

## An Analogue of a Hardy-Littlewood-Fejér Inequality for Upper Triangular Trace Class Operators

Allen L. Shields\*

Mathematics Department, University of Michigan, Ann Arbor Michigan 48109, U.S.A.

Let  $H$  denote a separable, infinite-dimensional Hilbert space. We shall consider operators (that is, bounded linear transformations) on  $H$ . An operator  $T$  is said to be in the Schatten class  $C_p$  if  $T$  is compact and if the eigenvalues of  $(T^*T)^{1/2}$  are in  $l^p$  ( $0 < p < \infty$ ); the norm in  $C_p$  is the  $l^p$  norm of these eigenvalues. The class  $C_2$  is called the Hilbert-Schmidt class, and  $C_1$  is called the trace class. By  $C_\infty$  we denote the class of all compact operators. For information on these classes see, for example, [23], [4, Chap. XI], [8, Chap. III], [19], [1].

Although there is an obvious analogy with the sequence spaces  $l^p$ , there is a deeper analogy with the function spaces  $L^p$  on the unit circle (however, the containment relations are reversed from the  $L^p$  case, namely,  $C_p \subset C_r$  for  $p < r$ ). This analogy serves as a source for new conjectures in operator theory. Motivated by this we establish an operator analogue of a coefficient inequality for the function space  $H^1$ . First we survey some known results in order to emphasize the nature and depth of the analogy.

Fix an orthonormal basis  $\{e_n\}$  ( $n=1, 2, \dots$ ) in  $H$ , and let  $E_n = \text{span}\{e_1, \dots, e_n\}$ . Each operator  $T$  on  $H$  has a matrix representation:

$$T \sim \{\hat{T}(j, k)\} \quad (j, k = 1, 2, \dots)$$

where  $\hat{T}(j, k) = (Te_k, e_j)$ . We shall say that  $T$  has the upper triangular form if  $\hat{T}(j, k) = 0$  for  $j > k$ . This is equivalent to saying that each of the subspaces  $E_n$  is invariant for  $T$ :  $TE_n \subset E_n$  ( $n \geq 1$ ). It can be shown that if  $T$  has the upper triangular form and is compact, then the spectrum of  $T$  is just the set of matrix entries on the main diagonal, together with 0. If  $T$  is not compact this need not be true as shown, for example, by the backward shift operator.

We define the upper triangular projection,  $P_u$ , on the vector space of all infinite matrices  $A(j, k)$  ( $j, k \geq 1$ ) as follows:

$$(P_u A)(j, k) = \begin{cases} A(j, k) & \text{if } j \leq k \\ 0 & \text{otherwise.} \end{cases}$$

---

\* This research was supported in part by the National Science Foundation  
1980 Mathematics Subject Classification number: 47B10.

We shall use the same symbol  $P_u$  to denote the restriction of this operator to various subspaces of matrices, for example to  $C_p$  (strictly speaking we should write  $\tilde{C}_p$ ).

We view the projection  $P_u$  as the analogue of the “analytic projection”  $P_a$ , defined on formal trigonometric series as follows:

$$P_a \sum_{-\infty}^{\infty} a_n e^{int} = \sum_0^{\infty} a_n e^{int}.$$

Along with the analytic projection one has the conjugate function operator in harmonic analysis, defined on formal series by the formula

$$\sum = \sum_{-\infty}^{\infty} a_n e^{int} \mapsto \tilde{\sum} = \sum_{-\infty}^{\infty} (-i)(\operatorname{sgn} n) a_n e^{int}.$$

where  $\operatorname{sgn} n$  is equal to 0 for  $n=0$ , +1 for  $n>0$ , and  $-1$  for  $n<0$ . Then  $\sum + i\tilde{\sum}$  is a formal power series. There is a simple relationship between this operator and the analytic projection operator:

$$\sum + i\tilde{\sum} + a_0 = 2P_a(\sum).$$

In particular, one is bounded (say on  $L^p$ ) if and only if the other is bounded. For matrices one can define an analogous operation,  $A \mapsto \tilde{A}$ , by multiplying the entries above the main diagonal by  $-i$ , those below by  $i$ , and those on the diagonal by 0. Again there is a simple relationship between this operator and the upper triangular projection (this relationship involves also the projection onto the main diagonal). And this operator is bounded (say on  $C_p$ ) if and only if  $P_u$  is bounded.

One can form the complex conjugate of a function or of a formal trigonometric series. The matrix analogue is to take the conjugate transpose:  $A^*(j, k) = \overline{A(k, j)}$ . The matrix  $A$  is called *Hermitian* (or “real”) if  $A = A^*$ ; every matrix has a unique decomposition,  $A = \operatorname{Re} A + i(\operatorname{Im} A)$ , where  $\operatorname{Re} A$  and  $\operatorname{Im} A$  are Hermitian. If  $A$  is an upper triangular matrix with real diagonal entries, then  $(\operatorname{Re} A)^{\sim} = \operatorname{Im} A$ .

