

Conditions for linear dependence of two operators

Bojan Kuzma, Gorazd Lešnjak, Chi-Kwong Li, Tatjana
Petek and Leiba Rodman

Dedicated to Israel Gohberg on the occasion of his 80th birthday

Abstract. The linear dependence property of two Hilbert space operators is expressed in terms of equality of size of values of certain sesquilinear and quadratic forms associated with the operators. The forms are based on q -numerical ranges.

Mathematics Subject Classification (2000). Primary 47A12; Secondary 47A99.

Keywords. Hilbert space, linear operators, linear dependence, numerical values, generalized numerical range.

1. Introduction

Let \mathbf{H} be a complex Hilbert space with the inner product $\langle \cdot, \cdot \rangle$, and let $L(\mathbf{H})$ be the algebra of bounded linear operators on \mathbf{H} . If \mathbf{H} is of finite dimension n , we will identify \mathbf{H} with \mathbb{C}^n , the complex vector space of n -component column vectors with the standard inner product $\langle x, y \rangle = y^*x$, $x, y \in \mathbb{C}^n$, and will identify $L(\mathbf{H})$ with M_n , the algebra of $n \times n$ complex matrices. We assume throughout that \mathbf{H} has dimension at least 2.

In this paper, we are interested in studying the following problem.

This work started during the visit of B. Kuzma, G. Lešnjak and T. Petek at College of William and Mary in Williamsburg. They are very grateful to the Department of Mathematics for the very warm hospitality they enjoyed during that visit. Research of C.-K. Li and L. Rodman was supported by the William and Mary Plumeri Awards and by NSF grants DMS 0600859 and DMS-0456625, respectively. B. Kuzma, G. Lešnjak, and T. Petek were supported by the grant BI-US/06-07-001.

Problem 1.1. *Suppose $A, B \in L(\mathbf{H})$. Find effective criteria to show that $\{A, B\}$ is linearly independent. In connection with this problem, when do $\{I, A\}$ and $\{I, B\}$ generate the same algebras in $L(\mathbf{H})$?*

Of course, if A and B are given, this is an easy question to answer. But in Problem 1.1 we assume that only partial information about the operators is given. For example, it is well known that if there is $\mu \in \mathbb{C}$ such that $\langle Ax, x \rangle = \mu \langle Bx, x \rangle$ for all normalized elements $x \in \mathbf{H}$, then $A = \mu B$. We will prove that for two nonzero operators $A, B \in L(\mathbf{H})$, at least one of the pairs A and B or A and B^* is linearly dependent if and only if there is $r > 0$ such that $|\langle Ax, x \rangle| = r |\langle Bx, x \rangle|$ for all normalized elements $x \in \mathbf{H}$; see Theorem 2.2 below. (Alternatively, we can say $|\langle Ax, x \rangle|/|\langle Bx, x \rangle|$ assumes only a single value whenever $|\langle Bx, x \rangle| \neq 0$.) The analysis turns out to be more involved, and inspires other equivalent conditions for $\{A, B\}$ to be linearly dependent.

Instead of comparing the absolute values of quadratic forms $\langle Ax, x \rangle$ and $\langle Bx, x \rangle$, we can also deduce our conclusion by considering the absolute values of general sesquilinear forms $\langle Ax, y \rangle$ and $\langle Bx, y \rangle$ for normalized elements $x, y \in \mathbf{H}$ with inner product $\langle x, y \rangle = q$ for a fixed value $q \in [0, 1]$. Clearly, if $q = 1$, we have $x = y$, and we are back to the quadratic form result. In general, we can show that two nonzero operators A and B are linearly dependent if and only if there is $r > 0$ and $q \in (0, 1)$ such that $|\langle Ax, y \rangle| = r |\langle Bx, y \rangle|$ for all normalized elements $x, y \in \mathbf{H}$ with $\langle x, y \rangle = q$; see Theorem 2.2. We also show that if $q = 0$, the above condition forces linear dependence of $A - aI$ and $B - bI$, for some $a, b \in \mathbb{C}$. Thus, $\{I, A\}$ and $\{I, B\}$ will generate the same algebra.

Note that $A \rightarrow \langle Ax, y \rangle$ can be viewed as the image of A under a (bounded) linear functional. The results we described above can be reformulated in the following form: two nonzero operators A and B are linearly dependent if and only if there is $r > 0$ such that $|f(A)| = r |f(B)|$ for all linear functionals in a certain specific class \mathcal{S} .

Clearly, to see whether $|f(A)| = r |f(B)|$ for all the linear functionals in a set \mathcal{S} , we can detect r by finding a linear functional $\hat{f} \in \mathcal{S}$ such that $\hat{f}(B) \neq 0$, and set $r = |\hat{f}(A)|/|\hat{f}(B)|$. Then we can replace B by B/r and test the condition $|f(A)| = |f(B)|$ for all $f \in \mathcal{S}$. We will do that in our subsequent discussion.

2. Main results

We consider Problem 1.1 in the context of numerical values of an operator defined by a constrained sesquilinear form, namely, q -numerical ranges. For a fixed q , $0 \leq q \leq 1$, consider the q -numerical range

$$W_q(A) = \{\langle Ax, y \rangle : x, y \in \mathbf{H}, \langle x, x \rangle = \langle y, y \rangle = 1, \langle x, y \rangle = q\}$$

of an operator $A \in L(\mathbf{H})$. The q -numerical range has been extensively studied during last twenty years or so; see, for example, [11, 10, 13, 2] among many works

on the subject. If $q = 1$, then $W_1(A)$ coincides with the classical numerical range

$$W(A) = \{\langle Ax, x \rangle : x \in \mathbf{H}, \langle x, x \rangle = 1\}.$$

Let $q = \cos t$, $0 \leq t \leq \pi/2$. Then it is easy to see that we have

$$W_q(A) = \{\langle Ax, (\cos t)x + (\sin t)y \rangle : x, y \in \mathbf{H}, (x, y) \text{ orthonormal pair}\}.$$

We say that the numbers $\langle Ax, (\cos t)x + (\sin t)y \rangle$, where (x, y) varies through the set of orthonormal pairs, form the q -numerical values of the operator A . The characterization of operators having the same q -numerical values is easy to obtain (and the case $q = 1$ is well known):

Proposition 2.1. *Fix $q = \cos t$, $0 \leq t \leq \pi/2$. Two operators $A, B \in L(\mathbf{H})$ have the property that*

$$\langle Ax, (\cos t)x + (\sin t)y \rangle = \langle Bx, (\cos t)x + (\sin t)y \rangle \quad (2.1)$$

for every orthonormal pair (x, y) , $x, y \in \mathbf{H}$, if and only if $A = B$ in case $t < \pi/2$, or $A - B$ is scalar (i.e., a scalar multiple of I) in case $t = \pi/2$.

Proof. The “if” part is obvious. For the “only if” part, if $t = \pi/2$, then for every nonzero $x \in \mathbf{H}$, the element $(A - B)x$ is orthogonal to $\text{span}\{x\}^\perp$, and therefore $(A - B)x$ is a scalar multiple of x : $(A - B)x = \lambda_x x$ for some $\lambda_x \in \mathbb{C}$; a priori λ_x may depend on x , but the additivity of $A - B$ easily implies that in fact λ_x is independent of x . Assume now $t < \pi/2$. The condition (2.1) implies that for a fixed orthonormal pair (x, y) , the two circles in the complex plane

$$\{\langle Ax, x + (\tan t)ye^{i\theta} \rangle : 0 \leq \theta \leq 2\pi\}, \{\langle Bx, x + (\tan t)ye^{i\theta} \rangle : 0 \leq \theta \leq 2\pi\}$$

coincide, therefore their centers are the same: $\langle Ax, x \rangle = \langle Bx, x \rangle$. Since this equality holds for every normalized element x , we have $A = B$, as required. \square

In this paper we consider $A, B \in L(\mathbf{H})$ for which we require only equality in size in (2.1):

$$\begin{aligned} |\langle Ax, (\cos t)x + (\sin t)y \rangle| &= |\langle Bx, (\cos t)x + (\sin t)y \rangle|, \\ \forall \text{ orthonormal pairs } (x, y), \quad x, y \in \mathbf{H}. \end{aligned} \quad (2.2)$$

Besides independent interest, the problem of characterization of operators A and B satisfying (2.2) came up (for $t = 0$) in the study of norm preservers of Jordan products [9].

A complete characterization of such A and B is given in our main result:

Theorem 2.2. *Fix $q = \cos t$, $0 \leq t \leq \pi/2$. Two operators $A, B \in L(\mathbf{H})$ have the property (2.2) if and only if*

- (1) $A = \mu B$ or $A = \mu B^*$ for some $\mu \in \mathbb{C}$, $|\mu| = 1$ in case $t = 0$;
- (2) $A = \mu B$ for some $\mu \in \mathbb{C}$, $|\mu| = 1$ in case $0 < t < \pi/2$;
- (3) $A = \mu B + \nu I$ for some $\mu, \nu \in \mathbb{C}$, $|\mu| = 1$ in case $t = \pi/2$.

Remark 2.3. *It is interesting to observe that the case $t = 0$ fails if one replaces the modulus by the real or imaginary part or by the argument of the complex number. To see this, pick any two positive definite $A, B \in L(\mathbf{H})$ and note that $\operatorname{Re}(\langle iAx, x \rangle) = 0 = \operatorname{Re}(\langle iBx, x \rangle)$, and $\operatorname{Arg}(\langle Ax, x \rangle) = 0 = \operatorname{Arg}(\langle Bx, x \rangle)$ for any normalized $x \in \mathbf{H}$.*

Corollary 2.4. *Let $A, B \in L(\mathbf{H})$. Then*

$$|\langle Ax, y \rangle| = |\langle Bx, y \rangle|, \quad \forall x, y \in \mathbf{H}$$

if and only if $A = \mu B$ for some $\mu \in \mathbf{C}$, $|\mu| = 1$.

Proof. The part “if” is obvious, and the part “only if” is immediate from Theorem 2.2, the case $0 < t < \pi/2$.

However, Corollary 2.4 is actually used to prove Theorem 2.2, so we will deduce the corollary from the case $t = \pi/2$ of Theorem 2.2. Indeed, we have $A = \mu B + \nu I$ for some $\mu, \nu \in \mathbf{C}$, $|\mu| = 1$, and hence

$$|\langle \mu Bx, y \rangle + \nu \langle x, y \rangle| = |\langle Bx, y \rangle|, \quad \forall x, y \in \mathbf{H}.$$

Assuming $\nu \neq 0$, and taking y orthogonal to Bx we see that y is also orthogonal to x . Thus, $(\operatorname{span}(Bx))^\perp \subseteq (\operatorname{span} x)^\perp$, and so Bx is a scalar multiple of x : $Bx = \lambda_x x$, $\lambda_x \in \mathbf{C}$, for every $x \in \mathbf{H}$. Linearity of B easily implies that B is scalar, and now clearly $A = \mu' B$ for some $\mu' \in \mathbf{C}$, $|\mu'| = 1$. \square

Sections 3, 4, and 5 will be devoted to the proof of Theorem 2.2. In the last Section 6 we extend Proposition 2.1 to functionals given by trace class operators, and formulate an open problem and a conjecture concerning extension of Theorem 2.2 to such functionals.

