Proyecciones Journal of Mathematics
Vol. 30, N^o 1, pp. 77-90, May 2011.
Universidad Católica del Norte
Antofagasta - Chile
DOI: 10.4067/S0716-09172011000100008

Numerical range of a pair of strictly upper triangular matrices

Wen Yan
Tuskegee University, U. S. A.

Received : January 2011. Accepted : March 2011

Abstract

Given two strictly upper triangular matrices $X, Y \in C_{m\times m}$, we study the range $W_Y(X) = \{trnXn^{-1}Y^* : n \in N\}$, where N is the group of unit upper triangular matrices in $C_{m\times m}$. We prove that it is either a point or the whole complex plane. We characterize when it is a point.

We also obtain some convexity result for a similar range, where N is replaced by any ball of C^k $(k = m(m-1)/2)$ embedded in N, $m \leq 4$.

2000 Mathematics Subject Classification : *Primary 15A60.*

Key Words and Phrases : *Numerical range, unit upper triangular matrices, strictly upper triangular matrices.*

1. Introduction

Let $C_{m\times m}$ be the space of all $m \times m$ complex matrices. The classical numerical range of $A \in C_{m\times m}$ is defined as

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

The celebrated Toeplitz-Hausdorff theorem [9] asserts that $W(A)$ is a compact convex subset of C. There are numerous generalizations [5, 1, 4, 8, 7, 10, 11, 12, 14] and our references are far from complete. One important view is to deem the numerical range as the image of an orbit under the linear functional [2] determined by A, that is,

$$W(A) = \{\operatorname{tr} Axx^* : x \in C^m,\ x^*x = 1\}.$$

The set

$$\{xx^* : x \in C^m,\ x^*x = 1\} = O(E_{11}) := \{UE_{11}U^* : U \in U(m)\}$$

is viewed as an orbit of the matrix $E_{11} := \operatorname{diag}(1, 0, \ldots, 0)$ under the conjugation action of $U(m)$, where $U(m)$ denotes the unitary group in $C_{m\times m}$. In general, if $C \in C_{m\times m}$, then denote by

$$O(C) := \{UCU^* : U \in U(m)\}$$

the orbit of C under the conjugation action of $U(n)$. The C-numerical range of A [13, 3] is defined to be the set

$$W_C(A) := \{\operatorname{tr} AY : Y \in O(C)\}.$$

If $C = \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)$, ($k$ 1's), it becomes Halmos's k-numerical range [7] of A

$$W_k(A) = \{\sum_{j=1}^{k} x_j^*Ax_j : x_1, \ldots, x_k \in C^m \quad \text{are orthonormal } \}.$$

If $C = \operatorname{diag}(c_1, \ldots, c_m)$ (c's are real), the C-numerical range of A becomes Westwick's c-numerical range [14] of A

$$W_c(A) = \{\sum_{j=1}^{m} c_jx_j^*Ax_j : x_1, \ldots, x_m \in C^m \quad \text{are orthonormal } \}.$$

Westwick's theorem [14] asserts that the c-numerical range of A is convex. The orbital point of view leads to several generalizations of the numerical range. Moreover the convexity result has been successfully extended in the context of compact Lie groups [11] and most real classical semisimple Lie algebras [8, 4, 12]. Usually the groups involved in the relevant orbital generalizations are compact (for example $U(m)$ is compact in the setting of the c-numerical range).

In this note we consider the group of $m\times m$ unit upper triangular matrices which is non-semisimple and noncompact. By a unit upper triangular matrix, we mean an upper triangular with diagonal entries all ones. Let N be the group of unit upper triangular matrices in $C_{m\times m}$. It is a unipotent (noncompact) Lie group whose Lie algebra n is the set of strictly upper triangular matrices in $C_{m\times m}$. Given $X \in n$, denote by

$$O(X) := \{nXn^{-1} : n \in N\} \subset n$$

the orbit of X under the conjugation action of the group N. Let $X, Y \in n$. The *numerical range* of the pair (X, Y) is defined as

$$W_Y(X) := \{\operatorname{tr} nXn^{-1}Y^* : n \in N\}.$$

It may be interpreted as the image of the orbit $O(X)$ under the linear functional determined by Y. In Section 2 we prove that $W_Y(X)$ is either a point (not necessarily the origin) or C. In Section 3, given $r > 0$, $c_{ij} \in C$, $1 \le i < j \le m$, we consider a compact subset of N:

$$N_1 := \{n := (n_{ij}) \in N : \sum_{1\le i<j\le m} |n_{ij} - c_{ij}|^2 = r^2\}.$$

In other words, the ball of radius r (with respect to the 2-norm) centered at c of C^s is embedded as $N_1 \subset N$, where $s = m(m-1)/2$. We consider the restricted range:

$$W_Y^1(X) := \{\operatorname{tr} nXn^{-1}Y^* : n \in N_1\}.$$

When $m = 2, 3, 4$ we prove that $W_Y^1(X)$ is a convex set. When $m > 4$ convexity of $W_Y^1(X)$ is unknown.

