
Operator-valued Extensions of Matrix-norm Inequalities

G.J.O. Jameson

Abstract. We describe a rather striking extension of a wide class of inequalities. Some famous classical inequalities, such as those of Hardy and Hilbert, equate to the evaluation of the norm of a matrix operator. Such inequalities can be presented in two versions, linear and bilinear. We show that in all such inequalities, the scalars can be replaced by operators on a Hilbert space, with the conclusions taking the form of an operator inequality in the usual sense. With careful formulation, a similar extension applies to the Cauchy–Schwarz inequality.

1. INTRODUCTION. The English mathematician G.H. Hardy (1877–1947) established a number of significant inequalities, each known, in its own context, simply as “Hardy’s inequality”. One of them states the following: *if x_1, x_2, \dots are numbers (real or complex), and such that $\sum_{k=1}^{\infty} |x_k|^2$ is convergent, and*

$$y_j = \frac{1}{j}(x_1 + x_2 + \cdots + x_j) \quad (j \geq 1),$$

then

$$\sum_{j=1}^{\infty} |y_j|^2 \leq 4 \sum_{k=1}^{\infty} |x_k|^2.$$

(and 4 is the best constant). Here $y = Ax$, where x, y are infinite sequences and A is the “Cesaro matrix” given by $a_{j,k} = \frac{1}{j}$ for $k \leq j$ and 0 for $k > j$. Though Hardy did not state it this way, his inequality amounts to the evaluation of the norm of this matrix as a linear operator.

In general, let $A = (a_{j,k})$ be any matrix (finite or infinite) of real or complex numbers. Let $\|A\|$ denote the usual operator norm of A as an operator on ℓ_2^n or ℓ_2 (for readers needing it, this notation is reviewed below). For vectors x, y , we then have the *linear* inequality $\|Ax\| \leq \|A\| \|x\|$ and the *bilinear* inequality $|\langle Ax, y \rangle| \leq \|A\| \|x\| \|y\|$. Written out explicitly, these inequalities say:

- (N1) if x_1, x_2, \dots are (real or complex) scalars and $y_j = \sum_k a_{j,k} x_k$ for each j , then

$$\sum_j |y_j|^2 \leq \|A\|^2 \sum_k |x_k|^2;$$

- (N2) for scalars x_j, y_k ,

$$\left| \sum_j \sum_k a_{j,k} x_k \bar{y}_j \right| \leq \|A\| \left(\sum_j |x_j|^2 \right)^{1/2} \left(\sum_k |y_k|^2 \right)^{1/2}.$$

So, for example, for the Cesaro matrix A , Hardy's inequality states that $\|A\| = 2$. There are other classical inequalities that express the norm of a matrix, for example Hilbert's inequality (to which we return later).

Our topic here is the rather striking fact that in all such inequalities, the scalars x_j, y_k can be replaced by operators on a Hilbert space H , with the conclusions taking the form of operator inequalities in the usual sense: $A \geq 0$, for self-adjoint A , means $\langle Ax, x \rangle \geq 0$ for $x \in H$. In the case of (N1), this is achieved by simply replacing $|x_k|^2$ by $X_k^* X_k$ and $|y_j|^2$ by $Y_j^* Y_j$. So, for example, the operator-valued version of Hardy's inequality takes the form: $\sum_{j=1}^{\infty} Y_j^* Y_j \leq 4 \sum_{k=1}^{\infty} X_k^* X_k$.

The operator-valued extension of the bilinear inequality (N2) requires a slightly more subtle reformulation. We will show how to derive it from the linear version with the help of an operator-valued version of the Cauchy–Schwarz inequality.

We will confine ourselves to proving the statements for finite matrices and sequences. The extensions to the infinite case are essentially routine and will present no difficulty to experts.

Operator-valued versions of scalar inequalities have been a growth industry in recent years. See, for example, the articles [6, 8] and the book [1]. Some of these results are quite deep. The theorems we present here are at the easy end of the scale: they will be obtained by simple and (arguably) elegant methods well within the scope of a basic course on Hilbert spaces. Despite this, they have not, to the author's knowledge, been explicitly stated elsewhere. Some results on the *norms* of bilinear combinations of operators are given in [9] and [10], but not the underlying operator inequalities. Most of their results can be derived as consequences of ours.

