MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 118, APRIL 1972

Uniform Approximation Through Partitioning

By S. E. Weinstein

Abstract. In this paper, the problem of best uniform polynomial approximation to a continuous function on a compact set X is approached through the partitioning of X and the definition of norms corresponding to the partition and each of the standard L_p norms $1 \leqq p < \infty$. For computational convenience, a pseudo norm is defined corresponding to each partition. When the partition is chosen appropriately, the corresponding best approximations (using both the norms and the pseudo norm) are arbitrarily close to a best uniform approximation. A chracterization theorem for best pseudo norm approximation is presented, along with an alternation theorem for best pseudo norm approximation to a univariate function.

1. Introduction. The central problem of this paper is the best uniform polynomial approximation to a continuous function on a compact set. The solution of this problem is approached through the partitioning of the set and defining a norm corresponding to the partition. The unique best approximations in these norms are used as approximations to the desired best uniform approximation.

In Section 2, it is shown that the partitions can be chosen so that the corresponding norm is close to the uniform norm and such that the corresponding best approximation is close to a best uniform approximation.

In Section 3, best approximation in these norms is characterized.

In Sections 4 and 5, we consider a pseudo norm which corresponds to the partition of the compact set. This pseudo norm has computational advantages when compared to the norms of Sections 2 and 3.

2. Norms Defined Through Partitioning. Let X be a compact metrizable set, and let μ be a strictly positive measure on X such that all continuous functions on X are measurable. Let $X = \bigcup_{i=1}^{k} E_i^k$ such that

$$\mu(E_i \cap E_j) = 0, \qquad i \neq j, \tag{2.1}$$

and

$$\mu(E_i) > 0, \qquad i = 1, 2, \cdots, k. \tag{2.2}$$

Such a measure-wise decomposition of X will be called an acceptable partition of X.

Definition 2.1. Let $U_k = \{E_i\}_{i=1}^{k}$ be an acceptable partition of X. For $1 \leqq p < \infty$ and $f \in L_p(X)$, define

$$r_p(E_i, f) = \left\{\frac{1}{\mu(E_i)} \int_{E_i} |f|^p \, d\mu\right\}^{1/p} \tag{2.3}$$

Received April 12, 1971.

AMS 1969 *subject classifications*. Primary 4115, 4140, 6520; Secondary 6530, 6533.

Key words and phrases. Approximation, uniform approximation, multivariate approximation.

and define a norm on $L_p(X)$ by

$$(2.4)\qquad ||f||_{U_k,p} = \max_{i=1,2,\cdots,k} r_p(E_i, f).$$

The case $p = 2$ is considered in Harris [2] and Weinstein [4] and [5].

Let $\{\varphi_i\}_{i=1}^n$ be a linearly independent set of functions in $C(X)$, and define

$$(2.5)\qquad \Pi_n = \left\{P_n(A; \cdot) = \sum_{i=1}^n a_i\varphi_i(\cdot): A = (a_1, \cdots, a_n) \in R^n\right\}.$$

Definition 2.2. Let U_k be an acceptable partition of X. For $1 \leqq p < \infty$ and $f \in L_p(X)$, there exists a polynomial $P_n(A(U_k, p); \cdot) \in \Pi_n$ which minimizes $||f - P_n(A; \cdot)||_{U_k,p}$ over all $P_n(A; \cdot) \in \Pi_n$. Such a polynomial is called a best $||\cdot||_{U_k,p}$ approximation to f.

We next show the uniqueness of the best $||\cdot||_{U_k,p}$ approximation for $1 < p < \infty$.

LEMMA 2.3. *Suppose* $1 < p < \infty$. *If* $||f||_{U_k,p} = ||g||_{U_k,p} = 1$ *and* $||\frac{1}{2}(f + g)||_{U_k,p} = 1$, *then* $f \equiv g$ *a.e. on all subsets* $E_i \subset U_k$ *such that* $r_p(E_i, \frac{1}{2}(f + g)) = 1$.

Proof. This is an immediate consequence of the strict convexity of the L_p norms for $1 < p < \infty$.

THEOREM 2.4. *Let* U_k *be an acceptable partition of* X, *and suppose that*

$$P_n(A; x) = P_n(B; x), \quad \text{a.e. on some } E_i \subset U_k,$$

implies that $A = B$.

Then, for $1 < p < \infty$, *each* $f \in L_p(X)$ *has a unique best* $||\cdot||_{U_k,p}$ *approximation from* Π_n.