We now list three further constructs used in harmonic analysis, together with their matrix analogues. Here  $f, g$  are trigonometric polynomials, and  $A, B$  are matrices with only finitely many non-zero entries. Then

$$\begin{aligned} fg &\leftrightarrow AB, \\ \int f\bar{g} &\leftrightarrow \operatorname{tr}(B^*A) = \operatorname{tr}(AB^*), \\ \text{convolution: } f * g &\leftrightarrow \text{Schur product: } A * B. \end{aligned}$$

Here  $\operatorname{tr}$  denotes the trace:  $\operatorname{tr} A = \sum A(i, i)$ . And the Schur product of  $A, B$  is the matrix with entries  $A(i, j)B(i, j)$ . This product has also been called the “Hadamard product” – see, for example, [10], p. 144 in the first edition, and § 85,

p. 174 in the second edition of the book. Prof. Halmos informs me that this name was suggested in an off-hand comment by von Neumann. (Perhaps he used this name because of the analogy with the Hadamard product of power series.) For more on this product see [3], and for a historical survey see [25].

These concepts are developed in greater generality in [9]. The Banach space aspects of the analogy with  $L^p$  are investigated in [1]. The following results are (more or less) analogous to well-known results in harmonic analysis.

1. If  $1 < p \leq \infty$ , if  $A \in C_p$  and  $B \in C_q$  where  $p^{-1} + q^{-1} = 1$ , then  $AB \in C_1$  and  $\|AB\|_1 \leq \|A\|_p \|B\|_q$ . Also, the trace functional extends to  $C_1$  with  $|\text{tr}(T)| \leq \|T\|_1$ . Finally,  $C_q$  may be isometrically identified with the dual space of  $C_p$  by means of the trace functional:  $\text{tr}(AB)$ . (See [8, Chap. III, §7.2, and also Theorem 12.3].)

2. If  $A$  is an  $n \times n$  matrix, then

$$c \log n \leq \|P_u A\| \leq \log_2(2n) \|A\|.$$

(See [16], Propositions 1.1 and 1.2, where the results are also shown to hold for other norms, in particular for the  $C_1$  norm.)

3. The upper triangular projection  $P_u$  is a bounded operator on  $C_p$  ( $1 < p < \infty$ ), but is not bounded on  $C_1$  or on  $C_\infty$  (see [16], Corollaries 1.2 and 1.3 and the related discussion).

4. If  $T$  is compact and quasi-nilpotent (that is,  $T$  is a Volterra operator in the terminology of Gohberg and Krein [8, 9]), and if  $\text{Im } T \in C_p$  for some  $p$ ,  $1 < p < \infty$ , then  $\text{Re } T \in C_p$ . This is a result of Macaev [18] (see also [9, Theorem III.6.2]). If  $\text{Im } T \in C_1$ , then  $\text{Re } T \in C_p$  for all  $p > 1$  (see [9, Theorem III.2.1] where a more precise result is obtained). In a similar vein we have the following result of W. Kahan [15]: If  $T$  is an  $n \times n$  matrix with real spectrum then

$$\|\text{Im } T\| \leq (\log_2 n + 0.038) \|\text{Re } T\|.$$

5. As an analogue of the Hausdorff-Young inequality we have the following (when  $p = \infty$  the norm on the left side is the  $l^\infty$  norm):

$$\begin{aligned} \left\{ \sum \|Te_n\|^p \right\}^{1/p} &\leq \|T\|_p & (2 \leq p \leq \infty), \\ \|T\|_p &\leq \left\{ \sum \|Te_n\|^p \right\}^{1/p} & (1 \leq p \leq 2). \end{aligned}$$

The first of these is a consequence of the interpolation theorem for linear operators on  $C_p$  spaces. (See [9, Chap. III, Theorem 5.2.]) The second follows by duality. (See [9, Chap. III, Theorem 3.3]. See also [4, Chap. XI, Lemma 9.32], where the result is incorrectly stated to hold in the range  $2 \leq p \leq \infty$ ; their proof works for  $0 < p \leq 2$ , with an additional constant factor on the right side. Thus the result is valid also for  $p < 1$ .)

The Hardy-Littlewood-Fejér inequality referred to in our title says that

$$\sum_{n=0}^{\infty} \frac{|\hat{f}(n)|}{n+1} \leq \pi \|f\|_1 \quad (f \in H^1),$$

with equality only for the zero function. Here  $H^1$  denotes the subspace of  $L^1$  consisting of those functions  $f$  for which the Fourier coefficients of negative index vanish:  $\hat{f}(n)=0$  ( $n<0$ ). This inequality is sometimes called “Hardy’s inequality” (see, for example, [14, Chap. 5, p. 70], or [7, Chap. II, Prob. 8, p. 93], or [5, §3.6, p. 48 – but see the notes at the end of the chapter on p. 52]). This is not correct. The inequality seems to have appeared for the first time in 1927 in separate papers in the *Mathematische Annalen*, by Fejér [6, Satz VI, p. 122–123, and equations (48), (49), p. 117], and by Hardy and Littlewood [12, p. 163, also Theorem 15, p. 206, and Theorem 16, p. 208]. Fejér obtains the precise constant  $\pi$  (he states the result in an equivalent form involving derivatives); Hardy and Littlewood prove a more general coefficient inequality for functions of class  $H^p$ ,  $0<p\leq 2$ , but when  $p=1$  they do not obtain the constant  $\pi$ . Zygmund [27, Notes to Chap. VII, Theorem 8.6, 8.7, p. 382], and Hardy, Littlewood and Polya [13, §9.7, pp. 236–237] refer to both papers when discussing this inequality.