We use notation \mathbf{e}_j for the j th standard unit vector in \mathbf{C}^n . $\operatorname{Re}(z)$ and $\operatorname{Im}(z)$ stand for the real and imaginary parts of the complex number $z = \operatorname{Re}(z) + i\operatorname{Im}(z)$. We denote by X^{tr} the transpose of a matrix or vector X . The (block) diagonal matrix or operator with diagonal matrix or operator blocks X_1, \dots, X_p (in that order) will be denoted $\operatorname{diag}(X_1, \dots, X_p)$.

3. Proof of Theorem 2.2: $t = 0$

For the proof of (1) we need preliminary results in matrix analysis which are of independent interest. We state and prove them first.

We start with the following known facts:

Proposition 3.1. (a) *If $T \in M_n$ is not the zero matrix, then there exists a unitary U such that the diagonal entries of UTU^* are all nonzero.*

(b) *If $R, S \in M_n$ are such that U^*RU and U^*SU have the same diagonal for every unitary $U \in M_n$, then $R = S$.*

Proof. Part (b) is obvious because under the hypotheses of part (b) we have $\langle Rx, x \rangle = \langle Sx, x \rangle$ for every $x \in \mathbb{C}^n$.

Part (a). Note that every matrix is unitarily equivalent to a matrix with equal entries on the main diagonal, see [7, Problem 3, p. 77]. So we are done if $\text{trace}(A) \neq 0$. Assume $\text{trace}(A) = 0$. Due to $A \neq 0$ there exists a unit vector x_1 with $\mu_1 := \langle Ax_1, x_1 \rangle \neq 0$. Choose any unitary U_1 with $U_1 \mathbf{e}_1 = x_1$. Then, the first diagonal entry of $U_1^* A U_1$ is $\mu_1 \neq 0$. Due to $\text{trace}(U_1^* A U_1) = \text{trace}(A) = 0$, the main lower-right $(n-1) \times (n-1)$ submatrix \widehat{A} of $U_1^* A U_1$, occupying the rows/columns $2, 3, \dots, n$ has a nonzero trace. By an induction argument, there exists an $(n-1) \times (n-1)$ unitary V such that $V \widehat{A} V^*$ has all diagonal entries equal and nonzero. Then, the unitary $U := (1 \oplus V) U_1$ does the job. \square

We denote by $\text{diagv } A$ the diagonal vector of $A \in M_n$: If $A = [a_{ij}]_{i,j=1}^n$, then $\text{diagv } A = [a_{11} \ a_{22} \ \dots \ a_{nn}]^{\text{tr}} \in \mathbb{C}^n$.

Theorem 3.2. *Let $A, B \in M_n$, where $n \geq 2$. Then the following three statements are equivalent:*

(i)
$$|\langle Ax, x \rangle| = |\langle Bx, x \rangle| \quad \text{for all } x \in \mathbb{C}^n. \quad (3.1)$$

(ii) *For each unitary V there exists a unimodular number $\gamma(V)$, and a map $h_V : \mathbb{C} \rightarrow \mathbb{C}$ which is either identity or complex conjugation, such that*

$$\text{diagv}(VBV^*) = \gamma(V) \text{diagv}(VAV^*)^{h_V}.$$

(iii) $B = \gamma A$ or $B = \gamma A^*$ for some unimodular number γ .

Proof of Theorem 3.2. The proof consists of several steps.

Step 1. (iii) \implies (ii) Trivial. The implication (ii) \implies (i) is also immediate: By scaling, it suffices to prove (i) only for vectors x of norm one; then apply (ii) with unitary V whose first row is x^* .

Step 2. We prove (ii) \implies (iii), for $n \geq 3$. If $A = 0$, the result follows immediately from Proposition 3.1(b). We assume therefore that $A \neq 0$.

We first show that map h_V is independent of the unitary V . So assume, to reach a contradiction, that

$$\text{diagv}(V_0 B V_0^*) = \gamma_0 \text{diagv}(V_0 A V_0^*) \notin \{e^{i\theta} \cdot \overline{\text{diagv}(V_0 A V_0^*)} : 0 \leq \theta \leq 2\pi\}, \quad (3.2)$$

for some unitary U_0 and unimodular γ_0 , while

$$\text{diagv}(V_1 B V_1^*) = \gamma_1 \overline{\text{diagv}(V_1 A V_1^*)} \notin \{e^{i\theta} \cdot \text{diagv}(V_1 A V_1^*), : 0 \leq \theta \leq 2\pi\} \quad (3.3)$$

for some other unitary U_1 and unimodular γ_1 . Choose hermitian S_0, S_1 with $e^{iS_0} = V_0$ and $e^{iS_1} = V_1$. Then, $V_t := e^{i(tS_1 + (1-t)S_0)}$ is a path that connects V_0 and V_1 in the set of unitaries. Clearly, V_t and $V_t^* = e^{-i(tS_1 + (1-t)S_0)}$ are both analytic functions of the real variable $t \in [0, 1]$. Moreover, $\mathbf{f}(t) := \text{diagv}(V_t A V_t^*)$, as well as $\mathbf{g}(t) := \text{diagv}(V_t B V_t^*)$ are also analytic vector-valued functions of real variable t . It is implicit in Eqs. (3.2)–(3.3) that $\mathbf{f}(0) \neq \mathbf{0}$ and $\mathbf{f}(1) \neq \mathbf{0}$. So at

least one, say the first one $a_1(t)$, component of a vector-valued function $\mathbf{f}(t)$ is not identically zero. Now, being analytic, $a_1(t)$ has only finitely many zeros on $[0, 1]$. In view of hypothesis (ii), the zeros of $a_1(t)$ precisely match those of $b_1(t)$, the first component of $\mathbf{g}(t)$. Moreover, at each t off the set Λ of their common zeros, one of $\gamma(t) := b_1(t)/a_1(t)$ and $\gamma_1(t) := b_1(t)/\overline{a_1(t)}$ is unimodular. Clearly then, both are unimodular for all t off the common zeros. Then, however, they must have only removable singularities at common zeros, so both $\gamma(t)$ and $\gamma_1(t)$ are analytic functions of $t \in [0, 1]$.

We next rewrite hypothesis (ii) into

$$\|\mathbf{g}(t) - \gamma(t)\mathbf{f}(t)\|^2 \cdot \|\mathbf{g}(t) - \gamma_1(t)\overline{\mathbf{f}(t)}\|^2 \equiv 0, \quad t \in [0, 1] \setminus \Lambda. \quad (3.4)$$

Both factors in the left hand side of (3.4) are analytic functions of a real variable t . We therefore conclude that at least one of them must vanish identically. Suppose the first one does, i.e. $\mathbf{g}(t) - \gamma(t)\mathbf{f}(t) \equiv 0$. Then, however, $\text{diagv}(V_t B V_t^*) = \gamma(t) \text{diagv}(V_t B V_t^*)$ for each t , contradicting Eq. (3.3). Likewise we get a contradiction if $(\mathbf{g}(t) - \gamma_1(t)\overline{\mathbf{f}(t)}) \equiv 0$.

If necessary, we replace B with B^* . In doing so, we can now guarantee that for each unitary V ,

$$\text{diagv}(V B V^*) = \gamma(V) \text{diagv}(V A V^*), \quad |\gamma(V)| = 1. \quad (3.5)$$

We next show the unimodular factor $\gamma(V)$ is independent of V . If the trace of A is nonzero, this is obvious: $\gamma(V) = \text{trace}(B)/\text{trace}(A)$, by (3.5). Thus, assume $\text{trace}(A) = 0$. By Proposition 3.1, there is a unitary $U \in M_n$ such that $U A U^*$ has nonzero diagonal entries μ_1, \dots, μ_n . We may assume that $U = I$; otherwise, replace (A, B) by $(U A U^*, U B U^*)$. Let $A = (a_{ij})$ and $B = (b_{ij})$. The hypothesis (ii), and the above consideration, now imply $\text{diagv} B = \gamma \text{diagv} A$. We may assume that $\gamma = 1$; otherwise, replace B by $\bar{\gamma} B$. Thus,

$$\text{diagv} B = \text{diagv} A = [\mu_1, \dots, \mu_n]^{\text{tr}}, \quad \mu_1, \dots, \mu_n \in \mathbb{C} \setminus \{0\}. \quad (3.6)$$

For $k = 1, \dots, n$, let $A_k \in M_{n-1}$ be the submatrix of A obtained from it by removing its k th row and k th column. Similarly, we construct the matrices B_1, \dots, B_n .

We claim that A_k and B_k are the same for all $k = 1, \dots, n$. It will then follow that A and B are the same (the hypothesis that $n \geq 3$ is used here), and, in particular, $\gamma(V) = 1$ for all unitary V , which will complete the proof. To prove our claim, let $V \in M_n$ be a unitary matrix with (k, k) entry equal to 1. Since $V A V^*$ and $V B V^*$ have the same nonzero (k, k) entry μ_k (by (3.6)), we see from Eq. (3.5), that the two matrices actually have the same corresponding diagonal entries. As a result, $U A_k U^*$ and $U B_k U^*$ have the same diagonal entries for all unitary matrices $U \in M_{n-1}$. So $\text{diagv}(U(A_k - B_k)U^*) = 0$ for all unitaries, which implies numerical range of $A_k - B_k$ consists only of zero. Thus, $A_k = B_k$ for $k = 1, 2, \dots, n$, as required.

Step 3. (i) \implies (iii), for $n = 2$.

If $A = 0$ then $\langle Bx, x \rangle = 0$ for all $x \in \mathbb{C}^2$, so $B = 0$, and we are done. Otherwise, by (a) of Proposition 3.1, there exists a unitary U such that all diagonal entries of U^*AU are nonzero. Obviously,

$$|\langle U^*AUx, x \rangle| = |\langle A(Ux), Ux \rangle| = |\langle B(Ux), Ux \rangle| = |\langle U^*BUx, x \rangle|.$$

Consequently, we may replace (A, B) with (U^*AU, U^*BU) without changing the validity of assumptions (i) and conclusion (iii). This way, $a_{11} \neq 0$ (we denote by a_{11} , resp., b_{11} , the top left entry of A , resp., B). Choose a vector $x := \mathbf{e}_1$ to deduce $|a_{11}| = |b_{11}|$. We may, thus, further replace (A, B) with $(1/a_{11}U, \gamma B/a_{11})$ where $\gamma := a_{11}/b_{11}$ is unimodular. In doing so, we can assume $a_{11} = 1 = b_{11}$. Hence it remains to see that $B = A$ or $B = A^*$.

