JOURNAL OF APPROXIMATION THEORY 80, 423–449 (1995)

Constrained L_p-Approximation by Generalized n-Convex Functions Induced by ECT-Systems

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 A. Micchelli

Received June 25, 1993; accepted in revised form February 11, 1994

The problem of finding a best L_p-approximation ($1 \leqslant p < \infty$) to a function in L_p from a special subcone of generalized n-convex functions induced by an ECT-system is considered. Tchebycheff splines with a countably infinite number of knots are introduced and best approximations are characterized in terms of local best approximations by these splines. Various properties of best approximations and their uniqueness in L_1 are investigated. Some special results for generalized monotone and convex cases are obtained.

1. INTRODUCTION

Let $I = (a, b)$. A set of functions $U_n = \{u_i\}_{i=0}^{n-1}$ in $C^{n-1}(I)$, $n \geqslant 1$, is called an ECT-system on I in its canonical form if there exist positive weight functions w_i in $C^{n-i}[a, b]$, $0 \leqslant i \leqslant n-1$, such that for all $x \in I$ [7, 15]

$$\begin{aligned} u_0(x) &= w_0(x), \\ u_1(x) &= w_0(x) \int_a^x w_1(t_1)\, dt_1, \\ &\ \vdots \\ u_{n-1}(x) &= w_0(x) \int_a^x w_1(t_1) \int_a^{t_1} w_2(t_2) \cdots \int_a^{t_{n-2}} w_{n-1}(t_{n-1})\, dt_{n-1} \cdots dt_1. \end{aligned} \tag{1.1}$$

0021-9045/95 $6.00

The space $V_n = \operatorname{span} U_n$ is called an ECT-space; it is a Haar space of dimension n, i.e., any function in V_n has no more than $n-1$ zeros in I. For this reason, functions in V_n are called generalized polynomials of degree at most $n-1$.

A real-valued function k on I is said to be a generalized n-convex function induced by U_n or a U_n-convex function if for any $n+1$ points $x_0 < x_1 < \cdots < x_n$ in I,

$$\det(u_i(x_j))_{i,j=0}^{n} \geqslant 0, \tag{1.2}$$

with $u_n = k$. Similarly, a generalized n-concave function is defined with $\geqslant$ replaced by $\leqslant$. Let $K_n = K_n(I)$ be the set of all generalized n-convex functions. The functions in K_1 (resp. K_2) are also called generalized monotone (resp. convex) functions. We remark that if $w_i = 1$ for $0 \leqslant i \leqslant n-1$, then $U_n = \{(x-a)^i\}_{i=0}^{n-1}$, and the corresponding K_n is the set of ordinary n-convex functions. It is easy to see that K_n is a convex cone considered as a subset of real functions on I.

For $1 \leqslant p < \infty$, let $L_p = L_p(I)$ with the norm $\|\cdot\|_p$. Given $K \subset L_p$ and $f \in L_p$, a function $g \in K$ is said to be a best L_p-approximation to f from K if $\|f-g\|_p = \inf\{\|f-k\|_p : k \in K\}$. The dual cone K^0 of a set $K \subset L_p$ is defined by

$$K^0 = \left\{h \in L_q : \int_a^b hk \leqslant 0 \text{ for all } k \in K\right\}, \qquad 1/p + 1/q = 1. \tag{1.3}$$

The dual cone is known to play a significant role in approximation [4, 20, 23].

In this work, we are concerned with a special convex subcone $K_n(S) = K_n(I, S)$ of K_n where $S \subset I$. This cone will be defined later. We let $K_{n,p}(S) = K_n(S) \cap L_p$, $1 \leqslant p < \infty$. This is a set from which we seek a best approximation to $f \in L_p$. When $S \neq I$, $K_{n,p}(S)$ is a proper *constrained* subcone of K_n in L_p; the problem is unconstrained if $S = I$. Such subcones arose naturally in the problems of best constrained approximation ([2] or [3]) which in turn arose from smoothing and interpolation problems (e.g., [10, 11]). With this motivation (for further details see [5]) the case of L_p-approximation by the subcone of ordinary n-convex functions was considered in [5]; here a characterization of the dual cone of n-convex functions was fundamental in establishing the main results. In this article, we investigate the structure and characterization of a best L_p-approximation to f in L_p from $K_{n,p}(S)$, the subcone of generalized n-convex functions, for $1 \leqslant p < \infty$, and the uniqueness of the approximation for $p = 1$. All the main results (included in Sections 3–6) are new even when they are specialized to L_p-approximation by n-convex functions. In Section 2, we extend certain results in [5] on dual cone, existence, and basic characterization of a best approximation to our framwork, and using [9], obtain some properties of U_n-convex functions. In Section 3, we develop the new concept of Tchebycheff splines with a countably infinite number of knots, and in Section 4 we apply it to obtain alternate characterizations of best approximations which are different from those available for monotone and n-convex problems in [5, 17, 18]. In Section 5, we investigate the spline-like structure, boundedness, and uniqueness (in L_1) of best approximations under certain conditions. In Section 6, we establish additional characterizations and uniqueness (in L_1) for best approximations by generalized monotone and convex functions.

We now define the subcone $K_n(S)$ introduced above. Let μ denote the Lebesgue–Stieltjes complete measure generated by a nondecreasing function g on I. Then, for each Borel set $A \subset I$, we have $\mu(A) = \inf\{\sum_{i=1}^{\infty} (g(b_i - g(a_i)) \colon A \subset \bigcup_{i=1}^{\infty} (a_i, b_i), (a_i, b_i) \subset I\}$, and μ is the completion of this measure on the Borel sets [12]. We denote by D^-, D^+, and D, respectively, the left and right derivative and the derivative of a function. For $k \in K_n$, we define

$$l_{n-1}^+ k = \frac{1}{w_{n-1}} D^+ \frac{1}{w_{n-2}} D \cdots \frac{1}{w_1} D\left(\frac{k}{w_0}\right), \qquad \text{for } n \geqslant 2, \tag{1.4}$$

and $(l_0^+ k)(t) = k(t^+)/w_0(t)$, for $n=1$. We define l_{n-1}^- and l_0^- analogously with $+$ replaced by $-$. Note that $l_{n-1}^+ k$ (resp. $l_{n-1}^- k$) is right-continuous (resp. left-continuous) and nondecreasing [7]. Let $S \subset I$ be any Borel set and $S' = I \backslash S$. We denote by $\mu_{k,n}$ the Lebesgue–Stieltjes measure generated by $l_{n-1}^+ k$ on I. Define a convex subcone of K_n by $K_n(S) = K_n(I, S) = \{k \in K_n \colon \mu_{k,n}(S') = 0\}$. 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. In particular, $K_n = K_n(I)$, $K_n(\varnothing) = V_n$, and $K_n(\pi)$ is the set of all U_n-convex Tchebycheff splines on I with simple knots at $\pi = \{t_1 < t_2 < \cdots < t_m\} \subset I$, where $\varnothing$ denotes the empty set.

The problems of unconstrained L_p-approximation by ordinary n-convex functions and generalized convex functions defined by a nonlinear family are considered in [6, 9, 17–19, 22, 23]. (Recall that 1-convex and 2-convex functions are monotone nondecreasing and convex, respectively.)

2. Existence and Basic Characterizations of a Best Approximation

In this section we obtain some properties of U_n-convex functions which will be used in our analysis. We also extend some of the main results of [5] regarding the dual cone, and the existence and characterization of a best approximation to our framework as a starting point of this article.

First, we present alternative definitions of U_n-convex functions which will be used in the sequel. If $U_n = \{u_i\}_{i=0}^{n-1}$, as in (1.1), is a canonical ECT-system of n functions on I, then there exists a function u_n so that $U_{n+1} = \{u_i\}_{i=0}^{n}$ forms a canonical ECT-system of $n+1$ functions on I [15, Theorem 9.4]. For our purpose, we choose $w_n = 1$ and let

$$u_n(x) = w_0(x) \int_a^x w_1(t_1) \int_a^{t_1} w_2(t_2) \cdots \int_a^{t_{n-1}} w_n(t_n)\, dt_n \cdots dt_1.$$

For $a < x_0 < x_1 < \cdots < x_n < b$, we define the nth order divided difference with respect to U_{n+1} by $[x_0, x_1, ..., x_n] f = \det(v_i(x_j))_{i,j=0}^n / \det(u_i(x_j))_{i,j=0}^n$, where $v_i = u_i$ for $0 \leqslant i \leqslant n-1$ and $v_n = f$ [15, p. 368]. Since U_{n+1} is an ECT-system, $\det(u_i(x_j))_{i,j=0}^n$ in the denominator of the right hand side of the above expression is strictly positive. Hence, k is U_n-convex if and only if, for any $n+1$ points $x_0 < x_1 < \cdots < x_n$ in I,

$$[x_0, x_1, ..., x_n] k \geqslant 0. \tag{2.1}$$

Furthermore, expanding the determinant (1.2) as in [14, p. 250, Theorem E], we may show the following: k is U_n-convex if and only if, whenever $x_1 < x_2 < \cdots < x_n$ are n points in I and $h \in V_n$ satisfies $h(x_i) = g(x_i)$, $1 \leqslant i \leqslant n$, then

$$(-1)^{n+i-1} (k(t) - h(t)) \geqslant 0, \qquad t \in (x_{i-1}, x_i), \quad 1 \leqslant i \leqslant n+1, \tag{2.2}$$

where $x_0 = a$ and $x_{n+1} = b$. See also [9, Sect. 1] for other definitions similar to (2.2). We observe that the interpolating function h appearing in the above definition is unique since $\det(u_i(x_j))_{i,j=0}^n \neq 0$.

We now introduce some notation and terminology. The Green's function associated with U_n is defined by

$$\begin{aligned} G_{n-1}(x, t) &= w_0(x) \int_t^x w(t_1) \\ &\quad \times \int_t^{t_1} w_2(t_2) \cdots \int_t^{t_{n-2}} w_{n-1}(t_{n-1})\, dt_{n-1} \cdots dt_1, \qquad t \leqslant x < b, \\ &= 0, \qquad a < x < t. \end{aligned}$$

For $n = 1$, $G_0(x, t) = w_0(x)$ for $t \leqslant x < b$ and zero elsewhere. We construct a B-spline by

$$B_{0,n}(t) = [x_0, x_1, ..., x_n]\, G_{n-1}(x, t). \tag{2.3}$$

A proof as in [15, Lemma 4.24] shows $B_{0,n}(t) > 0$ if $t \in (x_0, x_n)$ and $B_{0,n}(t) = 0$, otherwise. It follows that $G_{n-1}(x, t)$ is a U_n-convex function of

x for each t. Let f be a continuous function on I. Following [9], we say that f has r alternating local extrema if there exist points $x_1 < x_2 < \cdots < x_r$ in I such that exactly one of the following two conditions hold: (i) Points x_i with odd (resp. even) indices are local maxima (resp. minima) with $(-1)^i f(x_{i-1}) > (-1)^i f(x_i)$ for $2 \leqslant i \leqslant r$. (ii) Points x_i with odd (resp. even) indices are local minima (resp. maxima) with $(-1)^i f(x_{i-1}) < (-1)^i f(x_i)$ for $2 \leqslant i \leqslant r$. A constant function has zero alternating local extrema. If f has r local extrema at $\{x_i\}_{i=1}^r$, then f is monotone (i.e., nondecreasing or nonincreasing) on each (x_{i-1}, x_i), $1 \leqslant i \leqslant r+1$, where $x_0 = a$ and $x_{r+1} = b$.