Several different proofs may be found in the references above (and also in [7, Chap. VI, Prob. 8, p. 273]). One proof uses the fact that each  $H^1$  function of norm one is the product of two  $H^2$  functions of norm one to deduce the inequality from the bound for the (first) Hilbert matrix. This is the matrix with entries  $(1+n+m)^{-1}$  ( $n, m=0, 1, \dots$ ). This matrix represents a bounded operator on  $l^2$  with bound  $\pi$ , and this bound is not attained. (See, for example, [5, the Corollary to Theorem 3.14, p. 48], or [13, §8.12, Theorem 294, p. 212, and Chap. IX where an extensive survey of the older literature is given, and Appendix III]; a very simple proof of a generalization due to Schur is given in [21], and further generalizations are in [22]. A different proof, based on the so-called Schur test, is given in [11, Problems 37 and 38, pp. 22, 23].)

Our proof for the matrix analogue of this inequality follows the lines of the proof referred to above. However, we need the second Hilbert matrix rather than the first. This is the matrix with entries  $(n-m)^{-1}$  when  $n\neq m$ , and 0 when  $n=m$  ( $n, m=1, 2, \dots$ ). This matrix is also a bounded operator on  $l^2$  with bound  $\pi$ , which is not attained (see [21] for a very simple proof; further references are given below). We outline a simple proof that both Hilbert matrices are bounded by  $\pi$ . This proof has been known for some time; Sheldon Axler pointed out to us that one can use it to obtain a stronger result, due to Schur.

The proof begins by considering the  $L^\infty$  function  $\varphi$  with Fourier coefficients:  $\hat{\varphi}(0)=0$ ,  $\hat{\varphi}(n)=-n^{-1}$  ( $n\neq 0$ ). (This function has been used before: see [6, the footnote on p. 117], or [27, the proof of Theorem 8.7 in Chap. VII, pp. 286–287].) This is the function  $\varphi=i\psi$ , where

$$\psi(t)=\begin{cases} t+\pi, & -\pi\leq x<0, \\ t-\pi, & 0<x\leq\pi. \end{cases}$$

Then  $\|\varphi\|_\infty=\pi$ , but  $|\varphi|<\pi$  almost everywhere. Next one introduces the operator  $M_\varphi$  of multiplication by  $\varphi$  on  $L^2$ . The norm of this operator is  $\pi$  but it is not attained:  $\|M_\varphi f\|<\|f\|$ , for  $f\neq 0$ .

If we write  $L^2 = H^2 \oplus (H^2)^\perp$ , then  $M_\varphi$  can be written as a  $2 \times 2$  operator matrix

$$M_\varphi = \begin{pmatrix} A & X \\ B & Y \end{pmatrix}. \tag{1}$$

Here  $A$  maps  $H^2$  into itself and  $B$  maps  $H^2$  into  $(H^2)^\perp$ . Since  $M_\varphi$  is bounded by  $\pi$  which is not attained, the same must be true for both  $A$  and  $B$  (and for  $X$  and  $Y$ ). If we use the basis  $\{(2\pi)^{-1/2} \exp(int)\}$  ( $-\infty < n < \infty$ ) in  $L^2$ , then  $A$  is represented by the second Hilbert matrix and  $B$  by the first.

Schur proved the stronger result that  $A^*A + B^*B \leq \pi^2$  (see [24, § 5, pp. 16, 17], or [13, § 8.12, Theorem 294, pp. 212, 213]). This follows from the above considerations as observed by Axler. Namely,  $M_\varphi^* M_\varphi$  is bounded by  $\pi^2$ , and from (1) we see that the upper left entry in the  $2 \times 2$  matrix for this operator is  $A^*A + B^*B$ .

We now state our analogue of the inequality of Hardy, Littlewood and Fejér.

**Theorem 1.** *Let  $T \in C_1$  have the upper triangular form with respect to the orthonormal basis  $\{e_n\}$  ( $n = 1, 2, \dots$ ). Then*

$$\sum_{k=1}^{\infty} \sum_{j=1}^k \frac{|\hat{T}(j, k)|}{1+k-j} \leq \pi \|T\|_1,$$

with equality only when  $T = 0$ .

This result was obtained in 1972. We wish to thank D. Hadwin, P.R. Halmos, and E. Nordgren for helpful discussions of this material. And we would especially like to thank W. Arveson who pointed out the crucial importance of a result like Lemma 2.

We require three lemmas. Peter Weinberger pointed out that the second lemma, in the special case when the operator  $P$  is an invertible  $n \times n$  matrix, is really a standard fact from numerical analysis (it is called the Cholesky decomposition, see the comments following the proof). Throughout our discussion the orthonormal basis  $\{e_n\}$ ,  $n = 1, 2, \dots$  will be fixed; upper triangularity will always be with respect to this basis.

In general, operator multiplication is not weakly continuous; for example, if  $U$  is the unilateral shift then  $U^n \rightarrow 0$  weakly and  $U^{*n} \rightarrow 0$  strongly, but  $U^{*n} U^n = I$  for all  $n$ . The product is weakly continuous, however, if the operators on the right have the upper triangular form. We only require the case  $p = 2$  of the following result.

