Differential Properties of the Numerical Range Map of Pairs of Matrices

J. A. Hillman
School of Mathematics & Statistics
Sydney University, NSW, 2006, Australia

and

B. R. F. Jefferies, W. J. Ricker, and B. Straub*
School of Mathematics
The University of New South Wales
Sydney, NSW, 2052, Australia

Submitted by Peter Lancaster

ABSTRACT

Let $A = (A_1, A_2)$ be a pair of Hermitian operators in $\mathbb{C}^n$ and $A = A_1 + iA_2$. We investigate certain differential properties of the numerical range map $n_A : x \mapsto \langle A_1 x, x \rangle, \langle A_2 x, x \rangle$ with the aim of better understanding the nature of the numerical range $W(A)$ of $A$. For example, the joint eigenvalues of $A$ correspond to the stationary points of $n_A$ (i.e. points where the derivative $n'_A$ vanishes). Moreover, points $x$ where rank $n'_A(x) = 2$ get mapped by $n_A$ into the interior $W(A)^{\circ}$ of $W(A)$. For $n = 2$, it turns out that if $A_1$ and $A_2$ have no common invariant subspace, then the image under $n_A$ of the set $\Sigma_1 (A)$ consisting of those points $x$ with rank $n'_A(x) = 1$ is precisely the boundary $\partial W(A)$ of $W(A)$, and the image under $n_A$ of the rank 2 points for $n'_A$ is precisely $W(A)^{\circ}$; there are no rank 0 points for $n'_A$. As a consequence (for $n = 2$) we have that $A_1 A_2 = A_2 A_1$ iff $\Sigma_1 (A) \neq n_A^{-1}(\partial W(A))$. © 1997 Elsevier Science Inc.

*This author gratefully acknowledges the support of the Australian Research Council.

1. INTRODUCTION

Let $\mathbb{C}^n$ be the standard $n$-dimensional unitary space equipped with its usual inner product $\langle \cdot , \cdot \rangle$ and norm $\| \cdot \|$, where $n \geq 1$ is an integer. The space of all linear operators of $\mathbb{C}^n$ into itself is denoted by $\mathcal{L}(\mathbb{C}^n)$. Associated with each element $A \in \mathcal{L}(\mathbb{C}^n)$ are the Rayleigh quotient $x \mapsto \langle Ax, x \rangle / \langle x, x \rangle$, for $x \neq 0$, and its range $W(A)$, which is known as the numerical range of $A$. Since the quotient $\langle Ax, x \rangle / \langle x, x \rangle$ is invariant if $x$ is replaced by $\alpha x$ ($\alpha \neq 0$), it is clear that $W(A)$ is also given by

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

where $S(\mathbb{C}^n) = \{ x \in \mathbb{C}^n : \| x \| = 1 \}$. Both the Rayleigh quotient and the numerical range of $A$ have been extensively studied (see, e.g., [4, 5, 8, 9, 11, 13]) as they relate to certain important properties of the operator $A$. For example:

(P1) A stationary value of the Rayleigh quotient exists iff $A$ and its adjoint $A^*$ have an eigenvector in common. Moreover, all the eigenvectors of $A$ occur in this way iff all eigenvectors of $A$ are also eigenvectors of $A^*$, which is equivalent to $A$ being normal, i.e. $AA^* = A^*A$ (see [13, Section III.18]).

(P2) The set $W(A)$ is always compact and convex in $\mathbb{C}$. Thus, its boundary $\partial W(A)$ can have at most countably many points at which it is not differentiable. Any nondifferentiable points $\lambda \in \partial W(A)$ are necessarily eigenvalues of $A$ [4, Theorem 1].

In this paper we take a different approach. As is well known, there exists a uniquely determined pair $A = (A_1, A_2)$ of Hermitian operators in $\mathcal{L}(\mathbb{C}^n)$ such that $A = A_1 + iA_2$. Let $(u, v) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ be identified with $u + iv \in \mathbb{C}^n$, and $S(\mathbb{R}^{2n})$ be the unit sphere of $\mathbb{R}^{2n}$. Then the Rayleigh quotient of $A$, restricted to $S(\mathbb{C}^n)$, can be identified (see Section 2) with the numerical range map $n_A : S(\mathbb{R}^{2n}) \to \mathbb{R}^2$ associated with the pair $A$, where

$$n_A((u, v)) = (\langle A_1(u + iv), u + iv \rangle, \langle A_2(u + iv), u + iv \rangle),$$

$(u, v) \in S(\mathbb{R}^{2n})$.

Hence, the numerical range $W(A)$ can also be obtained as the range of $n_A$.

Our aim is to study certain differential properties of the map $n_A$. More precisely, since $n_A$ is a map from $S(\mathbb{R}^{2n})$ into $\mathbb{R}^2$, its derivative at a point
(u, v) ∈ S(ℝ^{2n}), considered as a linear map from the (2n - 1)-dimensional tangent space at (u, v) into ℝ^2, may have rank 0, rank 1, or rank 2. This partitions the domain S(ℝ^{2n}) of n_A into three pairwise disjoint sets Σ_0(A), Σ_1(A), and Σ_2(A), respectively. As it turns out, the stationary points of n_A are exactly the joint eigenvectors of A_1 and A_2 (which are obviously also the joint eigenvectors of the operator A = A_1 + iA_2 and its adjoint A^*). Moreover, all points (u, v) ∈ Σ_2(A) are mapped into the interior of W(A), and the points of Σ_1(A) turn out to be intimately connected with ∂W(A).

Our motivation for this approach is as follows. Given a collection of d Hermitian operators in C^n, say A_1, . . . , A_d, the Weyl calculus for the d-tuple A = (A_1, . . . , A_d) is an L(C^n)-valued distribution which represents a particular rule allowing the construction of certain functions of the matrices A_1, . . . , A_d. For ξ = (ξ_1, . . . , ξ_d) ∈ ℝ^d, the operator ⟨ξ, A⟩ = ∑_{j=1}^d ξ_j A_j is again Hermitian and hence ∥e^{i⟨ξ, A⟩}∥ = 1. Let L(ℝ^d) denote the Schwartz space of C-valued, rapidly decreasing functions on ℝ^d. More precisely then, the Weyl calculus for A (see [2, 12, 14]) is the L(C^n)-valued distribution T(A) defined by

\[ T(A)f = (2π)^{-n/2} \int_{ℝ^d} e^{i⟨ξ, A⟩} \hat{f}(ξ) \, dξ, \quad f ∈ L(ℝ^d); \]

here \( \hat{f} \) denotes the Fourier transform of \( f \). By virtue of a formula due to E. Nelson [12]; see also [7]) it is possible to write the distribution T(A) as a certain matrix-valued differential operator applied to the measure μ \circ n_A^{-1}, defined in ℝ^d, where μ is the uniform probability measure on S(ℂ^n) and n_A: x ↦ (⟨A_1 x, x⟩, . . . , ⟨A_d x, x⟩), for x ∈ S(ℂ^n), is the numerical range map. Accordingly, it is to be expected (and hoped) that information about differential properties of n_A will lead to a better understanding about analytic properties of the Weyl functional calculus for systems of matrices. The present note makes an initial step in this direction. The importance of differential properties of n_A is already apparent from classical matrix theory in the case d = 2.

