J. OPERATOR THEORY
42(1999), 379–404

THE BERGER-SHAW THEOREM IN THE HARDY MODULE OVER THE BIDISK

RONGWEI YANG

Communicated by William B. Arveson

Abstract. It is well known that the Hardy space over the bidisk $\mathbb{D}^2$ is an $A(\mathbb{D}^2)$ module and that $A(\mathbb{D}^2)$ is contained in $H^2(\mathbb{D}^2)$. Suppose $(h) \subset A(\mathbb{D}^2)$ is the principal ideal generated by a polynomial h, then its closure $[h](\subset H^2(\mathbb{D}^2))$ and the quotient $H^2(\mathbb{D}^2) \ominus [h]$ are both $A(\mathbb{D}^2)$ modules. We let R_z, R_w be the actions of the coordinate functions z and w on $[h]$, and let S_z, S_w be the actions of z and w on $H^2(\mathbb{D}^2) \ominus [h]$. In this paper, we will show that R_z and R_w, as well as S_z and S_w, essentially doubly commute. Moreover, both $[R_w^*, R_z]$ and $[S_w^*, S_z]$ are actually Hilbert-Schmidt.

Keywords: *Bidisk, cross commutator, Hilbert-Schmidt, submodule.*

MSC (2000): Primary 47B38; Secondary 47A45.

0. INTRODUCTION

The Berger-Shaw theorem says that the self-commutator of a multicyclic hyponormal operator is trace class ([1]). It is interesting to study the multivariate analogue of this theorem. In [6], the authors reformulated the theorem in an algebraic language and showed that if the spectrum of a finite rank hyponormal module is contained in an algebraic curve then the module is *reductive*. They also gave examples showing that it is generally not the case if the spectrum of the module is of higher dimension. However, many examples show that the *cross* commutators do not seem to have a close relation with the spectra of modules and are generally "small". This suggests that the following general questions may have positive answers.

QUESTIONS. Suppose T_1, T_2 are two doubly commuting operators acting on a separable Hilbert space H and R_1, R_2 are the restrictions of them to a jointly invariant subspace that is finitely generated by T_1, T_2.

(1) Is the cross commutator $[R_1^*, R_2]$ in some Schatten p-class?

(2) Is the product $[R_1^*, R_1][R_2^*, R_2]$ also small?

(3) What about the compressions of T_1, T_2 to the orthogonal complement of M?

A special case of the first question was studied by Curto, Muhly and Yan in [3]. The second question was raised by R. Douglas. The third one appears naturally from the study of essentially reductive quotient modules. Note that when $T_1 = T_2$ the first two questions are answered positively by the Berger-Shaw Theorem.

In this paper we will make a study of these questions in the case $H = H^2(\mathbb{D}^2)$, the Hardy space over the bidisk, and T_1, T_2 are the multiplications by the two coordinate functions z and w. Then a closed subspace of $H^2(\mathbb{D}^2)$ is jointly invariant for T_1 and T_2 if and only if it is an $A(\mathbb{D}^2)$ submodule. We will have a look at the third question first because it turns out to be the easiest. The answer to the second question is a consequence of the answer to the first one. Some related questions will also be studied in this paper.

We now begin the study by doing some preparations.

Throughout this paper we let E', E be two separable Hilbert spaces of infinite dimension and $\{\delta_j' : j \geqslant 0\}$, $\{\delta_j : j \geqslant 0\}$ are orthonormal bases for E' and E respectively. We let $H^2(E)$ denote the E-valued Hardy space, i.e.

$$H^2(E) := \Big\{ \sum_{j=0}^{\infty} z^j x_j : |z| = 1, \sum_{j=0}^{\infty} \|x_j\|_E^2 < \infty \Big\}.$$

It is well known that every function in $H^2(E)$ has an analytic continuation to the whole unit disk $\mathbb{D}$. For our convenience, we will not distinguish the functions of $H^2(E)$ from their extensions to $\mathbb{D}$. We let T_z be the Toeplitz operator on $H^2(E)$ such that for any $f \in H^2(E)$,

$$T_z f(z) = z f(z).$$

One sees that T_z is a shift operator of infinite multiplicity.

A $B(E', E)$-valued analytic function $\theta(z)$ on $\mathbb{D}$ is called *left-inner (inner)* if its boundary values on the unit circle $\mathbb{T}$ are almost everywhere isometries (unitaries) from E' into E. Therefore, multiplication by a left-inner θ defines an isometry from $H^2(E')$ into $H^2(E)$.

A closed subspace $M \subset H^2(E)$ is called *invariant* if

$$T_z M \subset M.$$

The Lax-Halmos Theorem gives a complete discription of invariant subspaces in terms of left-inner functions.

Theorem 0.1. (Lax-Halmos) *M is a nontrivial invariant subspace of $H^2(E)$ if and only if there is a closed subspace $E' \subset E$ and a $B(E', E)$-valued left-inner function θ such that*

$$M = \theta H^2(E'). \tag{0.1}$$

The representation is unique in the sense that

$$\theta H^2(E') = \theta' H^2(E'') \Leftrightarrow \theta = \theta' V,$$

where V is a unitary from E' onto E''.

In order to make a study of the Hardy modules over the bidisk, we identify the space E with another copy of the Hardy space. Then $H^2(E) = H^2(\mathbb{D}) \otimes E$ will be identified with $H^2(\mathbb{D}) \otimes H^2(\mathbb{D}) = H^2(\mathbb{D}^2)$. We do this in the following way.

Let u be the unitary map from E to $H^2(\mathbb{D})$ such that

$$u\delta_j = w^j, \quad j \geqslant 0.$$

Then $U = I \otimes u$ is a unitary from $H^2(\mathbb{D}) \otimes E$ to $H^2(\mathbb{D}) \otimes H^2(\mathbb{D})$ such that

$$U(z^i \delta_j) = z^i w^j, \quad i, j \geqslant 0.$$

It is not hard to see that $M \subset H^2(E)$ is invariant if and only if $UM \subset H^2(\mathbb{D}^2)$ is invariant under multiplication by the coordinate function z. This identification enables us to use the Lax-Halmos theorem to study certain properties of sub-Hardy modules over the bidisk which we will do in Section 1. Throughout this paper, we will let $\mathrm{d}|z|$ denote the normalized Lebesgue measure on the unit circle $\mathbb{T}$ and $\mathrm{d}|z|\,\mathrm{d}|w|$ be the product measure on the torus $\mathbb{T}^2$.

1. HILBERT-SCHMIDT OPERATORS

In this section we prove two technical lemmas and an important corollary.

Suppose θ is left inner with values in $B(E', E)$ and δ is any fixed element of E. We now define an operator N from $\theta E'$ to the Hardy space $H^2(\mathbb{D})$ over the unit disk as the following:

$$N\Big(\theta(z)\sum_{j=0}^{\infty}\alpha_j\delta_j'\Big) := \Big\langle \theta(z)\sum_{j=0}^{\infty}\alpha_j\delta_j', \delta\Big\rangle_E, \tag{1.1}$$

where $\sum_{j=0}^{\infty}\alpha_j\delta_j'$ is any element in E'.

LEMMA 1.1. *N is Hilbert-Schmidt and*

$$\mathrm{tr}(N^*N) = \int_{\mathbb{T}} \|\theta^*(z)\delta\|_{E'}^2 \,\mathrm{d}|z|. \tag{1.2}$$

Proof. Since θ is left inner, $\{\theta\delta_j \mid j \geqslant 0\}$ is an orthonormal basis for $\theta E'$. To prove the lemma, one suffices to show that $\sum_{j=0}^{\infty}\langle N^*N\theta\delta_j', \theta\delta_j'\rangle_{\theta E'}$ is finite. In fact,

$$\begin{aligned}
\sum_{j=0}^{\infty}\langle N^*N\theta\delta_j', \theta\delta_j'\rangle_{\theta E'} &= \sum_{j=0}^{\infty}\langle N\theta\delta_j', N\theta\delta_j'\rangle_{H^2} = \sum_{j=0}^{\infty}\int_{\mathbb{T}} |\langle\theta(z)\delta_j', \delta\rangle_E|^2 \,\mathrm{d}|z| \\
&= \sum_{j=0}^{\infty}\int_{\mathbb{T}} |\langle\delta_j', \theta^*(z)\delta\rangle_{E'}|^2 \,\mathrm{d}|z| \\
&= \int_{\mathbb{T}}\sum_{j=0}^{\infty} |\langle\delta_j', \theta^*(z)\delta\rangle_{E'}|^2 \,\mathrm{d}|z| = \int_{\mathbb{T}} \|\theta^*(z)\delta\|_{E'}^2 \,\mathrm{d}|z|. \qquad \blacksquare
\end{aligned}$$

So in general

$$\mathrm{tr}(N^*N) \leqslant \|\delta\|^2,$$

and the equality holds when θ is inner.