LEMMA 2.1. *Let $v \in V_n$, $n \geqslant 2$. Then v/w_0 has at most $n-2$ alternating local extrema in I.*

Proof. The derivatives $(u_i/w_0)'$ of u_i/w_0, $1 \leqslant i \leqslant n-1$, span an ECT-space V_{n-1} of dimension $n-1$. If $v \in V_n$ then $(v/w_0)' \in V_{n-1}$ and hence, $(v/w_0)'$ has at most $n-2$ zeros. A subset of these zeros clearly forms the alternating local extrema of v/w_0. ∎

Letting $\bar{K}_n = \{k/w_0 : k \in K_n\}$, $\bar{U}_n = \{u_i/w_0\}_{i=0}^{n-1}$, and dividing both sides of (2.2) by $w_0(t)$, we conclude that functions in $\bar{K}_n$ are $\bar{U}_n$-convex on I. By Lemma 2.1, therefore, all the results of [9] hold for $\bar{K}_n$, since the existence of alternating local extrema is a condition in that article. The next proposition gives four properties of U_n-convex functions, the first property for n-convex functions is also observed in [22, p. 236].

PROPOSITION 2.2. *Let k be a U_n-convex function, $n \geqslant 1$.*

(1) *There exist an integer r, $1 \leqslant r \leqslant n$, and points $\{x_i\}$ with $a = x_0 < x_1 < \cdots < x_r = b$ such that the following holds: If $r = n$, then $(-1)^{n+i} k/w_0$ is nondecreasing on (x_{i-1}, x_i) for all $1 \leqslant i \leqslant n$. If $r < n$, then $(-1)^{r+i} k/w_0$ (or equivalently $(-1)^i k/w_0$) is nondecreasing on (x_{i-1}, x_i) for all i or nonincreasing on (x_{i-1}, x_i) for all i. The integer r and points $\{x_i\}$ depend upon k. Hence, k has at most n sign-changes in I.*

(2) *$k(a^+)$ and $k(b^-)$ exist and are possibly infinite. If $|k(a^+)| = \infty$ (resp. $|k(b^-)| = \infty$), then $(-1)^n k(a^+) = \infty$ (resp. $k(b^-) = \infty$).*

(3) *Let $[c, d] \subset I$, then k is bounded on $[c, d]$ and $-\infty < l_{n-1}^+ k(c) \leqslant l_{n-1}^- k(d) < \infty$.*

(4) *If $a < x_0 < x_1 < \cdots < x_n < b$, then $[x_0, x_1, ..., x_n]k = 0$ if and only if $k \in V_n$ on (x_0, x_n).*

Proof. (1) If $n = 1$, then k/w_0 is nondecreasing and hence the statements hold with $r = n = 1$. For $n \geqslant 2$, by Lemma 2.1, this is a restatement of [9, Theorem 2.1(b)] as applied to $\bar{K}_n$.

(2) By (1), k/w_0 is monotone (nonincreasing or nondecreasing) on some interval (a, y) where $y \in I$. Therefore $k(a^+)/w_0(a^+)$ and hence, $k(a^+)$ exists and is possibly infinite. If x_i and h are as in the definition (2.2) of k, then $(-1)^n k \geqslant (-1)^n h$ on (a, x_1). But since $|h| < \infty$ on I, if $k(a^+) = \infty$, then we must have $(-1)^n k(a^+) = \infty$. The proof for point b is similar.

(3) If $n = 1$, then $-\infty < k(c)/w_0(c) \leqslant k(t)/w_0(t) \leqslant k(d)/w_0(d) < \infty$, $t \in [c, d]$, and since w_0 is continuous we conclude that k is bounded on $[c, d]$. For $n \geqslant 2$, the boundedness of k on $[c, d]$ follows from the continuity of k on $[c, d]$. To prove the remaining statement, let $0 < \varepsilon < \min\{c - a, b - d\}$. Then, by [7, Chap XI, Theorem 2.3], there exists a U_n-convex function $k(t, \varepsilon)$ such that $-\infty < l_{n-1}^+ k(a, \varepsilon)$ and $k(t, \varepsilon) = k(t)$ for all $t \in (a + \varepsilon, b - \varepsilon)$. Then $-\infty < l_{n-1}^+ k(a, \varepsilon) \leqslant l_{n-1}^+ k(c, \varepsilon) = l_{n-1}^+ k(c)$. By a symmetric argument, we have $l_{n-1} k(d) < \infty$.

(4) Let $n = 1$. Then $[x_0, x_1]k = 0$ if and only if $k(x_0)/w_0(x_0) = k(x_1)/w_0(x_1)$. This is equivalent to $k = \lambda w_0$ for some real λ since k/w_0 is nondecreasing. Now suppose that $n \geqslant 2$ and let $J = (x_0, x_n)$. We assert that for some $v \in V_n$ and for all $x_0 \leqslant x \leqslant x_n$,

$$k(x) = v(x) + \int_J G_{n-1}(x, t)\, d\mu_{k,n}(t). \tag{2.4}$$

By (3), $l_{n-1}^+ k(x_0) > -\infty$. Hence, by [7, Chap. XI, Lemma 2.2(b)] with $a = x_0$ and $b = x_n$, we conclude that (2.4) holds for all $x_0 < x < x_n$. Now the set of functions $\{G_{n-1}(x, t) : t \in J\}$ is equicontinuous in the variable x. Also, k and v are continuous. Thus, (2.4) holds at x_0 and x_n proving our assertion. Now by applying the linear functional of divided difference to (2.4) and using (2.3), we obtain $[x_0, x_1, ..., x_n]k = \int_J B_{0,n}(t)\, d\mu_{k,n}(t)$, since $[x_0, x_1, ..., x_n]v = 0$. If $[x_0, x_1, ..., x_n]k = 0$, then $\int_J B_{0,n}(t)\, d\mu_{k,n}(t) = 0$. However, $B_{0,n}(t) > 0$ for all $t \in J$. Hence, $\mu_{k,n}(J) = 0$, and, by (2.4), we have $k = v$ on $[x_0, x_n]$. Conversely, if $k = v$, then, clearly, $[x_0, x_1, ..., x_n]k = 0$. ∎

For $A \subset L_p$, we denote by $cc(A)$ and $\overline{cc}_p(A)$, respectively, the smallest convex cone and the smallest L_p-closed convex cone containing A. Clearly, $\overline{cc}_p(A)$ is the L_p-closure of $cc(A)$. If $K \subset L_p$ is closed convex cone, then a proper subset M of K is called a generating basis for K if $K = \overline{cc}_p(M)$. The following set of functions $M_n(S)$ of variable x will be shown to generate $K_{n,p}(S)$ if S is closed: $M_n(S) = \{\pm u_i(x) : 0 \leqslant i \leqslant n-1\} \cup \{G_{n-1}(x, t) : t \in S\}$. In the rest of this section, we state five theorems. Their proofs are similar to those of [5, Theorems 3.2, 3.5, 4.1, 4.3, and 4.5] and hence are omitted.

THEOREM 2.3. $K_{n,p}(S) \subset \overline{cc}_p(M_n(S))$, $1 \leqslant p < \infty$, $n \geqslant 1$.

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 I for some sequence $\{f_j\}$ in P. Such sets have found applications in proving the existence of a best approximation [9, 22]. The definition of $\overline{\overline{P}}$ given here is as in [9] but weaker than the one in [22]; however, it will be seen that all the results of [22] hold with this change.

THEOREM 2.4. *Let $1 \leqslant p < \infty$ and $n \geqslant 1$. The following six statements are equivalent.*

(1) *S is closed in I (i.e., in the relativized topology for I.*

(2) *If $\{k_j\}$ is a sequence in $K_n(S)$, such that k_j converges pointwise to a real function k on I, then $k \in K_n(S)$.*

(3) *$K_{n,p}(S) = K_n(S) \cap L_p = \overline{\overline{K_n(S)}} \cap L_p$. (This implies that $K_{n,p}(S)$, $1 < p < \infty$, is a Tchebycheff set.)*

(4) *$K_{n,p}(S)$ is proximinal in L_p.*

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

(6) *$K_{n,p}(S) = \overline{cc}_p(M_n(S))$.*

For $h \in L_1$, we let $h^{[0]}(x) = h(x)$, and $h^{[i]} = \int_a^x w_{i-1}(t)\, h^{[i-1]}(t)\, dt$, $1 \leqslant i \leqslant n$. Recall that the dual cone is defined by (1.3). The following theorem gives a characterization $(K_{n,p}(S))^0$, the dual cone of $K_{n,p}(S)$.

THEOREM 2.5. *For $n \geqslant 1$, $1 \leqslant p < \infty$, and all $S \subset I$,*

$$\begin{aligned} K^0_{n,p}(S) &= (M_n(S))^0 \\ &= \{h \in L_q : h^{[i]}(b) = 0,\ 1 \leqslant i \leqslant n,\ (-1)^n h^{[n]}(t) \leqslant 0,\ t \in S\}, \end{aligned}$$

where $1/p + 1/q = 1$.

The next two theorems give characterizations of a best L_p-approximation to $f \in L_p$ from $K_{n,p}(S)$, for $1 < p < \infty$ and $p = 1$, respectively. For $1 < p < \infty$, uniform convexity of L_p ensures uniqueness of a best approximation.

THEOREM 2.6. *Let $1 < p < \infty$, $n \geqslant 1$, $f \in L_p \backslash K_{n,p}(S)$, $g \in K_{n,p}(S)$, and $e = |f - g|^{p-1} \operatorname{sgn}(f - g)$. Define*

$$E^- = \{t \in I : (-1)^n e^{[n]}(t) < 0\}. \tag{2.5}$$

Then, g is the best L_p-approximation to f from $K_{n,p}(S)$ if and only if the following three statements hold.

(1) $e^{[i]}(b)=0,\ 1\leqslant i\leqslant n.$

(2) $(-1)^n e^{[n]}(t)\leqslant 0,\ t\in S.$

(3) *g is a generalized polynomial of degree at most $n-1$ on each component of the open set E^{-}, or $\int_I eg=0$.*

THEOREM 2.7. *Let $p=1$, $n\geqslant 1$, $f\in L_1\backslash K_{n,1}(S)$, and $g\in K_{n,1}(S)$. Define*

$$D(f-g)=\{e\in L_\infty : e=\operatorname{sgn}(f-g)\ a.e., \text{ where } f-g\neq 0\}. \tag{2.6}$$

Then, g is a best L_1-approximation to f from $K_{n,1}(S)$ if and only if there exists $e\in D(f-g)$, such that the statements (1), (2), and (3) of Theorem 2.6 are satisfied, where E^{-} is defined by (2.5) using this e.

3. TCHEBYCHEFF SPLINES WITH COUNTABLE KNOTS

In this section, we develop the new concept of Tchebycheff splines with a countably infinite number of knots. In the next section we show that such splines arise naturally in the characterization of a best approximation from $K_{n,p}(S)$. To attain compatibility with later sections, we develop these concepts on an arbitrary interval $J=(c,d)\subset I$. If f is a function on I, we denote its restriction to J by $f\,|\,J$. Similarly, if F is a set of functions on I, then $F\,|\,J$ denotes the set $\{f\,|\,J\colon f\in F\}$. Assume that I_i is a finite or countably infinite family of disjoint open intervals such that

$$[c,d]=\operatorname{cl}\Big(\bigcup_i I_i\Big), \tag{3.1}$$

where cl denotes the closure operation on the reals. Let π be the set consisting of the endpoints of these intervals; π is called a partition of J. If π is finite then the endpoints c and d are in π; however, this may not be the case if π is infinite (see Lemma 3.1 below). If π is infinite and for all arbitrarily small $\varepsilon>0$, the interval $(c+\varepsilon, d-\varepsilon)$ contains only a finite number of points of π, then π is called a regular infinite partition of J. If π is finite or infinite, we let $T=T(J,\pi)=\{k\in C^{n-2}(J)\colon k\in V_n\,|\,I_i \text{ for all } i\}$. Clearly, T is a vector space. A function in T will be described as a Tchebycheff spline (relative to U_n) of order $n-1$ on J with simple knots at the points of $J\cap\pi$. (Since J is open, its endpoints are not in J, although they may be in π). If π is finite, then T is the space of ordinary Tchebycheff splines much investigated in the literature [15]. If π is regular infinite, then a k in T is a spline with a countable number of knots; furthermore, for any

$\varepsilon > 0$, k is an ordinary spline on $(c+\varepsilon, d+\varepsilon)$ with a finite number of knots at the points of $(c+\varepsilon, d-\varepsilon) \cap \pi$. We let

$$T_p(J, \pi) = T(J, \pi) \cap L_p(J), \qquad 1 \leqslant p < \infty, \tag{3.2}$$

where $L_p(J) = L_p | J$ is the usual L_p-space of functions on J. We now define a subspace $T'_p(J, \pi)$ which is of significance in our analysis. Let

$$F_n = \{u_i, 0 \leqslant i \leqslant n-1\} \cup \{G_{n-1}(\cdot, t): t \in J \cap \pi\}, \tag{3.3}$$

and

$$T'_p(J, \pi) = \mathrm{cl}_p(\mathrm{span}(F_n | J)), \tag{3.4}$$

where cl_p denotes the closure operation in $L_p(J)$. Clearly, F_n or $F_n | J$ is a linearly independent set; this follows immediately as in the case when π is finite [15]. It will be seen in Theorem 3.2 below that $T'_p(J, \pi)$ is a closed subspace of $T_p(J, \pi)$. We remark that if π is regular infinite, then $T(J, \pi)$ is the pointwise closure of $\mathrm{span}(F_n | J)$. To see this let $k \in T(J, \pi)$ and $J' = (c+\varepsilon, d-\varepsilon)$ where $\varepsilon > 0$. Then, as was observed before, $k | J'$ is an ordinary spline with a finite number of knots in $J' \cap \pi$. Thus $k | J' \in \mathrm{span}(F_n | J')$ and the result follows. We now collect some properties of π in the following lemma. Its simple proof is left to the reader.

