J. OPERATOR THEORY
18(1987), 99–113



# NORMS FOR MATRICES AND OPERATORS

C. K. FONG, HEYDAR RADJAVI and PETER ROSENTHAL

## 1. INTRODUCTION

We consider several norms on (finite or infinite, complex) matrices: Schur norms (associated with the Schur product operation), maximal column norms, maximal row norms, and maximal entry norms. Precise definitions are given in Section 2 below. If $T$ is an operator on a (finite or infinite dimensional, complex) Hilbert space, then $T$ has a matrix representation with respect to each orthonormal basis. Each matrix representation of $T$ has each of the above matrix norms. Our main results concern the variation of these matrix norms as the basis varies.

General questions that are partially answered include the following: for a given matrix norm, for which operators $T$ is the matrix norm always equal to the operator norm? What is the supremum, over all bases, of the matrix norms of $T$? The infimum?

We also give sufficient conditions that certain of these matrix norms be equal to each other.

Our study was motivated by previous work on these questions. The paper [10] of Stout contains several related results; we present a somewhat different approach in Section 4 below. In [7], Ong investigates these questions in the finite dimensional case. We subsume Ong's main results in Section 3 below.

A number of unanswered subquestions are discussed at the end of this paper.

## 2. PRELIMINARIES

Consider the set $\mathscr{M}_{m,n}$ of $m\times n$ complex matrices, where $m$ and $n$ are positive integers or $\infty$. We will be concerned with several different norms on $\mathscr{M}_{m,n}$.

Let $A=(\alpha_{ij})\in\mathscr{M}_{m,n}$. Then the familiar *operator norm* of $A$, the norm o $A$ regarded as an operator from $\ell^2$ to $\ell^2$, is given by

$$\|A\| = \sup\{|\sum\alpha_{ij}x_i\bar{y}_j| : \sum|x_i|^2 = \sum|y_j|^2 = 1\}.$$

Recall that the *Schur product* (sometimes called the *Hadamard product*) of two matrices of the same size is their entry-wise product; i.e., if $A = (\alpha_{ij})$ and $B = (\beta_{ij})$, then $A * B = (\alpha_{ij}\beta_{ij})$. We define the *Schur norm* of $A$, to be denoted $\|A\|_s$, by $\|A\|_s = \sup\{\|A * B\| : \|B\| \leqslant 1\}$; i.e., $\|A\|_s$ is the norm of the operator on $\mathcal{M}_{m,n}$ corresponding to Schur multiplication by $A$. Schur norms have been studied in many papers — see Stout [10] and references given there.

We consider three other norms: the *maximal column norm* of $A = (\alpha_{ij})$ is defined by $\|A\|_c = \sup_j(\sum_i |\alpha_{ij}|^2)^{1/2}$, the *maximal row norm* by $\|A\|_r = \sup_i(\sum_j |\alpha_{ij}|^2)^{1/2}$, and the *maximal entry norm* by $\|A\|_e = \sup_{i,j} |\alpha_{ij}|$. (The maximal entry norm was introduced in [10].)

2.1. PROPOSITION. *For any matrix $A$,*

(i) *$\|A\|_c$ is the norm of $A$ as an operator from $\ell^1$ to $\ell^2$;*

(ii) *$\|A\|_r$ is the norm of $A$ as an operator from $\ell^2$ to $\ell^\infty$;*

(iii) *$\|A\|_e$ is the norm of $A$ as an operator from $\ell^1$ to $\ell^\infty$.*

*Proof.* (i) If $\{x_i\} \in \ell^1$, and if $\{e_1, e_2, \ldots\}$ denotes the standard basis for $\ell^1$, then

$$\|A\{x_i\}\|_2 = \|\sum x_i A e_i\|_2 \leqslant \sum |x_i| \cdot \|Ae_i\|_2 \leqslant$$
$$\leqslant (\sup\|Ae_i\|_2) \sum_i |x_i| = \|A\|_c \cdot \|\{x_i\}\|_1.$$

Thus the norm of $A$ as an operator from $\ell^1$ to $\ell^2$ is at most $\|A\|_c$. Clearly it cannot be smaller than $\|A\|_c$.

(ii) This is reduced to (i) by considering the transpose of $A$ as an operator from $\ell^1$ to $\ell^2$ and observing that $\|A\|_r = \|A^t\|_c$.

(iii) With notation as in (i),

$$\|A\{x_i\}\|_\infty = \|\sum x_i A e_i\|_\infty \leqslant \sum |x_i| \cdot \|Ae_i\|_\infty \leqslant$$
$$\leqslant (\sup \|Ae_i\|_\infty) \sum_i |x_i| = \|A\|_e \cdot \|\{x_i\}\|_1.$$

Hence $\|A\|_e$ is the norm of $A$ as an operator from $\ell^1$ to $\ell^\infty$. ▨

If $T$ is an operator on a finite or infinite dimensional Hilbert space $\mathcal{H}$ and $\mathcal{E} = \{e_j\}$ is an orthonormal basis for $\mathcal{H}$, then $T$ has a matrix representation with respect to $\mathcal{E}$; we let

$$[T]_{\mathcal{E}} = \begin{bmatrix} (Te_1, e_1) & (Te_2, e_1) & \cdots \\ (Te_1, e_2) & (Te_2, e_2) & \cdots \\ (Te_1, e_3) & \cdots & \\ \vdots & \vdots & \end{bmatrix}.$$

If $|||\cdot|||$ is any of the above norms on $\mathcal{M}_{m,n}$, we define $|||T|||_{\mathscr{E}}$ to be $|||[T]_{\mathscr{E}}|||$. Of course $\|T\|_{\mathscr{E}} = \|T\|$ for all $\mathscr{E}$, but the other norms of $T$ vary with the basis $\mathscr{E}$. For any $|||\cdot|||$, we define $|||T|||$ to be

$$\sup\{|||T|||_{\mathscr{E}} : \mathscr{E} \text{ is an orthonormal basis}\}.$$

Then $|||\cdot|||$ is a norm on the space $\mathscr{B}(\mathscr{H})$ of bounded operators on $\mathscr{H}$. These norms and $\inf\{|||T|||_{\mathscr{E}} : \mathscr{E}$ is an orthonormal basis$\}$ are studied in Section 4 below.

