2012

Compressions of linearly independent selfadjoint operators

Leiba Rodman
William & Mary

Ilya M. Spitkovsky
William & Mary, imspitkovsky@gmail.com

Linear Algebra and its Applications 436 (2012) 3757–3766

Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Compressions of linearly independent selfadjoint operators

Leiba Rodman, Ilya M. Spitkovsky *

Department of Mathematics, College of William & Mary, Williamsburg, VA 23187-8795, USA

ARTICLE INFO

Article history:
Received 24 October 2011
Accepted 31 October 2011
Available online 3 December 2011

Submitted by Peter Šemrl

AMS classification:
47A12

Keywords:
Generalized numerical range
Selfadjoint operators
Compressions
Linear independence

ABSTRACT

The following question is considered: What is the smallest number $\gamma(k)$ with the property that for every family $\{X_1, \ldots, X_k\}$ of k selfadjoint and linearly independent operators on a real or complex Hilbert space $\mathcal{H}$ there exists a subspace $\mathcal{H}_0 \subset \mathcal{H}$ of dimension $\gamma(k)$ such that the compressions of $X_1, \ldots, X_k$ to $\mathcal{H}_0$ are still linearly independent? Upper and lower bounds for $\gamma(k)$ are established for any k, and the exact value is found for $k = 2, 3$. It is also shown that the set of all $\gamma(k)$-dimensional subspaces $\mathcal{H}_0$ with the desired property is open and dense in the respective Grassmannian. The $k = 3$ case is used to prove that the ratio numerical range $W(A/B)$ of a pair of operators on a Hilbert space either has a non-empty interior, or lies in a line or a circle.

1. General problem and statements of main results

Let $\mathcal{H}$ be a Hilbert space over $\mathbb{F}$ with the inner product $\langle \cdot, \cdot \rangle$, where $\mathbb{F}$ stands for either the filed $\mathbb{R}$ of real or $\mathbb{C}$ of complex numbers. Denote by $L(\mathcal{H})$ the Banach algebra of linear bounded operators on $\mathcal{H}$, and let $X[\mathcal{H}_0]$ stand for the compression of a selfadjoint operator $X \in L(\mathcal{H})$ onto a (closed) subspace $\mathcal{H}_0$ of $\mathcal{H}$.

For every $k = 1, 2, \ldots,$ define the positive integer $\gamma_{\mathbb{F}}(k)$ as follows: For every linearly independent k-tuple of bounded selfadjoint operators $X_1, \ldots, X_k$ on $\mathcal{H}$, there is a subspace $\mathcal{H}_0$ of $\mathcal{H}$ of dimension $\gamma_{\mathbb{F}}(k)$ such that the compressions $X_1[\mathcal{H}_0], \ldots, X_k[\mathcal{H}_0]$ are linearly independent; and there exists a linearly independent k-tuple of bounded selfadjoint operators $X'_1, \ldots, X'_k$ on $\mathcal{H}$ such that the compressions $X'_1[\mathcal{H}'_0], \ldots, X'_k[\mathcal{H}'_0]$ are linearly dependent for every subspace $\mathcal{H}'_0$ of $\mathcal{H}$ of dimension $\gamma_{\mathbb{F}}(k) - 1$.

* Corresponding author.
E-mail addresses: lxrodm@math.wm.edu (L. Rodman), ilya@math.wm.edu, imspitkovsky@gmail.com (I.M. Spitkovsky).

0024-3795/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2011.10.038

For example, it is easy to see that

$$\gamma_{\mathbb{C}}(1) = \gamma_{\mathbb{R}}(1) = 1. \tag{1.1}$$

In this paper we address the problem of identifying the integers $\gamma_{\mathbb{C}}(k)$, $\gamma_{\mathbb{R}}(k)$. We came to this problem from studies of generalized numerical ranges for Hilbert space operators [1–4]. It seems to us however that the problem is independently interesting.

Bounds on $\gamma_{\mathbb{C}}(k)$, $\gamma_{\mathbb{R}}(k)$ can be easily given:

Proposition 1. *For $k = 1, 2, \ldots$,*

$$2k - 1 \geqslant \gamma_{\mathbb{F}}(k) \geqslant \begin{cases} \left\lceil \frac{k+1}{2} \right\rceil & \text{if } \mathbb{F} = \mathbb{C}, \\ k & \text{if } \mathbb{F} = \mathbb{R}, \end{cases} \tag{1.2}$$

where $\lceil x \rceil$ stands for the smallest integer greater than or equal to x.

Proof. Consider the linearly independent k-tuple of selfadjoint operators $X_1, \ldots, X_k \in L(\mathbb{F}^k)$ given in the following matrix form:

$$X_1 = \begin{bmatrix} 1 & 0_{1\times(k-1)} \\ 0_{(k-1)\times 1} & 0_{(k-1)\times(k-1)} \end{bmatrix}, \quad X_2 = \begin{bmatrix} 0 & e_1^T \\ e_1 & 0_{(k-1)\times(k-1)} \end{bmatrix},$$

$$\ldots, X_k = \begin{bmatrix} 0 & e_{k-1}^T \\ e_{k-1} & 0_{(k-1)\times(k-1)} \end{bmatrix},$$

where $e_1, \ldots, e_{k-1}$ are the standard unit coordinate vectors. Any ℓ-dimensional subspace $\mathcal{H}_0$ has intersection with Span $\{e_2, \ldots, e_k\} \subset \mathbb{F}^k$ of dimension at least $\ell - 1$. Thus, in a suitable orthonormal basis for $\mathcal{H}_0$, the compressions $\widehat{X}_j = X_j[\mathcal{H}_0]$ have the matrix form

$$\widehat{X}_j = \begin{bmatrix} * & * \\ * & 0_{(\ell-1)\times(\ell-1)} \end{bmatrix}, \quad j = 1, 2, \ldots, k.$$

