ON THE STAMPFLI POINT OF SOME OPERATORS AND MATRICES

THANIN QUARTZ AND ILYA M. SPITKOVSKY

(*Communicated by H. J. Woerdeman*)

Abstract. The center of mass of an operator A (denoted $\operatorname{St}(A)$, and called in this paper as the *Stampfli point* of A) was introduced by Stampfli in his Pacific J. Math (1970) paper as the unique $\lambda \in \mathbb{C}$ delivering the minimum value of $\|A - \lambda I\|$. We derive some results concerning the location of $\operatorname{St}(A)$ for several classes of operators, including 2-by-2 block operator matrices with scalar diagonal blocks and 3-by-3 matrices with repeated eigenvalues. We also show that for almost normal A its Stampfli point lies in the convex hull of the spectrum, which is not the case in general. Some relations between the property $\operatorname{St}(A) = 0$ and Roberts orthogonality of A to the identity operator are established.

1. Introduction

Let A be a bounded linear operator acting on a Hilbert space $\mathscr{H}$. In case $\dim \mathscr{H} = n < \infty$ we will identify $\mathscr{H}$ with $\mathbb{C}^n$ and A with its n-by-n matrix representation in the standard basis $\{e_1, \ldots, e_n\}$ of $\mathbb{C}^n$. We will write $A \in \mathfrak{B}(\mathscr{H})$ when the dimension of $\mathscr{H}$ is irrelevant and $A \in \mathbb{C}^{n \times n}$ to emphasize that it is finite.

We will denote the norm, the spectrum, and the numerical range of A as $\|A\|$, $\sigma(A)$ and $W(A)$, respectively. Recall that the latter (a.k.a. the *field of values*, or the *Hausdorff set* of A) is defined as

$$W(A) = \{\langle Ax, x \rangle \colon x \in \mathscr{H}, \ \|x\| = 1\}. \tag{1.1}$$

This notion goes back to classical papers by Toeplitz [17] and Hausdorff [9]; [7] is a more recent standard reference for the properties of $W(A)$. It is known in particular that $W(A)$ is convex (Toeplitz-Hausdorff theorem), its closure $\operatorname{cl} W(A)$ contains $\sigma(A)$, and thus the convex hull of it:

$$\operatorname{cl} W(A) \supseteq \operatorname{conv} \sigma(A). \tag{1.2}$$

Operators for which the equality in (1.2) holds are called *convexoid*; this class includes in particular all normal (and even hyponormal) operators. On the other hand, already

for non-normal $A \in \mathbb{C}^{2 \times 2}$ the inclusion in (1.2) is strict: $\operatorname{conv} \sigma(A)$ is then the line segment connecting the eigenvalues λ_1, λ_2 of A while $W(A)$ is an elliptical disk with the foci at λ_1, λ_2 (the Elliptical Range Theorem).

A variation of (1.1) is the so called *maximal* numerical range $W_0(A)$ consisting of the limits of all convergent sequences $\langle Ax_n, x_n \rangle$ with unit vectors $x_n \in \mathcal{H}$ such that $\|Ax_n\| \to \|A\|$. This notion was introduced by J. Stampfli in [15], where it was also observed that $W_0(A)$ is a closed convex subset of $\operatorname{cl} W(A)$. For $A \in \mathbb{C}^{n \times n}$, $W_0(A) = W(B)$, where B is the compression of A onto the eigenspace of A^*A corresponding to its maximal eigenvalue.

In the same paper [15], J. Stampfli introduced the *center of mass* of A as the (unique) value of $\lambda \in \mathbb{C}$ at which the minimum of $\|A - \lambda I\|$ is attained. Since this term is overused, and to give credit where it is due, we will call this value of λ the *Stampfli point* of A and denote it $\operatorname{St}(A)$. In this notation, according to [15, Corollary]:

$$\operatorname{St}(A) = \lambda \text{ if and only if } 0 \in W_0(A - \lambda I). \tag{1.3}$$

Note that the statements in (1.3) are **not** equivalent to $\lambda \in W_0(A)$, since the maximal numerical range does not behave nicely under shifts.

Several other observations made in [15] are as follows:

If A is normal (or even hyponormal), then $\operatorname{St}(A)$ is the center of the smallest circle circumscribing $\sigma(A)$. In general, $\operatorname{St}(A)$ lies in the closure of the numerical range of A but not necessarily in the convex hull of its spectrum. It is also mentioned in passing that the respective examples exist already when A is nilpotent and $\dim \mathcal{H} = 3$, with yet another example delivered by a (quasinilpotent) Volterra operator V, but no specifics were provided.

In a much more recent paper [6] it was established that $\operatorname{St}(V) = 1/\alpha^2$, where α is the root of $\alpha + \tan \alpha = 0$ lying in $(\pi/2, \pi)$. We are not aware of any other work in this direction.

In this paper we further explore properties of the Stampfli point. Section 2 provides an explicit formula for $\operatorname{St}(A)$ when A is unitarily similar to 2-by-2 block operator matrices with the diagonal blocks being scalar multiples of the identity operator I. This covers in particular quadratic operators, as well as tridiagonal matrices with constant main diagonal. In Section 3 the property $\operatorname{St}(A) \in \operatorname{conv} \sigma(A)$ is extended from normal to so called almost normal operators. Sections 4–6 are devoted to 3-by-3 matrices. An explicit procedure for computing $\operatorname{St}(A)$ when $\sigma(A)$ is a singleton $\{\lambda\}$ is outlined in Section 5, based on some auxiliary results established in Section 4. One of these results is the criterion for $\operatorname{St}(A)$ to coincide with λ. As a generalization of the latter, in Section 6 we characterize matrices $A \in \mathbb{C}^{3 \times 3}$ with a doubleton spectrum and $\operatorname{St}(A)$ coinciding with the multiple eigenvalue. Section 7 contains some observations on the relation between the Stampfli point of A and the Roberts orthogonality of A to I. Finally, several figures illustrating results of Sections 3–5 are presented in the Appendix.

The authors are thankful to the anonymous referee for the helpful comments.

2. Operators with scalar diagonal blocks

Let us start with the simplest possible case, in which the answer is explicit and can be obtained directly by a straightforward computation.

PROPOSITION 1. *Let $A \in \mathbb{C}^{2 \times 2}$. Then* $\mathrm{St}(A) = \mathrm{trace}\, A/2$.

Proof. Using a unitary similarity, put A in an upper-triangular form:

$$A = \begin{bmatrix} \lambda_1 & c \\ 0 & \lambda_2 \end{bmatrix},$$

where $\{\lambda_1, \lambda_2\} = \sigma(A)$, $c \geqslant 0$. Then, for any $\lambda \in \mathbb{C}$:

$$(A - \lambda I)^*(A - \lambda I) = \begin{bmatrix} |\lambda_1 - \lambda|^2 & c(\overline{\lambda_1} - \overline{\lambda}) \\ c(\lambda_1 - \lambda) & c^2 + |\lambda_2 - \lambda|^2 \end{bmatrix},$$

and so

$$2\|A - \lambda I\|^2 = |\lambda_1 - \lambda|^2 + |\lambda_2 - \lambda|^2 + c^2$$
$$+ \sqrt{(|\lambda_1 - \lambda|^2 - |\lambda_2 - \lambda|^2)^2 + c^4 + 2c^2(|\lambda_1 - \lambda|^2 + |\lambda_2 - \lambda|^2)}. \quad (2.1)$$

Elementary planar geometry shows that the minimum with respect to λ in the right hand side of (2.1) is attained when λ is a midpoint of $[\lambda_1, \lambda_2]$. □

This proposition was implicitly stated in [16], as was also observed in [5], and independently established in [1] as part of the proof of their Lemma 5(b). Note that it can also be proven based on (1.3) and the explicit formula from [8] for $W_0(A)$ in the case of 2-by-2 matrices.

THEOREM 2. *Let $A \in \mathbb{C}^{n \times n}$ be unitarily similar to a matrix of the form*

$$\begin{bmatrix} a_1 I_{n_1} & X \\ Y^* & a_2 I_{n_2} \end{bmatrix}, \quad (2.2)$$

with $XY^* \in \mathbb{C}^{n_1 \times n_1}$, $Y^*X \in \mathbb{C}^{n_2 \times n_2}$ being normal. Then $\mathrm{St}(A) = (a_1 + a_2)/2$.

