Applied Analysis and Optimization
Volume 2, Number 2, 2018, 199–218

Yokohama Publishers
ISSN 2189-1664 Online Journal

SOME APPLICATIONS OF THE HAHN-BANACH SEPARATION THEOREM

FRANK DEUTSCH, HEIN HUNDAL, AND LUDMIL ZIKATANOV*

We dedicate this paper to our friend and colleague Professor Yair Censor on the occasion of his 75th birthday.

Abstract. We show that a single special separation theorem (namely, a consequence of the geometric form of the Hahn-Banach theorem) can be used to prove Farkas-type theorems, existence theorems for numerical quadrature with positive coefficients, and detailed characterizations of best approximations from certain important cones in Hilbert space.

1. Introduction

We show that a single separation theorem—the geometric form of the Hahn-Banach theorem—has a variety of different applications.

In section 2 we state this general separation theorem (Theorem 2.1), but note that only a special consequence of it is needed for our applications (Theorem 2.2). The main idea in Section 2 is the notion of *a functional being positive relative to a set of functionals* (Definition 2.3). Then a useful characterization of this notion is given in Theorem 2.4. Some applications of this idea are given in Section 3. They include a proof of the existence of numerical quadrature with positive coefficients, new proofs of Farkas type theorems, an application to determining best approximations from certain convex cones in Hilbert space, and a specific application of the latter to determine best approximations that are also *shape-preserving*. Finally, in Section 4, we note that the notion of a functional *vanishing relative to a set of functionals* has a similar characterization.

2. The key theorem

The classical Hahn-Banach separation theorem (see, e.g., [7, p. 417]) may be stated as follows. (We shall restrict our attention throughout this paper to *real* linear spaces although the general results have analogous versions in complex spaces as well.)

Theorem 2.1 (Separation Theorem)**.** *If K_1 and K_2 are disjoint closed convex subsets of a* (real) *locally convex linear topological space L, and K_1 is compact, then*

2010 *Mathematics Subject Classification.* 46N10,41A65, 52A27.
Key words and phrases. Separation theorem, functional positive relative to a set of functionals, Farkas-type theorems, quadrature with positive coefficients, best approximation by convex cones.
*The work of L. Zikatanov was supported in part by NSF awards DMS-1720114 and DMS-1522615.

there exists a continuous linear functional f on L such that

$$\sup_{x \in K_2} f(x) < \inf_{y \in K_1} f(y). \tag{2.1}$$

The main tool of this paper (Theorem 2.2) is the special case of Theorem 2.1 when K_1 is a single point, $L = X^*$ is the dual space of the Banach space X, and X^* is endowed with the weak* topology. In the latter case, the weak* continuous linear functionals on X^* are precisely those of the form $\hat{x}$, for each $x \in X$, defined on X^* by

$$\hat{x}(x^*) := x^*(x) \quad \text{for each } x^* \in X^* \tag{2.2}$$

(see, e.g., [7, p. 422]). It is well-known that X and $\widehat{X} := \{\hat{x} \mid x \in X\} \subset X^{**}$ are isometrically isomorphic. X is called reflexive if $\widehat{X} = X^{**}$.

Thus the main tool of this paper can be stated as the following corollary of Theorem 2.1.

Theorem 2.2 (Main Tool)**.** *Let X be a (real) normed linear space, Γ a weak* closed convex cone in X^*, and $x^* \in X^* \setminus \Gamma$. Then there exists $x \in X$ such that*

$$\sup_{y^* \in \Gamma} y^*(x) = 0 < x^*(x). \tag{2.3}$$

Proof. By Theorem 2.1 (with $L = X^*$ endowed with its weak* topology, $K_1 = \Gamma$, and $K_2 = \{x^*\}$), we deduce that there exists $x \in X$ such that

$$\sup_{y^* \in \Gamma} y^*(x) < x^*(x). \tag{2.4}$$

Since Γ is a cone, $0 \in \Gamma$ so that $\sup_{y^* \in \Gamma} y^*(x) \geq 0$. But if $\sup_{y^* \in \Gamma} y^*(x) > 0$, then there exists $y_0^* \in \Gamma$ such that $y_0^*(x) > 0$. Since Γ is a cone, $ny_0^* \in \Gamma$ for each $n \in \mathbb{N}$ and hence $\lim_{n \to \infty} ny_0^*(x) = \infty$. But this contradicts the fact that this expression is bounded above by $x^*(x)$ from the inequality (2.4). This proves (2.3). $\square$

A cone is a set C such that $\rho c \in C$ for each $c \in C$ and $\rho \geq 0$.

Recall that the **dual cone (annihilator) in X^* of a set** $S \subset X$, denoted $S^\ominus$ ($S^\perp$), is defined by

$$S^\ominus := \{x^* \in X^* \mid x^*(s) \leq 0 \quad \text{for each } s \in S\}$$

$$(S^\perp := S^\ominus \cap [-S^\ominus] = \{x^* \in X^* \mid x^*(s) = 0 \quad \text{for each } s \in S\}).$$

Clearly, $S^\ominus$ ($S^\perp$) is a weak* closed convex cone (subspace) in X^*. Similarly, if $\Gamma \subset X^*$, then the **dual cone (annihilator) in** X **of** Γ, denoted $\Gamma_\ominus$ ($\Gamma_\perp$), is defined by

$$\Gamma_\ominus := \{x \in X \mid x^*(x) \leq 0 \text{ for all } x^* \in \Gamma\}$$

$$(\Gamma_\perp := \Gamma_\ominus \cap [-\Gamma_\ominus] = \{x \in X \mid x^*(x) = 0 \text{ for all } x^* \in \Gamma\}).$$

Clearly, $\Gamma_\ominus$ ($\Gamma_\perp$) is a closed convex cone (subspace) in X. The **conical hull** of a set $S \subset X$, denoted cone$\,(S)$, is the smallest convex cone that contains S, i.e., the intersection of all convex cones that contain S. Equivalently,

$$\operatorname{cone}(S) := \left\{ \sum_1^n \rho_i s_i \mid \rho_i \geq 0, s_i \in S, n < \infty \right\}. \tag{2.5}$$

The (norm) closure of cone (S) will be denoted by $\overline{\text{cone}}\,(S)$. If $S \subset X^*$, then the weak* closure of cone (S) will be denoted by $w^*\!-cl(\text{cone}\,(S))$.

Definition 2.3. Let Γ be a subset of X^*. An element $x^* \in X^*$ is said to be **positive relative to** Γ if $x \in X$ and $y^*(x) \geq 0$ for all $y^* \in \Gamma$ imply that $x^*(x) \geq 0$.

Similarly, by replacing both "$\geq$" signs in Definition 2.3 by "$\leq$" signs, we obtain the notion of x^* being **negative relative to** Γ. The following theorem governs this situation.

Theorem 2.4. *Let X be a normed linear space, $\Gamma \subset X^*$, and $x^* \in X^*$. Then the following statements are equivalent:*

(1) *x^* is positive relative to Γ.*
(2) *x^* is negative relative to Γ.*
(3) *$\Gamma_\ominus \subset (x^*)_\ominus$.*
(4) *$x^* \in w^*\!-cl(\text{cone}\,(\Gamma))$, the weak* closed conical hull of Γ.*
(5) *$x^* \in \overline{\text{cone}}\,\Gamma$.*

Moreover, if X is reflexive, then each of these statements is equivalent to (5) *$x^* \in \overline{\text{cone}}\,\Gamma$, the* (*norm*) *closed conical hull of Γ.*

Proof. (1) $\Rightarrow$ (2). Suppose (1) holds, $z \in X$, and $y^*(z) \leq 0$ for all $y^* \in \Gamma$. Then $y^*(-z) \geq 0$ for all $y^* \in \Gamma$. By (1), $x^*(-z) \geq 0$ or $x^*(z) \leq 0$. Thus (2) holds.

(2) $\Leftrightarrow$ (3). Suppose (2) holds. If $x \in \Gamma_\ominus$, then $y^*(x) \leq 0$ for all $y^* \in \Gamma$. By (2), $x^*(x) \leq 0$ so $x \in (x^*)_\ominus$. Thus (3) holds. Conversely, if (3) holds, then (2) clearly holds.

(3) $\Rightarrow$ (4). If (4) fails, then $x^* \notin w^*\!-cl(\text{cone}\,(\Gamma))$. By Theorem 2.2, there exists $x \in X$ such that

$$\sup\{y^*(x) \mid y^* \in \text{cone}\,(\Gamma)\} = 0 < x^*(x). \tag{2.6}$$

In particular, $y^*(x) \leq 0$ for all $y^* \in \Gamma$, but $x^*(x) > 0$. Thus x^* is not negative relative to Γ. That is, (2) fails.