2. MAIN RESULTS

Let A = (A_1, A_2) be a pair of Hermitian operators in L(C^n), and set A = A_1 + iA_2. For each k = 1, 2, . . . , let J_k: ℝ^k × ℝ^k → C^k be the bijection J_k(u, v) = u + iv. The Cartesian product ℝ^k × ℝ^k is identified with ℝ^{2k} in the obvious way. By N_A, we denote the map

\[ (u, v) ↦ J_1^{-1}(⟨A(u + iv), u + iv⟩), \quad u, v ∈ ℝ^n, \]
from \( \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n} \) into \( \mathbb{R}^2 \). Note that, because \( A_1, A_2 \) are Hermitian, both \( \langle A_1 x, x \rangle \) and \( \langle A_2 x, x \rangle \) are real numbers for every \( x \in \mathbb{C}^n \). Hence,

\[
N_A(u, v) = \left( \langle A_1(u + iv), u + iv \rangle, \langle A_2(u + iv), u + iv \rangle \right), \quad u, v \in \mathbb{R}^n.
\]

The derivative \( N_A'(p, q) \) of \( N_A \) at the point \( (p, q) \in \mathbb{R}^{2n} \) is the usual linear map from the Euclidean space \( \mathbb{R}^{2n} \) into \( \mathbb{R}^2 \) defined in the sense of standard multivariable calculus; i.e., it is specified by the condition

\[
\lim_{(h, k) \to 0} \frac{N_A(p + h, q + k) - N_A(p, q) - N_A'(p, q)(h, k)}{\| (h, k) \|} = 0.
\]

**Lemma 1.** Let \( p, q \in \mathbb{R}^n \) and \( x = p + iq \in \mathbb{C}^n \). Then the derivative \( N_A'(p, q): \mathbb{R}^{2n} \to \mathbb{R}^2 \) of \( N_A \) at \( (p, q) \in \mathbb{R}^{2n} \) is given by

\[
(h, k) \mapsto N_A'(p, q)(h, k)
\]

\[
= f_1^{-1}(\langle A_1 x, h + ik \rangle + \langle A_1(h + ik), x \rangle)
\]

\[
= 2(\text{Re}(\langle A_1 p, h \rangle + \langle A_1 q, k \rangle) + \text{Im}(\langle A_1 p, k \rangle - \langle A_1 q, h \rangle)).
\]

\[
\text{Re}(\langle A_2 p, h \rangle + \langle A_2 q, k \rangle) + \text{Im}(\langle A_2 p, k \rangle - \langle A_2 q, h \rangle).
\]

**Proof.** We have \( N_A(p, q) = (\langle A_1 x, x \rangle, \langle A_2 x, x \rangle) \). Since

\[
\langle A_j(p + iq), p + iq \rangle = \langle A_j p, p \rangle + \langle A_j q, q \rangle + 2 \text{Im}(\langle A_j p, q \rangle),
\]

for \( j \in \{1, 2\} \), the result follows by differentiating with respect to \( p \) and \( q \). 

By restricting \( N_A: \mathbb{R}^{2n} \to \mathbb{R}^2 \) to the unit sphere \( S(\mathbb{R}^{2n}) = \{ x \in \mathbb{R}^{2n} : \sum_{j=1}^{2n} x_j^2 = 1 \} \) we obtain the numerical range map \( n_A \) associated with the pair \( A = (A_1, A_2) \); cf. Section 1. Hence, \( n_A: S(\mathbb{R}^{2n}) \to \mathbb{R}^2 \) is a smooth map from the \((2n - 1)\)-dimensional \( C^\infty \) manifold \( S(\mathbb{R}^{2n}) \) into \( \mathbb{R}^2 \). Its derivative \( n_A'(u) \) at a point \( u \in S(\mathbb{R}^{2n}) \) is therefore a linear map from the tangent space \( T_u S(\mathbb{R}^{2n}) \) of \( S(\mathbb{R}^{2n}) \) at \( u \) into \( \mathbb{R}^2 \). Moreover, \( T_u S(\mathbb{R}^{2n}) \) is the orthogonal complement \( \{ u \}^\perp \) of \( \text{span}(u) \) in \( \mathbb{R}^{2n} \). So, if \( u = (p, q) \) with \( p, q \in \mathbb{R}^n \), then the linear map \( n_A'(u) \) is the restriction of the linear map \( N_A'(p, q): \mathbb{R}^{2n} \to \mathbb{R}^2 \) to \( \{ (p, q) \}^\perp \). In particular, given \( u \in S(\mathbb{R}^{2n}) \), the linear map \( n_A'(u) \) has rank 0, rank 1, or rank 2.
DEFINITION 1. Let \( u \in \mathcal{S}(\mathbb{R}^{2n}) \).

(a) The point \( u \) is called a **stationary point** or **rank 0 critical point** of \( n_A \) if \( n_A'(u) = 0 \). The set of all stationary points of \( n_A \) is denoted by \( \Sigma_0(A) \).

(b) The point \( u \) is called a **rank 1 critical point** of \( n_A \) if \( n_A'(u) \) has rank 1; the set consisting of all such points is denoted by \( \Sigma_1(A) \).

(c) We call \( u \) a **regular point** of \( n_A \) if \( n_A'(u) \) has rank 2. The set of all regular points of \( n_A \) is denoted by \( \Sigma_2(A) \).

We can now formulate the result corresponding to (P1) in terms of the map \( n_A \). A point \( \lambda = (\lambda_1, \lambda_2) \in \mathbb{R}^2 \) is called a **joint eigenvalue** of \( A \) if there exists \( x \in \mathcal{S}(\mathbb{C}^n) \), called a **joint eigenvector** corresponding to \( \lambda \), such that \( A_1x = \lambda_1x \) and \( A_2x = \lambda_2x \).

**PROPOSITION 1.** Let \( A = (A_1, A_2) \) be a pair of Hermitian operators in \( \mathbb{C}^n \), and \( A = A_1 + iA_2 \). Then a vector \( x = p + iq \in \mathcal{S}(\mathbb{C}^n) \approx \mathcal{S}(\mathbb{R}^{2n}) \) is a joint eigenvector of \( A \) iff \( (p, q) \) is a stationary point of \( n_A \). The corresponding joint eigenvalue is \( n_A(p, q) \).

**Proof.** If \( (p, q) \) is a stationary point of \( n_A \), then Lemma 1 implies

\[
\langle Ax, h + ik \rangle + \langle A(h + ik), x \rangle = 0
\]

for all \( (h, k) \in \{(p, q)\}^\perp \). Let \( (h, k) \in \{(p, q)\}^\perp \) and \( v = h + ik \). From Equation (1) and its conjugate we have

\[
\langle Ax, v \rangle + \langle A^*x, v \rangle = 0 \quad \text{and} \quad \langle A^*x, v \rangle + \langle Ax, v \rangle = 0.
\]

Adding the equations in (2) gives \( \text{Re} \langle A_1x, v \rangle = 0 \). Since

\[
\{(p, q)\}^\perp - \{(h, k) : \text{Re} \langle x, h + ik \rangle = 0 \},
\]

it follows that \( A_1x \in J_n((p, q)^{\perp\perp}) = \text{span}(x) \). Hence, \( A_1x = \lambda_1x \) for some \( \lambda_1 \in \mathbb{R} \), that is, \( x = p + iq \) is an eigenvector of \( A_1 \) with eigenvalue \( \lambda_1 \). A similar argument applies to show that \( x \) is also an eigenvector of \( A_2 \) and hence, \( x \) is a joint eigenvector of \( A = (A_1, A_2) \). It is routine to check that \( (\lambda_1, \lambda_2) = n_A(p, q) \).

