

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

BY  
GARY WEISS<sup>1</sup>

**ABSTRACT.** We prove the following statements about bounded linear operators on a separable, complex Hilbert space: (1) Every normal operator  $N$  that is similar to a Hilbert-Schmidt perturbation of a diagonal operator  $D$  is unitarily equivalent to a Hilbert-Schmidt perturbation of  $D$ ; (2) For every normal operator  $N$ , diagonal operator  $D$  and bounded operator  $X$ , the Hilbert-Schmidt norms (finite or infinite) of  $NX - XD$  and  $N^*X - XD^*$  are equal; (3) If  $NX - XN$  and  $N^*X - XN^*$  are Hilbert-Schmidt operators, then their Hilbert-Schmidt norms are equal; (4) If  $X$  is a Hilbert-Schmidt operator and  $N$  is a normal operator so that  $NX - XN$  is a trace class operator, then  $\text{Trace}(NX - XN) = 0$ ; (5) For every normal operator  $N$  that is a Hilbert-Schmidt perturbation of a diagonal operator, and every bounded operator  $X$ , the Hilbert-Schmidt norms (finite or infinite) of  $NX - XN$  and  $N^*X - XN^*$  are equal. The main technique employs the use of a new concept which we call 'generating functions for matrices'.

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

Consider the following statements:

(1) For every normal operator  $N$  and  $\epsilon > 0$ , there exist a diagonal operator  $D$  and a Hilbert-Schmidt operator  $K_\epsilon$  with  $\|K_\epsilon\|_2 < \epsilon$  for which  $N \cong D + K_\epsilon$  ( $\cong$  denotes unitary equivalence).

(2) For every normal operator  $N$ , there exist a diagonal operator  $D$  and a  $K \in C_2$  for which  $N \cong D + K$ .

(3) For every normal operator  $N$  and bounded operator  $X$ ,  $\|NX - XN\|_2 = \|N^*X - XN^*\|_2$ .

---

Received by the editors March 3, 1977 and, in revised form, August 29, 1977.

AMS (MOS) subject classifications (1970). Primary 47A05, 47A55, 47B10, 47B15, 47B47; Secondary 05A15, 05B20.

*Key words and phrases.* Fuglede's Theorem, commutators, normal operators, perturbations, Hilbert Schmidt operators, trace class, trace, generating functions for matrices, Laurent matrices, normal derivations.

<sup>1</sup> This research was partially supported by a grant from the N.S.F.

© American Mathematical Society 1979

(4) For every normal operator  $N$  and bounded operator  $X$ ,  $NX - XN \in C_2$  implies  $N^*X - XN^* \in C_2$ .

(5) For every normal operator  $N$  and bounded operator  $X$ , if  $NX - XN \in C_2$  and  $N^*X - XN^* \in C_2$ , then  $\|NX - XN\|_2 = \|N^*X - XN^*\|_2$ .

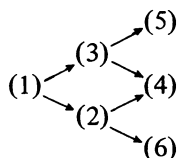
(6) If  $N$  is normal,  $X \in C_2$ , and  $NX - XN \in C_1$ , then  $\text{Trace}(NX - XN) = 0$ .

In [11] Weyl proved that every selfadjoint operator is a compact perturbation of a diagonalizable operator, and that the perturbation may be chosen with an arbitrarily small operator norm. In [8], von Neumann proved that the perturbation could be chosen to be in the Hilbert-Schmidt class and with arbitrarily small Hilbert-Schmidt norm. In [1], I. D. Berg generalized Weyl's result to normal operators, and proved that if the spectrum of the normal operator is 'thin enough', then the compact perturbation can also be chosen to be a Hilbert-Schmidt operator with an arbitrarily small Hilbert-Schmidt norm. He asked whether or not the von Neumann result generalizes to all normal operators (that is, statements (1) and (2)). These questions remain open. He conjectured that the full generalization fails and that he believes a barrier preventing a normal operator from having the representation (1) or (2) is that its absolutely continuous part have a spectrum of positive 2-dimensional Lebesgue measure. At present, not a single such normal operator is known which can be represented as in (1) or (2).

The 1970s has seen a flurry of deep results on the perturbation theory of operators and the theory of commutators. Besides Berg's paper [1], some of the well-known papers relating perturbation theory to commutators are Berger and Shaw [2], Brown, Douglas and Fillmore [3], Carey and Pincus [4], and Helton and Howe [7].

The connection between (3)–(6) and the Berg problem (2) is clear from the next remarks.

The following implications hold true.



Their proofs are elementary and fairly well known so we omit them (see [10] or [9, pp. 154–162]).

We shall prove that (5) and (6) are true ((6) settles a question in the negative in [9, p. 162]), and we shall obtain as corollaries that  $(3) \leftrightarrow (4)$  and  $(2) \rightarrow (3)$ . We shall also obtain related results. The above diagram is made current in the summary at the end of this paper.

**DEFINITION.** A *Laurent operator* is an operator of the form  $M_\phi$  acting on  $L^2(T)$ , where  $\phi(z) \in L^\infty(T)$  and  $T$  denotes the unit circle.

DEFINITION. If  $N$  is a normal operator and  $\phi(z) \in L^\infty(T)$ , then  $M_\phi$  is called a *Laurent part* of  $N$  provided  $M_\phi$  has no eigenvectors and there exists a diagonal operator  $D$  such that  $N = M_\phi \oplus D$ .

LEMMA 1. *Every normal operator is the direct sum of a diagonalizable operator and a Laurent part.*

PROOF. Let  $N$  be any normal operator and let  $\mathfrak{N}$  denote the closed linear span of the set of its eigenvectors. Then  $\mathfrak{N}$  reduces  $N$  and  $N|_{\mathfrak{N}} = D$  is a diagonal operator with the same set of eigenvectors as  $N$ . Let  $N_1 = N|_{\mathfrak{N}^\perp}$ . Then  $N = D \oplus N_1$  and  $N_1$  has no eigenvectors. Using the spectral theorem, we obtain  $N_1 \cong M_\psi$  acting on  $L^2(\mu)$ , with  $\psi \in L^\infty(\mu)$ , where  $\mu$  acts on a finite measure space (this may be accomplished since  $H$  is separable). Because  $N_1$  has no eigenvectors, it is clear that the underlying measure space can have no atoms. However, it is well known [6] that every finite nonatomic probability measure space can be realized as 1-dimensional Lebesgue measure on  $[0, 1]$  or equivalently, on the unit circle  $T$ . In other words, without loss of generality, we can insure that  $N_1 \cong M_\phi$  acting on  $L^2(T)$ , with  $\phi \in L^\infty(T)$ . Let  $D_1$  denote  $D$  under this unitary transformation. Then  $N \cong D_1 \oplus M_\phi$ . Q.E.D.

This lemma provides us with a crucial canonical form for the commutator  $NX - XN$ . Letting  $X$  be any operator in  $L(H)$ , relative to  $H = \mathfrak{N} \oplus \mathfrak{N}^\perp$ , we obtain

$$N = \begin{pmatrix} D & 0 \\ 0 & M_\phi \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} X_1 & X_2 \\ X_3 & X_4 \end{pmatrix}.$$

A computation then shows that

$$NX - XN = \begin{pmatrix} DX_1 - X_1D & DX_2 - X_2M_\phi \\ M_\phi X_3 - X_3D & M_\phi X_4 - X_4M_\phi \end{pmatrix} \quad (\text{I})$$

and

$$N^*X - XN^* = \begin{pmatrix} D^*X_1 - X_1D^* & D^*X_2 - X_2M_\phi^* \\ M_\phi^*X_3 - X_3D^* & M_\phi^*X_4 - X_4M_\phi^* \end{pmatrix}. \quad (\text{II})$$

Clearly then

$$\begin{aligned} \|NX - XN\|_2^2 &= \|DX_1 - X_1D\|_2^2 + \|DX_2 - X_2M_\phi\|_2^2 \\ &\quad + \|M_\phi X_3 - X_3D\|_2^2 + \|M_\phi X_4 - X_4M_\phi\|_2^2; \end{aligned} \quad (\text{I}')$$

and

$$\begin{aligned} \|N^*X - XN^*\|_2^2 &= \|D^*X_1 - X_1D^*\|_2^2 + \|D^*X_2 - X_2M_\phi^*\|_2^2 \\ &\quad + \|M_\phi^*X_3 - X_3D^*\|_2^2 + \|M_\phi^*X_4 - X_4M_\phi^*\|_2^2. \end{aligned} \quad (\text{II}')$$

The following theorem relates (I'), (II') and (I) to statements (3)–(6).