2. REVIEW OF NOTATION AND RESULTS ASSUMED. We review very briefly the ideas and results required; details can be seen in any introductory text on the subject, such as [4, 11].

Throughout the sequel, in any Hilbert space considered, the inner product will be denoted by $\langle \cdot, \cdot \rangle$. The derived norm is defined by $\|x\| = \langle x, x \rangle^{1/2}$. The *Cauchy–Schwarz inequality* states $|\langle x, y \rangle| \leq \|x\| \|y\|$. The space ℓ_2^n is \mathbb{C}^n (or \mathbb{R}^n) with inner product defined by $\langle x, y \rangle = \sum_{j=1}^n x_j \bar{y}_j$, and ℓ_2 is the obvious extension to infinite sequences with $\sum_{j=1}^{\infty} |x_j|^2$ convergent.

Given a Hilbert space H , the space H^n is the space consisting of the n -tuples $x = (x_1, x_2, \dots, x_n)$ with $x_j \in H$, with inner product defined by $\langle x, y \rangle = \sum_{j=1}^n \langle x_j, y_j \rangle$, hence norm defined by $\|x\|^2 = \sum_{j=1}^n \|x_j\|^2$.

A linear operator A between Hilbert spaces H, K is *bounded* if there is a constant M such that $\|Ax\| \leq M\|x\|$ for all x in H , and its *norm* $\|A\|$ is the least such M . With the Cauchy–Schwarz inequality, this implies that $|\langle Ax, y \rangle| \leq \|A\| \|x\| \|y\|$ for all $x \in H$ and $y \in K$. The adjoint operator A^* (from K to H) is defined by $\langle Ax, y \rangle = \langle x, A^*y \rangle$. It satisfies $A^{**} = A$, $(AB)^* = B^*A^*$ and $\|A^*\| = \|A\|$. If A is given by the matrix $(a_{j,k})$, then A^* is given by $(\bar{a}_{k,j})$. The space of bounded linear operators from H to K will be denoted by $L(H, K)$, and by $L(H)$ when $H = K$.

For a self-adjoint operator, $\langle Ax, x \rangle$ is real for all x , and we write $A \geq 0$ if $\langle Ax, x \rangle \geq 0$ for all x (such operators are “positive semi-definite”; we will just call them “positive”). We write $A \geq B$ if $A - B \geq 0$. Since $\langle A^*Ax, x \rangle = \|Ax\|^2$, we have $A^*A \geq 0$ for any A , and $\|A^*A\| = \|A\|^2$. Also, $A^*A \leq B^*B$ is equivalent to $\|Ax\| \leq \|Bx\|$ for all $x \in H$. If $0 \leq A \leq B$, then $\|A\| \leq \|B\|$.

Finally, a positive operator A has a unique positive square root $A^{1/2}$, and $0 \leq A \leq B$ implies $A^{1/2} \leq B^{1/2}$. (However, $0 \leq A \leq B$ does not imply $A^2 \leq B^2$ unless $AB = BA$.) For any operator A , the *right* and *left* moduli are defined by $|A|_R =$

$(A^*A)^{1/2}$ and $|A|_L = (AA^*)^{1/2}$, so that $|A|_L = |A^*|_R$, and the two coincide for self-adjoint A .

3. THE CAUCHY–SCHWARZ INEQUALITY FOR OPERATORS. Before proceeding to the generalizations of (N1) and (N2), we describe an operator version of the Cauchy–Schwarz inequality. It will be needed for the extension of (N2). It is also of interest in itself, and has found many other applications.

A simple-minded analogy with the scalar case might suggest something like the following: given operators X_j, Y_j , let $S = \sum_{j=1}^n X_j^* Y_j$. Then

$$S^* S \leq \left(\sum_{j=1}^n X_j^* X_j \right) \left(\sum_{j=1}^n Y_j^* Y_j \right).$$

However, this is a nonstarter, because the operator on the right-hand side need not even be self-adjoint: in fact, if A and B are self-adjoint, then AB is self-adjoint if and only if $AB = BA$.