Thus, for these compressions to be linearly independent, it is necessary that k does not exceed the real dimension of the space of all selfadjoint operators in $L(\mathbb{F}^\ell)$ having the form $\begin{bmatrix} * & * \\ * & 0_{(\ell-1)\times(\ell-1)} \end{bmatrix}$. This real dimension is equal to ℓ in the real case and to $2\ell - 1$ in the complex case. So

$$k \leqslant \begin{cases} 2\ell - 1 & \text{if } \mathbb{F} = \mathbb{C}, \\ \ell & \text{if } \mathbb{F} = \mathbb{R}. \end{cases} \tag{1.3}$$

These inequalities are satisfied in particular for some subspace $\mathcal{H}_0$ of dimension $\gamma_{\mathbb{F}}(k)$ for which $\widehat{X}_1, \ldots, \widehat{X}_k$ are linearly independent. Letting $\ell = \gamma_{\mathbb{F}}(k)$ in (1.3), the inequalities in (1.2) on the right follow.

To prove the inequalities on the left, let $X_1, \ldots, X_k \in L(\mathcal{H})$ be linearly independent selfadjoints, and let

$$X_j = \left[\alpha_{p,q}^{(j)}\right]_{p,q\in K}, \quad j = 1, 2, \ldots, k,$$

be the matrix representations of the X_j's with respect to an orthonormal basis $\{e_p\}_{p\in K}$ in $\mathcal{H}$ indexed by K. The linear independence of the X_j's now implies that the (possibly infinite) linear system

$$\sum_{j=1}^{k} x_j \alpha_{p,q}^{(j)} = 0, \quad p, q \in K,$$

where $x_1, \ldots, x_k$ are real unknowns, has only the trivial solution. Thus, there is a subsystem of k equations

$$\sum_{j=1}^{k} x_j \alpha_{p_u,q_v}^{(j)} = 0, \quad u, v = 1, \ldots, k, \tag{1.4}$$

that has only the trivial solution. Letting $\mathcal{H}_0$ be the subspace of $\mathcal{H}$ spanned by $e_{p_1}, \ldots, e_{p_k}, e_{q_1}, \ldots, e_{q_k}$ we see that $X_1[\mathcal{H}_0], \ldots, X_k[\mathcal{H}_0]$ are linearly independent, while $\dim \mathcal{H}_0 \leqslant 2k$.

If there are any repetitions in the sequence $\{p_1, \ldots, p_k, q_1, \ldots, q_k\}$, then $\dim \mathcal{H}_0 < 2k$ and we are done. Otherwise, let us go through another iteration of the procedure described above, with $\mathcal{H}$ replaced by $\mathcal{H}_0$, X_j by $X_j[\mathcal{H}_0]$, and $\{e_p\}$ by an orthonormal basis in which the matrix of X_1 is diagonal, and moreover $\alpha_{1,1}^{(1)} \neq 0$. The respective subsystem (1.4) of k equations can then be chosen in such a way that $p_0 = q_0 = 1$, thus guaranteeing that $\dim \mathcal{H}_0 \leqslant 2k - 1$. $\square$

In this paper we prove that the lower bounds for $\gamma_{\mathbb{F}}(k)$ specified in (1.2) are actually attained for $k = 2, 3$ (note that the case $k = 1$ is trivial):

Theorem 2. *Let $X_1, X_2 \in L(\mathcal{H})$ be linearly independent selfadjoint operators. Then there exists a 2-dimensional subspace $\mathcal{H}_0 \subseteq \mathcal{H}$ such that $X_1[\mathcal{H}_0], X_2[\mathcal{H}_0]$ are linearly independent (over $\mathbb{R}$, or equivalently, over $\mathbb{C}$).*

Theorem 3. *Let $X_1, X_2, X_3 \in L(\mathcal{H})$ be linearly independent selfadjoint operators. Then there exists a 2-dimensional subspace $\mathcal{H}_0 \subseteq \mathcal{H}$ in the complex case, and a 3-dimensional subspace $\mathcal{H}_0 \subseteq \mathcal{H}$ in the real case, such that the compressions $X_1[\mathcal{H}_0], X_2[\mathcal{H}_0], X_3[\mathcal{H}_0]$ are linearly independent.*

Determination of $\gamma_{\mathbb{F}}(k)$ for $k \geqslant 4$ remains an open problem.

Theorems 2 and 3 are proved in Sections 3 and 2, respectively. We return to the general problem in Section 4 where it is proved that the set of subspaces $\mathcal{H}_0$ of dimension $\gamma_{\mathbb{F}}(k)$ for which $X_1[\mathcal{H}_0], \ldots, X_k[\mathcal{H}_0]$ are linearly independent is open and dense. Finally, in Section 5 we apply Theorem 3 to prove a geometric property of the ratio numerical range (Theorem 6).

2. Proof of the main result: triples of operators

In order not to interrupt the exposition flow and for convenience of reference, we first prove an elementary auxiliary result, which is needed for the proof of the complex version only.

Lemma 4. *Let x_1, x_2, x_3 be linearly independent (over $\mathbb{C}$) vectors in a complex inner product space. Then for every nonzero vector y_1 lying in $\mathcal{L} = \operatorname{Span}\{x_1, x_2, x_3\}$ there exists $y_2 \in \mathcal{L} \setminus \{0\}$, $y_2 \perp y_1$, such that the orthogonal projections of x_j onto $\operatorname{Span}\{y_1, y_2\}$ are linearly independent over $\mathbb{R}$.*

Proof. Choose an orthonormal basis in $\mathcal{L}$ with its first vector collinear with y_1. The first coordinate of at least one vector x_j in this basis must be different from zero, because otherwise they would be lying in a 2-dimensional subspace and therefore linearly dependent. Scaling and renumbering if needed, we may without loss of generality suppose that the first coordinate of x_1 is one. Further adjusting x_2 by adding real multiples of x_1 and x_3 by adding real linear combinations of x_1, x_2 (which has no influence on their linear dependence or independence over $\mathbb{R}$) we may arrange for the first coordinate of x_2 to equal i or 0 while the first coordinate of x_3 is 0.

