DAVIS-WIELANDT SHELLS OF OPERATORS

CHI-KWONG LI, YIU-TUNG POON, AND NUNG-SING SZE

Dedicated to Professor Yik-Hoi Au-Yeung for his 70th birthday.

ABSTRACT. Basic properties of Davis-Wielandt shells are presented. Conditions on two operators A and B with the same Davis-Wielandt shells are analyzed. Special attention is given to the case when B is a compression of A, and when $B = A^*$, A^t, or $(A^*)^t$, where A^t is the transpose of A with respect to an orthonormal basis. The results are used to study the point spectrum, approximate spectrum, and residual spectrum of the sum of two operators. Relation between the geometrical properties of the Davis-Wielandt shells and algebraic properties of operators are obtained. Complete descriptions of the Davis-Wielandt shells are given for several classes of operators.

AMS Subject Classification 47A10, 47A12, 47B15
Keywords Davis-Wielandt shells, generalized numerical ranges, spectra, operators

1. Introduction

Let $\mathcal{B}(\mathcal{H})$ be the algebra of bounded linear operators acting on the Hilbert space $\mathcal{H}$. We identify $\mathcal{B}(\mathcal{H})$ with M_n if $\mathcal{H}$ has dimension n. The *numerical range* of $A \in \mathcal{B}(\mathcal{H})$ is defined by

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

see [4, 5, 6]. The numerical range is useful in studying matrices and operators. In particular, the geometrical properties of $W(A)$ often provide useful information on the algebraic or analytic properties of A. For instance, $W(A) = \{\mu\}$ if and only if $A = \mu I$; $W(A) \subseteq \mathbb{R}$ if and only if $A = A^*$; $W(A)$ has no interior point if and only if there are $a, b \in \mathbb{C}$ with $a \neq 0$ such that $aA + bI$ is self-adjoint; see [4, 5, 6]. Moreover, there are nice connections between $W(A)$ and the spectrum $\sigma(A)$ of the operator A. For example, the closure of $W(A)$, denoted by $\mathbf{cl}\,(W(A))$, always contains $\sigma(A)$. If A is normal, then $\mathbf{cl}\,(W(A)) = \mathbf{conv}\sigma(A)$, where $\mathbf{conv}\sigma(A)$ denotes the convex hull of $\sigma(A)$. However, $\mathbf{cl}\,(W(A)) = \mathbf{conv}\sigma(A)$ does not imply that A is normal; see Problem 10 in [6, p.14].

Motivated by theoretical study and applications, there have been many generalizations of the numerical range; see [4, 5, 6]. One of these generalizations is the *Davis-Wielandt shell* of $A \in \mathcal{B}(\mathcal{H})$ defined by

$$DW(A) = \{(\langle Ax, x\rangle, \langle Ax, Ax\rangle) : x \in \mathcal{H}, \langle x, x\rangle = 1\};$$

see [2, 3, 9]. Evidently, the projection of the set $DW(A)$ on the first coordinate is $W(A)$. So, $DW(A)$ captures more information about the operator A than $W(A)$. For example, in the finite dimensional case, normality of operators can be completely determined by the geometrical shape of their Davis-Wielandt shells, namely, $A \in M_n$ is normal if and only if $DW(A)$ is a polyhedron in $\mathbb{C} \times \mathbb{R}$ identified with $\mathbb{R}^3$. In [8], it was shown that the Davis-Wielandt shell is a useful tool for characterizing the eigenvalues of matrices in the set

$$\{U^*AU + V^*BV : U, V \in M_n \text{ are unitary}\}$$

for given $A, B \in M_n$.

In this paper, we establish more results showing that the Davis-Wielandt shell is useful in studying operators. In Section 2, we present some basic results for the Davis-Wielandt shell. In Section 3, we obtain conditions on the operators A and B such that $DW(A) = DW(B)$. We also compare the sets $DW(A)$, $DW(A^*)$, $DW(A^t)$ and $DW((A^*)^t)$, where A^t is the transpose of $A \in \mathcal{B}(\mathcal{H})$ with respect to a fixed orthonormal basis. In Section 4, we obtain relations between $DW(A)$ and different kinds of spectra of $A \in \mathcal{B}(\mathcal{H})$. Furthermore, for $A, B \in \mathcal{B}(\mathcal{H})$, we use the Davis-Wielandt shell to study the point spectrum, the approximate spectrum and the residual spectrum of an operator of the form $U^*AU + V^*BV$, where $U, V \in \mathcal{B}(\mathcal{H})$ are unitary operators. In Section 5, we study the relation between the geometrical properties of $DW(A)$ and the algebraic properties of $A \in \mathcal{B}(\mathcal{H})$. Complete descriptions are obtained for the Davis-Wielandt shells of several classes of operators.

2. Basic properties

We begin with the following observations.

Theorem 2.1. *Let $A \in \mathcal{B}(\mathcal{H})$.*

(a) $(\mu, r) \in DW(A)$ *if and only if there is an orthonormal pairs of vectors $x, y \in \mathcal{H}$ such that $Ax = \mu x + \sqrt{r - |\mu|^2}y$.*

(b) *The set $DW(A)$ is bounded. In particular, $DW(A) \subseteq \mathcal{P}(A)$ with*

$$\mathcal{P}(A) = \{(\mu, r) \in \mathbb{C} \times [0, \infty) : |\mu|^2 \le r \le \|A\|^2\}.$$

(c) *$DW(A) = DW(U^*AU)$ for any unitary $U \in \mathcal{B}(\mathcal{H})$.*

(d) *For any $\alpha, \beta \in \mathbb{C}$, $DW(\alpha A + \beta I)$ equals*

$$\{(\alpha\mu + \beta, |\alpha|^2\nu + 2\mathrm{Re}\,(\alpha\bar{\beta}\mu) + |\beta|^2) : (\mu, \nu) \in DW(A)\},$$

which is the image of $DW(A)$ under a real affine transform.

(e) *Suppose $A \in \mathcal{B}(\mathcal{H})$ is a direct sum of $A_1 \oplus \cdots \oplus A_m$. Then*

$$DW(A) = \mathbf{conv}\{DW(A_1) \cup \cdots \cup DW(A_m)\}.$$

(f) *The set $DW(A)$ is closed if $\dim \mathcal{H}$ is finite.*

Proof. (a) The implication ($\Leftarrow$) is clear. Suppose $(\mu, r) \in DW(A)$. Then there is unit vector $x \in \mathcal{H}$ such that $\langle Ax, x\rangle = \mu$. Thus, $Ax = \mu x + \nu y$ for some unit vector $y \in x^\perp$ with $\nu \geq 0$. As $r = \|Ax\|^2 = |\mu|^2 + \nu^2$, the assertion holds.

