MATRICES WITH CIRCULAR SYMMETRY ON THEIR UNITARY ORBITS AND C-NUMERICAL RANGES

CHI-KWONG LI AND NAM-KIU TSING

(Communicated by Louis J. Ratliff, Jr.)

Abstract. We give equivalent characterizations for those $n \times n$ complex matrices $A$ whose unitary orbits $\mathcal{U}(A)$ and $C$-numerical ranges $W_C(A)$ satisfy $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ or $e^{i\theta} W_C(A) = W_C(A)$ for some real $\theta$ (or for all real $\theta$). In particular, we show that they are the block-cyclic or block-shift operators. Some of these results are extended to infinite-dimensional Hilbert spaces.

1. INTRODUCTION

Let $C^n$ be equipped with the standard inner product $\langle \cdot, \cdot \rangle$ defined by $\langle x, y \rangle = y^* x$ for all $x, y \in C^n$. For any $A \in C^{n \times n}$, the unitary orbit of $A$ is the set

$$\mathcal{U}(A) = \{UAU^*: U \in C^{n \times n}, \ U^*U = I\}.$$ 

If we regard $A$ as a linear operator on $C^n$, then $\mathcal{U}(A)$ is the collection of all matrix representations of $A$ with respect to different orthonormal bases of $C^n$. In order to understand the properties of $A$, it is useful to study its unitary orbit. This is especially true when those "orthonormal bases-free" properties are concerned. For example, to conclude that $A, B \in C^{n \times n}$ are not the same linear operator represented in different orthonormal bases, one has to show that $\mathcal{U}(A) \neq \mathcal{U}(B)$.

Another tool for studying the properties of $A \in C^{n \times n}$ is the (classical) numerical range of $A$, which is defined and denoted by

$$W(A) = \{x^*Ax: x \in C^n, \ x^*x = 1\}.$$ 

There is extensive literature on the numerical range, and many nice results have been obtained (e.g., see [3] and its references). In particular, many of these results show the interesting interplay between the geometrical properties of

---

Received by the editors November 22, 1989.
1980 Mathematics Subject Classification (1985 Revision). Primary 15A60.
Key words and phrases. $C$-numerical range, unitary orbit, linear operator.
The research of the first author was supported in part by the National Science Foundation under Grant DMS 89-00922.
The research of the second author was supported in part by the National Science Foundation under Grant DMC 84-51515 and by the National Science Foundation’s Engineering Research Centers Program: NSF CDR 88-03012.
$W(A)$ and the algebraic and analytic properties of the matrix $A$. For examples, $W(A)$ is a singleton set if and only if $A$ is a scalar multiple of the identity matrix; $W(A)$ is a segment on the real line if and only if $A$ is hermitian; if $A$ is normal then $W(A)$ is just the convex hull of its eigenvalues; etc.

Suppose we use the same symbol $\langle \cdot, \cdot \rangle$ to denote the standard inner product on $\mathbb{C}^{n \times n}$ defined by $\langle A, B \rangle = \text{tr}(B^*A)$, and let $E_{11}$ be the $n \times n$ matrix with 1 as the $(1, 1)$ entry and 0 elsewhere. Then

$W(A) = \{ \text{tr}(E_{11}UAU^*) : U^*U = I \} = \{ (X, E_{11}) : X \in \mathcal{H}(A) \}.$

Hence $W(A)$ can be identified with the image of $\mathcal{H}(A)$ when projected onto the subspace spanned by $E_{11}$. As $W(A)$ already carries a great deal of information on $A$, one would naturally consider other rank-1 projections and expect that the images of $\mathcal{H}(A)$ under these projections would carry more information on $A$ or $\mathcal{H}(A)$. This gives rise to the following definition of the $C$-numerical range of $A$, which is introduced by Goldberg and Straus [4]; for $C \in \mathbb{C}^{n \times n}$,

