Numer Algor (2006) 42: 309–323
DOI 10.1007/s11075-006-9046-2

The many proofs of an identity on the norm of oblique projections

Daniel B. Szyld

Received: 22 February 2006 / Accepted: 8 August 2006 /
Published online: 28 October 2006

Abstract Given an oblique projector P on a Hilbert space, i.e., an operator satisfying $P^2 = P$, which is neither null nor the identity, it holds that $\|P\| = \|I - P\|$. This useful equality, while not widely-known, has been proven repeatedly in the literature. Many published proofs are reviewed, and simpler ones are presented.

Keywords idempotent operators · oblique projections · angle between subspaces · gap between subspaces

Mathematics Subject Classifications (2000) 15A24 · 15A60 · 46C99 · 46E99 · 46N40

1 Introduction

This paper concerns a nice and useful identity on an inner-product space between the norm of a projection and that of its complementary projection. Namely that

$$\|P\| = \|I - P\|, \tag{1.1}$$

where P is a projection which is neither null nor the identity, and the operator norm is the standard norm induced by the vector norm which is defined by the inner product. This identity is apparently not widely known in the numerical analysis community, nor in the linear algebra community, and almost all books on these subjects fail to mention it; one exception is [34, Exercise 5.9.9]. One purpose of this paper is to call attention to the identity (1.1) since we believe it can be very useful in the analysis of certain numerical algorithms. Indeed, in the last few years the equality (1.1) was applied in a few numerical analysis papers [4, 5, 15, 33, 37, 43]; see also [16].

Communicated by Michelle Benzi.

D. B. Szyld (✉)
Department of Mathematics, Temple University, Philadelphia, PA 19122, USA
e-mail: szyld@temple.edu

The identity (1.1) is apparently well-known among functional analysts, but it is not generally found in functional analysis texts. One can find it in [18], where Ljance [32] is given as its source; see also, e.g., [38]. It is also mentioned as an exercise in [25], with a reference in a footnote to Kato [24]. Generalizations to C^*-algebras are found in [27, 28]. There is an earlier proof of the identity (1.1) by Del Pasqua [10]. In the latter paper it is also pointed out that the identity (1.1) is not valid in general Banach spaces; the fact that the norm is associated with an inner product is essential. Indeed, in [20] it is shown that if the identity (1.1) holds for all projections P, this implies that the Banach space is an inner-product space. A simple example where the identity (1.1) does not hold when the norm is not associated with an inner product is given at the end of the next section.

Thus, we have three authors proving the same result within a five-year period, independently of each other: Del Pasqua [10], Ljance [32], and Kato [24], although by 1976 Kato was aware of Del Pasqua's paper [25, p.568]. None of the papers [10, 32], or [24], have been much cited, and in the few cases found, the citation was mostly for results other than the identity (1.1).

Other authors have proved and reproved the identity (1.1) in the half century since the first proof appeared, and in more than one occasion these new proofs were done without knowledge of the earlier results; see, e.g., [35]. One purpose of this paper to review many of the proofs found in the literature, and present new ones.

In the next section we state our notation and the precise theorem which is the object of this paper. We also present two simple proofs, which we believe are not available elsewhere. We present next, in Sections 3 and 4, two other proofs, an early one by Kato [24], and a more recent one. These four proofs have in common that only tools from the structure of the Hilbert space are used, such as the vector and operator norms, and the inner product.

All other proofs use more geometric concepts, such as angles, or the gap between subspaces. These concepts are reviewed in Section 5, and we show how these concepts relate to one another. For completeness we reproduce some of the proofs of these relations.

In Section 6 we present the proofs of the identity (1.1) by Del Pasqua [10] and Ljance [32], and complete the picture collecting all the relations among the geometric concepts mentioned in this paper in a single identity. We complete the review of some of the published proofs of the identity (1.1) in Section 7.

2 Preliminaries, statement of the main theorem, and simple proofs

Consider a Hilbert space $\mathcal{H}$ with inner product $\langle x, y\rangle$, and its associated norm $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$ (for example $\mathbb{R}^n$ and the Euclidean inner product). All the results in this paper apply to complex spaces, and this is assumed throughout; of course the results equally apply to real spaces. The vector norm induces an operator norm in the usual manner, i.e., for $A : \mathcal{H} \to \mathcal{H}$, one has

$$\|A\| = \sup_{\|u\|=1} \|Au\| = \sup_{\|u\|\le 1} \|Au\| = \sup_{\|u\|\le 1,\ \|v\|\le 1} |\langle v, Au\rangle|. \tag{2.1}$$

The adjoint operator A^* is such that for every $u, v \in \mathcal{H}$, $\langle Au, v\rangle = \langle u, A^*v\rangle$, and it holds that $\|A^*\| = \|A\|$. It follows from Eq. (2.1), that for any operator A on $\mathcal{H}$, $\|A\|^2 = \|A^*A\|$. Indeed, we have that

$$\|A\|^2 = \sup_{\|u\|\le 1} \|Au\|^2 = \sup_{\|u\|\le 1} \langle Au, Au\rangle = \sup_{\|u\|\le 1} \langle u, A^*Au\rangle$$
$$\le \sup_{\|u\|\le 1,\ \|v\|\le 1} |\langle v, A^*Au\rangle| = \|A^*A\| \le \|A\|^2,$$

and equality holds throughout. An operator A on $\mathcal{H}$ is continuous if and only if it is bounded, i.e., $\|A\| < \infty$. Given any subspace $\mathcal{X}$ of $\mathcal{H}$ we define its orthogonal complement by

