ON BEST APPROXIMATIONS TO COMPACT OPERATORS

DEBMALYA SAIN

Abstract. We study best approximations to compact operators between Banach spaces and Hilbert spaces, from the point of view of Birkhoff-James orthogonality and semi-inner-products. As an application of the present study, some distance formulae are presented in the space of compact operators. The special case of bounded linear functionals as compact operators is treated separately and some applications to best approximations in reflexive, strictly convex and smooth Banach spaces are discussed. An explicit example is presented in ℓ_p^n spaces, where $1 < p < \infty$, to illustrate the applicability of the methods developed in this article. A comparative analysis of the results presented in this article with the well-known classical duality principle in approximation theory is conducted to demonstrate the advantage in the former case, from a computational point of view.

1. Introduction

The purpose of this article is to study best approximations in the space of compact operators between Banach spaces and Hilbert spaces, and to present some distance formulae in certain special cases. Let us first establish the notations and the terminologies to be used throughout the article.

The letters $\mathbb{X}, \mathbb{Y}, \mathbb{Z}$ stand for Banach spaces and the letter $\mathbb{H}$ is used to denote a Hilbert space. The symbol $\langle \ , \ \rangle$ is used to denote the inner product on $\mathbb{H}$. Let θ denote the zero vector of any vector space, other than the scalar field. We work with both real and complex Hilbert spaces and *only* real Banach spaces. Let $B_{\mathbb{X}} = \{x \in \mathbb{X} \ : \|x\| \leq 1\}$ and $S_{\mathbb{X}} = \{x \in \mathbb{X} \ : \|x\| = 1\}$ be the unit ball and the unit sphere of $\mathbb{X}$, respectively. $\mathbb{X}$ is said to be strictly convex if every element of $S_{\mathbb{X}}$ is an extreme point of $B_{\mathbb{X}}$. Let $K(\mathbb{X}, \mathbb{Y})$ denote the Banach space of all compact operator from $\mathbb{X}$ to $\mathbb{Y}$, endowed with the usual operator norm and let $\mathbb{X}^*$ denote the dual space of $\mathbb{X}$. For the sake of brevity, we write $K(\mathbb{X}, \mathbb{Y}) = K(\mathbb{X})$, whenever $\mathbb{X} = \mathbb{Y}$. Given $T \in K(\mathbb{X}, \mathbb{Y})$, we use the notations $\mathcal{R}(T)$ and $\mathcal{N}(T)$ to denote the range of T and the kernel of T, respectively. The study of best approximation(s) to a given element out of a given subspace is a classical area of research in functional analysis. Let us recall the following basic definition in this context:

Definition 1.1. Given an element $x \in \mathbb{X}$ and a subspace $\mathbb{Y}$ of $\mathbb{X}$, let $dist\{x, \mathbb{Y}\} = \inf\{\|x - z\| : z \in \mathbb{Y}\}$ denote the distance between x and $\mathbb{Y}$. An element $y \in \mathbb{Y}$ is said to be a best approximation to x out of $\mathbb{Y}$ if $\|x - y\| = \min\{\|x - z\| : z \in \mathbb{Y}\}$.

It is well-known that in general neither the existence nor the uniqueness of best approximation is guaranteed. However, best approximation(s) always exist for

finite-dimensional subspaces, and more generally for compact subsets of infinite-dimensional subspaces. Moreover, best approximation is unique in a strictly convex Banach space, provided it exists. Clearly, the above definition makes sense in the space of operators between Banach (Hilbert) spaces, and it is worth studying only when $x \notin \mathbb{Y}$. For the study of best approximations of linear operators in various contexts and under additional assumptions, one may consult [1, 2, 18] and the references therein. The study of best approximations in Banach spaces is intimately connected to the concepts of Birkhoff-James orthogonality and semi-inner-products. Given $x, y \in \mathbb{X}$, we say that x is Birkhoff-James orthogonal to y, written as $x \perp_B y$, if $\|x + \lambda y\| \geq \|x\|$ for all scalars λ. It is easy to observe that Birkhoff-James orthogonality is homogeneous, i.e., $x \perp_B y$ implies that $\alpha x \perp_B \beta y$ for all scalars α, β. Moreover, we note that in a Hilbert space, the Birkhoff-James orthogonality relation $\perp_B$ coincides with the usual orthogonality relation $\perp$ induced by the underlying inner product $\langle\ ,\ \rangle$. Following [13], we say that $y \in x^+ (y \in x^-)$ if $\|x + \lambda y\| \geq \|x\|$ for all $\lambda \geq 0 (\lambda \leq 0)$. We use the notation $x^\perp = \{y \in \mathbb{X} : x \perp_B y\}$ to denote the Birkhoff-James orthogonality set of the vector x. We refer the readers to the pioneering articles [4, 7, 8] for the basic applications of Birkhoff-James orthogonality in understanding the geometry of Banach spaces, and to [3, 13, 15, 16, 17] for some of the more recent works in Banach spaces, involving the said notion of orthogonality. It is easy to observe that $y \in \mathbb{Y}$ is a best approximation to x out of $\mathbb{Y}$ if and only if $(x - y) \perp_B \mathbb{Y}$, i.e., $(x - y) \perp_B z$ for all $z \in \mathbb{Y}$. We next mention the concept of semi-inner-products in Banach spaces, which is integral to the theme of this article.

Definition 1.2. Let $\mathbb{X}$ be a real Banach space. A function $[\ ,\ ] : \mathbb{X} \times \mathbb{X} \longrightarrow \mathbb{R}$ is a semi-inner-product (s.i.p.) if for any $\alpha,\ \beta \in \mathbb{R}$ and for any $x,\ y,\ z \in \mathbb{X}$, it satisfies the following:
(a) $[\alpha x + \beta y, z] = \alpha[x, z] + \beta[y, z]$,
(b) $[x, x] > 0$, whenever $x \neq 0$,
(c) $|[x, y]|^2 \leq [x, x][y, y]$,
(d) $[x, \alpha y] = \alpha[x, y]$.

It was proved in [5] (see also [14] for a rigorous proof of the same) that every Banach space $(\mathbb{X}, \|.\|)$ can be represented as an s.i.p. space $(\mathbb{X}, [\ ,\ ])$ such that for all $x \in \mathbb{X}$, it holds that $[x, x] = \|x\|^2$. Whenever we speak of an s.i.p. $[\ ,\ ]$ in the context of a Banach space $\mathbb{X}$, we implicitly assume that $[\ ,\ ]$ is compatible with the norm, i.e., for all $x \in \mathbb{X}$, we have, $[x, x] = \|x\|^2$. In general, there can be many compatible s.i.p. corresponding to a given norm. As observed by Lumer in [10], there exists a unique s.i.p. on a normed space if and only if the space is smooth. We recall that $\mathbb{X}$ is said to be smooth if there exists a unique supporting hyperplane to $B_\mathbb{X}$ at each point of $S_\mathbb{X}$.

