Aequationes Mathematicae 41 (1991) 70–78

Some inequalities for commutators and an application to spectral variation

Rajendra Bhatia, Chandler Davis and Fuad Kittaneh

Dedicated to the memory of Alexander M. Ostrowski on the occasion of the 100th anniversary of his birth

Summary. If $\Gamma - I$ is a positive semidefinite operator and A and B are either both Hermitian or both unitary, then every unitarily invariant norm of $A - B$ is shown to be bounded by that of $A\Gamma - \Gamma B$. Some related inequalities are proved. An application leads to a generalization of the Lidskii–Wielandt inequality to matrices similar to Hermitian.

1. Introduction

Let $\mathscr{B}(\mathscr{H})$ be the space of bounded linear operators on a complex separable Hilbert space $\mathscr{H}$. The main problem considered in this paper is that of comparing the norm of the operator $A - B$ with that of a *commutator* $A\Gamma - \Gamma B$, where A, B, Γ are elements of $\mathscr{B}(\mathscr{H})$ satisfying some additional conditions. Problems of this kind have been studied by several authors and the results obtained have been found useful in numerical analysis and physics. See, e.g., [5], [12].

Apart from the usual operator norm $\|\cdot\|$ on $\mathscr{B}(\mathscr{H})$ we are interested in *unitarily invariant* or *symmetric* norms. Properties of such norms may be found in [2] or [7]. When the space $\mathscr{H}$ is infinite-dimensional, a unitarily invariant norm $|||\cdot|||$ is defined only on a *norm ideal* associated with it. We will make no explicit mention of this ideal, it being understood that when we talk of $|||A|||$ we are assuming that A belongs to the norm ideal associated with $|||\cdot|||$.

Of special interest are the Schatten p-norms defined as

$$\|A\|_p = \left(\sum_j s_j^p(A)\right)^{1/p}, \qquad 1 \leqslant p \leqslant \infty, \tag{1}$$

where $s_j(A)$ are the *singular values* of the compact operator A arranged in

AMS (1980) subject classification: Primary 15A42, 15A57, 15A60, Secondary 65F99.

Manuscript received February 8, 1990, and in final form, May 30, 1990.

decreasing order $s_1(A) \geqslant s_2(A) \geqslant \cdots$. When $p = \infty$, the norm $\|A\|_\infty$ coincides with the operator norm $\|A\| = s_1(A)$. The norm $\|A\|_2$ is called the *Hilbert–Schmidt norm* or the *Frobenius norm*; the norm $\|A\|_1$ is called the *trace norm*.

We shall prove the following:

THEOREM 1. *Let A, B be Hermitian operators and let Γ be a positive (semi-definite) operator satisfying $\Gamma \geqslant \gamma I$ for the positive real number γ. Then, for every unitarily invariant norm $|||\cdot|||$, we have*

$$|||A\Gamma - \Gamma B||| \geqslant \gamma |||A - B|||. \tag{2}$$

THEOREM 2. *Let A, B be Hermitian operators and let Γ be a positive operator such that $\Gamma \geqslant \pm(A - B)$. Then, for every unitarily invariant norm $|||\cdot|||$, we have*

$$|||A\Gamma - \Gamma B||| \geqslant |||(A - B)^2|||. \tag{3}$$

We will also obtain some extensions of these results. One of them says that the inequality (2) is true also when A, B are unitary instead of being Hermitian. We conjecture that this is so even when A, B are arbitrary normal operators. When the norm $|||\cdot|||$ is chosen to be the Hilbert–Schmidt norm $\|\cdot\|_2$ then the inequality (2) does remain true whenever A, B are normal. This was proved by J. G. Sun [11], and our methods lead to a different proof of this result.

An important consequence of Theorem 1 is a generalization of the Lidskii–Wielandt inequality for spectral variation of Hermitian matrices, which we now recall. If A, B are Hermitian matrices, we may write them as $A = UD_1U^{-1}$, $B = VD_2V^{-1}$, where U, V are unitary and where $D_1 = \operatorname{diag}(\alpha_1, \ldots, \alpha_n)$ and $D_2 = \operatorname{diag}(\beta_1, \ldots, \beta_n)$ are diagonal matrices whose diagonal entries are the eigenvalues of A and B respectively. The Lidskii–Wielandt inequality [2, Chapter 3] says that

$$\min_\Pi |||D_1 - \Pi D_2 \Pi^{-1}||| \leqslant |||A - B|||, \tag{4}$$

for every unitarily invariant norm, where Π ranges over the class of permutation matrices.