Back to the $H^2(\mathbb{D}^2)$ case, this lemma has an important corollary. Let us first introduce some operators.

For any bounded function f we let $T_f := Pf$ be the Toeplitz operator on $H^2(\mathbb{D}^2)$, where P is the projection from $L^2(\mathbb{T}^2)$ to $H^2(\mathbb{D}^2)$. For every non-negative integer j and $\lambda \in \mathbb{D}$, we let operators N_j and N_λ from $H^2(\mathbb{D}^2)$ to $H^2(\mathbb{D})$ be such that for any $f(z, w) = \sum_{k=0}^{\infty} f_k(z)w^k \in H^2(\mathbb{D}^2)$

$$N_j f(z) = f_j(z), \quad N_\lambda f(z) = f(z, \lambda).$$

Then one verifies that N_j is a contraction for each j and $\|N_\lambda\| = (1-|\lambda|^2)^{-1/2}$. Furthermore,

$$\sum_{k=0}^{\infty} T_{w^k} N_k = I \quad \text{on } H^2(\mathbb{D}^2), \tag{1.3}$$

$$N_\lambda = \sum_{k=0}^{\infty} \lambda^k N_k. \tag{1.4}$$

In what follows we will be mainly interested in the restrictions of N_k, N_λ to certain subspaces and will use the same notations to denote these restrictions.

Corollary 1.2. *For any $A(\mathbb{D}^2)$ submodule $M \subset H^2(\mathbb{D}^2)$, N_j and N_λ are Hilbert-Schmidt operators restricting on $M \ominus zM$ for each $j \geqslant 0$ and $\lambda \in \mathbb{D}$, and*

$$\operatorname{tr}(N_j^* N_j) \leqslant 1,$$

$$\Big\| p_\perp \frac{1}{1-\overline{\lambda} w} \Big\|^2 \leqslant \operatorname{tr}(N_\lambda^* N_\lambda) \leqslant (1-|\lambda|^2)^{-1},$$

where $p_\perp$ is the projection from $H^2(\mathbb{D}^2)$ onto $M \ominus zM$.

Proof. Because M is invariant under the multiplication by z, U^*M is invariant under T_z, where U is defined in the last paragraph of Section 0, and hence

$$U^*M = \theta H^2(E')$$

for some Hilbert space E' and a left inner function θ. Then

$$U^*(M \ominus zM) = \theta H^2(E') \ominus z\theta H^2(E') = \theta(H^2(E') \ominus zH^2(E')) = \theta E'.$$

Let us first deal with the operator N_λ.

In Lemma 1.1, if we choose $\delta = \sum_{j=0}^{\infty} \overline{\lambda}^j \delta_j \in E$, then for any $f(z,w) = \sum_{j=0}^{\infty} f_j(z) w^j$ inside $M \ominus zM$, $U^*f = \sum_{j=0}^{\infty} f_j(z)\delta_j$ is in $\theta E'$, and

$$NU^*f(z) = N\Big(\sum_{j=0}^{\infty} f_j(z)\delta_j\Big) = \Big\langle \sum_{j=0}^{\infty} f_j(z)\delta_j, \delta \Big\rangle = \sum_{j=0}^{\infty} f_j(z)\lambda^j = N_\lambda f(z).$$

So $N_\lambda = NU^*$, hence is Hilbert-Schmidt by Lemma 1.1, and

$$\operatorname{tr}(N_\lambda^* N_\lambda) = \operatorname{tr}(U^* N^* N U) = \operatorname{tr}(N^* N).$$

The inequality

$$\operatorname{tr}(N_\lambda^* N_\lambda) \leqslant (1-|\lambda|^2)^{-1}$$

comes from the remarks following the proof of Lemma 1.1. We now show the inequality

$$\Big\| p_\perp \frac{1}{1-\overline{\lambda} w} \Big\|^2 \leqslant \operatorname{tr}(N_\lambda^* N_\lambda).$$

Let $\{g_0, g_1, g_2, \ldots\}$ be an orthonormal basis for $M \ominus zM$. Then

$$N_\lambda g_k(z) = g_k(z,\lambda) = \int\limits_{\mathbb{T}} \frac{g_k(z,w)}{1-\lambda\overline{w}} \,\mathrm{d}|w|,$$

and therefore

$$\begin{aligned}\operatorname{tr}(N_\lambda^* N_\lambda) &= \sum_{k=0}^{\infty} \int\limits_{\mathbb{T}} \Big| \int\limits_{\mathbb{T}} \frac{g_k(z,w)}{1-\lambda\overline{w}} \,\mathrm{d}|w| \Big|^2 \mathrm{d}|z| \geqslant \sum_{k=0}^{\infty} \Big| \int\limits_{\mathbb{T}} \int\limits_{\mathbb{T}} \frac{g_k(z,w)}{1-\lambda\overline{w}} \,\mathrm{d}|w| \,\mathrm{d}|z| \Big|^2 \\ &= \sum_{k=0}^{\infty} |\langle g_k, (1-\overline{\lambda} w)^{-1}\rangle|^2 = \Big\| p_\perp \frac{1}{1-\overline{\lambda} w} \Big\|^2.\end{aligned}$$

For operators N_j, $j = 0,1,2,\ldots$, we choose δ to be δ_j, $j = 0,1,2,\ldots$ correspondingly in Lemma 1.1. Similar calculations will establish the assertion and the inequalities. ∎

If $\mathcal{L}^2$ denotes the collection of all the Hilbert-Schmidt operators acting on some Hilbert space K, then for any a, b in $\mathcal{L}^2$,

$$\langle a, b\rangle \stackrel{\text{def}}{=} \operatorname{trace}(b^* a)$$

defines an inner product which turns $(\mathcal{L}^2, \langle\cdot,\cdot\rangle)$ into a Hilbert space. If $|\cdot|$ is the norm induced from this inner product, then

$$|xay| \leqslant \|x\|\,\|y\|\,|a|, \tag{1.5}$$

for any $a \in \mathcal{L}^2$ and any bounded operators x and y ([7], p. 79), where $\|\cdot\|$ is the operator norm.

LEMMA 1.3. *Suppose A, B are two contractions such that $[A,B] = AB - BA$ is Hilbert-Schmidt and $f(z) = \sum\limits_{j=0}^{\infty} c_j z^j$ is any holomorphic function over the unit disk such that $\sum\limits_{j=0}^{\infty} j|c_j|$ converges, then $[f(A), B]$ is also Hilbert-Schmidt.*

Proof. We observe that for any positive interger n,

$$\begin{aligned}
&[A^n, B] \\
&= A^n B - BA^n \\
&= A^n B - A^{n-1}BA + A^{n-1}BA - BA^n \\
&= A^{n-1}[A, B] + [A^{n-1}, B]A \\
&\vdots \\
&= A^{n-1}[A, B] + A^{n-2}[A, B]A + \cdots + A[A, B]A^{n-2} + [A, B]A^{n-1},
\end{aligned}$$

hence

$$|[A^n, B]| \leqslant n|[A, B]|$$

by inequality (1.5). If we let $f_n(z) = \sum_{j=0}^{n} c_j z^j$ then $[f_n(A), B]$ is in $\mathcal{L}^2$ and

$$\begin{aligned}
|[f_n(A), B] - [f(A), B]| &= \Big|\Big[\sum_{j=n+1}^{\infty} c_j A^j, B\Big]\Big| \\
&\leqslant \sum_{j=n+1}^{\infty} |c_j|\,|[A^j, B]| \leqslant \sum_{j=n+1}^{\infty} j|c_j|\,|[A, B]|.
\end{aligned}$$

From the assumption on f,

$$\lim_{n\to\infty} \sum_{j=n+1}^{\infty} j|c_j|\,|[A, B]| = 0,$$

hence $[f(A), B]$ is also in $\mathcal{L}^2$, i.e. Hilbert-Schmidt. ∎

Corollary 1.2 is crucial for the rest of the sections and Lemma 1.3 will enable us to get around some technical difficulties.

2. Decomposition of cross commutators

In this section we will define the compression operators and decompose their cross commutators. We begin by introducing some notations.