Bhatia and Šemrl studied Birkhoff-James orthogonality of matrices (viewed as operators on a finite-dimensional Hilbert space) and obtained some useful distance formulae in their seminal article [3]. Indeed, Theorem 1.1 of [3], also known as the Bhatia-Šemrl Theorem, gives a complete characterization of the Birkhoff-James orthogonality of matrices. For an analogous study of orthogonality of operators between real Banach spaces, we refer the readers to [13, 15, 16]. For the study of orthogonality and best approximations in the space of matrices (a special case of compact operators on a Hilbert space), one should see [6, 11]. The notion of the

norm attainment set of an operator plays a central role in the study of orthogonality of operators. Given $T \in K(\mathbb{X}, \mathbb{Y})$, let $M_T = \{x \in S_{\mathbb{X}} : \|Tx\| = \|T\|\}$ denote the norm attainment set of the operator T. Observe that whenever $\mathbb{X}$ is reflexive, it follows that $M_T \neq \emptyset$.

Our aim in the present article is to further build upon the ideas presented in the above mentioned works, in the space of compact operators. We obtain a complete characterization of best approximations to a given compact operator out of a given finite-dimensional subspace, separately for Hilbert spaces and reflexive Banach spaces. In particular, distance formula for a compact operator and a one-dimensional subspace are presented in both the cases. As the most important part of this article, we treat the special case of functionals as compact operators and present an efficient algorithm to study the best approximation problems in reflexive smooth and strictly convex Banach spaces. Explicit examples are presented in the setting of ℓ_p^n spaces to illustrate the applicability of the methods developed here. In the short final section of this article, we make a comparative analysis of our results with the classical duality principle in approximation theory. Indeed, we show that from a purely computational point of view, the use of orthogonality can evidently strengthen the well-known duality principle.

2. Best approximations to compact operators

We begin with the observation that Bhatia-Šemrl type theorems for compact operators on a reflexive Banach space immediately give a complete characterization of best approximations out of a one-dimensional subspace.

Proposition 2.1. *Let $\mathbb{X}$ be a reflexive Banach space and let $\mathbb{Y}$ be any Banach space. Let $T, A \in K(\mathbb{X}, \mathbb{Y})$ be linearly independent and let $\lambda_0 \in \mathbb{R}$. Then the following are equivalent:*
(i) $(T - \lambda_0 A) \perp_B A$,
(ii) *there exist $x, y \in M_{T-\lambda_0 A}$ and s.i.p. $[\ ,\ ]_1, [\ ,\ ]_2$ on $\mathbb{Y}$ such that $[Ax, Tx - \lambda_0 Ax]_1 \geq 0$ and $[Ay, Ty - \lambda_0 Ay]_2 \leq 0$,*
(iii) $\lambda_0 A$ *is a best approximation to T out of span$\{A\}$.*

Proof. The equivalence of (i) and (iii) follows from the definitions of Birkhoff-James orthogonality and best approximations. Let us first prove that (i) implies (ii). It follows from Theorem 2.1 of [16] that there exist $x, y \in M_{T-\lambda_0 A}$ such that $Ax \in ((T - \lambda_0 A)x)^+$ and $Ay \in ((T - \lambda_0 A)y)^-$. Applying Theorem 2.4 of [17], we deduce that (ii) holds true. In similar spirit, applying the converses of these two theorems, we obtain that (ii) implies (i). $\qquad\square$

If $M_{T-\lambda_0 A}$ is of a particularly nice form then we have a refinement of the above observation. The proof is omitted as it follows directly from Theorem 2.2 of [17] and the observation that $(T - \lambda_0 A)x \perp_B Ax$ if and only if there exists an s.i.p. $[\ ,\ ]$ on $\mathbb{Y}$ such that $[Ax, Tx - \lambda_0 Ax] = 0$.

Proposition 2.2. *Let $\mathbb{X}$ be a reflexive Banach space and let $\mathbb{Y}$ be any Banach space. Let $T, A \in K(\mathbb{X}, \mathbb{Y})$ be linearly independent and let $\lambda_0 \in \mathbb{R}$. Also assume that $M_{T-\lambda_0 A} = \pm D$, where D is a connected subset of $S_{\mathbb{X}}$. Then the following are equivalent:*
(i) $(T - \lambda_0 A) \perp_B A$,

(ii) *there exists $x \in M_{T-\lambda_0 A}$ and an s.i.p. $[\ ,\ ]$ on $\mathbb{Y}$ such that $[Ax, Tx - \lambda_0 Ax] = 0$,*
(iii) $\lambda_0 A$ *is a best approximation to T out of $span\{A\}$.*

In case of compact operators on a Hilbert space, we have yet another refinement of both the above results.

Proposition 2.3. *Let $\mathbb{H}$ be a Hilbert space and let $T, A \in K(\mathbb{H})$ be linearly independent. Let $\lambda_0 \in \mathbb{C}$. Then the following are equivalent:*
(i) $(T - \lambda_0 A) \perp_B A$,
(ii) *there exists $x \in M_{T-\lambda_0 A}$ such that $\langle Tx, Ax \rangle = \lambda_0 \|Ax\|^2$,*
(iii) $\lambda_0 A$ *is a best approximation to T out of $span\{A\}$.*

Proof. To prove the equivalence of (i) and (ii), we first observe that the inner product $\langle\ ,\ \rangle$ is the only s.i.p. on $\mathbb{H}$. Applying Remark 3.1 of [3] (also see Theorem 2.2 of [17] and Theorem 2.2 of [15] for the real case), and the compactness of the operator $T - \lambda_0 A$, we deduce that $(T - \lambda_0 A) \perp_B A$ if and only if there exists $x \in M_{T-\lambda_0 A}$ such that $\langle Ax, Tx - \lambda_0 Ax \rangle = 0$. Now the desired conclusion follows from the conjugate symmetry and additivity properties of the inner product $\langle\ ,\ \rangle$. □

Remark 2.4. In view of the above results, a natural question arises regarding the uniqueness of best approximation to a compact operator out of a one-dimensional subspace. This can be answered in terms of a strengthening of Birkhoff-James orthogonality introduced in [12]. Given $x, y \in \mathbb{X}$, we say that x is strongly orthogonal to y in the sense of Birkhoff-James, written as $x \perp_{SB} y$, if $\|x + \lambda y\| > \|x\|$ for all $\lambda \neq 0$. In each of the above propositions, it is easy to observe that $\lambda_0 A$ is the unique best approximation to T out of $span\{A\}$ if and only if $(T - \lambda_0 A) \perp_{SB} A$.

In Theorem 2.9 of [17], a distance formula has been presented for a compact operator and a one-dimensional subspace, under additional assumptions on norm attainment of certain operators. We would like to observe that the norm attainment condition can be relaxed, without any essential changes to the argument presented there. To this end, we first prove the following modification of Theorem 2.5 of [17].

