J. OPERATOR THEORY
5(1981), 3–16

THE FUGLEDE COMMUTATIVITY THEOREM MODULO THE HILBERT-SCHMIDT CLASS AND GENERATING FUNCTIONS FOR MATRIX OPERATORS. II

GARY WEISS

Let $\mathscr{H}$ denote a separable, complex Hilbert space and let $\mathscr{L}(\mathscr{H})$ denote the class of all bounded linear operators acting on $\mathscr{H}$. Let $\mathscr{K}(\mathscr{H})$ denote the class of compact operators in $\mathscr{L}(\mathscr{H})$ and let C_p denote the Schatten p-class $(0 < p < \infty)$ with $\|\cdot\|_p$ $(1 \leqslant p < \infty)$ denoting the associated p-norm. Hence C_2 is the Hilbert-Schmidt class and C_1 is the trace class.

In [5] we pointed out connections between a problem of I. D. Berg [1], namely "Is every normal operator the sum of a diagonalizable operator and a Hilbert-Schmidt operator?", and several statements regarding normal operators, Hilbert-Schmidt operators and trace class operators. Some of these statements were proven and some were left open questions. Here we settle the main question [5, statement (3)] and obtain a generalization, and we ask several new questions.

THEOREM 1. *If N_1, N_2 are normal operators and X is a bounded operator then*

$$\|N_1X - XN_2\|_{C_2} = \|N_1^*X - XN_2^*\|_{C_2}.$$

In particular, $N_1X - XN_2 \in C_2$ implies $N_1^*X - XN_2^* \in C_2$.

We give a proof of this theorem which blends two earlier proofs. The first proof used generating functions and a kind of distribution theory. The second proof was entirely operator theoretic. The first proof was the original proof and suggests certain methods and generalizations. The second proof was a more recent proof that the author constructed from the first proof at the urging of Dan Voiculescu. It was felt that an operator theoretic proof was important.

In [5, Theorem 2c] we proved that to prove Theorem 1, it suffices to assume $N_1 = N_2 = M_\varphi$ (the operator of multiplication by φ) where $\varphi \in L^\infty(\mathbf{T}), M_\varphi$ acts on $L^2(\mathbf{T})$, and for every complex number c, $m\{z : \varphi(z) = c\} = 0$.

Proof. In [5, The Main Construction] we defined the generating function for the matrix operator $X = (x_{ij})$ to be the formal Fourier series $F(z, w) =$

$= \sum_{i,j=-\infty}^{\infty} x_{ij}z^i w^j$. In other words, the entries of the matrix operator are precisely the Fourier coefficients of its corresponding generating function. In addition, when $\varphi(z) = \sum_{n=-\infty}^{\infty} \varphi_n z^n$ is the Fourier series for $\varphi(z)$, we defined the formal product $\varphi(z)*F(z,w)$ in the canonical way to be the formal Fourier series $\sum_{i,j=-\infty}^{\infty}\left(\sum_{n=-\infty}^{\infty} \varphi_n x_{i-n,j}\right) z^i w^j$; and we defined $\varphi(w)*F(z,w)$ to be $\sum_{i,j}\left(\sum_{n} \varphi_n x_{i,j-n}\right) z^i w^j$. (As expected, when $F(z,w) \in L^2(\mathbf{T}^2)$, the formal product is just the function product.) We then showed that the generating function for $M_\varphi X$ is $\varphi(\bar{z})*F(z,w)$, that for XM_φ is $\varphi(w)*F(z,w)$, that for $M_\varphi^* X$ is $\overline{\varphi(\bar{z})}*F(z,w)$, and that for XM_φ^* is $\overline{\varphi(\bar{w})}*F(z,w)$. Therefore the generating function for $M_\varphi X - XM_\varphi$ is $(\varphi(z) - \varphi(w))*F(z,w)$, and that for $M_\varphi^* X - XM_\varphi^*$ is $\overline{(\varphi(z) - \varphi(w))}*F(z,w)$.

The hypothesis that $M_\varphi X - XM_\varphi \in C_2$ is equivalent to $(\varphi(\bar{z}) - \varphi(w))*F(z,w) \in L^2(\mathbf{T}^2)$. Indeed

$$\|M_\varphi X - XM_\varphi\|_{C_2}^2 = \iint_{\mathbf{T}^2} |(\varphi(\bar{z}) - \varphi(w))*F(z,w)|^2.$$

Replacing φ by $a\varphi + b$, it is clear that without loss of generality, we may assume that Range$\varphi \subset Q$ ($Q = [0,1) \times [0,1)$ considered as embedded in the complex plane).

Now $(\varphi(\bar{z}) - \varphi(w))*F(z,w)$ is a function in $L^2(\mathbf{T}^2)$ and $\{(z,w): \varphi(z) = \varphi(w)\}$ has 2-dimensional Lebesgue measures 0 in $\mathbf{T}^2$. (This follows from the fact that for every complex number c, $\{z: \varphi(z) = c\}$ has linear measure 0. See [5, Proof of Theorem 4] for the details.) Therefore, we have that $\overline{(\varphi(\bar{z}) - \varphi(w))}/(\varphi(\bar{z}) - \varphi(w)) \in L^\infty(\mathbf{T}^2)$ with modulus 1 almost everywhere in $\mathbf{T}^2$, and so

$$(\dagger) \equiv \frac{\overline{\varphi(z) - \varphi(w)}}{\varphi(z) - \varphi(w)} \cdot [(\varphi(\bar{z}) - \varphi(w))*F(z,w)] =$$

$$= \frac{1}{\varphi(\bar{z}) - \varphi(w)} \cdot \overline{(\varphi(\bar{z}) - \varphi(w))} \cdot [(\varphi(\bar{z}) - \varphi(w))*F(z,w)] =$$

$$= \frac{1}{\varphi(z) - \varphi(w)} \cdot (\overline{(\varphi(\bar{z}) - \varphi(w))}*[(\varphi(\bar{z}) - \varphi(w))*F(z,w)]).$$

