JOURNAL OF APPROXIMATION THEORY 71, 213–238 (1992)

Constrained Best Approximation in Hilbert Space, II

CHARLES K. CHUI,*,† FRANK DEUTSCH,‡ AND JOSEPH D. WARD*,†

† *Department of Mathematics, Texas A & M University, College Station, Texas 77843, U.S.A.; and*
‡ *Department of Mathematics, Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.*

Communicated by Allan Pinkus

Received August 13, 1990; accepted September 18, 1991

The problem considered is that of characterizing the best approximation, to a given x in a Hilbert space, from a set which is the intersection of a closed convex cone and a closed linear variety. This problem is shown to be equivalent to the (generally much simpler) problem of characterizing best approximations to a certain perturbation of x from the cone alone (or a subcone of the cone). Several applications to shape-preserving interpolation are given. © 1992 Academic Press, Inc.

1. INTRODUCTION

This work represents a continuation of our previous paper [5]. As pointed out there, constrained approximation problems arise when an approximant to a specified problem is required to preserve certain shapes such as positivity, monotonicity, and/or convexity. Moreover, in various specific formulations, this problem is often posed in data analysis, computer-aided geometric design, and mathematical modeling. Many such problems may be formulated as the study of existence, uniqueness, charcterization, and computational aspects of the solution to the extremal problem

$$\inf\{\|x\| \mid x \in C \cap A^{-1}(d)\},$$

where x is in a Hilbert space X, C is a closed convex cone that defines the constraint, A is a bounded linear operator from X into a finite-dimensional Hilbert space Y, and d is a given "data" vector in Y.

The objective of this paper is twofold. Our primary goal is to characterize the minimum norm interpolant, for *arbitrary* admissible data, under

* Supported by the National Science Foundation under Grants DMS-8602337 and DMS-8901345.

0021-9045/92 $5.00
Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

the assumption that C is a closed convex cone. We had resolved this problem in [5] in the important special case when d is in the *interior* of the data cone $A(C)$. There we showed that the minimum norm interpolant can be characterized as a best approximation in C to a linear combination of the finitely many vectors which define the interpolation conditions (see also Theorem 2.3 below). In particular, this leads to an easier approximation problem in C (rather than in $C \cap A^{-1}(d)$) which involves only *finitely* many parameters. Moreover, algorithms were developed (e.g., in [15, 18]) that applied to this situation. Here we show that an analogous result holds when d is in the boundary of the data cone except that, in this situation, the characterization involves a certain *subcone* of C rather than C itself (see Theorem 2.6 below). In particular, this minimization problem involves only *finitely* many parameters, despite the fact that the constraint set $C \cap A^{-1}(d)$ is infinite-dimensional in general. These subcones depend upon the minimal face of the data cone which contains the data vector. If this face is exposed, the subcone is very often readily computable. Examples are given to illustrate many of these points.

The second objective is to extend the results of [5] and [18] in order to characterize the minimum norm solution in case the set C described above is an arbitrary closed convex set, not necessarily a cone. Our characterization can be obtained provided that the data vector is an interior point of the data set $A(C)$.

A review of an earlier draft of this paper pointed out a large body of results related to our work, many of which had been motivated, at least in part, by [17]. Indeed, the problem

$$\inf\{\|x\| \mid x \in C,\ Ax = d\}$$

can easily be cast as a certain optimization problem whose "optimal" solution can be studied via Fenchal duality [2, 3, 13, 16, 19].

Interpolation problems have also been considered from a control theory perspective, providing another set of results [10]. In particular a practical motivation for the study of such constrained problems arises in connection with L_2 spectral estimation [12].

This paper consists of four sections. Following the Introduction, the general characterization of the minimal norm interpolant from a flat intersecting a convex cone is derived in Section 2 (Theorem 2.6). Some of these results have appeared in the context of optimization theory [2, 3, 16], where criteria for the minimal norm interpolant are given in terms of subdifferentials and intersections of conex [3] as well as subcones of C [4]. Our criteria describes the minimal norm interpolant of x in terms of the metric projection of a perturbation of x onto a subcone C_F of C. In particular this formulation seems more convenient for approximation theorists and, in addition, has led to some numerical algorithms seeking the optimal solution [10, 15]. Also in Proposition 2.8, we show that the subcone C_F may be described by finitely many parameters in case the minimal face of the data cone containing d is exposed. The notion of exposed faces in optimization to study certain minimization problems has already arisen [3, 16], but the usage here (Proposition 2.8 and the example following) appears to be new.

Perhaps the main results of the paper occur in Section 3. In our opinion, the three most important cones that arise in applications are the cones of positive functions, increasing functions, and convex functions. For these examples, in the case of boundary data, the subcones are shown to be related to certain subsets of the underlying interval. While this has already been recognized in case of the positive cone [17, 5, 2], these results are new for the cones of increasing and convex functions. Examples are given to illustrate the theory.

Finally, Section 4 deals with the minimization problem described above in the case that C is an arbitrary closed convex set. The main results in this section are Theorems 4.6 and 4.7.

2. Best Approximation from Constrained Cones

Throughout this section, X and Y will denote fixed Hilbert spaces with Y being finite-dimensional. Let A be a bounded linear operator from X into Y, C a closed convex *cone* in X, $d \in AC := \{Ax \mid x \in C\}$, and

$$K = C \cap A^{-1}(d) = \{x \in C \mid Ax = d\}.$$

Then K is a nonempty closed convex set in X, and consequently is a Chebyshev set. That is, each $x \in X$ has a unique best approximation $P_K(x)$ in K. Our problem in this section is to obtain "useful" *characterizations* of $P_K(x)$.

In [5] it was seen that "property CHIP" played an important role in this regard. For the pair of sets $\{C, A^{-1}(d)\}$, property CHIP is equivalent to the statement that

$$\overline{\text{con}}[(C-y) \cap A^{-1}(0)] = \overline{\text{con}}(C-y) \cap A^{-1}(0) \qquad (*)$$

for each $y \in C \cap A^{-1}(d) = K$, where "$\overline{\text{con}}$" denotes the "closed conical hull of". Clearly, $C - y \subset \overline{C-C}$ and $\overline{C-C}$ is a closed subspace in X. Hence, for the purpose of verifying whether or not $\{C, A^{-1}(d)\}$ has property CHIP, we may assume without loss of generality that $X = \overline{C-C}$.

Since A is continuous, it follows that

$$AC - AC \subset A(\overline{C-C}) \subset \overline{A(C-C)} = \overline{AC - AC} = AC - AC,$$

where the last equality holds due to the fact that $AC - AC$ is a subspace (since AC is a cone) in a finite-dimensional space, and is therefore closed. Thus,

$$A\overline{(C-C)} = AC - AC = \text{aff } AC,$$

where aff AC denotes the affine hull of AC, which, in turn, is the linear span of AC since AC is a convex cone.

The upshot of all this is as follows. To verify that $\{C, A^{-1}(d)\}$ has property CHIP, we may assume that $X = \overline{C - C}$ and that $R(A)$, the range of A, is aff AC.

Further, by replacing Y with $R(A)$, we may also assume that A is surjective and thus $Y = \text{aff } AC$. In [5; Lemma 3.1] we showed that if $d \in \text{int } AC$, then $\{C, A^{-1}(d)\}$ has property CHIP. But $d \in \text{int } AC$ if and only if $d \in \text{int } AC$ relative to aff AC; or equivalently, $d \in \text{ri } AC$, where ri AC denotes the relative interior of AC.

The next lemma is now immediate.

LEMMA 2.1. *If* $d \in \text{ri } AC$, *then* $\{C, A^{-1}(d)\}$ *has property* CHIP.

We recall the following well-known characterization of best approximations from convex cones (see e.g. [5; Proposition 2.1]).

LEMMA 2.2. *Let* $x \in X$ *and* $y \in C$. *Then* $y = P_C(x)$ *if and only if*

$$x - y \in (C - y)^0 = C^0 \cap y^{\perp}.$$

Here S^0 denotes the dual cone of S. That is,

$$S^0 = \{y \in X \mid \langle x, y \rangle \leqslant 0 \text{ for all } x \in S\}.$$

Now we can state the first of the two main results of this section. It characterizes best approximations from K when d is a relative interior point of AC.