Lemma 2.5. *Let $\mathbb{X}$ be a reflexive Banach space and let $\mathbb{Y}$ be any Banach space. Let $T, A \in K(\mathbb{X}, \mathbb{Y})$ be such that $T \perp_B A$ and $M_T = \pm D$, where D is a connected subset of $S_{\mathbb{X}}$. Then there exists an s.i.p. $[\ ,\ ]$ on $\mathbb{Y}$ such that*

$$\|T\| = \max\{|[Tx, y]| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B Ax\}$$
$$= \max\{|[Tx, y]| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, [Ax, y] = 0\}.$$

Proof. It follows from the defining properties of s.i.p. that given any s.i.p. $[\ ,\ ]$ on $\mathbb{Y}$, the following holds true:

$$\max\{|[Tx, y]| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B Ax\} \leq \|T\|.$$

On the other hand, Theorem 2.2 of [17] implies that there exists $x_0 \in M_T$ such that $Tx_0 \perp_B Ax_0$. If $Tx_0 = \theta$ then T is the zero operator and the result follows trivially. Let us assume that $Tx_0 \neq \theta$. Now, taking $x = x_0$ and $y = \frac{Tx_0}{\|Tx_0\|}$, we obtain by the homogeneity of Birkhoff-James orthogonality that $x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}$ and $y \perp_B Ax$. Since

$$|[Tx, y]| = |[Tx_0, \frac{Tx_0}{\|Tx_0\|}]| = \|Tx_0\| = \|T\|,$$

the first equality follows. To prove the second equality, we only need to observe that $y \perp_B Ax$ if and only if there exists an s.i.p. $[\,,\,]$ on $\mathbb{Y}$ such that $[Ax, y] = 0$. Indeed, applying Theorem 2.1 of [7], there exists $f \in S_{\mathbb{Y}^*}$ such that $f(y) = \|y\| = 1$ and $f(Ax) = 0$. Following the argument presented by Lumer in Theorem 2 of [10], we can construct an s.i.p. $[\,,\,]$ on $\mathbb{Y}$ such that $[Ax, y] = f(Ax) = 0$. This establishes the lemma. $\qquad \square$

Remark 2.6. We would like to note that the above lemma holds true for compact operators on a Hilbert space, real or complex, with the s.i.p. replaced by the underlying inner product $\langle\,,\,\rangle$. This can be verified easily by an easy application of Remark 3.1 of [3].

Applying the above lemma, we next obtain an improvement of Theorem 2.9 of [17].

Theorem 2.7. *Let $\mathbb{X}$ be a reflexive Banach space and let $\mathbb{Y}$ be any Banach space. Let $T, A \in K(\mathbb{X}, \mathbb{Y})$ be such that $T \notin span\{A\}$. Let $\lambda_0 A$ be a best approximation to T out of $span\{A\}$. Also assume that $M_{T-\lambda_0 A} = \pm D$, where D is a connected subset of $S_{\mathbb{X}}$. Then there exists an s.i.p. $[\,,\,]$ on $\mathbb{Y}$ such that*

$$dist(T, span\{A\}) = \max\{|[Tx, y]| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, [Ax, y] = 0\}.$$

Proof. As $T - \lambda_0 A, A$ are compact, $(T - \lambda_0 A) \perp_B A$ and $M_{T-\lambda_0 A} = \pm D$, where D is a connected subset of $S_{\mathbb{X}}$, it follows from Lemma 2.5 that there exists an s.i.p. $[\,,\,]$ on $\mathbb{Y}$ such that $\|T - \lambda_0 A\| = \max\{|[(T - \lambda_0 A)x, y]| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, [Ax, y] = 0\}$. Now, the desired result follows from the properties of s.i.p. and the fact that $\|T - \lambda_0 A\| = dist(T, span\{A\})$. This establishes the theorem. $\qquad \square$

The above distance formula becomes more useful from a computational point of view when $[\,,\,]$ is the unique s.i.p. on $\mathbb{Y}$, or, equivalently, when $\mathbb{Y}$ is smooth. In particular, given any $A \in K(\mathbb{X}, \mathbb{Y})$, the distance of a compact operator $T \in K(\mathbb{X}, \mathbb{Y})$ from $span\{A\}$ can be obtained by computing the maximum of the quantities $[Tx, y]$, where $x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}$, and $[Ax, y] = 0$, provided the condition on the norm attainment set is satisfied. The interesting thing to observe in this context is that we do not require the smoothness condition on $\mathbb{X}$. This is illustrated in the following example:

Example 2.8. Let $T, A \in K(\ell_\infty^2, \ell_2^2)$ be given by

$$T(u, v) = (u + 2v, 5u + 5v), A(u, v) = (u, 0) \text{ for all } (u, v) \in \ell_\infty^2.$$

Let us compute $dist(T, span\{A\})$, by applying Theorem 2.7. Given any $B \in K(\ell_\infty^2, \ell_2^2)$, we begin with the following two basic observations:

(i) $\|B\|_{\infty,2} = \max\{\|B(1, 1)\|_2, \|B(1, -1)\|_2\}$,
(ii) if $\|B(1, 1)\|_2 \neq \|B(1, -1)\|_2$ then M_B is necessarily of the form $M_B = \pm D$, where D is a connected subset of $S_{\ell_\infty^2}$. In fact, it is easy to show by using the convexity of norm that in this case D must be singleton.

Now, taking $B = T - \lambda A$, where $\lambda \in \mathbb{R}$, an easy computation reveals that $\|B(1, 1)\|_2 = \|B(1, -1)\|_2$ if and only if $\lambda = \lambda_1 = 13.5$. Let $\lambda_0 A$ be a best approximation to T out of $span\{A\}$. Since $\|T - \lambda_1 A\|_{\infty,2} = 14.5 > \|T\|_{\infty,2} = \sqrt{109}$, hence $\lambda_1 \neq \lambda_0$, so for $B_0 = T - \lambda_0 A$, we have $\|B_0(1, 1)\|_2 \neq \|B_0(1, -1)\|_2$, and

we conclude that the norm attainment condition on the operator $T - \lambda_0 A$ in Theorem 2.7 is satisfied. Taking any $x = (u, v) \in S_{\ell_\infty^2}$, we note that Ax is a scalar multiple of $(1, 0) \in S_{\ell_2^2}$. Therefore, by the homogeneity of Birkhoff-James orthogonality, if $y \in S_{\ell_2^2}$ is such that $y \perp_B Ax$, then $y = \pm(0, 1)$. In particular, $|[Tx, y]| = |[(u + 2v, 5u + 5v), (0, 1)]| = |5u + 5v|$, where $[\ ,\ ] = \langle\ ,\ \rangle$, the usual inner product on $\mathbb{R}^2$. Maximizing over $(u, v) \in S_{\ell_\infty^2}$, we obtain by applying Theorem 2.7,

$$dist(T, span\{A\}) = 10.$$

Remark 2.9. The above example shows that Theorem 2.7 is a proper improvement of Theorem 2.9 of [17]. Indeed, Theorem 2.9 of [17] is not applicable in case of the above example, as it can be easily checked that $M_{T-\lambda_1 A} = \{\pm(1, 1), \pm(1, -1)\}$, which is not of the desired form.

In case of compact operators on a Hilbert space, we have the following corollary to Theorem 2.7, the proof of which is omitted as it is now obvious in light of Remark 2.6.

Corollary 2.10. Let $\mathbb{H}$ be a Hilbert space and let $T, A \in K(\mathbb{H})$ be such that $T \notin span\{A\}$. Then

$$dist(T, span\{A\}) = \max\{|\langle Tx, y \rangle| : x, y \in S_{\mathbb{H}}, y \perp Ax\}.$$

Remark 2.11. When $\mathbb{H}$ is finite-dimensional, the above corollary was proved in [3], for the special case $A = I$, where I is the identity operator on $\mathbb{H}$. In this context, the significance of Theorem 2.7 is to illustrate that the concerned distance formula is valid for compact operators on a reflexive Banach space, with certain natural modifications. Moreover, the corresponding statement in case of Hilbert spaces follows directly from Theorem 2.7.