The following theorem generalizes this result to the case where A and B are diagonalizable matrices with all real eigenvalues. In the statement, we use the familiar concept of the *condition number* of an invertible matrix P, defined by

$$\operatorname{cond}(P) = \|P\| \, \|P^{-1}\|. \tag{5}$$

THEOREM 3. *Let A, B be matrices such that $A = PD_1P^{-1}$, $B = QD_2Q^{-1}$, for some invertible matrices P, Q and real diagonal matrices D_1, D_2. Then for every unitarily invariant norm $|||\cdot|||$ we have*

$$\min_{\Pi} |||D_1 - \Pi D_2 \Pi^{-1}||| \leqslant \operatorname{cond}(P) \operatorname{cond}(Q) |||A - B|||. \tag{6}$$

Notice that when A, B are Hermitian then P, Q are unitary and hence have condition numbers 1. Thus (6) is indeed a generalization of (4).

Sun [11] has proved the inequality (6) for the special case of the Frobenius norm $\|\cdot\|_2$ but without the restriction that A and B have all real eigenvalues. Sun proves this by using the inequality analogous to (2) valid in this situation. The passage from the first result to the second is effected by him using the Hoffman–Wielandt inequality for normal matrices [8]; we achieve this by the use of (4).

If our conjecture that the inequality (2) is true when A, B are any two normal operators can be proved, then known spectral variation results for normal matrices [2], [4], [5] will yield inequalities corresponding to (6) for general diagonalizable matrices.

Proofs of the above theorems and some corollaries are given in Section 2; some related matters are discussed in Section 3.

2. Proofs and extensions

Proof of Theorem 1. Let $\Gamma \geqslant \gamma I \geqslant 0$. Then it follows from known results that, for every operator X and for all unitarily invariant norms, we have

$$2\gamma |||X||| \leqslant |||X\Gamma + \Gamma X|||. \tag{7}$$

See, e.g., [5, Theorem 3.3], or [12]. Let now A, B be Hermitian operators and let $T = A\Gamma - \Gamma B$. Then

$$T + T^* = (A - B)\Gamma + \Gamma(A - B).$$

Hence, from (7) we get

$$2\gamma |||A - B||| \leqslant |||T + T^*||| \leqslant 2|||T||| = 2|||A\Gamma - \Gamma B|||.$$

This proves Theorem 1. □

Partly for the convenience of the matrix theorist who might be primarily interested in Theorem 3 and partly because of its intrinsic interest another "*ab initio*" proof of Theorem 1 is given below.

We need two well-known theorems of Ky Fan. One, called the Fan Dominance Theorem, says that for two operators T_1 and T_2 the following two conditions (8) and (9) are equivalent:

$$|||T_1||| \leqslant |||T_2||| \qquad \text{for all unitarily invariant norms,} \tag{8}$$

$$\sum_{j=1}^{k} s_j(T_1) \leqslant \sum_{j=1}^{k} s_j(T_2) \qquad \text{for all } k = 1, 2 \dots . \tag{9}$$

The other theorem [7, p. 47], called Fan's Maximum Principle, says that

$$\sum_{j=1}^{k} s_j(T) = \max \left| \sum_{j=1}^{k} \langle Te_j, f_j \rangle \right|, \tag{10}$$

where the maximum is taken over all choices of orthonormal k-tuples $e_1, \dots, e_k$ and $f_1, \dots, f_k$ in $\mathscr{H}$.

Another Proof of Theorem 1. Assume, without loss of generality, that $\gamma = 1$. We will prove the Theorem when A, B are $n \times n$ matrices. (The same proof works for compact operators and can be extended to arbitrary operators as in [5].)

