Product of operators and numerical range preserving maps

by

CHI-KWONG LI (Williamsburg, VA) and NUNG-SING SZE (Hong Kong)

To Professor Miroslav Fiedler on the occasion of his 80th birthday

Abstract. Let $V$ be the $C^*$-algebra $B(H)$ of bounded linear operators acting on the Hilbert space $H$, or the Jordan algebra $S(H)$ of self-adjoint operators in $B(H)$. For a fixed sequence $(i_1, \ldots, i_m)$ with $i_1, \ldots, i_m \in \{1, \ldots, k\}$, define a product of $A_1, \ldots, A_k \in V$ by $A_1 \cdots \cdots A_k = A_{i_1} \cdots A_{i_m}$. This includes the usual product $A_1 \cdots \cdots A_k = A_1 \cdots A_k$ and the Jordan triple product $A*B = ABA$ as special cases. Denote the numerical range of $A \in V$ by $W(A) = \{(Ax, x) : x \in H, (x, x) = 1\}$. If there is a unitary operator $U$ and a scalar $\mu$ satisfying $\mu^2 = 1$ such that $\phi : V \to V$ has the form $A \mapsto \mu U^*AU$ or $A \mapsto \mu U^*A^t U$, then $\phi$ is surjective and satisfies

$$W(A_1 \cdots \cdots A_k) = W(\phi(A_1) \cdots \cdots \phi(A_k)) \quad \text{for all } A_1, \ldots, A_k \in V.$$ 

It is shown that the converse is true under the assumption that one of the terms in $(i_1, \ldots, i_m)$ is different from all other terms. In the finite-dimensional case, the converse can be proved without the surjectivity assumption on $\phi$. An example is given to show that the assumption on $(i_1, \ldots, i_m)$ is necessary.

1. Introduction. Let $H$ be a Hilbert space having dimension at least 2. Denote by $B(H)$ the $C^*$-algebra of bounded linear operators acting on $H$, and $S(H)$ the Jordan algebra of self-adjoint operators in $B(H)$. If $H$ has dimension $n < \infty$, then $B(H)$ is identified with the algebra $M_n$ of $n \times n$ complex matrices and $S(H)$ is identified with the set $S_n$ of $n \times n$ complex Hermitian matrices. Define the numerical range of $A \in B(H)$ by

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

2000 Mathematics Subject Classification: 47A12, 47B15, 47B49, 15A60, 15A04, 15A18.

Key words and phrases: numerical range, Jordan triple product.

Li is an honorary professor of the Heilongjiang University and an honorary professor of the University of Hong Kong. His research was partially supported by a USA NSF grant and a HK RCG grant.

The research of Sze was partially supported by a HKU studentship and a HK RGC grant. He would like to thank Dr. Jor-Ting Chan for his guidance and encouragement.
Let $U \in B(H)$ be a unitary operator, and define a mapping $\phi$ on $B(H)$ or $S(H)$ by

$$A \mapsto U^*AU \quad \text{or} \quad A \mapsto U^*A^tU,$$

where $A^t$ is the transpose of $A$ with respect to a fixed orthonormal basis. (We will always use this interpretation of $A^t$ in our discussion.) Then $\phi$ is a bijective linear map preserving the numerical range, i.e., $W(\phi(A)) = W(A)$ for all $A$.

There has been considerable interest in studying the converse of the above statement. Pellegrini [8] obtained an interesting result on numerical range preserving maps on a general $C^*$-algebra, which implies that a surjective linear map $\phi : B(H) \to B(H)$ preserving the numerical range must be of the above form. Furthermore, by the result in [7], the same conclusion also holds for linear maps $\phi$ acting on $S(H)$. In [6], the author showed that additive preservers of the numerical range of matrices must be linear and have the standard form $A \mapsto U^*AU$ or $A \mapsto U^*A^tU$. In [2], it was shown that a multiplicative map $\phi : M_n \to M_n$ satisfies $W(\phi(A)) = W(A)$ for all $A \in M_n$ if and only if $\phi$ has the form $A \mapsto U^*AU$ for some unitary matrix $U \in M_n$. In [5], the authors replaced the condition that “$\phi$ is multiplicative and preserves the numerical range” on the surjective map $\phi : B(H) \to B(H)$ by the condition that “$W(AB) = W(\phi(A)\phi(B))$ for all $A, B$”, and showed that such a map has the form $A \mapsto \pm U^*AU$ for some unitary operator $U \in B(H)$. They also showed that a surjective map $\phi : B(H) \to B(H)$ satisfies $W(ABA) = W(\phi(A)\phi(B)\phi(A))$ for all $A, B \in B(H)$ if and only if $\phi$ has the form $A \mapsto \mu U^*AU$ or $A \mapsto \mu U^*A^tU$ for some unitary operator $U \in B(H)$ and $\mu \in \mathbb{C}$ with $\mu^3 = 1$. Similar results for mappings on $S(H)$ were also obtained. Recently, Gau and Li [3] obtained a similar result for surjective maps $\phi : \mathcal{V} \to \mathcal{V}$, where $\mathcal{V} = B(H)$ or $S(H)$, preserving the numerical range of the Jordan product, i.e., $W(AB + BA) = W(\phi(A)\phi(B) + \phi(B)\phi(A))$ for all $A, B \in \mathcal{V}$. Specifically, they showed that such a map must be of the form $A \mapsto \pm U^*AU$ or $A \mapsto \pm U^*A^tU$ for some unitary operator $U \in B(H)$. Moreover, the surjectivity assumption can be removed in the finite-dimensional case.

It is interesting that all the results mentioned in the preceding paragraph illustrate that under some mild assumptions, a numerical range preserving map $\phi$ is a $C^*$-isomorphism on $B(H)$ or a Jordan isomorphism on $S(H)$ up to a scalar multiple. Following this line of study, we consider a product of matrices involving $k$ matrices with $k \geq 2$ which includes the usual product $A_1 \cdots A_k = A_1 \cdots A_k$, and the Jordan triple product $A \ast B = ABA$. We prove the following result.