For any $h \in H^2(\mathbb{D}^2)$, we let

$$[h] := \overline{A(\mathbb{D}^2)h}^{H^2}$$

denote the submodule generated by h. Here we note that h is called *inner* if

$$|h(z, w)| = 1 \quad \text{a.e. on } \mathbb{T}^2.$$

It is not hard to see that

$$[h] = hH^2(\mathbb{D}^2)$$

when h is inner. Further, h is called *outer in the sense of Helson* (H) if

$$[h] = H^2(\mathbb{D}^2).$$

Given any submodule M, we can decompose $H^2(\mathbb{D}^2)$ as

$$H^2(\mathbb{D}^2) = (H^2(\mathbb{D}^2) \ominus M) \oplus M,$$

and let

$$p : H^2(\mathbb{D}^2) \to M,$$
$$q : H^2(\mathbb{D}^2) \to H^2(\mathbb{D}^2) \ominus M$$

be the projections. For any $f \in H^\infty(\mathbb{D}^2)$, we let S_f and R_f be the compressions of the operator T_f to $H^2(\mathbb{D}^2) \ominus M$ and M respectively, i.e.

$$S_f = qfq, \quad R_f = pfp.$$

In Sections 3 and 4 we will prove that when $M = [h]$ with h a polynomial, the cross commutators $[S_w^*, S_z]$ and $[R_w^*, R_z]$ are both Hilbert-Schmidt. To avoid the technical difficulties, we prove the assertion for the operators $[S_{\varphi_\lambda}^*, S_z]$ and $[R_{\varphi_\lambda}^*, R_z]$ first, where $\varphi_\lambda(w) = \frac{w-\lambda}{1-\bar{\lambda}w}$ with some $\lambda \in \mathbb{D}$ such that $h(z, \lambda) \neq 0$ for all $z \in \mathbb{T}$, and then apply Lemma 1.3.

First we need to have a better understanding of the two cross commutators $[S_w^*, S_z]$ and $[R_w^*, R_z]$. In view of the decomposition

$$H^2(\mathbb{D}^2) = (H^2(\mathbb{D}^2) \ominus M) \oplus M,$$

we can decompose the Toeplitz operators on $H^2(\mathbb{D}^2)$ correspondingly.

If we regard φ_λ as a multiplication operator on $H^2(\mathbb{D}^2)$, then

$$T_{\varphi_\lambda} = \begin{pmatrix} q\varphi_\lambda q & 0 \\ p\varphi_\lambda q & p\varphi_\lambda p \end{pmatrix},$$

$$T_z = \begin{pmatrix} qzq & 0 \\ pzq & pzp \end{pmatrix},$$

and

$$T_{\varphi_\lambda}^* T_z - T_z T_{\varphi_\lambda}^* = \begin{pmatrix} q\overline{\varphi}_\lambda qzq + q\overline{\varphi}_\lambda pzq - qzq\overline{\varphi}_\lambda q & q\overline{\varphi}_\lambda pzp - qzq\overline{\varphi}_\lambda p \\ p\overline{\varphi}_\lambda pzq - pzq\overline{\varphi}_\lambda q & p\overline{\varphi}_\lambda pzp - pzq\overline{\varphi}_\lambda p - pzp\overline{\varphi}_\lambda p \end{pmatrix}.$$

It is well known that T_z doubly commutes with T_w on $H^2(\mathbb{D}^2)$. Because φ_λ is a function of w only, it is then not hard to verify that

$$T^*_{\varphi_\lambda} T_z - T_z T^*_{\varphi_\lambda} = 0,$$

so we have that

$$q\overline{\varphi}_\lambda qzq + q\overline{\varphi}_\lambda pzq - qzq\overline{\varphi}_\lambda q = 0,$$

and

$$p\overline{\varphi}_\lambda pzp - pzq\overline{\varphi}_\lambda p - pzp\overline{\varphi}_\lambda p = 0,$$

i.e.

$$q\overline{\varphi}_\lambda qzq - qzq\overline{\varphi}_\lambda q = -q\overline{\varphi}_\lambda pzq,$$

$$p\overline{\varphi}_\lambda pzp - pzp\overline{\varphi}_\lambda p = pzq\overline{\varphi}_\lambda p.$$

Thus we have a following:

Proposition 2.1.

$$S^*_{\varphi_\lambda} S_z - S_z S^*_{\varphi_\lambda} = -q\overline{\varphi}_\lambda pzq, \tag{2.1}$$

$$R^*_{\varphi_\lambda} R_z - R_z R^*_{\varphi_\lambda} = pzq\overline{\varphi}_\lambda p. \tag{2.2}$$

3. Essential commutativity of S^*_w and S_z

In this section we will prove the essential commutativity of S^*_w and S_z on $H^2(\mathbb{D}^2)\ominus[h]$ when h is a polynomial. As we noted in the last section, we first prove the assertion for $S^*_{\varphi_\lambda}$ and S_z.

We first observe that for any $f \in H^2(\mathbb{D}^2) \ominus [h]$ and any $g \in [h]$,

$$\langle pzf, zg\rangle_{H^2} = \langle zf, zg\rangle_{H^2} = \langle f, g\rangle_{H^2} = 0.$$

So pz actually maps $H^2(\mathbb{D}^2) \ominus [h]$ into $[h] \ominus z[h]$. Therefore, $S^*_{\varphi_\lambda} S_z - S_z S^*_{\varphi_\lambda}$ can be decomposed as

$$H^2(\mathbb{D}^2) \ominus [h] \xrightarrow{-pz} [h] \ominus z[h] \xrightarrow{q\overline{\varphi}_\lambda} H^2(\mathbb{D}^2) \ominus [h]. \tag{3.1}$$

This observation has an interesting corollary when h is inner.

COROLLARY 3.1. *If h is inner, then $S_w^*S_z - S_zS_w^*$ is at most of rank 1 on $H^2(\mathbb{D}^2) \ominus [h]$.*

Proof. First we note that when $\lambda = 0$, $\varphi_\lambda(w) = w$. If h is inner,

$$[h] = hH^2(\mathbb{D}^2),$$

and $\{w^n h \mid n = 0, 1, 2, \ldots\}$ is an orthonormal basis for $[h] \ominus z[h]$. For any function $f(z,w) = \sum_{j=0}^{\infty} c_j w^j h$ inside $[h] \ominus z[h]$,

$$q\overline{w}f = q\overline{w}c_0 h + q\Big(\sum_{j=1}^{\infty} c_j w^{j-1} h\Big) = c_0 q\overline{w}h.$$

This shows that $q\overline{w}$ is at most of rank one and hence $S_w^*S_z - S_zS_w^* = -q\overline{w}pz$ is at most of rank one. ∎

This corollary enables us to give an operator theoretical proof of an interesting fact first noticed by W. Rudin in a slightly different context ([11], p. 123).

COROLLARY 3.2. *$h(z,w) = z - w$ has no inner-outer* (H) *factorization.*

Proof. As before, we let S_z, S_w be the compressions of T_z, T_w to $H^2(\mathbb{D}^2) \ominus [h]$ and set

$$e_n = \frac{1}{\sqrt{n+1}}\,(z^n + z^{n-1}w + \cdots + zw^{n-1} + w^n), \quad n = 0, 1, 2, \ldots.$$

One verifies that $\{e_n \mid n = 0, 1, 2, \ldots\}$ is an orthonormal basis for $H^2(\mathbb{D}^2) \ominus [z-w]$. Experts will know that $H^2(\mathbb{D}^2) \ominus [z-w]$ is actually the Bergman space over the unit disk. One then easily checks that

$$\begin{aligned} S_z &= S_w, \\ S_w e_n &= \frac{\sqrt{n+1}}{\sqrt{n+2}}\, e_{n+1}, \\ S_w^* e_n &= \frac{\sqrt{n}}{\sqrt{n+1}}\, e_{n-1}, \quad n \geqslant 1. \end{aligned}$$

Therefore,

$$[S_w^*, S_w]e_n = \frac{1}{n(n+1)}, \quad n = 0, 1, 2, \ldots.$$

If $z - w$ had an inner-outer factorization, then $[z-w] = gH^2(\mathbb{D}^2)$ for some inner function g and

$$[S_w^*, S_w] = [S_w^*, S_z]$$

would be at most a rank one operator which conflicts with the above computation. ∎

Similar methods can be used to show that the functions like $z - \mu w^n$, for $|\mu| < 1$ and n a nonnegative integer, have no inner-outer (H) factorization.

We now come to the main theorem of this section.

Theorem 3.3. *If $h \in H^\infty(\mathbb{D}^2)$ and there is a fixed $\lambda \in \mathbb{D}$ and a positive constant L such that*

$$L \leqslant |h(z, \lambda)| \tag{3.2}$$

for almost every $z \in \mathbb{T}$ then $S_w^* S_z - S_z S_w^*$ on $H^2(\mathbb{D}^2) \ominus [h]$ is Hilbert-Schmidt.

Proof. We first show that $S_{\varphi_\lambda}^* S_z - S_z S_{\varphi_\lambda}^*$ is Hilbert-Schmidt. From (3.1), it will be sufficient to show that

$$q\overline{\varphi}_\lambda : [h] \ominus z[h] \to H^2(\mathbb{D}^2) \ominus [h]$$

is Hilbert-Schmidt.

Let us recall that the operator N_λ from $[h] \ominus z[h]$ to $H^2(\mathbb{D})$ is defined by

$$N_\lambda g = g(\,\cdot\,, \lambda),$$

and it is Hilbert-Schmidt by Corollary 1.2. Suppose

$$hf_0, hf_1, hf_2, \ldots$$

is an orthonormal basis for $[h] \ominus z[h]$.