To point the way to a correct statement, consider the case $n = 1$. For $u \in H$, we have

$$\langle Y^* X X^* Y u, u \rangle = \langle X X^* Y u, Y u \rangle \leq \|X X^*\| \|Y u\|^2 = \|X\|^2 \langle Y^* Y u, u \rangle,$$

so

$$(Y^* X)(X^* Y) \leq \|X\|^2 Y^* Y. \tag{1}$$

We mention that X and Y cannot be interchanged on the right-hand side, and no liberties can be taken with the order of the operators. However, it can be rewritten, for example, as $(Y^* X^*)(X Y) \leq \|X\|^2 Y^* Y$. Further, we can take square roots in this to deduce $|XY|_R \leq \|X\| |Y|_R$.

We are now ready for the result in question. Various proofs have been given. Here we present a pleasantly simple one, following a method used in [8].

Theorem 1. *Let X_j, Y_j ($1 \leq j \leq n$) be operators on a Hilbert space H . Let $S = \sum_{j=1}^n X_j^* Y_j$. Then*

$$S^* S \leq \left\| \sum_{j=1}^n X_j^* X_j \right\| \left\| \sum_{j=1}^n Y_j^* Y_j \right\|.$$

Proof. Define $X : H \rightarrow H^n$ by $Xu = (X_1 u, X_2 u, \dots, X_n u)$ (and Y similarly from Y_j). Then it is easily checked that $X^* : H^n \rightarrow H$ is given by: $X^*(u_1, u_2, \dots, u_n) = \sum_{j=1}^n X_j^* u_j$. Hence $X^* X = \sum_{j=1}^n X_j^* X_j$ and $X^* Y = S$. The result now follows from (1). ■

As mentioned already, we cannot tamper with the order of the terms. However, the statement can, of course, be rewritten as follows: if $T = \sum_{j=1}^n X_j Y_j$, then $T^* T \leq \left\| \sum_{j=1}^n X_j X_j^* \right\| \left(\sum_{j=1}^n Y_j^* Y_j \right)$. Also, we can deduce the following norm inequality (which was stated, for example, in [7]), in which the asymmetry between X_j and Y_j is suppressed:

$$\left\| \sum_{j=1}^n X_j^* Y_j \right\|^2 \leq \left\| \sum_{j=1}^n X_j^* X_j \right\| \left\| \sum_{j=1}^n Y_j^* Y_j \right\|.$$

Theorem 1 is nontrivial even when the X_j are replaced by scalars λ_j (more exactly, by $\overline{\lambda_j}I$). After also writing X_j for Y_j and taking square roots, this delivers a version looking very much like the original Cauchy–Schwarz inequality for scalars:

$$\left| \sum_{j=1}^n \lambda_j X_j \right|_R \leq \left(\sum_{j=1}^n |\lambda_j|^2 \right)^{1/2} \left(\sum_{j=1}^n X_j^* X_j \right)^{1/2}.$$

4. VECTOR-VALUED VERSIONS OF (N1) AND (N2) We return to (N1) and (N2). Before giving the promised operator-valued extensions, we formulate vector-valued versions.

Theorem 2. *Let $A = (a_{j,k})$ be an $m \times n$ matrix, and let H be a Hilbert space.*

(i) *If $x_1, \dots, x_n \in H$ and $z_j = \sum_{k=1}^n a_{j,k} x_k$ for $1 \leq j \leq m$, then*

$$\sum_{j=1}^m \|z_j\|^2 \leq \|A\|^2 \sum_{k=1}^n \|x_k\|^2. \quad (2)$$

Equivalently, if A_H is the operator from $H^n \rightarrow H^m$ defined by $A_H(x_1, \dots, x_n) = (z_1, \dots, z_m)$, with z_j as above, then $\|A_H\| \leq \|A\|$.

