

On the field of values of oblique projections

Valeria Simoncini and Daniel B. Szyld

Report 10-03-22
March 2010

Department of Mathematics
Temple University
Philadelphia, PA 19122

This report is available in the World Wide Web at
<http://www.math.temple.edu/~szyld>

ON THE FIELD OF VALUES OF OBLIQUE PROJECTIONS*

VALERIA SIMONCINI[†] AND DANIEL B. SZYLD[‡]

Abstract. We highlight some properties of the field of values (or numerical range) $W(P)$ of an oblique projector P on a Hilbert space, i.e., of an operator satisfying $P^2 = P$. If P is neither null nor the identity, we present a direct proof showing that $W(P) = W(I - P)$, i.e., the field of values of an oblique projection coincides with that of its complementary projection. We also show that $W(P)$ is an elliptical disk with foci at 0 and 1 and eccentricity $1/\|P\|$. These two results combined provide a new proof of the identity $\|P\| = \|I - P\|$. We discuss the relation between the minimal canonical angle between the range and the null space of P and the shape of $W(P)$. In the finite dimensional case, we show a relation between the eigenvalues of matrices related to these complementary projections and present a second proof to the fact that $W(P)$ is an elliptical disk.

Key words. Idempotent operators. Oblique Projections. Field of Values. Numerical Range.

AMS subject classifications. 15A60, 46C99, 46E99

1. Introduction. Oblique projections, i.e., idempotent operators P such that

$$P^2 = P, \tag{1.1}$$

are ubiquitous in the analysis and construction of numerical methods for the solution of large linear systems of equations [2], [8], of eigenvalue problems [1], [5], [18], of multigrid methods [9], and in particular of discretizations of partial differential equations, [16], [22], [30]. In fact, the norms of such projections often play a crucial role in the analysis of these methods. See also the influential book [27].

On the other hand, the field of values of linear operators (formally defined below) plays an important role in the analysis of convergence of certain iterative methods for the solution of algebraic linear systems [4], [7], [13], [17], [21] [25], [26].

Thus, it is natural to ask what can one say about properties of the field of values of oblique projections P . As it turns out, one can completely characterize these sets, and their shape depends only on its spectral norm $\|P\|$. While this characterization is not specifically mentioned in the literature, and it is not widely known in the matrix analysis and numerical analysis communities, it would come as no surprise to researchers in the field operator theory, since it can be obtained using the canonical forms of quadratic operators¹ and their field of values; see [19], [23], [29]. See also [15] for appropriate canonical forms in the finite dimensional case. In this paper, we show our results directly, without the use of canonical forms.

Consider a Hilbert space \mathcal{H} with inner product $\langle x, y \rangle$, and its associated norm

$$\|x\| = \langle x, x \rangle^{\frac{1}{2}} \tag{1.2}$$

(for example \mathbb{C}^n and the Euclidean inner product). We denote by $W(A)$ the field of values (or numerical range) of the operator $A : \mathcal{H} \rightarrow \mathcal{H}$, i.e., the set in the complex

*This version dated 22 March 2010

[†]Dipartimento di Matematica, Università di Bologna, Piazza di Porta S. Donato, 5, I-40127 Bologna, Italy; and CIRSA, Ravenna, Italy (valeria@dm.unibo.it).

[‡]Department of Mathematics, Temple University, Philadelphia, PA 19122, USA (szyld@temple.edu). Research supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672.

¹A quadratic operator A satisfies $A^2 + \alpha A + \beta I = O$, for some scalars α, β .

plane defined as

$$W(A) = \{\langle Au, u \rangle, u \in \mathcal{H}, \langle u, u \rangle = 1\}. \quad (1.3)$$

The field of values is very useful to study the behavior of the operator, and in particular the closure of $W(A)$ contains the spectrum of the operator. The quantity $r(A) = \sup\{|z|, z \in W(A)\}$ is known as the numerical radius of A . For further details; see, e.g., the monograph [12] and the extensive bibliography therein, or [14, Chapter 1] for many results on the finite dimensional case.

In this article, we highlight an identity relating the field of values of a projection and that of its complementary projection, namely

$$W(P) = W(I - P). \quad (1.4)$$

Here we provide a direct proof of this result; see Section 2. We also make some observations on the form of $W(A)$: it is an elliptical disk with foci on the real axis, and thus, it is symmetric with respect to that axis. This is well-known in the operator theory literature; see [19], [23], [29]. Here, we also show this directly, without going through the canonical forms. It follows from those results that $W(P^*) = W(P)$. We make a connection between the minimal canonical angle between $\mathcal{X} = \mathcal{R}(P)$ (the range of P) and $\mathcal{Y} = \mathcal{N}(P)$ (the null space of P) and the shape of this ellipse.

