

# REMARKS ON INNER FUNCTIONS AND OPTIMAL APPROXIMANTS

CATHERINE BÉNÉTEAU, MATTHEW FLEEMAN, DMITRY KHAVINSON,  
DANIEL SECO, AND ALAN A. SOLA

ABSTRACT. We discuss the concept of inner function in reproducing kernel Hilbert spaces with an orthogonal basis of monomials and examine connections between inner functions and optimal polynomial approximants to  $1/f$ , where  $f$  is a function in the space. We revisit some classical examples from this perspective, and show how a construction of Shapiro and Shields can be modified to produce inner functions.

## 1. INTRODUCTION

The notion of inner function is a central concept in operator-theoretic function theory, and has played a significant role in the description of the invariant subspaces of the shift operator. The Hardy space  $H^2$  of the disk consists of all the functions  $f$  analytic in the open unit disk  $\mathbb{D}$  that satisfy the norm boundedness condition

$$\|f\|_{H^2}^2 = \sup_{0 < r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta < \infty.$$

It is well-known (see, e.g., [7]) that such functions have non-tangential boundary values on the unit circle  $\mathbb{T}$  almost everywhere. A bounded analytic function  $f: \mathbb{D} \rightarrow \mathbb{C}$  is called *inner* if its boundary values satisfy

$$|f(e^{i\theta})| = 1 \quad a.e. \theta \in [0, 2\pi). \quad (1)$$

Beurling [5] showed that every closed  $z$ -invariant subspace is generated by an inner function. Implicit in his proof is the fact that if  $G$  is an inner function, then  $G$  is orthogonal to  $z^j G$  for every integer  $j \geq 1$ . In his analysis of  $z$ -invariant subspaces in the Dirichlet space  $D$ , which is the space of all analytic functions in the disk whose derivative is square integrable with respect to area measure, Richter [22] showed that as in the case of the Hardy space, the invariant subspaces are generated by a single function which satisfies the same orthogonality relationships.

---

*Date:* July 18, 2017.

*2010 Mathematics Subject Classification.* Primary 46E22; Secondary 30J05.

The situation in the Bergman spaces turned out to be considerably more complicated. The Bergman space  $A^2$  is the set of all analytic functions in the disk whose modulus squared is integrable with respect to area measure. In 1991, Hedenmalm [15] noticed that if  $M$  is a  $z$ -invariant subspace, then the function that is a solution to a particular extremal problem related to  $M$  plays the same role as that of the  $H^2$  inner functions. Moreover, this function is a so-called *contractive divisor*. Duren, Khavinson, Shapiro and Sundberg [8] extended this result to all  $A^p$  spaces of the disk. In [19], Korenblum coined the term “ $A^2$ -inner” to describe Hedenmalm’s extremal function  $G$ , and in [1], Aleman, Richter and Sundberg proved an analogue of Beurling’s theorem for the Bergman space, namely, that  $z$ -invariant subspaces  $M$  of the Bergman space are generated by the “wandering subspace”  $M \ominus zM$ . They defined an  $A^2$ -inner function as a function  $G$  that has norm 1 and is such that  $G$  is orthogonal to  $z^j G$  for every integer  $j \geq 1$ . More generally, such orthogonality conditions appear naturally in connection with the study of wandering subspaces in operator theory [13, 18, 20, 6], and more concretely in many papers dealing with the shift operator acting on spaces of analytic functions.

Another nice property that functions in the Hardy space satisfy is that any function factors as the product of an inner and an outer function. Outer functions  $f \in H^2$  are defined by the condition that  $\log |f(0)| = \frac{1}{2\pi} \int_0^{2\pi} \log |f(e^{i\theta})| d\theta$ , and in the Hardy space, such functions are always *cyclic* (and vice versa), that is, their polynomial multiples generate the whole space. This factorization was an important tool in Beurling’s characterization of the invariant subspaces of the Hardy space. In [19], Korenblum defined a notion of an  $A^2$ -outer function based on the concept of domination, and he proved that cyclic functions are outer in this sense, and the converse was proved in [1]. However, cyclic functions in the Bergman space are still not well-understood, and it is an open problem to characterize cyclic functions in other spaces of analytic functions such as the Dirichlet space.

In [2], the authors proposed to investigate cyclic functions in a large class of Hilbert spaces of analytic functions via the study of *optimal polynomial approximants*, that is, polynomials  $p$  minimizing the norm  $\|pf - 1\|$  among all polynomials in the space  $\mathcal{P}_n$  of polynomials of degree less or equal to  $n$ . This was done in the generality described here in [12]. Optimal polynomial approximants are given as the unique solution to a linear system of the form  $Mc = b$  where  $M$  is a matrix whose elements are given as

$$M_{j,k} = \langle z^j f, z^k f \rangle, \quad (2)$$

$\langle \cdot, \cdot \rangle$  denotes the inner product in the space  $H$ ,  $c$  is the vector of unknown coefficients of the optimal polynomial approximant  $p_n^*$ ; and  $b$  is the vector given by

$$(b_j)_{j=0}^n = (\langle 1, z^j f \rangle)_{j=0}^n = (\overline{f(0)}, 0, \dots, 0). \quad (3)$$

Optimal approximants were further studied in a subsequent series of papers [3, 4], and it seems worthwhile to isolate additional properties of functions  $G$  that satisfy the orthogonality relations

$$\langle z^j G, G \rangle = 0, \quad j = 1, 2, \dots \quad (4)$$

from this perspective.

The goal of this paper is to discuss the notion of inner function in a wide class of Hilbert spaces, with a special focus on describing the optimal polynomial approximants associated with such functions. We point out that certain functions, which we dub Shapiro-Shields functions after the authors in whose papers they first appeared, are inner and can be viewed as analogues of finite Blaschke products. Finally, we show how distances between 1 and subspaces generated by inner functions can be computed using elementary linear algebra methods.