(ii) *For elements x_k, y_j of H , we have*

$$\left| \sum_{j=1}^m \sum_{k=1}^n a_{j,k} \langle x_k, y_j \rangle \right| \leq \|A\| \left(\sum_{k=1}^n \|x_k\|^2 \right)^{1/2} \left(\sum_{j=1}^m \|y_j\|^2 \right)^{1/2}. \quad (3)$$

Proof. For (i), choose an orthonormal basis e_1, e_2, \dots, e_R of H (or of the subspace generated by the elements x_k), and write $x_k = \sum_{r=1}^R x_k(r) e_r$. Then for each r , we have $z_j(r) = \sum_{k=1}^n a_{j,k} x_k(r)$, so by (N1),

$$\sum_{j=1}^m |z_j(r)|^2 \leq \|A\|^2 \sum_{k=1}^n |x_k(r)|^2.$$

Summation over r gives (2). This clearly implies $\|A_H\| \leq \|A\|$. (Needless to say, equality actually holds here, but this is not important for our purposes.)

Statement (ii) now follows at once, because the sum in (3) is $\langle A_H x, y \rangle$, where $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_m)$. ■

Curiously, quite lengthy proofs of Theorem 2 have appeared in the literature (I will refrain from listing them explicitly).

Remark. Enthusiasts for tensor products will recognize that (i) is also equivalent to the statement $\|I_H \otimes A\| = \|A\|$. Nonenthusiasts can ignore this remark.

5. THE OPERATOR-VALUED EXTENSIONS. The promised operator version of (N1) follows very easily from Theorem 2:

Theorem 3. *Let $A = (a_{j,k})$ be an $m \times n$ matrix and let H be a Hilbert space. Suppose that $X_1, \dots, X_n \in L(H)$ and $Y_j = \sum_{k=1}^n a_{j,k} X_k$ for $1 \leq j \leq m$. Then we have the operator inequality*

$$\sum_{j=1}^m Y_j^* Y_j \leq \|A\|^2 \sum_{k=1}^n X_k^* X_k. \quad (4)$$

Proof. The statement is equivalent to the assertion that

$$\sum_{j=1}^m \|Y_j u\|^2 \leq \|A\|^2 \sum_{k=1}^n \|X_k u\|^2$$

for all $u \in H$. But $Y_j u = \sum_{k=1}^n a_{j,k} X_k u$ for each j , so this is a case of (2). ■

Remark. By a straightforward adaptation of the proofs of Theorems 2 and 3, one can establish the following analogous result, reproducing *positivity* of A rather than its norm: *Suppose that the $n \times n$ matrix A is positive, so that $\sum_{j=1}^n \sum_{k=1}^n a_{j,k} \bar{x}_j x_k \geq 0$ for all scalars x_j . Then $\sum_{j=1}^n \sum_{k=1}^n a_{j,k} X_j^* X_k \geq 0$ for $X_j \in L(H)$.* We leave it to sufficiently interested readers to write out the details.

Theorems 1 and 3 fit together neatly to deliver the operator version of (N2).

Theorem 4. *Let $A = (a_{j,k})$ be an $m \times n$ matrix and let H be a Hilbert space. Let X_j ($1 \leq j \leq m$) and Y_k ($1 \leq k \leq n$) be elements of $L(H)$, and let*

$$T = \sum_{j=1}^m \sum_{k=1}^n a_{j,k} X_j^* Y_k.$$

Then

$$T^* T \leq \|A\|^2 \left\| \sum_{j=1}^m X_j^* X_j \right\| \left\| \sum_{k=1}^n Y_k^* Y_k \right\|. \tag{5}$$

Proof. For each j , let $Z_j = \sum_{k=1}^n a_{j,k} Y_k$. Then $T = \sum_{j=1}^m X_j^* Z_j$, and by Theorem 3,

$$\sum_{j=1}^m Z_j^* Z_j \leq \|A\|^2 \sum_{k=1}^n Y_k^* Y_k.$$

By Theorem 1,

$$T^* T \leq \left\| \sum_{j=1}^m X_j^* X_j \right\| \left\| \sum_{j=1}^m Z_j^* Z_j \right\|.$$

