JOURNAL OF APPROXIMATION THEORY 80, 180–203 (1995)

Dual Cones, Constrained n-Convex L_p-Approximation, and Perfect Splines

FRANK DEUTSCH

Department of Mathematics, 417 McAllister Building, Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.

VASANT A. UBHAYA

Department of Computer Science and Operations Research, 258 IACC Building, North Dakota State University, Fargo, North Dakota 58105, U.S.A.

AND

YUESHENG XU

Department of Mathematics, 300 Minard Hall, North Dakota State University, Fargo, North Dakota 58105, U.S.A.

Communicated by Charles K. Chui

Received January 25, 1993; accepted October 27, 1993

A generating basis and the dual cone of n-convex functions satisfying certain constraints are derived. As applications, the existence and characterization of a best L_p-approximation ($1 \leqslant p < \infty$) from such subcones to a function in L_p are established. The relationship between a best L_1-approximation and perfect splines is developed under certain conditions. © 1995 Academic Press, Inc.

INTRODUCTION

Recently, there has been considerable interest in best L_p-approximation, $1 \leqslant p < \infty$, by n-convex functions (e.g., [8, 12, 30, 34, 27]). In this article, we consider a constrained L_p-approximation problem in which the approximating set is a convex subcone of n-convex functions determined by certain constraints. This problem was seen to arise naturally when one considers best constrained approximation (see [2] or [3]), which in turn arises from smoothing and interpolation problems (see, e.g., [4, 16]). A main problem of [3], for example, was to characterize best constrained approximations to elements x in a Hilbert space X from the set

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

0021-9045/95 $6.00

where C is a closed convex cone in X, A is a bounded linear operator from X into a Hilbert space Y, and $b \in Y$. It was seen there that this problem reduced to the generally simpler problem of determining best approximations to a perturbation of x from a certain *subcone* of the cone C. In the important cases when the cone C is the cone of positive functions, the increasing functions, the convex functions, or, more generally, the cone of n-convex functions, it was seen in [3] that the subcones that arise are precisely of the form that we consider in this paper (in the more general framework of the L_p-space). We establish the existence of a best L_p-approximation and its characterization by first determining a generating basis and then the dual cone of the subcone. This approach, based on duality, leads to simplicity of both methods and results, and particularly, a simple proof for the characterization of a best approximation. We consider L_2-approximation by nondecreasing functions, a special case of the above problem, in some detail and extend an earlier result of [21]. We also explore the relationship between a best L_1-approximation from the subcone and perfect splines.

Let X be a real normed linear space and X^* its topological dual with its usual norm. Let $K \subset X$ be a closed convex cone, i.e., a closed subset of X which satisfies the condition that $\lambda f + \mu h \in K$ whenever $f, h \in K$, $\lambda \geqslant 0$ and $\mu \geqslant 0$. Given $f \in X$, let

$$P_K(f) = \{g \in K : \|f-g\| = \inf\{\|f-k\| : k \in K\}\},$$

where $\|\cdot\|$ is the norm on X. $P_K(f)$ is called the set of best approximations to f from K. Define the dual (or polar, or conjugate) cone K^0 of K by

$$K^0 = \{x^* \in X^* : x^*(k) \leqslant 0 \text{ for all } k \in K\}.$$

The dual cone plays a significant role in the characterization of a best approximation as follows.

THEOREM 1.1. *Let $f \in X \setminus K$ and $g \in K$. Then $g \in P_K(f)$ if and only if*

$$K^0 \cap g^\perp \cap D(f-g) \neq \varnothing,$$

where $g^\perp = \{x^* \in X^* : x^*(g) = 0\}$ and

$$D(h) = \{x^* \in X^* : \|x^*\| = 1, x^*(h) = \|h\|\}, \qquad h \in X.$$

This result is a special case of a general characterization of best approximations from any convex set established independently in [5, 24]. (See [26, p. 362] for an accessible reference to these papers. See also [6, 32, 35] for further results on duality.) For $A \subset X$, we denote by $\text{cc}(A)$ the smallest convex cone containing A or, equivalently, the set of all non-negative linear combinations of elements of A. We denote by $\overline{\mathrm{cc}}(A)$ the smallest closed convex cone containing A. Since the closure of a cone is a cone, this is the closure of $\mathrm{cc}(A)$. A proper subset M of K is called a generating basis for K if $K=\overline{\mathrm{cc}}(M)$.

In this article, we let $X=L_p(I)$, $\|\cdot\|=\|\cdot\|_p$, $1\leqslant p<\infty$, where $I=[a,b]$ is a compact real interval with Legesgue measure, and let $K=K_{n,p}(S)$, $n\geqslant 1$, be the convex subcone of the n-convex functions in L_p, to be defined below. In Section 2, we find a generating basis for K and characterize the dual cone K^0. These results are derived from earlier known work on generalized convex functions induced by Extended Tchebycheff systems, also called the ET systems [9, 10]. In Section 3, we use the results of [34] to establish the existence of a best L_p-approximation from K. Using the results of Section 2, we obtain a characterization of a best L_p-approximation in Section 4. In Section 5, we consider the case of 1-convex (i.e., nondecreasing) functions with $p=2$, and extend a characterization of a best approximation to a bounded function [21] to any function in L_2. In Section 6, under certain conditions, we characterize the unique best L_1-approximation by n-convex functions in terms of a unique perfect spline.

We now present the notation and terminology used in this article in detail. We first state the following two equivalent definitions of a real function k which is *n-convex* on an interval $J\subset I$, where $n\geqslant 1$; additional definitions appear in [22].

(1) For all $n+1$ points $s_0<s_1<\cdots<s_n$ in J, the nth order divided difference $[s_0,s_1,...,s_n]k$ of k is nonnegative.

(2) For all n points $s_1<s_2<\cdots<s_n$ in J, $(-1)^{n+i+1}(P(s)-k(s))\geqslant 0$ for all s in (s_i,s_{i+1}), $0\leqslant i\leqslant n$, where $P(s)$ is the unique Lagrange interpolating polynomial of degree at most $(n-1)$ passing through the points $(s_i,k(s_i))$, $1\leqslant i\leqslant n$, and s_0 and s_{n+1} are the left and right endpoints of J.

It is known that a function k which is n-convex on $J=(a,b)$ has at most n monotone segments. This result may be derived from [20] (or see [34, p. 236, property (2)]); it is extended to generalized convex functions in [12]. Hence, k is monotone on the intervals $(a,a+\varepsilon)$ and $(b-\varepsilon,b)$ for some $\varepsilon>0$. Consequently, we let $k(a)=k(a+)$ and $k(b)=k(b-)$, where these limits may be $\pm\infty$. We let K_n, $n\geqslant 1$, denote the set of all functions on I which are n-convex on (a,b) and are so extended to the endpoints. We point out that the functions which are n-convex on I are a proper subset of K_n; the former, by definition, are necessarily finite at the endpoints of I. A best approximation to an $f\in L_p$ may not exist from the former class, but always exists from K_n [34].

Let μ_g denote the Lebesgue–Stieltjes complete measure generated on (a,b) by a real nondecreasing and possibly unbounded function g on (a,b), which is not necessarily right continuous. Then, for each Borel set $A \subset (a, b)$, we have

$$\mu_g(A) = \inf\left\{ \sum_{i=1}^{\infty} (g(b_i) - g(a_i)) : A \subset \bigcup_{i=1}^{\infty} (a_i, b_i), (a_i, b_i) \subset (a, b) \right\},$$

and μ_g is the completion of this measure on the Borel sets [17]. Let $S \subset (a, b)$ be any Borel set and $S' = (a, b) \setminus S$. For $k \in K_n$, let $k_R^{(n-1)}$ denote the right continuous nondecreasing right derivative of the $(n-2)$nd derivative of k defined on (a, b), where $k_R^{(0)}(t) = k_R(t) = k(t+)$. (See Section 2 for the justification of the existence of these derivatives.) Let $\mu_{k,n} = \mu_{k_R^{(n-1)}}$, i.e., $\mu_{k,n}$ denote the Lebesgue–Stieltjes measure generated by $k_R^{(n-1)}$ on (a, b). Note that $\mu_{k,1}$, which is generated by k_R, is identical to μ_k, which is generated by k [17, p. 160, Proposition 3.9]. Define

$$K_n(S) = \{k \in K_n : \mu_{k,n}(S') = 0\}.$$

In particular, since $\mu_{k,1} = \mu_k$, we have $K_1(S) = \{k \in K_1 : \mu_k(S') = 0\}$. Note that each k in K_n generates a distinct $\mu_{k,n}$ and an associated sigma-field. However, S' is measurable relative to each $\mu_{k,n}$ since it is a Borel set; thus $K_n(S)$ is well defined. It is a convex subcone of K_n. Clearly, $K_n = K_n((a, b))$, and $K_n(\emptyset)$ is the set of all polynomials of degree at most $n-1$ on I. In addition, if $S = \Pi = \{t_1 < t_2 < \cdots < t_m\}$, then $K_n(\Pi)$ is the set of all n-convex splines of degree at most $n-1$ with simple knots at t_i. Thus, this framework covers several important cases of interest.

We define

$$K_{n,p}(S) = K_n(S) \cap L_p, \qquad 1 \leqslant p < \infty,$$

where $L_p = L_p(I)$. This is a cone (a subcone of $K_n(S)$ and hence of K_n) in L_p from which we seek best approximations. When $S \neq (a, b)$, $K_n(S)$ is a proper "constrained" subcone of K_n. Such sets arose naturally, but implicitly, in the study of constrained approximation in [15, 16] for $n = 1$, and explicitly in [2, 3] for $n = 1, 2$. In this notation, $(K_{n,p}(S))^0$ is the dual cone of $K_{n,p}(S)$ in L_p^*. For brevity, we let $K_{n,p} = K_{n,p}((a, b)) = K_n \cap L_p$, and $K_{n,p}^0$ its dual cone.