The last equality holds since both are $L^2(\mathbf{T}^2)$ functions, and so the function product and the formal product are the same. Let $[A,B] = AB - BA$. Then clearly $\overline{(\varphi(z) - \varphi(w))}*[(\varphi(\bar{z}) - \varphi(w))*F(z,w)]$ is the generating function for the operator $[M_\varphi^*, [M_\varphi, X]]$, while $(\varphi(\bar{z}) - \varphi(w))*[\overline{(\varphi(\bar{z}) - \varphi(w))}*F(z,w)]$ is the generating function for the operator $[M_\varphi, [M_\varphi^*, X]]$. It is well-known and straightforward to show that these two operators are equal. Since one is a Hilbert-Schmidt operator, they both are, and their generating functions are in $L^2(\mathbf{T}^2)$ and are equal almost everywhere in $\mathbf{T}^2$. Therefore, almost everywhere in $\mathbf{T}^2$,

$$(\dagger) = \frac{1}{\varphi(\bar{z}) - \varphi(w)}((\varphi(\bar{z}) - \varphi(w))*[\overline{(\varphi(\bar{z}) - \varphi(w))}*F(z,w)]).$$

Thus far, the ideas expressed here are similar to those in [5, Proof of Theorem 4]. However, at this point we want to remove the first * from this last expression, but cannot, as we don't even know that the expression in brackets is a function. If we could remove this *, we could cancel $\varphi(\bar{z}) - \varphi(w)$, with itself, and we would be done. What we do is introduce a kind of test function which will allow more of this algebraic manipulation.

THE TEST FUNCTION

Let $Q_{ij}^n = \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right) \times \left[\frac{j-1}{2^n}, \frac{j}{2^n}\right)$ for $1 \leqslant i, j \leqslant 2^n$, and let $E_{ij}^n = \varphi^{-1}(Q_{ij}^n)$. Clearly $\dot{\bigcup} Q_{ij}^n = Q$ and $\dot{\bigcup} E_{ij}^n = \mathbf{T}$. Let $\mathscr{S}$ denote the collection of quadruples of integers (i, j, p, q) for which $1 \leqslant i, j, p, q \leqslant 2^n$ and $|i-p| + |j-q| \geqslant 3$. Define the test function $t_n(z, w)$ for each positive integer n, as follows

$$t_n(z, w) = \sum_{(i,j,p,q) \in \mathscr{S}} \chi_{E_{ij}^n}(\bar{z}) \chi_{E_{pq}^n}(w).$$

We shall now show that

$$t_n(z, w) \cdot (\dagger) =$$

$$= \frac{t_n(z, w)}{\varphi(\bar{z}) - \varphi(w)}((\varphi(\bar{z}) - \varphi(w))*(\overline{\varphi(\bar{z}) - \varphi(w))}*F(z, w)) =$$

$$= t_n(z, w)*\overline{(\varphi(\bar{z}) - \varphi(w))}*F(z, w),$$

considered as formal Fourier series. The reader should realize that the first * in this last expression has not yet been defined. However, it is clear that since $\chi_{E_{ij}^n}(z)$ and $\chi_{E_{pq}^n}(z) \in L^\infty(\mathbf{T})$, the generating function for $M_{\chi_{E_{ij}^n}} X M_{\chi_{E_{pq}^n}}$ is given by $\chi_{E_{ij}^n}(\bar{z})*$ $*(\chi_{E_{pq}^n}(w)*F(z, w))$ (for every bounded operator matrix X with generating function F).

For simplicity we shall simplify notation as follows. Let $\chi_{ij}(z) = \chi_{E_{ij}^n}(z)$.

Clearly $*$ is distributive over $+$, since it reflects the corresponding algebraic operations of operator multiplication ($M_\varphi X$, XM_φ, etc.) and operator addition. Therefore, to prove the last equality, it suffices to show that for every (i, j, p, q) for which $|i-p|+|j-q| \geqslant 3$, we have

$$\frac{\chi_{ij}(\bar{z})\chi_{pq}(w)}{\varphi(\bar{z})-\varphi(w)}((\varphi(\bar{z})-\varphi(w))*[\overline{(\varphi(\bar{z})-\varphi(w))}*F(z,w)]) =$$

$$= \chi_{ij}(\bar{z})*(\chi_{pq}(w)*[\overline{(\varphi(\bar{z})-\varphi(w))}*F(z,w)]),$$

as formal Fourier series.