To see this, write

$$A := \begin{bmatrix} 1 & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B := \begin{bmatrix} 1 & b_{12} \\ b_{21} & b_{22} \end{bmatrix},$$

and choose a vector $x := \begin{bmatrix} 1 \\ r + is \end{bmatrix}$, where r, s are real. Then,

$$\begin{aligned} \langle Ax, x \rangle &= 1 + (r + is)a_{12} + \overline{(r + is)}a_{21} + |(r + is)|^2 a_{22} \\ &= 1 + r\operatorname{Re}(a_{12} + a_{21}) + s\operatorname{Im}(a_{21} - a_{12}) + (r^2 + s^2)\operatorname{Re}(a_{22}) \\ &\quad + i(r\operatorname{Im}(a_{12} + a_{21}) + s\operatorname{Re}(a_{12} - a_{21}) + (r^2 + s^2)\operatorname{Re}(a_{22})). \end{aligned} \quad (3.7)$$

Tedious, but straightforward computation shows that

$$\begin{aligned} |\langle Ax, x \rangle|^2 &= 1 + (r^2 + s^2)^2 |a_{22}|^2 + 2\operatorname{Re}(\overline{a_{22}}(a_{12} + a_{21})) r^3 \\ &\quad - 2\operatorname{Im}(\overline{a_{22}}(a_{12} - a_{21})) r^2 s + (|a_{12} + a_{21}|^2 + 2\operatorname{Re} a_{22}) r^2 \\ &\quad - 4\operatorname{Im}(\overline{a_{21}}a_{12}) rs + 2(\operatorname{Re}(a_{12} + a_{21})) r \\ &\quad + 2\operatorname{Re}(\overline{a_{22}}(a_{12} + a_{21})) r s^2 - 2\operatorname{Im}(\overline{a_{22}}(a_{12} - a_{21})) s^3 \\ &\quad - 2(\operatorname{Im}(a_{12} - a_{21})) s + (|a_{12} - a_{21}|^2 + 2\operatorname{Re} a_{22}) s^2. \end{aligned}$$

Comparing the coefficients with the corresponding formula for $|\langle Bx, x \rangle|^2$ gives the following set of equations:

$$|b_{22}|^2 = |a_{22}|^2 \quad (3.8)$$

$$\operatorname{Re}(\overline{b_{22}}(b_{12} + b_{21})) = \operatorname{Re}(\overline{a_{22}}(a_{12} + a_{21})) \quad (3.9)$$

$$\operatorname{Im}(\overline{b_{22}}(b_{12} - b_{21})) = \operatorname{Im}(\overline{a_{22}}(a_{12} - a_{21})) \quad (3.10)$$

$$\operatorname{Re}(b_{12} + b_{21}) = \operatorname{Re}(a_{12} + a_{21}) \quad (3.11)$$

$$\operatorname{Im}(b_{12} - b_{21}) = \operatorname{Im}(a_{12} - a_{21}) \quad (3.12)$$

$$|b_{12} + b_{21}|^2 + 2\operatorname{Re} b_{22} = |a_{12} + a_{21}|^2 + 2\operatorname{Re} a_{22} \quad (3.13)$$

$$|b_{12} - b_{21}|^2 + 2\operatorname{Re} b_{22} = |a_{12} - a_{21}|^2 + 2\operatorname{Re} a_{22} \quad (3.14)$$

$$\operatorname{Im}(\overline{b_{21}}b_{12}) = \operatorname{Im}(\overline{a_{21}}a_{12}). \quad (3.15)$$

Subtracting (3.14) from (3.13) gives, after an easy simplification,

$$4\operatorname{Re}(\overline{a_{21}}a_{12}) = 4\operatorname{Re}(\overline{b_{21}}b_{12}). \quad (3.16)$$

Decompose now $a_{12} = z_1 + iz_2$, and $a_{21} = y_1 + iy_2$, with z_1, z_2, y_1, y_2 real, and $b_{12} = \tilde{z}_1 + i\tilde{z}_2$, etc. Then, Eqs. (3.11)–(3.12), (3.15)–(3.16) give:

$$\tilde{y}_1 + \tilde{z}_1 = z_1 + y_1$$

$$\tilde{z}_2 - \tilde{y}_2 = z_2 - y_2$$

$$\tilde{y}_1\tilde{z}_1 + \tilde{y}_2\tilde{z}_2 = z_1y_1 + z_2y_2 \quad (3.17)$$

$$z_2y_1 - z_1y_2 = \tilde{y}_1\tilde{z}_2 - \tilde{y}_2\tilde{z}_1. \quad (3.18)$$

From the first two equations we get

$$\tilde{y}_1 = y_1 + z_1 - \tilde{z}_1 \quad \text{and} \quad \tilde{y}_2 = y_2 - z_2 + \tilde{z}_2. \quad (3.19)$$

Substitute this into (3.17), (3.18), and simplify, to get

$$(y_1 - \tilde{z}_1)(\tilde{z}_1 - z_1) = (z_2 - \tilde{z}_2)(y_2 + \tilde{z}_2) \quad (3.20)$$

$$z_2(y_1 - \tilde{z}_1) + y_2(\tilde{z}_1 - z_1) = (y_1 + z_1 - 2\tilde{z}_1)\tilde{z}_2. \quad (3.21)$$

We are now facing two possibilities:

Possibility 1. $\tilde{z}_1 = z_1$. Then, the last two equations further simplify into $(z_2 - \tilde{z}_2)(y_2 + \tilde{z}_2) = 0$, respectively, $(y_1 - z_1)z_2 = (y_1 - z_1)\tilde{z}_2$. So, either $\tilde{z}_2 = z_2$ or else $(y_1, y_2) = (z_1, -\tilde{z}_2)$. In the former case, Eq. (3.19) brings $(\tilde{y}_1, \tilde{y}_2) = (y_1, y_2)$, so

$$b_{12} = \tilde{z}_1 + i\tilde{z}_2 = a_{12}, \quad b_{21} = \tilde{y}_1 + i\tilde{y}_2 = a_{21}.$$

In the latter case, we similarly deduce $y_1 = z_1 = \tilde{z}_1 = \tilde{y}_1$ and $y_2 = -\tilde{z}_2$, and $\tilde{y}_2 = -z_2$. Therefore,

$$b_{12} = \overline{a_{21}}, \quad b_{21} = \overline{a_{12}}. \quad (3.22)$$

Possibility 2. $\tilde{z}_1 \neq z_1$. Then, (3.20) gives

$$y_1 = \frac{(\tilde{z}_1 - z_1)\tilde{z}_1 + (z_2 - \tilde{z}_2)(y_2 + \tilde{z}_2)}{\tilde{z}_1 - z_1}.$$

This simplifies the remaining (3.21) into

$$\frac{\left((z_1 - \tilde{z}_1)^2 + (z_2 - \tilde{z}_2)^2\right)(y_2 + \tilde{z}_2)}{z_1 - \tilde{z}_1} = 0.$$

Note that the sum of squares in the left factor is nonzero because $z_1 - \tilde{z}_1 \neq 0$. Hence, $\tilde{z}_2 = -y_2$. From the previous equation we now read $\tilde{z}_1 = y_1$. Moreover, Eq. (3.19) brings $\tilde{y}_1 = z_1$ and $\tilde{y}_2 = -z_2$. Therefore, $b_{12} = \overline{a_{21}}$ and $b_{21} = \overline{a_{12}}$, as in Eq. (3.22).

It only remains to compare b_{22} with a_{22} . Now, since

$$|\langle Bx, x \rangle| = |(\langle Bx, x \rangle)^*| = |\langle B^*x, x \rangle|,$$

we are free to replace (A, B) with (A, B^*) . This way, we can always assume $b_{12} = a_{12}$ and $b_{21} = a_{21}$. Hence, we are done if we also show $b_{22} = a_{22}$.

To this end, Eq. (3.14) immediately gives $\operatorname{Re} b_{22} = \operatorname{Re} a_{22}$, while from $|b_{22}| = |a_{22}|$ we deduce that either $b_{22} = a_{22}$ or else $b_{22} = \overline{a_{22}} \neq a_{22}$. In the former case we are done. In the latter case, consider Eqs. (3.9)–(3.10) with $b_{22} := \overline{a_{22}}$. Simplifying these equations yields

$$\operatorname{Im} (a_{12} + a_{21}) \operatorname{Im} a_{22} = 0 = \operatorname{Re} (a_{12} - a_{21}) \operatorname{Im} a_{22}.$$

We may divide by nonzero $\operatorname{Im} a_{22}$. What we end up with easily simplifies into $a_{12} = \overline{a_{21}}$. Then, however,

$$A^* = \begin{bmatrix} 1 & \overline{a_{21}} \\ a_{21} & a_{22} \end{bmatrix}^* = \begin{bmatrix} 1 & a_{21} \\ a_{21} & \overline{a_{22}} \end{bmatrix} = B,$$

which completes the proof of Step 3.

Step 4. Assuming (i) holds, we will prove that there exists a unimodular complex number γ such that either $\operatorname{diagv}(A) = \gamma \operatorname{diagv}(B)$ or else $\operatorname{diagv}(A) = \overline{\gamma} \operatorname{diagv}(B)$.

Let $A = [a_{ij}]_{i,j=1}^n$, $B = [b_{ij}]_{i,j=1}^n$. Choose any pair of distinct indices (i, j) , and let $x := \lambda \mathbf{e}_i + \mu \mathbf{e}_j$ be in the subspace spanned by $\mathbf{e}_i, \mathbf{e}_j$. Then, $\langle Ax, x \rangle = \langle A_{ij} z, z \rangle$, where $z := [\lambda, \mu]^{\operatorname{tr}}$, and A_{ij} is the 2×2 matrix formed by the i th and j th rows and columns of A . The identity (3.1) then reduces to

$$|\langle A_{ij} z, z \rangle| \equiv |\langle B_{ij} z, z \rangle|.$$

Here, B_{ij} is the 2×2 matrix formed by the i th and j th rows and columns of B . By Step 3, $B_{ij} \in \{\gamma A_{ij}, \overline{\gamma} \overline{A_{ij}}\}$, where γ is a unimodular number. Considering diagonal entries yields

$$(b_{ii}, b_{jj}) = \gamma(a_{ii}, a_{jj}) \quad \text{or} \quad (b_{ii}, b_{jj}) = \overline{\gamma}(\overline{a_{ii}}, \overline{a_{jj}}). \quad (3.23)$$

Consequently, either $\operatorname{diagv}(A) = 0 = \operatorname{diagv}(B)$ or else both diagonals have at least one nonzero entry. In the former case we are done.

In the latter case, we assume for simplicity the $(1, 1)$ entries of A and B are nonzero. Since $|a_{11}| = |b_{11}|$ we may replace (A, B) with $(A/a_{11}, \gamma B/a_{11})$ where $\gamma := a_{11}/b_{11}$ is unimodular. The identity (3.1) as well as the end result will not change. This way we achieve $a_{11} = 1 = b_{11}$. Moreover, when $i = 1$ Eq. (3.23) yields

$$(1, b_{jj}) \in \{(1, a_{jj}), \overline{(1, a_{jj})}\}.$$

Hence, it remains to see that $\operatorname{diagv}(A) = \operatorname{diagv}(B)$ or $\operatorname{diagv}(A) = \overline{\operatorname{diagv}(B)}$.