**Theorem 1.1.** Let $(\mathbb{F}, \mathcal{V}) = (\mathbb{C}, B(H))$ or $(\mathbb{R}, S(H))$. Fix a positive integer $k$ and a finite sequence $(i_1, \ldots, i_m)$ such that $\{i_1, \ldots, i_m\} = \{1, \ldots, k\}$
and there is an $i_r$ not equal to $i_s$ for all other $s$. For $A_1, \ldots, A_k \in V$, let

$$A_1 \ast \cdots \ast A_k = A_{i_1} \cdots A_{i_m}.$$ 

A surjective map $\phi : V \to V$ satisfies

$$W(\phi(A_1) \ast \cdots \ast \phi(A_k)) = W(A_1 \ast \cdots \ast A_k) \quad \text{for all } A_1, \ldots, A_k \in V$$

if and only if there exist a unitary operator $U \in B(H)$ and a scalar $\mu \in \mathbb{F}$ with $\mu^m = 1$ such that one of the following conditions holds:

(a) $\phi$ has the form $A \mapsto \mu U^* AU$.

(b) $r = (m+1)/2$, $(i_1, \ldots, i_m) = (i_m, \ldots, i_1)$, and $\phi$ has the form $A \mapsto \mu U^* A^t U$.

(c) $V = S_2$, $(i_{r+1}, \ldots, i_m, i_1, \ldots, i_{r-1}) = (i_{r-1}, \ldots, i_1, i_m, \ldots, i_{r+1})$ and $\phi$ has the form $A \mapsto \mu U^* A^t U$.

Here $A^t$ denotes the transpose of $A$ with respect to a certain orthonormal basis of $H$. Furthermore, if the dimension of $H$ is finite, then the surjectivity assumption on $\phi$ can be removed.

Note that the assumption that there is $i_r \notin \{i_1, \ldots, i_{r-1}, i_{r+1}, \ldots, i_m\}$ is necessary. For example, if $A \ast B = ABBA$, then mappings $\phi$ satisfying $W(\phi(A) \ast \phi(B)) = W(A \ast B)$ may not have nice structure. For instance, $\phi$ can send all involutions, i.e., those operators $X \in B(H)$ such that $X^2 = I_H$, to a fixed involution, and $\phi(X) = X$ for other $X$.

For the usual product $A_1 \ast \cdots \ast A_k = A_1 \cdots A_k$ and the Jordan triple product $A \ast B = ABA$, Hou and Di [5] have also obtained the result for $B(H)$ in Theorem 1.1 with the surjectivity assumption. Our result is stronger when $H$ is finite-dimensional.

It turns out that Theorem 1.1 can be deduced from the following special case.

**Theorem 1.2.** Let $(\mathbb{F}, V) = (\mathbb{C}, B(H))$ or $(\mathbb{R}, S(H))$. Suppose $r$, $s$ and $m$ are nonnegative integers such that $m - 1 = r + s > 0$. A surjective map $\phi : V \to V$ satisfies

$$W(\phi(A)^r \phi(B) \phi(A)^s) = W(A^r BA^s) \quad \text{for all } A, B \in V$$

if and only if there exist a unitary operator $U \in B(H)$ and a scalar $\mu \in \mathbb{F}$ with $\mu^m = 1$ such that one of the following conditions holds:

(a) $\phi$ has the form $A \mapsto \mu U^* AU$.

(b) $r = s$ and $\phi$ has the form $A \mapsto \mu U^* A^t U$.

(c) $V = S_2$ and $\phi$ has the form $A \mapsto \mu U^* A^t U$.

Here $A^t$ denotes the transpose of $A$ with respect to a certain orthonormal basis of $H$. Furthermore, if the dimension of $H$ is finite, then the surjectivity assumption on $\phi$ can be removed.
We present some auxiliary results in Section 2, and the proofs of the theorems in Section 3.

2. Auxiliary results. For any $x, y \in H$, denote by $xy^*$ the rank one operator $(xy^*)z = (z, y)x$ for all $z \in H$. Then for any operator $A \in B(H)$ with finite rank, $A$ can be written as $x_1y_1^* + \cdots + x_ky_k^*$ for some $x_i, y_i \in H$. Define the trace of $A$ by

$$\text{tr}(A) = (x_1, y_1) + \cdots + (x_k, y_k).$$

If $H$ is finite-dimensional, $\text{tr}(A)$ is equivalent to the usual matrix trace, i.e., the sum of all diagonal entries of the matrix $A$. For each positive integer $m$, let

$$\mathcal{R}^m = \{\mu xx^* : \mu \in \mathbb{F} \text{ and } x \in H \text{ with } (x, x) = 1 = \mu^m\}.$$ 

Note that $\mathcal{R}^1$ is the set of Hermitian rank one idempotents, and for all $m > 1, \mathcal{R}^1 \subseteq \mathcal{R}^m$.

**Proposition 2.1.** Let $V = B(H)$ or $S(H)$ and $\mathbb{F} = \mathbb{C}$ or $\mathbb{R}$ accordingly. Suppose $m$ is a positive integer with $m > 1$, and $\phi : V \to V$ is a map satisfying

$$(2.1) \quad \text{tr}(\phi(A)^{m-1}\phi(B)) = \text{tr}(A^{m-1}B) \quad \text{for all } A \in \mathcal{R}^m \text{ and } B \in \mathcal{V}.$$ 

If $H$ is finite-dimensional, then $\phi$ is an invertible $\mathbb{F}$-linear map. If $H$ is infinite-dimensional and $\phi(\mathcal{R}^m) = \mathcal{R}^m$, then $\phi$ is $\mathbb{F}$-linear.