We briefly review some related literature. If $f \in L_p$, $1 < p < \infty$, then the existence of the unique best approximation follows since $K_{n,p}$ is closed and convex [34, Theorem 3.1] and L_p is uniformly convex. In L_1, the existence follows by the same theorem in [34] or by [11]. We observe that 1-convex and 2-convex functions are, respectively, the nondecreasing and convex functions. More complex cases of n-convex functions occur for $n \geqslant 3$. There is much literature on L_p-approximation by *unconstrained* n-convex functions, particularly for $n = 1$. For characterization and properties of best approximants see [11, 27, 29–31, 37] and other references given there. Best constrained approximation in Hilbert spaces was investigated in [2, 3, 16]. Constrained approximation by nonnegative functions in L_p spaces was investigated in [15]. Certain interesting relationships between best L_1-approximation from the linear space of splines and perfect splines were obtained in [13, 28].

2. Preliminaries

In this section we obtain several preliminary results on n-convex functions and Lebesgue–Stieltjes measures. These results are needed in the analysis to follow.

We first state some basic facts about n-convexity. Let $k^{(i)}$ denote the ith derivative of a function k, where $k^{(0)} = k$.

Lemma 2.1. *Let $n \geqslant 1$ and $k \in K_n$.*

(1) *Every function in K_n, $n \geqslant 2$, is continuous on (a, b) [1].*

(2) *$k^{(i)}$ exists on (a, b) and $k^{(i)} \in K_{n-i}$, $1 \leqslant i \leqslant n-2$ [1, Corollary 15].*

(3) *$k^{(n-2)}$ is convex on (a, b).*

(4) *The left (resp., right) derivative $k_{\mathrm{L}}^{(n-1)}$ (resp., $k_{\mathrm{R}}^{(n-1)}$) of $k^{(n-2)}$ exists on (a, b), is nondecreasing, and is left (resp., right) continuous [22, 23].*

(5) *$k_{\mathrm{L}}^{(n-1)} = k_{\mathrm{R}}^{(n-1)}$ a.e., and, hence, $k^{(n-1)}$ exists a.e. on (a, b).*

Lemma 2.2. *Let k be a real nondecreasing and possibly unbounded function on (a, b) (i.e., $k \in K_1$). If $\mu = \mu_k$ is the Lebesgue–Stieltjes measure generated by k on (a, b) (as in Section 1), then, for any choice of $c < d$ in (a, b), the following hold [17].*

(1) $\mu\{c\} = k(c+) - k(c-)$.

(2) $\mu(c, d) = k(d-) - k(c+)$.

(3) $\mu[c, d] = k(d+) - k(c-)$.

(4) $\mu[c, d] = k(d-) - k(c-)$.

(5) $\mu(c, d] = k(d+) - k(c+)$.

Following the usual conventions, let $a_+ = \max\{a, 0\}$, $a_- = a_+ - a = \max\{-a, 0\}$, $(s-t)_+^{n-1} = ((s-t)_+)^{n-1}$ and $(s-t)_-^{n-1} = ((s-t)_-)^{n-1}$ for $n \geqslant 2$. Also define

$$
\begin{aligned}
(s-t)_+^0 &= 0, && \text{if } s<t,\\
&= 1, && \text{if } s \geqslant t,\\
(s-t)_-^0 &= 1, && \text{if } s<t,\\
&= 0, && \text{if } s \geqslant t.
\end{aligned}
$$

These functions will be used in this and the next section.

Let $k \in K_n(S)$ and $0<\delta<(b-a)/2$. For $0<\varepsilon<\delta$, define as in [10, p. 391], $\rho(\cdot;\varepsilon)=\rho_k(\cdot;\varepsilon)$ by

$$
\begin{aligned}
\rho(t;\varepsilon) &= k_{\mathrm{R}}^{(n-1)}(a+\varepsilon), && \text{if } t \in (a, a+\varepsilon),\\
&= k_{\mathrm{R}}^{(n-1)}(t), && \text{if } t \in [a+\varepsilon, b-\varepsilon),\\
&= k_{\mathrm{R}}^{(n-1)}(b-\varepsilon), && \text{if } t \in [b-\varepsilon, b).
\end{aligned}
\tag{2.1}
$$

Recall that if $k \in K_1(S)$, then $k_{\mathrm{R}}^{(0)}(t)=k_{\mathrm{R}}(t)=k(t+)$. Also define

$$
\begin{aligned}
k(t;\varepsilon) &= \left[\int_a^b (t-x)_+^{n-1}\, d\rho(x;\varepsilon) + \sum_{i=0}^{n-1} a_i(\varepsilon)\, t^i\right] \Big/ (n-1)!, && \text{if } n \geqslant 2,\\
&= \rho(t;\varepsilon), && \text{if } n=1,
\end{aligned}
\tag{2.2}
$$

where numbers $a_i(\varepsilon)$ are chosen so that $k(\cdot;\varepsilon)=k$ on $(a+\varepsilon, b-\varepsilon)$. The following lemmas collect some useful properties of the function $k(\cdot,\varepsilon)$ which play a significant role in our later developments. Recall from Section 1 that if $k \in K_n$, then $\mu_{k,n}$ is the measure generated by $k_{\mathrm{R}}^{(n-1)}$ and $\mu_{k,1}=\mu_k$.

LEMMA 2.3. *Let $k \in K_n(S)$, $n \geqslant 1$, and $0<\delta<(b-a)/2$. For $0<\varepsilon<\delta$, let $\rho(\cdot,\varepsilon)$ and $k(\cdot;\varepsilon)$ be defined by* (2.1) *and* (2.2). *Also, let μ be the Lebesgue–Stieltjes measure generated by $\rho(\cdot;\varepsilon)$. Then* (1)–(6) *below hold for $n \geqslant 2$. If $n=1$, then* (1)–(4) *hold verbatim*; (5) *and* (6) *hold with the function k there replaced by k_{R}.*

(1) $\mu(S')=0$.

(2) *μ is the measure generated by $k_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon)$.*

(3) *$k^{(i)}(a+;\varepsilon)$, $k^{(i)}(b-;\varepsilon)$ for $0 \leqslant i \leqslant n-2$, and $k_{\mathrm{R}}^{(n-1)}(a+;\varepsilon)$ and $k_{\mathrm{R}}^{(n-1)}(b-;\varepsilon)$ exist and are finite.*

(4) $k(\cdot;\varepsilon) \in K_{n,p}(S)$, $1 \leqslant p < \infty$.

(5) *$k(\cdot;\varepsilon)=k$ on $(a+\varepsilon, b-\varepsilon)$, $k(\cdot;\varepsilon) \leqslant k$ on $[b-\varepsilon, b)$, and $(-1)^n k(\cdot;\varepsilon) \leqslant (-1)^n k$ on $(a, a+\varepsilon]$.*

(6) *For each fixed t in $(a, a+\delta)$ (resp., $(b-\delta, b)$), $k(t;\varepsilon)$ (resp., $(-1)^n k(t;\varepsilon)$) is a nonincreasing function of ε for $0<\varepsilon<\delta$. Furthermore, $k(\cdot;\varepsilon) \uparrow k$* on *$(b-\delta, b)$, $(-1)^n k(\cdot;\varepsilon) \uparrow (-1)^n k$ on $(a, a+\delta)$, as $\varepsilon \downarrow 0$.*

Proof. To show (1), we use the right continuity of $\rho(\cdot;\varepsilon)$ and its continuity at $a+\varepsilon$. Suppose $n\geqslant 2$. Then $\rho(\cdot;\varepsilon)=k_{\rm R}^{(n-1)}$ on $(a+\varepsilon, b-\varepsilon]=J$, say, which gives $\mu=\mu_{k,n}$ on J (i.e., for measurable subsets of J). Since $k\in K_n(S)$, we have $\mu_{k,n}(S'\cap J)=0$ and hence $\mu(S'\cap J)=0$. If $n=1$, then $\rho(\cdot;\varepsilon)=k_{\rm R}$ on $(a+\varepsilon, b-\varepsilon]$, which gives $\mu=\mu_{k_{\rm R}}=\mu_k$ on J. Hence, as before, $\mu(S'\cap J)=0$. Now for all $n\geqslant 1$, we have $\mu(a, a+\varepsilon]=\mu(b-\varepsilon, b)=0$. We conclude that $\mu(S')=0$, which is (1).

To prove the remaining parts, we apply [10, Chap. XI, Theorem 2.3] with $w_i\equiv 1$ and n replaced by $n-1$. Suppose $n\geqslant 2$. We differentiate (2.2) $n-1$ times as justified in [10, p. 392] and obtain

$$\begin{aligned} k_{\rm R}^{(n-1)}(t;\varepsilon)&=\int_a^b (t-x)_+^0\, d\rho(x;\varepsilon)+a_{n-1}(\varepsilon)\\ &=\rho(t;\varepsilon)-\rho(a+;\varepsilon)+a_{n-1}(\varepsilon), \end{aligned}\tag{2.3}$$

by the right continuity of $\rho(\cdot;\varepsilon)$. Thus $k_{\rm R}^{(n-1)}(\cdot;\varepsilon)$ and $\rho(\cdot;\varepsilon)$ differ by a constant and (2) follows. To show (3) we observe that $\rho(\cdot;\varepsilon)$ is nondecreasing and bounded. Hence, again by (2.3), we conclude that $k_{\rm R}^{(n-1)}(a+;\varepsilon)$ and $k_{\rm R}^{(n-1)}(b+;\varepsilon)$ exist and are finite. It follows that $k^{(i)}(a+;\varepsilon)$ and $k^{(i)}(b-;\varepsilon)$ exist and are finite for $0\leqslant i\leqslant n-2$. If $n=1$, then (2) and (3) follow immediately. By (3), $k(\cdot;\varepsilon)$ is bounded and, hence, in L_p. Now, by the theorem in [10] cited above, and (1) and (2), we conclude that (4) holds; again (5) and (6) hold by the same theorem. The proof is complete.

LEMMA 2.4. *Let $k\in K_{n,p}(S)$, $n\geqslant 1$ and $1\leqslant p<\infty$. Then*

(1) $\|k(\cdot;\varepsilon)-k\|_p\to 0$ *as* $\varepsilon\downarrow 0$.

