Strict convexity of the singular value sequences

Bernard Aupetit, Endre Makai, jr.*, and Jaroslav Zemánek

* Research (partially) supported by Hungarian National Foundation for Scientific Research, grant no. T-016094

Abstract. If A and B are compact operators on a Hilbert space, with singular values satisfying $s_j(A) = s_j(B) = s_j((A+B)/2)$, for all $j = 1, 2, ...$, then $A = B$. Two proofs, geometric and analytic, are given.

1991 Mathematics Subject Classification: Primary 47B06; Secondary 47B07, 31A05.

0. Introduction

Let A be a compact linear operator from a Hilbert space H into a Hilbert space K. The *singular values* $s_1(A) \geq s_2(A) \geq \ldots$ are the eigenvalues of $|A| := (A^*A)^{1/2}$. We refer to [3] for other equivalent definitions and basic properties. In this note we offer two proofs, geometric and analytic, of the following uniqueness property of compact operators between Hilbert spaces.

Theorem. *If A and B are compact operators such that $s_j(A) = s_j(B) = s_j(t_0A + (1-t_0)B)$ for some $0 < t_0 < 1$ and all $j = 1, 2, \ldots$, then $A = B$.*

The assumption of the Theorem implies that the Ky Fan norms ([3], p. 37)

$$\sigma_k(A) := s_1(A) + s_2(A) + \ldots + s_k(A)$$

are constant on the segment $\{tA + (1-t)B : 0 \leq t \leq 1\}$. Therefore, each $s_j(\cdot)$ is constant on this segment as well, for $j = 1, 2, \ldots$. The analytic proof given below shows that the latter property extends to the whole real line connecting A and B. In particular, for $j = 1$, the norm $\|tA + (1-t)B\|$ is bounded for $t \to \infty$, which is impossible unless $A = B$.

If H is finite-dimensional, the squares of the singular values in question are the roots of the characteristic equation of $|tA + (1-t)B|^2$. Since the roots and the leading coefficient of this equation are constant for $0 \le t \le 1$, all its coefficients are constant for $0 \le t \le 1$. But these coefficients are polynomials in t, so they are constant for all $t \in \mathbf{R}$ and, consequently, the roots are constant for all t as well. In particular, the largest root $\|tA + (1-t)B\|^2$ is constant for $t \to \infty$, which implies that $A = B$. The infinite-dimensional situation, however, requires deeper analytic tools to derive the same conclusion (see the second proof).

If the singular values given are p-summable for some $1 < p < \infty$, the conclusion comes from the strict convexity of the Schatten p-norms [4]. So the interesting case is when the singular values in the Theorem are not p-summable for such p. The geometric proof is based on the Spectral Theorem ([3], p. 28).

The Theorem above also shows that no three distinct points of the unitary orbit of a compact operator can lie on a real line in the operator space; another proof in the finite dimensional case can be found in [5].

1. First proof

Let us denote $C := t_0A + (1-t_0)B$, and $s_j := s_j(A) = s_j(B) = s_j(C)$. The compact operators A, B, C have the representations (cf. [3], p. 28)

$$\begin{aligned} Ax &= \sum s_j \langle x, a_j' \rangle a_j, \\ Bx &= \sum s_j \langle x, b_j' \rangle b_j, \\ Cx &= \sum s_j \langle x, c_j' \rangle c_j, \end{aligned}$$

where $(a_j'), (b_j'), (c_j')$, and $(a_j), (b_j), (c_j)$ are suitable orthonormal sequences in H and K, respectively.