**Lemma 1.** *Let  $R$  denote either the space of all operators on  $H$ , with the weak operator topology, or any of the Banach spaces  $C_p$  ( $1 < p < \infty$ ) with its weak topology. If  $\{A_n\}$ ,  $\{B_n\} \subset R$ , with  $A_n \rightarrow A$  and  $B_n \rightarrow B$  weakly, and if each  $B_n$  has the upper triangular form, then  $A_n B_n \rightarrow AB$  weakly.*

*Proof.* We write our operators as matrices. One verifies that if  $\{A_n\} \subset R$  then  $A_n \rightarrow A$  weakly if and only if  $\{\|A_n\|_R\}$  is a bounded sequence, and  $\hat{A}_n(i, j) \rightarrow \hat{A}(i, j)$  for all  $i, j$ . Thus to complete the proof we must show that a)  $\|A_n B_n\|_R$  are

bounded, and b)  $d_n(i, j) \rightarrow d(i, j)$  for all  $i, j$ . Here (for typographical reasons) we let  $d_n$  and  $d$  denote the matrix entries of  $A_n B_n$  and of  $AB$ .

As to the first point we recall that in  $C_p$  we have

$$\|A_n B_n\|_p \leq \|A_n B_n\|_{p/2} \leq \|A_n\|_p \|B_n\|_p$$

(see [8, Chap. III, (7.4) and (7.5)]). Thus  $\{\|A_n B_n\|_R\}$  is a bounded sequence.

For the second point we note that

$$d_n(i, j) = \sum_{k=1}^j \widehat{A}_n(i, k) \widehat{B}_n(k, j)$$

since  $B_n$  is upper-triangular. A similar equation holds for  $d(i, j)$ , and thus, for each fixed choice of  $i, j$ ,  $d_n(i, j) \rightarrow d(i, j)$ .

**Lemma 2.** *Let  $P$  be a positive semi-definite operator in  $C_1$ , with  $\|P\|_1 = 1$ . Then there exists  $B \in C_2$  with  $\|B\|_2 = 1$  such that  $B$  has the upper triangular form and  $P = B^* B$ .*

*Proof.* We first prove the lemma under the additional assumption that  $P$  is one-to-one on each of the subspaces  $E_n$  ( $n \geq 1$ ). Then  $P^{1/2}$  is also one-to-one on each of these spaces. Let  $F_n = P^{1/2} E_n$ . Then  $F_1 \subset F_2 \subset \dots$ , and  $\dim F_n = n$  for all  $n$ . Hence there is an orthonormal set  $\{f_k\}$  such that  $F_n = \text{span}\{f_1, \dots, f_n\}$ . Define an operator  $V$  by:  $Vf_n = e_n$  ( $n \geq 1$ ), and  $V = 0$  on the orthogonal complement of the span of  $\{f_n\}$ . Then  $V$  is a partial isometry. Let  $B = VP^{1/2}$ . Then  $BE_n \subset E_n$  and so  $B$  has the upper triangular form. Also,  $B^* B = P^{1/2} V^* VP^{1/2} = P$ , since  $V^* V$  is the projection onto the span of  $\{f_n\}$ , which contains the range of  $P^{1/2}$ . Finally,  $\|B\|_2 \leq \|V\| \|P^{1/2}\|_2 = 1$ , and therefore  $1 = \|P\|_1 \leq \|B^*\|_2 \|B\|_2 \leq 1$ , which completes the proof in this case.

Now suppose that  $P$  is not one-to-one. Let  $S$  be a fixed positive trace-class operator with trivial kernel (for example, a diagonal matrix with strictly positive diagonal entries in  $l^1$ ). Let  $P_n = (P + n^{-1} S) d_n$ , where  $d_n = \|P + n^{-1} S\|_1^{-1}$ . Then  $P_n$  has norm one, and  $P_n \rightarrow P$ , in  $C_1$ . By the result proved above, there is a sequence  $\{B_n\}$  of operators in  $C_2$ , having the upper triangular form, with  $P_n = B_n^* B_n$ ,  $\|B_n\|_2 = 1$ , for all  $n$ . By passing to a subsequence we may assume that  $\{B_n\}$  is weakly convergent in  $C_2$ :  $B_n \rightarrow B$  for some  $B$  in the unit ball of  $C_2$ . The limit operator  $B$  must have the upper triangular form and, by Lemma 1, we have  $B^* B = P$ . From this we have  $\|B\|_2 \geq 1$  and hence the norm must equal unity. This completes the proof.

*Remark.* The representation  $P = B^* B$ , with  $B$  upper triangular, is known in numerical analysis as the Cholesky decomposition of  $P$  (here  $P$  is a strictly positive  $n \times n$  matrix). See, for example, [18, Chap. 9, Exercise 1.6, p. 174], or [26, Chap. 4, §42, p. 229]. One can arrange, in addition, that  $B$  has non-negative diagonal entries (this is done by multiplying the matrix  $B$  found in Lemma 2 on the left by a suitable diagonal unitary matrix);  $B$  is then uniquely determined by  $P$ .

**Lemma 3.** *Let  $T \in C_1$  have the upper triangular form, with  $\|T\|_1 = 1$ . Then there exist upper triangular operators  $A, B \in C_2$  with  $T = AB$  and  $\|A\|_2 = \|B\|_2 = 1$ .*

*Proof.* We first prove the lemma with the additional assumption that  $T$  is one-to-one on each of the spaces  $E_n$ ; this is equivalent to requiring that all diagonal matrix entries are different from 0:  $(Te_j, e_j) \neq 0$  for all  $j$ .