Conversely, if \( x = p + iq \in \mathcal{S}(\mathbb{C}^n) \) is an eigenvector of \( A_1 \), then (2) holds for all \( (h, k) \in \{(p, q)\}^\perp \) with \( v = h + ik \), as does (1). So, if \( x \) is a
joint eigenvector of \( \mathbf{A} \), then (1) holds for the Hermitian matrix \( \mathbf{A}_2 \) also, and \((p, q)\) is a stationary point of \( n_\mathbf{A} \).

So the image of the rank 0 points in \( S(\mathbb{R}^{2n}) \) under the numerical range map \( n_\mathbf{A} \) identifies a rather important subset of \( W(\mathbf{A}) \) [of \( J^{-1}_1(W(\mathbf{A})) \), to be precise], namely, the joint eigenvalues of \( \mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2) \). This suggests that a closer look at the differential structure of \( n_\mathbf{A} \) might lead to the identification of other natural subsets of \( W(\mathbf{A}) \).

**Proposition 2.** Let \( \mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2) \) be a pair of Hermitian operators in \( \mathbb{C}^n \) and \( \mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2 \). Suppose that \( x = p + iq \), with \( p, q \in \mathbb{R}^n \), is a point of \( S(\mathbb{C}^n) \) such that the derivative \( n'_\mathbf{A}(p, q) \) at \((p, q)\) has rank 2. Then \( (J_1 \circ n_\mathbf{A})(p, q) = \langle \mathbf{A}x, x \rangle \) belongs to the interior \( W(A)^0 \) of the numerical range \( W(A) \).

**Proof.** Let \( U \) be an open subset of \( \mathbb{R}^{2n-1} \), and let \( \varphi : U \to S(\mathbb{R}^{2n}) \) be a diffeomorphism onto an open neighborhood \( \varphi(U) \) of \((p, q)\). Let \( w \) be the element of the open set \( U \) such that \( \varphi(w) = (p, q) \). Then \( n_\mathbf{A} \circ \varphi : U \to \mathbb{R}^2 \) is a smooth map for which the derivative \( (n_\mathbf{A} \circ \varphi)'(w) \) of \( n_\mathbf{A} \circ \varphi \) at \( w \) has rank 2. Let \( F = (n_\mathbf{A} \circ \varphi)'(w) \). By the implicit function theorem, there exists an open set \( W \subseteq \mathbb{R}^{2n-1} \) and a diffeomorphism \( h : W \to U \) onto an open neighborhood \( V \) of \( w \) such that \( (F \circ h)(y_1, y_2, \ldots, y_{2n-1}) = (y_1, y_2) \) for all \( y \in W \). It follows that \( (n_\mathbf{A} \circ \varphi)(V) \) is an open subset of \( \mathbb{R}^3 \) contained in the set \( n_\mathbf{A}(S(\mathbb{R}^3)) = J^{-1}_1(W(\mathbf{A})) \), that is, \( (J_1 \circ n_\mathbf{A})(p, q) = \langle \mathbf{A}x, x \rangle \) is an interior point of \( W(\mathbf{A}) \).

We remark that the conclusion is also a consequence of the immersion theorem, [1, Theorem 3.5.7], which is perhaps not as well known as the implicit function theorem.

As a simple consequence of Proposition 2 we have the following result. Recall that the boundary \( \partial K \) of a subset \( K \) of \( \mathbb{R}^2 = \mathbb{C} \) is the set \( \overline{K} \cap \overline{K}^c \), where the bar indicates closure. The spectrum of \( \mathbf{A} \in \mathcal{L}(\mathbb{C}^n) \) is denoted by \( \sigma(\mathbf{A}) \). For the standard basis vectors in \( \mathbb{C}^n \) and \( \mathbb{R}^{2n} \) we use the notation \( e_j \) (\( j = 1, \ldots, n \)) and \( \mathbf{e}_j \) (\( j = 1, \ldots, 2n \)) respectively.

**Proposition 3.** Let \( \mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2) \) be a pair of Hermitian operators in \( \mathbb{C}^n \), and \( \mathbf{A} = \mathbf{A}_1 + i\mathbf{A}_2 \). Every element of \( J_1^{-1}(\partial W(\mathbf{A}) \cap \sigma(\mathbf{A})) \) is a joint eigenvalue of \( \mathbf{A} \).

**Proof.** Let \( \lambda = \alpha + ib \) be an element of \( \partial W(\mathbf{A}) \cap \sigma(\mathbf{A}) \), where \( \alpha, b \in \mathbb{R} \), and \((p, q) \in S(\mathbb{R}^{2n})\) satisfy \( \mathbf{A}(p + iq) = \lambda(p + iq) \). Then \( n_\mathbf{A}(p, q) = (\alpha, b) \) and \( n'_\mathbf{A}(p, q) \) has rank at most 1, by Proposition 2.
Choose a unitary operator $U \in \mathcal{L}(\mathbb{C}^n)$ such that $Ue_1 = p + iq$. Set $R = J_{n-1}^{-1} \circ U \circ J_n$. Then $(n_A \circ R)(\tilde{e}_1) = (a, b)$ and $(n_A \circ R)'(\tilde{e}_1)$ has rank at most 1. Note that $n_A \circ R$, which has domain $S(\mathbb{R}^{2n})$, is the numerical range map associated with the pair $(U^*A_1U, U^*A_2U)$. Moreover, $e_1$ is an eigenvector of $U^*AU$ with eigenvalue $\lambda$. Hence, it suffices to assume that $p = (1, 0, \ldots, 0)'$ (where $t$ denotes transpose) and $q = 0$, both being in $\mathbb{R}^n$, for then $p + iq = e_1$ is a joint eigenvector of the pair $(U^*A_1U, U^*A_2U)$ and so $Ue_1$ is a joint eigenvector of $A$.

