J. Austral. Math. Soc. (Series A) **65** (1998), 267–283

ON THE NUMERICAL RANGE MAP

M. JOSWIG and B. STRAUB

(Received 25 November 1997; revised 26 May 1998)

Communicated by J. A. Hillman

Abstract

Let $A \in \mathscr{L}(\mathbb{C}^n)$ and A_1, A_2 be the unique Hermitian operators such that $A = A_1 + iA_2$. The paper is concerned with the differential structure of the numerical range map $n_A : x \longmapsto (\langle A_1x, x\rangle, \langle A_2x, x\rangle)$ and its connection with certain natural subsets of the numerical range $W(A)$ of A. We completely characterize the various sets of critical and regular points of the map n_A as well as their respective images within $W(A)$. In particular, we show that the plane algebraic curves introduced by R. Kippenhahn appear naturally in this context. They basically coincide with the image of the critical points of n_A.

1991 *Mathematics subject classification* (*Amer. Math. Soc.*): 15A22, 15A60.

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 $\|\cdot\|$. By $\mathscr{L}(\mathbb{C}^n)$ we denote the space of all linear operators of $\mathbb{C}^n$ into itself. Associated with each operator $A \in \mathscr{L}(\mathbb{C}^n)$ is the set

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

where $\mathbb{S}(\mathbb{C}^n) = \{x \in \mathbb{C}^n \mid \|x\| = 1\}$, known as the *numerical range* or the *field of values* of A. This set has been studied extensively (see, for example, [6, 7, 8, 10, 12, 13, 14]) for it reflects certain important properties of A.

One of the main results concerning the geometrical shape of $W(A)$ is due to Kippenhahn [10] (see also [13] and [6]). His investigations in 1951 revealed that

The first and the second author gratefully acknowledge partial and full support, respectively, by the Australian Research Council; this work was carried out at the University of New South Wales.

the numerical range of A coincides with the convex hull of a certain plane algebraic curve C_A associated with A.

On the other hand, the numerical range $W(A)$ is the image of the *Rayleigh quotient* $R_A : x \longmapsto \dfrac{\langle Ax, x\rangle}{\langle x, x\rangle}$, for $x \neq 0$, of A. As $W(A) = R_A(\mathbb{S}(\mathbb{C}^n))$, this yields, for example, that $W(A)$ is always closed and bounded (see, for example, [12, Section 6.6]). However, more attention has been given to the Rayleigh quotient for its differential properties. It is well-known that R_A has a stationary value if and only if the operator A and its adjoint A^* have an eigenvector in common. In particular, all eigenvectors of A occur in this way if and only if the operator A is normal (see, for example, [14, Section III.18]). In this case, the set of stationary values of R_A coincides with the curve C_A associated with A as both consist precisely of the n eigenvalues of A. This suggests a possible link between the differential structure of R_A and the geometrical shape of $W(A)$, especially the curve C_A.

The aim of this paper is to investigate this connection for an arbitrary (that is, not necessarily normal) operator $A \in \mathscr{L}(\mathbb{C}^n)$. As in [8], we consider the *numerical range map* $n_A : \mathbb{S}(\mathbb{R}^{2n}) \longrightarrow \mathbb{R}^2$ of A which is given by

$$n_A(p,q) = (\langle A_1(p+iq), p+iq\rangle, \langle A_2(p+iq), p+iq\rangle) \quad \text{for all } (p,q) \in \mathbb{S}(\mathbb{R}^{2n}),$$

Here, A_1 and A_2 are the uniquely determined Hermitian operators in $\mathscr{L}(\mathbb{C}^n)$ such that $A = A_1 + iA_2$. By identifying a point $(p,q) \in \mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^{2n}$ with the vector $p + iq \in \mathbb{C}^n$, the map n_A can be viewed as the Rayleigh quotient of A restricted to $\mathbb{S}(\mathbb{C}^n)$.

Following the approach in [8], we split the domain $\mathbb{S}(\mathbb{R}^{2n})$ of n_A into the three subsets $\Sigma_0(A)$, $\Sigma_1(A)$ and $\Sigma_2(A)$. Here, $\Sigma_j(A)$ is the set of all points $(p,q) \in \mathbb{S}(\mathbb{R}^{2n})$ such that the derivative $n'_A(p,q)$ of n_A at (p,q), which is a linear map from the $(2n-1)$-dimensional tangent space to $\mathbb{S}(\mathbb{R}^{2n})$ at (p,q) into $\mathbb{R}^2$, has rank j. A new characterization of the sets $\Sigma_j(A)$ in terms of properties of the operators A_1 and A_2 then allows to extend and refine the results in [8]. For example, it turns out that $n_A(\Sigma_2(A))$, in fact, coincides with the interior of $W(A)$. However, our main result is that the image of the critical points of n_A, that is, the points where n'_A does not have the full rank, consists precisely of the curve C_A together with all line segments joining points on C_A at which C_A has the same tangent. This confirms the link between the differential properties of n_A and the shape of $W(A)$.

2. The numerical range map

Let $A \in \mathscr{L}(\mathbb{C}^n)$. We define Hermitian operators A_1 and A_2 in $\mathscr{L}(\mathbb{C}^n)$ as

$$A_1 = \frac{A + A^*}{2} \qquad \text{and} \qquad A_2 = \frac{A - A^*}{2i};$$

these are the unique Hermitian operators subject to $A = A_1 + iA_2$.

For each $k = 1, 2, \ldots$, let $J_k : \mathbb{R}^k \times \mathbb{R}^k \longrightarrow \mathbb{C}^k$ be the bijection $J_k(p, q) = p + iq$. The Cartesian product $\mathbb{R}^k \times \mathbb{R}^k$ and the Euclidean space $\mathbb{R}^{2k}$ are identified in the obvious way.

By N_A we denote the map

$$(p, q) \longmapsto J_1^{-1}(\langle A(p + iq), p + iq\rangle), \qquad p, q \in \mathbb{R}^n$$

from $\mathbb{R}^n \times \mathbb{R}^n \cong \mathbb{R}^{2n}$ into $\mathbb{R}^2$. Because A_1 and A_2 are Hermitian, both $\langle A_1 x, x\rangle$ and $\langle A_2 x, x\rangle$ are real numbers for every $x \in \mathbb{C}^n$. It follows that

$$N_A(p, q) = (\langle A_1(p + iq), p + iq\rangle, \langle A_2(p + iq), p + iq\rangle), \qquad p, q \in \mathbb{R}^n.$$

As the maps $(p, q) \longmapsto \langle A_j(p + iq), p + iq\rangle$ for $j = 1, 2$ are polynomials in the coordinates of p and q with coefficients in $\mathbb{R}$, it is clear that $N_A \in C^\infty(\mathbb{R}^{2n}, \mathbb{R}^2)$. The derivative $N_A'(p, q)$ of N_A at a point $(p, q) \in \mathbb{R}^{2n}$ is the linear map from $\mathbb{R}^{2n}$ into $\mathbb{R}^2$ specified by the condition

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

Setting $x = p + iq$, a straightforward calculation shows that $N_A'(p, q)$ is given by

$$\begin{aligned} N_A'(p, q)(h, k) &= 2\,(\mathrm{Re}\,(\langle A_1 p, h\rangle + \langle A_1 q, k\rangle) + \mathrm{Im}\,(\langle A_1 p, k\rangle - \langle A_1 q, h\rangle), \\ &\qquad \mathrm{Re}\,(\langle A_2 p, h\rangle + \langle A_2 q, k\rangle) + \mathrm{Im}\,(\langle A_2 p, k\rangle - \langle A_2 q, h\rangle)) \\ &= 2\,(\mathrm{Re}\langle A_1 x, h + ik\rangle, \mathrm{Re}\langle A_2 x, h + ik\rangle) \end{aligned}$$

for all $(h, k) \in \mathbb{R}^{2n}$.

DEFINITION 2.1. The *numerical range map* n_A associated with the operator A in $\mathscr{L}(\mathbb{C}^n)$ is the restriction of N_A to the unit sphere $\mathbb{S}(\mathbb{R}^{2n}) = \{y \in \mathbb{R}^{2n} : \|y\| = 1\}$ of $\mathbb{R}^{2n}$.