There is an alternate description of the Schur norm that is sometimes useful. Let $\|(\alpha_{ij})\|_1$ denote the trace norm of the matrix $(\alpha_{ij})$; i.e., $\|(\alpha_{ij})\|_1$ is the trace of $\sqrt{(\alpha_{ij})^*(\alpha_{ij})}$.

2.2. THEOREM. *For any matrix* $A = (\alpha_{ij})$,

(i) $\|A\|_s = \sup\{\|(\alpha_{ij}\xi_i\eta_j)\|_1 : \sum|\xi_i|^2 = \sum|\eta_j|^2 = 1\}$;

(ii) $\|A\|_s = \sup\{\|(\alpha_{ij}\xi_i\eta_j)\|_1 : \xi_i \geqslant 0, \ \ \eta_j \geqslant 0, \ \sum\xi_i^2 = \sum\eta_j^2 = 1\}$.

*Proof.* The first equality is a consequence of the relations

$$\|(\alpha_{ij})\|_s = \sup\{\|(\alpha_{ij}x_{ij})\| : \|(x_{ij})\| \leqslant 1\} =$$

$$= \sup\{|\sum_{i,j}\alpha_{ij}x_{ij}\xi_i\eta_j| : \|(x_{ij})\| \leqslant 1, \ \sum|\xi_i|^2 = \sum|\eta_j|^2 = 1\} =$$

$$= \sup\{|\operatorname{tr}[(\alpha_{ij}\xi_i\eta_j)(x_{ji})]| : \|(x_{ij})\| \leqslant 1, \ \sum|\xi_i|^2 = \sum|\eta_j|^2 = 1\} =$$

$$= \sup\{\|(\alpha_{ij}\xi_i\eta_j)\|_1 : \sum|\xi_i|^2 = \sum|\eta_j|^2 = 1\}.$$

To prove the second equality, observe that $x_{ij}\xi_i\eta_j$ can be expressed as $x_{ij}\alpha_i\beta_j|\xi_i|\cdot|\eta_j|$ with $|\alpha_i| = |\beta_j| = 1$ for all $i$ and $j$, and furthermore

$$\|(x_{ij}\alpha_i\beta_j)\| = \|\operatorname{diag}(\alpha_1, \alpha_2, \ldots)(x_{ij})\operatorname{diag}(\beta_1, \beta_2, \ldots)\| =$$

$$= \|(x_{ij})\| \leqslant 1.$$ ▨

The next result is well-known ([7], [9], [10]).

2.3. COROLLARY. *For any matrix* $A$, $\|A\|_s \leqslant \|A\|_c$ *and* $\|A\|_s \leqslant \|A\|_r$.

*Proof.* Let $A = (\alpha_{ij})$. If $\sum|\xi_i|^2 = \sum|\eta_j|^2 = 1$, then

$$\|(\alpha_{ij}\xi_i\eta_j)\|_1 = \|\operatorname{diag}(\xi_1, \xi_2, \ldots)(\alpha_{ij}\eta_j)\|_1 \leqslant$$

$$\leqslant \|\operatorname{diag}(\xi_1, \xi_2, \ldots)\|_2 \cdot \|(\alpha_{ij}\eta_j)\|_2 =$$

$$= (\sum_{ij}|\alpha_{ij}\eta_j|^2)^{1/2} = (\sum_j(\sum_i|\alpha_{ij}|^2)|\eta_j|^2)^{1/2} \leqslant \|A\|_c.$$

Thus $\|A\|_s \leqslant \|A\|_c$. The proof of the second inequality is similar. ▨

Note that clearly $\|A\|_e \leqslant \|A\|_c$, $\|A\|_r \leqslant \|A\|$, and $\|A\|_e \leqslant \|A\|_s$, so 2.3 immediately yields the following.

2.4. COROLLARY. *For any matrix $A$,*

$$\|A\|_e \leqslant \|A\|_s \leqslant \|A\|_c, \ \|A\|_r \leqslant \|A\|.$$ ▨

There are some cases where the inequalities of Corollary 2.4 become equalities.

2.5. PROPOSITION. *If $A$ is a matrix of rank 1, then $\|A\|_s = \|A\|_e$.*

*Proof.* Since $A$ has rank 1, it is of the form $(\alpha_i\beta_j)$, where $\{\alpha_i\}$ and $\{\beta_j\}$ are $\ell^2$ sequences. If $B = (x_{ij})$, then

$$\|A * B\| = \|(\alpha_i\beta_j \cdot x_{ij})\| =$$

$$= \|\operatorname{diag}(\alpha_1, \alpha_2, \ldots)(x_{ij})\operatorname{diag}(\beta_1, \beta_2 \ldots)\| \leqslant$$

$$\leqslant \max|\alpha_i| \cdot \max|\beta_j| \cdot \|(x_{ij})\| = \max|\alpha_i\beta_j| \cdot \|(x_{ij})\| = \|A\|_e \cdot \|B\|.$$

Alternatively, this result can be obtained as a corollary of Theorem 2.2. ▨

The following characterizes $\|A\|_s = \|A\|_c$ in the finite-dimensional case (see Theorem 3.1 below).