(4) $\Rightarrow$ (1). If (4) holds, then there is a net $(y^*_\alpha) \in \text{cone}\,(\Gamma)$ such that $x^*(x) = \lim_\alpha y^*_\alpha(x)$ for all $x \in X$. If $z \in X$ and $y^*(z) \geq 0$ for all $y^* \in \Gamma$, then, in particular, $y^*_\alpha(z) \geq 0$ for all α implies that $x^*(z) = \lim_\alpha y^*_\alpha(z) \geq 0$. That is, x^* is positive relative to Γ. Hence (1) holds, and the first four statements are equivalent.

Finally, suppose that X is reflexive. It suffices to show that $\overline{\text{cone}}\,(\Gamma) = w^*\!-cl(\text{cone}\,(\Gamma))$. Since X is reflexive, the weak topology and the weak* topology agree on X^* (see, e.g., [8, Proposition 3.113]). But a result of Mazur (see, e.g., [8, Theorem 3.45]) implies that a convex set is weakly closed if and only if it is norm closed. $\square$

In a Hilbert space H, we denote the inner product of x and y by $\langle x, y\rangle$ and the norm of x by $\|x\| = \sqrt{\langle x, x\rangle}$. Then, owing to the Riesz Representation Theorem which allows one to identify H^* with H, Definition 2.3 may be restated as follows.

Definition 2.5. A vector x in a Hilbert space H is said to be **positive relative to the set** $\Gamma \subset H$ if $y \in H$ and $\langle z, y\rangle \geq 0$ for all $z \in \Gamma$ imply that $\langle x, y\rangle \geq 0$.

Similarly, in a Hilbert space H, we need only one notion of a dual cone (annihilator). Namely, if $S \subset H$, then

$$S^{\ominus} := \{x \in H \mid \langle x, y\rangle \leq 0 \text{ for all } y \in S\}$$

$$\left(S^{\perp} = S^{\ominus} \cap (-S^{\ominus}) = \{x \in H \mid \langle x, y\rangle = 0 \text{ for all } y \in S\}\right).$$

Since a Hilbert space is reflexive, we obtain the following immediate consequence of Theorem 2.4.

Corollary 2.6. *Let H be a Hilbert space, $\Gamma \subset H$, and $x \in H$. Then the following statements are equivalent:*

(1) *x is positive relative to Γ.*
(2) *x is negative relative to Γ.*
(3) *$\Gamma^{\ominus} \subset (x)^{\ominus}$.*
(4) *$x \in \overline{\text{cone}}\,\Gamma$.*

Well-known examples of reflexive spaces are finite-dimensional spaces, Hilbert spaces, and the L_p spaces for $1 < p < \infty$. (The spaces L_1 and $C(T)$, for T compact, are never reflexive unless they are finite-dimensional.)

For a general convex set, we have the following relationship.

Lemma 2.7. *Let K be a convex subset of a normed linear space X. Then*

$$(K^{\ominus})_{\ominus} = \overline{\text{cone}}\,(K) \tag{2.7}$$

Proof. By definition, $K^{\ominus} = \{x^* \in X^* \mid x^*(K) \leq 0\}$. Thus

$$\begin{aligned} (K^{\ominus})_{\ominus} &= \{x \in X \mid x^*(x) \leq 0 \text{ for all } x^* \in K^{\ominus}\} \\ &= \{x \in X \mid x^*(x) \leq 0 \text{ for each } x^* \text{ such that } x^*(K) \leq 0\} \\ &\supset K. \end{aligned}$$

Since $(K^{\ominus})_{\ominus}$ is a closed convex cone, it follows that $(K^{\ominus})_{\ominus} \supset \overline{\text{cone}}\,(K)$. If the lemma were false, then there would exist $x \in (K^{\ominus})_{\ominus} \setminus \overline{\text{cone}}\,(K)$. By Theorem 2.1, there exists $x^* \in X^*$ such that $\sup x^*[\overline{\text{cone}}\,(K)] < x^*(x)$. Arguing as in the proof of Theorem 2.2, we deduce that $\sup x^*[\overline{\text{cone}}\,(K)] = 0 < x^*(x)$. But this contradicts the fact that $x^* \in K^{\ominus}$ and $x \in (K^{\ominus})_{\ominus}$. □

Corollary 2.8. *If C is a nonempty subset of X, then C is a closed convex cone in X if and only if*

$$C = (C^{\ominus})_{\ominus}. \tag{2.8}$$

It follows that *every* closed convex cone has the same special form. More precisely, we have the following easy consequence.

Lemma 2.9. *Let X be a normed linear space and let C be a nonempty subset of X. Then the following statements are equivalent:*

(1) *C is a closed convex cone.*

(2) *There exists a set* $\Gamma \subset X^*$ *such that*

$$C = \{x \in X \mid y^*(x) \le 0 \text{ for each } y^* \in \Gamma\}.$$

(*In fact,* $\Gamma = C^{\ominus}$ *works.*)

(3) *There exists a set* $\widetilde{\Gamma} \subset X^*$ *such that*

$$C = \{x \in X \mid y^*(x) \ge 0 \text{ for each } y^* \in \widetilde{\Gamma}\}.$$

(*In fact,* $\widetilde{\Gamma} = -C^{\ominus}$ *works.*)

We will need the following fact that goes back to Minkowski (see, e.g., [5, Lemma 6.33]).

Fact 2.10. If a nonzero vector x is a positive linear combination of the vectors $x_1, x_2, \ldots, x_n$, then x is a positive linear combination of a linearly independent subset of $\{x_1, x_2, \ldots, x_n\}$.

Theorem 2.11. *Let* X *be a reflexive Banach space and* $\Gamma \subset X^*$ *be weakly compact. Suppose there exists* $y \in X$ *such that* $y^*(y) > 0$ *for each* $y^* \in \Gamma$ *and* $\dim(\Gamma) = n$ (*so* Γ *contains a maximal set of* n *linearly independent vectors*). *Then each nonzero* $x^* \in \overline{\text{cone}}\,(\Gamma)$ *has a representation as*

$$x^* = \sum_1^m \rho_i y_i^*, \tag{2.9}$$

where $m \le n$, $\rho_i > 0$ *for* $i = 1, 2, \ldots, m$, *and* $\{y_1^*, y_2^*, \ldots, y_m^*\}$ *is a linearly independent subset of* Γ.

Proof. Let $\delta := \inf\{y^*(y) \mid y^* \in \Gamma\}$. If $\delta = 0$, then there exists a sequence (y_n^*) in Γ such that $\lim y_n^*(y) = 0$. By the Eberlein-Šmulian Theorem (see, e.g., [8, p. 129]), Γ is weakly sequentially compact, so there is a subsequence $(y_{n_k}^*)$ which converges weakly to $y^* \in \Gamma$ and, in particular, $0 = \lim y_{n_k}^*(y) = y^*(y) > 0$, which is absurd. Thus $\delta > 0$.

Let $x^* \in \overline{\text{cone}}\,(\Gamma) \setminus \{0\}$. Then there exists a sequence $(x_N^*)_1^\infty$ in $\text{cone}\,(\Gamma)$ such that $\|x_N^* - x^*\| \to 0$. Since $x^* \ne 0$, we may assume that $x_N^* \ne 0$ for all N. Then (x_N^*) is bounded, say $c := \sup_N \|x_N^*\| < \infty$, and

$$x_N^* = \sum_{i \in F_N} \rho_{N,i} x_{N,i}^* \tag{2.10}$$

for some scalars $\rho_{N,i} \ge 0$, $x_{N,i}^*$ in Γ, and F_N is finite. By the hypothesis $\dim(\Gamma) = n$ and Fact 2.10, we may assume that $F_N = \{1, 2, \ldots, n\}$. Thus we have that

$$x_N^* = \sum_1^n \rho_{N,i} x_{N,i}^* \tag{2.11}$$

where $\rho_{N,i} \ge 0$ for all i. Now

$$x_N^*(y) = \sum_{i=1}^N \rho_{N,i} x_{N,i}^*(y) \ge \sum_{i=1}^n \rho_{N,i}\delta \ge \rho_{N,i}\delta \text{ for each } i = 1, 2, \ldots, n. \tag{2.12}$$

Thus, for each $i \in \{1, \dots, n\}$, we have

$$\rho_{N,i} \le (\delta)^{-1} x_N^*(y) \le (\delta)^{-1} \|x_N^*\| \|y\| \le (\delta)^{-1} \|y\| c < \infty. \tag{2.13}$$

This shows that, for each $i = 1, 2, \dots, n$, the sequence of scalars $(\rho_{N,i})$ is bounded. By passing to a subsequence, we may assume that there exist $\rho_i \ge 0$ such that $\rho_{N,i} \to \rho_i$ for each i.