It follows that n_A is a smooth map from the $(2n - 1)$-dimensional C^∞-manifold $\mathbb{S}(\mathbb{R}^{2n})$ into $\mathbb{R}^2$. Its derivative $n_A'(y)$ at a point $y \in \mathbb{S}(\mathbb{R}^{2n})$ is therefore a linear map from the tangent space $T_y\mathbb{S}(\mathbb{R}^{2n})$ of $\mathbb{S}(\mathbb{R}^{2n})$ at y into $\mathbb{R}^2$. Note that $T_y\mathbb{S}(\mathbb{R}^{2n})$ is the orthogonal complement $\{y\}^\perp$ of the 1-dimensional subspace $\mathbb{R}y$ in $\mathbb{R}^{2n}$. More precisely then, the map $n_A'(y)$ is the restriction of the linear map $N_A'(y) : \mathbb{R}^{2n} \longrightarrow \mathbb{R}^2$ to $\{y\}^\perp$. In particular, given $y \in \mathbb{S}(\mathbb{R}^{2n})$, the linear map $n_A'(y)$ has rank 0, rank 1 or rank 2. We split $\mathbb{S}(\mathbb{R}^{2n})$ accordingly and set $\Sigma_r(A) = \{y \in \mathbb{S}(\mathbb{R}^{2n}) \mid \operatorname{rank} n_A'(y) = r\}$ for $r = 0, 1, 2$.

DEFINITION 2.2. Let $y \in \mathbb{S}(\mathbb{R}^{2n})$. If $y \in \Sigma_2(A)$, then we call y a *regular point* of n_A. Otherwise y is a *critical point* of n_A, and we add *rank* 0 (*rank* 1) if we want to specify that y belongs to $\Sigma_0(A)$ ($\Sigma_1(A)$, respectively). A rank 0 critical point is also called a *stationary point* of n_A.

REMARK 2.3. An *affine transformation* of $\mathbb{R}^2 = \mathbb{C}$ is a map $\tau_{\alpha\beta\gamma} : \mathbb{C} \longrightarrow \mathbb{C}$ with parameters $\alpha, \beta, \gamma \in \mathbb{C}$ such that $\alpha\beta \neq 0$ and $\alpha\beta^{-1} \notin i\mathbb{R}$, which sends $\lambda \in \mathbb{C}$ to $\alpha \operatorname{Re}\lambda + i\beta \operatorname{Im}\lambda + \gamma$. In results about the numerical range of an operator A, frequently the corresponding affine transformations $\tau_{\alpha\beta\gamma}(A) = \alpha A_1 + i\beta A_2 + \gamma I$ of A are considered (see [10, §2]) for the numerical range $W\left(\tau_{\alpha\beta\gamma}(A)\right)$ of $\tau_{\alpha\beta\gamma}(A)$ is given by $\tau_{\alpha\beta\gamma}\,(W(A))$, [10, §2.4]. For the derivative of the numerical range map $n_{\tau_{\alpha\beta\gamma}(A)}$ of $\tau_{\alpha\beta\gamma}(A)$ at a point $y \in \mathbb{S}(\mathbb{R}^{2n})$ we obtain

$$n'_{\tau_{\alpha\beta\gamma}(A)}(y)\left(T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n})\right) = J_1^{-1}\left(\tau_{\alpha\beta 0}\left(\left(J_1 \circ n'_A(y)\right)\left(T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n})\right)\right)\right).$$

This is an immediate consequence of the representation of $n'_{\tau_{\alpha\beta\gamma}(A)}(y)$. It follows that $\Sigma_j\left(\tau_{\alpha\beta\gamma}(A)\right) = \Sigma_j(A)$ for $j = 0, 1, 2$.

The proposition below already characterizes the set of all critical points of n_A.

PROPOSITION 2.4. *Let $(c, s) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Then a point $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a critical point of n_A such that*

$$n'_A(p, q)\left(T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n})\right) \subseteq \mathbb{R}(c, s)$$

if and only if the vector $x = p + iq \in \mathbb{S}(\mathbb{C}^n)$ is an eigenvector of the Hermitian operator $sA_1 - cA_2$.

PROOF. Assume that (p, q) is a critical point of n_A with

$$n'_A(p, q)(h, k) \in \mathbb{R}(c, s) \qquad \text{for all } (h, k) \in T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n}) = \{(p, q)\}^\perp.$$

Since $n'_A(p, q)(h, k) = 2\left(\operatorname{Re}\langle A_1x, h + ik\rangle, \operatorname{Re}\langle A_2x, h + ik\rangle\right)$ and $\mathbb{R}(c, s) = \{(s, -c)\}^\perp$, this is equivalent to

$$s \operatorname{Re}\langle A_1x, h + ik\rangle - c \operatorname{Re}\langle A_2x, h + ik\rangle = 0 \qquad \text{for all } (h, k) \in \{(p, q)\}^\perp$$

which can be written as

$$(1) \qquad \operatorname{Re}\langle (sA_1 - cA_2)x, h + ik\rangle = 0 \qquad \text{for all } (h, k) \in \{(p, q)\}^\perp$$

because c and s are real numbers. However, since

$$\{(p, q)\}^\perp = \left\{(h, k) \in \mathbb{R}^{2n} : \operatorname{Re}\langle x, h + ik\rangle = 0\right\},$$

it follows that condition (1) holds if and only if $(sA_1 - cA_2)x \in J_n\left(\{(p, q)\}^{\perp\perp}\right) = \mathbb{R}x$, that is, if and only if x is an eigenvector of $sA_1 - cA_2$. □

Proposition 2.4 allows us to distinguish between the two different types of critical points. A point $\lambda = (\beta, \gamma) \in \mathbb{R}^2$ is called a *joint eigenvalue* of a pair (B, C) of operators in $\mathbb{C}^n$ if there exists $x \in \mathbb{C}^n \setminus \{0\}$ such that $Bx = \beta x$ and $Cx = \gamma x$. In this case, the vector x is called a *joint eigenvector* of (B, C) corresponding to the joint eigenvalue λ.

COROLLARY 2.5 (see [8, Proposition 1]). *The point $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a stationary point of n_A if and only if the vector $x = p + iq$ is a joint eigenvector of some pair $(cA_1 + sA_2, dA_1 + tA_2)$ with $(c, s), (d, t) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ such that $ct - sd \neq 0$. The corresponding joint eigenvalue is the point $(\langle (cA_1 + sA_2)x, x\rangle, \langle (dA_1 + tA_2)x, x\rangle)$.*

PROOF. By definition, the point (p, q) is a stationary point if and only if

$$n'_A(p, q)\left(T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n})\right) = \{(0, 0)\}.$$

Since $\{(0, 0)\} = \mathbb{R}(-s, c) \cap \mathbb{R}(-t, d)$, Proposition 2.4 yields that this is equivalent to x being an eigenvector of both operators $cA_1 + sA_2$ and $dA_1 + tA_2$. □

We note that if one of the conditions of Corollary 2.5 is satisfied, then the vector $x = p + iq$ is a joint eigenvector of all the pairs $(cA_1 + sA_2, dA_1 + tA_2)$ with $(c, s), (d, t) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. In particular, then x is a joint eigenvector of the pair (A_1, A_2) corresponding to the joint eigenvalue $n_A(p, q)$.

Combining Proposition 2.4 and Corollary 2.5 gives the following.

COROLLARY 2.6. *The point $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a rank 1 critical point of n_A if and only if the vector $x = p + iq$ is an eigenvector of $cA_1 + sA_2$ for some point $(c, s) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ which is unique up to multiplication by $\gamma \in \mathbb{R} \setminus \{0\}$.*

REMARK 2.7. From its definition it follows that, given $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ and $x = p + iq$, the numerical range map n_A is constant on the subset $J_n^{-1}(\mathbb{C}x \cap \mathbb{S}(\mathbb{C}^n))$ of $\mathbb{S}(\mathbb{R}^{2n})$. As a simple consequence of the results above and the fact that multiplication by a complex number α leaves the (complex) eigenspaces of an operator $B \in \mathscr{L}(\mathbb{C}^n)$ invariant, we obtain for the derivative of n_A at any of the points $J_n^{-1}(\alpha x)$ with $\alpha \in \mathbb{S}(\mathbb{C})$ that

$$n'_A\left(J_n^{-1}(\alpha x)\right)\left(T_{J_n^{-1}(\alpha x)}\mathbb{S}(\mathbb{R}^{2n})\right) = n'_A(p, q)\left(T_{(p,q)}\mathbb{S}(\mathbb{R}^{2n})\right).$$

In particular, if $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a rank 0 (rank 1) critical point of n_A, then for every $\alpha = \alpha_1 + i\alpha_2 \in \mathbb{S}(\mathbb{C})$, the point $J_n^{-1}(\alpha(p + iq)) = (\alpha_1 p - \alpha_2 q, \alpha_1 q + \alpha_2 p)$ is also a rank 0 (rank 1) critical point of n_A.