We next obtain a complete characterization of the best approximation to a compact operator out of an $n-$dimensional subspace, where $n \geq 2$, under an additional condition on the norm attainment set.

Theorem 2.12. *Let $\mathbb{X}$ be a reflexive Banach space, let $\mathbb{Y}$ be any Banach space and let $n \geq 2$. Let $T, A_1, A_2, \ldots, A_n \in K(\mathbb{X}, \mathbb{Y})$ be such that $A_1, A_2, \ldots, A_n$ are linearly independent and $T \notin \mathbb{Z}$, where $\mathbb{Z} = span\{A_1, A_2, \ldots, A_n\}$. Also assume that $M_{T-S} = \pm D_S$, where D_S is a connected subset of $S_{\mathbb{X}}$, for every $S \in \mathbb{Z}$. Let $\alpha_i \in \mathbb{R}$, where $1 \leq i \leq n$. Then $\sum_{i=1}^{n} \alpha_i A_i$ is a best approximation to T out of $\mathbb{Z}$ if and only if given any n scalars $\beta_1, \beta_2, \ldots, \beta_n$, there exist scalars $\gamma_1, \gamma_2, \ldots, \gamma_n$ and s.i.p. $[\ ,\ ]_1, [\ ,\ ]_2$ on $\mathbb{Y}$ such that the following holds true:*

$$\left\| T - \sum_{i=1}^{n} \beta_i A_i \right\| = \max\left\{ \left| \left[\left(T - \sum_{i=1}^{n} \beta_i A_i\right)x, y\right]_1 \right| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B \sum_{i=1}^{n} \gamma_i A_i x \right\}$$

$$\geq \max\left\{ \left| \left[\left(T - \sum_{i=1}^{n} \alpha_i A_i\right)x, y\right]_2 \right| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B \sum_{i=1}^{n} \gamma_i A_i x \right\}$$

$$= \left\| T - \sum_{i=1}^{n} \alpha_i A_i \right\|.$$

Proof. As the sufficient part of the theorem follows trivially, we prove only the necessary part. An easy application of the Hahn-Banach Theorem shows that $(T - \sum_{i=1}^{n} \beta_i A_i)^{\perp}$ contains a subspace of codimension one in $K(\mathbb{X}, \mathbb{Y})$. Since $n \geq 2$,

it follows that $(T - \sum_{i=1}^{n} \beta_i A_i)^{\perp} \bigcap \mathbb{Z} \neq \{\theta\}$. Let $\sum_{i=1}^{n} \gamma_i A_i \in (T - \sum_{i=1}^{n} \beta_i A_i)^{\perp} \bigcap \mathbb{Z}$ be non-zero, where $\gamma_i (1 \leq i \leq n)$ are scalars. Since $(T - \sum_{i=1}^{n} \beta_i A_i) \perp_B \sum_{i=1}^{n} \gamma_i A_i$, and $M_{(T - \sum_{i=1}^{n} \beta_i A_i)}$ is of the desired form, applying Lemma 2.5, we obtain that

$$\|T - \sum_{i=1}^{n} \beta_i A_i\| = \max\{|[(T - \sum_{i=1}^{n} \beta_i A_i)x, y]_1| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B \sum_{i=1}^{n} \gamma_i A_i x\},$$

for some s.i.p. $[\ ,\ ]_1$ on $\mathbb{Y}$, proving the first equality. On the other hand, we also have that $(T - \sum_{i=1}^{n} \alpha_i A_i) \perp_B \sum_{i=1}^{n} \gamma_i A_i \in \mathbb{Z}$, as $\sum_{i=1}^{n} \alpha_i A_i$ is a best approximation to T out of $\mathbb{Z}$. Therefore, applying similar arguments, we obtain that

$$\|T - \sum_{i=1}^{n} \alpha_i A_i\| = \max\{|[(T - \sum_{i=1}^{n} \alpha_i A_i)x, y]_2| : x \in S_{\mathbb{X}}, y \in S_{\mathbb{Y}}, y \perp_B \sum_{i=1}^{n} \gamma_i A_i x\},$$

for some s.i.p. $[\ ,\ ]_2$ on $\mathbb{Y}$, proving the last equality. The remaining inequality follows from the fact that $\|T - \sum_{i=1}^{n} \beta_i A_i\| \geq \|T - \sum_{i=1}^{n} \alpha_i A_i\|$. This establishes the theorem. □

In view of the above theorem, we note that in case $\mathbb{Y}$ is smooth, $[\ ,\ ]_1 = [\ ,\ ]_2$. Regarding the uniqueness of best approximation, we make the following remark.

Remark 2.13. $\sum_{i=1}^{n} \alpha_i A_i$ is the unique best approximation to T out of $\mathbb{Z} = span\{A_1, A_2, \ldots, A_n\}$, if and only if the inequality in Theorem 2.12 is strict. This can be verified easily by following the proof of the said theorem.

In case of compact operators on a Hilbert space, we have the following corollary to Theorem 2.12. The proof is an easy adaptation of the proof of Theorem 2.12 since there is a unique s.i.p. in the Hilbert space.

Corollary 2.14. Let $\mathbb{H}$ be a Hilbert space and let $T, A_1, A_2, \ldots, A_n \in K(\mathbb{H})$ be such that $A_1, A_2, \ldots, A_n$ are linear independent and $T \notin span\{A_1, A_2, \ldots, A_n\}$. Let $\alpha_i \in \mathbb{C}$, where $1 \leq i \leq n$. Then $\sum_{i=1}^{n} \alpha_i A_i$ is a best approximation to T out of $span\{A_1, A_2, \ldots, A_n\}$ if and only if given any n scalars $\beta_1, \beta_2, \ldots, \beta_n$, there exist scalars $\gamma_1, \gamma_2, \ldots, \gamma_n$ such that the following holds true:

$$\|T - \sum_{i=1}^{n} \beta_i A_i\| = \max\{|\langle (T - \sum_{i=1}^{n} \beta_i A_i)x, y \rangle| : x, y \in S_{\mathbb{H}}, y \perp \sum_{i=1}^{n} \gamma_i A_i x\}$$

$$\geq \max\{|\langle (T - \sum_{i=1}^{n} \alpha_i A_i)x, y \rangle| : x, y \in S_{\mathbb{H}}, y \perp \sum_{i=1}^{n} \gamma_i A_i x\}$$

$$= \|T - \sum_{i=1}^{n} \alpha_i A_i\|.$$

Moreover, the best approximation is unique if and only if the above inequality is strict.

3. BEST APPROXIMATIONS IN REFLEXIVE SPACES

In light of the results obtained so far in this article, it is evident that for best approximations to compact operators, norm attainment set plays a central role. By virtue of Remark 1 in [15], the norm attainment set of a compact operator on a Hilbert space $\mathbb{H}$ is necessarily the unit sphere of some subspace of $\mathbb{H}$. Since such a nicety is no longer present in case of compact operators between Banach spaces, we

require additional conditions in that case. However, for the class of bounded linear functionals on a reflexive Banach space, we do have an additional advantage as mentioned in the following proposition. The proof is easy, and is therefore omitted.

