On numerical range of $\mathfrak{sp}(2n, \mathbb{C})$

Consider the complex symplectic Lie algebra \cite[p.3]{3} which is simple for $n \geq 1$:

$$\mathfrak{sp}(2n, \mathbb{C}) := \mathfrak{sp}(2n) \oplus i\mathfrak{sp}(2n) = \begin{cases} \begin{pmatrix} A_1 & A_2 \\ A_3 & -A_1^T \end{pmatrix} : A_1, A_2, A_3 \in \mathbb{C}^{n \times n}, A_2^T = A_2, A_3^T = A_3 \end{cases}.$$

The compact group $K = \text{Sp}(2n, \mathbb{C}) \cap U(2n)$ \cite{3} consists of the matrices

$$\begin{pmatrix} U & -V \\ V & U \end{pmatrix} \in U(2n).$$
It is known that for any $B \in \mathfrak{sp}(2n, \mathbb{C})$, there is $U \in K$ such that $UBU^* \in b \subset \mathfrak{sp}(2n, \mathbb{C})$, where

$$b := \left\{ \begin{pmatrix} A_1 & A_2 \\ 0 & -A_1^T \end{pmatrix}, \ A_1 \in \mathbb{C}_{n \times n} \text{ is upper triangular, } A_2^T = A_2 \right\}$$

(2.1)

is a Borel subalgebra of $\mathfrak{sp}(2n, \mathbb{C})$. The eigenvalues of $A \in \mathfrak{sp}(2n, \mathbb{C})$ occur in pairs but opposite in sign as we can see it from (2.1). By Toeplitz-Hausdorff theorem, the numerical range $W(A)$ of a matrix $A \in \mathfrak{sp}(2n, \mathbb{C})$ is convex. When $n = 1$, $W(A)$ is the elliptical disk with foci $\pm \lambda$ and the minor axis length $\frac{1}{2} |a_2|$, where $\pm \lambda$ are the eigenvalues of $A$.

**Lemma 2.1.** If $A \in \mathfrak{sp}(2n, \mathbb{C})$, then $0 \in W(A)$.

**Proof.** Let $A \in \mathfrak{sp}(2n, \mathbb{C})$. The eigenvalues $\pm \lambda_1, \ldots, \pm \lambda_n$ of $A$ are in the numerical range $W(A)$. Since the set $W(A)$ is convex,

$$0 = \frac{1}{2} \lambda_1 + \frac{1}{2} (-\lambda_1) \in W(A).$$

$\blacksquare$

Now we explore the relationship between the numerical range of $A$ and that of $A_1$, the singular values of $A_2$.

**Theorem 2.2.** Let $A = \begin{pmatrix} A_1 & A_2 \\ 0 & -A_1^T \end{pmatrix} \in b$. If the largest singular value of $A_2$ is $s$, then the circular disk $D_{s/2}$ centered at the origin and with the radius $r = \frac{s}{2}$ is a subset of the numerical range $W(A)$, i.e., $D_{s/2} \subset W(A)$.

**Proof.** By Takagi’s factorization of complex symmetric matrices, $A_2 = U \text{diag} (s_1, \ldots, s_n) U^T$ with $U \in U(n)$ and $s_1 \geq \cdots \geq s_n$ the singular values of $A_2$. Then

$$U^* A_2 U = \text{diag} (s_1, \ldots, s_n).$$

Let $X_1$ be the first column of $U$ and $X = \frac{1}{\sqrt{2}} \begin{pmatrix} X_1 \\ X_1 \end{pmatrix}$. Direct computation shows that

$$W(A) \ni X^* A X = \frac{1}{2} (X_1^*, X_1^T) \begin{pmatrix} A_1 & A_2 \\ 0 & -A_1^T \end{pmatrix} \begin{pmatrix} X_1 \\ X_1 \end{pmatrix}$$

$$= \frac{1}{2} (X_1^* A_1 X_1 - X_1^* A_1^T X_1 + X_1^* A_2 X_1)$$

$$= \frac{1}{2} \left[ X_1^* A_1 X_1 - (X_1^* A_1 X_1)^T + X_1^* A_2 X_1 \right]$$

$$= \frac{1}{2} X_1^* A_2 X_1$$

$$= \frac{1}{2} s_1.$$

Replace $X_1$ with $e^{-i\theta/2} X_1$, $\theta \in [0, 2\pi]$ in above equation, we can see that $\frac{1}{2} s_1 e^{i\theta} \in W(A)$. So the circle $\{ \frac{1}{2} e^{i\theta} s_1 : \theta \in [0, 2\pi] \} \subset W(A)$. Since $W(A)$ is convex, $D_{s/2} \subset W(A)$. $\blacksquare$