Our main concern in this paper is the connection between certain subsets of the numerical range $W(A)$ of A or, more precisely, of $J_1^{-1}(W(A)) = n_A(\mathbb{S}(\mathbb{R}^{2n}))$, and the differential structure of the numerical range map n_A. A first result in this direction was Corollary 2.5. There we have seen that $n_A(\Sigma_0(A))$ is the set of joint eigenvalues of the pair (A_1, A_2).

Suppose $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a point such that $n_A(p, q)$ belongs to the boundary $J_1^{-1}(\partial W(A))$ of $J_1^{-1}(W(A))$. Let $\{(a, b) \in \mathbb{R}^2 \mid ca + sb = \lambda\}$ be a supporting line to $J_1^{-1}(W(A))$, that is, a line which intersects $J_1^{-1}(W(A))$ in a non-empty subset of its boundary, passing through the point $n_A(p, q)$. Such a line exists because $W(A)$ is convex. Then $x = p + iq \in \mathbb{S}(\mathbb{C}^n)$ is a vector such that $\langle (cA_1 + sA_2)x, x\rangle$ is the maximum or minimum of the set $\{\langle (cA_1 + sA_2)z, z\rangle \mid z \in \mathbb{S}(\mathbb{C}^n)\}$. It follows that x is an eigenvector of the operator $cA_1 + sA_2$. Combining this well-known fact (see, for example, [2, Proposition 2.1]) with Proposition 2.4 yields the subsequent result.

LEMMA 2.8. $(J_1 \circ n_A)^{-1}(\partial W(A)) \subseteq \Sigma_0(A) \cup \Sigma_1(A)$.

REMARK 2.9. In the case that $A \in \mathscr{L}(\mathbb{C}^2)$, the inclusion of Lemma 2.8 becomes, in fact, an equality (see also [8, Proposition 4]). For if $(p, q) \in \mathbb{S}(\mathbb{R}^{2n})$ is a critical point of n_A, then $x = p + iq$ is an eigenvector of some operator $cA_1 + sA_2$ with $(c, s) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ and the corresponding eigenvalue, say α, is necessarily the maximal or minimal eigenvalue of $cA_1 + sA_2$. By [10, §3.9–10], the line $\{(a, b) \in \mathbb{R}^2 \mid ca + sb = \alpha\}$ is a supporting line to $J_1^{-1}(W(A))$. Since $n_A(p, q)$ is a point on this line, it therefore belongs to $J_1^{-1}(\partial W(A))$.

As an immediate consequence of Lemma 2.8 we obtain that $J_1 \circ n_A$ maps all regular points into the interior $W(A)^o$ of $W(A)$ (see also [8, Proposition 2]). However, as shown below, an even stronger result holds. We make use of the following fact.

REMARK 2.10. Let X be a non-trivial k-dimensional subspace of $\mathbb{C}^n$ with the induced norm and inner product, and let P be the orthogonal projection of $\mathbb{C}^n$ onto X. By $\tilde{X}$ we denote the $2k$-dimensional subspace $J_n^{-1}(X)$ of $\mathbb{R}^{2n}$ corresponding to X. Then the operator $PA|_X$ obtained by restricting the operator PA to X, belongs to $\mathscr{L}(X)$, and $PA_1|_X$ and $PA_2|_X$ are the unique Hermitian operators in $\mathscr{L}(X)$ such that $PA|_X = PA_1|_X + iPA_2|_X$. The numerical range map $n_{PA|_X}$ of $PA|_X$ is a smooth map from the unit sphere $S(\tilde{X})$ of $\tilde{X}$ into $\mathbb{R}^2$. In fact, $n_{PA|_X}$ is the restriction of n_A to the $(2k-1)$-dimensional C^∞-submanifold $\mathbb{S}(\tilde{X})$ of $\mathbb{S}(\mathbb{R}^{2n})$. It follows that the derivative $n'_{PA|_X}(y)$ of $n_{PA|_X}$ at a point $y \in \mathbb{S}(\tilde{X})$ is the restriction of $n'_A(y)$ to the tangent space $T_y\mathbb{S}(\tilde{X})$ of $\mathbb{S}(\tilde{X})$ at y. That is, the linear map $n'_{PA|_X}(y)$ is the restriction of $N'_A(y)$ to the orthogonal complement $\{y\}^\perp \cap \tilde{X}$ of $\mathbb{R}y$ in $\tilde{X}$. In particular, $\operatorname{rank} n'_{PA|_X}(y) \le \operatorname{rank} n'_A(y)$ for all $y \in \mathbb{S}(\tilde{X})$.

We write $l_{c,s,\lambda}$ for the line $\{(a,b) \in \mathbb{R}^2 \mid ca + sb = \lambda\}$ in $\mathbb{R}^2$ and, as this requires implicitly that $(c, s) \neq (0, 0)$, we adopt as a convention that the coordinates (c, s, λ) satisfy $c = \cos\varphi$, $s = \sin\varphi$ for some angle $\varphi \in [0, \pi)$.

PROPOSITION 2.11. $n_A(\Sigma_2(A)) = J_1^{-1}(W(A)^o)$.

PROOF. It remains to show that every point in $J_1^{-1}(W(A)^o)$ which is the image of a critical point is also the image of a regular point of n_A. Note that such a point can not exist if $\mathbb{C}^n$ has dimension $n \leq 2$ (see Remark 2.9).

Let $\mu_0 \in J_1^{-1}(W(A)^o) \cap n_A(\Sigma_0(A) \cup \Sigma_1(A))$ and let L be the set of lines $l_{c,s,\lambda}$ through the point μ_0 with coordinates (c, s, λ) such that λ is an eigenvalue of the operator $cA_1 + sA_2$. We pick a line l_{c_0,s_0,λ_0} through the point μ_0 such that $l_{c_0,s_0,\lambda_0} \notin L$, if such a line exists. Otherwise we take a line $l_{c_0,s_0,\lambda_0} \in L$ satisfying $l_{c_0,s_0,\lambda_0} \cap n_A(\Sigma_0(A)) \subseteq \{\mu_0\}$.

Without loss of generality we can assume that $c_0 = 1$ and $s_0 = 0$. If not, consider the operator $B = e^{-i\varphi_0}A$, where $\varphi_0 \in [0, \pi)$ is the angle such that $\cos\varphi_0 = c_0$ and $\sin\varphi_0 = s_0$. By Remark 2.3, we have $\Sigma_j(B) = \Sigma_j(A)$ for $j = 0, 1, 2$.

The line $l_{1,0,\lambda_0}$ intersects the boundary of $J_1^{-1}(W(A))$ in two points μ_1 and μ_2. For $j = 1, 2$, take $(p_j, q_j) \in n_A^{-1}(\{\mu_j\})$ and put $x_j = p_j + iq_j$. Note that, by Lemma 2.8 and the choice of $l_{1,0,\lambda_0}$, the two points (p_1, q_1) and (p_2, q_2) belong to $\Sigma_1(A)$. Let X be the 2-dimensional subspace of $\mathbb{C}^n$ spanned by the vectors x_1 and x_2. Then, using the notation of Remark 2.10, the numerical range $J_1^{-1}(W(PA|_X))$ of $PA|_X$ is either an ellipse or a line segment.

If $J_1^{-1}(W(PA|_X)) = n_{PA|_X}(\mathbb{S}(\tilde{X}))$ is an ellipse, then, by Remark 2.9, its interior is the image of the set of regular points of $n_{PA|_X}$. In particular, there exists a regular point (p_0, q_0) of $n_{PA|_X}$ in $\mathbb{S}(\tilde{X}) \subseteq \mathbb{S}(\mathbb{R}^{2n})$ such that $n_{PA|_X}(p_0, q_0) = \mu_0$. By Remark 2.10, (p_0, q_0) is also a regular point of n_A and $n_A(p_0, q_0) = \mu_0$.

Suppose $J_1^{-1}(W(PA|_X))$ is a line segment. More precisely then, it is the line segment joining the points μ_1 and μ_2 and contained in the line $l_{1,0,\lambda_0}$. This implies $PA_1|_X = \lambda_0 I_X$ because $\langle (PA_1|_X - \lambda_0 I_X)x, x\rangle = 0$ for all $x \in X$. It follows further that every point $(p_0, q_0) \in n_{PA|_X}^{-1}(\{\mu_0\})$ is a rank 1 critical point of $n_{PA|_X}$ and, by Proposition 2.4, that $n'_{PA|_X}(p_0, q_0)\big(T_{(p_0,q_0)}\mathbb{S}(\tilde{X})\big) = \mathbb{R}(0, 1) = l_{1,0,0}$. Remark 2.10 yields that $n'_A(p_0, q_0)\left(T_{(p_0,q_0)}\mathbb{S}(\mathbb{R}^{2n})\right)$ contains at least the line $l_{1,0,0}$.