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

A projection P is an idempotent operator, i.e., such that

$$P^2 = P. \tag{2.2}$$

The operator P is a projection along (or parallel to) its null space $\mathcal{Y} = \mathcal{N}(P)$ onto its range $\mathcal{X} = \mathcal{R}(P)$. If these subspaces are orthogonal, the projection is called orthogonal, and this is the case if and only if P is Hermitian. Otherwise, it is called an oblique projection. When we want to emphasize that a projection onto $\mathcal{X}$ is orthogonal, we denote it by $\Pi_{\mathcal{X}}$. In both the oblique and the orthogonal case, $I - P$ is also idempotent, and it is a projection along $\mathcal{X} = \mathcal{N}(I - P) = \mathcal{R}(P)$ onto $\mathcal{Y} = \mathcal{R}(I - P) = \mathcal{N}(P)$. Since $P = P^2$ one directly has that $(I - P)P = P(I - P) = O$, and that $\|P\| \ge 1$, with equality only in the case of an orthogonal projection. Thus, in the case of orthogonal projections, the identity (1.1) is trivial since $\|\Pi_{\mathcal{X}}\| = \|I - \Pi_{\mathcal{X}}\| = 1$.

From Eq. (2.2) it follows that if $x \in \mathcal{X} = \mathcal{R}(P)$, then $x = Px$. Indeed, let $x = Py$, then $x = Py = P^2y = Px$. 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 $x \in \mathcal{H}$ can be written as $x = Px + (I - P)x$.

Some early results on oblique projections, which foreshadowed the derivation of the identity (1.1) can be found in [1]; see also [3, 9, 12, 19, 35, 38, 42], for some additional properties of oblique projections not discussed here.

We state now the result which this paper highlights.

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

We begin by presenting two simple proofs, which we believe are new. In fact, they can be used in undergraduate courses. The first was reported to us in 2004 after we had a similar proof restricted to the finite dimensional case. The second was communicated to us after a first version of this paper was disseminated.

Proof of Theorem 2.1 (Krainer, 2004, personal communication) Let $u \in \mathcal{H}$, $\|u\| = 1$ be arbitrary. Let $x = Pu \in \mathcal{X}$ and $y = (I - P)u \in \mathcal{Y}$. Then

$$\|u\|^2 = \|x\|^2 + \|y\|^2 + 2Re\langle x, y\rangle = 1. \tag{2.3}$$

We will show that $\|Pu\| \leq \|I - P\|$ which would imply $\|P\| \leq \|I - P\|$. The theorem will then follow by symmetry. If either $x = 0$ or $y = 0$ we are done, since in the first case we have $Pu = 0$, and in the second $\|Pu\| = 1$. We assume then that $x \neq 0$ and $y \neq 0$. Let us define another element in $\mathcal{H}$. Let $w = \tilde{x} + \tilde{y}$, with

$$\tilde{x} = \frac{\|y\|}{\|x\|} x \in \mathcal{X}, \quad \tilde{y} = \frac{\|x\|}{\|y\|} y \in \mathcal{Y}.$$

Then we have, using Eq. (2.3) that $\|w\|^2 = \|y\|^2 + \|x\|^2 + 2Re\langle x, y\rangle = 1$ and

$$\|Pu\| = \|x\| = \|\tilde{y}\| = \|(I - P)w\| \leq \|I - P\|.$$

□

We illustrate in figure 1 the situation described in this proof in the two-dimensional case. The complete symmetry between P and $I - P$ can be appreciated. The following proof, can also be easily interpreted in the two-dimensional case, by choosing the appropriate basis, and expressing the operators P and $I - P$ in this basis.

Proof of Theorem 2.1 (Corach, 2006, personal communication) We consider $\mathcal{H} = \mathcal{X} \oplus \mathcal{X}^{\perp}$, and we write every linear operator on $\mathcal{H}$ as

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where $a : \mathcal{X} \to \mathcal{X}$, $b : \mathcal{X}^{\perp} \to \mathcal{X}$, $c : \mathcal{X} \to \mathcal{X}^{\perp}$, and $d : \mathcal{X}^{\perp} \to \mathcal{X}^{\perp}$. In this form we write

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & b \\ 0 & 0 \end{bmatrix}.$$

In the latter case, if $b = 0$ we have $\Pi_{\mathcal{X}}$. We can now simply compute

$$I - P = \begin{bmatrix} 0 & -b \\ 0 & -1 \end{bmatrix}.$$

It follows that $\|P\|^2 = \|P^*P\| = 1 + \|b^*b\|$, and $\|I - P\|^2 = \|(I - P)^*(I - P)\| = 1 + \|b^*b\|$. The theorem follows. □

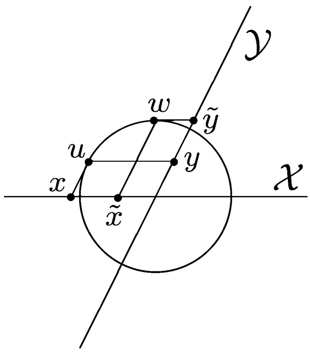

Figure 1 Two-dimensional illustration of a simple proof.

For an example of a Banach space, where Theorem 2.1 does not hold, consider the space $\mathbb{R}^2$ with the max norm. Let $\mathcal{X} = \{(x, y) \mid y = 0\}$ and $\mathcal{Y} = \{(x, y) \mid x = y\}$. Thus

$$P = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}, \qquad I - P = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix},$$

and we have that $\|P\|_\infty = 2$, while $\|I - P\|_\infty = 1$.

3 An early proof

We begin our presentation of the different proofs encountered in the literature with that of Kato [24], which is similar in flavor to our first simple proof in the previous section, and can also be easily interpreted in the case of two dimensions, as illustrated in figure 2.