Now, arguing by contradiction, suppose that

$$(1, b_{i_0 i_0}) = (1, a_{i_0 i_0}) \neq \overline{(1, a_{i_0 i_0})}, \quad (1, b_{i_1 i_1}) = \overline{(1, a_{i_1 i_1})} \neq (1, a_{i_1 i_1}),$$

for two different indices i_0 and i_1 . This is possible only when $b_{i_0 i_0} = a_{i_0 i_0}$ and $b_{i_1 i_1} = \overline{a_{i_1 i_1}}$ are both nonreal (hence also nonzero). Now, by Eq.(3.23),

$$(b_{i_0 i_0}, b_{i_1 i_1}) \in \{\gamma(a_{i_0 i_0}, a_{i_1 i_1}), \overline{\gamma}(\overline{a_{i_0 i_0}}, \overline{a_{i_1 i_1}})\}$$

implies

$$\frac{b_{i_0 i_0}}{b_{i_1 i_1}} \in \left\{ \frac{a_{i_0 i_0}}{a_{i_1 i_1}}, \overline{\left(\frac{a_{i_0 i_0}}{a_{i_1 i_1}} \right)} \right\}. \quad (3.24)$$

On the other hand, $\frac{b_{i_0 i_0}}{b_{i_1 i_1}} = \frac{a_{i_0 i_0}}{a_{i_1 i_1}}$, and in view of (3.24) we obtain either $a_{i_1 i_1} = \overline{a_{i_0 i_0}}$ or else $a_{i_0 i_0} = \overline{a_{i_0 i_0}}$. This is the desired contradiction.

Therefore, either $(1, b_{jj}) = (1, a_{jj})$ for all j or else $(1, b_{jj}) = \overline{(1, a_{jj})}$ for all j . In the first case, $\text{diagv}(A) = \text{diagv}(B)$ while in the second one, $\text{diagv}(A) = \overline{\text{diagv}(B)}$.

Step 5. (i) \implies (iii), for $n \geq 3$.

Fix any unitary U and consider $(A_U, B_U) := (U^*AU, U^*BU)$. Clearly,

$$|\langle A_U x, x \rangle| = |\langle A(Ux), Ux \rangle| = |\langle B(Ux), Ux \rangle| = |\langle B_U x, x \rangle|.$$

Then apply the result of Step 4 to (A_U, B_U) . We see that

$$\text{diagv}(U^*BU) = \gamma(U)\text{diagv}(U^*AU) \quad \text{or} \quad \text{diagv}(U^*BU) = \overline{\gamma(U)\text{diagv}(U^*AU)}$$

for each unitary U . By Step 2, $B = \gamma A$ or else $B = \overline{\gamma}A^*$, as required.

This completes the proof of Theorem 3.2. \square

Proof of Theorem 2.2 in case $t = 0$. We prove the nontrivial ‘‘only if’’ part. We may assume $A \neq 0$, $B \neq 0$. Multiplying A and B by nonzero complex numbers of the same absolute value, we may further suppose that

$$\langle Ae, e \rangle = \langle Be, e \rangle = 1 \quad (3.25)$$

for a fixed normalized element $e \in \mathbb{H}$. If $A \neq B$ and also $A \neq B^*$, then we have

$$\langle Af_1, f_1 \rangle \neq \langle Bf_1, f_1 \rangle \quad \text{and} \quad \langle Af_2, f_2 \rangle \neq \langle B^*f_2, f_2 \rangle \quad (3.26)$$

for some elements $f_1, f_2 \in \mathbb{H}$.

On the other hand, let P be the selfadjoint projection on the finite dimensional subspace $\mathbb{H}_1 \subset \mathbb{H}$, generated by e and f_1, f_2 , and let $\widehat{A} := PAP$ and $\widehat{B} := PBP$ be the operators acting on \mathbb{H}_1 . Clearly, $\langle Ag, g \rangle = \langle \widehat{A}g, g \rangle$ for any element $g \in \mathbb{H}_1$; likewise for \widehat{B} . Hence, by the assumptions, $|\langle \widehat{A}g, g \rangle| = |\langle \widehat{B}g, g \rangle|$ for every $g \in \mathbb{H}_1$. Therefore, by Theorem 3.2, we must have $\widehat{B} = \gamma\widehat{A}$ or $\widehat{B} = \overline{\gamma}\widehat{A}^*$. Actually, $\gamma = 1$, by Eq. (3.25). Then however,

$$\langle Af_1, f_1 \rangle = \langle \widehat{A}f_1, f_1 \rangle = \langle \widehat{B}f_1, f_1 \rangle = \langle Bf_1, f_1 \rangle$$

(respectively, $\langle Af_1, f_1 \rangle = \langle B^*f_2, f_2 \rangle$, if $\widehat{B} = \widehat{A}^*$), a contradiction with (3.26). \square

4. Proof of Theorem 2.2: $t = \pi/2$

Assume $A, B \in L(\mathbf{H})$ are such that

$$|\langle Ax, y \rangle| = |\langle Bx, y \rangle|, \quad \forall \text{ orthonormal pairs } (x, y), \quad x, y \in \mathbf{H}. \quad (4.1)$$

We proceed in steps to derive (3) of Theorem 2.2.

Step 1. Suppose that the implication (4.1) \implies Theorem 2.2(3) has been proved for \mathbb{C}^2 and \mathbb{C}^3 . We will prove the implication for general Hilbert space \mathbf{H} .

We may assume B is not scalar (otherwise Ax is orthogonal to $\text{span}\{x\}^\perp$ so $Ax = \lambda_x x$, and we are done as in the proof of Proposition 2.1). Therefore, there exists a normalized element $x \in \mathbf{H}$ such that Bx is not a scalar multiple of x , and hence there is an orthonormal pair (x, y) such that $\langle Bx, y \rangle \neq 0$. Let $\Omega := \{x, y, z\}$ be an orthonormal triple, where x and y are fixed, and let P be the orthogonal projection on $\text{span}\Omega$. By considering operators PAP and PBP on $\text{span}\Omega$ and using the supposition, we see that

$$PAP = \mu_\Omega PBP + \nu_\Omega P, \quad \mu_\Omega, \nu_\Omega \in \mathbb{C}, \quad |\mu_\Omega| = 1. \quad (4.2)$$

In fact, μ_Ω and ν_Ω are independent of Ω . Indeed, for two orthonormal triples Ω and Ω' we have in view of (4.2):

$$\begin{bmatrix} \mu_\Omega \langle Bx, x \rangle + \nu_\Omega & \mu_\Omega \langle By, x \rangle \\ \mu_\Omega \langle Bx, y \rangle & \mu_\Omega \langle By, y \rangle + \nu_\Omega \end{bmatrix} = \begin{bmatrix} \mu_{\Omega'} \langle Bx, x \rangle + \nu_{\Omega'} & \mu_{\Omega'} \langle By, x \rangle \\ \mu_{\Omega'} \langle Bx, y \rangle & \mu_{\Omega'} \langle By, y \rangle + \nu_{\Omega'} \end{bmatrix}.$$

Since $\langle Bx, y \rangle \neq 0$, we obtain $\mu_\Omega = \mu_{\Omega'}$ and $\nu_\Omega = \nu_{\Omega'}$. Thus,

$$PAP = \mu PBP + \nu P, \quad \mu, \nu \in \mathbb{C}, \quad |\mu| = 1. \quad (4.3)$$

Since any element $z \in \mathbf{H}$ can be included in the range of P , for some orthonormal triple Ω , we obtain from (4.3):

$$\langle Az, z \rangle = \mu \langle Bz, z \rangle + \nu \langle z, z \rangle, \quad \forall z \in \mathbf{H},$$

and (3) of Theorem 2.2 follows.

Step 2. We prove the implication (4.1) \implies Theorem 2.2(3) for \mathbb{C}^2 and \mathbb{C}^3 .

Applying simultaneous unitary similarity and addition of scalar matrices to A and to B we may assume that

$$A = [a_{ij}]_{i,j=1}^n, \quad B = [b_{ij}]_{i,j=1}^n, \quad a_{ij}, b_{ij} \in \mathbb{C},$$

where A is upper triangular, $a_{11} = 0$, a_{12}, \dots, a_{1n} are nonnegative and $b_{11} = 0$. (We need only the cases $n = 2, 3$, but this transformation can be applied for $L(\mathbb{C}^n)$ for any integer $n \geq 2$.) Applying (4.1) with $x = \mathbf{e}_i$, $y = \mathbf{e}_j$, $i < j$, we see that B is also upper triangular. Applying (4.1) with $x = \mathbf{e}_i$, $y = \mathbf{e}_j$, $i > j$, we see that $|b_{ij}| = |a_{ij}|$ for all $i < j$.

We proceed with $n = 2$. If $a_{12} = 0$ then also $b_{12} = 0$ in which case $A = \text{diag}(0, a_{22})$ and $B = \text{diag}(0, b_{22})$. With orthonormal

$$x = \begin{bmatrix} \cos t \\ \sin t \end{bmatrix}, \quad y = \begin{bmatrix} -\sin t \\ \cos t \end{bmatrix},$$

(4.1) easily gives $|a_{22}| = |b_{22}|$, and we are done. If $a_{12} \neq 0$ we further assume (replacing B with $e^{is}B$ for some real s) that $b_{12} = a_{12}$. So under $n = 2$ we are left with

Case (a).

$$A = \begin{bmatrix} 0 & a_{12} \\ 0 & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & a_{12} \\ 0 & b_{22} \end{bmatrix},$$

where $a_{12} > 0$.

Multiplying (4.1) with suitable scalar we see it holds for any (possibly not normalized) orthogonal vectors x, y . Apply (4.1) with $x = \begin{bmatrix} z \\ 1 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ -\bar{z} \end{bmatrix}$, $z \in \mathbb{C}$. We obtain:

$$(a_{12} - a_{22}z)(a_{12} - \overline{a_{22}z}) = (a_{12} - b_{22}z)(a_{12} - \overline{b_{22}z}), \quad \forall z \in \mathbb{C},$$

which yields $|a_{22}| = |b_{22}|$ and $a_{12}a_{22} = a_{12}b_{22}$. So, $a_{22} = b_{22}$ hence $A = B$, which proves case (a).

Next assume $n = 3$. If $a_{12} = 0 = a_{13}$ (and hence also $b_{12} = 0 = b_{13}$) then Corollary 2.4 is applicable to the already proven case of 2×2 matrices $[a_{ij}]_{i,j=2}^3$ and $[b_{ij}]_{i,j=2}^3$, and we are done using induction on n . Thus, we can assume that not both a_{12}, a_{13} are zeros, and letting a_{1r} be the first positive number among a_{12}, a_{13} , we further assume (replacing B with $e^{is}B$ for some real s) that $b_{1r} = a_{1r}$. So we are left with the following two cases to consider:

Case (b).

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & a_{12} & b_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{bmatrix},$$

where $a_{12} > 0$, $a_{13} \geq 0$, $|b_{13}| = a_{13}$, $|b_{23}| = |a_{23}|$;

Case (c).

$$A = \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & a_{13} \\ 0 & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{bmatrix},$$

where $a_{13} > 0$, $|b_{23}| = |a_{23}|$.

