Math. Ann. 281, 7–12 (1988)

Mathematische Annalen

On the Clarkson-McCarthy Inequalities

Rajendra Bhatia[1] and John A. R. Holbrook[2,*]

[1] Indian Statistical Institute, New Delhi 110016, India
[2] University of Guelph, Guelph, Ontario, NIG 2W1, Canada

1. Introduction

C. A. McCarthy proved in [7], among several other results, the following inequalities for Schatten p-norms of Hilbert space operators:

$$2(\|A\|_p^p+\|B\|_p^p)\leqq\|A+B\|_p^p+\|A-B\|_p^p\leqq 2^{p-1}(\|A\|_p^p+\|B\|_p^p) \tag{1}$$

for $2\leqq p<\infty$, and

$$2^{p-1}(\|A\|_p^p+\|B\|_p^p)\leqq\|A+B\|_p^p+\|A-B\|_p^p\leqq 2(\|A\|_p^p+\|B\|_p^p) \tag{2}$$

for $1\leqq p\leqq 2$.

These are non-commutative analogues of some inequalities of Clarkson for the classical Banach spaces and constitute one half of the "Clarkson-McCarthy Inequalities." These estimates have been found to be very powerful tools in operator theory and in mathematical physics. (See, e.g., Simon [11].)

Here we formulate and prove a more general version of these inequalities. Our analysis extends these inequalities to a wider class of norms which includes the p-norms and at the same time leads to a proof which is much simpler than McCarthy's original proof or some later proofs. Indeed, it appears to be simpler and more elementary than any other proof of which we are aware; see the discussion in [11].

Let $\mathscr{B}(\mathscr{H})$ denote the space of all bounded linear operators on a Hilbert space $\mathscr{H}$. For convenience, we take $\mathscr{H}$ to be infinite-dimensional. If an operator A is compact, we enumerate the eigenvalues of the positive operator $(A^*A)^{1/2}$ as $s_1(A)\geqq s_2(A)\geqq\dots$. These are called the *singular values* of A. An operator A is said to belong to the class $\mathscr{I}_p$ if $\sum_{j=1}^{\infty}(s_j(A))^p<\infty$, where p is a real number, $1\leqq p<\infty$. If $A\in\mathscr{I}_p$ then the Schatten p-norm of A is the number $\|A\|_p=(\sum(s_j(A))^p)^{1/p}$. It is well known that $\mathscr{I}_p$ is an ideal in $\mathscr{B}(\mathscr{H})$, that $\|A\|_p$ defines a norm on it, and that it is

* Supported in part by Indian Statistical Institute, New Delhi, and NSERC of Canada (under operating grant A 8745)

complete with respect to this norm. See Gohberg and Krein [5], Schatten [10], or [11].

These norms are special examples of *symmetric norms* or *unitarily invariant norms* each of which arises as a "*symmetric gauge function*" of the singular values. (See [10] for definitions.) Each such norm $|||\cdot|||$ is defined on a natural subclass $\mathcal{I}_{|||\cdot|||}$ of $\mathcal{B}(\mathcal{H})$ called the *norm ideal* associated with the norm $|||\cdot|||$ and satisfies the invariance property $|||UAV||| = |||A|||$ for all A in this ideal and for all unitary operators U, V. The usual operator norm $\|\cdot\|$ is also such a norm defined on all of $\mathcal{B}(\mathcal{H})$ and, for compact A, $\|A\| = s_1(A)$. It is hence conventional to denote $\|A\|$ by $\|A\|_\infty$.

Let $2 \leqq p \leqq \infty$ and let $r = p/2$. Note that $\|A\|_p^2 = \|A^*A\|_r$. This is a special instance of a more general phenomenon. We say that a (unitarily invariant) norm $|||\cdot|||$ is a *Q-norm* if there exists some other unitarily invariant norm $|||\cdot|||'$ such that $|||A|||^2 = |||A^*A|||'$. See Bhatia [2] for more examples of such norms and for an approximation theorem involving them.