The fact that $W(P)$ is an elliptical disk reinforces the notion that non-trivial idempotent operators behave like operators in a two-dimensional space; one dimension corresponding to the subspace \mathcal{X} and the other to its orthogonal complement \mathcal{X}^\perp ; cf. [6] and [14, Theorem 1.3.6] that show that the field of values of operators in a two-dimensional space are elliptical disks. See Lemma 2.2 below where the projection P is given explicitly in terms of the decomposition $\mathcal{H} = \mathcal{X} \oplus \mathcal{X}^\perp$, and note also that this decomposition is used in the proofs of our theorems.

In Section 3, for the finite dimensional case, we compute the eigenvalues of the Hermitian part of $e^{i\varphi}P$, from which the identity (1.4) will follow as well. We also present a different proof of the fact that $W(P)$ is an elliptical disk, in part to highlight the use of characterizations of the field of values other than (1.3), and to show how the explicit representation of P in terms of orthogonal bases of subspaces can be helpful.

We end this introduction with some further properties of the Hilbert space and of the oblique projections. The vector norm (1.2) induces an operator norm in the usual manner, i.e., for $A : \mathcal{H} \rightarrow \mathcal{H}$, one has $\|A\| = \sup_{\|u\|=1} \|Au\|$. The adjoint operator A^* is such that for every $u, v \in \mathcal{H}$, $\langle Au, v \rangle = \langle u, A^*v \rangle$. Given any subspace \mathcal{X} of \mathcal{H} we define its orthogonal complement by

$$\mathcal{X}^\perp = \{z \in \mathcal{H}, \langle z, x \rangle = 0 \text{ for all } x \in \mathcal{X}\}.$$

It follows directly from (1.1) that $\|P\| \geq 1$. The operator P is a projection along (or parallel to) its null space $\mathcal{Y} = \mathcal{N}(P)$ onto its range $\mathcal{X} = \mathcal{R}(P)$. The operator $I - P$ is also idempotent, and it is a projection along $\mathcal{X} = \mathcal{N}(I - P) = \mathcal{R}(P)$ onto $\mathcal{Y} = \mathcal{R}(I - P) = \mathcal{N}(P)$. It is called the complementary projection to P . From (1.1) it also follows that if $x \in \mathcal{X} = \mathcal{R}(P)$, then $x = Px$. Indeed, let $x = Pu$, then $x = Pu = P^2u = Px$. Using this, one can show that for a continuous projection P , the subspaces $\mathcal{X} = \mathcal{R}(P)$ and $\mathcal{Y} = \mathcal{N}(P)$ are closed sets. These two subspaces are also complementary, i.e., $\mathcal{X} \oplus \mathcal{Y} = \mathcal{H}$. This follows from the fact that any $u \in \mathcal{H}$ can be written as $u = Pu + (I - P)u$. As a consequence of this discussion, it follows that the spectrum of P consists of two points, namely

$$\Lambda(P) = \{0, 1\}, \quad (1.5)$$

with \mathcal{X} being the invariant subspace corresponding to the eigenvalue 1, and of course \mathcal{Y} that corresponding to the eigenvalue 0. Furthermore, this also trivially implies that the spectrum of an oblique projection coincides with that of its complementary projection, i.e.,

$$\Lambda(P) = \Lambda(I - P).$$

Another useful identity relating the norm of a projection with that of its complementary projection is

$$\|P\| = \|I - P\| \geq 1, \quad (1.6)$$

as long as P is neither null nor the identity. This identity was rediscovered and proved many times; see [3] and [28] for many of these proofs, references, and historical remarks. As we shall see, the characterization of $W(P)$ discussed in this paper provides a new proof of (1.6).

2. The general (infinite dimensional) case. We begin by presenting a direct proof of (1.4) in the spirit of some proofs of (1.6) in [28].

THEOREM 2.1. *Let P be a continuous projection on a Hilbert space \mathcal{H} , such neither $\mathcal{X} = \mathcal{R}(P)$ nor $\mathcal{Y} = \mathcal{N}(P)$ is the whole space. Then $W(P) = W(I - P)$.*

Proof. We will show that $W(P) \subseteq W(I - P)$, and the theorem will then follow by symmetry ($I - P$ is also an oblique projection). To that end, let $u \in \mathcal{H}$ with $\langle u, u \rangle = 1$ be arbitrary. We consider $\mathcal{H} = \mathcal{X} \oplus \mathcal{X}^\perp$. We can thus write $u = x + z$, $x \in \mathcal{X}$, $z \in \mathcal{X}^\perp$, and $\langle u, u \rangle = \langle x, x \rangle + \langle z, z \rangle = 1$, since $\langle x, z \rangle = 0$.