Consider Case (b). Applying the proof of Case (a) to the upper left 2×2 submatrices of A and B , we see that $a_{22} = b_{22}$. Now use (4.1) with

$$x = \begin{bmatrix} 0 \\ 1 \\ z \end{bmatrix}, \quad y = \mathbf{e}_1.$$

We obtain:

$$|a_{12} + a_{13}z|^2 = |a_{12} + b_{13}z|^2, \quad \forall z \in \mathbb{C},$$

and hence $a_{13}a_{12} = b_{13}a_{12}$. Since $a_{12} \neq 0$, we have $a_{13} = b_{13}$. Analogous consideration of (4.1) with

$$x = \mathbf{e}_3, \quad y = \begin{bmatrix} 1 \\ \bar{z} \\ 0 \end{bmatrix}$$

yields $a_{23}a_{13} = b_{23}a_{13}$. Thus, either $a_{13} = b_{13} = 0$, or $a_{13} = b_{13} \neq 0$ and $a_{23} = b_{23}$.

Therefore, one of the following three situations occurs:

- (b1) $a_{13} = b_{13} \neq 0$, $a_{23} = b_{23}$;
- (b2) $a_{13} = b_{13} = 0$, $a_{23} \neq 0$ (then also $b_{23} \neq 0$);
- (b3) $a_{13} = b_{13} = 0$, $a_{23} = b_{23} = 0$.

If (b1) holds, then the proof of Case (a) applied to the 2×2 principal submatrices of A and B in the 1st and 3rd rows and columns yields $a_{33} = b_{33}$, i.e., $A = B$. If (b3) holds, then we apply (4.1) with

$$x = \begin{bmatrix} w \\ 1 \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ 0 \\ -\bar{w} \end{bmatrix}, \quad w \in \mathbb{C},$$

resulting in

$$|a_{12} - a_{33}w|^2 = |a_{12} - b_{33}w|^2, \quad \forall w \in \mathbb{C}.$$

It follows that $|a_{33}| = |b_{33}|$ and

$$\operatorname{Re}(b_{33}a_{12}w - a_{33}a_{12}w) = 0, \quad \forall w \in \mathbb{C}.$$

Since $a_{12} \neq 0$, we conclude $a_{33} = b_{33}$, thus $A = B$. Finally, if (b2) holds, then we apply (4.1) with

$$x = \begin{bmatrix} -w \\ 1 \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ \bar{w} \\ 0 \end{bmatrix}, \quad w \in \mathbb{C}.$$

It follows that

$$|a_{12} + (a_{22} + a_{23})w|^2 = |a_{12} + (a_{22} + b_{23})w|^2.$$

In particular,

$$\operatorname{Re}((a_{22} + a_{23})wa_{12} - (a_{22} + b_{23})wa_{12}) = 0,$$

and since this equality holds for all $w \in \mathbb{C}$, we obtain $a_{23} = b_{23}$. Now we apply the proof of Case (a) to the lower right 2×2 submatrices of $A - a_{22}I$ and $B - a_{22}I$, and the equality $A = B$ follows.

This concludes our consideration of Case (b).

Consider now Case (c). Applying the proof of Case (a) to the 2×2 submatrices of A and B generated by the 1st and 3rd rows and columns, we see that $b_{33} = a_{33}$. Next, apply (4.1) with

$$x = \begin{bmatrix} -zp \\ z \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ \bar{p} \\ 0 \end{bmatrix}, \quad z, p \in \mathbb{C}.$$

It follows that

$$|a_{13} + (a_{22}z + a_{23})p|^2 = |a_{13} + (b_{22}z + b_{23})p|^2, \quad \forall z, p \in \mathbb{C}.$$

Consider this as a polynomial of the real and imaginary parts of p , with z as a parameter. In particular,

$$\operatorname{Re}((a_{22}z + a_{23})pa_{13} - (b_{22}z + b_{23})pa_{13}) = 0.$$

Since $a_{13} \neq 0$ and the equality holds for all $p \in \mathbb{C}$, we have

$$a_{22}z + a_{23} = b_{22}z + b_{23}, \quad \forall z \in \mathbb{C}.$$

Clearly, $a_{22} = b_{22}$ and $a_{23} = b_{23}$. This completes the proof of Step 2. \square

5. Proof of Theorem 2.2: $0 < t < \pi/2$

Again, we prove only the nontrivial ‘‘only if’’ part.

Step 1. Assume the dimension of H is at least 3.

Let (x, y) be an orthonormal pair. Then

$$|\langle Ax, x + (\tan t)ye^{is} \rangle| = |\langle Bx, x + (\tan t)ye^{is} \rangle|, \quad \forall s \in [0, 2\pi). \quad (5.1)$$

Consider the two circles

$$C_A := \{\langle Ax, x + (\tan t)ye^{is} \rangle : 0 \leq s < 2\pi\},$$

$$C_B := \{\langle Bx, x + (\tan t)ye^{is} \rangle : 0 \leq s < 2\pi\}$$

with centers and radii $\langle Ax, x \rangle$, $(\tan t)|\langle Ax, y \rangle|$, and $\langle Bx, x \rangle$, $(\tan t)|\langle Bx, y \rangle|$, respectively. Condition (5.1) implies that

$$\min_{z \in C_A} |z| = \min_{z \in C_B} |z| \quad \text{and} \quad \max_{z \in C_A} |z| = \max_{z \in C_B} |z|,$$

and therefore

$$|\langle Ax, x \rangle| + (\tan t)|\langle Ax, y \rangle| = |\langle Bx, x \rangle| + (\tan t)|\langle Bx, y \rangle|$$

and

$$| |\langle Ax, x \rangle| - (\tan t)|\langle Ax, y \rangle| | = | |\langle Bx, x \rangle| - (\tan t)|\langle Bx, y \rangle| |.$$

We see that one of the two possibilities holds: either

(a)

$$|\langle Ax, x \rangle| = |\langle Bx, x \rangle| \quad \text{and} \quad |\langle Ax, y \rangle| = |\langle Bx, y \rangle| \quad (5.2)$$

(this happens if the origin is not situated inside one of the circles C_A and C_B and outside of the other circle);

or

(b) there exist positive numbers $p \neq q$ such that $|\langle Ax, x \rangle| = p$, $|\langle Bx, x \rangle| = q$, $(\tan t)|\langle Ax, y \rangle| = q$, $(\tan t)|\langle Bx, y \rangle| = p$ (this happens if the origin is situated inside one circle and outside of the other).

Clearly, for every fixed normalized $x \in \mathbf{H}$, either (a) holds for all $y \in \mathbf{H}$ such that (x, y) form an orthonormal pair, or (b) holds for all such y . We claim that (b) is not possible (here we use the hypothesis that $\dim \mathbf{H} \geq 3$). Indeed, under (b) we have

$$|\langle Ax, y \rangle| = |\langle Bx, x \rangle|(\tan t)^{-1} \neq 0 \quad (5.3)$$

for every normalized y orthogonal to x . If y_1, y_2 are orthonormal elements both orthogonal to x , then there is a nonzero linear combination of y_1, y_2 which is orthogonal to Ax , a contradiction with (5.3). Thus, we have (a) for every orthonormal pair (x, y) , $x, y \in \mathbf{H}$, and by the part of Theorem 2.2 for the cases $t = 0$ and $t = \pi/2$, we obtain $B = \mu A$ or $B = \mu A^*$ for some $\mu \in \mathbf{C}$, $|\mu| = 1$, as well as $B = \gamma A + \nu I$ for some $\gamma, \nu \in \mathbf{C}$ with $|\gamma| = 1$.

We claim that $B = \mu' A$, for some $\mu' \in \mathbf{C}$ with $|\mu'| = 1$, always holds. Indeed, suppose $B = \mu A^*$, $|\mu| = 1$. Without loss of generality we may take $\mu = 1$. Taking squares in (5.1), and using (5.2), we obtain for every orthonormal pair (x, y) and every s , $0 \leq s < 2\pi$:

$$\operatorname{Re} \left(\langle Ax, ye^{is} \rangle \cdot \overline{\langle Ax, x \rangle} \right) = \operatorname{Re} \left(\langle Bx, ye^{is} \rangle \cdot \overline{\langle Bx, x \rangle} \right).$$

Thus,

$$\langle Ax, y \rangle \cdot \overline{\langle Ax, x \rangle} = \langle Bx, y \rangle \cdot \overline{\langle Bx, x \rangle}.$$

Substituting in this equality $B = \gamma A + \nu I$, we have

$$\langle Ax, y \rangle \cdot \overline{\langle Ax, x \rangle} = \gamma \langle Ax, y \rangle \overline{(\gamma \langle Ax, x \rangle + \nu)}. \quad (5.4)$$

If x is not an eigenvector of A , then we can take $y \not\perp Ax$, and (5.4) gives

$$\langle Ax, x \rangle = \langle Ax, x \rangle + \bar{\gamma}\nu,$$

thus $\nu = 0$ and we are done. If every normalized $x \in \mathbf{H}$ is an eigenvector of A , then $A = zI$, $z \in \mathbf{C}$, and

$$B = A^* = \bar{z}I = \frac{\bar{z}}{z}A,$$

and we are done again (the case $z = 0$ is trivial).

Step 2. Assume $\mathbf{H} = \mathbf{C}^2$.

We need to show that, for fixed $A, B \in M_2$, the equality

$$|x^* Ax + (\tan t)y^* Ax| = |x^* Bx + (\tan t)y^* Bx| \quad (5.5)$$

for every orthonormal pair (x, y) , $x, y \in \mathbb{C}^2$, implies

$$A = \mu B \quad \text{for some unimodular } \mu. \quad (5.6)$$

We consider a special case first.

Case 1. Suppose

$$A = \begin{bmatrix} 1 & a_1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \nu & b_1 \\ 0 & 0 \end{bmatrix}, \quad |\nu| = 1,$$

and (5.5) holds. We may assume $\nu = 1$. Write

$$a_1 = a'_1 + ia''_1, \quad b_1 = b'_1 + ib''_1,$$

where a'_1, a''_1, b'_1, b''_1 are real, and let $\tau = \tan t$. Applying (5.5) to the orthogonal pair

$$x = \begin{bmatrix} u + iv \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 1 \\ -u + iv \end{bmatrix}, \quad u, v \in \mathbb{R},$$

(note that x and y have equal lengths, and therefore (5.5) is applicable), we obtain

$$|u^2 + v^2 + (u - iv)a_1 + \tau(u + iv) + \tau a_1| = |u^2 + v^2 + (u - iv)b_1 + \tau(u + iv) + \tau b_1|.$$

Taking squares in this equality, and expressing the modulus squared of a complex number as the sum of squares of its real and imaginary parts, yields

$$\begin{aligned} & (u^2 + v^2 + ua'_1 + va''_1 + \tau u + \tau a'_1)^2 + (ua''_1 - va'_1 + \tau v + \tau a''_1)^2 \\ &= (u^2 + v^2 + ub'_1 + vb''_1 + \tau u + \tau b'_1)^2 + (ub''_1 - vb'_1 + \tau v + \tau b''_1)^2. \end{aligned} \quad (5.7)$$

This equality holds for all real u, v , and both sides are polynomials in u, v . Equating the coefficients of u^3 in both sides of (5.7) gives $2(a'_1 + \tau) = 2(b'_1 + \tau)$, and equating the coefficients of v^3 gives $2a''_1 = 2b''_1$. Thus, $a_1 = b_1$, as required.