LEMMA 3.1. *Let $J = (c, d) \subset I$ and π be a partition of J. If π is finite then $c, d \in \pi$ and π has the form $c = t_0 < t_1 < \cdots < t_m = d$ with $I_i = (t_{i-1}, t_i)$, $1 \leqslant i \leqslant m$. If π is infinite, then π is regular if and only if the set of accumulation points of π is nonempty and is contained in $\{c, d\}$. Furthermore, c or d is an accumulation point of π if and only if it is not a point of π. In this case π has one of the following three possible forms.*

(1) $\pi: \cdots < t_{-i} < \cdots < t_0 < t_1 < \cdots < t_i < \cdots$, *where* $c, d \notin \pi$; $t_{-i} \downarrow c$, $t_i \uparrow d$ *as* $i \to \infty$; *and* $I_i = (t_{i-1}, t_i)$, $-\infty < i < \infty$.

(2) $\pi: c = t_0 < t_1 < \cdots < t_i < \cdots$, *where* $c \in \pi$, $d \notin \pi$; $t_i \uparrow d$ *as* $i \to \infty$; *and* $I_i = (t_{i-1}, t_i)$, $1 \leqslant i < \infty$.

(3) $\pi: \cdots < t_{-i} < \cdots < t_{-1} < t_0 = d$, *a case symmetric to the above.*

Recall that $V_n = \mathrm{span}\{u_i: 0 \leqslant i \leqslant n-1\}$ and let $X_n = \mathrm{cl}_p(\mathrm{span}\{G_{n-1}(\cdot, t): t \in J \cap \pi\})$. The next theorem gives several properties of $T_p(J, \pi)$ and $T'_p(J, \pi)$. For the proof of parts (1)–(3), we use the methods of [22].

THEOREM 3.2. *Let $J = (c, d) \subset I$ and π be a finite or regular infinite partition of J. Let also $1 \leqslant p < \infty$, and $\|\cdot\|_p$ denote the norm in $L_p(J)$. Then the following holds.*

(1) *Let* (k_m) *be a sequence of functions in* $T_p(J, \pi)$ *such that* $\|k_m\|_p \leqslant D$ *for all* j *and some* $D > 0$. *Then there exists a subsequence* (h_m) *of* (k_m) *and* $h \in T_p(J, \pi)$ *such that* h_m *converges pointwise to* h *on* J *and* $\|h\|_p \leqslant D$.

(2) $T_p(J, \pi)$ *is proximinal in* $L_p(J)$ *and hence closed.*

(3) $T'_p(J, \pi)$ *is a closed subspace of* $T_p(J, \pi)$ *and proximinal in* $L_p(J)$.

(4) $T'_p(J, \pi) = V_n | J + X_n | J$.

Proof. If π is finite then the theorem holds by the theory of ordinary splines [15] and $T'_p(J, \pi) = T_p(J, \pi)$. Hence we prove the results when π is regular infinite.

(1) Since π is regular infinite, $[c, d] = \mathrm{cl}(\bigcup_{i=1}^{\infty} I_i)$ for some disjoint open intervals $I_i \subset J$. Since $k_m | I_1$ is a generalized polynomial, i.e., is in $V_n | I_1$, we have $k_m = \sum_{i=0}^{n-1} c_{m,i} u_i$ on $\mathrm{cl}(I_1)$, for some numbers $c_{m,i}$. Now $(\|k_m\|_p)$ is a bounded sequence and, by Lemma 2.1, each k_m/w_0 has at most $n-2$ alternating local extrema in I_1 and thus satisfies condition (2) of [22, p. 224]. By [22, Lemma 2.2], there exists some interval $[u, v] \subset I_1$ and $M > 0$ such that $|k_m(t)| \leqslant M$ for all $t \in [u, v]$ and all m. (In fact, take $[u, v]$ to be one of $[u_i, v_i]$ in [22, Lemma 2.2].) Let x_j, $0 \leqslant j \leqslant n-1$, be n distinct point in (u, v). Then, $|k_m(x_j)| \leqslant M$, $0 \leqslant j \leqslant n-1$, for all m. Since $\det(u_i(x_j))_{i,j=0}^{n-1} \neq 0$, the n values $k_m(x_j)$, $0 \leqslant j \leqslant n-1$, uniquely determine the coefficients $c_{m,i}$, $0 \leqslant i \leqslant n-1$, in the expression for k_m. It follows that $|c_{m,i}| \leqslant N$ for some $N > 0$, for all m and i. Hence, there exists a subsequence m_j of integers and numbers c_i such that $c_{m_j,i} \to c_i$ for $0 \leqslant i \leqslant n-1$. Denote the subsequence (k_{m_j}) of (k_m) by $f_{1,j}$ and define $h = \sum_{i=0}^{n-1} c_i u_i$ on $\mathrm{cl}(I_1)$. Then, $f_{1,j} \to h$ pointwise on $\mathrm{cl}(I_1)$. Since $c_{m_j,i}$ converge, we conclude that the ith derivatives $f_{1,j}^{(i)} \to k^{(i)}$ pointwise on $\mathrm{cl}(I_1)$ for $1 \leqslant i \leqslant n-2$. (At the endpoints of I_1, we consider the one-sided derivatives.) Again, by the same argument applied to the interval I_2, we conclude that a subsequence $(f_{1,j})$ of $(f_{2,j})$ converges pointwise on $\mathrm{cl}(I_2)$ to a polynomial, say h, defined on $\mathrm{cl}(I_2)$. Also, $f_{2,j}^{(i)} \to h^{(i)}$, $1 \leqslant i \leqslant n-2$, pointwise on $\mathrm{cl}(I_2)$. If I_1 and I_2 have a common endpoint x, then clearly h is uniquely defined at x; also, $h^{(i)}(x)$, $1 \leqslant i \leqslant n-2$, exist. Thus, $f_{2,j}$ and its first $n-2$ derivatives respectively converge pointwise to h and its corresponding derivatives on $\mathrm{cl}(I_1 \cup I_2)$. Applying this argument to each interval I_m, we obtain a subsequence $(f_{m,j})$ of (k_m) and a function h defined on $\mathrm{cl}(\bigcup_{i=1}^{m} I_i)$ which have properties as above. Then the diagonal sequence $(h_m = f_{m,m})$ converges pointwise on J to h with $h_m^{(i)}$, $1 \leqslant i \leqslant n-2$, also converging to $h^{(i)}$. Clearly, $h \in T(J, \pi)$.

We show that $\|h\|_p \leqslant D$. Let χ_m denote the indicator function of $\bigcup_{i=1}^{m} I_i$. Then $\|h_j \chi_m\|_p \leqslant \|h_j\|_p \leqslant D$. Since h_j and h are polynomials on each I_i and $h_j \to h$ on I_i, we conclude that $\|h \chi_m\|_p \leqslant D$. By the monotone convergence theorem $\|h \chi_m\|_p \uparrow \|h\|_p$ and, hence, $\|h\|_p \leqslant D$.

(2) Let $f \in L_p(J)$ and $k_m \in T_p(J, \pi)$ with $\|f-k_m\|_p \to \Delta = \inf\{\|f-h\|_p: h \in T_p(J, \pi)\}$. Then, $\|k_m\|_p$ is a bounded sequence and, hence by (1), there exists a subsequence (h_m) of (k_m) and $h \in T_p(J, \pi)$ with the properties stated there. Since by Fatou's Lemma, $\|f-h\|_p \leqslant \liminf \|f-h_m\|_p = \Delta$, we conclude that $\|f-h\|_p = \Delta$ and $T_p(J, \pi)$ is proximinal. Finally, proximinality implies closedness.

(3) Since $F_n|J \subset T_p(J, \pi)$ and, by (2), $T_p(J, \pi)$ is closed in L_p, we conclude that $T'_p(J, \pi) \subset T_p(J, \pi)$. Now proximinality of $T'_p(J, \pi)$ follows as in (2).

(4) Clearly $T'_p = T'_p(J, \pi) \supset V_n|J$ and $T'_p \supset X_n|J$. Hence $T'_p \supset V_n|J + X_n|J$. Now $V_n|J$ is finite dimensional and $X_n|J$ is closed in $L_p(J)$. Hence by [8, p. 68, Problem O], we have that $V_n|J + X_n|J$ is closed. Since $F_n|J \subset V_n|J + X_n|J$, we conclude that $T'_p \subset V_n|J + X_n|J$. ∎

4. Alternative Characterizations of a Best Approximation

In this section we obtain characterizations of a best approximation from $K_{n,p}(S)$ in terms of Tchebycheff splines with countable knots defined in the last section. These characterizations are new and are different from those of Section 2, and will be used in subsequent sections.

Using the characterization Theorems 2.6 and 2.7 of Section 2 as a starting point, we introduce some notation which will be used in the rest of this paper. Let $f \in L_p \backslash K_{n,p}(S)$ and $g \in K_{n,p}(S)$. If $1 < p < \infty$, then define $e = |f-g|^{p-1} \operatorname{sgn}(f-g)$ as in Theorem 2.6. If $p=1$, let $D(f-g)$ as in (2.6) and $e \in D(f-g)$. For all $1 \leqslant p < \infty$, define E^- by (2.5) and let

$$E^+ = \{t \in I: (-1)^n e^{[n]}(t) > 0\}, \tag{4.1}$$

$$A = \{x \in I: e^{[n]}(t) \neq 0, \text{ for } t \in ((x-\delta, x) \cup (x, x+\delta)) \cap S \text{ for some } \delta > 0\}, \tag{4.2}$$

$$B = \{x \in I: e^{[n]}(t) \neq 0, \text{ for } t \in (x-\delta, x) \cup (x, x+\delta) \text{ for some } \delta > 0\}, \tag{4.3}$$

and $B_0 = \{t \in B: e^{[n]}(t) = 0\}$. An open interval (c, d) is called a component of an open set G if $(c, d) \subset G$ and $c, d \notin G$. We leave the simple proof of the next lemma to the reader.