(b) Suppose $(\mu, r) \in DW(A)$. Evidently, $r \leq \|A\|^2$. By (a), $|\mu|^2 \leq r$. So, $DW(A) \subseteq \mathcal{P}(A)$ as asserted.

(c) – (f) can be verified readily. □

Next, we give a description of the $DW(A)$ for $A \in M_2$.

Theorem 2.2. *Let $A \in M_2$ with eigenvalues a_1, a_2. Then $DW(A) \subseteq \mathbb{C} \times \mathbb{R}$ (identified with $\mathbb{R}^3$) is an ellipsoid without the interior centered at $(\mathrm{tr}\, A, \mathrm{tr}\, A^*A)/2$ with a (vertical) principal axis*

$$\{(\mathrm{tr}\, A/2, r) : r \geq 0,\ |r - \mathrm{tr}\, A^*A/2| \leq \|A^*A - (\mathrm{tr}\, A^*A)I/2\|\};$$

the projection of $DW(A)$ on the first co-ordinate equals the elliptical disk $W(A)$ with foci a_1, a_2 and minor axis of length $\sqrt{\mathrm{tr}\, A^*A - |a_1|^2 - |a_2|^2}$. Consequently, $DW(A)$ is convex if and only if A is normal. In such case, $DW(A)$ is a line segment joining $(a_1, |a_1|^2)$ and $(a_2, |a_2|^2)$.

Using the convexity properties of the joint numerical range, see [7] for example, we have the following result.

Theorem 2.3. *Suppose $A \in \mathcal{B}(\mathcal{H})$ with $\dim \mathcal{H} \geq 3$. Then $DW(A)$ is convex.*

Let $A \in \mathcal{B}(\mathcal{H})$ and $\mathcal{E} = \{e_i : i \in I\}$ be an orthonormal basis of $\mathcal{H}$. The transpose of A with respect to $\mathcal{E}$ is the operator $A^t_{\mathcal{E}} \in \mathcal{B}(\mathcal{H})$ defined by

$$\langle A^t_{\mathcal{E}} e_i, e_j\rangle = \langle Ae_j, e_i\rangle.$$

We claim that

$$DW\left(A^t_{\mathcal{E}}\right) = \{(\overline{\mu}, r) : (\mu, r) \in DW(A^*)\}.$$

To see this, let $x = \sum_i x_i e_i \in \mathcal{H}$. Define $\overline{x} = \sum_i \overline{x_i} e_i$, where $\overline{x_i}$ is the complex conjugate of x_i. For $x = \sum_i x_i e_i,\ y = \sum_i y_i e_i \in \mathcal{H}$, we have

$$\langle A^t_{\mathcal{E}} x, y\rangle = \sum_i \sum_j x_i \overline{y_j} \langle A^t_{\mathcal{E}} e_i, e_j\rangle = \sum_j \sum_i \overline{y_j} x_i \langle A e_j, e_i\rangle = \langle A\overline{y}, \overline{x}\rangle\,.$$

Therefore, we have

$$\langle A^t_{\mathcal{E}} x, x\rangle = \langle A\overline{x}, \overline{x}\rangle = \overline{\langle A^*\overline{x}, \overline{x}\rangle},$$

and

$$\begin{aligned}\langle A^t_{\mathcal{E}} x, A^t_{\mathcal{E}} x\rangle &= \langle A\overline{\left(A^t_{\mathcal{E}} x\right)}, \overline{x}\rangle = \langle \overline{\left(A^t_{\mathcal{E}} x\right)}, A^*\overline{x}\rangle \\ &= \overline{\langle A^t_{\mathcal{E}} x, \overline{A^*\overline{x}}\rangle} = \overline{\langle AA^*\overline{x}, \overline{x}\rangle} = \langle A^*\overline{x}, A^*\overline{x}\rangle.\end{aligned}$$

Hence, our claim follows.

By the above discussion, we see that $DW\left(A^t_{\mathcal{E}}\right)$ is independent of $\mathcal{E}$ so that have the following.

Theorem 2.4. *Let $A \in \mathcal{B}(\mathcal{H})$, and let A^t be the transpose of A with respect to any orthonormal basis of $\mathcal{H}$. Then*

$$DW\left(A^t\right) = \{(\overline{\mu}, r) : (\mu, r) \in DW(A^*)\}.$$

3. Comparison of Davis-Wielandt shells of two operators

In the following, we compare $DW(A)$ and $DW(B)$ for A and B acting on two (possibly different) Hilbert spaces. We then apply the results to compare $DW(A)$, $DW(A^*)$, $DW(A^t)$, and $DW((A^*)^t)$. Let us begin with the easy case when $A, B \in M_2$.

Theorem 3.1. *For $A, B \in M_2$, the following conditions are equivalent.*

(a) $DW(A) = DW(B)$

(b) $W(A) = W(B)$

(c) $A = U^*BU$ *for some unitary* $U \in M_2$.

Moreover, we have $DW(A) = DW(A^t)$ and

$$DW(A^*) = DW((A^*)^t) = \{(\bar{\mu}, \nu) : (\mu, \nu) \in DW(A)\}.$$

The situation for higher dimensions is more intricate. We need the following notation in our discussion. For $A \in \mathcal{B}(\mathcal{H})$, let

$$\mathcal{L}_\mu(A) = \{r : (\mu, r) \in DW(A)\} \subseteq [0, \|A\|^2].$$

The *upper boundary* of $DW(A)$ is the set

$$\{(\mu, r) : \mu \in W(A),\ r = \sup \mathcal{L}_\mu(A)\}.$$

Similarly, we can define the *lower boundary* of $DW(A)$.

The following result is obvious.

Theorem 3.2. *Let A, B be bounded linear operators acting on two Hilbert spaces, which may be different. Then $DW(A) \subseteq DW(B)$ if and only if $\mathcal{L}_\mu(A) \subseteq \mathcal{L}_\mu(B)$ for each $\mu \in W(A)$.*

Next, we compare $DW(A)$ with $DW(B)$ for $B = A^t, A^*$, or $(A^*)^t$.

Theorem 3.3. *Let $A \in \mathcal{B}(\mathcal{H})$. Then for each $\mu \in W(A)$, the following conditions hold:*

(a) $\mathcal{L}_\mu(A^t) = \mathcal{L}_{\bar{\mu}}(A^*)$

(b) $\mathcal{L}_\mu(A) = \mathcal{L}_{\bar{\mu}}((A^*)^t)$

(c) $\sup \mathcal{L}_\mu(A) = \sup \mathcal{L}_\mu(A^t) = \sup \mathcal{L}_{\bar{\mu}}(A^*) = \sup \mathcal{L}_{\bar{\mu}}((A^*)^t)$.

Moreover, any one of the suprema is attained if and only if all the suprema are attained.