(2) $\int_a^b k(\cdot,\varepsilon)h\to\int_a^b kh$ *as* $\varepsilon\downarrow 0$ *for all* $h\in L_q$, *where* $1/p+1/q=1$.

Proof. Suppose $n\geqslant 2$. By Lemma 2.3(5), for $0<\varepsilon<\delta$ we have $|k-k(\cdot;\varepsilon)|\leqslant|k-k(\cdot;\delta)|\in L_p$. By Lemma 2.3(6), $k(\cdot;\varepsilon)\to k$ pointwise as $\varepsilon\downarrow 0$. Hence, by the bounded convergence theorem [7], we conclude that (1) holds. For $n=1$, since $k=k_{\rm R}$ a.e., by the same argument (1) holds. Now (2) follows immediately from (1) by an application of Holder's inequality [7]. The proof is complete.

A family F of real functions is said to be *equi-Lipschitzian* on a compact subinterval J of (a, b) if $|f(s)-f(t)|\leqslant c\,|s-t|$ holds for all f in F, all s, t in J, and some $c>0$. Parts of Theorem 2.5, below, are extensions of similar results for convex functions [23, Sect. 10]; others are contained in [34]. Results similar to parts (1) and (3) appeared in [36]. It was shown in [12] that Theorem 2.5 is also true in a more general framework of generalized convex functions relative to a nonlinear family under certain conditions.

THEOREM 2.5. *Let $n \geqslant 2$, $1 \leqslant p \leqslant \infty$, and (k_j) be a sequence in K_n.*

(1) *If a sequence in K_n converges pointwise to some real function k, then k is in K_n and the convergence is uniform on every compact subinterval of (a, b).*

(2) *If $(\|k_j\|_p)$ is bounded, then (k_j) is pointwise bounded on (a, b).*

(3) *If (k_j) is pointwise bounded on (a, b), then (k_j) is equi-Lipschitzian on every compact subinterval of (a, b) and (k_j) contains a subsequence which converges pointwise on (a, b) to some function in K_n.*

PROPOSITION 2.6. *Let (k_j) be a sequence in K_n, $n \geqslant 2$, such that $k_j \to k$ pointwise on (a, b) for some k in K_n. Then $k_j^{(i)} \to k^{(i)}$ pointwise on (a, b) uniformly on every compact subinterval J of (a, b) for all $0 \leqslant i \leqslant n-2$. (For the case $i = n-1$ see the remark following the proof below.)*

Proof. We first establish the result for $i = 1$ when $n \geqslant 3$; it holds for $i = 0$ by hypothesis and Theorem 2.5(1). Let (g_j) be any subsequence of $(k_j^{(1)})$. We show that this in turn contains a subsequence converging pointwise to $k^{(1)}$ uniformly on every J. This will prove the assertion. Since (k_j) is pointwise bounded on (a, b), by Theorem 2.5, there exists $c > 0$ such that $|k_j(s) - k_j(t)| \leqslant c\,|s - t|$ for s, t in J. Consequently, $|k_j^{(1)}(s)| \leqslant c$ for s in J. Thus $k_j^{(1)}$ is pointwise bounded on (a, b). Since $k_j^{(1)} \in K_{n-1}$, by Theorem 2.5, $(k_j^{(1)})$ contains a convergent subsequence. Hence, assume that g_j itself is convergent to some g in K_{n-1}. We show that $g = k^{(1)}$. Let (h_j) be the subsequence of (k_j) such that $g_j = h_j^{(1)}$. Let $t \in (a, b)$ and let $J' = [u, v] \subset (a, b)$ with $u < s < t < v$. Then since h_j is Lipschitzian and, hence, absolutely continuous on J', we have $h_j(t) = \int_s^t h_j^{(1)} + h_j(s)$. Since $|h_j^{(1)}| \leqslant c'$ for all j for some $c' > 0$, using the bounded convergence theorem and passing to limits, we obtain $k(t) = \int_s^t g + k(s)$. Since $k^{(1)}$ and g are continuous, we have $k^{(1)} = g$ on J' and, hence, on (a, b). By Theorem 2.5(1), $k_j^{(1)}$ converges to $k^{(1)}$ uniformly on J. Now since $k_j^{(1)}$ is in K_{n-1}, we apply the same argument to prove the assertion for $(k_j^{(2)})$, etc. The proof is complete.

Note that $k_j^{(n-1)}$ is the derivative of the convex function $k_j^{(n-2)}$. Its convergence is covered in [23, Theorem 25.7]. Recall from Section 1 that $g_{\mathrm{R}}(t) = g(t+)$. We define $g_{\mathrm{L}}(t) = g(t-)$.

LEMMA 2.7. *Let g, $k \in K_1$ and $g = k$ a.e. on (a, b). Then the following hold.*

(1) *$g_{\mathrm{R}} = k_{\mathrm{R}}$ and $g_{\mathrm{L}} = k_{\mathrm{L}}$ on (a, b).*

(2) *The sets of discontinuities of g and k are identical.*

(3) *g and k generate identical Lebesgue–Stieltjes measures on (a, b).*

Proof. (1) Let E be the set of continuity points of both g and k. Since a real nondecreasing function (possibly unbounded) has countable discontinuities, we have $\lambda(E')=0$, where $E'=(a,b)\setminus E$ and λ is the Lebesgue measure. Since $g=k$ a.e., $g=k$ on E. Now suppose that $s\in E'$. Then since $\lambda(E')=0$, given $\delta>0$ there exists $t\in E$ with $s<t<s+\delta$ so that $g(t)=k(t)$. It follows that $g(s+)=k(s+)$. Similarly, $g(s-)=k(s-)$. This gives (1). Since g is discontinuous at s if and only if $g(s+)-g(s-)>0$, (2) is established. Now g and g_{R} generate the same measure on (a,b) [17, Proposition 3.9]. Hence (3) follows from (2). The proof is complete.

LEMMA 2.8. *Let (k_j) be a sequence in K_1 such that $k_j\to k$ pointwise on (a,b) for some k in K_1. Let μ_j and μ be the Lebesgue–Stieltjes measures generated by k_j and k on (a,b). Let $c<d$ in (a,b) be any two points of continuity of k. Then,*

(1) $k_j(c+)\to k(c)$ *and* $k_j(c-)\to k(c)$;

(2) $\mu_j(c,d)\to\mu(c,d)$.

Proof. Let $\varepsilon>0$. There exists $N>0$ such that $k(c)-\varepsilon\leqslant k_j(c)\leqslant k_j(c+)$ for $j\geqslant N$. Hence $k(c)\leqslant\liminf k_j(c+)$. Now let $s>c$. Then there exists $N>0$ such that $k_j(c+)\leqslant k_j(s)\leqslant k(s)+\varepsilon$ for $j\geqslant N$. Hence $\limsup k_j(c+)\leqslant k(s)$. By continuity we have $\limsup k_j(c+)\leqslant k(c)$. This shows that $k_j(c+)\to k(c)$. Similarly, we have $k_j(c-)\to k(c)$, and (1) is established. Part (2) follows from Lemma 2.2(2) applied to k_j and k. The proof is complete.

The following slight generalization of Helly's selection theorem [18, p. 221, Lemma 2] is needed for our purpose.

LEMMA 2.9. *Let (k_j) be a sequence in K_1 which is bounded uniformly in j on every compact subset of (a,b). Then there exists a subsequence which converges pointwise on (a,b) to a function in K_1 which is bounded on every compact subset of (a,b).*

Proof. Let $0<\varepsilon<(b-a)/2$ and $I_m=[a+\varepsilon/m, b-\varepsilon/m]$. By Helly's theorem, there exists a subsequence $(g_{1,j})$ of (k_j) which converges at every point of I_1. Again, by the same theorem, there exists a subsequence $(g_{2,j})$ of $(g_{1,j})$ which converges at every point of I_2. Repeating this argument for each I_m, we finally let $(g_j)=(g_{j,j})$, the diagonal sequence which converges. Clearly, the limit function is in K_1 and is bounded on every compact subset of (a,b). The proof is complete.

3. Generating Basis for $K_{n,p}(S)$ and Existence of Best Approximations

In this section we obtain a generating basis for $K_n(S)$ and $K_{n,p}(S)$ from earlier known results [9, 10], and establish the existence of a best approximation from $K_{n,p}(S)$.

The following set of functions, $M_n(S)$ or $M'_n(S)$, of the variable s will be shown to generate $K_{n,p}(S)$, $n \geqslant 1$, if S is (relatively) closed in (a, b).

$$M_n(S) = \{\pm s^i : 0 \leqslant i \leqslant n-1\} \cup \{(s-t)_+^{n-1} : t \in S\},$$

$$M'_n(S) = \{\pm s^i : 0 \leqslant i \leqslant n-1\} \cup \{(-1)^n (s-t)_-^{n-1} : t \in S\}.$$

Note that $(s-t)_-^0$ and $(s-t)_+^0$ are right continuous.

We collect a few more facts for ease of reference.

Lemma 3.1. (1) *$k \in K_n$, $n \geqslant 1$, if and only if it is the $(n-1)$st indefinite integral of a nondecreasing function* [1, *Corollary* 8(a)].

(2) *$k(s) = (s-t)_+^{n-1}$ (resp., $(-1)^n (s-t)_-^{n-1}$ is n-convex.*

(3) *If $k(s) = (s-t)_+^{n-1}$, then $k_{\mathrm{R}}^{(n-1)}(s) = (n-1)!\,(s-t)_+^0$, and $\mu_{k,n}$ is zero on $(a, t) \cup (t, b)$. Hence $k \in K_n(S)$ if $t \in S$.*

(4) *$s^i \in K_n(S)$, $0 \leqslant i \leqslant n-1$.*

(5) *$M_n(S) \subset K_{n,p}(S)$ and $M'_n(S) \subset K_{n,p}(S)$, $1 \leqslant p < \infty$.*

Proof. (2) This follows from (1) since $(s-t)_+^{n-1}$ (resp., $(-1)^n (s-t)_-^{n-1}$) is the $(n-1)$st indefinite integral of the nondecreasing function $(n-1)!\,(s-t)_+^0$ (resp., $-(n-1)!\,(s-t)_-^0$) plus a polynomial of degree at most $n-2$.