If the first coordinate of x_2 is i, then we may choose y_2 as any vector in $\mathcal{L}_1^\perp$ which is not orthogonal to x_3. Otherwise (that is, if the first coordinates of x_2 and x_3 are both equal to zero) the linear independence over $\mathbb{R}$ of the projections of x_1, x_2, x_3 onto Span$\{y_1, y_2\}$ will be guaranteed if $y_2 \perp y_1$ and the scalar products $\langle x_2, y_2\rangle$, $\langle x_3, y_2\rangle$ are $\mathbb{R}$-*independent*, that is, they are both nonzero and their ratio is not real. In the notation

$$x_2 = \begin{bmatrix} 0 \\ \xi_1 \\ \xi_2 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ \eta_1 \\ \eta_2 \end{bmatrix}, \quad y_2 = \begin{bmatrix} 0 \\ \zeta_1 \\ \zeta_2 \end{bmatrix},$$

$\mathbb{R}$-independence of $\langle x_2, y_2\rangle$, and $\langle x_3, y_2\rangle$ holds whenever $\xi_1 + \xi_2 z$ and $\eta_1 + \eta_2 z$ are $\mathbb{R}$-independent, where $z = \overline{\zeta_2}/\overline{\zeta_1}$, $\zeta_1 \neq 0$, and also for $\zeta_1 = 0$, $\zeta_2 \neq 0$, provided that ξ_2 and η_2 are $\mathbb{R}$-independent. Since x_2 and x_3 are linearly independent, there is a plenitude of vectors y_2 satisfying these requirements. $\square$

We now start the proof of Theorem 3. Let us consider separately two scenarios, depending on whether or not the operators X_j are *locally linearly independent*, that is, whether or not there exist $u \in \mathcal{H}$ such that the vectors $X_j u$, $j = 1, 2, 3$, are linearly independent (over $\mathbb{C}$). The interested reader may consult, e.g., [5] and references therein for some general structural results on systems of locally linearly dependent operators and the history of the subject. For our purposes, however, we merely need the definition.

Case 1. Locally linearly dependent X_j, $j = 1, 2, 3$. Then

$$\text{for all } u \in \mathcal{H}, \text{ the vectors } X_1u, X_2u, X_3u \text{ are linearly dependent.} \tag{2.1}$$

According to Proposition 1, it suffices to consider the case of finite (at most 5-) dimensional $\mathcal{H}$. Of course, we may also suppose that $\dim \mathcal{H} \geqslant 3$, because otherwise there is nothing to prove.

Denote by m be the maximal rank of all operators in the real span of X_1, X_2, X_3. Passing to a different basis of this span if needed, we may without loss of generality suppose that X_1 has rank m. With an appropriate choice of an orthonormal basis $E = \{e_j\}$ in $\mathcal{H}$, X_1 can be represented in the block matrix form as

$$X_1 = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix},$$

where $\Lambda = \text{diag}[\lambda_1, \ldots, \lambda_m]$, the diagonal matrix with real nonzero $\lambda_1, \ldots, \lambda_m$ on the main diagonal. Let

$$X = \begin{bmatrix} A & B \\ B^* & D \end{bmatrix} \tag{2.2}$$

be the matching block matrix representation of any X from the real span of X_j. Then for sufficiently small t the left upper block $\Lambda + tA$ of $X_1 + tX$ is invertible along with Λ, and from the Schur complement formula we conclude that $X_1 + tX$ is congruent to the direct sum of $\Lambda + tA$ with

$$Z(t) = tD - t^2B^*(\Lambda + tA)^{-1}B.$$

Since the rank of $X_1 + tX$ should not exceed m, this is only possible if $Z(t) = 0$ for all t in some neighborhood of zero. In particular, $D = 0$.

Consequently, the (i, j) entries of X_2 and X_3 are equal to zero if $i, j > m$.

Subcase 1a. There exists a pair (i, j) such that the (i, j) entries of the matrices of X_2 and X_3 (in the same basis E that diagonalizes X_1) are $\mathbb{R}$-independent. Then $i \leqslant m$ or $j \leqslant m$ (since otherwise both entries are zeros) and $i \neq j$ (since otherwise they are both real). Let $\mathcal{H}_0 = \text{Span}\{e_i, e_j\}$. Then the matrix of $X_1[\mathcal{H}_0]$ is nonzero and diagonal, while the matrices of $X_2[\mathcal{H}_0]$ and $X_3[\mathcal{H}_0]$ have $\mathbb{R}$-independent off-diagonal entries. This makes the triple $\{X_j[\mathcal{H}_0] : j = 1, 2, 3\}$ linearly independent, which settles the subcase under consideration.

In the rest of the proof for Case 1 we may therefore suppose that the (i, j) entries of X_2 and X_3 are $\mathbb{R}$-dependent for any given pair (i, j).

Subcase 1b. The matrix of at least one of the operators X_2, X_3 in the basis $\{e_j\}$ is not diagonal.

There must be a nonzero entry in the off-diagonal part of the block $\begin{bmatrix} A \\ B^* \end{bmatrix}$ in representation (2.2) of either X_2 or X_3. Relabeling vectors $e_1, \dots, e_m$, switching between X_2 and X_3 if needed, and denoting the matrix of X_3 by (ξ_{ij}), we may without loss of generality suppose that $\xi_{i1} \neq 0$ for some $i > 1$. Then the $(i, 1)$ entry of X_2 equals $r\xi_{i1}$ for some real r. Substituting X_2 by $X_2 - rX_3$ we may suppose that this entry is actually equal to zero. But then, from linear dependence of $X_je_1, j = 1, 2, 3$ we conclude that all entries in the first column of X_2 (except maybe for its $(1, 1)$-entry) are equal to zero. Due to selfadjointness, this also holds for the elements in the first row of X_2. Finally, subtracting from X_2 an appropriate real scalar multiple of X_1, we can set its $(1, 1)$ entry to equal zero as well.