Proposition 1 is of course a particular case of Theorem 2 (corresponding to $n_1 = n_2 = 1$), but at the same time also the main ingredient of its proof.

Proof. As was shown in [3] (see the proof of Theorem 2.1 there), matrices under consideration are unitarily similar to direct sums of $\min\{n_1, n_2\}$ two-dimensional blocks A_k, all having a_1, a_2 as its diagonal entries, with $|n_1 - n_2|$ one-dimensional blocks, equal a_1 or a_2. According to Proposition 1, $\mathrm{St}(A_k) = (a_1 + a_2)/2$ does not depend on $k = 1, \ldots, \min\{n_1, n_2\}$. Since for any $\lambda \in \mathbb{C}$, $\|A_k - \lambda I\| \geqslant |a_j - \lambda|$ $(j = 1, 2)$, the value of $\mathrm{St}(A)$ for the whole matrix A coincides with that of its blocks A_k. □

The normality of XY^*, Y^*X holds in a trivial way if $Y = 0$, i.e., A is unitarily similar to

$$\begin{bmatrix} \lambda_1 I & Z \\ 0 & \lambda_2 I \end{bmatrix}, \tag{2.3}$$

with $\{\lambda_1, \lambda_2\} = \sigma(A)$. This happens if and only if A satisfies the equation

$$A^2 + pA + qI = 0 \tag{2.4}$$

with $p = -(\lambda_1 + \lambda_2)$, $q = \lambda_1 \lambda_2$.

Here is an infinite-dimensional analogue of this situation (see also [1, Lemma 5]).

THEOREM 3. *Let $A \in \mathfrak{B}(\mathscr{H})$ be a quadratic operator, i.e., (2.4) holds for some $p, q \in \mathbb{C}$. Then $\mathrm{St}(A) = -p/2$.*

Proof. As was observed in [18], for an operator A satisfying (2.4) there exists a partition $\mathscr{H} = \mathscr{H}_1 \oplus \mathscr{H}_2$ with respect to which A takes the form (2.3). But then for any $\lambda \in \mathbb{C}$

$$\|A - \lambda I\| = \|A_0 - \lambda I\|, \text{ where } A_0 = \begin{bmatrix} \lambda_1 & \|Z\| \\ 0 & \lambda_2 \end{bmatrix} \in \mathbb{C}^{2 \times 2},$$

and so $\mathrm{St}(A) = \mathrm{St}(A_0)$. It remains to invoke Proposition 1. $\square$

Note that in the setting of Theorem 3, $\sigma(A) = \{\lambda_1, \lambda_2\}$, and according to [18, Theorem 2.1] $W(A)$ is an elliptical disk with foci λ_1, λ_2 (possibly degenerating into the line segment $[\lambda_1, \lambda_2]$). So, for quadratic A the position of $\mathrm{St}(A)$ is defined by $\sigma(A)$ uniquely, and is indeed at the center of both $\sigma(A)$ and $W(A)$. This justifies to some extent the "center of mass" term for $\mathrm{St}(A)$ coined by Stampfli.

Finally, if in (2.2) $a_1 = a_2$, then no conditions on X, Y are needed and, moreover, the formula for $\mathrm{St}(A)$ holds in the infinite dimensional setting.

THEOREM 4. *Let $A \in \mathfrak{B}(\mathscr{H})$. If there exists a subspace $\mathscr{L}$ of $\mathscr{H}$ such that compressions of A onto $\mathscr{L}$ and its orthogonal complement $\mathscr{L}^\perp$ are both multiples of the identity by the same scalar a, then $\mathrm{St}(A) = a$.*

Proof. It suffices to consider the case $a = 0$. The operator A then can be represented as $\begin{bmatrix} 0 & X \\ Y^* & 0 \end{bmatrix}$ with respect to the decomposition $\mathscr{H} = \mathscr{L} \oplus \mathscr{L}^\perp$.

We need to show that the norm of $\begin{bmatrix} \lambda I & X \\ Y^* & \lambda I \end{bmatrix}$ attains its minimum with respect to λ at $\lambda = 0$. Equivalently, the rightmost point of the spectrum (or, which is the same in this case, the numerical range) of

$$\begin{bmatrix} \overline{\lambda} I & Y \\ X^* & \overline{\lambda} I \end{bmatrix} \begin{bmatrix} \lambda I & X \\ Y^* & \lambda I \end{bmatrix} = |\lambda|^2 I + \begin{bmatrix} YY^* & \overline{\lambda} X + \lambda Y \\ \lambda X^* + \overline{\lambda} Y^* & X^* X \end{bmatrix}$$

should be the smallest when $\lambda = 0$.

But this is indeed the case, simply because the norm of any block operator matrix $\begin{bmatrix} H_1 & Z \\ Z^* & H_2 \end{bmatrix}$ with fixed positive semi-definite diagonal blocks is minimal when its off-diagonal blocks are equal to zero. □

COROLLARY 1. *Let $A \in \mathfrak{B}(\mathcal{H})$ be such that the entries a_{ij} of its matrix in some orthonormal basis $\{f_j\}$ have the property: $a_{ij} = 0$ if $i - j$ is even and different from zero; $a_{jj} := a$ is independent of j. Then $\mathrm{St}(A) = a$.*

Indeed, such A meets conditions of Theorem 4 with $\mathcal{L} = \mathrm{Span}\{f_j : j \text{ even}\}$.

This corollary covers in particular tridiagonal matrices A with constant main diagonal.

3. Almost normal operators

We adopt the definition of almost normality which for $A \in \mathbb{C}^{n \times n}$ was introduced in [10] as having at least $n - 1$ pairwise orthogonal eigenvectors. We will therefore say that $A \in \mathfrak{B}(\mathcal{H})$ is *almost normal* if it has an invariant subspace $\mathcal{L}$ of codimension one such that $A|\mathcal{L}$ is normal.

THEOREM 5. *Let $A \in \mathfrak{B}(\mathcal{H})$ be an almost normal operator. Then*

$$\mathrm{St}(A) \in \mathrm{conv}\,\sigma(A). \tag{3.1}$$

Proof. According to the definition of almost normality, A can be represented in the matrix form

$$A = \begin{bmatrix} N & b \\ 0 & \mu \end{bmatrix} \tag{3.2}$$

with respect to the partition $\mathcal{H} = \mathcal{L} \oplus \mathbb{C}$. Here $N \in \mathfrak{B}(\mathcal{L})$ is a normal operator, b can be identified with a vector in $\mathcal{L}$, and $\mu \in \mathbb{C}$. Observe that $\sigma(A) = \sigma(N) \cup \{\mu\}$.

Suppose (3.1) does not hold. By shifting, rotating and scaling A we may then without loss of generality assume that $\mathrm{St}(A) = 0$ while $\sigma(A) \subset \{z: \mathrm{Re}\, z \geqslant 1\}$.

Any unit vector $x \in \mathcal{H}$ can, up to inconsequential unimodular scalar multiple, be written as $x = [\sqrt{t}\xi, \sqrt{1-t}]^T$. Here ξ is a unit vector in $\mathcal{L}$ and $t \in [0, 1]$. Then

$$Ax = [\sqrt{t}N\xi + \sqrt{1-t}\,b, \sqrt{1-t}\mu]^T,$$

and

$$\|Ax\|^2 = t\langle N^*N\xi, \xi\rangle + (1-t)(\|b\|^2 + |\mu|^2) + 2\sqrt{t(1-t)}\,\mathrm{Re}\langle N\xi, b\rangle. \tag{3.3}$$

Using the symbolic calculus for normal operators (see e.g. [14, Chapter 12]),

$$N = \int_{\sigma(N)} \zeta\, dE(\zeta), \quad N^*N = \int_{\sigma(N)} |\zeta|^2\, dE(\zeta),$$

where E is the spectral decomposition of N, so (3.3) can be rewritten as

$$\|Ax\|^2 = t\int_{\sigma(N)} |\zeta|^2 \, d\langle E(\zeta)\xi,\xi\rangle + (1-t)(\|b\|^2 + |\mu|^2)$$
$$+ 2\sqrt{t(1-t)}\,\mathrm{Re}\langle \int_{\sigma(N)} \zeta\, dE(\zeta)\xi, b\rangle.$$

Considering vectors of the form $U\xi$ along with ξ, where

$$U = \int_{\sigma(N)} \phi(\zeta)\, dE(\zeta), \quad |\phi| = 1 \text{ on } \sigma(N),$$

are unitary operators commuting with N, we conclude that $\|A\|^2$ is the supremum of

$$t\int_{\sigma(N)} |\zeta|^2 d\langle E(\zeta)\xi,\xi\rangle + (1-t)(\|b\|^2 + |\mu|^2) + 2\sqrt{t(1-t)} \int_{\sigma(N)} |\zeta| |\langle dE(\zeta)\xi,b\rangle|$$

(3.4)

taken over $t \in [0,1]$ and unit vectors $\xi \in \mathscr{L}$.