Proof. Lemma 2.3 implies that if

$$||f - P_n(A; \cdot)||_{U_k,p} = ||f - P_n(B; \cdot)||_{U_k,p},$$

then $P_n(A; x) = P_n(B; x)$ a.e. on some $E_i \subset U_k$. Therefore, $A = B$.

If Π_n is a Haar subspace of $C(X)$, then, since $\mu(E_i) > 0$ for all $E_i \subset U_k$, the hypothesis of Theorem 2.4 holds.

Let $||\cdot||_\infty$ denote the uniform norm defined on $C(X)$ by $||f||_\infty = \max_{x\in X} |f(x)|$, $f \in C(X)$.

Theorem 2.5. *Let* U_k *be an acceptable partition of* X *and let* U_l, $l > k$, *be an acceptable refinement of* U_k. *Then, for* $1 \leqq p < \infty$ *and* $f \in C(X)$,

$$(2.6)\qquad ||f||_{U_k,p} \leqq ||f||_{U_l,p} \leqq ||f||_\infty.$$

Proof. Suppose $U_k = \{E_i\}_{i=1}^k$ and $U_l = \{D_i\}_{i=1}^l$. To prove the first inequality, it suffices to show that, if $E_\nu = \bigcup_{i=1}^m D_i$, then

$$\max_{i=1,2,\cdots,m} r_p(D_i, f) \geqq r_p(E_\nu, f).$$

Suppose that

$$r_p(D_i, f) < r_p(E_\nu, f), \quad \text{for } i = 1, 2, \cdots, m.$$

Then

$$\frac{1}{\mu(D_i)}\int_{D_i} |f|^p \, d\mu < \frac{1}{\mu(E_\nu)}\int_{E_\nu} |f|^p \, d\mu$$

or

$$\int_{D_i} |f|^p \, d\mu < \frac{\mu(D_i)}{\mu(E_\nu)} \left\{ \int_{D_1} |f|^p \, d\mu + \cdots + \int_{D_m} |f|^p \, d\mu \right\}, \qquad i = 1, \cdots, m.$$

Sum both sides of this inequality over i to obtain the following contradiction:

$$\int_{D_1} |f|^p \, d\mu + \cdots + \int_{D_m} |f|^p \, d\mu < \int_{D_1} |f|^p \, d\mu + \cdots + \int_{D_m} |f|^p \, d\mu.$$

For each $D_i \in U_l$ and $1 \leqq p < \infty$, there exists a point $x_i(p) \in D_i$ such that

$$r_p(D_i, f) = \left\{ \frac{1}{\mu(D_i)} \int_{D_i} |f|^p \, d\mu \right\}^{1/p} \leqq |f(x_i(p))|.$$

Therefore,

$$||f||_{U_l,p} = \max_{i=1,\cdots,l} |f(x_i(p))| \leqq ||f||_\infty.$$

We next consider the closeness of $||f||_{U_k,p}$ and $||f||_\infty$.

$r_p(E_j, \cdot)$ is a weighted L_p norm on E_j, $j = 1, \cdots, k$. By Pólya's algorithm, $r_p(E_j, f) \to \sup_{x \in E_j} |f(x)|$ as $p \to \infty$. This immediately implies the following:

THEOREM 2.6. *Let $f \in C(X)$, and let U_k be an acceptable partition of X. Then*

$$||f||_{U_k,p} \to ||f||_\infty, \quad \text{as } p \to \infty. \tag{2.7}$$

We are now concerned with the closeness of $||f||_{U_k,p}$ and $||f||_\infty$ for fixed values of p.

Definition 2.7. Let

$$U_k^* = \{E_i \in U_k \text{: there exists a point } e_i \in E_i \text{ such that } |f(e_i)| = ||f||_\infty\} \tag{2.8}$$

and define

$$\gamma(U_k^*) = \min_{E_i \in U_k^*} \sup_{x,y \in E_i} \sigma(x, y), \tag{2.9}$$

where σ is the metric on X.

THEOREM 2.8. *Given $f \in C(X)$, U_k an acceptable partition of X, and* $1 \leqq p < \infty$,

$$||f||_{U_k,p} \leqq ||f||_\infty \leqq ||f||_{U_k,p} + \omega(\gamma(U_k^*)), \tag{2.10}$$

where ω is the modulus of continuity of f on X.

Proof. The first inequality follows from Theorem 2.5.

