ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Functional Analysis 213 (2004) 380–411
JOURNAL OF Functional Analysis
http://www.elsevier.com/locate/jfa

Defect operators, defect functions and defect indices for analytic submodules

Kunyu Guo

Department of Mathematics, Fudan University, Shanghai 200433, China

Received 10 June 2003; revised 2 September 2003; accepted 3 November 2003

Communicated by G. Pisier

Abstract

This paper mainly concerns defect operators and defect functions of Hardy submodules, Bergman submodules over the unit ball, and Hardy submodules over the polydisk. The defect operator (function) carries key information about operator theory (function theory) and structure of analytic submodules. The problem when a submodule has finite defect is attacked for both Hardy submodules and Bergman submodules. Our interest will be in submodules generated by polynomials. The reason for choosing such submodules is to understand the interaction of operator theory, function theory and algebraic geometry.

MSC: 47A13; 47A20; 46H25; 46C99

Keywords: Defector operators; Submodules; Finite defect problem

1. Introduction

The classical Beurling's theorem [Beu] says that for each invariant subspace M of the Hardy space $H^2(\mathbb{D})$ on the unit disk $\mathbb{D}$ there is an inner function η such that $M = \eta H^2(\mathbb{D})$. Since $H^2(\mathbb{D})$ admits a natural $C[z]$-module structure coming from multiplication by polynomials, we will call the Hardy space as the Hardy module, and an invariant subspace as a submodule (over the polynomial ring $C[z]$). Let P_M denote the orthogonal projection from $H^2(\mathbb{D})$ onto M. Then Beurling's theorem means

$$P_M = M_\eta M_\eta^*.$$

E-mail address: kyguo@fudan.edu.cn.

0022-1236/$ - see front matter
doi:10.1016/j.jfa.2003.11.004

To recover the inner function η from the representation of P_M we let K_λ and K_λ^M be the reproducing kernels of $H^2(\mathbb{D})$ and M, respectively. Then

$$K_\lambda^M = P_M K_\lambda = \overline{\eta(\lambda)}\eta K_\lambda,$$

and hence

$$|\eta(\lambda)|^2 = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = ||P_M k_\lambda||^2, \quad \lambda \in \mathbb{D}, \tag{1.1}$$

where $k_\lambda = K_\lambda/||K_\lambda||$ is the normalized reproducing kernel. Moreover, it is easy to see

$$||P_M k_\lambda||^2 = \langle (\eta \otimes \eta) K_\lambda, K_\lambda \rangle. \tag{1.2}$$

To generalize the operator-theoretic aspects of function theory on the unit disk to multi-variable operator theory, one considers the Hardy space $H^2(\mathbb{D}^d)$ on the unit polydisk in the d-dimensional complex space $\mathbb{C}^d$. We endow $H^2(\mathbb{D}^d)$ with the $C[z_1, \dots, z_d]$-module structure coming from multiplication of polynomials. In Douglas and Paulsen's Hilbert module language [DP], we will call $H^2(\mathbb{D}^d)$ the Hardy module over the polydisk. By a submodule M of $H^2(\mathbb{D}^d)$ we mean that M is closed, and invariant under multiplication by polynomials. A natural problem is to consider the structure of submodules of $H^2(\mathbb{D}^d)$. However, one quickly sees that a Beurling-like characterization is impossible [DP,Ru1], and hence attention is directed to find intrinsic notions of characterizing higher dimensional submodules. Along this line, many efforts were made by several authors (cf. [CG,DP,DPSY,DY,Fa,Guo1, Guo2,GY,Ya,Ya1,Ya2,Ya3,Ya4]).

Motivated by (1.1), we introduce defect function of a submodule M as follows:

$$D_M(\lambda) = ||K_\lambda^M||^2/||K_\lambda||^2 = ||P_M k_\lambda||, \quad \lambda \in \mathbb{D}^d, \tag{1.3}$$

where K_λ and k_λ are the reproducing kernel and the normalized reproducing kernel of $H^2(\mathbb{D}^d)$, respectively. Then one finds that the defect function D_M is a complete invariant in the sense of function theory, and its comparison to (1.1) shows that defect function is a higher dimensional function-theoretic counterpart of an inner function which expresses a submodule of $H^2(\mathbb{D})$. Moreover, motivated by (1.2), one easily check that there is a unique bounded linear operator, denoted by Δ_M, such that

$$D_M(\lambda) = \langle \Delta_M K_\lambda, K_\lambda \rangle, \quad \lambda \in \mathbb{D}^d. \tag{1.4}$$

This operator is called the defect operator of M, and its exact form will be given in Section 2. The defect operator is an invariant of submodules in the sense of operator theory. One will see that defect operator (function) carries key information about operator theory (function theory) and structure of submodules. We define the defect index of a submodule M as the rank of the defect operator Δ_M. This numerical variant for submodules plays a role of multiplicity of a submodule in an appropriate sense. By Beurling's theorem, each submodule of $H^2(\mathbb{D})$ has defect index 1. To understand higher dimensional submodules better, we are naturally led to ask when a submodule of $H^2(\mathbb{D}^d)$ $(d\geqslant 2)$ has finite defect, that is, when the defect operator Δ_M has finite rank for a higher dimensional submodule M. This is the so-called "Finite defect problem" for Hardy submodules over the polydisk.

We also will be concerned with the two most common Hilbert modules on the unit ball of $\mathbb{C}^d$, namely, the Hardy module and the Bergman module. There are several reasons for studying defect operators of submodules on the unit ball. One reason is that the theory of defect operators (functions) relies heavily on geometry of domains on which submodules are defined. As one knows, the ball is the prototype of two important classes of regions that have been studied in depth, namely, the strictly pseudoconvex domains and the bounded symmetric domains. Another reason is that the theory of defect operators (functions) is closely related to reproducing kernel theory. Other reasons will become apparent later.

In the present paper, we are mainly concerned with submodules generated by polynomials. The reason for choosing such submodules is to understand the interaction of operator theory, function theory and algebraic geometry.

Section 2 considers defect operators, defect functions and defect indices of Hardy submodules over the polydisk. Actually one finds that defect operators (functions) reveal rigidity of submodules. In Section 3, our interest is in "Finite defect problem" for Hardy submodules over the unit ball. Using the theories of algebraic variety and analytic variety it is shown that a submodule generated by polynomials has finite defect only if the submodule has finite codimension. Section 4 concerns defect operators (functions) of Bergman submodules over the unit ball.

2. Defect operators for Hardy submodules over the polydisk

2.1. Definitions and examples

Given an invariant subspace M of the Hardy space $H^2(\mathbb{D})$ over the unit disk $\mathbb{D}$, the Beurling's theorem [Beu] implies that there is an inner function η such that $M=\eta H^2(\mathbb{D})$. As in Introduction, let K_λ and K_λ^M be the reproducing kernels of $H^2(\mathbb{D})$ and M, respectively. Then

$$K_\lambda^M = P_M K_\lambda = \overline{\eta(\lambda)}\eta K_\lambda,$$

and hence

$$|\eta(\lambda)|^2 = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = ||P_M k_\lambda||^2, \tag{2.1}$$

here $k_\lambda = K_\lambda/||K_\lambda||$. This shows that the inner function can be captured from (2.1).

Recall that the Hardy space $H^2(\mathbb{D}^d)$ over the polydisk $\mathbb{D}^d$ is a functional Hilbert space consisting of some analytic functions on the polydisk $\mathbb{D}^d$ whose reproducing kernel and the normalized reproducing kernel are, respectively

$$K_\lambda(z) = \frac{1}{\prod_{k=1}^d (1-\overline{\lambda}_k z_k)}, \quad k_\lambda(z) = \frac{K_\lambda}{||K_\lambda||} = \frac{\prod_{k=1}^d (1-|\lambda_k|^2)^{1/2}}{\prod_{k=1}^d (1-\overline{\lambda}_k z_k)}.$$

Noticing that the Hardy space $H^2(\mathbb{D}^d)$ admits a natural $C[z_1, \ldots, z_d]$-module structure coming from multiplication by polynomials, we thus call the Hardy space $H^2(\mathbb{D}^d)$ as the Hardy module over the polydisk. By a submodule M of $H^2(\mathbb{D}^d)$ we mean that M is a closed invariant subspace under multiplication by polynomials. Given a submodule M, the defect function $D_M(\lambda)$ of M is defined by

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = ||P_M k_\lambda||^2, \quad \lambda \in \mathbb{D}^d, \tag{2.2}$$

where $K_\lambda^M = P_M K_\lambda$ be the reproducing kernel of the submodule M and P_M is the orthogonal projection from $H^2(\mathbb{D}^d)$ onto M. Letting $R_i = P_M M_{z_i} P_M$ be the restriction of M_{z_i} to the submodule M, then R_i is an isometry on M for $1 \leqslant i \leqslant d$. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ of nonnegative integers, let $R^\alpha = R_1^{\alpha_1} \cdots R_d^{\alpha_d}$, and, as usual, $|\alpha| = \alpha_1 + \cdots + \alpha_d$. Then from the definition of the defect function D_M, it is easy to check that there exists a unique bounded linear operator, denoted by Δ_M, such that

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = \prod_{k=1}^d (1-|\lambda_k|^2)||P_M K_\lambda||^2 = \langle \Delta_M K_\lambda, K_\lambda \rangle, \tag{2.3}$$

where

$$\Delta_M = \sum_{0 \leqslant \alpha \leqslant (1,\ldots,1)} (-1)^{|\alpha|} R^\alpha R^{*\alpha} = \sum_{0 \leqslant \alpha \leqslant (1,\ldots,1)} (-1)^{|\alpha|} M_z^\alpha P_M M_z^{*\alpha}. \tag{2.4}$$

We call Δ_M the defect operator of the submodule M. The reader also notice that the defect operator Δ_M and the projection P_M is connected by

$$P_M K_\lambda = K_\lambda \, \Delta_M K_\lambda, \quad \lambda \in \mathbb{D}^d. \tag{2.5}$$

In particular, for a submodule M of the Hardy module $H^2(\mathbb{D}^2)$ on the bidisk, its defect operator is

$$\Delta_M = P_M - M_{z_1} P_M M_{z_1}^* - M_{z_2} P_M M_{z_2}^* + M_{z_1 z_2} P_M M_{z_1 z_2}^*.$$

Similar to the case of the symmetric Fock space H_d^2 [Guo3], we will see that a submodule is uniquely determined by its defect function, and hence is uniquely determined by its defect operator.

Proposition 2.1. *For two submodules* M *and* N *of* $H^2(\mathbb{D}^d)$, *if* $D_M(\lambda) = D_N(\lambda)$, $\forall \lambda \in \mathbb{D}^d$, *then* $M = N$, *and therefore, if* $\Delta_M = \Delta_N$, *then* $M = N$.

To prove the Proposition we need the following lemma. The proof of the lemma appeared in [Eng]. Of course, the lemma can also be proved by using Taylor expansion.

Lemma 2.2. *Let* Ω *be a bounded complete Reinhart domain* (*i.e. a bounded domain with the property that for* $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_d) \in \Omega$, *if* $|\mu_i| \leqslant 1$, $i = 1, 2 \dots, d$, *then* $(\mu_1\lambda_1, \mu_2\lambda_2, \dots, \mu_d\lambda_d) \in \Omega$). *Suppose a function* $f(\lambda, z)$ *is defined on* $\Omega \times \Omega$, *and it is analytic in* z, *and co-analytic in* λ. *If* $f(\lambda, \lambda) = 0$ *for any* $\lambda \in \Omega$, *then* $f = 0$.