If $l_{1,0,\lambda_0} \notin L$, then we can deduce from Proposition 2.4 that every point $(p_0, q_0) \in n_{PA|_X}^{-1}(\{\mu_0\})$ is a regular point of n_A. Hence, assume that $l_{1,0,\lambda_0} \in L$. Since $PA_2|_X$ is Hermitian and $\mu_1 \neq \mu_2$, the vectors x_1 and x_2 are eigenvectors of $PA_2|_X$ and form an orthonormal basis of X. Thus, a point $(p, q) \in n_{PA|_X}^{-1}(\{\mu_0\})$ is of the form $J_n^{-1}(\alpha x_1 + \beta x_2)$ for constants $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha|^2 + |\beta|^2 = 1$ and $n_{PA|_X}(p, q) = (\lambda_0, |\alpha|^2\langle A_2x_1, x_1\rangle + |\beta|^2\langle A_2x_2, x_2\rangle) = \mu_0$. It suffices to show that there exists $(h_0, k_0) \in \tilde{X}^\perp$ such that $\mathrm{Re}\langle A_1x_1, h_0 + ik_0\rangle \neq 0$. Then we can find $\alpha_0, \beta_0 \in \mathbb{C}$

satisfying $(p_0, q_0) = J_n^{-1}(\alpha_0 x_1 + \beta_0 x_2) \in n_{PA|_X}^{-1}(\{\mu_0\})$ and $\mathrm{Re}\langle A_1(\alpha_0 x_1 + \beta_0 x_2), h_0 + ik_0\rangle \neq 0$. As $(h_0, k_0) \in \{(p_0, q_0)\}^\perp$ and $n'_A(p_0, q_0)$ is given by $n'_A(p_0, q_0)(h, k) = 2\,(\mathrm{Re}\langle A_1(\alpha_0 x_1 + \beta_0 x_2), h + ik\rangle, \mathrm{Re}\langle A_2(\alpha_0 x_1 + \beta_0 x_2), h + ik\rangle)$ for all $(h, k) \in T_{(p_0,q_0)}\mathbb{S}(\mathbb{R}^{2n}) = \{(p_0, q_0)\}^\perp$, this implies that $n'_A(p_0, q_0)\left(T_{(p_0,q_0)}\mathbb{S}(\mathbb{R}^{2n})\right)$ contains a point (r_1, r_2) with $r_1 \neq 0$ in addition to the line $l_{1,0,0}$, and therefore the whole plane. In other words, (p_0, q_0) is a regular point of n_A such that $n_A(p_0, q_0) = \mu_0$.

Now, since (p_1, q_1) is a rank 1 critical point of n_A and $l_{1,0,\lambda_0}$ is not the supporting line to $J_1^{-1}(W(A))$ at μ_1, there exists a pair $(c, s) \in \mathbb{S}(\mathbb{R}^2)$ with $s > 0$ such that $n'_A(p_1, q_1)\left(T_{(p_1,q_1)}\mathbb{S}(\mathbb{R}^{2n})\right) = l_{c,s,0}$. On the other hand, $(p_1, q_1) \in \Sigma_0(PA|_X)$ which gives $n'_A(p_1, q_1)\left(T_{(p_1,q_1)}\mathbb{S}(\tilde{X})\right) = \{(0, 0)\}$. As $T_{(p_1,q_1)}\mathbb{S}(\tilde{X}) = \{(p_1, q_1)\}^\perp \cap \tilde{X}$, there exists a point $(h_0, k_0) \in \tilde{X}^\perp$ and a real number $\gamma \neq 0$ such that $n'_A(p_1, q_1)(h_0, k_0) = \gamma(-s, c)$. In particular, $\mathrm{Re}\langle A_1 x_1, h_0 + ik_0\rangle = -\gamma s \neq 0$. This completes the proof. □

The worst case in the proof of Proposition 2.11 is indeed possible, that is, it can happen that every line through an interior point which is the image of a critical point, belongs to L and that the corresponding operator $PA|_X$ has a segment of the line for its numerical range. We illustrate this with an example.

EXAMPLE 2.12. Consider the operator $A = A_1 + iA_2 \in \mathscr{L}(\mathbb{C}^3)$, where A_1 and A_2 are the Hermitian operators

$$A_1 = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The numerical range $W(A)$ of A is the ellipse $\{\lambda \in \mathbb{C} : 2(\mathrm{Re}\,\lambda)^2 + (\mathrm{Im}\,\lambda)^2 \leq 1\}$, and the numerical range map n_A is given by

$$n_A(p, q) = \left(p_2(p_1 + p_3) + q_2(q_1 + q_3),\, p_1^2 - p_3^2 + q_1^2 - q_3^2\right) \quad \text{for } (p, q) \in \mathbb{S}(\mathbb{R}^6).$$

A direct calculation shows that $\Sigma_1(A)$ consists of the two sets $S_1 = \{(av, bv) \mid v \in \mathbb{S}(\mathbb{R}^3),\ v_2^2 = 2v_1v_3,\ (a, b) \in \mathbb{S}(\mathbb{R}^2)\}$ and $S_2 = \{(av, bv) \mid v \in \mathbb{S}(\mathbb{R}^3),\ v_1 = -v_3,\ (a, b) \in \mathbb{S}(\mathbb{R}^2)\}$. All other points $(p, q) \in \mathbb{S}(\mathbb{R}^6)$ belong to $\Sigma_2(A)$. In particular, $\Sigma_0(A) = \emptyset$. The image of S_1 under n_A is precisely the boundary of $J_1^{-1}(W(A))$, whereas $n_A(S_2) = \{(0, 0)\}$. We note that 0 is an eigenvalue of every operator $cA_1 + sA_2$ with $(c, s) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, that is, all lines through the point $(0, 0)$ belong to the set L of Proposition 2.11.

Any line l through the point $(0, 0) \in \mathbb{R}^2$ has a representation $\left\{(a, b) \in \mathbb{R}^2 \mid (2t - 1)a - \sqrt{2t(1 - t)}b = 0\right\}$ for some $0 < t \leq 1$. It intersects $J_1^{-1}(\partial W(A))$ in the two points $\mu_1(t) = \left(\sqrt{2t(1 - t)}, 2t - 1\right)$ and $\mu_2(t) = -\mu_1(t)$. For $j = 1, 2$, we have

$n_A^{-1}(\{\mu_j\}) = J_3^{-1}\left(\mathbb{C}x_j(t) \cap \mathbb{S}(\mathbb{C}^3)\right)$, where the vectors $x_1(t), x_2(t) \in \mathbb{C}^3$ are given by

$$x_1(t) = \begin{pmatrix} t \\ \sqrt{2t(1-t)} \\ 1-t \end{pmatrix} \quad \text{and} \quad x_2(t) = \begin{pmatrix} 1-t \\ -\sqrt{2t(1-t)} \\ t \end{pmatrix}.$$

Let $X(t) = \mathbb{C}x_1(t) + \mathbb{C}x_2(t)$, $\tilde{X}(t) = J_3^{-1}(X(t))$ and let $P(t)$ be the orthogonal projection of $\mathbb{C}^3$ onto $X(t)$. For $j, k = 1, 2$, the tangent space $T_{J_3^{-1}(x_j(t))}\mathbb{S}(\tilde{X}(t)) = \{J_3^{-1}(x_j(t))\}^{\perp} \cap \tilde{X}(t)$ of $\mathbb{S}(\tilde{X}(t))$ at $J_3^{-1}(x_j(t))$ is spanned by the three points in $J_3^{-1}\left(\{ix_j(t), x_{3-j}(t), ix_{3-j}(t)\}\right)$, and $\operatorname{Re}\langle A_k x_j(t), x\rangle = 0$ for every vector $x \in \{ix_j(t), x_{3-j}(t), ix_{3-j}(t)\}$. It follows that the points $J_3^{-1}(x_1(t))$ and $J_3^{-1}(x_2(t))$ in $\mathbb{S}(\tilde{X}(t))$ are stationary points of the numerical range map $n_{P(t)A|_{X(t)}}$ of $P(t)A|_{X(t)}$, and, by Remark 2.7, that $n_A^{-1}(\{\mu_1(t)\}) \cup n_A^{-1}(\{\mu_2(t)\}) = \Sigma_0\left(P(t)A|_{X(t)}\right)$. Hence, $J_1^{-1}\left(W\left(P(t)A|_{X(t)}\right)\right)$ is the line segment joining the points $\mu_1(t)$ and $\mu_2(t)$.