Lemma 4.1. (1) *A and B are open sets.*

(2) *If $J = (c, d) \subset I$ and $J \cap S$ contains only a finite number of zeros (possibly none) of $e^{[n]}$, then J is contained in some component of A.*

(3) *$B \subset A$, and hence a component of B is contained in some component of A.*

(4) *Let $J=(c,d)$ be a component of A (resp. B) and Z be the set of zeros of $e^{[n]}$ in $J\cap S$ (resp. J). Assume Z is infinite. Then the set of accumulation points of Z is nonempty and is contained in $\{c,d\}$. Furthermore, Z is countable.*

(5) *B_0 is a countable set, $E^-\cup E^+\cup B_0=B$, and $\mathrm{cl}(E^-\cup E^+)=\mathrm{cl}(E^-)\cup \mathrm{cl}(E^+)=\mathrm{cl}(B)$.*

The next proposition follows immediately from Lemma 3.1 and 4.1 (4).

PROPOSITION 4.2. *Let $1\leqslant p<\infty$, $n\geqslant 1$, $f\in L_p\backslash K_{n,p}(S)$, and g be a best L_p-approximation to f from $K_{n,p}(S)$. If $J=(c,d)$ is any component of A or B, then there exists a finite or regular infinite partition $\pi=\pi_J$ of J having one of the forms described in Lemma 3.1 such that the points of $J\cap\pi$ are the zeros of $e^{[n]}$ in $J\cap S$.*

We now define U_n-convex functions on an interval $J=(c,d)\subset I$. Let $U_n=\{u_i\}_{i=0}^{n-1}$ given by (1.1) be an ECT-system. A real valued function k on J is said to be U_n-convex (on J) if (1.2) holds for any $n+1$ points $x_0<x_1<\cdots<x_n$ in J. Alternative definitions are obtained by replacing the interval I in definitions (2.1) and (2.2) by J. Given w_i, $0\leqslant i\leqslant n-1$, as in Section 1, we may define functions $\hat{u}_i$, $0\leqslant i\leqslant n-1$, by (1.1) with a and b there replaced by c and d, respectively. Then $\hat{U}_n=\{\hat{u}_i\}_{i=0}^{n-1}$ is an ECT-system on J. If $\hat{V}_n=\operatorname{span}\hat{U}_n$, then $\hat{V}_n$ is precisely all functions in V_n restricted to J, i.e., $\hat{V}_n=V_n|J$. It may be easily verified that k on J is U_n-convex if and only if it is $\hat{U}_n$-convex (i.e., (1.2) holds with u_i replaced there by $\hat{u}_i$). It is easy to see that if k defined on I is U_n-convex on I, then it is U_n-convex on J. However, if k defined on J is U_n-convex on J, then there may not exist a U_n-convex function on I whose restriction to J equals k. This situation arises if $k(x)\to\pm\infty$ as $x\downarrow c$ or $x\uparrow d$. We denote the set of all U_n-convex functions on J by $K_n(J)$. As before we may defined $K_n(J,S)$ as follows. For $k\in K_n(J)$, let $\hat{\mu}_{k,n}$ denote the Lebesgue–Stieltjes measure generated by l_{n-1}^+k on J. Then $K_n(J,S)=\{k\in K_n(J)\colon \hat{\mu}_{k,n}(J\backslash S)=0\}$. We let $K_{n,p}(J,S)=K_n(J,S)\cap L_p(J)$, $1\leqslant p<\infty$. Recall that $T_p(J,\pi)$ and $T'_p(J,\pi)$ are defined by (3.2)–(3.4).

PROPOSITION 4.3. *Let $J=(c,d)\subset I$ and π be a finite or a regular infinite partition of J. Then $T_p(J,\pi)\cap K_{n,p}(J,S)\subset T'_p(J,\pi)$, $1\leqslant p<\infty$, $n\geqslant 1$.*

Proof. (See [5, Proof of Lemmas 2.3 and 2.4]). Let $k\in T_p(J,\pi)\cap K_{n,p}(J,S)$ and $\rho(t)=l_{n-1}^+k(t)$, where l_{n-1}^+ is defined by (1.4). Then, ρ is right continuous and nondecreasing [7]. Let $0<\delta<(d-c)/2$ and for $0<\varepsilon<\delta$, define as in [7, p. 391], $\rho(t,\varepsilon)=\rho(c+\varepsilon)$, $t\in(c,c+\varepsilon)$, $=\rho(t)$, $t\in[c+\varepsilon,d-\varepsilon)$, and $=\rho(d-\varepsilon)$, $t\in[d-\varepsilon,d)$. Also define $k(t,\varepsilon)$, $t\in J$, by $k(t,\varepsilon)=\int_{J\cap S}G_{n-1}(t,x)\,d\rho(x,\varepsilon)+\sum_{i=0}^{n-1}a_i(\varepsilon)\,u_i(t)$, $n\geqslant 2$, and $=\rho(t,\varepsilon)$,

$n=1$, where numbers $a_i(\varepsilon)$ are chosen so that $k(\cdot,\varepsilon)=k$ on $(c+\varepsilon, d-\varepsilon)$. Let $\hat{\mu}_{k,n}$ and μ be the Lebesgue–Stieltjes measure generated by ρ and $\rho(\cdot,\varepsilon)$, respectively, on J. Clearly, $\hat{\mu}_{k,n}=\mu$ on $(c+\varepsilon, d-\varepsilon]$. Since $\rho(\cdot,\varepsilon)$ is constant on each of $(c, c+\varepsilon]$ and $(d-\varepsilon, d)$, we have that $\mu(J\backslash(c+\varepsilon, d-\varepsilon])=0$. Hence, $\int_{J\cap S} G_{n-1}(t,x)\,d\rho(x,\varepsilon)=\int_{(c+\varepsilon,d-\varepsilon]\cap S} G_{n-1}(t,x)\,d\rho(x)$. If $[c,d]=\mathrm{cl}(\bigcup_i I_i)$ as in (3.1), then since $k\in T_p(J,\pi)$, we conclude that $k\,|\,I_i\in V_n\,|\,I_i$, i.e., it is generalized polynomial for all i. Hence ρ is constant on each I_i and $\hat{\mu}_{k,n}(J\backslash\pi)=\hat{\mu}_{k,n}(\bigcup_i I_i)=0$. Again, by the regularity of π, $\pi'=(c+\varepsilon, d-\varepsilon]\cap S\cap\pi$ is finite. Hence $\int_{J\cap S} G_{n-1}(t,x)\,d\rho(x,\varepsilon)=\sum_{x\in\pi'} G_{n-1}(t,x)\,\hat{\mu}_{k,n}\{x\}$. It follows that $k(\cdot,\varepsilon)\in\mathrm{span}(F_n\,|\,J)$ for all $\varepsilon>0$. By [5, Lemma 2.4] as applied to $L_p(J)$, we conclude that $k(\cdot,\varepsilon)\to k$ in $L_p(J)$ as $\varepsilon\to 0$. Hence $k\in\mathrm{cl}_p(\mathrm{span}(F_n\,|\,J))=T'_p(J,\pi)$. ∎

LEMMA 4.4. *Let $1\leq p<\infty$, $n\geq 1$, $f\in L_p\backslash K_{n,p}(S)$, and g be a best L_p-approximation to f from $K_{n,p}(S)$. If $J=(c,d)\subset I$ and $e^{[i]}(c)=e^{[i]}(d)=0$, $1\leq i\leq n$, then $g\,|\,J$ is a best $L_p(J)$-approximation to $f\,|\,J$ from $K_{n,p}(J,S)$.*

Proof. We prove the result for $1<p<\infty$, the proof for $p=1$ is similar. Let $\hat{f}=f\,|\,J$ and $\hat{g}=g\,|\,J$. If α corresponds to e when approximating $\hat{f}$ by functions in $K_{n,p}(J,S)$, then $\alpha=|\hat{f}-\hat{g}|^{p-1}\,\mathrm{sgn}(\hat{f}-\hat{g})=e$ on J. Define $\alpha_c^{[i]}(t)=\int_c^t w_{i-1}(s)\,\alpha_c^{[i-1]}(s)\,ds$, $t\in J$, where $\alpha_c^{[0]}=\alpha$.

We show that the three conditions of Theorem 2.6 hold for α. We first assert that

$$\alpha_c^{[i]}(t)=e^{[i]}(t), \qquad t\in J, \quad 1\leq i\leq n. \tag{4.4}$$

Clearly, $\alpha_c^{[0]}=\alpha=e=e^{[0]}$. Hence by integration, $\alpha_c^{[1]}(t)=e^{[1]}(t)-e^{[1]}(c)=e^{[1]}(t)$, since $e^{[1]}(c)=0$. Now $e^{[i]}(c)=0$, $1\leq i\leq n$, hence, repeating this process successively or by induction we conclude that (4.4) holds. Also

$$\{t\in J\colon (-1)^n\,\alpha_c^{[n]}(t)<0\}=J\cap E^-. \tag{4.5}$$

Since g is the best approximation to f from $K_{n,p}(S)$, the three conditions of Theorem 2.6 hold for e. Now by (4.4), (4.5), and the hypothesis, we find that $\alpha_c^{[i]}(d)=0$, $1\leq i\leq n$, $(-1)^n\,\alpha_c^{[n]}(t)\leq 0$ for $t\in J\cap S$ and $g\,|\,J$ is a generalized polynomial on each component of the open set $J\cap E^-$. Hence, by Theorem 2.6, the required conclusion follows. ∎

LEMMA 4.5. *Let $x\in I$ and (x_j) be a sequence in $I\backslash\{x\}$ such that $x_j\to x$. Suppose that $n\geq 1$ and $h\in L_1$ satisfies $h^{[n]}(x_j)=0$ for all j. Then $h^{[i]}(x)=0$, $1\leq i\leq n$. Furthermore, if $h^{[1]}$ is differentiable at x then $h(x)=0$.*

Proof. Clearly, $h^{[n]}(x)=0$ by the continuity of $h^{[n]}$. Suppose that $n\geq 2$. By successive applications of Rolle's theorem we conclude that, for each $1 \leqslant i \leqslant n-1$, there exists a sequence $(x_j^{(i)} \subset I \backslash \{x\}$ such that $h^{[i]}(x_j^{(i)})=0$ and $x_j^{(i)} \to x$ as $j \to \infty$. To prove the required result by induction assume that $h^{[i]}(x)=0$ for some $2 \leqslant i \leqslant n$. Then $(h^{[i]}(x_j^{(i)})-h^{[i]}(x))/(x_j^{(i)}-x)=0$ for all j. Letting $j \to \infty$ we conclude that $w_{i-1}(x)\, h^{[i-1]}(x)=0$ or $h^{[i-1]}(x)=0$. The last statement of the lemma follows as above. ∎

We now state and prove our main characterization theorem. The following notation will be useful in computations. Analogous to l_{n-1}^+ and l_{n-1}^- (see 1.4)), for a U_n-convex function k define the differential operators

$$l_i k = \frac{1}{w_i} D \frac{1}{w_{i-1}} D \frac{1}{w_{i-2}} \cdots \frac{1}{w_1} D\left(\frac{k}{w_0}\right), \qquad 0 \leqslant i \leqslant n-2, \quad n \geqslant 2, \tag{4.6}$$

where D denotes the derivative of a function as in Section 1 and $l_0 k = k/w_0$. Clearly $D(l_{i-1}k)=w_i(l_i k)$. Recall that if $1<p<\infty$, then the best approximations under consideration are unique.

THEOREM 4.6. *Let $1<p<\infty$, $n \geqslant 1$, $f \in L_p \backslash K_{n,p}(S)$, and $g \in K_{n,p}(S)$. Let $e=|f-g|^{p-1}\operatorname{sgn}(f-g)$, and the sets A and B be defined by* (4.2) *and* (4.3). *Then the following five statements are equivalent.*

(1) *g is the best L_p-approximation to f from $K_{n,p}(S)$.*

(2) *$g=f$ a.e. on $I\backslash A$, and, on each component J of A, $g|J$ is the best $L_p(J)$-approximation to $f|J$ from $K_{n,p}(J,S)$.*

(3) *Reiterate the above statement with A replaced by B.*

(4) *$g=f$ a.e. on $I\backslash A$, and, on each component J of A, $g|J$ is the best $L_p(J)$-approximation to $f|J$ from each of $K_{n,p}(J,S)$ and $T_p'(J,\pi)$, where $\pi=\pi_J$ denotes the finite or regular infinite partition of J formed by the zeros of $e^{[n]}$ in $J \cap S$ (see Proposition 4.2).*

(5) *Reiterate the above statement with A replaced by B.*

Proof. We prove the results for A, i.e., the equivalence of (1), (2), and (4). The proof of the equivalence of (1), (3), and (5), the results for B, is similar. However, since $B \subset A$ (Lemma 4.1(3)), some results for B follow immediately from those of A.