**Proof.** Suppose $H$ is finite-dimensional. We use an argument similar to that in the proof of Proposition 1.1 in [1]. Let $V = M_n$ or $S_n$. For every $X = (x_{ij}) \in V$, let $R_X$ be the $n^2$ row vector

$$R_X = (x_{11}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n}, \ldots, x_{n1}, \ldots, x_{nn}),$$

and $C_X$ the $n^2$ column vector

$$C_X = (x_{11}, x_{21}, \ldots, x_{1n}, x_{12}, x_{22}, \ldots, x_{n2}, \ldots, x_{1n}, \ldots, x_{nn})^t.$$ 

Then we deduce from (2.1) that for all $A \in \mathcal{R}^m$ and $B \in \mathcal{V}$,

$$(2.2) \quad R_{\phi(A)^{m-1}C_{\phi(B)}} = \text{tr}(\phi(A)^{m-1}\phi(B)) = \text{tr}(A^{m-1}B) = R_{A^{m-1}C_B}.$$ 

Note that we can choose $A_1, \ldots, A_{n^2}$ in $\mathcal{R}^m$ such that $\{A_1^{m-1}, \ldots, A_{n^2}^{m-1}\}$ forms a basis for $V$. Let $\Delta$ and $\Delta_{\phi}$ be $n^2 \times n^2$ matrices having rows $R_{A_1^{m-1}}, \ldots, R_{A_{n^2}^{m-1}}$ and $R_{\phi(A_1)^{m-1}}, \ldots, R_{\phi(A_{n^2})^{m-1}}$, respectively. By (2.2),

$$\Delta_{\phi}C_{\phi(B)} = \Delta C_B \quad \text{for all } B \in \mathcal{V}.$$ 

Now take a basis $\{B_1, \ldots, B_{n^2}\}$ in $V$ and let $\Omega$ and $\Omega_{\phi}$ be the $n^2 \times n^2$ matrices having columns $C_{B_1}, \ldots, C_{B_{n^2}}$ and $C_{\phi(B_1)}, \ldots, C_{\phi(B_{n^2})}$, respectively. Then $\Delta_{\phi}C_{\phi(B)} = \Delta C_B$. Note that both $\Delta$ and $\Omega$ are invertible, hence so is $\Delta_{\phi}$.
Therefore, for any $B \in \mathbf{V}$, 
\[ C_{\phi(B)} = \Delta_{\phi}^{-1} \Delta C_B. \]
Hence, $\phi$ is invertible and $\mathbb{F}$-linear.

Next, suppose $H$ is infinite-dimensional and $\phi(\mathcal{R}^m) = \mathcal{R}^m$. Take any $X, Y \in \mathbf{V}$. For any $x \in H$ with $(x, x) = 1$, since $\mathcal{R}^1 \subseteq \mathcal{R}^m = \phi(\mathcal{R}^m)$, there is $A \in \mathcal{R}^m$ such that $\phi(A) = xx^*$. Then $\phi(A)^{m-1} = xx^*$ and
\[
(\phi(X + Y)x, x) = \text{tr}(xx^* \phi(X + Y)) = \text{tr}(\phi(A)^{m-1} \phi(X + Y)) = \text{tr}(A^{m-1}X + \phi(A)^{m-1}Y) = \text{tr}(\phi(A)^{m-1} \phi(X)) + \text{tr}(\phi(A)^{m-1} \phi(Y)) = (\phi(X)x, x) + (\phi(Y)x, x).
\]
Since this is true for all unit vectors $x \in H$, it follows that $\phi(X + Y) = \phi(X) + \phi(Y)$. Similarly, we can show that $\phi(\lambda X) = \lambda \phi(X)$ for all $\lambda \in \mathbb{F}$ and $X \in \mathbf{V}$. $
$
It is well known that if $A \in M_2$ then $W(A)$ is an elliptical disk with the eigenvalues of $A$ as foci. Moreover, if $A \in B(H)$ is unitarily similar to $A_1 \oplus A_2$ then $W(A)$ is the convex hull of $W(A_1) \cup W(A_2)$. In particular, if $A$ has rank one, then $A$ is unitarily similar to $C \oplus 0$, where $C$ has a matrix representation \((a \ b)\); hence $W(A) = W(C)$ is an elliptical disk with 0 as a focus. These facts are used in the proof of the following lemma, which is an extension of a result in [5].

**Lemma 2.2.** Let $r$ and $s$ be two nonnegative integers with $r + s > 0$. For any $B \in B(H)$, $B$ has rank one if and only if for all $A \in B(H)$, $W(A^r BA^s)$ is an elliptical disk with zero as one of the foci.

*Proof.* Let $B \in B(H)$. If $B$ is rank one, then so is $A^r BA^s$. Therefore $W(A^r BA^s)$ is an elliptical disk with 0 as a focus by the discussion before the lemma.

Conversely, suppose $B$ has rank at least 2. Then there exist $x, y \in H$ such that $\{Bx, By\}$ is an orthonormal set. Let $C = x(Bx)^* - y(By)^*$. Then $BC = Bx(Bx)^* - By(By)^*$ has numerical range $[-1, 1]$. Suppose $r = 0$. Since $C$ has rank two, it has an operator matrix of the form $C_1 \oplus 0$, where $C_1 \in M_k$ with $2 \leq k \leq 4$, with respect to an orthonormal basis of $H$. Let $D$ have operator matrix $\text{diag}(1, \ldots, k) \oplus 0$ with respect to the same basis. Then $C + \nu D$ has operator matrix $(C_1 + \nu D_1) \oplus 0$, where $D_1 = \text{diag}(1, \ldots, k)$. Except for finitely many $\nu \in \mathbb{R}$, $C_1 + \nu D_1$ has distinct eigenvalues so that there is $A_\nu$ satisfying $A_\nu = C + \nu D$, and $W(BA_\nu^s) = W(BC + \nu BD)$. By [4, Problem 2.20], the mapping $\nu \mapsto \text{Closure}(W(BC + \nu BD))$ is continuous. Since $W(BC) = [-1, 1]$, there is a sufficiently small $\nu > 0$ such that $W(BA_\nu^s)$ is not an elliptical disk with 0 as a focus. If $s = 0$, we can fix an orthonormal basis of $H$, and apply the above argument to $B^\ell$ to show that there exists
A such that \( W(A^r B) = W(B^t (A^t)^r) \) is not an elliptical disk with 0 as a focus.

