e. $S(F^n) = \{ u \in F^n : uu^* = 1 \}$). If $F = R$ and $n \geq 3$ or $F = C$ or $H$ and $n = 2$, then the image of the mapping $f: S(F^n) \rightarrow S(R^2)$ defined by

$$f(u) = \left( \begin{array}{c} uAu^* \\ (uAu^*)^2 + (uBu^*)^2 \end{array} \right)$$

is a closed circular arc of length $< \pi$.

Theorem 2. Let $A$ and $B$ be two $n \times n$ hermitian matrices with elements in $F$ such that $(uAu^*)^2 + (uBu^*)^2 > 0$ for all $u (\neq 0)$ in $F^n$. If $F = R$ and $n \geq 3$ or $F = C$ or $H$ and $n \geq 2$, then $A$ and $B$ can be diagonalized simultaneously (i.e. there exists a nonsingular $n \times n$ matrix $U$ with elements in $F$ such that $UAU^*$ and $UBU^*$ are diagonal matrices).

For the case $F = R$ and $n \geq 3$, Theorem 2 has been proved by Greub [1, pp. 231–237], and Calabi [2] by different methods.

Received by the editors October 27, 1967.

1 The author wishes to thank Professor Y. C. Wong for suggesting the problem and for his advice and encouragement during the preparation of this paper.

2 The author is grateful to the referee for drawing his attention to this paper.
2. A mapping from a sphere to the unit circle. We first prove a lemma.

**Lemma 1.** For any real numbers \(a_0, \ldots, a_4, b_0, \ldots, b_4\) the following system of equations

\[
\begin{align*}
  a_0(x^2 - y^2) + a_1xy + a_2xz + a_3yz + a_4z^2 &= 0, \\
  b_0(x^2 - y^2) + b_1xy + b_2xz + b_3yz + b_4z^2 &= 0,
\end{align*}
\]

has nontrivial real solutions.

**Proof.** There are two cases.

*Case 1.* \(a_0b_1 - a_1b_0 = 0\). In this case we put \(z = 0\) in (1), and (1) becomes

\[
\begin{align*}
  a_0(x^2 - y^2) + a_1xy &= 0, \\
  b_0(x^2 - y^2) + b_1xy &= 0.
\end{align*}
\]

Since \(a_0b_1 - a_1b_0 = 0\), (2) has nontrivial real solutions. Hence (1) has nontrivial real solutions.

*Case 2.* \(a_0b_1 - a_1b_0 \neq 0\). In this case we put \(z = 1\) in (1), and (1) becomes

\[
\begin{align*}
  a_0(x^2 - y^2) + a_1xy + a_2x + a_3y + a_4 &= 0, \\
  b_0(x^2 - y^2) + b_1xy + b_2x + b_3y + b_4 &= 0.
\end{align*}
\]

Since \(a_0b_1 - a_1b_0 \neq 0\), (3) is equivalent to

\[
\begin{align*}
  x^2 - y^2 + a'_1 x + a'_2 y + a'_4 &= 0, \\
  xy + b'_1 x + b'_2 y + b'_4 &= 0,
\end{align*}
\]

where \(a'_1, a'_2, b'_1, b'_2, b'_1, b'_2\) and \(b'_4\) are some real numbers. But (4) obviously has real solutions. Hence (1) has nontrivial real solutions.

We now prove Theorem 1. Since \(f\) is continuous and \(S(F^n)\) is connected and compact, \(f(S(F^n))\) is a closed circular arc. Therefore, Theorem 1 will be proved if we can show that, for any \((a, b)\) in \(S(R^2)\), \((a, b)\) and \((-a, -b)\) cannot both belong to \(f(S(F^n))\).

Assume that there exists \((a_0, b_0)\) in \(S(R^2)\) such that \((a_0, b_0)\) and \((-a_0, -b_0)\) both belong to \(f(S(F^n))\). Then by the definition of \(f\) there exist \(u_1\) and \(u_2\) in \(F^n \setminus \{0\}\) such that \((a_0, b_0) = (u_1Au_1^*, u_1Bu_1^*)\) and \((-a_0, -b_0) = (u_2Au_2^*, u_2Bu_2^*)\). Obviously, \(u_1\) and \(u_2\) are linearly independent over \(F\). Since \(F = R\) and \(n \geq 3\) or \(F = C\) or \(H\) and \(n \geq 2\), there exists \(u_3\) in \(F^n\) such that \(u_1, u_2\) and \(u_3\) are linearly independent over \(R\). Now, for any \((x, y, z)\) in \(R^3\) we have