2. The shape of $W_Y(X)$

Theorem 1. Let $X, Y \in n$. When $m = 2$, $W_Y(X) = \{\operatorname{tr} nXn^{-1}Y^* : n \in N\}$ is a singleton set $\{x\bar{y}\}$ if

$$X = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix}.$$

When $m > 2$, $W_Y(X)$ is either a point or the whole complex plane C. If $W_Y(X)$ is a point, then the point is $\sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell}$. More precisely, $W_Y(X) = C$ if and only if one of the following is true.

(i) $x_{jk}\bar{y}_{i\ell} \neq 0$ for some i, j, k and ℓ such that

  (a) $1 \le i < j < k < \ell \le m$, or
  (b) $1 \le i = j < k < \ell - 1 \le m - 1$, or
  (c) $2 \le i + 1 < j < k = \ell \le m$.

(ii) $x_{jk}\bar{y}_{i\ell} = 0$ for all $1 \le i < j < k < \ell \le m$, but there exist i, ℓ such that $i < \ell - 1$, $x_{i,\ell-1}\bar{y}_{i\ell} \neq 0$ and $x_{i,\ell-1}\bar{y}_{i\ell} \neq x_{\ell t}\bar{y}_{\ell-1,t}$ for all $\ell < t \le m$, or $x_{i+1,\ell}\bar{y}_{i\ell} \neq 0$ and $x_{i+1,\ell}\bar{y}_{i\ell} \neq x_{ti}\bar{y}_{t,i+1}$ for all $1 \le t < i$.

Proof. The case $m = 2$ is trivial. Suppose $m > 2$. Let $n = (n_{ij}) \in N$. Clearly $M := n^{-1}$ is upper triangular. Because of the upper triangular form of n, X, Y, M, we have

$$\operatorname{tr} nXn^{-1}Y^* = \sum_{1\le i\le j<k\le \ell\le m} n_{ij}x_{jk}M_{k\ell}\bar{y}_{i\ell}.$$

Notice that the (k, ℓ) entry of M is

$$M_{kl} = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{if } k > 1 \\ (-1)^{k+l} \det \begin{pmatrix} n_{k,k+1} & n_{k,k+2} & n_{k,k+3} & \cdots & n_{k,l-1} & n_{k,l} \\ 1 & n_{k+1,k+2} & n_{k+1,k+3} & \cdots & n_{k+1,l-1} & n_{k+1,l} \\ 0 & 1 & n_{k+2,k+3} & \cdots & n_{k+2,l-1} & n_{k+2l} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 & n_{l-1,l} \end{pmatrix} & \\ & \text{if } k < l \end{cases}$$

Notice that $M_{k\ell}$ is a polynomial in the variables n_{st}, $k \le s < t \le \ell$. Moreover the exponent of each n_{st} in the expression (2.1) of $M_{k\ell}$ is either 0 or 1.

Evidently $\operatorname{tr} nXn^{-1}Y^*$ is a polynomial of n_{ij}, $1 \le i < j \le m$. Since n_{ij} does not appear in the polynomial $M_{k\ell}$ for $i \le j < k \le \ell$, the exponent of any n_{ij} $(i < j)$ in $\operatorname{tr} nXn^{-1}Y^*$ is either 0 or 1. We use n_1, ..., n_r to denote those n_{ij} $(i < j)$ which appear in the polynomial $\operatorname{tr} nXn^{-1}Y^*$. Let

$$f_0(n_1, n_2, ..., n_r) := \operatorname{tr} nXn^{-1}Y^*.$$

1. If f_0 is a constant polynomial. Then $\{\operatorname{tr} nXn^{-1}Y^* : n \in N\}$ is a point.

2. Otherwise, we can rewrite f_0 as

$$f_0(n_1, ..., n_r) = n_1 f_1(n_2, ..., n_r) + f_2(n_2, ..., n_r),$$

where f_1 is either a nonconstant polynomial in $n_2, n_3 ..., n_r$ or a nonzero constant number c. In either case we can choose complex numbers $c_2, ..., c_r$ for $n_2, ..., n_r$ such that $f_1(c_2, ..., c_r) \neq 0$. By the fundamental theorem of algebra $\{f_0(n_1, c_2, ..., c_r) : n_1 \in C\} = C$. Hence $W_Y(X) = C$.