Now, if the ith diagonal entry of X_2 (which we will denote by μ_i) is nonzero, then for $\mathcal{H}_0 = \operatorname{Span}\{e_1, e_i\}$ the matrices of $X_j[\mathcal{H}_0]$ for $j = 1, 2, 3$ take the form

$$\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_i \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & \mu_i \end{bmatrix}, \quad \begin{bmatrix} * & \xi_{1i} \\ \xi_{i1} & * \end{bmatrix}$$

respectively, and are therefore linearly independent.

We will show that the alternative (that is, the situation when $\mu_i = 0$ whenever $\xi_{i1} \neq 0$) leads to a contradiction, and therefore does not materialize. To this end, consider first $i \leqslant m$ and apply (2.1) to $u = e_i$. Since $X_1u = \lambda_ie_i$ and X_3u having the first coordinate $\overline{\xi_{1i}}$ are then linearly independent, it follows that X_2u must be their linear combination. Making use of the fact that the first and ith coordinates of X_2e_i are equal to zero, we conclude that the whole ith column of X_2 is equal to zero.

If $i > m$, we reach the same conclusion invoking (2.1) with $u = e_1 + te_i$ and varying t. Indeed, then $X_1u = \lambda_1e_1$ while X_3u has a nonzero ith coordinate for all except for maybe one value of t. Since $X_2u = tX_2e_i$, the ith column of X_2 lies in the span of e_1 and a vector with nonzero ith coordinate. This is again possible only if the combination is trivial.

We have thus shown that the ith column (and therefore the ith row) of X_2 is zero for any i such that $\xi_{i1} \neq 0$.

Suppose that the matrix of X_2 is nevertheless not diagonal. Then it must contain a nonzero entry say in (s, j) position with $1 < j \leqslant m, j \neq i$ and $s \neq 1, i, j$. (Here $i \neq 1$ is any index such that $\xi_{i1} \neq 0$.) Let us invoke (2.1) once again, this time for $u = e_1 + te_j$. For $t \neq 0$ the vector $X_2u = tX_2e_j$ has a nonzero sth entry, while $X_1u = \lambda_1e_1 + t\lambda_je_j$. Consequently, $X_3u = X_3e_1 + tX_3e_j$ should be a linear combination of X_2e_j and $\lambda_1e_1 + t\lambda_je_j$, and therefore its ith coordinate is zero. However, for t small enough it is different from zero.

The contradiction obtained shows that the matrix of X_2 in the basis $\{e_j\}$ is $\operatorname{diag}[0, \mu_2, \dots, \mu_m, 0, \dots, 0]$. Note that $\mu_j \neq 0$ for at least one value of j. From (2.1) with $u = e_1 + te_j$ we conclude that for $t \neq 0$, $X_3u = X_3e_1 + tX_3e_j$ must be a linear combination of the linearly independent vectors $X_1u = \lambda_1e_1 + t\lambda_je_j$ and $X_2u = t\mu_je_j$. Let $i \neq 1$ be any index such that $\xi_{i1} \neq 0$. Then $\mu_i = 0$, hence $j \neq i$. Thus, the ith coordinate of X_1u and $X_2u = t\mu_je_j$ is zero, while for X_3u it differs from zero. The contradiction obtained concludes the consideration of this subcase.

The only remaining situation in Case 1 is as follows.

Subcase 1c. The matrices of all X_j in the basis E are diagonal.

Consistently with the notation and conventions made earlier, let X_1 and X_2 have the matrices $\operatorname{diag}[\lambda_1, \dots, \lambda_m, 0, \dots, 0]$ and $\operatorname{diag}[0, \mu_2, \dots, \mu_m, 0, \dots, 0]$, respectively. By changing the order of vectors in E if necessary, we may suppose that $\mu_2 \neq 0$. Subtracting then a suitable linear combination of X_1 and X_2 from X_3, we can reduce the matrix of the latter to

$$\operatorname{diag}[0, 0, \nu_3, \ldots, \nu_m, 0, \ldots, 0].$$

Now take $u = e_1 + e_2 + e_k$ with $k \neq 1, 2$. Since $X_1 u = \lambda_1 e_1 + \lambda_2 e_2 + \lambda_k e_k$, $X_2 u = \mu_2 e_2 + \mu_k e_k$ and $X_3 u = \nu_k e_k$, property (2.1) for this choice of u implies that $\nu_k = 0$. In other words, $X_3 = 0$, which is in contradiction with linear independence of the set $\{X_1, X_2, X_3\}$.

This concludes the consideration of Case 1. Note that in this case we were able to come up with 2-dimensional compressions, both in real and in complex setting.

Case 2. Locally linearly independent X_j, $j = 1, 2, 3$ i.e. $X_1 u, X_2 u, X_3 u$ are linearly independent for some $u \in \mathcal{H}$.