Therefore, since $Px = x$, we have that $Pu = x + Pz$. Thus

$$\langle Pu, u \rangle = \langle x + Pz, x + z \rangle = \langle x, x \rangle + \langle Pz, x \rangle \quad (2.1)$$

We want to construct $w \in \mathcal{H}$, with $\langle w, w \rangle = 1$, such that $\langle (I - P)w, w \rangle = \langle Pu, u \rangle$. We will do this for three different cases. First, if $x = 0$, i.e., if $u \in \mathcal{X}^\perp$, then since $Pu \in \mathcal{X}$, $\langle Pu, u \rangle = 0$, and we let $w \in \mathcal{X}$, with $\|w\| = 1$. Thus, $(I - P)w = 0$. Second, if $z = 0$, i.e., if $u \in \mathcal{X}$, we have $\langle Pu, u \rangle = \langle u, u \rangle = 1$. We then choose $w \in \mathcal{Y}$, with $\|w\| = 1$. Thus, $(I - P)w = w$, and $\langle (I - P)w, w \rangle = \langle w, w \rangle = 1$.

Let us finally assume that $x \neq 0$ and $z \neq 0$. Consider then

$$w = \tilde{x} + \tilde{z}, \quad \text{where} \quad \tilde{x} = -\frac{\|z\|}{\|x\|}x \in \mathcal{X}, \quad \tilde{z} = \frac{\|x\|}{\|z\|}z \in \mathcal{X}^\perp.$$

Then, it follows that $\langle w, w \rangle = \|z\|^2 + \|x\|^2 = 1$, and since $(I - P)x = 0$, then we have that $(I - P)w = \tilde{z} - P\tilde{z}$. Thus, using the fact that $\langle \tilde{z}, \tilde{x} \rangle = 0$, and $\langle P\tilde{z}, \tilde{z} \rangle = 0$, we have

$$\begin{aligned} \langle (I - P)w, w \rangle &= \langle \tilde{z} - P\tilde{z}, \tilde{x} + \tilde{z} \rangle = \\ &= \langle \tilde{z}, \tilde{z} \rangle - \langle P\tilde{z}, \tilde{x} \rangle = \frac{\|x\|^2}{\|z\|^2} \langle z, z \rangle + \langle Pz, x \rangle = \|x\|^2 + \langle Pz, x \rangle. \end{aligned}$$

Comparing with (2.1) the theorem follows. \square

Figure 2.1 shows in the two-dimensional case, some elements used in the proof of Theorem 2.1. We mention that an immediate corollary of this theorem is that

$$r(P) = r(I - P). \quad (2.2)$$

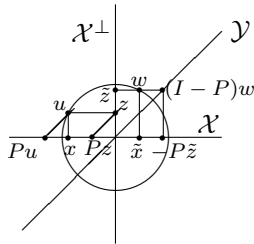


FIG. 2.1. Two-dimensional illustration of some elements of the proof of Theorem 2.1.

We analyze now the shape of $W(P)$. We begin with an auxiliary result which can be found, e.g., in [3, Example 5.8] [10, Eq. (1.11)], [28, §2].

LEMMA 2.2. *Let $P : \mathcal{H} \rightarrow \mathcal{H}$ be continuous and such that $P^2 = P$, with $\mathcal{R}(P) = \mathcal{X}$, and consider $\mathcal{H} = \mathcal{X} \oplus \mathcal{X}^\perp$. In term of this decomposition, we can then write*

$$P = \begin{bmatrix} I & B \\ 0 & 0 \end{bmatrix},$$

where $B = P|_{\mathcal{X}^\perp} : \mathcal{X}^\perp \rightarrow \mathcal{X}$ is the restriction of P to \mathcal{X}^\perp . Then $\|P\|^2 = 1 + \|B\|^2$.

Proof. It follows that $\|P\|^2 = \|P^*P\| = \|I + B^*B\| = 1 + \|B\|^2$. \square

We restate now [29, Theo. 2.1] for the particular case we have here that P is continuous and $\Lambda(P) = \{0, 1\}$, and offering a direct proof inspired in part by [6].

THEOREM 2.3. *Let P be continuous and such that $P^2 = P$. Then $W(P)$ is either the closed segment $[0, 1]$ or the (open or closed) elliptical disk with foci at 0 and 1, major axis $\|P\|$ and minor axis $(\|P\|^2 - 1)^{1/2}$.*

Proof. As in the proof of Theorem 2.1, we consider $\mathcal{H} = \mathcal{X} \oplus \mathcal{X}^\perp$. In terms of this decomposition, we can write $u \in \mathcal{H}$, $\|u\| = 1$ as $u = \alpha w + \beta v$, with $w \in \mathcal{X}$, and $v \in \mathcal{X}^\perp$ of unit norm, $\alpha, \beta \in \mathbb{C}$ such that $|\alpha|^2 + |\beta|^2 = 1$.