We note that the point $(0, 0) \in \mathbb{R}^2$ is, for example, the image under n_A of the regular point $\left(\begin{pmatrix} \frac{\sqrt{2}}{2} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \frac{\sqrt{2}}{2} \end{pmatrix}\right)$. This point was obtained by taking $t = 1$ in the above and proceeding then as in the proof of Proposition 2.11.

In the remaining part of the paper, we will have a closer look at the image of the rank 1 critical points of n_A. Motivated by the observations of Remark 2.7, we take a geometric approach.

3. Geometric properties

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $m \in \mathbb{N}$. The *Grassmannian* $G_{m,k}\mathbb{K}$ is the set of all k-dimensional $\mathbb{K}$-subspaces of the vector space $\mathbb{K}^m$. It is a compact analytic $\mathbb{K}$-manifold of dimension $(m-k)k$. The space $G_{m,k}\mathbb{K}$ has a natural topology which is induced by the differential structure of the manifold. This topology is, for example, determined by the metric h on $G_{m,k}\mathbb{K}$ with

$$h(U, V) = \sup_{v \in V \cap \mathbb{S}(\mathbb{K}^m)} \inf_{u \in U \cap \mathbb{S}(\mathbb{K}^m)} \|u - v\| \qquad \text{for all } U, V \in G_{m,k}\mathbb{K}.$$

The metric h corresponds to the Hausdorff metric (see [5, 4.5.23]) on the set $B_{m,k} = \left\{U \cap \mathbb{S}(\mathbb{K}^m) \;\middle|\; U \in G_{m,k}\mathbb{K}\right\}$. Other ways of describing the topology on $\mathrm{G}_{m,k}\mathbb{K}$ can be found, for example, in [17, Chapter 13] or [4, VII.8.2].

The $(m-1)$-dimensional *projective space* $\mathrm{PG}(\mathbb{K}^m)$ over $\mathbb{K}$ is defined as

$$\mathrm{PG}(\mathbb{K}^m) = \bigcup_{0 \le k \le m} G_{m,k}\mathbb{K}.$$

There are geometric operations $\vee$ and $\wedge$ defined on $\mathrm{PG}(\mathbb{K}^m)$, where $U \vee V$ and $U \wedge V$ denote the span and the intersection of the subspaces $U, V \leq \mathbb{K}^m$, respectively.

Usually, the 1-, 2-, $(m-1)$-dimensional subspaces of $\mathbb{K}^m$ are called the *points*, *lines*, and *hyperplanes* in $\mathrm{PG}(\mathbb{K}^m)$, respectively. The span of two points is a line and the intersection of two hyperplanes is an element of $G_{m,m-2}\mathbb{K}$.

By common abuse of notation we introduce *homogeneous coordinates* for the points in $\mathrm{PG}(\mathbb{K}^m)$ as $(u_1 : \ldots : u_m) = \mathbb{K}(u_1, \ldots, u_m)$ whenever it is clear which field $\mathbb{K}$ is involved. The coordinates of a vector in $\mathbb{K}^m$ are always expressed with respect to a fixed (standard) basis of $\mathbb{K}^m$.

A *polarity* π of $\mathrm{PG}(\mathbb{K}^m)$ is a bijection on $\mathrm{PG}(\mathbb{K}^m)$ which reverses the inclusion of subspaces and satisfies $\pi^2 = \mathrm{id}$. Throughout the paper, π denotes the *standard polarity* of $\mathrm{PG}(\mathbb{K}^m)$ which is characterized by

$$u^\pi = \left\{ v \in \mathbb{K}^m \;\middle|\; \sum_{j=1}^{m} u_j v_j = 0 \right\} \qquad \text{for all } u \in G_{m,1}\mathbb{K}.$$

It follows that $u^\pi \in G_{m,m-1}\mathbb{K}$. Using the polarity π we can also introduce homogeneous coordinates on $G_{m,m-1}\mathbb{K}$ by setting $[v_1 : \ldots : v_m] = (v_1 : \ldots : v_m)^\pi$.

In the special case where $m = 3$, the projective space $\mathrm{PG}(\mathbb{K}^3)$ is called the *projective plane* over $\mathbb{K}$. Here lines are the same as hyperplanes.

From Remark 2.7 it is clear that the numerical range map induces a mapping which is defined on a complex projective space of the proper dimension.

DEFINITION 3.1. For $A \in \mathscr{L}(\mathbb{C}^n)$, we define the map ν_A from $G_{n,1}\mathbb{C}$ into $G_{3,1}\mathbb{R}$ by

$$\nu_A(\mathbb{C}x) = \mathbb{R}\left(\langle A_1 x, x\rangle, \langle A_2 x, x\rangle, \|x\|^2\right).$$

As $\|x\| \neq 0$ for $x \neq 0$, we infer that the image of ν_A is contained in an affine subplane F of $\mathrm{PG}(\mathbb{R}^3)$ with the point set $\{(\alpha_1 : \alpha_2 : 1) \mid (\alpha_1, \alpha_2) \in \mathbb{R}^2\}$. Let $\tau : F \longrightarrow \mathbb{R}^2$ be the bijection $\tau(\alpha_1 : \alpha_2 : 1) = (\alpha_1, \alpha_2)$. By Remark 2.7, the numerical range map n_A from $\mathbb{S}(\mathbb{R}^{2n}) \cong \mathbb{S}(\mathbb{C}^n)$ to $\mathbb{R}^2$ factors through the projective space $\mathrm{PG}(\mathbb{C}^n)$. The factoring map is given by ν_A. More precisely, we have

$$n_A = \tau \circ \nu_A \circ \sigma,$$

where $\sigma : \mathbb{S}(\mathbb{R}^{2n}) \longrightarrow G_{n,1}\mathbb{C}$ is the map $\sigma(p,q) = \mathbb{C}(p + iq)$ for all $(p,q) \in \mathbb{S}(\mathbb{R}^{2n})$.

A non-empty subset C of $G_{3,1}\mathbb{K}$ is called a *plane* $\mathbb{K}$*-algebraic curve* if it is the zero locus of a homogeneous 3-variate polynomial over $\mathbb{K}$. We note that the defining polynomial of C is not uniquely determined; for example, if f defines the curve, then so does f^k for any $k \geq 1$. A curve is said to be *irreducible* if it has an irreducible defining polynomial. Since a polynomial ring over a field is a unique factorization domain, each algebraic curve C is the union of finitely many irreducible curves. If $C_1, \ldots, C_k$ are the irreducible components of C with irreducible defining polynomials $f_1, \ldots, f_k$, then the polynomial $f = f_1 \cdots f_k$ is a defining polynomial of C of minimal degree. We call f a *minimal polynomial of* C. It is unique up to a constant factor. Note that an irreducible real algebraic curve is not necessarily connected.

Let $C = \{u \in G_{3,1}\mathbb{K} \mid f(u) = 0\}$ be an algebraic curve and let f be a minimal polynomial of C. A point $u \in C$ is called *singular* or a *singularity* of C if $(\partial f/\partial u_j)(u) = 0$ for $j = 1, 2, 3$. Observe that C has only finitely many singular points. These are the singular points of the irreducible components of C together with the points of intersection of any two of these components. A non-singular point $u \in C$ is called a *simple* point of C. The curve C is the topological closure of its simple points. Also, if $u \in C$ is simple, then there exists a neighborhood of u in which C admits a smooth parameterization.

Let C be an irreducible plane algebraic curve and f be its minimal polynomial. At each simple point $u \in C$, we have a unique tangent line to C which is given by

$$\mathscr{T}_u C = \left[\frac{\partial f}{\partial u_1}(u) : \frac{\partial f}{\partial u_2}(u) : \frac{\partial f}{\partial u_3}(u) \right].$$

If C is not a projective line or a point, then it is well known that the set $\{(\mathscr{T}_u C)^\pi \mid u \in C \text{ simple}\}$ is contained in a unique irreducible algebraic curve C^*, the so-called *dual curve* of C. As an algebraic curve has at most finitely many singularities, the dual curve is the topological closure of the set $\{(\mathscr{T}_u C)^\pi \mid u \in C \text{ simple}\}$. We have $C^{**} = C$. If C is a projective line, then $\{(\mathscr{T}_u C)^\pi \mid u \in C\}$ consists of a single point u in $\mathrm{PG}(\mathbb{K}^3)$. In this case, we set $C^* = \{u\}$ and define C^{**} to be the image under π of the set of all lines in $\mathrm{PG}(\mathbb{K}^3)$ which pass through u. Then again we have $C^{**} = C$. The dual curve of a general plane algebraic curve C is the union of the dual curves of its irreducible components. In particular, C and C^* have the same number of irreducible components.