THEOREM 2.3. *If* $d \in \text{ri } AC$, *then for each* $x \in X$, *there exists a* $y_0 \in Y$ *such that*

$$A[P_C(x + A^*y_0)] = d. \tag{2.3.1}$$

Moreover,

$$P_K(x) = P_C(x + A^*y) \tag{2.3.2}$$

for any $y \in Y$ *satisfying* (2.3.1).

Proof. Let $M=\overline{C-C}$, so that M is a closed subspace of X which contains C and hence also K.

The proof of the case $X=M$ is virtually identical to that of Theorem 3.2 of [5] so we omit it.

Now assume $X \neq M$. Using the first part of the proof (applied to $P_M(x)$ instead of x), we obtain that, for each $x \in X$,

$$P_K(P_M(x)) = P_C(P_M(x) + A^*y) \tag{2.3.3}$$

for any $y \in Y$ which satisfies

$$A[P_C(P_M(x) + A^*y)] = d. \tag{2.3.4}$$

Next we recall the well-known facts that P_M is linear, $P_M^2 = P_M$, and

$$P_C = P_C \circ P_M, \qquad P_K = P_K \circ P_M \tag{2.3.5}$$

(cf. [6]). Using these facts, we deduce that

$$\begin{aligned} P_C(P_M(x) + A^*y) &= P_C[P_M(P_M(x) + A^*y)] \\ &= P_C[P_M(x) + P_M(A^*y)] \\ &= P_C[P_M(x + A^*y)] = P_C(x + A^*y). \end{aligned}$$

That is,

$$P_C(P_M(x) + A^*y) = P_C(x + A^*y). \tag{2.3.6}$$

Hence, from (2.3.3)–(2.3.6), we obtain

$$P_K(x) = P_K(P_M(x)) = P_C(P_M(x) + A^*y) = P_C(x + A^*y)$$

for any $y \in Y$ which satisfies

$$A[P_C(x + A^*y)] = d.$$

This proves (2.3.1) and (2.3.2). ∎

Now we turn to the case when $d \notin \operatorname{ri} AC$. Here we will find that there are results analogous to Theorem 2.3 except that C must now be replaced by a certain *subcone* of C.

DEFINITION 2.4. Let F denote the minimal convex extremal subset of AC which contains d, and let

$$C_F := C \cap A^{-1}(F).$$

Some of the following statements have already been observed in the literature [3, 19].

LEMMA 2.5. *The following statements hold:*

(1) *F is a convex cone.*

(2) *C_F is an extremal convex cone in C.*

(3) $AC_F = F$.

(4) $d \in \operatorname{ri} AC_F$.

(5) $K = C_F \cap A^{-1}(d)$.

(6) $\{C_F, A^{-1}(d)\}$ *has property* CHIP.

(7) *If* $d \in \operatorname{ri} AC$, *then* $C_F = C$.

Proof. (1) Let $y \in F$ and $0 < \lambda < 1$. Then $(1-\lambda)0 + \lambda(\lambda^{-1}y) = y \in F$ and $0, \lambda^{-1}y \in AC$ imply $0, \lambda^{-1}y \in F$ since F is extremal. For each $\rho > 1$, set $\lambda = \rho^{-1} \in (0, 1)$ and note $\rho y = \lambda^{-1}y \in F$. Thus, $\rho y \in F$ for each $\rho \geqslant 1$. If $0 \leqslant \rho \leqslant 1$, then the convexity of F implies that

$$\rho y = \rho y + (1-\rho)0 \in F.$$

Thus $\rho y \in F$ for any $\rho \geqslant 0$, and hence F is a cone.

(2) C_F is a convex cone since the inverse image, $A^{-1}(F)$, of a convex cone is a convex cone.

To see C_F is extremal in C, let x, $y \in C$, $0 < \lambda < 1$, and $\lambda x + (1-\lambda)y \in C_F$. Then $\lambda x + (1-\lambda)y \in A^{-1}(F)$ implies $\lambda Ax + (1-\lambda)Ay \in F$ and Ax, $Ay \in AC$ which implies that Ax and Ay are in F by the extremality of F. Thus x, $y \in A^{-1}(F)$, so that x, $y \in C \cap A^{-1}(F) = C_F$; that is, C_F is extremal.

(3) Let $A_C := A|_C$. Then $A_C\colon C \to AC$ is surjective so that

$$AC_F = A_C C_F = A_C[C \cap A^{-1}(F)] = A_C[A_C^{-1}(F)] = F.$$

(4) By (3), it suffices to show that $d \in \operatorname{ri} F$. Suppose $d \in F \backslash \operatorname{ri} F$. We will work in the finite-dimensional Hilbert space span $F = \operatorname{aff} F$. Then

$$\operatorname{int} F = \operatorname{ri} F \neq \varnothing.$$

By the Eidelheit separation theorem, there exists a $z \in \operatorname{span} F \backslash \{0\}$ (cf. [19]) such that

$$\sup_{y \in \operatorname{int} F} \langle z, y \rangle \leqslant \langle z, d \rangle.$$

Since F is a convex cone by (1) and int F is dense in F (cf. [19]), it follows that

$$\sup_{y \in F} \langle z, y \rangle = 0 = \langle z, d \rangle.$$

If $z \in F^{\perp}$, then $z \in (\text{span } F)^{\perp} = \{0\}$ which contradicts $z \neq 0$. Thus $\langle z, y_0 \rangle < 0$ for some $y_0 \in F$. Set

$$H = \{ y \in \text{span } F \mid \langle y, z \rangle = 0 \}.$$

This H is a closed hyperplane (in span F) which supports F at d. Thus $E := H \cap F$ is a convex extremal subset of F which does not equal F (since $y_0 \in F \backslash E$). But this contradicts the minimality of F.

(5) Since $C_F \subset C$, we have

$$C_F \cap A^{-1}(d) \subset C \cap A^{-1}(d) = K.$$

Conversely, if $x \in K$, then $x \in C$ and $x \in A^{-1}(d) \subset A^{-1}(F)$. Thus $x \in C \cap A^{-1}(F) = C_F$.

(6) This follows from Lemma 2.1 (with C replaced by C_F).

(7) Suppose $d \in \text{ri } AC$. Since $d \in \text{ri } AC_F = \text{ri } F$ by (3) and (4), and since AC and F are both convex extremal subsets of AC, it follows that $AC = F$ (cf. [20]). Thus

$$C_F = C \cap A^{-1}(F) = C \cap A^{-1}(AC) \supset C \cap C = C \supset C_F$$

implies $C_F = C$. ∎

It is perhaps worth noting that parts of Lemma 2.5 are still valid if F is replaced by any extremal convex subset E of AC which contains d. Denoting $C_E := C \cap A^{-1}(E)$, one can verify that

(1) E is a convex cone,

(2) C_E is an extremal convex cone in C,

(3) $AC_E = E$,

(4) $K = C_E \cap A^{-1}(d)$, and

(5) If $d \in \text{int } AC$, then $C_E = C$.

The following main result of this section, which generalizes Theorem 2.3, is a consequence of Theorem 2.3 and Lemma 2.5.

THEOREM 2.6. *For each $x \in X$,*

$$P_K(x) = P_{C_F}(x + A^* y) \tag{2.6.1}$$

for any $y \in Y$ *which satisfies*

$$A[P_{C_F}(x + A^*y)] = d. \tag{2.6.2}$$

Moreover, if $d \in \operatorname{ri} AC$, *then* $C_F = C$.

Proof. Since $K = C_F \cap A^{-1}(d)$ and $d \in \operatorname{ri} AC_F$ by Lemma 2.5, we apply Theorem 2.3 (with C replaced by C_F) to obtain (2.6.1) and (2.6.2). Also, if $d \in \operatorname{ri} AC$, then $C_F = C$ by Lemma 2.5. ∎

In practice, it may be difficult or impractical to determine the minimal face F of AC which contains d, and then the subcone $C_F = C \cap A^{-1}(F)$.

One approach that greatly helps to identify C_F is the case that F is an exposed face.

DEFINITION 2.7. A face E of a convex set D is called an *exposed* face if $E = D \cap H$ for some closed hyperplane H.

Remarks. (i) In case the convex set D is a cone, the hyperplane H contains the origin [20]; i.e., $H = \ker \lambda$ for some continuous linear functional λ.

(ii) If $E = D \cap \ker \lambda$ and $x \in D \backslash E$, then $\lambda(x) \neq 0$. In fact, $\lambda(x)$ has the same sign for all $x \in D \backslash E$.