For brevity let s_j denote $s_j(A - B)$. Label the eigenvalues of $A - B$ as λ_j and eigenvectors as x_j in such a way that $s_j = |\lambda_j|$, the x_j are orthonormal, and $(A - B)x_j = \lambda_j x_j$, $j = 1, 2, \dots, n$. Define $y_j = \pm x_j$, the sign being chosen as positive if $\lambda_j \geqslant 0$ and negative if $\lambda_j < 0$. Then $(A - B)x_j = s_j y_j$ and the y_j are also orthonormal.

By Fan's Dominance Theorem and the Maximum Principle, our Theorem will be proved if we show

$$\left| \sum_{j=1}^{k} \langle (A\Gamma - \Gamma B)y_j, x_j \rangle \right| \geqslant \sum_{j=1}^{k} s_j, \tag{11}$$

for $k = 1, 2, \dots, n$.

Note first that

$$(A\Gamma - \Gamma B) - (A - B)\Gamma = B\Gamma - \Gamma B,$$

which is a skew-Hermitian matrix. Hence, for every vector x

$$\mathrm{Re}\langle (A\Gamma - \Gamma B)x, \pm x \rangle = \mathrm{Re}\langle (A - B)\Gamma x, \pm x \rangle.$$

Therefore,

$$\begin{aligned}\operatorname{Re}\sum_{j=1}^{k}\langle (A\Gamma-\Gamma B)y_j, x_j\rangle &= \operatorname{Re}\sum_{j=1}^{k}\langle (A-B)\Gamma y_j, x_j\rangle\\ &= \sum_{j=1}^{k}\langle \Gamma y_j, s_j y_j\rangle\\ &= \sum_{j=1}^{k} s_j\langle \Gamma y_j, y_j\rangle\\ &\geqslant \sum_{j=1}^{k} s_j,\end{aligned}$$

because $\Gamma \geqslant I$. This proves (11) as required. □

Proof of Theorem 2. Let X be a Hermitian operator and let $\Gamma \geqslant \pm X$. It follows from [3, Theorem 2] that

$$|||X\Gamma + \Gamma X||| \geqslant 2|||X^2|||. \tag{12}$$

The inequality (3) can be obtained from (12) just as (2) was obtained from (7) in our first proof of Theorem 1 above. □

REMARK. The two theorems of Fan we used in our second proof of Theorem 1 were also used in [3] to prove the result from which the inequality (12) above follows.

We can extend these results to non-Hermitian operators by a familiar device of going to 2×2 operator matrices.

THEOREM 4. *Let $A, B \in \mathscr{B}(\mathscr{H})$ and let Γ be a positive operator, $\Gamma \geqslant \gamma I \geqslant 0$. Then for every unitarily invariant norm on $\mathscr{B}(\mathscr{H} \oplus \mathscr{H})$ we have*

$$|||(A\Gamma - \Gamma B) \oplus (A^*\Gamma - \Gamma B^*)||| \geqslant \gamma |||(A-B) \oplus (A-B)|||. \tag{13}$$

Proof. The operators with block decompositions $\begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}$ are Hermitian. Apply Theorem 1 taking these in place of A and B and $\begin{bmatrix} \Gamma & 0 \\ 0 & \Gamma \end{bmatrix}$ in place of Γ.

COROLLARY 5. *Let A, B be unitary operators and let $\Gamma \geqslant \gamma I \geqslant 0$. Then, for every unitarily invariant norm,*

$$|||A\Gamma - \Gamma B||| \geqslant \gamma |||A - B|||. \tag{14}$$

Proof. If A, B are unitary then for every operator X we have

$$s_j(AX - XB) = s_j(X - A^*XB) = s_j(A^*X - XB^*).$$

Hence the operator $(A\Gamma - \Gamma B) \oplus (A^*\Gamma - \Gamma B^*)$ has the same singular values as those of $A\Gamma - \Gamma B$ but each with twice the multiplicity. The inequality (14) now follows from (13) and the Fan Dominance Theorem. □

COROLLARY 6 (J. G. Sun). *Let A, B be normal operators and let $\Gamma \geqslant \gamma I \geqslant 0$. Then*

$$\|A\Gamma - \Gamma B\|_2 \geqslant \gamma \|A - B\|_2. \tag{15}$$

Proof. We use the Fuglede–Putnam Theorem modulo the Hilbert–Schmidt Operators [13]. This says

$$\|AX - XB\|_2 = \|A^*X - XB^*\|_2 \tag{16}$$

for normal operators A, B and for all X. The inequality (15) now follows from (13) and (16). □

Sun's proof of (15) is much different from ours.