We first show that $h(z, w) f_k(z, \lambda) \in [h]$ for every k. In fact,

$$\int_{\mathbb{T}} |f_k(z, \lambda)|^2 \, \mathrm{d}|z| \leqslant L^{-2} \int_{\mathbb{T}} |h(z, \lambda) f_k(z, \lambda)|^2 \, \mathrm{d}|z| = L^{-2} \|N_\lambda(hf_k)\|^2 < \infty,$$

i.e. $f_k(z, \lambda) \in H^2(\mathbb{D})$ and hence $h(z, w) f_k(z, \lambda) \in [h]$ since h is bounded. Furthermore,

$$\|h(\,\cdot\,,\cdot\,) f_k(\,\cdot\,, \lambda)\|^2 \leqslant \|h\|_\infty^2 \|f_k(\,\cdot\,, \lambda)\|^2 \leqslant \|h\|_\infty^2 L^{-2} \|N_\lambda(hf_k)\|^2. \tag{3.3}$$

Next, we observe that

$$q\overline{\varphi}_\lambda h f_k = q\overline{\varphi}_\lambda h(f_k - f_k(\,\cdot\,, \lambda)) + q\overline{\varphi}_\lambda h f_k(\,\cdot\,, \lambda). \tag{3.4}$$

Since $f_k(z, w) - f_k(z, \lambda)$ vanishes at $w = \lambda$ for every $z \in \mathbb{D}$, it has $\varphi_\lambda(w)$ as a factor, and hence

$$q\overline{\varphi}_\lambda h(f_k - f_k(\,\cdot\,, \lambda)) = 0. \tag{3.5}$$

Combining (3.3) and (3.4),

$$\begin{aligned}\sum_{k=0}^{\infty}\|q\overline{\varphi}_\lambda h f_k\|^2_{H^2(\mathbb{D}^2)} &= \sum_{k=0}^{\infty}\|q\overline{\varphi}_\lambda h(f_k - f_k(\,\cdot\,,\lambda)) + q\overline{\varphi}_\lambda h f_k(\,\cdot\,,\lambda)\|^2_{H^2(\mathbb{D}^2)}\\ &= \sum_{k=0}^{\infty}\|q\overline{\varphi}_\lambda h f_k(\,\cdot\,,\lambda)\|^2_{H^2(\mathbb{D}^2)} \leqslant \sum_{k=0}^{\infty}\|h(\,\cdot\,,\cdot\,)f_k(\,\cdot\,,\lambda)\|^2_{H^2(\mathbb{D}^2)}\\ &\leqslant \|h\|^2_\infty L^{-2}\sum_{k=0}^{\infty}\|h(\,\cdot\,,\lambda)f_k(\,\cdot\,,\lambda)\|^2_{H^2(\mathbb{D})} = \|h\|^2_\infty L^{-2}\operatorname{tr}(N^*_\lambda N_\lambda).\end{aligned}$$

This shows that $q\overline{\varphi}_\lambda$, and hence $[S^*_{\varphi_\lambda}, S_z]$ is Hilbert-Schmidt.

Assuming $\widehat{\varphi}_\lambda(w) = \overline{\varphi_\lambda(\overline{w})}$, one verifies that

$$S^*_{\varphi_\lambda} = \widehat{\varphi}_\lambda(S^*_w).$$

The fact that

$$\widehat{\varphi}_\lambda(\widehat{\varphi}_\lambda(w)) = w$$

and an application of Lemma 1.3 with $f = \widehat{\varphi}_\lambda$ then imply that $[S^*_w, S_z]$ is Hilbert-Schmidt. ∎

In Theorem 3.3, if h is continuous on the boundary of $\mathbb{D}\times\mathbb{D}$, then the inequality (3.2) will hold once there is a $\lambda\in\mathbb{D}$ such that $h(z,\lambda)$ has no zero on $\mathbb{T}$. This idea leads to the assertion that $S^*_w S_z - S_z S^*_w$ is Hilbert-Schmidt on $H^2(\mathbb{D}^2)\ominus[h]$ for any polynomial h in two complex variables. But we need to recall some knowlege from complex analysis before we can prove it.

Suppose G is a bounded open set in the complex plane $\mathbb{C}$. We let $A(G)$ denote the collection of all the functions that are holomorphic on G and are continuous to the boundary of G; $Z(f)$ denotes the zeros of f.

To make a study of zero sets of polynomials, we need a classical theorem in several complex variables.

THEOREM 3.4. *Let*

$$h(z,w) = z^n + a_1(w)z^{n-1} + \cdots + a_n(w)$$

be a pseudopolynomial without multiple factors, where the $a_j(w)$'s are all in $A(G)$. Further let

$$D_h := \{w\in G \mid \Delta_h(w) = 0\},$$

where $\Delta_h(w)$ is the discriminant of h. Then for any $w_0 \in G - D_h$ there exists an open neighborhood of $U(w_0) \subset G - D_h$ and holomorphic functions $f_1, f_2, \ldots, f_n$ on U with $f_i(w) \neq f_j(w)$ for $i \neq j$ and $w \in U$, such that

$$h(z,w) = (z - f_1(w))(z - f_2(w)) \cdots (z - f_n(w))$$

for all $w \in U$ and all complex number z.

This theorem is taken from [8], but similar theorems can be found in other standard books on several complex variables. It reveals some information on the zero sets of polynomials which we state as

Corollary 3.5. *For any polynomial $p(z,w)$ not having $z - \lambda$ with $|\lambda| = 1$ as a factor, the set*

$$Y_p = \{w \in \mathbb{C} \mid p(z,w) = 0 \textit{ for some } z \in \mathbb{T}\}$$

has no interior.

Proof. We first assume that p is irreducible and write $p(z,w)$ as

$$p(z,w) = a_0(w)z^n + a_1(w)z^{n-1} + \cdots + a_n(z)$$

with $a_j(w)$ polynomials of one variable and $a_0(w)$ not identically zero. Then on $\mathbb{C} \setminus Z(a_0)$, we have

$$p(z,w) = a_0(w)\left(z^n + \frac{a_1(w)}{a_0(w)}\, z^{n-1} + \cdots + \frac{a_n(w)}{a_0(w)}\right).$$

Let Δ_p be the discriminant (see [8] for the definition) of p. If p is irreducible, Δ_p is not identically zero, and so neither is the discriminant of

$$q(z,w) = z^n + \frac{a_1(w)}{a_0(w)}\, z^{n-1} + \cdots + \frac{a_n(w)}{a_0(w)}.$$

This implies that the pseudopolynomial $q(z,w)$ has no multiple factor either.

We now prove the corollary for the irreducible polynomial p. We do it by showing that given any open disk $B \subset \mathbb{C}$, there is a $w \in B$ which is not in Y_p.

Given any small open disk B and a point w_0 in $B \setminus \{Z(\Delta_p) \cup Z(a_0)\}$, the above theorem shows the existence of an open neighborhood $U \subset B$ of w_0 and holomorphic functions $f_1, f_2, \ldots, f_n$ on U with $f_i(w) \neq f_j(w)$ for $i \neq j$ and $w \in U$ such that

$$p(z,w) = a_0(w)(z - f_1(w))(z - f_2(w)) \cdots (z - f_n(w)), \tag{3.6}$$

for all $z \in \mathbb{C}$. Then $f_1(w)$ can not be a constant λ of modulus 1 because p does not have factors of the form $z - \lambda$ from the assumption. So we can choose a smaller open disk $B_1 \subset U$ such that $f_1(B_1) \cap \mathbb{T}$ is empty. Carrying the same argument out for f_2 on B_1, we have an open disk $B_2 \subset B_1$ such that $f_2(B_2) \cap \mathbb{T}$ is empty. Continuing this procedure, we have disks $B_1, B_2, \ldots, B_n$ such that $B_j \subset B_{j-1}$ for $j = 2, 3, \ldots, n$. Then for any $w \in B_n$, $p(z, w)$ will have no zero on $\mathbb{T}$ and hence w is not in Y_p.