Now, suppose that \( rs > 0 \). Let \( H_0 \) be the subspace of \( H \) spanned by \( \{x, y, Bx, By\} \), which has dimension \( p \in \{2, 3, 4\} \). Suppose \( B_0 \in M_p \) is the compression of \( B \) on \( H_0 \). Then \( B_0 = PU \) for some positive semidefinite \( P \in S_p \) with rank at least 2, and a unitary matrix \( U \in M_p \). Let \( V \in M_p \) be a unitary matrix such that \( V^* UV \) is in diagonal form. Then \( V^* PV \) is positive semidefinite with rank at least 2. Note that the \( 2 \times 2 \) principal minors of \( V^* PV \) are nonnegative, and their sum is the 2-elementary symmetric function of the eigenvalues of \( V^* PV \), which is positive. So, at least one \( 2 \times 2 \) principal minor of \( V^* PV \) is nonzero. Since \( V^* B_0 V \) is the product of \( V^* PV \) and the diagonal unitary matrix \( V^* UV \), the \( 2 \times 2 \) principal minors of \( V^* B_0 V \) are unit multiples of those of \( V^* PV \). It follows that at least one \( 2 \times 2 \) principal minor of \( V^* B_0 V \) is nonzero. Hence, there exists a two-dimensional subspace \( H_1 \) of \( H_0 \) such that the compression \( B_1 \) of \( B \) on \( H_1 \) is invertible. Suppose \( \{u, v\} \) is an orthonormal basis of \( H_1 \) such that \( B_1 = auu^* + bvv^* + cuv^* \). Then \( \det(B_1) = ac \neq 0 \). Let \( A = \alpha uu^* + \beta vv^* \) so that \( \alpha^r s a = 1 \) and \( \beta^{r+s} c = -1 \). Then \( A^r BA^s = uu^* - vv^* + \alpha^r \beta^s bvv^* \) and \( W(A^r BA^s) \) is an elliptical disk with foci \( 1, -1 \).

Note that the analog of the above result for \( V = S(H) \) does not hold if \( H \) has dimension at least 3. For example, if \( A \ast B = ABA \) and \( B = uu^* + vv^* \) for some orthonormal set \( \{u, v\} \) in \( H \), then \( W(ABA) \) is always a line segment with 0 as an end point. To prove our main theorems, we need a characterization of elements in \( R^m \) when \( V = S(H) \).

**Lemma 2.3.** Let \( r, s \) and \( m \) be nonnegative integers such that \( m - 1 = r + s > 0 \). Suppose \( X \in S(H) \) is such that \( W(X^m) = [0, 1] \). Then \( X \in R^m \) if and only if the following holds:

\[
\text{(i) For any } Y \in S(H) \text{ satisfying } W(Y^m) = [0, 1] = W(X^r Y X^s), \text{ we have }
\{ Z \in S(H) : W(Z^m) = [0, 1], Y^r Z Y^s = 0_H \}
\subseteq \{ Z \in S(H) : W(Z^m) = [0, 1], X^r Z X^s = 0_H \}.
\]

**Proof.** Since \( W(X^m) = [0, 1] \), \( X \) has an eigenvalue \( \mu \) satisfying \( \mu^m = 1 \) with a unit eigenvector \( u \). Assume that \( X \neq \mu uu^* \). Then \( X = \begin{bmatrix} \mu & 0 \\ 0 & X_2 \end{bmatrix} \) on \( H = \text{span}\{u\} \oplus \{u\}^\perp \), where \( X_2 \) is nonzero. Let \( Y = \begin{bmatrix} \mu & 0 \\ 0 & I_{\{u\}} \end{bmatrix} \). Then \( W(Y^m) = [0, 1] = W(X^r Y X^s) \). Note that the operator \( Z = \begin{bmatrix} 0 & 0 \\ 0 & X_2^{m-1} \end{bmatrix} \) satisfies \( W(Z^m) = [0, 1] \) and \( Y^r Z Y^s = 0_H \) but \( X^r Z X^s = [0] \oplus X_2^{m-1} \neq 0_H \).

Conversely, suppose \( X = \mu uu^* \) on \( H = \text{span}\{u\} \oplus \{u\}^\perp \). For any \( Y \in S(H) \) satisfying \( W(Y^m) = [0, 1] = W(X^r Y X^s) \), we have \( Y = \begin{bmatrix} \mu & Y_1 \\ 0 & I_{\{u\}} \end{bmatrix} \) and
W(Y_1^m) \subseteq [0, 1]. Suppose

\[ Z = \begin{pmatrix} \alpha & z_1^* \\ z_1 & Z_2 \end{pmatrix} \]
on span\{u\} \oplus \{u\}^\perp satisfying W(Z^m) = [0, 1] and Y^r Z Y^s = 0_H. If rs > 0 then \( \alpha = 0 \); if rs = 0 then \( \alpha = 0 \) and \( z_1 = 0 \). In both cases, we see that \( X^r Z X^s = 0_H. \) \( \blacksquare \)

3. Proofs of the main theorems

3.1. Proof of Theorem 1.2. We need the following lemma.

Lemma 3.1. Let \( V = M_n \) or \( S_n \), and let \( \phi : V \to V \) be a map satisfying (1.2). Then

(3.1) \( \phi(R^m) \subseteq R^m. \)

Proof. Each matrix \( A \in R^m \) can be written as \( \mu U^* E_{11} U \) for some unitary matrix \( U \) and \( \mu \in \mathbb{F} \) with \( \mu^m = 1 \). It suffices to prove that \( \phi(E_{11}) \in R^m. \)

For the other cases, we may replace the map \( \phi \) by the map \( A \mapsto \phi(\mu U^* AU). \)

We first consider the case when \( V = S_n \). For \( i = 1, \ldots, n \), let \( F_i = \phi(E_{ii}). \)

Since \( E_{ii}^r E_{jj} E_{ii}^s = 0_n \) for all \( i \neq j \), we have

\[ W(F_i^r F_j F_i^s) = W(E_{ii}^r E_{jj} E_{ii}^s) = W(0_n) = \{0\}. \]

It follows that \( F_i^r F_j F_i^s = 0_n \) for all \( i \neq j \).

We claim that \( F_i F_j = F_j F_i = 0_n \) for all \( i \neq j \). If the claim holds, then there are \( \alpha_1, \ldots, \alpha_n \in \mathbb{R} \) and a unitary matrix \( V \) such that \( F_i = \alpha_i V^* E_{ii} V \). Furthermore, as \( W(F_i^m) = W(E_{ii}^m) = [0, 1] \), \( \alpha_i^m = 1 \). Therefore, \( \phi(E_{11}) = F_1 = \alpha V^* E_{11} V \in R^m \) and the result follows.