In Section 2, we give the relevant definitions and prove a characterization of inner functions in terms of optimal polynomial approximants: the optimal approximant of all orders of an inner function turn out to be equal to a single constant. In Section 3, we use a slightly modified version of a construction of Shapiro and Shields to give examples of inner functions in Hilbert spaces of weighted Hardy type vanishing on prescribed finite sets. We also compute distances between the function 1 and invariant subspaces generated by an inner function. Finally, in Section 4, we discuss some examples of inner functions.

**Acknowledgements.** The authors would like to thank Constanze Liaw for many helpful conversations. Bénéteau and Khavinson are grateful to Stockholm University for supporting their visit during work on this project. The work of Seco is supported by Ministerio de Economía y Competitividad Project MTM2014-51824-P.

## 2. CHARACTERIZATION OF INNER FUNCTIONS VIA OPTIMAL APPROXIMANTS

From now on, given a sequence of real positive numbers  $\omega = \{\omega_k\}_{k \in \mathbb{N}}$ , let  $H = H_\omega^2$  be the Hilbert space of holomorphic functions  $f: \mathbb{D} \rightarrow \mathbb{C}$

with Taylor coefficients  $\{a_k\}_{k \in \mathbb{N}}$  endowed with norm

$$\|f\|_\omega := \left( \sum_{k=0}^{\infty} |a_k|^2 \omega_k \right)^{1/2} < \infty, \quad (5)$$

and equipped with the inner product  $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_\omega$  induced by the norm. Without loss of generality we will assume  $\omega_0 = 1$ . Furthermore, let us restrict ourselves to the class of weights  $\omega$  with

$$\lim_{k \rightarrow \infty} \frac{\omega_k}{\omega_{k+1}} = 1. \quad (6)$$

Condition (6) ensures that functions that are holomorphic on a disk of radius strictly greater than 1 are elements of  $H$ , that all elements of  $H$  are holomorphic on  $\mathbb{D}$ , and that both the forward and backward shifts are bounded operators. Note that the Hardy space, the Bergman space, and the Dirichlet space are all examples of spaces  $H_\omega^2$  with appropriate choice of weights  $\omega$ . See [7, 25, 16, 10, 11] for treatments of these spaces.

Spaces defined as in (5) are examples of *reproducing kernel Hilbert spaces* (from now on, *RKHS*) over the disk  $\mathbb{D}$ . This means that for any point  $z_0 \in \mathbb{D}$ , point evaluation is a bounded functional, and there exists a function  $k_{z_0} \in H_\omega^2$  such that for any function  $g \in H_\omega^2$  we have the reproducing property

$$g(z_0) = \langle g, k_{z_0} \rangle. \quad (7)$$

Using the usual representation of a reproducing kernel in terms of an orthonormal basis, we find that

$$k_{z_0}(z) = \sum_{k=0}^{\infty} \frac{\overline{z_0^k} z^k}{\omega_k}. \quad (8)$$

See [21] for a primer on the theory of RKHS.

**Definition 2.1.** A function  $f \in H$  is called *H-inner* (or simply, *inner*) if  $\|f\|_H = 1$  and for all  $j > 0$ :

$$\langle z^j f, f \rangle_H = 0. \quad (9)$$

Note that the standard Hardy inner functions are inner in this sense, as are Bergman-inner functions. We will say a closed subspace  $M \subset H$  is *z-invariant* in  $H$  if  $zM \subset M$ . For a function  $f \in H$ , we will write  $[f]$  for the subspace generated by  $f$  under the unilateral shift, that is, the closure of all polynomial multiples of  $f$  in the norm of  $H$ .

A related concept is that of cyclicity. A function is called *cyclic* (in  $H$ , for the operator of multiplication by  $z$ ) if  $[f]$  is equal to the whole

space  $H$ . Because polynomials are dense in  $H$ , the function 1 is cyclic, and hence a function  $f$  is cyclic if and only if  $1 \in [f]$ . That is, we can define  $f$  to be cyclic if there exists a sequence  $\{p_n\}_{n \in \mathbb{N}}$  of polynomials such that

$$\lim_{n \rightarrow \infty} \|p_n f - 1\|_H^2 = 0. \quad (10)$$

The following theorem is known for many spaces and appears in different forms in several places, including in [22, 1], but we include it for completeness and to emphasize that it holds thanks to general principles governing these Hilbert spaces, rather than to particular characteristics of the individual spaces.

**Theorem 2.2.** *Let  $M$  be  $z$ -invariant in  $H$ , and assume that there exists a function in  $M$  that does not vanish at 0. Then there is a unique solution  $G = h/\sqrt{h(0)}$  to the extremal problem*

$$\sup \{ \operatorname{Re}(g(0)) : g \in M, \|g\| \leq 1 \}, \quad (11)$$

where  $h$  is the orthogonal projection of 1 onto  $M$ . Moreover,  $G$  is an  $H$ -inner function. Conversely, if  $G$  is a (non-constant)  $H$ -inner function, then  $G$  generates a proper  $z$ -invariant subspace and solves (up to multiplication by a unimodular constant) the extremal problem (11) for  $M = [G]$ .

*Proof.* Let  $M$  be  $z$ -invariant in  $H$  and let  $h$  be the orthogonal projection of 1 onto  $M$ . Then since  $1 - h \perp h$ , we have that  $\|1 - h\|^2 + \|h\|^2 = 1$ ,  $\langle 1 - h, 1 - h \rangle = 1 - h(0)$ , and  $0 = \langle z^j h, 1 - h \rangle = -\langle z^j h, h \rangle$ ,  $j = 1, 2, \dots$ . Therefore  $\|h\|^2 = h(0)$ , so letting  $G(z) = h(z)/\sqrt{h(0)}$ , we see that  $G$  has norm 1 and is inner.