Now $|i-p|+|j-q| \geqslant 3$ implies one of $|i-p|$ or $|j-q| \geqslant 2$. Say $|i-p| \geqslant 2$. Then either $i<p$ or $i>p$, say $i<p$. The reader will see, after reading this proof, that the proofs for the other three cases are essentially identical. Let $c = \frac{i-1}{2^n} + \mathrm{i}\frac{j-1}{2^n}$. Then $(\bar{z}, w) \in E_{ij}^n \times E_{pq}^n$ implies $\varphi(\bar{z}) \in Q_{ij}^n$, $\varphi(w) \in Q_{pq}^n$, $\varphi(w)-c \neq 0$, and finally that $\left|\frac{\varphi(\bar{z})-c}{\varphi(w)-c}\right| < 1$. Therefore $\frac{\chi_{ij}(\bar{z})\chi_{pq}(w)}{\varphi(\bar{z})-\varphi(w)} = 0$ if $(\bar{z}, w) \notin E_{ij}^n \times E_{pq}^n$ and otherwise

$$\frac{\chi_{ij}(\bar{z})\chi_{pq}(w)}{\varphi(\bar{z})-\varphi(w)} = \frac{1}{(\varphi(\bar{z})-c)-(\varphi(w)-c)} =$$

$$= -(\varphi(w)-c)^{-1}\left[1-\frac{\varphi(\bar{z})-c}{\varphi(w)-c}\right]^{-1} =$$

$$= -(\varphi(w)-c)^{-1}\sum_{k=0}^{\infty}\left(\frac{\varphi(\bar{z})-c}{\varphi(w)-c}\right)^k.$$

If we now let

$$\varphi_k(z) = \chi_{ij}(z)(\varphi(z)-c)^k$$

and

$$\psi_k(w) = \chi_{pq}(w)(\varphi(w)-c)^{-(k+1)},$$

we have $\varphi_k \in L^\infty(\mathbf{T})$. Also since $p > i$ and $|i-p| \geqslant 2$, we have $\psi_k(w) = 0$ if $w \notin E_{pq}^n$ and otherwise $\varphi(w) \in Q_{pq}^n$, which implies

$$|\varphi(w)-c| = \left|\varphi(w)-\left(\frac{i-1}{2^n}+\mathrm{i}\frac{j-1}{2^n}\right)\right| \geqslant \frac{p-i}{2^n} \geqslant \frac{2}{2^n}.$$

Therefore $\psi_k \in L^\infty(\mathbf{T})$ also. In addition, if $(\bar{z}, w) \in E^n_{ij} \times E^n_{pq}$, then $\varphi(\bar{z}) \in Q^n_{ij}$ and $\varphi(w) \in Q^n_{pq}$, and so $|\varphi(\bar{z}) - c| \leqslant \frac{\sqrt{2}}{2^n}$ and $|\varphi(w) - c| \geqslant \frac{2}{2^n}$. Therefore

$$\sum_k \|\varphi_k\|_{L^\infty(\mathbf{T})} \|\psi_k\|_{L^\infty(\mathbf{T})} \leqslant 2^{n-1} \sum_k 2^{-k/2} < \infty.$$

We now claim that if $G(z, w)$ is the generating function for any bounded matrix operator for which $(\varphi(\bar{z}) - \varphi(w)) * G(z, w) \in L^2(\mathbf{T}^2)$, then

$$\frac{\chi_{ij}(\bar{z})\chi_{pq}(w)}{\varphi(\bar{z}) - \varphi(w)} (\varphi(\bar{z}) - \varphi(w)) * G(z, w) = \\ = \chi_{ij}(\bar{z}) * \chi_{pq}(w) * G(z, w).$$

The proof of this claim requires some work and we give it next. The above functions φ_k, ψ_k can now be used to say

$$\frac{\chi_{ij}(\bar{z})\chi_{pq}(w)}{\varphi(\bar{z}) - \varphi(w)} = - \sum_{k=0}^{\infty} \varphi_k(\bar{z})\psi_k(w)$$

where the sum converges in $L^\infty(\mathbf{T}^2)$. Now

$$\varphi_k(\bar{z})\, \psi_k(w)[(\varphi(\bar{z}) - \varphi(w)) * G(z, w)] = \\ = \varphi_k(\bar{z})\, \psi_k(w) * [(\varphi(\bar{z}) - \varphi(w)) * G(z, w)] = \\ = [\varphi_k(\bar{z})\, \psi_k(w) * (\varphi(\bar{z}) - \varphi(w))] * G(z, w).$$

This last equality follows since if we let G denote the matrix operator with generating function $G(z, w)$, the former expression is the generating function for $M_{\varphi_k}[M_\varphi, G]M_{\psi_k}$ and the latter expression is the generating function for $M_{\varphi_k}M_\varphi G M_{\psi_k} - M_{\varphi_k} G M_\varphi M_{\psi_k}$, which are equal as operators. Therefore

$$\frac{\chi_{ij}(\bar{z})\, \chi_{pq}(w)}{\varphi(\bar{z}) - \varphi(w)} [(\varphi(\bar{z}) - \varphi(w)) * G(z, w)] = \\ = - \sum_{k=0}^{\infty} [\varphi_k(\bar{z})\psi_k(w) * (\varphi(\bar{z}) - \varphi(w))] * G(z, w)$$

where the convergence is in $L^2(\mathbf{T}^2)$. But considering the corresponding finite sum of operators,

$$- \sum_{k=0}^{K} (M_{\varphi_k} M_\varphi G M_{\psi_k} - M_{\varphi_k} G M_\varphi M_{\psi_k}) = - \sum_{k=0}^{K} (M_{\varphi_k} M_{(\varphi - c)} G M_{\psi_k} - M_{\varphi_k} G M_{(\varphi - c)} M_{\psi_k}) =$$

$$= - \sum_{k=0}^{K} (M_{\chi_{ij}(\varphi - c)^k} M_{(\varphi - c)} G M_{\chi_{pq}(\varphi - c)^{-(k+1)}} - M_{\chi_{ij}(\varphi - c)^k} G M_{(\varphi - c)} M_{\chi_{pq}(\varphi - c)^{-(k+1)}}) =$$

$$= - \sum_{k=0}^{K} (M_{\chi_{ij}(\varphi - c)^{k+1}} G M_{\chi_{pq}(\varphi - c)^{-(k+1)}} - M_{\chi_{ij}(\varphi - c)^k} G M_{\chi_{pq}(\varphi - c)^{-k}}) =$$

$$= - M_{\chi_{ij}(\varphi - c)^{K+1}} G M_{\chi_{pq}(\varphi - c)^{-(K+1)}} + M_{\chi_{ij}} G M_{\chi_{pq}}.$$