If p is an arbitary polynomial not having $z - \lambda$ with $|\lambda| = 1$ as a factor, we factorize p into a product of irreducible polynomials as

$$p(z, w) = p_1^{d_1} p_2^{d_2} \cdots p_m^{d_m}.$$

If we let

$$Y_j = \{w \in \mathbb{C} \mid p_j(z, w) = 0 \text{ for some } z \in \mathbb{T}\},$$

then $Y_p \subset \bigcup_{j=1}^{m} Y_j$, hence it has no interior. ∎

We feel it may be interesting to have a closer look at the set Y_p, but that is not the purpose of this paper. The result in Corollary 3.5 is good enough for us to state

THEOREM 3.6. *For any polynomial h, $S_w^* S_z - S_z S_w^*$ is Hilbert-Schmidt on $H^2(\mathbb{D}^2) \ominus [h]$.*

Proof. Suppose h is any polynomial. If h is of the form $(z - \lambda)g$ for some polynomial g and some λ of modulus 1, then $[h] = [g]$ because $z - \lambda$ is outer (H). So without loss of generality, we assume that h does not have this kind of factor. Then from the above corollary, $h(z, \mu)$ has no zeros on $\mathbb{T}$ for any $\mu \in \mathbb{D} \setminus Y_h$. Theorem 3.3 and the observations immediately after it then imply that $[S_w^*, S_z]$ is Hilbert-Schmidt. ∎

For any function $f \in A(\mathbb{D}^2)$, we can define an operator S_f by

$$S_f x \stackrel{\text{def}}{=} q f x$$

for any $x \in H^2(\mathbb{D}^2) \ominus [h]$, where q is the projection from $H^2(\mathbb{D}^2)$ onto $H^2(\mathbb{D}^2) \ominus [h]$. One checks that this turns $H^2(\mathbb{D}^2) \ominus [h]$ into a Hilbert $A(\mathbb{D}^2)$ quotient module. The module is called *essentially reductive* if S_f is essentially normal for every $f \in A(\mathbb{D}^2)$. It is easy to see that $H^2(\mathbb{D}^2) \ominus [h]$ is *essentially reductive* if and only if both $[S_z^*, S_z]$ and $[S_w^*, S_w]$ are compact. Currently we do not know how to characterize those functions h for which $H^2(\mathbb{D}^2) \ominus [h]$ is essentially reductive, even though some partial results are available. [4] and [5] are good references on this topic. However, if we consider $H^2(\mathbb{D}^2) \ominus [h]$ as a module over the the subalgebra $A(\mathbb{D}) \subset A(\mathbb{D}^2)$, Theorem 3.7 yields the following

COROLLARY 3.7. *Assume h is a polynomial. If there is a $g \in A(\mathbb{D})$ and a $f \in [h] \cap H^\infty(\mathbb{D}^2)$, such that*

$$z = g(w) + f(z, w),$$

then $H^2(\mathbb{D}^2) \ominus [h]$ is an essentially reductive module over $A(\mathbb{D})$ with the action defined by

$$f \cdot x \stackrel{\text{def}}{=} f(S_z)x$$

for all $f \in A(\mathbb{D})$ and all $x \in H^2(\mathbb{D}^2) \ominus [h]$.

Proof. It suffices to show that S_z is essentially normal. From the assumption on f, S_f is equal to 0. Since $z - g(w) = f(z, w)$, we have that

$$S_z = S_g = g(S_w).$$

Suppose $\{p_n\}$ is a sequence of polynomials which converges to g in supremum norm, then from Lemma 1.3, $[S_z^*, p_n(S_w)]$ is compact for each n and it is also not hard to see that $[S_z^*, p_n(S_w)]$ converges to $[S_z^*, g(S_w)]$ in the operator norm, and hence $[S_z^*, S_z] = [S_z^*, g(S_w)]$ is compact. ∎

This corollary shows in particular that $H^2(\mathbb{D}^2) \ominus [h]$ is essentially reductive over $A(\mathbb{D}^2)$ when h is linear.

4. ESSENTIAL COMMUTATIVITY OF R_w^* AND R_z

In the last section we proved that the module actions of the two coordinate functions z, w on the quotient module $H^2(\mathbb{D}^2) \ominus [h]$ essentially doubly commute when h is a polynomial. It is then natural to ask if there is a similar phenomenon in the case of submodules. A result due to Curto, Muhly and Yan ([3]) answered the question affirmatively in a special case and Curto asked if it is true for any polynomially generated submodules ([2]). Since $C[z, w]$ is Noetherian, one only needs to look at the submodules generated by a finite number of polynomials. In this section we will answer Curto's question partially and a complete answer will be given in Section 6.

At first, we thought that the submodule case should be easier to deal with than the quotient module case because z, w act as isometries on submodules. But it turns out that the submodule case is more subtle and needs a finer analysis.

Let us now get down to details.

Suppose M is a submodule and R_w and R_z are the module actions by coordinate functions z and w. It is obvious R_w and R_z are commuting isometries. In

[3], Curto, Muhly and Yan made a study of the essential commutativity of operators R_w^*, R_z in the case that M is generated by a finite number of homogeneous polynomials. They were actually able to show that $[R_w^*, R_z]$ is Hilbert-Schmidt. In this section we will show that this is also true when M is generated by an arbitrary polynomial. The same result for the case that M is generated by a finite number of polynomials is a corollary of this result and will be treated in Section 6.

We suppose h is a polynomial that does not have a factor $z-\mu$ with $|\mu|=1$. Then from Corollary 3.5 there is a $\lambda \in \mathbb{D}$ such that $h(z,\lambda)$ is bounded away from 0 on $\mathbb{T}$. As in Section 3, we will see that this is crucial in the development of the proofs.

For a bounded analytic function $f(z,w)$ over the unit bidisk, we recall that R_f is the restriction of the Toeplitz operator T_f onto $[h]$ and by Proposition 2.1,

$$R^*_{\varphi_\lambda} R_z - R_z R^*_{\varphi_\lambda} = pzq\overline{\varphi}_\lambda p.$$

We let

$$p_1 : H^2(\mathbb{D}^2) \to \varphi_\lambda[h], \quad q_1 : H^2(\mathbb{D}^2) \to [h] \ominus \varphi_\lambda[h]$$

be the projections; then $p = p_1 + q_1$. It is not hard to see that

$$(R^*_{\varphi_\lambda} R_z - R_z R^*_{\varphi_\lambda}) p_1 = pzq\overline{\varphi}_\lambda p_1 = 0.$$

Moreover, by the remarks preceding Proposition 2.1,

$$T_z T_{\overline{\varphi}_\lambda} = T_z T^*_{\varphi_\lambda} = T^*_{\varphi_\lambda} T_z = T_{\overline{\varphi}_\lambda} T_z,$$

and hence,

$$\begin{aligned} R^*_{\varphi_\lambda} R_z - R_z R^*_{\varphi_\lambda} &= pzq\overline{\varphi}_\lambda (p_1 + q_1) = pzq\overline{\varphi}_\lambda q_1 \\ &= pz(P-p)\overline{\varphi}_\lambda q_1 = pT_z T_{\overline{\varphi}_\lambda} q_1 - pzp\overline{\varphi}_\lambda q_1 \\ &= pT_{\overline{\varphi}_\lambda} T_z q_1 - pzp\overline{\varphi}_\lambda q_1 = p\overline{\varphi}_\lambda z q_1 - pzp\overline{\varphi}_\lambda q_1, \end{aligned} \tag{4.1}$$

where P is the projection from $L^2(\mathbb{T}^2)$ to $H^2(\mathbb{D}^2)$. For any $f \in [h] \ominus \varphi_\lambda[h]$ and $g \in [h]$,

$$\langle p\overline{\varphi}_\lambda f, g\rangle = \langle f, \varphi_\lambda g\rangle = 0,$$

i.e.

$$p\overline{\varphi}_\lambda q_1 = 0. \tag{4.2}$$

Combining equations (4.1) and (4.2) we have that

$$R^*_{\varphi_\lambda} R_z - R_z R^*_{\varphi_\lambda} = p\overline{\varphi}_\lambda z q_1.$$

Furthermore, equation (4.2) also implies that

$$p\overline{\varphi}_\lambda z q_1 = p\overline{\varphi}_\lambda (p_1 + q_1) z q_1 = p\overline{\varphi}_\lambda p_1 z q_1 + p\overline{\varphi}_\lambda q_1 z q_1 = p\overline{\varphi}_\lambda p_1 z q_1.$$

Since $p\overline{\varphi}_\lambda$ acts on $\varphi_\lambda[h]$ as an isometry, the above observations then yield.

Proposition 4.1. *$[R^*_{\varphi_\lambda}, R_z]$ is Hilbert-Schmidt on $[h]$ if and only if $p_1 z q_1$ is Hilbert-Schmidt and*

$$\mathrm{tr}([R^*_{\varphi_\lambda}, R_z]^*[R^*_{\varphi_\lambda}, R_z]) = \mathrm{tr}((p_1 z q_1)^*(p_1 z q_1))(p_1 z q_1)).$$

We further observe that, for any $f \in [h] \ominus \varphi_\lambda[h]$ and $g \in \varphi_\lambda[h]$,

$$\langle p_1 z f, zg\rangle = \langle f, g\rangle = 0.$$

So the range of operator $p_1 z q_1$ is a subspace of $\varphi_\lambda[h] \ominus z\varphi_\lambda[h]$. If we let $p_\perp$ be the projection from $\varphi_\lambda[h]$ onto $\varphi_\lambda[h] \ominus z\varphi_\lambda[h]$ then

$$p_1 z q_1 = p_\perp z q_1. \tag{4.3}$$

We will prove that $p_\perp z q_1$ is Hilbert-Schmidt after some preparation.

