The Davis-Wielandt Shell

GUPTA Pranav

1 Introduction

This project aims to explore the concept of Davis-Wielandt (DW) Shells in linear algebra literature and prove two key theorems regarding the convexity of a DW shell for an n x n matrix, and the separation of the DW shells of two different n x n matrices. For simplicity, the scope of this project is limited to square matrices in M_n with $n \geq 3$ that exist on the Hilbert Space $\mathbb{H}$.

The numerical range $W(A)$ of a square matrix $A \in M_n$ is a geometric object in the complex plane that contains all the possible values of the quadratic $\langle Ax, x\rangle$, where $x \in \mathbb{C}^n$ and $\langle x, x\rangle = 1$. [5] There are several practical use cases for knowing the numerical range of a matrix in engineering. For instance, it is well known that the numerical range of a matrix contains all of the eigenvalues of the matrix. [4] This can provide an estimate of the location and the spread of the eigenvalues for the matrix, which can be used in stability analysis. For example, if the numerical range of the matrix lies within the left half of the complex plane, then the system is guaranteed to be asymptotically stable. The numerical range is also useful in understanding important geometric properties of the given matrix. For instance, one can yield the numerical radius by defining the minimum and maximum magnitude values in the numerical range, which can be used in perturbation and convergence problems. [5] The numerical range of A $\in M_n$ is defined by the equation:

$$W(A) = \{\langle Ax, x\rangle : \langle x, x\rangle = 1\}. \tag{1}$$

The DW Shell is a generalization of the numerical range. It extends a third real and non-negative dimension from the complex plane of the numerical range, which is given by $\langle Ax, Ax\rangle$ and in practical applications can represent the 'gain' of the output for the matrix A with respect to different possible input x vectors. The DW Shell of $A \in M_n$ is defined by the equation:

$$DW(A) = \{(\langle Ax, x\rangle, \langle Ax, Ax\rangle) : \langle x, x\rangle = 1\}. \tag{2}$$

It is known that $W(A) \subset \mathbb{C}$, and is compact and contains the spectrum of A. According to the Toeplitz-Hausdorff theorem, $W(A)$ is convex. [2] The convexity of $DW(A)$ is a crucial property for understanding the geometric properties of $DW(A)$ and how it interacts with other operators. The next section proves that $DW(A)$ is convex for all n x n matrices where $n \geq 3$.

2 Convexity of Davis-Wielandt Shells

As defined above, the DW shell is a three-dimensional geometric shape. The first two coordinates in this 3D space correspond to the real and imaginary parts of each possible value in the numerical range $W(A)$ of a given square matrix A. The third coordinate represents the respective L2 norm of Ax for the respective unit vector $x \in \mathbb{C}$. In the definition of the numerical range described in Equation 1, we see that the range is defined as the quadratic $\langle Ax, x\rangle$, which can also be written as:

$$W(A) = \{x^*Ax : \forall x \in \mathbb{C}^n, \langle x, x\rangle = 1\}. \tag{3}$$

Using Hausdorff's [3] result that proves that $W(A)$ defined in Equation 3 is convex, Au-Yeung [1] proved that the set $W(A_1, A_2)$ defined in Equation 6 is indeed convex for Hermitian matrices $A_1, A_2 \in \mathbb{H}_n(R)$ and $n > 2$. In [1], Au-Yeung and Tsing demonstrate that the set $W_c(A_1, A_2, A_3)$ is convex for $n > 2$ and for any $A_1, A_2, A_3 \in \mathbb{H}_n(R)$ and $(c_1, ..., c_n) \in \mathbb{R}^n$, where the set $W_c(A_1, A_2, A_3)$ is written as:

$$W_c(A_1, A_2, A_3) = \{(tr[CUA_1U]^*, tr[CUA_2U^*], tr[CUA_3U^*]) : \forall U \in \mathbb{U}_n\}, \tag{4}$$

where C is a diagonal n x n matrix with the diagonal entries being $(c_1, ..., c_n)$, and U_n is the set of n x n unitary matrices.