For each $E_i \in U_k^*$, let $x_i(p) \in E_i$ be as defined in Theorem 2.5, and let $e_i \in E_i$ be as defined in Definition 2.7. Then,

$$\begin{aligned} ||f||_\infty - ||f||_{U_k,p} &\leqq |f(e_i)| - |f(x_i(p))| \leqq |f(e_i) - f(x_i(p))| \\ &\leqq \omega\left(\sup_{x,y \in E_i} \sigma(x, y) \right). \end{aligned}$$

The proof is completed by minimizing the right-hand side of this inequality over all $E_i \in U_k^*$, and noting that

$$\omega(\gamma(U_k^*)) = \min_{E_i \in U_k^*} \omega\left(\sup_{x,y \in E_i} \sigma(x, y) \right).$$

Definition 2.9. Let $U_k = \{E_i\}_{i=1}^k$ be an acceptable partition of X. Define the mesh width for U_k by

$$\delta(U_k) = \max_{E_i \in U_k} \sup_{x, y \in E_i} \sigma(x, y). \tag{2.11}$$

COROLLARY 2.10. *If $f \in C(X)$, and $\{U_k\}_{k=1}^{\infty}$ is a sequence of acceptable refinements of X such that*

$$\delta(U_k) \to 0, \quad as\ k \to \infty, \tag{2.12}$$

then, for $1 \leqq p < \infty$,

$$||f||_{U_k, p} \uparrow ||f||_\infty \quad as\ k \to \infty. \tag{2.13}$$

Proof. It suffices to note that

$$\delta(U_k) \geqq \gamma(U_k^*),$$

and to apply Theorem 2.8.

COROLLARY 2.11. *If $f \in C(X)$, and f has an extreme point x^* in the interior of X, then, given any fixed positive integer k, $1 \leqq p < \infty$ and $\epsilon > 0$, there exists an acceptable partition U_k of X so that*

$$||f||_{U_k, p} \leqq ||f||_\infty \leqq ||f||_{U_k, p} + \epsilon. \tag{2.14}$$

Proof. Let $\delta = \delta(\epsilon)$ correspond to the definition of the continuity of f on X. The set $E_1 = \{x \in X : \sigma(x, x^*) < \delta(\epsilon)\}$ has positive measure. Let U_k contain the set E_1. Then

$$\gamma(U_k^*) \leqq \delta \quad \text{and} \quad \omega(\gamma(U_k^*)) \leqq \epsilon.$$

We apply Theorem 2.8 to complete the proof.

Theorems 2.6, 2.8, Corollaries 2.10 and 2.11 motivate the use of the norm $||\cdot||_{U_k, p}$ as a substitute for the uniform norm. This in turn motivates the use of $P_n(A(U_k, p); \cdot)$ as an approximation to f which is nearly a best uniform approximation.

The following theorem found in [3] shows that, if $||f - P_n(A; \cdot)||$ is nearly minimal, then $P_n(A; \cdot)$ is close to a best $||\cdot||$ approximation to f.

THEOREM 2.12. *Let $||\cdot||$ be any norm on $C(X)$ and let $f \in C(X)$. Let*

$$\rho = \inf_{P_n(A; \cdot) \in \Pi_n} ||f - P_n(A; \cdot)||$$

and define

$$\mathcal{A}^* = \{A^* \in R^n : ||f - P_n(A^*; \cdot)|| = \rho\}. \tag{2.15}$$

Given any $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that

$$||f - P_n(A; \cdot)|| \leqq \rho + \delta \tag{2.16}$$

implies

$$\sigma_n(A, \mathcal{A}^*) = \inf_{A^* \in \mathcal{A}^*} \sigma(A, A^*) < \epsilon, \tag{2.17}$$

where σ_n is the Euclidean metric on R^n.

Theorem 2.12 and Corollary 2.10 can be combined to show that, if $\delta(U_k)$ is small, then $P_n(A(U_k, p); \cdot)$ is nearly a best uniform approximation to f on X.

THEOREM 2.13. *Given $f \in C(X)$, $1 \leqq p < \infty$ and any $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that for any acceptable partition U_k of X*

$$\delta(U_k) \leqq \delta \tag{2.18}$$

implies

$$\sigma_n(A(U_k, p), \mathfrak{A}^*) < \epsilon, \tag{2.19}$$

where $\mathfrak{A}^*$ is the set of parameters of the best uniform approximations to f on X.

Proof. Harris [2] proves the case $p = 2$. The proof for $1 \leqq p < \infty, p \neq 2$, is exactly the same. Thus, the details are omitted.

Theorem 2.13 provides the basis for an algorithm for the computation of a best uniform approximation to f on X. This algorithm is discussed in [5]. However, the necessity of increasing k is a disadvantage.