When \( m \) is odd, as \( W(\phi(I_n)^m) = W(I_n^m) = \{1\} \), we have \( \phi(I_n) = I_n. \)

Then for any \( i = 1, \ldots, n \),

\[ W(F_i) = W(\phi(I_n)^r \phi(E_{ii}) \phi(I_n)^s) = W(I_n^r E_{ii} I_n^s) = W(E_{ii}) = [0, 1]. \]

Thus, \( F_i \) is positive semidefinite. Now for any \( i \neq j \), as \( F_i^r F_j F_i^s = 0_n \), we deduce that \( F_i F_j = F_j F_i = 0_n. \)

When \( m \) is even, since \( W(\phi(I_n)^m) = \{1\} \), the eigenvalues of \( \phi(I_n) \) can be either 1 or \(-1\) only. Write \( \phi(I_n) = V^* (I_p \oplus -I_q) V \) for some unitary matrix \( V \) and nonnegative integers \( p \) and \( q \) such that \( p + q = n \). Then for any \( i = 1, \ldots, n \),

\[ W(\phi(I_n)^r \phi(E_{ii}) \phi(I_n)^s) = W(I_n^r E_{ii} I_n^s) = W(E_{ii}) = [0, 1]. \]

Since one of \( r \) and \( s \) is odd while the other one must be even, either \( \phi(I_n) F_i \) or \( F_i \phi(I_n) \) is positive semidefinite. In both cases, we conclude that \( F_i = V^* (P_i \oplus -Q_i) V \) for some positive semidefinite matrices \( P_i \in H_p \) and \( Q_i \in H_q \). Since \( F_i^r F_j F_i^s = 0_n \), we have \( P_i^r P_j P_i^s = 0_p \) and \( Q_i^r Q_j Q_i^s = 0_q \) for
all \( i \neq j \). Then we conclude that \( P_i P_j = P_j P_i = 0_p \) and \( Q_i Q_j = Q_j Q_i = 0_q \) and hence \( F_i F_j = F_j F_i = 0_n \).

So, our claim is proved and the lemma follows if \( V = S_n \).

Next, we turn to the case when \( V = M_n \). We divide the proof into a sequence of assertions.

**Assertion 1.** Let \( D = \text{diag}(0, e^{i\theta_2}, \ldots, e^{i\theta_n}) \) with \( 0 < \theta_2 < \cdots < \theta_n < \pi/m \).

Then

\[ \phi(D) = V^*([0] \oplus T)V \]

for some unitary matrix \( V \in M_n \) and invertible upper triangular matrix \( T \in M_{n-1} \).

Proof. Note that \( D^m \) has \( n \) distinct eigenvalues and \( W(D^m) \) is a polygon with \( n \) vertices including zero. Since \( W(\phi(D)^m) = W(D^m) \), it follows that \( \phi(D)^m \) has \( n \) distinct eigenvalues, including one zero eigenvalue. Then so has \( \phi(D) \). Therefore, we may write

\[ \phi(D) = V^* \begin{pmatrix} 0 & x^* \\ 0 & T \end{pmatrix} V \]

for some \( x \in \mathbb{C}^{n-1} \), a unitary matrix \( V \) and an upper triangular matrix \( T \in M_{n-1} \) such that all eigenvalues of \( T \) are nonzero. Then \( T \) is invertible. Since \( W(\phi(D)^m) \) is a polygon with \( n \) vertices, \( \phi(D)^m \) is a normal matrix. Note that an upper triangular matrix is normal if and only if it is diagonal. Observe that

\[ \phi(D)^m = V^* \begin{pmatrix} 0 & x^* T^{m-1} \\ 0 & T^m \end{pmatrix} V. \]

It follows that \( x = 0 \) as \( T \) is invertible, i.e., \( \phi(D) = V^*([0] \oplus T)V \). The proof of the assertion is complete.

**Assertion 2.** The inclusion (3.1) holds if \( rs = 0 \).

Proof. Suppose \( r = 0 \). Then, as \( E_{11}D^s = 0_n \), where \( D \) is the matrix defined in Assertion 1, \( \phi(E_{11})\phi(D)^s = 0_n \). It follows that only the first column of \( V^*\phi(E_{11})V \) is nonzero, where \( V \) is the unitary matrix defined in Assertion 1. Hence, \( \phi(E_{11}) \) is a rank one matrix. Note that \( W(\phi(E_{11})^m) = W(E_{11}^m) = [0, 1] \), and since a rank one matrix \( A \in M_n \) satisfies \( W(A^m) = [0, 1] \) if and only if \( A \in \mathcal{R}m \), we conclude that \( \phi(E_{11}) \in \mathcal{R}m \). The proof for \( s = 0 \) is similar. Thus, our assertion is true.

**Assertion 3.** Suppose \( rs > 0 \). For any nonzero

\[ A = \begin{pmatrix} a & w^* \\ z & 0_{n-1} \end{pmatrix} \in M_n, \]
Numerical range preserving maps

we have

$$\phi\left(\begin{pmatrix} a & w^* \\ z & 0_{n-1} \end{pmatrix}\right) = V^*\begin{pmatrix} \alpha & x^* \\ y & 0_{n-1} \end{pmatrix}V$$

for some $\alpha \in \mathbb{C}$ and $x, y \in \mathbb{C}^{n-1}$, where $V$ is the unitary matrix defined in Assertion 1. Furthermore, if $A^m \neq 0$ is Hermitian, then $x = \beta y$ for some nonzero $\beta \in \mathbb{C}$.

Proof. Let $D$ be the matrix defined in Assertion 1. Since $D^rAD^s = 0_n$, it follows that $\phi(D)^r\phi(A)\phi(D)^s = 0_n$. Thus

$$\phi(A) = V^*\begin{pmatrix} \alpha & x^* \\ y & 0_{n-1} \end{pmatrix}V$$

for some $\alpha \in \mathbb{C}$ and $x, y \in \mathbb{C}^{n-1}$, where $V$ is defined in Assertion 1. If $A^m \neq 0$ is Hermitian, then $\phi(A)^m$ is Hermitian as well. Clearly, if $x$ or $y$ is the zero vector, say $x = 0$, then $\alpha \neq 0$ as $A^m \neq 0_n$. Therefore, $y$ must also be zero. Thus the assertion holds.

Now we assume that both $x$ and $y$ are nonzero vectors. By induction, we have

$$\phi(A)^k = V^*\begin{pmatrix} a_{k+1} & a_kx^* \\ a_ky & a_{k-1}y^*x^* \end{pmatrix}V$$

for all $k = 1, 2, \ldots$, where the sequence $\{a_k\}$ satisfies $a_{k+1} = \alpha a_k + x^*y a_{k-1}$ with $a_0 = 0, a_1 = 1$ and $a_2 = \alpha$.