By hypothesis, we have $\|A\| = \|B\| = \|C\| = s_1$; we may assume $s_1 > 0$. From the above representation, we see that C attains its norm at $c_1' \in H$, satisfying $\|c_1'\| = 1$. Hence $\|C\| = \|Cc_1'\| \le t_0\|Ac_1'\| + (1-t_0)\|Bc_1'\| \le t_0\|A\| + (1-t_0)\|B\| = \|C\|$. This implies that Ac_1' and Bc_1' are linearly dependent (with a positive proportionality factor), further that both A and B attain their norms at c_1'. In particular, $\|Ac_1'\| = \|Bc_1'\|$, hence actually $Ac_1' = Bc_1'$, which implies $Ac_1' = Bc_1' = Cc_1' = s_1c_1$.

Let $k \geq 1$ be defined by $s_1 = \ldots = s_k > s_{k+1}$ (notice that the singular values of a compact operator converge to zero). Remembering that $\|c_1'\| = 1$ and $\|Ac_1'\| = \|A\| = s_1$, the above representation for A and the Parseval inequality yield

$$0 < s_1^2 = \|Ac_1'\|^2 = \sum s_j^2 |\langle c_1', a_j' \rangle|^2 \leq \sum s_1^2 |\langle c_1', a_j' \rangle|^2 \leq s_1^2.$$

Consequently, we get that $\langle c_1', a_j' \rangle = 0$ for $j \geq k+1$, and that $\sum |\langle c_1', a_j' \rangle|^2 = 1 = \|c_1'\|^2$. It follows that c_1' belongs to the linear span M' of $a_1', \ldots, a_k'$. Similarly, the same holds for c_j' with $j \leq k$.

Let T denote the invertible linear transformation from M' onto $M := \text{span}\{a_1, \ldots, a_k\}$ defined by $Ta_j' = a_j$ ($j = 1, \ldots, k$). Let V' be a unitary operator on M', and let V be the unitary operator $TV'T^{-1}$ on M. Then for $y \in M'$ we have

$$\begin{aligned}
\sum_{j=1}^{k} s_j \langle y, a_j' \rangle a_j &= s_1 T \left(\sum_{j=1}^{k} \langle y, a_j' \rangle a_j' \right) = s_1 T(y) \\
&= s_1 T \left(\sum_{j=1}^{k} \langle y, V' a_j' \rangle V' a_j' \right) \\
&= s_1 \sum_{j=1}^{k} \langle y, V' a_j' \rangle V T a_j' \\
&= \sum_{j=1}^{k} s_j \langle y, V' a_j' \rangle V a_j.
\end{aligned}$$

Let now $x \in H$ be arbitrary. Applying the last formula for y, the orthogonal projection of x to M', we obtain

$$Ax = \sum_{j=1}^{k} s_j \langle x, V' a_j' \rangle V a_j + \sum_{j>k} s_j \langle x, a_j' \rangle a_j.$$

In other words, in the original representation of A we can replace a_j' by $V'a_j'$, and a_j by Va_j, for $j = 1, \ldots, k$, retaining all further a_j' and a_j, for $j > k$. So, by a suitable choice of the unitary operator V' we can have $V'a_j' = c_j' \in M'$, for $j = 1, \ldots, k$. To simplify notation, we will assume that in the original representation of A we have $a_j' = c_j'$, for $j = 1, \ldots, k$. Similarly, we will assume that $b_j' = c_j'$, for $j = 1, \ldots, k$.

Let $1 \le j \le k$. We know that c'_j belongs to $M' = \text{span}\{a'_1, \dots, a'_k\} = \text{span}\{c'_1, \dots, c'_k\} = \text{span}\{b'_1, \dots, b'_k\}$, hence $\langle c'_j, a'_\ell \rangle = \langle c'_j, b'_\ell \rangle = 0$ for $\ell > k$. Remembering that $Ac'_j = Bc'_j = s_j c_j$, we get $s_j a_j = Ac'_j = s_j c_j = Bc'_j = s_j b_j$, hence $a_j = c_j = b_j$, because $s_j > 0$. The last equalities also imply that $\langle c_j, a_\ell \rangle = \langle c_j, b_\ell \rangle = 0$ for $\ell > k$.