We also recall that each symmetric gauge function has an "*associate*" [10] or a "*conjugate*" [5] symmetric gauge function and through this duality each unitarily invariant norm has a conjugate norm associated with it. The norm $\|\cdot\|_p$ is conjugate to $\|\cdot\|_q$ if $\frac{1}{p} + \frac{1}{q} = 1$. We will say that a unitarily invariant norm is a *Q^*-norm* if it is conjugate to a Q-norm. The class of such norms includes the p-norms for $1 \leqq p \leqq 2$.

The following questions are thus natural. If an inequality $\|A\|_p \leqq c\|B\|_p$ holds for $1 \leqq p \leqq \infty$ (with the same constant c) then does it hold for all unitarily invariant norms? If such an inequality holds for $2 \leqq p \leqq \infty$ then does it hold for all Q-norms and if it holds for $1 \leqq p \leqq 2$ then does it hold for all Q^*-norms? There are several well-known results in operator theory (see, e.g. Marshall and Olkin [6]) where the first question has a positive answer. See [2] for a recent result in which the second question has a positive answer.

We will obtain extensions of (1) and (2) to Q-norms and Q^*-norms respectively. However, to do this we need to recast them in a form such that the constants occurring in them become independent of p. We will see that such a recasting also leads to a better understanding of the original inequalities.

2. Main Results

Let $\mathbb{R}_+^\infty$ be the space of all sequences of positive real numbers. Given two elements $\{x_j\}$ and $\{y_j\}$ of this space define another element by setting $\{x_j\} \vee \{y_j\} = \{x_1, y_1, x_2, y_2, \ldots\}$. Let $|||\cdot|||$ be any unitarily invariant norm on $\mathcal{B}(\mathcal{H})$ and let Φ be the associated symmetric gauge function on $\mathbb{R}_+^\infty$, i.e., $|||A||| = \Phi(\{s_j(A)\})$.

Given two operators A and B we define

$$|||A \oplus B||| = \Phi(\{s_j(A)\} \vee \{s_j(B)\}).$$

This quantity is simply the $|||\cdot|||$-norm of $A \oplus B$ regarded as the operator $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$ in $\mathcal{B}(\mathcal{H} \oplus \mathcal{H})$. Note that

$$\|A \oplus B\| = \max(\|A\|, \|B\|),$$

$$\|A \oplus B\|_p = (\|A\|_p^p + \|B\|_p^p)^{1/p} \quad \text{for} \quad 1 \leqq p < \infty,$$

and in particular

$$\|A\oplus A\|_p = 2^{1/p}\|A\|_p \quad \text{for} \quad 1\leqq p<\infty.$$

Extensions to direct sums involving more than two operators are obtained in the same way. If the operator ideal $\mathscr{I}_\Phi$ is normed by the symmetric gauge function Φ then so is $\mathscr{I}_\Phi\oplus\mathscr{I}_\Phi$ by the above procedure.

The proof of (1) and (2) in [7] goes via the following inequalities, which are of independent interest. If A, B are positive operators in $\mathscr{I}_p$ for any $p\geqq 1$, then

$$2^{1-p}\|A+B\|_p^p \leqq \|A\|_p^p + \|B\|_p^p \leqq \|A+B\|_p^p. \tag{3}$$

Note that in the notations defined above this can be rewritten as

$$\tfrac{1}{2}\|(A+B)\oplus(A+B)\|_p \leqq \|A\oplus B\|_p \leqq \|(A+B)\oplus 0\|_p. \tag{4}$$

Thus the following theorem (its history is outlined in the next section) includes a generalisation of (3).