$W_C(A) = \{ \text{tr}(CUAU^*) : U^*U = I \}.$

As $W_C(A) = \{ (X, C^*) : X \in \mathcal{H}(A) \}$, $W_C(A)$ can be viewed as the (linearly scaled) image of $\mathcal{H}(A)$ when projected onto the subspace spanned by $C^*$. As a matter of fact, many interesting results, which relate the geometrical shape of $W_C(A)$ with the algebraic or analytic properties of $A$, have been obtained for this generalization of the numerical range. For example, it has been shown that $W_C(A)$ is a singleton set if and only if $A$ or $C$ is a scalar multiple of $I$; $W_C(A)$ is a nondegenerate line segment in the complex plane if and only if each of $A$ and $C$ is normal (but not a scalar multiple of $I$) and has collinear eigenvalues [5,7]. Some other related results can be found in [6,8]. In this paper we do research along a similar line. We characterize in §2 those matrices $A$ for which

$e^{i\theta}W_C(A) = W_C(A)$ for all $C \in \mathbb{C}^{n \times n}$

for some real $\theta$ (or for all real $\theta$). There are several alternative descriptions for such matrices. For example, they are precisely those matrices $A$ whose unitary orbits satisfy $e^{i\theta}\mathcal{H}(A) = \mathcal{H}(A)$, and they are the block-cyclic or block-shift operators. Here, a block-cyclic (or block-shift, respectively) operator $A$ on $\mathbb{C}^n$ is one that satisfies $A(S_i) \subset S_{i+1}$ for $i = 1, \ldots, m - 1$, and $A(S_m) \subset S_1$ (or $A(S_m) = 0$, respectively) for some direct sum $\mathbb{C}^n = S_1 \oplus \cdots \oplus S_m$ of mutually orthogonal subspaces. This includes the weighted cyclic (or weighted shift, respectively) operators $A$ defined by $A(x_i) = a_ix_{i+1}$, $i = 1, \ldots, n - 1$, and $A(x_n) = a_nx_1$ (or 0, respectively), for some orthonormal basis $\{ x_1, \ldots, x_n \}$ of $\mathbb{C}^n$ and $a_1, \ldots, a_n \in \mathbb{C}$. In §3, we consider the same problem in infinite-dimensional Hilbert spaces instead of $\mathbb{C}^n$. Some of the results in §2 are extended there.
2. Main results (finite-dimensional case)

The result below gives equivalent characterizations for matrices which have circular symmetry on their unitary orbits or C-numerical ranges. In particular, it shows that these matrices are exactly the block-shift operators.

Theorem 2.1. Let $A \in \mathbb{C}^{n \times n}$. The following conditions are equivalent.

(a) $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ for all $\theta \in \mathbb{R}$.

(b) $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ for some $\theta \in \mathbb{R}$ with $e^{ik\theta} \neq 1$ for all $k = 1, \ldots, n$.

(c) $e^{i\theta} W_C(A) = W_C(A)$ for all $C \in \mathbb{C}^{n \times n}$ and all $\theta \in \mathbb{R}$.

(d) $e^{i\theta} W_{A^*}(A) = W_{A^*}(A)$ for some $\theta \in \mathbb{R}$ with $e^{ik\theta} \neq 1$ for all $k = 1, \ldots, n$.

(e) $A$ is unitarily similar to a matrix $M = (M_{kk})_{1 \leq k, \ell \leq m}$ in block form such that all $M_{kk}$'s are square blocks and $M_{kk} = 0$ if $\ell + 1 \neq k$.

(f) $\mathbb{C}^n$ can be factored into a direct sum of mutually orthogonal subspaces $\mathbb{C}^n = S_1 \oplus \cdots \oplus S_m$ such that $A(S_k) \subseteq S_{k+1}$ for $k = 1, \ldots, m-1$, and $A(S_m) = \{0\}$.

Proof. That (a) implies (b) and (c) implies (d) are clear. Also, the equivalence of (e) and (f) is not hard to see. We shall show in the following that (d) $\Rightarrow$ (b) $\Rightarrow$ (e) $\Rightarrow$ (a) $\Rightarrow$ (c), so that all six conditions are equivalent.