Thus, every element of $W(P)$ is of the form

$$\begin{aligned} \langle Pu, u \rangle &= \langle \alpha Pw + \beta Pv, \alpha w + \beta v \rangle = \langle \alpha w, \alpha w \rangle + \langle \beta Pv, \alpha w \rangle \\ &= |\alpha|^2 + \beta \bar{\alpha} \langle Pv, w \rangle = |\alpha|^2 + |\alpha| |\beta| e^{i\omega} \langle Pv, w \rangle, \end{aligned}$$

where ω depends only on the arguments of α and β . As ω varies we obtain the circle with center $t = |\alpha|^2$ ($0 \leq t \leq 1$) and radius $r = \sqrt{t(1-t)} |\langle Pv, w \rangle|$. That is, the circle in the (ξ, η) plane described by the equation is

$$F(t) = (\xi - t)^2 + \eta^2 - (t - t^2)m^2 = 0, \quad (2.3)$$

where $m = |\langle Pv, w \rangle|$. Conversely, using the same argument, every point in the circle (2.3) is in $W(P)$. To find the envelope of this family of circles, parametrized by $0 \leq t \leq 1$, one takes the derivative of $F(t)$ with respect to t , equates to zero and obtains $t = (m^2 + 2\xi)/(2 + 2m^2)$. Substituting this in the above equation of the circle yields (after some algebra) the ellipse

$$\frac{4}{1+m^2} \left(\xi - \frac{1}{2} \right)^2 + \frac{4}{m^2} \eta^2 = 1. \quad (2.4)$$

This ellipse has foci at 0 and 1, with major axis $a = \sqrt{1+m^2}$ and minor axis $b = m$, i.e., with eccentricity $1/\sqrt{1+m^2}$.

Observe now that for different values of $m = |\langle Pv, w \rangle|$, we have different ellipses (2.4) with the same foci, and thus their union is just the largest of them. Since by Lemma 2.2, $\sup\{|\langle Pv, w \rangle|, w \in \mathcal{X}, v \in \mathcal{X}^\perp, \|w\| = \|v\| = 1\} = \|P|_{\mathcal{X}^\perp}\| = \sqrt{\|P\|^2 - 1}$, the theorem follows. If this supremum is attained, i.e., if there is an element v of unit norm so that $\|Pv\| = \|P\|$, then this elliptical disk is closed. Otherwise, it is open. \square

Theorem 2.3 indicates in particular that $W(P)$ is symmetric with respect to the real axis. Therefore we have that

$$W(P^*) = W(P). \quad (2.5)$$

Recall that P^* is in fact the oblique projection onto \mathcal{Y}^\perp along \mathcal{X}^\perp ; see, e.g., [28, §5]. The identity (2.5) can be shown directly, and we do so in the Appendix.

As a corollary of Theorem 2.3 it also follows that

$$r(P) = \|P\|/2 + 1/2. \quad (2.6)$$

Indeed, since the foci are on the real axis, $r(P)$ is given by the right-most ξ value of the ellipse in (2.4), with $m = \sqrt{\|P\|^2 - 1}$, and this is obtained for $\eta = 0$.

Note also that Theorem 2.3 together with (1.6) imply Theorem 2.1, providing a different proof of the latter. Conversely, it follows from (2.2) and (2.6) that $\|P\| = \|I - P\|$, providing yet another proof of the identity (1.6).

We end this section with an observation on how the minimal canonical angle between the subspaces $\mathcal{X} = \mathcal{R}(P)$ and $\mathcal{Y} = \mathcal{N}(P)$ determines the shape of the elliptical disk $W(P)$. Recall that the minimal canonical angle $0 \leq \theta_{\min}(\mathcal{X}, \mathcal{Y}) \leq \pi/2$ between two nonzero subspaces \mathcal{X} and \mathcal{Y} of a Hilbert space can be defined as

$$\cos \theta_{\min}(\mathcal{X}, \mathcal{Y}) = \sup_{\substack{x \in \mathcal{X}, y \in \mathcal{Y} \\ \|x\|=1, \|y\|=1}} |\langle x, y \rangle|.$$

It turns out that

$$\frac{1}{\|P\|} = \sin \theta_{\min}(\mathcal{X}, \mathcal{Y})$$

(see, e.g., [11, §VI.5.4], [20]) and thus the eccentricity of the elliptical disk $W(P)$ is precisely the sine of the minimal canonical angle between \mathcal{X} and \mathcal{Y} . As a consequence we have the following observation.