Now let  $g \in M$ . Then  $0 = \langle g, 1 - h \rangle = g(0) - \langle g, h \rangle$ , or  $\langle g, h \rangle = g(0)$ , so in fact  $h(z) = K_M(z, 0)$ , the reproducing kernel of  $M$  at 0. If in addition  $\|g\| \leq 1$ , then by the Cauchy-Schwarz inequality,  $\operatorname{Re}(g(0)) \leq \|h\| = \sqrt{h(0)}$ . Therefore, by the above,  $G$  solves the extremal problem (11).

Uniqueness follows by a standard argument in extremal problems in Hilbert spaces (see, e.g, [10]): it is easy to see that the extremal problem (11) is equivalent to the problem of minimizing the norm of all functions  $g \in M$  such that  $g(0) = 1$ . This last collection of functions forms a convex set in  $H$ , and convex sets in Hilbert spaces have unique elements of minimal norm.

Conversely, suppose  $G$  is a (non-constant) inner function. Suppose the invariant subspace  $[G] = H$ . Then  $1 \in [G]$ , and therefore there exist a sequence of polynomials  $p_n$  such that  $p_n G \rightarrow 1$  in  $H$ . Since  $G$  is inner,  $\langle G, p_n G \rangle = \overline{p_n(0)}$ . Taking limits of both sides as  $n \rightarrow \infty$  forces

$G(0) = 1/\overline{G(0)}$ , or  $|G(0)|^2 = 1$ . Since  $\|G\| = 1$ , this can only happen if  $G$  is a constant. Thus non-constant inner functions generate proper invariant subspaces.

Moreover, for any polynomial  $p$ , since  $G$  is inner,  $\langle (p-p(0))G, G \rangle = 0$ , or  $\langle pG, G \rangle = p(0)$ , that is,  $\overline{G(0)}G(z)$  is the reproducing kernel at 0 for the space  $[G]$ . Thus, by the above discussion,  $G$  is a unimodular multiple of the extremal solution to problem (11).  $\square$

Now let us examine how inner functions relate to optimal approximants. The study of optimal approximants in the context of the Dirichlet-type spaces  $D_\alpha$  was initiated by the authors in [2], who were interested in cyclic functions. They proposed to examine cyclicity (in a certain family of spaces) via the study of *optimal polynomial approximants*, that is, polynomials  $p$  minimizing the norm  $\|pf - 1\|$  among all polynomials in the space  $\mathcal{P}_n$  of polynomials of degree less or equal to  $n$ . However, this definition makes sense for *all* functions in the space, not only the cyclic functions. Indeed, the theorem below indicates what these optimal approximants are for *inner* functions. If we denote by  $\mathcal{P}_n$  the space of polynomials of degree less or equal to  $n$ , it is clear that if  $p_n^*$  is the optimal approximant of degree  $n$ , then  $p_n^*f$  is the orthogonal projection of 1 onto the space  $\mathcal{P}_n f$ .

**Theorem 2.3.** *Let  $f \in H$ , not identically 0, and for each  $n$ , let  $p_n^*$  be the optimal approximant of degree  $n$  of  $1/f$ .*

- (a) *If  $h$  is the orthogonal projection of 1 onto  $[f]$ , then  $h$  is the unique function such that  $\|p_n^*f - h\| \rightarrow 0$  as  $n \rightarrow \infty$ .*
- (b) *If  $f$  is inner, then all the optimal approximants are constants:  $p_n^*(z) = \overline{f(0)}$ .*
- (c) *If  $f$  is inner, then  $\|p_n^*f - 1\|^2 = \text{dist}_H^2(1, [f]) = 1 - |f(0)|^2$ .*

*Proof.* To prove (a), notice that  $\bigcup_{n \in \mathbb{N}} \mathcal{P}_n f$  is dense in  $[f]$ . Since  $H$  is a Hilbert space, the orthogonal projection of 1 onto  $\mathcal{P}_n f$ ,  $p_n^*f$ , must converge to the orthogonal projection of 1 onto  $[f]$ , that is,  $h$ . (This fact was already noticed by the authors in [4].)

To prove (b), notice that for an inner function  $f$ , the matrix  $M$  whose entries  $M_{j,k}$  are given by (2) has zeros in all positions of the first column and row (except for the first position,  $j = k = 0$ , where we have  $M_{0,0} = 1$ ). Therefore, the inverse  $M^{-1}$  of  $M$  also satisfies this property. This tells us that the optimal approximants  $p_n^*$  (for all  $n \in \mathbb{N}$ ) are all the same constant, given by

$$p_n^*(z) = p_0^*(z) = \overline{f(0)}. \quad (12)$$

But then any  $H$ -inner functions  $f$  and its corresponding optimal polynomial approximants  $p_n^*$  have for all  $n \in \mathbb{N}$  the property that

$$\|p_n^*f - 1\|^2 = 1 - |f(0)|^2, \quad (13)$$

and by definition of  $p_n^*$ , this quantity equals  $\text{dist}_H^2(1, [f])$ , which proves (c).  $\square$

Note that functions in  $H$  always admit a kind of a weak factorization, where the “outer” factor can be expressed in terms of the optimal approximants. More specifically, we have the following.