Proof of Theorem 2.1 (Kato [24]) If $\|P\| = 1$, then $\|I - P\| = 1$, and there is nothing to prove. We therefore consider the case where $\|P\| > 1$. For any α such that

$$\|P\| > \alpha > 1, \tag{3.1}$$

there exists $u \neq 0$ such that

$$\|Pu\| \geq \alpha \|u\| > 0. \tag{3.2}$$

Consider now

$$v = u - \frac{\langle u, Pu \rangle}{\|Pu\|^2} Pu, \tag{3.3}$$

i.e., $v = \Pi_{\mathcal{X}^\perp} u$, so that $\langle v, Pu \rangle = 0$, and therefore

$$\|v\|^2 = \|u\|^2 - \frac{|\langle u, Pu \rangle|^2}{\|Pu\|^2}. \tag{3.4}$$

A direct calculation, using the fact that $P^2 = P$ shows that $v - Pv = u - Pu$, and therefore, from Eq. (3.2), and $v \neq 0$ we have that

$$\frac{\|(I - P)v\|^2}{\|v\|^2} = \frac{\|(I - P)u\|^2 \|Pu\|^2}{\|u\|^2 \|Pu\|^2 - |\langle u, Pu \rangle|^2} \geq \frac{\|Pu\|^2}{\|u\|^2} \geq \alpha^2, \tag{3.5}$$

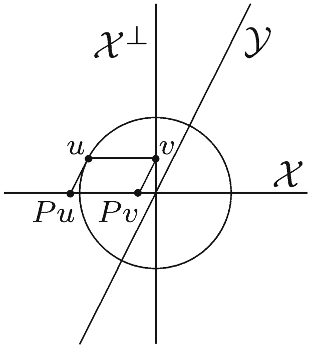

Figure 2 Two-dimensional illustration of Kato's proof.

where the equality follows from Eq. (3.4), and the middle inequality from the following identity

$$\|u\|^2\|(I-P)u\|^2 - \left(\|u\|^2\|Pu\|^2 - |\langle u, Pu\rangle|^2\right)$$
$$= \left(\|u\|^2 - |\langle u, Pu\rangle|\right)^2 + 2\|u\|^2\left(|\langle u, Pu\rangle| - Re\langle u, Pu\rangle\right) \geq 0.$$

Thus, Eq. (3.5) indicates that $\|I-P\| \geq \alpha$. Since α was arbitrarily chosen satisfying Eq. (3.1), we have that $\|I-P\| \geq \|P\|$. By symmetry, the theorem follows. □

4 A recent proof

We continue our survey of the different proofs with the most recently published, from 2003, where in the first part, the space is assumed to be two-dimensional. Observe the similitude with the second simple proof in Section 2.

Proof of Theorem 2.1 (Xu and Zikatanov [43]) We first prove the theorem assuming that $\dim \mathcal{H} = 2$. In that case, both P and $I-P$ have rank 1, namely $Pv = \langle b, v\rangle a$ and $(I-P)v = \langle d, v\rangle c$ for some fixed nonzero $a, b, c, d \in \mathcal{H}$ such that $\langle a, b\rangle = \langle c, d\rangle = 1$. (Note that $b \in \mathcal{Y}^\perp$ and $d \in \mathcal{X}^\perp$). Thus for any $v \in \mathcal{H}$,

$$v = Pv + (I-P)v = \langle b, v\rangle a + \langle d, v\rangle c.$$

By replacing v with d and b in the last expression, we have that

$$\|a\|^2\|b\|^2 = \|c\|^2\|d\|^2 = 1 - \langle a, c\rangle\langle b, d\rangle.$$

The theorem for the two-dimensional case follows from the following equalities:

$$\|P^*P\| = \|a\|^2\|b\|^2, \quad \|(I-P)^*(I-P)\| = \|c\|^2\|d\|^2.$$

For the general case, consider any $v \in \mathcal{H}$, $\|v\| = 1$, and the subspace

$$\mathcal{V} = \mathrm{span}\{v, Pv\}.$$

Observe that $\mathcal{V}$ is invariant under P and $I-P$. If $\dim \mathcal{V} = 1$, then one has either $Pv = 0$ or $(I-P)v = 0$. In the first case $\|(I-P)v\| = \|v\| = 1 \leq \|P\|$. In the second, $\|(I-P)v\| = 0 \leq \|P\|$. If $\dim \mathcal{V} = 2$, by the previous part $\|(I-P)v\|_{\mathcal{V}} \leq \|I-P\|_{\mathcal{V}} = \|P\|_{\mathcal{V}}$, where we have used the norm restricted to the subspace $\mathcal{V}$. Therefore,

$$\|(I-P)v\| = \|(I-P)v\|_{\mathcal{V}} \leq \|P\|_{\mathcal{V}} \leq \|P\|,$$

and since v was arbitrary of unit norm, we conclude that $\|I-P\| \leq \|P\|$. By symmetry, the theorem is proved. □

5 The gap, canonical angles, and other geometric ingredients

We begin this section with a characterization of the adjoint of a continuous oblique projection P. Let $\mathcal{X} = \mathcal{R}(P)$, $\mathcal{Y} = \mathcal{N}(P)$. Let Q be the oblique projection such that $\mathcal{R}(Q) = \mathcal{Y}^\perp$ and $\mathcal{N}(Q) = \mathcal{X}^\perp$, cf. figure 3. It holds that

$$P = Q^*$$

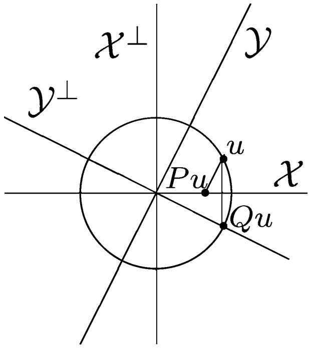

Figure 3 Four subspaces: $\mathcal{X} = \mathcal{R}(P)$, $\mathcal{Y} = \mathcal{N}(P)$, $\mathcal{R}(P)^\perp = \mathcal{N}(Q)$, $\mathcal{N}(P)^\perp = \mathcal{R}(Q)$.