Relative to our situation, recall that F is the minimal convex extremal subset of AC which contains d. We assume that $Y = l_2(n)$ and hence

$$Ax = (\langle x, x_1 \rangle, ..., \langle x, x_n \rangle), \qquad x \in X,$$

for some linearly independent set $\{x_1, x_2, ..., x_n\}$ in X. Suppose that F is an exposed face of AC. Then $F = AC \cap \ker \lambda$ for some linear functional λ on Y. Then we can identify λ with an element $(\lambda_1, \lambda_2, ..., \lambda_n) \in l_2(n)$. Set

$$C_\lambda = \left\{ x \in C \,\middle|\, \left\langle x, \sum_1^n \lambda_i x_i \right\rangle = 0 \right\}.$$

PROPOSITION 2.8. *If* F *is an exposed face of* AC, *then* $C_F = C_\lambda$.

Proof. $x \in C_F$ if and only if $x \in C$ and $Ax \in F$ if and only if $x \in C$ and $(\langle x, x_1 \rangle, ..., \langle x, x_n \rangle) \in F$ if and only if $x \in C$ and $\langle x, \sum_1^n \lambda_i x_i \rangle = \sum_1^n \lambda \langle x, x_i \rangle = 0$ if and only if $x \in C_\lambda$. ∎

As an application we prove, for the case $p = 2$, Theorem 2.1 of [17] which was stated but not proved there. (The authors of [17] omitted the proof since certain details were quite technical). That is, we discuss the spline problem which, after standard recasting, corresponds to

$$\min \left\{ \|x\| \mid x \in L_2[0, 1], x \geqslant 0, \text{ and } \int_0^1 xM_i = d_i,\ i = 1, ..., n-k \right\},$$

where $M_i = M_{i,k}$ is the kth order B-spline supported on $[t_i, t_{i+k}]$, and $0 \leqslant t_1 < \cdots < t_n \leqslant 1$.

Let $\Omega = \{i \mid d_i > 0\}$ so that $\Omega^c = \{i \mid d_i = 0\}$.

CLAIM 1. *d is an interior point relative to $AC \cap \{(x_1, ..., x_{n-k}) \mid x_i = 0$ if $i \in \Omega^c\}$.*

Proof. Note that for some $g_0 \geqslant 0$ a.e., $d = (\int_0^1 g_0 M_1, ..., \int_0^1 g_0 M_{n-k})$. Let $\rho_N = \{x \in [0, 1] \mid g_0(x) \geqslant 1/N\}$ and $j \in \Omega$. Then for sufficiently large N, there exist points $\{t_i \mid i \in \Omega,\ t_i \in \rho_N\}$ so that

$$M_j(t_j) > 0, \qquad j \in \Omega.$$

By the Schoenberg–Whitney theorem (cf. [1]), the matrix $[M_i(t_j)]$ is invertible; i.e., $\{M_j \mid j \in \Omega\}$ is a linearly independent set over $\{t_j \mid j \in \Omega\}$ so that $\{M_j \mid j \in \Omega\}$ is linearly independent over ρ_N for sufficiently large N.

So for sufficiently small $\varepsilon > 0$ and any vector $\hat{d}$ with $\|\hat{d}\| < \varepsilon$, there exist scalars $\{a_i\}$ that satisfy

$$g_1 := g_0 + \left(\sum_{j \in \Omega} a_i M_i\right) \chi_{\rho_N} \text{ interpolates } d + \hat{d} \text{ with } g_1 \geqslant 0 \text{ a.e.}$$

Next, let $R = \{(x_1, ..., x_{n-k}) \mid x_i \geqslant 0,\ x_i = 0 \text{ if } i \in \Omega^c\}$ and let

$$F = R \cap AC.$$

CLAIM 2. *F is the minimal face of AC containing d.*

Proof. We establish this by showing that any face E containing d also contains F. Let y be any element of F. For sufficiently small λ, we have

$$\frac{1}{1-\lambda} d - \frac{\lambda}{1-\lambda} y \in \text{ball}(d, \varepsilon) \subset AC,$$

so that

$$d = \lambda y + (1-\lambda)\left[\frac{1}{1-\lambda} d - \frac{\lambda}{1-\lambda} y\right]$$

with $y \in AC$ and $1/(1-\lambda)d - \lambda/(1-\lambda)\, y \in AC$. Since E is a face containing d, we have $y \in E$. Since y is arbitrary in F, it follows that $F \subset E$ and so F is minimal.

We next establish that F is exposed. Let $\lambda(x) := \langle x, u \rangle$ where $u = (u_1, ..., u_{n-k})$ with

$$u_i = \begin{cases} 1 & \text{if } i \in \Omega^c \\ 0 & \text{if } i \in \Omega \end{cases}$$

and $H = \ker \lambda$. Since, under the current hypotheses, $d \in AC$ implies $d_i \geq 0$, it follows that $H \cap C = R \cap C = F$. By Theorem 2.6 and Proposition 2.8, the solution has the form

$$P_{C_F}\left(\sum_1^{n-k} a_i x_i\right) = P_{C_\lambda}\left(\sum_1^{n-k} a_i x_i\right),$$

where $C_\lambda = \{x \in C \mid \langle x, \sum_1^{n-k} u_i M_i \rangle = 0\}$. Since x and $\sum u_i M_i$ are nonnegative, x must be zero on $\bigcup_{j=1}^{n-k} \{(t_j, t_{j+k}) \mid d_j = 0\}$.

An example illustrating certain boundary data relative to the cone of increasing functions will be given in the next section.

As another approach that allows one to avoid the computation of C_F directly, it is often possible to seek subcones C_Ω of C and bounded linear operators $B: X \to Y$ such that $K = C_\Omega \cap B^{-1}(d')$, where $d' \in \operatorname{int} BC_\Omega$. The next theorem governs this situation.

THEOREM 2.9. *Let C_Ω be a closed convex subcone of C, B a bounded linear operator from X into a finite-dimensional Hilbert space Z, and $d' \in Z$ be such that $K = C_\Omega \cap B^{-1}(d')$ and $d' \in \operatorname{ri} BC_\Omega$. Then for any $x \in X$, there exists a $z_0 \in Z$ for which*

$$B[P_{C_\Omega}(x + B^* z_0)] = d'. \tag{2.9.1}$$

In addition,

$$P_K(x) = P_{C_\Omega}(x + B^* z) \tag{2.9.2}$$

for any $z \in Z$ satisfying (2.9.1).

Proof. This result follows directly from Theorem 2.3 by replacing A, Y, C, and d with B, Z, C_Ω, and d', respectively. ∎

3. BEST CONSTRAINED APPROXIMATION FROM THE POSITIVE, INCREASING, AND CONVEX CONES

In this section we give three applications of the results in Section 2. The first application concerns the case when $X = L_2(\mu)$, C is the cone of positive functions in X, and A is a bounded linear operator from X into $l_2(n)$. The second application corresponds to the case $X = L_2(a, b)$, C is the cone of increasing functions in X, and A is a bounded linear operator from X into $l_2(n)$. The third case deals with the cone of convex functions. The first application was originally proved (for a special case) in [17] using variational methods, and also in [5] using an approach based on property CHIP. Our reason for including it here is to show how it can be deduced from Theorem 2.9. The second and third applications of Theorem 2.9 seem to be new. Throughout this section, various properties will be associated with equivalence classes of functions. This will mean that some representative within a given equivalence class possesses that property. For example, $x \in L_2(\mu)$ and $x \geqslant 0$ means there is some function x_0 where $x_0(t) \geqslant 0$ for every t and $x_0(t) = x(t)$ a.e. (μ).