(d) $\Rightarrow$ (b): It suffices to consider the case of $A \neq 0$. Suppose (d) holds for some $\theta \in \mathbb{R}$ and nonzero $A$. Let $\|\cdot\|$ be the norm induced by the standard inner product on $\mathbb{C}^{n \times n}$ described in §1. Observe that $W_{A^*}(A) = \{ (UAU^*, A) : U \text{ unitary} \}$, and $\|UAU^*\| = \|A\|$ for all unitary $U$. Putting $U = I$ we get $\|A\|^2 \in W_{A^*}(A)$. Therefore $e^{i\theta} \|A\|^2 \in e^{i\theta} W_{A^*}(A) = W_{A^*}(A)$. Let $V$ be unitary such that $e^{i\theta} \|A\|^2 = \langle VAV^*, A \rangle$. Since $\|VAV^*, A\| = |e^{i\theta} \|A\|^2| = \|A\|^2 = \|VAV^*\| \|A\|$, by the Cauchy-Schwarz inequality, we have $VAV^* = zA$ for some $z \in \mathbb{C}$. Thus $e^{i\theta} \|A\|^2 = \langle VAV^*, A \rangle = \langle zA, A \rangle = z \|A\|^2$. As $\|A\| \neq 0$, this implies $z = e^{i\theta}$. Hence $e^{i\theta} A = VAV^* \in \mathcal{U}(A)$. This is equivalent to $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$.

(b) $\Rightarrow$ (e): Suppose (b) holds. Then there exists a unitary $U$ such that $e^{i\theta} A = UAU^*$, or, equivalently, $UA = e^{i\theta} U$. Diagonalize $U$ to become $U = WDW^*$ where $W$ is unitary and

$$D = \text{diag}(z_1, \ldots, z_1, z_2, \ldots, z_2, \ldots, z_m, \ldots, z_m),$$

where $z_1, \ldots, z_m$ are distinct complex scalars of moduli 1. Since $e^{ik\theta} \neq 1$ for all $k = 1, \ldots, n$, the order of $z_1, \ldots, z_m$ can be arranged to satisfy

$$z_kz_{\ell} = e^{i\theta} \quad \text{only if } \ell + 1 = k(1 \leq k, \ell \leq m).$$

Let $n_k$ be the multiplicity of $z_k$ in $D$, and $W^*AW = M = (M_{kk})$ be in block form such that each $M_{kk}$ has order $n_k \times n_k$. Since $DM = W^*WDW^*AW = W^*UAUW = W^*(e^{i\theta} AU)W = e^{i\theta} W^*AWDW^*W = e^{i\theta} MD$, this implies that
$z_k M_{kl} = e^{i\theta} z_l M_{kl}$, or equivalently, $(z_k z_l - e^{i\theta}) M_{kl} = 0$, for all $1 \leq k, \ell \leq m$. By (1), we have $M_{kl} \neq 0$ only if $\ell + 1 = k$. Hence $W^* A W = (M_{kl})$ is of the required form in (e).

(e) $\Rightarrow$ (a): Let $U$ be unitary such that $M = (M_{kl}) = U A U^*$ is of the form satisfying condition (e). For any $\theta \in \mathbb{R}$, consider a unitary matrix $D = (D_{kl})$ which has the same block form as $(M_{kl})$ and is defined as

$$D_{kl} = \begin{cases} e^{i\theta} 1_{n_k}, & \text{if } k = \ell; \\ 0, & \text{otherwise}. \end{cases}$$

Then

$$D U A U^* D^* = D M D^* = (e^{i\theta} e^{-i\theta} M_{kl}) = e^{i\theta} M.$$

Hence

$$e^{i\theta} A = e^{i\theta} (U^* M U) = U^* (e^{i\theta} M) U = U^* (D U A U^* D^*) U = (U^* D U) A (U^* D U)^* \in \mathcal{Z}(A).$$

This implies $e^{i\theta} \mathcal{Z}(A) \subseteq \mathcal{Z}(A)$.

(a) $\Rightarrow$ (c): Suppose (a) holds. Since, for any $C$, $W_C(A)$ is the image of $\mathcal{Z}(A)$ under a linear map, we have $e^{i\theta} W_C(A) = W_C(A)$ for all $\theta \in \mathbb{R}$.

Matrices that satisfy any of the conditions in Theorem 2.1 can be described in terms of the exact shape of their $C$-numerical ranges or the symmetry of their $C$-numerical ranges about a certain axis on the complex plane, as shown in the following corollary.

**Corollary 2.2.** A matrix $A \in \mathbb{C}^n \times n$ satisfies any of the conditions in Theorem 2.1 if and only if one of the following conditions holds.