Naturally, the general theory of real algebraic curves differs quite a bit from the complex theory. However, the above statements on the duality of real algebraic curves immediately follow from the corresponding statements in the complex case.

The details and further information on complex algebraic curves can be found, for example, in [16]. The literature for the real case is somewhat less easy to access. As a general reference to the theory of real algebraic geometry, see [3].

Let $A \in \mathscr{L}(\mathbb{C}^n)$. We define an algebraic curve C_A in the real projective plane by setting its dual curve to be

$$D_A = \{(c : s : \lambda) \in G_{3,1}\mathbb{R} \mid \det(cA_1 + sA_2 + \lambda I) = 0\}.$$

We write f_A for the defining polynomial of D_A, that is, $f_A(c, s, \lambda) = \det(cA_1 + sA_2 + \lambda I)$ for $(c : s : \lambda) \in G_{3,1}\mathbb{R}$, and denote the minimal polynomial of D_A by m_A.

Kippenhahn [10] showed that the curve $C_A = D_A^*$ is contained in the affine subplane $F = \{(\alpha_1 : \alpha_2 : 1) \mid (\alpha_1, \alpha_2) \in \mathbb{R}^2\}$ and that the convex hull $\mathrm{co}(C_A)$ of C_A is exactly the image of ν_A, that is, $\tau(\mathrm{co}(C_A))$ is the numerical range of A. Note that C_A and D_A are the real parts of the curves in $\mathrm{PG}(\mathbb{C}^3)$ considered in [10]. Our aim is to show that the curve C_A itself is contained in the image under $\tau^{-1} \circ n_A = \nu_A \circ \sigma$ of the set $\Sigma_0(A) \cup \Sigma_1(A)$ of critical points of n_A.

Every point $u \in D_A$ has a representation $(\cos\varphi : \sin\varphi : \lambda)$ for some $\varphi \in [0, \pi)$ and $\lambda \in \mathbb{R}$. As u is a zero of f_A, it follows that $-\lambda$ is an eigenvalue of the operator $\mathscr{A}(\varphi) = \cos\varphi A_1 + \sin\varphi A_2$.

LEMMA 3.2. *Let the map $\mathscr{A} : \mathbb{R} \longrightarrow \mathscr{L}(\mathbb{C}^n)$ be given by $\mathscr{A}(\varphi) = \cos\varphi A_1 + \sin\varphi A_2$ for $\varphi \in \mathbb{R}$. Let $\varphi_0 \in \mathbb{R}$ and λ_0 be an eigenvalue with multiplicity r of the operator $\mathscr{A}(\varphi_0)$. Then there exists a neighborhood T of φ_0 and regular analytic functions $\lambda_j : T \longrightarrow \mathbb{R}$ and $x_j : T \longrightarrow \mathbb{S}(\mathbb{C}^n)$ with $1 \le j \le r$, such that $\lambda_j(\varphi_0) = \lambda_0$, $\mathscr{A}(\varphi)x_j(\varphi) = \lambda_j(\varphi)x_j(\varphi)$ and $\langle x_j(\varphi), x_k(\varphi)\rangle = \delta_{jk}$ for every $\varphi \in T$ and $1 \le j, k \le r$. Moreover, for every $1 \le j \le r$, the derivative of $\lambda_j(\cdot)$ at φ_0 is given by $\lambda_j'(\varphi_0) = \langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle$.*

PROOF. The map $\mathscr{A}$ is a regular analytic function from $\mathbb{R}$ into the Hermitian operators on $\mathbb{C}^n$. We can thus apply [15, Satz 1] and obtain the functions $\lambda_j(\cdot)$ and $x_j(\cdot)$ with the desired properties.

For $\varphi \in T$ we can express $\lambda_j(\varphi)$ as $\langle \mathscr{A}(\varphi)x_j(\varphi), x_j(\varphi)\rangle$. Then

$$\begin{aligned}
\lambda_j'(\varphi_0) &= \lim_{h\to 0} \frac{\langle \mathscr{A}(\varphi_0 + h)x_j(\varphi_0 + h), x_j(\varphi_0 + h)\rangle - \langle \mathscr{A}(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle}{h} \\
&= \lim_{h\to 0} \frac{\sin h}{h} \langle (-\sin\varphi_0 A_1 + \cos\varphi_0 A_2)x_j(\varphi_0 + h), x_j(\varphi_0 + h)\rangle \\
&\quad + \lim_{h\to 0} \frac{\cos h - 1}{h} \langle \mathscr{A}(\varphi_0)x_j(\varphi_0 + h), x_j(\varphi_0 + h)\rangle \\
&\quad + \lim_{h\to 0} \frac{\langle \mathscr{A}(\varphi_0)x_j(\varphi_0 + h), x_j(\varphi_0 + h)\rangle - \langle \mathscr{A}(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle}{h} \\
&= L_1 + L_2 + L_3.
\end{aligned}$$

The continuity of $x_j(\cdot)$ in φ_0 gives $L_1 = \langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle$ and $L_2 = 0$.

As $\mathscr{A}(\varphi_0) = \mathscr{A}(\varphi_0) + i0$, we can apply the results of Section 2 to the third limit. Let $P_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the projection onto the first coordinate and let $p(\varphi) = J_n^{-1}\left(x_j(\varphi)\right) \in \mathbb{R}^{2n}$ for $\varphi \in T$. Then

$$\begin{aligned}
&\frac{\langle \mathscr{A}(\varphi_0)x_j(\varphi_0 + h), x_j(\varphi_0 + h)\rangle - \langle \mathscr{A}(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle}{h} \\
&\quad = P_1\left(\frac{n_{\mathscr{A}(\varphi_0)}(p(\varphi_0 + h)) - n_{\mathscr{A}(\varphi_0)}(p(\varphi_0))}{h}\right)
\end{aligned}$$

$$= P_1\left(\frac{n'_{\mathscr{A}(\varphi_0)}(p(\varphi_0))\,(p(\varphi_0+h)-p(\varphi_0)) + o\,(p(\varphi_0+h)-p(\varphi_0))}{h}\right),$$

where o is the Landau symbol. Since $x_j(\varphi_0)$ is a joint eigenvector of the pair $(\mathscr{A}(\varphi_0), 0)$, it follows by Corollary 2.5 that $n'_{\mathscr{A}(\varphi_0)}(p(\varphi_0)) = 0 \in \mathscr{L}(\mathbb{R}^{2n}, \mathbb{R}^2)$. Hence

$$L_3 = P_1\left(\lim_{h\to 0}\frac{o\,(p(\varphi_0+h)-p(\varphi_0))}{\|p(\varphi_0+h)-p(\varphi_0)\|}\frac{\|p(\varphi_0+h)-p(\varphi_0)\|}{h}\right) = 0$$

and

$$\lambda_j'(\varphi_0) = \langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle. \qquad \square$$

We can now prove the first main result in this section.

PROPOSITION 3.3. *Let $A \in \mathscr{L}(\mathbb{C}^n)$. Then the curve C_A is contained in the image under $\nu_A \circ \sigma$ of the critical points of n_A.*

PROOF. Let M be the (finite) set of singularities of D_A, and $u_0 = (\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$ be a point in $D_A \setminus M$. Then u_0 belongs to a uniquely determined irreducible component D of D_A, and there exists a neighborhood N of u_0 such that every point $u \in D_A \cap N$ is contained in $D \setminus M$. As u_0 is a zero of f_A, $-\lambda_0$ is an eigenvalue of $\mathscr{A}(\varphi_0)$, say with multiplicity r. By Lemma 3.2, there exist r functions $\varphi \longmapsto \lambda_j(\varphi)$ defined and regular analytic in a neighborhood T of φ_0, such that $\lambda_j(\varphi_0) = \lambda_0$ and $-\lambda_j(\varphi)$ is an eigenvalue of the operator $\mathscr{A}(\varphi)$ for every $\varphi \in T$ and $1 \le j \le r$. It follows that in a neighborhood $N_0 \subseteq N$ of u_0, the curve D has the r parameterizations $\varphi \mapsto (\cos\varphi : \sin\varphi : \lambda_j(\varphi))$ for $\varphi \in T_0 \subseteq T$. As none of the points in $D \cap N_0$ belongs to M, this gives $\lambda_j(\varphi) = \lambda_1(\varphi)$ for all $\varphi \in T_0$ and $1 \le j \le r$. In particular, for every $\varphi \in T_0$ the eigenspace of $\mathscr{A}(\varphi)$ corresponding to $-\lambda_1(\varphi)$ has dimension r. Then the tangent to D_A at the point $(\cos\varphi_0 : \sin\varphi_0 : \lambda_1(\varphi_0))$ is given by the formula

$$\begin{aligned}&(\cos\varphi_0 : \sin\varphi_0 : \lambda_1(\varphi_0)) \vee (-\sin\varphi_0 : \cos\varphi_0 : \lambda_1'(\varphi_0))\\ &\quad= [-\cos\varphi_0\lambda_1(\varphi_0) + \sin\varphi_0\lambda_1'(\varphi_0) : -\sin\varphi_0\lambda_1(\varphi_0) - \cos\varphi_0\lambda_1'(\varphi_0) : 1].\end{aligned} \tag{2}$$

By Lemma 3.2, there exists a regular analytic function $x : T_0 \longrightarrow \mathbb{S}(\mathbb{C}^n)$ such that $\mathscr{A}(\varphi)x(\varphi) = -\lambda_1(\varphi)x(\varphi)$ for all $\varphi \in T_0$. Moreover, we have $\lambda_1(\varphi_0) = -\langle \mathscr{A}(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle$ and $\lambda_1'(\varphi_0) = -\langle \mathscr{A}'(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle$.