Example (The Cone of Positive Functions)

We first deal with the cone of positive functions. Let $(T, \mathscr{S}, \mu)$ be a measure space, $L_2(\mu) := L_2(T, \mathscr{S}, \mu)$, and we assume that μ is chosen so that $L_2(\mu)$ is a separable Banach space. Furthermore, let

$$C = \{x \in L_2(\mu) \mid x \geqslant 0 \text{ on } T\},$$

$x_i \in L_2(\mu)$, $d_i \in \mathbb{R}$ $(i = 1, 2, ..., n)$, and

$$K = \{x \in C \mid \langle x, x_i \rangle = d_i \ (i = 1, 2, ..., n)\}.$$

Assume $K \neq \varnothing$. It is no loss of generality to assume that $\{x_1, x_2, ..., x_n\}$ is linearly independent. Defining A on $L_2(\mu)$ by

$$Ax = (\langle x, x_1 \rangle, \langle x, x_2 \rangle, ..., \langle x, x_n \rangle),$$

we see that A is a bounded linear operator from $L_2(\mu)$ onto $l_2(n)$ and

$$K = C \cap A^{-1}(d).$$

Moreover, $A^*: l_2(n) \to X$ is given by

$$A^*y = \sum_1^n y(i)\, x_i, \qquad y = (y(1), y(2), ..., y(n)).$$

Let $\tilde{K}$ be a countable dense subset of K and let Ω_k and Ω be the subsets of T defined by

$$\Omega_k := \{t \in T \mid k(t) > 0\}$$

$$\Omega := \bigcup_{k \in \tilde{K}} \Omega_k.$$

That is, Ω is the (countable) union of the supports of the elements in $\tilde{K}$ and hence is measurable. It is easily established that for any $k \in K$, $\Omega_k = \Omega \cap \Omega_k$ a.e. (μ).

For any measurable subset S of T, we define

$$C_S := \{y \in C \mid y = 0 \text{ on } T \backslash S\}.$$

Note that $C = C_T$ and $K = C_\Omega \cap A^{-1}(d)$.

LEMMA 3.1. *The following statements hold*:

(1) *C_S is a closed convex subcone of C.*

(2) *$C_S^0 = \{y \in L_2(\mu) \mid y \leqslant 0$ on $S\}$.*

(3) *For each $x \in L_2(\mu)$, $P_{C_S}(x) = x_+ \chi_S$, where $x_+ = \max\{x, 0\}$. In particular, $P_C(x) = x_+$.*

Proof. Statements (1) and (2) are obvious. To verify (3), we fix any $x \in L_2(\mu)$ and $y \in C_S$. Then, by Lemma 2.2, $y = P_{C_S}(x)$ if and only if $x - y \in C_S^0 \cap y^\perp$. Equivalently, $x - y \leqslant 0$ on S and $\int_T (x-y)\, y\, d\mu = 0$. Since $y \geqslant 0$ and $y = 0$ off S, this statement is equivalent to $x \leqslant y$ on S and $y(t) = 0$ whenever $t \in S$ and $x(t) < y(t)$. That is, $y = x_+ \chi_S$. In particular, taking $S = T$, we obtain $C_S = C$ and the result follows. ∎

THEOREM 3.2. *For any $x \in L_2(\mu)$, there exist scalars $\alpha_1, ..., \alpha_n$ such that*

$$\left\langle \left(x + \sum_1^n \alpha_i x_i\right)_+ \chi_\Omega, x_j \right\rangle = d_j \qquad (j = 1, 2, ..., n). \tag{3.2.1}$$

In addition,

$$P_K(x) = \left(x + \sum_1^n \beta_i x_i\right)_+ \chi_\Omega \tag{3.2.2}$$

for any choice of scalars β_i chosen to satisfy (3.2.1). *Moreover, if the set $\{x_1, x_2, ..., x_n\}$ is linearly independent over Ω, then the factor χ_Ω may be deleted from* (3.2.1) *and* (3.2.2).

The proof is similar to that of Theorem 3.4 below which we prove in detail. Theorem 3.2 was first proved by Micchelli *et al.* [17] (in the special case when $x = 0$) using variational methods. The result as stated here, using a similar approach, was established in [5]. It was also shown in [17] that the characteristic function χ_Ω cannot be deleted, in general, from (3.2.1) and (3.2.2).

Example (*The Cone of Increasing Functions*)

We next deal with the case where the cone consists of increasing functions. Let $I = (a, b)$ and let $L_2(I)$ denote the space of square-integrable Lebesgue measurable functions on I,

$$C = \{x \in L_2(I) \mid x \text{ is increasing on } I\},$$

$x_i \in L_2(I)$, $d_i \in \mathbb{R}$ $(i = 1, 2, ..., n)$, and $K = \{x \in C \mid \langle x, x_i \rangle = d_i$ $(i = 1, 2, ..., n)\}$. Increasing functions are real-valued on I, but may be unbounded. Assume $K \neq \emptyset$. It is no loss of generality to assume that $\{x_1, x_2, ..., x_n\}$ is linearly independent. Defining A on $L_2(I)$ by

$$Ax = (\langle x, x_1 \rangle, \langle x, x_2 \rangle, ..., \langle x, x_n \rangle),$$

we see that A is a bounded linear map from $L_2(I)$ onto $l_2(n)$ and

$$K = C \cap A^{-1}(d).$$

Moreover, $A^*: l_2(n) \to L_2(I)$ is given by

$$A^*y = \sum_1^n y(i)\, x_i, \qquad y = (y(1), ..., y(n)).$$

For any $y \in C$, let μ_y denote the *Lebesgue–Stieltjes measure* induced on the Borel sets in I by y. By modifying y on a set of Lebesgue measure zero, we may assume that y is continuous from the right on I. Thus

$$\mu_y(E) = \inf\left\{\sum_1^\infty [y(b_i) - y(a_i)] \,\middle|\, E \subset \bigcup_1^\infty (a_i, b_i)\right\}$$

for any Borel set $E \subset I$. The *support* of any Borel measure μ on I is the (relatively) closed set

$$\operatorname{supp} \mu := I \backslash U_\mu,$$

where

$$U_\mu := \bigcup \{U \mid U \text{ open}, \mu(U) = 0\}.$$

That is, $\operatorname{supp} \mu$ is the smallest (relatively) closed set in I such that μ vanishes on its complement.

For any measurable subset $S \subset I$, define

$$C_S := \{y \in C \mid \mu_y(I \backslash S) = 0\}.$$

In particular, $C = C_I$. For any $x \in L_2(I)$, define the "indefinite integral" of x by

$$x^{[1]}(t) := \int_a^t x(s)\, ds, \qquad t \in [a, b].$$

Note that $x^{[1]}$ is continuous on $[a, b]$ and $x^{[1]}(a) = 0$. A function $y \in L_2(I)$ is said to be *constant* on a subset S of I provided that y is continuous at each point of S and $y(t_1) = y(t_2)$ for all points t_1, t_2 in S.

The following lemma is from [8].

LEMMA 3.3. (1) *C_S is a closed convex subcone of C*

(2) $C_S^0 = \{y \in L_2(I) \mid y^{[1]}(b) = 0,\ y^{[1]}(t) \geqslant 0 \text{ on } S\}$

(3) *If $x \in L_2(I)$ and $y \in C_S$, then $y = P_{C_S}(x)$ if and only if the following conditions hold*:

(i) $y^{[1]}(b) = x^{[1]}(b)$,

(ii) $y^{[1]}(t) \leqslant x^{[1]}(t)$ *for all $t \in S$, and*

(iii) *if $y^{[1]}(s) < x^{[1]}(s)$ for some $s \in S$, then y is constant in the "component" of S (i.e., the largest open interval) which contains s.*

(4) *If $x \in L_2(I)$ and $y \in C$, then $y = P_C(x)$ if and only if $y^{[1]}$ is the greatest convex minorant of $x^{[1]}$. (That is, $y^{[1]}$ is the pointwise supremum of all convex functions which are below $x^{[1]}$.)*

Next we define a subset Ω of I as follows: set

$$\Omega = \overline{\bigcup \{\operatorname{supp} \mu_y \mid y \in K\}}.$$

Then Ω is closed, hence measurable, $C_\Omega = \{y \in C \mid \mu_y(I \backslash \Omega) = 0\}$, and $K = \{y \in C_\Omega \mid \langle y, x_i \rangle = d_i,\ i = 1, 2, ..., n\}$.

For each i, let x_i^* denote the functional on X whose representer is x_i. That is,

$$x_i^*(x) := \langle x, x_i \rangle, \qquad x \in X.$$

Now we can state the main characterization theorem governing this application.

THEOREM 3.4. *For each $x \in L_2(I)$, there exist scalars $\alpha_{1,0}, ..., \alpha_{n,0}$ such that*

$$\left\langle P_{C_\Omega}\left(x + \sum_1^n \alpha_{i,0} x_i\right), x_j \right\rangle = d_j \qquad (j = 1, 2, ..., n). \tag{3.4.1}$$

In addition,

$$P_K(x) = P_{C_\Omega}\left(x + \sum_1^n \alpha_i x_i\right) \tag{3.4.2}$$

for any set of scalars α_i chosen to satisfy (3.4.1). *Moreover, if $\{x_1^*, x_2^*, ..., x_n^*\}$ is linearly independent over C_Ω, then C_Ω may be replaced by C in* (3.4.1) *and* (3.4.2).