- (g) $W_C(A)$ is a circular disc on the complex plane centered at the origin for all $C \in \mathbb{C}^n \times n$.
- (h) $W_C(A)$ is a circular disc on the complex plane centered at the origin.
- (i) $W_C(A)$ is symmetric about the real axis (or any axis passing through the origin) on the complex plane for all $C \in \mathbb{C}^n \times n$.

**Proof.** Suppose $A$ satisfies any of the equivalent conditions in Theorem 2.1. Then there exists some unitary $U$ such that $U A U^* = M = (M_{kl})$ satisfies condition (e) in Theorem 2.1. For any $C \in \mathbb{C}^n \times n$, let $V$ be unitary such that $V^* C V$ is in lower triangular form. Then, since $M$ is lower triangular with all diagonal elements zero, we have $0 = \text{tr}((V^* C V) M) = \text{tr}((V^* C V) (U A U^*)) = \text{tr}(C V U) A (V U)^* \in W_C(A)$. As the unitary group $\{U \in \mathbb{C}^n \times n : U^* U = I\}$ is compact and connected, so is $W_C(A)$. Now that $W_C(A)$ is compact and connected, contains the origin, and satisfies condition (c) of Theorem 2.1 (which is equivalent to all other conditions in the theorem), it follows that $W_C(A)$ is a compact circular disc on the complex plane centered at the origin. Therefore condition (g) holds. By taking $C = A^*$, we then have condition (h). That (h) implies condition (d) of Theorem 2.1 is obvious.
Also, it is clear that (g) implies (i). Conversely, if condition (i) holds, then, by considering \( e^{i\theta} C \) instead of \( C \) in (i) for various \( \theta \in \mathbb{R} \), (g) also holds.

In conditions (b) and (d) of Theorem 2.1, the real scalar \( \theta \) satisfies \( e^{ik\theta} \neq 1 \) for all \( k = 1, \ldots, n \). One may consider the case when \( \theta \) is such that \( e^{im\theta} = 1 \) for some \( 1 \leq m \leq n \). By arguments similar to those in the proof of Theorem 2.1, we obtain the following characterizations for those matrices whose unitary orbits or \( C \)-numerical ranges have some weaker symmetry properties. It turns out that, instead of the block-shift operators in the case of Theorem 2.1, these matrices are the block-cyclic operators.

**Theorem 2.3.** Let \( A \in \mathbb{C}^{n \times n} \). Suppose \( \theta \in \mathbb{R} \) is such that \( e^{im\theta} = 1 \) for some \( 1 \leq m \leq n \). Then the following conditions are equivalent.

(a) \( e^{it} \mathcal{U}(A) = \mathcal{U}(A) \).
(b) \( e^{it} W_C(A) = W_C(A) \) for all \( C \in \mathbb{C}^{n \times n} \).
(c) \( e^{it} W_{A^*}(A) = W_{A^*}(A) \).
(d) \( A \) is unitarily similar to a matrix \( M = (M_{k\ell})_{1 \leq k, \ell \leq m} \) in block form such that all \( M_{kk} \)'s are square blocks and \( M_{k\ell} = 0 \) if \( \ell + 1 \neq k \) and \( k, \ell \neq (1, m) \).
(e) \( \mathbb{C}^n \) can be factored into a direct sum of mutually orthogonal subspaces \( \mathbb{C}^n = S_1 \oplus \cdots \oplus S_m \) such that \( A(S_k) \subset S_{k+1} \) for \( k = 1, \ldots, m-1 \), and \( A(S_1) = S_1 \).

**Proof.** The proof is exactly the same as that of Theorem 2.1 except that the previous condition (1) is replaced by "\( z_k \overline{z}_k = e^{i\theta} \) only if \( k+1 \equiv k \pmod{m} \)."

There are several questions related to the above results that we would like to pose:

(I) In conditions (c) of Theorem 2.1, (g) of Corollary 2.2, and (b) of Theorem 2.3, is it possible to replace \( \mathbb{C}^{n \times n} \) by a smaller subset (which, unlike the subset \( \{ A^* \} \) in conditions (d) of Theorem 2.1, (h) of Corollary 2.2, and (c) of Theorem 2.3, does not depend on \( A \)) such as the collection of all hermitian matrices, unitary matrices, or normal matrices?