Theorem 1. *Let A, B be any two positive operators belonging to the norm ideal associated with a unitarily invariant norm $|||\cdot|||$. Then*

$$\tfrac{1}{2}|||(A+B)\oplus(A+B)||| \leqq |||A\oplus B||| \leqq |||(A+B)\oplus 0|||. \tag{5}$$

To recast (1) and (2) in a similar mould we need to go to quadruplets instead of pairs. Thus, for example, the second inequality in (1) can be rewritten first as

$$2^{1/p}(\|A+B\|_p^p + \|A-B\|_p^p)^{1/p} \leqq 2(\|A\|_p^p + \|B\|_p^p)^{1/p}$$

and then as

$$\|(A+B)\oplus(A+B)\oplus(A-B)\oplus(A-B)\|_p \leqq 2\|A\oplus 0\oplus B\oplus 0\|_p$$

for $2\leqq p<\infty$.

Note that the first inequality in (1) can be obtained from the second one by replacing A and B by $A+B$ and $A-B$ respectively, and vice versa. Similar considerations apply to the pair of inequalities in (2). The following two theorems, then, are the promised generalisations of (1) and (2).

Theorem 2. *Let A and B be two operators belonging to the ideal $\mathscr{I}_Q$ associated with a Q-norm $\|\cdot\|_Q$. Then*

$$\|(A+B)\oplus(A+B)\oplus(A-B)\oplus(A-B)\|_Q \leqq 2\|A\oplus 0\oplus B\oplus 0\|_Q. \tag{6}$$

Theorem 3. *Let A and B be two operators belonging to the ideal $\mathscr{I}_{Q^*}$ associated with a Q^*-norm $\|\cdot\|_{Q^*}$. Then*

$$2\|A\oplus 0\oplus B\oplus 0\|_{Q^*} \leqq \|(A+B)\oplus(A+B)\oplus(A-B)\oplus(A-B)\|_{Q^*}. \tag{7}$$

3. Proofs of the Results

In finite dimensions Theorem 1 is a restatement of a majorisation result due to Thompson [12]. One proof of Thompson's result given in Ando [1] goes through, without any change, to infinite dimensions. For the convenience of the reader we reproduce this short and elegant proof.

To prove the first inequality in (5) write $(A+B)\oplus(A+B)=(A\oplus B)+(B\oplus A)$ and note that the two terms on the right-hand side have the same norm. To prove the second, write

$$\begin{pmatrix} A+B & 0 \\ 0 & 0 \end{pmatrix} = XX^*, \quad \text{where} \quad X=\begin{pmatrix} A^{1/2} & B^{1/2} \\ 0 & 0 \end{pmatrix}.$$

Then note

$$X^*X=\begin{pmatrix} A & A^{1/2}B^{1/2} \\ B^{1/2}A^{1/2} & B \end{pmatrix}.$$

By general properties of unitarily invariant norms (see, e.g., [5]), $|||XX^*|||=|||X^*X|||$ and $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, being a "pinching" of X^*X, has smaller norm than X^*X. This proves the second inequality in (5).

Recently, in [3], we have begun a study of "weakly unitarily invariant" norms. These are norms on spaces of finite-dimensional operators that are invariant under unitary conjugations $A\to U^*AU$. The pinching inequality extends readily to these norms and, in a finite-dimensional setting, X^*X and XX^* are unitarily conjugate. Thus the proof above shows that the inequalities (5) are valid for this extended class of norms. Such a norm τ' gives rise to a unitarily invariant norm τ by the same procedure as we have used to define Q-norms: $\tau(A)=(\tau'(A^*A))^{1/2}$ (provided this τ satisfies the triangle inequality). On this basis analogues of Theorems 2 and 3 may be formulated and proved in the new setting.