2.3. PROPOSITION. *Let $A = (\xi_i u_{ij} \eta_j^{-1})$, where $\{\xi_i\} \in \ell^2$ with $\xi_i > 0$ for all $i$, $(u_{ij})$ is a unitary matrix, and $\eta_j = (\sum_i \xi_i^2 |u_{ij}|^2)^{1/2}$. Then $\|A\|_s = \|A\|_c = 1$.*

*Proof.* Note that for each $j$ there exists some $i$ with $u_{ij} \neq 0$, so $\eta_j > 0$ Each column of $A$ is a unit vector, because, for each $j$,

$$\sum_i |\xi_i u_{ij} \eta_j^{-1}|^2 = (\sum_j \xi_i^2 |u_{ij}|^2)\eta_j^{-2} = 1.$$

Thus $\|A\|_c = 1$ and it remains to show that $\|A\|_s \geqslant 1$. Since $\sum \eta_j^2 = \sum \xi_i^2$, and since normalizing $\{\xi_i\}$ does not change $A$, we can assume that $\sum \xi_i^2 = \sum \eta_j^2 = 1$. By Theorem 2.2,

$$\|A\|_s \geqslant \|(\xi_i^2 u_{ij})\|_1 = \|\operatorname{diag}(\xi_1^2, \xi_2^2, \ldots)(u_{ij})\|_1 =$$

$$= \|\operatorname{diag}(\xi_1^2, \xi_2^2, \ldots)\|_1 = 1.$$

## 3. FINITE-DIMENSIONAL RESULTS

In this section we consider finite matrices only. We begin with a description of the matrices with equal Schur and column norms.

3.1. THEOREM. *If* $A = (\alpha_{ij})$ *and* $\|A\|_c = t$, *then* $\|A\|_s = t$ *if and only if there is a unitary matrix* $(u_{ij})$ *and vectors* $\{\xi_i\}$, $\{\eta_j\}$ *with* $\xi_i \geqslant 0$, $\eta_j \geqslant 0$ *and* $\sum \xi_i^2 = \sum \eta_j^2 = 1$ *such that* $\alpha_{ij}\eta_j = t\xi_i u_{ij}$ *for all* $i, j$.

*Proof.* With no loss of generality we can replace $t$ by 1.

First assume there are vectors $\{\xi_i\}$, $\{\eta_j\}$ and a unitary matrix $(u_{ij})$ as described. Then

$$\sum_{i,j} |\alpha_{ij}\eta_j|^2 = \sum_{i,j} |\xi_i u_{ij}|^2$$

implies $\sum_j \eta_j^2(\sum_i |\alpha_{ij}|^2) = \sum_i \xi_i^2 = 1$, and since $\sum_i |\alpha_{ij}|^2 \leqslant 1$ for every $j$ by assumption, the equation $\sum \eta_j^2 = 1$ implies that $\sum_i |\alpha_{ij}|^2 = 1$ for all $j$. Now

$$\eta_j^2 = \eta_j^2 \sum_i |\alpha_{ij}|^2 = \sum_i |\alpha_{ij}\eta_j|^2 = \sum_i |\xi_i u_{ij}|^2 = \sum_i \xi_i^2 |u_{ij}|^2,$$

so that the equation $\|A\|_s = 1$ follows from Proposition 2.6.

To show the converse assume $\|A\|_s = 1$. By Theorem 2.2 and a compactness argument there are vectors $\{\xi_i\}$ and $\{\eta_j\}$ with $\xi_i \geqslant 0$, $\eta_j \geqslant 0$, $\sum \xi_i^2 = 1 = \sum \eta_j^2$ and such that

$$\|(\alpha_{ij}\xi_i\eta_j)\|_1 = \|A\|_s = 1.$$

Let $V = (v_{ij})$ be the unitary matrix in the polar decomposition of the matrix $(\alpha_{ij}\xi_i\eta_j)$. Then

$$1 = \|(\alpha_{ij}\xi_i\eta_j)\|_1 = \mathrm{tr}[(v_{ij})(\alpha_{ij}\xi_i\eta_j)] =$$

$$= \sum_{i,j} (\xi_i v_{ji})(\alpha_{ij}\eta_j) \leqslant (\sum_{i,j} \xi_i^2 |v_{ij}|^2)^{1/2} (\sum_{i,j} \eta_j^2 |\alpha_{ij}|^2)^{1/2} =$$

$$= (\sum_{i,j} \eta_j^2 |\alpha_{ij}|^2)^{1/2} \leqslant \|A\|_c = 1.$$

Thus the Schwarz inequality above is in fact an equality and, hence, there is a constant $\theta$ of absolute value 1 such that $\alpha_{ij}\eta_j = \theta\xi_j \cdot \bar{v}_{ji}$. Then $(\theta\bar{v}_{ji})$ is the desired unitary matrix.

Our next result is a strengthening of a theorem of Ong [7] characterizing $\|A\|_s = \|A\|$.

3.2. THEOREM. *For a finite matrix $A$, $\|A\|_s = \|A\|_r = \|A\|_c = t$ if and only if there is a unitary matrix $U$ such that by permuting the rows and permuting the columns, $A$ can be rewritten in the form $t(U \oplus B)$, where $\|B\|_r \leq 1$ and $\|B\|_c \leq 1$.*

*Proof.* We can assume again that $t = 1$. Since a permutation of either columns or rows leaves $\|A\|_s$, $\|A\|_r$ and $\|A\|_c$ unchanged, these norms are all equal to 1 if $A$ is of the form stated above. To show the converse, let $A = (\alpha_{ij})$ and apply Theorem 2.2 to get nonnegative vectors $\{\xi_i\}$ and $\{\eta_j\}$ with $\sum \xi_i^2 = \sum \eta_j^2 = 1$ and

$$1 = \|(\alpha_{ij})\|_s = \|(\alpha_{ij}\xi_i\eta_j)\|_1.$$

By the proof of Theorem 3.1 there is a unitary matrix $U = (u_{ij})$ such that $\alpha_{ij}\eta_j = \xi_i u_{ij}$ for all $i$ and $j$. A similar argument, this time using $\|A\|_s = \|A\|_r$, shows the existence of a unitary matrix $V = (v_{ij})$ with $\xi_i\alpha_{ij} = v_{ij}\eta_j$. Hence

$$(\alpha_{ij})\mathrm{diag}(\eta_1, \eta_2, \ldots) = \mathrm{diag}(\xi_1, \xi_2, \ldots)(u_{ij}),$$

$$(v_{ij})\mathrm{diag}(\eta_1, \eta_2, \ldots) = \mathrm{diag}(\xi_1, \xi_2, \ldots)(\alpha_{ij}).$$