A similar reasoning applied to $A - I$ in place of A yields the conclusion that $\|A - I\|^2$ is the supremum of

$$t\int_{\sigma(N)} |\zeta - 1|^2 d\langle E(\zeta)\xi,\xi\rangle + (1-t)(\|b\|^2 + |\mu - 1|^2)$$
$$+ 2\sqrt{t(1-t)}\int_{\sigma(N)} |\zeta - 1| |\langle dE(\zeta)\xi,b\rangle| \quad (3.5)$$

over the same set.

Since $\mathrm{Re}\,\zeta, \mathrm{Re}\,\mu \geqslant 1$, we have

$$|\zeta|^2 - |\zeta - 1|^2 \geqslant 1, \quad |\mu|^2 - |\mu - 1|^2 \geqslant 1, \quad |\zeta| \geqslant |\zeta - 1|,$$

and so the difference between (3.4) and (3.5) is not smaller than

$$t\int_{\sigma(N)} d\langle E(\zeta)\xi,\xi\rangle + 1 - t = 1.$$

But then $\|A\| > \|A - I\|$, in contradiction with $\mathrm{St}(A) = 0$. $\quad\square$

For the convenience of readers interested in finite-dimensional setting only, let us provide a (shorter and more elementary) adaptation of this proof to the case of almost normal $A \in \mathbb{C}^{n\times n}$.

Proof of Theorem 5 in the finite dimensional setting. Via an appropriate unitary similarity, for an almost normal $A \in \mathbb{C}^{n\times n}$ its representation (3.2) can be further specified to

$$A = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 & b_1 \\ 0 & \lambda_2 & \dots & 0 & b_2 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & b_{n-1} \\ 0 & \dots & \dots & 0 & \mu \end{bmatrix}, \quad (3.6)$$

with $b_j \geqslant 0$, $j = 1, \ldots, n-1$.

For $\xi = [\xi_1, \ldots, \xi_n]^T \in \mathbb{C}^n$, denote $|\xi_j| = t_j$, and without loss of generality set $\xi_n = t_n$. Since the j-th entry of $A\xi$ is $\lambda_j \xi_j + b_j t_n$, its maximal absolute value is attained when

$$\arg \xi_j = -\arg \lambda_j \text{ if } \lambda_j b_j t_n \neq 0, \quad j = 1, \ldots, n-1. \tag{3.7}$$

So, condition (3.7) holds for ξ maximazing $\|A\xi\| / \|\xi\|$. Therefore, for such ξ

$$\langle A\xi, \xi \rangle = \sum_{j=1}^{n-1} \left(\lambda_j t_j^2 + \overline{\xi_j} b_j t_n \right) + \mu t_n^2 \tag{3.8}$$

is a linear combination of the points in $\sigma(A)$ and $\overline{\xi_j}$ with non-negative coefficients not all equal to zero.

Suppose now that (3.1) does not hold. Shifting and rotating A (as it was done in the proof for the general setting) we may without loss of generality assume that $\text{St}(A) = 0$ while $\lambda_1, \ldots, \lambda_{n-1}, \mu$ all have positive real parts. According to (3.7) and (3.8), $\langle A\xi, \xi \rangle$ then also has a positive real part and is therefore different from zero. This is in contradiction with (1.3). $\square$

Recall that an almost normal matrix is *pure* if it is unitarily irreducible. According to [12, Theorem 2.1] this is the case if and only if in (3.6) all λ_j are distinct and $b_j \neq 0$, $j = 1, \ldots n-1$. The following refinement of the inclusion (3.1) in the finite dimensional case is of some interest.

COROLLARY 2. *Let A be as in (3.6) with all b_j different from zero. Then $\text{St}(A)$ lies in the relative interior of* $\text{conv}\{\lambda_1, \ldots, \lambda_{n-1}, \mu\}$.

Proof. We will continue using the notation $|\xi_j| = t_j$ along with the convention $\xi_n = t_n$. Since $\|A\| \geqslant \sqrt{|\lambda_j|^2 + b_j^2}$, under the condition imposed on A we have $\|A\| > |\lambda_j|$ for all $j = 1, \ldots, n-1$. So, ξ cannot be orthogonal to e_n, i.e., $t_n > 0$.

Moreover, since

$$A^*A = \begin{bmatrix} |\lambda_1|^2 & 0 & \ldots & 0 & & \overline{\lambda_1} b_1 \\ 0 & |\lambda_2|^2 & \ldots & 0 & & \overline{\lambda_2} b_2 \\ \vdots & & \ddots & & & \vdots \\ 0 & \ldots & 0 & |\lambda_{n-1}|^2 & & \overline{\lambda_{n-1}} b_{n-1} \\ \lambda_1 b_1 & \lambda_2 b_2 & \ldots & \lambda_{n-1} b_{n-1} & b_1^2 + \ldots + b_{n-1}^2 + |\mu|^2 \end{bmatrix},$$

the j-th entry of $A^*A\xi$ is $|\lambda_j|^2 \xi_j + \overline{\lambda_j} b_j t_n$ ($j = 1, \ldots, n-1$). So, t_j can only equal zero if $\lambda_j = 0$.

With these observations at hand, suppose that $\text{St}(A)$ is not in the relative interior of $\text{conv}\{\lambda_1, \ldots, \lambda_{n-1}, \mu\}$. As in the proof above, we may assume that $\text{St}(A) = 0$, but instead of $\text{Re}\,\lambda_j, \text{Re}\,\mu$ being positive we have a weaker condition $\text{Re}\,\lambda_j \geqslant 0, \text{Re}\,\mu \geqslant 0$,

with at least one of the inequalities strict. Since now all t_j and b_j are non-zero, we still may conclude from (3.8) that $\operatorname{Re}\langle A\xi, \xi \rangle > 0$, in contradiction with (1.3). $\square$

Note that all $A \in \mathbb{C}^{2 \times 2}$ are either normal or pure almost normal, while Proposition 1 shows that for such matrices the statement of Corollary 2 holds in both cases. However, already for $n = 3$ there exist unitarily reducible and almost normal (or even normal) matrices A with $\operatorname{St}(A)$ lying on the relative boundary of $\operatorname{conv}\sigma(A)$.

4. 3-by-3 matrices with singleton spectra. Auxiliary statements

Let $A \in \mathbb{C}^{3 \times 3}$ be such that its spectrum is a singleton: $\sigma(A) = \{\lambda\}$. Then, up to a unitary similarity,

$$A = \begin{bmatrix} \lambda & x & y \\ 0 & \lambda & z \\ 0 & 0 & \lambda \end{bmatrix}. \tag{4.1}$$

PROPOSITION 6. *The matrix* (4.1) *has* $\operatorname{St}(A) = \lambda$ *if and only if* $xyz = 0$.

Note that for matrices (4.1) $\operatorname{conv}\sigma(A) = \{\lambda\}$. So, condition $xyz = 0$ is actually the criterion for $St(A) \in \operatorname{conv}\sigma(A)$ to hold in this case.

Proof. To make use of (1.3), compute

$$(A - \lambda I)^*(A - \lambda I) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & |x|^2 & \overline{x}y \\ 0 & x\overline{y} & |y|^2 + |z|^2 \end{bmatrix}.$$

If $xy \neq 0$, then the maximal eigenvalue of $(A - \lambda I)^*(A - \lambda I)$ is simple, and the respective eigenvector is $\xi = [0, \xi_2, \xi_3]^T$ with $\xi_2, \xi_3 \neq 0$. But then $\langle (A - \lambda I)\xi, \xi \rangle = z\overline{\xi_2}\xi_3$. So, condition $W_0(A - \lambda I) \ni 0$ holds in this case if and only if $z = 0$.