Let  $T = UP$  be the polar decomposition of  $T$ ; then  $P = (T^*T)^{1/2}$  is a positive operator of norm one in  $C_1$ , and  $U$  maps the range of  $P$  isometrically onto the range of  $T$ . Since  $\|Pf\| = \|Tf\|$  for all  $f$  we see that  $P$  has the same kernel as  $T$ , therefore  $P$  is one-to-one on each of the spaces  $E_n$ . By Lemma 2,  $P = B^*B$  where  $B$  is an upper triangular operator of norm one in  $C_2$ . We see that  $B$  must be one-to-one on each of the spaces  $E_n$ . Now let  $A = UB^*$ . Then  $A$  is in the unit ball of  $C_2$ , and  $AB = T$ . From this we see that  $\|A\|_2 = 1$ . To show that  $A$  has the upper triangular form we must show that it maps each space  $E_n$  into itself. Since  $B$  is one-to-one on  $E_n$  and  $E_n$  is finite dimensional we have  $E_n = BE_n$ . Hence

$$AE_n = ABE_n = TE_n = E_n.$$

Now suppose that  $T$  is not one-to-one on each of the spaces  $E_n$ , that is, some diagonal matrix entries are 0. Let  $S$  be a diagonal matrix  $C_1$  operator, with non-zero diagonal entries precisely in those places where  $T$  has a zero. Let  $T_n = (T + n^{-1}S)d_n$ , where  $d_n = \|T + n^{-1}S\|_1^{-1}$ . Then  $T_n$  satisfies all the conditions of the lemma, and in addition is one-to-one on each of the spaces  $E_n$ . By what was proved above there are upper triangular operators  $A_n, B_n$  in the unit ball of  $C_2$  with  $T_n = A_nB_n$ . By passing to a subsequence we may assume that the sequences  $\{A_n\}$  and  $\{B_n\}$  are weakly convergent in  $C_2$ :  $A_n \rightarrow A, B_n \rightarrow B$ , where  $A, B \in \text{Ball}(C_2)$ . By Lemma 1 we have  $A_nB_n \rightarrow AB$ , and so  $T = AB$ . This completes the proof since weak convergence preserves the upper triangular form.

*Remarks.* 1. Is it really necessary to break the proof into two parts? The construction given in the first part should work in general, even when  $T$  is not one-to-one; the difficulty is to prove that  $A$  has the upper triangular form.

2. If the upper triangular form were not involved the lemma would follow immediately from the polar decomposition. Indeed, if  $\|T\|_1 = 1$  then  $T = (UP)^{1/2}P^{1/2}$  factors  $T$  into the product of two  $C_2$  operators of norm one. This is analogous to the situation for functions: it is easy to write an  $L^1$  function of norm one as the product of two  $L^2$  functions of norm one, but more difficult to do this for power series, that is, for  $H^1$  and  $H^2$ .

*Proof of Theorem.* Without loss of generality we may assume that  $\|T\|_1 = 1$ . Then by Lemma 3 there are upper triangular operators  $A, B$  of norm one in  $C_2$  such that  $T = AB$ . Let  $t_{ij}, a_{ij}, b_{ij}$  denote the matrix entries of  $\hat{T}, \hat{A}, \hat{B}$  respectively. The following summations are written with each variable going from 1 to  $\infty$ . Because of the upper triangularity, however, the terms are equal to zero if  $j > k$ , or if  $j > r$ , or if  $r > k$ . Thus we really have  $1 \leq j \leq r \leq k < \infty$ . We use the boundedness of the second Hilbert matrix (which was discussed ear-

lier); then we use the Cauchy(-Buniakovsky-Schwarz) inequality. We have:

$$\begin{aligned} \sum_{j,k} \frac{|t_{jk}|}{1+k-j} &\leq \sum_{j,k} \sum_r \frac{|a_{jr}b_{rk}|}{1+k-j} = \sum_r \sum_{j,k} \frac{|a_{jr}b_{rk}|}{1+k-j} \\ &\leq \pi \sum_r (\sum_j |a_{jr}|^2)^{1/2} (\sum_k |b_{rk}|^2)^{1/2} \\ &\leq \pi (\sum_r \sum_j |a_{jr}|^2)^{1/2} (\sum_r \sum_k |b_{rk}|^2)^{1/2} = \pi. \end{aligned}$$

We have strict inequality because the bound  $\pi$  for the second Hilbert matrix is not attained.

We would like to indicate a somewhat different proof of Theorem 1 (and of the Hardy-Littlewood-Fejér inequality) which establishes a more general result. This result could also be obtained by our previous method if we replace the second Hilbert matrix by an appropriate matrix. (This proof is well known in the function-theoretic case.) We require the following lemma whose proof uses Lemma 3. We state both the function theory and the operator theory versions.