It is impossible to have both $a_m$ and $a_{m-1}$ equal to zero, since then $a_{m+1} = 0$, and hence $\phi(A)^m = 0_n$. Then $W(A^m) = W(\phi(A)^m) = \{0\}$, which contradicts our assumption that $A^m \neq 0_n$. Thus, $a_m$ or $a_{m-1}$ must be nonzero. In both cases, as $A^m$ is Hermitian, we must have $x = \beta y$ for some nonzero $\beta \in \mathbb{C}$. The proof of our assertion is complete.

Assertion 4. The inclusion (3.1) holds if $rs > 0$.

Proof. For $i = 1, \ldots, n$, let $H_i = \frac{1}{2}(E_{1i} + E_{i1})$. Then $H_i^m$ is Hermitian and $H_i^m \neq 0_n$. By Assertion 3, we write

$$\phi(H_i) = V^*\begin{pmatrix} \alpha_i & \beta_i z_i^* \\ z_i & 0_{n-1} \end{pmatrix}V$$

for some $\alpha_i, \beta_i \in \mathbb{C}$ and $z_i \in \mathbb{C}^{n-1}$ with $\beta_i \neq 0$. Denote by $Z_i$ the $n \times 2$ matrix $\begin{pmatrix} 1 & 0 \\ 0 & z_i \end{pmatrix}$ and by $K_i$ the $2 \times 2$ matrix $\begin{pmatrix} \alpha_i & \beta_i \\ 1 & 0 \end{pmatrix}$. Then

$$\phi(H_i) = V^*\begin{pmatrix} 1 & 0 \\ 0 & z_i \end{pmatrix}\begin{pmatrix} \alpha_i & \beta_i \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & z_i^* \end{pmatrix}V = V^*Z_iK_iZ_i^*V.$$
Observe that for any distinct \( i < j \), \( H_i^*H_jH_i^* = 0_n \). Setting \( R_{ij} = Z_i^*Z_j \), we have

\[
0_n = \phi(H_i)^r \phi(H_j)\phi(H_i)^s = V^*Z_i (K_i R_{ii})^{r-1} K_i [R_{ij} K_j R_{ij}] K_i (R_{ii} K_i)^{s-1} Z_i^* V.
\]

Thus, we have \( 0_n = \phi(H_j)^m = W(H_j^m) \neq \{0\} \), we have \( z_j \neq 0 \).

To see this, suppose \( z_i \neq 0 \). Then the \( n \times 2 \) matrix \( Z_i \) has rank 2 and hence the \( 2 \times 2 \) matrix \( Z_i^* Z_i \) is invertible. Also both \( K_i \) and \( K_j \) are invertible. Then (3.2) holds only when

\[
\begin{pmatrix} 1 & 0 \\ z_i^* z_j & 1 \end{pmatrix} \begin{pmatrix} \alpha_j & \beta_j \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & z_j^* z_i \end{pmatrix} = (Z_i^* Z_j) K_j (Z_j^* Z_i) = 0_2.
\]

Thus, \( \beta_j z_j^* z_i = z_i^* z_j = \alpha_j = 0 \). Finally, since \( W(\phi(H_j)^m) = W(H_j^m) \neq \{0\} \), we have \( z_j \neq 0 \).

Now we must have \( z_1 = 0 \). Otherwise, \( \alpha_j = z_i^* z_j = 0 \) and \( z_j \neq 0 \) for all \( j = 2, \ldots, n \). We can then further deduce that \( z_i^* z_j = 0 \) for all \( i \neq j \). Thus, we have \( n \) nonzero orthogonal vectors \( z_1, \ldots, z_n \) in \( \mathbb{C}^{n-1} \), which is impossible. Therefore, \( z_1 = 0 \) and hence \( \alpha_1 \neq 0 \). Finally, as \( W(\phi(H_1)^m) = W(H_1^m) = [0, 1] \), we have \( \alpha_1^m = 1 \). So \( \phi(E_{11}) = \phi(H_1) = \alpha_1 V^*E_{11} V \in H^m \) and the result follows. The proof of our assertion is complete.

Combining the assertions, we get the result for \( V = M_n \).

Proof of Theorem 1.2. First, consider the sufficiency part. If (a) or (b) holds, then clearly \( \phi \) satisfies (1.2). Suppose (c) holds. Then for any \( A, B \in S_2 \), there is a unitary \( V \in M_2 \) such that \( V^* A V = D \) is a real diagonal matrix, and \( V^* B V = C \) is a real symmetric matrix. Thus,

\[
\phi(A^r B A^s) = W(D^r C D^s) = W(D^r C D^s) = W((D^t)^r C^t (D^t)^s) = W(\phi(A)^r \phi(B) \phi(A)^s).
\]

Next we turn to the necessity. Suppose \( V = B(H) \) or \( S(H) \). Assume that \( \phi : V \to V \) satisfies (1.2), and that \( \phi \) is surjective if \( H \) is infinite-dimensional. We divide the proof into several steps.

Step 1. We show that \( \phi(R_1^m) = R_1^m \) and \( \phi \) is linear. Suppose \( H \) is finite-dimensional with no surjectivity assumption on \( \phi \). By Lemma 3.1, \( \phi(R_1^m) \subseteq R_1^m \). Suppose \( H \) is infinite-dimensional. For \( V = S(H) \), we have \( \phi(R_1^m) = R_1^m \) by Lemma 2.3 and the surjectivity of \( \phi \). For \( V = B(H) \), by Lemma 2.2 and the surjectivity of \( \phi \), we see that \( \phi \) maps the set of rank one operators onto itself; since a rank one operator \( A \in B(H) \) satisfies \( W(A^m) = [0, 1] \) if and only if \( A \in R_1^m \), we also have \( \phi(R_1^m) = R_1^m \).
Now, for any $A \in \mathcal{R}^m$ and $B \in \mathcal{V}$, both $A^r BA^s$ and $\phi(A)^r \phi(B) \phi(A)^s$ have rank at most one. As a result, $W(A^r BA^s)$ is an elliptical disk with foci $\text{tr}(A^r BA^s)$ and 0, and $W(\phi(A)^r \phi(B) \phi(A)^s)$ is an elliptical disk with foci $\text{tr}(\phi(A)^r \phi(B) \phi(A)^s)$ and 0. Since $W(A^r BA^s) = W(\phi(A)^r \phi(B) \phi(A)^s)$, we conclude that