We show that (1) implies (4). Let g be the best approximation to f. Let $x \in I\backslash A$. It follows directly from the definition of A that there exists a sequence $(x_j) \subset I\backslash\{x\}$ such that $x_j \to x$ and $e^{[n]}(x_j)=0$ for all j. By Lemma 4.5 with $h=e$ it follows that $e^{[i]}(x)=0$ for $1 \leqslant i \leqslant n$, and $e(x)=0$ if $e^{[1]}$ is differentiable at x. Now since $e^{[1]}$ is differentiable a.e. on I, we conclude that $e=0$ a.e. on $I\backslash A$. This gives $g=f$ a.e. on $I\backslash A$.

Let $J=(c,d)$ be a component of A. Since $c, d \in I \backslash A$, as shown above we have $e^{[i]}(c)=e^{[i]}(d)=0$, $1 \leqslant i \leqslant n$. By Lemma 4.4, $g|J$ is the best approximation to $f|J$ from $K_{n,p}(J, S)$.

It now remains to show that $\hat{g}=g|J$ is the best approximation to $\hat{f}=f|J$ from $T'_p(J, \pi)$. To this end, we first show that $\mu_{g,n}(E^-)=0$. By Theorem 2.6, g is a generalized polynomial of degree $n-1$ on each component (x, y) of E^-. Hence, by the definition of $\mu_{g,n}$, we have $\mu_{g,n}(x, y)=(l^-_{n-1}g)(y)-(l^+_{n-1}g)(x)=0$. Now E^- is the union of a countable number of its components. Hence, $\mu_{g,n}(E^-)=0$. Since $g \in C^{(n-2)}(I)$, and $[c, d]=\mathrm{cl}(\bigcup_i I_i)$ as in (3.1), to establish $\hat{g} \in T(J, \pi)$, it suffices to show that g is a generalized polynomial of degree $n-1$ on each $I_i=(c_i, d_i)$, say. By Theorem 2.6, we have $(-1)^n e^{[n]} \leqslant 0$ on S. Since $I_i \cap S$ contains no zeros of $e^{[n]}$, we conclude that $(-1)^n e^{[n]}<0$ on $I_i \cap S$. It follows that $I_i \cap S \subset E^-$. Since $\mu_{g,n}(E^-)=0$ as shown above, we have $\mu_{g,n}(I_i \cap S)=0$. Again, since $\mu_{g,n}(S')=0$, we have $\mu_{g,n}(I_i)=0$, i.e., $(l^+_{n-1}g)(c_i)=(l^-_{n-1}g(d_i)$. It follows that g is a generalized polynomial of degree $n-1$ on I_i. Thus $\hat{g} \in T(J, \pi)$, and since $g \in L_p$ we find $\hat{g} \in T_p(J, \pi)$. Again since $\hat{g} \in K_{n,p}(J, S)$, by Proposition 4.3, we conclude that $\hat{g} \in T'_p(J, \pi)$.

We now show that $\hat{g}$ is the best approximation to $\hat{f}$ from $T'_p(J, \pi)$ or equivalently, $\int_J \alpha h=0$ for all $h \in T'_p(J, \pi)$, where $\alpha=|\hat{f}-\hat{g}|^{p-1}\,\mathrm{sgn}(\hat{f}-\hat{g})$ on J (see the proof of Lemma 4.4). Now $\alpha=e$ on J, $T'_p(J, \pi)=\mathrm{cl}_p(\mathrm{span}(F_n|J))$, and $x^*(h)=\int_J \alpha h=\int_J eh$ for $h \in L_p(J)$ is a continuous linear functional on $L_p(J)$. Hence $\hat{g}$ is the best approximation to $\hat{f}$ if and only if e satisfies

$$\int_J eu_i=0, \qquad 0 \leqslant i \leqslant n-1, \tag{4.7}$$

and

$$\int_J eG_{n-1}(\cdot, t)=0, \qquad t \in J \cap \pi. \tag{4.8}$$

We show that (4.7) and (4.8) hold by using the fact that $e^{[i]}(c)=e^{[i]}(d)=0$, $1 \leqslant i \leqslant n$, which is established above, and $e^{[n]}(t)=0$ for $t \in J \cap \pi$ which follows from Proposition 4.2. Integration by parts yields

$$\int_J eu_i=e^{[1]}(d)\frac{u_i(d)}{w_0(d)}-e^{[1]}(c)\frac{u_i(c)}{w_0(c)}-\int_J e^{[1]}w_1(l_1u_i)=-\int_J e^{[1]}w_1(l_1u_i),$$

where l_i are defined by (4.6). Since $l_iu_i=1$, successive integration by parts gives for $0 \leqslant i \leqslant n-1$, $(-1)^i \int_J eu_i=\int_J e^{[i]}w_i(l_iu_i)=\int_J e^{[i]}w_i=(e^{[i+1]}(d)-e^{[i+1]}(c))=0$. Similarly, since $l_{n-1}G_{n-1}(\cdot, t)=1$ on $[t, d)$ and 0, elsewhere, we have

$$(-1)^{n-1}\int_J eG_{n-1}(\cdot, t) = \int_J e^{[n-1]}w_{n-1}l_{n-1}G_{n-1}(\cdot, t)$$

$$= \int_{[t, d)} e^{[n-1]}w_{n-1} = (e^{[n]}(d) - e^{[n]}(t)) = 0.$$

We have shown that (1) implies (4).

Clearly, (4) implies (2). We show that (2) implies (1). Let $k \in K_{n,p}(S)$ and J be any component of A. Then $g|J$ is the best approximation to $f|J$ from $K_{n,p}(J, S)$. Since $k|J \in K_{n,p}(J, S)$, we have $\|(f-g)\chi_J\|_p \leqslant \|(f-k)\chi_J\|_p$ or $\|(f-g)\chi_J\|_p^p \leqslant \|(f-k)\chi_J\|_p^p$, where χ_F denotes the indicator function of the set F. Summing the latter inequality over all (countable) components J of A we obtain $\|(f-g)\chi_A\|_p^p \leqslant \|(f-k)\chi_A\|_p^p$. Since $g = f$ a.e. in $I \backslash A$, we conclude that $\|f-g\|_p \leqslant \|f-k\|_p$. Hence g is the best approximation to f from $K_{n,p}(S)$. ∎

The following analogue of the above theorem for $p = 1$ may be similarly proved using Theorem 2.7. Since $p = 1$, a best approximation is not necessarily unique.

THEOREM 4.7. *Let* $p = 1$, $n \geqslant 1$, $f \in L_1 \backslash K_{n,1}(S)$, *and* $g \in K_{n,1}(S)$. *Let* $D(f-g)$ *be defined by* (2.6). *Then* g *is a best* L_1*-approximation to* f *from* $K_{n,1}(S)$ *if and only if there exists* $e \in D(f-g)$ *satisfying the statements* (2)–(5) *of Theorem* 4.6 *with "the best* $L_p(J)$*-approximation," in those statements replaced by "a best* $L_1(J)$*-approximation,"* p *replaced by* 1, *and the sets* A *and* B *defined by* (4.2) *and* (4.3) *using this* e.

Finally, we remark that in Theorem 4.6, another equivalent statement (2′) may be obtained from statement (2) by letting $J = (c, d)$ in that statement, where $c < d$ are two endpoints of (possibly different) components of A (and, in particular J may be a component of A). This may be justified by Lemma 4.4. Similar changes may be made in the central part of statement (4) pertaining to $K_{n,p}(J, S)$, and statements (3) and (5). Similar remarks apply to Theorem 4.7.

5. PROPERTIES OF BEST APPROXIMATIONS

Throughout this section unless otherwise stated, we assume that $1 \leqslant p < \infty$, $n \geqslant 1$, $f \in L_p \backslash K_{n,p}(S)$, and g is a best L_p-approximation to f from $K_{n,p}(S)$. We investigate the structure and properties of g, and also its uniqueness in L_1. We use the notation of the last section; in particular, e, $e^{[i]}$, and E^- have the meanings given there.

A real function k on I is said to be strictly U_n-convex on $J = (c, d) \subset I$ if $[x_0, x_1, ..., x_n]k > 0$ for all $c < x_0 < x_1 < \cdots < x_n < d$. The following lemma is an extension of [17, Lemma 3.2(ii)] and may be proved similarly. A similar result for convex functions has appeared in [21].

LEMMA 5.1. *Let $k \in K_n(S)$ and $J = (c, d) \subset I$. Assume that k is not strictly U_n-convex on any subinterval of J. Then there exists an infinite (not necessarily regular) partition π of J such that $k \mid J \in T(J, \pi)$.*

PROPOSITION 5.2. *Let $J = (c, d) \subset I$.*

(1) *If g is strictly U_n-convex on J, then $g = f$ a.e. on J.*

(2) *If $g \neq f$ a.e. on J, then $g \in T(J, \pi)$ for some infinite (not necessarily regular) partition π of J.*

Proof. (1) By Theorems 2.6 and 2.7, g is a generalized polynomial on each component of E^-. Hence we have $J \cap E^- = \varnothing$ and, consequently, $(-1)^n e^{[n]} \geqslant 0$ on J. If, for some $x \in J$, we have $(-1)^n e^{[n]}(x) > 0$, then there exists an open interval $J' = (u, v) \subset J$ such that $(-1)^n e^{[n]} > 0$ on J'. Then, by Theorems 2.6 and 2.7, we have $J' \subset S' = I \backslash S$. Since $\mu_{g,n}(S') = 0$, $0 = \mu_{g,n}(J') = l_{n-1}^-(v) - l_{n-1}^+(u)$. Hence, g is a generalized polynomial on J', which is a contradiction. It follows that $e^{[n]} = 0$ on J, and hence, $e = 0$ a.e. on J. This gives $f = g$ a.e. on J.

(2) It follows that $g \neq f$ a.e. on any subinterval of J and hence, by (1), g is not strictly U_n-convex on any subinterval on J. Now the result follows from Lemma 5.1. ∎

A point $t \in S$ is called a two-sided limit point of S if there exist sequences (s_m) and (t_m) in S with $s_m < t < t_m$ such that $s_m \to t$ and $t_m \to t$. Let $\hat{S} \subset S$ denote all the two-sided limit points of S. If h is a function on I and $J \subset I$ then we let $Z_J(h) = \{t \in J : h(t) = 0\}$.

LEMMA 5.3. *Let $n \geqslant 2$ and $J = (c, d) \subset I$. Then $Z_J(e^{[n]}) \cap \hat{S} \subset Z_J(e^{[n-1]})$.*

Proof. Let $t \in Z_J(e^{[n]}) \cap \hat{S}$. Then there exist sequences (s_m) and (t_m) in S with $s_m < t < t_m$ such that $s_m \to t$ and $t_m \to t$. By Theorems 2.6 and 2.7 we have $(-1)^n e^{[n]} \leqslant 0$ on S; in particular, $(-1)^n e^{[n]}(s_m) \leqslant 0$. Since $t \in Z_J(e^{[n]})$ we have $e^{[n]}(t) = 0$, and hence, $(-1)^n (e^{[n]}(t) - e^{[n]}(s_m))/(t - s_m) \geqslant 0$. Letting $m \to \infty$ we obtain $(-1)^n e^{[n-1]}(t) \geqslant 0$. Similarly, using the sequence (t_m) we obtain $(-1)^n e^{[n-1]}(t) \leqslant 0$, which gives $e^{[n-1]}(t) = 0$. ∎

THEOREM 5.4. *Let $J = (c, d) \subset I$. In each of the following cases, the splines are the ordinary Tchebycheff splines (i.e., with a finite number of knots) of degree $n-1$, and the knots are the zeros of $e^{[n]}$ in $J \cap S$. ($\lfloor x \rfloor$ denotes the largest integer no larger than x.)*

(1) *If $f>g$ a.e. on J or $f<g$ a.e. on J then $g|J$ is a spline with at most n knots in J.*

(2) *Assume that $Z_J(e^{[n]}) \subset \hat{S}$. Then the following holds.*

(i) *If $f>g$ a.e. on J, then $g|J$ is a spline with at most $\lfloor (n-1)/2 \rfloor$ knots in J.*

(ii) *If $f<g$ a.e. on J, then $g|J$ is a spline with at most $\lfloor n/2 \rfloor$ knots in J.*

If $J \subset S$, then $Z_J(e^{[n]}) \subset \hat{S}$ and (i) *and* (ii) *above hold.*

Proof. We provide proofs for $1<p<\infty$; the case $p=1$ may be similarly argued.