**Theorem 2.4.** *Let  $f \in H$  with  $f(0) \neq 0$  and for each  $n$ , let  $p_n^*$  be the optimal approximants of degree  $n$  of  $1/f$ . Let  $G$  be the solution to the extremal problem*

$$\sup \{ \text{Re}(g(0)) : g \in [f], \|g\| \leq 1 \}. \quad (14)$$

*Then there exists a function  $F$ , analytic in the disk, such that  $f(z) = G(z)F(z)$ , and*

$$F(z) = \lim_{n \rightarrow \infty} \frac{\overline{f(0)}}{p_n^*(z)}.$$

*Proof.* Suppose  $f \in H$ , and let  $G$  be the solution to the extremal problem (14). Then  $[G] \subset [f]$ , and therefore all the zeros of  $G$  must also be zeros of  $f$ , and so  $F := f/G$  is an analytic function in the disk. Letting  $p_n^*$  be the optimal approximants of degree  $n$  of  $1/f$ , by Theorem 2.3,  $\|p_n^*f - h\| \rightarrow 0$  as  $n \rightarrow \infty$ , where  $h$  is the orthogonal projection of 1 onto  $[f]$ . By theorem 2.2,  $G(z) = h(z)/\sqrt{h(0)}$ . Therefore, since norm convergence implies pointwise convergence,

$$\lim_{n \rightarrow \infty} p_n^*(z)G(z)F(z) = \sqrt{h(0)}G(z) = \overline{f(0)}G(z),$$

and therefore the conclusion follows.  $\square$

*Remark 2.5.* It is natural to ask whether the function  $F$  appearing in Theorem 2.4 is truly “outer”, in the sense of being cyclic in  $H$ ; in particular, we would require that  $F \in H$ . This, unfortunately, need not be the case in general: it is known, see [17], that factorization fails in some weighted Bergman spaces. While  $G$  has precisely the same zeros as  $f$  inside the disk, it may well happen that  $G(\zeta) = 0$  for some  $\zeta \in \mathbb{T}$ , leading to rapid growth in  $F$  at this boundary point, thus preventing  $F$  from belonging to  $H$ . See [17] for further details.

### 3. SHAPIRO-SHIELDS FUNCTIONS AND DISTANCES TO INVARIANT SUBSPACES GENERATED BY INNER FUNCTIONS

We now describe a method for constructing inner functions in the setting of a general  $H_\omega^2$ -space. The functions in question can be viewed as analogues of finite Blaschke products in the Hardy space or of contractive divisors associated with finite zero sets in the Bergman space.

Let  $Z = \{z_i\}_{i=1}^n$  be a collection of  $n$  distinct points in  $\mathbb{D} \setminus \{0\}$  and let  $k_w$  denote the reproducing kernel of  $H = H_\omega^2$  at the point  $w \in \mathbb{D}$ . We define  $K_Z$  to be the  $n \times n$  matrix with elements given by

$$K_{i,j} = \langle k_{z_j}, k_{z_i} \rangle, \quad (15)$$

for  $i, j = 1, \dots, n$ . In what follows, we write  $|M|$  for the determinant of a square matrix  $M$ .

In their study of the classical Dirichlet space and its zero sets, Shapiro and Shields [24] exhibited functions that have a prescribed finite set of zeros in the disk and maximize a certain functional, and a modification of their construction yields inner functions. With a standard Hilbert space argument, their construction can be extended to infinite zero sets (see [24], and see for instance [11] for further developments).

**Definition 3.1.** Let  $Z = \{z_i\}_{i=1}^n \subset \mathbb{D}$ , let  $k_{z_i}$  denote the reproducing kernel at  $z_i$  in  $H_\omega^2$ , and set

$$f_Z(z) := \begin{vmatrix} 1 & 1 & \cdots & 1 \\ (k_{z_i}(z))_{i=1}^n & K_Z & & \end{vmatrix}. \quad (16)$$

We define the *Shapiro-Shields function* in  $H_\omega^2$ , for the set  $Z$ , as

$$g_Z(z) := \frac{f_Z(z)}{\|f_Z\|_\omega}. \quad (17)$$

It is readily seen that  $g_Z(z_i) = 0$  for  $i = 1, \dots, n$  since the first and  $(i+1)$ th column in the determinant defining  $f_Z$  are then equal. In order to show that Shapiro-Shields functions  $g_Z$  are indeed  $H$ -inner, we need two auxiliary results.

**Lemma 3.2.**

$$\|f_Z\|^2 = \overline{|K_Z|} f_Z(0).$$

*Proof.* Consider the matrices  $(B_t)_{t=0}^n$  where

$$(B_t)_{i,j} = \begin{cases} K_{i,j} & \text{if } i \neq t \\ 1 & \text{otherwise.} \end{cases} \quad (18)$$

In particular,  $B_0 = K_Z$ .

We first expand the determinant defining  $f_Z$  in terms of the  $B_t$ 's:

$$f_Z(z) = |B_0| + \sum_{t=1}^n k_{z_t}(z) |B_t| (-1)^t. \quad (19)$$

This expresses the norm  $\|f_Z\|$  in terms of a useful linear combination:

$$\langle f_Z, f_Z \rangle = \overline{|B_0|} \langle f_Z, 1 \rangle + \sum_{t=1}^n \overline{|B_t|} (-1)^t \langle f_Z, k_{z_t} \rangle.$$

Two observations finish the proof: first, by the definition of the norm in  $H$ ,  $\langle f_Z, 1 \rangle = f_Z(0)$  and second, by the reproducing property of the kernels,  $\langle f_Z, k_{z_t} \rangle = f_Z(z_t)$  is the determinant of a matrix where two columns ( $j = 0$  and  $j = t$ ) are identical, and hence  $f_Z(z_t) = 0$ .  $\square$

The previous Lemma tells us, in particular, that  $g_Z$  is well defined provided  $|B_0| \neq 0$ . The next Lemma shows that this is indeed the case.

**Lemma 3.3.** *Let  $Z = \{z_i\}_{i=1}^n \subset \mathbb{D}$  be a set of  $n$  distinct points. Then  $K_Z$  is invertible.*

*Proof.* If  $n = 1$ ,  $K_Z = k_{z_1}(z_1) = \|k_{z_1}\|^2 > 0$ .