So $W_Y(X)$ is either a point or C.

We are going to show that $W_Y(X) = C$ if either (i) or (ii) holds. Suppose (i)(a) is true, that is, there exists $x_{j_0k_0}\bar{y}_{i_0\ell_0} \neq 0$ for some $1 \leq i_0 < j_0 < k_0 < \ell_0 \leq m$. Define

$$n(s) := (n_{ij}) = I_m + sE_{i_0,j_0} + sE_{k_0,\ell_0} \in N, \quad s \in C,$$

and E_{ij} is the matrix with 1 as the (i, j) entry and zeros elsewhere. So $M := n(s)^{-1} = I_m - sE_{i_0,j_0} - sE_{k_0,\ell_0}$. Then

$$f(s) := \operatorname{tr} n(s)Xn(s)^{-1}Y^* = \sum_{1\leq i\leq j<k\leq \ell\leq m} n_{ij}x_{jk}M_{kl}\bar{y}_{i\ell}$$

is a quadratic polynomial in s, and the leading term of $f(s)$ is $n_{i_0j_0}x_{j_0k_0}M_{k_0\ell_0}\bar{y}_{i_0\ell_0} = -x_{j_0k_0}\bar{y}_{i_0\ell_0}s^2$. Therefore

$$C = \{f(s) : s \in C\} \subset W_Y(X) \subset C.$$

We now insert a lemma.

Lemma 2. Suppose (i)(a) is not true.

1. If there exist $1 \leq i_0 < k_0 < \ell_0 \leq m$ such that $x_{i_0k_0}\bar{y}_{i_0\ell_0} \neq 0$, then $x_{ik_0}\bar{y}_{i\ell_0} = 0$ for all $i \neq i_0$.

2. If there exist $1 \leq i_0 < j_0 < \ell_0 \leq m$ such that $x_{j_0\ell_0}\bar{y}_{i_0\ell_0} \neq 0$, then $x_{j_0\ell}\bar{y}_{i_0\ell} = 0$ for all $\ell \neq \ell_0$.

Proof. (1) If there exists $i_1 \neq i_0$ such that $1 \leq i_1 < k_0$ and $x_{i_1k_0}\bar{y}_{i_1\ell_0} \neq 0$, then we have the following two cases.

(a) if $i_0 < i_1$, then $x_{i_1k_0}\bar{y}_{i_0\ell_0} \neq 0$ with $1 \leq i_0 < i_1 < k_0 < \ell_0$,

(b) if $i_0 > i_1$, then $x_{i_0k_0}\bar{y}_{i_1\ell_0} \neq 0$ with $1 \leq i_1 < i_0 < k_0 < \ell_0$.

Both are under case (i)(a). The proof of (2) is analogous. □

Suppose (i)(b) is true. Let i_0, j_0, k_0 and ℓ_0 be such that $1 \leq i_0 = j_0 < k_0 < \ell_0 - 1 \leq m - 1$ and $x_{j_0k_0}\bar{y}_{i_0\ell_0} = x_{i_0k_0}\bar{y}_{i_0\ell_0} \neq 0$. Let $n(s) := (n_{ij}) \in N$ be defined as follows:

(2.1) $\quad n_{k,k+1} = s, \quad k = k_0, \ldots, \ell_0 - 1, \quad n_{ij} = 0$ for all other $i < j$.

Set

$$g(s) := \operatorname{tr} n(s)Xn(s)^{-1}Y^* = \sum_{1 \leq i \leq j < k \leq \ell \leq m} n_{ij}x_{jk}M_{k\ell}\bar{y}_{i\ell},$$

where $M := n(s)^{-1}$ and

$$M_{kl} = \begin{cases} (-1)^{k+l} \ s^{l-k} & \text{if } \ k_0 \leq k \leq l_0, \\ 1 \text{ if } k = l, \\ 0 \text{ for all other } k, l. \end{cases}$$