(II) If \( B \in \mathbb{C}^{n \times n} \) is of the form \( A + \nu I \) where \( \nu \in \mathbb{C} \) and \( A \) satisfies any of the equivalent conditions in Theorem 2.1 and Corollary 2.2, then \( W_C(B) \) is a circular disc (the center of which may not be the origin) for all \( C \in \mathbb{C}^{n \times n} \). Does the converse hold?

(III) By a result of Westwick [10] (see also [9]), if \( C = \alpha K + \beta I \) where \( \alpha, \beta \in \mathbb{C} \) and \( K \) is hermitian, then \( W_C(X) \) is convex for any \( X \in \mathbb{C}^{n \times n} \). By the discussion in (II), if \( C = A + \nu I \) where \( A \) satisfies any of the equivalent conditions in Theorem 2.1 and Corollary 2.2, then \( W_C(X) \) is also convex for any \( X \in \mathbb{C}^{n \times n} \). It would be interesting to identify more matrices \( C \) for which \( W_C(X) \) is convex for all \( X \).
Let $H$ be a Hilbert space and $\mathcal{B}(H)$ the set of all continuous linear operators on $H$. As the case of $\dim H < \infty$ (i.e., $H$ is isomorphic to $\mathbb{C}^n$ for some $n$) is already treated in §2, we may well assume that $H$ is infinite-dimensional in this section, though all the arguments below are still valid even if $H$ is of finite dimension. Let $A$ and $A_t$ ($t > a$) be in $\mathcal{B}(H)$. We write $\text{s.lim}_{t \searrow a} A_t = A$ if $\lim_{t \searrow a} \| (A_t - A) x \| = 0$ for all $x \in H$. An operator $U \in \mathcal{B}(H)$ is called unitary if $U^* U = I$. Similar to that in the previous sections, the unitary orbit of $A \in \mathcal{B}(H)$ is defined as $\mathcal{U}(A) = \{ UAU^* : U \text{ is unitary} \}$. For a given real $\theta$, we shall study the condition $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ and try to give equivalent characterization for those $A$ which satisfy it.

The following spectral resolution formula for unitary operators on $H$ is well known (e.g., see [11]).

**Lemma 3.1.** Let $U \in \mathcal{B}(H)$ be unitary. Then it admits a spectral resolution $U = \int_0^{2\pi} e^{i t} dP_t$, where the integral is the norm-convergence limit of the corresponding Riemann sums, and $\{ P_t : t \in [0, 2\pi] \}$ is a family of orthogonal projections on $H$ which satisfy

\begin{align*}
(2a) & \quad P_0 = 0, \quad P_{2\pi} = I, \\
(2b) & \quad P_a P_b = P_{\min(a, b)} \quad \text{for all } a, b \in [0, 2\pi], \\
(2c) & \quad \text{s.lim}_{t \searrow a} P_t = P_a \quad \text{for all } a \in [0, 2\pi].
\end{align*}

Moreover, $U^n = \int_0^{2\pi} e^{int} dP_t$ for all integers $n$. In particular, $U^* = \int_0^{2\pi} e^{-it} dP_t$.

Let $\{ P_t : t \in [0, 2\pi] \}$ be a family of orthogonal projections on $H$ which satisfies (2). For any $0 < a < b < 2\pi$, define $P(a, b) = P_b - P_a$. The following lemma is easy to verify.

**Lemma 3.2.** Let $\{ P_t : t \in [0, 2\pi] \}$ be a family of orthogonal projections on $H$ that satisfies (2).

(a) For any partition $0 = a_0 < a_1 < \cdots < a_m = 2\pi$, the family of operators $P(a_{k-1}, a_k)$, $k = 1, \ldots, m$, forms a complete set of mutually orthogonal projections on $H$.

(b) If $(a, b) \subset (c, d) \subset (0, 2\pi)$, then $P(a, b) P(c, d) = P(c, d) P(a, b) = P(a, b)$.

We remark that some of the operators $P(a_{k-1}, a_k)$ in Lemma 3.2(a) may be zero.

**Lemma 3.3.** Let $U$ be a unitary operator on $H$ which admits the spectral resolution described in Lemma 3.1. Then for any $0 \leq a < b \leq 2\pi$,

(a) $U^n P(a, b) = P(a, b) U^n$ for all integers $n$;

(b) $\| P(a, b) (I - e^{-ia} U) \| = \| P(a, b) (I - e^{ia} U^*) \| \leq (b - a)$. 