Proposition 3.1. *Let $\mathbb{X}$ be a reflexive Banach space and let $f \in X^*$. Then M_f is of the form $\pm D$, where D is a connected subset of $S_{\mathbb{X}}$ (D is a face of $B_{\mathbb{X}}$).*

Let us now present an analogous result to Proposition 2.1 and Theorem 2.7, for bounded linear functionals on a Banach space.

Theorem 3.2. *Let $\mathbb{X}$ be a reflexive Banach space and let $f, g \in \mathbb{X}^*$ be linearly independent. Let $\lambda_0 \in \mathbb{R}$. Then the following are equivalent:*

(i) $(f - \lambda_0 g) \perp_B g$,
(ii) $M_{f - \lambda_0 g} \bigcap \mathcal{N}(g) \neq \emptyset$,
(iii) $\lambda_0 g$ *is a best approximation to f out of $span\{g\}$.*
Moreover,

$$dist(f, span\{g\}) = \max\{|f(x)| : x \in \mathcal{N}(g) \bigcap S_{\mathbb{X}}\} = \|f|_{\mathcal{N}(g)}\|.$$

Proof. The equivalence of (i) and (iii) follows from the corresponding definitions, as before. To prove that each of them is equivalent to (ii), we apply Theorem 2.13 of [16] to conclude that there exist $z, w \in M_{f - \lambda_0 g}$ such that $(f - \lambda_0 g)z.g(z) \geq 0$ and $(f - \lambda_0 g)w.g(w) \leq 0$. By virtue of Proposition 3.1, $M_{f - \lambda_0 g} = \pm D$, where D is a connected subset of $S_{\mathbb{X}}$. It is now easy to deduce that there exists $x \in M_{f - \lambda_0 g}$ such that $(f - \lambda_0 g)x.g(x) = 0$. Clearly, $f - \lambda_0 g \neq \theta$. Therefore, we conclude that $g(x) = 0$, or, equivalently, $x \in M_{f - \lambda_0 g} \bigcap \mathcal{N}(g)$. On the other hand, if (ii) holds, then taking $x \in M_{f - \lambda_0 g} \bigcap \mathcal{N}(g)$, we obtain that for any $\lambda \in \mathbb{R}$,

$$\|(f - \lambda_0 g) + \lambda g\| \geq |(f - (\lambda_0 - \lambda)g)x| = |(f - \lambda_0 g)x| = \|f - \lambda_0 g\|.$$

This proves that $(f - \lambda_0 g) \perp_B g$, and completes the proof of the first part of the theorem. The second part follows from Theorem 2.7. We just need to observe that since $\mathbb{Y} = \mathbb{R}$, the only s.i.p. on $\mathbb{Y}$ is given by the usual multiplication of real numbers. Therefore, if $x \in S_{\mathbb{X}}$ is such that $g(x).y = 0$ for some $y = \pm 1$, then $x \in \mathcal{N}(g)$. This establishes the theorem. $\square$

Remark 3.3. Slightly digressing from our main topic of interest, we would like to note that Proposition 3.1 allows us to improve Theorem 2.13 of [16] by proving that the condition of strict convexity in the said theorem is redundant.

Our next result on best approximations to functionals is analogous to Theorem 2.12. The proof is omitted as it can be completed using similar arguments as before.

Theorem 3.4. *Let $\mathbb{X}$ be a reflexive Banach space and let $n \geq 2$. Let $f, g_1, g_2, \ldots, g_n \in \mathbb{X}^*$ be such that $g_1, g_2, \ldots, g_n$ are linearly independent and $f \notin \mathbb{Z}$, where $\mathbb{Z} = span\{g_1, g_2, \ldots, g_n\}$. Let $\alpha_i \in \mathbb{R}$, where $1 \leq i \leq n$. Then $\sum_{i=1}^{n} \alpha_i g_i$ is a best approximation to f out of $\mathbb{Z}$ if and only if given any n scalars $\beta_1, \beta_2, \ldots, \beta_n$, there exist*

scalars $\gamma_1, \gamma_2, \ldots, \gamma_n$ *such that the following holds true:*

$$\|f - \sum_{i=1}^{n} \beta_i g_i\| = \max\{|(f - \sum_{i=1}^{n} \beta_i g_i)x| : x \in \mathcal{N}(\sum_{i=1}^{n} \gamma_i g_i) \bigcap S_{\mathbb{X}}\}$$

$$\geq \max\{|(f - \sum_{i=1}^{n} \alpha_i g_i)x| : x \in \mathcal{N}(\sum_{i=1}^{n} \gamma_i g_i) \bigcap S_{\mathbb{X}}\}$$

$$= \|f - \sum_{i=1}^{n} \alpha_i g_i\|.$$

If in addition, $\mathbb{X}$ is strictly convex, then we have the following refinement of the above theorem:

Theorem 3.5. *Let* $\mathbb{X}$ *be a reflexive strictly convex Banach space and let* $n \in \mathbb{N}$. *Let* $f, g_1, g_2, \ldots, g_n \in \mathbb{X}^*$ *be such that* $g_1, g_2, \ldots, g_n$ *are linearly independent and* $f \notin \mathbb{Z}$, *where* $\mathbb{Z} = span\{g_1, g_2, \ldots, g_n\}$. *Let* $\alpha_i \in \mathbb{R}$, *where* $1 \leq i \leq n$. *Then* $\sum_{i=1}^{n} \alpha_i g_i$ *is a best approximation to* f *out of* $\mathbb{Z}$ *if and only if* $\bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap M_{f - \sum_{i=1}^{n} \alpha_i g_i} = \{\pm x_0\}$, *for some* $x_0 \in S_{\mathbb{X}}$.

Proof. To prove the sufficient part of the theorem, simply observe that for any $1 \leq j \leq n$, and for any $\lambda \in \mathbb{R}$, we have the following:

$$\|(f - \sum_{i=1}^{n} \alpha_i g_i) + \lambda g_j\| \geq |\{(f - \sum_{i=1}^{n} \alpha_i g_i) + \lambda g_j\}x_0|$$

$$= |(f - \sum_{i=1}^{n} \alpha_i g_i)x_0|$$

$$= \|f - \sum_{i=1}^{n} \alpha_i g_i\|.$$

This shows that for any $1 \leq j \leq n$, we have that $(f - \sum_{i=1}^{n} \alpha_i g_i) \perp_B g_j$. As $\mathbb{X}$ is reflexive and strictly convex, it is easy to see that $\mathbb{X}^*$ is smooth and consequently, Birkhoff-James orthogonality is right additive in $\mathbb{X}^*$. Therefore, it follows that $(f - \sum_{i=1}^{n} \alpha_i g_i) \perp_B \mathbb{Z}$. This is clearly equivalent to the desired conclusion that $\sum_{i=1}^{n} \alpha_i g_i$ is a best approximation to f out of $\mathbb{Z}$. Let us now prove the necessary part of the theorem. Clearly, for each $1 \leq j \leq n$, it follows that $(f - \sum_{i=1}^{n} \alpha_i g_i) \perp_B g_j$. Therefore, Theorem 3.2 implies that for each $1 \leq j \leq n$, $M_{f - \sum_{i=1}^{n} \alpha_i g_i} \bigcap \mathcal{N}(g_j) \neq \emptyset$. We also note that since $\mathbb{X}$ is strictly convex, there exists a unique $x_0 \in S_{\mathbb{X}}$ such that $M_{f - \sum_{i=1}^{n} \alpha_i g_i} = \{\pm x_0\}$. Combining these two observations, we obtain that $\bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap M_{f - \sum_{i=1}^{n} \alpha_i g_i} = \{\pm x_0\}$. This establishes the theorem. $\square$

As an immediate application of the above theorem, we have the following elegant distance formula in the dual of a reflexive strictly convex Banach space.