It follows that the two diagonal matrices above have the same rank and thus $I = \{i : \xi_i \neq 0\}$ and $J = \{j : \eta_j \neq 0\}$ have the same number of elements. Permuting both columns and rows if necessary, we can assume that $I = J = \{1, 2, \ldots, k\}$. The matrix equations above then take the form

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} Y & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix},$$

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}\begin{pmatrix} Y & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} X & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where $X = \mathrm{diag}(\xi_1, \ldots, \xi_k)$, $Y = \mathrm{diag}(\eta_1, \ldots, \eta_k)$, and the matrices $A$, $U$ and $V$ have been partitioned accordingly. Since $X$ and $Y$ are invertible, it follows immediately that $A_{21}$, $A_{12}$, $U_{12}$ and $V_{21}$ are all zero. Thus $U_{11}$ and $V_{11}$ are unitary; this forces $U_{21}$ and $V_{12}$ to be zero also. Now the equations $A_{11}Y = XU_{11}$ and $V_{11}Y = XA_{11}$ imply $X^2U_{11} = V_{11}Y^2$, and since $X^2$ and $Y^2$ are positive, the uniqueness of polar decomposition implies $U_{11} = V_{11}$. Thus $Y^2 = U_{11}^{-1}X^2U_{11}$ and hence $Y = U_{11}^{-1}XU_{11}$, because $X$ and $Y$ are positive. We conclude that $A_{11} = U_{11}$. Clearly $\|A_{22}\|_r \leq 1$ and $\|A_{22}\|_c \leq 1$.

3.3. COROLLARY (Ong [7]). *If $A$ is a finite matrix, then $\|A\|_s = \|A\| = t$ if and only if, by permuting the rows and permuting the columns, $A$ can be re-written in the form $t(U \oplus B)$ where $U$ is unitary and $\|B\| \leq 1$.*

*Proof.* If $\|A\|_s = \|A\|$, then $\|A\|_s = \|A\|_r = \|A\|_c$ by Corollary 2.3, so that $A$ is of the desired form by Theorem 3.2. The converse is obvious. ▨

3.4. THEOREM. *For a finite matrix $A$, $\|A\|_c = \|A\|$ if and only if $\|A\|^2$ is on the main diagonal of $A^*A$.*

*Proof.* Assume with no loss of generality that $\|A\| = 1$. Then the $(i, i)$ entry in $A^*A$ is 1 if and only if the $i^{\text{th}}$ column of $A$ has norm 1. ▨

We now consider the problem of determining the operators whose various norms are equal to their operator norms. We shall use the following fact more than once in the sequel: an $n \times n$ matrix with trace $t$ is unitarily equivalent to a matrix whose main diagonal entries are all equal to $t/n$. See Fillmore [3] and Halmos [6, p. 109] for this and related results.

3.5. THEOREM (Ong [7]). *If $T$ is an operator on a finite-dimensional space, then the following are equivalent:*

(1) $\|T\|_{c,\mathscr{E}} = \|T\|$ *for every orthonormal basis* $\mathscr{E}$.

(2) $\|T\|_{s,\mathscr{E}} = \|T\|$ *for every orthonormal basis* $\mathscr{E}$.

(3) *$T$ is a multiple of a unitary operator.*

*Proof.* The implications (3) ⇒ (2) ⇒ (1) are easy to verify. To prove (1) ⇒ (3) we may assume $\|T\| = 1$. Let $P = T^*T$. By Theorem 3.4 the main diagonal of $[P]_{\mathscr{E}}$ contains 1 for every basis $\mathscr{E}$. Now there exists a basis relative to which $[P]_{\mathscr{E}}$ has constant main diagonal. This makes all diagonal entries equal to 1. Since $\|P\| = 1$, the nondiagonal entries are forced to be zero. Thus $P = I$, so that $T$ is unitary.

There are few operators whose maximal entry norms are equal to their operator norms.

3.6. THEOREM. *For $T$ an operator on a finite dimensional space, $\|T\|_{c,\mathscr{E}} = \|T\|$ for all orthonormal bases $\mathscr{E}$ if and only if $T$ is a multiple of the identity.*

*Proof.* If $T$ is not a multiple of the identity, then, by [8], there exists an orthonormal basis $\mathscr{E}$ such that every entry of $[T]_{\mathscr{E}}$ is nonzero. Thus the absolute value of each entry is less than the norm of the column containing it, which is no more than $\|T\|$. Thus $\|T\|_{c,\mathscr{E}} < \|T\|$. Hence $\|T\|_{c,\mathscr{E}} = \|T\|$ for every $\mathscr{E}$ forces $T$ to be a multiple of the identity. The converse is trivial. ▨

Recall that the Hilbert-Schmidt norm, $\|T\|_2$, of an operator is the square root of the sum of the squares of the entries in any matrix representation of $T$.

3.7. THEOREM. *If $T$ is an operator on an $n$-dimensional space, then* $\inf\{\|T\|_{c,\mathscr{E}} : \mathscr{E}$ *is an orthonormal basis*$\} = \|T\|_2 / \sqrt{n}$.

*Proof*. By definition we have $\|T\|_{c,\mathcal{E}} \geq \|T\|_2/\sqrt{n}$ for every orthonormal basis $\mathcal{E}$. Now there is a basis $\mathcal{E}$ such that $[T^*T]_{\mathcal{E}}$ has constant diagonal entries. This means that the columns of $[T]_{\mathcal{E}}$ all have the same norm; thus $\|T\|_{c,\mathcal{E}} = \|T\|_2/\sqrt{n}$ for this basis. ∎

## 4. INFINITE-DIMENSIONAL RESULTS

In this section we consider bounded linear operators on separable, infinite-dimensional, complex Hilbert spaces.