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
Proof. (a) It suffices to prove the particular case of \( n = 1 \), and the general result will follow by repeated application of the particular case. Write \( U \) as the (norm-convergence) limit of the Riemann sums \( \sum_{k=1}^{m} e^{iak} P(a_{k-1}, a_k) \), and then apply Lemma 3.2(b).

(b) The first equality holds because

\[
(P(a, b)(I - e^{-ia} U))^* = P(a, b)(I - e^{ia} U^*).
\]

For the second inequality, notice that

\[
\|P(a, b)(I - e^{ia} U^*)\| = \left| \int_a^b (1 - e^{ia}) dP_t \right| \\
\leq |1 - e^{ia}| \left| \int_a^b dP_t \right| \\
\leq |1 - e^{ia}| \leq |a - b|.
\]

Let \( 0 \leq a < b \leq 2\pi \) and \( \theta \in [0, 2\pi] \). We identify \( P(a + \theta, b + \theta) \) with \( P(a', b') \), where \( a', b' \in [0, 2\pi] \) with \( a' \equiv a + \theta \) (mod \( 2\pi \)) and \( b' \equiv b + \theta \) (mod \( 2\pi \)), and use the convention that \( P(a' b') = P(a', 2\pi) + P(0, b') = I + P_{b'} - P_{a'} \) if \( b' \leq a' \).

Theorem 3.4. Let \( A \in \mathfrak{B}(H) \) and \( \theta \in [0, 2\pi] \) be such that \( e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A) \). Then there exists a family \( \{P_t : t \in [0, 2\pi]\} \) of orthogonal projections on \( H \) which satisfies (2), such that

\[
\text{Im} AP(a, b) \subset \text{Im} P(a + \theta, b + \theta) \quad \text{for all} \quad 0 \leq a < b \leq 2\pi.
\]

Proof. Let \( e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A) \) for some \( \theta \in [0, 2\pi] \). Then \( e^{i\theta} A + UAU^* \) for some unitary \( U \). By Lemma 3.1, \( U = \int_0^{2\pi} e^{it} dP_t \) for some family \( \{P_t : t \in [0, 2\pi]\} \) of orthogonal projections satisfying (2). We are now going to prove that (3) holds. We need only to prove the case of \( (a, b) \neq (0, 2\pi) \), for if \( (a, b) = (0, 2\pi) \) then \( P(a + \theta, b + \theta) = I \). We divide the proof into 3 steps.

Step 1. Show that for \( (a, b), (c, d) \subset (0, 2\pi) \), if \( (b - a) + (d - c) < |1 - e^{i(\theta + a - c)}| \) then \( P(c, d)AP(a, b) = 0 \). Since \( A - e^{i\theta} U^* AU = 0 \), we have

\[
0 = P(c, d)A - e^{i\theta} U^* AU)P(a, b) \\
= P(c, d)AP(a, b) - e^{i\theta} P(c, d)U^* AU P(a, b) \\
= S_1 + S_2 + S_3,
\]

where

\[
S_1 = (1 - e^{i(\theta + a - c)}) P(c, d)AP(a, b), \\
S_2 = e^{i(\theta + a - c)} P(c, d)AP(a, b) - e^{-i\theta} U P(a, b), \\
S_3 = e^{i(\theta - c)} P(c, d) - e^i P(c, d)U^* AU P(a, b).
\]
Now
\[\|S_2\| = \|P(c, d)A(P(a, b) - e^{-ia}UP(a, b))\|
= \|P(c, d)AP(a, b)P(a, b)(I - e^{-ia}U)\|
\leq \|P(c, d)AP(a, b)\| \|P(a, b)(I - e^{-ia}U)\|
\leq (b - a)\|P(c, d)AP(a, b)\|.
\]
Here we have used Lemma 3.3(a) and (b). Similarly, we have
\[\|S_3\| = \|\{P(c, d) - e^{ic}P(c, d)U^*\}AUP(a, b)\|
= \|P(c, d)(I - e^{ic}U^*)P(c, d)AP(a, b)U\|
\leq \|P(c, d)(I - e^{ic}U^*)\| \|P(c, d)AP(a, b)\| \|U\|
\leq (d - c)\|P(c, d)AP(a, b)\|.
\]
Then
\[0 = \|S_1 + S_2 + S_3\| \geq \|S_1\| - \|S_2\| - \|S_3\|
\geq (\|1 - e^{i(\theta + a - c)}\| - (b - a) - (d - c)\|P(c, d)AP(a, b)\|).
\]
As \(|1 - e^{i(\theta + a - c)}| > (b - a) + (d - c)\), we must have \(P(c, d)AP(a, b) = 0\).