REMARK 2.4. *The smaller the minimal canonical angle between two subspaces, the larger the field of values of the oblique projection onto one subspace along the other subspace. Moreover, the deviation of $\|P\|$ from one provides a measure of the degeneracy of the ellipse: the closer $\|P\|$ to one, the more the ellipse tends to be the segment $[0, 1]$. On the other hand, for large $\|P\|$, the ellipse axis values become increasingly close to each other, approximating a circle.*

3. The finite dimensional case. In this section we consider the finite dimensional case. In this way, we can find some relations of spectra, and as a consequence provide different proofs of Theorems 2.1 and 2.3 with the explicit form of the projection matrix. The projector onto \mathcal{X} along \mathcal{Y} can be written as

$$P = U(V^*U)^{-1}V^* \in \mathbb{C}^{n \times n}$$

for some full rank matrices U, V with the same number of columns. Thus, $\mathcal{X} = \mathcal{R}(U)$ and $\mathcal{Y} = \mathcal{N}(V^*)$. Without loss of generality we can assume that their columns of U and V are orthonormal.

Given a matrix $A \in \mathbb{C}^{n \times n}$, the finite dimension field of values (or numerical range) is defined as $W(A) = \{x^*Ax : \|x\| = 1, x \in \mathbb{C}^n\}$, where $\|x\|$ is the Euclidean vector norm.² We denote by $\Re\zeta$ and $\Im\zeta$ the real and imaginary parts of $\zeta \in \mathbb{C}$. We first recall the following property.

THEOREM 3.1. [14, Theo. 1.5.12] *For each matrix $A \in \mathbb{C}^{n \times n}$ and each $\varphi \in [0, 2\pi)$ let λ_φ be the largest eigenvalue of the Hermitian part of $e^{i\varphi}A$. Let $H_\varphi = e^{-i\varphi}\{\zeta : \Re\zeta \leq \lambda_\varphi\}$. Then*

$$W(A) = \bigcap_{0 \leq \varphi < 2\pi} H_\varphi.$$

THEOREM 3.2. *Let $P = U(V^*U)^{-1}V^* \in \mathbb{C}^{n \times n}$. For any $\varphi \in [0, 2\pi)$, it holds that $\Lambda(e^{i\varphi}P + (e^{i\varphi}P)^*) \setminus \{0\} = \Lambda(e^{i\varphi}(I - P) + (e^{i\varphi}(I - P))^*) \setminus \{\Re e^{i\varphi}\}$. As a consequence, $W(P) = W(I - P)$.*

Proof. Let $\rho = e^{i\varphi}$ with $0 \leq \varphi < 2\pi$, and let

$$[U, V] = Q \begin{bmatrix} I & R_1 \\ 0 & R_2 \end{bmatrix}, \quad R_1 = U^*V,$$

be the skinny QR decomposition of $[U, V]$. Then we have

$$\frac{1}{2}(\rho P + \bar{\rho}P^*) = Q \begin{bmatrix} \Re(\rho)I & \frac{1}{2}\bar{\rho}R_1^{-*}R_2^* \\ \frac{1}{2}\rho R_2 R_1^{-1} & \end{bmatrix} Q^*.$$

The nonzero (and real) eigenvalues of $\frac{1}{2}(\rho P + \bar{\rho}P^*)$ are thus the eigenvalues of the inner block matrix. Explicit computation shows that they are given by

$$\begin{aligned} \lambda_i^{(\pm)}(\rho) &= \frac{\Re\rho}{2} \pm \frac{\sqrt{(\Re\rho)^2 + 4|\rho|^2\sigma_i(\frac{1}{2}R_2R_1^{-1})^2}}{2} \\ &= \frac{\Re\rho}{2} \pm \frac{\sqrt{(\Re\rho)^2 + \sigma_i^2}}{2}, \quad \sigma_i = \sigma_i(R_2R_1^{-1}), \end{aligned}$$

where $\sigma_i(B)$ is a nonzero singular value of B . In particular,

$$\lambda_i^{(-)}(\rho) = \Re\rho - \lambda_i^{(+)}(\rho). \quad (3.1)$$

Since $\Lambda(\frac{1}{2}(\rho(I - P) + \bar{\rho}(I - P)^*)) = \Re\rho - \Lambda(\frac{1}{2}(\rho P + \bar{\rho}P^*))$, the first result follows from property (3.1).