On the other hand, if $xy = 0$, then one of the standard basis vectors e_2 or e_3 maximizes the norm of $A - \lambda I$ while $\langle (A - \lambda I)e_j, e_j \rangle = 0$. $\square$

Our goal in the next Section 5 is to compute $\operatorname{St}(A)$ for matrices (4.1) with $xyz \neq 0$. An intermediate step in this direction is a characterization of such matrices with $\lambda = 1$ and $\operatorname{St}(A) = 0$. By an additional (in this case, diagonal) unitary similarity we can arrange that $x, z > 0$.

PROPOSITION 7. *Let the matrix A be given by* (4.1) *in which* $\lambda = 1$ *and* $x, z > 0$. *Let also* $\operatorname{St}(A) = 0$. *Then* $y < 0$ *and*

$$\det \begin{bmatrix} a_1 & 0 & b_1 & 0 & 0 \\ a_2 & a_1 & b_2 & b_1 & 0 \\ a_3 & a_2 & b_3 & b_2 & b_1 \\ a_4 & a_3 & 0 & b_3 & b_2 \\ 0 & a_4 & 0 & 0 & b_3 \end{bmatrix} = 0, \tag{4.2}$$

where

$$a_1 = 2x, \; a_2 = -3xz, \; a_3 = xz^2, \; a_4 = 2y - xz, \; b_1 = 4x^2 + y^2 + z^2 - xyz,$$
$$b_2 = -(z^3 + x^2z + y^2z + 6xy - xyz^2), \; b_3 = x^2 + 4y^2 + z^2 - xyz. \tag{4.3}$$

Proof. According to (1.3), there exists a unit vector $\xi = [\xi_1, \xi_2, \xi_3]^T$ such that $\|A\xi\| = \|A\|$ and $\langle A\xi, \xi \rangle = 0$. Let us denote $|\xi_j| = t_j$; without loss of generality $\xi_3 \geqslant 0$ and therefore $\xi_3 = t_3$.

A direct computation shows that

$$A\xi = \begin{bmatrix} \xi_1 + x\xi_2 + y\xi_3 \\ \xi_2 + z\xi_3 \\ \xi_3 \end{bmatrix}$$

and so

$$\|A\xi\|^2 = t_3^2 + |\xi_1 + x\xi_2 + y\xi_3|^2 + |\xi_2 + z\xi_3|^2. \tag{4.4}$$

With ξ_2, ξ_3 and $|\xi_1|$ being fixed, the right hand side of (4.4) is maximal when

$$\overline{\xi_1}(x\xi_2 + y\xi_3) \geqslant 0, \tag{4.5}$$

and thus can be rewritten as

$$t_1^2 + t_3^2 + |x\xi_2 + y\xi_3|^2 + |\xi_2 + z\xi_3|^2 + 2t_1 |x\xi_2 + y\xi_3|.$$

At the same time

$$\langle A\xi, \xi \rangle = 1 + \overline{\xi_1}(x\xi_2 + y\xi_3) + z\overline{\xi_2}\xi_3 = 1 + \overline{\xi_1}(x\xi_2 + yt_3) + z\overline{\xi_2}t_3,$$

which due to (4.5) implies

$$\langle A\xi, \xi \rangle = 1 + t_1 |x\xi_2 + yt_3| + z\overline{\xi_2}t_3. \tag{4.6}$$

Since $z > 0$ and $t_3 \geqslant 0$, condition $\langle A\xi, \xi \rangle = 0$ can hold only if $t_3 > 0$, $\xi_2 < 0$ (and thus $\xi_2 = -t_2$).

Furthermore, $\xi = [\xi_1, -t_2, t_3]^T$ is an eigenvector of A^*A corresponding to its eigenvalue $\sigma := \|A\|^2$. Since

$$A^*A = \begin{bmatrix} 1 & x & y \\ x & x^2 + 1 & xy + z \\ \overline{y} & x\overline{y} + z & |y|^2 + z^2 + 1 \end{bmatrix}, \tag{4.7}$$

equating the second entries of $A^*A\xi$ and $\sigma\xi$ we have in particular

$$x\xi_1 - t_2(x^2 + 1) + t_3(xy + z) = -\sigma t_2. \tag{4.8}$$

Suppose for a moment that $y > 0$. Then (4.7) shows that the matrix A^*A is entry-wise positive. By the Perron theorem, its maximal eigenvalue is simple and the respective eigenvectors have entries with coinciding arguments. This is in contradiction with ξ_2, ξ_3 having opposite signs.

Consequently, y cannot be a positive real number. In particular,

$$x\xi_2 + y\xi_3 (= -xt_2 + yt_3) \neq 0.$$

Comparing the first entries of $A^*A\xi$ and $\sigma\xi$ we conclude from here that also $\xi_1 \neq 0$.

Due to (4.5), then $\xi_1 = \mu(-xt_2 + yt_3)$ with some $\mu > 0$. Plugging this into (4.8) and taking the imaginary parts:

$$\mathrm{Im}\,(xyt_3(\mu + 1)) = xt_3(\mu + 1)\,\mathrm{Im}\,y = 0.$$

So, $y \in \mathbb{R}$. As it was shown above, y cannot be positive. It is also different from zero, according to Proposition 6. Thus, the first assertion of the proposition is proven: $y < 0$.

Returning to (4.5) again, we see that $\xi_1 < 0$. So finally $\xi = [-t_1, -t_2, t_3]^T$ with all t_j positive.

Plugging this into (4.6) and recalling that $t_1^2 + t_2^2 + t_3^2 = 1$:

$$t_1^2 + t_2^2 + t_3^2 + xt_1t_2 - yt_1t_3 - zt_2t_3 = 0. \tag{4.9}$$

The collinearity of $A^*A\xi$ and ξ yields two more homogeneous second degree equations in t_j:

$$xt_1^2 - xt_2^2 + x^2t_1t_2 - (xy+z)t_1t_3 + yt_2t_3 = 0 \tag{4.10}$$

and

$$yt_1^2 - yt_3^2 + (xy+z)t_1t_2 - (y^2+z^2)t_1t_3 + xt_2t_3 = 0. \tag{4.11}$$

Dividing (4.9)–(4.11) by t_3^2 and introducing new variables $u = t_1/t_3$, $v = t_2/t_3$:

$$\begin{aligned}
u^2 + v^2 + xuv - yu - zv + 1 &= 0, \\
xu^2 - xv^2 + x^2uv - (xy+z)u + yv &= 0, \\
yu^2 + (xy+z)uv - (y^2+z^2)u + xv - y &= 0.
\end{aligned} \tag{4.12}$$

From the first two equations of (4.12) it follows that

$$u = \frac{-2xv^2 + (xz+y)v - x}{z}, \tag{4.13}$$

while the first, multiplied by y and subtracted from the third, becomes

$$yv^2 - zuv + z^2u - (x+yz)v + 2y = 0. \tag{4.14}$$

Plugging (4.13) into (4.14) and the first equation of (4.12) yields the following system of two equations in one variable v:

$$\begin{aligned}
2xv^3 - 3xzv^2 &+ xz^2v + 2y - xz = 0, \\
4x^2v^4 - (4xy + 6x^2z)\,v^3 &+ (4x^2 + y^2 + z^2 + 5xyz + 2x^2z^2)\,v^2 \\
- (3x^2z + y^2z + z^3 + xyz^2 + 2xy)\,v &+ (x^2 + z^2 + xyz) = 0.
\end{aligned} \tag{4.15}$$

By some additional equivalent transformations aimed at lowering the degrees of polynomials involved, the system (4.15) can be reduced to

$$2xv^3 - 3xzv^2 + xz^2v + 2y - xz = 0,$$
$$\left(4x^2 + y^2 + z^2 - xyz\right)v^2 - \left(z^3 + x^2z + y^2z + 6xy - xyz^2\right)v \tag{4.16}$$
$$+ \left(x^2 + 4y^2 + z^2 - xyz\right) = 0.$$

It remains to observe that the determinant in (4.2) is the resultant of the left hand sides of (4.16) and thus its equality to zero is equivalent to the system (4.16) being consistent. $\square$

5. 3-by-3 matrices with singleton spectra. Main result

Admittedly, Proposition 7 does not look constructive. Nevertheless, it serves as the main ingredient in the construction of a procedure allowing to compute $\mathrm{St}(A)$ for all matrices of the form (4.1). Due to Proposition 6, only the case $xyz \neq 0$ needs to be considered.