As an alternative to increasing k, we can adjust the partition U_k, while keeping k fixed, so that the resulting best $||\cdot||_{U_k,p}$ approximation is nearly a best uniform approximation. An algorithm to achieve this is presented in [5]. This procedure is motivated by the following:

Definition 2.14. A subset $S = \{x_i\}_{i=1}^{\nu}$ of the extremal points of $f - P^*$, where P^* is a best uniform polynomial approximation to f on X, is said to be a critical point set if P^* is a best uniform approximation to f on S, but is not a best uniform approximation to f on any proper subset of S.

THEOREM 2.15. *Let P^* be a best uniform approximation to $f \in C(X)$ and let $S = \{x_i\}_{i=1}^{\nu}$ be a critical point set for $f - P^*$, and, for $\delta > 0$,*

$$E_i(\delta) = \{x \in X : \sigma(x, x_i) < \delta\}, \qquad i = 1, 2, \cdots, \nu. \tag{2.20}$$

Given any positive integer $k > \nu$, $1 \leqq p < 0$ and $\epsilon > 0$, there exists an acceptable partition $U_k = \{E_i\}_{i=1}^{k}$ of X, such that

$$\sigma_n(A(U_k, p), \mathfrak{A}^*) < \epsilon, \tag{2.21}$$

where $\mathfrak{A}^*$ is the set of parameters of the best uniform approximations to f on X.

Proof. The case $p = 2$ is proved in [4]. The basic idea of this proof is identical to that given there. Thus, we shall omit the details.

3. Characterization of a Best $||\cdot||_{U_k,p}$ Approximation. The following theorem appears in Cheney [1].

THEOREM 3.1. *Let U be a compact subset of R^n. A necessary and sufficient condition that the system of linear inequalities $\langle u, z\rangle > 0$ $(u \in U)$ be inconsistent is that the origin in R^n belongs to $K(U)$, where $K(U)$ denotes the convex hull of U.*

This theorem is used to characterize a best $||\cdot||_{U_k,p}$ approximation, for p even.

THEOREM 3.2 (CHARACTERIZATION). *Let $U_k = \{E_i\}_{i=1}^{k}$ be an acceptable partition of X, and let $f \in C(X)$. Let p be an even positive integer.*

$||f - P_n(A; \cdot)||_{U_k,p}$ is a minimum over all $P_n(A; \cdot) \in \Pi_n$, if and only if the origin in R^n belongs to the set

$$K\Big\{C^i + 2\sum_{i_2=1}^{n} a_{i_2}C^i_{i_2} + 3\sum_{i_2=1;\, i_3=1}^{n} a_{i_2}a_{i_3}C^i_{i_2 i_3} + \cdots + (p-1)\sum_{i_2=1,\cdots,i_{p-1}=1}^{n} a_{i_2}\cdots a_{i_{p-1}}C^i_{i_2\cdots i_{p-1}} + p\sum_{i_2=1,\cdots,i_p=1}^{n} a_{i_2}\cdots a_{i_p}D^i_{i_2\cdots i_p}: r_p(E_i, f - P_n(A;\cdot)) = ||f - P_n(A;\cdot)||_{U_k,p}\Big\}, \tag{3.1}$$

where for $i_1 = 1, \cdots, n$, $i_2 = 1, \cdots, n, \cdots, i_p = 1, \cdots, n$, *and* $l = 1, \cdots, p-1$,

$$c^j_{i_1 \cdots i_l} = \frac{\binom{P}{l}}{\mu(E_j)} \int_{E_j} f^{p-l} \varphi_{i_1} \cdots \varphi_{i_l} \, d\mu, \tag{3.2}$$

$$C^j = (c^j_1, c^j_2, \cdots, c^j_n), \tag{3.3}$$

$$C^j_{i_2 \cdots i_l} = (c^j_{1, i_2, \cdots, i_l}, c^j_{2, i_2, \cdots, i_l}, \cdots, c^j_{n, i_2, \cdots, i_l}), \tag{3.4}$$

$$d^j_{i_1 \cdots i_p} = \frac{1}{\mu(E_j)} \int_{E_j} \varphi_{i_1} \cdots \varphi_{i_p} \, d\mu, \tag{3.5}$$

$$D^j_{i_2 \cdots i_p} = (d^j_{1, i_2, \cdots, i_p}, d^j_{2, i_2, \cdots, i_p}, \cdots, d^j_{n, i_2, \cdots, i_p}). \tag{3.6}$$

Proof. The proof is essentially the same as the proof of the characterization theorem on p. 73 of Cheney [1]. Thus, we prove only the sufficiency, and omit the details of the necessity. p is assumed even, so that the definition of $r_p(E_i, f)$ does not include an absolute value under the integral sign.