Proof. (a) follows from Theorem 2.4 and (b) follows from (a).

For (c), if dim $\mathcal{H} \le 2$, then A is unitarily similar to A^t and the result follows.

Suppose dim $\mathcal{H} \ge 3$. Let $(\mu, r) \in DW(A)$. Then there exists a unit vector $x \in \mathcal{H}$ such that $\mu = \langle Ax, x\rangle$ and $r = \langle Ax, Ax\rangle$. Therefore, $Ax = \mu x + \nu y$ for some unit vector $y \in x^\perp$ and $\nu = \sqrt{r - |\mu|^2}$ so that there are $a_{12}, a_{22} \in \mathbb{C}$ satisfying

$$\begin{bmatrix} \langle Ax, x\rangle & \langle Ay, x\rangle \\ \langle Ax, y\rangle & \langle Ay, y\rangle \end{bmatrix} = \begin{bmatrix} \mu & a_{12} \\ \nu & a_{22} \end{bmatrix}$$

and

$$\begin{bmatrix} \langle A^*x, x\rangle & \langle A^*y, x\rangle \\ \langle A^*x, y\rangle & \langle A^*y, y\rangle \end{bmatrix} = \begin{bmatrix} \overline{\mu} & \nu \\ \overline{a_{12}} & \overline{a_{22}} \end{bmatrix}.$$

Since every matrix $X \in M_2$ is unitarily similar to its transpose, there is a unitary $V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \in M_2$ such that

$$V^* \begin{bmatrix} \overline{\mu} & \nu \\ \overline{a_{12}} & \overline{a_{22}} \end{bmatrix} V = \begin{bmatrix} \overline{\mu} & \overline{a_{12}} \\ \nu & \overline{a_{22}} \end{bmatrix}.$$

Then for $v = v_{11}x + v_{21}y$, we have

$$\overline{\mu} = \langle A^*v, v\rangle \qquad \text{and} \qquad s = \|A^*v\|^2 \ge \nu^2 + |\mu|^2 = r.$$

Hence, $(\overline{\mu}, s) \in DW(A^*)$, and $\sup \mathcal{L}_\mu(A) \le \sup \mathcal{L}_{\bar\mu}(A^*)$. The reverse inequality can be proved similarly. ¿From the proof, it is clear one of the suprema is attained if and only if both suprema are attained.

The other equalities in (c) follow from (a), (b) and $\sup \mathcal{L}_\mu(A) = \sup \mathcal{L}_{\bar\mu}(A^*)$.
□

We have the following corollary.

Corollary 3.4. *Let $A \in \mathcal{B}(\mathcal{H})$, and let $T : \mathbb{C}\times\mathbb{R} \to \mathbb{C}\times\mathbb{R}$ be the involution map $(\mu, r) \mapsto (\bar\mu, r)$.*

(a) *T will transform $DW(A)$ to $DW((A^*)^t)$.*

(b) *T will transform $DW(A^t)$ to $DW(A^*)$.*

(c) *$DW(A)$ and $DW(A^t)$ have the same upper boundary.*

(d) *$DW(A^*)$ and $DW((A^*)^t)$ have the same upper boundary.*

Note that $DW(A) \ne DW(A^t)$ and the map $(\mu, r) \mapsto (\bar\mu, r)$ does not transform $DW(A)$ to $DW(A^*)$. Here is an example.

Example 3.5. *Let $\mathcal{H}$ be an infinite dimensional Hilbert space with an orthonormal basis $\{e_n\}_{n=1}^\infty$, and let $S \in \mathcal{B}(H)$ be the unilateral shift such that $S(e_n) = e_{n+1}$ for $n \ge 1$. Then $S^* = S^t$ is the backward shift such that $S^t(e_1) = 0$ and $S^t(e_n) = e_{n-1}$ for $n > 1$. We have*

$$DW(S) = \{(\mu, 1) : \mu \in \mathbb{C},\ |\mu| < 1\}$$

and

$$DW(S^t) = DW(S) \cup \{(\mu, r) : \mu \in \mathbb{C},\ |\mu|^2 \le r < 1\}.$$

Thus, the map T defined by $(\mu, r) \mapsto (\bar\mu, r)$ does not transform $DW(S)$ to $DW(S^t)$.

Verification. Note that $|\langle Sx, x\rangle| < 1$ and $\langle Sx, Sx\rangle = 1$ for all unit vector $x \in \mathcal{H}$. Thus,

$$DW(S) \subseteq \{(\mu, 1) : \mu \in \mathbb{C},\ |\mu| < 1\}. \tag{1}$$

Since the projection of $DW(S)$ to the first co-ordinate is $W(S)$ and

$$\{\mu \in \mathbb{C} : |\mu| \le 1\} = \sigma(S) \subseteq \mathbf{cl}\,(W(S)),$$

we see that the set inclusion (1) is actually a set equality.

Next, observe that $\|S^t\| = 1$. By Theorem 2.1 (b) and Corollary 3.4 (c), $DW(S) \subseteq DW(S^t)$ and

$$DW(S^t) \subseteq DW(S) \cup \{(\mu, r) \in \mathbb{C}\times[0,\infty) : |\mu|^2 \le r < 1\}. \tag{2}$$

For any $\mu \in \mathbb{C}$ with $|\mu| < 1$, let $x = \sum_{k=1}^{\infty} \mu^k e_k$. Then $S^t x = \mu x$ so that $(\mu, |\mu|^2) \in DW(S^t)$. Hence the convex set $DW(S^t)$ has lower boundary $\{(\mu, |\mu|^2) : \mu \in \mathbb{C}, |\mu| < 1\}$. Consequently, the set inclusion in (2) is actually a set equality. □

Next we compare $DW(A)$ and $DW(A^*)$, it suffices to compare $\mathcal{L}_\mu(A)$ and $\mathcal{L}_{\bar{\mu}}(A^*)$. We have the following result.

Theorem 3.6. *Let $A \in \mathcal{B}(\mathcal{H})$. Then*

$$DW(A^*) \subseteq \{(\bar{\mu}, r) : (\mu, r) \in DW(A)\}$$

if and only if $\mathcal{L}_{\bar{\mu}}(A^*) \subseteq \mathcal{L}_\mu(A)$ for every $\mu \in W(A)$. As a result,

$$DW(A^*) = \{(\bar{\mu}, r) : (\mu, r) \in DW(A)\}$$

if and only if for every unit vector $x \in \mathcal{H}$, there exist unit vectors y and $z \in \mathcal{H}$ such that the following conditions are satisfied:

$$\langle Ax, x\rangle = \langle Ay, y\rangle \quad \text{with} \quad \|Ay\| \le \|A^*x\|$$

and

$$\langle Ax, x\rangle = \langle Az, z\rangle \quad \text{with} \quad \|A^*z\| \le \|Ax\|.$$

Proof. Note that

$$DW(A) = \cup_{\mu \in W(A)} \{(\mu, r) : r \in \mathcal{L}_\mu(A)\}$$

and

$$DW(A^*) = \cup_{\mu \in W(A)} \{(\bar{\mu}, r) : r \in \mathcal{L}_{\bar{\mu}}(A^*)\}.$$

Hence, $DW(A^*) \subseteq \{(\bar{\mu}, \nu) : (\mu, \nu) \in DW(A)\}$ if and only if $\mathcal{L}_{\bar{\mu}}(A^*) \subseteq \mathcal{L}_\mu(A)$ for each $\mu \in W(A)$.