It follows that the tangent to D_A at the point $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$ is given by

$$\begin{aligned}&[\cos\varphi_0\langle \mathscr{A}(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle - \sin\varphi_0\langle \mathscr{A}'(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle :\\ &\quad \sin\varphi_0\langle \mathscr{A}(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle + \cos\varphi_0\langle \mathscr{A}'(\varphi_0)x(\varphi_0), x(\varphi_0)\rangle : 1]\\ &\quad = [\langle A_1x(\varphi_0), x(\varphi_0)\rangle : \langle A_2x(\varphi_0), x(\varphi_0)\rangle : 1]\end{aligned}$$

Its image under π, that is, the point $(\langle A_1x(\varphi_0), x(\varphi_0)\rangle : \langle A_2x(\varphi_0), x(\varphi_0)\rangle : 1) \in C_A$, is the image under $\nu_A \circ \sigma = \tau^{-1} \circ n_A$ of the critical point $p(\varphi_0) = J_n^{-1}(x(\varphi_0))$.

Let S be the set of points of C_A which correspond to the tangents to D_A at simple points. Note that S contains all components of C_A which consist of single points. We have shown above that all points $u \in S$ belong to $(\nu_A \circ \sigma)(\Sigma_0(A) \cup \Sigma_1(A))$. However, as D_A has only finitely many singular points and none of these is a component of D_A, the set $C_A \setminus S$ is at most finite. It follows that $\tau(C_A)$ is the closure of $\tau(S)$ in $\mathbb{R}^2$. Since $n_A(\Sigma_0(A) \cup \Sigma_1(A))$ is a closed set in $\mathbb{R}^2$ which contains $\tau(S)$, $\tau(C_A)$ must be a subset of $n_A(\Sigma_0(A) \cup \Sigma_1(A))$ as well. Hence $C_A \subseteq (\nu_A \circ \sigma)(\Sigma_0(A) \cup \Sigma_1(A))$. □

From the proof of Proposition 3.3 it is clear that the curve C_A coincides with the image of the critical points of n_A if $f_A = m_A$ and D_A is a curve without singularities. In order to clarify the general situation we need the following lemma.

LEMMA 3.4. *Suppose* $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$ *is a point of* D_A. *Let* E *be the eigenspace of the operator* $\mathscr{A}(\varphi_0)$ *corresponding to its eigenvalue* $-\lambda_0$ *and* P *be the orthogonal projection of* $\mathbb{C}^n$ *onto* E. *Let* K *be the set of points of* C_A *at which the line* $[\cos\varphi_0 : \sin\varphi_0 : \lambda_0]$ *is tangent to* C_A. *Then* $(J_1 \circ \tau)(K)$ *is the set of eigenvalues of the operator* $PA|_E$, *and* $(n_A \circ J_n^{-1})(E)$ *is the convex hull of* $\tau(K)$.

PROOF. Let r be the dimension of E, and let $\lambda_j(\cdot)$ and $x_j(\cdot)$ with $1 \le j \le r$ be the regular analytic eigenvalue and eigenvector functions as given by Lemma 3.2. Then in a neighborhood of $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$, the curve D_A has the r (not necessarily distinct) parameterizations $\Phi_j : \varphi \longmapsto (\cos\varphi : \sin\varphi : -\lambda_j(\varphi))$. Following the argument in the proof of Proposition 3.3, we infer that for $1 \le j \le r$, the projective line $[\langle A_1x_j(\varphi_0), x_j(\varphi_0)\rangle : \langle A_2x_j(\varphi_0), x_j(\varphi_0)\rangle : 1]$ is the tangent to D_A at the point $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$ along the parameterization Φ_j. The duality of the algebraic curves C_A and D_A yields that the line $[\cos\varphi_0 : \sin\varphi_0 : \lambda_0]$ is a tangent to the curve C_A at the point $(\langle A_1x_j(\varphi_0), x_j(\varphi_0)\rangle : \langle A_2x_j(\varphi_0), x_j(\varphi_0)\rangle : 1)$. This can be verified by computing the tangent directly using that, in a neighborhood of the point $(\langle A_1x_j(\varphi_0), x_j(\varphi_0)\rangle : \langle A_2x_j(\varphi_0), x_j(\varphi_0)\rangle : 1)$, the curve C_A admits the parameterization

$$\varphi \longmapsto (\cos\varphi\lambda_j(\varphi) - \sin\varphi\lambda_j'(\varphi) : \sin\varphi\lambda_j(\varphi) + \cos\varphi\lambda_j'(\varphi) : 1)$$

(see equation (2)). In particular, the duality of C_A and D_A yields that K is the set $\{(\langle A_1x_j(\varphi_0), x_j(\varphi_0)\rangle : \langle A_2x_j(\varphi_0), x_j(\varphi_0)\rangle : 1) \mid 1 \le j \le r\}$.

It follows that the r vectors $x_j(\varphi_0)$, which form an orthonormal basis of E, belong to $(J_n \circ n_A^{-1} \circ \tau)(K)$. They are eigenvectors of the operator $P\mathscr{A}'(\varphi_0)|_E$. In fact, by Lemma 3.2, we have $\lambda_j'(\varphi_0) = \langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_j(\varphi_0)\rangle$ for every $1 \le j \le r$. By [11, Theorem 7], the derivatives $\lambda_j'(\varphi_0)$ are the eigenvalues of the $(r \times r)$-matrix $\left(\langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_k(\varphi_0)\rangle\right)_{1 \le j,k \le r}$. Hence $\langle \mathscr{A}'(\varphi_0)x_j(\varphi_0), x_k(\varphi_0)\rangle = 0$ for all $j \ne k$.

This implies $n'_B\left(J_n^{-1}(x_j(\varphi_0))\right) = 0$ for all $1 \le j \le r$, where B denotes the operator $e^{-i\varphi_0}PA|_E = P\mathscr{A}(\varphi_0)|_E + iP\mathscr{A}'(\varphi_0)|_E$. By Corollary 2.5, the r vectors $x_j(\varphi_0)$ are joint eigenvectors of the pairs $(P\mathscr{A}(\varphi_0)|_E, P\mathscr{A}'(\varphi_0)|_E)$ and $(PA_1|_E, PA_2|_E)$. Since $(n_A \circ J_n^{-1})(E) = J_1^{-1}(W(PA|_E))$, this completes the proof. □

From Lemma 3.4 we can deduce that if the curve D_A has a unique tangent at the point $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$, then $n_A \circ J_n^{-1}$ maps the eigenspace E of the operator $\mathscr{A}(\varphi_0)$ corresponding to its eigenvalue $-\lambda_0$ to a single point in $\mathbb{R}^2$ which belongs to $\tau(C_A)$. This follows from the fact that then K consists of a single point; the condition holds, in particular, for all simple points of D_A. If there exists more than one line which is tangent to D_A at $(\cos\varphi_0 : \sin\varphi_0 : \lambda_0)$ along some local parameterization of D_A, then the set K contains more than one point. In this case, $(n_A \circ J_n^{-1})(E)$ is the segment of the line $l_{\cos\varphi_0, \sin\varphi_0, -\lambda_0}$ which joins all points in $\tau(K)$. We have thus proved the following theorem.