The Proof of Proposition 2.1. As done for the symmetric Fock space H_d^2 [Guo3], to obtain the desired conclusion, considering functions

$$G_M(\lambda, z) = \langle \Delta_M K_\lambda, K_z \rangle, \quad G_N(\lambda, z) = \langle \Delta_N K_\lambda, K_z \rangle,$$

then $G_M(\lambda, z)$ and $G_N(\lambda, z)$ are analytic in z, and co-analytic in λ, respectively. By (2.3)

$$G_M(\lambda, \lambda) = D_M(\lambda) = D_N(\lambda) = G_N(\lambda, \lambda).$$

Applying Lemma 2.2 gives

$$G_M(\lambda, z) = G_N(\lambda, z),$$

and hence $\Delta_M = \Delta_N$. By (2.5)

$$\begin{aligned}(1 - \bar{\lambda}_1 z_1) \cdots (1 - \bar{\lambda}_d z_d) \langle P_M K_\lambda, K_z \rangle \\ = \langle \Delta_M K_\lambda, K_z \rangle = \langle \Delta_N K_\lambda, K_z \rangle \\ = (1 - \bar{\lambda}_1 z_1) \cdots (1 - \bar{\lambda}_d z_d) \langle P_N K_\lambda, K_z \rangle,\end{aligned}$$

and hence, $M = N$. □

Remark 2.3. In our joint paper [GY], we study the core operators and the core functions of Hardy submodules on the bidisk $\mathbb{D}^2$. In fact, the core operator for a submodule just is the defect operator introduced in this section. The core function for a submodule M is defined by $G_M(\lambda, z) = \langle \Delta_M K_\lambda, K_z \rangle$ for $\lambda, z \in \mathbb{D}^2$, and hence the defect function $D_M(\lambda) = G_M(\lambda, \lambda)$.

First let us record some results from [GY] which are proved in the case of the bidisk $\mathbb{D}^2$.

Proposition 2.4. *Let M be a submodule of $H^2(\mathbb{D}^d)$, then we have*

1. *the defect function $D_M(\lambda)$ is subharmonic, and for almost all $z \in \mathbb{T}^d$ with respect to the measure $d\theta_1 \cdots d\theta_d$, $D_M(\lambda) \to 1$ as $\lambda \to z$, where $\mathbb{T}^d = \{z = (z_1, \dots, z_d) : |z_1| = \cdots = |z_d| = 1\}$ is the distinguished boundary of $\mathbb{D}^d$;*
2. *the defect operator $\Delta_M \geqslant 0$ if and only if there is an inner function η such that $M = \eta H^2(\mathbb{D}^d)$;*
3. *if Δ_M is in trace class, then* Trace $\Delta_M = 1$.

Proof. It is easy to see that the proof for term (1) is completely parallel to the case of the bidisk in [GY]. But term (2) was not mentioned in [GY]. Here we verify term (2). In fact, if there is an inner function η such that $M = \eta H^2(\mathbb{D}^d)$, then $P_M = M_\eta M_\eta^*$, and hence $\Delta_M = \eta \otimes \eta \geqslant 0$. To verify the opposite direction, notice that there exist a sequence $\{\phi_n\} \subset M$ such that

$$\Delta_M = \sum_n \phi_n \otimes \phi_n. \quad (SOT)$$

By (2.2) and (2.3)

$$D_M(\lambda) = ||P_M k_\lambda||^2 = \langle \Delta_M K_\lambda, K_\lambda \rangle = \sum_n |\phi_n(\lambda)|^2 \leqslant 1,$$

and by the term (1), $\sum_n |\phi_n(z)|^2 = 1$ on $\mathbb{T}^d$. This implies that

$$\text{Trace}(\Delta_M) = \sum_n ||\phi_n||^2 = \frac{1}{(2\pi)^d} \sum_n \int_{\mathbb{T}^d} |\phi_n(z)|^2 d\theta_1 \cdots d\theta_d = 1.$$

This shows that Δ_M is trace class. Similarly to the proof of Corollary 3.4 in [DY], we have that

$$\mathbb{S} = M \ominus \overline{(z_1 M + \cdots + z_d M)} \neq \{0\}$$

and it is easy to check that every function from $\mathbb{S}$ is eigenvector of Δ_M with the corresponding eigenvalue 1. Taking a function η from $\mathbb{S}$ with $||\eta|| = 1$, then $\Delta_M - \eta \otimes \eta \geqslant 0$. Note that

$$\text{Trace}(\Delta_M - \eta \otimes \eta) = 0$$

and hence $\Delta_M = \eta \otimes \eta$. By Proposition 2.4(1), η is an inner function. Applying Proposition 2.1 we see $M = \eta H^2(\mathbb{D}^d)$. Term (3) was proved in the bidisk case in [GY]. In fact, term (3) directly comes from (1). To see this, let μ_n, $n = 1, 2, \dots$ be eigenvalues of Δ_M counting multiplicity, and ϕ_n be the corresponding unital eigenvectors. Then $\Delta_M = \sum_n \mu_n \phi_n \otimes \phi_n$. By term (1) and (2.3), we have

$$\text{Trace}\, \Delta_M = \sum_n \mu_n = \frac{1}{(2\pi)^d} \sum_n \mu_n \int_{\mathbb{T}^d} |\phi_n(z)|^2 d\theta_1 \cdots d\theta_d$$
$$= \frac{1}{(2\pi)^d} \int_{\mathbb{T}^d} D_M(z) d\theta_1 \cdots d\theta_d = 1.$$

Now turn to general isometric Hilbert modules. Let $T = (T_1, \dots, T_d)$ be a tuple of commuting operators acting on a Hilbert space H. Then, one naturally makes H into a Hilbert module over the polynomial ring $C[z_1, \dots, z_d]$. The $C[z_1, \dots, z_d]$-module structure is define by

$$p \cdot \xi = p(T_1, \dots, T_d)\xi, \quad p \in C[z_1, \dots, z_d], \xi \in H.$$

We say that two modules H_1 and H_2 over the polynomial ring $C[z_1, \dots, z_d]$ are unitarily equivalent if there is a unitary module map U from H_1 onto H_2, that is, $U : H_1 \to H_2$ is a unitary operator, and $Up \cdot f = p \cdot Uf$ for any polynomial p and $f \in H_1$. By an isometric Hilbert module H over $C[z_1, \dots, z_d]$ we mean H's canonical operators $T_1, \dots, T_d$ are isometries. Then each submodules of $H^2(\mathbb{D}^d)$ is an isometric Hilbert modules. As done for Hardy submodules, by following (2.4) we define the defect operator Δ_H of an isometric module H as:

$$\Delta_H = \sum_{0 \leqslant \alpha \leqslant (1,\dots,1)} (-1)^{|\alpha|} T^\alpha T^{*\alpha}, \tag{2.6}$$

and define the defect index of the module H as the rank of the defect operator Δ_H. It follows that if two isometric modules H_1 and H_2 are unitarily equivalent, then their defect operators are unitarily equivalent, and hence two modules have the same defect index. The next examples will show that defect index play a role of multiplicity of an isometric module in an appropriate sense.

Example 1. Let H be an isometric module over $C[z]$ whose canonical operator S is a pure isometry (that is, S satisfies $S^{*n} \to 0$ in the strong operator topology). Then by the Von Neumann–Wold Decomposition theorem, H is unitarily equivalent to $H^2(\mathbb{D}) \otimes \mathbb{C}^n$, where $n = \text{rank}(I - SS^*)$ is the defect index of H.

Example 2. Let H be an isometric module over $C[z_1, z_2]$ whose canonical operators $S = (S_1, S_2)$ are pure isometries. By [AEM], $\Delta_H \geqslant 0$ if and only if H is unitarily equivalent to $H^2(\mathbb{D}^2) \otimes \mathbb{C}^n$, where $n = \text{rank}(\Delta_H)$ is the defect index of H.

Let us make a simple comment about defect indices of Hardy submodules. By Proposition 2.4(1), a submodule M of $H^2(\mathbb{D}^d)$ has defect index 1 if and only if $\Delta_M \geqslant 0$, if and only if $M = \eta H^2(\mathbb{D}^d)$ for some inner function η. For a submodule M

of $H^2(\mathbb{D}^d)$, since

$$\mathbb{S} = M \ominus \overline{(z_1 M + \cdots + z_d M)} \neq \{0\}$$

and every function from $\mathbb{S}$ is eigenvector of Δ_M with the corresponding eigenvalue 1, combining this fact with Proposition 2.4(3) implies that there is not a submodule M whose defect index is 2. Recently, Yang studied a class of submodules over the bidisk, the so-called M_q-type submodules [Ya4]. A M_q-type submodule M is

$$M = \eta_1(z)H^2(\mathbb{D}^2) + \eta_2(w)(\mathbb{D}^2),$$

where η_1 and η_2 are nontrivial inner functions. It is shown that such a M is closed, and its defect operator is

$$\Delta_M = \eta_1(z) \otimes \eta_2(z) + \eta_2(w) \otimes \eta_2(w) - \eta_1(z)\eta_2(w) \otimes \eta_1(z)\eta_2(w),$$

and hence M_q-type submodules have defect index 3 (see [Ya4]). Moreover, a careful verification shows that the submodule $[z-w, zw]$ generated by $z-w$ and zw on the bidisk has defect index 5.

The following example will show that defect operator capture key information about operator theory.

Example 3. Let H be an isometric module over $C[z_1, \dots, z_d]$ $(d \geqslant 2)$ with canonical operators $S_1, \dots, S_d$. Assume some S_i has finite multiplicity $\operatorname{rank}(I - S_i S_i^*) < \infty$. Then there is no submodule M of $H^2(\mathbb{D}^d)$ such that M is unitarily equivalent to H. To show this, we may assume $\operatorname{rank}(I - S_d S_d^*) < \infty$, and write P for $I - S_d S_d^*$. Then

$$\Delta_H = \sum_{0 \leqslant \alpha \leqslant (1,\dots,1)} (-1)^{|\alpha|} S^\alpha S^{*\alpha} = \sum_{0 \leqslant \alpha' \leqslant (1,\dots,1)} (-1)^{|\alpha'|} \underline{S}^{\alpha'} P \underline{S}^{*\alpha'},$$

where $\alpha' = (\alpha_1, \dots, \alpha_{d-1})$ and $\underline{S} = (S_1, \dots, S_{d-1})$. We thus have

$$\operatorname{Trace} \Delta_H = 0.$$

From Proposition 2.4(3), there is no submodule M of $H^2(\mathbb{D}^d)$ such that M is unitarily equivalent to H.

Example 3 also shows that each M_{z_i}, restricted on a submodule M, has infinite multiplicity. This fact also was noticed by Fang [Fa]. We also find that study for defect operators of isometric modules is relevant in the theory of operator models in the polydisk [AEM,CV]. $\square$

2.2. Finite defect problem

Beurling's theorem shows that each submodule of $H^2(\mathbb{D})$ has defect index 1. To understand higher dimensional submodules better, we are naturally led to ask when a submodule of $H^2(\mathbb{D}^d)$ $(d \geq 2)$ has finite defect, that is, when the defect operator Δ_M has finite rank for a higher dimensional submodule M. This is the so-called "Finite defect problem". Clearly, if M is a finite codimensional submodule, then M has finite defect. From this it is deduced that if M is unitarily equivalent to a finite codimensional submodule, then M has finite defect. Notice that if M is unitarily equivalent to a finite codimensional submodule, say, N, then by [ACD], there is an inner function η such that $M = \eta N$. However, as shown by M_q-type submodules, this never contain all submodules with finite defect indices. Moreover, In [GY], it was shown that for a homogenous submodule M of $H^2(\mathbb{D}^2)$ (i.e. a submodule generated by homogeneous polynomials), M has finite defect if and only if there are a monomial $z_1^s z_2^t$ and a finite codimensional submodule N such that $M = z_1^s z_2^t N$, that is, M is unitarily equivalent to a finite codimensional submodule.

Combining the above facts and the next example in dimension 3 will show that the study for defect operators depends strongly on the dimension of the polydisk.

