JOURNAL OF APPROXIMATION THEORY 32, 306–315 (1981)

Weighted Simultaneous Chebyshev Approximation

YING-GUANG SHI

Computing Center, Chinese Academy of Sciences, Peking, People's Republic of China

Communicated by E. W. Cheney

Received March 20, 1980

In this paper we discuss the problem of weighted simultaneous Chebyshev approximation to functions $f_1,..., f_m \in C(X)$ $(1 \leqslant m \leqslant \infty)$, i.e., we wish to minimize the expression $\|\{\sum_{j=1}^m \lambda_j |f_j - q|^p\}^{1/p}\|_\infty$, where $\lambda_j > 0$, $\sum_{j=1}^m \lambda_j = 1$, $p \geqslant 1$. For this problem we establish the main theorems of the Chebyshev theory, which include the theorems of existence, alternation, de La Vallée Poussin, uniqueness, strong uniqueness, as well as that of continuity of the best approximation operator, etc.

1. INTRODUCTION

Let X be a compact subset of an interval $[a, b]$ containing at least $n + 1$ points, where n is a fixed natural number. In the space $C(X)$ of all real-valued continuous functions f, we define the uniform norm $\|f\| = \max_{x \in X} |f(x)|$. Let M be an n-dimensional Haar subspace of $C[a, b]$, and $h_1^*,..., h_n^*$ a basis for M.

Let $F \equiv (f_1,..., f_m)$, where $m \leqslant \infty$ and $f_j \in C(X)$. Suppose the corresponding weight of f_j is $\lambda_j > 0$, satisfying $\sum \lambda_j = 1$, where $\sum$ denotes the summation for $j = 1,..., m$. In case $m = \infty$, we shall further assume that F satisfies the condition: the series $\sum \lambda_j |f_j(x)|^p$ is uniformly convergent, or equivalently, is continuous. The set of all such F is denoted by $\mathscr{F}$. We are concerned about approximating F simultaneously with weights by an element $q \in M$ in the sense of minimization of the expression $\|E_q\| = \|\{\sum \lambda_j |f_j - q|^p\}^{1/p}\|$ $(1 \leqslant p < \infty)$.

If there exists an element $h \in M$ which satisfies

$$\|E_h\| = \inf_{q \in M} \|E_q\| \tag{1}$$

then h is said to be a weighted simultaneous best approximation to F, or simply a best approximation to F.

W. H. Ling [1] investigated the case when $p = 1$, $m = 2$, $\lambda_1 = \lambda_2$ and showed that approximating in the "sum" norm is, with one restriction,

0021-9045/81/080306–10$02.00/0

equivalent to approximating the arithmetic mean. But it is rather difficult for the cases $m \geqslant 3$ to get similar results by means of his method. In the present paper, we use another method, i.e., generalization of the definition of an alternation system (Cf. Section 2), which enables us to obtain a number of interesting results for the general case stated above, and indeed to establish the main theorems of the Chebyshev theory, including the theorems of existence, alternation, de La Vallée Poussin, uniqueness, strong uniqueness, as well as the theorem of continuity of the best approximation operator.

2. NOTATION

Letting d be any real number, we still write $E_d(x) = \{\sum \lambda_j |f_j(x) - d|^p\}^{1/p}$. Now let $x \in X$ be fixed. Put $E^*(x) = \inf_d E_d(x)$ and $D_x = \{d: E_d(x) = E^*(x)\}$. Since $E_d(x)$ is a continuous function of d [2, p. 26] and $E_d(x) \to \infty$ if $|d| \to \infty$, the infimum of $E_d(x)$ is attained by some point d. Hence D_x is a non-empty closed set. We put $f^+(x) = \max D_x$ and call it an upper critical function, and likewise put $f^-(x) = \min D_x$, a lower critical function.

Let $F \in \mathscr{F}$ be fixed and $h \in M$ be arbitrary. Define

$$X_h = \{x \in X: E_h(x) = \|E_h\|\},$$

$$X_+ = \{x \in X_h: f^-(x) > h(x)\}, \qquad X_- = \{x \in X_h: f^+(x) < h(x)\},$$

$$\bar{X}_+ = \{x \in X_h: f^-(x) \geqslant h(x)\}, \qquad \bar{X}_- = \{x \in X_h: f^+(x) \leqslant h(x)\},$$

$$X_0 = \{x \in X_h: f^-(x) \leqslant h(x) \leqslant f^+(x)\}.$$

$$\begin{aligned} \sigma(x) &= 1, \qquad f^-(x) \geqslant h(x), \\ &= -1, \qquad f^+(x) \leqslant h(x). \end{aligned}$$

A system of $n+1$ ordered points

$$x_1 < x_2 < \cdots < x_{n+1} \tag{2}$$

in $X_+ \cup X_- [\bar{X}_+ \cup \bar{X}_-]$ is said to be an alternation system [a generalized alternation system] of h to F if it satisfies $\sigma(x_{i+1}) = -\sigma(x_i)$ $(i = 1, 2,..., n)$.