Suppose that $\|f - P_n(A; \cdot)\|_{U_k, p}$ is not a minimum. Then, there exists a $B \in R^n$ such that

$$\|f - P_n(A - B; \cdot)\|_{U_k, p} < \|f - P_n(A; \cdot)\|_{U_k, p}.$$

Then, for all j such that

$$r_p(E_j, f - P_n(A; \cdot)) = \|f - P_n(A; \cdot)\|_{U_k, p}, \tag{3.7}$$

$$r_p(E_j, f - P_n(A - B; \cdot)) < r_p(E_j, f - P_n(A; \cdot)).$$

Also, for $0 < \lambda \leqq 1$,

$$r_p(E_j, f - P_n(A - \lambda B; \cdot)) < r_p(E_j, f - P_n(A; \cdot)).$$

Using the notation of (3.2)–(3.6), this is equivalent to

$$\begin{aligned} r^p_p(E_j, f - P_n(A - \lambda B; \cdot)) &= r^p_p(E_j, f - P_n(A; \cdot)) \\ &\quad - \lambda \sum_{i_1=1}^n b_{i_1} \Big\{ c^j_{i_1} + 2 \sum_{i_2=1}^n a_{i_2} c^j_{i_1 i_2} + 3 \sum_{i_2=1; i_3=1}^n a_{i_2} a_{i_3} c^j_{i_1 i_2 i_3} + \cdots \\ &\qquad + (p-1) \sum_{i_2=1, \cdots, i_{p-1}=1}^n a_{i_2} \cdots a_{i_{p-1}} c^j_{i_1 \cdots i_{p-1}} \\ &\qquad + p \sum_{i_2=1, \cdots, i_p=1}^n a_{i_2} \cdots a_{i_p} d^j_{i_1 \cdots i_p} \Big\} + O(\lambda^2) \\ &= r^p_p(E_j, f - P_n(A; \cdot)) \\ &\quad - \lambda B \cdot \Big(C^j + 2 \sum_{i_2=1}^n a_{i_2} C^j_{i_2} + 3 \sum_{i_2=1; i_3=1}^n a_{i_2} a_{i_3} C^j_{i_2 i_3} + \cdots \\ &\qquad + (p-1) \sum_{i_2=1, \cdots, i_{p-1}=1}^n a_{i_2} \cdots a_{i_{p-1}} C^j_{i_2 \cdots i_{p-1}} \\ &\qquad + p \sum_{i_2=1, \cdots, i_p=1}^n a_{i_2} \cdots a_{i_p} D_{i_2 \cdots i_p} \Big) \\ &\quad + O(\lambda^2) < r^p_p(E_j, f - P_n(A; \cdot)). \end{aligned}$$

For this inequality to hold for $0 < \lambda \leqq 1$ and λ sufficiently small, it is necessary for the inner product appearing there to be positive for all j satisfying (3.7). Theorem 3.1 completes the sufficiency.

4. Uniform Approximation Through the Use of a Pseudo Norm. Theorem 3.2 does not readily lead to a computational algorithm for best $\|\cdot\|_{U_k,p}$ approximation. As an alternative, Harris [2] suggests the use of a pseudo norm defined as follows:

Definition 4.1. Let $U_k = \{E_i\}_{i=1}^k$ be an acceptable partition of X. For $f \in C(X)$, define

$$s(E_i, f) = \frac{1}{\mu(E_i)} \int_{E_i} f \, d\mu \tag{4.1}$$

and define a pseudo norm by

$$ps\|f\|_{U_k} = \max_{i=1,\cdots,k} |s(E_i, f)|. \tag{4.2}$$

There exists a best $ps\|\cdot\|_{U_k}$ polynomial approximation to f. However, the best approximation is not in general unique.

Several of the results of Section 2 for $\|\cdot\|_{U_k}$ extend to $ps\|\cdot\|_{U_k}$.