**THEOREM 2.** (a) For every diagonalizable operator  $D$ , and  $X \in L(H)$ ,  $\|DX - XD\|_2 = \|D^*X - XD^*\|_2$ .

(b) For every normal operator  $N$ , diagonalizable operator  $D$ , and  $X \in L(H)$ ,  $\|DX - XN\|_2 = \|D^*X - XN^*\|_2$  and  $\|XD - NX\|_2 = \|XD^* - N^*X\|_2$ .

(c) To prove any of the statements (1)–(6), it is necessary and sufficient to prove the corresponding statement for the special case when  $N = M_\phi$ , where  $\phi(z) \in L^\infty(T)$  and  $H = L^2(T)$ .

**PROOF.** A simple computation proves part (a). It may also be found in [9, p. 147] or [10]. To prove part (c), consider separately each of the statements (1)–(6). To obtain (1)–(2), use the Laurent decomposition for a normal operator. To obtain (3)–(5) consider (I') and (II') and to obtain (6) consider (I). The proof of part (b) is not so easy. We give the proof next. First of all, it clearly suffices to prove the first equality, since the second equality follows from the first one by using adjoints.

In the basis which diagonalizes  $D$ , let  $\langle d_n \rangle_{n=1}^\infty$  denote the diagonal entries of  $D$ , and let  $N = (n_{ij})$  be the matrix for  $N$  and  $X = (x_{ij})$  be the matrix for  $X$ . Then

$$\|DX - XN\|_2^2 = \sum_{i,j=1}^\infty \left| d_i x_{ij} - \sum_k x_{ik} n_{kj} \right|^2;$$

and

$$\|D^*X - XN^*\|_2^2 = \sum_{i,j=1}^\infty \left| \bar{d}_i x_{ij} - \sum_k x_{ik} n_{kj}^* \right|^2,$$

where  $n_{kj}^* = \overline{n_{jk}}$  are the matrix entries of  $N^*$ . To show that these quantities are equal, it suffices to show that for every fixed  $i$ , we can obtain

$$\sum_{j=1}^\infty \left| d_i x_{ij} - \sum_k x_{ik} n_{kj} \right|^2 = \sum_{j=1}^\infty \left| \bar{d}_i x_{ij} - \sum_k x_{ik} n_{kj}^* \right|^2.$$

To prove this, let  $x^{(i)} = (x_{i1}, x_{i2}, x_{i3}, \dots)$  denote the  $i$ th row vector of  $(x_{ij})_{j=1}^\infty$ . Let

$$X^{(i)} = \begin{pmatrix} 0 & \begin{pmatrix} x_{i1}x_{i2} & \cdots \\ 0 & 0 & \cdots \\ \cdot & \cdot & \cdot \end{pmatrix} \\ 0 & 0 \end{pmatrix} \text{ and } N^{(i)} = \begin{pmatrix} d_i I & 0 \\ 0 & N \end{pmatrix}.$$

Then straightforward computations show that  $\sum_{j=1}^\infty |d_i x_{ij} - \sum_k x_{ik} n_{kj}|^2$  equals the square of the  $l^2$ -norm of the row vector  $d_i x^{(i)} - x^{(i)}N$ , and this equals the square of the Hilbert-Schmidt norm of  $N^{(i)}X^{(i)} - X^{(i)}N^{(i)}$ . Similarly,  $\sum_{j=1}^\infty |\bar{d}_i x_{ij} - \sum_k x_{ik} n_{kj}^*|^2 = \|N^{(i)*}X^{(i)} - X^{(i)}N^{(i)*}\|_2^2$ . In this case, however,  $X^{(i)}$  is a rank 1 matrix with its only nonzero row equal to  $x^{(i)} \in l^2$  (since  $X \in L(H)$ ). Therefore  $X^{(i)} \in C_2$ . Also  $N^{(i)}$  is clearly normal. In [9, p. 147, Theorem 8e] and in [10] we showed that under these circumstances, we have

$$\|N^{(i)}X^{(i)} - X^{(i)}N^{(i)}\|_2 = \|N^{(i)*}X^{(i)} - X^{(i)}N^{(i)*}\|_2.$$

Briefly, this is true since  $N^{(i)}$  is a normal operator and therefore it must be the uniform limit of diagonalizable operators. The latter equality is true replacing  $N^{(i)}$  by a diagonalizable operator, by part (a) of this theorem. Then we can take the  $l^2$ -norm limit of these replacement commutators due to the fact that if  $A_n \rightarrow A$  uniformly and  $Y \in C_2$  then  $A_n Y \rightarrow AY$  and  $YA_n \rightarrow YA$  in the Hilbert-Schmidt norm. This proves that the necessary sums are equal, and thus proves the theorem. Q.E.D.

Before developing the main technique, we are able to obtain a corollary to Theorem 2 that bears directly on the problem of I. D. Berg (statement (2)). Recall that  $\cong$  denotes unitary equivalence and  $\sim$  denotes similarity. Berg's problem asks if for every normal operator  $N$ , there exists a diagonalizable operator  $D$ , and  $K \in C_2$  such that  $N \cong D + K$ . The next corollary shows that Berg's problem is equivalent to the corresponding problem relative to similarity.

**COROLLARY 3.** *If  $N$  is a normal operator and  $D$  is a diagonalizable operator, then  $N \cong D + K$  for some  $K \in C_2$  if and only if  $N \sim D + K_1$  for some  $K_1 \in C_2$ .*

**PROOF.** One implication is trivial. For the other implication, assume  $N \sim D + K_1$ , with  $K_1 \in C_2$ . Then there exists an invertible operator  $S$  so that  $N = S^{*-1}(D + K_1)S^*$ , equivalently

$$DS^* - S^*N \in C_2. \quad (\dagger)$$

Applying Theorem 2(b), we obtain  $D^*S^* - S^*N^* \in C_2$ , or equivalently

$$NS - SD \in C_2. \quad (\dagger\dagger)$$

By  $(\dagger)$ , we obtain

$$DS^*S - S^*NS = D|S|^2 - S^*NS \in C_2. \quad (\dagger\dagger\dagger)$$

Applying  $(\dagger\dagger)$  to this, we get  $D|S|^2 - S^*SD = D|S|^2 - |S|^2D \in C_2$ .