We claim that then there exists a 3-dimensional subspace $\mathcal{H}_1$ of $\mathcal{H}$ and $v \in \mathcal{H}_1$ such that the vectors $X_j[\mathcal{H}_1]v$, $j = 1, 2, 3$, are linearly independent. Indeed, if $u \in \operatorname{Span}\{X_1 u, X_2 u, X_3 u\}$, we may simply take

$$\mathcal{H}_1 = \operatorname{Span}\{X_1 u, X_2 u, X_3 u\} \text{ and } v = u.$$

Otherwise, let v be a perturbation of u so small that the set $\{v, X_1 v, X_2 v, X_3 v\} =: \Gamma$ is still linearly independent, while in addition $\langle X_1 v, v\rangle \neq 0$ (such a perturbation exists, since for any nonzero $A \in L(\mathcal{H})$ the set of vectors x for which $\langle Ax, x\rangle \neq 0$ is dense in $\mathcal{H}$). Let Z be the 4×3 matrix whose jth column equals the coordinate vector of $X_j v$ in the orthonormal basis $\{v_1, \ldots, v_4\}$ obtained from Γ by the Gram–Schmidt procedure. Since the rank of Z equals three and its (1,1) entry is nonzero, for some $j_0 = 2, 3$ or 4 the deletion of the jth row of Z yields a nonsingular 3×3 submatrix. It remains to choose $\mathcal{H}_1$ as the span of v_j, $j \in \{1, 2, 3, 4\} \setminus \{j_0\}$.

In the real setting, $\mathcal{H}_0 = \mathcal{H}_1$ is the desired 3-dimensional subspace. In the complex setting, we need to go one step further. Namely, let us apply Lemma 4 to $x_j = X_j[\mathcal{H}_1]v$, $j = 1, 2, 3$, and $y_1 = v$. We end up with a 2-dimensional subspace $\mathcal{H}_0 = \operatorname{Span}\{y_1, y_2\}$ and an orthonormal basis in it such that the first columns of the matrices of $X_j[\mathcal{H}_0]$ with respect to this basis are linearly independent over $\mathbb{R}$. But then $X_j[\mathcal{H}_0]$ themselves also are linearly independent (over $\mathbb{R}$ or equivalently, over $\mathbb{C}$).

3. Proof of the main result: pairs of operators

In the complex setting, if $X_1, X_2 \in L(\mathcal{H})$ are linearly independent selfadjoints, we simply adjoin a selfadjoint $X_3 \in L(\mathcal{H})$ so that X_1, X_2, X_3 are linearly independent, and use Theorem 3 for X_1, X_2, X_3.

Consider now the real case. Let $X_1, X_2 \in L(\mathcal{H})$ be linearly independent selfadjoints, where $\mathcal{H}$ is a real Hilbert space. By Proposition 1, we may (and do) assume that $\mathcal{H}$ is at most 3-dimensional. In fact, we may suppose that $\dim \mathcal{H} = 3$, since otherwise $\dim \mathcal{H} = 2$ and there is nothing to prove.

Step 1. Let us first show that X_1 and X_2 are locally linearly independent.

As in Section 2, with respect to a suitable orthonormal basis $\{e_j\}$,

$$X_1 = \begin{bmatrix} \Lambda & 0 \\ 0 & 0 \end{bmatrix}, \quad X_2 = \begin{bmatrix} A & B \\ B^* & 0 \end{bmatrix},$$

where $\Lambda = \operatorname{diag}[\lambda_1, \ldots, \lambda_m]$ with real nonzero $\lambda_1, \ldots, \lambda_m$. Arguing by contradiction, assume that for every $u \in \operatorname{Span}\{e_1, \ldots, e_m\} \setminus \{0\}$ there exists a real number μ_u such that $X_2 u = \mu_u X_1 u$. Take linearly independent $u, v \in \operatorname{Span}\{e_1, \ldots, e_m\}$ (we leave aside for the time being the case $m = 1$), and write:

$$X_2 u = \mu_u X_1 u, \quad X_2 u = \mu_v X_1 v, \quad X_2(u + v) = \mu_{u+v} X_1(u + v).$$

Subtracting the first two equalities from the third, we obtain

$$0 = (\mu_{u+v} - \mu_u) X_1 u + (\mu_{u+v} - \mu_v) X_1 v.$$

Linear independence of u and v and injectivity of X_1 on Span $\{e_1, \ldots, e_m\}$ yield

$$\mu_{u+v} - \mu_u = 0, \quad \mu_{u+v} - \mu_v = 0.$$

Thus, $\mu_u = \mu_v$, and since u, v is an arbitrary linearly independent pair of vectors in Span $\{e_1, \ldots, e_m\}$, it follows that in fact $\mu := \mu_u$ is independent of $u \in$ Span $\{e_1, \ldots, e_m\} \setminus \{0\}$. The same conclusion obviously holds also in the case $m = 1$. We now have $A = \mu\Lambda$, $B = 0$, and $X_2 = \mu X_1$, a contradiction with the linear independence of X_1 and X_2.

Step 2. Analogously to the proof in Section 2, the real setting of Case 2, it follows from the linear independence of $\{X_1 u, X_2 u\}$ that there exist a 2-dimensional subspace $\mathcal{H}_1$ of $\mathcal{H}$ and $v \in \mathcal{H}_1$ such that $X_1[\mathcal{H}_1]v$ and $X_2[\mathcal{H}_1]v$ are linearly independent. Thus, $X_1[\mathcal{H}_1]$ and $X_2[\mathcal{H}_1]$ are linearly independent, and the proof of Theorem 2 is complete.