The statement follows. ■

6. NORM INEQUALITIES. Since $0 \leq X \leq Y$ implies $\|X\| \leq \|Y\|$, we can deduce the following inequality of norms in Theorem 3:

$$\left\| \sum_{j=1}^m Y_j^* Y_j \right\| \leq \|A\|^2 \left\| \sum_{k=1}^n X_k^* X_k \right\|. \tag{6}$$

Similarly, in Theorem 4, we have the following norm inequality, suppressing the inherent asymmetry between X_j and Y_k :

$$\|T\|^2 \leq \|A\|^2 \left\| \sum_{j=1}^m X_j^* X_j \right\| \left\| \sum_{k=1}^n Y_k^* Y_k \right\|. \tag{7}$$

This norm inequality was given in [10], without any mention of the underlying operator inequality we have stated.

Now $\|\sum_{j=1}^m X_j^* X_j\| \leq \sum_{j=1}^m \|X_j\|^2$, so (7) holds with the right-hand side replaced by $(\sum_{j=1}^m \|X_j\|^2)(\sum_{k=1}^n \|Y_k\|^2)$. This inequality was given in [9], again without the underlying operator inequality.

Both [9] and [10] actually present further inequalities for other norms, the “trace class” or “Schatten ideal” norms. We will not embark on this topic here, except to mention that their results (or most of them) can be derived as consequences of Theorems 3 and 4.

Let us explore alternative types of norm inequalities a little further. Reverting to the setting of Theorem 3, let $Y_j = \sum_{k=1}^n a_{j,k} X_k$. Another possible generalization from the scalar case would be an inequality of the form

$$\sum_{j=1}^m \|Y_j\|^2 \leq C^2 \sum_{k=1}^n \|X_k\|^2, \quad (8)$$

By comparison with (6), we now have a larger quantity on both sides. Now $\|Y_j\| \leq \sum_{k=1}^n |a_{j,k}| \|X_k\|$, so by the original scalar inequality (N1), statement (8) holds with C equal to the norm of the matrix $(|a_{j,k}|)$ (which in general is larger than $\|A\|$). As the following simple example shows, (8) does not always hold with $C = \|A\|$.

Example 1. Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

so that $y_1 = x_1 + x_2$ and $y_2 = x_1 - x_2$. Then $|y_1|^2 + |y_2|^2 = 2|x_1|^2 + 2|x_2|^2$, so $\|A\| = \sqrt{2}$. Now take the 2×2 matrices $X_1 = \text{diag}(1, 1)$ and $X_2 = \text{diag}(1, -1)$. Then $\|X_1\| = \|X_2\| = 1$, and clearly $\|Y_1\| = \|Y_2\| = 2$, so (8) only holds with $C \geq 2$.

7. OPERATOR-VALUED VERSIONS OF SOME CLASSICAL INEQUALITIES. We finish by recording the operator-valued generalizations of some famous classical inequalities of this type. These inequalities describe the norms of the infinite matrices in question, thereby equating to the best constant when stated for all finite truncations (in fact, exact determination of the norm of the $n \times n$ truncation is generally not easy).

Example 2. *Hardy’s inequality* was mentioned in the Introduction. More generally, as an operator on ℓ_p , the norm of the Cesaro matrix A is $p/(p-1)$. A proof can be seen in [5, pp. 239–242]. For the case we want, $p = 2$, a very neat proof appeared in the MONTHLY article [3], where it is attributed to [2]. It is so short that we can repeat it here. One can easily verify that AA^* is the matrix having $1/\max(j, k)$ in place (j, k) , and hence that $A + A^* = AA^* + D$, where D is the diagonal matrix with entries $(1, \frac{1}{2}, \frac{1}{3}, \dots)$. Then $\|I - D\| = 1$ and

$$(I - A)(I - A^*) = I - A - A^* + AA^* = I - D,$$

so $\|(I - A)(I - A^*)\| = \|I - A\|^2 = 1$, hence $\|A\| \leq 2$.