Example 4. Consider the homogeneous submodule $M = [z_2, z_3]$ of the Hardy module $H^2(\mathbb{D}^3)$ generated by z_2, z_3. It is not difficult to verify that the reproducing kernel of the submodule M is

$$K_\lambda^M(z) = \frac{1}{(1-\bar{\lambda}_1 z_1)(1-\bar{\lambda}_2 z_2)(1-\bar{\lambda}_3 z_3)} - \frac{1}{(1-\bar{\lambda}_1 z_1)}.$$

By (2.5) we have

$$\begin{aligned}\Delta_M K_\lambda(z) &= \bar{\lambda}_2 z_2 + \bar{\lambda}_3 z_3 - \overline{\lambda_2 \lambda_3} z_2 z_3 \\ &= [(z_2 \otimes z_2 + z_3 \otimes z_3 - z_2 z_3 \otimes z_2 z_3) K_\lambda](z),\end{aligned}$$

and hence

$$\Delta_M = z_2 \otimes z_2 + z_3 \otimes z_3 - z_2 z_3 \otimes z_2 z_3.$$

However, by Theorem 3.1 in [Guo2] the submodule $M = [z_2, z_3]$ is not unitarily equivalent to any finite codimensional submodule.

From the above several observations, the answer to the following question may be difficult.

Finite defect problem. Can one completely characterize those higher dimensional submodules which have finite defect indices?

In the dimension $d = 2$, we can completely characterize those submodules generated by polynomials that have finite defect indices.

For this we need some preliminaries.

Let I be an ideal of the polynomials ring $C[z_1, \ldots, z_d]$. Since the polynomial ring $C[z_1, \ldots, z_d]$ is Noetherian [ZS], the ideal I is generated by finitely many polynomials. This implies that I has a greatest common divisor p, and so, I can be uniquely written as $I = pL$, which is called the Beurling form of I (cf. [Guo2]). For a polynomial p with $Z(p) \cap \mathbb{D}^d \neq \emptyset$, we decompose $p = p_1 p_2$ such that the zero set of each prime factor of p_1 meets $\mathbb{D}^d$ nontrivially, and $Z(p_2) \cap \mathbb{D}^n = \emptyset$. Define $L(p)$ on $C[z_1, \dots, z_d]$ as follows: $L(p) = 1$ if $Z(p) \cap \mathbb{D}^d = \emptyset$; $L(p) = p_1$ if $Z(p) \cap \mathbb{D}^d \neq \emptyset$. Set

$$\mathscr{V}_d = \{z \in \mathbb{C}^d : |z_i| > 1, \text{ for } i = 1, 2, \dots, d\}.$$

For an ideal I of the polynomial ring $C[z_1, \dots, z_d]$, as usual, we write $[I]$ for the submodule generated by I, that is, $[I] = \overline{I}$.

Theorem 2.5. *Let I be an ideal of the polynomial ring $C[z_1, z_2]$, and let $I = pL$ be its Beurling form. Then $[I]$, as a submodule of $H^2(\mathbb{D}^2)$, has finite defect index if and only if $Z(L(p)) \cap \mathscr{V}_2 = \emptyset$, and in this case there are a rational inner function η and a finite codimensional submodule N such that $[I] = \eta N$.*

Theorem 2.5 comes from the next Theorem 2.6 whose proof is long. We will place the proof of Theorem 2.6 at the end of this section.

Theorem 2.6. *Let I be an ideal of $C[z_1, \dots, z_d]$, and let $I = pL$ be its Beurling form. If $[I]$ has finite defect, then there exists a polynomial r satisfying $Z(r) \cap \mathbb{D}^d = \emptyset$ such that $|p| = |r|$ on $\mathbb{T}^d$.*

Let r be polynomial in Theorem 2.6 satisfying $Z(r) \cap \mathbb{D}^d = \emptyset$. Then by [Guo2, Proposition 2.9], the principal submodule $[r]$ generated by r equals $H^2(\mathbb{D}^d)$, and therefore, r is outer in Rudin's sense [Ru1]. Consequently, the condition $|p| = |r|$ on $\mathbb{T}^d$ implies that p/r is a rational inner function. Note that r is uniquely determined by p except for a modular 1 constant because outer function r is uniquely determined by the restriction of $|r|$ to $\mathbb{T}^d$. Writing η_p for the rational inner function p/r, then we have

Corollary 2.7. *Let I be an ideal of $C[z_1, \dots, z_d]$, and let $I = pL$ be its Beurling form. If $[I]$ has finite defect, then $[I] = \eta_p [L]$.*

Recall a theorem, due to Rudin [Ru1, Theorem 5.2.6]. This theorem says that a polynomials p is the numerator of a rational inner function on $\mathbb{D}^d$ if and only if p has no zero in $\mathscr{V}_d$. Combining Rudin's theorem, Theorem 2.6 and fact that $[q] = H^2(\mathbb{D}^d)$ if a polynomial q satisfies $Z(q) \cap \mathbb{D}^d = \emptyset$, we have

Corollary 2.8. *Let I be an ideal of $C[z_1, \dots, z_d]$, and let $I = pL$ be its Beurling form. If $Z(L(p)) \cap \mathscr{V}_d \neq \emptyset$, then $[I]$ is of infinite defect.*

By Corollary 2.8 we can obtain the following.

Corollary 2.9. *Let p is a polynomial. Then the principal submodule* $[p]$ *has finite defect if and only if* $Z(L(p)) \cap \mathscr{V}_d = \emptyset$, *and in this case there is a rational inner function* η_p *such that* $[p] = \eta_p H^2(\mathbb{D}^d)$.

Proof. The necessariness is from Corollary 2.8. To achieve the opposite direction, assume that $Z(L(p)) \cap \mathscr{V}_d = \emptyset$. Then by Rudin's theorem mentioned above, there is a polynomial r satisfying $Z(r) \cap \mathbb{D}^d = \emptyset$ such that $L(p)/r$ is a rational inner function. It follows that

$$[p] = \frac{L(p)}{r}[r] = \eta_p H^2(\mathbb{D}^d).$$

This gives the desired conclusion. □

The proof of Theorem 2.5. If $[I]$ has finite defect, then applying Corollary 2.8 gives $Z(L(p)) \cap \mathscr{V}_2 = \emptyset$. To reach at the opposite direction, by Rudin's theorem mentioned above, there is a polynomial r satisfying $Z(r) \cap \mathbb{D}^2 = \emptyset$ such that $L(p)/r$ is a rational inner function. Since the greatest common divisor of L is 1, and hence by [Ya1], L is a finite codimensional ideal of $C[z_1, z_2]$ and it follows that the submodule $[L]$ is finite codimensional. Note that $[rL] = [L]$ by [Ge]. This means that

$$[I] = \frac{L(p)}{r}[rL] = \eta_p[L],$$

and hence $[I]$ is unitarily equivalent to the finite codimensional submodule $[L]$, and so, $[I]$ has finite defect.

Let us see an example.

Example 5. We consider the submodule $[z_1 + z_2 + \alpha]$ of $H^2(\mathbb{D}^2)$, where α is constant. If $|\alpha| \geqslant 2$, then by [Ge],

$$[z_1 + z_2 + \alpha] = H^2(\mathbb{D}^2),$$

and hence in this case, the submodule has defect index 1. If $|\alpha| < 2$, then $Z(z_1 + z_2 + \alpha) \cap \mathscr{V}_2 \neq \emptyset$, and hence by Corollary 2.8, the submodule $[z_1 + z_2 + \alpha]$ has infinite defect.

Now turning to the proof of Theorem 2.6 we need two lemmas. The next lemma comes from [Guo2].

Lemma 2.10. *Let* $f = p/q$ *be a rational function, where p and q are without common factors. If f is analytic on* $\mathbb{D}^d$, *then* $Z(q) \cap \mathbb{D}^d = \emptyset$.

The following lemma is key for the proof of Theorem 2.6.

Lemma 2.11. *Let I be an ideal of* $C[z_1, \ldots, z_d]$. *If the submodule* $[I]$ *has finite defect index l, then there are polynomials* $p_1, \ldots, p_s, q_1, \ldots, q_t$ *(here* $s + t = l$*) and r satisfying the greatest common divisor*

$$GCD(p_1, \ldots, p_s\, q_1, \ldots, q_t, r) = 1$$

and $Z(r) \cap \mathbb{D}^d = \emptyset$ *such that*

1. *the rational functions* $p_1/r, \ldots, p_s/r, q_1/r, \ldots, q_t/r$ *belong to* $[I]$, *and* $\{p_1, \ldots, p_s, q_1, \ldots, q_t\}$ *is a generating set of* $[I]$;
2. $\Delta_{[I]} = \sum_{i=1}^{s} \frac{p_i}{r} \otimes \frac{p_i}{r} - \sum_{j=1}^{t} \frac{q_j}{r} \otimes \frac{q_j}{r}$;
3. $\sum_{i=1}^{s} |p_i(z)|^2 - \sum_{j=1}^{t} |q_j(z)|^2 = |r(z)|^2$ *for* $z \in \mathbb{T}^d$.

Proof. Since $\Delta_{[I]}$ is adjoint, there exist mutually orthogonal vectors $\phi_1, \phi_2, \ldots, \phi_s$ $(s \geqslant 1)$, $\psi_1, \psi_2, \ldots, \psi_t$ $(t \geqslant 0)$ in $\Delta_{[I]} H^2(\mathbb{D}^d)$ (here $s + t = l$) such that

$$\Delta_{[I]} = (\phi_1 \otimes \phi_1 + \phi_2 \otimes \phi_2 + \cdots \phi_s \otimes \phi_s) - (\psi_1 \otimes \psi_1 + \psi_2 \otimes \psi_2 + \cdots \psi_t \otimes \psi_t).$$

By (2.5), one sees

$$P_{[I]} K_\lambda = \left[\sum_{i=1}^{s} \overline{\phi_i(\lambda)} \phi_i - \sum_{j=1}^{t} \overline{\psi_j(\lambda)} \psi_j \right] K_\lambda. \tag{2.7}$$

This implies that the vectors $\{\phi_1, \phi_2, \ldots, \phi_s, \psi_1, \psi_2, \ldots, \psi_t\}$ is a generating set of $[I]$. For any $\phi \in H^2(\mathbb{D}^d)$, consider the densely defined operator M_ϕ in $H^2(\mathbb{D}^d)$ whose domain of definition contains $C[z_1, \ldots, z_d]$ by $M_\phi h = \phi h$. From [Ru5], we see that M_ϕ^* is densely defined, and

$$Dom(M_\phi^*) \supseteq C[z_1, \ldots, z_d].$$

According to (2.7), for any polynomial h and each $\lambda \in \mathbb{D}^d$ the following holds

$$\begin{aligned} &\left[P_{[I]} - \left(\sum_{i=1}^{s} M_{\phi_i} M_{\phi_i}^* - \sum_{j=1}^{t} M_{\psi_j} M_{\psi_j}^* \right) \right] h(\lambda) \\ = \quad &\left\langle \left[P_{[I]} - \left(\sum_{i=1}^{s} M_{\phi_i} M_{\phi_i}^* - \sum_{j=1}^{t} M_{\psi_j} M_{\psi_j}^* \right) \right] h, K_\lambda \right\rangle \\ &= \left\langle h, P_{[I]} K_\lambda - \left[\sum_{i=1}^{s} \overline{\phi_i(\lambda)} \phi_i - \sum_{j=1}^{t} \overline{\psi_j(\lambda)} \psi_j \right] K_\lambda \right\rangle \\ &= 0. \end{aligned}$$

It follows that

$$P_{[I]}h = \left(\sum_{i=1}^{s} M_{\phi_i} M_{\phi_i}^* - \sum_{j=1}^{t} M_{\psi_j} M_{\psi_j}^* \right) h$$

for any polynomial h. It is easy to see that for each $\phi \in H^2(\mathbb{D}^d)$ and every polynomial p, $M_\phi^* p$ also is a polynomial since $H^2(\mathbb{D}^d)$ enjoys an orthonormal basis $\{z^\alpha\}_{\alpha \in \mathbb{Z}_+^d}$. Picking polynomials $q_1, \dots, q_{s+t}$ in I, then

$$\sum_{i=1}^{s} M_{\phi_i} M_{\phi_i}^* q_k - \sum_{j=1}^{t} M_{\psi_j} M_{\psi_j}^* q_k = q_k, \quad k = 1, \dots, s+t. \tag{2.8}$$