Suppose

$$h = \sum_{j=0}^{m} a_j(z) w^j$$

is a polynomial and that

$$|h(z, \lambda)| \geqslant \varepsilon, \tag{4.4}$$

for some fixed positive ε and all $z \in \mathbb{T}$. Assume $\mathcal{H}$ to be the L^2-closure of $\mathrm{span}\{h(z, w) z^j \mid j \geqslant 0\}$, then $\mathcal{H} \subset [h]$ and we have the following

Lemma 4.2. *$\mathcal{H} = \{h(z, w) f(z) \mid f \in H^2(\mathbb{D})\} = hH^2(\mathbb{D})$.*

Proof. It is not hard to check that $hH^2(\mathbb{D}) \subset \mathcal{H}$.

For the other direction, we assume hf is any function in $\mathcal{H}$ and need to show that $f \in H^2(\mathbb{D})$. In fact, if $p_n(z)$, $n \geqslant 1$ is a sequence of polynomials such that $h(z, w)p_n(z)$, $n \geqslant 1$, converges to $h(z, w)f(z, w)$ in $L^2(\mathbb{T}^2)$, then $h(z, \lambda)p_n(z)$, $n \geqslant 1$, converges to $h(z, \lambda)f(z, \lambda)$ in $L^2(\mathbb{T})$ by the boundedness of N_λ. Our assumption on h then implies that $p_n(z)$, $n \geqslant 1$, converges to $f(z, \lambda)$ in $L^2(\mathbb{T})$, and in particular, $f(z, \lambda) \in H^2(\mathbb{D})$. This in turn implies that $h(z, w)p_n(z)$, $n \geqslant 1$, converges to $h(z, w)f(z, \lambda)$ in $L^2(\mathbb{T}^2)$ since h is a bounded function. Hence by the uniqueness of the limit,

$$h(z, w)f(z, w) = h(z, w)f(z, \lambda),$$

and therefore

$$f(z, w) = f(z, \lambda). \quad ∎$$

It is interesting to see from this lemma and Corollary 3.5 that $hH^2(\mathbb{D})$ is actually closed in $H^2(\mathbb{D}^2)$ for any polynomial h not having a factor $z - \mu$ with $|\mu| = 1$.

LEMMA 4.3. *The operator $V : [h] \to \mathcal{H}$ defined by*

$$V(hf) = h(z,w)f(z,\lambda)$$

is bounded.

Proof. First of all $h(z,\lambda)f(z,\lambda) = N_\lambda(hf)$ is in $H^2(\mathbb{D})$ and hence so is $f(z,\lambda)$ since $|h(z,\lambda)| \geqslant \varepsilon$ on $\mathbb{T}$. So V is indeed a map from $[h]$ to $\mathcal{H}$.

Next we choose a number M sufficiently large such that

$$\int_{\mathbb{T}} |h(z,w)|^2 \,\mathrm{d}|w| \leqslant M\varepsilon^2 \leqslant M|h(z,\lambda)|^2$$

for all $z \in \mathbb{T}$. Then for any $h(z,w)f(z,w) \in [h]$,

$$\|V(hf)\|^2 = \int_{\mathbb{T}^2} |h(z,w)f(z,\lambda)|^2 \,\mathrm{d}|z|\,\mathrm{d}|w| = \int_{\mathbb{T}} \Big(\int_{\mathbb{T}} |h(z,w)|^2 \,\mathrm{d}|w| \Big) |f(z,\lambda)|^2 \,\mathrm{d}|z|$$
$$\leqslant M \int_{\mathbb{T}} |h(z,\lambda)f(z,\lambda)|^2 \,\mathrm{d}|z| \leqslant M(1-|\lambda|^2)^{-1}\|hf\|^2. \quad \blacksquare$$

This lemma enables us to reduce the problem further.

For any $h(z,w)f(z,w) \in [h] \ominus \varphi_\lambda[h]$,

$$p_\perp zhf = p_\perp zV(hf) + p_\perp z(hf - Vhf).$$

But

$$zh(z,w)f(z,w) - zV(hf)(z,w) = zh(z,w)(f(z,w) - f(z,\lambda)),$$

and since $f(z,w) - f(z,\lambda)$ vanishes at $w = \lambda$ for every z, it has φ_λ as a factor, hence $z(hf - V(hf)) \in z\varphi_\lambda[h]$. Therefore by the definition of $p_\perp$,

$$p_\perp zhf = p_\perp zV(hf) + p_\perp z\varphi_\lambda hg = p_\perp zV(hf). \tag{4.5}$$

To prove that $p_\perp zq_1$ is Hilbert-Schmidt, one then suffices to show that $p_\perp z$ restricted to $\mathcal{H}$ is Hilbert-Schmidt. Before proving it, we make another observation and state a lemma.

Since $h(z,w)$ is a polynomial and

$$\int_{\mathbb{T}} |h(z,w)|^2 \,\mathrm{d}|w| = \sum_{k=0}^{m} |a_k(z)|^2,$$

the Riesz-Fejér theorem implies that there is a polynomial $Q(z)$ such that

$$|Q(z)|^2 = \int_{\mathbb{T}} |h(z,w)|^2 \,\mathrm{d}|w|$$

on $\mathbb{T}$. If Q vanishes at some $\mu \in \mathbb{T}$, then $a_k(\mu) = 0$ for each k, and hence h has a factor $(z-\mu)$. But this contradicts our assumption on h. So we can find a positive constant, say η, such that

$$|Q(z)| \geqslant \eta, \tag{4.6}$$

for all $z \in \mathbb{T}$.

Suppose $\{h(z,w)f_n(z) \mid n \geqslant 0\}$ is an orthonormal basis for $\mathcal{H}$, then

$$\begin{aligned}
\delta_{i,j} &= \int_{\mathbb{T}^2} h(z,w)f_i(z)\overline{h(z,w)f_j(z)}\,\mathrm{d}|z|\,\mathrm{d}|w| \\
&= \int_{\mathbb{T}} \Big(\int_{\mathbb{T}} |h(z,w)|^2\,\mathrm{d}|w|\Big) f_i(z)\overline{f_j(z)}\,\mathrm{d}|z| \\
&= \int_{\mathbb{T}} Q(z)f_i(z)\overline{Q(z)f_j(z)}\,\mathrm{d}|z|.
\end{aligned}$$

So $\{Q(z)f_k(z) \mid k \geqslant 0\}$ is orthonormal in $H^2(\mathbb{D})$, but of course it may not be complete.

Lemma 4.4. *The linear operator* $J : \overline{\operatorname{span}\{Qf_k \mid k \geqslant 0\}} \to H^2(\mathbb{D})$ *defined by*

$$J(Qf_k) = f_k, \quad k \geqslant 0,$$

is bounded.

Proof. By inequality (4.6), for any function $Qf \in \overline{\operatorname{span}\{Qf_k \mid k \geqslant 0\}}$,

$$\int_{\mathbb{T}} |f(z)|^2\,\mathrm{d}|z| \leqslant \eta^{-2} \int_{\mathbb{T}} |Q(z)f(z)|^2\,\mathrm{d}|z|. \quad \blacksquare$$

Now we are in the position to prove

Proposition 4.5. *$p_\perp z$ restricted to $\mathcal{H}$ is Hilbert-Schmidt.*

Proof. Assume $\{g_k \mid k \geqslant 0\} \subset [h] \ominus z[h]$ is an orthonormal basis and, as above, $\{h(z,w)f_n(z) \mid n \geqslant 0\}$ is an orthonormal basis for $\mathcal{H}$. Since φ_λ is inner,

$\{\varphi_\lambda(w)g_k(z,w) \mid k \geqslant 0\}$ is an orthonormal basis for $\varphi_\lambda[h] \ominus z\varphi_\lambda[h]$. Therefore, by identity (1.3) and the expression of h,

$$p_\perp zhf_n = \sum_{k=0}^{\infty} \langle zhf_n, \varphi_\lambda g_k \rangle \varphi_\lambda g_k = \sum_{k=0}^{\infty} \Big\langle \sum_{i=0}^{m} za_i w^i f_n, \varphi_\lambda \sum_{j=0}^{\infty} T_{w^j} N_j g_k \Big\rangle \varphi_\lambda g_k.$$