Notice that $\deg g(s) = \ell_0 - k_0$. Only $M_{k_0\ell_0} = (-1)^{\ell_0+k_0}s^{\ell_0-k_0}$ of M has the highest degree. Moreover n_{ij} in $n_{ij}x_{jk_0}s^{\ell_0-k_0-1}\bar{y}_{i,\ell_0-1} = n_{ij}x_{jk_0}M_{k_0,\ell_0-1}\bar{y}_{i,\ell_0-1}$ $(i \leq j < k_0 < \ell_0 - 1)$ or $n_{ij}x_{j,k_0+1}s^{\ell_0-k_0-1}\bar{y}_{i\ell_0} = n_{ij}x_{j,k_0+1}M_{k_0+1,\ell_0}\bar{y}_{i\ell_0}$ $(i \leq j < k_0 + 1 < \ell_0)$ cannot be s, by (2.1). So the leading term of $g(s)$ is

$$(-1)^{k_0+\ell_0}\left[\sum_{1 \leq i \leq j < k_0} n_{ij}x_{jk_0}\bar{y}_{i\ell_0}\right]s^{\ell_0-k_0} = (-1)^{k_0+\ell_0}\left[\sum_{1 \leq i < k_0} x_{ik_0}\bar{y}_{i\ell_0}\right]s^{\ell_0-k_0}.$$
(2.2)

If (i)(a) is not true, then by Lemma 2(1), (2.2) becomes

$$(-1)^{k_0+\ell_0}x_{i_0k_0}\bar{y}_{i_0\ell_0}s^{\ell_0-k_0}.$$

Therefore $\{g(s) : s \in C\} = C$ and hence $W_Y(X) = C$.

If (i)(c) is true, then there exist $2 \leq i_0 + 1 < j_0 < \ell_0 \leq m$ such that $x_{j_0\ell_0}\bar{y}_{i_0\ell_0} \neq 0$. Let $n(s) := (n_{ij}) = I_m + sE_{i_0,j_0} \in N$, $s \in C$. Then $M := n(s)^{-1} = I_m - sE_{i_0,j_0}$. We may assume that (i)(a) is not true. Then $x_{j_0\ell}\bar{y}_{i_0\ell} = 0$ for all $\ell \neq \ell_0$ by Lemma 2(2). Thus the only possible nonconstant term in the polynomial

$$h(s) := \operatorname{tr} n(s)Xn(s)^{-1}Y^* = \sum_{1 \leq i \leq j < k \leq \ell \leq m} n_{ij}x_{jk}M_{k\ell}\bar{y}_{i\ell}$$

is

$$\sum_{j_0<\ell\le m} n_{i_0j_0}x_{j_0\ell}M_{\ell\ell}\bar{y}_{i_0\ell}+\sum_{1\le i<i_0} n_{ii}x_{ii_0}M_{i_0j_0}\bar{y}_{ij_0}=x_{j_0\ell_0}\bar{y}_{i_0\ell_0}s-\sum_{1\le i<i_0} x_{ii_0}\bar{y}_{ij_0}s.$$

Since $i_0+1<j_0$, if there exists $x_{ii_0}\bar{y}_{ij_0}\neq 0$ for some $i<i_0$, then this becomes case (i)(b) and $W_Y(X)=C$. Otherwise $x_{ii_0}\bar{y}_{ij_0}=0$ for all $1\le i<i_0$, then the leading term of $h(s)$ is $x_{j_0\ell_0}\bar{y}_{i_0\ell_0}s$ with nonzero coefficient. Therefore $\{h(s): s\in C\}=C$ and hence $W_Y(X)=C$.

Suppose condition (ii) holds. Then there exist $1\le i_0<\ell_0-1\le m-1$ such that (1) $x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0}\neq 0$ and $x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0}\neq x_{\ell_0t}\bar{y}_{\ell_0-1,t}$ for all $t>\ell_0$, or (2) $x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0}\neq 0$ and $x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0}\neq x_{ti_0}\bar{y}_{t,i_0+1}$ for all $1\le t<i_0$. We may assume that condition (i) does not hold.

(1) Define $n(s):=(n_{ij})=I_m+sE_{\ell_0-1,\ell_0}\in N$, $s\in C$. So $M:=n(s)^{-1}=I_m-sE_{\ell_0-1,\ell_0}$. Let

$$u(s):=\operatorname{tr} n(s)Xn(s)^{-1}Y^*=\sum_{1\le i\le j<k\le\ell\le m} n_{ij}x_{jk}M_{k\ell}\bar{y}_{i\ell}.$$