\[
(xu_1 + yu_2 + zu_3)A(xu_1 + yu_2 + zu_3)^* = a_0(x^2 - y^2) + a_1xy + a_2xz + a_3yz + a_4z^2,
\]

and
\[(x_1 + y_2 + z_3)B(x_1 + y_2 + z_3)^* = b_0(x^2 - y^2) + b_1xy + b_2xz + b_3yz + b_4z^2,\]

where the a's and b's are all real numbers; for example, \(a_1 = u_1Au_\mathbf{1}^* + u_2Au_\mathbf{2}^*, \ b_1 = u_1Bu_\mathbf{1}^* + u_2Bu_\mathbf{2}^*. \) Therefore, by Lemma 1, there exists \((x_0, y_0, z_0) \neq (0, 0, 0)\) in \(R^3\) such that

\[ (x_0u_1 + y_0u_2 + z_0u_3)A(x_0u_1 + y_0u_2 + z_0u_3)^* = 0, \]

and

\[ (x_0u_1 + y_0u_2 + z_0u_3)B(x_0u_1 + y_0u_2 + z_0u_3)^* = 0. \]

Since \(u_1, u_2\) and \(u_3\) are linearly independent over \(R\), we obtain a contradiction to the hypothesis that \((u_1u_\mathbf{1}^*)^2 + (u_1u_\mathbf{2}^*)^2 > 0\) for all \(u(\neq 0)\) in \(F^n\). Hence Theorem 1 is proved.


Suppose that the conditions of Theorem 2 are satisfied. By Theorem 1, \(f(S(F^n))\) is a closed circular arc of length \(<\pi\). Let \((a, b)\) be the midpoint of this circular arc. Then if \(u\) is any point in \(S(F^n)\) and if \(\theta < \pi/2\) is the angle between the radii of \(S(R^2)\) with end points \((a, b)\) and \(f(u)\), we have

\[
a \frac{uAu^*}{((uAu^*)^2 + (uBu^*)^2)^{1/2}} + b \frac{uBu^*}{((uAu^*)^2 + (uBu^*)^2)^{1/2}} = \cos \theta > 0.
\]

Therefore, \(aA + bB\) is positive definite, and Theorem 2 is proved by the following lemma:

**Lemma 2.** If \(A\) and \(B\) are two \(n \times n\) hermitian matrices with elements in \(F\) such that \(aA + bB\) is positive definite for some \((a, b)\) in \(R^2\), then \(A\) and \(B\) can be diagonalized simultaneously.

**Proof.** Since \(aA + bB\) is a positive definite hermitian matrix, one of the \(a\) and \(b\), say \(a\), is not zero and there exists a unitary matrix \(U_1\) such that

\[ U_1(aA + bB)U_1^* = \text{diag}(a_1, \ldots, a_n), \]

where \(a_1, \ldots, a_n\) are positive real numbers. (This is well known if \(F = R\) or \(C\); for example, see [3, pp. 12–13]. For \(F = H\), it is proved in [4] and [5].) Let \(U_2 = \text{diag}(1/\sqrt{a_1}, \ldots, 1/\sqrt{a_n})\). Then

\[ U_2U_1(aA + bB)U_1^*U_2^* = I, \]

where \(I\) is the \(n \times n\) identity matrix. Since \(U_2U_1B U_1^*U_2^*\) is a hermitian matrix, there exists a unitary matrix \(U_3\) such that
(6) \[ UBU^* = \text{diagonal matrix}, \quad \text{where} \quad U = U_3 U_2 U_1. \]

From (5) and (6) it follows that

\[ UAU^* = \frac{1}{a}(I - bUBU^*) = \text{diagonal matrix}. \]

Thus Lemma 2 is proved.

**Added in proof.** The author has just learned that, for the case \( F = \text{real closed field and } n \geq 3 \), Theorem 2 has been proved by Wonenburger [J. Math. Mech. 15 (1966), 617–622]; and for the case \( F = \mathbb{R} \) and \( n \geq 3 \) or \( F = \mathbb{C} \) and \( n \geq 2 \) by Kraljević [Glasnik Mat. Ser. III 1 (21) (1966), 57–63]. Their methods of proof are quite different from that of the author.

**References**


**University of Hong Kong**