LEMMA 3.5. $C_\Omega^0 \cap K^\perp = C_\Omega^\perp$.

Proof. Let $z \in C_\Omega^0 \cap K^\perp$. Then by Lemma 3.3, $z^{[1]}(b)=0$, and $z^{[1]}(t) \geqslant 0$ for all $t \in \Omega$. For any bounded $y \in C_\Omega$, an integration by parts [11, p. 154] yields

$$\langle z, y\rangle = \int_a^b zy = y(b^-)\, z^{[1]}(b) - y(a^+)\, z^{[1]}(a) - \int_a^b z^{[1]}\, d\mu_y$$

$$= -\int_a^b z^{[1]}\, d\mu_y = -\int_\Omega z^{[1]}\, d\mu_y.$$

Now suppose $y \in C_\Omega$ is unbounded. Let

$$y_n(t) = \begin{cases} n & \text{if } \; y(t) > n \\ y(t) & \text{if } \; |y(t)| \leqslant n \\ -n & \text{if } \; y(t) < -n. \end{cases}$$

Then $y_n \in C_\Omega$, y_n is bounded, $|y_n| \leqslant |y|$ for all n, and $y_n(t) \to y(t)$ for all t. Moreover, using standard arguments, we deduce that

$$\int_a^b f\, d\mu_y = \lim_n \int_a^b f\, d\mu_{y_n} \tag{3.4.3}$$

for each nonnegative μ_y-measurable function f. Hence the above argument yields

$$\langle z, y_n\rangle = -\int_\Omega z^{[1]}\, d\mu_{y_n} \qquad \text{for each } n. \tag{3.4.4}$$

Since $z^{[1]}$ is continuous and nonnegative on Ω, we obtain from (3.4.3) that

$$\lim_n \left[-\int_\Omega z^{[1]}\, d\mu_{y_n} \right] = -\int_\Omega z^{[1]}\, d\mu_y. \tag{3.4.5}$$

Also, the dominated convergence theorem implies that

$$\lim_n \langle z, y_n\rangle = \langle z, y\rangle. \tag{3.4.6}$$

Combining (3.4.4)–(3.4.6), we get

$$\langle z, y\rangle = -\int_\Omega z^{[1]}\, d\mu_y \tag{3.4.7}$$

for each $y \in C_\Omega$.

Then for each $y \in K$, it follows that (since $z \in K^\perp$)

$$0 = \langle z, y \rangle = -\int_\Omega z^{[1]}\, d\mu_y.$$

Since $z^{[1]} \geqslant 0$ on Ω, it follows that $z^{[1]} = 0$ a.e. (μ_y) on Ω for each $y \in K$. Thus $z^{[1]} = 0$ on $\bigcup \{\text{supp}\, \mu_y \mid y \in K\}$. Since $z^{[1]}$ is continuous, $z^{[1]} = 0$ on Ω. Using (3.4.7), we obtain that $\langle z, y \rangle = 0$ for each $y \in C_\Omega$. That is, $z \in C_\Omega^\perp$.

We have shown that $C_\Omega^0 \cap K^\perp \subset C_\Omega^\perp$. For the reverse inclusion, note that $C_\Omega^\perp \subset C_\Omega^0$ and since $C_\Omega \supset K$, $C_\Omega^\perp \subset K^\perp$. Thus $C_\Omega^\perp \subset C_\Omega^0 \cap K^\perp$. This completes the proof. ∎

Proof of Theorem 3.4. If $x_i^*(y) = 0$ for all $y \in C_\Omega$ and all i, then

$$\begin{aligned} K &= \{ y \in C_n \mid \langle y, x_i \rangle = 0 \ (i = 1, 2, ..., n)\} \\ &= C_\Omega. \end{aligned}$$

Claim. $P_{C_\Omega}(x) = P_{C_\Omega}(x + \sum_1^n \alpha_i x_i)$ for every set of scalars $\alpha_1, ..., \alpha_n$.

For let $y = P_{C_\Omega}(x)$. Then (by Lemma 2.2) $x - y \in C_\Omega^0 \cap y^\perp$. For any scalars $\alpha_1, ..., \alpha_n$, we see that

$$\sum_1^n \alpha_i x_i \in y^\perp \cap C_\Omega^\perp \subset C_\Omega^0 \cap y^\perp$$

so $x + \sum_1^n \alpha_i x_i - y \in C_\Omega^0 \cap y^\perp$ also since $C_\Omega^0 \cap y^\perp$ is a convex cone. Hence (again by Lemma 2.2) $y = P_{C_\Omega}(x + \sum_1^n \alpha_i x_i)$. This proves the claim.

The claim shows that the theorem holds when $x_i^*(y) = 0$ for all $y \in C_\Omega$ and all $i = 1, 2, ..., n$.

Thus we can assume $x_i^*(y) \neq 0$ for some i and some $y \in C_\Omega$. By reindexing, we may let $\{x_1^*, ..., x_m^*\}$, $1 \leqslant m \leqslant n$, be a maximal subset of $\{x_1^*, ..., x_n^*\}$ which is linearly independent over C_Ω. Then for any $i = m+1, ..., n$, there are scalars $\lambda_1, ..., \lambda_m$ so that

$$x_i^*(y) = \sum_{j=1}^m \lambda_j x_j^*(y) \qquad \text{for all} \quad y \in C_\Omega.$$

In particular, for $y \in K$,

$$\begin{aligned} d_i = \langle y, x_i \rangle = x_i^*(y) &= \sum_{j=1}^m \lambda_j x_j^*(y) \\ &= \sum_{j=1}^m \lambda_j \langle y, x_j \rangle = \sum_{j=1}^m \lambda_j d_j. \end{aligned}$$

It follows that

$$\begin{aligned} K &= \{ y \in C_\Omega \mid \langle y, x_j \rangle = d_j \ (j = 1, 2, ..., n) \} \\ &= \{ y \in C_\Omega \mid \langle y, x_j \rangle = d_j \ (j = 1, 2, ..., m) \} \\ &= C_\Omega \cap B^{-1}(d'), \end{aligned}$$

where $B: L_2(I) \to l_2(m)$ is defined by

$$Bx := (\langle x, x_1 \rangle, \langle x, x_2 \rangle, ..., \langle x, x_m \rangle)$$

and

$$d' = (d_1, d_2, ..., d_m) \in l_2(m).$$

Claim. $C_\Omega^0 \cap K^\perp \cap \mathscr{R}(B^*) = \{0\}$.

By Lemma 3.4, it suffices to show that

$$C_\Omega^\perp \cap \mathscr{R}(B^*) = \{0\}. \tag{3.4.8}$$

Suppose $z \in C_\Omega^\perp \cap \mathscr{R}(B^*)$. Then $z = \sum_1^m \beta_i x_i$ and $\langle z, y \rangle = 0$ for all $y \in C_\Omega$. Thus, for all $y \in C_\Omega$,

$$\begin{aligned} \sum_1^m \beta_i x_i^*(y) &= \sum_1^m \beta_i \langle y, x_i \rangle = \left\langle y, \sum_1^m \beta_i x_i \right\rangle \\ &= \langle y, z \rangle = 0. \end{aligned}$$

Since $\{x_1^*, ..., x_m^*\}$ is linearly independent over C_Ω, $\beta_1 = \cdots = \beta_m = 0$. That is, $z = 0$. This proves the claim.

By [5], $d' \in \operatorname{int} BC_\Omega$ and the result follows by applying Theorem 2.9.

Finally, if $\{x_1^*, ..., x_n^*\}$ is linearly independent over C_Ω, the proof of the last claim shows that $C_\Omega^0 \cap K^\perp \cap \mathscr{R}(A^*) = \{0\}$. Since $C_\Omega \subset C$, it follows that $C_\Omega^0 \supset C^0$ and hence $C^0 \cap K^\perp \cap \mathscr{R}(A^*) = \{0\}$. By [5], $d \in \operatorname{int} AC$ and the result follows by Theorem 2.6. ∎

An Example of Theorem 3.4

As an example of the above theory, let $0 = t_1 < \cdots < t_n = 1$ be n arbitrarily space points on [0, 1] and let $M_i = M_{i,k}$, $i = 1, ..., n-k$, be the corresponding normalized B-splines. If C corresponds to the cone of increasing functions, then any possible data sequence d is clearly increasing. Moreover, if $d_i = d_{i+1}$ for some i, then x (the increasing function which interpolates d) must be constant on $\operatorname{supp} M_i \cup \operatorname{supp} M_{i+1} = [t_i, t_{i+k+1}]$ and such data vectors d must also be boundary data.