(3) This is clear.

(4) The $(n-1)$st derivative of $k(s) = s^i$, $0 \leqslant i \leqslant n-1$, is constant so that $\mu_{k,n} = 0$. Thus $k \in K_n(S)$.

(5) By (3) and (4) we have $M_n(S) \subset K_n(S)$. Since functions in $M_n(S)$ are bounded, we have $M_n(S) \subset L_p$, and the first inclusion in (5) follows. A similar proof establishes the second inclusion.

The proof is complete.

Let S_n denote the set of all polynomial spline functions of degree $n-1$ with a finite number of simple variable knots in (a, b) [25]. It is then easy to see that $M_n = M_n((a, b)) \subset S_n$ and S_n is spanned by $M_n \cup \{-M_n\}$. Similar results hold for $M'_n = M'_n((a, b))$. Recall that if $A \subset L_p$, then $\overline{\mathrm{cc}}_p(A)$ denotes the closure of $\mathrm{cc}(A)$ in L_p.

Theorem 3.2. *$K_{n,p}(S) \subset \overline{\mathrm{cc}}_p(M_n(S)) = \overline{\mathrm{cc}}_p(M'_n(S))$ for all $n \geqslant 1$ and $1 \leqslant p < \infty$.*

Proof. Clearly, $(-1)^n (s-t)_-^{n-1} + (s-t)^{n-1} = (s-t)_+^{n-1}$ for all $n \geq 1$. Hence, $\mathrm{cc}(M_n(S)) = \mathrm{cc}(M'_n(S))$ and $\overline{\mathrm{cc}}_p(M_n(S)) = \overline{\mathrm{cc}}_p(M'_n(S))$. Now let $k \in K_{n,p}(S)$ and $0 < \delta < (b-a)/2$. For $0 < \varepsilon < \delta$, construct $k(\cdot\,; \varepsilon)$ as in (2.2). Then by Lemma 2.3(4), $k(\cdot\,; \varepsilon) \in K_{n,p}(S)$, and by Lemma 2.4(1), $\|k(\cdot\,; \varepsilon) - k\|_p \to 0$ as $\varepsilon \downarrow 0$. Let μ be the measure generated by $\rho(\cdot\,; \varepsilon)$ on (a, b). First suppose that $n \geq 2$. Let $f(t) = \int_S (t-x)_+^{n-1}\, d\rho(x; \varepsilon)$, $t \in (a, b)$. Note that $\rho(\cdot\,; \varepsilon)$ is bounded, and the family $F = \{(t-x)_+^{n-1} : t \in I\}$ of functions of the variable x, is equi-continuous on I, i.e., given $\theta > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \theta$ whenever $|x - y| < \delta$ for all $f \in F$. Let $a = x_0 < x_1 < \cdots < x_{m+1} = b$ be a partition of (a, b) such that $x_i - x_{i-1} < \delta$ for $1 \leq i \leq m+1$. For convenience of notation, let $\rho(b; \varepsilon) = \rho(b-; \varepsilon)$ and $c = \rho(b-; \varepsilon) - \rho(a+; \varepsilon)$. Since $\mu(S') = 0$, by the right continuity of $\rho(\cdot\,; \varepsilon)$, we have $\mu((x_{i-1}, x_i] \cap S) = \mu(x_{i-1}, x_i] = \rho(x_i; \varepsilon) - \rho(x_{i-1}; \varepsilon) = \lambda_i$, say. Then $\lambda_i \geq 0$. Let $D = \{1 \leq i \leq m+1 : \lambda_i > 0\}$. If $i \in D$ then $(x_{i-1}, x_i] \cap S \neq \emptyset$. Now choose $y_i \in (x_{i-1}, x_i] \cap S$ arbitrarily for $i \in D$, and define $g(t) = \sum_{i \in D} (t - y_i)_+^{n-1} \lambda_i$. Then, by construction, $|f(s) - g(s)| \leq \theta c$ for all s in (a, b) since $\sum_{i \in D} \lambda_i = c$. Clearly, $g \in \mathrm{cc}(M_n(S))$, and hence, $f \in \overline{\mathrm{cc}}_p(M_p(S))$. We conclude that $k(\cdot\,; \varepsilon)$ is in $\overline{\mathrm{cc}}_p(M_n(S))$. Thus $k \in \overline{\mathrm{cc}}_p(M_n(S))$ and the result is established for $n \geq 2$.

Now suppose that $n = 1$ and, for convenience, let $f = k(\cdot\,; \varepsilon)$ and $\theta > 0$. Then $f \in K_{1,p}(S)$. Note that f is bounded and right continuous. Let j be the smallest integer with $j + 1 \geq (f(b) - f(a))/\theta$. Let $I_i = \{s \in I : f(s) \geq f(a) + i\theta\}$, $0 \leq i \leq j$. Since f is right continuous, I_i has the form $[s_i, b)$, where $a = s_0 \leq s_1 \leq \cdots \leq s_{j+1} = b$. Let $a = x_0 < x_1 < \cdots < x_{m+1} = b$ be all distinct elements among s_i. (If f has a jump at a point t then some of the s_i may be identical to t.) Then $f(x_{i-1}) < f(x_i)$ for $1 \leq i \leq m$. Now, as before, $\mu((x_{i-1}, x_i] \cap S) = f(x_i) - f(x_{i-1}) > 0$ for $1 \leq i \leq m$. Define $y_i \in S$ with $a = y_0 < y_1 < \cdots < y_{m+1} = b$ as follows. If $\mu\{x_i\} = f(x_i) - f(x_i^-) > 0$ for $1 \leq i \leq m$ then $x_i \in S$, and let $y_i = x_i$. Otherwise, if $f(x_i) = f(x_i^-)$, choose y_i arbitrarily in $(x_{i-1}, x_i] \cap S$, which is nonempty since its μ-measure is positive. Now define $g(s) = f(a) + \sum_{i=1}^m (f(y_i) - f(y_{i-1}))(s - y_i)_+^0$. Then $g \in \mathrm{cc}(M_1(S))$ and, by construction, $|f(s) - g(s)| \leq 2\theta$ for $s \in (a, b)$. Hence, $f \in \overline{\mathrm{cc}}_p(M_1(S))$. Consequently, by Lemma 2.4(1), $k \in \overline{\mathrm{cc}}_p(M_1(S))$ establishing the result for $n = 1$. The proof is complete.

PROPOSITION 3.3. *Let $n \geq 1$ and $1 \leq p < \infty$. Assume that S is not (relatively) closed in (a, b). Then $\overline{\mathrm{cc}}_p(M_n(S)) \backslash K_{n,p}(S)$ is not empty.*

Proof. There exist $t \in (a, b) \backslash S$ and a sequence (t_j) in S such that $t_j \to t$. Define $k_j(s) = (s - t_j)_+^{n-1}$ and $k(s) = (s-t)_+^{n-1}$. Clearly, $k_j \in M_n(S)$, and $k_j \to k$ a.e. on (a, b). Since k_j and k are bounded by $(b-a)^{n-1}$, using the dominated convergence theorem [7], we have $\|k_j - k\|_p \to 0$. Thus $k \in \overline{\mathrm{cc}}_p(M_n(S))$. Let μ be the Lebesgue–Stieltjes measure generated by

$k_{\mathrm{R}}^{(n-1)}(s)=(n-1)!\,(s-t)_{+}^{0}$. Then $\mu\{t\}=(n-1)!\neq 0$, and $k\notin K_{n,p}(S)$. The proof is complete.

Let H denote the set of all extended real-valued functions on I. For $P\subset H$ we define $\overline{\overline{P}}$ to be the set of all functions f in H such that $f_j\to f$ pointwise on (a,b) for some sequence (f_j) in P. Such sets find applications in proving the existence of a best approximation [12, 34]. The definition of $\overline{\overline{P}}$ given here is as in [12] but weaker than the one in [34]; however, it will be seen that all the results of [34] hold with this change. The following results from [34] will be used in our proofs.

THEOREM 3.4. (1) $K_{n,p}=K_n\cap L_p=\overline{\overline{K_n}}\cap L_p$.

[2] *If* $P\subset K_n$ *is nonempty with* $P\cap L_P=\overline{\overline{P}}\cap L_p$, *then* $\overline{\overline{P}}\cap L_p$ *is proximinal in* L_p. *In particular,* $K_{n,p}$ *is proximinal in* L_p.

Now we state the main result of this section.

THEOREM 3.5. *Let* $1\leqslant p<\infty$. *The following statements are equivalent.*

(1) S *is (relatively) closed in* (a,b).

(2) *If* (k_j) *is a sequence in* $K_n(S)$, $n\geqslant 1$, *such that* k_j *converges pointwise to a real function* k *on* (a,b), *then* $k\in K_n(S)$.

(3) $K_{n,p}(S)=K_n(S)\cap L_p=\overline{\overline{K_n(S)}}\cap L_p$.

(4) $K_{n,p}(S)$ *is proximinal in* L_p, $n\geqslant 1$. *(Hence* $K_{n,p}(S)$, $1<p<\infty$, *is Chebychev.)*

(5) $K_{n,p}(S)$ *is closed in* L_p, $n\geqslant 1$.

(6) $K_{n,p}(S)=\overline{\mathrm{cc}}_p(M_n(S))=\overline{\mathrm{cc}}_p(M_n'(S))$, $n\geqslant 1$.