[29], [18, Section VI.5.3]. This can be seen by considering arbitrary elements $u, v \in \mathcal{H}$, and writing them as $u = u_\mathcal{X} + u_\mathcal{Y}$, $u_\mathcal{X} \in \mathcal{X}$, $u_\mathcal{Y} \in \mathcal{Y}$, and $v = v_\mathcal{W} + v_\mathcal{Z}$, $v_\mathcal{W} \in \mathcal{W} = \mathcal{Y}^\perp$, $v_\mathcal{Z} \in \mathcal{Z} = \mathcal{X}^\perp$ and thus

$$\langle Pu, v\rangle = \langle u_\mathcal{X}, v\rangle = \langle u_\mathcal{X}, v_\mathcal{W}\rangle = \langle u, v_\mathcal{W}\rangle = \langle u, Qv\rangle.$$

The *minimal gap* between two arbitrary closed subspaces $\mathcal{X}$ and $\mathcal{Y}$ of a Hilbert space can be defined as

$$g(\mathcal{X}, \mathcal{Y}) = \inf_{x\in\mathcal{X}} \frac{\operatorname{dist}(x, \mathcal{Y})}{\|x\|}, \quad \text{where } \operatorname{dist}(x, \mathcal{Y}) = \inf_{y\in\mathcal{Y}} \|x - y\| = \|x - \Pi_\mathcal{Y} x\|; \quad (5.1)$$

see, e.g., [11], and cf. [25, p. 219]. The equivalent definition

$$g(\mathcal{X}, \mathcal{Y}) = \inf_{\substack{x\in\mathcal{X} \\ \|x\|=1}} \operatorname{dist}(x, \mathcal{Y}) \quad (5.2)$$

is given in [10] where it is called *index of disjointness* of the subspaces.

The minimal gap (5.2) should not be confused with the usual concept of *gap* between two arbitrary subspaces $\mathcal{X}$ and $\mathcal{W}$ of a Hilbert space (which perhaps should be called the *maximal gap*, and sometimes it is also referred as *opening* or *aperture*), defined as

$$G(\mathcal{X}, \mathcal{W}) = \sup_{\substack{x\in\mathcal{X} \\ \|x\|=1}} \operatorname{dist}(x, \mathcal{W}) = \sup_{\substack{w\in\mathcal{W} \\ \|w\|=1}} \operatorname{dist}(w, \mathcal{X}). \quad (5.3)$$

This implies that $G(\mathcal{X}, \mathcal{W}) = \|(I - \Pi_\mathcal{W})\Pi_\mathcal{X}\| = \|(I - \Pi_\mathcal{X})\Pi_\mathcal{W}\| = \|\Pi_\mathcal{X} - \Pi_\mathcal{W}\|$ and it is equal to the sine of the maximal canonical angle between the subspaces, i.e.,

$$G(\mathcal{X}, \mathcal{W}) = \left(1 - \cos^2 \theta_{\max}(\mathcal{X}, \mathcal{W})\right)^{1/2}, \quad (5.4)$$

where

$$\cos \theta_{\max}(\mathcal{X}, \mathcal{W}) = \inf_{\substack{x\in\mathcal{X},\ w\in\mathcal{W} \\ \|x\|=1,\ \|w\|=1}} |\langle x, w\rangle|;$$

see, e.g., [2, Section 34], [8, Section 2.5.1], [18, Section VI.5.3], [25, Section IV.2.1], [32], [39, Section II.4.1]. The concept of the gap (5.3) was apparently first introduced in [29]. In the finite dimensional case, one usually specifies that $\mathcal{X}$ and $\mathcal{W}$ have the same dimension.

The minimal canonical angle $0 \le \theta_{\min}(\mathcal{X}, \mathcal{Y}) \le \pi/2$ between two nonzero subspaces $\mathcal{X}$ and $\mathcal{Y}$ of a Hilbert space can be defined as (see, e.g., [11, 13], [18, Section VI.5.4], [32])

$$\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) = \sup_{\substack{x\in\mathcal{X},\ y\in\mathcal{Y}\\ \|x\|=1,\ \|y\|=1}} |\langle x, y\rangle|. \tag{5.5}$$

It turns out that we can express this angle as the norm of the product of the two orthogonal projections onto the subspaces $\mathcal{X}$ and $\mathcal{Y}$. The following result is given in [22] for finite dimensions. The proof we present for the general case is almost identical to that of [22].

Lemma 5.1 Let $\mathcal{X}$ and $\mathcal{Y}$ be nonzero closed subspaces of a Hilbert space, then

$$\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) = \|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\| = \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}\|. \tag{5.6}$$

Proof If $\|u\| \le 1$, then $\Pi_{\mathcal{X}}u \in \mathcal{X}$ and $\|\Pi_{\mathcal{X}}u\| \le \|\Pi_{\mathcal{X}}\|\|u\| = \|u\| \le 1$.

$$\begin{aligned}\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) &= \sup_{\substack{x\in\mathcal{X},\ y\in\mathcal{Y}\\ \|x\|=1,\ \|y\|=1}} |\langle x, y\rangle| = \sup_{\substack{x\in\mathcal{X},\ y\in\mathcal{Y}\\ \|x\|\le1,\ \|y\|\le1}} |\langle x, y\rangle|\\ &= \sup_{\|u\|\le1,\ \|v\|\le1} |\langle \Pi_{\mathcal{X}}u, \Pi_{\mathcal{Y}}v\rangle| = \sup_{\|u\|\le1,\ \|v\|\le1} |\langle u, \Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}v\rangle|\\ &= \|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\|,\end{aligned}$$

where in the last two equalities we have used the fact that an orthogonal projection is Hermitian, and the identity (2.1). To complete the proof we have $\|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\| = \|(\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}})^*\| = \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}\|$. □

Just as we have the relation (5.4), it turns out the sine of the minimal angle is equal to the minimal gap, as shown next.