In Theorem 2.2, if $n = 1$, $A = \begin{pmatrix} a & b \\ 0 & -a \end{pmatrix}$. The singular value of the $(2, 1)$ block is $|b|$ and the numerical range $W(A)$ is the elliptical disk with foci $a, -a$ and minor axis length $|b|$. Clearly the circular disk $D_{|b|/2}$ is a subset of $W(A)$.

**Lemma 2.3.** Let $A = \begin{pmatrix} A_1 & A_2 \\ 0 & -A_1^T \end{pmatrix} \in \mathfrak{sp}(2n, \mathbb{C})$. The numerical range of $A$ is the union of all elliptical disks with foci $X^* A_1 X \in W(A_1)$ and $-Y^* A_1 Y \in -W(A_1)$ and minor axis length $|X^* A_2 Y|$, where $X, Y \in \mathbb{C}^n$ are unit vectors.
Proof. For any unit vector $Z \in \mathbb{C}^{2n}$, let $Z = \left( e^{i\alpha} \sin \theta X, e^{i\beta} \cos \theta Y \right)$, where $X$, $Y$ are unit vectors in $\mathbb{C}^n$.

\[
Z^* AZ = X^* A_1 X \sin^2 \theta + e^{i(a-b)X} A_2 Y \sin \theta \cos \theta - Y^T A_1^T Y \cos^2 \theta
\]

which generates the required ellipse containing $Z^* AZ$ if we let $\theta$, $\alpha$, $\beta$ run over all the values in $[0, 2\pi]$. By the convexity property of numerical range, the union of all these closed elliptical disks is the numerical range of $A$.

\[\square\]

Actually, Lemma 2.3 holds for any matrix $A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$, where $A_1$ and $A_3$ are square matrix.

For any square matrix $A$ and $s > 0$, define the set

\[W(A, s) := \{ a + re^{i\theta} : a \in W(A), 0 \leq \theta \leq 2\pi, \text{ and } 0 \leq r \leq s \} .\]

Clearly $W(A) = W(A, 0) \subset W(A, s)$. The set $W(A, s)$ is the region obtained by expanding the region $W(A)$ by $s$ units in all directions. The following is the main result of this paper.

**Theorem 2.4.** For any $A = \begin{pmatrix} A_1 & A_2 \\ 0 & -A_1^T \end{pmatrix} \in b$, define $B = A_1 \oplus (-A_1^T)$. Then

1. $W(A) \subseteq W(B, \frac{1}{2}s)$, where $s$ is the largest singular values of $A_2$.
2. If $A_1$ is skew symmetric and $A_2 = sI_n$, then $W(A) = W(B, \frac{1}{2}s)$.
3. If $W(A) = W(B, \frac{1}{2}s)$, then $W(A_1) = W(-A_1)$, furthermore for any point $x$ on the boundary $\partial W(A_1)$ that is not convex combination of any other two points in $W(A_1)$, there is $X \in \mathbb{C}^n$ such that $x = X^* A_1 X = -X^* A_1^T X$.

**Proof.** (1) Let $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \in \mathbb{C}^{2n}$ be any unit vector with $X_1 = (x_1, \ldots, x_n)^T$ and $X_2 = (x_{n+1}, \ldots, x_{2n})^T$. Then

\[
X^* A X = X^* B X + X_1^* A_1 X_2
= X^* B X + X_1^* A_2 X_2
= X^* B X + (UX_1)^* U A_2 U^T (U^T X_2)
\text{for any } U \in U(n) \tag{2.2}
\]