THEOREM 4.2. *Let U_k be an acceptable partition of X, and let U_l, $l > k$, be an acceptable refinement of U_k. Then, for $f \in C(X)$,*

$$ps\|f\|_{U_k} \leqq ps\|f\|_{U_l} \leqq \|f\|_\infty \tag{4.3}$$

and

$$\|f\|_\infty \leqq ps\|f\|_{U_k} + \omega(\gamma(U_k^*)). \tag{4.4}$$

Proof. Suppose $U_k = \{E_i\}_{i=1}^k$ and $U_l = \{D_i\}_{i=1}^l$. Consider $E_\nu = \bigcup_{i=1}^m D_i$. Suppose that

$$|s(D_i, f)| < |s(E_\nu, f)|, \quad \text{for } i = 1, \cdots, m.$$

Then

$$\left|\frac{1}{\mu(D_i)} \int_{D_i} f \, d\mu\right| < \left|\frac{1}{\mu(E_\nu)} \int_{E_\nu} f \, d\mu\right|$$

or

$$\left|\int_{D_i} f \, d\mu\right| < \frac{\mu(D_i)}{\mu(E_\nu)} \left|\int_{D_1} f \, d\mu + \cdots + \int_{D_m} f \, d\mu\right|.$$

Sum both sides of this inequality over i to obtain the following contradiction:

$$\left|\int_{D_1} f \, d\mu\right| + \cdots + \left|\int_{D_m} f \, d\mu\right| < \left|\int_{D_1} f \, d\mu + \cdots + \int_{D_m} f \, d\mu\right|.$$

The remainder of the proof is exactly as in Theorems 2.5 and 2.8. Thus, the details are omitted.

COROLLARY 4.3. *If $f \in C(X)$ and $\{U_k\}_{k=1}^\infty$ is a sequence of acceptable refinements of X such that*

$$\delta(U_k) \to 0 \quad \text{as } k \to \infty, \tag{4.5}$$

then

$$ps\|f\|_{U_k} \uparrow \|f\|_\infty \quad as\ k \to \infty\,. \tag{4.6}$$

COROLLARY 4.4. *If $f \in C(X)$ and f has an extreme point x^* in the interior of X, then, given any positive integer k and $\epsilon > 0$, there exists an acceptable partition U_k of X so that*

$$ps\|f\|_{U_k} \leqq \|f\|_\infty \leqq ps\|f\|_{U_k} + \epsilon. \tag{4.7}$$

The following theorem characterizes best polynomial approximation to f in $ps\|\cdot\|_{U_k}$.

THEOREM 4.5. *In order that the coefficients $a_1, \cdots, a_n$ shall render $ps\|r\|_{U_k}$ a minimum where $r \equiv f - \sum_{i=1}^n a_i\varphi_i$, it is necessary and sufficient that the origin of an n-space shall lie in the convex hull of the point set*

$$H = \left\{\int_{E_j} r\, d\mu\, \hat{x}_j : |s(E_j, r)| = ps\|r\|_{U_k}\right\}, \tag{4.8}$$

where $\hat{x}_j$ denotes the n-tuple $[\int_{E_j} \varphi_1\, d\mu, \int_{E_j} \varphi_2\, d\mu, \cdots, \int_{E_j} \varphi_n\, d\mu]$.

Proof. The proof is essentially identical to that given in Cheney [1, p. 73]. We therefore present only the sufficiency.

Suppose $ps\|r - Q\|_{U_k} < ps\|r\|_{U_k}$, where $Q = \sum_{i=1}^n b_i\varphi_i$. For all j such that $|s(E_j, r)| = ps\|r\|_{U_k}$,

$$|s(E_j, r - Q)| < |s(E_j, r)| \quad \text{or} \quad s^2(E_j, r - Q) < s^2(E_j, r),$$

from which we obtain the following system of linear inequalities to be satisfied by $B = (b_1, \cdots, b_n)$:

$$\int_{E_j} r\, d\mu \sum_{i=1}^n b_i \int_{E_j} \varphi_i\, d\mu = \int_{E_j} r\, d\mu\, \langle B, \hat{x}_j\rangle > 0.$$

By Theorem 3.1, $\theta \notin K(H)$.

Define

$$c_j = \frac{1}{\mu(E_j)} \int_{E_j} f\, d\mu \quad \text{and} \quad d_{ij} = \frac{1}{\mu(E_j)} \int_{E_j} \varphi_i\, d\mu, \tag{4.9}$$
$$i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, k.$$

Then,

$$s\left(E_j, f - \sum_{i=1}^n a_i\varphi_i\right) = c_j - \sum_{i=1}^n a_i d_{ij}. \tag{4.10}$$

Therefore, the problem of minimizing $ps\|f - \sum_{i=1}^n a_i\varphi_i\|_{U_k}$ is equivalent to the problem of finding $a_1, \cdots, a_n$ and ρ(min) so that

$$-\rho \leqq c_j - \sum_{i=1}^n a_i d_{ij} \leqq \rho, \qquad j = 1, 2, \cdots, k. \tag{4.11}$$

This is a linear programming problem. This approach has been used to minimize $ps\|f - \sum_{i=1}^n a_i\varphi_i\|_{U_k}$ in [2].