Now since

$$\|\chi_{ij}(\varphi - c)^{K+1}\|_{L^\infty(\mathrm{T})} \leqslant \left(\frac{\sqrt{2}}{2^n}\right)^{K+1}$$

and

$$\|\chi_{pq}(\varphi - c)^{-(K+1)}\|_{L^\infty(\mathrm{T})} \leqslant \left(\frac{2^n}{2}\right)^{K+1},$$

we have that

$$\|M_{\chi_{ij}(\varphi-c)^{K+1}} G M_{\chi_{pq}(\varphi-c)^{-(K+1)}}\|_{\mathscr{L}(\mathscr{H})} \leqslant \left(\frac{1}{\sqrt{2}}\right)^{K+1} \|G\| \to 0.$$

Hence the partial sum of the operators converges in the uniform norm to $M_{\chi_{ij}} G M_{\chi_{pq}}$ with the generating function $\chi_{ij}(\bar{z})\chi_{pq}(w) * G(z, w)$. The uniqueness of the generating function (equivalently, the uniqueness of the matrix in the standard basis) gives us that

$$-\sum_{k=0}^{\infty} [\varphi_k(\bar{z})\psi_k(w) * (\varphi(\bar{z}) - \varphi(w))] * G(z, w) =$$

$$= \chi_{ij}(\bar{z})\, \chi_{pq}(w) * G(z, w),$$

and we have proved the claim. Thus we have that

$$t_n(z, w)\ (\dagger) = t_n(z, w) * \overline{[(\varphi(\bar{z}) - \varphi(w)) * F(z, w)]}$$

where $|t_n(z, w)| \leqslant 1$ for all z, w, and $(\dagger)$ is a function in $L^2(\mathbf{T}^2)$. In other words

$$t_n(z, w) * \overline{[(\varphi(\bar{z}) - \varphi(w)) * F(z, w)]}$$

is a sequence of functions in $L^2(\mathbf{T}^2)$ with L^2-norm uniformly bounded by the L^2-norm of $(\varphi(\bar{z}) - \varphi(w)) * F(z, w)$. Therefore for every n, the corresponding operator statement is

$$\|\sum_{(i,j,p,q)\in\mathscr{S}} M_{\chi_{ij}}[M_\varphi^*, X] M_{\chi_{pq}}\|_{C_2} \leqslant \|[M_\varphi, X]\|_{C_2}.$$

It is elementary to show that if $\{T_n\}$ is a sequence of bounded operators such that $T_n \to T$ in the weak operator topology and $\|T_n\|_{C_2} \leqslant M$ for every n, then $T \in C_2$ and $\|T\|_{C_2} \leqslant M$. Therefore our proof will be finished if we can show that $T_n \to [M_\varphi^*, X]$ in the weak operator topology, where

$$T_n = \sum_{(i,j,p,q)\in\mathscr{S}} M_{\chi_{ij}}[M_\varphi^*, X] M_{\chi_{pq}}.$$

To do this, let us recall some facts. Let $f \otimes g$ for $f, g \in L^2(\mathrm{T})$ denote the rank one operator $f \otimes g\ (h) = (h, g)f$. It is elementary to verify that if $A \in \mathscr{L}(\mathscr{H})$ with matrix entries $(a_{ij})_{i,j=-\infty}^{\infty}$ relative to the standard basis $\{z^n\}_{n=-\infty}^{\infty}$, then $a_{ij} = (Az^j, z^i)$. Furthermore, the matrix for $z^j \otimes z^i$ has a 1 in the j, i entry and 0 elsewhere and so a simple verification shows that $a_{ij} = \text{trace } A(z^j \otimes z^i)$. That is, $(Az^j, z^i) = \text{trace } A(z^j \otimes z^i)$. From this it follows directly that if $f, g \in L^2(\mathbf{T})$, then $(Af, g) = \text{trace } A(f \otimes g)$. The reader should take care to realize that (Af, g), as well as $f \otimes g$, are linear in f and conjugate linear in g.

Now to compute $[M_\varphi^*, X] - T_n$, note that $\sum_{1 < i, j < 2^n} M_{\chi_{ij}} = I$. Define

$$\mathscr{S}' = \{(i, j, p, q) \notin \mathscr{S} \text{ and } 1 \leqslant i, j, p, q \leqslant 2^n\}.$$

Clearly also

$$\mathscr{S}' = \{(i, j, p, q) \colon 1 \leqslant i, j, p, q \leqslant 2^n \text{ and } |i-p| + |j-q| \leqslant 2\}.$$

Then

$$[M_\varphi^*, X] - T_n = \sum_{(i, j, p, q) \in \mathscr{S}'} M_{\chi_{ij}} [M_\varphi^*, X] M_{\chi_{pq}}.$$