In light of this, consider the minimization problem

$$\min\left\{\|y\| \mid y\in L_2[0,1],\ y \text{ is increasing, and } \int_0^1 yM_i=d_i,\ i=1,...,n-k\right\},$$

where $d_i<d_{i+1}$ if $i=1, ..., n-k-2$, and $d_{n-k-1}=d_{n-k}$.

We also assume that the problem is feasible and that there is a C^1 interpolate y_0 to d which satisfies $\text{meas}\{\text{supp}\, y_0'\cap[t_i, t_{i+k}]\}>0$ for $i=1, ..., n-k-2$. The C^1 condition insures that the integration by parts formula applies to $\int_0^1 y_0 M_i$. Under these assumptions, one can show that d is an interior point to $AC\cap\{c\in R^{n-k} \mid c_{n-k-1}=c_{n-k}\}$. More specifically, one can show, using arguments similar to those following Theorem 2.6, that the set $\{M_1, ..., M_{n-k-2}\}$ is linearly independent over

$$\Omega_N := \{x\in[0,1] : y_0'(x)\geqslant 1/N\}$$

for some N. Hence if $M_i^{[1]}(x) := \int_0^x M_i(t)\,dt$, the set $\{M_1^{[1]}, ..., M_{n-k-2}^{[1]}\}$ is also linearly independent over Ω_N. So for sufficiently small ε_i, one finds a ρ,

$$\rho := \left(\sum_1^{n-k-2} a_i M_i\right)\chi_{\Omega_N},$$

for which $\int_0^1 \rho M_i^{[1]} = -\varepsilon_i$, $i=1, ..., n-k-2$.

Also, since $\Omega_N\subset[0, t_{n-k-2}]$, $\rho\equiv 0$ on $[t_{n-k-1}, 1]$. Hence if $\rho^{[1]}(x) := \int_0^x \rho(t)\,dt$, we have

$$\int_0^1 (g_0+\rho^{[1]})\,M_i = d_i + \int_0^1 \rho^{[1]} M_i = d_i+\varepsilon_i+\rho^{[1]}(1)$$

for all i. So for an appropriate constant c, $y_0+\rho^{[1]}+c$ is an increasing function which interpolates any point in a sufficiently small neighborhood about d. As before it follows that

$$AC\cap\{c\in R^{n-k} \mid c_{n-k-1}=c_{n-k}\}$$

is a minimal face containing d as a relative interior point. Also if $\lambda(d)=\langle u, d\rangle$, where

$$u_i=\begin{cases}1 & i=n-k-1\\ -1 & i=n-k\\ 0 & \text{elsewhere,}\end{cases}$$

then for any $d \in AC$, $\lambda(d) = 0$ if and only if $d_{n-k-1} = d_{n-k}$ since d is an increasing vector. Thus, by Theorem 2.6, we find that the minimal norm solution is given by $P_{C_\lambda}(\sum_1^{n-k} u_i M_i)$ where

$$\begin{aligned} C_\lambda &= \{ y \mid y \text{ increasing}, \langle y, M_{n-k-1} - M_{n-k} \rangle = 0 \} \\ &= \{ y \mid y \text{ increasing}, y \text{ constant on } [t_{n-k-1}, 1] \}. \end{aligned}$$

Example (*The Cone of Convex Functions*)

We now turn our attention to the cone of convex functions. Let $I = (a, b)$,

$$C = \{ x \in L_2(I) \mid x \text{ is convex on } I \},$$

$x_i \in L_2(I)$, $d_i \in \mathbb{R}$ $(i = 1, 2, ..., n)$, and

$$K = \{ x \in C \mid \langle x, x_i \rangle = d_i \ (i = 1, 2, ..., n) \}.$$

Assume $K \neq \emptyset$. It is no loss of generality to assume that $\{x_1, x_2, ..., x_n\}$ is linearly independent. Defining A on $L_2(I)$ by

$$Ax = (\langle x, x_1 \rangle, \langle x, x_2 \rangle, ..., \langle x, x_n \rangle),$$

we see that A is a bounded linear operator from $L_2(I)$ onto $l_2(n)$ and

$$K = C \cap A^{-1}(d),$$

where $d = (d_1, d_2, ..., d_n)$.

For any $y \in C$, the derivative y' exists, is continuous, and is increasing except on a countable set. By defining y' on this countable set by taking right-hand limits, we may assume that y' is defined on all I. Just as in a previous example, let $\mu_{y'}$ denote the Lebesgue–Stieltjes measure induced by y', and let supp $\mu_{y'}$ denote the support of $\mu_{y'}$.

Given any measurable subset S of I, define

$$C_S = \{ y \in C \mid \mu_{y'}(I \backslash S) = 0 \}.$$

In particular, $C = C_I$. The first and second indefinite integrals of any $x \in L_2(I)$ are defined on I by

$$x^{[1]}(t) = \int_a^t x(s)\,ds, \qquad x^{[2]}(t) = \int_a^t x^{[1]}(s)\,ds.$$

The following lemma is from [9].

LEMMA 3.6. (1) *C_S is a closed convex subcone of C.*

(2) $C_S^0 = \{x \in L_2(I) \mid x^{[1]}(b) = x^{[2]}(b) = 0,\ x^{[2]} \leqslant 0 \text{ on } S\}$

(3) *Let $x \in L_2(I)$ and $y \in C_S$. Then $y = P_{C_S}(x)$ if and only if*

(i) $y^{[1]}(b) = x^{[1]}(b),\ y^{[2]}(b) = x^{[2]}(b)$,

(ii) $y^{[2]}(s) \geqslant x^{[2]}(s)$ *for all $s \in S$, and*

(iii) *if $y^{[2]}(s) > x^{[2]}(s)$ for some $s \in S$, then y is linear in the component of S which contains s.*

Next, as before, let $\tilde{K}$ be a countable dense subset of K and let $\Omega \subset I$ be given by

$$\Omega = \bigcup \{\text{supp } \mu_{y'} \mid y \in \tilde{K}\}.$$

Then Ω is measurable,

$$C_\Omega = \{y \in C \mid \mu_{y'}(I \backslash \Omega) = 0\},$$

and

$$K = \{y \in C_\Omega \mid \langle y_i, x_i \rangle = d_i \ (i = 1, 2, ..., n)\}.$$

The theorem characterizing best approximations from K can be stated as follows:

THEOREM 3.7. *For each $x \in L_2(I)$, there exist scalars $\alpha_{1,0}, ..., \alpha_{n,0}$ such that*

$$\left\langle P_{C_\Omega}\left(x + \sum_1^n \alpha_{i,0} x_i\right), x_j \right\rangle = d_j \qquad (j = 1, 2, ..., n). \tag{3.7.1}$$

In addition,

$$P_K(x) = P_{C_\Omega}\left(x + \sum_1^n \alpha_i x_i\right) \tag{3.7.2}$$

for any set of scalars α_i chosen to satisfy (3.7.1). *Moreover, if $\{x_1^*, x_2^*, ..., x_n^*\}$ is linearly independent over C_Ω, then C_Ω may be replaced by C in* (3.7.1) *and* (3.7.2).

The proof is similar to Theorem 3.4 and is omitted. In fact, there are analogues of Theorems 3.4 and 3.7 for the convex cone of N-convex functions, $N = 1, 2, ...,$ which were inspired by these examples (see [9]).

4. Best Interpolation from a Convex Set

The main results in this section are Theorems 4.6 and 4.7 which extend to arbitrary closed convex sets certain results obtained in [5] for convex cones. These theorems also constitute a generalization of [18, Theorem 2.2]. Our results are then applied to an example which was derived in [15] using different methods. In this section X and Y will denote Hilbert spaces, C any closed convex *set* in X, A a bounded linear operator from X into Y, $d \in Y$, and

$$K = C \cap A^{-1}(d) = \{x \in X \mid x \in C, Ax = d\}.$$

THEOREM 4.1. *Suppose* $\{C, A^{-1}(d)\}$ *has property* CHIP. *Then for any* $x \in X$ *and* $k_0 \in K$, *the following statements are equivalent.*

(1) $k_0 = P_K(x)$;

(2) $x - k_0 \in \overline{(C - k_0)^0 + R(A^*)}$, *where* $R(A^*)$ *denotes the range of the adjoint of* A.