Lemma 5.2 [42] Let $\mathcal{X}$ and $\mathcal{Y}$ be nonzero closed subspaces of a Hilbert space, then

$$g(\mathcal{X}, \mathcal{Y}) = (1 - \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}\|^2)^{1/2} = \sin\theta_{\min}(\mathcal{X}, \mathcal{Y}). \tag{5.7}$$

Proof By definition

$$g(\mathcal{X}, \mathcal{Y}) = \inf_{\substack{x\in\mathcal{X}\\ \|x\|=1}} \operatorname{dist}(x, \mathcal{Y}) = \inf_{\substack{x\in\mathcal{X}\\ \|x\|=1}} \|(I - \Pi_{\mathcal{Y}})x\| = \inf_{\substack{x\in\mathcal{X}\\ \|x\|=1}} \|(I - \Pi_{\mathcal{Y}})\Pi_{\mathcal{X}}x\|.$$

For any $x \in \mathcal{X}$ with $\|x\| = 1$, we have that $\|(I - \Pi_{\mathcal{Y}})\Pi_{\mathcal{X}}x\|^2 + \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}x\|^2 = 1$. Therefore

$$g(\mathcal{X}, \mathcal{Y})^2 = 1 - \sup_{\substack{x \in \mathcal{X} \\ \|x\|=1}} \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}x\|^2 = 1 - \|\Pi_{\mathcal{Y}}\Pi_{\mathcal{X}}\|^2.$$

□

Ljance [32] relates the *minimal* canonical angle between complementary subspaces $\mathcal{X}$ and $\mathcal{Y}$ and the *maximal* canonical angle between $\mathcal{X}$ and the orthogonal complement of $\mathcal{Y}$. Namely, he shows that

$$\theta_{\min}(\mathcal{X}, \mathcal{Y}) + \theta_{\max}(\mathcal{X}, \mathcal{Y}^{\perp}) = \frac{\pi}{2};$$

cf. figure 3, and see also [18, Section VI.5.3], [29]. As a consequence, we can relate the minimal gap between $\mathcal{X}$ and $\mathcal{Y}$ to the (maximal) gap between $\mathcal{X}$ and $\mathcal{Y}^{\perp}$.

Lemma 5.3 [32] Let $\mathcal{X}$ and $\mathcal{Y}$ be nonzero complementary closed subspaces of a Hilbert space, then

$$\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) = \sin\theta_{\max}(\mathcal{X}, \mathcal{Y}^{\perp}), \tag{5.8}$$

i.e.,

$$\sqrt{1 - g(\mathcal{X}, \mathcal{Y})^2} = G(\mathcal{X}, \mathcal{Y}^{\perp}) = \|\Pi_{\mathcal{X}} - \Pi_{\mathcal{Y}^{\perp}}\|.$$

Proof Let $\mathcal{W} = \mathcal{Y}^{\perp}$. Since

$$\text{dist}(x, \mathcal{W}) = \sqrt{\|x\|^2 - \|\Pi_{\mathcal{W}}x\|^2},$$

from Eq. (5.3) it follows that for any $\varepsilon > 0$, there is an $x = x_\varepsilon \in \mathcal{X}$, $\|x\| = 1$, such that

$$\sqrt{1 - \|\Pi_{\mathcal{W}}x\|^2} \;>\; \sin\theta_{\max}(\mathcal{X}, \mathcal{W}) - \varepsilon.$$

Consider now $y = y_\varepsilon = (x - \Pi_{\mathcal{W}}x)/\|x - \Pi_{\mathcal{W}}x\|$. We thus have $y \in \mathcal{Y}$, $\|y\| = 1$, and

$$\langle x, y\rangle = \sqrt{1 - \|\Pi_{\mathcal{W}}x\|^2} \;>\; \sin\theta_{\max}(\mathcal{X}, \mathcal{W}) - \varepsilon,$$

which by Eq. (5.5) implies that $\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) > \sin\theta_{\max}(\mathcal{X}, \mathcal{W}) - \varepsilon$. Since ε was arbitrary, we have $\cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) \geq \sin\theta_{\max}(\mathcal{X}, \mathcal{W})$.

To prove the inequality on the other direction, from Eq. (5.5) it follows that for any $\varepsilon > 0$, there are $x = x_\varepsilon \in \mathcal{X}$, $\|x\| = 1$, and $y = y_\varepsilon \in \mathcal{Y}$, $\|y\| = 1$, such that $\langle x, y\rangle > \cos\theta_{\min}(\mathcal{X}, \mathcal{Y}) - \varepsilon$, where if necessary, one vector x, or y, have been multiplied by $e^{i\alpha}$ for some real α, so that $\langle x, y\rangle$ is real and positive. Consider now $z = z_\varepsilon = (x - \Pi_{\mathcal{W}}x)/\lambda \in \mathcal{Y}$, where $\lambda = \|x - \Pi_{\mathcal{W}}x\| = (1 - \|\Pi_{\mathcal{W}}x\|^2)^{1/2}$, so that $\|z\| = 1$. With this construction, $x = \lambda z + \Pi_{\mathcal{W}}x$ and since $\Pi_{\mathcal{W}}x \perp \mathcal{Y}$, $\langle x, z\rangle = \lambda$ and

$$\langle x, y\rangle = \lambda\langle z, y\rangle = \langle x, z\rangle\langle z, y\rangle \leq \langle x, z\rangle\, \|z\|\, \|y\| = \langle x, z\rangle.$$

Therefore

$$\sin\theta_{\max}(\mathcal{X},\mathcal{W}) \geq \sqrt{1-\|\Pi_{\mathcal{W}}x\|^2} = \lambda = \langle x, z\rangle \geq \langle x, y\rangle > \cos\theta_{\min}(\mathcal{X},\mathcal{Y}) - \varepsilon.$$

Since ε was arbitrary, the lemma follows. □

6 The other early proofs

In this section we present the two other proofs from the 1950s in the order in which they were published. They both make use of the geometric concepts developed in Section 5, and in particular on the fact that

$$g(\mathcal{X},\mathcal{Y}) = g(\mathcal{Y},\mathcal{X}); \tag{6.1}$$

see Eqs. (5.7) and (5.6).