Theorem 3.6. *Let* $\mathbb{X}$ *be a reflexive strictly convex Banach space and let* $n \in \mathbb{N}$. *Let* $f, g_1, g_2, \ldots, g_n \in \mathbb{X}^*$ *be such that* $g_1, g_2, \ldots, g_n$ *are linearly independent and* $f \notin \mathbb{Z}$, *where* $\mathbb{Z} = span\{g_1, g_2, \ldots, g_n\}$. *Then*

$$dist(f, \mathbb{Z}) = \max\{|f(x)| : x \in \bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap S_{\mathbb{X}}\}.$$

Proof. A standard compactness argument shows that there exists $\alpha_i \in \mathbb{R}$, where $1 \leq i \leq n$, such that $\sum_{i=1}^{n} \alpha_i g_i$ is a best approximation to f out of $\mathbb{Z}$. Using the necessary part of Theorem 3.5, we obtain that $\bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap M_{f-\sum_{i=1}^{n} \alpha_i g_i} = \{\pm x_0\}$, for some $x_0 \in S_{\mathbb{X}}$. We also note that $\|f - \sum_{i=1}^{n} \alpha_i g_i\| = |(f - \sum_{i=1}^{n} \alpha_i g_i)x_0| = |f(x_0)| = \max\{|f(x)| : x \in \bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap S_{\mathbb{X}}\}$, where the last equality is a consequence of the fact that $f = f - \sum_{i=1}^{n} \alpha_i g_i$, when restricted to $\bigcap_{i=1}^{n} \mathcal{N}(g_i)$. Indeed, if there exists $y_0 \in \bigcap_{i=1}^{n} \mathcal{N}(g_i) \bigcap S_{\mathbb{X}}$ such that $|f(y_0)| > |f(x_0)|$ then we have that

$$\|f - \sum_{i=1}^{n} \alpha_i g_i\| \geq |(f - \sum_{i=1}^{n} \alpha_i g_i)y_0| = |f(y_0)| > |f(x_0)| = \|f - \sum_{i=1}^{n} \alpha_i g_i\|,$$

a contradiction. As $dist(f, \mathbb{Z}) = \|f - \sum_{i=1}^{n} \alpha_i g_i\|$, this completes the proof of the theorem. $\qquad\square$

The above theorem, in conjunction with Theorem 3.5, turns out to be extremely useful in studying best approximations in a reflexive, smooth and strictly convex Banach space $\mathbb{X}$. To illustrate the main idea, let us begin with the following standard best approximation problem:

Problem: Let $\mathbb{X}$ be a reflexive, smooth and strictly convex Banach space. Let $x_0, y_1, \ldots, y_n \in \mathbb{X}$, where $n \geq 1$, be such that $y_1, \ldots, y_n$ are linearly independent and $x_0 \notin \mathbb{Y} = span\{y_1, \ldots, y_n\}$. Find the (unique) best approximation to x_0 out of $\mathbb{Y}$ and compute $dist(x_0, \mathbb{Y})$.

Theorem 3.5 and Theorem 3.6 allow us to approach the above problem in the following three steps:

Step 1: We identify $\mathbb{X}$ with its double dual $\mathbb{X}^{**}$ via the canonical isometric isomorphism ψ. Let $\psi(x_0) = f_0$ and $\psi(y_i) = g_i$, where $1 \leq i \leq n$. Let $\mathbb{Z} = span\{g_1, \ldots, g_n\}$. The original problem is clearly equivalent to finding the best approximation to f_0 out of $\mathbb{Z}$ and computing $dist(f_0, \mathbb{Z})$. As $\mathbb{X}^*$ is reflexive, smooth and strictly convex, we are in a position to apply Theorem 3.5 and Theorem 3.6.

Step 2: Let $\mathcal{W} = \bigcap_{i=1}^{n} \mathcal{N}(g_i)$. Applying Theorem 3.6, we obtain that

$$dist(x_0, \mathbb{Y}) = dist(f_0, \mathbb{Z}) = \max\{|f_0(x)| : x \in \mathcal{W} \bigcap S_{\mathbb{X}^*}\}.$$

Step 3: Since $\mathbb{X}^*$ is strictly convex, so is $\mathcal{W}$. Therefore, $f_0\big|_{\mathcal{W}}$ attains norm at only one pair of points, say, $\pm h_0 \in S_{\mathcal{W}}$. Applying Theorem 3.5, $\sum_{i=1}^{n} \alpha_i g_i$ is the unique best approximation to f_0 out of $\mathbb{Z}$ if and only if $M_{f_0-\sum_{i=1}^{n} \alpha_i g_i} = \{\pm h_0\}$. We also note that since $\mathbb{X}^*$ is smooth, there exists a unique $\xi_0 \in \mathbb{X}^{**}$ such that $M_{\xi_0} = \{\pm h_0\}$. Therefore, we must have that $f_0 - \sum_{i=1}^{n} \alpha_i g_i = \xi_0$, which completely describes the unique best approximation to f_0 out of $\mathbb{Z}$ by means of the following equation:

$$\sum_{i=1}^{n} \alpha_i g_i = f_0 - \xi_0.$$

We have therefore obtained a complete answer to the original problem in light of the fact that $\sum_{i=1}^{n} \alpha_i y_i$ is the unique best approximation to x_0 out of $\mathbb{Y}$.

Remark 3.7. Unlike Theorem 3.5 and Theorem 3.6, we have made use of the strict convexity of $\mathbb{X}^*$ in the above algorithm. Since $\mathbb{X}$ is reflexive, this is ensured by (and is in fact equivalent to) the smoothness of $\mathbb{X}$.

The above algorithm can be applied in a more efficient manner from computational point of view, in case of ℓ_p^n spaces, where $1 < p < \infty$. This is because of the well-known identification of ℓ_p^n with $\ell_q^{n^*}$, where $\frac{1}{p} + \frac{1}{q} = 1$, under the mapping $\Omega : \ell_p^n \longrightarrow \ell_q^{n^*}$ given by

$$\Omega(\widetilde{x}) = f \qquad \forall\, \widetilde{x} = (x_1, x_2, \ldots, x_n) \in \ell_p^n,$$

where $f : \ell_q^n \longrightarrow \mathbb{R}$ is given by

$$f(e_i) = x_i \,\forall\, i \in \mathbb{N},$$

$\{e_i : 1 \leq i \leq n\}$ being the standard ordered basis for ℓ_q^n. Let us explain this in more detail in the following:

Problem: Let $\widetilde{x}, \widetilde{y_1}, \ldots, \widetilde{y_m} \in \ell_p^n$, where $1 < p < \infty$ and $1 \leq m < n$, be such that $\widetilde{y_1}, \ldots, \widetilde{y_m}$ are linearly independent and $\widetilde{x} \notin \mathbb{Y} = span\{\widetilde{y_1}, \ldots, \widetilde{y_m}\}$. Compute $dist(\widetilde{x}, \mathbb{Y})$.