4. Openness and denseness

Let $\mathcal{H}$ be real or complex Hilbert space. Fix a positive integer $\ell \leqslant \dim \mathcal{H}$. It is well known (the basic references in the context of Banach space are [6–8], see also the English translation [9] of [7]) that the set $\mathrm{Gr}_\ell\, \mathcal{H}$ of all ℓ-dimensional subspaces of $\mathcal{H}$ is a connected complete metric space, called the *Grassmannian*. In the Hilbert space setting, the standard metric on the Grassmannian is given by the *gap* $\theta(\mathcal{M}, \mathcal{N})$ between two subspaces $\mathcal{M}, \mathcal{N} \subseteq \mathcal{H}$:

$$\theta(\mathcal{M}, \mathcal{N}) = \|P_{\mathcal{M}} - P_{\mathcal{N}}\|,$$

where $P_{\mathcal{M}}$ stands for the orthogonal projection onto $\mathcal{M}$. Equivalently, the metric topology in $\mathrm{Gr}_\ell\, \mathcal{H}$ can be given in terms of orthonormal bases: If $f_1, \ldots, f_\ell$ is an orthonormal basis in $\mathcal{M} \in \mathrm{Gr}_\ell\, \mathcal{H}$, then a basis of open neighborhoods of $\mathcal{M}$ is given by

$$\{\mathrm{Span}\{f_1', \ldots, f_\ell'\} \colon \{f_1', \ldots, f_\ell'\} \text{ is orthonormal}, \quad \|f_j' - f_j\| < 1/m,\ j = 1, 2, \ldots, \ell\},$$

for $m = 1, 2, \ldots$. Indeed, if $\{f_j\}_{j=1}^{\ell}$, $\{f_j'\}_{j=1}^{\ell}$ are orthonormal bases for $\mathcal{M}, \mathcal{M} \in \mathrm{Gr}_\ell\, \mathcal{H}$, respectively, and if $\|f_j' - f_j\| < \alpha, j = 1, 2, \ldots, \ell$, for some positive α, then the formulas

$$P_{\mathcal{M}} x = \sum_{j=1}^{\ell} \langle x, f_j \rangle f_j, \quad P_{\mathcal{M}'} x = \sum_{j=1}^{\ell} \langle x, f_j' \rangle f_j', \quad x \in \mathcal{H}$$

easily lead to the inequality $\|P_{\mathcal{M}} - P_{\mathcal{M}'}\| \leqslant 2\ell\alpha$. Conversely, if $\|P_{\mathcal{M}} - P_{\mathcal{M}'}\| \leqslant \beta$ for $\mathcal{M}' \in \mathrm{Gr}_\ell\, \mathcal{H}$ and for some positive β, then let $u_j = P_{\mathcal{M}'} f_j, j = 1, 2, \ldots, \ell$, where $\{f_1, \ldots, f_\ell\}$ is a given orthonormal basis for $\mathcal{M}$. Using the inequalities $\left|\langle u_i, u_j \rangle - \langle f_i, f_j \rangle\right| \leqslant 2\beta$, we see that if β is sufficiently small, then the Gram matrix of $\{u_1, \ldots, u_\ell\}$ is invertible, hence $\{u_1, \ldots, u_\ell\}$ are linearly independent. Now the Gram–Schmidt procedure applied to $\{u_1, \ldots, u_\ell\}$ yields an orthonormal basis $\{f_j'\}_{j=1}^{\ell}$ for $\mathcal{M}'$ satisfying $\|f_j' - f_j\| \leqslant C\beta$, where the positive constant C depends on ℓ only, as it follows from the well-known formulas for this procedure, see e.g. [10, Theorem 9.1].

Theorem 5. *Let $X_1, \ldots, X_k \in L(\mathcal{H})$ be linearly independent selfadjoint operators, where $\mathcal{H}$ is an $\mathbb{F}$-Hilbert space, $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Then the set* $\mathrm{LinInd}_{\mathbb{F}}\, \{X_1, \ldots, X_k\}$ *of all $\gamma_{\mathbb{F}}(k)$-dimensional subspaces $\mathcal{H}_0 \subseteq \mathcal{H}$ such that the compressions $X_1[\mathcal{H}_0], X_2[\mathcal{H}_0], \ldots, X_k[\mathcal{H}_0]$ are linearly independent is open and dense in $\mathrm{Gr}_{\gamma_{\mathbb{F}}(k)} \mathcal{H}$.*

Proof. We prove the result for the complex case only; the proof in the real case is completely analogous.
We begin with denseness. Let

$$\mathcal{H}_1 \in \mathrm{LinInd}_{\mathbb{C}}\, \{X_1, \ldots, X_k\},$$

and let $\mathcal{H}_0$ be any subspace of $\mathcal{H}$ of dimension $\ell := \gamma_{\mathbb{C}}(k)$. Assuming the intersection $\mathcal{H}_1 \cap \mathcal{H}_0$ is p-dimensional, let $f_1, \ldots f_\ell$ and $g_1, \ldots, g_\ell$ be orthonormal bases in $\mathcal{H}_1$ and $\mathcal{H}_0$, respectively, such that

$f_1 = g_1, \ldots, f_p = g_p$ is a basis for $\mathcal{H}_1 \cap \mathcal{H}_0$. (The situation when $\mathcal{H}_1 \cap \mathcal{H}_0 = \{0\}$ is not excluded; in this case the proof goes through with obvious changes.) Consider

$$f_j(t) := \begin{cases} f_j & j = 1, 2, \ldots, p; \quad -\infty < t < \infty, \\ \\ tf_j + (1-t)g_j & j = p+1, \ldots, \ell; \quad -\infty < t < \infty. \end{cases}$$

It is easy to see that $\{f_1(t), \ldots, f_\ell(t)\}$ are linearly independent for every $t \in \mathbb{R}$. Let $\{h_1(t), \ldots, h_\ell(t)\}$ be an orthonormal system obtained from $\{f_1(t), \ldots, f_\ell(t)\}$ by using the Gram–Schmidt process. Clearly,

$$\{h_1(1), \ldots, h_\ell(1)\} = \{f_1, \ldots, f_\ell\} \text{ and } \{h_1(0), \ldots, h_\ell(0)\} = \{g_1, \ldots, g_\ell\}.$$