Therefore, to prove the that $DW(A)$ is convex for all n x n matrices where $n \geq 3$, we go through the following steps.

Using Toeplitz decomposition, every $A \in \mathbb{C}_n$ can be written uniquely as

$$A = A_r + iA_c,$$

in which A_r is the matrix of the real parts of all entries in A and A_c is the matrix of the magnitudes of the complex parts of all entries in A. A can be decomposed by using the following definition:

$$A = \frac{1}{2}(A + A^*) + i[\frac{1}{2i}(A - A^*)]. \tag{5}$$

We can observe that indeed $A_r = \frac{1}{2}(A + A^*)$ and $A_c = \frac{1}{2i}(A - A^*)$ are real-valued Hermitian Matrices. Therefore, the numerical range in Equation 3 can be defined as:

$$W(A_r, A_c) = \{(x^*A_rx, x^*A_cx) : \forall A_r, A_c \in \mathbb{H}_n, \forall x \in \mathbb{C}^n | \langle x, x\rangle = 1\}. \tag{6}$$

The third dimension in the DW Shell is given by $\langle Ax, Ax\rangle$, which can be rewritten as $\{x^*A^*Ax\}$. Thus, we can let $A_d = A^*A$, rewriting the third dimension values as $\{x^*A_dx\}$. Notice that since A_d is a product of $A \in \mathbf{M_n}(\mathbf{C})$ with its conjugate transpose. Hence, A_d is Hermitian, as shown below:

$$A_d^* = (A^*A)^* = (A^*)^*A^* = AA^* = A^*A = A_d.$$

Therefore, we can rewrite Equation 4 as:

$$W_c(A_r, A_c, A_d) = \{(tr[CUA_rU^*], tr[CUA_cU^*], tr[CUA_dU^*]) : \forall U \in \mathbb{U}_n\}. \tag{7}$$

For all $U \in \mathbb{U}_n(C)$, it is known that the columns of U form an orthonormal basis in $\mathbb{C}^n$. This means each column vector u_k of U (for $k = 1, 2, \ldots, n$) satisfies:

$$\|u_k\|_2 = 1 \quad \text{and} \quad \langle u_i, u_j\rangle = 0 \quad \text{for} \quad i \neq j,$$

where $\|u_k\|_2$ is the ℓ_2-norm of u_k.

We can therefore partition the unitary matrix U in Equation 7 as $[U1 \mid U2]$, where $U1$ is the first column vector in U, and $U2$ is the remaining n x $n-1$ matrix. Moreover, since the diagonal matrix C in Equation 7 can contain any value $c \in \mathbb{R}$ along its diagonal, we can let $c_{11} = 1$ and all other entries $c_{ij} = 0 : i \neq j \neq 1$.

$$C = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}, U = \begin{bmatrix} u_{11} & * & \ldots & * \\ u_{21} & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & * & \ldots & * \end{bmatrix}.$$

This way for each matrix M in (A_r, A_c, A_d), the coordinate value which, according to Equation 7, is given as $tr[CU^*MU]$.

$$CU^*MU = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} \begin{bmatrix} u_{11} & u_{21} & \ldots & u_{n1} \\ * & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \ldots & * \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1n} \\ m_{21} & m_{22} & \ldots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \ldots & m_{nn} \end{bmatrix} \begin{bmatrix} u_{11} & * & \ldots & * \\ u_{21} & * & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ u_{n1} & * & \ldots & * \end{bmatrix},$$

$$tr[CU^*MU] = tr(\begin{bmatrix} U_1^*MU_1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}) = U_1^*MU_1.$$