Then

$$\begin{aligned}
u(s) &= \sum_{1\le i\le\ell_0-1} n_{ii}x_{i,\ell_0-1}M_{\ell_0-1,\ell_0}\bar{y}_{i\ell_0}+\sum_{\ell_0<\ell\le m} n_{\ell_0-1,\ell_0}x_{\ell_0\ell}M_{\ell\ell}\bar{y}_{\ell_0-1,\ell}\\
&\quad+\sum_{1\le i<\ell\le m} n_{ii}x_{i\ell}M_{\ell\ell}\bar{y}_{i\ell}\\
&= -x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0}s+\sum_{\ell_0<t\le m} x_{\ell_0t}\bar{y}_{\ell_0-1,t}s+\sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell}.
\end{aligned}$$

The last equality is due to Lemma 2(1) which implies $x_{i,\ell_0-1}\bar{y}_{i\ell_0}=0$ for all $i\neq i_0$. Therefore, if $x_{\ell_0t}\bar{y}_{\ell_0-1,t}=0$ for all $t>\ell_0$, then

$$u(s)=-x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0}s+\sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell}.$$

Otherwise by Lemma 2(2) there is only one t, say t_0, such that $x_{\ell_0t}\bar{y}_{\ell_0-1,t}\neq 0$. Hence

$$u(s)=(x_{\ell_0t_0}\bar{y}_{\ell_0-1,t_0}-x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0})s+\sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell},$$

where $x_{i_0,\ell_0-1}\bar{y}_{i_0\ell_0}\neq x_{\ell_0t_0}\bar{y}_{\ell_0-1,t_0}$ by (ii). In both cases, the polynomial $u(s)$ is linear. Thus $W_Y(X)=C$.

(2) If there exist $x_{i_0+1\ell_0}\bar{y}_{i_0\ell_0} \neq 0$ with $1 \leq i_0 < \ell_0 - 1 \leq m-1$ and $x_{i_0+1\ell_0}\bar{y}_{i_0\ell_0} \neq x_{ti_0}\bar{y}_{ti_0+1}$ for all $1 \leq t < i_0$.

Define $n(s) := (n_{ij}) = I_m + sE_{i_0,i_0+1} \in N$, $s \in C$. Then $M := n(s)^{-1} = I_m - sE_{i_0,i_0+1}$. Let

$$v(s) := \operatorname{tr} n(s)Xn(s)^{-1}Y^*.$$

By Lemma 2(2), $x_{i_0+1,\ell}\bar{y}_{i_0\ell} = 0$ for all $\ell \neq \ell_0$. Thus

$$\begin{aligned} v(s) &= \textstyle\sum_{i_0+1<\ell\leq m} n_{i_0,i_0+1}x_{i_0+1,\ell}M_{\ell\ell}\bar{y}_{i_0\ell} + \sum_{1\leq i<i_0} n_{ii}x_{ii_0}M_{i_0,i_0+1}\bar{y}_{i,i_0+1} \\ &\quad + \textstyle\sum_{1\leq i\leq\ell\leq m} x_{i\ell}M_{\ell\ell}\bar{y}_{i\ell} \\ &= x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0}s - \textstyle\sum_{1\leq t<i_0} x_{ti_0}\bar{y}_{t,i_0+1}s + \sum_{1\leq i\leq\ell\leq m} x_{i\ell}\bar{y}_{i\ell}. \end{aligned}$$

Therefore, if $x_{ti_0}\bar{y}_{t,i_0+1} = 0$ for all $t < i_0$, then

$$v(s) = x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0}s + \sum_{1\leq i\leq\ell\leq m} x_{i\ell}\bar{y}_{i\ell}.$$

Otherwise by Lemma 2(1), there is only one t, denoted by t_0, such that $x_{ti_0}\bar{y}_{t,i_0+1} \neq 0$. Hence

$$v(s) = (x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0} - x_{t_0i_0}\bar{y}_{t_0,i_0+1})s + \sum_{1\leq i\leq\ell\leq m} x_{i\ell}\bar{y}_{i\ell},$$

where $x_{i_0+1,\ell_0}\bar{y}_{i_0\ell_0} \neq x_{t_0i_0}\bar{y}_{t_0,i_0+1}$ by (ii). In both cases, the polynomial $v(s)$ is linear. Therefore $W_Y(X) = C$.

So either (i) or (ii) implies $W_Y(X) = C$.