If we now fix $f, g \in L^2(\mathbf{T})$, then

$$(([M_\varphi^*, X] - T_n) f, g) = \sum_{(i, j, p, q) \in \mathscr{S}'} (M_{\chi_{ij}} [M_\varphi^*, X] M_{\chi_{pq}} f, g) =$$

$$= \sum_{(i, j, p, q) \in \mathscr{S}'} \text{trace } (M_{\chi_{ij}} [M_\varphi^*, X] M_{\chi_{pq}}) (f \otimes g).$$

Setting $Y = [M_\varphi^*, X] M_{\chi_{pq}} (f \otimes g)$, we obtain $Y \in C_1$ and hence $\text{trace} M_{\chi_{ij}} Y = \text{trace} Y M_{\chi_{ij}}$. Therefore

$$\text{trace}(M_{\chi_{ij}} [M_\varphi^*, X] M_{\chi_{pq}}) (f \otimes g) = \text{trace}[M_\varphi^*, X] M_{\chi_{pq}} (f \otimes g) M_{\chi_{ij}}.$$

It is easy to see that if R is any trace class operator (in particular, $R = M_{\chi_{pq}}(f \otimes g) M_{\chi_{ij}}$) and $A, B \in \mathscr{L}(\mathscr{H})$, then $\text{trace}[A, B]R = \text{trace}[R, A]B$. Indeed,

$$\text{trace}([A, B]R - [R, A]B) = \text{trace}(ABR - RAB + ARB - BAR) =$$

$$= \text{trace}[AB, R] + \text{trace}[AR, B] = 0,$$

since R and $AR \in C_1$. Hence,

$$|(([M_\varphi^*, X] - T_n) f, g)| = \Big| \sum_{(i, j, p, q) \in \mathscr{S}'} \text{trace}[M_{\chi_{pq}} (f \otimes g) M_{\chi_{ij}}, M_\varphi^*] X \Big| \leqslant$$

$$\leqslant \sum_{(i, j, p, q) \in \mathscr{S}'} |\text{trace}[M_\varphi^*, M_{\chi_{pq}} (f \otimes g) M_{\chi_{ij}}] X| \leqslant$$

$$\leqslant \sum_{(i, j, p, q) \in \mathscr{S}'} \|X\|_{\mathscr{L}(\mathscr{H})} \|[M_\varphi^*, M_{\chi_{pq}} (f \otimes g) M_{\chi_{ij}}]\|_{C_1}.$$

Note that the commutator in this last expression is the difference between two rank one operators and hence is a rank 2 operator. It is well-known that there exists $C > 0$ such that $\|A\|_{C_1} \leqslant C\|A\|_{C_2}$ for every rank 2 operator A. This allows us to obtain an upper bound on the trace norm of this commutator by looking at its Hilbert-Schmidt norm. That is,

$$\|[M_\varphi^*, M_{\chi_{pq}}(f \otimes g) M_{\chi_{ij}}]\|_{C_1} \leqslant C\|[M_\varphi^*, M_{\chi_{pq}}(f \otimes g) M_{\chi_{ij}}]\|_{C_2}.$$

Next, we need to know the generating function for this commutator. We start by considering the generating function for $f \otimes g$.

The matrix for $z^i \otimes z^j$ has a 1 in the i, j entry and 0 elsewhere. Hence its generating function is $z^i w^j$. It follows that if $f, g \in L^2(\mathbf{T})$, then $f \otimes g = \sum \hat{f}(i)\overline{\hat{g}(j)} z^i \otimes z^j$. Therefore the generating function of $f \otimes g$ is

$$\sum_{i,j} \hat{f}(i)\overline{\hat{g}(j)} z^i w^j = (\sum_i \hat{f}(i) z^i)(\sum_j \overline{\hat{g}(j)}\, w^j) = f(z)\overline{g(\overline{w})}.$$

Earlier remarks then give that the generating function for $M_{\chi_{pq}}(f \otimes g)M_{\chi_{ij}}$ is $\chi_{pq}(\overline{z})\chi_{ij}(w) f(z)\overline{g(\overline{w})}$, and finally the generating function for $[M_\varphi^*, M_{\chi_{pq}}(f \otimes g)M_{\chi_{ij}}]$ is $(\varphi(\overline{z}) - \varphi(w))\chi_{pq}(\overline{z})\chi_{ij}(w) f(z)\overline{g(\overline{w})}$. Therefore after a change of variables of $\overline{z}$ to z,

$$\|[M_\varphi^*, M_{\chi_{pq}}(f \otimes g) M_{\chi_{ij}}]\|_{C_2}^2 = \iint_{E_{pq}^n \times E_{ij}^n} |(\varphi(z) - \varphi(w))\, f(\overline{z}) g(\overline{w})|^2 \leqslant$$

$$\leqslant \left(\frac{3\sqrt{2}}{2^n}\right)^2 \iint_{E_{pq}^n \times E_{ij}^n} |f(\overline{z}) g(\overline{w})|^2.$$

This last equality follows by $(z, w) \in E_{pq}^n \times E_{ij}^n$ where $(i, j, p, q) \in \mathscr{S}'$ implies $|i - p| + |j - q| \leqslant 2$. This puts $\varphi(z) \in Q_{pq}^n$, $\varphi(w) \in Q_{ij}^n$ and so $|\varphi(z) - \varphi(w)| \leqslant \frac{3\sqrt{2}}{2^n}$ by the location of these squares relative to each other.