Therefore, by using the method above, and the fact that $\langle U_1, U_1\rangle = \langle x, x\rangle = 1$, we get that $tr[CUA_rU^*] = U_1^*A_rU_1 = x^*A_rx$, $tr[CUA_cU^*] = U_1^*A_cU_1 = x^*A_cx$, $tr[CUA_dU^*] = U_1^*A_dU_1 = x^*A_dx$. Equation 7 can hence be rewritten as:

$$W_{(1,0,\ldots,0)}(A_r, A_c, A_d) = \{(x^*A_rx, x^*A_cx, x^*A_dx) : \forall x \in \mathbf{C^n}, \langle x, x\rangle = 1\}, \tag{8}$$

or,

$$W_{(1,0,\ldots,0)}(A_r, A_c, A_d) = \{(\langle Ax, x\rangle, \langle Ax, Ax\rangle) : \langle x, x\rangle = 1\} = DW(A). \tag{9}$$

Therefore, we can prove that $DW(A)$ is in fact equal to $W_{(1,0,\ldots,0)}(A_r, A_c, A_d)$. Since $W_c(A_r, A_c, A_d)$ in Equation 7 is convex for $n \geq 3$ [1], we have proven that the DW shell of $A : \forall A \in \mathbf{M_n}$ is indeed convex for $n \geq 3$.

3 The Separation of Davis-Wielandt Shells

Theorem:

Let $A, B \in \mathbb{M}_n$. The following are equivalent. [6]

(a) $det(U^*AU + V^*BV) \neq 0$ for any unitary matrices $U, V \in \mathbb{U}_n$ ($\mathbb{U}_n$ is the set of unitary matrices in $\mathbb{M}_n$).

(b) $DW(A) \cap DW(-B) = \emptyset$.

(c) There is $\xi \in \mathbb{C}$ such that the singular values of $A + \xi I_n$ and $B - \xi I_n$ lie in two disjoint closed intervals in $[0, \infty)$.

Lemma 1:

For a matrix $A \in \mathbb{M}_n$ and a unit vector $x \in \mathbb{C}^n$ such that $\langle Ax, x\rangle = \mu$, Ax can be rewritten as:

$$Ax = \mu x + \nu y$$

for some unit vector $y \in x^\perp$ (the orthogonal complement of x) with $\nu > 0$.

Proof.

Consider the projection of Ax onto x. Since x is a unit vector, the projection is given by:

$$\text{Proj}_x(Ax) = \langle Ax, x\rangle x = \mu x.$$

We can decompose Ax into two components: one along x and one orthogonal to x. Let $P_x = xx^*$ be the projection operator onto the span of x (where x^* is the Hermitian transpose of x). Then:

$$Ax = P_xAx + (I - P_x)Ax.$$

The component along x is μx:

$$P_xAx = xx^*Ax = \langle Ax, x\rangle x = \mu x.$$

The remaining component, orthogonal to x, is:

$$(I - P_x)Ax = Ax - \mu x.$$

Let:

$$y = \frac{(Ax - \mu x)}{\|Ax - \mu x\|}$$

if $Ax \neq \mu x$. Notice y is in $x^\perp$ and is a unit vector. Then:

$$Ax - \mu x = \nu y$$

where $\nu = \|Ax - \mu x\|$ is the norm of the orthogonal component. Note that if $Ax = \mu x$, then the orthogonal component is zero, so $\nu = 0$ and y can be any unit vector in $x^\perp$.

Therefore Ax can be rewritten as:

$$Ax = \mu x + \nu y$$

To prove that (a) $\implies$ (b):

Assume (a) is true: $\det(U^*AU + V^*BV) \neq 0$ for any unitary matrices U and V. This means that for any choice of unitary matrices U and V, the matrix sum $U^*AU + V^*BV$ is non-singular.

Suppose $DW(A) \cap DW(-B) \neq \emptyset$. For all unit vectors $x \in \mathbb{C}^n$, the $DW(A)$ for $A \in \mathbb{M}_n$ is defined as:

$$DW(A) = \{(\langle Ax, x\rangle, \langle Ax, Ax\rangle) : x \in \mathbb{C}^n, \langle x, x\rangle = 1\},$$

where every point (μ, z) contained in the set $DW(A)$ is on the $\mathbb{C} \times \mathbb{R}$ space, such that:

1. $\mu = \langle Ax, x\rangle = x^*Ax$,
2. $z = \langle Ax, Ax\rangle = x^*A^*Ax = \|Ax\|^2$.