Suppose $DW(A^*) = \{(\bar{\mu}, r) : (\mu, r) \in DW(A)\}$, or equivalently, $DW(A) = \{(\mu, r) : (\bar{\mu}, r) \in DW(A^*)\}$. Suppose $x \in \mathcal{H}$ is a unit vector so that $\bar{\mu} = \langle A^*x, x\rangle$ and $r = \|A^*x\|^2$. Then $r \in \mathcal{L}_\mu(A)$ so that there is a unit vector $y \in \mathcal{H}$ satisfying $\langle Ax, x\rangle = \langle Ay, y\rangle$ and $\|Ay\|^2 \le r = \|A^*x\|^2$. On the other hand, if $\tilde{r} = \|Ax\|^2$, then $\tilde{r} \in \mathcal{L}_{\bar{\mu}}(A^*)$. So, there exists a unit vector $z \in \mathcal{H}$ such that $\langle Ax, x\rangle = \langle Az, z\rangle$ and $\|A^*z\|^2 \le \tilde{r} = \|Ax\|^2$.

Conversely, suppose for any unit vector $x \in \mathcal{H}$ there are unit vectors y and z satisfying the said conditions. Then $\inf \mathcal{L}_\mu(A) = \inf \mathcal{L}_{\bar{\mu}}(A^*)$, and any of the infima is attained if and only if both infima are attained. By Theorem 3.3, $\sup \mathcal{L}_\mu(A) = \sup \mathcal{L}_{\bar{\mu}}(A^*)$, and any of the suprema is attained if and only if both suprema are attained. Thus, $\mathcal{L}_\mu(A) = \mathcal{L}_{\bar{\mu}}(A^*)$ for every $\mu \in W(A)$. Hence, $DW(A^*) = \{(\bar{\mu}, r) : (\mu, r) \in DW(A)\}$. □

In the finite dimensional case, we have the following.

Theorem 3.7. *Let $A \in M_n$. Then*

$$DW(A^*) = \{(\overline{\mu}, r) : (\mu, r) \in DW(A)\}. \tag{3}$$

Consequently, $DW(A) = DW(A^t)$ and $DW(A^*) = DW((A^*)^t)$.

Proof. The result for $n = 2$ follows easily from Theorem 3.1. For $n \geq 3$, $DW(A)$ is always convex by Theorem 2.3. Suppose our assertion is not true. Assume that there is $(\mu, r) \in DW(A)$ such that $(\overline{\mu}, r) \notin DW(A^*)$. Then there is a unit vector x such that $(\overline{x^*Ax}, x^*A^*Ax) = (x^*Hx - ix^*Gx, x^*A^*Ax) \notin DW(A^*)$, where $H = (A + A^*)/2$ and $G = (A - A^*)/2i$. By the separation theorem, there are $a, b, c \in \mathbb{R}$ such that

$$x^*(aA^*A + bH + cG)x > y^*(aAA^* + bH + cG)y$$

for all unit vector y. We may perturb a if necessary to assume that $a \neq 0$. Then for $\mu = (b + ic)/(2a)$,

$$x^*(A + \mu I)^*(A + \mu I)x = x^*(A^*A + (b/a)H + (c/a)G + |\mu|^2 I)x$$

$$\notin W(AA^* + (b/a)H + (c/a)G + |\mu|^2 I) = W((A + \mu I)(A + \mu I)^*).$$

This is a contradiction because

$$W((A + \mu I)^*(A + \mu I)) = W((A + \mu I)(A + \mu I)^*)$$

is the line segment $[s_n^2, s_1^2]$, where s_1 and s_n are the largest and smallest singular values of $A + \mu I$. Thus, equation (3) holds.

The last assertion follows from Corollary 3.4 and (3). □

Now, we compare $DW(B)$ and $DW(A)$ when B is a compression of A to a closed subspace $\mathcal{H}_1$ of $\mathcal{H}$, i.e., $B = X^*AX$ where $X : \mathcal{H}_1 \to \mathcal{H}$ such that $X^*X = I_{\mathcal{H}_1}$. For the classical numerical range, we have $W(B) \subseteq W(A)$. On the contrary, $DW(B)$ may not be contained in $DW(A)$. For instance, if $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $B = [0]$, then

$$DW(B) = \{(0, 0)\} \not\subseteq \{(t, 1) : t \in [-1, 1]\} = DW(A).$$

Nevertheless, we have the following.

Theorem 3.8. *Suppose B is a compression of $A \in \mathcal{B}(\mathcal{H})$ on the closed subspace $\mathcal{H}_1$. Then $DW(B) \subseteq DW(A)$ if any one of the following holds.*

(a) *$\mathcal{H}_1$ is an invariant subspace of A.*

(b) *$\mathcal{H}$ is finite dimensional and $\mathcal{H}_1^\perp$ is an invariant subspace of A.*

Proof. If (a) holds, then A has operator matrix of the form $\begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$, where $B \in \mathcal{B}(\mathcal{H}_1)$ and $D \in \mathcal{B}(\mathcal{H}_1^\perp)$. For any unit vector $y \in \mathcal{H}_1$, let $x = y \oplus 0 \in \mathcal{H}_1 \oplus \mathcal{H}_1^\perp$. Then $\langle Ax, x\rangle = \langle By, y\rangle$ and $\langle Ax, Ax\rangle = \langle By, By\rangle$. Therefore, $(\langle By, y\rangle, \langle By, By\rangle) \in DW(A)$.

If (b) holds, then A has operator matrix of the form $\begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$. By Theorem 3.7 and part (a), we have

$$DW(B) = DW(B^t) \subseteq DW(A^t) = DW(A)$$

as asserted. □

In Theorem 3.8 (b), the assumption on the dimension of $\mathcal{H}$ is indispensable as shown by the following example.