Claim. *There exist polynomials $q_1, \dots, q_{s+t}$ in I such that*

$$r = \begin{vmatrix} M_{\phi_1}^* q_1 & \dots & M_{\phi_s}^* q_1 & M_{\psi_1}^* q_1 & \dots & M_{\psi_t}^* q_1 \\ M_{\phi_1}^* q_2 & \dots & M_{\phi_s}^* q_2 & M_{\psi_1}^* q_2 & \dots & M_{\psi_t}^* q_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ M_{\phi_1}^* q_{s+t} & \dots & M_{\phi_s}^* q_{s+t} & M_{\psi_1}^* q_{s+t} & \dots & M_{\psi_t}^* q_{s+t} \end{vmatrix} \not\equiv 0, \tag{2.9}$$

that is, $r(z)$ is a nonzero polynomial. To get a contradiction assume that the above determinant $r(z) \equiv 0$ for any polynomials $q_1, \dots, q_{s+t}$ in I. Then we have $r(0) = 0$. Since

$$r(0) = \begin{vmatrix} \langle q_1, \phi_1 \rangle & \dots & \langle q_1, \phi_s \rangle & \langle q_1, \psi_1 \rangle & \dots & \langle q_1, \psi_t \rangle \\ \langle q_2, \phi_1 \rangle & \dots & \langle q_2, \phi_s \rangle & \langle q_2, \psi_1 \rangle & \dots & \langle q_2, \psi_t \rangle \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \langle q_{s+t}, \phi_1 \rangle & \dots & \langle q_{s+t}, \phi_s \rangle & \langle q_{s+t}, \psi_1 \rangle & \dots & \langle q_{s+t}, \psi_t \rangle \end{vmatrix} = 0, \tag{2.10}$$

and all ϕ_i and ψ_j are in $[I]$, there are polynomial sequences in I, $\{q_1^{(n)}\}, \dots, \{q_s^{(n)}\}$ and $\{p_1^{(n)}\}, \dots, \{p_t^{(n)}\}$ such that $q_i^{(n)}$ converge to ϕ_i, and $p_j^{(n)}$ converge to ψ_j in the norm of $H^2(\mathbb{D}^d)$. Noticing that vectors $\phi_1, \phi_2, \dots, \phi_s, \psi_1, \psi_2, \dots, \psi_t$ are mutually orthogonal, this yields the following:

$$||\phi_1||^2 \cdots ||\phi_s||^2 ||\psi_1||^2 \cdots ||\psi_t||^2 = 0.$$

This contradiction implies that the Claim is true.

By the Claim there exist polynomials $q_1, \dots, q_{s+t}$ in I such that the determinant $r(z) \not\equiv 0$ in (2.9). Noticing that $r(z)$ is a nonzero polynomial, and applying Cramer's rule to solve the corresponding $s+t$ equations (2.8), then there exist polynomials $q_1, r_1, q_2, r_2, \dots, q_{s+t}, r_{s+t}$ satisfying $GCD(q_j, r_j) = 1$ for $j = 1, \dots, s+t$ such that

$$\phi_1 = q_1/r_1, \dots, \phi_s = q_s/r_s, \quad \psi_1 = q_{s+1}/r_{s+1}, \dots, \psi_t = q_{s+t}/r_{s+t}.$$

Since all ϕ_i and ψ_j are analytic on $\mathbb{D}^d$, applying Lemma 2.10 gives

$$Z(r_j) \cap \mathbb{D}^d = \emptyset, \quad j = 1, 2, \dots, s + t.$$

Set

$$q'_j = q_j r_1 \cdots r_{j-1} r_{j+1} \cdots r_{s+t}, \quad j = 1, 2, \dots, s + t, \quad r' = r_1 r_2 \cdots r_{s+t}.$$

Let $q = GCD(q'_1, q'_2, \dots, q'_{s+t}, r')$ and

$$p_1 = q'_1/q, \;\; p_2 = q'_2/q, \dots, p_{s+t} = q'_{s+t}/q, \;\; r = r'/q.$$

Then polynomials $p_1, p_2, \dots, , p_{s+t}, r$ satisfy

$$GCD(p_1, p_2, \dots, p_{s+t}, r) = 1$$

and

$$\phi_1 = p_1/r, \dots, \phi_s = p_s/r, \;\; \psi_1 = p_{s+1}/r, \dots, \psi_t = p_{s+t}/r.$$

Since $\{p_1/r, \dots, p_{s+t}/r\}$ is a generating set of $[I]$, by [Guo2, Proposition 2.9], this insures that $\{p_1, \dots, p_{s+t}\}$ is a generating set of $[I]$. Furthermore, from Proposition 2.4(1) and (2.3) we have

$$\sum_{i=1}^{s} |p_i(z)|^2 - \sum_{j=1}^{t} |p_{s+j}(z)|^2 = |r(z)|^2, \quad z \in \mathbb{T}^d.$$

This completes the proof. $\square$

The proof of Theorem 2.6. We use Lemma 2.11 to complete the proof of the theorem. Considering Lemma 2.11(3)

$$\sum_{i=1}^{s} |p_i(z)|^2 - \sum_{j=1}^{t} |q_j(z)|^2 = |r(z)|^2, \quad z \in \mathbb{T}^d, \tag{2.11}$$

we take a large natural number N satisfying that $\overline{p_i(z)}z^N$, $\overline{q_j(z)}z^N$, $\overline{r(z)}z^N$ become polynomial for $i = 1, 2, \dots, s;\; j = 1, 2, \dots, t$, where $z^N = z_1^N z_2^N \cdots z_d^N$. Multiplying two side of (2.11) by z^N gives

$$\sum_{i=1}^{s} p_i(z)\overline{p_i(z)}z^N - \sum_{j=1}^{t} q_j(z)\overline{q_j(z)}z^N = r(z)\overline{r(z)}z^N. \tag{2.12}$$

Setting

$$q = GCD(p_1, p_2, \dots, p_s, q_1, q_2, \dots, q_t), \tag{2.13}$$

Lemma 2.11 says $GCD(q,r)=1$. Combining this fact with the equality (2.12) shows that there is a polynomial h such that

$$\overline{r(z)}z^N = q(z)h(z),$$

and hence the following is true

$$r(z) = z^N\overline{q(z)}\ \overline{h(z)}, \quad z\in\mathbb{T}^d.$$

From the above equality we see that $(N,N,\dots,N)$ can be written as

$$(N,N,\dots,N) = (k_1,k_2,\dots,k_d) + (k_1',k_2',\dots,k_d')$$

such that $F(z) = z_1^{k_1}z_2^{k_2}\cdots z_d^{k_d}\overline{q(z)}$ and $G(z) = z_1^{k_1'}z_2^{k_2'}\cdots z_d^{k_d'}\overline{h(z)}$ become polynomials. By $Z(r)\cap\mathbb{D}^d = \emptyset$, we see

$$Z(F)\cap\mathbb{D}^d = \emptyset. \tag{2.14}$$

Let $I = pL$ be the Beurling form of the ideal I. Decompose $p = p'p''$ such that the zero set of each prime factor of p' meets $\mathbb{D}^d$ nontrivially, and $Z(p'')\cap\mathbb{D}^d = \emptyset$. Then by [DPSY] or [Guo2]

$$[I]\cap C[z_1,\dots,z_d]\subseteq[p']\cap C[z_1,\dots,z_d] = p'\ C[z_1,\dots,z_d].$$

From the above inclusion and (2.13), there exists a polynomial p''' such that $q = p'p'''$. Since

$$F(z) = z_1^{k_1}z_2^{k_2}\cdots z_d^{k_d}\overline{q(z)} = z_1^{k_1}z_2^{k_2}\cdots z_d^{k_d}\overline{p'(z)p'''(z)}$$

is a polynomial, this means that $(k_1,k_2,\dots,k_d)$ can be written as

$$(k_1,k_2,\dots,k_d) = (i_1,i_2,\dots,i_d) + (j_1,j_2,\dots,j_d)$$

such that $H(z) = z_1^{i_1}z_2^{i_2}\cdots z_d^{i_d}\overline{p'(z)}$ and $Q(z) = z_1^{j_1}z_2^{j_2}\cdots z_d^{j_d}\overline{p'''(z)}$ become polynomials. From (2.14), we have

$$Z(H)\cap\mathbb{D}^d = \emptyset.$$

Setting

$$R(z) = H(z)p''(z) = z_1^{i_1}z_2^{i_2}\cdots z_d^{i_d}\overline{p'(z)}p''(z),$$

then $R(z)$ is a polynomial satisfying $Z(R)\cap\mathbb{D}^d = \emptyset$. Note that on the distinguished boundary $\mathbb{T}^d$ of $\mathbb{D}^d$, we have $|p(z)| = |R(z)|$, and the theorem follows. $\square$

3. Defect operators of Hardy submodules over the unit ball

This section will study the defect operators and the defect functions for Hardy submodules over the unit ball $\mathbb{B}_d$ of $\mathbb{C}^d$. One will find that structures of submodules is highly distinct from the case of the polydisk. Furthermore, in Section 3.2, it is shown that a submodule generated by polynomials has finite defect if and only if such a submodule has finite codimension. The proof of the result is based on the theories of analytic variety and algebraic variety.

3.1. Preliminaries

We begin by recalling some notions. The Hardy space $H^2(\mathbb{B}_d)$ consists of analytic functions f in the unit ball $\mathbb{B}_d$ satisfying

$$||f||^2 = \sup_{0<r<1} \int_{\partial\mathbb{B}_d} |f(r\xi)|^2 \, d\sigma < \infty,$$

where $d\sigma$ is the natural rotation-invariant probability measure on $\partial\mathbb{B}_d$. As one knows, each function f in $H^2(\mathbb{B}_d)$ has a non-tangential limit $f(\xi)$ at almost every point $\xi \in \partial\mathbb{B}_d$ with respect to $d\sigma$. Furthermore,

$$||f||^2 = \int_{\partial\mathbb{B}_d} |f(\xi)|^2 \, d\sigma.$$

The Hardy space $H^2(\mathbb{B}_d)$ is a reproducing function space. The reproducing kernel and the normalized reproducing kernel are, respectively,

$$K_\lambda(z) = \frac{1}{(1 - \overline{\lambda_1} z_1 - \cdots - \overline{\lambda_d} z_d)^d}, \quad k_\lambda = \frac{K_\lambda}{||K_\lambda||}. \tag{3.1}$$

Since $H^2(\mathbb{B}_d)$ is invariant under multiplication by polynomials, the space $H^2(\mathbb{B}_d)$ naturally admits a $C[z_1, \dots, z_d]$-module structure. Given a submodule M of $H^2(\mathbb{B}_d)$, as done in Section 2, we introduce the defect function $D_M(\lambda)$ and the defect operator Δ_M for the submodule M. The defect function $D_M(\lambda)$ is defined by

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = ||P_M k_\lambda||^2, \tag{3.2}$$

where $K_\lambda^M = P_M K_\lambda$ be the reproducing kernel of the submodule M, P_M is the orthogonal projection from $H^2(\mathbb{B}_d)$ onto M. From the definition of the defect function D_M, it is easy to check that there exists a unique bounded linear operator, denoted by Δ_M, such that

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = \left[1 - \sum_{j=1}^{d} |\lambda_j|^2\right]^d ||P_M K_\lambda||^2 = \langle \Delta_M K_\lambda, K_\lambda \rangle, \tag{3.3}$$

where

$$\Delta_M = P_M + \sum_{k=1}^{d} (-1)^k \frac{d!}{k!(d-k)!} \sum_{i_1,i_2,\dots,i_k \in \{1,2,\dots,d\}} M_{z_{i_1} z_{i_2} \cdots z_{i_k}} P_M M^*_{z_{i_1} z_{i_2} \cdots z_{i_k}}.$$

We call Δ_M defect operator of the submodule M, and rankΔ_M defect index of M. The reader easily verifies that the defect operator Δ_M and the projection P_M is connected by

$$P_M K_\lambda = K_\lambda \, \Delta_M K_\lambda, \quad \lambda \in \mathbb{B}_d. \tag{3.4}$$

In the dimension $d = 2$, the defect operator of a submodule M is

$$\begin{aligned}\Delta_M &= P_M - 2(M_{z_1} P_M M^*_{z_1} + M_{z_2} P_M M^*_{z_2}) + 2M_{z_1 z_2} P_M M^*_{z_1 z_2} \\ &\quad + M_{z_1^2} P_M M^*_{z_1^2} + M_{z_2^2} P_M M^*_{z_2^2}.\end{aligned}$$

As the same in Sections 2, submodule is uniquely determined by its defect function, and hence is uniquely determined by its defect operator.