We claim  $D|S|^2 - |S|^2D \in C_2$  implies  $D|S| - |S|D \in C_2$ , assuming  $S$  is invertible. (The proof of this fact works even if  $D$  is an arbitrary operator in  $L(H)$ .) Apply the Weyl-von Neumann Theorem [8] to the positive operator  $|S|$  to obtain  $|S| \cong D(\lambda_n) + K_\epsilon$ , where  $0 < \lambda_n < \|S\|$  and  $D(\lambda_n)$  denotes the diagonal matrix with the nonnegative diagonal sequence  $\langle \lambda_n \rangle_{n=1}^\infty$ , and  $K_\epsilon \in C_2$  with  $\|K_\epsilon\|_2 < \epsilon$  ( $\epsilon$  remains to be chosen). Hence  $|S|^2 \cong D(\lambda_n^2) + K'$ , where  $K' \in C_2$ . Because  $S$  is invertible,  $|S|$  is invertible, and so, bounded below. Choose  $\epsilon > 0$  so that  $|S|$  is bounded below by  $2\epsilon$ . Then since  $|S| \cong D(\lambda_n) + K_\epsilon$ , we obtain  $D(\lambda_n) \cong |S| - K'_\epsilon$  with  $\|K'_\epsilon\|_2 < \epsilon$  and

$$\begin{aligned} \lambda_n &= \| |S|e_n - K'_\epsilon e_n \|_H \geq \| |S|e_n \|_H - \| K'_\epsilon e_n \|_H \\ &\geq 2\epsilon - \| K'_\epsilon \|_{L(H)} \geq 2\epsilon - \| K'_\epsilon \|_2 \geq \epsilon. \end{aligned}$$

Let  $D = (D_{ij})$  denote the matrix for  $D$  with respect to that basis which diagonalizes  $D(\lambda_n)$ . Then  $D|S|^2 - |S|^2D = DD(\lambda_n^2) - D(\lambda_n^2)D + DK' -$

$K'D$ , and so  $D|S|^2 - |S|^2D \in C_2$  and  $K' \in C_2$  imply  $DD(\lambda_n^2) - D(\lambda_n^2)D \in C_2$ . Therefore,

$$\begin{aligned} \infty &> \|DD(\lambda_n^2) - D(\lambda_n^2)D\|_2^2 = \sum_{i,j=1}^{\infty} |(\lambda_i^2 - \lambda_j^2)D_{ij}|^2 \\ &= \sum_{i,j=1}^{\infty} (\lambda_i + \lambda_j)^2 |(\lambda_i - \lambda_j)D_{ij}|^2 \geq \sum_{i,j=1}^{\infty} (2\varepsilon)^2 |(\lambda_i - \lambda_j)D_{ij}|^2 \\ &= 4\varepsilon^2 \|DD(\lambda_n) - D(\lambda_n)D\|_2^2. \end{aligned}$$

Therefore  $D|S| - |S|D - (DK'_\varepsilon - K'_\varepsilon D) = DD(\lambda_n) - D(\lambda_n)D \in C_2$ . Since  $K'_\varepsilon \in C_2$ , we get  $D|S| - |S|D \in C_2$ , which proves the claim.

Let  $S = U|S|$  be the polar decomposition for  $S$ . The invertibility of  $S$  guarantees that  $|S|$  is invertible and  $U$  is unitary. Substituting this in (†) we obtain  $D|S|U^* - |S|U^*N \in C_2$ . But we now also have that  $D|S| - |S|D \in C_2$ . Therefore  $|S|DU^* - |S|U^*N \in C_2$ . Multiplying by  $|S|^{-1}$ , we obtain  $DU^* - U^*N \in C_2$  or equivalently,  $N - UDU^* \in C_2$ , with  $U$  a unitary operator. Q.E.D.

**The main construction.** In this construction, we use the notation that was introduced earlier.

By virtue of Theorem 2(c) we devote our attention to  $M_\phi X - XM_\phi$ , where  $\phi \in L^\infty(T)$  and  $X \in L(L^2(T))$ . In addition, if  $NX - XN$  is a trace class operator and  $X \in C_2$ , then from the matrix computation I, it is easy to see that (for  $N = D \oplus M_\phi$ )  $M_\phi X_4 - X_4 M_\phi$  must be a trace class operator with  $X_4 = P_{\mathcal{O}\mathbb{R}^\perp} X P_{\mathcal{O}\mathbb{R}^\perp} \in C_2$ , and  $\text{Trace}(NX - XN) = \text{Trace}(M_\phi X_4 - X_4 M_\phi)$  (since  $\text{Trace}(DX_1 - X_1 D) = 0$ ).

What is the matrix for  $M_\phi$ ? Let  $\phi(z) = \sum_{n=-\infty}^{\infty} \phi_n z^n$  denote the Fourier series for  $\phi$ . Then  $(M_\phi)_{i,j} = (M_\phi z^j, z^i) = \int_T \phi(z) z^{j-i} = \phi_{j-i}$ . The  $k$ th diagonal ( $k = 0, \pm 1, \pm 2, \dots$ ) in this 2-way infinite matrix is described by the set of all entries  $(i, j)$  for which  $j - i = k$ . In other words, the matrix for  $M_\phi$  is a Laurent matrix. Its entries are constant on the diagonals, and those constants are the Fourier coefficients of  $\phi$ .

Let us now introduce generating functions for matrix operators. They are related to Schwartz kernels in distribution theory.

**DEFINITION.** Let  $X = (x_{ij}) \in L(L^2(T))$ . The *generating function* for  $X$  is defined as the formal Fourier series given by  $F(z, w) = \sum_{i,j=-\infty}^{\infty} x_{ij} z^i w^j$ .

It is well known in the theory of distribution that since  $|x_{ij}| \leq \|X\|$ , the uniform boundedness of  $x_{ij}$  allows us to view  $F$  as a distribution on  $C^\infty(T)$ , where  $T^2$  denotes the torus. In particular,  $\langle F, z^i w^j \rangle = x_{ij}$ . It is also well-known that if  $\phi(z) \in C^\infty(T)$ , then  $\phi(z)F(z, w)$  (the formal power series product) is also a distribution on  $C^\infty(T^2)$ . We need to extend this definition to include all functions  $\phi(z)$  or  $\phi(w)$  in the larger class  $L^\infty(T)$ . By way of motivation, suppose  $F$  could be thought of as a function. For example, suppose  $F \in L^2(T^2)$  (equivalently  $X \in C_2$ ). Then formal computations hold

true and yield

$$\begin{aligned}\phi(z)F(z, w) &= \left( \sum_n \phi_n z^n \right) \left( \sum_{i,j} x_{ij} z^i w^j \right) \\ &= \sum_{i,j} \left( \sum_{n+k=i} \phi_n x_{kj} z^n z^k \right) w^j = \sum_{i,j} \left( \sum_n \phi_n x_{i-n,j} \right) z^i w^j\end{aligned}$$

and, similarly,  $\phi(w)F(z, w) = \sum_{i,j} (\sum_n \phi_n x_{i,j-n}) z^i w^j$ . Since  $\langle \phi_n \rangle$  and the columns of  $(x_{ij})$  are sequences in  $l^2$  (whether or not  $X \in C_2$ ), the expression  $\sum_n \phi_n x_{i,j-n}$  is a well-defined absolutely convergent series. In other words, the following operation is well defined.

DEFINITION. Let  $\phi, \psi \in L^\infty(T)$  where  $\phi(z) = \sum_n \phi_n z^n$  and  $\psi(z) = \sum_n \psi_n z^n$ , and let  $X \in L(L^2(T))$  so that  $F(z, w) = \sum_{i,j} x_{ij} z^i w^j$  is the generating function for  $X$ . Define the binary operation  $*$  as follows

$$[\phi(z) + \psi(w)] * F(z, w) = \sum_{i,j} \left( \sum_n (\phi_n x_{i-n,j} + \psi_n x_{i,j-n}) \right) z^i w^j.$$

It is helpful to recognize that  $*$  simply denotes the formal product of these power series and that this same symbol is used to denote formal products in some computer languages.

Also the reader should take care not to confuse this symbol with the symbol for operator adjoints.

Note. It is clear that  $(\phi(z) + \psi(w)) * F = \phi(z) * F + \psi(w) * F$ , where the sums in this equation are well-defined formal sums.