Example 3.9. *Let $\mathcal{H}$ be an infinite dimensional Hilbert space with an orthonormal basis $\{e_n\}_{n=1}^\infty$ and $S \in \mathcal{B}(H)$ be the unilateral shift such that $S(e_n) = e_{n+1}$ for $n \geq 1$. Let $\mathcal{H}_1$ be the subspace spanned by e_1. Then with $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_1^\perp$, S has the form $\begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$ with $B = [0]$ so that $(0,0) \in DW(B)$. By Example 3.5, $DW(S) = \{(\mu, 1) : \mu \in \mathbb{C},\ |\mu| < 1\}$ does not contain $(0,0)$.*

4. Spectra of operators

Denote by $\mathbf{cl}\,(S)$ and ∂S the closure and the boundary of a set S. Let $A \in \mathcal{B}(H)$. Recall that the *point spectrum* of A is the set $\sigma_p(A)$ of eigenvalues of A. The *residual point spectrum* of A is the set $\sigma_r(A)$ of complex number $\lambda \in \mathbb{C}$ such that the range of $\lambda I - A$ is not dense in $\mathcal{H}$. The *approximate point spectrum* of A is the set $\sigma_a(A)$ of complex number $\lambda \in \mathbb{C}$ such that there exists a sequence of unit vectors $\{x_n\}_1^\infty$ in $\mathcal{H}$ such that $\lim_{n\to\infty} \|(\lambda I - A)x_n\| = 0$. We have

$$\sigma_p(A) \subseteq \sigma_a(A), \quad \sigma(A) = \sigma_a(A) \cup \sigma_r(A)\,.$$

Theorem 4.1. *Suppose $A \in \mathcal{B}(\mathcal{H})$ so that $DW(A) \subseteq \mathcal{P}$ with*

$$\mathcal{P} = \{(\mu, r) \in \mathbb{C} \times [0, \infty) : |\mu|^2 \leq r\}.$$

Then the following conditions hold.

(a) $(\mu, r) \in DW(A) \cap \partial\mathcal{P}$ *if and only if* $\mu \in \sigma_p(A)$ *and* $r = |\mu|^2$.
(b) $(\mu, r) \in \mathbf{cl}\,(DW(A)) \cap \partial\mathcal{P}$ *if and only if* $\mu \in \sigma_a(A)$ *and* $r = |\mu|^2$.
(c) $(\mu, r) \in DW(A^*) \cap \partial\mathcal{P}$ *if and only if* $\overline{\mu} \in \sigma_r(A)$ *and* $r = |\mu|^2$.

(d) *If $DW(A)$ is closed, then $\sigma_p(A) = \sigma_a(A)$.*

Proof. (a) Suppose $(\mu, r) \in DW(A) \cap \partial\mathcal{P}$. Then $r = |\mu|^2$. Thus, there is a unit vector $x \in \mathcal{H}$ such that $\mu = \langle Ax, x\rangle$ and $|\mu| = \|Ax\|$. So, $|\langle Ax, x\rangle| = \|Ax\|$, and hence $Ax = \mu x$, i.e., $\mu \in \sigma_p(A)$. Conversely, if $Ax = \mu x$, then $(\mu, |\mu|^2) \in DW(A) \cap \partial\mathcal{P}$.

To prove (b), suppose $(\mu, |\mu|^2) \in \mathbf{cl}\,(DW(A))$. Then there is a sequence of unit vectors $\{x_n\}$ in $\mathcal{H}$ such that $(\langle Ax_n, x_n\rangle, \|Ax_n\|^2) \to (\mu, |\mu|^2)$. Suppose $Ax_n = \mu_n x_n + \nu_n y_n$, where $y_n \in x_n^\perp$ is a unit vector and $\mu_n, \nu_n \in \mathbb{C}$ such that $\nu_n \geq 0$. Then $\mu_n \to \mu$ and $|\mu_n|^2 + |\nu_n|^2 \to |\mu|^2$. We see that $\|Ax_n - \mu x_n\| \to 0$. Thus, $\mu \in \sigma_a(A)$. For the converse, suppose $\mu \in \sigma_a(A)$. Then there is a sequence of unit vectors $\{x_n\}$ in $\mathcal{H}$ such that $\|Ax_n - \mu x_n\| \to 0$. It follows that $(\langle Ax_n, x_n\rangle, \|Ax_n\|^2) \to (\mu, |\mu|^2)$. Thus, $(\mu, |\mu|^2) \in \mathbf{cl}\,(DW(A)) \cap \partial\mathcal{P}$.

Condition (c) follows from (a) and the fact that $\sigma_r(A^*) = \overline{\sigma_p(A)}$; see for example [5, Chapter 9].

Condition (d) follows from (a) and (b). □

The following example illustrates Theorem 4.1 and shows that $DW(A)$ is more useful than $W(A)$ in the study of the spectrum of A.

Example 4.2. *Let $A = \mathrm{diag}\,(-1, 1, 1/2, 1/3, \cdots, 1/n, \cdots)$. Then*

$$\sigma_p(A) = \{-1\} \cup \{1/k : k = 1, 2, \dots\}, \quad \sigma(A) = \sigma_a(A) = \{0\} \cup \sigma_p(A),$$

$$W(A) = [-1, 1], \quad DW(A) = \mathbf{conv}\{(\mu, \mu^2) : \mu \in \sigma_p(A)\},$$

$$DW(A) \cap \partial\mathcal{P} = \{(\mu, \mu^2) : \mu \in \sigma_p(A)\},$$

and

$$\mathbf{cl}\,(DW(A)) \cap \partial\mathcal{P} = \{(\mu, \mu^2) : \mu \in \sigma(A)\}.$$

Let $A \in \mathcal{B}(H)$. The *unitary similarity orbit* of A is denoted by

$$\mathcal{U}(A) = \{U^*AU : U^*U = I = UU^*\}.$$

In the following, we study different kinds of spectra of operators in

$$\mathcal{U}(A) + \mathcal{U}(B) = \{X + Y : (X, Y) \in \mathcal{U}(A) \times \mathcal{U}(B)\}.$$

In [8], we considered $A, B \in M_n$ and proved that a complex number μ is an *eigenvalue* of some $C \in \mathcal{U}(A) + \mathcal{U}(B)$ if and only if $DW(A) \cap DW(\mu I - B) \neq \emptyset$. Here, we extend the result to infinite dimensional operators, and consider the different types of spectra. Let

$$\sigma(A, B) = \{\mu \in \mathbb{C} : \mu \in \sigma(C) \text{ for some } C \in \mathcal{U}(A) + \mathcal{U}(B)\}.$$

Similarly, define $\sigma_p(A,B)$, $\sigma_a(A,B)$ and $\sigma_r(A,B)$. Evidently,

$$\sigma(A,B) = \sigma_a(A,B) \cup \sigma_r(A,B)\,. \tag{4}$$

Theorem 4.3. *Let $A, B \in \mathcal{B}(\mathcal{H})$. Then*

$$\sigma_p(A,B) = \{\mu \in \mathbb{C} : DW(A) \cap DW(\mu I - B) \neq \emptyset\} \tag{5}$$

and

$$\sigma_r(A,B) = \sigma_p(A^*,B^*) = \{\mu \in \mathbb{C} : DW(A^*) \cap DW((\mu I - B)^*) \neq \emptyset\}. \tag{6}$$