Since $A_2$ is symmetric, there exists a $U \in U(n)$ such that $UA_2 U^T = \text{diag} \left( s_1, \ldots, s_n \right)$, where $s_1 \geq \cdots \geq s_n$ are singular values of $A_2$. From now on, we always assume that $A_2 = \text{diag} \left( s_1, \ldots, s_n \right)$ unless otherwise specified. Let $Y_1 = (y_1, \ldots, y_n)^T = UX_1$ and $Y_2 = (y_{n+1}, \ldots, y_{2n})^T = U^T X_2$. In (2.2),

\[
| (UX_1)^* U A_2 U^T (U^T X_2) | = | Y_1^* \text{diag} \left( s_1, \ldots, s_n \right) Y_2 |
\]

Since $s_1 \geq \cdots \geq s_n$ are singular values of $A_2$, From now on, we always assume that $A_2 = \text{diag} \left( s_1, \ldots, s_n \right)$ unless otherwise specified. Let $Y_1 = (y_1, \ldots, y_n)^T = UX_1$ and $Y_2 = (y_{n+1}, \ldots, y_{2n})^T = U^T X_2$. In (2.2),

\[
| (UX_1)^* U A_2 U^T (U^T X_2) | = | Y_1^* \text{diag} \left( s_1, \ldots, s_n \right) Y_2 |
\]

by

\[
\leq \sum_{k=1}^{n} | y_k | s_k | y_{n+k} |
\]

\[
\leq \sum_{k=1}^{n} \frac{1}{2} s_k (| y_k |^2 + | y_{n+k} |^2) \tag{2.3}
\]

\[
\leq \frac{1}{2} s_1 \sum_{k=1}^{n} (| y_k |^2 + | y_{n+k} |^2) = \frac{1}{2} s.
\]
Thus $X^*AX \in W(B, \frac{1}{2}s)$. Hence $W(A) \subset W(B, \frac{1}{2}s)$. The equality in (2.3) holds if and only if

$$Y_2 = e^{i\theta}Y_1$$

for some $\theta$ and $s_j = s$ for all $1 \leq j \leq n$ where $y_j \neq 0$.

(2) Assume that $A_1$ is skew symmetric and $A_2 = sI_n$, we will show that the boundary of $W(B, \frac{1}{2}s)$ is a subset of $W(A)$. Let $p = Z' = BZ + \frac{1}{2}s e^{i\theta}$ be any point on the boundary of $W(B, s)$, where $Z = (\sin aX_1^\top, \cos aX_1^\top)^\top \in \mathbb{C}^{2n}$ for some unit vectors $X_1, X_2 \in \mathbb{C}^n$ and $a \in [0, 2\pi]$.

$$p = \sin^2 aX_1^\top A_1 X_1 + \cos^2 aX_2^\top A_1 X_2 + \frac{1}{2}s e^{i\theta}.$$

Since the numerical range $W(A_1)$ is convex, the convex combination $\sin^2 aX_1^\top A_1 X_1 + \cos^2 aX_2^\top A_1 X_2 = X_2 A_1 X_2 \in W(A_1)$ for some unit vector $X_3 \in \mathbb{C}^n$. Let $Z_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{-i\theta/2} X_3 \\ e^{i\theta/2} X_3 \end{pmatrix}$. Then

$$W(A) \supset Z_1^\top A_1 Z_1 = \frac{1}{2} X_2^\top A_1 X_2 + \frac{1}{2} X_2^\top A_1 X_2 + \frac{1}{2} e^{i\theta} X_3^\top s I_n X_3 = p.$$

So the boundary of $W(B, \frac{1}{2}s)$ is a subset of $W(A)$. Thus $W(B, \frac{1}{2}s) \subset W(A)$ since $W(A)$ is convex. Therefore $W(A) = W(B, \frac{1}{2}s)$.