Proof. (1)$\Rightarrow$(2) Let (k_j) and k be as in (2). Then $k_j\in K_n$ and hence $k\in K_n$. By Proposition 2.6, $k_j^{(n-2)}\to k^{(n-2)}$ pointwise on (a,b). Again, by Theorem 2.5, the sequence $(k_j^{(n-2)})$ of convex functions is equi-Lipschitzian on compact subsets of (a,b). Hence, the sequence $(g_j=k_{j,\mathrm{R}}^{(n-1)})$ of non-decreasing functions is bounded on compact subsets of (a,b). By Lemma 2.9, there exists a subsequence of (g_j) converging to some g in K_1 on (a,b). Assume, for convenience, that (g_j) itself converges to g. Now let E be the subset of (a,b) on which $k^{(n-1)}$ exists, i.e., $k_{\mathrm{R}}^{(n-1)}=k_{\mathrm{L}}^{(n-1)}$ holds. Then $\lambda((a,b)\setminus E)=0$, where λ is the Lebesgue measure on (a,b) [22]. By a known result, e.g., [23, Theorem 25.7], we conclude that $g_j(s)\to k_{\mathrm{R}}^{(n-1)}(s)$ for $s\in E$. It follows that $k_{\mathrm{R}}^{(n-1)}=g$ a.e. Let μ_j, μ, and μ' be the Lesbesgue–Stieltjes measures generated by g_j, g, and $k_{\mathrm{R}}^{(n-1)}$, respectively. Then, by Lemma 2.7, we have $\mu=\mu'$. Let (u,v) be a component (maximal open subinterval) of the open set $S'=(a,b)\setminus S$. Since $g\in K_1$, we can find sequences (c_i) and (d_i) of continuity points of g such that $u<c_i<d_i<v$ and

$c_i \downarrow u$, $d_i \uparrow v$. Then $\mu_j(c_i, d_i) \leqslant \mu_j(u, v) = 0$. Letting $j \to \infty$, by Lemma 2.8, we obtain $\mu(c_i, d_i) = 0$ for all i, which gives $\mu'(u, v) = \mu(u, v) = 0$. Thus $\mu'(S') = 0$ and $k \in K_n(S)$.

(2) ⇒ (3) Clearly $K_{n,p}(S) \subset \overline{\overline{K_n(S)}} \cap L_p$. Suppose now that $k \in \overline{\overline{K_n(S)}} \cap L_p$. Then there exists a sequence (k_j) in $K_n(S)$ such that $k_j \to k$ pointwise on (a, b). Since $\overline{\overline{K_n(S)}} \cap L_p \subset \overline{\overline{K_n}} \cap L_p = K_n \cap L_p$ by Proposition 3.3, we conclude that $k \in K_n$ and, hence, is real-valued on (a, b). By (2), $k \in K_n(S)$.

(3) ⇒ (4) This follows by Theorem 3.4 with $P = K_n(S)$.

(4) ⇒ (5) Proximality implies closedness.

(5) ⇒ (6) Since $K_{n,p}(S)$ is closed in L_p, we have $K_{n,p}(S) \supset \overline{\text{cc}}_p(M_n(S))$.

The converse follows by Theorem 3.2.

(6) ⇒ (1) This follows by Proposition 3.3.

The proof is complete.

4. Characterization of $(K_{n,p}(S))^0$ and Best L_p-Approximation by $K_{n,p}(S)$

In this section we apply the results of Section 2 to characterize the dual cone $(K_{n,p}(S))^0$ and a best approximation to f in L_p from $K_{n,p}(S)$.

For $h \in L_1$, we define

$$h^{[0]} = h, \qquad h^{[i]}(s) = \int_a^s h^{[i-1]}(t)\, dt, \qquad s \in [a, b), \quad i \geqslant 1.$$

Thus $h^{[i]}(a) = 0$, for $i \geqslant 1$. Note that L_p^*, $1 \leqslant p < \infty$, is identified with L_q, where $q = p/(p-1)$ if $p > 1$, and $q = \infty$ if $p = 1$.

Theorem 4.1. *For $n \geqslant 1$, $1 \leqslant p < \infty$, and all $S \subset (a, b)$, the following hold.*

(1) $(K_{n,p}(S))^0 = (M_n(S))^0 = (M_n'(S))^0$.

(2) $(K_{n,p}(S))^0 = \{h \in L_q : h^{[i]}(b) = 0,\ 1 \leqslant i \leqslant n, \text{ and } (-1)^n h^{[n]}(t) \leqslant 0,\ t \in S\}$.

Proof. (1) By Theorem 3.2, we have $M_n(S) \subset K_{n,p}(S) \subset \overline{\text{cc}}_p(M_n(S))$. Hence $(M_n(S))^0 \supset (K_{n,p}(S))^0 \supset (\overline{\text{cc}}_p(M_n(S)))^0$. Since $(M_n(S))^0 = (\overline{\text{cc}}_p(M_n(S)))^0$, as may be easily verified, the result follows.

(2) Suppose first that $h \in (K_{n,p}(S))^0$. Then, by (1), $\int_a^b hk \leqslant 0$ for all $k \in M_n(S)$. We first prove that $h^{[i]}(b) = 0$ for $1 \leqslant i \leqslant n$. We proceed by

induction on i. Substituting $k(s) = \pm 1$ in $\int_a^b hk \leqslant 0$, we at once obtain $h^{[1]}(b) = 0$. Next assume that $h^{[i]}(b) = 0$ for $1 \leqslant i \leqslant m$, where $m \leqslant n-1$. Then, since $h \in L_1$ and $k(s) = \pm s^m$ is in $M_n(S)$, we integrate by parts to obtain

$$0 = \int_a^b s^m h(s)\, ds$$

$$= [s^m h^{[1]}(s)]_a^b - m \int_a^b s^{m-1} h^{[1]}(s)\, ds$$

$$= -m \int_a^b s^{m-1} h^{[1]}(s)\, ds.$$

Hence, $\int_a^b s^{m-1} h^{[1]}(s)\, ds = 0$. Applying the above step successively, we obtain $\int_a^b s^0 h^{[m]}(s)\, ds = 0$, which gives $h^{[m+1]}(b) = 0$. Hence $h^{[i]}(b) = 0$, $1 \leqslant i \leqslant n$. Again, substituting $k(s) = (s-t)_+^{n-1}$ with $t \in S$ in $\int_a^b hk \leqslant 0$ and integrating by parts, we may easily verify that

$$0 \geqslant \int_a^b (s-t)_+^{n-1} h(s)\, ds = \int_t^b (s-t)^{n-1} h(s)\, ds = (-1)^n (n-1)!\, h^{[n]}(t).$$

This gives $(-1)^n h^{[n]}(t) \leqslant 0$, $t \in S$.

Conversely, if $h \in L_q$ and satisfies $h^{[i]}(b) = 0$ for $1 \leqslant i \leqslant n$, and $(-1)^n h^{[n]}(t) \leqslant 0$ for $t \in S$, then we may show as above that $\int_a^b hk \leqslant 0$ for all $k \in M_n(S)$. Thus $h \in (K_{n,p}(S))^0$ and the proof is complete.

Next we obtain one preliminary result needed for characterization of a best approximation.

LEMMA 4.2. *Assume $g \in K_{n,p}(S)$, and let $\rho_g(\cdot\,; \varepsilon)$ and $g(\cdot\,; \varepsilon)$ be as defined by* (2.1) *and* (2.2) *for this g. Also, let $h \in L_q$, where $1 \leqslant p < \infty$, $1/p + 1/q = 1$, and $h^{[i]}(a) = h^{[i]}(b) = 0$, $0 \leqslant i \leqslant n$. Then, for $n \geqslant 1$,*

$$\int_a^b hg(\cdot\,; \varepsilon) = (-1)^n \int_{(a+\varepsilon, b-\varepsilon]} h^{[n]}\, dg_{\mathrm{R}}^{(n-1)}.$$

Proof. Suppose first that $n \geqslant 2$. By Lemma 2.3(4), $g(\cdot\,; \varepsilon) \in K_{n,p}(S)$. Hence, $g^{(i)}(\cdot\,; \varepsilon)$ is obtained by successive indefinite integrations of $g_{\mathrm{R}}^{(n-1)}(\cdot\,; \varepsilon)$. By Lemma 2.3(3), we conclude that $g^{(i)}(\cdot\,; \varepsilon)$, $0 \leqslant i \leqslant n-2$, is bounded on (a, b); it is also absolutely continuous since it is an indefinite integral. Again, $h^{[i]}$, $i \geqslant 1$, is absolutely continuous on I. Problem 3.3.6 of [17, p. 166] as extended to the Lebesgue–Stieltjes signed measure generated by $g(\cdot;\varepsilon)$ gives us $\int_a^b h^{[1]}\,dg(\cdot;\varepsilon)=\int_a^b h^{[1]}g^{(1)}(\cdot;\varepsilon)$. Similarly, $\int_a^b g(\cdot;\varepsilon)\,dh^{[1]}=\int_a^b hg(\cdot;\varepsilon)$. By [7, Theorem III.6.22], we have

$$\begin{aligned}\int_a^b hg(\cdot;\varepsilon)&=\int_a^b g(\cdot;\varepsilon)\,dh^{[1]}\\&=h^{[1]}(b-)\,g(b-;\varepsilon)-h^{[1]}(a+)\,g(a+;\varepsilon)-\int_{(a,b)}h^{[1]}\,dg(\cdot;\varepsilon)\\&=-\int_a^b h^{[1]}g^{(1)}(\cdot;\varepsilon).\end{aligned}$$

The above argument applied successively gives

$$\begin{aligned}\int_a^b hg(\cdot;\varepsilon)&=(-1)^{n-1}\int_a^b h^{[n-1]}g^{(n-1)}(\cdot;\varepsilon)\\&=(-1)^{n-1}\int_a^b h^{[n-1]}g_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon),\end{aligned}$$

since $g^{(n-1)}(\cdot;\varepsilon)=g_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon)$ a.e. Again, arguing as above we obtain $\int_a^b hg(\cdot;\varepsilon)=(-1)^n\int_{(a,b)}h^{[n]}\,dg_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon)$. Now, by definition, $\rho(\cdot;\varepsilon)$ is constant on $(a,a+\varepsilon]$ and $(b-\varepsilon,b)$, continuous at $a+\varepsilon$, and $\rho(\cdot;\varepsilon)=g_{\mathrm{R}}^{(n-1)}$ on $(a+\varepsilon,b-\varepsilon]$. As in the proof of (2.3), $g_{\mathrm{R}}^{(n-1)}(\cdot;\,)=\rho(\cdot;\varepsilon)+c=g_{\mathrm{R}}^{(n-1)}+c$ on $(a+\varepsilon,b-\varepsilon]$ for some constant c depending on ε. Hence,

$$\begin{aligned}\int_{(a,b)}h^{[n]}\,dg_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon)&=\int_{(a+\varepsilon,b-\varepsilon]}h^{[n]}\,dg_{\mathrm{R}}^{(n-1)}(\cdot;\varepsilon)\\&=\int_{(a+\varepsilon,b-\varepsilon]}h^{[n]}\,dg_{\mathrm{R}}^{(n-1)}.\end{aligned}$$

The required result is established for $n\geqslant 2$. If $n=1$, then the result may be derived as above by using the results on integration by parts. The proof is complete.