Since Γ is weakly sequentially compact, by passing to a further subsequence, say (N') of (N), we may assume that for each $i = 1, 2, \dots, n$, there exist $y_i^* \in \Gamma$ such that $x_{N',i}^* \to y_i^*$ weakly. Thus, for all $x \in X$, we have

$$x^*(x) = \lim_{N'} \sum_{i=1}^{n} \rho_{N',i} x_{N',i}^*(x) = \sum_{i=1}^{n} \rho_i y_i^*(x). \tag{2.14}$$

That is, $x^* = \sum_1^n \rho_i y_i^*$. By appealing to Fact 2.10, we get the representation (2.9) for x^*. □

Again, in the case of Hilbert space, this result reduces to the following fact that was first established by Tchakaloff [16], who used it to prove the existence of quadrature rules having positive coefficients (see also Theorem 3.1 below).

Corollary 2.12. *Let H be a Hilbert space and $\Gamma \subset H$ be weakly compact. Suppose there exists $e \in H$ such that $\langle y, e \rangle > 0$ for each $y \in \Gamma$ and $\dim(\Gamma) = n$ (so Γ contains a maximal set of n linearly independent vectors). Then each nonzero $x \in \overline{\text{cone}}\,(\Gamma)$ has a representation as*

$$x = \sum_1^m \rho_i y_i, \tag{2.15}$$

where $m \le n$, $\rho_i > 0$ for $i = 1, 2, \dots, m$, and $\{y_1, y_2, \dots, y_m\}$ is a linearly independent subset of Γ.

3. Some applications of theorem 2.4

In this section we show the usefulness of Theorem 2.4 by exhibiting a variety of different applications.

3.1. **An Application to the Existence of Positive Quadrature Rules.** In the first application, we show the existence of quadrature rules that are exact for polynomials of degree at most n, are based on a set of $n+1$ points, and have positive coefficients. Let $\mathcal{P}_n$ denote the set of polynomials of degree (at most) n regarded as a subspace of $C[a, b]$. That is, $\mathcal{P}_n$ is endowed with the norm $\|x\| = \max\{|x(t)| \mid a \le t \le b\}$. Define the linear functionals x^* and x_t^* on $X := \mathcal{P}_n$ by

$$x^*(x) := \int_a^b x(t)\, dt \quad \text{for all } x \in X \tag{3.1}$$

and

$$x_t^*(x) := x(t) \quad \text{for all } x \in X. \tag{3.2}$$

Theorem 3.1 (Numerical Quadrature)**.** *Let $X = \mathcal{P}_n$. Then there exists $m \le n+1$ points $a \le t_1 < t_2 < \cdots < t_m \le b$ and m scalars $w_i > 0$ such that $x^* = \sum_1^m w_i x^*_{t_i}$. More explicitly,*

$$\int_a^b x(t)\,dt = \sum_1^m w_i x(t_i) \textit{ for all } x \in X. \tag{3.3}$$

Proof. First note that x^* is positive relative to the set $\Gamma := \{x^*_t \mid t \in [a,b]\}$, since a function that is nonnegative at each point in $[a,b]$ must have a nonnegative integral. Since X is finite-dimensional, it is reflexive. By Theorem 2.4(4), we have $x^* \in \overline{\text{cone}}\,(\Gamma)$. Next note that for the identically 1 function e on $[a,b]$, we have $x^*_t(e) = 1$ for all $t \in [a,b]$. Further, it is easy to check that Γ is a closed and bounded subset of X^*, hence is compact since in a finite-dimensional space X all linear vector space topologies on X^* coincide (see, e.g., [8, Corollary 3.15]). Finally, since $\dim X^* = \dim X = n+1$, we can apply Theorem 2.11 to get the result. □

3.2. **Applications Related to Farkas Type Results.** In this section we note that the so-called Farkas Lemma is a consequence of Theorem 2.4. According to Wikipedia,

> Farkas' lemma is a solvability theorem for a finite system of linear inequalities in mathematics. It was originally proven by the Hungarian mathematician Gyula Farkas [9]. Farkas' lemma is the key result underpinning the linear programming duality and has played a central role in the development of mathematical optimization (alternatively, mathematical programming). It is used amongst other things in the proof of the Karush-Kuhn-Tucker theorem in nonlinear programming.

Since the setting for this result is in a Hilbert space, we will be appealing to the Hilbert space version of Theorem 2.4, namely, Theorem 2.6.

Theorem 3.2. *Let H be a Hilbert space and $\{b, a_1, a_2, \ldots, a_m\} \subset H$. Then exactly one of the following two systems has a solution:*

System 1: $\sum_1^m y_i a_i = b$ for some $y_i \ge 0$.

System 2: There exists $x \in H$ such that $\langle a_i, x\rangle \le 0$ for $i = 1, \ldots, m$ and $\langle b, x\rangle > 0$.

Proof. Letting $\Gamma := \{a_1, a_2, \ldots, a_m\}$, we see that the cone generated by Γ is finitely generated and, as is well-known, must be closed (see, e.g., [5, Theorem 6.34]). Hence $\overline{\text{cone}}\,(\Gamma) = \{\sum_1^m \rho_i a_i \mid \rho_i \ge 0\}$.

Clearly, system 1 has a solution if and only if $b \in \overline{\text{cone}}\,\{a_1, a_2, \ldots, a_m\}$. By Theorem 2.6, system 1 has a solution if and only if b is negative relative to $\Gamma := \{a_1, a_2, \ldots, a_m\}$.

But obviously, system 2 has a solution if and only if b is not negative relative to Γ. This completes the proof. □

If this theorem is given in its (obviously equivalent) matrix formulation, then it can be stated as in the following theorem. This is the version given by Gale, Kuhn, and Tucker [10] (where vector inequalities are interpreted componentwise).

Theorem 3.3. *Let A be an $m \times n$ matrix and $b \in \mathbb{R}^n$. Then exactly one of the following two systems has a solution:*
System 1: $A^T y = b$ and $y \geq 0$ for some $y \in \mathbb{R}^m$.
System 2: There exists $x \in R^n$ such that $Ax \leq 0$ and $\langle b, x\rangle > 0$.

The next theorem extends a result of Hiriart-Urruty and Lemaréchal [12, Theorem 4.3.4] who called it a *generalized Farkas theorem*.

Theorem 3.4. *Let J be a index set and (b, r) and $(s_j, p_j) \in \mathbb{R}^n \times \mathbb{R}$ for all $j \in J$. Suppose that the system of inequalities*

$$\langle s_j, x\rangle \leq p_j \tag{3.4}$$

has a solution $x \in \mathbb{R}^n$. Then the following statements are equivalent:

(1) *$\langle b, x\rangle \leq r$ for all x that satisfy relation* (3.4).
(2) *$(b, r) \in \overline{\text{cone}}\,\{(s_j, p_j) \mid j \in J\}$.*
(3) *$(b, r) \in \overline{\text{cone}}\,(\{(0, 1)\} \cup \{(s_j, p_j) \mid j \in J\})$.*

Proof. First note that $x \in \mathbb{R}^n$ is a solution to (3.4) if and only if $(x, -1) \in \mathbb{R}^n \times \mathbb{R}$ is a solution to

$$\langle (s_j, p_j), (x, -1)\rangle \leq 0 \text{ for all } j \in J. \tag{3.5}$$

Using this fact, we see that statement (1) holds $\iff$

$$\langle b, x\rangle - r \leq 0 \text{ for all } x \text{ that satisfy (3.5)} \tag{3.6}$$

$\iff$

$$\langle (b, r), (x, -1)\rangle \leq 0 \text{ for all } x \text{ such that } \langle (s_j, p_j), (x, -1)\rangle \leq 0 \text{ for all } j \in J.$$

But the last statement just means that (b, r) is negative relative to the set $\{(s_j, p_j) \mid j \in J\}$. By Corollary 2.6, this is equivalent to statement (2). Thus we have proved (1) $\iff$ (2).

Clearly (2) implies (3) since the conical hull in (3) is larger than that of (2). Finally, the same proof of (3) implies (1) as given in [12, Theorem 4.3.4] works here. □

Remark 3.5. Hirriart-Urruty and Lemaréchal [12, Theorem 4.3.4] actually proved the equivalence of statements (1) and (3) of Theorem 3.4. The sharper equivalence of statements (1) and (2) proven above was seen to be a simple consequence of Corollary 2.6.

3.3. **An Application to Best Approximation.** In this section we give an application to a problem of best approximation from a convex cone in a Hilbert space. We will need a special case of the following well-known characterization of best approximations from convex sets. This characterization goes back at least to Aronszajn [1] in 1950 (see also [5, Theorem 4.1]). The fact that every closed convex subset C of a Hilbert space H admits unique nearest points (best approximations) to each $x \in H$ is due to Riesz [15]. If $x \in H$, we denote its unique best approximation in C by $P_C(x)$.