The first result that we consider is due to Anderson [1]. Anderson's proof relies on several deep theorems. A stronger theorem than ours is given by Stout [10]; our proof appears to be more direct. Recall that the essential numerical range of $T$, denoted $W_e(T)$, is the numerical range of the image of $T$ in the Calkin algebra (see [2, p. 127]).

4.1. THEOREM ([1]). *If* $0 \in W_e(T)$ *and* $p > 1$, $\varepsilon > 0$, *then there is an orthonormal basis* $\{e_n\}$ *such that* $\sum_{n=1}^{\infty} |(Te_n, e_n)|^p < \varepsilon$.

*Proof*. Since $0 \in W_e(T)$, there is an orthonormal sequence $\{f_n\}$ such that $(Tf_n, f_n) \to 0$. Choosing a subsequence we may assume that $\sum_{n=1}^{\infty} |(Tf_n, f_n)|^p < \varepsilon/2^p$. Let $D$ be the diagonal operator given by

$$Dx = \sum_{n=1}^{\infty} (Tf_n, f_n)(x, f_n) f_n .$$

Then $D$ is in the Schatten class $\mathcal{C}_p$ and $\|D\|_p^p < \varepsilon/2^p$. The operator $S = T - D$ now satisfies $(Sf_n, f_n) = 0$ for all $n$. Extend $\{f_n\}$ to an orthonormal basis $\mathcal{E}$ of the space. By permuting $\mathcal{E}$, if necessary, we can assume that $[S]_{\mathcal{E}}$ has the block-matrix form $(S_{ij})_{i,j=1}^{\infty}$, where each diagonal block $S_{kk}$ is an $n_k \times n_k$ matrix whose main diagonal is $(\alpha_k, 0, \ldots, 0)$. The integers $n_k$ can be prescribed arbitrarily. We choose them so that

$$(\sup |\alpha_k|^p) \sum_{k=1}^{\infty} n_k^{1-p} < \varepsilon/2^p .$$

Now each $S_{kk}$ is unitarily equivalent to a matrix whose main-diagonal entries are all $\alpha_k/n_k$. Thus there is an orthonormal basis $\{e_n\}$ such that

$$\sum_{n=1}^{\infty} |(Se_n, e_n)|^p = \sum_{k=1}^{\infty} n_k(|\alpha_k|/n_k)^p \leq$$

$$\leq \sup |a_k|^p \cdot \sum_{k=1}^{\infty} n_k^{1-p} \leq \varepsilon/2^p .$$

Since $T = S + D$, and since, by [5, p. 94],

$$\sum_{n=1}^{\infty} |(De_n, e_n)|^p \leqslant \|D\|_p^p < \varepsilon/2^p,$$

we conclude that $\sum_{n=1}^{\infty} |(Te_n, e_n)|^p < \varepsilon.$ ▨

We use $\sigma_{re}(T)$ to denote the right essential spectrum of $T$; i.e., to say that $0 \in \sigma_{re}(T)$ is to say that the image of $T$ in the Calkin algebra has no right inverse.

4.2. COROLLARY. *If $0 \in \sigma_{re}(T)$, $p > 2$ and $\varepsilon > 0$, then there is an orthonormal basis $\{e_n\}$ such that $\sum_{n=1}^{\infty} \|Te_n\|^p < \varepsilon$.*

*Proof.* Let $T = VP$ be the polar decomposition of $T$. From $0 \in \sigma_{re}(T)$ we have $0 \in \sigma_{re}(P)$ and hence $0 \in W_e(P)$. Applying Theorem 4.1 to $P^2$ gives a basis $\{e_n\}$ such that

$$\sum (P^2e_n, e_n)^{p/2} < \varepsilon,$$

so that $\sum \|Te_n\|^p \leqslant \sum \|Pe_n\|^2 < \varepsilon.$ ▨

We require two lemmas about Schur norms.

4.3. LEMMA. *If $A = (A_{ij})$ is a block matrix, then*

$$\|A\|_S^2 \leqslant \sum_{i,j} \|A_{ij}\|_S^2.$$

*Proof.* Let $B = (B_{ij})$ be any conforming block matrix with $\|B\| \leqslant 1$. Then

$$\alpha_{ij} \equiv \|A_{ij} * B_{ij}\| \leqslant \|A_{ij}\|_S$$

and thus $\|(\alpha_{ij})\|^2 \leqslant \|(\alpha_{ij})\|_2^2 \leqslant \sum_{i,j} \|A_{ij}\|_S^2$. Since

$$\|A * B\|^2 = \|(A_{ij} * B_{ij})\|^2 \leqslant \|(\alpha_{ij})\|^2,$$

it follows that $\|A\|_S^2 \leqslant \|(\alpha_{ij})\|^2 \leqslant \sum_{i,j} \|A_{ij}\|_S^2$. ▨

4.4. LEMMA. *Fix a positive integer $n$. Let $w_n = \exp(2\pi i/n)$ and let $U_n$ denote the $n\times n$ unitary matrix with $w_n^{ij}/\sqrt{n}$ in position $(i, j)$. If $A$ is an $m\times n$ matrix* *of the form*

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ \beta_2 & 0 & \dots & 0 \\ \beta_3 & & & \\ \vdots & \vdots & & \vdots \\ \beta_m & 0 & \dots & 0 \end{pmatrix},$$

*then* $\|U_m^* A U_n\|_s \leqslant \left(\frac{1}{m}\sum_{i=1}^{n} |\alpha_i|^2\right)^{1/2} + \left(\frac{1}{n}\sum_{j=2}^{m} |\beta_j|^2\right)^{1/2}.$