Let $A_j = S_j + iT_j$, $j \in \{1, 2\}$, where each $S_j$ is real symmetric and each $T_j$ is real skew-symmetric. Then $Ae_1 = \lambda e_1$, together with the fact that $S_j$ and $T_j$ $(j = 1, 2)$ have real entries, implies that

$$\begin{align*}
(S_1 - T_2)e_1 &= \alpha e_1 & \text{and} & & (S_2 + T_1)e_1 &= \beta e_1.
\end{align*}$$

The tangent space to $S(\mathbb{R}^{2n})$ at $\tilde{e}_1$ is spanned by $\{\tilde{e}_2, \ldots, \tilde{e}_{2n}\}$. By Lemma 1 the matrix of $n_A'(\tilde{e}_1)$ with respect to this basis for the tangent space and the standard basis of $\mathbb{R}^2$ is the $2 \times (2n - 1)$ matrix (with real entries)

$$Q = 2 \begin{bmatrix}
(PS_1 e_1)' & (T_1 e_1)' \\
(PS_2 e_1)' & (T_2 e_1)'
\end{bmatrix},$$

where $P : \mathbb{R}^n \to \mathbb{R}^{n-1}$ is the projection onto the last $n - 1$ coordinates, and the elements $S_j e_1$ and $T_j e_1$ (of $\mathbb{C}^n$) are considered as points of $\mathbb{R}^n$, since all entries are real. The matrix $Q$ is assumed to have rank at most 1, and so the two vectors

$$\begin{bmatrix}
PS_1 e_1 \\
T_1 e_1
\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}
PS_2 e_1 \\
T_2 e_1
\end{bmatrix}$$

are linearly dependent in $\mathbb{R}^{2n-1}$. From (3) we have

$$\begin{align*}
(PS_1 - PT_2)e_1 &= 0 & \text{and} & & (PS_2 + PT_1)e_1 &= 0.
\end{align*}$$

The linear dependence of (4) and the equations (5) are satisfied only if $PS_1 e_1 = PS_2 e_1 = PT_1 e_1 = PT_2 e_1 = 0$, that is, only if $e_1$ is an eigenvector of $S_1, S_2, T_1, T_2$. Then $e_1$ is also an eigenvector of $A_j = S_j + iT_j$ for $j \in \{1, 2\}$. Because $A_1$ and $A_2$ are Hermitian, $T_1 e_1 = 0 = T_2 e_1$. Hence, $e_1$ is a joint eigenvector of $A$. ■
Remark 1.

(I) Note that by (P2), Proposition 3 applies in particular to all points \( \lambda \in \partial W(A) \) at which the boundary is not a differentiable arc.

(II) In the spirit of this paper, our proof of Proposition 3 is based on differential properties of the map \( n_A \) (i.e. on Lemma 1 and Proposition 2). It should be noted that Proposition 3 also follows, via different methods, from a result of R. Kippenhahn (cf. comments at the top of p. 201 in [S]) which states that if \( \lambda \in \partial W(A) \cap \sigma(A) \), then \( A \) is unitarily reducible, that is, there exists a unitary matrix \( U \) such that

\[
U^*AU = \begin{bmatrix} \lambda & 0 \\ 0 & B \end{bmatrix}.
\]

Then

\[
U^*A^*U = \begin{bmatrix} \lambda & 0 \\ 0 & B^* \end{bmatrix}.
\]

In the notation of the proof of Proposition 3, it follows that \( f = Ue_1 \neq 0 \) satisfies \( Af = \lambda f \) and \( A^*f = \bar{\lambda}f \), which implies that

\[
A_1f = \frac{\lambda + \bar{\lambda}}{2}f = af \quad \text{and} \quad A_2f = \frac{\lambda - \bar{\lambda}}{2i}f = bf.
\]

So, \((a, b)\) is a joint eigenvalue of \( A \).

It may be of interest to note that the result of Kippenhahn used above follows from an earlier theorem in his paper [8, Theorem 12] which states that if \( A \) is not unitarily reducible, then \( \sigma(A) \subseteq W(A)^\circ \). Noting that always \( \sigma(A) \subseteq W(A) \), we can deduce Kippenhahn's theorem directly from Proposition 3 above, since the existence of a joint eigenvalue of \( A \) easily implies that \( A = A_1 + iA_2 \) is unitarily reducible.

Proposition 2 suggests that the boundary \( \partial W(A) \) may be the image under \( J_1 \circ n_A \) of the points in \( S(\mathbb{R}^{2n}) \) where \( n_A' \) has rank 1; it is easy to produce examples where this fails. The problem, in general, is that a point of \( W(A) \) can be the image (under \( J_1 \circ n_A \)) of distinct points from \( S(\mathbb{R}^{2n}) \) at which \( n_A' \) has different ranks. This occurs, for example, when there are \( A \)-invariant subspaces and we have a decomposition \( A = B \oplus C \). The simplest such
A-invariant subspace is the joint eigenspace corresponding to a joint eigenvalue of A.

**Example 1.** Let

\[
B_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad B_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};
\]

these are Hermitian operators in \( \mathbb{C}^2 \) with no joint invariant subspace. The numerical range \( W(B) \) of \( B = B_1 + iB_2 \) is the closed unit disc \( \mathbb{D} \) in \( \mathbb{C} \). Let

\[
A_1 = \begin{bmatrix} 0 & 0 \\ 0 & B_1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 2 & 0 \\ 0 & B_2 \end{bmatrix},
\]

considered as Hermitian operators in \( \mathbb{C}^3 = \mathbb{C} \oplus \mathbb{C}^2 \). Then the numerical range \( W(A) \) of \( A = A_1 + iA_2 \) is the convex hull in \( \mathbb{R}^2 \) of \( \{0 + 2i\} \cup \mathbb{D} \), and the numerical range map \( n_A \) of \( A \) is given by

\[
n_A(p, q) = \left( p_2^2 + q_2^2 - p_3^2 - q_3^2, 2\left(p_1^2 + q_1^2 + p_2p_3 + q_2q_3\right)\right),
\]

\((p, q) \in S(\mathbb{R}^3 \times \mathbb{R}^3)\).

The critical points of \( n_A \) can be found by direct computation. For every \((p, q) \in S(\mathbb{R}^3 \times \mathbb{R}^3)\), one takes a basis \( \{u_1, \ldots, u_5\} \) of \( T_{(p, q)} S(\mathbb{R}^6) = \{(p, q)\}^\perp \) and computes the dimension of \( n_A'(p, q)(\{u_1, \ldots, u_5\}) \), where \( n_A'(p, q) \) is the Jacobi matrix of \( n_A \) at \((p, q)\), i.e.

\[
n_A'(p, q) = \begin{bmatrix}
0 & 2p_2 & -2p_3 & 0 & 2q_2 & -2q_2 \\
4p_1 & 2p_3 & 2p_2 & 4q_1 & 2q_3 & 2q_2
\end{bmatrix}.
\]

This yields that \( \{(av, bv) : v \in S(\mathbb{R}^3), v_1 = 0, (a, b) \in S(\mathbb{R} \times \mathbb{R}) \} \subseteq \Sigma_1(A) \). In particular,

\[
u = \begin{bmatrix}
0 \\
\sqrt{2}/2 \\
2 \\
\sqrt{2}/2 \\
2
\end{bmatrix}, \begin{bmatrix}
0 \\
0
\end{bmatrix}
\]
is a rank 1 critical point of \( n_\Lambda \), and \( n_\Lambda(u) = (0, 1) \in J_1^{-1}(W(\Lambda)^c) \). However, \((0, 1)\) is also the image under \( n_\Lambda \) of the regular point
\[
\begin{pmatrix}
\frac{\sqrt{2}}{2} & 0 \\
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
-\frac{1}{2}
\end{pmatrix}.
\]

Perhaps more surprising is that, as the following examples show, such points may even occur in the absence of \( \Lambda \)-invariant subspaces.