The equalities above also show that $[\Re\rho - \lambda_{\max}^{(+)}(\rho), \lambda_{\max}^{(+)}(\rho)]$ is the spectral interval (including zero and $\Re\rho$) of both Hermitian matrices. In light of Theorem 3.1, the result $W(P) = W(I - P)$ immediately follows, since we also have that $\frac{1}{2}\lambda_{\max}(e^{i\varphi}A + (e^{i\varphi}A)^*)$, $0 \leq \varphi < 2\pi$, is the same for $A = P$ and $A = I - P$. \square

LEMMA 3.3. *It holds that $\sigma_i(R_2R_1^{-1})^2 = 1/\sigma_{k-i+1}(U^*V)^2 - 1$. In particular, $\sigma_1(R_2R_1^{-1})^2 = 1/\sigma_{\min}(U^*V)^2 - 1 = \|P\|^2 - 1$.*

Proof. Since $R_2^*R_2 = V^*(I - UU^*)(I - UU^*)V$, the eigenvalue problem $R_2^*R_2z = \sigma^2R_1^*R_1z$ is written as $V^*(I - UU^*)Vz = \sigma^2V^*UU^*Vz$ or, equivalently, as $(\sigma^2 + 1)^{-1}z = V^*UU^*Vz$, from which the result follows. \square

²This definition is standard in the matrix analysis literature (see, e.g., [14, Ch. 1]) and it is consistent with the general definition (1.3) if one considers the Euclidean inner product in \mathbb{C}^n as $\langle x, y \rangle = y^*x$.

We can now explicitly write down the equation of the ellipse enclosing $W(P)$ in the finite dimensional case. This result clearly also implies symmetry with respect to the real axis, foci at $\{0, 1\}$, together with the values of the major and minor axes $a = \|P\|$ and $b = \sqrt{\|P\|^2 - 1}$.

THEOREM 3.4. *Each point $x + iy \in \partial W(P)$ satisfies*

$$\frac{4}{\|P\|^2} \left(x - \frac{1}{2}\right)^2 + \frac{4}{\|P\|^2 - 1} y^2 = 1,$$

Proof. We recall that a point on $\partial W(A)$ is given by ([14, Th. 1.5.11])

$$p_\varphi = x_\varphi^* P x_\varphi$$

where x_φ is the unit eigenvector corresponding to $\lambda_{\max}(\rho)$, $\rho = \exp(i\varphi)$. We can thus write $x_\varphi = Qz_\varphi$ for some unit vector z_φ . From $\frac{1}{2}(\rho P + \bar{\rho}P^*)Qz_\varphi = Qz_\varphi\lambda_{\max}$ we obtain that the vector $z_\varphi = [z^{(1)}; z^{(2)}]$ satisfies $z^{(2)} = \frac{\rho}{2\lambda_{\max}} R_2 R_1^{-1} z^{(1)}$ and $z^{(1)}$ is such that $(R_2 R_1^{-1})^* (R_2 R_1^{-1}) z^{(1)} = \sigma_1^2 z^{(1)}$. Moreover, since $R_2^* R_2 = I - R_1^* R_1$, we also obtain $\|R_2 R_1^{-1} z^{(1)}\|^2 = \|R_1^{-1} z^{(1)}\|^2 - \|z^{(1)}\|^2$.

After some simple algebraic computations we see that

$$V^* Q = [R_1^*, -R_1^* R_1 R_2^{-1} + R_2^{-1}], \quad Q^* U = [I : 0],$$

which yields, after some little more algebra,

$$\begin{aligned} p_\varphi &= z_\varphi^* Q^* P Q z_\varphi = \left(1 - \frac{\rho}{2\lambda_{\max}}\right) \|z^{(1)}\|^2 + \frac{\rho}{2\lambda_{\max}} \|R_1^{-1} z^{(1)}\|^2 \\ &= \|z^{(1)}\|^2 + \frac{\rho}{2\lambda_{\max}} \|R_2 R_1^{-1}\|^2 \|z^{(1)}\|^2. \end{aligned}$$

Let $\sigma_1 = \sigma_1(R_2 R_1^{-1}) = \|R_2 R_1^{-1}\|$. Since $1 = \|z^{(1)}\|^2 + \|z^{(2)}\|^2 = \left(1 + \frac{\sigma_1^2}{4\lambda_{\max}^2}\right) \|z^{(1)}\|^2$, we have that $\|z^{(1)}\|^2 = 1 / \left(1 + \frac{\sigma_1^2}{4\lambda_{\max}^2}\right) =: t$. Therefore

$$x + iy := p_\varphi = t \left(1 + \frac{\sigma_1^2}{2\lambda_{\max}} \Re \rho\right) + it \frac{\sigma_1^2}{2\lambda_{\max}} \Im \rho.$$

We write $\Re \rho, \Im \rho$ in terms of x and y as follows

$$\Re \rho = \frac{1}{t} \frac{2\lambda_{\max}}{\sigma_1^2} (x - t), \quad \Im \rho = \frac{1}{t} \frac{2\lambda_{\max}}{\sigma_1^2} y.$$