Note that a_i's and f_n are functions of z only, so $\sum\limits_{i=0}^{m} za_i w^i f_n$ is orthogonal to $\sum\limits_{j=m+1}^{\infty} w^j \varphi_\lambda N_j g_k$ because the later has the factor w^{m+1}. It then follows that

$$\begin{aligned}
p_\perp zhf_n &= \sum_{k=0}^{\infty} \Big\langle \sum_{i=0}^{m} za_i w^i f_n, \varphi_\lambda \sum_{j=0}^{\infty} T_{w^j} N_j g_k \Big\rangle \varphi_\lambda g_k \\
&= \sum_{k=0}^{\infty} \Big\langle \sum_{i=0}^{m} za_i w^i f_n, \sum_{j=0}^{m} \varphi_\lambda w^j N_j g_k \Big\rangle \varphi_\lambda g_k \\
&= \sum_{k=0}^{\infty} \sum_{i,j=0}^{m} \varphi_\lambda g_k \Big(\int_{\mathbb{T}} za_i(z) f_n(z) \overline{N_j g_k(z)} \,\mathrm{d}|z| \Big) \Big(\int_{\mathbb{T}} w^i \overline{\varphi_\lambda(w) w^j} \,\mathrm{d}|w| \Big) \\
&= \sum_{k=0}^{\infty} \Big(\sum_{i,j=0}^{m} c_{ij} \langle f_n, T^*_{za_i} N_j g_k \rangle_{H^2(\mathbb{D})} \Big) \varphi_\lambda g_k,
\end{aligned}$$

where

$$c_{ij} = \int_{\mathbb{T}} w^i \overline{\varphi_\lambda(w) w^j} \,\mathrm{d}|w|.$$

If $c := \max\{|c_{ij}| \mid 0 \leqslant i, j \leqslant m\}$, then the Cauchy inequality yields

$$\begin{aligned}
\|p_\perp zhf_n\|^2 &= \sum_{k=0}^{\infty} \Big| \sum_{i,j=0}^{m} c_{ij} \langle f_n, T^*_{za_i} N_j g_k \rangle_{H^2(\mathbb{D})} \Big|^2 \\
&\leqslant (mc)^2 \sum_{k=0}^{\infty} \sum_{i,j=0}^{m} |\langle f_n, T^*_{za_i} N_j g_k \rangle_{H^2(\mathbb{D})}|^2 \\
&= (mc)^2 \sum_{k=0}^{\infty} \sum_{i,j=0}^{m} |\langle J(Qf_n), T^*_{za_i} N_j g_k \rangle_{H^2(\mathbb{D})}|^2 \\
&= (mc)^2 \sum_{k=0}^{\infty} \sum_{i,j=0}^{m} |\langle Qf_n, J^* T^*_{za_i} N_j g_k \rangle_{H^2(\mathbb{D})}|^2,
\end{aligned}$$

where J is the operator defined in Lemma 4.4. Therefore, by the fact that $\{Qf_n \mid n \geqslant 0\}$ is orthogonal in $H^2(\mathbb{D})$ and the fact that N_j is Hilbert-Schmidt on $[h] \ominus z[h]$

for each j,

$$\begin{aligned}\sum_{n=0}^{\infty}\|p_{\perp}zhf_n\|^2 &\leqslant (mc)^2\sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\sum_{i,j=0}^{m}|\langle Qf_n, J^*T_{za_i}^*N_jg_k\rangle_{H^2(\mathbb{D})}|^2\\ &= (mc)^2\sum_{k=0}^{\infty}\sum_{i,j=0}^{m}\sum_{n=0}^{\infty}|\langle Qf_n, J^*T_{za_i}^*N_jg_k\rangle_{H^2(\mathbb{D})}|^2\\ &\leqslant (mc)^2\sum_{k=0}^{\infty}\sum_{i,j=0}^{m}\|J^*T_{za_i}^*N_jg_k\|_{H^2(\mathbb{D})}^2\\ &= (mc)^2\sum_{i,j=0}^{m}\|J^*T_{za_i}^*\|^2\sum_{k=0}^{\infty}\|N_jg_k\|_{H^2(\mathbb{D})}^2\\ &= (mc)^2\sum_{i,j=0}^{m}\|J^*T_{za_i}^*\|^2\,\mathrm{tr}(N_j^*N_j)<\infty. \quad \blacksquare\end{aligned}$$

Theorem 4.6. *$[R_w^*, R_z]$ is Hilbert-Schmidt on $[h]$ for any polynomial h.*

Proof. If $h=(z-\lambda)h_1$ for some polynomial h_1 and $\lambda\in\mathbb{T}$, then $[h]=[h_1]$. If h_1 is a nonzero constant then $[h_1]=H^2(\mathbb{D}^2)$ and hence

$$R_w = T_w, \quad R_z = T_z.$$

Therefore $[R_w^*, R_z]=0$. So without loss of generality, we may assume h does not have a factor $z-\lambda$ for some $\lambda\in\mathbb{T}$. Propositions 4.1, 4.5 and Equality (4.3) together imply that $[R_{\varphi_\lambda}^*, R_z]$ is Hilbert-Schmidt. An argument similar to that in the end of the proof of Theorem 3.3 establishes our assertion. ∎

5. OPERATOR $[R_z^*, R_z][R_w^*, R_w]$ ON $[h]$

In this section we are going to use the result of the last section to prove the following:

Theorem 5.1. *The operator $[R_z^*, R_z][R_w^*, R_w]$ is Hilbert-Schmidt on $[h]$ when h is a polynomial.*

Proof. For the same reason as in the proof of Theorem 4.6, we assume that h does not have a factor $z-\mu$ for $\mu\in\mathbb{T}$. Then by Corollary 3.5, $h(z,\lambda)$ is bounded away from zero on $\mathbb{T}$ for some $\lambda\in\mathbb{D}$. To make our computations clearer, we assume that $h(z,0)$ is bounded away from 0 on $\mathbb{T}$. Then one sees that for any $hf \in [h]$, $h(f - f(\cdot\,,0))$ is a function in $w[h]$. Therefore,

$$
\begin{aligned}
[R_w^*, R_w]hf &= hf - R_w R_w^* hf = hf - R_w R_w^* h(f - f(\cdot\,,0) + f(\cdot\,,0)) \\
&= hf - h(f - f(\cdot\,,0)) - R_w R_w^* hf(\cdot\,,0) \\
&= hf(\cdot\,,0) - R_w R_w^* hf(\cdot\,,0) = [R_w^*, R_w]h(\cdot\,,\cdot)f(\cdot\,,0).
\end{aligned} \tag{5.1}
$$

Similarly,

$$
\begin{aligned}
[R_z^*, R_z]hf(\cdot\,,0) &= hf(\cdot\,,0) - R_z R_z^* hf(\cdot\,,0) \\
&= hf(\cdot\,,0) - R_z R_z^* h(f(\cdot\,,0) - f(0,0) + f(0,0)) \\
&= hf(\cdot\,,0) - h(f(\cdot\,,0) - f(0,0)) - R_z R_z^* hf(0,0) \\
&= hf(0,0) - f(0,0) R_z R_z^* h = f(0,0)[R_z^*, R_z]h.
\end{aligned} \tag{5.2}
$$

By the essential commutativity of R_z^* and R_w, and Equalities (5.1), (5.2),

$$
\begin{aligned}
[R_z^*, R_z][R_w^*, R_w]hf &= [R_z^*, R_z][R_w^*, R_w]h(\cdot\,,\cdot)f(\cdot\,,0) \\
&= [R_w^*, R_w][R_z^*, R_z]h(\cdot\,,\cdot)f(\cdot\,,0) + Khf(\cdot\,,0) \\
&= f(0\,,0)[R_w^*, R_w][R_z^*, R_z]h + Khf(\cdot\,,0),
\end{aligned} \tag{5.3}
$$

where K a Hilbert-Schmidt operator from Theorem 4.6. If we let A, B be operators from $[h]$ to itself such that for any $hf \in [h]$

$$
Ahf = f(0,0)h; \quad Bhf = h(\cdot\,,\cdot)f(\cdot\,,0),
$$

then the above computation shows that

$$
[R_z^*, R_z][R_w^*, R_w] = [R_w^*, R_w][R_z^*, R_z]A + KB.
$$

We observe that A is a rank one operator with kernel $\overline{z[h] + w[h]}$ and one verifies that $[h] \ominus (\overline{z[h] + w[h]})$ is one dimensional, hence A is a bounded. Thus to prove that $[R_z^*, R_z][R_w^*, R_w]$ is Hilbert-Schmidt, it suffices to check that B is bounded, but this is clear from our assumption on h and Lemma 4.3.

If $h(z, \lambda)$ is bounded away from zero on $\mathbb{T}$ for some non-zero $\lambda \in \mathbb{D}$, then similar computations will show that $[R_z^*, R_z][R_{\varphi_\lambda}^*, R_{\varphi_\lambda}]$ is Hilbert-Schmidt. Then applying Lemma 1.3 twice will establish the assertion. ∎

One sees that the proof of Theorem 5.1 depends heavily on the fact that R_z, R_w are isometries. A corresponding study for the product $[S_z^*, S_z][S_w^*, S_w]$ is thus expected to be harder and we plan to return to that at a later time.