Since $DW(A) \cap DW(-B) \neq \emptyset$, there exists at least one point (μ, z) such that $(\mu, z) \in DW(A) \cap DW(-B)$. Therefore, there must exist at two vectors $u_1, v_1 \in \mathbb{C}^n$ such that:

$$(\langle Au_1, u_1 \rangle, \|Au_1\|^2) = (\langle -Bv_1, v_1 \rangle, \|Bv_1\|^2).$$

This implies the following two points:

1. $\mu = \langle Au_1, u_1 \rangle = u_1^* Au_1 = \langle -Bv_1, v_1 \rangle = -v_1^* Bv_1,$
2. $z = \|Au_1\|^2 = \|Bv_1\|^2.$

Using the result of Lemma 1, we can rewrite Au_1 and Bv_1 as follows:

$$Au_1 = \mu u_1 + \nu u_2,$$

$$-Bv_1 = \mu v_1 + \nu v_2,$$

where $u_1 \perp u_2$ and $v_1 \perp v_2$.

Note that since $z = \|Au_1\|^2 = \|Bv_1\|^2$,

$$z = \|Au_1\|^2 = (\mu u_1 + \nu u_2)^*(\mu u_1 + \nu u_2) = |\mu|^2 u_1^* u_1 + \nu^2 u_2^* u_2 + \mu\nu u_1^* u_2 + \mu\nu u_2^* u_1,$$
$$z = \|Bv_1\|^2 = (\mu v_1 + \nu v_2)^*(\mu v_1 + \nu v_2) = |\mu|^2 v_1^* v_1 + \nu^2 v_2^* v_2 + \mu\nu v_1^* v_2 + \mu\nu v_2^* v_1.$$

Because $u_1 \perp u_2$, $u_2^* u_1 = u_1^* u_2 = 0$. As u_1 and u_2 are unit vectors, $u_1^* u_1 = u_2^* u_2 = 1$. Likewise, $v_2^* v_1 = v_1^* v_2 = 0$ and $v_1^* v_1 = v_2^* v_2 = 1$. Therefore, we can simplify the above expression as:

$$z = \|Au_1\|^2 = \|Bv_1\|^2 = |\mu|^2 + \nu^2,$$

or,

$$\nu = \sqrt{z - |\mu|^2}.$$

Let U and V be unitary matrices with their first two columns $U_{i1} = u_1$, $U_{i2} = u_2$ and $V_{i1} = v_1$, $V_{i2} = v_2$, respectively. For the matrix sum $U^*AU + V^*BV$, we aim to prove that since the first column is equal to zero, the matrix sum is singular.

Consider the matrices U^*AU and V^*BV:

$$U^*AU = \begin{bmatrix} u_{11}^* & u_{21}^* & \dots & u_{n1}^* \\ u_{12}^* & u_{22}^* & \dots & u_{n2}^* \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & * & \dots & * \\ u_{21} & u_{22} & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{n1} & u_{n2} & * & \dots & * \end{bmatrix} = \begin{bmatrix} u_1^* Au_1 & * & \dots & * \\ u_2^* Au_1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix},$$

$$V^*(-B)V = \begin{bmatrix} v_{11}^* & v_{21}^* & \dots & v_{n1}^* \\ v_{12}^* & v_{22}^* & \dots & v_{n2}^* \\ * & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & * \end{bmatrix} \begin{bmatrix} -b_{11} & -b_{12} & \dots & -b_{1n} \\ -b_{21} & -b_{22} & \dots & -b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -b_{n1} & -b_{n2} & \dots & -b_{nn} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & * & \dots & * \\ v_{21} & v_{22} & * & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{n1} & v_{n2} & * & \dots & * \end{bmatrix} = \begin{bmatrix} -v_1^* Bv_1 & * & \dots & * \\ -v_2^* Bv_1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix}.$$