**Example 2.** Consider the Hermitian operators
\[
A_1 = \begin{bmatrix}
0 & \frac{1}{2} & 0 \\
\frac{1}{2} & 0 & \frac{1}{2} \\
0 & \frac{1}{2} & 0
\end{bmatrix}
\quad \text{and} \quad
A_2 = \begin{bmatrix}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & -1
\end{bmatrix}
\]
on \( \mathbb{C}^3 \). This pair \( \Lambda = (A_1, A_2) \) has no joint eigenvalues and therefore no joint invariant subspace. The numerical range \( W(\Lambda) \) of \( \Lambda = A_1 + iA_2 \) is the ellipse \( \{ \lambda \in \mathbb{C} : 2(\text{Re}\, \lambda)^2 + (\text{Im}\, \lambda)^2 \leq 1 \} \), and the numerical range map \( n_\Lambda \) is given by
\[
n_\Lambda(p, q) = (p_2(\lambda_1 + p_3) + q_2(q_1 + q_3), p_1^2 - p_3^2 + q_1^2 - q_3^2).
\]
\[
(p, q) \in S(\mathbb{R}^3 \times \mathbb{R}^3).
\]

A direct calculation shows that \( \Sigma_1(\Lambda) \) consists of the two sets \( S_1 = \{(av_1, bv_1) : v \in S(\mathbb{R}^3), v_2^2 = 2v_1v_3, (a, b) \in S(\mathbb{R} \times \mathbb{R}) \} \) and \( S_2 = \{(av, bv) : v \in S(\mathbb{R}^3), v_1 = -v_3, (a, b) \in S(\mathbb{R} \times \mathbb{R}) \} \). All other points \( u \in S(\mathbb{R}^6) \) belong to \( \Sigma_2(\Lambda) \). Now, the image of \( S_1 \) under \( n_\Lambda \) is precisely the boundary of \( J_1^{-1}(W(\Lambda)) \), whereas \( n_\Lambda(S_2) = \{(0, 0)\} \). But the point \((0, 0) \in \mathbb{R}^2\) is also the image under \( n_\Lambda \) of the regular point
\[
\begin{pmatrix}
\frac{\sqrt{2}}{2} & 0 \\
0 & 0 \\
0 & \frac{\sqrt{2}}{2}
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
\frac{1}{2}
\end{pmatrix}.
\]

**Example 3.** Let
\[
A_1 = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & \sqrt{2}
\end{bmatrix}
\quad \text{and} \quad
A_2 = \begin{bmatrix}
0 & i/2 & 0 \\
-i/2 & 0 & i/2 \\
0 & -i/2 & 0
\end{bmatrix}.
\]
Again, this pair $A = (A_1, A_2)$ of Hermitian operators on $\mathbb{C}^3$ has no joint eigenvalues and so no joint invariant subspace. The numerical range $W(A)$ is the convex hull in $\mathbb{R}^2$ of the curve

$$\{(u, v) \in \mathbb{R}^2: -4\sqrt{2}uv^2 + 13u^2v^2 + 4u^4 - 7\sqrt{2}u - 10\sqrt{2}u^3 - 16v^2 + 2 + 18u^2 + 32v^4 = 0\}$$

(see also [8, Example 6.3]). The numerical range map $n_A$ is given by

$$n_A(p, q) = \left(\sqrt{2}(p_3^2 + q_3^2), q_2(p_1 - p_3) - p_2(q_1 - q_3)\right),$$

$$(p, q) \in S(\mathbb{R}^2 \times \mathbb{R}^2).$$

Direct calculations yield that $u \in S(\mathbb{R}^6)$ is a rank 1 critical point of $n_A$ iff $u$ belongs to one of the sets $S_1 = \{(v, w) \in S(\mathbb{R}^3 \times \mathbb{R}^3): v_3 = w_3 = 0\}$ or $S_2 = \{(v, w) \in S(\mathbb{R}^3 \times \mathbb{R}^3): v_1 = v_2 = w_1 = w_2 = 0\}$ or

$$S_3 = \left\{(v, w) \in S(\mathbb{R}^3 \times \mathbb{R}^3): (v, w) = J_3^{-1}\left(\alpha \begin{bmatrix} x_1 \\ \alpha x_2 \\ x_3 \end{bmatrix}\right)\right\},$$

for some $x_j \in \mathbb{R}$ ($j = 1, 2, 3$),

$$x_1 \neq 0, x_1 - x_3 - \frac{x_2^2}{x_1} = 0, \alpha \in S(\mathbb{C}).$$

In particular, the point

$$u = \begin{bmatrix} \sqrt{2}/2 \\ 0 \\ \sqrt{2}/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

belongs to $\Sigma_1(A)$. However, $n_A(u) = (\sqrt{2}/2, 0)$, a point in the interior of $J_1^{-1}(W(A))$, is also the image under $n_A$ of the regular point

$$v = \begin{bmatrix} 0 \\ \sqrt{2}/2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$
In each of these examples, the image under \( n_\Lambda \) of all the critical points coincides with a certain algebraic curve \( C(\Lambda) \) (using the notation of \([5]\)) together with any parts of \( J_1^{-1}(\partial W(\Lambda)) \) which do not already belong to \( C(\Lambda) \). The importance of the curves \( C(\Lambda) \) in determining the numerical range of a given operator \( \Lambda = \Lambda_1 + i\Lambda_2 \) was first observed by R. Kippenhahn \([8, \text{Theorem 10}]\) and, independently, by F. D. Murnaghan \([11]\). They showed that \( W(\Lambda) \) coincides with the convex hull (in \( \mathbb{R}^2 \)) of the curve \( C(\Lambda) \) and then studied \( C(\Lambda) \) for further information on \( \Lambda \) (see also \([5]\)). Since other examples exhibit the same phenomena, this suggests the following conjecture.

**Conjecture.** Let \( \Lambda = (\Lambda_1, \Lambda_2) \) be a pair of Hermitian operators in \( \mathbb{C}^n \), and \( \Lambda = \Lambda_1 + i\Lambda_2 \). If \( \Lambda \) has no joint invariant subspace, then

\[
n_\Lambda(\Sigma_1(\Lambda)) = C(\Lambda) \cup J_1^{-1}(\partial W(\Lambda)).
\]

(*)

For \( n = 2 \), the Conjecture is true. To see this we note that \( \Lambda \) has no joint invariant subspaces (or, equivalently, no joint eigenvalues) precisely when \( \Lambda_1 \Lambda_2 \neq \Lambda_2 \Lambda_1 \) (see \([6, \text{Proposition 10}]\)), that is, if \( \Lambda = \Lambda_1 + i\Lambda_2 \) is not a normal operator in \( \mathbb{C}^2 \). In this case \( C(\Lambda) \) is an ellipse (see, e.g., \([8, \text{Section 7}]\)) and therefore \( C(\Lambda) = J_1^{-1}(\partial W(\Lambda)) \). Then (*) follows from Proposition 4 below (a result of independent interest which determines a natural decomposition of \( W(\Lambda) \)).