(1) We prove the result for the case $f>g$ a.e. on J; the proof for the other case is similar. Since $f>g$ a.e. on J we have $w_0 e>0$ a.e. on J. This implies that $e^{[1]}$ is strictly increasing on J. Consequently, $e^{[1]}$ has at most one zero in J. Hence, its indefinite integral $e^{[2]}$ has at most two zeros in J. Applying this argument in succession or by induction, we conclude that $e^{[n]}$ has at most n zeros in J and consequently, in $J \cap S$. By Lemma 4.1(2), J is contained in some component J' of A. By Theorem 4.6(4), $g|J' \in T'_p(J', \pi)$, where π is the partition of J' such that points of $J' \cap \pi$ are the zeros of $e^{[n]}$ in $J' \cap S$ as in Proposition 4.2. Since $J \cap S \subset J' \cap S$ and contains at most n zeros of $e^{[n]}$, $g|J$ is a spline with at most n knots.

(2) First consider the case $n=1$. Note that, by Theorems 2.6 and 2.7, $e^{[1]} \geqslant 0$ on S. If $f>g$ a.e. on J, then as in (1), $e^{[1]}$ is strictly increasing on J. Hence, if for some $t \in J$, $e^{[n]}(t)=0$, then $e^{[1]}<0$ on (c, t). Since $e^{[1]} \geqslant 0$ on S, there does not exist a sequence $(s_m) \subset S$ with $s_m<t$ and $s_m \to t$. This is a contradiction to the hypothesis that $t \in \hat{S}$. Hence, $e^{[1]} \neq 0$ on J and $e^{[1]}$ has no zeros in $J \cap S$. If $f<g$ a.e. on J, then as above, we may show that $e^{[1]}$ has no zeros in $J \cap S$. Now using Lemma 4.1(2) and Theorem 4.6(4), and arguing as in (1) we conclude that $g|J$ is a spline with no knots. This proves (i) and (ii) for $n=1$.

Now consider the case $n \geqslant 2$. If $f>g$ a.e. on J, then, as in (1), the number of zeros of $e^{[n]}$ and $e^{[n-1]}$ in J cannot exceed n and $n-1$, respectively. Let r be the number of zeros of $e^{[n]}$ in J. By Rolle's theorem, there is a zero of $e^{[n-1]}$ strictly between each pair of adjacent zeros of $e^{[n]}$; hence there are $r-1$ such zeros in J. Again since $Z_J(e^{[n]}) \subset \hat{S}$, by Lemma 5.3, we have $Z_J(e^{[n]}) \subset Z_J(e^{[n-1]})$. Consequently, $e^{[n-1]}$ has $2r-1$ zeros in J. Now since $2r-1 \leqslant n-1$, we conclude that $r \leqslant \lfloor n/2 \rfloor$. If $f<g$ a.e. on J, then by an identical method, we may show that $r \leqslant \lfloor n/2 \rfloor$, where r is as above. Hence, the number of zeros of $e^{[n]}$ in $J \cap S$ cannot exceed $\lfloor n/2 \rfloor$. Now using Lemma 4.1(2) and Theorem 4.6(4) and arguing as in (1), we conclude that $g|J$ is as spline with at most $\lfloor n/2 \rfloor$ knots. If n is odd then

$\lfloor n/2 \rfloor = \lfloor (n-1)/2 \rfloor$. Hence we have established (i) where n is odd and (ii) for all n.

It remains to show (i) where n is even. Assume that n is even and $e^{[n]}$ has r zeros in J. As shown above $r \leqslant n/2$. Suppose now that $r = n/2$. For convenience denote the r zeros of $e^{[n]}$ by $z_1^n < z_3^n < z_5^n < \cdots < z_{n-1}^n$. As before, by Rolle's theorem and Lemma 5.3, $e^{[n-1]}$ has $2r-1 = n-1$ zeros in J denoted by $z_1^{n-1} < z_2^{n-1} < \cdots < z_{n-1}^{n-1}$, where $z_i^{n-1} = z_i^n$, $i = 1, 3, 5, ..., n-1$. Again, by Rolle's theorem, applied successively, we conclude that $e^{[j]}$, $1 \leqslant j \leqslant n-1$, has j zeros $z_1^j < z_2^j < \cdots < z_j^j$ in J where $z_i^{j+1} < z_i^j < z_{i+1}^{j+1}$, $1 \leqslant i \leqslant j$. Let $t_i = z_i^i$, $1 \leqslant i \leqslant n-1$, and $t_n = z_{n-1}^n$. Then, $t_1 < t_2 < \cdots < t_{n-1} = t_n$ and $e^{[i]}(t_i) = 0$, $0 \leqslant i \leqslant n$. Now $f > g$ a.e. on J, hence, $w_0 e > 0$ a.e. on J. Since $e^{[1]}(t_1) = 0$ we have $e^{[1]} > 0$ on (t_1, d). By repeated applications of this argument, we have $e^{[j]} > 0$ on (t_j, d), and in particular, $e^{[n]} > 0$ on (t_n, d). Since n is even by Theorem 2.6, $e^{[n]} \leqslant 0$ on S. Hence there does not exists a sequence $(s_m) \subset S$ with $t_n < s_m$ and $s_m \to t_n$. This contradicts the fact that $t_n \in Z_J(e^{[n]}) \subset \hat{S}$. Thus $e^{[n]}$ has no more than $n/2 - 1$ zeros in J, and, hence in $J \cap S$. Then arguing as before we conclude that $g|J$ is a spline with no more than $n/2 - 1 = \lfloor (n-1)/2 \rfloor$ knots. ∎

The following example shows that the bounds on the number of knots in Theorem 5.4 cannot be improved in general. Let $I = S = (0, 4)$, $n = p = 2$, and $w_0 = w_1 = 1$. Let $f(t) = -t + 1$, $1 < t \leqslant 2$, $= t - 3$, $2 < t \leqslant 3$, and $= 0$ elsewhere. Then g, the best L_2-approximation to f by $K_{2,2}(I)$, is given by $g(t) = -t/2 + 1/4$, $0 < t \leqslant 2$, and $= t/2 - 7/4$, $2 < t < 4$. Clearly, $f < g$ on $(3/2, 5/2)$ and g has one knot in this interval. Again, $f > g$ on $(1/2, 3/2)$ and also on $(5/2, 7/2)$, and g has no knots in each of these intervals. This verifies (i) and (ii) of Theorem 5.4(2).

We now investigate boundedness of a best approximation. We use the differential operators l_{n-1}^+ and l_i, $0 \leqslant i \leqslant n-2$, defined by (1.4) and (4.6), respectively.

THEOREM 5.5. *Let $n \geqslant 2$. If f is essentially bounded on a right neighborhood of a and a left neighborhood of b, then g is continuous and bounded on I, and can be extended to a continuous function on $[a, b]$. In particular, these conclusions about g hold if $f \in L_\infty$ or f is continuous and bounded on I.*

Proof. For convenience of notation we use the operator l_{n-1} to denote l_{n-1}^+ in this proof. By Proposition 2.2(2), $g(a^+)$ exists. We need to show that $|g(a^+)| < \infty$. Assume to the contrary that $|g(a^+)| = \infty$; we show a contradiction. Define $g_i = w_i(l_i g)$, $0 \leqslant i \leqslant n-1$, where $g_0 = w_0(l_0 g) = g$. Note that $(1/w_{n-1}) D^+(1/w_{n-2}) D \cdots (1/w_{i+1}) D g_i / w_i = l_{n-1} g$, which is nondecreasing since g is U_n-convex on I. It follows that g_i is U_{n-i}-convex on I, where $U_{n-i} = \{u_j\}_{j=i}^{n-1}$ [7, p. 386, Theorem 2.1]. Since g may be obtained from g_i by a succession of integral operations, $|g(a^+)| = \infty$

implies $|g_i(a^+)| = \infty$ for $0 \leqslant i \leqslant n-1$. By Proposition 2.2(2), we conclude that $(-1)^{n-i} g_i(a^+) = \infty$, $0 \leqslant i \leqslant n-1$. Since w_0 is continuous on $[a, b]$ and f is bounded in a neighborhood of a, there exists $c \in I$ such that $(-1)^{n-i} g_i(c) > 0$ or $(-1)^{n-i} (l_i g)(c) > 0$ for $0 \leqslant i \leqslant n-1$, and $(-1)^n g(c) > (-1)^n w_0(c) f/w_0$ a.e. on (a, c).

Analogous to $G_{n-1}(x, t)$, let for $0 \leqslant i \leqslant n-1$,

$$
\begin{aligned}
H_i(x, t) &= w_0(x) \int_x^t w_1(t_1) \int_{t_1}^t w_2(t_2) \cdots \int_{t_{i-1}}^t w_i(t_i)\, dt_i \cdots dt_2\, dt_1, \qquad a < x < t, \\
&= 0, \qquad t \leqslant x < b,
\end{aligned}
$$

where $H_0(x, t) = w_0(x)$ for $x < t$ and 0, otherwise; $H_0(\cdot, t)$ is thus right-continuous. Now define

$$
\begin{aligned}
h(x) &= \sum_{i=0}^{n-1} (-1)^i (l_i g)(c)\, H_i(x, c), \qquad a < x < c, \\
&= g(x), \qquad c \leqslant x < b.
\end{aligned} \tag{5.1}
$$

Note that h is a generalized polynomial of degree $n-1$ on (a, c). It is easy to verify that $(l_i h)(c)$ exists for $0 \leqslant i \leqslant n-1$ and equals $(l_i g)(c)$. Hence h is U_n-convex on I. Now $l_{n-1} h$ equals the constant $(l_{n-1} g)(c)$ on $(a, c]$ and hence, it is continuous at c. (Recall that $l_{n-1} h$ is right-continuous and non-decreasing.) Let μ be the Lebesgue–Stieltjes measure generated by $l_{n-1} h$. Then $\mu(a, c] = 0$, and hence $\mu(S' \cap (a, c]) = 0$. Again, since $g \in K_{n,p}(S)$ and $g = h$ on (c, b), we conclude that $\mu(S' \cap (c, b)) = 0$ giving $\mu(S') = 0$. Thus $h \in K_{n,p}(S)$.