*Proof.* Let

$$A_1 = \begin{pmatrix} \alpha_1 & \alpha_2 \dots \alpha_n \\ 0 & 0 \dots 0 \\ \vdots & \\ 0 & 0 \dots 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \dots 0 \\ \beta_2 & 0 \dots 0 \\ \vdots & \\ \beta_n & 0 \dots 0 \end{pmatrix}.$$

Then

$$A_1 U_n = \begin{pmatrix} \gamma_1 & \gamma_2 \cdots \gamma_n \\ 0 & 0 \dots 0 \\ \vdots & \\ 0 & 0 \dots 0 \end{pmatrix},$$

where $\gamma_j = \sum \alpha_k w^{kj}/\sqrt{n}$. Since $U_m^* A_1 U_n$ is a rank-one matrix, $\|U_m^* A_1 U_n\|_s = \|U_m^* A_1 U_n\|_e$ by Proposition 2.5. The entries of $U_m^* A_1 U_n$ are all of absolute value $|\gamma_j|/\sqrt{m}$, and since

$$|\gamma_j|^2 \leqslant \sum_k |\alpha_k|^2 \sum_k |w_n^{kj}|^2/n = \sum_k |\alpha_k|^2,$$

we obtain $\|U_m^* A_1 U_n\|^2 \leqslant \sum_k |\alpha_k|^2/m$. The proof of the inequality $\|U_m^* A_2 U_n\|_s^2 \leqslant$ $\leqslant \sum_j |\beta_j|^2/n$ is similar. ▨

4.5. THEOREM. *If* $0 \in W_e(T)$, $\dim \mathscr{H}$ *is infinite, and* $\varepsilon > 0$, *then there is a basis* $\mathscr{E}$ *such that* $\|T\|_{s,\mathscr{E}} < \varepsilon$.

*Proof.* Since $W_e(T)$ contains 0, a result of [4] implies that $T$ has a block matrix

$$\begin{pmatrix} * & * \\ * & D \end{pmatrix},$$

where $D$ acts on an infinite-dimensional subspace and $\|D\| < \varepsilon/2$. Let $S$ be an operator having the same matrix except that $D$ is replaced by 0. It suffices to find a basis $\mathscr{E}$ such that $\|S\|_{s,\mathscr{E}} < \varepsilon/2$. Observe that for any preassigned sequence $\{n_j\}$ of positive integers there is a basis $\mathscr{F}$ such that $[S]_{\mathscr{F}}$ is an infinite block matrix $(S_{ij})$ where each $S_{ij}$ is an $n_i \times n_j$ matrix of the form

$$\begin{pmatrix} \alpha_1(i,j) & \alpha_2(i,j) & \dots & \alpha_{n_j}(i,j) \\ \beta_2(i,j) & 0 & \dots & 0 \\ \vdots & & & \\ \beta_{n_i}(i,j) & 0 & \dots & 0 \end{pmatrix}.$$

Note that, for each $i$,

$$\sum_{j=1}^{\infty} \sum_{k=1}^{n_j} |\alpha_k(i,j)|^2 \leqslant \|S\|^2.$$

Let $U = U_{n_1} \oplus U_{n_2} \oplus \dots$, where each direct summand is defined as the unitary matrix in Lemma 4.4. Then the $(i,j)$ block of $U^*[S]_{\mathscr{F}}U$ is

$$A_{ij} \equiv U_{n_i}^* S_{ij} U_{n_j}$$

and, by Lemma 4.4,

$$\|A_{ij}\|_s \leqslant \left(\frac{1}{n_i} \sum_{k=1}^{n_j} |\alpha_k(i,j)|^2\right)^{1/2} + \left(\frac{1}{n_j} \sum_{k=2}^{n_i} |\beta_k(i,j)|^2\right)^{1/2} \leqslant$$

$$\leqslant \left(\frac{2}{n_i} \sum_{k=1}^{n_j} |\alpha_k(i,j)|^2 + \frac{2}{n_j} \sum_{k=2}^{n_i} |\beta_k(i,j)|^2\right)^{1/2}.$$

Therefore

$$\sum_{i,j} \|A_{ij}\|_s^2 \leqslant 4\|S\|^2 \left(\sum_{j=1}^{\infty} \frac{1}{n_j}\right).$$

An appropriate choice of $\{n_j\}$ yields

$$\|U^*[S]_{\mathscr{F}}U\|_s < \frac{\varepsilon}{2}$$

by Lemma 4.3. Thus we have found a basis $\mathscr{E}$ giving $\|S\|_{s,\mathscr{E}} < \varepsilon/2$ as desired. ▨

4.6. THEOREM. *For each* $T \in \mathscr{B}(\mathscr{H})$ *(with* $\dim \mathscr{H}$ *infinite),*

$$\inf_{\mathscr{E}} \|T\|_{e,\mathscr{E}} = \inf_{\mathscr{E}} \|T\|_{s,\mathscr{E}} = \text{dist}(0, W_e(T)).$$

*Proof.* Let $\lambda \in W_e(T)$ with $|\lambda| = \text{dist}(0, W_e(T))$. That

$$\inf_{\mathscr{E}} \|T\|_{s,\mathscr{E}} \geqslant \inf_{\mathscr{E}} \|T\|_{e,\mathscr{E}} \geqslant |\lambda|$$

is easy to see. Since $0 \in W_e(T - \lambda)$, by Theorem 4.5 there exists a basis $\mathscr{E}$ such that $\|T - \lambda\|_{s,\mathscr{E}} < \varepsilon$ and hence $\|T\|_{s,\mathscr{E}} < |\lambda| + \varepsilon$. ▨

4.7. THEOREM. *If* $T \in \mathscr{B}(\mathscr{H})$ *(with* $\dim \mathscr{H}$ *infinite), then*

$$\inf_{\mathscr{E}} \|T\|_{c,\mathscr{E}} = \inf_{\mathscr{E}} \sigma_e(T^*T)^{1/2}.$$