THEOREM 8. *Let A be given by* (4.1) *with $xyz \neq 0$. Then* $\mathrm{St}(A) = \lambda + \zeta$, *where*

$$\arg \zeta = \arg(x\bar{y}z) \tag{5.1}$$

and $|\zeta|$ is a positive root of the polynomial

$$\begin{aligned}
P_A(s) = {}& 4s^5(u^2 + v^2)(u^6 + 3u^4(v^2 + w^2) + (v^2 + w^2)^3 + 3u^2(v^4 - 7v^2w^2 + w^4)) \\
&+ 4s^4uvw(4u^6 + 6u^4(2v^2 - 3w^2) + (v^2 + w^2)^2(4v^2 + w^2) + 6u^2(2v^4 - 6v^2w^2 + w^4)) \\
&+ 3s^3u^2w^2(2u^6 + 7v^6 + 13v^4w^2 + 6v^2w^4 + u^4(11v^2 - 5w^2) + 2u^2(8v^4 - 14v^2w^2 + w^4)) \\
&+ s^2u^3vw^3(9u^4 + 7v^4 + 18v^2w^2 + 6w^4 + 4u^2(4v^2 - 3w^2)) \\
&+ su^4v^2w^4(-5v^2 + 3w^2) - 3u^5v^3w^5. \tag{5.2}
\end{aligned}$$

Here for the notational convenience we have set

$$u = |x|, \ v = |y|, \text{ and } w = |z|. \tag{5.3}$$

Proof. Denoting $\mathrm{St}(A) - \lambda := \zeta$, from Proposition 6 we have $\zeta \neq 0$. So, we may consider the matrix

$$B = -\zeta^{-1}(A - (\lambda + \zeta)I) = \begin{bmatrix} 1 & -x/\zeta & -y/\zeta \\ 0 & 1 & -z/\zeta \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.4}$$

for which $\mathrm{St}(B) = 0$. A diagonal unitary similarity can be applied to put B in the form

$$\begin{bmatrix} 1 & u/|\zeta| & -ye^{-i(\arg x + \arg z - 2\arg\zeta)}/\zeta \\ 0 & 1 & w/|\zeta| \\ 0 & 0 & 1 \end{bmatrix}, \tag{5.5}$$

thus making Proposition 7 applicable. Consequently, the right upper entry of the matrix (5.5) is negative. This is equivalent to (5.1) and also allows to rewrite this entry as $-v/|\zeta|$.

Replacing x, y, z in (4.3) by $u/s, -v/s$ and w/s, respectively, and expanding the determinant in (4.2), we arrive at (5.2). □

Note that the polynomial (5.2) has a negative constant term and a positive coefficient of s^2, while the coefficient of s^5 is strictly positive unless

$$u = v = w. \tag{5.6}$$

We will see from Theorem 9 below that if (5.6) holds, then P_A is a cubic polynomial with exactly one positive root which thus delivers the correct value of $|\zeta|$. On the other hand, if (5.6) does not hold, according to the Descartes' sign rule the 5-th degree polynomial P_A has one or three positive roots. In the latter case, the inclusion $W_0(B) \ni 0$ should be checked in order to choose the "right" one.

In all our numerical experiments this was the smallest positive root, but we do not have a proof that this is always the case. Numerical examples are shown below, one with a unique root for (5.2) and one where there are three. Direct computations of $St(X)$ for various matrices X throughout the paper were carried out by the Mathematica script *NMinimize*.

EXAMPLE 1. Consider A as follows,

$$A = \begin{bmatrix} \lambda & 8 & -1 \\ 0 & \lambda & 7 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Then (5.4) takes the form

$$B = \begin{bmatrix} 1 & 8/|\zeta| & -1/|\zeta| \\ 0 & 1 & 7/|\zeta| \\ 0 & 0 & 1 \end{bmatrix}.$$

There is only one positive root, $s \approx 0.7$, to the corresponding polynomial P_B, and indeed for this value of $|\zeta|$, $St(B) = 0$. Since $\arg(x\bar{y}z) = \pi$, we then have that $\zeta \approx -0.7$, implying that $St(A) \approx \lambda - 0.7$ which is consistent with the value computed directly.

EXAMPLE 2. Now let

$$A = \begin{bmatrix} \lambda & 8 & -1 \\ 0 & \lambda & 7.5 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Then (5.4) takes the form

$$B = \begin{bmatrix} 1 & 8/|\zeta| & -1/|\zeta| \\ 0 & 1 & 7/|\zeta| \\ 0 & 0 & 1 \end{bmatrix}.$$

The respective polynomial P_B has three positive roots, approximately equal 0.83, 1.37, and 2.1. Computations show that $\mathrm{St}(B)$ equals zero only for the matrix B corresponding to the minimal value of s. So, $|\zeta| \approx 0.83$. Since $\arg(x\bar{y}z) = \pi$, it follows that $\zeta \approx -0.83$, implying that $\mathrm{St}(A) \approx \lambda - 0.83$. As in Example 1, this is consistent with the value of $\mathrm{St}(A)$ computed directly.

Computation of the displacement ζ between the spectrum $\{\lambda\}$ of the matrix (4.1) and its Stampfli point becomes much simpler under an additional condition $|x| = |z|$. This covers in particular the case $x = z$ in which A is a triangular Toeplitz matrix.

THEOREM 9. *Let in* (4.1) $0 \neq |x| = |z| := u$. *Then* $\mathrm{St}(A) = \lambda + \zeta$, *where* $\arg \zeta$ *is determined by* (5.1), *while*

$$|\zeta| = \begin{cases} \dfrac{u^2 v}{u^2 - v^2} & \text{if } \dfrac{v}{u} \leqslant 2 - \sqrt{3}, \\[2ex] \dfrac{u^2(2\sqrt{6u^2 + v^2} - v)}{2(8u^2 + v^2)} & \text{otherwise,} \end{cases} \tag{5.7}$$

where $|y| := v$. *In particular,*

$$|\zeta| = \frac{2\sqrt{7} - 1}{18} u \text{ if } v = u. \tag{5.8}$$

Proof. Observe first of all that (5.7) means that $\zeta = 0$ if $y = 0$, which is in agreement with Proposition 6. So, only the case $y \neq 0$ is of interest.

Further, under the requirement $|x| = |z|$ in the notation (5.6) we have $u = w$. Polynomial (5.2) then simplifies greatly, and actually factors into the product of $(sv^2 - (s-v)u^2)^2$ by the cubic

$$4s^3 v^4 + 4s^2(9s + 2v)u^2 v^2 + s(32s^2 + 36sv + v^2)u^4 - 3(s-v)u^6.$$

So, it has a double root $u^2 v/(u^2 - v^2)$ (negative if $u < v$ and disappearing if $u = v$), and three simple roots

$$-\frac{u^2 v}{u^2 + v^2}, \quad \frac{u^2(-v \pm 2\sqrt{6u^2 + v^2})}{2(8u^2 + v^2)},$$

exactly one of which being positive. This justifies the second line of (5.7) when $u \leqslant v$, and its particular case (5.8) corresponding to $u = v$.

It remains to consider the case $u > v$ and show that the matrix

$$C = \begin{bmatrix} 1 & \frac{u^2 - v^2}{uv} & -\frac{u^2 - v^2}{u^2} \\ 0 & 1 & \frac{u^2 - v^2}{uv} \\ 0 & 0 & 1 \end{bmatrix},$$

obtained from (5.5) by plugging in ζ from the first line of (5.7) and replacing w with u, satisfies $W_0(C) \ni 0$ if and only if $\frac{v^2}{u^2} \leqslant 7 - 4\sqrt{3}$.