Since $(-1)^{n-i} (l_i g)(c) > 0$, $1 \leqslant i \leqslant n-1$, we obtain by (5.1), $(-1)^n h \geqslant (-1)^n (l_0 g)(c)\, H_0(\cdot, c) = (-1)^n g(c) w_0/w_0(c)$ on (a, c). Hence $(-1)^n h \geqslant (-1)^n f$ a.e. on (a, c). If $p(t) = (l_{n-1} g)(t)$, it is known that $g(x) = (-1)^n \int_x^c H_{n-1}(x, t)\, dp(t) + \sum_{i=0}^{n-1} (-1)^i (l_i g)(c)\, H_i(x, c)$, $a < x < c$. (See, e.g., [7, p. 388, Eq. (2.26)] with n there replaced by $n-1$; however, this expression should be corrected for minor errors in the exponents of (-1) within the summation.) This gives

$$
g(x) = (-1)^n \int_x^c H_{n-1}(x, t)\, dp(t) + h(x), \qquad a < x < c. \tag{5.2}
$$

Now $(-1)^n g(a^+) = \infty$. Hence, by (5.2) we conclude that $\mu_{g,n}(a, c) > 0$, where $\mu_{g,n}$ is the Lebesgue–Stieltjes measure generated by p. Again since $H_{n-1}(\cdot, t) > 0$ on (a, c), we obtain $(-1)^n g > (-1)^n h$ on (a, c). Thus $(-1)^n g > (-1)^n h \geqslant (-1)^n f$ on (a, c). Hence, $\int_a^c |f - h|^p < \int_a^c |f - g|^p$. Since $h = g$ on $[c, b]$, we conclude that $\|f - h\|_p < \|f - g\|_p$ and that g is not a best approximation, a contradiction. Thus $|g(a^+)| < \infty$. Similarly, using [7, p. 384, Eq. (2.9)], we may show that $|g(b^-)| < \infty$. Hence g is bounded on I; its continuity follows from U_n-convexity for $n \geq 2$. By letting $g(a) = g(a^+)$ and $g(b) = g(b^-)$, we have the resulting continuous extension to $[a, b]$. ∎

The following known result, which is needed for our next theorem, follows immediately from [7, p. 410, Theorem 5.5].

LEMMA 5.6. *Let* $1 \leq p < \infty$, $n \geq 1$, *and* $f \in L_p \backslash V_n$. *If* g *is a best* L_p*-approximation to* f *from* V_n, *then* e *has at least* n *sign changes in* I.

THEOREM 5.7. *Let* f *be* U_n*-concave on* I. *Then* $g \in V_n$, *and hence,* g *is also a best approximation to* f *from* V_n. *If* $p = 1$ *and* f *is continuous on* $[a, b]$, *then* g *is unique. (For* $1 < p < \infty$, g *is unique by the uniform convexity of* L_p.)

Proof. If $f = g$ a.e., then g is both U_n-convex and U_n-concave; thus $g \in V_n$. Now suppose that $f \neq g$ and hence, $e \neq 0$ on a set of positive measure. It follows that $e^{[n]} \neq 0$ on a set of positive measure. Clearly, $g - f$ is U_n-convex on I. By Proposition 2.2(1), $g - f$ and, hence, e has at most n sign changes in I. Since $e^{[1]}(a) = e^{[1]}(b) = 0$, by Rolle's theorem, we find that $e^{[1]}$ has at most $(n+2) - 1$ separated zeros in $[a, b]$. (The endpoints a and b, and possibly small intervals containing each endpoint constitute two separated zeros.) Applying this argument in succession or by induction, we conclude that $e^{[n]}$ has at most $(n+2) - n = 2$ separated zeros in $[a, b]$.

Now let $[a, c]$ and $[d, b]$ be the largest intervals such that $e^{[n]}(t) = 0$ for $t \in [a, c] \cup [d, b]$. Since $e^{[n]}(a) = e^{[n]}(b) = 0$ and $e^{[n]}$ is continuous, such intervals exist. Again $c < d$ since $e^{[n]} \neq 0$ on a set of positive measure. Let $J = (c, d)$. Then, by the above argument on zeros, $e^{[n]}$ has no separated zeros in J and, hence, no ordinary zeros in J. Since $e^{[n]} = 0$ on $[a, c] \cup [d, b]$, we see that J is a component of B. If π is the partition of J as in Proposition 4.2, then π: $c < d$, and $T'_p(J, \pi) = V_n | J$. By statement (5) of Theorems 4.6 and 4.7, we find that $g | J$ is a best $L_p(J)$-approximation to $f | J$ from $V_n | J$.

We show that $a = c$ and $b = d$. This will establish that $g \in V_n$ and is a best approximation to f from V_n. Suppose to the contrary that $a < c$. Since $e^{[n]} = 0$ on $[a, c]$, we have $e = 0$ a.e. or equivalently, $f - g = 0$ a.e. on (a, c). If $n = 1$, then $f - g$ is U_1-concave. Hence $(f - g)/w_0$ is nonincreasing and e cannot have any sign changes in J. By Lemma 5.6 with $n = 1$ as applied to J, we find that e has at least one sign change. This contradiction establishes that $a = c$. Similarly, $b = d$ in this case. Now suppose that $n \geq 2$. By Lemma 5.6, e and, hence, $f - g$ has at least n sign changes in J. Since $f - g = 0$ a.e. on (a, c), $h_1 = w_1 l_1(f - g)$ has n sign changes in J. However, $-h_1$ is

U_{n-1}-convex on J. This contradicts Proposition 2.2(1) that $-h_1$ has at most $n-1$ sign changes in J. Thus $a=c$ and similarly $b=d$.

If $p=1$ and f is continuous on $[a,b]$, then g is unique since V_n is a Tchebycheff space. ∎

Next, we obtain sufficient conditions to ensure that a best approximation to $f \in L_1$ from $K_{n,1}(S)$ is unique. Recall that Theorem 2.4 ensures the existence of a best approximation if S is closed in I.

THEOREM 5.8. *Let $n \geqslant 1$ and $f \in L_1 \backslash K_{n,1}(S)$. Assume that for all $J=(c,d) \subset I$ and for all finite or regular infinite partitions π of J, $f|J$ has a unique best L_1-approximation from $T_1(J,\pi)$. Then a best L_1-approximation to f from $K_{n,1}(S)$, if it exists, is unique.*

Proof. Let g and h be two best approximation to f from $K_{n,1}(S)$. We show that $g=h$ a.e. By statement (5) of Theorem 4.7 we find that there exists $e \in D(f-g)$ such that $g=f$ a.e. on $I\backslash B$ and on each component J of B, $g|J$ is a best $L_1(J)$-approximation to $f|J$ from $T_1'(J,\pi)$, where π is the finite or regular infinite partition of J formed by the zeros of $e^{[n]}$ in $J \cap S$. Then, the three statements of Theorem 2.7 hold and, in particular, we have $\int_I eg=0$. Now $h \in K_{n,1}(S)$ and, by a well known result (see, e.g., [4] or [5]), we have $e \in (K_{n,1}(S))^0$, the dual cone of $K_{n,1}(S)$. Hence $\int_I eh \leqslant 0$. Since $\|e\|_\infty=1$, we have $\|f-g\|_1=\int_I e(f-g)=\int_I ef \leqslant \int_I e(f-h) \leqslant \|f-h\|_1$. Since $\|f-g\|_1=\|f-h\|_1$, equality holds throughout in the above expression, showing that $\int_I eh=0$ and $\int_I e(f-h)=\|f-h\|_1$. Again since $\|e\|_\infty=1$, we conclude that $e=\operatorname{sgn}(f-h)$ a.e. on the set $\{t\colon f(t)-h(t) \neq 0\}$, i.e., $e \in D(f-h)$ which is defined by (2.6). As observed above, the first two statements of Theorem 2.7 hold for e, and $\int_I eh=0$; thus, all three statements hold for e and h. Now applying statement (5) of Theorem 4.7 for h and e, we conclude that $h=f$ a.e. on $I\backslash B$ and, on each component J of B, $h|J$ is a best $L_1(J)$-approximation to $f|J$ from $T_1'(J,\pi)$. By hypothesis, $g|J=h|J$ a.e. on each component J of B giving $g=h$ a.e. on B. Again, $g=f=h$ a.e. on $I\backslash B$, and hence, $g=h$ a.e. on I. ∎

THEOREM 5.9. *Let $n \geqslant 1$, f be continuous on $[a,b]$ and g be a best L_1-approximation to f from $K_{n,1}(S)$. Let e and B be as in statement (5) of Theorem 4.7. Assume that $f-g$ has a finite number of sign changes in each component of B or that S is finite. Then g is the unique best approximation to f from $K_{n,1}(S)$. In particular, g is unique if $f-g$ has a finite number of sign changes in I.*

Proof. Let $J=(c,d)$ be a component of B. Exactly as in the proof of Theorem 4.6, by applying Lemma 4.5, we may show that $e^{[i]}(c)=e^{[i]}(d)=0$, $1 \leqslant i \leqslant n$. If m is the finite number of sign changes of $f-g$ in J, then, by Rolle's theorem, $e^{[1]}$ has at most $(m+2)-1$ separated zeros in $[c, d]$. Applying this argument in succession, or by induction, we conclude that $e^{[n]}$ has at most $(m+2)-n$ separated zeros in $[c, d]$. Now, by Lemma 4.1(4), the set of zeros of $e^{[n]}$ in J is countable. Hence $e^{[n]}$ cannot be identically equal to zero on any nondegenerate subinterval of $[c, d]$. Therefore, $e^{[n]}$ has at most $m-n+2$ ordinary zeros in $[c, d]$, and at most $m-n$, i.e., a finite number of zeros in J. If π is the partition of J formed by the zeros of $e^{[n]}$ in $J \cap S$ as in Theorem 4.7, then π is finite. Hence $T_1'(J, \pi) = T_1(J, \pi)$ is the space of ordinary Tchebycheff splines on J with simple knots in $J \cap \pi$. Since f is continuous on $[c, d]$, by the theory of Property A [13] as applied to the ECT system and $T_1(J, \pi)$, we conclude that a best approximation to $f \mid J$ from $T_1(J, \pi)$ is unique. Then the result follows as in Theorem 5.7. If S is finite, then again π is finite and the above arguments establish the result. The last statement of the theorem also follows as above. ∎

6. Approximation by Generalized Monotone and Convex Functions

In this section we consider the problem of finding a best L_p-approximation to f from $K_{n,p}(S)$ for $n=1$ and 2. We derive stronger characterization theorems than those in Section 4 and establish the uniqueness of L_1-approximation. A function k in K_1 (K_n with $n=1$) is called a generalized monotone (nondecreasing) function and is defined by $k(x)/w_0(x) \leqslant k(y)/w_0(y)$ if $x \leqslant y$ as may be easily verified. The functions in K_2 are called generalized convex functions. Recall that E^- and E^+ are defined by (2.5) and (4.1), respectively, and $\hat{S}$ and Z_J are defined in Section 5.

THEOREM 6.1. *Let $1<p<\infty$, $n=1$ or 2, $f \in L_p \backslash K_{n,p}(S)$, and $g \in K_{n,p}(S)$. If $n=2$, assume additionally that $Z_I(e^{[2]}) \subset \hat{S}$. Then the following four statements are equivalent.*

(1) *g is the best L_p-approximation to f from $K_{n,p}(S)$.*

(2) *$e^{[i]}(b)=0$ for $1 \leqslant i \leqslant n$, $(-1)^n e^{[n]}(t) \leqslant 0$ for $t \in S$, and, on each component J of E^-, $g \mid J \in V_n \mid J$.*

(3) *$e^{[i]}(b)=0$ for $1 \leqslant i \leqslant n$, $(-1)^n e^{[n]}(t) \leqslant 0$ for $t \in S$, and, on each component J of E^-, $g \mid J$ is the best $L_p(J)$-approximation to $f \mid J$ from $V_n \mid J$.*

(4) *$g=f$ a.e. on $I \backslash (E^- \cup E^+)$, and, on each component J of E^- or E^+, $g \mid J$ is the best $L_p(J)$-approximation to $f \mid J$ from each of $V_n \mid J$ and $K_{n,p}(J, S)$.*

Proof. We first establish the theorem for $n=1$. The equivalence of (1) and (2) is a restatement of Theorem 2.6 for $n=1$. We simultaneously show that (1) or (2) implies (3) and (4). Let g be the best approximation to f. If $J=(c, d)$ is a component of E^-, then by (2), $g|J \in V_1|J$. If J is a component of E^+, then, since $e^{[1]} \geqslant 0$ on S by (2), we must have $J \subset I \backslash S = S'$. Hence $\mu_{g,1}(J)=\mu_{g,1}(S')=0$. Thus $g=\alpha w_0$ on J for some α, and $g|J \in V_1|J$. Now, let $J=(c, d)$ be a component of E^- or E^+. Then by the definition of E^- and E^+, we have $e^{[1]}(c)=e^{[1]}(d)=0$. Hence, by Lemma 4.4 with $n=1$, we conclude that $g|J$ is the best approximation to $f|J$ from $K_{1,p}(J, S)$. But since $g|J \in V_1|J \subset K_{1,p}(J, S)$, we find that $g|J$ is also the best approximation to $f|J$ from $V_1|J$. By Theorem 4.6(5), we have $g=f$ a.e. on $I \backslash B$. By Lemma 4.1(5), $B=E^- \cup E^+ \cup B_0$, where B_0 is a countable set and hence its Lebesgue measure is zero. It follows that $g=f$ a.e. on $I \backslash (E^- \cup E^+)$. Thus (3) and (4) are established. Clearly, (3) implies (2). Now we may show (4) implies (1) by a proof similar to the one used to establish that statement (2) implies statement (1) in Theorem 4.6. Thus, the four statements are equivalent for $n=1$.