Proposition 3.1. *For submodules M and N, if $D_M(\lambda) = D_N(\lambda), \forall \lambda \in \mathbb{B}_d$, then $M = N$. Therefore, if $\Delta_M = \Delta_N$, then $M = N$.*

However, unlike Hardy-submodules on the polydisk, in general, the defect functions for submodules of $H^2(\mathbb{B}_d)$ need not be subharmonic. Let us check an example.

Example 6. Set

$$H_0^2 = \{f \in H^2(\mathbb{B}_2) : f(0,0) = 0\}.$$

Then an easy computing shows that the defect function $D(z)$ for the submodule H_0^2 is given by

$$D(z) = 1 - (1 - |z_1|^2 - |z_2|^2)^2.$$

Since

$$\frac{\partial^2 D(z)}{\partial z_1 \partial \overline{z_1}} + \frac{\partial^2 D(z)}{\partial z_2 \partial \overline{z_2}} = 4 - 6(|z_1|^2 + |z_2|^2),$$

applying [Kr, Corollary 2.1.12] shows that $D(z)$ is not subharmonic.

Moreover, similarly to Proposition 2.4, we have

Proposition 3.2. *Let M be a submodule of $H^2(\mathbb{B}_d)$. Then*

1. *for almost all $z \in \partial\mathbb{B}_d$ with respect to the measure $d\sigma$, $D_M(\lambda) \to 1$ as $\lambda \to z$ non-tangentially*;
2. *the defect operator $\Delta_M \geqslant 0$ if and only if there is an inner function η such that $M = \eta H^2(\mathbb{D}^d)$*;
3. *if Δ_M is in trace class*, *then* Trace $\Delta_M = 1$.

Proof. The proof of term (1) is completely similar to that of Theorem 2.1 in [GY]. We omit the details. The proofs of terms (2) and (3) are parallel to (2) and (3) of Proposition 2.4. □

Remark. Concerning the symmetric Fock space H_d^2 over the unit ball $\mathbb{B}_d$ considered by Arveson [Arv1,Arv2], an analogous result to Proposition 3.2(1) was presented in [GRS] whose proof is completely different from Proposition 3.2(1).

For two submodules M and N of $H^2(\mathbb{B}_d)$, we call that they are unitarily equivalent if there exists a unitary operator $U : M \to N$ such that $Upf = pUf$ for any polynomial p and $f \in M$. It is easy to see that if M and N is unitarily equivalent, then their defect operators are unitarily equivalent, and hence M and N have the same defect.

3.2. Finite defect problem

As the same in Section 2, we are naturally led to study submodules with finite defect indices. By identity (3.4), the range $\Delta_M H^2(\mathbb{B}_d)$ is a generating set of M, and it follows that if M has finite defect, then M must be finitely generated. Clearly, if M is of finite codimension in $H^2(\mathbb{B}_d)$, then M has finite defect, and therefore, if M is unitarily equivalent to a finite codimensional submodule, then M has finite defect. Using function theory from [Ru2,Ru3], it is easy to prove that if M is unitarily equivalent to a finite codimensional submodule, say, N, then there is an inner function η such that $M = \eta N$. Concerning inner functions on the unit ball, Rudin posed the existence problem of inner functions over the unit ball in the sixties: Do there exist nonconstant inner functions in $H^\infty(\mathbb{B}_d)$ [Ru2]? This problem was affirmatively solved in 1982 by Aleksandrov [Ru3].

Before continuing let us see an example.

Example 7. Suppose M is homogeneous, that is, M is generated by homogeneous polynomials. Then M has finite defect only if M is of finite codimension in $H^2(\mathbb{B}_d)$. In fact, if M is homogeneous, then it is not difficult to check that M has an orthonormal basis consisting of homogeneous polynomials, and hence P_M maps polynomials to polynomials. Also, note that for any $\phi \in H^\infty(\mathbb{B}_d)$, M_ϕ^* maps polynomials to polynomials. This shows that the defect operator Δ_M maps polynomials to polynomials. If Δ_M is of finite rank l, then there exist polynomials

$p_1, \dots, p_l$ and real numbers $\alpha_1, \dots, \alpha_l$ such that

$$\Delta_M = \sum_{k=1}^{l} \alpha_k p_k \otimes p_k,$$

and by (3.4) we have that

$$P_M = \sum_{k=1}^{l} \alpha_k M_{p_k} M_{p_k}^*.$$

This gives

$$P_{H^2(\mathbb{B}_d) \ominus M} = I - \sum_{k=1}^{l} \alpha_k M_{p_k} M_{p_k}^*.$$

Combining (3.2) and Proposition 3.2(1) yields the following

$$1 - \sum_{k=1}^{l} \alpha_k |p_k(\xi)|^2 = 0$$

for any $\xi \in \partial\mathbb{B}_d$. According to the Coburn's theorem [Cob], the projection $P_{H^2(\mathbb{B}_d)\ominus M}$ is compact and hence M is finite codimensional in $H^2(\mathbb{B}_d)$. Recall that the Coburn's theorem says the following: Let $C^*[M_{z_1}, \dots, M_{z_d}]$ be the C^*-algebra generated by the operators $M_{z_1}, \dots, M_{z_d}$ on $H^2(\mathbb{B}_d)$, then $C^*[M_{z_1}, \dots, M_{z_d}]$ contains all compact operators $\mathscr{K}$, and there exists a $*$-isomorphism

$$C^*[M_{z_1}, \dots, M_{z_d}]/\mathscr{K} \cong C(\partial\mathbb{B}_d),$$

where the correspondence is given by $M_{z_i} + \mathscr{K} \mapsto z_i$ for $i = 1, \dots, d$.

Notice that this example stands in rather stark contrast with the case of the Hardy module $H^2(\mathbb{D}^d)$. Indeed, we conjecture that the answer to the following problem is yes.

Conjecture. *Let M be a nonzero submodule of $H^2(\mathbb{B}_d)$. Then M has finite defect only if there are a finite codimensional submodule N, and an inner function η such that $M = \eta N$.*

If we assume this conjecture, then the conjecture implies the following: if I is an ideal of $C[z_1, \dots, z_d]$, and the submodule $[I]$ has finite defect, then $[I]$ is finite codimensional in $H^2(\mathbb{B}_d)$. To see this, by the conjecture, $[I]$ is unitarily equivalent to a finite codimensional submodule N. Since each finite codimensional submodule is generated by polynomials, from [Guo2, Theorem 5.2] one has $[I] = N$.

Actually, we will prove this result.

Theorem 3.3. *Let I an ideal of $C[z_1, \dots, z_d]$. If $[I]$ has finite defect, then $[I]$ is finite codimensional in $H^2(\mathbb{B}_d)$.*

For this theorem, we will need several propositions and lemmas.

Lemma 3.4. *Let $f = p/q$ be a rational function, where p and q are without common factors. If f is analytic on $\mathbb{B}_d$, then $Z(q) \cap \mathbb{B}_d = \emptyset$.*

Proof. Similarly to the proof of [Guo2, Lemma 3.2]. □

Lemma 3.5. *Let q be a polynomial, and $Z(q) \cap \mathbb{B}_d = \emptyset$. Then $[q] = H^2(\mathbb{B}_d)$.*

Proof. See the Remark following Proposition 2.9 in [Guo2]. □

Let f be analytic on $\mathbb{B}_d$. For each $\xi \in \partial\mathbb{B}_d$, the slice function f_ξ is defined by $f_\xi(z) = f(z\xi)$, $\forall z \in \mathbb{D}$. The same as the proof of Lemma 3.7 in [Guo2] we have

Lemma 3.6. *Let f be analytic on $\mathbb{B}_d$. If for almost all $\xi \in \partial\mathbb{B}_d$, the slice function $f_\xi(z) = f(z\xi)$ is a polynomial, then f is a polynomial.*

We will need a result which comes from [Ru1, Theorem 14.3.3].

Lemma 3.7. *Assume $d \geqslant 2$. Let Ω be a bounded domain in $\mathbb{C}^d$, and let $A(\Omega) = C(\overline{\Omega}) \cap Hol(\Omega)$ be the so-called Ω-algebra. If $f \in A(\Omega), g \in A(\Omega)$, and $|f(\lambda)| \leqslant |g(\lambda)|$ for each boundary point λ of Ω, then $|f(z)| \leqslant |g(z)|$ for every $z \in \Omega$.*

Proposition 3.8. *Let I be an ideal of $C[z_1, \dots, z_d]$, and $[I]$ have finite defect index l. Then there exist polynomials $p_1, \dots, p_s$, $q_1, \dots, q_t$ (here $s + t = l$) and r satisfying the greatest common divisor*

$$GCD(p_1, \dots, p_s, q_1, \dots, q_t) = 1, \tag{3.5}$$

and $Z(r) \cap \mathbb{B}_d = \emptyset$ such that

1. *the rational functions $p_1/r, \dots, p_s/r, q_1/r, \dots, q_t/r$ belong to $[I]$, and $\{p_1, \dots, p_s, q_1, \dots, q_t\}$ is a generating set of $[I]$;*
2. $\Delta_{[I]} = \sum_{i=1}^{s} \frac{p_i}{r} \otimes \frac{p_i}{r} - \sum_{j=1}^{t} \frac{q_j}{r} \otimes \frac{q_j}{r}$;
3. $\sum_{i=1}^{s} |p_i(\xi)|^2 - \sum_{j=1}^{t} |q_j(\xi)|^2 = |r(\xi)|^2, \quad \xi \in \partial\mathbb{B}_d$.