(3) Suppose that $W(A) = W(B, \frac{1}{2}s)$. Assume that $W(A_1) \neq W(-A_1)$. Since $W(A_1)$ and $W(-A_1)$ are convex, we can pick a point $p = X_1^\top A_1 X_1, ||X_1|| = 1$ on the boundary of $W(A_1)$ that satisfies the following: the supporting line $\ell$ of $W(A_1)$ through $p$ does not intersect $W(-A_1)$ and both $W(A)$ and $W(-A_1)$ are on the same side of $\ell$. Choose a vector $pq = \frac{1}{2}se^{i\theta}$ that is perpendicular to $\ell$ and $q \notin W(A_1)$. So $q = p + \frac{1}{2}se^{i\theta}$. Because $W(B) = \text{conv} \{ W(A_1) \cup W(-A_1) \}$, the convex hull containing $W(A_1)$ and $W(-A_1)$, $q$ is on the boundary of $W(B, \frac{1}{2}s)$ and the distance between $q$ and $W(B)$ is $\frac{1}{2}s$. Since $W(A) = W(B, \frac{1}{2}s)$, there exists a unit vector $Z = \begin{pmatrix} \sin aY_1 \\ \cos aY_2 \end{pmatrix} \in \mathbb{C}^{2n}$ such that

$$q = Z^\top AZ = \sin^2 aY_1^\top A_1 Y_1 + \cos^2 a(-Y_2^\top A_1^\top Y_2) + \frac{1}{2}\sin(2a)Y_1^\top A_2 Y_2 \in W(A) \quad (2.4)$$

Because $\sin^2 aY_1^\top A_1 Y_1 + \cos^2 a(-Y_2^\top A_1^\top Y_2)$ is a convex combination of $Y_1^\top A_1 Y_1 \in W(A_1)$ and $-Y_2^\top A_1^\top Y_2 \in W(-A_1)$, hence it is in $W(B)$. Thus the distance between $q = Z^\top AZ$ and $W(B)$,

$$d(q, W(B)) \leq \frac{1}{2}\sin(2a)Y_1^\top A_2 Y_2 \leq \frac{1}{2}s \sin(2a) \quad \text{by inequalities (2.2) and (2.3)}.$$

For that the equality holds, $\sin 2a = 1, Y_2 = e^{i\theta} Y_1$ and $\sin^2 aY_1^\top A_1 Y_1 + \cos^2 a(-Y_2^\top A_1^\top Y_2) = p$. But if this is the case, then $Y^\top A_1 Y_1 = -Y_2^\top A_1^\top Y_2 = p$, that means $p$ is also in $W(-A_1)$, which is not true. So $W(A_1) = W(-A_1)$. Furthermore, if $x = p$ is a point on the boundary of $W(A_1) = W(-A_1)$ that is not a convex combination of any other points in $W(A_1)$, let $q$ be the point in Equation (2.4). Since the distance $d(q, W(B)) = \frac{1}{2}s$, $Y_1 = e^{i\theta} Y_2$ and $Y_1^\top A_1 Y_1 = -Y_2^\top A_1^\top Y_2 = p$. Let $X = Y_1$, then $X^\top A_1 X = -X^\top A_1^\top X = p = x$. 

Theorem 2.4 holds for more general matrices. Let $A = \begin{pmatrix} B_{m \times n} & C_{m \times n} \\ 0 & D_{n \times n} \end{pmatrix}$ be any $m \times n$ square matrix with $m \leq n$. Let $C = U_1 \Sigma U_2^\top$ be the singular value decomposition of $C$. So we can assume that the block $C = \Sigma$. If $W(A) = W(B, \frac{1}{2}s)$, then $W(B) = W(D)$ and for each $x$ on the boundary of $W(B)$ that is not a convex combination of another two points of $W(B)$, then there exists $X \in \mathbb{C}^m$ and $Y = \begin{pmatrix} X \\ 0 \end{pmatrix} \in \mathbb{C}^n$ such that $x = X^\top B X = Y^\top D Y$.

Clearly if $B = D$ and $C = sI_n$, then $W(A) = W(B, \frac{1}{2}s)$. We conjecture that $W(A) = W(B, \frac{1}{2}s)$ if and only if $A$ is unitarily similar to a matrix of the following form

$$\begin{pmatrix} B_1 & 0 & sI_k & 0 \\ 0 & B_2 & 0 & \Sigma_1 \\ 0 & 0 & B_1 & 0 \\ 0 & 0 & 0 & D_2 \end{pmatrix},$$

where $W(B_2) \cup W(D_2) \subset W(B_1)$ and the singular values of $\Sigma_1$ are less than or equal to $s$. 
Acknowledgement: The work of Jicheng Tao was supported by the provincial Natural Science Foundation of Zhejiang (No. LY16A010009).

References