Suppose (i) and(ii) are not true. Then the only nonzero terms among $x_{jk}\bar{y}_{i\ell}$, $1 \leq i \leq j < k \leq \ell \leq m$, are (1) $x_{i,\ell-1}\bar{y}_{i\ell}$ with $x_{i,\ell-1}\bar{y}_{i\ell} = x_{\ell t}\bar{y}_{\ell-1,t} \neq 0$ for some $t > \ell$, and (2) $x_{i+1,\ell}\bar{y}_{i\ell}$ with $x_{i+1,\ell}\bar{y}_{i\ell} = x_{ti}\bar{y}_{t,i+1} \neq 0$ for some $t < i$. Indeed for each case t is unique by Lemma 2. Thus

$$\begin{aligned} &\operatorname{tr} nXn^{-1}Y^* \\ = {} & \sum_{1\leq i\leq j<k\leq\ell\leq m} n_{ij}x_{jk}M_{k\ell}\bar{y}_{i\ell} \\ = {} & \sum_{1\leq i<\ell-1\leq m-1} n_{ii}x_{i,\ell-1}M_{\ell-1,\ell}\bar{y}_{i\ell} + \sum_{1\leq i<\ell-1\leq m-1} n_{i,i+1}x_{i+1,\ell}M_{\ell\ell}\bar{y}_{i\ell} \\ & + \sum_{1\leq i<\ell\leq m} n_{ii}x_{i\ell}M_{\ell\ell}\bar{y}_{i\ell} \qquad \text{(since (i) does not hold)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{1\le i<\ell-1\le m-1} -n_{\ell-1,\ell}x_{i,\ell-1}\bar{y}_{i\ell} + \sum_{1\le i<\ell-1\le m-1} n_{i,i+1}x_{i+1,\ell}\bar{y}_{i\ell} \\
&\quad + \sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell} \qquad \text{(since } M_{\ell-1,\ell} = -n_{\ell-1,\ell}\text{)} \\
&= \sum_{1\le i<\ell-1<t-1\le m-1} [-n_{\ell-1,\ell}x_{i,\ell-1}\bar{y}_{i\ell} + n_{\ell-1,\ell}x_{\ell t}\bar{y}_{\ell-1,t}] \\
&\quad + \sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell} \qquad \text{(since (ii) does not hold)} \\
&= \sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell}.
\end{aligned}$$

Therefore $W_Y(X) = \{\sum_{1\le i<\ell\le m} x_{i\ell}\bar{y}_{i\ell}\}$.

□

3. Convexity of $W_Y^1(X)$

Given $c_{ij} \in C$, $1 \le i < j \le m$, $r > 0$, let

$$N_1 := \{n := (n_{ij}) \in N : \sum_{1\le i<j\le m} |n_{ij} - c_{ij}|^2 = r^2\} \subset N.$$

In other words, N_1 is the embedding in N of the ball in C^s ($s = m(m-1)/2$) of radius r centered at $c = (c_{12}, \dots, c_{1n}, c_{23}, \dots, c_{2n}, \dots, c_{n-1,n})^T$. We define the range:

$$W_Y^1(X) := \{\operatorname{tr} nXn^{-1}Y^* : n \in N_1\} \subset W_Y(X).$$

Theorem 1. 1. When $m = 2$, $W_Y^1(X)$ is the singleton set $\{x\bar{y}\}$ if

$$X = \begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & y \\ 0 & 0 \end{pmatrix}.$$

2. When $m = 3$, for any $r > 0$, $c_1 := c_{12}, c_2 := c_{13}, c_3 := c_{23} \in C$, $1 \le i < j \le m$, $W_Y^1(X)$ is the circular disc in C centered at $\sum_{i=1}^3 x_i\bar{y}_i + x_3\bar{y}_2c_1 - x_1\bar{y}_2c_3$ with radius $r|y_2|\sqrt{|x_1|^2 + |x_3|^2}$, if

$$X = \begin{pmatrix} 0 & x_1 & x_2 \\ 0 & 0 & x_3 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & y_1 & y_2 \\ 0 & 0 & y_3 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbf{n}$$

Proof. The first statement is trivial. When $m = 3$, let

$$n = \begin{pmatrix} 1 & n_1 & n_2 \\ 0 & 1 & n_3 \\ 0 & 0 & 1 \end{pmatrix} \in N, \quad \text{such that} \quad \sum_{j=1}^{3} |n_j - c_j|^2 = r^2.$$

By direct computation

$$\begin{aligned} \operatorname{tr} nXn^{-1}Y^* &= \sum_{i=1}^{3} x_i\bar{y}_i + x_3\bar{y}_2 n_1 - x_1\bar{y}_2 n_3 \\ &= \sum_{i=1}^{3} x_i\bar{y}_i + x_3\bar{y}_2 c_1 - x_1\bar{y}_2 c_3 + x_3\bar{y}_2(n_1 - c_1) - x_1\bar{y}_2(n_3 - c_3). \end{aligned}$$