Proof of Theorem 2.1 (Del Pasqua [10]) We will show that

$$\|P\| = \frac{1}{g(\mathcal{X},\mathcal{Y})} \tag{6.2}$$

and the theorem will follow from Eq. (6.1). To that end, consider $x \in \mathcal{X}$, $\|x\| = 1$, and $y \in \mathcal{Y}$. Since $Py = 0$, we then have

$$1 = \|x\| = \|P(x-y)\| \leq \|P\|\,\|x-y\|.$$

Thus

$$\|x-y\| \geq \frac{1}{\|P\|},$$

and therefore, since $y \in \mathcal{Y}$ is arbitrary, by Eq. (5.2),

$$g(\mathcal{X},\mathcal{Y}) \geq \frac{1}{\|P\|}.$$

For the inequality in the other direction consider an arbitrary element $z \in \mathcal{H}$ written as $z = x - y$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$. We then have, by Eq. (5.1)

$$\|Pz\| = \|P(x-y)\| = \|x\| \leq \frac{1}{g(\mathcal{X},\mathcal{Y})}\|x-y\| = \frac{1}{g(\mathcal{X},\mathcal{Y})}\|z\|.$$

Since $z \in \mathcal{H}$ was arbitrary,

$$\|P\| \leq \frac{1}{g(\mathcal{X},\mathcal{Y})}.$$

□

Proof of Theorem 2.1 (Ljance [32]) We will show that

$$\|P\| = \frac{1}{\cos\theta_{\max}(\mathcal{X},\mathcal{Y}^{\perp})},$$

and the result will follow using Eq. (5.8), and the symmetry in Eq. (5.5); cf. Eq. (6.1). Let $\mathcal{W} = \mathcal{Y}^{\perp}$. First observe that

$$P = P\Pi_{\mathcal{W}}. \tag{6.3}$$

This is illustrated for two dimensions in figure 4.

To see that the identity (6.3) holds, consider any $u, v \in \mathcal{H}$, and we have that

$$\langle P\Pi_{\mathcal{W}}u, v\rangle = \langle u, \Pi_{\mathcal{W}}P^*v\rangle = \langle u, P^*v\rangle = \langle Pu, v\rangle,$$

where we have used that P^* is a projection onto $\mathcal{W} = \mathcal{Y}^{\perp}$. Therefore, since $\|\Pi_{\mathcal{W}}w\| \leq \|w\|$, we conclude that

$$\|P\| = \sup_{w\in\mathcal{W},\, \|w\|=1} \|Pw\|.$$

Next, we want to show that for any $w \in \mathcal{W}$, $\langle w, w - Pw\rangle = 0$, i.e., that w is the image in $\mathcal{W}$ of the orthogonal projection applied to Pw. In other words, we want to show that $\Pi_{\mathcal{W}}Pw = w$; see figure 4 for a two-dimensional representation of this. To see this operator identity in general, let $z \in \mathcal{H}$ be arbitrary, then

$$\langle \Pi_{\mathcal{W}}Pw, z\rangle = \langle w, P^*\Pi_{\mathcal{W}}z\rangle = \langle w, \Pi_{\mathcal{W}}z\rangle = \langle \Pi_{\mathcal{W}}w, z\rangle = \langle w, z\rangle.$$

Therefore

$$\|Pw\|^2 = \|w\|^2 + \|w - Pw\|^2 = \|w\|^2 + \|Pw\|^2 \left\| \frac{Pw}{\|Pw\|} - \frac{w}{\|Pw\|} \right\|^2.$$

Using again that any $w \in \mathcal{W}$ is the closest element in $\mathcal{W}$ to Pw, we have that

$$\left\| \frac{Pw}{\|Pw\|} - \frac{w}{\|Pw\|} \right\| = \operatorname{dist}\left(\frac{Pw}{\|Pw\|}, \mathcal{W}\right).$$

Therefore, for $w \in \mathcal{W}$ such that $\|w\| = 1$

$$\|Pw\| = \frac{1}{\sqrt{1 - \operatorname{dist}\left(\frac{Pw}{\|Pw\|}, \mathcal{W}\right)^2}},$$

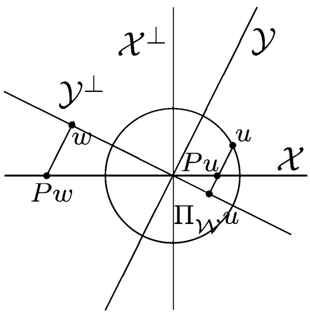

Figure 4 Two-dimensional illustration that for $\mathcal{W} = \mathcal{Y}^{\perp}$, it holds that $P = P\Pi_{\mathcal{W}}$, and for $w \in \mathcal{W}$, $\Pi_{\mathcal{W}}Pw = w$.

and consequently

$$\sup_{w\in\mathcal{W},\,\|w\|=1}\|Pw\| = \frac{1}{\sqrt{1-\sup \operatorname{dist}\left(\frac{Pw}{\|Pw\|},\mathcal{W}\right)^2}}.$$

To conclude the proof, we need to show that every element $x\in\mathcal{X}$, $\|x\|=1$ can be written as $Pw/\|Pw\|$, for some $w\in\mathcal{W}$. This follows from Eq. (6.3), since then

$$P\mathcal{W} = P\Pi_{\mathcal{W}}\mathcal{W} = P\Pi_{\mathcal{W}}\mathcal{H} = P\mathcal{H} = \mathcal{X}.$$

□

Let us pause and collect the geometric results derived.