Step 2. Show that for \((a, b), (c, d) \subset (0, 2\pi)\), if \(\inf\{|1 - e^{ia':c'}: a' \in (a + \theta, b + \theta], c' \in (c, d)\} > 0\), then \(P(c, d)AP(a, b) = 0\). Suppose
\[\inf\{|1 - e^{ia':c'}: a' \in (a + \theta, b + \theta], c' \in (c, d)\} = \epsilon > 0.
\]
Partition \([a, b], [c, d]\) into sub-intervals \([a', b'], [c', d']\), respectively, such that \((b' - a'), (d' - c') < \frac{\pi}{\epsilon}\). Then, by the result of Step 1, \(P(c', d')AP(a', b') = 0\). Hence
\[P(c, d)AP(a, b) = \sum P(c', d')A \sum P(a', b')
= \sum \sum P(c', d')AP(a', b') = 0.
\]

Step 3. Show that for \((a, b), (c, d) \subset (0, 2\pi)\), if \(\{e^{ia'}: a' \in (a + \theta, b + \theta]\} \cap \{e^{ic'}: c' \in (c, d]\} = \emptyset\), then \(P(c, d)AP(a, b) = 0\). Suppose \(\{e^{ia'}: a' \in (a + \theta, b + \theta]\} \cap \{e^{ic'}: c' \in (c, d]\} = \emptyset\). By the result of Step 2, we need only to consider the case of
\[\inf\{|1 - e^{ia':c'}: a' \in (a + \theta, b + \theta], c' \in (c, d)\} = 0,
\]
or, equivalently, \(d \equiv a + \theta \pmod{2\pi}\) or \(c \equiv b + \theta \pmod{2\pi}\). Without loss of generality, let \(d = a + \theta\). Then, by the result of Step 2, \(P(c, d)AP(t, b) = 0\) for all \(t \in (a, b)\). Since \(P(a, b) = \slim_{t \downarrow a} P(t, b)\), we have \(P(c, d)AP(a, b) = \slim_{t \downarrow a} P(c, d)AP(t, b) = 0\).

By the result of Step 3, and using Lemma 3.2(a), we can conclude that \((I - P(a + \theta, b + \theta))AP(a, b) = 0\). Hence \(\text{Im } AP(a, b) \subset \text{Im } P(a + \theta, b + \theta)\).

If \(\theta \in \mathbb{R}\) is a rational multiple of \(\pi\), then Theorem 3.4 implies the following result, which is analogous to Theorem 2.3.
Theorem 3.5. Let $A \in \mathcal{B}(H)$. Suppose $\theta \in \mathbb{R}$ is a rational multiple of $\pi$ that satisfies

$$e^{ik\theta} \neq 1 \quad \text{for } k = 1, \ldots, m - 1, \text{ and } e^{im\theta} = 1.$$ 

Then the following conditions are equivalent.

(a) $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$.

(b) $H$ can be factored into a direct sum of mutually orthogonal subspaces $H = S_1 \oplus \cdots \oplus S_m$, such that

$$A(S_k) \subset S_{k+1} \quad \text{for } k = 1, \ldots, m - 1, \text{ and } A(S_m) \subset S_1.$$ 

In other words, the operator matrix of $A$ with respect to the direct sum $S_1 \oplus \cdots \oplus S_m$ is of the form given in Theorem 2.3(d).

Proof. Suppose (a) holds. Then $e^{ik\theta} \mathcal{U}(A) = \mathcal{U}(A)$ for all integers $k$. By (4), $e^{i\theta}$ generates a cyclic group (under usual multiplication) of order $m$. Hence $e^{i2\pi/m} = e^{ik\theta}$ for some integer $k$. Therefore $e^{i2\pi/m} \mathcal{U}(A) = \mathcal{U}(A)$. By Theorem 3.4, there exists a family $\{P_t: t \in [0, 2\pi]\}$ of orthogonal projections on $H$ which satisfies (2) and (3) with $\theta = \frac{2\pi}{m}$. For $k = 1, \ldots, m$, define $S_k = \text{Im} P(k \frac{2\pi}{m}, (k + 1) \frac{2\pi}{m})$. Then $S_1, \ldots, S_m$ satisfy (5). Moreover, by Lemma 3.2(a), $S_1, \ldots, S_m$ form a complete set of mutually orthogonal subspaces of $H$. Hence (b) holds.