**Proposition 4.** Let \( \Lambda = (\Lambda_1, \Lambda_2) \) be a noncommuting pair of Hermitian operators in \( \mathbb{C}^2 \) and \( \Lambda = \Lambda_1 + i\Lambda_2 \). Then \( \Sigma_0(\Lambda) = \emptyset \). Moreover,

\[
n_\Lambda^{-1}(J_1^{-1}(\partial W(\Lambda))) = \Sigma_1(\Lambda) \quad \text{and} \quad n_\Lambda^{-1}(J_1^{-1}(W(\Lambda)^o)) = \Sigma_2(\Lambda).
\]

Before we establish Proposition 4, we collect some relevant facts which hold for all \( n \geq 1 \). Recall that if \( \lambda \in \mathbb{R}^2 \) belongs to \( n_\Lambda(\Sigma_2(\Lambda)) \) then, by Proposition 2, \( \lambda \in J_1^{-1}(W(\Lambda)^o) \). In particular, \( W(\Lambda)^o \neq \emptyset \) whenever \( \Sigma_0(\Lambda) \neq \emptyset \). If \( \lambda \in J_1^{-1}(W(\Lambda)) \) and \( n_\Lambda^{-1}(\lambda) \subseteq \Sigma_2(\Lambda) \), then we call \( \lambda \) a regular value of \( n_\Lambda \). Thus, a regular value of \( n_\Lambda \) necessarily lies in \( J_1^{-1}(W(\Lambda)^o) \).

**Fact 1.** The set of regular values of \( n_\Lambda \) is dense in \( J_1^{-1}(W(\Lambda)) \); this follows from Sard's theorem \([1, \text{Theorem 3.6.3}]\).

**Fact 2.** \( (J_1 \circ n_\Lambda)(\Sigma_2(\Lambda)) \) is dense in \( W(\Lambda) \) and is contained in \( W(\Lambda)^o \); cf. Fact 1.
FACT 3. \( n_A(\Sigma_2(A)) \) is an open subset of \( \mathbb{R}^2 \).

Suppose, in addition, that \( A_1 \) and \( A_2 \) have no joint eigenvalues.

FACT 4. \( S(\mathbb{R}^2^n) = \Sigma_1(A) \cup \Sigma_2(A) \), by Proposition 1.

FACT 5. \( \Sigma_2(A) \subseteq n_A^{-1}(J^{-1}_1(W(A)^*)), \) by Proposition 2.

FACT 6. \( n_A^{-1}(J^{-1}_1(\partial W(A))) \subseteq \Sigma_1(A) \) by Facts 4 and 5.

Proof of Proposition 4. Since \( A_1 A_2 \neq A_2 A_1 \), it follows by [6, Proposition 10] that \( A_1 \) and \( A_2 \) have no joint eigenvalues. Hence, by Proposition 1, \( \Sigma_0(A) = \emptyset \).

The numerical range map \( n_A \) is invariant under unitary transformations, so, after making a unitary change of basis for \( \mathbb{C}^2 \) if necessary, we may assume that

\[
\begin{bmatrix}
1 \\
0
\end{bmatrix}
\]

is an eigenvector for \( A = A_1 + iA_2 \), so that

\[
A = \begin{bmatrix}
\alpha & \beta \\
0 & \gamma
\end{bmatrix}
\]

for some complex numbers \( \alpha, \beta, \gamma \). Necessarily \( \beta \neq 0 \); otherwise \( A \) and \( A^* \) commute and then \( A_1A_2 = A_2A_1 \) follows, which contradicts our assumptions. The desired equalities are true iff they are true with \( A \) replaced by \( \lambda A \), for \( \lambda \neq 0 \). So, replacing \( A \) by \( B^{-1}A \), we may assume that \( \beta = 1 \). Then, with \( \mathcal{M}_\lambda \) denoting \( J^1 \circ N_\lambda \circ J_2^{-1} \), so that \( \mathcal{M}_\lambda(x) = \langle Ax, x \rangle \) for all \( x \in \mathbb{C}^2 \), we have

\[
\mathcal{M}_\lambda(x) = \alpha|x_1|^2 + \bar{x}_1 x_2 + \gamma|x_2|^2. \quad x = (x_1, x_2) \in \mathbb{C}^2.
\]

To simplify computations, we reduce the problem to two real dimensions. To do this, note that for any \( \lambda \in \mathbb{C} \) with \( |\lambda| = 1 \), the equality \( \mathcal{M}_\lambda(x) = \mathcal{M}_\lambda(\bar{x}) \) holds for all \( x \in \mathbb{C}^2 \). Any vector \( x = (x_1, x_2) \) in \( S(\mathbb{C}^2) \) with \( x_2 \neq 0 \) can be written in the form \( \lambda(|z|^2 + 1)^{-1/2}(z, 1) \) for some \( z \in \mathbb{C} \) and some \( \lambda \in \mathbb{C} \) such that \( |\lambda| = 1 \), so that

\[
\mathcal{M}_\lambda(x) = \mathcal{M}_\lambda((|z|^2 + 1)^{-1/2}(z, 1)). \quad (6)
\]
If \( x_2 = 0 \), then \( Ax = \alpha x \) and so \( \mathcal{M}_A(x) = \alpha \). Moreover, we have the identities \( \mathcal{M}_A(\{x \in S(C^2) : x_2 \neq 0\}) = W(A) \) and \( \mathcal{M}_A(\{x \in S(C^2) : x_2 = 0\}) = \{\alpha\} \).

For each \( z \in \mathbb{C} \), let \( \nu(z) \in \mathbb{C} \) be equal to the right-hand side of (6), in which case

\[
\nu(z) = \alpha + \frac{z + \gamma - \alpha}{|z|^2 + 1}, \quad z \in \mathbb{C}.
\]

Since \( \nu \) takes the value \( \alpha \) at the point \(-\overline{\gamma - \alpha}\), it is routine to check that

\[
\nu(\mathbb{C}) = \mathcal{M}_A(S(C^2)) = W(A).
\]

The problem thereby reduces to the calculation of the critical points of the map \( f^{-1}_1 \circ \nu \circ f_1 \) by the following lemma.

**Lemma 2.** The point \( \lambda \in f^{-1}_1(W(A)) \) is a regular value of \( n_A \) if it is a regular value of the map \( f^{-1}_1 \circ \nu \circ f_1 \).

**Proof.** Recall the complex projective line \( \mathbb{C}P^1 \) is the set of all complex lines in \( \mathbb{C}^2 \). It may be identified with the (compact) quotient manifold \( S(\mathbb{R}^4)/\sim \), where \( f_2^{-1}x \sim f_2^{-1}y \) for \( x, y \in S(C^2) \), if there exists \( \rho \in \mathbb{C} \) with \( |\rho| = 1 \) such that \( x = \rho y \). Let \( \pi : S(C^2) \rightarrow \mathbb{C}P^1 \) be the projection given by \( \pi(x) = \text{span}(x) \) for \( x \in S(C^2) \), and \( f_2 : S(\mathbb{R}^4) \rightarrow S(C^2) \) be the restriction of \( f_2 \) from \( \mathbb{R}^4 \) to \( S(\mathbb{R}^4) \). Then \( \tilde{\pi} = \pi \circ f_2 \) is a submersion from \( S(\mathbb{R}^4) \) onto \( \mathbb{C}P^1 \), where \( \mathbb{C}P^1 \) is considered here as a real two-dimensional manifold.