As was observed before, if $X=L_p$, $1\leqslant p<\infty$, then X^* is identified with L_q. Hence, if $1<p<\infty$, then $D_p(f)=(|f|/\|f\|_p)^{p-1}\operatorname{sgn}(f)\in L_q$, where $q=p/(p-1)$, and

$$D_1(f)=\{h:\|h\|_\infty=1,\ h=\operatorname{sgn}(f)\text{ a.e. where } f\neq 0\}\subset L_\infty.$$

We observe that e and E in Theorem 4.3, below, depend upon g; in fact they are uniquely determined by $f-g$. This, however, is not the case in Theorem 4.5. For simplicity of notation, we suppress any dependence. Similar remarks apply to other characterization theorems in subsequent sections.

THEOREM 4.3. *Let* $1<p<\infty$, $n\geq 1$, $S\subset(a,b)$, $K=K_{n,p}(S)$, $f\in L_p\backslash K$, $g\in K$, *and* $e=|f-g|^{p-1}\operatorname{sgn}(f-g)$. Define

$$E=\{t\in(a,b):(-1)^n e^{[n]}(t)<0\}. \tag{4.1}$$

Then the following four statements are equivalent.

(1) $g=P_K(f)$.

(2) (i) $e^{[i]}(b)=0$ *for* $1\leq i\leq n$ *and* $(-1)^n e^{[n]}(t)\leq 0$, $t\in S$.

(ii) $\int_a^b eg=0$.

(3) *Condition* (2)(i) *holds, and* g *is a polynomial of degree at most* $n-1$ *on each of the components* (= *maximal open subintervals*) *of the open set* E.

(4) *Condition* (2)(i) *holds, and* g *is a polynomial of degree at most* $n-1$ *on each of the components of* E *which contains an element of* S.

Proof. The equivalence of (1) and (2) follows immediately from Theorems 1.1 and 4.1.

Let $n\geq 2$. For convenience, let $J(s,t)=(-1)^n\int_{(s,t]} e^{[n]}\,dg_{\mathbb{R}}^{(n-1)}$, where $a\leq s<t<b$. Suppose now that (2) holds and (c,d) is a component of E. Let $s,t\in(c,d)$ and $0<\varepsilon<\min\{s-a,b-t\}$. Then $a+\varepsilon<s<t<b-\varepsilon$. Since $\mu_{g,n}(S')=0$ and $(-1)^n e^{[n]}\leq 0$ on S, we have $J(a+\varepsilon,b-\varepsilon)\leq J(s,t)\leq 0$. By Lemma 4.2 with $h=e$, we obtain $\int_a^b eg(\cdot;\varepsilon)=J(a+\varepsilon,b-\varepsilon)\leq J(s,t)\leq 0$. Using Lemma 2.4(2) and letting $\varepsilon\downarrow 0$ we find that $0=\int_a^b eg=J(s,t)$. Since $(-1)^n e^{[n]}<0$ on $(s,t]$, we conclude that $\mu_{g,n}(s,t]=g_{\mathbb{R}}^{(n-1)}(t)-g_{\mathbb{R}}^{(n-1)}(s)=0$. Hence $g_{\mathbb{R}}^{(n-1)}$ is constant on $[s,t]$. Since s, t are arbitrary, $g_{\mathbb{R}}^{(n-1)}$ is constant on (c,d). Thus (3) holds for $n\geq 2$. If $n=1$, we define $J(s,t)=(-1)^n\int_{(s,t]} e^{[n]}\,dg_{\mathbb{R}}$ and argue as above to conclude that (3) holds. Clearly, (3) implies (4).

Now suppose that (4) holds. If (c,d) is a component of E such that $(c,d)\cap S\neq\emptyset$, then $\mu_{g,n}(c,d)=0$. Hence, $\mu_{g,n}(S\cap E)=0$. Again, $\mu_{g,n}(S')=0$ and $e^{[n]}(t)=0$ for t in $S\backslash E$. Hence, $J(a+\varepsilon,b-\varepsilon)=0$ for all $0<\varepsilon(b-a)/2$. By Lemma 4.2 with $h=e$, we have $\int_a^b eg(\cdot;\varepsilon)=0$. Again, by Lemma 2.4, letting $\varepsilon\downarrow 0$ we conclude that $\int_a^b eg=0$. Thus (2) holds. The proof is complete.

If $p=2$, then the above theorem takes the following simpler form. Its proof is straightforward since $e=f-g$, as may be easily seen.

COROLLARY 4.4. *Let* $n\geq 1$, $S\subset(a,b)$, $K=K_{n,2}$, $f\in L_2\backslash K$, $g\in K$, *and*

$$E=E_g=\{t\in(a,b):f^{[n]}(t)<g^{[n]}(t)\}.$$

Then the following three statements are equivalent.

(1) $g = P_K((f)$.

(2) (i) $f^{[i]}(b) = g^{[i]}(b)$ *for* $1 \leqslant i \leqslant n$ *and* $(-1)^n f^{[n]}(t) \leqslant (-1)^n g^{[n]}(t)$, $t \in S$.

(ii) *g is a polynomial of degree at most $n-1$ on each of the components (=maximal open subintervals) of the open set E.*

(3) *Condition* (2)(i) *holds, and g is a polynomial of degree at most $n-1$ on each of the components of E which contains an element of S.*

The following result for $p=1$, which is analogous to Theorem 4.3, may be proved in the same way.

THEOREM 4.5. *Let $n \geqslant 1$, $S \subset (a, b)$, $K = K_{n,1}(S)$, $f \in L_1 \backslash K$, and $g \in K$. Then the following four statements are equivalent.*

(1) $g \in P_K(f)$.

(2) *There exists $e \in L_\infty$ satisfying*

(i) $\|e\|_\infty = 1$, $e = \operatorname{sgn}(f-g)$ *a.e., where* $f - g \neq 0$, $e^{[i]}(b) = 0$ *for* $1 \leqslant i \leqslant n$, $(-1)^n e^{[n]}(t) \leqslant 0$ *for* $t \in S$, *and*

(ii) $\int_a^b eg = 0$.

(3) *There exists $e \in L_\infty$ satisfying condition* (2)(i), *and g is a polynomial of degree at most $n-1$ on each of the components of the open set E defined by* (4.1).

(4) *There exists $e \in L_\infty$ satisfying condition* (2)(i), *and g is a polynomial of degree at most $n-1$ on each of the components of the open set E which contains an element of S.*

We remark that if $S = (a, b)$ (resp. $S = \Pi$ as defined in Section 1) then Theorems 4.3 and 4.5 reduce to the characterization of a best L_p-approximation from $K_{n,p}$ [30] (resp., n-convex splines in L_p of degree at most $n-1$ with simple knots [31]). These chararacterizations were obtained by using an extension of integration by parts. Our approach based on duality leads to a simpler yet more general proof.

5. L_2-APPROXIMATION BY NONDECREASING FUNCTIONS

In this section, we derive a more detailed characterization in the special case when $p=2$ and $n=1$. Recall that K_1 (resp., K_2) is the set of nondecreasing (resp., convex) functions. Let B (resp., C) be the set of all bounded (resp., continuous) functions on I (resp., on $[a, b]$). A function k in $K_2 \cap B$ is said to be the *greatest convex minorant* (*gcm*) of f in B if it is the largest convex function which does not exceed f at any point in I. Specifically,

$$k(s) = \sup\{h(s) : h \in K_2, h(t) \leqslant f(t), t \in I\}, \qquad s \in I.$$

Such a unique k clearly exists. It is shown in [33, Theorem 3.1] that if $f \in C$, then its gcm is also in C. For fixed $f \in C$ and $k \in C$ with $k \leqslant f$, define

$$E(k) = \{s \in I : k(s) < f(s)\}.$$

Then $E(k)$ is open in I. If k is the gcm of f, then it is shown in [33, Theorem 3.1] that $k(a) = f(a)$, $k(a) = f(b)$, and, hence, $E(k) \subset (a, b)$.

PROPOSITION 5.1. *Let $f \in C$, $k \in K_2 \cap C$ and $k \leqslant f$. Then k is the greatest convex minorant of f if and only if the following two conditions hold:*

(1) *$k(a) = f(a)$, and $k(b) = f(b)$.*

(2) *k is linear on each component of the open set $E(k)$.*

Proof. If k is the gcm of f, then the conditions follow by [33, Theorems 3.1 and 2.1(ii)].

Conversely, suppose that $g \in K_2 \cap C$, $g \leqslant f$, and the conditions hold for g. Also, let k be the gcm of f. We show that $g = k$. Note that $g \leqslant k \leqslant f$ and, by (1), $E(g) \subset (a, b)$. Let (c, d) be a component of $E(g)$. Then $g(c) = f(c)$. Also, $g(c) \leqslant k(c) \leqslant f(c)$. Hence, $g(c) = k(c)$. Similarly, $g(d) = k(d)$. Since g is linear on (c, d) and k is convex with $g \leqslant k$, we conclude that $g = k$ on (c, d). On $I \backslash E(g)$, we have $g = f$ and, hence, that $g = k = f$. The proof is complete.