Let us now compute the generating function for  $M_\phi X - XM_\phi$ .

$$\begin{aligned}(M_\phi X)_{i,j} &= ((\phi_j - i)(x_{ij}))_{i,j} = \sum_k \phi_{k-i} x_{kj} = \sum_n \phi_n x_{i+n,j} \quad \text{and} \\ (XM_\phi)_{i,j} &= \sum_k x_{ik} \phi_{j-k} = \sum_n \phi_n x_{i,j-n}.\end{aligned}$$

Also,  $M_\phi^* = M_{\phi^*}$ , where  $\phi^*(z) = \sum_n \bar{\phi}_{-n} z^n$ , and  $(M_\phi^*)_{i,j} = \bar{\phi}_{i-j}$ . This gives us the following information about  $M_\phi^* X - XM_\phi^*$ .

$$\begin{aligned}(M_\phi^* X)_{i,j} &= ((\bar{\phi}_{i-j})(x_{ij}))_{i,j} = \sum_k \bar{\phi}_{i-k} x_{kj} = \sum_n \bar{\phi}_n x_{i-n,j} \quad \text{and} \\ (XM_\phi^*)_{i,j} &= \sum_k x_{ik} \bar{\phi}_{k-j} = \sum_n \bar{\phi}_n x_{i,j+n}.\end{aligned}$$

So

$$\begin{aligned}(M_\phi X - XM_\phi)_{i,j} &= \sum_n \phi_n (x_{i+n,j} - x_{i,j-n}) \quad \text{and} \\ (M_\phi^* X - XM_\phi^*)_{i,j} &= \sum_n \bar{\phi}_n (x_{i-n,j} - x_{i,j+n}).\end{aligned}$$

Now regard  $F(z, w) = \sum x_{ij} z^i w^j$  as a distribution on  $C^\infty(T^2)$ . Then a

computation shows

$$\begin{aligned}x_{i+n,j} &= \langle \bar{z}^n F, z^i w^j \rangle, & x_{i,j-n} &= \langle w^n F, z^i w^j \rangle, \\x_{i-n,j} &= \langle z^n F, z^i w^j \rangle, & x_{i,j+n} &= \langle \bar{w}^n F, z^i w^j \rangle.\end{aligned}$$

An additional computation shows

$$(M_\phi X - XM_\phi)_{i,j} = \sum_n \phi_n \langle (\bar{z}^n - w^n) F, z^i w^j \rangle = \langle (\phi(\bar{z}) - \phi(w)) * F, z^i w^j \rangle,$$

and

$$(M_\phi^* X - XM_\phi^*)_{i,j} = \sum_n \bar{\phi}_n \langle (z^n - \bar{w}^n) F, z^i w^j \rangle = \langle \overline{(\phi(\bar{z}) - \phi(w))} * F, z^i w^j \rangle.$$

This says that *the generating function for  $M_\phi X - XM_\phi$  is  $(\phi(\bar{z}) - \phi(w)) * F(z, w)$  and the generating function for  $M_\phi^* X - XM_\phi^*$  is  $\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)$* . (Note, the equalities above are best proven by computing the last expressions first, in terms of  $\phi_n$  and  $x_{ij}$ .)

This completes the construction of the generating function for the commutators. We now apply them in two settings, namely with regard to statements (5) and (6). First we prove statement (5).

**Fuglede's theorem modulo  $C_2$ .**

**THEOREM 4.** *If  $N$  is normal,  $X \in L(H)$ , and  $NX - XN \in C_2$  and  $N^*X - XN^* \in C_2$ , then  $\|NX - XN\|_2 = \|N^*X - XN^*\|_2$*

**PROOF.** By Theorem 2(c) it suffices to prove Theorem 4 when  $N = M_\phi$  acting on  $L^2(T)$ , with  $\phi \in L^\infty(T)$ , and such that  $M_\phi$  has no eigenvalues. By the main construction, the assumption on  $M_\phi$  and  $X$  is equivalent to the assumptions that  $(\phi(\bar{z}) - \phi(w)) * F(z, w) \in L^2(T^2)$  and  $\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w) \in L^2(T^2)$ . Of course, here we are treating those formal Fourier series in  $z, w$  which have square summable coefficients as functions in  $L^2(T^2)$ . Therefore the entries of  $M_\phi X - XM_\phi$  are the coefficients of  $(\phi(\bar{z}) - \phi(w)) * F(z, w)$ , and by Bessel's equality, satisfy

$$\begin{aligned}\|M_\phi X - XM_\phi\|_2^2 &= \sum_{i,j} |(M_\phi X - XM_\phi)_{i,j}|^2 \\&= \iint_{T^2} |(\phi(\bar{z}) - \phi(w)) * F(z, w)|^2.\end{aligned}$$

Similarly,

$$\|M_\phi^* X - XM_\phi^*\|_2^2 = \iint_{T^2} |\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)|^2.$$

Now the assumption that  $M_\phi$  has no eigenvectors is needed. It guarantees that  $\phi(\bar{z}) \neq \phi(w)$  almost everywhere with respect to 2-dimensional Lebesgue measure on  $T^2$ . To see this, let  $E = \{(z, w) \in T^2: \phi(\bar{z}) = \phi(w)\}$ . If, on the



contrary,  $m \times m(E) \neq 0$ , then  $\int_{T^2} \chi_E(z, w) \neq 0$ . Fubini's Theorem guarantees that there exists  $w_0 \in T$  such that  $0 \neq \int \chi_E(z, w_0) = mE_{w_0} \equiv m\{z \in T: (z, w_0) \in E\}$ . But then  $\chi_{E_{w_0}}(z)$  is an eigenvector for  $M_\phi$  since  $M_\phi \chi_{E_{w_0}} = \phi(z) \chi_{E_{w_0}}(z) = \phi(w_0) \chi_{E_{w_0}}$ , which is a contradiction. Therefore we know that  $\overline{\phi(\bar{z}) - \phi(w)} / (\phi(\bar{z}) - \phi(w)) \neq 0$  almost everywhere in  $T^2$ , and thus  $\overline{(\phi(\bar{z}) - \phi(w))} / (\phi(\bar{z}) - \phi(w))$  is a measurable function in  $T^2$  with modulus 1 almost everywhere in  $T^2$ .

Using this function we obtain

$$\begin{aligned} (*) &= \iint_{T^2} |(\phi(\bar{z}) - \phi(w)) * F(z, w)|^2 \\ &= \iint_{T^2} \left| \frac{\overline{\phi(\bar{z}) - \phi(w)}}{\phi(\bar{z}) - \phi(w)} ((\phi(\bar{z}) - \phi(w)) * F(z, w)) \right|^2 \\ &= \iint_{T^2} \left| \frac{1}{\phi(\bar{z}) - \phi(w)} \overline{(\phi(\bar{z}) - \phi(w))} ((\phi(\bar{z}) - \phi(w)) * F(z, w)) \right|^2. \end{aligned}$$

In addition, for every normal operator  $N$ , the derivations  $\delta_N$  and  $\delta_{N^*}$  commute (the proof is simple algebra). Hence,  $M_\phi^*(M_\phi X - XM_\phi) - (M_\phi X - XM_\phi)M_\phi^* = M_\phi(M_\phi^* X - XM_\phi^*) - (M_\phi^* X - XM_\phi^*)M_\phi$ . The generating function for the left-hand side of this equality is given by  $\overline{(\phi(\bar{z}) - \phi(w))} * ((\phi(\bar{z}) - \phi(w)) * F(z, w))$ , which is the same formal Fourier series as  $(\phi(\bar{z}) - \phi(w))((\phi(\bar{z}) - \phi(w)) * F(z, w))$  because of the assumption that  $(\phi(\bar{z}) - \phi(w)) * F(z, w)$  is a function in  $L^2(T^2)$ . Similarly, the generating function for the right-hand side of the equality is given by

$$\begin{aligned} &(\phi(\bar{z}) - \phi(w)) * (\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)) \\ &= (\phi(\bar{z}) - \phi(w))(\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)). \end{aligned}$$

This last equality follows from the assumption that  $\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)$  is also a function in  $L^2(T^2)$ . Hence  $\overline{(\phi(\bar{z}) - \phi(w))}((\phi(\bar{z}) - \phi(w)) * F(z, w))$  is a power series identical to  $(\phi(\bar{z}) - \phi(w))(\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w))$ . Thus

$$\begin{aligned} (*) &= \iint_{T^2} \left| \frac{1}{\phi(\bar{z}) - \phi(w)} (\phi(\bar{z}) - \phi(w))(\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)) \right|^2 \\ &= \iint_{T^2} |\overline{(\phi(\bar{z}) - \phi(w))} * F(z, w)|^2. \end{aligned}$$

Whence  $\|M_\phi X - XM_\phi\|_2^2 = \|M_\phi^* X - XM_\phi^*\|_2^2$ . Q.E.D.