By the definition of Lemma 1, Au_1 is decomposed into its projections on two orthogonal unit vectors u_1 and u_2. For our purposes, u_1 and u_2 are defined to be the first two column vectors of the unitary matrix U. Note that for all unitary matrices, all the column vectors in a given unitary matrix are orthonormal. Since $Au_1 = \mu u_1 + \nu u_2$, Au_1 has no projections on other possible column vectors orthogonal to u_1 and u_2. This is reflected in the result of U^*AU above, where each entry in the first column, given as $u_j^*(Au_1)$, is equal to zero for all column vectors $u_j \perp u_1 \perp u_2$.

Likewise, Since $Bv_1 = \mu v_1 + \nu v_2$, Bv_1 has no projections on other possible column vectors orthogonal to v_1 and v_2. Hence in the result of $V^*(-B)V$ above, each entry in the first column, given as $v_j^*(-Bv_1)$, is equal to zero for all column vectors $v_j \perp v_1 \perp v_2$.

Since $\mu = u_1^* A u_1 = -v_1^* B v_1$, we can see that the first term in the first column of the matrix sum $U^*AU + V^*BV$ below will be equal to zero.

Since $Au_1 = \mu u_1 + \nu u_2$, multiplying u_2^* to both sides yields:

$$u_2^* A u_1 = u_2^*(\mu u_1 + \nu u_2) = \mu u_2^* u_1 + \nu u_2^* u_2.$$

Since μ is a complex number, and $u_1 \perp u_2$, $\mu u_2^* u_1 = 0$. Since u_2 is unit vector, $u_2^* u_2 = 1$. Hence:

$$u_2^* A u_1 = \nu.$$

Similarly:

$$v_2^*(-B)v_1 = v_2^*(\mu v_1 + \nu v_2) = \mu v_2^* v_1 + \nu v_2^* v_2 = \nu.$$

Since $u_2^* A u_1 = v_2^*(-B)v_1 = \nu$, we can see that the second term in the first column of the matrix sum $U^*AU + V^*BV$ will be equal to zero too.

Therefore,

$$U^*AU + V^*(B)V = \begin{bmatrix} u_1^* B u_1 & * & \dots & * \\ u_2^* B u_1 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix} + \begin{bmatrix} -(-v_1^* B v_1) & * & \dots & * \\ -(-v_2^* B v_1) & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix} = \begin{bmatrix} \mu & * & \dots & * \\ \nu & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix} + \begin{bmatrix} -\mu & * & \dots & * \\ -\nu & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix}.$$

Because the first column of the matrix sum is a zero vector:

$$U^*AU + V^*(B)V = \begin{bmatrix} 0 & * & \dots & * \\ 0 & * & \dots & * \\ 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \dots & * \end{bmatrix},$$

$$det(U^*AU + V^*(B)V) = 0.$$

Therefore, by contradiction, if $det(U^*AU + V^*BV) \neq 0$ for any unitary matrices $U, V \in \mathbb{U}_n$, $DW(A) \cap DW(-B) = \emptyset$.

To prove that (b) $\implies$ (c):

As proven in Section 2, $DW(A)$ is convex for all matrices $A \in \mathbb{M}_n$ where $n \geq 3$. Since $DW(A)$ and $DW(-B)$ are compact convex sets, and $DW(A) \cap DW(-B) = \emptyset$, then by the Separation Theorem, there exists a linear functional f such that:

$$f(\alpha) > f(\beta) \quad \text{for all } (\alpha, \beta) \in DW(A) \times DW(-B).$$

Using the definition of the DW Shell, and by the Separation Theorem, there are $a, b, c \in \mathbb{R}$ such that for the matrices $A, -B \in \mathbb{M}_n$ and a unit vector $x \in \mathbb{C}$:

$$x^*(aA^*A + bA_r + cA_c)x > x^*(a(-B)^*(-B) + b(-B_r) + c(-B_c))x.$$

Note that A_r and A_c refer to the real and imaginary parts of A as described in Equation 5. Likewise, B_r and B_c are the real and imaginary parts of B. This functional describes scalars a, b, c such that a hyperplane could separate the disjoint sets $DW(A)$ and $DW(-B)$.