We now turn to the proofs of Theorems 2 and 3. Since $A^*A\oplus 0$ and $B^*B\oplus 0$ are positive operators, the first inequality of (5) shows that, for any unitarily invariant norm $|||\cdot|||$,

$$2|||(A^*A+B^*B)\oplus 0\oplus(A^*A+B^*B)\oplus 0|||\leqq 4|||A^*A\oplus 0\oplus B^*B\oplus 0|||. \qquad (8)$$

By unitary invariance and the relation $2(A^*A+B^*B)=C^+ + C^-$ where $C^+=(A+B)^*(A+B)$, $C^-=(A-B)^*(A-B)$, the left side of (8) is $|||((C^+\oplus C^+)+(C^-\oplus C^-))\oplus(0\oplus 0)|||$. By the second inequality of (5) this is not less than $|||C^+\oplus C^+\oplus C^-\oplus C^-|||$. Thus we have

$$|||C^+\oplus C^+\oplus C^-\oplus C^-|||\leqq|||4A^*A\oplus 0\oplus 4B^*B\oplus 0|||,$$

for every unitarily invariant norm. Hence, the inequality (6) is true for all Q-norms.

We shall obtain (7) from (6), by duality. It is a central result of the Schatten theory (see [10]) that $\mathscr{I}_{Q^*}$ is the Banach space dual of $\mathscr{I}_Q$ under the bilinear pairing $\langle T,S\rangle=\operatorname{tr} TS$. We apply this to operators in $\mathscr{B}=\mathscr{B}\left(\bigoplus_1^4\mathscr{H}\right)$. For $T\in\mathscr{B}$ let $\Lambda(T)=\frac{T_1+T_3}{2}\oplus\frac{T_1+T_3}{2}\oplus\frac{T_1-T_3}{2}\oplus\frac{T_1-T_3}{2}$, where the T_k are the diagonal blocks in the 4×4 operator block matrix corresponding to T. Clearly Λ is a linear map on $\mathscr{B}$ and we claim that it is contractive with respect to $\|\cdot\|_Q$. To see this note that the pinching inequality ensures that $\|T\|_Q\geqq\|T_1\oplus T_2\oplus T_3\oplus T_4\|_Q$, while this is the same as $\|T_1\oplus -T_2\oplus T_3\oplus -T_4\|_Q$ by unitary invariance. Hence each is no less than $\|T_1\oplus 0\oplus T_3\oplus 0\|_Q$ which dominates $\|\Lambda(T)\|_Q$ by (6). On general grounds, then, the adjoint Λ^* is also contractive (with respect to $\|\cdot\|_{Q^*}$).

Now we claim that $\Lambda^*((A+B)\oplus(A+B)\oplus(A-B)\oplus(A-B))=2A\oplus 0\oplus 2B\oplus 0$. To see this we must check that for all $T\in\mathscr{I}_Q$

$$\operatorname{tr} T(2A\oplus 0\oplus 2B\oplus 0)=\operatorname{tr}\Lambda(T)((A+B)\oplus(A+B)\oplus(A-B)\oplus(A-B)),$$

that is

$$\begin{aligned}&\operatorname{tr}(2T_1A\oplus 0\oplus 2T_3B\oplus 0)\\ &=\operatorname{tr}\left(\left(\frac{T_1+T_3}{2}\right)(A+B)\oplus\left(\frac{T_1+T_3}{2}\right)(A+B)\right.\\ &\left.\oplus\left(\frac{T_1-T_3}{2}\right)(A-B)\oplus\left(\frac{T_1-T_3}{2}\right)(A-B)\right).\end{aligned}$$

Bearing in mind that $\operatorname{tr}(\oplus X_k)=\sum\operatorname{tr}X_k$, verification of this is routine. Since Λ^* is contractive, the inequality (7) follows. We have proved all the theorems stated in Sect. 2.

We recall that the inequalities

$$2(\|A\|_p^p+\|B\|_p^p)^{q/p}\leqq\|A+B\|_p^q+\|A-B\|_p^q \tag{9}$$

$\left(\text{for } 2\leqq p<\infty;\ \frac{1}{p}+\frac{1}{q}=1\right)$, and

$$\|A+B\|_p^q+\|A-B\|_p^q\leqq 2(\|A\|_p^p+\|B\|_p^p)^{q/p} \tag{10}$$

(for $1<p\leqq 2$) complement (1) and (2) to form the complete set of "Clarkson-McCarthy inequalities." We remark that, while (1) and (2) are commonly proved separately, they follow from (9) and (10) simply by the convexity properties of the power functions. Thus (1) is a consequence of (9) and the convexity of $t\to t^{p/q}$. It would therefore be doubly worthwhile to find a more direct proof (perhaps along the lines of our treatment of (1) and (2)) for the inequalities (9) and (10).