Proof. Combining Lemmas 3.4, 3.5 and the proof of Lemma 2.11 shows that Proposition 3.8(1), (2), (3) are true, and

$$GCD(p_1, \dots, p_s, q_1, \dots, q_t, r) = 1, \quad Z(r) \cap \mathbb{B}_d = \emptyset. \tag{3.6}$$

Below, we will combine (3.6) and Proposition 3.8(3) to prove (3.5). Since Proposition 3.8(3) will be used several times, it is singled out here,

$$\sum_{i=1}^{s} |p_i(\xi)|^2 - \sum_{j=1}^{t} |q_j(\xi)|^2 = |r(\xi)|^2, \quad \xi \in \partial\mathbb{B}_d. \tag{3.7}$$

To prove (3.5), let

$$p = GCD(p_1, \dots, p_s, q_1, \dots, q_t)$$

and

$$p_1 = pp'_1, \dots, p_s = pp'_s, \quad q_1 = pq'_1, \dots, q_t = pq'_t.$$

By (3.7), there exists a positive constant γ such that

$$|r(\xi)|^2 \leqslant \gamma\, |p(\xi)|^2, \quad \forall \xi \in \partial\mathbb{B}_d.$$

Combining this fact with $Z(r) \cap \mathbb{B}_d = \emptyset$, and applying Lemma 3.7 we have

$$Z(p) \cap \mathbb{B}_d = \emptyset.$$

For each $\xi \in \partial\mathbb{B}_d$, by (3.7)

$$|p_\xi(e^{i\theta})|^2 \left[\sum_{i=1}^{s} |p'_{i\xi}(e^{i\theta})|^2 - \sum_{j=1}^{t} |q'_{j\xi}(e^{i\theta})|^2 \right] = |r_\xi(e^{i\theta})|^2. \tag{3.8}$$

Set $N = \max\{\deg p_1, \dots, \deg p_s, \deg q_1 \dots, \deg q_t, \deg r\}$, where $\deg h$ denotes the homogeneous degree of a polynomial h. Then we have

$$\begin{aligned} &p_\xi(e^{i\theta})\overline{p_\xi(e^{i\theta})}e^{iN\theta} \left[\sum_{i=1}^{s} p'_{i\xi}(e^{i\theta})\overline{p'_{i\xi}(e^{i\theta})}e^{iN\theta} - \sum_{j=1}^{t} q'_{j\xi}(e^{i\theta})\overline{q'_{j\xi}(e^{i\theta})}e^{iN\theta} \right] \\ &\quad = r_\xi(e^{i\theta})\overline{r_\xi(e^{i\theta})}e^{2iN\theta}. \end{aligned} \tag{3.9}$$

Since polynomials $p_\xi(z), r_\xi(z)$ in the variable z have no zero point on the unit disk $\mathbb{D}$, they are outer functions in the Hardy space $H^2(\mathbb{D})$. Note that $\overline{p_\xi(e^{i\theta})}e^{iN\theta}, \overline{p'_{i\xi}(e^{i\theta})}e^{iN\theta}$, $\overline{q'_{j\xi}(e^{i\theta})}e^{iN\theta}$, $\overline{r_\xi(e^{i\theta})}e^{2iN\theta}$ are analytic polynomials in $e^{i\theta}$ for each $\xi \in \partial\mathbb{B}_d$. Therefore, for each $\xi \in \partial\mathbb{B}_d$, the outer factor of the right side in (3.9) equals $r_\xi^2(e^{i\theta})$, and the outer factor of the left side equals $p_\xi^2(e^{i\theta}) \times$ some polynomial. By the uniqueness of inner–outer decomposition of functions in $H^2(\mathbb{D})$ [Gar], the slice function $r_\xi^2(e^{i\theta})/p_\xi^2(e^{i\theta})$ is a polynomial for each $\xi \in \partial\mathbb{B}_d$. By Lemma 3.6, the function r^2/p^2 is a polynomial, and hence

$$p^2 = GCD(p^2, r^2).$$

Since $GCD(p,r)=1$, this implies that p^2 is a constant, and hence p is a constant. This shows that (3.5) is true. The proposition follows. $\square$

Proposition 3.9. *Let I be an ideal of the polynomial ring $C[z_1,\ldots,z_d]$. If $[I]$ has finite defect, then the ideal I has only finitely many zeros in $\mathbb{B}_d$, that is, $Z(I)\cap\mathbb{B}_d$ is a finite set.*

For this proposition we need some preliminaries involving the theories of analytic variety and algebraic variety.

An analytic variety in $\mathbb{C}^d$ is a closed set $V\subset\mathbb{C}^d$ with the following property: for each $z\in V$ correspond analytic functions $f_1,\ldots,f_r$, defined on some neighborhood O of z such that

$$V\cap O=Z(f_1)\cap\cdots\cap Z(f_r).$$

If, in addition, these r functions can be so chosen that their Jacobian matrix has rank r at z, then z is called a regular point of V at which V has complex dimension $d-r$. In symbols, $\dim_z V=d-r$. Notice that the complex dimension of V at a regular point is independent of choices of functions. If $\dim_z V=k$ at every regular point of V, then V is said to have pure dimension k. An analytic variety of pure dimension 1 is called an analytic curve. More generally, one defines the complex dimension of V as

$$\dim V=\sup\dim_z V,$$

where the supremum is taken over all regular points of V (cf. [Her]).

An algebraic variety V in $\mathbb{C}^d$ is the intersection of the zero-sets of finitely many polynomials. Since each ideal of $C[z_1,\ldots,z_d]$ is generated by finitely many polynomials, algebraic varieties are zero sets of ideals of $C[z_1,\ldots,z_d]$. A algebraic variety V is called irreducible if $V=V_1\cup V_2$ implies $V=V_1$ or $V=V_2$, where V_1 and V_2 are algebraic varieties. A basic result from algebraic geometry [Ken] states that each algebraic variety can be uniquely decomposed as a finite irredundant union of irreducible varieties. Similarly to analytic variety, one can define regular points, dimension for an algebraic variety. We refer the reader to [Ken, Chapter IV] for details. An algebraic variety of pure dimension 1 is called an algebraic curve.

Lemma 3.10. *Let S be an irreducible algebraic curve, and let S^* be the set of all regular points of S. Then S^* is connected.*

This lemma may be known by many algebraic geometricians, but we cannot locate a reference recording this result.

Proof. From [Her, p. 108, Corollary 1], for each irreducible analytic variety V in $\mathbb{C}^d$, V^*, the set of all regular points of V, is connected. Therefore, it is sufficient to show that S, as an analytic curve in $\mathbb{C}^d$, is irreducible. To get a contradiction we assume that S is reducible. Then by [Her, p. 111, Theorem 8], S can be decomposed as

$$S = S_1 \cup S_2 \cup S_3 \cup \dots,$$

where each S_i is irreducible analytic variety in $\mathbb{C}^d$, and $S_i \neq S$. Since S has pure dimension 1, by [Her, p.109, Corollary 3], every S_i has pure dimension $\leqslant 1$. Clearly, there exists at least one S_j such that S_j has pure dimension 1. Applying [Ru4, Theorem 2] to this S_j shows that S_j is an irreducible algebraic curve. Since S be an irreducible algebraic curve, this implies $S_j = S$. This contradiction shows that S is an irreducible analytic curve, completing the proof. □

Lemma 3.11. *Let S be an irreducible algebraic curve, and $S \cap \mathbb{B}_d \neq \emptyset$. Then $S \cap \partial\mathbb{B}_d$ is an infinite set.*

Proof. Let $S^{**} = S - S^*$ be the set of all singular points of S. Then, by [Ken, p. 190, Corollary] S^{**} is a proper subvariety of S, and hence $S^{**}$is a finite set. Since S is a connected and unbounded set [Ken, p. 191, Theorem 5.1], Lemma 3.10 implies that

$$S^* \cap \mathbb{B}_d \neq \emptyset, \quad S^* \cap \partial\mathbb{B}_d \neq \emptyset, \quad S^* \cap (\mathbb{C}^d - \overline{\mathbb{B}_d}) \neq \emptyset.$$

By Lemma 3.10, there is at least one $\zeta \in S^* \cap \partial\mathbb{B}_d$ such that to every $\varepsilon > 0$,

$$S^* \cap \{z \in \mathbb{C}^d : |z - \zeta| < \varepsilon\} \cap \mathbb{B}_d \neq \emptyset. \tag{3.10}$$

Since S has complex dimension 1 at the regular point ζ, then by Implicit complex-analytic mapping theorem [Ken, p. 49, Theorem 3.5], one can choose a open connected neighborhood O of the origin of the complex plane and a open neighborhood $O(\zeta)$ of ζ, and analytic functions $\phi_1, \dots, \phi_{d-1}$ defined on O satisfying $\phi_1(0) = \dots = \phi_{d-1}(0) = 0$ such that

$$S \cap O(\zeta) = \zeta + \{(z, \phi_1(z), \dots, \phi_{d-1}(z)) : z \in O\}.$$

Considering the subharmonic function

$$H(z) = |z + \zeta_1|^2 + |\phi_1(z) + \zeta_2|^2 + \dots + |\phi_{d-1}(z) + \zeta_d|^2, \quad z \in O,$$

then $H(0) = 1$. From (3.10) we see that there are infinitely many $w \in O$ satisfying $H(w) < 1$, and hence by the definition of subharmonic function [Gar], there exist infinitely many $z \in O$ such that $H(z) > 1$. Set

$$X = \{z \in O : H(z) < 1\}, \quad Y = \{z \in O : H(z) = 1\}, \quad Z = \{z \in O : H(z) > 1\}.$$

Note that O is connected, and $O - Y = X \dot{\cup} Z$ is not connected. This implies that Y is an infinite set, otherwise, $O - Y$ is connected. Since each point in $S \cap O(\zeta)$ is regular, we have

$$\zeta + \{(z, \phi_1(z), \dots, \phi_{d-1}(z)) \colon z \in Y\} \subset S^* \cap \partial \mathbb{B}_d.$$

By the fact that Y is an infinite set, the set $S \cap \partial \mathbb{B}_d$ is infinite. □

We now give the proof of Proposition 3.9.

The proof of Proposition 3.9. Let $p_1, \dots, p_s, q_1, \dots, q_t$ and r be the polynomials appeared in Proposition 3.8. Then by Proposition 3.8(3) we have

$$\sum_{i=1}^{s} |p_i(\xi)|^2 - \sum_{j=1}^{t} |q_j(\xi)|^2 = |r(\xi)|^2, \quad \xi \in \partial \mathbb{B}_d. \tag{3.11}$$

Set $V = Z(I) \cap \mathbb{B}_d$, and for each $\alpha \neq 0$, $V_\alpha = \{z \in V : r(z) = \alpha\}$.

Claim 1. *For each* $\alpha \neq 0$, $V_\alpha = \{z \in V : r(z) = \alpha\}$ *is a finite set. Indeed, by Proposition* 3.8 *the submodule* $[I]$ *is generated by* $\{p_1, \dots, p_s, q_1 \dots, q_t\}$, *and hence*

$$V = Z(p_1) \cap \cdots \cap Z(p_l) \cap Z(q_1) \cdots Z(q_t) \cap \mathbb{B}_d.$$

This means

$$V_\alpha = \{z \in \mathbb{B}_d : r(z) - \alpha = 0, p_i(z) = 0, q_j(z) = 0, i = 1, \dots, s, j = 1, \dots, t\}.$$

By (3.11) it is easily seen that V_α is a compact analytic variety in $\mathbb{B}_d$. Applying [Her, p. 92, Corollary 1] shows that V_α is a finite set. The Claim 1 follows.

To complete the proof, we assume that $V = Z(I) \cap \mathbb{B}_d$ is an infinite set. Then there exists an irreducible component S of $Z(I)$ such that $S \cap \mathbb{B}_d$, denoted by S_0, is an infinite set.

Claim 2. *The variety* S *has pure dimension* 1, *that is,* S *is an irreducible algebraic curve*.

In fact, since S is irreducible, by [Ken, p. 172, Theorem 2.9] each regular point of S has the same dimension, say, k. Note that the set S^* of all regular points of S is dense in S (cf. [Ken]). Therefore, $S^* \cap \mathbb{B}_d$ is dense in $S_0 = S \cap \mathbb{B}_d$. Taking a regular point ζ of S_0 then by Implicit complex-analytic mapping theorem [Ken, p. 49, Theorem 3.5], there exist a open neighborhood $O \subset \mathbb{C}^k$ about (0), and a open neighborhood $O(\zeta)$ of ζ, and analytic functions $\phi_1, \dots, \phi_{d-k}$ defined on O satisfying $\phi_1(0) = \cdots = \phi_{d-k}(0) = 0$ such that

$$S_0 \cap O(\zeta) = \zeta + \{(z_1, \dots, z_k, \phi_1(z), \dots, \phi_{d-k}(z)) : z \in O\}.$$

Consider analytic function $Q(z_1, \dots, z_k)$ defined on O by

$$Q(z_1, \dots, z_k) = r(\zeta + Z) - r(\zeta),$$

where $Z = (z_1, \dots, z_k, \phi_1(z_1, \dots, z_k), \dots, \phi_{d-k}(z_1, \dots, z_k))$. Since $r(\zeta) \neq 0$ and $Q(0) = 0$, Claim 1 implies that Q has only finitely many zero points in O. Since an analytic function in several variables has no isolated zero point (cf. [Kr]), this implies $k = 1$, and hence the variety S has pure dimension 1, that is, S is an irreducible algebraic curve.