Suppose  $n \geq 2$ . From (15) we see that  $K_Z$  is a Gram matrix, and therefore its determinant is non-zero if and only if the kernels  $\{k_{z_i}\}_{i=1}^n$  are linearly independent. Seeking to arrive at a contradiction, let us assume linear independence fails. Without loss of generality, suppose

$$k_{z_n} = \sum_{i=1}^{n-1} \lambda_i k_{z_i}. \quad (20)$$

Let  $L$  be the Lagrange interpolating polynomial that is equal to 1 at  $z_n$  and vanishes at  $z_i$  for  $1 \leq i \leq n-1$ . Taking the inner product of  $L$  with the left hand side of (20) then gives 1, while the inner product of  $L$  with the right hand side of (20) gives 0, a contradiction.  $\square$

We arrive at the main result of this Section.

**Theorem 3.4.** *Let  $Z = \{z_i\}_{i=1}^n \subset \mathbb{D} \setminus \{0\}$  be a set of  $n$  distinct points. Then the function  $g_Z$  is  $H$ -inner and given by*

$$g_Z(z) = \frac{f_Z(z)}{\sqrt{|K_Z|} f_Z(0)}. \quad (21)$$

*Proof.* By the definition of  $g_Z$  and by Lemma 3.2, we have  $\|g_Z\| = 1$ . It remains to verify that

$$\langle z^k f_Z, f_Z \rangle = 0, \quad k = 1, 2, \dots \quad (22)$$

From the definition of  $f_Z$  in (16) we see that

$$\langle z^k f_Z, f_Z \rangle = \overline{|B_0|} \langle z^k f_Z, 1 \rangle + \sum_{t=1}^n \overline{|B_t|} (-1)^t \langle z^k f_Z, k_{z_t} \rangle. \quad (23)$$

The reproducing property of the kernel implies that

$$\langle z^k f_Z, k_{z_t} \rangle = z_t^k f_Z(z_t) = 0,$$

where the second identity comes from evaluating  $f_Z(z_t)$  as a determinant, with columns  $j = 0$  and  $j = t$  being the same. Finally, observe that  $\langle z^k f_Z, 1 \rangle = 0$  for all  $k \geq 1$  by the definition of the norm in  $H$ .  $\square$

Property (c) in Theorem 2.2 relates inner functions to distances to invariant subspaces. Let us now compute the distance from the function 1 to the invariant subspace generated by a polynomial with zero set  $Z = \{z_1, \dots, z_n\} \subset \mathbb{D} \setminus \{0\}$ . Denote by  $K_Z^{-1}$  the inverse of the matrix  $K_Z$  and let  $v^*$  be the transpose conjugate of a vector  $v$ .

**Theorem 3.5.** *Let  $Z \subset \mathbb{D} \setminus \{0\}$  be finite, and let  $\text{dist}_H(1, [g_Z])$  denote the distance from 1 to  $[g_Z]$  in  $H$ , where  $g_Z$  is the Shapiro-Shields function associated with  $Z$  and let  $K_Z$  be defined as in (15).*

*Then, letting  $v = (1, \dots, 1)$ , we have*

$$\text{dist}_H^2(1, [g_Z]) = v K_Z^{-1} v^*. \quad (24)$$

*Proof.* By Lemma 3.3, the matrix  $K_Z$  is invertible and hence the right-hand side of (24) is well defined.

As mentioned in Theorem 2.3, the distance from the function 1 to the invariant subspace generated by an inner function  $f \in H$  is given by  $\sqrt{1 - |f(0)|^2}$ .

With the notation from (17), this means that we just need to compute  $1 - |g_Z(0)|^2$ . From the definition of  $g_Z$  and Lemma 3.2 it is clear that

$$|g_Z(0)|^2 = \frac{|f_Z(0)|^2}{\|f_Z\|^2} = \frac{\overline{f_Z(0)}}{|K_Z|}. \quad (25)$$

Since  $|g_Z(0)| \in \mathbb{R}$ , we can ignore the conjugation. From now on, denote by  $v$  an  $n$ -vector with coordinates  $(1, \dots, 1)$ , and  $v^*$  its transpose (as in the statement of the Theorem). Bearing in mind that  $k_z(0) = 1$  for any  $z \in \mathbb{D}$ , we obtain

$$1 - |g_Z(0)|^2 = 1 - \frac{\begin{vmatrix} 1 & v \\ v^* & K_Z \end{vmatrix}}{|K_Z|}. \quad (26)$$

Notice that  $|K_Z| = \begin{vmatrix} 1 & v \\ 0 & K_Z \end{vmatrix}$ . This, applied to both the denominator and the numerator of the right-hand side of (26), shows that

$$1 - |g_Z(0)|^2 = \frac{\begin{vmatrix} 1 & v \\ 0 & K_Z \end{vmatrix} - \begin{vmatrix} 1 & v \\ v^* & K_Z \end{vmatrix}}{\begin{vmatrix} 1 & v \\ 0 & K_Z \end{vmatrix}}, \quad (27)$$

and by elementary linear algebra, this yields

$$1 - |g_Z(0)|^2 = \frac{\begin{vmatrix} 0 & v \\ -v^* & K_Z \end{vmatrix}}{\begin{vmatrix} 1 & v \\ 0 & K_Z \end{vmatrix}}. \quad (28)$$

By Lemma 3.3, the system

$$\begin{pmatrix} 1 & v \\ 0 & K_Z \end{pmatrix} c = \begin{pmatrix} 0 \\ -v^* \end{pmatrix} \quad (29)$$

has a unique solution  $c := (c_0, \dots, c_n)^T$  and, by applying Cramer's rule, we see that the left-hand side of (28) is equal to  $c_0$ .