Step 1: Let $\widetilde{x} = (x_1, \ldots, x_n)$ and let $\widetilde{y_i} = (y_{i1}, \ldots, y_{in})$, for each $1 \leq i \leq m$. Let $\Omega(\widetilde{x}) = f$ and let $\Omega(\widetilde{y_i}) = g_i$, for each $1 \leq i \leq m$. Let $\mathbb{Z} = span\{g_1, \ldots, g_m\}$. The original problem is clearly equivalent to finding the best approximation to f out of $\mathbb{Z}$ and computing $dist(f, \mathbb{Z})$.

Step 2: Let $\mathcal{W} = \bigcap_{i=1}^m \mathcal{N}(g_i)$. Clearly, $\mathcal{W}$ is a subspace in $\mathbb{R}^n$ determined by the fact that $(\eta_1, \ldots, \eta_n) \in \mathcal{W}$ if and only if $(\eta_1, \ldots, \eta_n)$ satisfies the following system of equations:

$$\sum_{j=1}^n y_{kj}\eta_j = 0 \;;\; 1 \leq k \leq m.$$

Applying Theorem 3.6, we obtain that

$$dist(\widetilde{x}, \mathbb{Y}) = dist(f, \mathbb{Z}) = \max\{|f(x)| : x \in \mathcal{W} \bigcap S_{\ell_q^n}\}.$$

In particular, when $n = 2$, we have the following explicit distance formula:

Theorem 3.8. *Let* $\mathbb{X} = \ell_p^2$, *where* $1 < p < \infty$. *Let* $\widetilde{x} = (a, b) \in \mathbb{X}$, *let* $(c, d) \neq (0, 0)$ *and let* $\mathbb{Y} = span\{(c, d)\}$ *be a one-dimensional subspace in* $\mathbb{X}$ *be such that* $\widetilde{x} \notin \mathbb{Y}$. *Then*

$$dist(\widetilde{x}, \mathbb{Y}) = \frac{|ad - bc|}{(|c|^q + |d|^q)^{\frac{1}{q}}},$$

where $\frac{1}{p} + \frac{1}{q} = 1$.

Proof. Let $\widetilde{y} = (c, d)$. Let $\Omega : \ell_p^2 \longrightarrow \ell_q^{2^*}$ be the canonical isometric isomorphism. Let $\Omega(\widetilde{x}) = f$ and let $\Omega(\widetilde{y}) = g$. Let us first assume that $d \neq 0$. Now, we have that

$$\mathcal{N}(g) = \{(\eta_1, \eta_2) \in \mathbb{R}^2 : c\eta_1 + d\eta_2 = 0\}.$$

Since $d \neq 0$, it follows that $(1, -\frac{c}{d}) \in \mathcal{N}(g)$. Let $\alpha \in \mathbb{R}$ be such that $\|\alpha(1, -\frac{c}{d})\|_q = 1$. An easy calculation implies that $|\alpha| = \frac{1}{(1+|\frac{c}{d}|^q)^{\frac{1}{q}}}$. Applying Theorem 3.6, we obtain that

$$dist(\widetilde{x}, \mathbb{Y}) = |f(\alpha(1, -\frac{c}{d}))| = \frac{|ad - bc|}{|d|(1 + |\frac{c}{d}|^q)^{\frac{1}{q}}} = \frac{|ad - bc|}{(|c|^q + |d|^q)^{\frac{1}{q}}},$$

proving the first equality. Similarly, if $c \neq 0$, it can be proved that

$$dist(\widetilde{x}, \mathbb{Y}) = \frac{|ad - bc|}{(|c|^q + |d|^q)^{\frac{1}{q}}}.$$

This establishes the theorem.   □

The methods developed in the present article point to the fact that every best approximation problem in a reflexive smooth and strictly convex Banach space corresponds to a maximization problem in a particular subspace of the dual space. Moreover, in case of ℓ_p^n spaces, where $1 < p < \infty$, computation of the distance of a given point from a given subspace reduces to a trivial calculation when the subspace is of codimension one in ℓ_p^n. Indeed, in this case we just need to evaluate a particular functional on the unit sphere of a one-dimensional subspace of ℓ_q^n, where $\frac{1}{p} + \frac{1}{q} = 1$. This is illustrated in the following example:

Example 3.9. Let $\mathbb{X} = \ell_p^3$, where $1 < p < \infty$. Let $x_0 = (a, b, c)$ and let $\mathbb{Y}_0 = span\{(1, 0, -1), (1, 2, 1)\}$, where $a, b, c \in \mathbb{R}$. Let us compute $dist(x_0, \mathbb{Y}_0)$ by using the ideas developed above. Let $\Omega : \ell_p^3 \longrightarrow \ell_q^{3*}$ be the canonical isometric isomorphism, where $\frac{1}{p} + \frac{1}{q} = 1$. Let $y_1 = (1, 0, -1)$ and let $y_2 = (1, 2, 1)$. Let us assume that $\Omega(x_0) = f_0$, $\Omega(y_1) = g_1$, $\Omega(y_2) = g_2$. An easy computation shows that $\mathcal{N}(g_1) \bigcap \mathcal{N}(g_2) = span\{(1, -1, 1)\}$. Since $\|(1, -1, 1)\|_q = 3^{\frac{1}{q}}$, applying Theorem 3.6 we obtain that

$$dist(x_0, \mathbb{Y}_0) = |f_0(3^{\frac{-1}{q}}(1, -1, 1))| = 3^{\frac{-1}{q}}|a - b + c|.$$

Let us end this section with the following remark, emphasizing the utility of the methods developed here in studying best approximation problems.

Remark 3.10. The above example, when viewed purely in the ℓ_p^3 setting, amounts to solving the following minimization problem:

$$\min_{\alpha, \beta \in \mathbb{R}} \{|a - \alpha - \beta|^p + |b - 2\beta|^p + |c + \alpha - \beta|^p\}^{\frac{1}{p}},$$

which require some computational efforts. However, applying the duality theory of best approximations developed in this article, we can readily conclude via a trivial computation that the answer to the above minimization problem is given by $3^{\frac{-1}{q}}|a - b + c|$, where $\frac{1}{p} + \frac{1}{q} = 1$. Equivalently, this can be expressed by means of the following inequality:

$$|a - \alpha - \beta|^p + |b - 2\beta|^p + |c + \alpha - \beta|^p \geq 3^{1-p}|a - b + c|^p \ \forall \ \alpha, \beta, a, b, c \in \mathbb{R}.$$

Indeed, the methods developed in this article in view of best approximations in ℓ_p spaces give rise to a family of such inequalities, which seem not so easy to prove otherwise. It should be noted that Theorem 3.6 guarantees that each of these inequalities is optimal.

