SOME THEOREMS ON THE REAL PENCIL AND SIMULTANEOUS DIAGONALIZATION OF TWO HERMITIAN BILINEAR FUNCTIONS

YIK-HOI AU-YEUNG[1]

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. We shall consider a pair of hermitian bilinear functions in a left vector space V over F. Here V may be finite or infinite dimensional. If $F = R$, then the word "hermitian" merely means "symmetric."

The purpose of this note is: (i) to prove Theorem 1 on the image of a mapping f defined by two hermitian bilinear functions, and use it to prove Theorem 2 on the positive semidefiniteness of a nontrivial real linear combination of two hermitian bilinear functions, and (ii) to establish some relation between the image of the mapping f defined in Theorem 1 and the simultaneous diagonalization of two hermitian bilinear functions (Theorems 3, 4, 5 and 6), and use them to reconsider Theorem 1 for finite dimensional cases (Theorem 1').

THEOREM 1. *Let V be a left vector space over F (finite ($\geqq 2$) or infinite dimensional), ϕ and ψ two hermitian bilinear functions in V not both identical zero, $Z = \{u \in V : \phi(u, u) = \psi(u, u) = 0\}$, and $f : D = V \backslash Z \to S$, where S is the real unit circle, is the mapping defined by*

$$f(u) = \left(\frac{\phi(u, u)}{([\phi(u, u)]^2 + [\psi(u, u)]^2)^{1/2}}, \frac{\psi(u, u)}{([\phi(u, u)]^2 + [\psi(u, u)]^2)^{1/2}} \right).$$

Then the image $f(D)$ is of one of the following types:
(i) *Two diametrically opposite points in S;*
(ii) *S;*
(iii) *A circular arc of length $< \pi$ with or without one or both endpoints; and*
(iv) *A half circle with or without one or both endpoints.*

Here and in what follows a circular arc of length $< \pi$ includes the case that it is a singleton, and throughout this note, the symbols V, ϕ, ψ, f, D and S are defined as in Theorem 1. (A mapping slightly different from f has been considered by the author in a previous paper [1].)

THEOREM 1'. (a) *If V is finite dimensional, then case* (iii) *in Theorem 1 can be replaced by*

(iii)' *A closed circular arc of length* $<\pi$.

(b) *If V is 2-dimensional, then cases* (iii) *and* (iv) *in Theorem 1 can be replaced by*

(iii)' *A closed circular arc of length* $<\pi$.

(iv)' *An open half circle.*

THEOREM 2. *If in Theorem* 1, $f(D)\neq S$, *then there exists* (α, β) ($\neq(0, 0)$) *in* R^2 *such that* $\alpha\phi+\beta\psi$ *is positive semidefinite (i.e.* $\alpha\phi(u, u)$ $+\beta\psi(u, u)\geqq 0$ *for all* $u\in V$).

THEOREM 3. *If V is finite dimensional and* ϕ *and* ψ *can be diagonalized simultaneously, then* $f(D)$ *is closed.*

THEOREM 4. *If V is 2-dimensional and* ϕ *and* ψ *can be diagonalized simultaneously, then* $f(D)$ *is a set consisting of only two diametrically opposite points or is a (closed) circular arc of length* $<\pi$.

THEOREM 5. *If V is finite dimensional and if* $f(D)$ *is a set consisting of only two diametrically opposite points or is a (closed) circular arc of length* $<\pi$, *then* ϕ *and* ψ *can be diagonalized simultaneously.*

THEOREM 6. *If V is 3-dimensional and* $f(D)$ *is a closed half circle, then* ϕ *and* ψ *can be diagonalized simultaneously.*

For finite dimensional cases, we see from Theorems 3–6 and 1' that the behaviour of $f(D)$ determines the possibility of simultaneous diagonalization of ϕ and ψ except for the following cases: (a) V is of dimension >2 and $f(D)=S$; (b) V is of dimension >3 and $f(D)$ is a closed half circle. We shall show in §7 that in these cases ϕ and ψ may or may not be diagonalized simultaneously. In §7 we also give examples to show that all cases which appear in Theorems 1 or 1' really occur. Hence for finite dimensional cases, our information is complete.

2. **Proof of Theorems 1 and 2.** To prove Theorem 1 we first prove the following lemma.

LEMMA 1. *If* P_1, P_2 *are two points of* $f(D)$ *and the length of the circular arc* $(P_1P_2)^\vee$ *is* $<\pi$, *then* $(P_1P_2)^\vee\subset f(D)$.