Proof. In [8], (5) was proved for the finite dimensional case. The argument can be easily adapted to the general case as follows. If $\mu \in \sigma_p(A,B)$, then there are unitary operators $U, V \in \mathcal{B}(\mathcal{H})$ and a unit vector $x \in \mathcal{H}$ such that $(U^*AU + V^*BV)x = \mu x$. Let $U^*AUx = (\mu I - V^*BV)x = \alpha x + \beta y$ with $\alpha, \beta \in \mathbb{C}$ and unit vector $y \in x^\perp$. Thus

$$\begin{aligned} (\alpha, |\alpha|^2 + |\beta|^2) &\in DW(U^*AU) \cap DW(\mu I - V^*BV) \\ &= DW(A) \cap DW(\mu I - B). \end{aligned}$$

Conversely, if $(\nu, r) \in DW(A) \cap DW(\mu I - B)$, then there are unit vectors $x, \hat{x}, y, \hat{y} \in \mathcal{H}$ with $\hat{x} \in x^\perp$ and $\hat{y} \in y^\perp$ such that $Ax = \nu x + \sqrt{r - |\nu|^2}\hat{x}$ and $(\mu I - B)y = \nu y + \sqrt{r - |\nu|^2}\hat{y}$. Let $U = I$ and $V \in \mathcal{B}(\mathcal{H})$ be unitary such that $Vx = y$ and $V\hat{x} = \hat{y}$. Then $(U^*AU + V^*BV - \mu I)x = 0$ so that $\mu \in \sigma_p(A,B)$.

Next, suppose $\mu \in \mathbb{C}$. By the fact that $\sigma_p(T^*) = \overline{\sigma_r(T)}$ for any $T \in \mathcal{B}(\mathcal{H})$, we have

$$\begin{aligned} & \mu \in \sigma_r(A,B) \\ \Leftrightarrow\ & \mu \in \sigma_r(C) \text{ for some } C \in \mathcal{U}(A) + \mathcal{U}(B) \\ \Leftrightarrow\ & \overline{\mu} \in \sigma_p(C^*) \text{ for some } C^* \in \mathcal{U}(A^*) + \mathcal{U}(B^*) \\ \Leftrightarrow\ & DW(A^*) \cap DW((\mu I - B)^*) \neq \emptyset. \end{aligned}$$

□

For $\sigma_a(A,B)$, we have the following.

Theorem 4.4. *If $\mu \in \sigma_a(A,B)$ then $\mathbf{cl}\,(DW(A)) \cap \mathbf{cl}\,(DW(\mu I - B)) \neq \emptyset$.*

Proof. Suppose $\mu \in \sigma_a(A,B)$. Then there exist unitaries U and $V \in \mathcal{B}(\mathcal{H})$ and a sequence of unit vectors $\{x_n\}_1^\infty$ in $\mathcal{H}$ such that

$$\lim_{n\to\infty} \|(\mu I - (U^*AU + V^*BV))x_n\| = 0.$$

Let $u_n = Ux_n$ and $v_n = Vx_n$. Then

$$\begin{aligned} & (\langle U^*AUx_n, x_n\rangle, \langle U^*AUx_n, U^*AUx_n\rangle) \\ =\ & (\langle Au_n, u_n\rangle, \langle Au_n, Au_n\rangle) \in DW(A) \end{aligned}$$

and

$$\begin{aligned} & (\langle V^*(\mu I - B)Vx_n, x_n\rangle, \langle V^*(\mu I - B)Vx_n, V^*(\mu I - B)Vx_n\rangle) \\ = & (\langle(\mu I - B)v_n, v_n\rangle, \langle(\mu I - B)v_n, (\mu I - B)v_n\rangle) \in DW(\mu I - B)\,. \end{aligned}$$

Since

$$\begin{aligned} & |\langle V^*(\mu I - B)Vx_n, x_n\rangle - \langle U^*AUx_n, x_n\rangle| \\ = & |\langle(\mu I - (U^*AU + V^*BV))x_n, x_n\rangle| \\ \leq & \|(\mu I - (U^*AU + V^*BV))x_n\| \end{aligned}$$

and

$$\begin{aligned} & |\langle V^*(\mu I - B)Vx_n, V^*(\mu I - B)Vx_n\rangle - \langle U^*AUx_n, U^*AUx_n\rangle| \\ = & |\langle(\mu I - (U^*AU + V^*BV))x_n, V^*(\mu I - B)Vx_n\rangle \\ & \qquad + \langle U^*AUx_n, (\mu I - (U^*AU + V^*BV))x_n\rangle| \\ \leq & \|(\mu I - (U^*AU + V^*BV))x_n\|\|\mu I - B\| \\ & \qquad + \|A\|\|(\mu I - (U^*AU + V^*BV))x_n\|, \end{aligned}$$

we have

$$\begin{aligned} & \lim_{n\to\infty} |\langle V^*(\mu I - B)Vx_n, x_n\rangle - \langle U^*AUx_n, x_n\rangle| \\ = & \lim_{n\to\infty} |\langle V^*(\mu I - B)Vx_n, V^*(\mu I - B)Vx_n\rangle - \langle U^*AUx_n, U^*AUx_n\rangle| \\ = & 0\,. \end{aligned}$$

Passing to a subsequence of $\{x_n\}_1^\infty$, if necessary, we may assume that there are $c \in \mathbb{C}$ and $r \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \langle U^*AUx_n, x_n\rangle = c = \lim_{n\to\infty} \langle V^*(\mu I - B)Vx_n, x_n\rangle$$

and

$$\lim_{n\to\infty} \langle U^*AUx_n, U^*AUx_n\rangle = r = \lim_{n\to\infty} \langle V^*(\mu I - B)Vx_n, V^*(\mu I - B)Vx_n\rangle.$$

Hence, $(c, r) \in \mathbf{cl}\,(DW(A)) \cap \mathbf{cl}\,(DW(\mu I - B))$. □

Corollary 4.5. *Suppose $A, B \in \mathcal{B}(\mathcal{H})$ such that both $DW(A)$ and $DW(B)$ are closed. Then*

$$\sigma_p(A, B) = \sigma_a(A, B). \tag{7}$$

Proof. Suppose A, $B \in \mathcal{B}(\mathcal{H})$ such that both $DW(A)$ and $DW(B)$ are closed. Let $\mu \in \sigma_a(A, B)$. For a unit vector $x \in \mathcal{H}$, let $(\langle Bx, x\rangle, \langle Bx, Bx\rangle) =$