Fact 3.6. Let C be a closed convex set in a Hilbert space H, $x \in H$, and $x_0 \in C$. Then $x_0 = P_C(x)$ if and only if $x - x_0 \in (C - x_0)^{\ominus}$, i.e.,

$$\langle x - x_0, y - x_0 \rangle \leq 0 \text{ for each } y \in C. \tag{3.7}$$

In the special case when C is a closed convex cone, Moreau [13] showed, among other things, that this result could be sharpened to the following.

Fact 3.7. Let C be a closed convex cone in the Hilbert space H, $x \in H$, and $x_0 \in C$. Then $x_0 = P_C(x)$ if and only if $x - x_0 \in C^{\ominus} \cap x_0^{\perp}$, i.e.,

$$\langle x - x_0, y \rangle \leq 0 \text{ for all } y \in C \text{ and } \langle x - x_0, x_0 \rangle = 0. \tag{3.8}$$

Moreover, $H = C \boxplus C^{\ominus}$, which means that each $x \in H$ has a unique representation as $x = c + c'$ where $c \in C$, $c' \in C^{\ominus}$, and $\langle c, c' \rangle = 0$. In fact,

$$x = P_C(x) + P_{C^{\ominus}}(x) \text{ for each } x \in H. \tag{3.9}$$

(For proofs of these facts, see, e.g., [5, Theorems 4.1, 4.7, and 5.9].)

The main result of this section is Theorem 3.9. To provide some motivation, we exhibit a simple example.

Example 3.8. Let $X = \ell_2(2)$ denote Euclidean 2-space, K denote the line segment joining the two points $k_1 = (0, -1)$ and $k_2 = (1, 1)$, and $C = -K^{\ominus}$, i.e., $C = \{y \in X \mid \langle y, k_i \rangle \geq 0 \text{ for } i = 1, 2\}$, see Figure 1. Let $x = (2, 1)$ and $x_0 = (2, 0)$. Then $x_0 = x + k_1 = P_C(x)$.

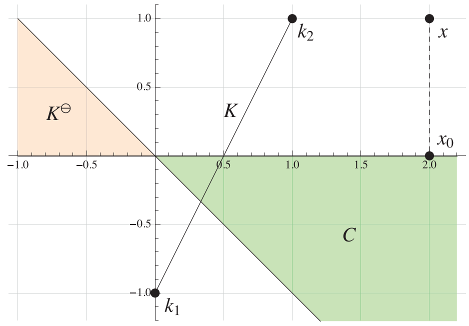

FIGURE 1

Theorem 3.9. *Let K be a compact set in the Hilbert space H, and suppose that there exists $e \in H$ such that*

(i) $\langle k, e \rangle > 0$ *for all* $k \in K$, *and*
(ii) $\dim K = n$.

Let C be (*the closed convex cone*) *defined by*

$$C := -K^{\ominus} = \{y \in H \mid \langle y, k \rangle \geq 0 \textit{ for all } k \in K\}. \tag{3.10}$$

Let $x \in H \setminus C$ and $x_0 \in C$. Then the following statements are equivalent:

(1) $x_0 = P_C(x)$;
(2)

$$x_0 = x + \sum_1^m \rho_i k_i, \tag{3.11}$$

where $1 \leq m \leq n$, $\rho_i > 0$, $k_i \in K$ for $i = 1, 2, \ldots, m$, where the vectors $k_1, k_2, \ldots, k_m$ are linearly independent, and $\langle k_i, x_0 \rangle = 0$ for $i = 1, 2, \ldots, m$.

Moreover, if $\dim H = n$ *and* $x_0 \neq 0$ *in any of the two statements, then* $m \leq n-1$.

Proof. (1) $\Rightarrow$ (2). If (1) holds, then by Fact 3.6, we have that $\langle x - x_0, y \rangle \leq 0$ for all $y \in C$ and $\langle x - x_0, x_0 \rangle = 0$. Thus if $y \in X$ and $\langle y, k \rangle \geq 0$ for all $k \in K$, then $y \in C$ so that $-\langle x_0 - x, y \rangle = \langle x - x_0, y \rangle \leq 0$, so $\langle x_0 - x, y \rangle \geq 0$. Thus $x_0 - x$

is positive relative to K. By Corollary 2.6, we see that $x_0 - x \in \overline{\text{cone}}\,(K)$. By Corollary 2.12, we have that $x_0 - x = \sum_1^m \rho_i k_i$, where $\rho_i > 0$ for all i, $m \leq n$, and the set $\{k_1, \ldots, k_m\}$ is linearly independent. Also, since $\langle x - x_0, x_0 \rangle = 0$, we see that

$$\sum_1^m \rho_i \langle k_i, x_0 \rangle = \left\langle \sum_1^m \rho_i k_i, x_0 \right\rangle = \langle x_0 - x, x_0 \rangle = 0 \tag{3.12}$$

which, since $\langle k_i, x_0 \rangle \geq 0$ and $\rho_i > 0$ for all i, implies that $\langle k_i, x_0 \rangle = 0$ for all i. Thus (2) holds.

(2) $\Rightarrow$ (1). If (2) holds, then $x_0 - x = \sum_1^m \rho_i k_i$ and $\langle k_i, x_0 \rangle = 0$ for all i. Thus for all $y \in C$ we have $\langle x_0 - x, y \rangle = \sum_1^m \rho_i \langle k_i, y \rangle \geq 0$ and $\langle x_0 - x, x_0 \rangle = \sum_1^m \rho_i \langle k_i, x_0 \rangle = 0$. In other words, $x - x_0 \in C^\circ \cap x_0^\perp$. By Fact 3.7, we see that $x_0 = P_C(x)$, i.e., (1) holds. This proves the equivalence of the two statements.

Finally, if $\dim X = n$ and $x_0 \neq 0$ in any of the two statements, then we see that $\langle x_0, k_i \rangle = 0$ for each $i = 1, \ldots, m$. But the null space of x_0, i.e., $x_0^\perp := \{y \in X \mid \langle x_0, y \rangle = 0\}$, is an $(n-1)$-dimensional subspace of the n-dimensional space X. Since $\{k_1, \ldots, k_m\}$ is a linearly independent set contained in $x_0^\perp$, we must have $m \leq n - 1$. □

Remark 3.10. It is worth noting that if either of the equivalent statements (1) or (2) holds in Theorem 3.9, then there exists at least one i such that $\langle x, k_i \rangle < 0$.

To see this, assume (2) holds. Then

$$\begin{aligned} \sum_1^m \rho_i \langle k_i, x \rangle &= \left\langle \sum_1^m \rho_k k_k, x \right\rangle = \langle x_0 - x, x \rangle = \langle x_0 - x, x - x_0 \rangle \\ &= -\|x_0 - x\|^2 < 0 \end{aligned}$$

which, since $\rho_i > 0$ for each i, implies that $\langle k_i, x \rangle < 0$ for some i.

The following corollary of Theorem 3.9 shows that in certain cases, one can even obtain an *explicit formula* for the best approximation to any vector.

Corollary 3.11. *Let $K = \{k_1, k_2, \ldots, k_n\}$ be an orthonormal subset of the Hilbert space H, and suppose that there exists $e \in H$ such that $\langle e, k_i \rangle > 0$ for each i. Let*

$$C := -K^\ominus = \{y \in H \mid \langle y, k_i \rangle \geq 0 \text{ for each } i\}, \tag{3.13}$$

and $x \in H$. Then

$$P_C(x) = x + \sum_{i=1}^n \max\{0, -\langle x, k_i \rangle\} k_i. \tag{3.14}$$

Proof. If $x \in C$, then $P_C(x) = x$ and $\langle x, k_i \rangle \geq 0$ for each i implies that $\max\{0, -\langle x, k_i \rangle\} = 0$ for each i and thus formula (3.14) is correct. Hence we may assume that $x \in H \setminus C$.

Let $x_0 = x + \sum_{i=1}^n \rho_i k_i$, where $\rho_i = \max\{0, -\langle x, k_i \rangle\}$. It suffices to show that $x_0 = P_C(x)$. Let $J = \{j \in \{1, 2, \ldots, n\} \mid \langle x, k_j \rangle < 0\}$. Since $x \notin C$, we see that J is not empty, $\rho_j = -\langle x, k_j \rangle$ for each $j \in J$, $\rho_i = 0$ for all $i \notin J$, and $x_0 = x + \sum_{j \in J} \rho_j k_j$.