Hence,

$$|(([M_\varphi^*, X] - T_n) f, g)| \leqslant$$

$$\leqslant \frac{3\sqrt{2}C}{2^n} \|X\| \sum_{(i,j,p,q) \in \mathscr{S}'} \left(\iint_{E_{pq}^n \times E_{ij}^n} |f(\overline{z}) g(w)|^2 \right)^{1/2} \leqslant$$

$$\leqslant \frac{3\sqrt{2}C}{2^n} \|X\| \left[\sum_{(i,j,p,q) \in \mathscr{S}'} \left(\iint_{E_{pq}^n \times E_{ij}^n} |f(\overline{z}) g(w)|^2 \right) \right]^{1/2} \Big[\sum_{(i,j,p,q) \in \mathscr{S}'} 1 \Big]^{1/2}.$$

This last inequality follows from Hölder's inequality. The last bracketed expression indicates that we need to compute the cardinality of $\mathscr{S}'$. Recall that $\mathscr{S}'$ is the set of all 4-tuples (i, j, p, q) for which $1 \leqslant i, j, p, q \leqslant 2^n$ and $|i-p| + |j-q| \leqslant 2$. The number for which $|i-p| + |j-q| = 0$ is clearly $2^n \cdot 2^n$. The number for which $|i-p| + |j-q| = 1$ is the number for which $i = p$ and $|j-q| = 1$, or $|i-p| = 1$ and $j = p$. This is $2 \cdot 2 \cdot 2^n(2^n - 1)$. Finally consider the number for which $|i-p| + |j-q| = 2$, that is $|i-p| = |j-q| = 1$, or $|i-p| = 2$ and $j = q$, or $i = p$ and $|j-q| = 2$. This is $2 \cdot 2 \cdot (2^n - 1)(2^n - 1) + 2(2^n \cdot 2 \cdot (2^n - 2))$. The crucial point is that the total of these quantities, namely the cardinality of $\mathscr{S}'$, is less than or equal to $k \cdot 4^n$, where k is some positive number independent of n.

Continuing on from this last inequality we obtain

$$|(([M_\varphi^*, X] - T_n)f, g)| \leqslant \frac{3\sqrt{2}}{2^n} C \|X\| \left[\iint\limits_{\dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n} |f(\bar{z})g(w)|^2 \right] [k4^n]^{1/2} \leqslant$$

$$\leqslant 3\sqrt{2}\, C\sqrt{k} \|X\| \left[\iint\limits_{\dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n} |f(\bar{z})g(w)|^2 \right]^{1/2}.$$

Now our proof will be complete if we can show that this last bracketed expression approaches 0 as $n \to \infty$. To see this, note that $(z, w) \in \dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n$ implies that $\varphi(z) \in Q_{pq}^n$, $\varphi(w) \in Q_{ij}^n$ and $|i-p| + |j-q| \leqslant 2$. Considering the location of these squares relative to each other, we obtain $|\varphi(z) - \varphi(w)| \leqslant \frac{3\sqrt{2}}{2^n}$. Hence the sets $\dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n$ are nested downward and

$$\bigcap_n \dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n \subset \{(z, w) : \varphi(z) = \varphi(w)\}.$$

This last set has Lebesgue measure 0. Set $\chi_n(z) = \chi_{\dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n}(z)$. Then $\chi_n(z) \downarrow 0$. Hence by the Lebesgue dominated convergence theorem,

$$\iint\limits_{\dot{\bigcup}_{(i,j,p,q) \in \mathscr{S}'} E_{pq}^n \times E_{ij}^n} |f(\bar{z})g(w)|^2 = \iint\limits_{\mathbf{T}^2} \chi_n |f(\bar{z})g(w)|^2 \to 0$$

as $n \to \infty$, since $\chi_n |f(\bar{z})g(w)|^2 \leqslant |f(\bar{z})g(w)|^2 \in L^2(\mathbf{T}^2)$. Q.E.D.

Historically, the original Fuglede Theorem was proved for $N_1 = N_2$. C. R. Putnam noted the generalization. We now have evidence of possible further generalizations. We ask the following questions. Let $\{M_n\}$ and $\{N_n\}$ denote any two sequences of commuting normal operators. (Employing the Putnam 2×2 matrix trick to the following question, the reader will see that without loss of generality we may also assume $M_n = N_n$ for every n.) Let $X \in \mathscr{L}(\mathscr{H})$.

QUESTION 1. a) Does $\sum_{n=1}^{N} M_n X N_n = 0$ imply $\sum_{n=1}^{N} M_n^* X N_n^* = 0$?

b) Does $\sum_{n=1}^{N} M_n X N_n \in C_2$ imply $\sum_{n=1}^{N} M_n^* X N_n^* \in C_2$ and are the Hilbert-Schmidt norms of both expressions equal?

QUESTION 2. Suppose $\sum_{n=1}^{\infty} \|M_n\| \|N_n\| < \infty$.

a) Does $\sum_{n=1}^{\infty} M_n X N_n = 0$ imply $\sum_{n=1}^{\infty} M_n^* X N_n^* = 0$?

b) Does $\sum_{n=1}^{\infty} M_n X N_n \in C_2$ imply $\sum_{n=1}^{\infty} M_n^* X N_n^* \in C_2$ and are their Hilbert-Schmidt norms equal?

Recently C. Apostol employed Theorem 1 to settle Question 1a and b in the affirmative in the case when $N = 2$. The other cases remain unsolved but we have some evidence. First of all, when generating functions are applied, the statements all appear "formally" to be true. Also, if all the M_n's and N_n's were simultaneously diagonal matrices, then straightforward matrix calculations prove all the statements in this case.