Moreover, $\mathcal{H}(t) := \mathrm{Span}\,\{h_1(t), \ldots, h_\ell(t)\}$ belongs to $\mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}$ if and only if the $k \times \ell^2$ matrix $Z(t)$ whose jth row is

$$\begin{aligned} &[\ \langle X_j h_1(t), h_1(t)\rangle, \ldots, \langle X_j h_\ell(t), h_\ell(t)\rangle, \\ &\ \Re\langle X_j h_1(t), h_2(t)\rangle, \Im\langle X_j h_1(t), h_2(t)\rangle, \ldots, \Re\langle X_j h_1(t), h_\ell(t)\rangle, \Im\langle X_j h_1(t), h_\ell(t)\rangle, \\ &\ \ldots, \Re\langle X_j h_{\ell-1}(t), h_\ell(t)\rangle, \Im\langle X_j h_{\ell-1}(t), h_\ell(t)\rangle\ ] \end{aligned}$$

has rank k. (We denote by $\Re z$ and $\Im z$ the real and imaginary parts of a complex number z.) Observe that the determinants of the $k \times k$ submatrices of $Z(t)$ are real analytic functions of the real variable t and that at least one of those determinants is nonzero for $t = 1$. Since the zeros of a real analytic function cannot accumulate within its domain of definition (unless the function is identically zero), it follows that the same determinant is nonzero for all $t \in (0, \epsilon)$, where $\epsilon > 0$ is sufficiently small. Thus,

$$\mathrm{Span}\{h_1(t), \ldots, h_\ell(t)\} \in \mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}, \quad t \in (0, \epsilon),$$

and

$$\lim_{t \to 0+} \theta(\mathcal{H}_0, \mathrm{Span}\{h_1(t), \ldots, h_\ell(t)\}) = 0,$$

which proves the denseness of $\mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}$.

Next, we prove that $\mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}$ is open.

Let $\mathcal{H}_1 \in \mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}$, and let $f_1, \ldots f_\ell$ be an orthonormal basis for $\mathcal{H}_1$. For any orthonormal system $\{h_1, \ldots, h_\ell\}$ we denote by $Z(\{h_1, \ldots, h_\ell\})$ the $k \times \ell^2$ matrix whose jth row is

$$\begin{aligned} &[\ \langle X_j h_1, h_1\rangle, \ldots, \langle X_j h_\ell, h_\ell\rangle, \Re\langle X_j h_1, h_2\rangle, \Im\langle X_j h_1, h_2\rangle, \ldots, \\ &\ \Re\langle X_j h_1, h_\ell\rangle, \Im\langle X_j h_1, h_\ell\rangle, \ldots, \Re\langle X_j h_{\ell-1}, h_\ell\rangle, \Im\langle X_j h_{\ell-1}, h_\ell\rangle\ ], \end{aligned}$$

for $j = 1, 2, \ldots, k$. Then at least one $k \times k$ minor of $Z(\{f_1, \ldots f_\ell\})$ is nonzero. Clearly, there exists $\delta > 0$ such that the determinant of the same $k \times k$ submatrix of $Z(\{h_1, \ldots, h_\ell\})$ is nonzero for every orthonormal system $\{h_1, \ldots, h_\ell\}$ satisfying $\|h_j - f_j\| < \delta, j = 1, 2, \ldots, k$. From our description of a basis of open neighborhoods given at the beginning of this section, the openness of $\mathrm{LinInd}_{\mathbb{C}}\,\{X_1, \ldots, X_k\}$ follows. $\square$

5. Application to ratio numerical ranges

For $A, B \in L(\mathcal{H})$, $B \neq 0$, the *ratio numerical range* $W(A/B)$ was introduced in [3] as

$$W(A/B) = \{\langle Ax, x\rangle / \langle Bx, x\rangle : \langle Bx, x\rangle \neq 0\}.$$

Note that $W(B/I)$ is simply the classical numerical range (or the *field of values*) $W(A)$ of A and that in the particular case of finite dimensional $\mathcal{H}$ and $0 \notin W(B)$ the set $W(A/B)$ was studied earlier in [1]. On the other hand, if $\dim \mathcal{H} < \infty$ and A, B are such that

$$\langle Bx, x\rangle = \langle Ax, x\rangle = 0 \implies x = 0, \tag{5.1}$$

then $W(B/A)$ coincides with the numerical range of the matrix pencil $\lambda B - A$ treated in [11,2].

From [11] it follows in particular that, under condition (5.1) and in the finite dimensional case, $W(A/B)$ is either contained in a line or a circle, or has a non-empty interior. We will show that condition (5.1) is redundant and the statement actually holds in any dimension.

Theorem 6. *The ratio numerical range $W(A/B)$ either lies in a line, or in a circle, or has a non-empty interior.*

Proof. Using the standard notation

$$\Re(X) = \frac{1}{2}(X + X^*), \quad \Im(X) = \frac{1}{2i}(X - X^*)$$

for $X \in L(\mathcal{H})$, let us denote by $r(A, B)$ the number of linearly independent (over $\mathbb{R}$) operators among $\Re(A), \Im(A), \Re(B), \Im(B)$. According to [3, Theorem 3.1] $W(A/B)$ lies in a line or a circle if and only if $r(A, B) \leqslant 2$. So, we need only to verify that if $r(A, B) \geqslant 3$, then $W(A/B)$ has non-empty interior.

By Theorem 3, it is possible to find a 2-dimensional subspace $\mathcal{H}_0$ such that for the compressions A_0, B_0 of A, B onto $\mathcal{H}_0$ we still have $r(A_0, B_0) \geqslant 3$. Since $W(A_0/B_0) \subset W(A/B)$, it suffices to consider the case $\dim \mathcal{H} = 2$.

Recall the obvious properties $W((A - \lambda B)/B) = W(A/B) - \lambda$ and $W(A/B) \setminus \{0\} = 1/W(B/A)$ due to which the interiors of $W(A/B)$ and $W(B/(A - \lambda B))$ are (or are not) empty simultaneously. Considering $W(B/(A - \lambda B))$ with $\lambda = \operatorname{tr}(A)/\operatorname{tr}(B)$ instead of $W(A/B)$ if $\operatorname{tr}(B) \neq 0$, we therefore may without loss of generality suppose that B is traceless.