Using the orthonormality of the set K, we see that for all $j \in J$,

$$\langle x_0, k_j\rangle = \langle x, k_j\rangle + \left\langle \sum_{i\in J} \rho_i k_i, k_j \right\rangle = \langle x, k_j\rangle + \rho_j = 0, \tag{3.15}$$

and for each $i \notin J$,

$$\langle x_0, k_i\rangle = \langle x, k_i\rangle + \sum_{j\in J} \rho_j \langle k_j, k_i\rangle = \langle x, k_i\rangle \geq 0. \tag{3.16}$$

The relations (3.15) and (3.16) together show that $x_0 \in C$. Finally, the equality (3.15) shows that Theorem 3.9(2) is verified. Thus $x_0 = P_C(x)$, and the proof is complete. □

We next consider *two* alternate versions of Theorem 3.9 which may be more useful for the actual computation of best approximations from finitely generated convex cones.

We consider the following scenario. Let H be a Hilbert space, $\{k_1, k_2, \ldots, k_m\}$ a finite subset of H, and C the convex cone generated by K:

$$C := \left\{ \sum_{i=1}^{m} \rho_i k_i \;\middle|\; \rho_i \geq 0 \text{ for all } i \right\} = \overline{\text{cone}}\,\{k_1, k_2, \ldots, k_m\}. \tag{3.17}$$

By definition of the dual cone, we have

$$\begin{aligned} C^{\ominus} &= \{y \in H \mid \langle y, c\rangle \leq 0 \text{ for all } c \in C\} \\ &= \{y \in H \mid \langle y, k_i\rangle \leq 0 \text{ for all } i = 1, \ldots, m\}. \end{aligned} \tag{3.18}$$

As an easy consequence of a theorem of the first author characterizing best approximations from a polyhedron ([5, Theorem 6.41]), we obtain the following.

Theorem 3.12. *Let $\{k_1, k_2, \ldots, k_m\}$ be a finite subset of the Hilbert space H, and let C be the finitely-generated cone defined by equation* (3.17). *Then for each $x \in H$,*

$$P_{C^{\ominus}}(x) = x - \sum_{1}^{m} \rho_i k_i \quad \textit{and} \quad P_C(x) = \sum_{1}^{m} \rho_i k_i \tag{3.19}$$

for any set of scalars ρ_i that satisfy the following three conditions:

$$\rho_i \geq 0 \qquad (i = 1, 2, \ldots, m) \tag{3.20}$$

$$\langle x, k_i\rangle - \sum_{j=1}^{m} \rho_j \langle k_j, k_i\rangle \leq 0 \qquad (i = 1, 2, \ldots, m) \tag{3.21}$$

and

$$\rho_i[\langle x, k_i\rangle - \sum_{j=1}^{m} \rho_j \langle k_j, k_i\rangle] = 0 \qquad (i = 1, 2, \ldots, m). \tag{3.22}$$

Moreover, if $x \in H$ and $x_0 \in C^{\ominus}$, then $x_0 = P_{C^{\ominus}}(x)$ if and only if

$$x_0 = x - \sum_{i\in I(x_0)} \rho_i k_i \textit{ for some } \rho_i \geq 0, \tag{3.23}$$

where $I(x_0) := \{i \mid \langle x_0, k_i\rangle = 0\}$.

Proof. In [5, Theorem 6.41], take $X = H$, $c_i = 0$ and $h_i = k_i$ for all $i = 1, 2, \ldots, m$, and note that $Q = \{y \in H \mid \langle y, k_i\rangle \leq 0\} = C^{\ominus}$. The conclusion of [5, Theorem 6.41] now shows that $P_{C^{\ominus}}(x) = P_Q(x) = x - \sum_1^m \rho_i k_i$, where the ρ_i satisfy the relations (3.20), (3.21), and (3.22). Finally, by Fact 3.7, we obtain that $P_C(x) = x - P_{C^{\ominus}}(x) = \sum_1^m \rho_i k_i$.

The last statement of the theorem follows from the last statement of [5, Theorem 6.41]. □

We will prove an alternate characterization of best approximations from finitely generated cones that yields detailed information of a different kind. But first we need to recall some relevant concepts.

For the remainder of this section, we assume that $T : H_1 \to H_2$ is a bounded linear operator between the Hilbert spaces H_1 and H_2 that has closed range. Then the adjoint mapping $T^* : H_2 \to H_1$ also has closed range (see, e.g., [5, Lemma 8.39]). We denote the *range* and *null* space of T by

$$\mathcal{R}(T) := \{T(x) \mid x \in H_1\}, \qquad \mathcal{N}(T) := \{x \in H_1 \mid T(x) = 0\}.$$

The following relationships between these concepts are well-known (see, e.g., [5, Lemma 8.33]):

$$\mathcal{N}(T) = \mathcal{R}(T^*)^{\perp}, \quad \mathcal{N}(T^*) = \mathcal{R}(T)^{\perp}, \quad \text{and} \tag{3.24}$$

$$\mathcal{N}(T)^{\perp} = \overline{\mathcal{R}(T^*)} = \mathcal{R}(T^*), \quad \mathcal{N}(T^*)^{\perp} = \overline{\mathcal{R}(T)} = \mathcal{R}(T). \tag{3.25}$$

Definition 3.13. For any $y \in H_2$, the set of **generalized solutions** to the equation $T(x) = y$ is the set

$$G(y) := \{x_0 \in H_1 \mid \|T(x_0) - y\| \leq \|T(x) - y\| \text{ for all } x \in H_1\}.$$

Since $\mathcal{R}(T)$ is closed, it is a Chebyshev set so $G(y)$ is not empty. For each $y \in H_2$, let $T^{\dagger}(y)$ denote the *minimal norm* element of $G(y)$. The mapping $T^{\dagger} : H_2 \to H_1$ thus defined is called the **generalized inverse** of T.

The following facts are well-known (see, e.g., [11] or [5, pp 177–185]).

Fact 3.14. (1) $T^{\dagger}$ is a bounded linear mapping.
(2) $(T^*)^{\dagger} = (T^{\dagger})^*$.
(3) $TT^{\dagger} = P_{\mathcal{R}(T)} = P_{\mathcal{N}(T^*)^{\perp}}$.
(4) $T^{\dagger}T = P_{\mathcal{N}(T)^{\perp}}$.
(5) $TT^{\dagger}T = T$.

As in the above Theorem 3.12, we again let $\{k_1, k_2, \ldots, k_m\}$ be a finite subset of the Hilbert space H and C be the convex cone generated by the k_i:

$$C = \overline{\text{cone}}\,\{k_1, k_2, \ldots, k_m\} = \left\{\sum_1^m \rho_i k_i \mid \rho_i \geq 0 \text{ for all } i\right\}. \tag{3.26}$$

It follows that

$$C^{\ominus} = \{y \in H \mid \langle y, c\rangle \leq 0 \text{ for all } c \in C\} \tag{3.27}$$

$$= \{y \in H \mid \langle y, k_i\rangle \leq 0 \text{ for all } i\}. \tag{3.28}$$

Let $S : \mathbb{R}^m \to H$ be the bounded linear operator defined by

$$S(\alpha) = \sum_1^m \alpha_i k_i \qquad \text{for all} \quad \alpha = (\alpha_1, \alpha_2, \dots, \alpha_m) \in \mathbb{R}^m. \tag{3.29}$$

If $S^* : H \to \mathbb{R}^m$ denotes the adjoint of S, then

$$\langle S^*(y), e_j\rangle = \langle y, S(e_j)\rangle = \langle y, k_j\rangle \text{ for all } j, \tag{3.30}$$

where e_j denote the canonical bases vectors in $\mathbb{R}^m$, i.e., $e_j = (\delta_{1j}, \delta_{2j}, \dots, \delta_{mj})$, and δ_{ij} is Kronecker's delta—the scalar which is 1 when $i = j$ and 0 otherwise.

As was noted in Fact 3.7, if C is a closed convex cone in a Hilbert space H, then $H = C \boxplus C^\ominus$, which means that each $x \in H$ has a unique orthogonal decomposition as $x = P_C(x) + P_{C^\ominus}(x)$. In the case of a finitely generated cone, we will strengthen and extend this even further by showing that $C^\ominus$ has an even *stronger* orthogonal decomposition as the sum of $N(S^*)$ and a certain subset of $\mathcal{N}(S^*)^\perp$. For a vector $\rho = (\rho_1, \dots, \rho_m) \in \mathbb{R}^m$, we write $\rho \geq 0$ to mean $\rho_i \geq 0$ for each i.