**Lemma 4.** a) *If  $f \in H^1$  then there exists  $f_1 \in H^1$  with non-negative Taylor coefficients, such that  $\|f_1\|_1 \leq \|f\|_1$ , and*

$$|\hat{f}(n)| \leq \hat{f}_1(n) \quad (n=0, 1, \dots).$$

b) *If  $T$  is an upper triangular trace-class operator then there is an upper triangular trace class operator  $T_1$  with non-negative matrix entries, such that  $\|T_1\|_1 \leq \|T\|_1$ , and*

$$|\hat{T}(i,j)| \leq \hat{T}_1(i,j) \quad (1 \leq i, j < \infty).$$

*Proof.* We prove a) since the proof of b) is entirely analogous. We may assume that  $\|f\|_1 = 1$ . Then we have  $f = gh$ , where  $g, h \in \text{Ball}(H^2)$ . Let  $g_1(z) = \sum |\hat{g}(n)| z^n$ , and define  $h_1$  similarly. Then  $g_1, h_1 \in \text{Ball}(H^2)$ , and the function  $f_1 = g_1 h_1$  has the desired properties.

The following result generalizes Theorem 1 (take  $A$  to be the second Hilbert matrix) and the function theoretic inequality (take  $\phi = -i\psi$  where  $\psi$  is the function defined earlier).

**Theorem 2.** a) *Let  $\phi \in L^\infty$  have  $\hat{\phi}(n) \geq 0$  for  $n \geq 0$ . Then*

$$\sum_0^\infty |\hat{f}(n)| \hat{\phi}(n) \leq \|\phi\|_\infty \|f\|_1 \quad (f \in H^1).$$

b) *Let  $A$  be a bounded operator on  $H$ , with  $\hat{A}(i,j) \geq 0$  for  $i \leq j$ . Then for all upper triangular trace class operators  $T$  we have*

$$\sum_{i,j} |\hat{T}(i,j)| \hat{A}(i,j) \leq \|A\| \|T\|_1.$$

*Proof.* We prove b); the proof of a) is similar, with the trace functional replaced by integration. By Lemma 4b) we may assume that  $T$  has non-



negative matrix entries. Then

$$\sum \hat{T}(i, j) \hat{A}(i, j) = \text{tr}(TA^*) \leq \|A\| \|T\|_1.$$

This completes the proof.

Note that the proofs of both theorems required Lemma 3, which is really the key result.

Next we mention some problems suggested by Theorem 1. Our definition of upper triangularity is equivalent to the requirement that the operator in question should map each of the spaces  $E_n$  into itself, that is, the family  $\{E_n\}$  is a chain of invariant subspaces for the operator. (By a “chain” of subspaces we mean a family of subspaces that is linearly ordered by inclusion.) Further, the family  $\{E_n\}$  forms a maximal chain. Suppose now that we replaced this chain by another maximal chain. This leads to a corresponding concept of upper triangularity; in fact most of the discussion in [8] and [9] takes place in this more general context. Is there an analogue of the Hardy-Littlewood-Fejér inequality for such chains? We discuss this problem for two special chains.

First, suppose we have an orthonormal basis  $\{e_n\}$  ( $-\infty < n < \infty$ ); let  $E_n = \text{span}\{e_k\}$  ( $-\infty < k \leq n$ ). Unlike our previous situation, these spaces are not finite-dimensional. If  $T \in C_1$ ,  $\|T\|_1 = 1$ , and if  $TE_n \subset E_n$  for all  $n$ , then do we have  $T = AB$  where  $A, B \in C_2$ ,  $\|A\|_2 = \|B\|_2 = 1$ , and both  $A$  and  $B$  leave each subspace  $E_n$  invariant? If this were true, then just as before we would have

$$\sum_{k=-\infty}^{\infty} \sum_{j=-\infty}^k \frac{|\hat{T}(j, h)|}{1+k-j} \leq \pi \|T\|_1.$$

Arveson [2, Cor. 2, p. 221] has established an analogue of the Cholesky decomposition for this chain: if  $P$  is a positive, invertible operator, then  $P = B^*B$  where  $B$  (and also  $B^{-1}$ ) leaves each of the subspaces  $E_n$  invariant. Of course an invertible operator cannot be in  $C_1$ , so this result is not really the analogue of Lemma 2 for this chain. For our purposes, however, we only need a weaker result. Indeed, Lemma 2 was only used to prove Lemma 3, and here the positive operator  $P$  is the positive part of an upper triangular trace-class operator. Thus we need the analogue of Lemma 2 only for such  $P$ .

As our second example we take the Hilbert space  $L^2(-\infty, \infty)$  with the chain of subspaces

$$E_a = \{f \in L^2 : f = 0 \text{ a.e. on } (a, \infty)\}, \quad -\infty < a < \infty. \tag{2}$$

If  $A$  is a Hilbert-Schmidt operator on  $L^2$  then it has a representation

$$(Af)(x) = \int_{-\infty}^{\infty} K(x, y) f(y) dy \quad (f \in L^2), \tag{3}$$

where  $K \in L^2(\mathbb{R}^2; dx dy)$  (see, for example, [8; Chap. III, 9.3, Theorem 1, p. 141], or [4; Chap. XI, §6, p. 1009]). The following result shows that “upper triangularity” of the kernel  $K$  is equivalent to the invariance of the subspaces  $E_a$ . This result is surely known, but we were unable to find a specific reference.

**Proposition 1.** *Let  $K \in L^2(\mathbb{R}^2)$  and let the operator  $A$  and the chain of subspaces  $\{E_a\}$  be defined by (3) and (2) above. Then the following are equivalent.*

1.  $AE_a \subset E_a$  for all  $a \in \mathbb{R}$ .
2.  $K(x, y) = 0$  a.e. in  $\{(x, y) : x > y\}$ .