4. A COMPARISON WITH THE CLASSICAL DUALITY PRINCIPLE

The aim of this section is to make a comparative analysis of some of our results, namely, Theorem 3.5 and Theorem 3.6, with the following well-known classical duality principle which is one of the very basic and foundational results in approximation theory:

$$dist(x, \mathbb{Y}) = \inf\{\|x - y\|_{\mathbb{X}} : y \in \mathbb{Y}\} = \sup\{f(x) : \|f\|_{\mathbb{X}^*} \leq 1, f|_{\mathbb{Y}} = 0\},$$

where $\mathbb{Y}$ is a subspace of $\mathbb{X}$.

We begin with the observation that neither Theorem 3.5 nor Theorem 3.6 of this article follows as a direct consequence of the above duality principle. In particular, the use of Birkhoff-James orthogonality in studying best approximations in a dual space allows us to approach the problem from a point of view which is computationally much more convenient. To explain this further, let us first have a closer look at the above distance formula given by the classical duality principle. To explicitly compute the distance between a given point x and a given subspace $\mathbb{Y}$ of the Banach space $\mathbb{X}$ by applying the above formula, the major hindrance is that in general we do not have an explicit description of all those functionals f in $B_{\mathbb{X}^*}$ which vanish on $\mathbb{Y}$. This can be remedied from a theoretical point of view in case $\mathbb{X}$ is reflexive and $\mathbb{Y}$ is of codimension one in $\mathbb{X}$. Since $\mathbb{X}$ is reflexive, it follows from Theorem 1 of [9] that there exists $z \in S_{\mathbb{X}}$ such that $z \perp_B \mathbb{Y}$. It is now easy to see that the following holds true:

$$\{f \in S_{\mathbb{X}^*} : f(\mathbb{Y}) = 0\} = \{f \in S_{\mathbb{X}^*} : |f(z)| = 1\}.$$

If in addition, $\mathbb{X}$ is smooth then there exists a unique $f_0 \in S_{\mathbb{X}^*}$ such that $f_0(z) = 1$. Therefore, by applying the classical duality principle, we obtain that

$$dist(x, \mathbb{Y}) = |f_0(x)|.$$

Unfortunately, Theorem 1 of [9] does not provide us with a method to compute z such that $z \perp_B \mathbb{Y}$. In fact, even in the finite-dimensional case, it is computationally not straightforward to explicitly find out such a z in a general Banach space, when the norm is not induced by an inner product. The reader is invited to verify this claim for ℓ_p^n spaces, where $n \geq 2$ and $p \in (1, \infty) \setminus \{2\}$. Indeed, to find out such a z such that $z \perp_B \mathbb{Y}$, it is readily seen from the expression of the unique s.i.p. on ℓ_p^n that the problem amounts to solving *nonlinear* equations, which is quite well-known to be computationally hard. Moreover, if the codimension of $\mathbb{Y}$ in $\mathbb{X}$ is strictly greater than one, then it is easy to see that $\mathbb{Y}$ cannot be written as $\mathbb{Y} = v^{\perp}$, for any $v \in S_{\mathbb{X}}$. Therefore, it becomes an even more difficult problem to have an explicit description of all those functionals f in $B_{\mathbb{X}^*}$ which vanish on $\mathbb{Y}$. This explains the computational difficulty in applying the classical duality principle in determining the distance between a given point x and a given subspace $\mathbb{Y}$ of the Banach space $\mathbb{X}$.

In stark contrast to the above, in case of a reflexive and strictly convex Banach space $\mathbb{X}$, our method allows us to approach the problem in a computationally efficient manner, by identifying $\mathbb{X}$ with $\mathbb{X}^{**}$ and then by applying Theorem 3.6. Indeed, in case $\mathbb{X}$ is a finite-dimensional strictly convex Banach space, and $\mathbb{Y}$ is of codimension one in $\mathbb{X}$, the problem of computing $dist(x, \mathbb{Y})$ reduces to a trivial computation involving the solution to a given system of *linear* equations. This

explains the computational advantage obtained in applying Theorem 3.6 over the classical duality principle and it can be further verified readily by trying to establish Theorem 3.8 or Example 3.9, by the later alone.

References

1. J. P. Aubin, *Best approximation of linear operators in Hilbert spaces*, SIAM J. Numer. Anal., **5** (3) (1968), 518–521.
2. L. Arambašić, R. Rajić, *Operator version of the best approximation problem in Hilbert C**-*modules*, J. Math. Anal. Appl., **413** (1) (2014), 311–320.
3. R. Bhatia and P. Šemrl, *Orthogonality of matrices and some distance problems*, Linear Algebra Appl., **287** (1999), no. 1–3, 77–85.
4. G. Birkhoff, *Orthogonality in linear metric spaces*, Duke Math. J., **1** (1935), 169–172.
5. J. R. Giles, *Classes of semi-inner-product spaces*, Trans. Amer. Math. Soc., **129** (1967), 436–446.
6. P. Grover, *Orthogonality to matrix subspaces, and a distance formula*, Linear Algebra Appl., **445** (2014), 280–288.
7. R. C. James, *Orthogonality and linear functionals in normed linear spaces*, Trans. Amer. Math. Soc., **61** (1947), 265–292.
8. R. C. James, *Inner product in normed linear spaces*, Bull. Amer. Math. Soc., **53** (1947), 559–566.
9. R. C. James, *Reflexivity and the sup of linear functionals*, Israel J. Math., **13** (1972), 289–300.
10. G. Lumer, *Semi-inner-product spaces*, Trans. Amer. Math. Soc., **100** (1961), 29–43.
11. K. K. Lau, W. O. J. Riha, *Characterization of best approximations in normed linear spaces of matrices by elements of finite-dimensional linear subspaces*, Linear Algebra Appl., **35** (1981), 109–120.
12. K. Paul, D. Sain, K. Jha, *On strong orthogonality and strictly convex normed linear spaces*, J Inequal. Appl., **242** (2013), doi.org/10.1186/1029-242X-2013-242.
13. D. Sain, *Birkhoff-James orthogonality of linear operators on finite dimensional Banach spaces*, J. Math. Anal. Appl., **447** (2017), 860–866.
14. D. Sain, On the Norm Attainment Set of a Bounded Linear Operator and Semi-Inner-Products in Normed Spaces, Indian J. Pure Appl. Math., **51** (2020), 179–186.
15. D. Sain, K. Paul, *Operator norm attainment and inner product spaces*, Linear Algebra Appl., **439** (2013), 2448–2452.
16. D. Sain, K. Paul, A. Mal, *A complete characterization of Birkhoff-James orthogonality in infinite dimensional normed space*, J. Operator Theory, **80** (2) (2018), 399–413.
17. D. Sain, A. Mal, K. Paul, *Some remarks on Birkhoff-James orthogonality of linear operators*, Expo. Math., **38** (2020), 138–147.
18. P. Wójcik, *Extensions of linear operators from hyperplanes and strong uniqueness of best approximation in* $\mathcal{L}(\mathcal{X}, \mathcal{W})$, J. Approx. Theory, **246** (2019), 28–42.

(Sain) Department of Mathematics, Indian Institute of Science, Bengaluru 560012, Karnataka, India

Email address: saindebmalya@gmail.com