The chief purpose in computing a best $ps\|\cdot\|_{U_k}$ approximation is to compute an approximation which is nearly a best uniform approximation. The corresponding analogues to Theorems 2.13 and 2.15 hold for best $ps\|\cdot\|_{U_k}$ approximations. This justifies their use as nearly best uniform approximations.

5. An Alternation Theorem. This section follows the development in Cheney [1, pp. 73–75].

Consider the case $X = [a, b]$, and choose

$$a = x_0 < x_1 < \cdots < x_k = b, \qquad k \geqq n. \tag{5.1}$$

Define

$$E_j = [x_{j-1}, x_j], \qquad j = 1, 2, \cdots, k. \tag{5.2}$$

Then $U_k = \{E_j\}_{j=1}^k$ is an acceptable partition of X.

Let $\{\varphi_i\}_{i=1}^n$ satisfy the Haar condition on $[a, b]$.

LEMMA 5.1. *Each determinant*

$$\operatorname{Det}(E_{j_1}, \cdots, E_{j_n}) = \begin{vmatrix} \int_{E_{j_1}} \varphi_1 \, d\mu & \int_{E_{j_1}} \varphi_2 \, d\mu & \cdots & \int_{E_{j_1}} \varphi_n \, d\mu \\ \vdots & \vdots & & \vdots \\ \int_{E_{j_n}} \varphi_1 \, d\mu & \int_{E_{j_n}} \varphi_2 \, d\mu & \cdots & \int_{E_{j_n}} \varphi_n \, d\mu \end{vmatrix} \neq 0, \tag{5.3}$$

where $\{E_{j_i}\}_{i=1}^n$ *consists of distinct subsets of* X *as defined by* (5.1) *and* (5.2).

Proof. If $\operatorname{Det}(E_{j_1}, \cdots, E_{j_n}) = 0$, then there exists a nontrivial linear combination of the columns which vanishes. Therefore, there exists a nontrivial polynomial $P = \sum_{i=1}^n a_i\varphi_i$ such that

$$\int_{E_{j_1}} P \, d\mu = \cdots = \int_{E_{j_n}} P \, d\mu = 0.$$

Therefore, P vanishes at least once in the interior of each interval E_{j_i}, $i = 1, 2, \cdots, n$, which contradicts the Haar condition for $\{\varphi_i\}_{i=1}^n$.

Definition 5.2. Let $E_{j_\nu} = [x_{j_\nu - 1}, x_{j_\nu}]$ and $E_{j_\mu} = [x_{j_\mu - 1}, x_{j_\mu}]$ as in (5.1) and (5.2). Then we say

$$E_{j_\nu} < E_{j_\mu} \quad \text{if } x_{j_\nu} \leqq x_{j_\mu - 1}. \tag{5.4}$$

LEMMA 5.3. *Let* (5.1) *and* (5.2) *hold. In addition, choose* $a = y_0 < y_1 < \cdots < y_k = b$ *and define* $D_j = [y_{j-1}, y_j]$ *for* $j = 1, 2, \cdots, k$. *If* $E_{j_1} < E_{j_2} < \cdots < E_{j_n}$ *and* $D_{l_1} < D_{l_2} < \cdots < D_{l_n}$, *then*

$$\operatorname{Det}(E_{j_1}, E_{j_2}, \cdots, E_{j_n}) \quad \textit{and} \quad \operatorname{Det}(D_{l_1}, D_{l_2}, \cdots, D_{l_n})$$

have the same sign.

Proof. Suppose

$$\operatorname{Det}(E_{j_1}, E_{j_2}, \cdots, E_{j_n}) < 0 < \operatorname{Det}(D_{l_1}, D_{l_2}, \cdots, D_{l_n}),$$

then there exists a $\lambda \in (0, 1)$, such that

$$\operatorname{Det}(\lambda E_{j_1} + (1 - \lambda)D_{l_1}, \cdots, \lambda E_{j_n} + (1 - \lambda)D_{l_n}) = 0,$$

where

$$\lambda E_{i_\nu} + (1 - \lambda) D_{l_\nu} = [\lambda x_{i_{\nu-1}} + (1 - \lambda) y_{l_{\nu-1}}, \lambda x_{i_\nu} + (1 - \lambda) y_{l_\nu}],$$
$$\nu = 1, 2, \cdots, n.$$