**COROLLARY 5.** *Statement (2)  $\rightarrow$  statement (3). Also statement (3)  $\leftrightarrow$  statement (4).*

**PROOF.** (2)  $\rightarrow$  (3): Since (2)  $\rightarrow$  (4), then  $NX - XN \in C_2$  implies  $N^*X -$

$XN^* \in C_2$ . Therefore, in the case  $NX - XN \in C_2$ , we have  $N^*X - XN^* \in C_2$  and Theorem 4 applies to give  $\|NX - XN\|_2 = \|N^*X - XN^*\|_2$ . In the case  $NX - XN \notin C_2$ , assuming statement (2) gives  $N^*X - XN^* \notin C_2$ , hence  $\|NX - XN\|_2 = \infty = \|N^*X - XN^*\|_2$ .

(3)  $\leftrightarrow$  (4): It clearly suffices to prove (4)  $\rightarrow$  (3). If  $N$  is normal and  $X \in L(H)$  so that  $NX - XN \in C_2$ , then assuming statement (4) gives  $N^*X - XN^* \in C_2$  and Theorem 4 gives equality of their Hilbert-Schmidt norms. Q.E.D.

REMARK. It is hoped that distribution theory techniques might be used to show that if  $(\phi(\bar{z}) - \phi(w)) * F(z, w) \in L^2(T^2)$ , then  $(\phi(\bar{z}) - \phi(w)) * F(z, w)$  may be viewed in some way as a function. This could lead to proofs of statements (3) and (4).

We next prove statement (6).

**The trace of  $NX - XN$ .** Here we assume  $N$  is a normal operator and  $X$  is a Hilbert-Schmidt operator for which  $NX - XN$  is a trace class operator. We shall prove that  $\text{trace}(NX - XN) = 0$ . Applying Theorem 2(c) we see that to prove statement (6), it suffices to prove it for the special case when  $N = M_\phi$  ( $\phi \in L^\infty(T)$ ) acting on  $H = L^2(T)$ . To reiterate, this is because  $NX - XN \in C_1$  implies  $M_\phi X_4 - X_4 M_\phi \in C_1$  and  $\text{trace}(NX - XN) = \text{trace}(M_\phi X_4 - X_4 M_\phi)$ , where  $M_\phi$  is a Laurent part of  $N$  from Lemma 1 and  $X_4$  is as described earlier in equation (I), with  $X_4 \in C_2$  when  $X \in C_2$ .

If  $X \in C_2$ , the generating function for  $X$ , namely  $F(z, w)$ , is contained in  $L^2(T^2)$ . Therefore the generating function for  $M_\phi X - XM_\phi$  is  $(\phi(\bar{z}) - \phi(w)) * F(z, w) = (\phi(\bar{z}) - \phi(w))F(z, w)$  (the ordinary product of functions in  $L^\infty(T^2)$  and  $L^2(T^2)$ , respectively).

We have  $M_\phi X - XM_\phi = K = (K_{ij}) \in C_1$ . This implies  $(\phi(\bar{z}) - \phi(w))F(z, w) = K(z, w)$  almost everywhere in  $T^2$ , where  $K(z, w)$  denotes the generating function for  $K$  (since  $K \in C_1 \subset C_2$ , we see that  $K(z, w) \in L^2(T^2)$  and so it is indeed a function having the same Fourier coefficients as  $(\phi(\bar{z}) - \phi(w))F(z, w)$ ). Then, as before,  $\phi(\bar{z}) - \phi(w) \neq 0$  almost everywhere in  $T^2$ , and for almost all  $(z, w) \in T^2$ ,

$$F(z, w) = \frac{K(z, w)}{\phi(\bar{z}) - \phi(w)} \in L^2(T^2).$$

Our program is as follows. We shall assume  $\text{trace } K \neq 0$ . This will imply certain facts about  $K(z, w)$ , whose form we shall have to choose carefully to be able to relate it to the trace condition on  $K$ . We shall then apply a function theoretic result (which we shall give next) to conclude that  $K(z, w)/(\phi(\bar{z}) - \phi(w)) \notin L^2(T^2)$  in order to obtain a contradiction.

THEOREM 6. If  $\phi(x) \in L^\infty[0, 1]$  then

$$\iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy = \infty.$$

*Note.* The author first found a proof in the real case (by applying the Weyl-von Neumann Theorem to generating functions). Then Hugh Montgomery gave a somewhat involved proof of the theorem for  $\phi$  complex-valued. Then Larry Brown suggested an easy and direct proof when  $\phi$  is real-valued. Finally, the author found a relatively easy proof for the general complex case.

**PROOF.** By way of motivation, and to illustrate a connection between the real phenomenon and the planar phenomenon, we present Larry Brown's proof for the case when  $\phi$  is real-valued. By replacing  $\phi$  by  $a\phi + b$  (for some real numbers  $a, b$ ), we may assume  $\text{range } \phi \subset [0, 1)$ . Fix a positive integer  $n$ . Let  $E_k = \{x \in [0, 1]: (k-1)/n \leq \phi(x) < k/n\}$ , for  $1 \leq k \leq n$ . Then  $\cup E_k = [0, 1)$  (the symbol  $\cup$  denotes a disjoint union) and so  $\sum_{k=1}^n mE_k = 1$ . Hölder's inequality yields  $1 = (\sum_{k=1}^n mE_k)^2 \leq \sum_{k=1}^n (mE_k)^2 \cdot \sum_{k=1}^n 1$ , that is  $\sum_{k=1}^n (mE_k)^2 \geq 1/n$ . Also if  $(x, y) \in E_k \times E_k$ , then  $|\phi(x) - \phi(y)| \leq 1/n$ . Hence

$$\begin{aligned} \iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy &\geq \iint_{\cup E_k \times E_k} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \\ &= \sum_{k=1}^n \iint_{E_k \times E_k} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \geq \sum_{k=1}^n \iint_{E_k \times E_k} \frac{1}{(1/n)^2} dx dy \\ &= n^2 \sum_{k=1}^n (mE_k)^2 \geq n^2 \cdot \frac{1}{n} = n. \end{aligned}$$

Since  $n$  is arbitrary, we have that

$$\iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy = \infty.$$

The case when  $\phi$  is complex-valued is somewhat involved. By replacing  $\phi$  by  $a\phi + b$  (for some complex numbers  $a, b$ ), we may assume  $\text{range } \phi \subset [0, 1) \times [0, 1)$  (where  $Q$  denotes  $[0, 1) \times [0, 1)$  viewed as imbedded in  $\mathbb{C}$ ). Let  $Q_{ij}^{(n)} = [(i-1)/2^n, i/2^n) \times [(j-1)/2^n, j/2^n)$  for  $1 \leq i, j \leq 2^n$ . Define  $E_{ij}^{(n)} = \phi^{-1}(Q_{ij}^{(n)})$ . Hence, for each  $n$ ,  $Q = \cup_{i,j} Q_{ij}^{(n)}$  and  $[0, 1) \times [0, 1) = \cup_{i,j} E_{ij}^{(n)}$ . Therefore  $\sum_{i,j} mE_{ij}^{(n)} = 1$  for each  $n$ . In addition, for each  $n$ ,  $\{Q_{ij}^{(n)}\}$  is a partition of  $Q$ , and the collection of these partitions is a nested sequence. Therefore for each  $n$ ,  $\{E_{ij}^{(n)}\}$  is a partition of  $[0, 1) \times [0, 1)$  and also the collection of these partitions is a nested sequence.