In the same spirit we can prove the following, in which $|X|$ stands for the positive operator $(X^*X)^{1/2}$.

THEOREM 7. *Let $A, B \in \mathscr{B}(\mathscr{H})$ and let Γ be a positive operator such that $\Gamma \geqslant |A - B|$ and $\Gamma \geqslant |A^* - B^*|$. Then*

$$|||(A\Gamma - \Gamma B) \oplus (A^*\Gamma - \Gamma B^*)||| \geqslant |||(|A - B|^2) \oplus (|A - B|^2)|||, \tag{17}$$

for every unitarily invariant norm on $\mathscr{B}(\mathscr{H} \oplus \mathscr{H})$.

Proof. This can be proved using Theorem 2 exactly in the same way as Theorem 4 was proved using Theorem 1. One must also recall that $|X|^2$ and $|X^*|^2$ have the same non-zero singular values. □

COROLLARY 8. *If A, B are normal operators and Γ is a positive operator such that $\Gamma \geqslant |A-B|$ and $\Gamma \geqslant |A^*-B^*|$ then*

$$\|A\Gamma - \Gamma B\|_2 \geqslant \| |A-B|^2 \|_2. \tag{18}$$

Proof. Use (16) and (17). □

We now turn to Theorem 3. To prove it, we need the following property of unitarily invariant norms: for any two operators X, Y we have [2, p. 29]

$$|||XY||| \leqslant |||X||| \, \|Y\|,$$

where $|||\cdot|||$ is any unitarily invariant norm and $\|\cdot\|$ is the operator bound norm. It follows that, if X and Z are invertible, then

$$|||Y||| \leqslant \|X^{-1}\| \, |||XYZ||| \, \|Z^{-1}\|. \tag{19}$$

Proof of Theorem 3. Using (19) we can write

$$\begin{aligned} |||A-B||| &= |||P(D_1P^{-1}Q - P^{-1}QD_2)Q^{-1}||| \\ &\geqslant \|P^{-1}\|^{-1} |||D_1P^{-1}Q - P^{-1}QD_2||| \, \|Q\|^{-1}. \end{aligned} \tag{20}$$

Now let $s_1 \geqslant \cdots \geqslant s_n$ denote the singular values of $P^{-1}Q$. We can find unitary matrices U and V such that $P^{-1}Q = USV$, where S is the diagonal matrix with diagonal entries $s_1, \ldots, s_n$. Note that $S \geqslant s_n I$ and $s_n > 0$. Using unitary invariance, Theorem 1, and the inequality (4), we find

$$\begin{aligned} |||D_1P^{-1}Q - P^{-1}QD_2||| &= |||D_1USV - USVD_2||| \\ &= |||U^*D_1US - SVD_2V^*||| \\ &\geqslant s_n |||U^*D_1U - VD_2V^*||| \\ &\geqslant s_n \min_{\Pi} |||D_1 - \Pi D_2 \Pi^{-1}|||. \end{aligned} \tag{21}$$

Note also that

$$s_n = \|S^{-1}\|^{-1} = \|Q^{-1}P\|^{-1} \geqslant \|Q^{-1}\|^{-1} \|P\|^{-1}. \tag{22}$$

Combining the inequalities (20), (21) and (22) we get the result (6). □

3. Remarks

(1) It is tempting to attempt a generalization of Theorem 1 which would say that for $\Gamma_1 \geqslant \Gamma_2 \geqslant 0$ and Hermitian A, B we have $|||A\Gamma_1 - \Gamma_1 B||| \geqslant |||A\Gamma_2 - \Gamma_2 B|||$. This is refuted by the example

$$\Gamma_1 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \quad \Gamma_2 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad A = B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

(2) The Lidskii–Wielandt inequality (4) clearly gives a tight bound. The bound (6) is less so when A, B are not Hermitian. Just look at the seemingly favourable case when A and B commute; still the left-hand side of (6) is generally smaller than $|||A - B|||$, and this is only aggravated by introducing the condition number coefficients.