PROOF. From the definition of f, we see that a point $(\alpha, \beta)\in S$ is in $f(D)$ if and only if there exists a $u\in V$ such that $(\phi(u, u), \psi(u, u))=(\alpha, \beta)$. Now let $P_1=(\phi(u_1, u_1), \psi(u_1, u_1))=(\alpha_1, \beta_1)$ and $P_2=(\phi(u_2, u_2), \psi(u_2, u_2))$ $=(\alpha_2, \beta_2)$ be two distinct points in $f(D)$ such that the length of $(P_1P_2)^\vee$ is $<\pi$. Then $\alpha_1\beta_2-\alpha_2\beta_1\neq 0$ and for any interior point (α, β) of $(P_1P_2)^\vee$

we have

(1) $$(\alpha_1\beta_2 - \alpha_2\beta_1)(\alpha_1\beta - \alpha\beta_1) > 0,$$

and

(2) $$(\alpha_2\beta_1 - \alpha_1\beta_2)(\alpha_2\beta - \alpha\beta_2) > 0.$$

We now show that for any interior point (α, β) of $(P_1P_2)^{\cup}$ there exists a $u \in V$ of the form $u = x_1u_1 + x_2u_2$, where $x_1, x_2 \in R$, such that $(\phi(u, u), \psi(u, u)) = (\alpha, \beta)$, and this means $(\alpha, \beta) \in f(D)$. In fact, consider the following system of equations with real coefficients:

(3)
$$\phi(u, u) \equiv \alpha_1x_1^2 + \alpha_2x_2^2 + \alpha_3x_1x_2 = \alpha,$$
$$\psi(u, u) \equiv \beta_1x_1^2 + \beta_2x_2^2 + \beta_3x_1x_2 = \beta,$$

where $\alpha_3 = \phi(u_1, u_2) + \phi(u_2, u_1)$ and $\beta_3 = \psi(u_1, u_2) + \psi(u_2, u_1)$. Since $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$, system (3) is equivalent to the following

$$(\alpha_2\beta_1 - \alpha_1\beta_2)x_2^2 + (\alpha_3\beta_1 - \alpha_1\beta_3)x_1x_2 = \alpha\beta_1 - \alpha_1\beta,$$
$$(\alpha_1\beta_2 - \alpha_2\beta_1)x_1^2 + (\alpha_3\beta_2 - \alpha_2\beta_3)x_1x_2 = \alpha\beta_2 - \alpha_2\beta,$$

which, on account of (1) and (2), has real solution for x_1 and x_2. Hence our lemma is proved.

Theorem 1 now follows immediately from Lemma 1. If $f(D)$ is disconnected, then $f(D)$ must consist of only two diametrically opposite points in S; if $f(D)$ is connected, then it is a circular arc; and, if this circular arc is of length $> \pi$, then it must be S itself. Hence Theorem 1 is proved.

If $f(D) \neq S$, by Theorem 1 there exists a half circle $\mathcal{C}$ such that $f(D) \subset \mathcal{C}$. Let (α, β) be the midpoint of $\mathcal{C}$, then $\alpha\phi + \beta\psi$ is positive semidefinite. Thus Theorem 2 is proved.

3. **Proof of Theorems 3 and 4.** To prove Theorems 3 and 4 we first prove the following lemma which is true for vector space of finite or infinite dimensions.

LEMMA 2. *If B is a base of V such that $\phi(u_i, u_j) = \psi(u_i, u_j) = 0$ for all $u_i \neq u_j$ in B, then the set $X = \{f(u) : u \in B \cap D\}$ determines $f(D)$ in the following manner:*

(a) *If X consists of only two diametrically opposite points, then $f(D) = X$;*

(b) *If X is contained in a half circle but not of case (a), then $f(D)$ is the smallest circular arc containing X (this circular arc may or may not contain one or both endpoints); and*

(c) *If X is not contained in any half circle, then $f(D) = S$.*

PROOF. Case (a) is obvious and case (c) follows from Theorem 1. We now prove case (b). Let $\mathcal{C}$ be the smallest circular arc containing X and let (α, β) and (γ, δ) be the anticlockwise and clockwise endpoints of $\mathcal{C}$. Then

$$(4) \qquad \alpha\psi(u_i, u_i) - \beta\phi(u_i, u_i) \leqq 0 \qquad \text{for all } u_i \in B \cap D.$$