The same proof as given in [5, Theorem 2.1] works here. There it was assumed that C is a closed convex cone. In that case, we also get that $(C - k_0)^0 = C^0 \cap k_0^{\perp}$.

LEMMA 4.2. *Consider the following statements*:

(1) d *is a "Slater point"*: $d \in A(\text{int } C)$ (*equivalently*, $(\text{int } C) \cap A^{-1}(d) \neq \emptyset$)

(2) $d \in \text{int } AC$;

(3) $\{C, A^{-1}(d)\}$ *has property* CHIP.

Then (2) $\Rightarrow$ (3). *In addition, if* A *is surjective, then* (1) $\Rightarrow$ (2).

The proof of this lemma is exactly the same as in [5, Lemma 3.1] where C was assumed to be a closed convex cone.

DEFINITION 4.3. Let D be a closed convex cone and M a closed subspace in a Hilbert space. The *inclination* between D and M is defined by

$$i(D, M) := \inf\{\|x - y\| \mid x \in D, y \in M, \|x\| = \|y\| = 1\}.$$

The *cosine* of the angle between D and M is defined by

$$c(D, M) := \sup\{|\langle x, y\rangle| \mid x \in D, y \in M, y \in M, \|x\| = \|y\| = 1\}.$$

LEMMA 4.4. *Let D be a closed convex cone and M a closed subspace in a Hilbert space X. Then the following statements hold*:

(1) $0 \leqslant i(D, M) \leqslant 2$.

(2) $0 \leqslant c(D, M) \leqslant 1$.

(3) *$i(D, M) > 0$ if and only if $c(D, M) < 1$.*

(4) *If $i(D, M) > 0$ (or $c(D, M) < 1$), then $D \cap M = \{0\}$ and $D + M$ is closed.*

(5) *If both D and M are closed subspaces, then $D \cap M = \{0\}$ and $D + M$ is closed if and only if $i(D, M) > 0$ if and only if $c(D, M) < 1$.*

Proof. (1) and (2) are obvious.

(3) For all x, y in X with $\|x\| = \|y\| = 1$, we have

$$\|x - y\|^2 = \|x\|^2 - 2\langle x, y\rangle + \|y\|^2 = 2[1 - \langle x, y\rangle]. \tag{4.4.1}$$

From (4.4.1) we deduce that $i(D, M) = 0$ if and only if $c(D, M) = 1$. Equivalently, $i(D, M) > 0$ if and only if $c(D, M) < 1$.

(4) Let $c = c(D, M) < 1$. Then $D \cap M = \{0\}$ and

$$|\langle x, y\rangle| \leqslant c\,\|x\|\,\|y\|$$

for all $x \in D$, $y \in M$. Let $z_n \in D + M$ and $z_n \to z$. We must show that $z \in D + M$. Now $z_n = x_n + y_n$ for some $x_n \in D$, $y_n \in M$. Since $\{z_n\}$ converges, it must be bounded, so that for some $\rho > 0$,

$$\begin{aligned}
\rho \geqslant \|z_n\|^2 &= \|x_n\|^2 + 2\langle x_n, y_n\rangle + \|y_n\|^2 \\
&\geqslant \|x_n\|^2 - 2\,|\langle x_n, y_n\rangle| + \|y_n\|^2 \\
&\geqslant \|x_n\|^2 - 2c\,\|x_n\|\,\|y_n\| + \|y_n\|^2 \\
&= (\|x_n\| - \|y_n\|)^2 + 2(1-c)\,\|x_n\|\,\|y_n\|.
\end{aligned}$$

Since $c < 1$, it follows that both sequences $\{\|x_n\| - \|y_n\|\}$ and $\{\|x_n\|\,\|y_n\|\}$ are bounded. From this we deduce that $\{\|x_n\|\}$ and $\{\|y_n\|\}$ are bounded. By passing to a subsequence if necessary, we may assume that $x_n \to x$ weakly and $y_n \to y$ weakly. Since D and M are weakly closed, we have $x \in D$ and $y \in M$. Thus $z_n = x_n + y_n \to x + y$ weakly. But $z_n \to z$, so that $z = x + y \in D + M$.

(5) Assume that both D and M are closed subspaces. Using (3) and (4), it suffices to verify that if $D \cap M = \{0\}$ and $D + M$ is closed, then $c(D, M) < 1$ (we may assume $D + M = X$). If the result were false, then $c(D, M) = 1$ and there exist $x_n \in D$, $y_n \in M$ with $\|x_n\| = \|y_n\| = 1$ such that

$\langle x_n, y_n \rangle \to 1$. By (4.4.1), $\|x_n - y_n\| \to 0$. Hence if Q denotes the projection of $D+M$ onto M along D, then

$$-y_n = Q(x_n - y_n) \to Q(0) = 0,$$

which contradicts $\|y_n\| = 1$. ∎

Remark. The proof of Lemma 4.4 is modeled after the analogous one in [7, Lemma 2.5] where both D and M were subspaces.

LEMMA 4.5. *Let $d \in AC$.*

(1) *If $d \in \operatorname{int} AC$, then*

$$(C-k)^0 \cap R(A^*) = \{0\} \tag{4.5.1}$$

for each $k \in C \cap A^{-1}(d)$.

(2) *If A is surjective and $d \in \operatorname{int} AC$, then* (4.5.1) *holds and $(C-k)^0 + R(A^*)$ is closed for each $k \in C \cap A^{-1}(d)$.*

(3) *If A is surjective and* $\operatorname{int} AC \neq \emptyset$, *then $d \in \operatorname{int} AC$ if and only if* (4.5.1) *holds for all $k \in C \cap A^{-1}(d)$.*

(4) *If Y is finite-dimensional and* (4.5.1) *holds, then $(C-k)^0 + R(A^*)$ is closed for all $k \in C \cap A^{-1}(d)$,*

Proof. (1) If $d \in \operatorname{int} AC$, then d cannot be separated from AC by a closed hyperplane. Thus for each $k \in C \cap A^{-1}(d)$, we have $\{y \in Y \mid \langle y, Ac \rangle \leqslant \langle y, d \rangle$ for all $c \in C\} = \{0\}$, which is equivalent to $\{y \mid \langle y, A(c-k) \rangle \leqslant 0$ for all $c \in C\} = \{0\}$, or $\{y \mid \langle A^*y, c-k \rangle \leqslant 0$ for all $c \in C\} = \{0\}$; that is, $\{y \mid A^*y \in (C-k)^0\} = \{0\}$. Hence, it follows that $\{A^*y \mid A^*y \in (C-k)^0\} = \{0\}$, or equivalently, $(C-k)^0 \cap R(A^*) = \{0\}$. This proves (4.5.1).

(2) Assume that A is surjective, $d \in \operatorname{int} AC$, and $k \in C \cap A^{-1}(d)$. By part (1), (4.5.1) holds. Since $R(A) = Y$ is closed, $R(A^*)$ is also closed. If the result were false, then by Lemma 4.4(4) (with $M = R(A^*)$ and $D = (C-k)^0$), there are sequences $\{x_n\}$ in $(C-k)^0$ and $\{z_n\}$ in $R(A^*)$ so that $\|x_n\| = \|z_n\| = 1$ and $\|x_n - z_n\| \to 0$. Since $R(A^*)$ is closed, the Open Mapping Theorem implies that there is a constant $R > 0$ and a sequence $\{y_n\}$ in Y so that $z_n = A^*y_n$ and $\|y_n\| \leqslant R\,\|z_n\| = R$. Moreover,

$$\rho := \frac{1}{\|A^*\|} = \frac{1}{\|A^*\|}\,\|z_n\| \leqslant \|y_n\|.$$

That is, $0 < \rho \leqslant \|y_n\| \leqslant R$ for all n. For each $c \in C$, since $x_n \in (C-k)^0$, we have

$$\begin{aligned}\langle A^*y_n, c-k \rangle &= \langle z_n, c-k \rangle \\ &= \langle z_n - x_n, c-k \rangle + \langle x_n, c-k \rangle \\ &\leqslant \|z_n - x_n\|\,\|c-k\|.\end{aligned}$$

Hence,

$$\langle y_n, A(c-k)\rangle = \langle A^*y_n, c-k\rangle \leq r' \|z_n - y_n\|$$

for each $c \in C$ with $c-k \in (C-k) \cap B(0, r')$, where $r' > 0$ is chosen so that $B(0, \delta) \subset A[(C-k) \cap B(0, r')]$ for some $\delta > 0$. The existence of δ and r' is guaranteed by the Baire Category Theorem [5, Lemma 3.1]. It follows that $\langle y_n, y'\rangle \leq r' \|z_n - x_n\|$ for all $y' \in B(0, r')$. This, in turn, implies that

$$0 < \delta'\rho \leq \delta' \|y_n\| = \sup\{\langle y_n, y'\rangle \mid y' \in B(0, \delta')\} \leq r' \|z_n - x_n\| \to 0,$$

which is absurd.