Theorem 3 gives the (linear) operator-valued version: Let $X_k \in L(H)$ ($1 \leq k \leq n$) and let $Y_j = \frac{1}{j}(X_1 + X_2 + \dots + X_j)$ for $1 \leq j \leq n$. Then $\sum_{j=1}^n Y_j^* Y_j \leq$

$4 \sum_{k=1}^n X_k^* X_k$. The best constant, valid for all n , is 4, since this is already the case for scalars. Of course, one can also write down a bilinear version as in Theorem 4.

Example 3. *Copson's inequality* is the dual of Hardy's inequality: the norm of the transposed matrix is also 2. We go straight to the statement of the operator-valued version: Let $X_k \in L(H)$ ($1 \leq k \leq n$) and let $Y_j = \sum_{k=j}^n X_k/k$ ($1 \leq j \leq n$). Then $\sum_{j=1}^n Y_j^* Y_j \leq 4 \sum_{k=1}^n X_k^* X_k$.

Example 4. According to [5], the result we know as *Hilbert's inequality* was simply included by Hilbert in his lectures, and it was actually published by Weyl in his dissertation in 1908. In matrix form, it states the following: let $h_{j,k} = 1/(j+k+1)$ for $j \geq 0, k \geq 0$. Then the norm of the matrix $(h_{j,k})$ equals π . This can, of course, be written out explicitly in both linear and bilinear versions as in (N1) and (N2). The bilinear version is often known as Hilbert's "double series theorem". Hilbert's beautiful proof can be seen in [5, pp. 235–236], and in generalized form in [11, Section 13.4]. (Choi's method, described in Example 2, gives a quick proof of the weaker bound 4, by comparison with AA^* .) So the (linear) operator-valued version is: Let $X_k \in L(H)$ ($0 \leq k \leq n$), and let $Y_j = \sum_{k=0}^n X_k/(j+k+1)$ for $0 \leq j \leq n$. Then $\sum_{j=0}^n Y_j^* Y_j \leq \pi^2 \sum_{k=0}^n X_k^* X_k$.

It is tempting to conjecture that Hilbert would have approved of this extension of his inequality to the spaces named after him.

REFERENCES

1. Bhatia, R. (2007) *Positive Definite Matrices*. Princeton, NJ: Princeton Univ. Press.
2. Brown, A., Halmos, P. R., Shields, A. L. (1965). Cesàro operators. *Acta Sci. Math.* 26(1–2): 125–137.
3. Choi, M. D. (1983). Tricks or treats with the Hilbert matrix. *Amer. Math. Monthly.* 90(5): 301–312.
4. Halmos, P. R. (1958). *Finite-Dimensional Vector Spaces*. Princeton, NJ: Van Nostrand.
5. Hardy, G. H., Littlewood, J., Polya, G. (1967). *Inequalities*, 2nd ed. Cambridge: Cambridge Univ. Press.
6. Jameson, G. J. O. (1996). Khinchin's inequality for operators. *Glasgow Math. J.* 38(3): 327–336.
7. Jocić, D. R. (1999). Cauchy–Schwartz inequality for elementary operators in Schatten ideals. *J. London Math. Soc.* 60(3): 925–934.
8. Lust-Piquard, F., Pisier, G. (1991). Non-commutative Khintchine and Paley inequalities. *Arkiv för Matematik* 29(2): 241–260.
9. Redheffer, R. M., Volkmann, P. (1983). Bilineare Funktionen mit Hilbertraum-Operatoren als Veränderlichen. *Math. Ann.* 262(1): 133–143.
10. Tonge, A. M. (1986). Bilinear mappings and trace class ideals. *Math. Ann.* 275(2): 281–290.
11. Young, N. J. (1988). *An Introduction to Hilbert Space*. Cambridge: Cambridge Univ. Press.

GRAHAM JAMESON graduated from Cambridge University in 1963 and received his doctorate from Edinburgh University in 1967. Most of his lecturing career was at Lancaster University, UK, where he is now emeritus. He is author of various articles on functional and classical analysis and analytic number theory, and of several books, for example *The Prime Number Theorem* (Cambridge, 2003). He continues to enjoy hiking in the mountains, especially in Scotland and Austria.

Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, United Kingdom.
g.jameson@lancaster.ac.uk