(3) In Section 1 we conjectured that Theorem 1 can be generalized to the situation when A and B are normal operators. Corollary 6 says this does hold for the Hilbert–Schmidt norm. For other Schatten p-norms we can get a weaker result. In a recent paper Abdessemed and Davies [1] have proved that there exist constants c_p, depending on p alone, such that whenever A, B are normal and X is any operator we have

$$\|AX - XB\|_p \geqslant c_p \|A^*X - XB^*\|_p,$$

for $2 \leqslant p < \infty$; and when $\dim \mathscr{H} < \infty$ also for $1 < p < 2$. Using this result we can conclude from Theorem 4 that there exist constants k_p, depending on p alone, such that, whenever A, B are normal and $\Gamma \geqslant \gamma I \geqslant 0$, then

$$\|A\Gamma - \Gamma B\|_p \geqslant k_p \gamma \|A - B\|_p,$$

for $2 \leqslant p < \infty$; and, when $\dim \mathscr{H} < \infty$, also for $1 < p < 2$.

(4) We take this opportunity to make two remarks on the results of Abdessemed and Davies [1]. First, following the arguments used in [9] the inequality of [1] quoted in Remark 3 above can be shown to be valid also when the operators A and B^* are subnormal. Second, in [10] and [6] inequalities relating the p-norm of a block operator matrix to the norms of the blocks have been obtained by two of the present authors. These results can be used to improve some of the estimates in [1].

Acknowledgement

R. Bhatia thanks the University of Toronto and NSERC (Canada) for support in the Summer of 1989 when some of this work was done.

REFERENCES

[1] ABDESSEMED, A. and DAVIES, E. B., *Some commutator estimates in the Schatten classes*. J. London Math. Soc. *39* (1989), 299–308.

[2] BHATIA, R., *Perturbation bounds for matrix eigenvalues*. Longman, Essex and Wiley, New York, 1987.

[3] BHATIA, R., *Some inequalities for norm ideals*. Commun. Math. Phys. *111* (1987), 33–39.

[4] BHATIA, R., DAVIS, CH. and KOOSIS, P., *An extremal problem in Fourier analysis with applications to operator theory*. J. Funct. Anal. *82* (1989), 138–150.

[5] BHATIA, R., DAVIS, CH. and MCINTOSH, A., *Perturbation of spectral subspaces and solution of linear operator equations*. Linear Algebra Appl. *52–53* (1983), 45–67.

[6] BHATIA, R. and KITTANEH, F., *Norm inequalities for partitioned operators and an application*. preprint.

[7] GOHBERG, I. C. and KREIN, M. G., *Introduction to the theory of linear nonselfadjoint operators*. American Mathematical Society, Providence, RI, 1969.

[8] HOFFMAN, A. J. and WIELANDT, H. W., *The variation of the spectrum of a normal matrix*. Duke Math. J., *20* (1953), 37–39.

[9] KITTANEH, F., *On generalized Fuglede–Putnam theorems of Hilbert–Schmidt type*. Proc. Amer. Math. Soc. *88* (1983), 293–298.

[10] KITTANEH, F., *On the continuity of the absolute value map in the Schatten classes*. Linear Algebra Appl. *118* (1989), 61–68.

[11] SUN, J. G., *On the perturbation of the eigenvalues of a normal matrix* (Chinese). Math. Numer. Sinica *6* (1984), 334–336.

[12] VAN HEMMEN, J. L. and ANDO, T., *An inequality for trace ideals*. Commun. Math. Phys. *76* (1980), 143–148.

[13] WEISS, G., *The Fuglede commutativity theorem modulo the Hilbert–Schmidt class and generating functions for matrix operators, I, II*. Trans. Amer. Math. Soc. *246* (1978), 193–209, J. Operator Theory *5* (1981), 3–16.

Indian Statistical Institute,
New Delhi 110 016,
India.

Department of Mathematics,
University of Toronto,
Toronto, Ontario,
M5S 1A1,
Canada.

Department of Mathematics,
Kuwait University,
Kuwait.