So the operators A and B coincide on M', map it to M, and, moreover, map the orthogonal complement of M' to the orthogonal complement of M. We can now repeat the preceding argument for the restrictions of A, B, C to the orthogonal complement of M' and proceed by induction to finish the proof.

2. Second proof

As noticed in the introduction, it is enough to show that the spectrum of $|tA+(1-t)B|^2$ is constant for $t \in \mathbf{R}$, knowing that it is constant for $0 \le t \le 1$. This comes from the following more general lemma applied to $K(\lambda)$, the spectrum of the complex polynomial with compact (self-adjoint) coefficients

$$(\lambda A^* + (1-\lambda)B^*)(\lambda A + (1-\lambda)B),$$

which for $\lambda \in \mathbf{R}$ coincides with the function in question. For the definition and basic properties of analytic multifunctions we refer to [1] or [2].

Lemma. *Let K be an analytic multifunction from a domain D of the complex plane to the complex plane. Suppose that $K(\lambda)$ is constant and countable on a subset I of strictly positive capacity (for instance, an interval). Then K is constant on D.*

Proof. (i) By [1], Theorem 7.2.8, K is countable on all D. Denote by K_0 the constant value of $K(\lambda)$ on I. Since I has strictly positive capacity, it is not countable so by [1], Theorem 7.2.13 (originally proved in [2]), we have $K_0 \subset K(\lambda)$, for every $\lambda \in D$.

(ii) By Evans's Theorem ([1], Theorem A.1.24), there exists ϕ subharmonic on all the complex plane such that $K_0 = \{z \in \mathbf{C} : \phi(z) = -\infty\}$. We set

$$\psi(\lambda) = \max_{z \in K(\lambda)} \phi(z), \quad \text{for } \lambda \in D.$$

By the definition of an analytic multifunction ([1], p. 143), the function ψ is subharmonic on D. Moreover, $\psi(\lambda) \equiv -\infty$ on I so, by H. Cartan's Theorem ([1], Theorem A.1.29), we have $\psi \equiv -\infty$ on D. This means that $K(\lambda) \subset K_0$, for all $\lambda \in D$. Consequently, $K(\lambda) = K_0$ on all D.

Remark. In the particular case studied in this paper we have $D = \mathbf{C}$, $K(\lambda)$ is the spectrum of $(\lambda A^* + (1-\lambda)B^*)(\lambda A + (1-\lambda)B)$ so, in fact, the Lemma is not used in its full strength. Part (i) can be replaced by the use of Liouville's Theorem ([1], Theorem 3.4.14), because $K(\lambda) \subset K_0$, for every $\lambda \in \mathbf{C}$, implies the constancy of the polynomial hull of $K(\lambda)$, which is $K(\lambda)$, because $K(\lambda)$ is countable.

References

[1] B. Aupetit, *A primer on spectral theory*, Springer (New York), 1991.

[2] B. Aupetit and J. Zemánek, On zeros of analytic multivalued functions, *Acta Sci. Math. (Szeged)*, **46** (1983), 311-316.

[3] H. König, *Eigenvalue distribution of compact operators*, Birkhäuser (Basel), 1986.

[4] C.A. McCarthy, c_p, *Israel J. Math.*, **5** (1967), 249-271.

[5] R.C. Thompson, The matrix numerical range, *Linear and Multilinear Algebra*, **21** (1987), 321-323.

Bernard Aupetit
Département de mathématiques et de statistique
Université Laval
Québec, Canada, G1K 7P4
e-mail: baupetit@mat.ulaval.ca

Endre Makai, jr.
Mathematical Institute
Hungarian Academy of Sciences
1364 Budapest, P.O. Box 127
Hungary
e-mail: makai@math-inst.hu

Jaroslav Zemánek
Institute of Mathematics
Polish Academy of Sciences
00-950 Warszawa, P.O. Box 137, Poland
e-mail: zemanek@impan.gov.pl