The following identities are easily verified:  $\cap_n \cup_{i,j} (Q_{ij}^{(n)} \times Q_{ij}^{(n)}) = \{(z, w) \in \mathbb{C} \times \mathbb{C}: z = w\}$ . Also, and following from this,  $\cap_n \cup_{i,j} (E_{ij}^{(n)} \times E_{ij}^{(n)}) = \{(x, y) \in [0, 1) \times [0, 1): \phi(x) = \phi(y)\}$  (call this set  $F$ ).

If  $F$  has positive measure, then clearly

$$\iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \geq \iint_F \frac{1}{|\phi(x) - \phi(y)|^2} dx dy = \infty.$$

Hence, we may assume  $m \times m(F) = 0$ .

As before, Hölder's inequality yields a useful fact, namely that

$$\sum_{i,j} (mE_{ij}^{(n)})^2 > \frac{(\sum_{i,j} mE_{ij}^{(n)})^2}{\sum_{1 \leq i,j \leq 2^n} 1} = \frac{1}{4^n}.$$

Note also that  $(x, y) \in E_{ij}^{(n)} \times E_{ij}^{(n)}$  implies  $\phi(x), \phi(y) \in Q_{ij}^{(n)}$ , and so  $|\phi(x) - \phi(y)| < \text{diameter } Q_{ij}^{(n)} = \sqrt{2}/2^n$ . If we now assume to the contrary that

$$\iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy < \infty,$$

and if for convenience we set  $E^{(n)} = \bigcup_{i,j} E_{ij}^{(n)} \times E_{ij}^{(n)}$ , then we have

$$\begin{aligned} \infty &> \iint_{[0,1] \times [0,1]} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \\ &\geq \iint_{E^{(n)}} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \\ &= \sum_{1 \leq i,j \leq 2^n} \iint_{E_{ij}^{(n)} \times E_{ij}^{(n)}} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy \\ &> \sum_{i,j=1}^{2^n} \iint_{E_{ij}^{(n)} \times E_{ij}^{(n)}} \frac{1}{(\sqrt{2}/2^n)^2} dx dy \\ &= \frac{4^n}{2} \sum_{i,j} (mE_{ij}^{(n)})^2 > \frac{4^n}{2} \cdot \frac{1}{4^n} = \frac{1}{2}. \end{aligned}$$

That is

$$\iint_{[0,1] \times [0,1]} \frac{\chi_{E^{(n)}}(x, y)}{|\phi(x) - \phi(y)|^2} dx dy = \iint_{E^{(n)}} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy > \frac{1}{2}$$

for every  $n$ . On the other hand,

$$\frac{\chi_{E^{(n)}}(x, y)}{|\phi(x) - \phi(y)|^2} \leq \frac{1}{|\phi(x) - \phi(y)|^2} \in L'([0, 1] \times [0, 1])$$

and

$$\frac{\chi_{E^{(n)}}(x, y)}{|\phi(x) - \phi(y)|^2} \rightarrow \frac{\chi_F(x, y)}{|\phi(x) - \phi(y)|^2} = 0$$

almost everywhere in  $[0, 1] \times [0, 1]$  (since  $m \times m(F) = 0$ ). Therefore, by the Lebesgue dominating convergence theorem,

$$\int_{[0,1] \times [0,1]} \frac{\chi_{E^{(n)}}(x, y)}{|\phi(x) - \phi(y)|^2} dx dy \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

and this is a contradiction, as we showed that these integrals are all bounded below by  $1/2$ . Q.E.D.

REMARK. If we view  $\phi$  as  $\phi: [0, 1] \rightarrow \mathbb{R}^n$  ( $n = 1$  in the real case, and  $n = 2$  in the complex case), then Hugh Montgomery has shown that if  $\phi \in L^\infty_{\mathbb{R}^n}(0, 1)$ , then it is possible to have, for  $n \geq 3$ ,

$$\iint \frac{1}{\|\phi(x) - \phi(y)\|_{\mathbb{R}^n}^2} dx dy < \infty.$$

Also, it is interesting to note that if we replace the power 2 by the power  $p$  in the integral of the theorem, then Larry Brown's proof for the real case works for all  $p > 1$ , but the proof for the complex case works only for  $p \geq 2$ . In addition, by modifying slightly the proof of the complex case, we obtain the real case for  $p = 1$ .

COROLLARY 7. *If  $E$  is any measurable subset of  $[0, 1]$  of positive measure and  $|\phi|$  is essentially bounded on  $E$ , then*

$$\iint_{E \times E} \frac{1}{|\phi(x) - \phi(y)|^2} dx dy = \infty.$$

PROOF. Use the proof of Theorem 6 and in it, replace the domain  $[0, 1]$  of  $\phi$  by  $E$ .

THEOREM 8. *If  $N$  is a normal operator,  $X \in C_2$  and  $NX - XN \in C_1$ , then  $\text{trace}(NX - XN) = 0$ .*

PROOF. Set  $NX - XN = K \in C_1$ . Then  $K$  has a Schmidt expansion [5]. That is, there exist orthonormal sequences  $\{f_n\}, \{g_n\}$  in  $H$  and a nonnegative, real-valued sequence  $\langle a_n \rangle \in l_+^1$  such that  $K = \sum a_n(f_n \otimes g_n)$ , where the series converges in the trace norm and  $f_n \otimes g_n$  denotes the rank one operator  $h \rightarrow (h, g_n)f_n$ . Note that the matrix of  $f_n \otimes g_n$  is  $(f_n(i)g_n(j))$  if  $\langle f_n(i) \rangle_{i=1}^\infty$  and  $\langle g_n(i) \rangle_{i=1}^\infty$  denote the sequences for  $f_n$  and  $g_n$  respectively.

The Schmidt expansion provides us with a useful form for the generating function for  $K$ . Since the Schmidt expansion converges in the trace norm to  $K$ , it must also converge in the Hilbert-Schmidt norm. Therefore the generating function for  $\sum_{n=1}^m a_n(f_n \otimes g_n)$  converges to  $K(z, w)$  in  $L^2(T^2)$ , as  $m \rightarrow \infty$ .

The generating function for  $f_n \otimes g_n$ , the matrix with entries  $(f_n(i)g_n(j))$ , is  $\sum_{i,j} f_n(i)g_n(j)z^i w^j = (\sum_i f_n(i)z^i)(\sum_j g_n(j)w^j)$ . Define  $f_n(z) \equiv \sum_i f_n(i)z^i$  and  $g_n(w) \equiv \sum_j g_n(j)w^j$ . Hence the generating function for  $f_n \otimes g_n$  is  $f_n(z)g_n(w)$  and  $\{f_n(z)\}$  and  $\{g_n(z)\}$  are orthonormal sequences in  $L^2(T)$ . Clearly then the generating function for  $\sum_{n=1}^m a_n f_n \otimes g_n$  is  $\sum_{n=1}^m a_n f_n(z)g_n(w)$ . Also it is clear that  $\sum_{n=1}^m a_n f_n(z)g_n(w)$  converges to  $\sum_n a_n f_n(z)g_n(w)$ , in  $L^2(T^2)$ . Therefore, the generating function for  $K$  can be written as  $K(z, w) = \sum a_n f_n(z)g_n(w)$ . That is, the Fourier coefficients of  $\sum a_n f_n(z)g_n(w)$  in  $L^2(T^2)$  are the matrix entries of  $K$ .