Theorem 6.1 *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*

$$\frac{1}{\|P\|} = \sin\theta_{\min}(\mathcal{X},\mathcal{Y}) = g(\mathcal{X},\mathcal{Y}) = \sqrt{1-\|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\|^2} \tag{6.4}$$

$$= \cos\theta_{\max}(\mathcal{X},\mathcal{Y}^{\perp}) = \sqrt{1-G(\mathcal{X},\mathcal{Y}^{\perp})^2} = \sqrt{1-\|\Pi_{\mathcal{X}}-\Pi_{\mathcal{Y}^{\perp}}\|^2}. \tag{6.5}$$

7 Other proofs

Most of the other published proofs of Theorem 2.1 show one or another identity in Eqs. (6.4) and (6.5):

Gohberg and Kreĭn [18, Section VI.5.2] show that

$$\|P\|^{-1} = \sin\theta_{\min}(\mathcal{X},\mathcal{Y}). \tag{7.1}$$

Labrousse [30] shows that

$$\|P\|^{-1} = \sqrt{1-\|\Pi_{\mathcal{X}}-\Pi_{\mathcal{Y}^{\perp}}\|^2} = \sqrt{1-G(\mathcal{X},\mathcal{Y}^{\perp})^2}.$$

Pták [35] shows that

$$\|P\|^{-1} = \sqrt{1-\|(I-\Pi_{\mathcal{X}})\Pi_{\mathcal{Y}}\|^2} = \sqrt{1-\|(I-\Pi_{\mathcal{Y}})\Pi_{\mathcal{X}}\|^2}.$$

Ipsen and Meyer [22] show Eq. (7.1). Their presentation is geared towards classroom use in Linear Algebra courses, and the proof is for the Euclidean space.

Buckholtz [7] shows that $\|P\|^{-1} = \sqrt{1-\|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\|^2} = \sin\theta_{\min}(\mathcal{X},\mathcal{Y})$.

Wimmer [42] proves Eqs. (5.7) and (6.2), and refers to Ljance [32].

Rakočević [36] shows that $\|P\|^{-1} = \sqrt{1-\|\Pi_{\mathcal{X}}\Pi_{\mathcal{Y}}\|^2}$. He is aware of most of the previous proofs.

We conclude our review by mentioning a more general result from 1989 by Lewkowitcz [31], from which Theorem 2.1 follows as a corollary, but only valid in finite dimensions, with the Euclidean norm. Recall that in that case $\|P\| = \sigma_1(P)$, the largest singular value [21].

Theorem 7.1 [31] *Let P be a projection onto $\mathcal{X} \subset \mathbb{C}^n$ of dimension $m < n$. Let σ_k be the kth singular value. Then, the following holds:*

$$\sigma_k(P) = \sigma_k(I - P) \geq 1, \quad k = 1, \ldots, m,$$
$$\sigma_k(P) = 0,\ \sigma_k(I - P) = 1, \quad k = m + 1, \ldots, n - m,$$
$$\sigma_k(P) = \sigma_k(I - P) = 0, \quad k = n - m + 1, \ldots, n.$$

We end the paper by noting that in finite dimensions, canonical or principal angles in between the minimal and maximal can be computed using singular values; see, e.g., [6, 21, 39], and two recent papers discussing the computation of these angles [14, 26]. We also mention a very recent paper [17] on recursive definitions of these intermediate angles, and a recent thesis extending these concepts to infinite dimensions [23]; see also [40, 41].

Acknowledgments The research for this paper was supported in part by the U.S. Department of Energy under grant DE-FG02-05ER25672. The author benefited from communications with Christopher Beattie, Michele Benzi, Gustavo Corach, Pavel Guerzhoy, Andrew Knyazev, J.J. Koliha, Thomas Krainer, Itzchak Lewkowitcz, Gerardo Mendoza, Carl D. Meyer, Vladimir Rakočević, Valeria Simoncini, Ludmil Zikatanov, and two anonymous referees. Their help, comments, suggestions, and encouragement are very much appreciated.

References

1. Afriat, S.N.: Orthogonal and oblique projectors and the characteristics of pairs of vector spaces. Proc. Camb. Philos. Soc. **53**, 800–816 (1957)
2. Akhiezer, N.I., Glazman, I.M.: Theory of Linear Operators in Hilbert Space. F. Ungar, New York (1961 and 1963) (Reprinted by Dover, New York (1993))
3. Andruchov, E., Corach, G.: Geometry of oblique projections. Stud. Math. **137**, 61–79 (1999)
4. Beattie, C.: Galerkin eigenvector approximations. Math. Comput. **69**, 1409–1434 (2000)
5. Beattie, C., Embree, M., Rossi, J.: Convergence of restarted Krylov subspaces to invariant subspaces. SIAM J. Matrix Anal. Appl. **25**, 1074–1109 (2004)
6. Björck, Å., Golub, G.H.: Numerical methods for computing angles between linear subspaces. Math. Comput. **27**, 579–594 (1973)
7. Buckholtz, D.: Hilbert space idempotents and involutions. Proc. Am. Math. Soc. **128**, 1415–1418 (1999)
8. Chatelin, F.. Spectral Approximation of Linear Operators. Academic, New York (1983)
9. Corach, G., Maestripieri, A., Stojanoff, D.: A classification of projectors. In: Jarosz, K., Sołtysiak, A. (eds.) Topological Algebras, their Applications, and Related Topics. Banach Center Publications, vol. 67, pp. 145–160. Polish Academy of Sciences, Warsaw (2005)
10. Del Pasqua, D.: Su una nozione di varietà lineari disgiunte di uno spazio di Banach (On a notion of disjoint linear manifolds of a Banach space). Rend. Mat. Appl. **5**, 406–422 (1955)
11. Deutsch, F.: The angle between subspaces of a Hilbert space. In: Singh, S. P. (ed.) Approximation Theory, Wavelets and Applications, Proceedings of the NATO Advanced Study Institute with the assistance of Antonio Carbone and B. Watson, pp. 107–130. Kluwer, Dordrecht, The Netherlands (1995)
12. Dirr, G., Rakočević, V., Wimmer, H.K.: Estimates for projections in Banach spaces and existence of direct complements. Stud. Math. **170**, 211–216 (2005)