Direct computations show that C^*C has the maximal eigenvalue u^2/v^2 of multiplicity two, and the respective eigenspace $\mathscr{L}$ is the span of

$$\xi_1 = \begin{bmatrix} -v^2/u^2 \\ 0 \\ 1 \end{bmatrix} \text{ and } \xi_2 = \begin{bmatrix} v/u \\ 1 \\ 0 \end{bmatrix}.$$

Since $\langle C\xi_2, \xi_2 \rangle = 2 \neq 0$, we need to figure out whether

$$\langle C(\xi_1 + t\xi_2), \xi_1 + t\xi_2 \rangle = 0 \tag{5.9}$$

for some $t \in \mathbb{C}$. But

$$\langle C(\xi_1 + t\xi_2), \xi_1 + t\xi_2 \rangle = 2|t|^2 + 1 + \frac{v^2}{u^2} - 2\operatorname{Re} t \frac{v}{u} + \bar{t} \frac{u^2 - v^2}{uv}.$$

So, in order for (5.9) to hold t must be real and satisfy

$$2t^2 + t\left(\frac{u}{v} - 3\frac{v}{u}\right) + \left(1 + \frac{v^2}{u^2}\right) = 0. \tag{5.10}$$

Such t exists if and only if the discriminant of the quadratic equation (5.10) is non-negative, which amount to

$$\frac{v^2}{u^2} - 14 + \frac{u^2}{v^2} \geqslant 0,$$

or, equivalently,

$$\frac{v^2}{u^2} \in [0, 7 - 4\sqrt{3}] \cup [7 + 4\sqrt{3}, +\infty).$$

Since $u > v > 0$, only the first interval in the right hand side is of relevance. $\square$

It is worth mentioning that according to (5.7) ζ is indeed the smallest positive root of the polynomial (5.2) in the case $|x| = |z|$. Here is a numerical example.

EXAMPLE 3. Consider the matrix A as follows,

$$A = \begin{bmatrix} \lambda & 4 & -2 \\ 0 & \lambda & 4 \\ 0 & 0 & \lambda \end{bmatrix}.$$

Since $\frac{v}{u} = \frac{1}{2} > 2 - \sqrt{3}$, by (5.7) we have that

$$|\zeta| = \frac{4^2(2\sqrt{6 \times 4^2 + 2^2} - 2)}{2(8 \times 4^2 + 2^2)} = \frac{12}{11}$$

and $\arg \zeta = \arg(-2) = \pi$ which implies that $\zeta = -\frac{12}{11}$. Hence, $St(A) = \lambda - \frac{12}{11}$.

6. On 3-by-3 matrices with a two-point spectrum

A 3-by-3 matrix A with a multiple eigenvalue is unitarily similar to

$$\begin{bmatrix} \mu & x & y \\ 0 & \mu & z \\ 0 & 0 & \lambda \end{bmatrix}. \tag{6.1}$$

This class of matrices is intermediate between (4.1) and the whole $\mathbb{C}^{3\times3}$. As such, it does not admit (to the best of our knowledge) an explicit procedure for computing $\mathrm{St}(A)$ similar to Theorem 8, but still allows for an extension of Proposition 6, delivering the criterion for $\mathrm{St}(A)$ to coincide with the multiple eigenvalue of A. We will use the same notation (5.6) as in Theorem 8, in addition abbreviating $|\lambda - \mu|$ to ρ.

THEOREM 10. *A matrix A unitary similar to* (6.1) *has* $\mathrm{St}(A)$ *coinciding with its repeated eigenvalue μ if and only if the following three conditions hold:*

$$\overline{x}y\overline{z}(\lambda - \mu) \leqslant 0, \tag{6.2}$$

$$|\rho v - uw| \leqslant \sqrt{(u^2 - \rho^2)(w^2 + \rho^2)}, \tag{6.3}$$

and

$$vw\rho^3 + uv^2\rho^2 + vw(v^2 + w^2 - u^2)\rho - uv^2w^2 = 0. \tag{6.4}$$

Note that condition (6.3) implicitly contains the requirement $u \geqslant \rho$.

Proof. If $\lambda = \mu$, then $\rho = 0$ and conditions (6.2), (6.3) hold automatically, while (6.4) boils down to $xyz = 0$. This is in full agreement with Proposition 6, so we need only to deal with $\lambda \neq \mu$.

Replacing A of the form (6.1) by

$$\frac{1}{\lambda - \mu}A - \frac{\mu}{\lambda - \mu}I = \begin{bmatrix} 0 & x/(\lambda - \mu) & y/(\lambda - \mu) \\ 0 & 0 & z/(\lambda - \mu) \\ 0 & 0 & 1 \end{bmatrix},$$

we may without loss of generality suppose that in (6.1) $\lambda = 1$ and $\mu = 0$. In other words, it suffices to prove the statement for matrices

$$A = \begin{bmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & 1 \end{bmatrix}, \tag{6.5}$$

where in addition we may (as was done earlier) assume $x, z \geqslant 0$. Respectively, conditions (6.2)–(6.4) for the purpose of this proof should be replaced by

$$x \geqslant 1, \ y \leqslant 0, \ |xz + y| \leqslant \sqrt{(x^2 - 1)(z^2 + 1)} \tag{6.6}$$

and

$$yz - xy^2 + yz(y^2 + z^2 - x^2) + xy^2z^2 = 0. \tag{6.7}$$

According to (6.5),

$$A^*A = [0] \oplus \begin{bmatrix} x^2 & xy \\ x\bar{y} & |y|^2 + z^2 + 1 \end{bmatrix}. \tag{6.8}$$

Eigenvectors ξ corresponding to the maximal eigenvalue σ of A^*A therefore have the first coordinate equal to zero. Consequently,

$$A\xi = [x\xi_2 + y\xi_3, z\xi_3, \xi_3]^T, \text{ and } \langle A\xi, \xi \rangle = \xi_3(z\bar{\xi_2} + \bar{\xi_3}). \tag{6.9}$$

Case 1. $x = 0$. From (6.8) we see that ξ is collinear with e_3, and (6.9) implies $\langle A\xi, \xi \rangle = |\xi_3|^2 \neq 0$. So, $\mathrm{St}(A) \neq 0$. Since condition (6.6) fails, the statement holds.

Case 2. $y = 0$. Then condition (6.7) holds while (6.6) simplifies to $x^2 \geqslant 1 + z^2$. On the other hand, in this case $A^*A = \mathrm{diag}[0, x^2, z^2 + 1]$. So, ξ is collinear to e_3 if $x^2 < 1 + z^2$ while $\xi = e_2$ is admissible otherwise. By (6.9), $W_0(A) \ni 0$ is equivalent to $x^2 \geqslant 1 + z^2$. Once again, the statement holds.

It remains to consider

Case 3. $x, y \neq 0$. The eigenvalue σ of A^*A is then simple, and the entries ξ_2, ξ_3 of the respective eigenvector ξ are different from zero. By scaling, without loss of generality let $\xi_2 = 1$. According to (6.9), condition $\langle A\xi, \xi \rangle = 0$ then holds if and only if $\xi_3 = -z$. Rewritten entry-wise, $A^*A\xi = \sigma\xi$ takes the form

$$x^2 - xyz = \sigma, \quad x\bar{y} - (|y|^2 + z^2 + 1)z = -\sigma z. \tag{6.10}$$

Since $\sigma > |y|^2 + z^2 + 1$, the second equality in (6.10) implies that $y < 0$, and (6.7) follows. Finally, $\tau = \mathrm{trace}(A^*A) - \sigma = 1 + y^2 + z^2 + xyz$ is the second non-zero eigenvalue of A^*A, and the upper bound on $|xz + y|$ in (6.6) is necessary and sufficient for τ not to exceed σ. $\square$

7. Roberts orthogonality

Two vectors x and y of a normed linear space X are called *Roberts orthogonal* (denoted: $x \perp_R y$) if for all scalars v:

$$\|x + vy\| = \|x - vy\|.$$

If X is an inner product space, Roberts orthogonality and usual orthogonality coincide: $x \perp_R y$ if and only if $\langle x, y \rangle = 0$. While this notion goes back to [13], a more recent treatment of the case when X is a unital C^*-algebra can be found in [2]. In particular, [2, Theorem 2.4] delivers the criterion for $A \in \mathfrak{B}(\mathscr{H})$ to be Roberts orthogonal to the identity operator I. To describe this result, recall that the *Davis-Wielandt shell* of A is the subset of $\mathbb{C} \times \mathbb{R}$ (identified with $\mathbb{R}^3$) defined by

$$DW(A) = \{(\langle Ax, x \rangle, \|Ax\|^2) : x \in \mathscr{H}, \|x\| = 1\}.$$

This set is a (possibly, degenerate) ellipsoid if $\dim \mathcal{H} = 2$, and a convex body otherwise, closed if $\dim \mathcal{H} < \infty$, and located between the horizontal planes $\Pi_0 = \mathbb{C} \times \{0\}$ and $\Pi_1 = \mathbb{C} \times \{\|A\|^2\}$.