The first equation in (29) tells us that

$$c_0 = -\sum_{j=1}^n c_j = -v \cdot c, \quad (30)$$

while the rest can be expressed in the simple form  $K_Z c = -v^*$ . Applying Lemma 3.3, we obtain

$$c = -K_Z^{-1} v^*. \quad (31)$$

Substituting the value of  $c_0$  obtained in (30) into (31) finishes the proof.  $\square$

#### 4. EXAMPLES

**Example 4.1** (Zero-based invariant subspaces in  $H^2$ ). The Shapiro-Shields functions associated with  $Z = \{z_1\}$ , a singleton, are straightforward to compute. In the case of the Hardy space  $H^2$ , the reproducing kernel is the Szegő kernel

$$k_{z_1}(z) = \frac{1}{1 - \bar{z}_1 z},$$

and hence

$$f_Z(z) = \frac{1}{1 - |z_1|^2} - \frac{1}{1 - \bar{z}_1 z} = \frac{\bar{z}_1}{1 - |z_1|^2} \frac{z_1 - z}{1 - \bar{z}_1 z}.$$

After normalizing, we obtain

$$g_Z(z) = \frac{\bar{z}_1}{|z_1|} \frac{z_1 - z}{1 - \bar{z}_1 z},$$

a classical Blaschke factor. The associated distance is

$$\text{dist}_{H^2}^2(1, [g_{\{z_1\}}]) = 1 - |z_1|^2.$$

Let us turn to zero sets containing two points,  $Z = \{z_1, z_2\}$ . After a somewhat lengthy computation that can be carried out by hand or using computer algebra, we obtain

$$f_Z(z) = \frac{1}{(1 - |z_1|^2)(1 - |z_2|^2)} \frac{|z_1 - z_2|^2}{|1 - \bar{z}_1 z_2|^2} \left( \frac{\bar{z}_1}{|z_1|} \frac{z_1 - z}{1 - \bar{z}_1 z} \right) \left( \frac{\bar{z}_2}{|z_2|} \frac{z_2 - z}{1 - \bar{z}_2 z} \right).$$

Note that, at this stage, it is not clear that the Shapiro-Shields function factors as a product of singleton Shapiro-Shields functions. After normalizing as in Theorem 3.4, however, we do arrive at

$$g_Z(z) = \left( \frac{\bar{z}_1}{|z_1|} \frac{z_1 - z}{1 - \bar{z}_1 z} \right) \left( \frac{\bar{z}_2}{|z_2|} \frac{z_2 - z}{1 - \bar{z}_2 z} \right),$$

a Blaschke product. It is immediate that

$$\text{dist}_{H^2}(1, [g_{\{z_1, z_2\}}]) = 1 - |z_1 z_2|^2.$$

**Example 4.2** (Zero-based invariant subspaces in  $A^2$ ). We next turn to the Bergman space, whose reproducing kernel is

$$k_{z_1}(z) = \frac{1}{(1 - \bar{z}_1 z)^2}.$$

Computing the corresponding  $2 \times 2$ -determinant in the definition of  $f_Z$  for the Bergman space, we find that

$$f_Z(z) = \frac{\bar{z}_1}{(1 - |z_1|^2)^2} \frac{z_1 - z}{1 - \bar{z}_1 z} \left( \frac{2 - \bar{z}_1 z - |z_1|^2}{1 - \bar{z}_1 z} \right).$$

Normalization as in Theorem 3.4 gives us

$$g_Z(z) = \frac{1}{\sqrt{2 - |z_1|^2}} \frac{\bar{z}_1}{|z_1|} \frac{z_1 - z}{1 - \bar{z}_1 z} \left( \frac{2 - \bar{z}_1 z - |z_1|^2}{1 - \bar{z}_1 z} \right).$$

This recovers the well-known single-point extremal function for the Bergman space, see [15, p.56]. In terms of distances,

$$\text{dist}_{A^2}^2(1, [g_{\{z_1\}}]) = 1 - |z_1|^2(2 - |z_1|^2) = (1 - |z_1|^2)^2.$$

It is a priori clear that the distance to a single-zero invariant subspace is smaller in the Bergman space than in  $H^2$ , but the above computations give us a quantitative comparison.

The invariant subspace generated by two simple zeros can also be handled. The function  $f_Z$  can be expressed as a linear combination of kernels, cf. also [15]. Moreover,

$$K_Z = \frac{1}{(1 - |z_1|^2)^2(1 - |z_2|^2)^2} \frac{|z_1 - z_2|^2}{|1 - z_1\bar{z}_2|^4} (2 - |z_1 + z_2|^2 + 2|z_1z_2|^2).$$

and

$$f_Z(0) = \frac{1}{(1 - |z_1|^2)^2(1 - |z_2|^2)^2} \left( 1 - (1 - |z_1|^2)^2 - (1 - |z_2|^2)^2 \right. \\ \left. + (1 - |z_1|^2)^2(1 - |z_2|^2)^2 \frac{2\operatorname{Re}(1 - \bar{z}_1z_2)^2 - 1}{|1 - z_1\bar{z}_2|^4} \right).$$

After simplifying, for instance by using computer algebra, we obtain the Bergman-inner function

$$g_Z(z) = C_Z B_Z(z) \left( 1 + \frac{1 - |z_1|^2}{1 - \bar{z}_1z} + \frac{1 - |z_2|^2}{1 - \bar{z}_2z} \right. \\ \left. + \frac{1 - |z_1|^2}{1 - \bar{z}_1z} \frac{1 - |z_2|^2}{1 - \bar{z}_2z} \frac{|1 - z_1\bar{z}_2|^2 - (1 - |z_1|^2)(1 - |z_2|^2)}{|1 - z_1\bar{z}_2|^2 + (1 - |z_1|^2)(1 - |z_2|^2)} \right),$$

where  $B_Z = b_{z_1}b_{z_2}$  is a Blaschke product and  $C_Z$  is the constant

$$C_Z = \left( 3 - |z_1|^2 - |z_2|^2 \right. \\ \left. + (1 - |z_1|^2)(1 - |z_2|^2) \frac{|1 - z_1\bar{z}_2|^2 - (1 - |z_1|^2)(1 - |z_2|^2)}{|1 - z_1\bar{z}_2|^2 + (1 - |z_1|^2)(1 - |z_2|^2)} \right)^{-1/2}$$

This is a special case of a more general result of Hansbo: in [14, Corollary 2.10], he obtains a formula for the  $A^p$ -extremal function associated with two zeros having arbitrary multiplicity. Note that, as is well-known (see, e.g., [10]), the rational function complementing the Blaschke product  $B_Z$  in the above formula is (up to a constant) the reproducing kernel at 0 in the weighted Bergman space with weight  $|B_Z|^2$ .