The locus of $x_3\bar{y}_2(n_1 - c_1) - x_1\bar{y}_2(n_3 - c_3)$, as n runs through N_1, is

$$L = \{r(|x_3\bar{y}_2|e^{i\xi_1}\cos\theta + |x_1\bar{y}_2|e^{i\xi_2}\sin\theta) : \ \theta, \xi_1, \xi_2 \in [0, \pi]\}.$$

It is the circular disc centered at the origin with radius $r\sqrt{|x_3\bar{y}_2|^2 + |x_1\bar{y}_2|^2}$.
□

To establish the 4×4 case, we need the following result of Gutiérrez and Medrano [6] which generalizes the Toeplitz-Hausdorff's theorem.

Theorem 2. [6] Let $A \in C_{m\times m}$ with $m \geq 2$. Given $\alpha, \beta, c \in C^m$, and $r > 0$. The set

$$\{z^*Az + \alpha^*z + z^*\beta : z \in C^m, \ (z - c)^*(z - c) = r^2\}$$

is a compact convex set in C.

Theorem 3. When $m = 4$, for any $r > 0$, $c_{ij} \in C$, $1 \leq i < j \leq m$, $W^1_Y(X)$ is a compact convex subset of C. In general it is not necessary a circular disk.

Proof. Let

$$n = \begin{pmatrix} 1 & n_1 & n_2 & n_3 \\ 0 & 1 & n_4 & n_5 \\ 0 & 0 & 1 & n_6 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in N_1.$$

Let

$$X = \begin{pmatrix} 0 & x_1 & x_2 & x_3 \\ 0 & 0 & x_4 & x_5 \\ 0 & 0 & 0 & x_6 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 & y_1 & y_2 & y_3 \\ 0 & 0 & y_4 & y_5 \\ 0 & 0 & 0 & y_6 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbf{n}$$

By direct computation

$$n^{-1} = \begin{pmatrix} 1 & -n_1 & n_1n_4 - n_2 & -n_1n_4n_6 + n_1n_5 + n_2n_6 - n_3 \\ 0 & 1 & -n_4 & n_4n_6 - n_5 \\ 0 & 0 & 1 & -n_6 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Hence
$nXn^{-}1$

$$= \begin{pmatrix} 0 & x_1 & x_2 + x_4n_1 - x_1n_4 & x_3 - x_4n_1n_6 + x_1n_4n_6 + x_5n_1 + x_6n_2 - x_1n_5 - x_2n_6 \\ 0 & 0 & x_4 & x_5 + x_6n_4 - x_4n_6 \\ 0 & 0 & 0 & x_6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} \operatorname{tr} nXn^{-1}Y^* &= \sum_{i=1}^{6} x_i\bar{y}_i - x_4\bar{y}_3n_1n_6 + x_1\bar{y}_3n_4n_6 + (x_4\bar{y}_2 + x_5\bar{y}_3)n_1 + x_6\bar{y}_3n_2 \\ &\quad +(x_6\bar{y}_5 - x_1\bar{y}_2)n_4 - x_1\bar{y}_3n_5 - (x_2\bar{y}_3 + x_4\bar{y}_5)n_6. \end{aligned} \tag{3.1}$$

Set

$$z := (n_1, n_2, n_3, n_4, n_5, \bar{n}_6)^*.$$

Set $A := (a_{ij})$, where $a_{16} = -x_4\bar{y}_3$, $a_{46} = x_1\bar{y}_3$, and $a_{ij} = 0$ otherwise.
Set

$$\alpha := (0, 0, 0, 0, 0, -(x_2\bar{y}_3 + x_4\bar{y}_5))^*,$$

and

$$\beta := (x_4\bar{y}_2 + x_5\bar{y}_3, x_6\bar{y}_3, 0, x_6\bar{y}_5 - x_1\bar{y}_2, -x_1\bar{y}_3, 0)^T.$$

Note that $\operatorname{tr} nXn^{-1}Y^* = z^*Az + \alpha^*z + z^*\beta$. Now

$$W_Y^1(X) = \{z^*Az + \alpha^*z + z^*\beta : z \in C^6,\ (z - c)^*(z - c) = r^2\}.$$

By Theorem 2, it is convex.