For each vertical line having a non-empty intersection with $\mathrm{cl}(DW(A))$ choose the uppermost point of this intersection, and call the union $DW_{ub}(A)$ of all such points the *upper boundary* of $DW(A)$. It was shown in [2] that $A \perp_R I$ if and only if $DW_{ub}(A) = DW_{ub}(-A)$; in other words, if and only if $DW_{ub}(A)$ is symmetric with respect to the vertical coordinate axis.

Note however that $\mathrm{cl}\,W(A)$ and $W_0(A)$ are the projections onto Π_0 of $DW_{ub}(A)$ and its intersection with Π_1, respectively. So, from $DW_{ub}(A) = DW_{ub}(-A)$ it immediately follows that both $\mathrm{cl}\,W(A)$ and $W_0(A)$ are central symmetric. The necessity of the former condition for A and I to be Roberts orthogonal (along with its sufficiency when $\dim \mathcal{H} = 2$) was observed in [2]. From the necessity of the latter and (1.3) we obtain

PROPOSITION 11. *Let $A \in \mathfrak{B}(\mathcal{H})$ be Roberts orthogonal to I. Then $\mathrm{St}(A) = 0$.*

Indeed, the set $W_0(A)$ is convex, and its central symmetry implies therefore that it contains zero.

THEOREM 12. *A quadratic operator A is Roberts orthogonal to the identity if and only if it is nilpotent or a scalar multiple of an involution.*

Proof. Necessity. Combining Proposition 11 with Theorem 3 we see that A satisfies (2.4) with $p = 0$. If in addition $q = 0$, then A is nilpotent; otherwise $(-q)^{-1/2}A$ is an involution.

Sufficiency. The operator in question is unitarily similar to (2.3) with $\lambda_1 = -\lambda_2 := a$. The observation $\|A - \lambda I\| = \|A_0 - \lambda I\|$ from the proof of Theorem 3 followed by the application of (2.1) to A_0 in place of A show that in our case

$$\|A - \lambda I\|^2 = |a|^2 + |\lambda|^2 + \|Z\|^2/2 + \sqrt{(|a|^2 + |\lambda|^2 + \|Z\|^2/2)^2 - |a^2 - \lambda^2|^2}.$$

The right hand side is indeed invariant under the change $\lambda \mapsto -\lambda$, and so $A \perp_R I$.   □

So, for quadratic operators A condition $\mathrm{St}(A) = 0$ is not only necessary but also sufficient for $A \perp_R I$. As a by-product we see that the central symmetry of $W(A)$ is yet another equivalent condition; the fact for $n = 2$ observed in [2, Proposition 2.7].

Moving to $A \in \mathbb{C}^{3 \times 3}$, we restrict our attention to matrices with circular numerical ranges.

THEOREM 13. *A matrix $A \in \mathbb{C}^{3 \times 3}$ with a circular numerical range is Roberts orthogonal to the identity if and only if it is nilpotent or unitarily similar to a direct sum of a 1-by-1 and a nilpotent 2-by-2 block.*

Proof. Necessity. Let $A \perp_R I$ while $W(A)$ is a circular disk. The central symmetry of $W(A)$ means that this disk is centered at the origin. According to [11, Corollary 2.5]

(see also [4]), such A is unitarily similar to

$$\begin{bmatrix} 0 & x & y \\ 0 & 0 & z \\ 0 & 0 & \lambda \end{bmatrix}, \tag{7.1}$$

where

$$x\bar{y}z = -\lambda(|y|^2 + |z|^2) \tag{7.2}$$

and

$$|x|^2 + |y|^2 + |z|^2 \geqslant 4|\lambda|^2. \tag{7.3}$$

In the notation of Theorem 10 condition (7.2) implies $uvw = \rho(v^2 + w^2)$. Plugging this into (6.4) (which holds, since $\mathrm{St}(A) = \mu = 0$) we conclude that $\lambda = 0$ or $z = 0$. If $\lambda = 0$, the matrix (7.1) is nilpotent. Otherwise, (7.2) implies that along with z also $y = 0$, and A is therefore unitarily similar to

$$[\lambda] \oplus B, \text{ where } B = \begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix}. \tag{7.4}$$

Sufficiency. If A is nilpotent, the circularity of its numerical range implies that in (7.1) not only $\lambda = 0$, but also $xyz = 0$, due to (7.2). For any $v \in \mathbb{C}$ the matrix $A - vI$ by a rotation through $-\arg v$ and an appropriate diagonal unitary similarity can be reduced to

$$\begin{bmatrix} |v| & |x| & |y| \\ 0 & |v| & |z| \\ 0 & 0 & |v| \end{bmatrix}.$$

So, $\|A - vI\|$ in this case does not depend on the argument of μ, only on its absolute value. In particular, $A \perp_R I$.

Finally, if A is unitarily similar to (7.4), the circularity of $W(A)$ implies that $2|\lambda| \leqslant |x|$. From here,

$$\|B - vI\|^2 = |v|^2 + \frac{|x|^2}{2} + \sqrt{|xv|^2 + \frac{|x|^4}{4}} \geqslant |v|^2 + 2|\lambda|^2 + 2|\lambda|\sqrt{|v|^2 + |\lambda|^2},$$

so $|\lambda - v| \leqslant \|B - vI\|$ for all $v \in \mathbb{C}$. Consequently,

$$\|A - vI\| = \max\{|\lambda - v|, \|B - vI\|\} = \|B - vI\|,$$

and the relation $A \perp_R I$ follows from $B \perp_R I$. $\square$

COROLLARY 3. *Let $A \in \mathbb{C}^{3 \times 3}$ have a circular numerical range. Then $A \perp_R I$ if and only if the disk $W(A)$ is centered at zero and $\mathrm{St}(A) = 0$.*

A particular example

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & -\frac{1}{2} \end{bmatrix}$$

was considered in [2]. It was shown there that $A \not\perp_R I$ via computing

$$\|A+I\| \approx 2.1617 \neq 2.1366 \approx \|A-I\|,$$

while $W(A)$ is a circular disk centered at the origin. This agrees with our Theorem 13 since this matrix is neither unitarily reducible nor nilpotent. Note also that $\mathrm{St}(A) \approx 0.02 \neq 0$, in agreement with Corollary 3.

Appendix

The figures below represent graphs of numerical ranges (bounded by green curves), spectra (blue dots) and Stampfli points (black dots) of several matrices[1]. The matrices in Fig. 1, 3 and 4 are nilpotent and unitarily irreducible, with the numerical range having each of the three possible shapes (circular, with a flat portion on the boundary, or ovular, as per [11]). The one with the circular numerical range (Fig. 1) satisfies conditions of Proposition 6, and indeed has the Stampfli point located at the origin. In Fig. 3 and 4, the Stampfli point differs from zero, and is positioned in agreement with Theorem 8. Finally, Fig. 2 illustrates that for almost normal matrices, $\mathrm{St}(A)$ lies in the interior of $\mathrm{conv}\,\sigma(A)$ (bounded by the red triangle), in agreement with Corollary 2.
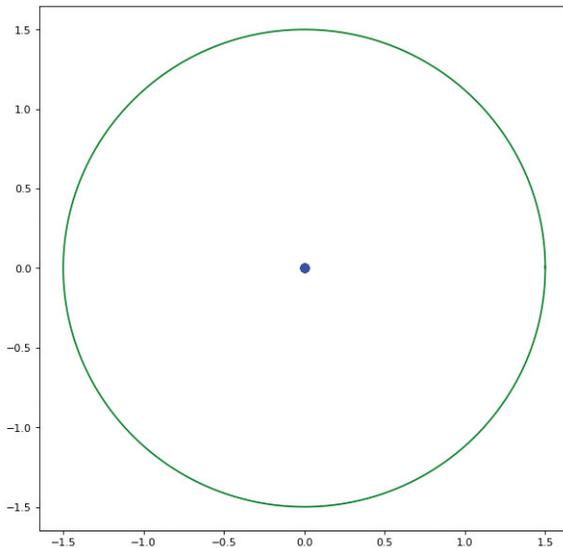

Figure 1: $A = \begin{bmatrix} 0 & 2-i & 0 \\ 0 & 0 & 2i \\ 0 & 0 & 0 \end{bmatrix}$; $St(A) = 0$.