(c_x, r_x). Then we have

$$\begin{aligned} \langle(\mu I - B)x, x\rangle &= \mu - \langle Bx, x\rangle = \mu - c_x \\ \langle(\mu I - B)x, (\mu I - B)x\rangle &= |\mu|^2 - \mu\langle x, Bx\rangle - \overline{\mu}\langle Bx, x\rangle + \langle Bx, Bx\rangle \\ &= |\mu|^2 - \mu\overline{c}_x - \overline{\mu}c_x + r_x\,. \end{aligned}$$

Therefore, $DW(\mu I - B)$ is a continuous image of the compact set $DW(B)$. Hence, $DW(\mu I - B)$ is closed. We have

$$\begin{aligned} & \mu \in \sigma_a(A, B) \\ \Rightarrow\quad & \mathbf{cl}\,(DW(A)) \cap \mathbf{cl}\,(DW(\mu I - B)) \neq \emptyset \\ \Rightarrow\quad & DW(A) \cap DW(\mu I - B) \neq \emptyset \\ \Rightarrow\quad & \mu \in \sigma_p(A, B). \end{aligned}$$

This establishes (7). □

The following example shows that the converse of Theorem 4.4 may not hold.

Example 4.6. *Let $\{e_n\}_1^\infty$ be an orthonormal basis of $\mathcal{H}$. Define $A, B \in \mathcal{B}(\mathcal{H})$ by $A = \operatorname{diag}(0, 1, 1, \dots)$, and $B = \operatorname{diag}(1, 1/2, 1/3, \dots)$ with respect to the basis $\{e_n\}_1^\infty$. Then*

$$(0, 0) \in DW(A) \cap \mathbf{cl}\,(DW(0I - B)) \qquad \textit{and} \qquad 0 \notin \sigma_a(A, B).$$

Verification. Clearly, $(0, 0) \in DW(A) \cap \mathbf{cl}\,(DW(0I - B))$. Suppose there exist a unitary U and a sequence $\{x_n\}$ of unit vectors such that

$$\lim_{n\to\infty} \|(A + U^*BU)x_n\| = 0.$$

Let

$$x_n = \sum_{k=1}^{\infty} x_{nk}\, e_k \qquad \text{and} \qquad Ue_1 = u = \sum_{k=1}^{\infty} u_k e_k.$$

Then

$$\begin{aligned} & \lim_{n\to\infty} \|(A + U^*BU)x_n\| = 0 \\ \Rightarrow\quad & \lim_{n\to\infty} \|(UA + BU)x_n\| = 0 \\ \Rightarrow\quad & \lim_{n\to\infty} \|(UA - U + U + BU)x_n\| = 0 \\ \Rightarrow\quad & \lim_{n\to\infty} \|U(A - I)x_n + (I + B)Ux_n\| = 0. \end{aligned}$$

Since

$$\begin{aligned} \|U(A - I)x_n + (I + B)Ux_n\| &\geq \|(I + B)Ux_n\| - \|U(A - I)x_n\| \\ &\geq \|Ux_n\| - \|x_{n1}u\| = 1 - |x_{n1}|\,, \end{aligned}$$

we have

$$\lim_{n\to\infty} |x_{n1}| = 1 \quad \Rightarrow \quad \lim_{n\to\infty} \|(x_n - x_{n1}e_1)\| = 0\,.$$

Since

$$\begin{aligned}\|(UA+BU)x_n\| &\geq \|(UA+BU)(x_{n1}e_1)\| - \|(UA+BU)(x_n - x_{n1}e_1)\| \\ &\geq |x_{n1}|\|Bu\| - \|(UA+BU)\|\|(x_n - x_{n1}e_1)\|\end{aligned}$$

and

$$\begin{aligned}\|(UA+BU)x_n\| &\leq \|(UA+BU)(x_{n1}e_1)\| + \|(UA+BU)(x_n - x_{n1}e_1)\| \\ &\leq \|Bu\| + \|(UA+BU)\|\|(x_n - x_{n1}e_1)\|\,,\end{aligned}$$

we have

$$\lim_{n\to\infty} \|(A+U^*BU)x_n\| = \|Bu\| = \sqrt{\sum_{k=1}^{\infty} \frac{|u_k|^2}{k^2}} > 0\,,$$

which is a contradiction. □

The next example shows that $DW(A) \cap DW(\mu I - B)$ may or may not be empty for $\mu \in \sigma_a(A,B)$.

Example 4.7. *Let $\mathcal{H}$ be an infinite dimensional Hilbert space with an orthonormal basis $\{e_n\}_{n=1}^{\infty}$ and $S^* \in \mathcal{B}(H)$ be the backward shift defined in Example 3.5. Then $(0,0) \in DW(S^*)$ and $(1,1) \notin DW(S^*)$. If $(A,B) = (S^*,0)$, then $0,1 \in \sigma_a(A,B)$. However, $DW(A) \cap DW(0I - B) \neq \emptyset$ while $DW(A) \cap DW(1I - B) = \emptyset$.*

Remark 4.8. *Example 4.6 and Example 4.7 show that a necessary and sufficient condition for $\mu \in \sigma_a(A,B)$ cannot be described by*

$$DW(A) \cap DW(\mu I - B) \neq \emptyset, \quad DW(A) \cap \mathbf{cl}\,(DW(\mu I - B)) \neq \emptyset$$

nor

$$\mathbf{cl}\,(DW(A)) \cap \mathbf{cl}\,(DW(\mu I - B)) \neq \emptyset.$$

5. Special operators

In this section, we study the relationship between the geometrical properties of $DW(A)$ and the algebraic properties of $A \in \mathcal{B}(\mathcal{H})$. In particular, we give complete descriptions of the Davis-Wielandt shells for several classes of operators.

It is well known that for $A \in B(\mathcal{H})$, $W(A) = \{\lambda\}$ if and only if $A = \lambda I$; $W(A)$ is a line segment if and only if A is essentially self-adjoint, i.e., there are $a, b \in \mathbb{C}$ with $a \neq 0$ such that $aA + bI$ is self-adjoint; see [4, 5, 6]. Using these results, we have the following.

Theorem 5.1. *Let $A \in \mathcal{B}(\mathcal{H})$.*

(a) *A is essentially self-adjoint if and only if $DW(A)$ is a subset of a plane perpendicular to the (x, y)-plane if we identify $\mathbb{C} \times \mathbb{R}$ with $\mathbb{R}^3$.*
(b) *There are $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$ such that $\alpha A + \beta I$ is a nonzero orthogonal projection $P \neq I$ if and only if $DW(A)$ is a nondegenerate line segment.*

Proof. (a) The implication ($\Rightarrow$) is clear. Conversely, if $DW(A)$ is a subset of a plane perpendicular to the (x, y)-plane, then $W(A)$ is a subset of a straight line in the complex plane. Then there are $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$ such that $W(\alpha A + \beta I) = \alpha W(A) + \beta \subseteq \mathbb{R}$. It follows that $\alpha A + \beta I$ is self-adjoint.