Now for any $u \in V$, u can be expressed as $u = \sum_{i=1}^{n} \lambda_i u_i$, where $\lambda_i \in F$ $u_i \in B$. From our assumption on ϕ and ψ we have

$$\alpha\psi(u, u) - \beta\phi(u, u) = \alpha\left[\sum_{i=1}^{n} \lambda_i\bar{\lambda}_i\psi(u_i, u_i)\right]$$

$$- \beta\left[\sum_{i=1}^{n} \lambda_i\bar{\lambda}_i\phi(u_i, u_i)\right]$$

$$= \sum_{i=1}^{n} \lambda_i\bar{\lambda}_i[\alpha\psi(u_i, u_i) - \beta\phi(u_i, u_i)]$$

where $\bar{\lambda}_i$ denote the conjugate of λ_i in F. From this equation and inequality (4) we have

$$\alpha\psi(u, u) - \beta\phi(u, u) \leqq 0 \qquad \text{for all } u \in D.$$

Similarly

$$\gamma\psi(u, u) - \delta\phi(u, u) \geqq 0 \qquad \text{for all } u \in D.$$

Furthermore, it can be easily proved that if an endpoint of $\mathcal{C}$ belongs to f then it belongs to X. Hence $f(D) \subset \mathcal{C}$. Conversely, from Theorem 1 and the fact that $X (\subset f(D))$ is not a set consisting of only two diametrically opposite points, we have $\mathcal{C} \subset f(D)$. Thus Lemma 2 is proved.

Theorems 3 and 4 now follow immediately from Lemma 2. If V is finite dimensional, say of F-dimension n, and if ϕ and ψ can be diagonalized simultaneously, then there exists a base $B = \{u_1, u_2, \cdots, u_n\}$ of V such that $\phi(u_i, u_j) = \psi(u_i, u_j) = 0$ for all $i \neq j$. Now since the set $X = \{f(u): u \in B \cap D\}$ is finite, $f(D)$ must be closed, and Theorem 3 is proved.

If V is of F-dimension 2 and if ϕ and ψ can be diagonalized simultaneously, then there exists a base $\{u_1, u_2\}$ of V such that $\phi(u_1, u_2) = \psi(u_1, u_2) = 0$. Since the set $X = \{f(u): u \in B \cap D\}$ contains at most two points, Theorem 4 follows from Lemma 2.

4. Proof of Theorem 5. If $f(D)$ consists of only two diametrically opposite points, then there exists $(\alpha, \beta)(\neq(0, 0)) \in R^2$ such that $\alpha\phi + \beta\psi = 0$. On the other hand it is known that any hermitian bilinear function in a finite dimensional vector space V can be diagonalized.

(For $F = R$ or C see [3, pp. 12–13] and for $F = H$ see [4] or [5].) Hence ϕ and ψ can be diagonalized simultaneously.

If $f(D)$ is a circular arc of length $< \pi$, then we consider a half circle $\mathcal{C}$ such that the closure of $f(D)$ lies entirely in the interior of $\mathcal{C}$. Let (α_1, β_1) and (α_2, β_2) be two distinct points in S sufficiently near the midpoint of $\mathcal{C}$. Then we have $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$ and $\alpha_1\phi + \beta_1\psi$ and $\alpha_2\phi + \beta_2\psi$ are positive semidefinite. But under these conditions ϕ and ψ can be diagonalized simultaneously, as has been proved recently by the author [2].

5. **Proof of Theorem 6.** To prove Theorem 6 we first prove the following two lemmas which hold for vector spaces of finite or infinite dimensions.

LEMMA 3. *Let ζ be a hermitian bilinear function in V and $u_0 \in V$. If $\zeta(u_0, u) + \zeta(u, u_0) \leqq 0$ (or $\geqq 0$) for all $u \in V$, then $\zeta(u_0, u) = 0$ for all $u \in V$.*

PROOF. Suppose that there exists $v \in V$ such that $\zeta(u_0, v) \neq 0$. Let $u = \zeta(u_0, v)v$. Then

$$\zeta(u_0, u) + \zeta(u, u_0) = \zeta(u_0, \zeta(u_0, v)v) + \zeta(\zeta(u_0, v)v, u_0)$$

$$= \zeta(u_0, v)\overline{\zeta(u_0, v)} + \zeta(u_0, v)\zeta(v, u_0)$$

$$= 2\zeta(u_0, v)\overline{\zeta(u_0, v)} > 0,$$

and this contradicts our assumption.