The map \( \psi : z \mapsto \text{span}((z, 1)) \) maps \( \mathbb{C} \) homeomorphically onto \( \mathbb{C}P^1 \setminus \{\text{span}(e_1)\} \). We can interpret \( \mathbb{C}P^1 \) as a one-point compactification of \( \mathbb{C} \), in which the point \( \infty \) at infinity is given by \( \text{span}(e_1) \), so that \( \psi \) is the restriction to \( \mathbb{C} \) of a homeomorphism from \( \mathbb{C} \cup \{\infty\} \) onto \( \mathbb{C}P^1 \). The set \( \pi^{-1}(\infty) \) is equal to \( \{\rho e_1 : \rho \in \mathbb{C}, |\rho| = 1\} \).

Let \( \tilde{\nu} : \mathbb{C}P^1 \rightarrow \mathbb{C} \) be the map induced by \( \nu \) under this correspondence, that is, \( (\tilde{\nu} \circ \psi)(z) = \nu(z) \) for all \( z \in \mathbb{C} \), and \( \tilde{\nu}(\text{span}(e_1)) = (f_1^{-1} n_A f_2^{-1})(e_1) = \alpha \). Then \( n_A(x) = (f_1^{-1} \circ \tilde{\nu} \circ \tilde{\pi})(x) \) for all \( x \in S(\mathbb{R}^4) \).

Because \( \tilde{\pi} \) is a submersion, it follows that rank \( n_A(x) = 2 \) iff rank \( \tilde{\nu}'(\tilde{\pi}(x)) = 2 \) for any \( x \in S(\mathbb{R}^4) \), where we regard \( \tilde{\nu} \) as a map between the real two-dimensional manifolds \( \mathbb{C}P^1 \) and \( \mathbb{R}^2 \).

If \( x = (x_1, x_2) \) belongs to \( S(C^2) \setminus \pi^{-1}(\infty) \), then for \( z = x_2^{-1}x_1 \in \mathbb{C} \) we have \( \psi(z) = \pi(x) \) and \( \tilde{\nu}'(\pi(x)) = (f_1^{-1} \circ \nu \circ f_1)'(f_2^{-1}z) \). Hence, \( n_A \) has rank 2 at \( f_2^{-1}x \) iff \( (f_1^{-1} \circ \nu \circ f_1)' \) has rank 2 at the corresponding point \( f_1^{-1}z \) in \( \mathbb{R}^2 \). It remains to check that \( n_A \) has rank 2 at \( \tilde{\pi}^{-1}(\infty) = f_2^{-1}(\pi^{-1}(\infty)) \). Let
\( \rho = \rho_1 + i\rho_2 \in \mathbb{C} \) have modulus one, where \( \rho_1, \rho_2 \in \mathbb{R} \). The tangent space \((J^{-1}_2(\rho e_1))^\perp \) to \( S(\mathbb{R}^4) \) at \((\rho_1, 0, \rho_2, 0) = J_1'(\rho e_1) \) is spanned by the vectors \((h, k) = (-a\rho_2, u, a\rho_1, v) \) in \( \mathbb{R}^4 \) as \( a, u, v \) vary in \( \mathbb{R} \). An appeal to Lemma 1 gives

\[
J_1(n'_A(\rho_1, 0, \rho_2, 0)(h, k)) = \langle A\rho e_1, h + ik \rangle + \langle A^*\rho e_1, h + ik \rangle
\]

\[
= \langle \rho \begin{bmatrix} \alpha \\ 0 \end{bmatrix}, h + ik \rangle + \langle \rho \begin{bmatrix} \bar{\alpha} \\ 1 \end{bmatrix}, h + ik \rangle
\]

\[
= \rho(u + iv) = J_1\begin{bmatrix} \rho_1 & \rho_2 \\ -\rho_2 & \rho_1 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix},
\]

showing that \( n'_A \) necessarily has rank 2 at every such point \( J_1'(\rho e_1) \).

It follows that for \( \lambda \in J^{-1}_1(W(A)) \), the derivative \( n'_A \) has rank 2 at every point in the preimage \( n^{-1}_A(\lambda) \) iff \((J^{-1}_1 \circ \nu \circ J_1)' \) has rank 2 at every point of \((J^{-1}_1 \circ \nu^{-1} \circ J_1)(\lambda) \). This proves Lemma 2.

Returning to the proof of Proposition 4, we note that

\[
\text{Re } \nu(u + iv) = \text{Re } \alpha + \frac{u - \text{Re}(\alpha - \gamma)}{u^2 + v^2 + 1} \quad \text{and} \quad \text{Im } \nu(u + iv) = \text{Im } \alpha + \frac{-v + \text{Im}(\alpha - \gamma)}{u^2 + v^2 + 1}
\]

for all \( u, v \in \mathbb{R} \), and so

\[
(J^{-1}_1 \circ \nu \circ J_1)'(u, v) = (u^2 + v^2 + 1)^{-2}
\]

\[
\times \begin{bmatrix} v^2 - u^2 + 1 + 2u \text{Re}(\alpha - \gamma) & 2v[ -u + \text{Re}(\alpha - \gamma)] \\ 2u[ v + \text{Im}(\alpha - \gamma)] & v^2 - u^2 - 1 + 2v \text{Im}(\alpha - \gamma) \end{bmatrix}.
\]

This matrix has determinant \((u^2 + v^2 + 1)^{-3}(u^2 + v^2 - 1 - 2u \text{Re}(\alpha - \gamma) + 2v \text{Im}(\alpha - \gamma)) \), and so the set of critical points of \( J^{-1}_1 \circ \nu \circ J_1 \) is precisely the circle \([u - \text{Re}(\alpha - \gamma)]^2 + [v + \text{Im}(\alpha - \gamma)]^2 = 1 + |\alpha - \gamma|^2 \), i.e., the set of critical points of \( \nu \) is the circle \(|z - (\alpha - \gamma)|^2 = 1 + |\alpha - \gamma|^2 \). The map \( \nu \) is at most two to one, since for each \( a, b \in \mathbb{R} \) the preimage \( \nu^{-1}(a + ib) \) is the intersection of the loci \( \text{Re } \nu(u + iv) = a \) and \( \text{Im } \nu(u + iv) = b \), which are distinct circles. In fact, \( \nu \) takes equal values at points related by inversion in the circle of critical points, i.e.,

\[
\nu(z) = \nu(w) \quad \text{if} \quad (w - (\alpha - \gamma))(z - (\alpha - \gamma)) = 1 + |\alpha - \gamma|^2.
\]
This may be verified by straightforward calculation. Note also that \( \nu(\alpha - \gamma) = \nu(\alpha - \gamma) \), corresponding to the fact that inversion in the circle interchanges its center with the point \( \infty \). Let \( \Gamma = \{ z \in \mathbb{C} : |z - (\alpha - \gamma)|^2 \leq 1 + |\alpha - \gamma|^2 \} \) be the closed disc bounded by the set \( \partial \Gamma \) of critical points of \( \nu \).