*Proof.* It is trivial that 2. implies 1. For the converse we assume, to obtain a contradiction, that the set of  $(x, y)$  with  $x > y$  and  $K(x, y) \neq 0$  has positive two-dimensional measure. Hence there is an  $\alpha > 0$  and a set  $S \subset \mathbb{R}^2$  of positive measure such that for all  $(x, y) \in S$  we have

- (i)  $x > y + \alpha$ ,
- (ii)  $K(x, y) \neq 0$ .

It follows from Lebesgue’s theorem on differentiating an integral that there is a square  $Q$  with sides parallel to the coordinate axes and side length  $< \alpha$ , contained in  $\{(x, y) : x > y + \alpha\}$ , such that  $\iint_Q K \neq 0$ . Then  $Q$  has the form

$$Q = [x_0, x_0 + \delta] \times [y_0, y_0 + \delta] = I \times J$$

for some  $\delta < \alpha$  and some  $(x_0, y_0)$  with  $x_0 > y_0 + \alpha$ .

Let  $f(y) = 1$  for  $y \in J$ , and  $f = 0$  otherwise. In particular,  $f(y) = 0$  for  $y > y_0 + \delta$ . To obtain a contradiction we now show that  $(Af)(x)$  does not vanish almost everywhere in  $y_0 + \delta < x < \infty$ . It will be sufficient to show that  $\int_I Af \neq 0$ . We have

$$\int_I (Af)(x) dx = \int_I \left( \int_I K(x, y) f(y) dy \right) dx = \iint_Q K dx dy \neq 0,$$

which completes the proof.

Suppose one could prove that every trace class operator of trace norm one that leaves all the subspaces  $E_a$  invariant is the product of two Hilbert-Schmidt operators, each of Hilbert-Schmidt norm one, and each leaving the spaces  $E_a$  invariant. Then, just as in the proof of Theorem 1 we would have

$$\int_{-\infty}^{\infty} \int_{-\infty}^y \frac{|K(x, y)|}{y - x} dx dy \leq c \|A\|_1.$$

The proof uses the boundedness of the Hilbert transform operator on  $L^2(\mathbb{R}^2)$  (in place of the boundedness of the Hilbert matrix).

David Larson [17] has shown recently that for the chain of subspaces  $\{E_a\}$  (or more generally, for any complete chain) if  $P$  is any positive invertible operator then we can find an operator  $B$  that leaves invariant all the subspaces in the chain, such that  $P = B^*B$ . The hypothesis that  $P$  is invertible cannot be completely omitted. For example, Larson shows that in  $L^2(0, 1)$  if  $P$  is the projection onto the constant functions, then  $P$  cannot be factored in the form  $B^*B$ , where  $B$  leaves invariant the subspaces of functions vanishing almost everywhere in  $(a, 1)$ , for  $0 < a < 1$ . Of course as remarked earlier we really only need to factor positive operators of the form  $P = (T^*T)^{1/2}$ , where  $T$  is a trace-class operator that leaves invariant the subspaces in the chain. The rank-one projection referred to above does not have this form (see the appendix).

These examples suggest the following problems. Here  $\mathcal{M}$  denotes a chain (not necessarily maximal) of subspaces of Hilbert space;  $P, T, A, B$  denote bounded linear transformations. We say that an operator  $S$  leaves  $\mathcal{M}$  invariant if  $SM \subset M$  for all  $M \in \mathcal{M}$ .

*Question 1.* If  $P = (T^*T)^{1/2}$  where  $T$  leaves  $\mathcal{M}$  invariant, then does there exist  $B$  leaving  $\mathcal{M}$  invariant such that  $P = B^*B$ ?

*Question 2.* If  $T$  leaves  $\mathcal{M}$  invariant and  $T$  is in the unit ball of the trace class, then do we have  $T = AB$ , where  $A$  and  $B$  are in the unit ball of the Hilbert-Schmidt class, and each of them leaves  $\mathcal{M}$  invariant?

These questions seem to be open even when  $\mathcal{M}$  contains just one subspace  $M$ , different from  $\{0\}$  and  $H$ .

**Appendix.** In the space  $L^2 = L^2(0, 1)$  let

$$E_a = \{f \in L^2(0, 1) : f = 0 \text{ a.e. in } (a, 1)\}.$$

If  $f, g \in L^2$ , then by  $f \otimes g$  we denote the rank-one operator defined by

$$(f \otimes g)(h) = (h, g) f.$$

Thus  $P = 1 \otimes 1$  is the (orthogonal) projection operator onto the constant functions.

**Proposition 2.** Let  $S = f \otimes g$  ( $f \neq 0, g \neq 0$ ). Then  $SE_a \subset E_a$  ( $0 < a < 1$ ) if and only if there exists  $\alpha$ ,  $0 < \alpha < 1$ , such that  $f \in E_\alpha$  and  $g \in E_\alpha^\perp$ .