THEOREM 3.5. *Let $A \in \mathscr{L}(\mathbb{C}^n)$. The image under $\nu_A \circ \sigma$ of the critical points of n_A consists of the curve C_A together with all line segments joining pairs of points of C_A at which C_A has the same tangent line.*

If (λ_1, λ_2) belongs to $n_A(\Sigma_0(A))$, then (λ_1, λ_2) is a joint eigenvalue of the pair (A_1, A_2). It follows that $(u\lambda_1 + v\lambda_2 + w)$ is a factor of f_A and so $\{(\lambda_1 : \lambda_2 : 1)\}$ is a component of C_A. From Theorem 3.5 one therefore obtains the image of the rank 1 critical points of n_A by subtracting all such components of C_A. However, as is well known, a component of C_A which consists only of a single point does not necessarily correspond to a joint eigenvalue of the pair (A_1, A_2). In Example 2.12, for instance, $\{(0 : 0 : 1)\}$ is a component of C_A, yet the set of stationary points is empty.

From the local representation of C_A given above, one can derive formulas for the curvature of $\tau(C_A)$ at all but a finite number of points.

PROPOSITION 3.6. *Let $(u_0, v_0, 1) \in C_A$, and $[\cos\varphi_0, \sin\varphi_0, \lambda_0]$ be a tangent to C_A at $(u_0, v_0, 1)$. If $(\cos\varphi_0, \sin\varphi_0, \lambda_0)$ is a simple point of D_A, then the signed curvature k_0 of $\tau(C_A)$ at the point (u_0, v_0) is given by*

$$k_0 = \left|\frac{2}{r}\operatorname{trace}\left(\mathscr{A}'(\varphi_0)S\mathscr{A}'(\varphi_0)P\right)\right|^{-1}, \tag{3}$$

where r is the multiplicity of the eigenvalue $-\lambda_0$ of $\mathscr{A}(\varphi_0)$, P is the orthogonal projection onto the eigenspace of $\mathscr{A}(\varphi_0)$ corresponding to $-\lambda_0$ and S denotes the reduced resolvent (or Moore-Penrose inverse) of $\mathscr{A}(\varphi_0) + \lambda_0$.

PROOF. Suppose $(\cos\varphi_0, \sin\varphi_0, \lambda_0)$ is a simple point of D_A. Then by Lemma 3.2 there exists a unique local eigenvalue function $\lambda(\cdot)$ such that $\lambda(\varphi_0) = -\lambda_0$. From

https://doi.org/10.1017/S1446788700034996 Published online by Cambridge University Press

equation (2) it follows that in a neighborhood of (u_0, v_0) the curve $\tau(C_A)$ admits the parameterization $\varphi \longmapsto (u(\varphi), v(\varphi))$, say for $\varphi \in T_0$, where $u(\varphi) = \cos\varphi\lambda(\varphi) - \sin\varphi\lambda'(\varphi)$ and $v(\varphi) = \sin\varphi\lambda(\varphi) + \cos\varphi\lambda'(\varphi)$.

The signed curvature $k(\varphi)$ at any point $(u(\varphi), v(\varphi))$ with $\varphi \in T_0$ can then be computed by the formula

$$k(\varphi) = \frac{u'(\varphi)v''(\varphi) - u''(\varphi)v'(\varphi)}{(u'(\varphi)^2 + v'(\varphi)^2)^{3/2}}.$$

This gives $k_0 = |\lambda(\varphi_0) + \lambda''(\varphi_0)|^{-1}$ (see [1, p. 295]).

From the perturbation results of [9, Section II.2], applied to the map $\psi \longmapsto \mathscr{A}(\psi + \varphi_0) = \cos(\psi)\mathscr{A}(\varphi_0) + \sin(\psi)\mathscr{A}'(\varphi_0)$, we obtain a power series expansion of the eigenvalue function $\psi \longmapsto \tilde{\lambda}(\psi) = \lambda(\psi + \varphi_0)$ at $\psi = 0$, that is, $\tilde{\lambda}(\psi) = \sum_{k=0}^{\infty} \Lambda_k \psi^k$. For the details on the coefficients Λ_k, we refer to [9, Section II.2]. It follows that $\lambda''(\varphi_0) = \tilde{\lambda}''(0) = 2\Lambda_2 = \frac{2}{r}\operatorname{trace}\left(-\frac{1}{2}\mathscr{A}(\varphi_0)P - \mathscr{A}'(\varphi_0)S\mathscr{A}'(\varphi_0)P\right)$ with P and S as above. This completes the proof. □

Note that k_0 is allowed to be infinity; this situation will occur, for example, if $\{(u_0, v_0, 1)\}$ is a component of C_A.

REMARK 3.7. Let $x(\cdot)$ be any of the r (local) eigenvector functions $x_j(\cdot)$ obtained in Lemma 3.2. Then we have $\lambda(\varphi) = \langle \mathscr{A}(\varphi)x(\varphi), x(\varphi)\rangle$ and $\lambda'(\varphi) = \langle \mathscr{A}'(\varphi)x(\varphi), x(\varphi)\rangle$ for all φ in the domain of $x(\cdot)$. Using this representation of $\lambda'(\varphi)$, an argument similar to the one applied in the proof of Lemma 3.2 yields that $\lambda''(\varphi) = -\lambda(\varphi) + 2\operatorname{Re}\langle \mathscr{A}'(\varphi)x(\varphi), x'(\varphi)\rangle$. Hence, the curvature of $\tau(C_A)$ at (x_0, y_0) is also given by

$$k_0 = \left|2\operatorname{Re}\langle \mathscr{A}'(\varphi_0)x(\varphi_0), x'(\varphi_0)\rangle\right|^{-1}.$$

REMARK 3.8. Proposition 3.6 yields in particular the curvature of the numerical range $W(A)$ at certain boundary points. Let $(u_0, v_0) \in \partial W(A)$. If there exists a supporting line $l_{c,s,\lambda}$ to $\partial W(A)$ at (u_0, v_0) such that $(c, s, -\lambda)$ is a simple point of D_A, then the curvature of $\partial W(A)$ at (u_0, v_0) is given by the formulas above. A sufficient but not necessary condition for $(c, s, -\lambda)$ to be a simple point of D_A is that λ is a simple eigenvalue of the operator $cA_1 + sA_2$. In this case, our formula for k_0 is the reciprocal of the corresponding formula for the radius of curvature derived from [6, Theorem 3.3]. Thus, Proposition 3.6 is the natural generalization of [6, Theorem 3.3].

References

[1] J. Bazer and D. H. Y. Yen, 'Lacunas of the Riemann matrix of symmetric-hyperbolic systems in two space variables', *Comm. Pure Appl. Math.* **22** (1969), 279–333.

[2] P. Binding and C-K. Li, 'Joint ranges of Hermitian matrices and simultaneous diagonalization', *Linear Algebra Appl.* **151** (1991), 157–167.

[3] J. Bochnak, M. Coste and M.-F. Roy, *Géométrie Algébrique Réelle* (Springer, New York, 1987).

[4] G. E. Bredon, *Topology and Geometry*, Graduate Texts in Mathematics 139 (Springer, New York, 1993).

[5] R. Engelking, *General Topology*, 2nd edition (Heldermann, Berlin, 1992).

[6] M. Fiedler, 'Geometry of the numerical range of matrices', *Linear Algebra Appl.* **37** (1981), 81–96.

[7] ———, 'Numerical range of matrices and Levinger's theorem', *Linear Algebra Appl.* **220** (1995), 171–180.

[8] J. A. Hillman, B. R. F. Jefferies, W. J. Ricker and B. Straub, 'Differential properties of the numerical range map of pairs of matrices', *Linear Algebra Appl.* **267** (1997), 317–334.

[9] T. Kato, *Perturbation Theory for Linear Operators*, Grundlehren der mathematischen Wissenschaften 132 (Springer, Berlin, 1980).

[10] R. Kippenhahn, 'Über den Wertevorrat einer Matrix', *Math. Nachr.* **6** (1951), 193–228.

[11] P. Lancaster, 'On eigenvalues of matrices dependent on a parameter', *Numer. Math.* **6** (1964), 377–387.

[12] ———, *Theory of Matrices* (Academic Press, New York, 1969).

[13] F. D. Murnaghan, 'On the field of values of a square matrix', *Proc. Nat. Acad. Sci. U.S.A.* **18** (1932), 246–248.

[14] M. C. Pease III, *Methods of Matrix Algebra* (Academic Press, New York, 1965).

[15] F. Rellich, 'Störungstheorie der Spektralzerlegung I', *Math. Ann.* **113** (1937), 600–619.

[16] I. R. Shafarevich, *Basic Algebraic Geometry* (Springer, New York, 1977).

[17] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, volume V, 2nd edition (Publish or Perish, Inc., Wilmington, 1979).

Fachbereich Mathematik
Technische Universität Berlin
Strasse des 17. Juni 136
D-10623 Berlin
Germany

School of Mathematics
The University of New South Wales
Sydney, NSW 2052
Australia