Before going on we require the following notion. For a prime ideal P of $C[z_1, \dots, z_d]$, the height of P, denoted by height(P), is defined as the maximal length l of any properly increasing chain of prime ideals

$$0 = P_0 \subset P_1 \cdots \subset P_l = P.$$

Since the polynomial ring $C[z_1, \dots, z_d]$ is Noetherian, every prime ideal has finite height and the height of an arbitrary ideal is defined as the minimum of the heights of its associated prime ideals. For an ideal J, one has

$$\dim_{\mathbb{C}} Z(J) = d - l, \tag{3.12}$$

where $l = \text{height}(J)$ is the height of J, and $\dim_{\mathbb{C}} Z(J)$ the complex dimension of the zero variety of J (cf. [Ken, p. 196]).

Put $I_1 = \{p \in C[z_1, \dots, z_d] : Z(p) \supset S\}$, and $I_2 = I_1 + rC[z_1, \dots, z_d]$. Then I_1 is prime, $Z(I_1) = S$ and by (3.12) $\text{height}(I_1) = d - 1$.

Claim 3. *height*$(I_2) = d$.

In fact, since $I_2 \supset I_1$, we have $\text{height}(I_2) \geqslant d - 1$. Now assume that $\text{height}(I_2) = d - 1$. Taking a finite irredundant primary decomposition of I_2

$$I_2 = J_1 \cap \cdots \cap J_n,$$

then there is a primary ideal J_s such that its radical ideal $\sqrt{J_s}$ has the height $d - 1$. Since $\sqrt{J_s}$ is prime, and $\sqrt{J_s} \supset I_1$, we get $I_1 = \sqrt{J_s}$. Since $r \in \sqrt{J_s}$, this forces $r \in I_1$, and hence $Z(r) \supset S$. This contradicts the fact that r has no zero point in $\mathbb{B}_d$. We conclude therefore that $\text{height}(I_2) = d$.

From Claim 3 and [Ken] we see that $Z(I_2) = S \cap Z(r)$ is a finite set. By (3.11) one has

$$S \cap \partial\mathbb{B}_d \subseteq Z(r) \cap S \cap \partial\mathbb{B}_d.$$

This shows that $S \cap \partial\mathbb{B}_d$ is a finite set. This contradicts Lemma 3.11, and it follows that I has only finitely many zero points in $\mathbb{B}_d$. The proof of Proposition 3.9 is completed. $\square$

To prove Theorem 3.3, we also need the following notions. Let $\mathscr{H}(\overline{\mathbb{B}_d})$ denote the space of analytic functions defined on neighborhoods of $\overline{\mathbb{B}_d}$, and $C^\infty(\overline{\mathbb{B}_d})$ the space of germs of smooth functions on $\overline{\mathbb{B}_d}$. Defining seminorms $\|\cdot\|_n$ on $C^\infty(\overline{\mathbb{B}_d})$, $n = 0, 1, 2, \dots$, by setting

$$\|f\|_n = \max\{|\overline{D}^\beta D^\alpha f(z)| : z \in \overline{\mathbb{B}_d}, |\alpha| + |\beta| \leqslant n\},$$

where $\alpha = (\alpha_1, \dots, \alpha_d)$, $\beta = (\beta_1, \dots, \beta_d)$ are ordered d-tuples of nonnegative integers, and $|\alpha| = \alpha_1 + \dots + \alpha_d$, $|\beta| = \beta_1 + \dots + \beta_d$, and

$$D^\alpha = \left(\frac{\partial}{\partial z_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial z_d}\right)^{\alpha_d}, \quad \overline{D}^\beta = \left(\frac{\partial}{\partial \overline{z}_1}\right)^{\beta_1} \cdots \left(\frac{\partial}{\partial \overline{z}_d}\right)^{\beta_d}.$$

If $|\alpha| = |\beta| = 0$, $D^\alpha f = \overline{D}^\beta f = f$.

This family of seminorms makes $C^\infty(\overline{\mathbb{B}_d})$ into a locally convex topological linear space. From the basic theory of functional analysis [Ru5], for each continuous linear functional F on $C^\infty(\overline{\mathbb{B}_d})$ there exist a positive constant C and a nonnegative integer m such that

$$F(f) \leqslant C\|f\|_m, \quad f \in C^\infty(\overline{\mathbb{B}_d}). \tag{3.13}$$

We now will give the proof of Theorem 3.3.

The proof of Theorem 3.3. By Proposition 3.9, $Z(I) \cap \mathbb{B}_d$ is a finite set, say, $Z(I) \cap \mathbb{B}_d = \{\lambda_1, \dots, \lambda_l\}$. From the basic theory of algebraic variety [Ken] the ideal I can be decomposed as $I = I_1 \cap I_2$ such that $Z(I_1) = \{\lambda_1, \dots, \lambda_l\}$, and $Z(I_2) \cap \mathbb{B}_d = \emptyset$. It follows that I_1 is finite codimensional in $C[z_1, \dots, z_d]$. Since $I_1 I_2 \subset I \subset I_2$ and

$$I_2/I = I_2/I_1 \cap I_2 \cong (I_1 + I_2)/I_1 \subset C[z_1, \dots, z_d]/I_1,$$

we see $\dim I_2/I < \infty$, and hence there exists a finite dimensional linear subspace $\mathscr{R}$ of I_2 such that

$$I_2 = I \dot{+} \mathscr{R}.$$

It is easy to see that $[I_2] = [I] + \mathscr{R}$. Since $\mathscr{R}$ is finite dimensional, the defect operator $\varDelta_{[I]}$ has finite rank if and only if $\varDelta_{[I_2]}$ has finite rank. Hence, to complete the proof, it is enough to show that the following is true: if $Z(I) \cap \mathbb{B}_d = \emptyset$ and $\varDelta_{[I]}$ has finite rank, then $[I] = H^2(\mathbb{B}_d)$. Below, we will use an idea from [PS]. By way of contradiction we suppose that I is not dense in $H^2(\mathbb{B}_d)$. Let $h \in H^2(\mathbb{B}_d)$, and $h \perp [I]$. Then h induces the linear functional $\langle \cdot, h \rangle_{H^2(\mathbb{B}_d)}$ on $\mathscr{H}(\overline{\mathbb{B}_d})$, which produces a continuous linear functional $F : \mathscr{H}(\overline{\mathbb{B}_d})/\overline{I} \to \mathbb{C}$, where $\overline{I}$ denotes the closure of I in $\mathscr{H}(\overline{\mathbb{B}_d})$. By Malgrange's (flatness and separation) Theorem on ideals of smooth functions [Mal], the extended ideal $\overline{I} \cdot C^\infty(\overline{\mathbb{B}_d})$ is closed in $C^\infty(\overline{\mathbb{B}_d})$, and $\overline{I} \cdot C^\infty(\overline{\mathbb{B}_d}) \cap \mathscr{H}(\overline{\mathbb{B}_d}) = \overline{I}$. This enables us to extend F (using Hahn–Banach's Theorem) to a continuous linear functional

$$\tilde{F}: C^{\infty}(\overline{\mathbb{B}_d})/\overline{I}\cdot C^{\infty}(\overline{\mathbb{B}_d}) \to \mathbb{C}.$$

Notice that $\tilde{F}$ is supported by $Z(I)\cap\overline{\mathbb{B}_d}$, and by (3.11)

$$Z(I)\cap\overline{\mathbb{B}_d} = Z(I)\cap\partial\mathbb{B}_d \subset Z(r)\cap\partial\mathbb{B}_d.$$

Therefore, by (3.13) there exist a nonnegative integer m and a constant C such that

$$|\tilde{F}(\tilde{f})| \leqslant C\|f\|_{m,Z(r)\cap\partial\mathbb{B}_d}, \quad f\in C^{\infty}(\overline{\mathbb{B}_d}),$$

where $\tilde{f} = f + \overline{I}\cdot C^{\infty}(\overline{\mathbb{B}_d})$, and

$$\|f\|_{m,Z(r)\cap\partial\mathbb{B}_d} = \max\{|\overline{D}^{\beta}D^{\alpha}f(z)| : z\in Z(r)\cap\partial\mathbb{B}_d, |\alpha|+|\beta|\leqslant m\}.$$

This insures that for any polynomial p,

$$\tilde{F}(r^{m+1}p) = 0.$$

Since

$$\langle r^{m+1}p, h\rangle_{H^2(\mathbb{B}_d)} = F(r^{m+1}p + \overline{I}) = \tilde{F}(r^{m+1}p) = 0$$

for any polynomial p, this implies $h\perp[r^{m+1}]$. By Lemma 3.5,

$$[r^{m+1}] = H^2(\mathbb{B}_d),$$

and hence $h=0$, reaching a contradiction. The proof of Theorem 3.3 is completed. □

4. Defect operators of Bergman submodules over the unit ball

In this section we will briefly mention the defect operators and the defect functions for submodules of the Bergman module $L_a^2(\mathbb{B}_d)$ over the unit ball $\mathbb{B}_d$. The Bergman space $L_a^2(\mathbb{B}_d)$ is the closed subspace of $L^2(\mathbb{B}_d, dv)$ consisting of analytic functions, where dv is the normalized volume measure with $v(\mathbb{B}_d)=1$. It is well known that the reproducing kernel and the normalized reproduce kernel of $L_a^2(\mathbb{B}_d)$ are, respectively

$$K_{\lambda}(z) = \frac{1}{(1-\overline{\lambda_1}z_1 - \overline{\lambda_2}z_2 - \cdots - \overline{\lambda_d}z_d)^{d+1}}, \quad k_{\lambda} = \frac{K_{\lambda}(z)}{\|K_{\lambda}\|}.$$

Since $L_a^2(\mathbb{B}_d)$ admits a natural $C[z_1,\dots,z_d]$-module structure coming from multiplication by polynomials we will call $L_a^2(\mathbb{B}_d)$ as Bergman module over the unit ball. As done in Sections 2 and 3, given a submodule we define the defect function $D_M(\lambda)$

as follows:

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = ||P_M k_\lambda||^2. \tag{4.1}$$

From the definition of the defect function D_M, it is easy to check that there is a unique bounded linear operator, denoted by Δ_M, such that

$$D_M(\lambda) = \frac{||K_\lambda^M||^2}{||K_\lambda||^2} = \left[1 - \sum_{j=1}^{d} |\lambda_j|^2\right]^{d+1} ||P_M K_\lambda||^2 = \langle \Delta_M K_\lambda, K_\lambda \rangle, \tag{4.2}$$

where

$$\Delta_M = P_M + \sum_{k=1}^{d+1} (-1)^k \frac{(d+1)!}{k!(d+1-k)!} \sum_{i_1, i_2, \dots, i_k \in \{1,2,\dots,d\}} M_{z_{i_1} z_{i_2} \cdots z_{i_k}} P_M M^*_{z_{i_1} z_{i_2} \cdots z_{i_k}}.$$

We then define the defect operator of the submodule M as Δ_M. In particular, in the dimension $d = 1$, for a submodule M of the Bergman module $L_a^2(\mathbb{D})$, its defect operator is

$$\Delta_M = P_M - 2M_z P_M M_z^* + M_{z^2} P_M M_{z^2}^*.$$

In the case of the Bergman module $L_a^2(\mathbb{D})$, Yang and Zhu have made some progress for study of defect operators of submodules [YZ]. In this section we will concentrate attention on dimension $d \geqslant 2$.

As in the previous sections, a submodule is uniquely determined by its defect function, and hence is uniquely determined by its defect operator. Also, the defect functions for submodules of $L_a^2(\mathbb{B}_d)$ need not be subharmonic, in general. An example is that the defect function of the submodule $\{f \in L_a^2(\mathbb{D}) : f(0) = 0\}$ is not subharmonic.