Choose 4×4 strictly upper triangular matrices X, Y such that $x_1 = x_6 = 0$ and $-x_4\bar{y}_3 = x_4\bar{y}_2 + x_5\bar{y}_3 = -(x_2\bar{y}_3 + x_4\bar{y}_5) = 1$. Set $c = 0$. So

$$W_Y^1(X) = \sum_{i=1}^{6} x_i\bar{y}_i + S,$$

where $S = \{\xi_1 + \xi_2 + \xi_1\xi_2 : \xi_1, \xi_2 \in C,\ |\xi_1|^2 + |\xi_2|^2 \le 1\}$. The set S is symmetric about the x-axis. By direct computation $S \cap R = [-1, \sqrt{2} + \frac{1}{2}]$. The set S is not a circular disk by considering the point $\sqrt{2}i - \frac{1}{2} \in S$ given by $\xi_1 = \xi_2 = i/\sqrt{2}$. $\square$

If one replaces the expression in Theorem 2 by the form $z^T Az + \alpha^T z + z^T\beta$ (clearly (3.1) is of this form), we may not have a convex set.

Example 4. Let $f(u) = u^2 + 2u + 1$, $u \in C$. If

$$A = \operatorname{diag}(1, 0, \ldots, 0) \in C_{m\times m}, \quad \alpha = (2, 0, \ldots, 0)^T, \beta = (0, \ldots, 0)^T \in C^m,$$

the set $W := \{z^T Az + \alpha^T z + z^T\beta + 1 : z \in C^m, z^*z = 1\} = \{f(u) : u \in C,\ u^*u = 1\}$ is not convex.

Proof. Let $u = (\cos\theta + i\sin\theta)$, and $-\pi \le \theta < \pi$. Then the elements of W are of the form

$$f(u) = \cos 2\theta + 2\cos\theta + 1 + i(\sin 2\theta + 2\sin\theta).$$

Clearly W is symmetric about the x-axis. By choosing $\theta = -2\pi/3$ and $2\pi/3$ respectively, we have $P_1 = -1/2 + i\sqrt{3}/2$, $P_2 = -1/2 - i\sqrt{3}/2 \in W$. The midpoint $-1/2$ of P_1 and P_2 is not contained in W. Therefore W is not convex. $\square$

References

[1] Y. H. Au-Yeung and N. K. Tsing, Some theorems on the generalized numerical range, Linear and Multilinear Algebra, **15**, pp. 3–11, (1984).

[2] C. Davis, The Toeplitz-Hausdorff theorem explained, Canad. Math. Bull., **14**, pp. 245–246, (1971).

[3] W. S. Cheung and N. K. Tsing, The C-numerical range of matrices is star-shaped, Linear and Multilinear Algebra, **41**, pp. 245–250, (1996).

[4] D. Z. Djokovic and T. Y. Tam, Some questions about semisimple Lie groups originating in matrix theory, Bull. Canad. Math. Soc., **46**, pp. 332–343, (2003).

[5] K. R. Gustafson and D. K. M. Rao, *Numerical Range: The Field of Values of Linear Operators and Matrices*, Universitext, Springer-Verlag, New York, (1997).

[6] V. G. Gutiérrez and S. L. de Medrano, An extension of the Toeplitz-Hausdorff theorem, Bol. Soc. Mat. Mexicana (3), **9**, pp. 273–278, (2003).

[7] P. R. Halmos, *A Hilbert Space Problem Book*, Springer-Verlag, New York, (1978).

[8] C. K. Li and T. Y. Tam, Numerical ranges arising from simple Lie algebras, J. Canad. Math. Soc., **52**, pp. 141–171, (2000),

[9] R. Raghavendran, Toeplitz-Hausdorff theorem on numerical ranges, Proc. Amer. Math. Soc., **20**, pp. 284–285, (1969).

[10] T. Y. Tam, An extension of a convexity theorem of the generalized numerical range associated with $SO(2n+1)$, Proc. Amer. Math. Soc., **127**, pp. 35–44, (1999).

[11] T. Y. Tam, Convexity of generalized numerical range associated with a compact Lie group, J. Austral. Math. Soc., **70**, pp. 57–66, (2002).

[12] T. Y. Tam, On the shape of numerical range associated with Lie groups, Taiwanese J. Math., **5**, pp. 497–506, (2001).

[13] N. K. Tsing, On the shape of the generalized numerical ranges, Linear and Multilinear Algebra, **10**, pp. 173–182, (1981).

[14] R. Westwick, A theorem on numerical range, Linear and Multilinear Algebra, **2**, pp. 311–315, (1975).

Wen Yan
Department of Mathematics,
Tuskegee University,
AL 36088
U. S. A.
e-mail : wyliamg@gmail.com