Using the identity $\Re \rho^2 + \Im \rho^2 = 1$ we obtain

$$(x - t)^2 + y^2 - t(1 - t)\sigma_1^2 = 0,$$

which shows that each point $p_\varphi = x + iy$ belongs to a circumference of a family of circles depending on the parameter t . To obtain the envelope of the family, to which p_φ belongs, we differentiate the left-hand side with respect to t , and setting this derivative equal to zero we obtain $t = (\sigma_1^2 + 2x) / (2 + 2\sigma_1^2)$. Substituting this value of t in the family we obtain, after some algebraic computations,

$$\frac{4}{1 + \sigma_1^2} \left(x - \frac{1}{2}\right)^2 + \frac{4}{\sigma_1^2} y^2 = 1$$

which, taking into account Lemma 3.3, yields the sought after ellipse equation. \square

4. Appendix. Here we show directly that $W(P) = W(P^*)$. We recall that if P is the projection onto \mathcal{X} along \mathcal{Y} , then P^* is the oblique projection onto \mathcal{Y}^\perp along \mathcal{X}^\perp . We present our result for the finite dimensional case, i.e., for $\mathcal{H} = \mathbb{C}^n$, but the proof we present goes over easily to any separable Hilbert space \mathcal{H} .

THEOREM 4.1. $W(P) = W(P^*)$.

Proof. Let $V = [V_1, V_2]$ be an orthonormal basis of \mathcal{H} such that the columns of V_1 are a basis of \mathcal{X} and those of V_2 are a basis of \mathcal{X}^\perp . Similarly, let $W = [W_1, W_2]$ be an orthonormal basis of \mathcal{H} such that the columns of W_1 are a basis of \mathcal{Y}^\perp (of the same dimension as \mathcal{X}) and those of W_2 are a basis of \mathcal{Y} . It follows then that we can write $W = BV$, where $B = WV^*$ is an isometry (a unitary matrix), so that $B^*B = BB^* = I$.

We want to show that $W(P) \subseteq W(P^*)$ and thus the theorem will follow by symmetry. To that end let $u \in \mathcal{H}$, $\|u\| = 1$. We write $u = x + z$, $x \in \mathcal{X}$, $z \in \mathcal{X}^\perp$, so that $x = V_1\alpha$, $z = V_2\beta$, for some α and β such that $\|\alpha\|^2 + \|\beta\|^2 = 1$, where the norms here are the Euclidean norms in the appropriate spaces. Thus

$$\langle Pu, u \rangle = \langle x, x \rangle + \langle Pz, x \rangle.$$

Consider now $v = Bx + y$, with $y = W_2\gamma \in \mathcal{Y}$, for some appropriate γ to be determined later. Observe that $Bx = BV_1\alpha = W_1\alpha \in \mathcal{Y}^\perp$. Thus, $\|v\|^2 = \|\alpha\|^2 + \|\gamma\|^2$. We compute now

$$\langle P^*v, v \rangle = \langle Bx, Bx \rangle + \langle P^*y, Bx \rangle = \langle x, x \rangle + \langle \gamma, W_2^*PBx \rangle.$$

We thus need to choose γ so that $\|\gamma\| = \|\beta\|$ and $\langle \gamma, W_2^*PBx \rangle = \langle Pz, x \rangle$. If $\gamma = a + ib$ is a scalar, and thus so is W_2^*PBx , then a and b can easily be found such that $\langle a + ib, W_2^*PBx \rangle = \langle Pz, x \rangle$ and $a^2 + b^2 = \|\beta\|^2$. If γ is a vector, then we have two equations to satisfy and more than two free parameters in γ . Thus we have $\|v\| = 1$ and $\langle P^*v, v \rangle = \langle Pu, u \rangle$. \square

Acknowledgment. We thank Andreas Frommer, Roger Horn, and Chi-Kwong Li for their comments on an earlier draft, and for pointers to some papers not known to us. We also thank Shiferaw Berhanu for some questions which led us to write the Appendix.

REFERENCES

- [1] Christopher Beattie. Galerkin eigenvector approximations. *Mathematics of Computation*, 69:1409–1434, 2000.
- [2] Christopher Beattie, Mark Embree, and John Rossi. Convergence of restarted Krylov subspaces to invariant subspaces. *SIAM Journal on Matrix Analysis and Applications*, 25:1074–1109, 2004.
- [3] Albrecht Böttcher and Ilya M. Spitkovsky. A gentle guide to the basics of two projections theory. *Linear Algebra and its Applications*, 432:1412–1459, 2010.
- [4] Thierry Braconnier and Nicholas J. Higham. Computing the field of values and pseudospectra using the Lanczos method with continuation. *BIT, Numerical Computation*, 36:422–440, 1996.
- [5] Françoise Chatelin. *Spectral Approximation of Linear Operators*. Academic Press, New York, 1983.
- [6] William F. Donoghue. On the numerical range of a bounded operator. *Michigan Mathematical Journal*, 4:261–263, 1957.
- [7] Michael Eiermann. Fields of values and iterative methods. *Linear Algebra and its Application*, 180:167–197, 1993.
- [8] Michael Eiermann and Oliver Ernst. Geometric aspects in the theory of Krylov subspace methods. *Acta Numerica*, pages 251–312, 2001.