Then \( \nu(\partial \Gamma) \cap \nu(\mathbb{C} \setminus \partial \Gamma) = \emptyset \), and \( \nu \) is injective when restricted to each of the two components of \( \mathbb{C} \setminus \partial \Gamma \).

We claim that \( \nu \mid_{\partial \Gamma} \) is also injective. This can be seen most easily by parametrizing \( \partial \Gamma \) as the circle \( z = (\alpha - \gamma) + re^{i\theta} \) where \( r = \sqrt{1 + |\alpha - \gamma|^2} \). Let \( \rho(\theta) = \nu((\alpha - \gamma) + re^{i\theta}) \). Then

\[
\rho(\theta) = \alpha + re^{-i\theta} \left[ \left| (\alpha - \gamma) + re^{i\theta} \right|^2 + 1 \right]^{-1} = \alpha + \left[ e^{i2\theta}(\alpha - \gamma) + 2re^{i\theta} + (\alpha - \gamma) \right]^{-1}.
\]

Now, if \( \rho(\theta) = \rho(\psi) \), then \( (e^{i2\theta} - e^{i2\psi})(\alpha - \gamma) = 2r(e^{i\psi} - e^{-i\theta}) \), and so either \( e^{i\theta} = e^{i\psi} \) or \( (e^{i\theta} + e^{i\psi})(\alpha - \gamma) = -2r \). But in the latter case we would have \( 2 > |e^{i\theta} + e^{i\psi}| = 2r|\alpha - \gamma|^{-1} > 2 \), which is impossible. Hence \( \nu \mid_{\partial \Gamma} \) is injective. It follows that \( \nu \) maps the compact set \( \Gamma \) homeomorphically onto \( W(A) \), so that \( \nu(\Gamma^\circ) = W(A)^o \) and \( \nu(\partial \Gamma) = \partial W(A) \). Since \( \partial \Gamma \) is the set of critical points for \( \nu \) and the regular values of \( \nu \mid_{\partial \Gamma} \) coincide, every point of \( W(A)^o \) is a regular value of \( \nu \), that is, \( n^{-1}_A(J^{-1}_1(W(A)^o)) \subseteq \Sigma_2(A) \). By Fact 5, the equality \( n^{-1}_A(J^{-1}_1(W(A)^o)) = \Sigma_2(A) \) holds; hence \( n^{-1}_A(J^{-1}_1(\partial W(A))) = \Sigma_2(A) \) by Fact 4.

**Remark 2.** We note that \( \Sigma_2(A) = n^{-1}_A(J^{-1}_1(W(A)^o)) \) actually holds for any pair \( A = (A_1, A_2) \) of Hermitian operators in \( \mathbb{C}^2 \). Indeed, in case \( A_1 A_2 = A_2 A_1 \), this is immediate from Proposition 4. If \( A_1 A_2 = A_2 A_1 \), then \( W(A) \) is a line segment (possibly a single point), and so \( W(A)^o = \emptyset \). Proposition 2 yields that also \( \Sigma_2(A) = n^{-1}_A(J^{-1}_1(W(A)^o)) = \emptyset \).

The above results lead to a characterization of commutativity for pairs of Hermitian matrices in \( \mathbb{C}^2 \) in terms of differential properties of \( n_A \).

**Proposition 5.** Let \( A = (A_1, A_2) \) be a pair of Hermitian operators in \( \mathbb{C}^2 \), and \( A = A_1 + iA_2 \). Then \( A_1 A_2 = A_2 A_1 \) if and only if \( n^{-1}_A(J^{-1}_1(\partial W(A))) \neq \Sigma_1(A) \).

**Proof.** If \( A_1 A_2 = A_2 A_1 \), then \( W(A) \) is a line segment (possibly a single point) and so \( \partial W(A) = W(A) \). Hence, \( n^{-1}_A(J^{-1}_1(\partial W(A))) = S(\mathbb{R}^4) \). Since
DIFFERENTIAL PROPERTIES

$n_A(\Sigma_0(A)) \neq \emptyset$, being the joint eigenvalues of $A = (A_1, A_2)$, we must have $\Sigma_0(A) \neq \emptyset$ and so $\Sigma_1(A) = S(\mathbb{R}^4) = n_A^{-1}(J^{-1}_1(\partial W(A)))$. The converse is a routine consequence of Proposition 4.

**Remark 3.**

(I) The criterion in Proposition 5, based on the differential structure of $n_A$, is different to other known criteria for commutativity of Hermitian operators in $\mathbb{C}^2$, such as those given in [3, Corollary 3.7; 6, Proposition 10; 9, p. 232, 10], for example.

(II) Propositions 4 and 5 fail for $n \geq 3$. Indeed, consider the Hermitian operators

$$A_1 = \begin{bmatrix} 0 & 0 \\ 0 & B_1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & B_2 \end{bmatrix}$$

on $\mathbb{C}^3$, where $B_1, B_2$ are the two Hermitian operators on $\mathbb{C}^2$ given in Example 1. Then $W(A) = \mathbb{D}$, the closed unit disc in $\mathbb{C}$, and $\lambda = (0, 1) \in J^{-1}_1(\partial W(A))$ is a joint eigenvalue of $A$. Since $e_1 \in \mathbb{C}^3$ is a joint eigenvector of $A$ corresponding to $\lambda$, we see that $x = J^{-1}_1(e_1) \in n_A^{-1}(J^{-1}_1(\partial W(A)))$. By Proposition 1 we have $n_A'(x) = 0$ and so rank $n_A'(x) = 0$, showing that $x \notin \Sigma_1(A)$. Hence, $n_A^{-1}(J^{-1}_1(\partial W(A))) \neq \Sigma_1(A)$. But $A_1A_2 \neq A_2A_1$. This shows that Proposition 5 fails. Proposition 4 also fails for this example, on two accounts: firstly, because $\Sigma_0(A) \neq \emptyset$, and secondly, because $n_A^{-1}(J^{-1}_1(\partial W(A))) \neq \Sigma_1(A)$. The reason is that, for $n \geq 3$, the condition $A_1A_2 \neq A_2A_1$ no longer implies $\Sigma_0(A) = \emptyset$ as it does for $n = 2$. So it is natural to ask if Proposition 4 remains valid in $\mathbb{C}^n$ if the hypothesis $A_1A_2 \neq A_2A_1$ is replaced by the stronger assumption of the Conjecture, i.e. that $A_1$ and $A_2$ have no (nontrivial) joint invariant subspace.

Even this fails. Indeed, consider the pair $A = (A_1, A_2)$ of Example 2, where $n = 3$. Then certainly $\Sigma_0(A) = \emptyset$ holds (which, for $n = 3$, is equivalent to $A_1$ and $A_2$ having no joint invariant subspace). However, $n_A^{-1}(J^{-1}_1(\partial W(A))) \neq \Sigma_1(A)$, since, in the notation of Example 2, $n_A^{-1}(J^{-1}_1(\partial W(A))) = S_1$ but $\Sigma_1(A) = S_1 \cup S_2$ (with $S_1$ and $S_2$ nonempty and disjoint). This shows that Proposition 4 really is particular to $n = 2$. Any analogue to higher dimensions will probably need to involve the curves $C(A)$.

REFERENCES


Received 10 September 1995; final manuscript accepted 13 January 1997