We next give the case for $N = 2$. Apostol's contribution was the use of projection operators to pass from the case where the normal operators are invertible to the case where they are 1-1.

COROLLARY 2. *Let $\{M_1, M_2\}$ and $\{N_1, N_2\}$ denote commuting pairs of normal operators and let $X \in \mathscr{L}(\mathscr{H})$. Then*

$$\|M_1 X N_1 + M_2 X N_2\|_{C_2} = \|M_1^* X N_1^* + M_2^* X N_2^*\|_{C_2}.$$

In particular, if one of these two expressions are in C_2, then the other is in C_2, and if one is 0 *then the other is* 0.

Proof. We first claim that the corollary holds true when M_2 or N_2 is 0 (the case $N = 1$ in Questions 1a and 1b). Indeed, applying Theorem 1 to $M_1 X N_1 \in C_2$, where M_1 is the normal operator and $X N_1$ is the bounded operator, we obtain $\|M_1 X N_1\|_{C_2} = \|M_1^* X N_1\|_{C_2}$. Similarly, $\|M_1^* X N_1\|_{C_2} = \|M_1^* X N_1^*\|_{C_2}$. Therefore, $\|M_1 X N_1\|_{C_2} = \|M_1^* X N_1^*\|_{C_2}$.

We now claim that in Questions 1 and 2 we may assume, without loss of generality, that $M_n = N_n$ for all n. Set $A_n = M_n \oplus N_n$ and $Y = \begin{pmatrix} 0 & X \\ 0 & 0 \end{pmatrix}$. Then A_n is normal and

$$\sum A_n Y A_n = \begin{pmatrix} 0 & \sum M_n X N_n \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \sum A_n^* X A_n^* = \begin{pmatrix} 0 & \sum M_n^* X N_n^* \\ 0 & 0 \end{pmatrix}.$$

From these equations, it is clear that to settle any of the Questions 1 or 2, it suffices to assume that $M_n = N_n$.

Therefore we must prove that if M and N are commuting normal operators and $X \in \mathscr{L}(\mathscr{H})$, then $\|MXM + NXN\|_{C_2} = \|M^*XM^* + N^*XN^*\|_{C_2}$.

The next reduction is to the case where M and N are 1-1. To indicate why this is important, note that if M and N were invertible, then $MXM + NXN \in C_2$ if and only if $N^{-1}MX + XNM^{-1} \in C_2$, and Theorem 1 would apply to this expression being that $N^{-1}M$ and $-NM^{-1}$ are normal operators. Let $M = M_1 \oplus 0$ be the orthogonal decomposition of M relative to $\mathscr{H} = (\ker M)^{\perp} \oplus \ker M$. Since N and N^* commute with M (that N^* commutes with M follows by the classical Fuglede Theorem), then $\ker M$ is an invariant subspace for N and N^* and therefore it is a reducing subspace for N; say $N = N_1 \oplus N_2$. If then $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, we have

$$MXM + NXN = \begin{pmatrix} M_1AM_1 + N_1AN_1 & N_1BN_2 \\ N_2CN_1 & N_2DN_2 \end{pmatrix}$$

and also the corresponding adjoint equation. But, for example, $\|N_1BN_2\|_{C_2} = \|N_1^*BN_2^*\|_{C_2}$ (and so on) by Theorem 1. In other words, the question reduces to the operator in the upper left corner. That is, if $M_1AM_1 + N_1AN_1 \in C_2$ where M_1 is 1-1, together with the earlier hypotheses, must $M_1^*AM_1^* + N_1^*AN_1^* \in C_2$? Similarly, we may assume without loss of generality, that N_1 is 1-1, and hence that M and N are 1-1.

If M and N are commuting normal operators that are 1-1, then using the spectral theorem we can obtain a sequence of projection operators P_n which commute with M and N such that $P_nM \to M$ and $P_nN \to N$ uniformly and P_nM is bounded below on $(\ker P_nM)^{\perp}$ and P_nN is bounded below on $(\ker P_nN)^{\perp}$. In other words, use the spectral theorem to obtain M_φ and M_ψ representing M and N respectively. Then use the projections determined by $\left\{|\varphi| > \frac{1}{n}\right\} \cap \left\{|\psi| > \frac{1}{n}\right\}$. Let E_n denote the orthogonal projection onto $L^2\left(\left\{|\varphi| \leqslant \frac{1}{n}\right\} \cup \left\{|\psi| \leqslant \frac{1}{n}\right\}\right)$. Then $P_nE_n=0$, and $P_nM + \frac{1}{n} E_n$ and $P_nN + \frac{1}{n} E_n$ are commuting normal operators that are invertible. Therefore

$$P_n(MXM + NXN)P_n = P_nM(P_nXP_n)P_nM + P_nN(P_nXP_n)P_nN =$$

$$= \left(P_nM + \frac{1}{n} E_n\right)(P_nXP_n)\left(P_nM + \frac{1}{n} E_n\right) +$$

$$+ \left(P_nN + \frac{1}{n} E_n\right)(P_nXP_n)\left(P_nN + \frac{1}{n} \cdot E_n\right).$$