Expanding with Equation 5 and dividing by a on both sides:

$$x^*(A^*A + \frac{b}{a}\frac{A + A^*}{2} + \frac{c}{a}\frac{A - A^*}{2i})x > x^*(B^*B + \frac{-b}{a}\frac{B + B^*}{2} + \frac{-c}{a}\frac{B - B^*}{2i})x.$$

Simplifying gives:

$$x^*(A^*A + \frac{b}{2a}(A + A^*) + \frac{-ic}{2a}(A - A^*))x > x^*(B^*B + \frac{-b}{2a}(B + B^*) + \frac{ic}{2a}(B - B^*))x,$$

$$x^*(A^*A + \frac{b-ic}{2a}A + \frac{b+ic}{2a}A^*)x > x^*(B^*B - \frac{b-ic}{2a}B - \frac{b+ic}{2a}B^*)x.$$

Let $\xi = \frac{b+ic}{2a}$. Hence, $\overline{\xi} = \frac{b-ic}{2a}$. Adding $|\xi|^2 I_n$ to both sides:

$$x^*(A^*A + \overline{\xi}A + \xi A^* + |\xi|^2 I_n)x > x^*(B^*B - \overline{\xi}B - \xi B^* + |\xi|^2 I_n)x.$$

Factorizing both sides yields:

$$x^*(A + \xi I_n)^*(A + \xi I_n)x > x^*(B - \xi I_n)^*(B - \xi I_n)x.$$

Therefore,

$$\|(A + \xi I_n)x\|^2 > \|(B - \xi I_n)x\|^2.$$

This implies that for any unit vector $x \in \mathbb{C}^n$, a complex number ξ can be chosen such that all the singular values of $(A + \xi I_n)$ are greater than all the singular values of $(B - \xi I_n)$. In other words, the functional describes ξ that can separate all the values α and β in sets $DW(A)$ and $DW(-B)$ respectively. And hence, if $DW(A) \cap DW(-B) = \emptyset$, there exists $\xi \in \mathbb{C}$ such that the singular values of $A + \xi I_n$ and $B - \xi I_n$ lie in two disjoint closed intervals in $[0, \infty)$.

To prove that (c) $\implies$ (a):

According to (c), there exists a $\xi \in \mathbb{C}$ such that the singular values of $A + \xi I_n$ and $B - \xi I_n$ lie in two disjoint closed intervals $[a_1, a_2]$ and $[b_1, b_2]$ in $[0, \infty)$, respectively.

Consider $A' = A + \xi I_n$ and $B' = B - \xi I_n$. Since $U^*AU + V^*BV$ can be rewritten as:

$$U^*AU + V^*BV = U^*(A' - \xi I_n)U + V^*(B' + \xi I_n)V,$$

$$= U^*A'U - \xi U^*U + V^*B'V + \xi V^*V.$$

Since U and V are unitary, $U^*U = I_n$ and $V^*V = I_n$:

$$U^*AU + V^*BV = U^*A'U + V^*B'V.$$

We can denote the singular values of A' as $\sigma_1(A') \geq \sigma_2(A') \geq \cdots \geq \sigma_n(A') > 0$ and of B' as $\sigma_1(B') \geq \sigma_2(B') \geq \cdots \geq \sigma_n(B') \geq 0$. Since U and V are unitary,

$$\sigma_i(U^*A'U) = \sigma_i(A'),$$
$$\sigma_i(V^*B'V) = \sigma_i(B').$$

Without loss of generalization, we can assume that

$$\sigma_i(A') > \sigma_i(B') \text{ for all } i.$$

Every matrix A' and B' can be decomposed using Singular Value Decomposition:

$$A' = P\Sigma_{A'}Q^*,$$

$$B' = R\Sigma_{B'}S^*,$$

where: P, Q, R, and S are unitary matrices. $\Sigma_{A'}$ and $\Sigma_{B'}$ are diagonal matrices with singular values $\sigma_i(A')$ and $\sigma_i(B')$ on the diagonal, respectively.