*Proof.* Let  $h_a$  denote the characteristic function of  $[0, a]$ . Then  $Sh_a = (h_a, g) f \in E_a$  for  $0 < a < 1$ . Let  $\alpha = \inf a$ , where  $f \in E_a$ . Then  $\alpha > 0$  and  $f \in E_\alpha$ . Hence for  $a < \alpha$  we must have  $\int_0^a g = (g, h_a) = 0$ ; it follows from Lebesgue's theorem on differentiating an integral that  $g = 0$  almost everywhere in  $[0, \alpha]$ , i.e.,  $g \in E_\alpha^\perp$ .

**Corollary.** If  $P = 1 \otimes 1$ , if  $B$  and  $T$  are operators with  $BE_a \subset E_a$  and  $TE_a \subset E_a$  ( $0 < a < 1$ ), then

- i)  $P \neq B^*B$ ,
- ii)  $P \neq (T^*T)^{1/2}$ .

*Proof.* Of course ii) is a consequence of i) since  $P^2 = P$ . To prove i), assume that  $P = B^*B$ . Since  $B$  must have rank one we have  $B = f \otimes g$  for some  $f, g \in L^2$  with  $f \neq 0, g \neq 0$ ; we may assume that  $\|f\| = 1$ . By Proposition 2 we have  $f \in E_\alpha$ ,  $g \in E_\alpha^\perp$  for some  $\alpha \in (0, 1)$ . A calculation shows that  $B^*B = g \otimes g$ . Thus  $g \otimes g = 1 \otimes 1$  and so  $g \equiv \text{const}$ . Hence  $\alpha = 0$ , which is a contradiction.

**Bibliography**

1. Arazy, J., Lindenstrauss, J.: Some linear topological properties of the spaces  $C_p$  of operators on Hilbert space. *Compositio Math.* **30**, 81-111 (1975)
2. Arveson, W.: Interpolation problems in nest algebras. *J. Functional Anal.* **20**, 208-233 (1975)
3. Bennett, G.: Schur multipliers. *Duke Math. J.* **44**, 603-639 (1977)

4. Dunford, N., Schwartz, J.: Linear operators, Part II: Spectral Theory. New York: Interscience 1963
5. Duren, P.L.: Theory of  $H^p$  spaces. New York-London: Academic Press 1970
6. Fejér, L.: Über gewisse Minimumprobleme der Funktionentheorie. Math. Ann. **97**, 104–123 (1927)
7. Garnett, J.: Bounded Analytic Functions. New York-London: Academic Press 1981
8. Gohberg, I.C., Kreĭn, M.G.: Introduction to the theory of linear non-selfadjoint operators in Hilbert space. Moscow: Izdatel'stvo Nauka 1965 [Russian]; English transl.: Translations of Mathematical Monographs **18**. Providence, Rhode Island: Amer. Math. Soc. 1969
9. Gohberg, I.C., Kreĭn, M.G.: Theory of Volterra operators on Hilbert space and its applications. Moscow: Izdatel'stvo Nauka 1967 [Russian]; English transl. Translations of Mathematical Monographs **24**. Providence, Rhode Island: Amer. Math. Soc. 1970
10. Halmos, P.R.: Finite dimensional vector spaces, second edition. Princeton-Toronto-London: Van Nostrand 1958
11. Halmos, P.R.: A Hilbert space problem book. Princeton-Toronto-London: Van Nostrand 1967
12. Hardy, G.H., Littlewood, J.E.: Some new properties of Fourier constants. Math. Ann. **97**, 159–209 (1926)
13. Hardy, G.H., Littlewood, J.E., Polya, G.: Inequalities. second edition, Cambridge University Press, 1954
14. Hoffman, K.: Banach spaces of analytic functions. Englewood Cliffs, New Jersey: Prentice-Hall, 1962
15. Kahan, W.: Every  $n \times n$  matrix  $Z$  with real spectrum satisfies  $\|Z - Z^*\| \leq \|Z + Z^*\| (\log_2 n + 0.038)$ . Proc. Amer. Math. Soc. **39**, 235–241 (1973)
16. Kwapień, S., Pełczyński, A.: The main triangle projection in matrix spaces and its applications. Studia Math. **34**, 43–68 (1970)
17. Larson, D.: Nest algebras and similarity transformations. Preprint
18. Macaev, V.I.: Volterra operators obtained from self-adjoint operators by perturbation. Dokl. Akad. Nauk SSSR **139**, 810–813 (1961) [Russian]; English transl. Soviet Math. Dokl. **2**, 1013–1016 (1961)
19. McCarthy, C.A.:  $C_p$ . Israel J. Math. **5**, 249–271 (1967)
20. Ortega, J.M.: Numerical analysis, a second course. New York-London: Academic Press 1972
21. Redheffer, R., Volkmann, P.: On Schur's generalization of Hilbert's inequality. Preprint
22. Redheffer, R., Volkmann, P.: Über eine Ungleichung von Schur. Preprint
23. Schatten, R.: Norm ideals of completely continuous operators. Berlin-Heidelberg-New York: Springer 1960
24. Schur, I.: Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen. J. Reine Angew. Math. **110**, 1–28 (1911)
25. Styan, G.P.H.: Hadamard products and multivariate statistical analysis. Linear Algebra Appl. **6**, 217–240 (1973)
26. Wilkinson, J.H.: The algebraic eigenvalue problem. Oxford: Clarendon Press 1965
27. Zygmund, A.: Trigonometric series, second edition. Cambridge: Cambridge University Press 1959

Received June 14, 1982