Hence here we have a situation where $AYA + BYB \in C_2$, A, B are commuting, normal, invertible operators, $Y \in \mathscr{L}(\mathscr{H})$. Let us assume that we have established the corollary in this case. That is, $\|AYA + BYB\|_{C_2} = \|A^*YA^* + B^*YB^*\|_{C_2}$. Then

$$\|P_n(M^*XM^* + N^*XN^*)P_n\|_{C_2} =$$

$$= \|(P_nM)^*P_nXP_n(P_nM)^* + (P_nN)^*P_nXP_n(P_nN)^*\|_{C_2} =$$

$$= \Bigg\|\left(P_nM + \frac{1}{n} E_n\right)^* P_nXP_n\left(P_nM + \frac{1}{n} E_n\right)^* +$$

$$+ \left(P_nN + \frac{1}{n} E_n\right)^* P_nXP_n\left(P_nN + \frac{1}{n} E_n\right)^* \Bigg\|_{C_2} =$$

$$= \Bigg\|\left(P_nM + \frac{1}{n} E_n\right) P_nXP_n\left(P_nM + \frac{1}{n} E_n\right) +$$

$$+ \left(P_nN + \frac{1}{n} E_n\right) P_nXP_n\left(P_nN + \frac{1}{n} E_n\right)\Bigg\|_{C_2} =$$

$$= \|P_n(MXM + NXN)P_n\|_{C_2} \to \|MXM + NXN\|_{C_2}.$$

Clearly if a sequence of operators $T_n \to T$ uniformly and $\|T_n\|_{C_2} \leqslant M$, then $\|T\|_{C_2} \leqslant M$ (recall this fact used in the proof of Theorem 1, where we only needed the hypothesis that $T_n \to T$ in the weak operator topology). Therefore

$$\|M^*XM^* + N^*XN^*\|_{C_2} = \|MXM + NXN\|_{C_2}.$$

Finally, to complete the proof it suffices to prove the case where M and N are invertible.

$$MXM + NXN = N(N^{-1}MX + XNM^{-1})M \in C_2.$$

Hence, $N^{-1}M$ and $-NM^{-1}$ are normal operators and $N^{-1}MX + XNM^{-1} \in C_2$. Applying Theorem 1 twice we obtain

$$\|MXM + NXN\|_{C_2} = \|N^*(N^{-1}MX + XNM^{-1})M^*\|_{C_2} =$$
$$= \|N^{-1}M(N^*XM^*) + (N^*XM^*)NM^{-1}\|_{C_2} =$$
$$= \|(N^{-1}M)^*(N^*XM^*) + (N^*XM^*)(NM^{-1})^*\|_{C_2} =$$
$$= \|M^*XM^* + N^*XN^*\|_{C_2}.$$

Q.E.D.

We do not yet see how to apply these techniques to settle Question 1 when $N = 3$. We do have some positive indications. For instance, if both expressions are in C_2, then we can prove that their Hilbert-Schmidt norms are equal.

REMARK. Donald Hadwin has proven that (2) ⇒ (1) in [5]. Namely, if a normal operator is the sum of a diagonalizable operator and a Hilbert-Schmidt operator, then an operator can be chosen with an arbitrary small Hilbert-Schmidt norm.

Since this paper was written, Dan Voiculescu has settled Berg's question in the affirmative [7]. This result implies Theorem 1, but does not settle Questions 1 or 2 nor does it provide another proof of Corollary 2.

Finally we wish to ask another question relating commutators of normal operators, Fuglede's Theorem modulo the *trace class*, and the trace.

In [5, Theorem 8] we proved that if N is a normal operator, $X \in C_2$ and $NX - XN \in C_1$, then trace$(NX - XN) = 0$. If we merely assume $X \in \mathscr{L}(\mathscr{H})$, the statement fails. For example, let N be the bilateral shift and X the unilateral shift viewed as a 2-way infinite matrix. Then trace $(NX - XN) = 1$. However, can we only assume X is a compact operator?

QUESTION 3. If N is a normal operator and $X \in \mathscr{K}(\mathscr{H})$ such that $NX - XN \in C_1$, must trace$(NX - XN) = 0$?

This relates to another open question. In [4] we asked if the Fuglede theorem was true modulo C_1. Suppose it were. That is, N normal, $X \in \mathscr{L}(\mathscr{H})$, and $NX - XN \in C_1$ implies $N^*X - XN^* \in C_1$. We claim then that this would settle Question 3 in the affirmative. To see this, note that $X^* + X$ and $X^* - X$ are diagonal, and so trace$[X^* + X, T] =$ trace$[X^* - X, T] = 0$.

REFERENCES

1. BERG, I. D., An extension of the Weyl-von Neumann Theorem to normal operators, *Trans. Amer. Math. Soc.*, **160**(1971), 365–371.
2. VON NEUMANN, J., *Charakterisierung des Spektrums eines Integraloperators*, Actualites Sci. Indust., no. 229, Hermann, Paris, 1935.

3. WEISS, G., *Commutators and operator ideals*, dissertation, University of Michigan, 1975.
4. WEISS, G., Fuglede's commutativity theorem modulo operator ideals, *Proc. Amer. Math. Soc.*, to appear.
5. WEISS, G., The Fuglede Theorem modulo the Hilbert-Schmidt class and generating functions for matrix operators. I, *Trans. Amer. Math. Soc.*, **246**(1978), 193–209.
6. WEYL, H., Uber beschrankte quadratischen Formen deren Differenz vollstetig ist, *Rend. Circ. Mat. Pelermo*, **27**(1909), 373–392.
7. VOICULESCU, D., Some results on norm-ideal perturbations of Hilbert space operators, *J. Operator Theory*, **2**(1979), 3–37.

GARY WEISS
Department of Mathematics,
University of Cincinnati,
Cincinnati, OH 45221,
U.S.A.

Received November 27, 1978; revised April 8, 1980.