Lemma 3.15. *The following orthogonal decomposition holds:*

$$C^\ominus = \mathcal{N}(S^*) \boxplus \mathcal{B}, \qquad \textit{where} \tag{3.31}$$

$$\mathcal{B} := \{z \in H \mid z = -(S^*)^\dagger(\rho), \quad \rho \in \mathcal{N}(S)^\perp, \quad \rho \geq 0\} \subset \mathcal{N}(S^*)^\perp. \tag{3.32}$$

In particular, each $c' \in C^\ominus$ has a unique representation as $c' = y + z$, where $y \in \mathcal{N}(S^*)$, $z \in \mathcal{B}$ and $\langle y, z\rangle = 0$.

Proof. We first show that $\mathcal{B} \subset \mathcal{N}(S^*)^\perp$. If $z \in \mathcal{B}$, then $z = -(S^*)^\dagger(\rho)$, where $\rho \in \mathcal{N}(S)^\perp$. Since $\mathcal{N}(S)^\perp = \mathcal{R}(S^*)$ by the relation (3.25), we can write $z = -(S^*)^\dagger S^*(u)$, for some $u \in H$. Further, by Fact 3.14(4), we see that $z = -P_{\mathcal{N}(S^*)^\perp}(u) \in \mathcal{N}(S^*)^\perp$, which proves $\mathcal{B} \subset \mathcal{N}(S^*)^\perp$ and thus verifies (3.32).

Let

$$\mathcal{D} := \mathcal{N}(S^*) \boxplus \{z \in H \mid z = -(S^*)^\dagger(\rho), \quad \rho \in \mathcal{N}(S)^\perp, \quad \rho \geq 0\}.$$

If we can show that $\mathcal{D} = C^\ominus$, then the last statement of the lemma will follow from this and relation (3.32). Thus to complete the proof, we need to show that $\mathcal{D} = C^\ominus$.

Let $y \in C^\ominus$. For each $j = 1, \dots, m$, let $\rho_j := -\langle y, k_j\rangle$. Since $y \in C^\ominus$, it follows that $\rho_j \geq 0$ and so $\rho \geq 0$. To see that $\rho \in \mathcal{N}(S)^\perp$, take any $\eta \in \mathcal{N}(S)$. Then $S(\eta) = 0$ and

$$\langle \rho, \eta\rangle = \sum_1^m \rho_j \eta_j = -\sum_1^m \langle y, k_j\rangle \eta_j \tag{3.33}$$

$$= -\langle y, \sum_1^m \eta_j k_j\rangle = -\langle y, S(\eta)\rangle = 0. \tag{3.34}$$

Since $\eta \in \mathcal{N}(S)$ was arbitrary, it follows that $\rho \in \mathcal{N}(S)^\perp$.

The definition of ρ_j yields

$$\langle \rho, e_j\rangle = \rho_j = -\langle y, k_j\rangle = -\langle y, S(e_j)\rangle = -\langle S^*(y), e_j\rangle = \langle -S^*(y), e_j\rangle.$$

Since this holds for all the basis vectors e_j, it follows that $\rho = -S^*(y)$. Further, by Fact 3.14(4), we see that $(S^*)^\dagger S^* = P_{\mathcal{N}(S^*)^\perp}$ and hence we can write

$$y = y - (S^*)^\dagger S^*(y) + (S^*)^\dagger S^*(y) = y_0 - (S^*)^\dagger(\rho),$$

where

$$y_0 := [I - (S^*)^\dagger S^*](y) = [I - P_{\mathcal{N}(S^*)^\perp}](y) = P_{\mathcal{N}(S^*)}(y) \in \mathcal{N}(S^*).$$

Thus $y \in \mathcal{D}$ and hence $C^\ominus \subset \mathcal{D}$.

For the reverse inclusion, suppose that $y \in \mathcal{D}$. Then $y = z_0 - (S^*)^\dagger(\rho)$ for some $z_0 \in \mathcal{N}(S^*)$ and $\rho \in \mathcal{N}(S)^\perp$ with $\rho \geq 0$. Then, for each $j = 1, 2, \ldots, m$, we have

$$\begin{aligned}
\langle y, k_j \rangle &= \langle z_0, k_j \rangle - \langle (S^*)^\dagger(\rho), k_j \rangle = \langle z_0, S(e_j) \rangle - \langle (S^*)^\dagger(\rho), k_j \rangle \\
&= \langle S^*(z_0), e_j \rangle - \langle (S^*)^\dagger(\rho), k_j \rangle = -\langle (S^*)^\dagger(\rho), S(e_j) \rangle \\
&= -\langle S^*(S^*)^\dagger(\rho), e_j \rangle = -\langle P_{\mathcal{N}(S)^\perp}(\rho), e_j \rangle \text{ (using Fact 3.14(3))} \\
&= -\langle \rho, e_j \rangle = -\rho_j \leq 0,
\end{aligned}$$

which implies that $y \in C^\ominus$ and hence $\mathcal{D} \subset C^\ominus$. Thus $\mathcal{D} = C^\ominus$ and the proof is complete. □

Based on this lemma, we can now give a detailed description of best approximations from C and $C^\ominus$ to any $x \in H$.

Theorem 3.16. *Let C and S be defined as in equations (3.26) and (3.29). For each $x \in H$, let $x_0 := x - (S^*)^\dagger S^*(x)$. Then $x_0 \in \mathcal{N}(S^*)$ and there exist ρ, $\eta \in \mathbb{R}^m$ such that*

1. $x = S(\rho) + x_0 - (S^*)^\dagger(\eta)$.
2. $\rho \geq 0,\ \eta \geq 0,\ \ \eta \in \mathcal{N}(S)^\perp,$ *and* $\langle \rho, \eta \rangle = 0$.
3. $P_{C^\ominus}(x) = x_0 - (S^*)^\dagger(\eta),$ *and* $\langle x_0, (S^*)^\dagger(\eta) \rangle = 0$.
4. $P_C(x) = S(\rho) = (S^*)^\dagger[S^*(x) + \eta]$.

Proof. Using Fact 3.14(5), we see that

$$S^*(x_0) = S^*(x) - S^*(S^*)^\dagger S^*(x) = S^*(x) - S^*(x) = 0$$

and hence $x_0 \in \mathcal{N}(S^*)$.

By Fact 3.7, we have that $x = P_C(x) + P_{C^\ominus}(x)$ and $\langle P_C(x), P_{C^\ominus}(x) \rangle = 0$. By definition of C, $P_C(x) = S(\rho)$ for some $\rho \in \mathbb{R}^m$ with $\rho \geq 0$. Also, since $P_{C^\ominus}(x) \in C^\ominus$, we use Lemma 3.15 to obtain that $P_{C^\ominus}(x) = y - (S^*)^\dagger(\eta)$ for some $y \in \mathcal{N}(S^*)$ and $-(S^*)^\dagger(\eta) \in \mathcal{N}(S^*)^\perp$ for some $\eta \in \mathcal{N}(S)^\perp$ with $\eta \geq 0$, and $\langle y, (S^*)^\dagger(\eta) \rangle = 0$. We can rewrite this as

$$P_{C^\ominus}(x) = x_0 - (S^*)^\dagger(\eta) + y - x_0. \tag{3.35}$$

Also observe that

$$x - P_{C^\ominus}(x) = \underbrace{(S^*)^\dagger S^*(x) + (S^*)^\dagger(\eta)}_{\in \mathcal{N}(S^*)^\perp} + \underbrace{x_0 - y}_{\in \mathcal{N}(S^*)} \tag{3.36}$$

since $(S^*)^\dagger(\eta) \in \mathcal{N}(S^*)^\perp$ and $S^*(x) \in \mathcal{R}(S^*) = \mathcal{N}(S)^\perp$.

Claim: $y = x_0$.

For if $y \neq x_0$, then by the Pythagorean theorem we obtain

$$\begin{aligned} \|x - P_{C^\ominus}(x)\|^2 &= \|(S^*)^\dagger S^*(x) + (S^*)^\dagger(\eta)\|^2 + \|x_0 - y\|^2 \\ &> \|(S^*)^\dagger S^*(x) + (S^*)^\dagger(\eta)\|^2 = \|x - z\|^2, \end{aligned}$$

where $z := x_0 - (S^*)^\dagger(\eta) \in C^\ominus$. This shows that z is a better approximation to x_0 from $C^\ominus$ than $P_{C^\ominus}(x)$ is, which is absurd and proves the claim.