THEOREM 5.2. *Let $K = K_{1,2}$, the set of nondecreasing functions in L_2, $f \in L_2 \backslash K$ and $g \in K$. Then $g = P_K(f)$ if and only if g is a.e. equal on I to the derivative of the greatest convex minorant of $f^{[1]}$ (the derivative exists a.e. on I).*

Proof. Since $e = f - g$, we have $e^{[1]} = f^{[1]} - g^{[1]}$. By Corollary 4.4 we find that $g^{[1]} \leqslant f^{[1]}$, $g^{[1]}(a) = f^{[1]}(a)$, $g^{[1]}(b) = f^{[1]}(b)$, and g is constant on each component of $G = \{s : g^{[1]}(s) < f^{[1]}(s)\}$. By Proposition 5.1, $g^{[1]}$ is the gcm of $f^{[1]}$. The proof is complete.

The above characterization was obtained in [21] for a *bounded* function f by methods of optimal control. We have thus generalized this result to any $f \in L_2$ by using duality methods.

6. L_1-APPROXIMATION AND PERFECT SPLINES

In this section, we characterize a best L_1-approximation to a continuous f from $K_{n,1}$ in terms of perfect splines of order n. Some interesting relations between best L_1-approximation from splines and perfect splines are investigated in [13, 28]. A *perfect spline* p of order n with knots at t_i, $1 \leqslant i \leqslant r$, with $a = t_0 < t_1 < \cdots < t_r < t_{r+1} = b$ is any function of the form [13]

$$p(t) = \sum_{i=0}^{n-1} a_i t^i + d \sum_{i=0}^{r-1} (-1)^i \int_{t_i}^{t_{i+1}} (t-s)_+^{n-1} \, ds.$$

Note that $p^{(n-1)}$ is continuous on (a, b) and $p^{(n)}(t) = (-1)^i (n-1)! \, d$ for all $t \in (t_i, t_{i+1})$, $0 \leqslant i \leqslant r$. We first establish a special characterization theorem for $p \geqslant 1$. Let $S_n(t_1, t_2, ..., t_r)$ denote the set of all polynomial spline functions of order n on I with simple knots at the points $t_1 < t_2 < \cdots < t_r$ in (a, b). By sign changes of a function we mean strong sign changes as in [25, p. 25, Definition 2.11].

THEOREM 6.1. *Let* $1 \leqslant p < \infty$, $n \geqslant 1$, $K = K_{n,p}(S)$, $f \in L_p \setminus K$, *and* $g \in K$. *Assume that* $f \neq g$ *a.e. on* (a, b) *and* $f - g$ *has* $m < \infty$ *sign changes in* (a, b). *Let* $e = |f-g|^{p-1} \operatorname{sgn}(f-g)$ *if* $1 \leqslant p < \infty$. *Let*

$$E = \{t \in (a, b) : (-1)^n e^{[n]}(t) < 0\}.$$

Then $e^{[n]}$ *has no more than* $m + n$ *distinct zeros in* (a, b).

The following two statements are equivalent.

(1) $g \in P_K(f)$.

(2) (i) $e^{[i]}(b) = 0$ *for* $1 \leqslant i \leqslant n$, *and* $(-1)^n e^{[n]}(t) \leqslant 0$, $t \in S$.

(ii) g *is a best* L_p-*approximation to* f *from* $S_n(t_1, t_2, ..., t_r)$, *where* t_i *are the distinct zeros of* $e^{[n]}$ *in* (a, b) *and* $r \leqslant m + n$. *(For* $p > 1$, *the function* g *is unique since* L_p *is uniformly convex.)*

Moreover, if $p = 1$ *and* f *is continuous on* $[a, b]$, *then the function* g *in* (1) *and* (2)(ii) *is unique.*

Proof. Let r be the number of distinct zeros of $e^{[n]}$ in (a, b). We show that $r \leqslant m + n$. Note that each $e^{[i]}$, $1 \leqslant i \leqslant n$, is continuous. By Rolle's theorem, $e^{[n-1]}$ has at least $r - 1$ zeros in (a, b). Repeating this argument we find that $e^{[1]}$ has at least $r - n + 1$ zeros in (a, b). Now if $c < d$ are two zeros of $e^{[1]}$, then $0 = e^{[1]}(d) - e^{[1]}(c) = \int_c^d e$. Since e and $f - g$ have the same sign changes, we conclude that e changes sign in (c, d). Thus the number of sign changes of e in (a, b) is at least $r - n$. Since $r - n \leqslant m$, the result follows.

Now we show the equivalence of (1) and (2). Let $g \in P_K(f)$. Then, (2)(i) holds by Theorems 4.3 and 4.5. Let the r zeros of $e^{[n]}$ in (a, b) be denoted by t_i as in (2)(ii). Clearly, $r \leqslant m+n$. Let $t_0 = a$, $t_{r+1} = b$, and $I_i = (t_i, t_{i+1})$, $0 \leqslant i \leqslant r$. By Theorems 4.3 and 4.5, g is a polynomial of degree at most $n-1$ on I_i. Hence, $g \in S_n(t_1, t_2, ..., t_r) = S_n$, say. Then using integration by parts as in Lemma 4.2 and the equalities, $e^{[i]}(a) = e^{[i]}(b) = 0$, $1 \leqslant i \leqslant n$, we obtain for $0 \leqslant i \leqslant r$,

$$\int_a^b e(t)(t-t_i)_+^{n-1}\,dt = (-1)^{n-1}(n-1)!\int_a^b (e^{[n]}(t))^{(1)}(t-t_i)_+^0\,dt$$

$$= (-1)^{n-1}(n-1)!\,(e^{[n]}(b) - e^{[n]}(t_i)) = 0.$$

It is known that $(t-t_i)_+^{n-1}$, $0 \leqslant i \leqslant r$, form a basis for S_n [25]. Hence, the above equation shows that $\int_a^b eh = 0$, for all $h \in S_n$. Therefore g is a best L_p-approximation to f from S_n. We have shown that (2) holds.

Conversely, let g satisfy (2). Then, g is a polynomial of degree at most $n-1$ on (t_i, t_{i+1}), $0 \leqslant i \leqslant r$, which are components of E. By Theorems 4.3 and 4.5, g is a best L_p-approximation to f from $K_{n,p}(S)$.

We now show the last statement. Note that $S_n(t_1, t_2, ..., t_r)$ is an A-space and a best L_1-approximation to a continuous f from this set is unique [19]. Hence g in (2) is unique. It remains to show that a best approximation from $K = K_{n,1}(S)$ is unique. Indeed, let $g, k \in P_K(f)$ and $e = \operatorname{sgn}(f-g)$. Then (2) holds and $g \in S_n(t_1, t_2, ..., t_r)$. Since $e \in K^0$, we have $\int_a^b ek \leqslant 0$. Hence, by a well known argument,

$$\|f-g\|_1 = \int_a^b e(f-g) = \int_a^b ef \leqslant \int_a^b e(f-k) \leqslant \|e\|_\infty\,\|f-k\|_1 = \|f-k\|_1.$$

Since $\|f-g\|_1 = \|f-k\|_1$, equality holds throughout and $\int_a^b ek = 0$. Then arguing as in the proof of Theorem 4.3 we conclude that k is a polynomial of degree at most $n-1$ on each component of E. Again arguing as in the part of the above proof which shows (1) implies (2), we obtain that $k \in S_n$, which is an A-space. Consequently, $g = k$ and the proof is complete.

Note that if $S_n = S_n(t_1, t_2, ..., t_r)$, then the above theorem shows that $d_p(f, K) = d_p(f, S_n) = d_p(f, K \cap S_n)$, where f and K are as in the theorem and $d_p(f, A)$ denotes the distance of f from A in L_p, $1 \leqslant p < \infty$. Now we state a theorem involving perfect splines.

THEOREM 6.2. *Let* $n \geqslant 1$, $K = K_{n,1}(S)$, $f \in L_1 \backslash K$, *and* $g \in K$. *Assume that* $f \neq g$ *a.e. on* (a, b) *and* $f-g$ *has* $m < \infty$ *sign changes in* (a, b) *at* s_i, $1 \leqslant i \leqslant m$, *where* $s_1 < s_2 < \cdots < s_m$. *Then the following two statements are equivalent.*

(1) $g \in P_K(f)$.

(2) *There is a perfect spline p of degree n with knots at s_i, $1 \leqslant i \leqslant m$, and distinct zeros at t_i, $1 \leqslant i \leqslant r$, in (a, b) with $t_1 < t_2 < \cdots < t_r$ such that the following four conditions hold.*

(i) $p^{(i)}(a) = p^{(i)}(b) = 0$, $0 \leqslant i \leqslant n-1$.

(ii) $p^{(n)} = (-1)^n \operatorname{sgn}(f-g)$ *a.e. in* (a, b).

(iii) $p(t) \leqslant 0$, $t \in S$.

(iv) *g is a best L_1-approximation to f from $S_n(t_1, t_2, ..., t_r)$.*

Moreover, if f is continuous on $[a, b]$, then the function g in (1) and (2)(iv) is unique.

Remark. The perfect spline p in (2) is given by $p = (-1)^n e^{[n]}$, where $e = \operatorname{sgn}(f-g)$).

Proof. Under the hypothesis, Theorem 6.1 applies with $p = 1$. Define $p(t) = (-1)^n e^{[n]}(t)$, where $e = \operatorname{sgn}(f-g)$. Let $s_0 = a$ and $s_{m+1} = b$. Then $p^{(n)}(t) = (-1)^n e(t) = \sigma(-1)^i$ for $t \in (s_i, s_{i+1})$, $0 \leqslant i \leqslant m$, where σ is the sign of $f-g$ on (s_0, s_1). With these arguments this theorem is a restatement of Theorem 6.1. The proof is complete.