Now let $n=2$ and $J=(c, d)$ be a component of E^- or E^+. By the definitions of E^- and E^+ we conclude that $e^{[2]}(c)=e^{[2]}(d)=0$. By Lemma 5.3 with $n=2$ we have $Z_I(e^{[2]}) \subset Z_I(e^{[1]})$. Hence $e^{[1]}(c)=e^{[1]}(d)=0$. Now, by Lemma 4.4 with $n=2$ we find that $g|J$ is the best approximation to f from $K_{2,p}(J, S)$. The rest of the proof is similar to the case $n=1$. ∎

The following analogue of Theorem 6.1 for $p=1$ may be proved by similar methods.

THEOREM 6.2. *Let $p=1$, $n=1$ or 2, $f \in L_1 \backslash K_{n,1}(S)$, and $g \in K_{1,1}(S)$. If $n=2$, assume additionally that $Z_I(e^{[2]} \subset \hat{S}$. Then g is a best L_1-approximation to f from $K_{n,1}(S)$ if and only if there exists $e \in D(f-g)$ satisfying one of the statements (2), (3), and (4) of Theorem 6.1 with "the best $L_p(J)$-approximation" in those statements replaced by "a best $L_1(J)$-approximation," p replaced by 1, and the sets E^- and E^+ defined by using this e.*

We remark that in Theorem 6.1 another equivalent statement (4′) may be obtained from statement (4) by replacing its part "on each component J of E^- or E^+, $g|J$ is the best $L_p(J)$-approximation to $f|J$ from $K_{n,p}(J, S)$" by the following: "If $J=(c, d)$, where $c<d$ are any two zeros of $e^{[n]}$ (and, in particular, if J is a component of E^- or E^+), then $g|J$ is the best $L_p(J)$-approximation to $f|J$ from $K_{n,p}(J, S)$." This may be justified by Lemma 4.4. The part of (4) pertaining to the best $L_p(J)$-approximation from $V_n|J$ remains unchanged. Similar remarks apply to Theorem 6.2 also.

Next we investigate the continuity properties of a best L_p-approximation from $K_{1,p}(S)$. Recall that the functions in $K_{n,p}(S)$, $n \geqslant 2$, are continuous on I. We observe that if g is a generalized monotone function then, by Proposition 2.2(2), $g(x^-)$ and $g(x^+)$ exist for each x in I. Given $x \in I$ and an open interval $J \subset I$ we write $J \to x$ and say "J shrinks to x" whenever

$J=(x-\delta_1, x+\delta_2)$ with $\delta_1 \geqslant 0$, $\delta_2 \geqslant 0$ and $\delta_1+\delta_2>0$ (so that $x \in \mathrm{cl}(J)$) and $\delta_1 \downarrow 0$, $\delta_2 \downarrow 0$ [1]. We denote by $\mathrm{int}(S)$ and $\mathrm{bd}(S)$, the interior and boundary of S. In the following λ is the Lebesgue measure on I.

THEOREM 6.3. *Let* $1 \leqslant p < \infty$, $f \in L_p \backslash K_{1,p}(S)$, *and* g *be a best* L_p*-approximation to* f *from* $K_{1,p}(S)$. *Then the following holds.*

(1) *g is continuous at each* $x \in S' = I \backslash S$. *If* $x \in S$ *and* $e^{[1]}(x) > 0$, *then* g *is continuous at* x.

(2) *Let* $x \in \mathrm{int}(S)$, $e^{[1]}(x)=0$, *and* $P \subset I$ *be any measurable set such that*

$$\lambda(P \cap J)/\lambda(J)^p \to 0, \qquad \text{as} \quad J \to x. \tag{6.1}$$

If $P' = I \backslash P$; *then*

$$\begin{aligned} &\lim \operatorname{ess\,inf}\{f(t) \colon t \in P', t \uparrow x\} \\ &\qquad \leqslant g(x^-) \leqslant g(x^+) \leqslant \lim \operatorname{ess\,sup}\{f(t) \colon t \in P', t \downarrow x\}. \end{aligned} \tag{6.2}$$

Remark. Condition (6.1) implies $\lambda(P' \cap J)/\lambda(J) \to 1$, hence, in (6.2) there exist $t \in P'$ with $t \uparrow x$ and $t \downarrow x$.)

Proof. (1) Suppose that g is discontinuous at $x \in I$. Then $g(x^+) > g(x^-)$. The Lebesgue–Stieltjes measure $\mu_{g,1}$ generated by $g(t^+)/w_0(t)$ coincides with that generated by $g(t)/w_0(t)$, and $\mu_{g,1}\{x\} = (g(x^+) - g(x^-))/w_0(x) > 0$ [12, Proposition 3.9]. Hence $x \in S$. If $e^{[1]}(x) > 0$ for some $x \in I$, then $x \in E^-$ and, by Theorem 6.1(2), there exists α such that $g = \alpha w_0$ on $(x-\delta, x+\delta)$ for some $\delta > 0$. Hence g is continuous at x.

(2) We outline the proof. If θ denotes the right side of (6.2), we assume $\theta < g(x^+)$. Then show that $e^{[1]}(s) < 0$ for all s sufficiently close to x. By Theorem 6.1(2), $s \in I \backslash S = S'$. Hence $x \notin \mathrm{int}(S)$, a contradiction. Thus $g(x^+) \leqslant \theta$. ∎

We say that f on I is p-approximately continuous at $x \in I$ if there exists a measurable set $P \subset I$ such that (6.1) holds and $f \mid P'$ is continuous where $P' = I \backslash P$. Note that 1-approximate continuity is identical to the well known approximate continuity [1]. Clearly, p-approximate continuity for $p > 1$ implies 1-approximate continuity. If f is continuous on I, then, by taking $P = \varnothing$ in (6.1), we see that f is p-approximately continuous. Theorem 6.3 then at one gives us the following.

THEOREM 6.4. *Assume the hypothesis of Theorem* 6.3. *If f is p-approximately continuous on* $\mathrm{int}(S)$, *then g is continuous on* $I \backslash \mathrm{bd}(S)$. *In particular, this conclusion holds if f is continuous on* $\mathrm{int}(S)$.

The following example show that the results of the above theorem cannot be strengthened. Let $I=(0, 1)$, $S=\{1/2\}$, $p=2$, $w_0=1$, and $f(x)=x$ on I. Then $K_{1,2}(S)$ is the set of nondecreasing functions which are constant on $(0, 1/2)$ and $(1/2, 1)$ with a possible jump at 1/2. Then, g, the best L_2-approximation to f from $K_{1,2}(S)$ is given by $g(x)=1/4$, $0<x<1/2$ and $g(x)=3/4$, $1/2 \leqslant x<1$.

For the problem of unconstrained L_1-approximation by nondecreasing functions, [16] shows that g is continuous if f is approximately continuous (their proof holds only for $p=1$). We are motivated by their result and proof to introduce the condition (6.1) and p-approximate continuity so that one single result (6.2) could be obtained for all $1 \leqslant p<\infty$ under uniform assumptions in our more general framework.

We next establish the uniqueness of best approximation from $K_{n,1}(S)$, $n=1, 2$.

THEOREM 6.5. *Let $n=1$ or 2, and f be continuous on $[a, b]$. If $n=2$, assume additionally that $Z_f(e^{[2]}) \subset \hat{S}$. Then a best L_1-approximation to f from $K_{n,1}(S)$, if it exists, is unique.*

Proof. Let J be any component of E^- or E^+ as in statement (3) of Theorem 6.2. Since f is continuous on cl(J), by the theory of Property A [13] applied to $V_n|J$ as in Theorem 5.8, a best approximation to $f|J$ from $V_n|J$ is unique. Then the required result follows by statement (3) of Theorem 6.2 in the same way as Theorems 5.7 and 5.8 follow from statement (5) of Theorem 4.7. ∎

REFERENCES

1. A. M. BRUCKNER, "Differentiation of Real Functions," Lecture Notes in Mathematics, Vol. 659, Springer-Verlag, New York, 1978.
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, *Approx. Theory* **71** (1992), 213–238.
4. F. R. DEUTSCH AND F. H. MASERICK, Applications of the Hahn-Banach theorem in approximation theory, *SIAM Rev.* **9** (1967), 516–530.
5. F. DEUTSCH, V. A. UBHAYA, AND Y. XU, Dual cones, constrained n-convex L_p-approximation and perfect splines, *J. Approx. Theory* **80** (1995), 180–203.
6. R. HOUTARI, R. LEGG, AND D. TOWNSEND, Existence of best n-convex approximants in L_1, *Approx. Theory Appl.* **5** (1989), 51–57.
7. S. KARLIN AND W. J. STUDDEN, "Tchebycheff Systems: With Applications in Analysis and Statistics," Interscience, New York, 1966.
8. J. L. KELLY AND I. NAMIOKA, "Linear Topological Spaces," Van Nostrand, Princeton, NJ, 1963.
9. R. M. MATHSEN AND V. A. UBHAYA, Generalized convex functions and best L_p-approximation, *Proc. Amer. Math. Soc.* **114** (1992), 733–740.

10. C. A. MICCHELLI, P. W. SMITH, J. J. SWETITS, AND J. D. WARD, Constrained L_p-approximation, *Constr. Approx.* **1** (1985), 93–102.
11. 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.
12. A. MUKHERJEA AND K. POTHOVEN, "Real and Functional Analysis, Part A, Real Analysis," Plenum, New York, 1984.
13. A. M. PINKUS, "On L^1-Approximation," Cambridge Univ. Press, Cambridge, 1989.
14. A. W. ROBERTS AND D. E. VARBER, "Convex Functions," Academic Press, New York, 1973.
15. L. L. SCHUMAKER, "Spline Functions: Basic Theory," Wiley, New York, 1981.
16. P. W. SMITH AND J. J. SWETITS, Best approximation by monotone functions, *J. Approx. Theory* **44** (1987), 389–403.
17. J. J. SWETITS, S. E. WEINSTEIN, AND Y. XU, On the characterization and computation of best monotone approximation in $L_p[0,1]$ for $1 \leqslant p < \infty$, *J. Approx. Theory* **60** (1990), 58–69.
18. 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.
19. J. J. SWETITS, S. E. WEINSTEIN, AND Y. XU, Best L_p approximation with multiple constraints for $1 \leqslant p < \infty$, *J. Approx. Theory* **65** (1991), 90–108.
20. V. A. UBHAYA, Duality in approximation and conjugate cones in normed linear spaces, *J. Math. Anal. Appl.* **58** (1977), 419-436.
21. 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.
22. V. A. UBHAYA, L_p approximation from nonconvex subsets of special classes of functions, *J. Approx. Theory* **57** (1989), 223–238.
23. V. A. UBHAYA, Duality and Lipschitzian selections in best approximation from nonconvex cones, *J. Approx. Theory* **64** (1991), 315–342.