Thus $P_{C^\ominus}(x) = x_0 - (S^*)^\dagger(\eta)$ and this proves statement (3). Altogether we have that $x = S(\rho) + x_0 - (S^*)^\dagger(\eta)$ and this proves statement (1). Statement (4) follows from (3) and Fact 3.7: $P_C(x) = x - P_{C^\ominus}(x)$. To verify statement (2), it remains to show that $\langle \rho, \eta \rangle = 0$. But

$$\begin{aligned} 0 &= \langle P_C(x), P_{C^\ominus}(x)\rangle = \langle S(\rho), x_0 - (S^*)^\dagger(\eta)\rangle \\ &= \langle \rho, S^*(x_0) - S^*(S^*)^\dagger(\eta)\rangle \\ &= \langle \rho, -P_{\mathcal{N}(S)^\perp}(\eta)\rangle \quad \text{(using } x_0 \in \mathcal{N}(S^*) \text{ and Fact 3.14(3))} \\ &= \langle \rho, -\eta\rangle \quad \text{(since } \eta \in \mathcal{N}(S)^\perp\text{)}. \end{aligned}$$

This completes the proof. □

Remark 3.17. Related to the work of this section, we should mention that Ekárt, Németh, and Németh [14] have suggested a "heuristic" algorithm for computing best approximations from finitely generated cones, in the case where the generators are linearly independent. While they did not prove the convergence of their algorithm, they stated that they numerically solved an extensive set of examples which seemed to suggest that their algorithm was both fast and accurate.

We believe that Theorems 3.9, 3.12, and 3.16 will assist us in obtaining an *efficient* algorithm for the actual computation of best approximations from finitely generated cones in Hilbert space. This will be the subject of a future paper.

3.4. **An Application to Shape-Preserving Approximation.** In this section, we give a class of problems related to "shape-preserving" approximation that can be handled by Theorem 3.9.

Given $x \in L_2[-1,1]$, we want to find its best approximation from the set of polynomials of degree at most n whose rth derivative in nonnegative:

$$C = C_{n,r} := \{p \in \mathcal{P}_n \mid p^{(r)}(t) \geq 0 \text{ for all } t \in [-1,1]\}. \tag{3.37}$$

It is not hard to show that C is a closed convex cone in $L_2[-1,1]$. The interest in such a set is to preserve certain *shape* features of the function being approximated. For example, if $r = 0, 1$, or 2, then C represents all polynomials of degree $\leq n$ that are nonnegative, increasing, or convex, respectively, on $[-1,1]$. It is natural, for example, to want to approximate a convex function in $L_2[-1,1]$ by a *convex* polynomial in $\mathcal{P}_n$

Choose an orthonormal basis $\{p_0, p_1, \ldots, p_n\}$ for $\mathcal{P}_n$. For definiteness, suppose these are the (normalized) Legendre polynomials. The first five Legendre polynomials are given by

(0) $p_0(t) = \frac{\sqrt{2}}{2}$,

(1) $p_1(t) = \frac{\sqrt{6}}{2}t$,

(2) $p_2(t) = \frac{\sqrt{10}}{4}(3t^2 - 1)$,

(3) $p_3(t) = \frac{\sqrt{14}}{4}(5t^3 - 3t)$,

(4) $p_4(t) = \frac{3\sqrt{2}}{16}(35t^4 - 30t^2 + 3)$.

Thus for each $p \in \mathcal{P}_n$ we can write its Fourier expansion as $p = \sum_0^n \langle p, p_i \rangle p_i$.

For each $\alpha \in [-1, 1]$, define

$$k_\alpha := \sum_{i=0}^{n} p_i^{(r)}(\alpha) p_i \tag{3.38}$$

and set

$$K := \{k_\alpha \mid \alpha \in [-1, 1]\}. \tag{3.39}$$

Lemma 3.18. *For each $\alpha \in [-1, 1]$ and $p \in \mathcal{P}_n$, we have*

$$\langle k_\alpha, p \rangle = p^{(r)}(\alpha). \tag{3.40}$$

In other words, k_α is the representer of the linear functional "the rth derivative evaluated at α" on the space $\mathcal{P}_n$.

Proof. Using the orthonormality of the p_i, we get

$$\begin{aligned} \langle k_\alpha, p \rangle &= \left\langle \sum_{i=0}^{n} p_i^{(r)}(\alpha) p_i, \sum_{j=0}^{n} \langle p, p_j \rangle p_j \right\rangle = \sum_{i=0}^{n} \sum_{j=0}^{n} p_i^{(r)}(\alpha) \langle p, p_j \rangle \langle p_i, p_j \rangle \\ &= \sum_{i=0}^{n} \langle p, p_i \rangle p_i^{(r)}(\alpha) = p^{(r)}(\alpha). \end{aligned}$$

□

Lemma 3.19. (1) *K is a compact set in $\mathcal{P}_n$.*

(2) *If $e(t) = t^r$, then $\langle e, k \rangle = r! > 0$ for all $k \in K$.*

(3) *If $C = C_{n,r}$ is defined as in eq.* (3.37)*, then*

$$C = \{p \in \mathcal{P}_n \mid \langle p, k_\alpha \rangle \geq 0 \textit{ for all } \alpha \in [-1, 1]\}. \tag{3.41}$$

Proof. (1) Let (x_m) be a sequence in K. Then there exist $\alpha_m \in [-1, 1]$ such that $x_m = k_{\alpha_m}$ for each m. Since the α_m are bounded, there is a subsequence $\alpha_{m'}$ which converges to some point $\alpha \in [-1, 1]$. Since k_α is a continuous function of α, it follows that $k_{\alpha_{m'}}$ converges to k_α. Thus K is compact.

(2) The rth derivative of t^r is the constant $r!$.

(3) This is an immediate consequence of Lemma 3.18. □

The following result was first proved in the unpublished thesis of the first author [4, Theorem 17].

Theorem 3.20. *Let r, n be integers with $0 \leq r < n$, $X = \mathcal{P}_n$, and*

$$C = C_{n,r} := \{p \in \mathcal{P}_n \mid p^{(r)}(t) \geq 0 \textit{ for all } t \in [-1, 1]\}. \tag{3.42}$$

Let $x \in X \setminus C$, $x_0 \in C$, and let k_α be defined as in (3.38)*. Then the following statements are equivalent:*

(1) *$x_0 = P_C(x)$;*

(2) $x_0 = x + \sum_1^m \rho_i k_{\alpha_i}$, *where* $m \leq n+1$, $\rho_i > 0$, $\alpha_i \in [-1,1]$ *and* $x_0^{(r)}(\alpha_i) = 0$ *for all* i, *and* $\{k_{\alpha_1}, k_{\alpha_2}, \ldots, k_{\alpha_m}\}$ *is linearly independent.*

Moreover, if $x_0^{(r)} \not\equiv 0$ *in any of the statements above, then*

$$m \leq \frac{1}{2}(n-r+2).$$

Proof. The equivalence of the statements (1) and (2) is a consequence of Theorem 3.9 along with Lemmas 3.18 and 3.19.

It remains to show that $m \leq (1/2)(n-r+2)$ when $x_0^{(r)} \not\equiv 0$. Since the vectors k_{α_i} are linearly independent, it follows that $\alpha_1 \ldots, \alpha_m$ are distinct points in $[-1,1]$. Now $x_0^{(r)}$ is a (nonzero) polynomial of degree at most $n-r$, so it has at most $n-r$ zeros. Since $x_0^{(r)}(\alpha_i) = 0$ for $i = 1, \ldots, m$, we must have $m \leq n-r$. If $x_0^{(r)}(\alpha) = 0$ for some α with $-1 < \alpha < 1$, then α cannot be a simple zero of $x_0^{(r)}$ (i.e., $x_0^{(r+1)}(\alpha) = 0$ also) since $x_0^{(r)}(t) \geq 0$ for all $-1 \leq t \leq 1$. It follows that $x_0^{(r)}$ can have at most $1/2(n-r)$ zeros in the open interval $(-1,1)$. If $x_0^{(r)}$ has a zero at one of the end points $t = \pm 1$, then $x_0^{(r)}$ can have at most $1 + 1/2(n-r-1) = 1/2(n-r+1)$ zeros in $[-1,1]$. Finally, if $x_0^{(r)}$ has zeros at both end points $t = \pm 1$, then we see that $x_0^{(r)}$ has at most $2 + 1/2(n-r-2) = 1/2(n-r+2)$ zeros in $[-1,1]$. In all possible cases, we see th at $x_0^{(r)}$ has at most $m \leq 1/2(n-r+2)$ zeros in $[-1,1]$. □

4. Elements vanishing relative to a set

Definition 4.1. Let X be a normed linear space and $\Gamma \subset X^*$. An element $x^* \in X^*$ is said to **vanish relative to** Γ if $x \in X$ and $y^*(x) = 0$ for all $y^* \in \Gamma$ imply that $x^*(x) = 0$.

Again, when X is a Hilbert space, the above definition reduces to the following form.

Definition 4.2. Let X be a Hilbert space and $\Gamma \subset X$. An element $x \in X$ is said to **vanish relative to** Γ if $z \in X$ and $\langle z, y\rangle = 0$ for all $y \in \Gamma$ imply that $\langle x, z\rangle = 0$.