Let W_n be the Sobolev space of real functions f on (a, b) such that $f^{(n-1)}$ exists and is absolutely continuous on (a, b), or, equivalently, $f^{(n)}$ exists a.e. on (a, b) and $f^{(n)} \in L_1$. We consider a problem on W_n equipped with the usual L_1 norm. Let $n \geqslant 1$, S be (relatively) closed in (a, b), $K = K_{n,1}(S)$, and $f \in W_n \backslash K$. Then, by Theorem 3.5, $P_K(f) \neq \varnothing$. Let $g \in P_K(f)$, and assume that $f \neq g$ a.e. on (a, b) and $f-g$ has $m < \infty$ sign changes in (a, b) at s_i, $1 \leqslant i \leqslant m$, where $s_1 < s_2 < \cdots < s_m$. Then, by Theorem 6.2, g is unique, and if $p_0 = ((-1)^n (f-g))^{[n]}$, then p_0 has $r \leqslant m+n$ zeros at $t_1 < t_2 < \cdots < t_r$ in (a, b). Let P_n denote the set of all perfect splines p of degree n with knots at s_i, $1 \leqslant i \leqslant m$, zeros at t_i, $1 \leqslant i \leqslant r$, and satisfying Theorem 6.2(2), conditions (i) and (iii). Then $p_0 \in P_n$. We consider the problem of finding $p_* \in P_n$ such that

$$\left| \int_a^b p_* f^{(n)} \right| \geqslant \left| \int_a^b p f^{(n)} \right|, \qquad \text{all} \quad p \in P_n.$$

The following theorem shows that $p_* = p_0$. We let $\Delta = \max\{|t_{i+1} - t_i| : 0 \leqslant i \leqslant r\}$, where $t_0 = a$ and $t_{r+1} = b$.

THEOREM 6.3. *For the above problem the following hold.*

(1) $\|f-g\|_1 = |\int_a^b p_0 f^{(n)}| \geqslant |\int_a^b p f^{(n)}|$, *for all* $p \in P_n$, *and*

(2) $\|f-g\|_1 \leqslant \min\{\Delta^n/(4n), (n-1)^{n-1} \Delta^n/(n!\, 2^n)\} \|f^{(n)}\|_1$.

Proof. (1) For convenience let $I_i = (t_i, t_{i+1})$, $0 \leqslant i \leqslant r$. For all $p \in P_n$, since $p^{(i)}(a) = p^{(i)}(b) = 0$, $0 \leqslant i \leqslant n-1$, integration by parts as in Lemma 4.2 yields

$$\int_a^b p^{(n)}(f-g) = (-1)^{n-1} \int_a^b p^{(1)}(f^{(n-1)} - g^{(n-1)}),$$

$$= (-1)^{n-1} \sum_{i=0}^{r} \int_{I_i} p^{(1)}(f^{(n-1)} - g^{(n-1)}).$$

Since $p(t_i) = 0$, again integration by parts gives

$$\int_{I_i} p^{(1)}(f^{(n-1)} - g^{(n-1)}) = -\int_{I_i} p(f^{(n)} - g^{(n)}).$$

By Theorem 6.2, $g \in S_n(t_1, t_2, ..., t_r)$. Consequently, $g^{(n)}(t) = 0$ for $t \in I_i$, $0 \leqslant i \leqslant r$. Also, $|p^{(n)}(t)| = 1$ for $t \neq t_i$. Hence we obtain, using the above equalities,

$$\left|\int_a^b pf^{(n)}\right| = \left|\int_a^b p(f^{(n)} - g^{(n)})\right| = \left|\int_a^b p^{(n)}(f-g)\right| \leqslant \|f-g\|_1.$$

Since $p_0^{(n)} = (-1)^n \operatorname{sgn}(f-g)$ a.e., we have

$$\int_a^b p_0 f^{(n)} = \int_a^b p_0^{(n)}(f-g) = (-1)^n \int_a^b |f-g| = (-1)^n \|f-g\|_1.$$

This establishes (1).

(2) By an estimate given in [14] we have $\|p_0\|_\infty \leqslant \Delta^n/(4n)\, \|p_0^{(n)}\|_\infty$, and

$$\|p_0\|_\infty \leqslant (n-1)^{n-1}\, \Delta^n/(n!\, 2^n)\, \|p_0^{(n)}\|_\infty.$$

Using (1) we obtain $\|f-g\|_1 \leqslant \|p_0\|_\infty \|f^{(n)}\|_1$. From these three inequalities and the fact that $\|p_0^{(n)}\|_\infty = 1$, we obtain (2).

The proof is complete.

References

1. P. S. Bullen, A criterion for n-convexity, *Pacific J. Math.* **36** (1971), 81–98.
2. C. K. Chui, F. Deutsch, and J. D. Ward, Constrained best approximation in Hilbert space, *Constr. Approx.* **6** (1990), 35–64.
3. C. K. Chui, F. Deutsch, and J. D. Ward, Constrained best approximation in Hilbert space, II, *J. Approx. Theory* **71** (1992), 213–238.

4. C. DEBOOR, The quasi-interpolant as a tool in elementary polynomial spline theory, *in* "Approximation Theory" (G. G. Lorentz, Ed.), pp. 269–276, Academic Press, New York, 1973.
5. F. R. DEUTSCH, "Some Applications of Functional Analysis to Approximation Theory," Doctoral dissertation, Brown University, 1965.
6. F. R. DEUTSCH AND P. H. MASERICK, Applications of the Hahn–Banach theorem in approximation theory, *SIAM Rev.* **9** (1967), 516–530.
7. N. DUNFORD AND J. T. SCHWARTZ, "Linear Operators, Part I," Interscience, New York, 1958.
8. R. HUOTARI, R. LEGG, AND D. TOWNSEND, Existence of best n-convex approximants in L_1, *Approx. Theory Appl.* **5** (1989), 51-57.
9. S. KARLIN AND A. NOVIKOFF, Generalized convex inequalities, *Pacific J. Math.* **13** (1963), 1251–1279.
10. S. KARLIN AND W. J. STUDDEN, "Tchebycheff Systems: With Applications in Analysis and Statistics," Interscience, New York, 1966.
11. D. LANDERS AND L. ROGGE, Isotonic approximation in L_S, *J. Approx. Theory* **31** (1981), 199–223.
12. R. M. MATHSEN AND V. A UBHAYA, Generalized convex functions and best L_p-approximation, *Proc. Amer. Math. Soc.* **114** (1992), 733–740.
13. C. A. MICCHELLI, Best L^2 approximation by weak Chebyshev systems and the uniqueness of interpolating perfect splines, *J. Approx. Theory* **19** (1977), 1–14.
14. C. A. MICCHELLI, T. J. RIVLIN, AND S. WINOGRAD, The optimal recovery of smooth functions, *Numer. Math.* **26** (1976), 191–200.
15. C. A. MICCHELLI, P. W. SMITH, J. SWETITS, AND J. D. WARD, Constrained L_p approximation, *Constr. Approx.* **1** (1985), 93–102.
16. C. A. MICCHELLI AND F. UTRERAS, Smoothing and interpolation in a convex subset of a Hilbert space, *SIAM J. Sci. Statist. Comput.* **9** (1988), 728–746.
17. A. MUKHERJEA AND K. POTHOVEN, "Real and Functional Analysis," Part A, "Real Analysis," Plenum, New York, 1984.
18. I. P. NATANSON, "Theory of Functions of a Real Variable," Ungar, New York, 1964.
19. A. M. PINKUS, "On L^1-Approximation," Cambridge Univ. Press, Cambridge, 1989.
20. T. POPOVICIU, "Les Fonctions Convexes," Hermann, Paris, 1944.
21. W. T. REID, A simple optimal control problem involving approximation by monotone functions, *J. Optim. Theory Appl.* **2** (1968), 365–377.
22. A. W. ROBERTS AND D. E. VARBERG, "Convex Functions," Academic Press, New York, 1973.
23. R. T. ROCKAFELLAR, "Convex Analysis," Princeton Univ. Press, Princeton, NJ, 1970.
24. G. SH. RUBINSTEIN, On an extremal problem in a linear normed space, *Sibirsk. Mat. Zh.* **6** (1965), 711–714. [in Russian]
25. L. L. SCHUMAKER, "Spline Functions: Basic Theory," Wiley, New York, 1981.
26. I. SINGER, "The Theory of Best Approximation and Functional Analysis," SIAM, Philadelphia, 1974.
27. P. W. SMITH AND J. J. SWETITS, Best approximation by monotone functions, *J. Approx. Theory* **44** (1987), 398–403.
28. H. STRAUSS, Best L_1-approximation, *J. Approx. Theory* **41** (1984), 297–308.
29. J. J. SWETITS, S. E. WEINSTEIN, AND Y. XU, On the characterization and computation of best monotone approximation in $L_p[0, 1]$ for $1 \leq p < \infty$, *J. Approx. Theory* **60** (1990), 58–69.
30. J. J. SWETITS, S. E. WEINSTEIN, AND Y. XU, Approximation in $L_p[0, 1]$ by n-convex functions, *Numer. Funct. Anal. Optim.* **11** (1990), 167–179.
31. J. J. SWETITS, S. E. WEINSTEIN, AND Y. XU, Best L_p-approximation with multiple constraints for $1 \leq p < \infty$, *J. Approx. Theory* **65** (1991), 90–108.

32. V. A. UBHAYA, Duality in approximation and conjugate cones in normed linear spaces, *J. Math. Anal. Appl.* **58** (1977), 419–436.
33. V. A. UBHAYA, An $O(n)$ algorithm for discrete n-point convex approximation with applications to continuous case, *J. Math. Anal. Appl.* **72** (1979), 338–354.
34. V. A. UBHAYA, L_p approximation from nonconvex subsets of special classes of functions, *J. Approx. Theory* **57** (1989), 223–238.
35. V. A. UBHAYA, Duality and Lipschitzian selections in best approximation from nonconvex cones, *J. Approx. Theory* **64** (1991), 315–342.
36. D. ZWICK, Existence of best n-convex approximations, *Proc. Math. Soc.* **97** (1986), 273–276.
37. D. ZWICK, Characterizing shape preserving L_1-approximation, *Proc. Amer. Math. Soc.* **103** (1988), 1139–1146.