Conversely, if (b) holds and $\theta$ satisfies (4), then define a unitary $U \in \mathcal{B}(H)$ by $U(\sum_{k=1}^m x_k) = \sum_{k=1}^m e^{ik\theta} x_k$, where $x_k \in S_k$ for $k = 1, \ldots, m$. By direct verification, we have $U A U^* = e^{i\theta} A$. Hence $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$.

Notice that some of the subspaces $S_k$ in Theorem 3.5(b) may be zero. With this in mind, one may apply Theorem 3.5 (which is true no matter whether $H$ is infinite-dimensional or not) to prove part of the statement in Theorem 2.1. For example, if condition (b) of Theorem 2.1 holds for an $A \in \mathbb{C}^{n \times n}$ and a real $\theta$ such that $e^{i\theta}$ has finite order $m > n$, then, by Theorem 3.5, $\mathbb{C}^n$ can be factored into a direct sum $\mathbb{C}^n = S_1 \oplus \cdots \oplus S_m$ of mutually orthogonal subspaces such that (5) holds. However, since $m > n$, at most $n$ of the $S_k$'s can be nonzero. By suitably renaming the $S_k$'s, one gets (f) of Theorem 2.1. Of course, in the finite-dimensional case, we would prefer the original proofs given in §2 for their simplicity and algebraic nature.

Finally, we point out the following open questions.

(1) One would naturally ask what might happen if $\theta$ is not a rational multiple of $\pi$ in Theorem 3.5. Consider the following conditions:

(a) $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ for some $\theta \in \mathbb{R}$ which is not a rational multiple of $\pi$.

(b) $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$ for all $\theta \in \mathbb{R}$.

(c) $\{S_k\}_{k \in \mathbb{Z}}$ is a family of mutually orthogonal subspaces of $H$ such that, if $P_k$ stands for the orthogonal projection of $H$ onto $S_k$, then $\lim_{m \to +\infty} \|x - \sum_{k=-m}^m P_k x\| = 0$ for all $x \in H$, and $A(S_k) \subset S_{k+1}$ for all integers $k$. 

It is not difficult to see that (c) $\Rightarrow$ (b) $\Rightarrow$ (a) is always true. Whether the converses are true in general is not known. Notice that, if $\mathcal{U}(A)$ is closed then one will have (a) $\Rightarrow$ (b). However, when $H$ is infinite-dimensional, $\mathcal{U}(A)$ may fail to be closed [1].

(2) Let $H$ be infinite-dimensional, and $f$ a bounded linear functional on the Banach space $\mathcal{B}(H)$, the norm in which is the operator norm. As a generalization of the $C$-numerical range in the finite-dimensional case, one may consider the image $f(\mathcal{U}(A))$ of $\mathcal{U}(A)$ under $f$. Clearly, if $e^{i\theta} \mathcal{U}(A) = \mathcal{U}(A)$, then $e^{i\theta} f(\mathcal{U}(A)) = f(\mathcal{U}(A))$ for all linear functionals $f$. Whether the converse is true is unknown. If $H$ is separable and one replaces $\mathcal{B}(H)$ by the symmetrically normed ideal $\mathcal{S}_2$ of all Hilbert-Schmidt operators, i.e., those compact operators $A$ on $H$ that satisfy $|\text{tr}(A^*A)| < \infty$ (see [2, Chapter 3] for more definitions and properties), then, since $\mathcal{S}_2$ is a Hilbert space by itself, $f(\mathcal{U}(A))$ is in the form of $W_C(A)$ for some $C \in \mathcal{S}_2$. Using the same proof in the finite-dimensional case, one obtains equivalent statements like those of Theorem 2.1(a)–(d), or Theorem 2.3(a)–(c).

ACKNOWLEDGMENT

Thanks are due to Professor Chandler Davis for his helpful discussion, and to Professor Leiba Rodman for providing the reference [1].

REFERENCES