*Proof.* Let $\lambda = \inf \sigma_e(T^*T)^{1/2}$, and let $T = VP$ be the polar decomposition of $T$ (so that $P = (T^*T)^{1/2}$). Then $0 \in \sigma_{re}V(P - \lambda)$ and hence, by Corollary 4.2, there is a basis $\mathscr{E}$ such that $\|V(P - \lambda)\|_{c,\mathscr{E}} < \varepsilon$ for any preassigned $\varepsilon > 0$. Thus

$$\|VP\|_{c,\mathscr{E}} \leqslant \|V(P - \lambda)\|_{c,\mathscr{E}} + \lambda \|V\|_{c,\mathscr{E}} <$$

$$< \varepsilon + \lambda \|V\| = \varepsilon + \lambda.$$

We have shown that $\inf_{\mathscr{E}} \|T\|_{c,\mathscr{E}} \leqslant \inf_{\mathscr{E}} \sigma_e(T^*T)^{1/2}$. The reverse inequality is easy to verify. ▨

4.8. THEOREM. *For* $T \in \mathscr{B}(\mathscr{H})$ *with* $\mathscr{H}$ *infinite-dimensional, the following are equivalent:*

(1) $\|T\|_{s,\mathscr{E}} = \|T\|$ *for every basis* $\mathscr{E}$,

(2) $\|T\|_{e,\mathscr{E}} = \|T\|$ *for every basis* $\mathscr{E}$,

(3) $T = \lambda + K$ *for a compact operator* $K$ *and a complex number* $\lambda$ *satisfying* $\|\lambda + K\| = |\lambda|$.

*Proof.* The implications (3) ⇒ (2) ⇒ (1) are easily seen to hold. To prove (1)⇒(3) let $\lambda \in W_e(T)$. By Theorem 4.5 there is a basis $\mathscr{E}$ with $\|T - \lambda\|_{s,\mathscr{E}}$ arbitrarily small. Now

$$\|T\| - |\lambda| = \|T\|_{s,\mathscr{E}} - |\lambda| \leqslant \|T - \lambda\|_{s,\mathscr{E}},$$

so that $|\lambda| = \|T\|$. Therefore $W_e(T)$ is a subset of the circle $\{z : |z| = \|T\|\}$. Since $W_e(T)$ is convex, it must be a singleton, say $\{\lambda\}$. Then $T = K + \lambda$ for some compact $K$ and $\|T\| = |\lambda|$.

4.9. THEOREM. *Let* $\tilde{T}$ *denote the image of the operator* $T$ *in the Calkin algebra. Then* $\|T\|_{c,\mathscr{E}} = \|T\|$ *for every basis* $\mathscr{E}$ *if and only if* $\tilde{T}$ *is a scalar multiple of an isometry.*

*Proof.* Assume $\tilde{T}$ is a multiple of an isometry. Then $\tilde{T}^*\tilde{T} = \alpha\tilde{I}$. Thus $T^*T = \alpha I + K$ for some compact $K$. If $\{e_n\}$ is any basis for $\mathscr{H}$, then

$$\sup_n \|Te_n\|^2 = \sup_n (T^*Te_n, e_n) \geqslant \alpha \geqslant \|T\|.$$

Since $\sup_n \|Te_n\|^2 \leqslant \|T\|$, we get $\|T\|_{c,\mathscr{E}} = \|T\|$.

To prove the converse suppose $\tilde{T}^*\tilde{T}$ is not a scalar. Then there exists a $\lambda \in \sigma_e(T^*T)$ with $\lambda < \|T\|^2$. By Theorem 4.1 there is a basis $\mathscr{E} = \{e_n\}$ with

$$((T^*T - \lambda)e_n, e_n) \leqslant \frac{1}{2} \cdot (\|T\|^2 - \lambda)$$

for all $n$, or $\|Te_n\|^2 \leqslant (\|T\|^2 + \lambda)/2 < \|T\|^2$, which implies $\|T\|_{c,\mathscr{E}} < \|T\|$. ▨

The following corollary is immediate.

4.10. COROLLARY. *Suppose* $\|T\| = 1$. *Then* $\|T\|_{c,\mathscr{E}} = \|T\|_{r,\mathscr{E}} = 1$ *for every basis if and only if T is essentially unitary.* ▨

Let $\|T\|_s = \sup_{\mathscr{E}} \|T\|_{s,\mathscr{E}}$. Then $\|T\|_s$ is a norm on $\mathscr{B}(\mathscr{H})$ with $\|T\|_s \leqslant \|T\|$ for all $T$. The question arises: is $\|T\|_s = \|T\|$ for all $T$?

4.11. EXAMPLE. Let $\mathscr{H}$ have on orthonormal basis $\{e_n\}_{n=1}^{\infty}$ and define $T$ by $Te_1 = Te_2 = e_1$ and $Te_j = 0$ for $j \geqslant 3$. Then $\|T\|_s \leqslant \sqrt{7}/2$, so that $\|T\|_s < \|T\| = \sqrt{2}$.

*Proof.* Since $T$ has rank 1, $\|T\|_{s,\mathscr{E}} = \|T\|_{e,\mathscr{E}}$ for every basis, by Proposition 2.5. Thus the norms $\|T\|_s$ and $\|T\|_e$ are equal. To show that $\|T\|_e \leqslant \sqrt{7}/2$ we must verify that neither of the two numbers

$$p = \sup\{|(Tx, x)| : \|x\| = 1\},$$

and

$$q = \sup\{|(Tx_1, x_2)| : \|x_1\| = \|x_2\| = 1, (x_1, x_2) = 0\}$$

exceeds $\sqrt{7}/2$. But $p$ is the numerical radius of $T$, which is also seen to be that of the matrix

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

This is not hard to compute: $p = (1 + \sqrt{2})/2$.

To complete the proof let $x_1$ and $x_2$ be any unit vectors with $(x_1, x_2) = 0$; we shall show that $|(Tx_1, x_2)| \leqslant \sqrt{7}/2$. We may assume, of course, that $Tx_1 \neq 0$.