\begin{equation}
\text{tr}(A^{r+s} B) = \text{tr}(A^r BA^s) = \text{tr}(\phi(A)^r \phi(B) \phi(A)^s) = \text{tr}(\phi(A)^{r+s} \phi(B))
\end{equation}

for all $A \in \mathcal{R}^m$ and $B \in \mathcal{V}$. By Proposition 2.1, $\phi$ is linear. Moreover, if $H$ is finite-dimensional, $\phi$ is invertible. Indeed, $\phi^{-1}$ also satisfies (1.2), and hence (3.1) and (3.3). So, $\phi(\mathcal{R}^m) = \mathcal{R}^m$.

**Step 2.** We show that $\phi(I_H) = \mu I_H$ with $\mu^m = 1$. For any $x \in H$ with $(x, x) = 1$, there are $y \in H$ and $\mu \in \mathbb{F}$ with $(y, y) = \mu^m = 1$ such that $\phi(\mu yy^*) = xx^*$. Then by (3.3),

\[
(\phi(I_H) x, x) = \text{tr}(xx^* \phi(I_H)) = \text{tr}((xx^*)^{m-1} \phi(I_H)) = \text{tr}(\mu yy^*)^{m-1} \phi(I_H)) = \text{tr}((\mu yy^*)^{m-1} I_H) = \mu^{m-1} (y, y) = \mu^{-1}.
\]

It follows that $W(\phi(I_H)) \subseteq \{ \mu^{-1} : \mu^m = 1 \} = \{ \mu : \mu^m = 1 \}$. By the convexity of numerical range, $W(\phi(I_H))$ is a singleton set. Thus, $\phi(I_H) = \mu I_H$ for some $\mu^m = 1$.

**Step 3.** We show that $\phi$ has the asserted form. Using the result in Step 2, and replacing $\phi$ by the map $A \mapsto \mu^{-1} \phi(A)$, we have $\phi(I_H) = I_H$. Furthermore,

\[W(\phi(A)) = W(\phi(I_H)^r \phi(A) \phi(I_H)^s) = W(I_H^r A I_H^s) = W(A) \quad \text{for all } A \in \mathcal{V}.
\]

Since $\phi$ is linear, by the results in [7, 8] the map $\phi$ has the form

\[A \mapsto U^* A U \quad \text{or} \quad A \mapsto U^* A^t U
\]

for some unitary operator $W \in B(H)$.

**Step 4.** It remains to show that $r = s$ when $\mathcal{V} \neq S_2$ and $\phi$ has the form $A \mapsto U^* A^t U$. For any $A, B \in \mathcal{V}$,

\[W(A^s BA^r) = W((A^t)^s B^t (A^t)^r) = W(U^*(A^t)^r B^t (A^t)^s U) = W(\phi(A)^r \phi(B) \phi(A)^s) = W(A^r BA^s).
\]

For $\mathcal{V} = B(H)$, let $\{ u, v \}$ be an orthonormal set in $H$, $A = uu^* + uv^* + vv^*$ and $B = vv^*$. Then

\[W(suv^* + vv^*) = W(A^s BA^r) = W(A^r BA^s) = W(ruv^* + vv^*).
\]

Thus, $r = s$ and the result follows.

Now consider $\mathcal{V} = S(H)$, where $H$ has dimension at least 3. Suppose $r \neq s$. Without loss of generality, we assume that $r > s$. Let $A, B \in S(H)$ be such that

\[A^{r-s} = D \oplus 0 \quad \text{and} \quad A^s B A^s = E \oplus 0,
\]
where
\[ D = \text{diag}(3, 2, 1) \quad \text{and} \quad E = \begin{pmatrix} 1 & 1 & i \\ 1 & 0 & 1 \\ -i & 1 & 0 \end{pmatrix} \]
with respect to a suitable orthonormal basis. Then
\[
W(DE \oplus 0) = W(A^*BA^*) = W(A^*BA^*) = W(ED \oplus 0)
\]
\[ = W(DE \oplus 0) = W(DE \oplus 0). \]
Therefore, \( W(DE \oplus 0) \) is symmetric about the real axis. But this is impossible as the eigenvalues of \( DE - ED \) are \( 2i, (\sqrt{3} - 1)i/2 \) and \( (-\sqrt{3} - 1)i/2 \). Hence \( \{\text{Im} \, z : z \in W(DE \oplus 0_{n-3})\} = [(-\sqrt{3} - 1)/2, 2] \) so that the two horizontal support lines of \( W(DE \oplus 0) \) are \( \{z : \text{Im} \, z = 2\} \) and \( \{z : \text{Im} \, z = (-\sqrt{3} - 1)/2\} \), which is a contradiction. Therefore, we must have \( r = s \).

The proof of our theorem is complete. \( \blacksquare \)

3.2. Proof of Theorem 1.1. If (a) holds then \( \phi \) clearly satisfies (1.1). Suppose (b) holds. Then for any \( A_1, \ldots, A_k \in \mathbf{V} \), we have
\[
W(\phi(A_1) \cdots \phi(A_k)) = W(\phi(A_{i_1}) \cdots \phi(A_{i_m})) = W(U^*A_{i_1}^t \cdots A_{i_m}^t U) = W((A_{i_m} \cdots A_{i_1})^t)
\]
\[ = W(A_{i_m} \cdots A_{i_1}) = W(A_{i_1} \cdots A_{i_m}) = W(A_1 \cdots \cdots A_k). \]
Suppose (c) holds. Note that \( X, Y \in M_2 \) have the same numerical range if and only if the two matrices have the same eigenvalues and the same Frobenius norm, equivalently, \( \text{tr}(X) = \text{tr}(Y) \), \( \det(X) = \det(Y) \) and \( \text{tr}(XX^*) = \text{tr}(YY^*) \). One readily checks that these conditions are satisfied for \( X = A_1 \cdots \cdots A_k \) and \( Y = \phi(A_1) \cdots \cdots \phi(A_k) \) for any \( A_1, \ldots, A_k \in S_2 \) if (c) holds. So, condition (1.1) follows.