We may suppose in addition that for some unit vector $x \in \mathcal{H}$, $\langle Ax, x\rangle = \langle Bx, x\rangle = 0$, since otherwise (5.1) holds. Choosing this x as the first vector of an orthonormal basis of $\mathcal{H}$, we obtain the following matrix representations:

$$A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} 0 & b_{12} \\ b_{21} & 0 \end{bmatrix}. \tag{5.2}$$

Note that $a_{22} \neq 0$ due to the condition imposed on $r(A, B)$.

In the definition of $W(A/B)$ only unit vectors in $\mathcal{H}$ both coordinates of which are nonzero will have an impact. Denoting by s the ratio of their absolute values and by θ the difference of arguments, we think of $W(A/B)$ as the range of the function

$$f(s, \theta) = \frac{a_{22}s + a_{12}e^{i\theta} + a_{21}e^{-i\theta}}{b_{12}e^{i\theta} + b_{21}e^{-i\theta}} \tag{5.3}$$

on the set $\{s > 0, \theta \neq \arg(-b_{21}/b_{12}) \mod \pi \text{ if } |b_{12}| = |b_{21}|\}$.

Suppose the interior of $W(A/B)$ is empty. Then the Jacobian of the mapping (5.3) is identically zero, that is, the ratio $\frac{\partial f}{\partial \theta} \Big/ \frac{\partial f}{\partial s}$ must be real for all admissible values of s and θ. A direct computation shows that this ratio is a linear function of s, namely

$$\frac{\partial f}{\partial \theta}(s, \theta) \Big/ \frac{\partial f}{\partial s}(s, \theta) = m(\theta)s + n(\theta),$$

where

$$m(\theta) = -i\frac{b_{12}e^{i\theta} - b_{21}e^{-i\theta}}{b_{12}e^{i\theta} + b_{21}e^{-i\theta}} \tag{5.4}$$

and

$$n(\theta) = \frac{i}{a_{22}}\left(a_{12}e^{i\theta} - a_{21}e^{-i\theta} - (a_{12}e^{i\theta} + a_{21}e^{-i\theta})\frac{b_{12}e^{i\theta} - b_{21}e^{-i\theta}}{b_{12}e^{i\theta} + b_{21}e^{-i\theta}}\right). \tag{5.5}$$

So, both $m(\theta)$ and $n(\theta)$ must be real for all θ such that $b_{12}e^{i\theta} + b_{21}e^{-i\theta} \neq 0$.

From (5.4) we conclude then that $|b_{12}| = |b_{21}|$. Applying a diagonal unitary similarity to A and B given by (5.2), which of course does not change $W(A/B)$ and $r(A, B)$, we may then arrange that $b_{12} = b_{21}$. Scaling both A and B by their inverse (which once again will not change the ratio numerical range and the number $r(A, B)$) we may even suppose that $b_{12} = b_{21} = 1$. Under this condition (5.5) simplifies to

$$\begin{aligned}
&\frac{i}{a_{22}}\left(a_{12}e^{i\theta} - a_{21}e^{-i\theta} - i(a_{12}e^{i\theta} + a_{21}e^{-i\theta})\tan\theta\right) \\
&= \frac{i}{a_{22}}\left(a_{12}e^{i\theta}(1 - i\tan\theta) - a_{21}e^{-i\theta}(1 + i\tan\theta)\right) \\
&= \frac{i}{a_{22}}(1 - i\tan\theta)e^{i\theta}(a_{12} - a_{21}),
\end{aligned}$$

which can be real for all the admissible (and thus, by continuity, for all) values of θ only if $a_{12} = a_{21}$. But then $\Re(B)(= B)$, $\Im(B)(= 0)$, $\Re(A)$ and $\Im(A)$ lie in the span of two matrices, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. This contradiction with $r(A, B) \geqslant 3$ completes the proof. $\square$

References

[1] E. Einstein, C.R. Johnson, B. Lins, I.M. Spitkovsky, The ratio field of values, Linear Algebra Appl. 434 (2011) 1119–1136.
[2] P.J. Psarrakos, Numerical range of linear pencils, Linear Algebra Appl. 317 (1–3) (2000) 127–141.
[3] L. Rodman, I.M. Spitkovsky, Ratio numerical ranges of operators, Integral Equations Operator Theory 71 (2011) 245–257.
[4] M.T. Chien, H. Nakazato, The numerical range of linear pencils of 2-by-2 matrices, Linear Algebra Appl. 341 (2002) 69–100.
[5] M. Brešar, P. Šemrl, On locally linearly dependent operators and derivations, Trans. Amer. Math. Soc. 351 (3) (1999) 1257–1275.
[6] I.C. Gohberg, A.S. Markus, Two theorems on the gap between subspaces of a Banach space, Uspehi Mat. Nauk 14 (5(89)) (1959) 135–140.
[7] I.C. Gohberg, M.G. Kreĭn, Fundamental aspects of defect numbers, root numbers and indexes of linear operators, Uspehi Mat. Nauk (N.S.) 12 (2(74)) (1957) 43–118.
[8] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, Berlin, 1995. (reprint of the 1980 edition).
[9] I.C. Gohberg, M.G. Kreĭn, The basic propositions on defect numbers, root numbers and indices of linear operators, Amer. Math. Soc. Trans. 13 (2) (1960) 185–264.
[10] I. Gohberg, S. Goldberg, Basic Operator Theory, Birkhäuser Verlag, Boston, Basel, Stuttgart, 1981.
[11] H. Nakazato, P. Psarrakos, On the shape of numerical range of matrix polynomials, Linear Algebra Appl. 338 (2001) 105–123.