To continue with the proof of Step 2, we bring a general fact. Given fixed $\alpha, \beta, \gamma, \delta \in \mathbb{C}$, assume the identity

$$|\alpha + e^{i\xi}\beta| = |\gamma + e^{i\xi}\delta|, \quad \xi \in \mathbb{R}, \quad (5.8)$$

holds. Note that (5.8) is equivalent to

$$|\alpha|^2 + |\beta|^2 - |\gamma|^2 - |\delta|^2 + 2\operatorname{Re}(e^{i\xi}(\beta\bar{\alpha} - \delta\bar{\gamma})) = 0. \quad (5.9)$$

Due to arbitrariness of $\xi \in \mathbb{R}$ (5.9) is further equivalent to

$$|\alpha|^2 + |\beta|^2 = |\gamma|^2 + |\delta|^2, \quad \beta\bar{\alpha} = \delta\bar{\gamma}. \quad (5.10)$$

Adding two times the absolute values of the second equation in (5.10) to the first one, and subtracting the same from the first equation in (5.10), we easily find that at least one of the following two conditions holds:

- (i) $(|\gamma|, |\delta|) = (|\alpha|, |\beta|)$;
- (ii) $(|\gamma|, |\delta|) = (|\beta|, |\alpha|)$;

Multiply $\beta\bar{\alpha} = \delta\bar{\gamma}$ with α and use either $\alpha\bar{\alpha} = \gamma\bar{\gamma}$ or $\alpha\bar{\alpha} = \delta\bar{\delta}$ to obtain

$$(i') \quad (\gamma, \delta) = \mu(\alpha, \beta), \quad \text{or} \quad (ii') \quad (\gamma, \delta) = \mu(\bar{\beta}, \bar{\alpha}); \quad \text{for some } \mu \in \mathbb{C}, |\mu| = 1. \quad (5.11)$$

Now, write $A = \sum a_{ij}E_{ij}$ and $B = \sum b_{ij}E_{ij}$, where E_{ij} are the standard matrix units in M_2 : E_{ij} has 1 in the (i, j) th position and zeros elsewhere. Let $\{\mathbf{e}_1, \mathbf{e}_2\}$ be the standard basis of unit vectors for \mathbb{C}^2 .

Case 2. Suppose A or B is zero, say, $A = 0$. Applying (5.5) with $(x, y) = (\mathbf{e}_1, \mathbf{e}_2 e^{i\xi})$ for every $\xi \in [0, 2\pi)$, we see that the first column of B is zero. Applying (5.5) with $(x, y) = (\mathbf{e}_2, \mathbf{e}_1 e^{i\xi})$ for every $\xi \in [0, 2\pi)$, we see that the second column of B is also zero. So, (5.6) holds.

Case 3. Suppose both A and B are nonzero nilpotent. Replacing (A, B) by (U^*AU, U^*BU) for a suitable unitary U , we may assume that $a_{11} = a_{21} = a_{22} = 0$. Applying (5.5) with $(x, y) = (\mathbf{e}_1, \mathbf{e}_2 e^{i\xi})$ for every $\xi \in [0, 2\pi)$, we see that the first column of B is zero. Since B is nilpotent, we see that $b_{22} = 0$. Applying (5.5) with $(x, y) = (\mathbf{e}_2, \mathbf{e}_1 e^{i\xi})$ for every $\xi \in [0, 2\pi)$, we see that $|a_{12}| = |b_{12}|$. So, (5.6) holds.

Case 4. Suppose A and B are nonzero, and at least one of them, say, A is not nilpotent. Replacing (A, B) by $(U^*AU/\gamma, U^*BU/\gamma)$ for a suitable unitary U and a suitable $\gamma \in \mathbb{C}$, we may assume that $(a_{11}, a_{21}) = (1, 0)$ and $|a_{22}| \leq 1$ (see the Schur unitary triangularization theorem [7, Theorem 2.3.1]).

Now, for $(x, y) = (c\mathbf{e}_1 + s\mathbf{e}_2, e^{-i\xi}(-s\mathbf{e}_1 + c\mathbf{e}_2))$ with $\xi, c, s \in \mathbb{R}$ such that $(c, s) = (\cos u, \sin u)$ for some $u \in \mathbb{R}$, equation (5.5) is valid. Hence,

$$\begin{aligned} & |c^2 + a_{12}cs + s^2a_{22} + e^{i\xi} \tan t(-cs(1 - a_{22}) - s^2a_{12})| \\ &= |b_{11}c^2 + (b_{12} + b_{21})cs + s^2b_{22} + e^{i\xi} \tan t(b_{21}c^2 - cs(b_{11} - b_{22}) - s^2b_{12})|. \end{aligned} \quad (5.12)$$

It follows (see the implication (5.8) \Rightarrow (5.11)) that for any pair $(c, s) = (\cos u, \sin u)$ with $c, s > 0$, at least one of the two pairs of equalities (i'') and (ii'') below holds:

$$\begin{aligned} c^2 + a_{12}cs + s^2a_{22} &= \mu_s(b_{11}c^2 + (b_{12} + b_{21})cs + s^2b_{22}) \\ \tan t(-(1 - a_{22})cs - s^2a_{12}) &= \mu_s \tan t(b_{21}c^2 - (b_{11} - b_{22})cs - s^2b_{12}) \end{aligned} \quad (i'')$$

for some unimodular $\mu_s \in \mathbb{C}$;

$$\begin{aligned} c^2 + a_{12}cs + s^2a_{22} &= \mu_s \tan t(\overline{b_{21}}c^2 - (\overline{b_{11}} - \overline{b_{22}})cs - s^2\overline{b_{12}}) \\ \tan t(-(1 - a_{22})cs - s^2a_{12}) &= \mu_s(\overline{b_{11}}c^2 + (\overline{b_{12}} + \overline{b_{21}})cs + s^2\overline{b_{22}}) \end{aligned} \quad (ii'')$$

for some unimodular $\mu_s \in \mathbb{C}$. Rewrite (i'') and (ii'') into equivalent forms

$$\begin{aligned} (1 - \mu_s b_{11}) \left(\frac{s}{c}\right)^{-1} + a_{12} - \mu_s(b_{12} + b_{21}) + \left(\frac{s}{c}\right)(a_{22} - \mu_s b_{22}) &= 0 \\ (-\mu_s b_{21}) \left(\frac{s}{c}\right)^{-1} - ((1 - a_{22}) - \mu_s(b_{11} - b_{22})) - \left(\frac{s}{c}\right)(a_{12} - \mu_s b_{12}) &= 0 \end{aligned} \quad (i''')$$

and

$$\begin{aligned} (1 - \mu_s \tau \overline{b_{21}}) \left(\frac{s}{c}\right)^{-1} + a_{12} + \mu_s \tau (\overline{b_{11}} - \overline{b_{22}}) + \left(\frac{s}{c}\right)(a_{22} + \mu_s \tau \overline{b_{12}}) &= 0 \\ (-\mu_s \overline{b_{11}}) \left(\frac{s}{c}\right)^{-1} - (\tau(1 - a_{22}) + \mu_s(\overline{b_{12}} + \overline{b_{21}})) - \left(\frac{s}{c}\right)(\tau a_{12} + \mu_s \overline{b_{22}}) &= 0 \end{aligned} \quad (\text{ii}'')$$

respectively, with $\tau := \tan t > 0$.

Fix a sequence of pairs of positive numbers (c_i, s_i) , with $c_i^2 + s_i^2 = 1$, converging to $(1, 0)$. Passing to a subsequence, we have that at least one of (i'') and (ii'') holds for all its members, and we may also assume that $\lim_{i \rightarrow \infty} \mu_{s_i} = \mu$ for some unimodular μ .

Suppose (i'') holds for all (c_i, s_i) . Clearly $(s_i/c_i)^{-1}$ converges to ∞ , while $|\mu_{s_i}| = 1$ is bounded. It follows from the first equation of (i'') that $\lim_{i \rightarrow \infty} (1 - \mu_{s_i} \overline{b_{11}}) = 0$, so $1 - \mu \overline{b_{11}} = 0$ and $\overline{b_{11}} = \mu^{-1}$. The second equation in (i'') yields that $\lim_{i \rightarrow \infty} (-\mu_{s_i} \overline{b_{21}}) = 0$, hence $\overline{b_{21}} = 0$. Now the second equation in (i'') takes the form

$$a_{22} - 1 + \mu_{s_i}(\mu^{-1} - \overline{b_{22}}) - \left(\frac{s_i}{c_i}\right)(a_{12} - \mu_{s_i} \overline{b_{12}}) = 0, \quad (5.13)$$

and passing to the limit when $i \rightarrow \infty$ gives

$$a_{22} - 1 + \mu(\mu^{-1} - \overline{b_{22}}) = 0, \quad (5.14)$$

i.e., $\overline{b_{22}} = \mu^{-1} a_{22}$. Next, substitute zero for $\overline{b_{21}}$ and μ^{-1} for $\overline{b_{11}}$ in the first equation in (i''), and pass to the limit. The result is

$$\lim_{i \rightarrow \infty} (1 - \mu_{s_i} \mu^{-1}) \left(\frac{s_i}{c_i}\right)^{-1} = -a_{12} + \mu \overline{b_{12}}. \quad (5.15)$$

On the other hand, substituting $\overline{b_{22}} = \mu^{-1} a_{22}$ into (5.13) yields, after some rearrangements

$$(a_{22} - 1) (1 - \mu_{s_i} \mu^{-1}) \left(\frac{s_i}{c_i}\right)^{-1} - (a_{12} - \overline{b_{12}} \mu_{s_i}) = 0.$$

Using (5.15) it follows after simplification that

$$\mu a_{22} \overline{b_{12}} - a_{12} a_{22} = 0.$$

Thus, either $\overline{b_{12}} = \mu^{-1} a_{12}$, and then (5.6) holds: $B = \mu^{-1} A$, or $a_{22} = 0$, and then (5.6) holds by virtue of Case 1.

Thus, the proof of Step 2 is complete if there is a sequence of positive numbers (c_i, s_i) with $c_i^2 + s_i^2 = 1$ converging to $(1, 0)$ such that (i'') holds for all (c_i, s_i) .