(b) The implication ($\Rightarrow$) is clear. Conversely, suppose that $DW(A)$ is a nondegenerate line segment. By (a), A is essentially self-adjoint. If $\sigma(A)$ has more than two points, then $\mathbf{cl}\,(DW(A))$ has more than two points in the set $\{(\mu, |\mu|^2) : \mu \in \mathbb{C}\}$ so that $DW(A)$ is not a line segment by convexity. Clearly, $\sigma(A)$ cannot be a singleton. Thus, $\sigma(A)$ has two points and the result follows. □

Corollary 5.2. *Let $A \in \mathcal{B}(\mathcal{H})$.*

(a) *A is self-adjoint if and only if $DW(A)$ is a subset of the (x, z)-plane if we identify $\mathbb{C} \times \mathbb{R}$ with $\mathbb{R}^3$.*
(b) *A is a non-scalar orthogonal projection if and only if $DW(A)$ is the line joining $(0, 0)$ and $(1, 1)$.*
(c) *$A = \lambda I$ if and only if $DW(A) = \{(\lambda, |\lambda|^2)\}$.*

Theorem 5.3. *Let $A \in \mathcal{B}(\mathcal{H})$. Then there are $\alpha, \beta \in \mathbb{C}$ such that $\alpha \neq 0$ and $\alpha A + \beta I$ is an isometry if and only if $DW(A)$ is a subset of a plane not perpendicular to the (x, y)-plane when $\mathbb{C} \times \mathbb{R}$ is identified with $\mathbb{R}^3$.*

Proof. Suppose $B = \alpha A + \beta I$ is an isometry for some $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$. Then

$$DW(B) \subseteq \{(\mu, 1) : \mu \in \mathbb{C}\}.$$

By Theorem 2.1 (d), $DW(A) = DW(\alpha^{-1}(B - \beta I))$ is a subset of

$$\{(\alpha^{-1}(\mu - \beta), |\alpha|^{-2}(|\beta|^2 - 2\mathrm{Re}\,(\mu\bar{\beta}) + 1)\,) : \mu \in \mathbb{C}\},$$

which is a subset of a plane not perpendicular to the (x, y)-plane if we identify $\mathbb{C} \times \mathbb{R}$ with $\mathbb{R}^3$.

Conversely, suppose $DW(A) \subseteq \{(x+iy, z) : ax+by+cz = d\}$ with $c \neq 0$. Let $\beta = (a+ib)/(2c)$. By Theorem 2.1 (d), $DW(A+\beta I)$ consists of points in $\mathbb{C} \times \mathbb{R}$ with the second co-ordinate equal to

$$(ax+by+cz)/c + |\beta|^2 = d/c + |\beta|^2,$$

where $(x+iy, z) \in DW(A)$. Thus, $\|(A+\beta I)v\|^2 = d/c + |\beta|^2$ for every unit vector $v \in \mathcal{H}$. Equivalently, $A + \beta I$ is a multiple of an isometry. □

We can say a little bit more for $DW(A)$ when $A \in \mathcal{B}(\mathcal{H})$ is an isometry.

Theorem 5.4. *Suppose $A \in \mathcal{B}(\mathcal{H})$ is an isometry. Then*

$$DW(A) = \{(\mu, 1) : \mu \in W(A)\}.$$

If A has a compression equal to the shift operator S defined in Example 3.5 so that $DW(S) = \{(\mu, 1) : \mu \in \mathbb{C}, |\mu| < 1\}$, then

$$DW(A) = DW(S) \cup \{(\mu, 1) : \mu \in \sigma_p(A)\}. \tag{8}$$

Proof. Suppose $A \in \mathcal{B}(\mathcal{H})$ is an isometry. The first assertion follows from the fact that the projection of $DW(A)$ to its first co-ordinate equals $W(A)$, and that $\|Ax\| = 1$ for all unit vectors $x \in \mathcal{H}$.

If S is a compression of A, then A has an operator matrix of the form

$$\begin{bmatrix} S & B \\ 0 & C \end{bmatrix}$$

as A is an isometry. By Theorem 3.8 (a),

$$DW(S) \subseteq DW(A) \subseteq DW(S) \cup \{(\mu, 1) : \mu \in \mathbb{C},\ |\mu| = 1\}.$$

Moreover, for $\mu \in \mathbb{C}$ with $|\mu| = 1$, we have $(\mu, 1) \in DW(A)$ if and only if $\mu \in \sigma_p(A)$ by Theorem 4.1 (a). We get (8). □

Acknowledgment

Li is an honorary professor of the University of Hong Kong. His research was supported by a USA NSF grant and a HK RCG grant. This research was conducted when the three authors were visiting the University of Hong Kong in the summer of 2007 sponsored by a HK RCG grant. They would like to express their thanks for the hospitality of the staff of the mathematics department of the University of Hong Kong. Also, the authors wish to express their thanks to the referee for a very careful reading of the manuscript and many helpful comments.

References

[1] J.B. Conway, A Course in Functional Analysis, Second edition, Graduate Texts in Mathematics, 96, Spring-Verlag, New York, 1990.

[2] C. Davis, The shell of a Hilbert-space operator, Acta Sci. Math.(Szeged) 29 (1968), 69-86.

[3] C. Davis, The shell of a Hilbert-space operator. II, Acta Sci. Math. (Szeged) 31 (1970), 301-318.

[4] K.E. Gustafson and D.K.M. Rao, Numerical ranges: The field of values of linear operators and matrices, Springer, New York, 1997.

[5] P. Halmos, A Hilbert Space Problem Book, Second edition, Graduate Texts in Mathematics, 19, Encyclopedia of Mathematics and its Applications, Spring-Verlag, New York, 1982.

[6] R.A. Horn and C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[7] C.K. Li and Y.T. Poon, Convexity of the joint numerical range, SIAM J. Matrix Analysis Appl. 21 (1999), 668-678.

[8] C.K. Li, Y.T Poon and N.S. Sze, Eigenvalues of the sum of matrices from unitary similarity orbits, SIAM J. Matrix Analysis Appl., to appear.

[9] H. Wielandt, On eigenvalues of sums of normal matrices, Pacific J. Math. 5 (1955), 633–638.

(Li) Department of Mathematics, College of William & Mary,
Williamsburg, VA 23185 (ckli@math.wm.edu).

(Poon) Department of Mathematics, Iowa State University,
Ames, IA 50011 (ytpoon@iastate.edu).

(Sze) Department of Mathematics, University of Connecticut,
Storrs, CT 06269 (sze@math.uconn.edu).