To see the connection between  $\text{trace } K$  and  $K(z, w)$ , recall that  $\sum_{n=1}^m a_n(f_n \otimes g_n) \rightarrow K$  in the trace norm as  $m \rightarrow \infty$ , and hence

$$\begin{aligned}\text{trace } K &= \lim_{m \rightarrow \infty} \text{trace} \sum_{n=1}^m a_n (f_n \otimes g_n) \\ &= \lim_{m \rightarrow \infty} \sum_{n=1}^m a_n \text{trace}(f_n \otimes g_n).\end{aligned}$$

Using the fact that the matrix for  $f_n \otimes g_n$  is  $(f_n(i) \overline{g_n(j)})$ , a simple calculation using Lebesgue integration on  $T$  shows that

$$\text{trace}(f_n \otimes g_n) = \int_T f_n(\bar{z}) g_n(z) dm.$$

Therefore

$$\text{trace } K = \lim_{m \rightarrow \infty} \int_T \left[ \sum_{n=1}^m a_n f_n(\bar{z}) g_n(z) \right] dm.$$

Also

$$\begin{aligned}\sum_{n=1}^{\infty} a_n \|f_n(\bar{z}) g_n(z)\|_{L^1(T)} &\leq \sum a_n \|f_n\|_{L^2(T)} \|g_n\|_{L^2(T)} \\ &= \sum a_n < \infty\end{aligned}$$

by Hölder's inequality, and so  $\sum_n a_n f_n(\bar{z}) g_n(z)$  converges in  $L^1(T)$ . Therefore  $\text{trace } K = \int_T [\sum_n a_n f_n(\bar{z}) g_n(z)] dm$ . In other words,

$$\text{trace } K = \int_T K(\bar{z}, z) dm.$$

If we now assume  $\text{trace } K = \int_T K(\bar{z}, z) dm \neq 0$ , we can conclude that  $\int_T |K(\bar{z}, z)| dm > |\int_T K(\bar{z}, z) dm| > 0$ . From elementary measure theory, it follows that there exist  $\varepsilon > 0$  and a measurable set  $A \subset T$  of strictly positive Lebesgue measure such that  $|K(\bar{z}, z)| > \varepsilon$  on  $A$ .

Our next objective is show that there exists a measurable set  $E \subset T$  of strictly positive Lebesgue measure such that  $|K(\bar{z}, w)| \geq \varepsilon/4$  on  $E \times E$ . The proof of this requires several steps, including the use of Egoroff's theorem, Lusin's theorem and convergence in measure.

Define  $K_m(z, w) = \sum_{n=1}^m a_n f_n(z) g_n(w)$ ,  $F_m(z) = \sum_{n=m+1}^{\infty} a_n |f_n(z)|^2$ , and  $G_m(z) = \sum_{n=m+1}^{\infty} a_n |g_n(z)|^2$  (recall that  $\langle a_n \rangle \in l_+^1$ ). Clearly then

$$\begin{aligned}|K(\bar{z}, w) - K_m(\bar{z}, w)| &= \left| \sum_{n=m+1}^{\infty} a_n f_n(\bar{z}) g_n(w) \right| \\ &\leq \sum_{n=m+1}^{\infty} a_n |f_n(\bar{z})| |g_n(w)| \\ &\leq \sum_{n=m+1}^{\infty} a_n [|f_n(\bar{z})|^2 + |g_n(w)|^2] \\ &= F_m(\bar{z}) + G_m(w).\end{aligned}$$

Now  $F_m, G_m \geq 0$  and  $\int F_m(\bar{z}) dm = \sum_{m+1}^{\infty} a_n = \int G_m(w) dm \rightarrow 0$  as  $m \rightarrow \infty$ , since  $f_n(\bar{z}), g_n(w) \in (L^2(T))_1$  and  $\langle a_n \rangle \in l_+^1$ . Therefore  $F_m(\bar{z}), G_m(w) \rightarrow 0$  in measure. Therefore there exists some subsequence  $m_k$  such that  $F_{m_k}(\bar{z}) \rightarrow 0$  almost everywhere in  $T$ . Clearly  $F_m(\bar{z}) \geq F_{m+1}(\bar{z}) \geq 0$  for all  $m$  and all  $z \in T$ . The two facts imply that  $F_m(\bar{z}) \rightarrow 0$  almost everywhere in  $T$ . Likewise  $G_m(w) \rightarrow 0$  almost everywhere in  $T$ . If we now apply Egoroff's theorem, we obtain a measurable set  $E_1$  and a positive integer  $m_1$  where  $m(T \setminus E_1) < mA/4$  and  $F_m(\bar{z}) < \varepsilon/16$  for every  $m \geq m_1$  and  $z \in E_1$ . Similarly we obtain a measurable set  $E_2$  and a positive integer  $m_2$  where  $m(T \setminus E_2) < mA/4$  and  $G_m(w) < \varepsilon/16$  for every  $m \geq m_2$  and  $w \in E_2$ . Let  $m_0 = \max(m_1, m_2)$  and  $E_0 = E_1 \cap E_2$ . Then  $m(T \setminus E_0) \leq m(T \setminus E_1) + m(T \setminus E_2) < mA/2$  and  $F_m(\bar{z}) + G_m(w) < \varepsilon/8$  on  $E_0 \times E_0$  for every  $m \geq m_0$ . Hence  $|K(\bar{z}, w) - K_m(\bar{z}, w)| \leq F_m(\bar{z}) + G_m(w) < \varepsilon/8$  on  $E_0 \times E_0$  for every  $m \geq m_0$ . In particular,  $|K(\bar{z}, w) - K_{m_0}(\bar{z}, w)| < \varepsilon/8$  on  $E_0 \times E_0$ .

We next apply Lusin's theorem to obtain information about  $K_{m_0}(\bar{z}, w)$ . The functions  $f_1(\bar{z}), \dots, f_{m_0}(\bar{z}), g_1(w), \dots, g_{m_0}(w) \in L^2(T)$ . If we apply Lusin's theorem to each one of these functions we see that for each such function, there is a set in  $T$  on which the function is continuous and whose complement has arbitrarily small measure. Hence we can insure the existence of a measurable set  $B \subset T$  on which  $f_1(\bar{z}), \dots, f_{m_0}(\bar{z})$  are continuous for all  $z \in B$  and  $g_1(w), \dots, g_{m_0}(w)$  are continuous for all  $w \in B$  and furthermore  $m(T \setminus B) < mA/2$ . Therefore  $\sum_{n=1}^{m_0} a_n f_n(\bar{z}) g_n(w) = K_{m_0}(\bar{z}, w)$  is continuous on  $B \times B$ .

Now we have  $m(T \setminus (E_0 \cap B)) \leq m(T \setminus E_0) + m(T \setminus B) < mA/2 + mA/2 = mA$ . From this it follows that  $m(A \cap E_0 \cap B) > 0$ . From elementary measure theory we can obtain a closed set  $F \subset A \cap E_0 \cap B$  of strictly positive measure. Hence  $|K(\bar{z}, w) - K_{m_0}(\bar{z}, w)| \leq \varepsilon/8$  on  $F \times F$ ,  $K_{m_0}(\bar{z}, w)$  is continuous on  $F \times F$ , and  $|K(\bar{z}, z)| \geq \varepsilon$  on  $F$ , with  $F$  closed and  $mF > 0$ . Therefore for every  $z \in F$ ,  $|K_{m_0}(\bar{z}, z)| > |K(\bar{z}, z)| - \varepsilon/8 > \varepsilon - \varepsilon/8 = 7\varepsilon/8$ . That is,  $|K_{m_0}(\bar{z}, z)| > \varepsilon/2$  on  $F$ .