This idea can be characterized in a useful way just as "positive relative to a set" was in Theorem 2.4.

Theorem 4.3. *Let* X *be a normed linear space,* $\Gamma \subset X^*$*, and* $x^* \in X^*$*. Then the following statements are equivalent:*

(1) x^* *vanishes relative to* Γ.
(2) $\Gamma_\perp \subset (x^*)_\perp$.
(3) $x^* \in w^*\!-cl(\mathrm{span}\,(\Gamma))$, *the weak* closed linear span of* Γ.

Moreover, if X *is reflexive, then each of these statements is equivalent to*

(4) $x^* \in \overline{\mathrm{span}}\,(\Gamma)$, *the (norm) closed linear span of* Γ.

Proof. (1) ⇒ (2). Suppose (1) holds. Let $x \in \Gamma_\perp$. Then $y^*(x) = 0$ for all $y^* \in \Gamma$. By (1), $x^*(x) = 0$. That is, $x \in (x^*)_\perp$. Hence (2) holds.

(2) ⇒ (3). If (3) fails, $x^* \notin w^*\!-cl(\mathrm{span}\,(\Gamma))$. By Theorem 2.1, there exists a weak* continuous linear functional f on X^* such that

$$\sup f(\mathrm{span}\,(\Gamma)) < f(x^*). \tag{4.1}$$

But $f = \hat{x}$ for some $x \in X$. Thus we can rewrite the inequality (4.1) as $\sup \hat{x}(\text{span}(\Gamma)) < \hat{x}(x^*)$, or

$$\sup\{y^*(x) \mid y^* \in \text{span}(\Gamma)\} < x^*(x). \tag{4.2}$$

Since $\text{span}(\Gamma)$ is a linear subspace, the only way the expression on the left side of (4.2) can be bounded above is if $y^*(x) = 0$ for each $y^* \in \Gamma$. In this case, it follows that $x^*(x) > 0$. Thus $x \in \Gamma_\perp \setminus (x^*)_\perp$ and (2) fails.

(3) $\Rightarrow$ (1). Let $x^* \in w^*\!-cl(\text{span}(\Gamma))$. If $x \in X$ and $y^*(x) = 0$ for all $y^* \in \Gamma$, then clearly $y^*(x) = 0$ for all $y^* \in \text{span}(\Gamma)$. Since $x^* \in w^*\!-cl(\text{span}(\Gamma))$, there exists a net $(y^*_\alpha) \in \text{span}(\Gamma)$ such that y^*_α weak* converges to x^*, i.e., $y^*_\alpha(z) \to x^*(z)$ for each $z \in X$. But $y^*_\alpha(x) = 0$ for all α, so $x^*(x) = 0$. That is, x^* vanishes relative to Γ and (1) holds.

If X is reflexive, then the same proof as in Theorem 2.4 works. □

Corollary 4.4. *Let X be a normed linear space, $\Gamma \subset X^*$, and $x^* \in X^*$. If x^* is positive relative to Γ, then x^* vanishes relative to Γ.*

Proof. By Theorem 2.4, we have $x^* \in w^*\!-cl(\text{cone}(\Gamma))$. A fortiori, $x^* \in w^*\!-cl(\text{span}(\Gamma))$. By Theorem 4.3, the result follows.

A simpler, more direct, proof goes as follows. For any subset S of X^*, we have $S_\ominus \cap (-S_\ominus) = S_\perp$. Hence if $\Gamma_\ominus \subset (x^*_\ominus)$, then $-\Gamma_\ominus \subset -(x^*_\ominus)$ and hence

$$\Gamma_\perp = \Gamma_\ominus \cap [-\Gamma_\ominus] \subset (x^*)_\ominus \cap [-(x^*)_\ominus] = (x^*)_\perp$$

In other words, using Theorem 4.3, x^* vanishes relative to Γ. □

The following simple example shows that the converse to this theorem is not valid.

Example 4.5. Let $X = \mathbb{R}$ with the absolute-value norm: $\|x\| := |x|$. Let $x = -1$ and $\Gamma = \{1\}$. Then x vanishes relative to Γ, but x is not positive relative to Γ.

Again, by the same argument as in Theorem 2.4, we note that in a reflexive Banach space X, a convex set in X^* is (norm) closed if and only if it is weak* closed. Thus we have the following result.

Corollary 4.6. *Let X be a reflexive Banach space, $\Gamma \subset X^*$, and $x^* \in X^*$. Then x^* vanishes relative to Γ if and only if $x^* \in \overline{\text{span}}(\Gamma)$, the (norm) closed linear span of Γ.*

Corollary 4.7. *Let H be a Hilbert space, $\Gamma \subset H$, and $x \in H$. Then x vanishes relative to Γ if and only if $x \in \overline{\text{span}}(\Gamma)$, the norm closed linear span of Γ.*

One important application of Theorem 4.3 is the following.

Lemma 4.8. *Let X be a normed linear space and $f, f_1, \ldots, f_n$ be in X^*. Then the following statements are equivalent:*

(1) *f vanishes relative to $\Gamma = \{f_1, f_2, \ldots, f_n\}$.*
(2) *If $x \in X$ and $f_i(x) = 0$ for each $i = 1, 2, \ldots, n$, then $f(x) = 0$.*
(3) *$\cap_{i=1}^n f_i^{-1}(0) \subset f^{-1}(0)$.*
(4) *$f \in \text{span}\{f_1, f_2, \ldots, f_n\}$.*

Proof. The equivalence of (1) and (2) is just a rewording of the definition, and the equivalence of (2) and (3) is obvious. Finally, (1) holds if and only if f lies in the weak* closed linear span of Γ by Theorem 4.3. But the linear span of Γ, being finite-dimensional, is weak* closed (see, eg., [8, Corollary 3.14]). That is, (4) holds. $\square$

This result—in even more general vector spaces—has proven useful in studying weak topologies on vector spaces (see, e.g., [7, p. 421] or [8, Lemma 3.21]).

References

[1] N. Aronszajn, *Introduction to the Theory of Hilbert Spaces, Vol. 1,* Research Foundation of Oklahoma A & M College, Stillwater, OK, 1950.
[2] E. W. Cheney, *Introduction to Approximation Theory*, McGraw-Hill, 1966.
[3] J. B. Conway, *A Course in Functional Analysis* (second edition), Springer-Verlag, New York, 1990.
[4] F. Deutsch, *Some applications of functional analysis to approximation theory*, Doctoral dissertation, Brown University, Providence, Rhode Island, July, 1965,
[5] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer-Verlag, New York, 2001.
[6] F. Deutsch and P. H. Maserick, *Applications of the Hahn-Banach theorem in approximation theory*, SIAM Review **6** (1967),516–530.
[7] N. Dunford and J. T. Schwartz, *Linear Operators, Part I: General Theory*, Interscience Publishers, New York, 1958.
[8] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler, *Banach Space Theory*, Springer, New York, 2011.
[9] J. Farkas, *Über die Theorie der einfachen Ungleichungen*, J. für die Reine und Angewandte Mathematik **124** (1902), 1–27.
[10] David Gale, Harold Kuhn, and Albert Tucker, *Linear programming and the theory of games*, Chapter 12 in Koopmans, *Activity Analysis of Production and Allocation*, Wiley, 1951.
[11] C. W. Groetsch, *Generalized Inverses of Linear Operators*, Marcel Dekker, Inc., New York, 1977.
[12] J-B. Hiriart-Urruty and Claude Lemaréchal, *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, New York, 1993.
[13] J. J. Moreau, *Décomposition orthogonale dans un espace hilbertien selon deux cônes mutuallement polaires*, C. R. Acad. Sci., Paris **255** (1962), 238–240.
[14] A. Ekárt, A. B. Németh, and S. Z. Németh, *Rapid heuristic projection on simplicial cones*, arXiv: 1001.1928v2 [math.OC] 19 Jan 2010.
[15] F. Riesz, *Zur Theorie des Hilbertschen Raumes,* Acta Sci. Math. Szeged, **7** (1934), 34–38.
[16] V. Tchakaloff, *Formules de cubatures mecaniques a coefficients non negatifs*, Bull. Sci. Math. **81** (1957), 123–134.

Manuscript received December 24 2017
revised February 22 2018

F. Deutsch
Department of Mathematics, The Pennsylvania State University, University Park, PA, 16802, USA
E-mail address: deutsch@math.psu.edu

H. Hundal
146 Cedar Ridge, Port Matilda, PA, 16870, USA
E-mail address: hundalhh@yahoo.com

L. Zikatanov
Department of Mathematics, The Pennsylvania State University, University Park, PA, 16802, USA
E-mail address: ltz@math.psu.edu