6. Concluding remarks

In this section we will generalize the major theorems obtained so far to the case when $[h]$ is replaced by submodules generated by a finite number of polynomials. Here we need a fact from commutative algebra which we state in a form that fits into our work. Readers may find more information in [9]. We thank Professor C. Sah for showing us his proof of the following statement.

Lemma 6.1. *Suppose* $p_1, p_2, \ldots, p_k$ *are polynomials in* $C[z, w]$ *such that the greatest common divisor* $\mathrm{GCD}(p_1, p_2, \ldots, p_k) = 1$*, then the quotient*

$$C[z, w]/(p_1, p_2, \ldots, p_k)$$

is finite dimensional.

Proof. First of all, $C[z, w]$ is a Unique Factorization Domain (UFD) of Krull dimension 2.

We denote the ideal $(p_1, p_2, \ldots, p_k)$ by I and suppose

$$I = \bigcap_{s=1}^{n} I_s$$

is the irredundant primary representation of I. If we let $J_s = \sqrt{I_s}$ be the radical of I_s, $s = 1, 2, \ldots, n$, then each J_s is prime and it is either maximal or minimal since the Krull dimension of $C[z, w]$ is 2. In an UFD, every minimal prime ideal is principal ([12], p. 238). Since $\mathrm{GCD}(p_1, p_2, \ldots, p_k) = 1$, the associated prime ideals $J_1, J_2, \ldots, J_s$ must all be maximal and hence each J_s must have the form $(z - z_s, w - w_s)$ with $(z_s, w_s) \in \mathbb{C}^2$, $s = 1, 2, \ldots, n$, mutually different. Therefore, we can choose an integer, say m, sufficiently large such that

$$J_s^m = (z - z_s, w - w_s)^m \subset I_s$$

for each s. Then,

$$\bigcap_{s=1}^{n} J_s^m \subset \bigcap_{s=1}^{n} I_s = I,$$

and therefore,

$$\dim(C[z, w]/I) \leqslant \dim\Big(C[z, w]/\Big(\bigcap_{s=1}^{n} J_s^m\Big)\Big).$$

By the Nullstellensatz, one easily checks that

$$J_i^m + J_j^m = C[z, w], \quad i \neq j.$$

The Chinese Remainder Theorem then implies that

$$C[z,w]/\Big(\bigcap_{s=1}^{n} J_s^m\Big) = \prod_{s=1}^{n} C[z,w]/J_s^m,$$

and hence

$$\dim(C[z,w]/I) \leqslant \prod_{s=1}^{n} \dim(C[z,w]/J_s^m) = \left(\frac{m(m+1)}{2}\right)^n. \quad ∎$$

It would be interesting to generalize this lemma to polynomial rings of higher Krull dimensions.

If $h_1, h_2, \ldots, h_k$ are polynomials and we set

$$G = \mathrm{GCD}(h_1, h_2, \ldots, h_k) \quad \text{and} \quad f_j = h_j/G, \tag{6.1}$$

$j = 1, 2, \ldots, k$; then

$$\mathrm{GCD}(f_1, f_2, \ldots, f_k) = 1.$$

If $\{e_1, e_2, \ldots, e_m\}$ is a basis for

$$C[z,w]/(f_1, f_2, \ldots, f_k),$$

then for any polynomial $g(z,w)$,

$$g(z,w) = \sum_{i=1}^{m} c_i e_i(z,w) + r(z,w)$$

with $r \in (f_1, f_2, \ldots, f_k)$ and some constants c_i, $i = 1, 2, \ldots, m$. Therefore,

$$G(z,w)g(z,w) = \sum_{i=1}^{m} c_i G(z,w) e_i(z,w) + G(z,w) r(z,w). \tag{6.2}$$

It is easy to see that $G(z,w)r(z,w) \in (h_1, h_2, \ldots, h_k)$ and hence $(G)/(h_1, h_2, \ldots, h_k)$ is also finite dimensional.

COROLLARY 6.2. *If M is a submodule of $H^2(\mathbb{D}^2)$ generated by a finite number of polynomials, then*

(i) $[S_z^*, S_w]$ *is Hilbert-Schmidt on* $H^2(\mathbb{D}^2) \ominus M$;

(ii) $[R_z^*, R_w]$ *is Hilbert-Schmidt on* M;

(iii) $[R_z^*, R_z][R_w^*, R_w]$ *is Hilbert-Schmidt on* M.

Proof. Suppose $h_1, h_2, \dots, h_k$ are polynomials and $M = [h_1, h_2, \dots, h_k]$ is the closed submodule generated by $h_1, h_2, \dots, h_k$. We assume G, f_i, $i = 1, 2, \dots, k$, and e_j, $j = 1, 2, \dots, m$ to be as in (6.1) and (6.2). Consider the space

$$\mathcal{K} := \operatorname{span}\{e_j \mid j = 1, 2, \dots, m\} + M.$$

It is closed because $\operatorname{span}\{e_j \mid j = 1, 2, \dots, m\}$ is finite dimensional. For any polynomial g, identity (6.2) implies that $Gg \in \mathcal{K}$, and hence $[G] \subset \mathcal{K}$. The inclusion

$$[G] \ominus M \subset \mathcal{K} \ominus M$$

then forces $[G] \ominus M$ to be finite dimensional. We let

$$\begin{aligned}
&p_G : H^2(\mathbb{D}^2) \to [G], \quad q_G : H^2(\mathbb{D}^2) \to H^2(\mathbb{D}^2) \ominus [G],\\
&p_M : H^2(\mathbb{D}^2) \to M, \quad q_M : H^2(\mathbb{D}^2) \to H^2(\mathbb{D}^2) \ominus M,\\
&p_\perp : H^2(\mathbb{D}^2) \to [G] \ominus M,
\end{aligned}$$

be the projections. Then $p_\perp$ is of finite rank and

$$p_G = p_M + p_\perp, \quad q_G = q_M - p_\perp.$$

One verifies that

$$\begin{aligned}
p_G z p_G &= p_M z p_M + p_M z p_\perp + p_\perp z p_M + p_\perp z p_\perp,\\
q_G z q_G &= q_M z q_M - q_M z p_\perp - p_\perp z q_M + p_\perp z p_\perp,
\end{aligned}$$

and consequently, $p_G z p_G - p_M z p_M$ and $q_G z q_G - q_M z q_M$ are of finite rank. Similarly, $q_G w q_G - q_M w q_M$ and $q_G w q_G - q_M w q_M$ are also of finite rank. The assertion in this corollary then follows easily from Theorems 3.6, 4.6 and 5.1. ∎

We conclude this paper by a conjecture suggested by Corollary 6.2.

Conjecture. The assertions in Corollary 6.2 still hold if M is replaced by any finitely generated submodule.

Acknowledgements. This paper constitutes part of the author's Ph.D. Dissertation written at the State University of New York at Stony Brook and the Texas A & M University under the direction of Prof. Ronald G. Douglas. The author would like to thank Prof. Douglas for many valuable discussions and suggestions.

REFERENCES

1. C. Berger, B. Shaw, Selfcommutators of multicyclic hyponormal operators are always trace class, *Bull. Amer. Math. Soc.* **79**(1973), 1193–1199.
2. R. Curto, Problems in multivariable operator theory, *Contemp. Math.* **120**(1991), 15–17.
3. R. Curto, P. Muhly, K. Yan, The C^*-algebra of an homogeneous ideal in two variables is type I, in *Current Topics in Operator Algebras*, Nara 1990, pp. 130–136.
4. R. Douglas, Invariants for Hilbert modules, *Proc. Sympos. Pure Math.* **51**(1990), 179–196.
5. R. Douglas, V. Paulsen, *Hilbert Modules over the Function Algebras*, Pitman Res. Notes Math. Ser., vol. 217, Longman Sci. Tech., Harlow 1989.
6. R. Douglas, K. Yan, A multivariable Berger-Shaw theorem, *J. Operator Theory* **27**(1992), 205–217.
7. I.C. Gohberg, M.G. Krain, *Introduction to the Theory of Linear Nonselfajoint Operators*, Transl. Math. Monographs, vol. 18, Amer. Math. Soc., Providence 1969.
8. H. Grauert, K. Fritzsche, *Several complex variables*, Graduate Texts in Math., Springer-Verlag, vol. 38, 1976.
9. K. Kendig, *Elementary Algebraic Geometry*, Springer-Verlag, New York 1977.
10. N.K. Nikolskii, *Treatise on the Shift Operator*, Grundlehren Math. Wiss., vol. 273, Springer-Verlag, Berlin–Heidelberg 1986.
11. W. Rudin, *Function Theory in Polydiscs*, New York, Benjamin 1969.
12. O. Zariski, P. Samuel, *Commutative Algebra*, vol. 1, Springer-Verlag, New York 1975.

RONGWEI YANG
Department of Mathematics
University of Georgia
Athens, GA 30602
USA

E-mail: ryang@math.uga.edu

Received November 4, 1997.