(3) Assume that A is surjective, int $AC \neq \varnothing$, and (4.5.1) holds. By (1), it suffices to verify that $d \in \operatorname{int} AC$. Now A^* is injective (as a consequence of the well-known relations: $\mathcal{N}(A^*) = \mathcal{N}(A^*)^{\perp\perp} = \overline{R(A)}^{\perp} = Y^{\perp} = \{0\}$). Hence $y = 0$ if and only if $A^*y = 0$. Thus the implications in the proof of (1) are reversible and we obtain that (4.5.1) holds. That is, d cannot be separated from AC by a closed hyperplane. Since int $AC \neq \varnothing$, this latter condition is equivalent to $d \in \operatorname{int} AC$.

(4) Assume that Y is finite-dimensional and (4.5.1) holds. Since $(C-k)^0$ is a closed convex cone and $R(A^*)$ is a finite-dimensional subspace, it follows from the "Dieudonne Separation Theorem" [14, p. 105] that $(C-k)^0 + R(A^*)$ is closed. ∎

We are now ready to state the two main results of this section.

THEOREM 4.6. *Let $\{C, A^{-1}(d)\}$ have property* CHIP *and suppose that for each $k \in K$, $(C-k)^0 + R(A^*)$ is closed. Then for every $x \in X$,*

$$P_K(x) = P_C(x + A^*y) \tag{4.6.1}$$

for every $y \in Y$ chosen so that

$$A[P_C(x + A^*y)] = d. \tag{4.6.2}$$

Proof. Let $x \in X$ and $k_0 \in K$. By Theorem 4.1, $k_0 = P_K(x)$ if and only if $x - k_0 \in \overline{(C-k_0)^0 + R(A^*)} = (C - k_0)^0 + R(A^*)$ if and only if $x + A^*y - k_0 \in (C-k_0)^0$ for some $y \in Y$ if and only if $k_0 = P_C(x + A^*y)$ for some $y \in Y$ (see [5, Theorem 2.1]). This verifies (4.6.1). The last statement follows by observing that (4.6.2) guarantees that the element $k_0 = P_C(x + A^*y)$ is in K. ∎

THEOREM 4.7. *Let $d \in \operatorname{int} AC$. If A is either a surjection or Y is finite-dimensional, then for any $x \in X$,*

$$P_K(x) = P_C(x + A^*y) \tag{4.7.1}$$

for any $y \in Y$ chosen so that

$$A[P_C(x+A^*y)]=d. \tag{4.7.2}$$

Proof. By Lemma 4.2, $\{C, A^{-1}(d)\}$ has property CHIP. By Lemma 4.5, $(C-k)^0+R(A^*)$ is closed for each $k \in K$. The result now follows from Theorem 4.6. ▮

Using variational methods, Micchelli and Utreras [18] proved a more restricted version of Theorem 4.7 under the more stringent conditions that d is a "Slater point" (i.e., int $C \cap A^{-1}(d) \neq \varnothing$), A is surjective, and $x=0$. (To verify that these conditions are more stringent, see Lemma 4.2). As an application of the above results we recall a result of [15] which had applications to shape preservation of data interpolation.

THEOREM [15]. *Let $l \geqslant 0$ and assume that there is an admissible y^* such that $\{\phi_1, ..., \phi_N\}$ is linearly independent over the support of (y^*-l). Then the unique solution y_0 to*

$$\inf\left\{\int_a^b y^2\,dt \mid y(t) \geqslant l(t) \text{ and } \int_a^b y\phi_i\,dt = d_i,\ i=1, ..., N\right\}$$

is characterized by

$$y_0=\left(\sum_{j=1}^N \alpha_j\phi_j - l\right)_+ + l,$$

where the coefficients $\alpha_1, ..., \alpha_N$ are determined by the requirement that y_0 satisfy the interpolation conditions.

Note that if $l \neq 0$, the constrained set is convex but not a convex cone (it is a translate of a cone). Moreover the constrained set has no interior point and hence, no Slater point condition holds, and the results of [18] do not apply.

Relative to our setting, it is easy to see that the linear independence condition above implies that the data point d is an interior point of the data cone. Additionally, y_0 is just $P_C(\sum_1^N \alpha_j\phi_j)$, which follows from [15, Lemma 3.1] or which can be verified directly. Hence the above theorem may be viewed as a consequence of Theorem 4.7.

REFERENCES

1. C. DE BOOR, "A Practical Guide to Splines," Springer-Verlag, New York, 1978.
2. J. M. BORWEIN AND A. S. LEWIS, Partially finite convex programming, *Math. Programming* (1991), to appear (in two parts); preprint, Dalhousie University (1988).

3. J. M. BORWEIN AND H. WOLKOWICZ, Facial reduction for a cone-convex programming problem, *J. Austral. Math. Soc. Ser. A* **30** (1981), 369–380.
4. J. M. BORWEIN AND H. WOLKOWICZ, A simple constraint qualification in infinite dimensional programming, *Math. Programming* **35** (1986), 83–96.
5. C. K. CHUI, F. DEUTSCH AND J. D. WARD, Constrained best approximation in Hilbert space, *Constr. Approx.* **6** (1990), 35–64.
6. F. DEUTSCH, "Some Applications of Functional Analysis to Approximation Theory," Ph.D. Dissertation, Brown University, 1966.
7. F. DEUTSCH, Rate of convergence of the method of alternating projections, *in* "Parametric Optimization and Approximation" (B. Brosowski and F. Deutsch, Eds.), ISNM, Vol. 72, pp. 96–107, Birkhäuser-Verlag, Basel, 1984.
8. F. DEUTSCH AND V. UBHAYA, Constrained isotone approximation, in progress.
9. F. DEUTSCH AND V. UBHAYA, Constrained n-convex approximation, in progress.
10. A. L. DONTCHEV, Duality methods for constrained best interpolation, *Math. Balkanica* **1** (1987), 96–105.
11. N. DUNFORD AND J. SCHWARTZ, "Linear Operators I," Interscience, New York, 1958.
12. B. K. GOODRICH AND A. STEINHARDT, L_2 spectral estimation. *SIAM J. Appl. Math.* **46** (1986), 417–428.
13. M. S. GOWDA AND M. TEBOULLE, A comparison of constraint qualifications in infinite-dimensional convex programming, *SIAM J. Control Optim.* **28** (1990), 925–935.
14. R. B. HOLMES, "Geometric Functional Analysis and Its Applications," Springer-Verlag, New York, 1975.
15. L. D. IRVINE, S. P. MARIN, AND P. W. SMITH, Constrained interpolation and smoothing, *Constr. Approx.* **2** (1986), 129–151.
16. H. MASSAM, Optimality conditions for a cone–convex programming problem, *J. Austral. Math. Soc. Ser. A* **27** (1979), 141–162.
17. C. A. MICCHELLI, P. W. SMITH, J. SWETITS, AND J. D. WARD, Constrained L_p Approximation, *Constr. Approx.* **1** (1985), 93–102.
18. C. A. MICCHELLI AND F. UTRERAS, Smoothing and interpolation in a convex set a Hilbert space, *SIAM J. Sci. Statist. Comput.* **9** (1988), 728–746.
19. R. T. ROCKAFELLAR, Duality and stability in extremum problems involving convex functions, *Pacific J. Math.* **21** (1967), 167–187.
20. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, 1970.

Printed in Belgium
Uitgever: Academic Press, Inc.
Verantwoordelijke uitgever voor België:
Hubert Van Maele
Altenastraat 20, B-8310 Sint-Kruis