Next, we turn to the necessity. Applying Theorem 1.2 with \( A_{is} = B \) and \( A_{ir} = A \) for all other \( s \neq r \), we conclude that there exist a unitary operator \( U \in B(H) \) and a scalar \( \mu \in \mathbb{F} \) with \( \mu^m = 1 \) such that one of the following holds:

(a) \( A \mapsto \mu U^*AU \) for all \( A \in \mathbf{V} \).
(b) \( r = (m + 1)/2 \) and \( \phi \) has the form \( A \mapsto \mu U^*A^tU \).
(c) \( \mathbf{V} = S_2 \) and \( \phi \) has the form \( A \mapsto \mu U^*A^tU \).

It remains to prove \( (i_{r+1}, \ldots, i_m, i_1, \ldots, i_{r-1}) = (i_{r-1}, \ldots, i_1, i_m, \ldots, i_{r+1}) \) if (b) or (c) holds.

Evidently, the result holds for \( k = 2 \) as we must have \( i_1 = \cdots = i_{r-1} = i_{r+1} = \cdots = i_m \) in this case. Now we assume that \( k \geq 3 \). Then we have
\[
W(A_{i_1} \cdots A_{i_m}) = W(\phi(A_{i_1}) \cdots \phi(A_{i_m})) = W(U^*A_{i_1}^t \cdots A_{i_m}^t U)
\]
\[ = W(A_{i_1}^t \cdots A_{i_m}^t) = W(A_{i_m} \cdots A_{i_1}). \]
By taking $A_{ir} = R$, where $R$ is a Hermitian rank one idempotent, and considering the foci of the elliptical disks for the above numerical ranges, we conclude that
\[
\text{tr}(A_{i+1} \cdots A_i A_{i+1} \cdots A_{i-1} R) = \text{tr}(A_{i+1} \cdots A_i R A_{i+1} \cdots A_{i-1}) = \text{tr}(A_{i+1} \cdots A_i R A_{i+1} \cdots A_i) = \text{tr}(A_{i+1} \cdots A_{i-1} A_i A_{i+1} \cdots A_{i-1} R).
\]
Since $R$ can be an arbitrary Hermitian rank one idempotent, by the fact that $X$ and $Y$ are equal if $\text{tr}(XR) = \text{tr}(YR)$ for all Hermitian rank one idempotent $R$, we deduce that
\[
A_{i+1} \cdots A_i A_{i+1} \cdots A_{i-1} = A_{i+1} \cdots A_i A_{i+1} \cdots A_{i-1}
\]
for all choices of $A_1, \ldots, A_k$.

We now use a similar argument to the one in the proof of [1, Theorem 2.1]. We give the details for the sake of completeness. For simplicity, we rename $(i_{r+1}, \ldots, i_m, i_1, \ldots, i_{r-1})$ as $(j_1, \ldots, j_{m-1})$ and we have to show that $(j_1, \ldots, j_{m-1}) = (j_{m-1}, \ldots, j_1)$. Suppose $(j_1, \ldots, j_{m-1}) \neq (j_{m-1}, \ldots, j_1)$. Let $1 \leq p \leq m/2$ be the smallest integer such that $j_p \neq j_{m-p}$. For any $\lambda > 0$, let $D = \text{diag}(\lambda, 1)$ and $S$ be some $2 \times 2$ symmetric matrix with positive entries. Fix a two-dimensional subspace $H_1$ in $H$ and take $A_{j_p} = D \oplus I_{H_1^\perp}$ and $A_{j_l} = S \oplus I_{H_1^\perp}$ for all other $j_l \neq j_p$ on $H = H_1 \oplus H_1^\perp$. Then
\[
A_{j_p} \cdots A_{j_{m-p}} = (D^{d_1} S^{s_1} D^{d_2} S^{s_2} \cdots D^{d_q} S^{s_q}) \oplus I_{H_1^\perp}
\]
for positive integers $d_i, s_i$. Note that
\[
D^{d_i} S^{s_i} = \begin{pmatrix} \lambda^{d_i} e_i & \lambda^{d_i} f_i \\ g_i & h_i \end{pmatrix} \quad \text{and} \quad S^{s_i} D^{d_i} = \begin{pmatrix} \lambda^{d_i} e_i & f_i \\ \lambda^{d_i} g_i & h_i \end{pmatrix}
\]
for some positive numbers $e_i, f_i, g_i, h_i$. We check that the $(1, 2)$ entry of $D^{d_1} S^{s_1} \cdots D^{d_q} S^{s_q}$ is a polynomial of degree $d_1 + \cdots + d_q$ in $\lambda$, while the $(1, 2)$ entry of $S^{s_q} D^{d_q} \cdots S^{s_1} D^{d_1}$ is a polynomial of degree $d_2 + \cdots + d_q$. So, there is $\lambda > 0$ such that
\[
A_{j_p} \cdots A_{j_{m-p}} = (D^{d_1} S^{s_1} \cdots D^{d_q} S^{s_q}) \oplus I_{H_1^\perp} \neq (S^{s_q} D^{d_q} \cdots S^{s_1} D^{d_1}) \oplus I_{H_1^\perp} = A_{j_{m-p}} \cdots A_{j_p}.
\]
It follows that $A_{j_1} \cdots A_{j_{m-1}} \neq A_{j_{m-1}} \cdots A_{j_1}$, which is a contradiction. Hence, $(j_1, \ldots, j_{m-1}) = (j_{m-1}, \ldots, j_1)$ as asserted. $\blacksquare$

**Acknowledgements.** We thank Professor Jinchuan Hou for sending us the preprint [5], and his inspiring presentation of the paper at the workshop on preserver problems at the University of Hong Kong in 2004. Thanks are also due to Dr. Jor-Ting Chan for his effort in organizing the two workshops on preserver problems at the University of Hong Kong in 2004 and 2005 that facilitated this research.
References


Department of Mathematics
College of William and Mary
Williamsburg, VA 23185, U.S.A.
E-mail: cki@math.wm.edu

Department of Mathematics
University of Hong Kong
Hong Kong
E-mail: NungSingSze@graduate.hku.hk

Received July 18, 2005
Revised version December 2, 2005