For any unit vectors u and v:

$$\|A'u\|^2 = \|P\Sigma_{A'}Q^*u\|^2 = (u^*Q\Sigma_{A'}^*P^*)(P\Sigma_{A'}Q^*u) = \|\Sigma_{A'}Q^*u\|^2,$$

$$\|B'v\|^2 = \|R\Sigma_{B'}S^*v\|^2 = (v^*S\Sigma_{B'}R^*)(R\Sigma_{B'}S^*v) = \|\Sigma_{B'}S^*v\|^2.$$

Since P, and Q, R, S are unitary, they preserve the norm of vectors. Thus, we have:

$$\|A'u\| = \|\Sigma_{A'}y\| \quad \text{where} \quad y = Q^*u,$$

$$\|B'v\| = \|\Sigma_{B'}z\| \quad \text{where} \quad z = S^*v.$$

Since y and z are still unit vectors (due to the unitary transformations), we can write:

$$\|\Sigma_{A'} y\|^2 = \sum_{i=1}^{n} \sigma_i^2(A')|y_i|^2,$$

$$\|\Sigma_{B'} z\|^2 = \sum_{i=1}^{n} \sigma_i^2(B')|z_i|^2.$$

Due to the strict inequality $\sigma_i(A') > \sigma_i(B')$, for any non-negative weights $|y_i|^2$ and $|z_i|^2$ summing to 1, we have:

$$\sum_{i=1}^{n} \sigma_i^2(A')|y_i|^2 > \sum_{i=1}^{n} \sigma_i^2(B')|z_i|^2,$$

$$\|A'u\| > \|B'v\|.$$

Now, consider the matrix sum $M = U^*A'U + V^*B'V$. We want to show that $\det(M) \neq 0$. Assume for contradiction that $\det(M) = 0$. This would imply that M is singular, meaning there exists a non-zero vector x such that:

$$Mx = (U^*A'U + V^*B'V)x = 0.$$

However, this implies:

$$U^*A'Ux = -V^*B'Vx.$$

Taking norms on both sides and considering the unitary nature of U and V:

$$\|U^*A'Ux\| = \|V^*B'Vx\|.$$

This contradicts the earlier derivation that for any unit vectors u and v, $\|A'u\| > \|B'v\|$. Specifically, if $x = u = v$, then:

$$\|A'x\| > \|B'x\|.$$

Thus, there cannot exist such a non-zero vector x, and $U^*A'U + V^*B'V$ cannot be singular. Therefore, for any unitary matrices $U, V \in \mathbb{U}_n$, if there exists $\xi \in \mathbb{C}$ such that the singular values of $A + \xi I_n$ and $B - \xi I_n$ lie in two disjoint closed intervals in $[0, \infty)$:

$$\det(U^*AU + V^*BV) \neq 0$$

References

[1] Yik-Hoi Au-Yeung and Nam-Kiu Tsing. An extension of the hausdorff toeplitz theorem on the numerical range. *Proceedings of the American Mathematical Society Volume 89 Number 2*, 1989.

[2] Chandler Davis. The toeplitz hausdorff theorem explained. *Canad. Math. Bull. Vol. 14 (2)*, 1971.

[3] F. Hausdorff, Der Wertvorrat einer Bilinearform, Math. Z. 3 (1919), 314-316.

[4] Chi-Kwong Li. Lecture notes on Numerical Range.

[5] Chi-Kwong Li, Yiu-Tung Poon, and Nung-Sing Sze. Davis wielandt shells of operators. *Operators and Matrices 3(3)*, 2008.

[6] Chi-Kwong Li, Yiu-Tung Poon, and Nung-Sing Sze. Eigenvalues of the sum of matricesfrom unitary similarity orbits. *SIAM Journal on Matrix Analysis and Applications 30(2):560-581*, 2008.