4. On an Inequality of Phillips

In [8] Phillips proved the following theorem, which is related to material in the preceding sections.

Theorem 4 (J. Phillips). *Let* $A\geqq B\geqq 0$ *and* $t\geqq 1$. *Then*

$$\|A^{1/t}-B^{1/t}\|_t^t\leqq\|A-B\|_1\,. \tag{11}$$

When $t=2$, this is a special case of the Powers-Størmer inequality [9] which is valid for any two positive operators A and B. Phillips gave an intricate proof of (11) and noted that a simpler proof would be possible if one had the inequality $\operatorname{tr}(A^t+B^t)\leqq\operatorname{tr}(A+B)^t$ for all positive operators A, B and real numbers $t\geqq 1$. But this is a fact proved by McCarthy in [7]. Indeed, the inequalities (3) are equivalent to the inequalities

$$2^{1-p}\operatorname{tr}(A+B)^p\leqq\operatorname{tr}A^p+\operatorname{tr}B^p\leqq\operatorname{tr}(A+B)^p \tag{12}$$

for all positive operators A, B and for all $p\geqq 1$.

Let us now indicate a short proof of (11) following Phillips. The map $A \to A^r$ is operator monotone for $0 \leqq r \leqq 1$ on the class of positive operators (see Donoghue [4]). So, the condition $A \geqq B \geqq 0$ implies $A^{1/t} \geqq B^{1/t} \geqq 0$ for all $t \geqq 1$. Using (12) write

$$\operatorname{tr} A = \operatorname{tr}(A^{1/t} - B^{1/t} + B^{1/t})^t \geqq \operatorname{tr}(A^{1/t} - B^{1/t})^t + \operatorname{tr} B .$$

I.e., $\operatorname{tr}(A-B) \geqq \operatorname{tr}(A^{1/t} - B^{1/t})^t$, which is the same as the inequality (11).

References

1. Ando, T.: Majorisation, doubly stochastic matrices and comparison of eigenvalues. Lecture Notes, Sapporo, 1982. Linear Algebra Appl. (to appear)
2. Bhatia, R.: Some inequalities for norm ideals. Commun. Math. Phys. 111, 33–39 (1987)
3. Bhatia, R., Holbrook, J.A.R.: Unitary invariance and spectral variation. Linear Algebra Appl. 95, 43–68 (1987)
4. Donoghue, W.: Monotone matrix functions and analytic continuation. Berlin, Heidelberg, New York: Springer 1974
5. Gohberg, I.C., Krein, M.G.: Introduction to the theory of linear nonselfadjoint operators. Am. Math. Soc. Translations of Mathematical Monographs, Vol. 18 (1969)
6. Marshall, A.W., Olkin, I.: Inequalities: Theory of majorisation and its applications. New York: Academic Press 1979
7. McCarthy, C.A.: c_p. Isr. J. Math. 5, 249–271 (1967)
8. Phillips, J.: Generalised Powers-Størmer inequalities. Talk at the Canadian Operator Theory Conference, Victoria, July 1986 (unpublished)
9. Powers, R.T., Stormer, E.: Free states of the canonical anticommutation relations. Commun. Math. Phys. 16, 1–33 (1970)
10. Schatten, R.: Norm ideals of completely continuous operators (Sec. Ed.). Berlin, Heidelberg, New York: Springer 1970
11. Simon, B.: Trace ideals and their applications. Cambridge University Press, 1979
12. Thompson, R.C.: Convex and concave functions of singular values of matrix sums. Pac. J. Math. 62, 285–290 (1976)

Received February 24, 1987