We now assume that (ii'') holds for all positive (c, s) with $c^2 + s^2 = 1$ and s sufficiently close to zero. It follows from the first equation of (ii'') that $\lim_{s \rightarrow 0} (1 - \mu_s \tau \overline{b_{21}}) = 0$. Denoting by μ any partial limit of μ_s as $s \rightarrow 0$, we have $1 - \mu \tau \overline{b_{21}} = 0$, or

$$\overline{b_{21}} = \mu \tau^{-1}. \quad (5.16)$$

(By the way this shows that μ is unique, i.e., $\mu = \lim_{s \rightarrow 0} \mu_s$.) The second equation in (ii'') yields $\lim_{s \rightarrow 0} (-\mu_s \overline{b_{11}}) = 0$, hence $\overline{b_{11}} = 0$. Letting $s \rightarrow 0$, the second equation in (ii'') gives

$$\tau(1 - a_{22}) + \mu(\overline{b_{12}} + \overline{b_{21}}) = 0. \quad (5.17)$$

Thus,

$$\overline{b_{12}} = \mu^{-1}(-\mu\overline{b_{21}} - \tau(1 - a_{22})),$$

and using $\overline{b_{21}} = \mu^{-1}\tau^{-1}$ we obtain

$$\overline{b_{12}} = \mu^{-1}(-\tau - \tau^{-1} + \tau a_{22}). \quad (5.18)$$

It follows from (5.16) and (5.18) that

$$\overline{b_{12}} + \overline{b_{21}} = \mu^{-1}\tau(-1 + a_{22}). \quad (5.19)$$

Substituting μ^{-1} for $\tau\overline{b_{21}}$ and zero for b_{11} in the first equation in (ii''), we find

$$(1 - \mu_s\mu^{-1})\left(\frac{s}{c}\right)^{-1} + a_{12} - \mu_s\tau\overline{b_{22}} + \left(\frac{s}{c}\right)(a_{22} + \mu_s\tau\overline{b_{12}}) = 0,$$

and passing to the limit when $s \rightarrow 0$, it follows that

$$\lim_{s \rightarrow 0}(\mu - \mu_s)\left(\frac{s}{c}\right)^{-1} = \mu^2\tau\overline{b_{22}} - \mu a_{12}. \quad (5.20)$$

On the other hand, using $b_{11} = 0$ and (5.17), the second equation in (ii'') reads

$$(\mu - \mu_s)(\overline{b_{12}} + \overline{b_{21}}) - \left(\frac{s}{c}\right)(\tau a_{12} + \mu_s\overline{b_{22}}) = 0.$$

In view of (5.20) we have

$$(\mu^2\tau\overline{b_{22}} - \mu a_{12})(\overline{b_{12}} + \overline{b_{21}}) = \tau a_{12} + \mu\overline{b_{22}}. \quad (5.21)$$

Using (5.19), we solve (5.21) for $\overline{b_{22}}$:

$$\overline{b_{22}} = \frac{a_{12}\tau a_{22}}{\mu\tau^2 a_{22} - \mu\tau^2 - \mu}. \quad (5.22)$$

Note that the condition $|a_{22}| \leq 1$ guarantees that the denominator in (5.22) is nonzero.

Next, we show that $b_{12} + b_{21} = b_{22} = 0$. Arguing by contradiction, let us suppose that

$$b_{12} + b_{21} \neq 0 \quad \text{or} \quad b_{22} \neq 0. \quad (5.23)$$

Then the second equation in (ii'') can be solved for μ_s (assuming s is close enough to zero):

$$\mu_s = \frac{\tau(1 - a_{22}) + w\tau a_{12}}{-\overline{b_{12}} - \overline{b_{21}} - w\overline{b_{22}}}, \quad w := \frac{s}{c}. \quad (5.24)$$

Substituting the right hand side of (5.24) into the first equation in (ii''), after some simple algebra, we obtain:

$$\begin{aligned} & (\overline{b_{22}}a_{22} - \tau^2\overline{b_{12}}a_{12})w^3 + \\ & ((\tau^2 + 1)\overline{b_{22}}a_{12} + \overline{b_{21}}a_{22} + \overline{b_{12}}((\tau^2 + 1)a_{22} - \tau^2))w^2 + \\ & ((\overline{b_{12}} + (\tau^2 + 1)\overline{b_{21}})a_{12} + \overline{b_{22}}(-a_{22}\tau^2 + \tau^2 + 1))w + \\ & \overline{b_{12}} + \overline{b_{21}}(-a_{22}\tau^2 + \tau^2 + 1) = 0 \end{aligned} \quad (5.25)$$

The equation holds for all w close to zero; equating coefficients of powers of w on the right hand and on the left hand sides of (5.25), the following equalities result

((5.26), ((5.27), ((5.28), ((5.29) correspond to the coefficients of w^3 , w^0 , w^2 , w^1 , respectively):

$$-a_{22}\overline{b_{22}} + \tau^2 a_{12}\overline{b_{12}} = 0, \quad (5.26)$$

$$\overline{b_{12}} + \overline{b_{21}} + \tau^2(1 - a_{22})\overline{b_{21}} = 0, \quad (5.27)$$

$$a_{12}\overline{b_{22}} + \tau^2 a_{12}\overline{b_{22}} + a_{22}(\overline{b_{12}} + \overline{b_{21}}) - \tau^2(1 - a_{22})\overline{b_{12}} = 0, \quad (5.28)$$

$$-\overline{b_{22}} - \tau^2 a_{12}\overline{b_{21}} + a_{12}(-\overline{b_{12}} - \overline{b_{21}}) - \tau^2(1 - a_{22})\overline{b_{22}} = 0. \quad (5.29)$$

Substituting the right hand sides of (5.22) and (5.18) for $\overline{b_{22}}$ and $\overline{b_{12}}$, respectively, in (5.26) yields after simplification:

$$a_{12}(a_{22} - 1)(-\tau^2 + (\tau^2 - 1)a_{22} - 1) = 0$$

Thus, at least one of the three equalities holds:

$$a_{12} = 0, \quad (5.30)$$

$$a_{22} = 1, \quad (5.31)$$

$$a_{22}(\tau^2 - 1) = \tau^2 + 1. \quad (5.32)$$

However, (5.32) is impossible because it contradicts $\tau > 0$ and $|a_{22}| \leq 1$. In the case (5.30) holds we have $b_{22} = 0$, by (5.26). Substitute $b_{22} = 0$ and the right hand sides of (5.18) and (5.19) for $\overline{b_{12}}$ and $\overline{b_{12}} + \overline{b_{21}}$, respectively, in (5.28), to obtain:

$$\frac{\tau(\tau^2 + 1)(a_{22} - 1)^2}{\mu} = 0,$$

and since $\tau > 0$ we have $a_{22} = 1$. But then $b_{12} + b_{21} = 0$ by (5.19), a contradiction with (5.23). So (5.30) cannot be true and hence we must have $a_{22} = 1$. Then $b_{12} + b_{21} = 0$. Now (5.28) gives

$$a_{12}\overline{b_{22}}(1 + \tau^2) = 0,$$

so either $a_{12} = 0$ or $b_{22} = 0$, and in either case a contradiction with (5.23) results.

Thus, (5.23) cannot hold, and we have $b_{12} + b_{21} = 0$ and $b_{22} = 0$. By (5.19) $a_{22} = 1$ and then by (5.22) $a_{12} = 0$. Keeping in mind (5.16), the result is that

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \mu \begin{bmatrix} 0 & -\cot t \\ \cot t & 0 \end{bmatrix}.$$

We now can finish the proof of Step 2 as follows. If already $A = \mu B$ for some $|\mu| = 1$, then we are done. Assume lastly $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \mu \begin{bmatrix} 0 & -\cot t \\ \cot t & 0 \end{bmatrix}$. Here we replace the pair (A, B) with

$$(A', B') := (UAU^*, UBU^*), \quad A' = \sum a'_{ij}E_{ij}, \quad B' = \sum b'_{ij}E_{ij},$$

for the unitary $U := \text{diag}(i, 1)$. Clearly, the new pair still satisfies the defining identity (5.5), and still $(a'_{11}, a'_{21}) = (1, 0)$ and $|a'_{22}| \leq 1$. This allows us to use the same arguments as above in Case 4. In particular, Eq. (5.12) with (a_{ij}, b_{ij}) replaced by (a'_{ij}, b'_{ij}) gives either $A' = \mu' B'$ ($|\mu'| = 1$) wherefrom $A = \mu' B$, or

else $A' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B' = \mu' \begin{bmatrix} 0 & -\cot t \\ \cot t & 0 \end{bmatrix}$. But the last case is contradictory, namely recall that

$$\mu' \begin{bmatrix} 0 & -\cot t \\ \cot t & 0 \end{bmatrix} = B' = UBU^* = \mu \begin{bmatrix} 0 & -i \cot t \\ -i \cot t & 0 \end{bmatrix},$$

giving $\mu' = 0 = \mu$, a contradiction.

This concludes the proof of Step 2. \square

6. Linear dependence in terms of trace functionals

If $C \in L(\mathbf{H})$ is a trace class operator, then the formula

$$W_C(A) = \{\text{trace}(CU^*AU) : U \in L(\mathbf{H}), U \text{ unitary}\}$$

defines the C -numerical range of an operator $A \in L(\mathbf{H})$. The C -numerical ranges also have been extensively studied, see [5, 8, 1, 12, 3], a representative sample of relevant works. In particular, C -numerical ranges of matrices have been applied recently in quantum computing and control [4, 14, 6]. It is easy to see that the q -numerical range is actually the C -numerical range with C given by

$$Cx = q\langle x, y \rangle y + \sqrt{1 - q^2} \langle x, z \rangle y, \quad x \in \mathbf{H}, \quad (6.1)$$

where (y, z) , $y, z \in \mathbf{H}$, is a fixed orthonormal pair. Note that every rank one operator is unitarily similar (after appropriate scaling) to an operator of the form (6.1); thus, the q -numerical ranges represent the C -numerical ranges with rank one operators C .

The result of Proposition 2.1 extends to C -numerical ranges, as follows.

Theorem 6.1. *Let f be the bounded linear functional on $L(\mathbf{H})$, given by a trace class operator C :*

$$f(X) = \text{trace}(CX), \quad X \in L(\mathbf{H}). \quad (6.2)$$

Assume that C is not scalar. Suppose $A, B \in L(\mathbf{H})$. Then

$$f(U^*AU) = f(U^*BU)$$

holds for every unitary U if and only if either (1) $\text{trace } C \neq 0$ and $A = B$, or (2) $\text{trace } C = 0$ and $A - B$ is scalar.

For the proof of Theorem 6.1 a few lemmas will be needed. We start with a simple observation.

Lemma 6.2. *An operator $A \in L(\mathbf{H})$ has the property that*

$$\langle Ax, x \rangle = \langle Ay, y \rangle \quad \forall \text{ orthonormal pairs } (x, y), \quad x, y \in \mathbf{H} \quad (6.3)$$

if and only if A is scalar.

Proof. The “if” part is trivial, and for the “only if” part note that if $z, w \in \mathbf{H}$ are normalized elements such that (z, y) and (w, y) are orthonormal pairs for some $y \in \mathbf{H}$, then

$$\langle Az, z \rangle = \langle Aw, w \rangle. \quad (6.4)$$

Thus, if the dimension of \mathbf{H} is at least 3, then (6.4) holds for any normalized z and w . Hence the numerical range of A is a singleton, and A is scalar. If the dimension of A is 2, then the statement of Lemma 6.2 can be easily verified by a straightforward computation: Subtracting from A a suitable scalar, we can assume that

$$\langle Ae_1, e_1 \rangle = \langle Ae_2, e_2 \rangle = 0.$$

So $A = \begin{bmatrix} 0 & a \\ b & 0 \end{bmatrix}$ for some $a, b \in \mathbb{C}$, and further consideration using property (6.3) shows that we must have $a = b = 0$. \square

We denote by $L_1(\mathbf{H})$ the ideal of trace operators in $L(\mathbf{H})$, and by $L_{10}(\mathbf{H})$ the (closed in the trace-class norm) subspace of trace operators with zero trace.