LEMMA 4. *Let $u_0 \in V$ such that $[\phi(u_0, u_0)]^2 + [\psi(u_0, u_0)]^2 \neq 0$. If $\phi(u_0, u_0)\psi(u, u) - \phi(u, u)\psi(u_0, u_0) \leqq 0$ (or $\geqq 0$) for all $u \in V$, then V can be decomposed into direct sum*

$$V = L\{u_0\} \oplus W,$$

where $L\{u_0\}$ is the subspace of V spanned by u_0 and $W = \{u \in V : \phi(u_0, u) = \psi(u_0, u) = 0\}$.

PROOF. Since $[\phi(u_0, u_0)]^2 + [\psi(u_0, u_0)]^2 \neq 0$, we have $L\{u_0\} \cap W = \{0\}$. Now for any $u \in V$ and for any positive real number ϵ, we have

$$\phi(u_0, u_0)\psi(u_0 + \epsilon u, u_0 + \epsilon u) - \phi(u_0 + \epsilon u, u_0 + \epsilon u)\psi(u_0, u_0)$$

$$= \epsilon[\zeta(u_0, u) + \zeta(u, u_0)] + \epsilon^2\zeta(u, u),$$

where $\zeta = \phi(u_0, u_0)\psi - \psi(u_0, u_0)\phi$. From the assumption of the lemma and the fact that ϵ is positive, we have

$$\zeta(u_0, u) + \zeta(u, u_0) + \epsilon\zeta(u, u) \leqq 0.$$

Since ϵ is arbitrary, we have

$$\zeta(u_0, u) + \zeta(u, u_0) \leqq 0$$

for all $u \in V$, and from Lemma 3 we have $\zeta(u_0, u) = 0$ for all u. This means

$$\phi(u_0, u_0)\psi(u_0, u) - \psi(u_0, u_0)\phi(u_0, u) = 0$$

for all $u \in V$. From this expression and the fact that $[\phi(u_0, u_0)]^2 + [\psi(u_0, u_0)]^2 \neq 0$ we see that for any $u \in V$ there exists $\lambda \in F$ such that $\phi(u_0, u) = \lambda\phi(u_0, u_0)$ and $\psi(u_0, u) = \lambda\psi(u_0, u_0)$. This means $\phi(u_0, u - \bar{\lambda}u_0) = \psi(u_0, u - \bar{\lambda}u_0) = 0$ (i.e. $u - \bar{\lambda}u_0 \in W$), and hence the lemma is proved. [The technique used in this proof is due to Professor Y. C. Wong.]

We now prove Theorem 6. First let $(\phi(u_1, u_1), \psi(u_1, u_1)) = (\alpha, \beta)$ and $(\phi(v_1, v_1), \psi(v_1, v_1)) = (-\alpha, -\beta)$ be the endpoints of $f(D)$ such that

(5) $\phi(u_1, u_1)\psi(u, u) - \phi(u, u)\psi(u_1, u_1) = \alpha\psi(u, u) - \beta\phi(u, u) \leqq 0$

for all $u \in V$. Then by Lemma 4 we have

$$V = L\{u_1\} \oplus W_1,$$

where $W_1 = \{u \in V: \phi(u_1, u) = \psi(u_1, u) = 0\}$. Next we consider the restrictions of ϕ and ψ on W_1 and prove that there exists $u_2 \in W_1$ such that

(6) $\phi(u_2, u_2)\psi(u, u) - \phi(u, u)\psi(u_2, u_2) \geqq 0$

for all $u \in W_1$. In fact let $v_1 = \lambda u_1 + u_2$, where $\lambda \in F$ and $u_2 \in W_1$. Then by the property of W_1 and v_1 we have

$$\phi(u_2, u_2) = -\alpha(1 + \lambda\bar{\lambda}), \qquad \psi(u_2, u_2) = -\beta(1 + \lambda\bar{\lambda}).$$

From these two equalities and inequality (5) we have inequality (6). Now by applying Lemma 4 once more we have

$$W_1 = L\{u_2\} \oplus W_2,$$

where $W_2 = \{u \in W_1: \phi(u_2, u) = \psi(u_2, u) = 0\}$. Therefore, if we take any $u_3(\neq 0) \in W_2$, then the base $\{u_1, u_2, u_3\}$ gives a simultaneous diagonalization of ϕ and ψ. Thus Theorem 6 is proved.