13. Dixmier, J.: Étude sur les variétés et les opérateurs de Julia, avec quelques applications (On Julia varieties and operators with applications). Bull. Soc. Math. Fr. **77**, 11–101 (1949)
14. Drmač, Z.: On principal angles between subspaces of Euclidean space. SIAM J. Matrix Anal. Appl. **22**, 173–194 (2000)
15. Eiermann, M., Ernst, O.: Geometric aspects in the theory of Krylov subspace methods. Acta Numer. **10**, 251–312 (2001)
16. Falgout, R.D., Vassilevski, P.S., Zikatanov, L.T.: On two-grid convergence estimates. Numer. Linear Algebra Appl. **12**, 471–494 (2005)
17. Galantái, A., Hegedüs, C.J.: Jordan's principal angles in complex vector spaces. Numer. Linear Algebra Appl. **13**, 589–598 (2006)
18. Gohberg, I.C., Kreĭn, M.G.: Introduction to the theory of nonselfadjoint operators. Nauka, Moscow, 1965. Translations of Mathematical Monographs, vol. 18, American Mathematical Society, Providence, Rhode Island (1969)
19. Gohberg, I.C., Lancaster, P., Rodman, L.: Invariant Subspaces of Matrices with Applications. Wiley, New York (1986) (Reprinted by SIAM, 2006. Classics in Applied Mathematics, vol. 51)
20. Gurarĭi, V.I.: Openings and inclinations of subspaces of a Banach space. Teor. Funkc. Funkc. Anal. ih Priloz. **1**, 194–204 (1965)
21. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. John Hopkins University Press, Baltimore (1996)
22. Ipsen, I.C.F., Meyer, C.D.: The angle between complementary subspaces. Am. Math. Mon. **102**, 904–911 (1995)
23. Jujunashvili, A.: Angles between infinite-dimensional subspaces. PhD thesis, Department of Applied Mathematics, University of Colorado, Denver (2005)
24. Kato, T.: Estimation of iterated matrices, with application to the von Neumann condition. Numer. Math. **2**, 22–29 (1960)
25. Kato, T.: Perturbation Theory for Linear Operators, 2nd edn. Springer, Berlin Heidelberg New York (1980)
26. Knyazev, A., Argentati, M.E.: Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates. SIAM J. Sci. Comput. **23**, 2008–2040 (2002)
27. Koliha, J.J.: Range projections of idempotents in C^*-algebras. Demonstr. Math. **34**, 91–103 (2001)
28. Koliha, J.J., Rakočević, V.: On the norm of idempotents in C^*-algebras. Rocky Mt. J. Math. **34**, 685–697 (2004)
29. Kreĭn, M.G., Krasnolsel'skiĭ, M.A., Mil'man, D.P.: On the defect numbers of linear operators in a Banach space and on some geometric questions. In: Akademiya Nauk Ukrainskoĭ RSR. Zbirnik Prac' Insttitut Matematiki. (Collection of Papers of the Institute of Mathematics of the Academy of Sciences of Ukraine), vol. 11, pp. 97–112 (1948)
30. Labrousse, J.-P.: Une caracterization topologique des générateurs infinitésimaux de semi-groupes analytiques et de contractions sur un espace de Hilbert (A topological characterization of the infinitesimal generators of analytic semigroups and of contractions on a Hilbert space). Atti Accad. Naz. Lincei, Cl. Sci. Fis. Mat. Nat., Rend. Lincei Suppl. **52**, 631–636 (1972)
31. Lewkowitcz, I.: Bounds for the singular values of a matrix with nonnegative eigenvalues. Linear Algebra Appl. **112**, 29–37 (1989)
32. Ljance, V.É.: Some properties of idempotent operators. Teor. Prikl. Matematica **1**, 16–22 (1958/1959)
33. Mandel, J., Dohrmann, C.R., Tezaur, R.: An algebraic theory for primal and dual substructuring methods by constraints. Appl. Numer. Math. **54**, 167–193 (2005)
34. Meyer, C.D.: Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia (2000)
35. Pták, V.: Extremal operators and oblique projections. Čas. Pěst. Mat. **110**, 343–350 (1985)
36. Rakočević, V.: On the norm of idempotent operators in a Hilbert space. Am. Math. Mon. **107**, 748–750 (2000)
37. Simoncini, V., Szyld, D.B.: On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods. SIAM Rev. **47**, 247–272 (2005)
38. Steinberg, J.: Oblique projections in Hilbert spaces. Integr. Equ. Oper. Theory **38**, 81–119 (2000)
39. Stewart, G.W., Sun, J.-G.: Matrix Perturbation Theory. Academic, San Diego, California, USA (1990)
40. Wedin, P.-Å.: On angles between subspaces of a finite dimensional inner product space. In: Kågström, B., Ruhe, A. (eds.) Matrix Pencils, Proceedings of a Conference Held at Pite Havsbad, Sweden, March 22–24, 1982. Lecture Notes in Mathematics, vol. 973, pp. 263–285. Springer, Berlin Heidelberg New York (1983)

41. Wimmer, H.K.: Canonical angles of unitary spaces and perturbations of direct complements. Linear Algebra Appl. **287**, 373–379 (1999)
42. Wimmer, H.K.: Lipschitz continuity of oblique projections. Proc. Am. Math. Soc. **128**, 873–876 (1999)
43. Xu, J., Zikatanov, L.: Some observations on Babuška and Brezzi theories. Numer. Math. **94**, 195–202 (2003)