---

[1]The color-coding refers to the online version of the paper.
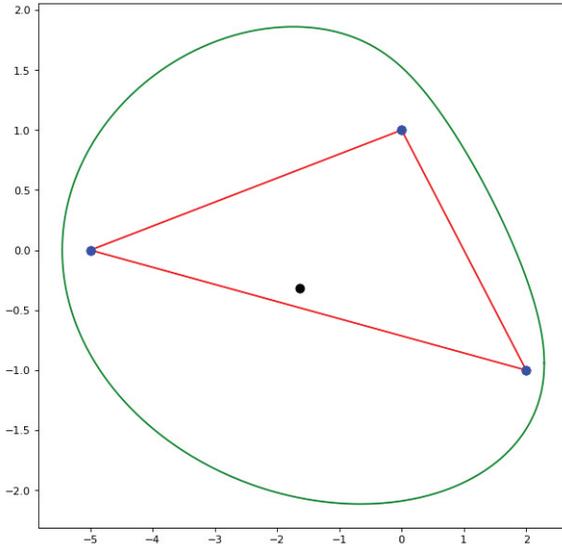
Figure 2: $A = \begin{bmatrix} 2-i & 0 & 2-2i \\ 0 & i & 2 \\ 0 & 0 & -5 \end{bmatrix}$; $St(A) \approx -1.6 - 0.3i$
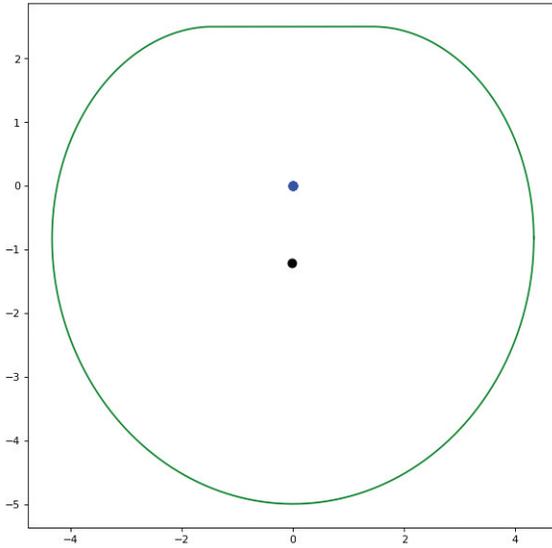

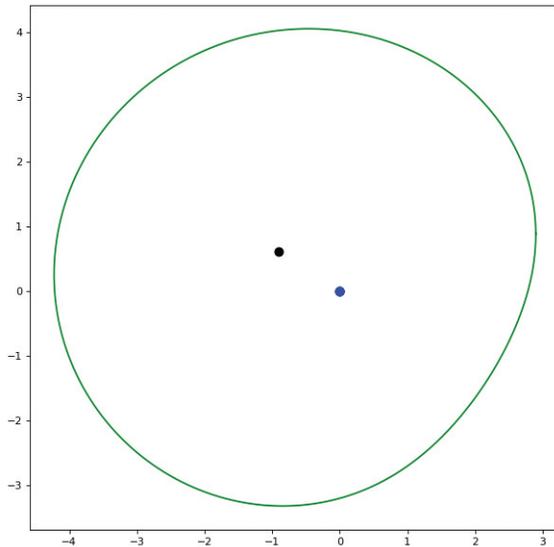Figure 3: $A = \begin{bmatrix} 0 & 3-4i & -5 \\ 0 & 0 & -4+3i \\ 0 & 0 & 0 \end{bmatrix}$; $St(A) \approx -1.2i$

Figure 4: $A = \begin{bmatrix} 0 & 1-4i & -3-2i \\ 0 & 0 & 1+5i \\ 0 & 0 & 0 \end{bmatrix}$; $St(A) \approx -0.9 + 0.6i$

REFERENCES

[1] AMER A.-O. AND P. Y. WU, *Scalar approximants of quadratic operators with applications*, Operators and Matrices **12** (2018), no. 1, 253–262.

[2] L. ARAMBAŠIĆ, T. BERIĆ, AND R. RAJIĆ, *Roberts orthogonality and Davis-Wielandt shell*, Linear Algebra Appl. **539** (2018), 1–13.

[3] E. BROWN AND I. SPITKOVSKY, *On matrices with elliptical numerical ranges*, Linear Multilinear Algebra **52** (2004), 177–193.

[4] M. CHIEN AND B. TAM, *Circularity of the numerical range*, Linear Algebra Appl. **201** (1994), 113–133.

[5] L. Z. GEVORGYAN, *On the minimal norm of a linear operator pencil*, Dokl. Nats. Akad. Nauk Armen. **110** (2010), no. 2, 97–104.

[6] L. Z. GEVORGYAN, *On the transcendental radius of the Volterra integration operator*, Ann. Funct. Anal. **6** (2015), no. 1, 54–58.

[7] K. E. GUSTAFSON AND D. K. M. RAO, *Numerical range. The field of values of linear operators and matrices*, Springer, New York, 1997.

[8] A. N. HAMED AND I. M. SPITKOVSKY, *On the maximal numeical range of some matrices*, Electron. J. Linear Algebra **34** (2018), 288–303.

[9] F. HAUSDORFF, *Der Wertvorrat einer Bilinearform*, Math. Z. **3** (1919), 314–316.

[10] KH. D. IKRAMOV, *On almost normal matrices*, Vestnik Moskov. Univ. Ser. XV, Vychisl. Mat. Kibernet. (2011), no. 1, 5–9, 56.

[11] D. KEELER, L. RODMAN, AND I. SPITKOVSKY, *The numerical range of* 3×3 *matrices*, Linear Algebra Appl. **252** (1997), 115–139.

[12] T. MORAN AND I. M. SPITKOVSKY, *On almost normal matrices*, Textos de Matemática **44** (2013), 131–144.

[13] D. B. ROBERTS, *On the geometry of abstract vector spaces*, Tôhoku Math. J. **39** (1934), 42–59.

[14] W. RUDIN, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill, Inc., New York, 1991.

[15] J. G. STAMPFLI, *The norm of a derivation*, Pacific J. Math. **33** (1970), 737–747.

[16] E. L. STOLOV, *The Hausdorff set of a matrix*, Izv. Vyssh. Uchebn. Zaved. Mat. (1979), no. 10, 98–100, English translation in Soviet Math. (Iz. VUZ) **23** (1979), no. 10, 85–87.

[17] O. TOEPLITZ, *Das algebraische Analogon zu einem Satze von Fejér*, Math. Z. **2** (1918), no. 1–2, 187–197.

[18] SHU-HSIEN TSO AND PEI YUAN WU, *Matricial ranges of quadratic operators*, Rocky Mountain J. Math. **29** (1999), no. 3, 1139–1152.

Thanin Quartz
Division of Science and Mathematics
New York University Abu Dhabi (NYUAD)
Saadiyat Island, P.O. Box 129188 Abu Dhabi, United Arab Emirates
e-mail: thanin.quartz@nyu.edu; tquartz@uwaterloo.ca

Ilya M. Spitkovsky
Division of Science and Mathematics
New York University Abu Dhabi (NYUAD)
Saadiyat Island, P.O. Box 129188 Abu Dhabi, United Arab Emirates
e-mail: ims2@nyu.edu; ilya@math.wm.edu;
imspitkovsky@gmail.com

Operators and Matrices
www.ele-math.com
oam@ele-math.com