Lemma 6.3. *Let $C \in L_1(\mathbf{H})$ be a nonzero operator with zero trace. Then $X \in L(\mathbf{H})$ satisfies the property that $\text{trace}(UCU^*X) = 0$ for every unitary U if and only if X is scalar.*

The statement and proof of this and the following lemma is inspired by [15] (these lemmas are essentially proved in [15] in the case \mathbf{H} is finite dimensional).

Proof. The “if” part being trivial, we prove the “only if” part. Suppose the operator UCU^*X has zero trace for every unitary U but X is not scalar. We may replace C by any (finite) nonzero linear combination of operators in the unitary orbit of C . By doing so, we may (and do) assume without loss of generality that, for some orthonormal pair (x, y) , $x, y \in \mathbf{H}$, and with respect to the orthogonal decomposition

$$\mathbf{H} = (\text{span } x) \oplus (\text{span } y) \oplus (\text{span } \{x, y\})^\perp, \quad (6.5)$$

the operator C has the following matrix form:

$$C = \text{diag}(c_1, c_2, C_0), \quad (6.6)$$

where $c_1, c_2 \in \mathbb{C}$ and $c_1 \neq c_2$. Indeed, let $x \in \mathbf{H}$ be a normalized element such that $\langle Cx, x \rangle \neq 0$; the condition that C has zero trace guarantees that there exists a normalized y orthogonal to x such that $\langle Cy, y \rangle \neq \langle Cx, x \rangle$. Now let U_1, U_2, U_3 be self-adjoint unitary operators given by

$$U_1 = \text{diag}[1, -1, I], \quad U_2 = \text{diag}[-1, 1, I], \quad U_3 = \text{diag}[-1, -1, I], \quad I = I_{(\text{span } \{x, y\})^\perp},$$

with respect to the decomposition (6.5). It is easy to see that the operator

$$C + U_1CU_1 + U_2CU_2 + U_3CU_3$$

has the desired form (6.6). Independently, X can be also replaced by V^*XV , for any unitary V . Since X is not scalar, $\langle Xx', x' \rangle \neq \langle Xy', y' \rangle$ for some orthonormal

pair (x', y') by Lemma 6.2. Applying a transformation $X \rightarrow V^*XV$, we may assume $(x', y') = (x, y)$. So

$$X = \begin{bmatrix} x_1 & * & * \\ * & x_2 & * \\ * & * & X_0 \end{bmatrix}, \quad x_1, x_2 \in \mathbb{C}, \quad x_1 \neq x_2, \quad X_0 \in L((\text{span}\{x, y\})^\perp),$$

with respect to (6.5). Now

$$0 = \text{trace}(CX) = c_1x_1 + c_2x_2 + \text{trace}(C_0X_0), \quad (6.7)$$

and letting

$$U = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & I_{(\text{span}\{x, y\})^\perp} \end{bmatrix},$$

we also have

$$0 = \text{trace}(UCUX) = c_2x_1 + c_1x_2 + \text{trace}(C_0X_0). \quad (6.8)$$

Comparing (6.7) and (6.8) we see that

$$(c_1 - c_2)(x_1 - x_2) = 0,$$

a contradiction with $c_1 \neq c_2$, $x_1 \neq x_2$. \square

The result of the next lemma was proved in [15, 5] in case \mathbf{H} is finite dimensional.

Lemma 6.4. *Let $C \in L_1(\mathbf{H})$ be a nonzero operator. Then the closure (in the trace norm) of the linear span of operators of the form UCU^* , U unitary, coincides with $L_1(\mathbf{H})$ if $\text{trace } C \neq 0$, and coincides with $L_{10}(\mathbf{H})$ if $\text{trace } C = 0$.*

Proof. Denote by $\mathcal{U}(C)$ the closure of the linear span of operators of the form UCU^* , U unitary. Suppose $\text{trace } C = 0$, and arguing by contradiction, assume $\mathcal{U}(C) \neq L_{10}(\mathbf{H})$. Then (because $L(\mathbf{H})$ is the dual of $L_1(\mathbf{H})$) there exists $X \in L(\mathbf{H})$ such that $\text{trace}(TX) = 0$ for every $T \in \mathcal{U}(C)$ but $\text{trace}(T_0X) \neq 0$ for some $T_0 \in L_{10}(\mathbf{H})$. Being nonscalar, $C \neq 0$, so by Lemma 6.3, the first condition implies that X is scalar, which contradicts the second condition.

Next, suppose $\text{trace } C \neq 0$. Since C is not scalar, we have $\langle Cx, x \rangle \neq \langle Cy, y \rangle$ for some orthonormal pair (x, y) by Lemma 6.2; hence $\widehat{C} := C - VCV^* \neq 0$ for some unitary V . Clearly $\text{trace } \widehat{C} = 0$ and $\mathcal{U}(C) \supseteq \mathcal{U}(\widehat{C})$. By the first part of the lemma we have $\mathcal{U}(\widehat{C}) = L_{10}(\mathbf{H})$, hence $\mathcal{U}(C) \supseteq L_{10}(\mathbf{H})$. On the other hand, since $C \in \mathcal{U}(C)$ and $\text{trace } C \neq 0$, we have $\mathcal{U}(C) \neq L_{10}(\mathbf{H})$, hence $\mathcal{U}(C) = L_1(\mathbf{H})$. \square

Proof of Theorem 6.1. The “if” part is trivial. We prove the “only if” part. The condition implies that $\text{trace}(AUCU^*) = \text{trace}(BUCU^*)$, i.e., $\text{trace}((A - B)UCU^*) = 0$, for every unitary U . Since the closure of the linear span of $\{UCU^* : U \text{ unitary}\}$ is either $L_1(\mathbf{H})$ or $L_{10}(\mathbf{H})$ by Lemma 6.4, we see that (1) or (2) holds. \square

We were not able to prove a generalization of the result of Theorem 2.2 to the framework of trace functionals. Therefore the following open problem is suggested:

Open Problem 6.5. *Suppose f is a bounded linear functional on $L(\mathbf{H})$ given by (6.2), where the trace class operator C is not scalar. Characterize pairs $A, B \in L(\mathbf{H})$ such that*

$$|f(U^*AU)| = |f(U^*BU)| \quad \forall \text{ unitary } U \in L(\mathbf{H}). \quad (6.9)$$

By analogy with Theorem 2.2, we conjecture:

Conjecture 6.6. *Under the hypotheses of the open problem, (6.9) holds if and only if:*

- (1) $\text{trace } C = 0$, $C = C^*$, and either $A = \mu B + \nu I$ or $A = \mu B^* + \nu I$ for some $\mu, \nu \in \mathbf{C}$, $|\mu| = 1$;
- (2) $\text{trace } C = 0$, $C \neq C^*$, and $A = \mu B + \nu I$ for some $\mu, \nu \in \mathbf{C}$, $|\mu| = 1$;
- (3) $\text{trace } C \neq 0$, $C = C^*$, and either $A = \mu B$ or $A = \mu B^*$ for some $\mu \in \mathbf{C}$, $|\mu| = 1$;
- (4) $\text{trace } C \neq 0$, $C \neq C^*$, and $A = \mu B$ for some $\mu \in \mathbf{C}$, $|\mu| = 1$.

Theorem 2.2 proves the conjecture in the case when C is any rank one operator.

References

- [1] W.-S. Cheung and N.-K. Tsing, The C -numerical range of matrices is star-shaped, *Linear and Multilinear Algebra* **41** (1996), 245–250.
- [2] M.-T. Chien and N. Nakazato, Davis-Wielandt shell and q -numerical range, *Linear Algebra Appl.* **340** (2002), 15–31.
- [3] M.-D. Choi, C.-K. Li, and T.-Y. Poon, Some convexity features associated with unitary orbits, *Canadian J. Mathematics* **55** (2003), 91–111.
- [4] G. Dirr, U. Helmke, H. Kleinsteuber, and Th. Schulte-Herbrüggen, Relative C -numerical ranges for applications in quantum control and quantum information, *Linear Multilinear Algebra* **56** (2008), 27–51.
- [5] M. Goldberg and E. G. Straus, Norm properties of C -numerical radii, *Linear Algebra Appl.* **24** (1979), 113–131.
- [6] U. Helmke, K. Hüper, J. B. Moore, and Th. Schulte-Herbrüggen, Gradient flows computing the C -numerical range with applications in NMR spectroscopy, *J. Global Optim.* **23** (2002), 283–308.
- [7] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [8] M. S. Jones, A note on the shape of the generalized C -numerical range, *Linear and Multilinear Algebra* **31** (1992), 81–84.
- [9] B. Kuzma, C.-K. Li, G. Lešnjak, T. Petek, and L. Rodman, Norm preservers of Jordan products, paper in preparation.
- [10] C.-K. Li, A survey on linear preservers of numerical ranges and radii. *Taiwanese J. Math.* **5** (2001), 477–496.

- [11] C.-K. Li, P. P. Mehta, and L. Rodman, A generalized numerical range: the range of a constrained sesquilinear form, *Linear and Multilinear Algebra* **37** (1994), 25–49.
- [12] C.-K. Li and L. Rodman, Multiplicative preservers of C -numerical ranges and radii *Linear and Multilinear Algebra* **52** (2004), 265–279.
- [13] C.-K. Li, B.-S. Tam, and N.-K. Tsing, Linear operators preserving the (p, q) -numerical range, *Linear Algebra Appl.* **110** (1988), 75–89.
- [14] Th. Schulte-Herbrüggen, G. Dirr, U. Helmke, and S. J. Glaser, The significance of the C -numerical range and the local C -numerical range in quantum control and quantum information, *Linear and Multilinear Algebra* **56** (2008), 3–26.
- [15] B.-S. Tam, A simple proof of the Goldberg - Straus theorem on numerical radii, *Glasgow Math. J.* **28** (1986), 139–141.

Bojan Kuzma
Inst. of Mathematics, Physics, and Mechanics
University of Ljubljana
Jadranska 19
1000 Ljubljana, Slovenia
e-mail: kuzma@fmf.uni-lj.si

Gorazd Lešnjak
Department of Electrical Engineering and Computer Science
University of Maribor
Smetanova 17
2000 Maribor, Slovenia
e-mail: gorazd.lesnjak@uni-mb.si

Chi-Kwong Li
Department of Mathematics
College of William and Mary
Williamsburg, VA 23187-8795, USA
e-mail: ckli@math.wm.edu

Tatjana Petek
FERI
University of Maribor
Smetanova 17
2000 Maribor, Slovenia
e-mail: tatjana.petek@uni-mb.si

Leiba Rodman
Department of Mathematics
College of William and Mary
Williamsburg, VA 23187-8795, USA
e-mail: lxrodm@math.wm.edu