We now employ the continuity of  $K_{m_0}(\bar{z}, w)$  on  $F \times F$ . For every  $z_0 \in F$ ,  $|K_{m_0}(\bar{z}_0, z_0)| > \varepsilon/2$  and the continuity of  $K_{m_0}(\bar{z}, w)$  on  $F \times F$  implies that there exists an open interval  $N(z_0)$  containing  $z_0$  such that  $|K_{m_0}(\bar{z}, w) - K_{m_0}(\bar{z}_0, z_0)| < \varepsilon/8$  on  $[N(z_0) \cap F] \times [N(z_0) \cap F]$ . The collection  $\{N(z): z \in F\}$  forms an open cover of  $F$ . Since  $F$  is closed and contained in  $T$ , it is compact. Therefore, we can extract a finite subcover  $N(z_1), \dots, N(z_k)$ . But  $mF = m(\cup_{n=1}^k (N(z_n) \cap F)) \leq \sum_{n=1}^k m(N(z_n) \cap F)$ . Since  $mF > 0$ , there is some  $n_0$  such that  $m(N(z_{n_0}) \cap F) > 0$ . Set  $E = N(z_{n_0}) \cap F$ .

We claim that  $E$  accomplishes our objective, namely that  $|K(\bar{z}, w)| \geq \varepsilon/4$  on  $E \times E$  and  $mE > 0$ . To see this we apply the last remarks. Since  $E \subset F$ , we have that  $|K(\bar{z}, w) - K_{m_0}(\bar{z}, w)| \leq \varepsilon/8$  on  $E \times E$ . Also,  $|K_{m_0}(\bar{z}, w) - K_{m_0}(\bar{z}_{n_0}, z_{n_0})| < \varepsilon/8$  on  $E \times E$ , by the construction of  $N(z_{n_0})$  and  $E$ .

Furthermore, since  $z_{n_0} \in E \subset F$ , we have that  $|K_{m_0}(\bar{z}_{n_0}, z_{n_0})| > \varepsilon/2$ . The last three inequalities imply that for every  $(z, w) \in E \times E$ ,  $|K(\bar{z}, w)| \geq |K_{m_0}(\bar{z}, w)| - \varepsilon/8 \geq |K_{m_0}(\bar{z}_{n_0}, z_{n_0})| - \varepsilon/8 - \varepsilon/8 \geq \varepsilon/2 - \varepsilon/8 - \varepsilon/8 = \varepsilon/4$ . That is, we have  $|K(\bar{z}, w)| \geq \varepsilon/4$  on  $E \times E$  and  $mE > 0$ .

The rest of the proof is straightforward. By Theorem 2c, it suffices to prove Theorem 8 for the case  $N = M_\phi$  ( $\phi \in L^\infty(T)$ ) where  $M_\phi$  has no eigenvectors, and  $X$  is any Hilbert-Schmidt operator contained in  $L(L^2(T))$ . Suppose  $M_\phi X - XM_\phi = K \in C_1$ . Let  $F(z, w)$  denote the generating function for  $X$  and let  $K(z, w)$  denote the special generating function of  $K$ , namely  $\sum a_n f_n(z) g_n(w)$  considered earlier in this proof. Then by earlier remarks, the generating function for  $M_\phi X - XM_\phi$ , which is  $(\phi(\bar{z}) - \phi(w))F(z, w)$ , is the same as the generating function for  $K$ , which is  $K(z, w)$ . That is,  $(\phi(\bar{z}) - \phi(w))F(z, w) = K(z, w)$  almost everywhere in  $T^2$ . Also we have seen that since  $M_\phi$  has no eigenvectors,  $\phi(\bar{z}) - \phi(w) \neq 0$  almost everywhere in  $T^2$ . Hence  $F(z, w) = K(z, w)/(\phi(\bar{z}) - \phi(w)) \in L^2(T^2)$ . Substituting  $\bar{z}$  for  $z$ , we obtain  $K(\bar{z}, w)/(\phi(\bar{z}) - \phi(w)) \in L^2(T^2)$ .

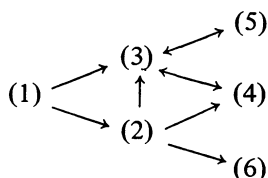
We now prove Theorem 8 by contradiction. If we assume to the contrary that  $\text{trace } K \neq 0$ , we can apply the previous result that there exists a measurable set  $E$  with  $mE > 0$  and  $\varepsilon > 0$  such that  $|K(\bar{z}, w)| > \varepsilon$  on  $E \times E$ . Finally, if we identify  $T$  with  $[0, 1]$  in the usual way, then we can apply Corollary 7 to obtain

$$\begin{aligned} \infty &> \iint_{T^2} \left| \frac{K(\bar{z}, w)}{\phi(\bar{z}) - \phi(w)} \right|^2 \geq \iint_{E \times E} \frac{|K(\bar{z}, w)|^2}{|\phi(\bar{z}) - \phi(w)|^2} \\ &\geq \varepsilon^2 \iint_{E \times E} \frac{1}{|\phi(\bar{z}) - \phi(w)|^2} = \infty, \end{aligned}$$

which is a contradiction. Q.E.D.

ADDENDUM. The author has recently proved that statement (3) (equivalently (4)) is true. That is, the Fuglede Commutativity Theorem Modulo the Hilbert-Schmidt class holds true. In fact, for every normal operator  $N$  and  $X \in L(H)$ ,  $\|NX - XN\|_2 = \|N^*X - XN^*\|_2$ .

In summary, statements (3)–(6) are true. In this paper, the following implications were proven to hold true.



ACKNOWLEDGMENTS. I should like to acknowledge the following mathematicians for their helpful suggestions and encouragements: Allen Shields, Jeffrey Rauch, Hugh Montgomery and Joel Anderson.



## REFERENCES

1. I. D. Berg, *An extension of the Weyl-von Neumann Theorem to normal operators*, Trans. Amer. Math. Soc. **160** (1971), 365–371.
2. C. A. Berger and B. I. Shaw, *Self-commutators of multicyclic hyponormal operators are always trace class*, Bull. Amer. Math. Soc. **79** (1973), 1193–1199.
3. L. G. Brown, R. G. Douglas and P. A. Fillmore, *Unitary equivalence modulo the compact operators and extensions of  $C^*$ -algebras*, Lecture Notes in Math., no. 345, Springer-Verlag, Berlin and New York, 1973, pp. 58–128.
4. R. W. Carey and J. D. Pincus, *Perturbation by trace class operators*, Bull. Amer. Math. Soc. **80** (1974), 758–759.
5. I. C. Gohberg and M. G. Krein, *Introduction to the theory of linear nonselfadjoint operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence, R. I., 1969.
6. P. R. Halmos, *Measure theory*, Van Nostrand, New York, 1950, pp. 171–174.
7. J. W. Helton and R. E. Howe, *Integral operators: commutators, traces, index and homology*, Lecture Notes in Math., no. 345, Springer-Verlag, Berlin and New York, 1973, pp. 141–209.
8. J. von Neumann, *Charakterisierung des Spektrums eines Integraloperators*, Actualités Sci. Indust., no. 229, Hermann, Paris, 1935.
9. Gary Weiss, *Commutators and operator ideals*, Dissertation, Univ. of Michigan, Ann Arbor, Mich., 1975.
10. ———, *Fuglede's commutativity theorem modulo operator ideals* (preprint).
11. H. Weyl, *Über beschränkte quadratischen Formen deren Differenz vollstetig ist*, Rend. Circ. Mat. Palermo **27** (1909), 373–392.

DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF CINCINNATI, CINCINNATI, OHIO 45221