Moreover, similarly to Proposition 3.2 we have

Proposition 4.1. *Let M be a submodule of $L_a^2(\mathbb{B}_d)$. We have the following*:

1. *the defect operator $\Delta_M \geqslant 0$ only if $M = L_a^2(\mathbb{B}_d)$*;
2. *if M contains a function from $H^2(\mathbb{B}_d)$, then for almost all $z \in \partial\mathbb{B}_d$ with respect to the measure $d\sigma$, $D_M(\lambda) \to 1$ as $\lambda \to z$ non-tangentially.*

Proof. (1). First note that

$$\mathbb{S} = M \ominus \overline{(z_1 M + z_2 M + \cdots z_d M)} \neq \{0\}$$

and it is easy to check that every function from $\mathbb{S}$ is eigenvector of Δ_M with the corresponding eigenvalue 1. Taking a function g from $\mathbb{S}$ with $||g|| = 1$, then $\Delta_M -$

$g \otimes g \geqslant 0$. By (4.1) and (4.2), we have

$$\langle \Delta_M K_\lambda, K_\lambda \rangle - \langle g \otimes g K_\lambda, K_\lambda \rangle = ||P_M k_\lambda||^2 - |g(\lambda)|^2 \geqslant 0,$$

and it follows that

$$\int_{\mathbb{B}_d} ||P_M k_\lambda||^2 \, dv(\lambda) - 1 \geqslant 0.$$

From the above inequality it is deduced that

$$||P_M k_\lambda|| = 1, \quad \lambda \in \mathbb{B}_d,$$

and hence $M = L_a^2(\mathbb{B}_d)$.

(2) Take a nonzero $f \in M \cap H^2(\mathbb{B}_d)$, and set $[f] = \overline{fC[z_1, \dots, z_d]}$, which is contained in M. If we can show $D_{[f]}(z) = 1$ a.e. on $\partial\mathbb{B}_d$, then the conclusion will follow from the fact

$$1 \geqslant ||P_M k_\lambda||^2 \geqslant ||P_{[f]} k_\lambda||^2 = D_{[f]}(\lambda).$$

The same argument as in [GY] enables us to get the following:

$$1 \geqslant D_{[f]}(\lambda) \geqslant \frac{|f(\lambda)|^2}{||fk_\lambda||^2}.$$

Notice that

$$\begin{aligned} ||fk_\lambda||^2 &= \int_{\mathbb{B}_d} |f(z)|^2 |k_\lambda(z)|^2 \, dv = \int_{\mathbb{B}_d} |f \circ \phi_\lambda(z)|^2 \, dv \\ &\leqslant \int_{\partial\mathbb{B}_d} |f \circ \phi_\lambda(\xi)|^2 \, d\sigma = \int_{\partial\mathbb{B}_d} P(\lambda, \xi) |f(\xi)|^2 \, d\sigma, \end{aligned}$$

where ϕ_λ is the canonical automorphism of the ball $\mathbb{B}_d$ that maps the origin to λ (cf. [Ru2, Chapter 2]), and $P(\lambda, \xi)$ is the invariant Poisson kernel at λ for the unit ball (cf. [Ru2, Chapter 3]). We thus obtain that

$$1 \geqslant D_{[f]}(\lambda) \geqslant \frac{|f(\lambda)|^2}{||fk_\lambda||^2} \geqslant \frac{|f(\lambda)|^2}{\int_{\partial\mathbb{B}_d} P(\lambda, \xi) |f(\xi)|^2 d\sigma}.$$

Since for almost all $z \in \partial\mathbb{B}_d$, $\int_{\partial\mathbb{B}_d} P(\lambda, \xi) |f(\xi)|^2 \, d\sigma$ converges to $|f(z)|^2$ as $\lambda \to z$ non-tangentially, this insures that $D_{[f]}(\lambda) \to 1$ as $\lambda \to z \in \partial\mathbb{B}_d$ non-tangentially. The required conclusion follows. □

A natural problem arises here. Does Proposition 4.1(2) hold for any nonzero Bergman submodule?

We now turn to studying submodules with finite defect indices. In the dimension $d = 1$, this problem was considered by Yang and Zhu [YZ]. In particular, they found that the defect operator for the submodule generated by the singular inner function $S(z) = \exp(-(1+z)/(1-z))$ has rank 2. Therefore, in the case of the dimension $d = 1$, one cannot even give a reasonable guess for finite defect problem. In dimension $d \geqslant 2$, we conjecture that M has finite defect only if M has finite codimension in $L_a^2(\mathbb{B}_d)$. Indeed, following the proof of Theorem 3.3, we can prove the following:

Proposition 4.2. *Let I an ideal of $C[z_1, \ldots, z_d]$. If the submodule $[I]$ has finite defect, then $[I]$ is finite codimensional in $L_a^2(\mathbb{B}_d)$.*

Remark 1. The results in this section can be generalized to the weighted Bergman space $L_a^2(\mathbb{B}_d, dv_\alpha), \alpha > -1$, where $dv_\alpha(z) = c_\alpha(1 - |z|^2)^\alpha dv(z)$, and c_α is a positive constant such that $v_\alpha(\mathbb{B}_d) = 1$. For the weighted Bergman module $L_a^2(\mathbb{B}_d, dv_\alpha)$, its reproducing is given by

$$K_\lambda(z) = \frac{1}{(1 - \langle z, \lambda \rangle)^{\alpha+d+1}}.$$

Similarly to the case of the Bergman module $L_a^2(\mathbb{B}_d)$, one can define defect operators (functions) for submodules of $L_a^2(\mathbb{B}_d, dv_\alpha)$.

Remark 2. In [Arv1,Arv2], Arveson developed the theory of d-contractions. We refer the reader to references mentioned above for a far-reaching operator-algebraic development of this theory. This theory especially concerns a typical function space on $\mathbb{B}_d$, the so-called symmetric Fock space H_d^2 on $\mathbb{B}_d$, which is induced by the reproducing kernel

$$K_\lambda(z) = \frac{1}{1 - \langle z, \lambda \rangle}.$$

Noticing that H_d^2 admits a natural $C[z_1, \ldots, z_d]$-module structure coming from multiplication by polynomial, then H_d^2 is a typical Hilbert module over the unit ball. Concerning submodules of H_d^2, a natural problem was asked by Arveson in [Arv2]: in dimension $d \geqslant 2$, must the defect index of each nonzero submodule of H_d^2, which has infinite codimension in H_d^2, be infinite? The paper [Guo3] shows that the answer to this problem is yes. However, the techniques used in [Guo3] are completely different from the present paper, because, as one has seen, the theory of defect operators relies heavily on expressions of reproducing kernels.

Acknowledgments

The author is deeply indebted to the referee for helpful suggestions which make this paper more readable. The author thanks Dr. X. Fang and Professors K. Izuchi and J. Yang for their helps. The author is partially supported by NNSFC, Shuguang project in Shanghai, the Teaching and Research Award Program for Outstanding Young Teachers in Higher Education Institutions of MOE, China.

References

[ACD] O. Agrawal, D. Clark, R. Douglas, Invariant subspaces in the polydisk, Pacific J. Math. 121 (1986) 1–11.
[AEM] C. Ambrozie, M. Englis, V. Muller, Operator tuples and analytic models over general domains in $\mathbb{C}^n$, J. Oper. Theory 47 (2002) 287–302.
[Arv1] W. Arveson, Subalgebras of C^*-algebras III: multivariable operator theory, Acta Math. 181 (1998) 159–228.
[Arv2] W. Arveson, The curvature invariant of a Hilbert module over $C[z_1, \ldots, z_n]$, J. Reine Angew. Mat. 522 (2000) 173–236.
[Beu] A. Beurling, On two problems concerning linear transformations in Hilbert space, Acta Math. 81 (1949) 239–255.
[CG] X. Chen, K. Guo, Analytic Hilbert Modules, in: Chapman & Hall/CRC Research Notes in Mathematics, Vol. 433, Chapman & Hall, London, 2003.
[Cob] L. Coburn, Singular integral operators and Toeplitz operators on odd spheres, Indiana Univ. Math. J. 23 (1973) 433–439.
[CV] R. Curto, F.-H. Vasilescu, Standard operator models in the polydisk, Indiana Univ. Math. J. 42 (1993) 791–810.
[DP] R.G. Douglas, V.I. Paulsen, Hilbert modules over function algebras, in: Pitman Research Notes in Mathematics, Vol. 217, Longman, Harlow, 1989.
[DPSY] R. Douglas, V. Paulsen, C. Sah, K. Yan, Algebraic reduction and rigidity for Hilbert modules, Amer. J. Math. 117 (1995) 75–92.
[DY] R. Douglas, R. Yang, Operator theory in the Hardy space over the bidisk, Integral Equations Oper. Theory 38 (2000) 207–227.
[Eng] M. Englis, Density of algebras generated by Toeplitz operators on Bergman spaces, Ark. Mat. 30 (1992) 227–240.
[Fa] X. Fang, Hilbert polynomials and Arveson's curvature invariant, J. Funct. Anal. 198 (2003) 445–464.
[Gar] J. Garnett, Bounded Analytic Functions, Academic Press, New York, 1981.
[Ge] R. Gelca, Rings with topologies induced by spaces of functions, Houston J. Math. 21 (1995) 395–405.
[Guo1] K. Guo, Characteristic spaces and rigidity for analytic Hilbert modules, J. Funct. Anal. 163 (1999) 133–151.
[Guo2] K. Guo, Equivalence of Hardy submodules generated by polynomials, J. Funct. Anal. 178 (2000) 343–371.
[Guo3] K. Guo, Defect operators for submodules of H_d^2, preprint.
[GY] K. Guo, R. Yang, The core function of submodules over the bidisk, Indiana Univ. Math. J., to appear.
[GRS] D. Greene, S. Richter, C. Sundberg, The structure of inner multiplications on spaces with complete Nevanlinna Pick kernels, J. Funct. Anal. 194 (2002) 311–331.
[Her] M. Hervé, Several Complex Variables: Local Theory, 2nd Edition, Oxford University Press, Oxford, 1987.

[Ken] K. Kendig, Elementary Algebraic Geometry, Springer, New York, 1977.

[Kr] S. Krantz, Function Theory of Several Complex Variables, Wiley, New York, 1982.

[Mal] B. Malgrange, Ideals of Differentiable Functions, Oxford University Press, Oxford, 1966.

[PS] M. Putinar, N. Salinas, Analytic transversality and nullstellensatz in Bergman space, Contemp. Math. 137 (1992) 367–381.

[Ru1] W. Rudin, Function Theory in Polydisks, Benjamin, New York, 1969.

[Ru2] W. Rudin, Function Theory in the Unit Ball of $\mathbb{C}^n$, Springer, New York, 1980.

[Ru3] W. Rudin, New construction of functions holomorphic in the unit ball of $\mathbb{C}^n$, Conference Board of the Mathematical Science Regional Conference Series in Mathematics, Vol. 63, 1986.

[Ru4] W. Rudin, A geometric criterion for algebraic varieties, J. Math. Mech. 17 (1968) 671–683.

[Ru5] W. Rudin, Functional Analysis, McGraw-Hill, New York, 1973.

[Ya] R. Yang, Beurling's phenomenon in two variables, International Equations Oper. Theory, to appear.

[Ya1] R. Yang, The Berger-Shaw theorem in the Hardy module over the bidisk, J. Operator Theory 42 (1999) 379–404.

[Ya2] R. Yang, Operator theory in the Hardy space over the bidisk (II), Integral Equations Oper. Theory 42 (2002) 99–124.

[Ya3] R. Yang, Operator theory in the Hardy space over the bidisk (III), J. Funct. Anal. 186 (2001) 521–545.

[Ya4] R. Yang, M_q-type submodules and issues of unitary equivalence, preprint.

[YZ] R. Yang, K. Zhu, The root operators on invariant subspaces of the Bergman space, Illinois J. Math., to appear.

[ZS] O. Zariski, P. Samuel, Commutative algebra, Vol. (I), (II), Van Nostrand, Princeton, NJ, 1958/1960.