- [9] Robert D. Falgout, Panyot S. Vassilevski, and Ludmil T. Zikatanov. On two-grid convergence estimates. *Numerical Linear Algebra with Applications*, 12:471–494, 2005.
- [10] Israel Feldman, Naum Krupnik and Alexander Markus. On the norm of polynomials of two adjoint projections. *Integral Equations and Operator Theory*, 14:69–91, 1991.
- [11] Israel C. Gohberg and Mark G. Krein. *Introduction to the Theory of Nonselfadjoint Operators*, Nauka, Moscow, 1965. In Russian. American Mathematical Society, Translations of Mathematical Monographs, vol. 18, Providence, R.I., 1969.
- [12] Karl E. Gustafson and Duggirra K.M. Rao. *Numerical Range: The Field of Values of Linear Operators and Matrices*. Springer, New York, 1997.
- [13] Michiel E. Hochstenbach, David A. Singer, and Paul F. Zachelin. On the field of values and pseudospectra of the inverse of a large matrix. CASA report 07-06, Department of Mathematics and Computer Science, TU Eindhoven, The Netherlands, February 2007.
- [14] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.
- [15] Roger A. Horn and Vladimir V. Sergeichuk. Canonical forms for unitary congruence and $*$ -congruence. *Linear and Multilinear Algebra*, 57:777–815, 2009.
- [16] Tosio Kato. Estimation of iterated matrices, with application to the von Neumann condition. *Numerische Mathematik*, 2:22–29, 1960.
- [17] Axel Klawonn and Gerhard Starke. Block triangular preconditioners for nonsymmetric saddle point problems: Field-of-values analysis. *Numerische Mathematik*, 81:577–594, 1999.
- [18] Andrew Knyazev and Merico E. Argentati. Principal angles between subspaces in an A -based scalar product: algorithms and perturbation estimates. *SIAM Journal on Scientific Computing*, 23:2008–2040, 2002.
- [19] Chi-Kwong Li, Yiu-Tung Poon, and Nung-Sing Sze. Elliptical range theorems for generalized numerical ranges of quadratic operators. *Rocky Mountain Journal of Mathematics*, in press, 2010.
- [20] Vladyslav É. Ljance. Some properties of idempotent operators. *Teoreticheskaia i Prikladnaia Matematika*, 1:16–22, 1958/1959. In Russian.
- [21] Daniel Loghin, Martin van Gijzen, and Eline Jonkersb. Bounds on the eigenvalue range and on the field of values of non-Hermitian and indefinite finite element matrices. *Journal of Computational and Applied Mathematics*, 189:304–323, 2006.
- [22] Jan Mandel, Clark R. Dohrmann, and Radek Tezaur. An algebraic theory for primal and dual substructuring methods by constraints. *Applied Numerical Mathematics*, 54:167–193, 2005.
- [23] Leiba Rodman and Ilya M. Spitkovsky. On generalized numerical ranges of quadratic operators. *Operator Theory: Advances and Applications*, 179:241–256, 2008.
- [24] Valeria Simoncini and Daniel B. Szyld. On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods. *SIAM Review*, 47:247–272, 2005.
- [25] Valeria Simoncini and Daniel B. Szyld. New conditions for non-stagnation of minimal residual methods. *Numerische Mathematik*, 109:477–487, 2008.
- [26] Gerhard Starke. Field-of-values analysis of preconditioned iterative methods for nonsymmetric elliptic problems. *Numerische Mathematik*, 78:103–117, 1997.
- [27] G. W. (Pete) Stewart and Ji-guang Sun. *Matrix Perturbation Theory*. Academic Press, San Diego and London, 1990.
- [28] Daniel B. Szyld. The many proofs of an identity on the norm of oblique projections. *Numerical Algorithms*, 42:309–323, 2006.
- [29] Shu-Hsien Tso and Pei Yuan Wu. Matricial ranges of quadratic operators. *Rocky Mountain Journal of Mathematics*, 29:1139–1152, 1999.
- [30] Jinchao Xu and Ludmil Zikatanov. Some observations on Babuška and Brezzi theories. *Numerische Mathematik*, 94:195–202, 2003.