For  $A^2$  then, the distance to a two-zero invariant subspace is

$$\operatorname{dist}_{A^2}^2(1, [g_{\{z_1, z_2\}}]) = 1 - |z_1z_2|^2 \left( 3 - |z_1|^2 - |z_2|^2 \right. \\ \left. + (1 - |z_1|^2)(1 - |z_2|^2) \frac{|1 - \bar{z}_1z_2|^2 - (1 - |z_1|^2)(1 - |z_2|^2)}{|1 - \bar{z}_1z_2|^2 + (1 - |z_1|^2)(1 - |z_2|^2)} \right).$$

One immediately sees that not only the radial position but also the angle influences the distance function in the Bergman space: for two zeros on the same radius, the distance is maximized by placing the zeros at the same point, and minimized by placing them antipodally.

**Example 4.3** (Optimal approximants for a product function). The following example should be compared with Theorem 2.3, part (a).

For  $\lambda \in \mathbb{D}$ , let

$$f(z) = (1 - z)b_\lambda(z) = (1 - z)\frac{\lambda - z}{1 - \bar{\lambda}z}.$$

Since  $f$  contains a Blaschke factor,  $[f]$  is a proper closed invariant subspace of  $H^2$ , but  $f$  is not inner in the sense of Definition 2.1.

It is natural to ask what the optimal approximants  $p_n^*$  to  $1/f$  look like. By choosing coefficients in  $p = \sum_{k=0}^n c_k z^k$  in a way that minimizes the norm expression  $\|pf - 1\|_{H^2}$ , we obtain

$$p_0^* = \frac{1}{2}\bar{\lambda}, \quad p_1^* = \frac{2}{3}\bar{\lambda}\left(1 + \frac{1}{2}z\right), \quad \text{and} \quad p_2^* = \frac{3}{4}\bar{\lambda}\left(1 + \frac{2}{3}z + \frac{1}{3}z^2\right).$$

We now recognize the  $p_n^*$  for  $n = 0, 1, 2$  as  $\bar{\lambda}$ -multiples of the optimal approximants associated with the function  $1 - z$ , as computed in [2].

This is in fact the case for all  $n$ , as we will now prove. First, note that for any polynomial  $p$ ,

$$\|pf - 1\|_{H^2} = \|(1 - z)p - b_\lambda^{-1}\|_{H^2}.$$

Next, we expand  $b_\lambda^{-1}$  in a Laurent series: there is a constant term  $\bar{\lambda}$ , and the remaining powers of  $z$  are all negative. By orthogonality then, minimizing  $\|pf - 1\|_{H^2}$  over polynomials of degree  $n$  is equivalent to minimizing  $\|(1 - z)p - \bar{\lambda}\|_{H^2}$ . The equality

$$\|(1 - z)p - \bar{\lambda}\|_{H^2} = |\lambda| \left\| \frac{p(1 - z)}{\bar{\lambda}} - 1 \right\|_{H^2}$$

shows that the optimal approximants to  $1/f$  are indeed given by  $\bar{\lambda}$  times the optimal approximants to  $1/(1 - z)$  for all  $n$ .

A straight-forward modification of the above argument identifies the  $H^2$ -optimal approximants to  $f = (1 - z)B_\lambda$ , where  $B_\lambda = \prod_{k=1}^N b_{\lambda_k}$  is a finite Blaschke product: the  $n$ th-order optimal approximant is

$$p_n^* = \left( \prod_{k=1}^N \bar{\lambda}_k \right) \cdot q_n^*,$$

where  $q_n^*$  is the optimal approximant of order  $n$  to  $1/(1 - z)$ .

Despite its simplicity, this example illustrates the fact that in  $H^2$ , optimal approximants are essentially determined by the outer part, since the linear system giving the coefficients of the polynomials depends only on the outer part, and the inner part only affects the end result by multiplying the independent term by a constant.

**Example 4.4** (Singular inner functions). The construction of Shapiro and Shields does not produce inner functions with singular factors. However, it is instructive to examine distances associated with such functions as well.

Let us focus on the atomic case. For any  $\sigma > 0$ ,

$$S_\sigma(z) = \exp\left(-\sigma \frac{1+z}{1-z}\right)$$

is an inner function for the Hardy space  $H^2$ , and we have

$$\text{dist}_{H^2}^2(1, [S_\sigma]) = 1 - |S_\sigma(0)|^2 = 1 - e^{-2\sigma}.$$

The function  $S_\sigma$  is not  $A^2$ -inner, but Duren, Khavinson, Shapiro, and Sundberg computed the Bergman extremal function for the subspace generated by  $S_\sigma$  in [9]. Using a limiting argument, they obtained

$$G_\sigma(z) = \frac{1}{\sqrt{1+2\sigma}} \left(1 + \frac{2\sigma}{1-z}\right) S_\sigma(z).$$

Hence

$$\text{dist}_{A^2}^2(1, [S_\sigma]) = 1 - (1+2\sigma)e^{-2\sigma}.$$

As with the case of a single zero, the Bergman distance is smaller than the Hardy distance, but this time merely by a correction of the coefficient multiplying  $e^{-2\sigma}$ .

*Remark 4.5.* As can be seen from the preceding two examples, there are pairs of invariant subspaces, one zero-based and one associated with a singular inner function, that are equidistant to 1.