Lemma 5.1 implies that, for some $\nu < \eta$,

$$\lambda E_{i_\eta} + (1 - \lambda) D_{l_\eta} = \lambda E_{i_\nu} + (1 - \lambda) D_{l_\nu},$$

which implies that the right endpoint of these two intervals must be the same. Therefore,

$$\lambda x_{i_\eta} + (1 - \lambda) y_{l_\eta} = \lambda x_{i_\nu} + (1 - \lambda) y_{l_\nu},$$

which implies

$$\lambda(x_{i_\eta} - x_{i_\nu}) = (1 - \lambda)(y_{l_\nu} - y_{l_\eta}).$$

Since both $\lambda > 0$ and $1 - \lambda > 0$, $x_{i_\eta} - x_{i_\nu}$ and $y_{l_\eta} - y_{l_\nu}$ have opposite signs, which is a contradiction.

LEMMA 5.4. *Let the hypothesis of Lemma* 5.3 *hold, with the additional hypothesis that we choose* $\{E_{i_i}\}_{i=0}^n$ *from* U_k *so that*

$$E_{i_0} < E_{i_1} < \cdots < E_{i_n}. \tag{5.5}$$

Let $\lambda_0, \cdots, \lambda_n$ *be nonzero constants. In order that the origin lie in the convex hull of the n-tuples* $\lambda_0 \hat{x}_{i_0}, \cdots, \lambda_n \hat{x}_{i_n}$, *it is necessary and sufficient that the* λ'*s alternate in sign*: $\lambda_j \lambda_{j-1} < 0$ *for* $j = 1, 2, \cdots, n$.

Proof. The proof is exactly the same as the proof of the lemma on p. 74 of [1], with the exception that Lemma 5.3 replaces the first paragraph of that proof.

THEOREM 5.5 (ALTERNATION THEOREM). *Let* $\{\varphi_i\}_{i=1}^n$ *be a system of elements of* $C[a, b]$ *satisfying the Haar condition, and let* X *be any closed subset of* $[a, b]$. *Choose an acceptable partition* U_k *of* X *into* $k > n$ *subintervals. In order that* $P = \sum_{i=1}^n a_i \varphi_i$ *shall be a best* $ps\|\cdot\|_{U_k}$ *approximation to a given* $f \in C(X)$, *it is necessary and sufficient that the function* $s(\cdot, r)$ *exhibit at least* $n + 1$ "*alternations*" *where* $r = f - P$. *That is*

$$s(E_{i_\nu}, r) = -s(E_{i_{\nu-1}}, r) = \pm ps\|r\|_{U_k}, \qquad \nu = 1, 2, \cdots, n, \tag{5.6}$$

with

$$E_{i_0} < E_{i_1} < \cdots < E_{i_n} \quad \text{and} \quad E_{i_\nu} \in U_k, \qquad \nu = 0, 1, \cdots, n. \tag{5.7}$$

Proof. The proof is exactly the same as the proof of the alternation theorem on p. 75 of [1]. In particular, it relies on the Characterization Theorem 4.5 and Lemma 5.4. We shall omit the details.

6. **Conclusions.** The norms of Sections 2 and 3 and the pseudo norms of Sections 4 and 5 provide good approximations to the uniform norm. The pseudo norms are particularly advantageous for the computation of best approximations through the use of either linear programming or an exchange algorithm based on the Alternation Theorem of Section 5. Theorems 2.13, 2.15 and their analogues for best pseudo norm approximation suggest that through proper modifications of the partitions, the approximations defined in this paper are arbitrarily close to best uniform approxima-

tions. Algorithms to achieve this are discussed in [5]. These procedures are of particular interest when X is multidimensional.

Department of Mathematics
The University of Utah
Salt Lake City, Utah 84112

1. E. W. CHENEY, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966. MR **36** #5568.
2. R. H. HARRIS, *Uniform Approximation of Functions: Approximation by Partitioning*, Thesis, University of Utah, Salt Lake City, Utah, 1970.
3. S. E. WEINSTEIN, "Approximations of functions of several variables: Product Chebychev approximations. I," *J. Approximation Theory*, v. 2, 1969, pp. 433–447. MR **40** #7683.
4. S. E. WEINSTEIN, "Uniform approximation of functions through optimal partitioning," *SIAM J. Numer. Anal.* (To appear.)
5. S. E. WEINSTEIN, "Computation of best uniform approximations through partitioning." (In preparation.)