6. **Proof of Theorem 1′.** If $f(D)$ is a circular arc of length $< \pi$, then by Theorem 5 ϕ and ψ can be diagonalized simultaneously and by Theorem 3, $f(D)$ must be closed.

If V is 2-dimensional and $f(D)$ is a circular arc of length $\leqq \pi$ containing at least one endpoint, then by the argument we used in the

proof of Theorem 6, we can prove that ϕ and ψ can be diagonalized simultaneously, and hence by Theorem 4 the length of $f(D)$ must $< \pi$.

7. **Some examples.** It can easily be seen from Lemma 2 that for an infinite dimensional vector space all the cases which appear in Theorem 1 really occur. In Example 1, we shall construct ϕ and ψ such that $f(D)$ is a circular arc of length $\leqq \pi$ without one endpoint. Examples of other types can be constructed similarly.

EXAMPLE 1. Let V be an infinite dimensional left vector space over F with a countable base $B = \{u_0, u_1, u_2, \cdots\}$, $(PQ)^{\smile}$ a circular arc in S (not reduced to one point) of length $\leqq \pi$, and $P_1, P_2, \cdots$ a sequence of points in the interior of $(PQ)^{\smile}$ which tends to P. Define ϕ and ψ by

$$\phi(u_i, u_j) = \psi(u_i, u_j) = 0 \quad \text{for all } i \neq j,$$

$$(\phi(u_0, u_0), \psi(u_0, u_0)) = Q, \qquad (\phi(u_k, u_k), \psi(u_k, u_k)) = P_k \quad \text{for all } k \geqq 1.$$

Then by Lemma 2 $f(D)$ is the circular arc$(PQ)^{\smile}$ but without P.

In the case of finite dimensional V of F-dimension n, Lemma 2 also shows that for $n \geqq 2$, cases (i), (iii)$'$ and for $n \geqq 3$ cases (ii) and (iv) where $f(D)$ is a closed half circle in Theorem 1$'$ really occur (in fact in these cases ϕ and ψ can be so constructed that they can be diagonalized simultaneously). We now show by construction that the remaining cases:

(1) $n = 2$ and $f(D) = S$;

(2) $n \geqq 2$ and $f(D)$ is an open half circle; and

(3) $n > 2$ and $f(D)$ is a half circle with one endpoint only, also really occur.

(In the following examples we shall identify a hermitian bilinear function with a matrix.)

EXAMPLE 2. Let $V = R^2$ and

$$\phi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then $f(D) = S$ and hence by Theorem 4, ϕ and ψ cannot be diagonalized simultaneously.

EXAMPLE 3. Let $V = R^2$ and

$$\phi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then $f(D)$ is the upper half circle without both endpoints and hence by Theorem 3, ϕ and ψ cannot be diagonalized simultaneously.

EXAMPLE 4. Let $V = R^3$ and

$$\phi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $f(D)$ is the upper half circle without the endpoint $(-1, 0)$ and hence by Theorem 3, ϕ and ψ cannot be diagonalized simultaneously.

Finally, we give two examples which together with the remark just preceding Example 2, show that Theorems 3–6 in fact give "all" the possible relations between the behaviour of $f(D)$ and simultaneous diagonalization of two hermitian bilinear functions.

EXAMPLE 5. Let $V = R^3$ and

$$\phi = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then $f(D) = S$, but ϕ and ψ cannot be diagonalized simultaneously.

EXAMPLE 6. Let $V = R^4$ and

$$\phi = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \psi = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then $f(D)$ is the upper half circle with both endpoints, but ϕ and ψ cannot be diagonalized simultaneously.

REFERENCES

1. Yik-Hoi Au-Yeung, *A theorem on a mapping from a sphere to the circle and the simultaneous diagonalization of two hermitian matrices*, Proc. Amer. Math. Soc. 20(1969), 545–548.

2. ———, *A note on some theorems on simultaneous diagonalization of two hermitian matrices* (to appear).

3. C. Chevalley, *Theory of Lie groups*, Princeton Univ. Press, Princeton, N. J. 1964.

4. H. C. Lee, *Eigenvalues and canonical forms of matrices with quaternion coefficients*, Proc. Roy. Irish Acad. Sect. A 52(1949), 253–260.

5. J. Radon, *Linear Scharen orthogonaler Matrizen*, Abh. Math. Sem. Univ. Hamburg 1(1922), 1–14.

UNIVERSITY OF HONG KONG