## REFERENCES

- [1] ALEMAN, A., RICHTER, S., and SUNDBERG, C. Beurling's theorem for the Bergman space, *Acta Math.* **177** (1996), 275-310.
- [2] BÉNÉTEAU, C., CONDORI, A. A., LIAW, C., SECO, D., and SOLA, A. A., Cyclicity in Dirichlet-type spaces and extremal polynomials, *J. Anal. Math.* **126** (2015), 259-286.
- [3] BÉNÉTEAU, C., KHAVINSON, D., LIAW, C., SECO, D., and SOLA, A. A., Orthogonal polynomials, reproducing kernels, and zeros of optimal approximants, *J. London Math. Soc.* **94** (2016), 726-746.
- [4] BÉNÉTEAU, C., KHAVINSON, D., LIAW, C., SECO, D., and SIMANEK, B., Zeros of optimal polynomial approximants: Jacobi matrices and Jentzsch-type theorems, *preprint*, <http://arxiv.org/abs/1606.08615>.
- [5] BEURLING, A., On two problems concerning linear transformations in Hilbert space, *Acta Math.* **81** (1949), 239-255.
- [6] CARSWELL, B., DUREN, P., and STESSIN, M., Multiplication invariant subspaces of the Bergman space, *Indiana Univ. Math. J.* **51** (2002), no. 4, 931-961.
- [7] DUREN, P., *Theory of  $H^p$  Spaces*, Academic Press, New York-London 1970; Second Edition, Dover Publications, Mineola, N.Y., 2000.

- [8] DUREN, P., KHAVINSON, D., SHAPIRO, H., and SUNDBERG, C., Contractive zero-divisors in Bergman spaces, *Pacific J. Math.* **157** (1993), 37-56.
- [9] DUREN, P., KHAVINSON, D., SHAPIRO, H., and SUNDBERG, C., Invariant subspaces in Bergman spaces and the biharmonic equation, *Michigan Math. J.* **41** (1994), 247-259.
- [10] DUREN, P. and SCHUSTER, A., *Bergman Spaces*, American Mathematical Society, Providence, R.I., 2004.
- [11] EL-FALLAH, O., KELLAY, K., MASHREGHI, J., and RANSFORD, T., *A Primer on the Dirichlet Space*, Cambridge Tracts in Math. **203**, Cambridge University Press, 2014.
- [12] FRICAÏN, E., MASHREGHI, J., and SECO, D., Cyclicity in reproducing kernel Hilbert spaces of analytic functions, *Comput. Methods Funct. Theory* (2014) Issue 14, 665-680.
- [13] HALMOS, P.R., Shifts on Hilbert spaces, *J. Reine Angew. Math.* **208** (1961), 102-112.
- [14] HANSBO, J., Reproducing kernels and contractive divisors in Bergman spaces, *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **232** (1996), Issled. po Linein. Oper. i Teor. Funktsii. 24, 174-198.
- [15] HEDENMALM, H., A factorization theorem for square area-integrable analytic functions, *J. Reine Angew. Math.* **422** (1991), 45-68.
- [16] HEDENMALM, H., KORENBLUM, B., and ZHU, K., *Theory of Bergman Spaces*, Graduate Texts in Mathematics, Springer-Verlag, New York, 2000.
- [17] HEDENMALM, H. and ZHU, K., On the failure of optimal factorization for certain weighted Bergman spaces, *Complex Variables Theory Appl.* **19** (1992), 165-176.
- [18] KHAVINSON, D., LANCE, T., and STESSIN, M., Wandering property in the Hardy space, *Michigan Math. J.* **44** (1997), no. 3, 597-606.
- [19] KORENBLUM, B., Outer functions and cyclic elements in Bergman spaces, *J. Funct. Anal.* **115** (1993), 104-118.
- [20] LANCE, T. , and STESSIN, M., Multiplication invariant subspaces of Hardy spaces, *Canad. J. Math.* **49** (1997), no. 1, 100-118.
- [21] PAULSEN, V. and RAGHUPATHI, M., *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, Cambridge Studies in Advanced Mathematics **152**, Cambridge University Press, 2016.
- [22] RICHTER, S., Invariant subspaces of the Dirichlet shift, *J. Reine Angew. Math.* **386** (1988), 205-220.
- [23] SECO, D., Some problems on optimal approximants, *Recent progress on operator theory and approximation in spaces of analytic functions*, *Contemp. Mathematics* **679** (2016) 193-205, AMS.
- [24] SHAPIRO, H., and SHIELDS, A. L., On the zeros of functions with finite Dirichlet integral and some related function spaces, *Math. Z.* **80** (1962) 217-299.
- [25] SHIELDS, A.L. Weighted shift operators and analytic function theory, *Topics in operator theory (C. Pearcy, ed.)*, Math. Surveys **13** (1974),49-128 .

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH FLORIDA, 4202 E. FOWLER AVENUE, TAMPA, FL 33620, USA.

*E-mail address:* `cbenetea@usf.edu`

DEPARTMENT OF MATHEMATICS, BAYLOR UNIVERSITY, ONE BEAR PLACE #97328, WACO, TX 76798-7328, USA.

*E-mail address:* `Matthew_Fleeman@baylor.edu`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH FLORIDA, 4202 E. FOWLER AVENUE, TAMPA, FL 33620, USA.

*E-mail address:* `dkhavins@usf.edu`

INSTITUTO DE CIENCIAS MATEMÁTICAS, CALLE NICOLÁS CABRERA, UAM, 28049 MADRID, SPAIN.

*E-mail address:* `dseco@mat.uab.cat`

DEPARTMENT OF MATHEMATICS, STOCKHOLM UNIVERSITY, SE-106 91 STOCKHOLM, SWEDEN.

*E-mail address:* `sola@math.su.se`