We may further assume that $Tx_1$ is a linear combination of $x_1$ and $x_2$. Otherwise, $Tx_1 = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$ with $x_3$ a unit vector orthogonal to both $x_1$ and $x_2$ and with $\alpha_3 \neq 0$. Now normalize the vector $\alpha_2 x_2 + \alpha_3 x_3$ to get a unit vector $y$ orthogonal to $x_1$. Then

$$|(Tx_1, y)| = (|\alpha_2|^2 + |\alpha_3|^2)^{1/2} > |\alpha_2| = |(Tx_1, x_2)|,$$

so that $x_2$ may be replaced by $y$ to make $Tx_1$ a linear combination of $x_1$ and $x_2$.

We extend $\{x_1, x_2\}$ to an orthonormal basis $\{x_n\}_{n=1}^{\infty}$ for $\mathscr{H}$. Since $T$ has rank 1, $Tx_n$ is a multiple of $Tx_1$ for every $n$, so that

$$1 = \operatorname{tr} T = \sum (Tx_n, x_n) = (Tx_1, x_1) + (Tx_2, x_2).$$

Thus at least one of the two numbers $(Tx_1, x_1)$ and $(Tx_2, x_2)$ has absolute value $\geqslant 1/2$. Now

$$|(Tx_1, x_2)|^2 + |(Tx_1, x_1)|^2 + |(Tx_2, x_2)|^2 \leqslant \sum |(Tx_i, x_j)|^2 \leqslant \|T\|_2^2 = 2$$

implies that $|(Tx_1, x_2)|^2 \leqslant 2 - 1/4$ or $|(Tx_1, x_2)| \leqslant \sqrt{7}/2$.

## 5. REMARKS AND UNSOLVED PROBLEMS

The necessary and sufficient condition that all the Schur norms of an operator be equal to the operator norm is very different in the finite-dimensional (Theorem 3.5) and infinite-dimensional (Theorem 4.8) cases. A related question that we have been unable to answer in either case is: if all the Schur norms of an operator are equal to each other, must they be equal to the operator norm? If not, what are necessary and sufficient conditions that all Schur norms of an operator be equal? Similar questions can be asked for the other matrix norms.

It is obvious that $\sup_{\mathscr{E}}\{\|T\|_{c,\mathscr{E}}\}$ and $\sup_{\mathscr{E}}\{\|T\|_{r,\mathscr{E}}\}$ are both equal to the operator norm.

We have been unable to compute $|||T|||$ in general for $|||T|||$ the supremum, over all bases $\mathscr{E}$, of any other matrix norm of $[T]_{\mathscr{E}}$. Example 4.11 shows that $|||\cdot|||$ is not the operator norm in the Schur and maximal-entry cases. The norms $|||\cdot|||$ are unitarily invariant in the sense that $|||U^{-1}TU||| = |||T|||$ for all unitary $U$. Note that $|||T||| \geqslant (1/2)\|T\|$ for each $|||\cdot|||$. This follows since the numerical radius of $T$ is at least $(1/2)\|T\|$, and if $(Tf, f)$ is within $\varepsilon$ of the numerical radius of $T$, then $\|T\|_{e,\mathscr{E}}$ is at least the numerical radius of $T$ minus $\varepsilon$ whenever $\mathscr{E}$ contains $f$. Thus $(1/2)\|T\| \leqslant |||T||| \leqslant \|T\|$ for each $|||\cdot|||$, and each $|||\cdot|||$ is equivalent to the operator norm. Can any of the $|||\cdot|||$ be computed in terms of other invariants of the operator? Can the $|||\cdot|||$ be used along with other invariants to give a complete set of unitary invariants for some classes of operators?

## REFERENCES


1. ANDERSON, J., *Deviations, commutators and the essential numerical range*, Ph.D. Thesis, Indiana University, 1971.
2. BONSALL, F. F.; DUNCAN J., *Numerical ranges. II*, London Math. Soc. Lecture Note Series, Vol. **10**, Cambridge, 1973.
3. FILLMORE, P. A., On similarity and the diagonal of a matrix, *Amer. Math. Monthly*, **76**(1969), 167–169.
4. FILLMORE, P. A.; STAMPFLI, J. G.; WILLIAMS, J. P., On the essential numerical range, the essential spectrum and a problem of Halmos, *Acta Sci. Math. (Szeged)*, **33**(1972), 179–192.
5. GOHBERG, I. C.; KREIN, M. G., *Introduction to the theory of linear nonselfadjoint operators*, Transl. Amer. Math. Soc., **18**, 1969.
6. HALMOS, P. R., *Finite-dimensional vector spaces*, D. Van Nostrand, Princeton, 1958.
7. ONG, S. C., On the Schur multiplier norm of matrices, *Linear Algebra Appl.*, **56**(1984), 45–55.
8. RADJAVI, H.; ROSENTHAL, P., Matrices for operators and generators of $\mathscr{B}(\mathscr{H})$, *J. London Math. Soc. (2)*, **2**(1970), 557–560.
9. SCHUR, I., Bemerkungen zur Theorie der beschrankten Bilinearformen mit unendlich vielen Verander-lichen, *J. Reine Angew. Math.*, **140**(1911), 1–28.
10. STOUT, Q. F., Schur products of operators and the essential numerical range, *Trans. Amer. Math. Soc.*, **264**(1981), 39–47.
11. WALTER, M. E., On the norm of a Schur product, to appear.



*C. K. FONG*
*Department of Mathematics,*
*University of Ottawa,*
*Ottawa, Ontario,*
*Canada.*

*HEYDAR RADJAVI*
*Department of Mathematics,*
*Dalhousie University,*
*Halifax, Nova Scotia,*
*Canada.*

*PETER ROSENTHAL*
*Department of Mathematics,*
*University of Toronto,*
*Toronto, Ontario,*
*Canada.*


Received April 20, 1986.