We should point out that the previous notation simplifies in case $p > 1$ (since $f^+ = f^-$, see Remark 2 below).

3. EXISTENCE AND CHARACTERIZATION

By the usual compactness arguments the existence of best approximations may directly be obtained.

THEOREM 1 (Existence). *To each $F \in \mathscr{F}$ there exists at least one best approximation from M.*

In what follows, we shall first show several lemmas for the preliminaries. However, we need only to mention that using the inequality of Minkowski [4, p. 123], it is easy to prove the following lemma.

LEMMA 1. *E_q is a convex function of q.*

LEMMA 2. (a) *For each $x \in X$, $E_d(x)$ is a monotone increasing [decreasing] function of d on interval $[f^+(x), \infty)[(-\infty, f^-(x)]]$ and satisfies $E_d(x) = E^*(x)$ on the interval $[f^-(x), f^+(x)]$;*

(b) *$f^+(x)$ and $f^-(x)$ are respectively upper and lower semicontinuous:*

$$\overline{\lim_{y \to x}} f^+(y) \leqslant f^+(x) \qquad \text{and} \qquad \underline{\lim_{y \to x}} f^-(y) \geqslant f^-(x).$$

Proof. (a) From Lemma 1 and the definitions of f^+ and f^-, it follows immediately.

(b) It is easy to see that $E_d(x)$ is a continuous function of d and x. Now let $y_k \to x$ and $d_k = f^+(y_k) \to d$. Thus, $E_{d_k}(y_k) = E^*(y_k) \leqslant E_c(y_k), \forall c$. By continuity, we have $E_d(x) \leqslant E_c(x), \forall c$. Hence $E_d(x) = E^*(x)$ and $d \leqslant f^+(x)$. By the arbitrariness of d_k, we have $\overline{\lim}_{y \to x} f^+(y) \leqslant f^+(x)$. Similarly, it may be shown that $\underline{\lim}_{y \to x} f^-(y) \geqslant f^-(x)$.

Remark 1. As $p = 1$ we may assume another equivalent definition

$$D_x = \left\{d: \sum_{f_j(x) \leqslant d} \lambda_j \geqslant \tfrac{1}{2}\right\} \cap \left\{d: \sum_{f_j(x) \geqslant d} \lambda_j \geqslant \tfrac{1}{2}\right\}.$$

Remark 2. As $p > 1$ we have $f^+ = f^- \in C(X)$. Actually, it is sufficient to prove that for each x, D_x is a set of a single element. Let $d_1, d_2 \in D_x$, namely, $E_{d_1}(x) = E_{d_2}(x) = E^*(x)$. Then $E_{(1/2)(d_1+d_2)}(x) = \frac{1}{2}E_{d_1}(x) + \frac{1}{2}E_{d_2}(x)$. According to the condition under which the equal sign holds in the inequality of Minkowski, there exists a number λ for which $f_j(x) - d_1 = \lambda[f_j(x) - d_2], \forall j$. It follows that either $f_j(x) = \text{const.}, \forall j$ or $\lambda = 1$. From either of them we obtain that $d_1 = d_2$.

LEMMA 3. *If $s \in M$ satisfies*

$$\mu_i \equiv (-1)^i s(x_i) \geqslant 0 \ (\leqslant 0), \qquad i = 1, 2, \ldots, n+1, \tag{3}$$

where $x_1 < x_2 < \cdots < x_{n+1}$, then $s = 0$.

The proof of this lemma may be completed using a well-known method, and therefore is omitted.

LEMMA 4. *If there exists a generalized alternation system of $h \in M$ to F, then h is the unique best approximation to F.*

Proof. Let (2) be such a system and let $q \in M$ satisfy $\|E_q\| \leqslant \|E_h\|$. If $x_i \in \bar{X}_+$, then $f^-(x_i) \geqslant h(x_i)$; we shall have $q(x_i) \geqslant h(x_i)$, for otherwise, by part (a) of Lemma 2, from $q(x_i) < h(x_i)$ it would follow $E_q(x_i) > E_h(x_i) = \|E_h\|$. Similarly, we can show that if $x_i \in \bar{X}_-$, then $q(x_i) \leqslant h(x_i)$. Thus $s \equiv q - h$ satisfies (3) at the points of the generalized alternation system (2). By Lemma 3, $s = 0$, viz., $q = h$. This proves that h is the unique best approximation to F.

THEOREM 2. *Let $h \in M$. If $X_0 \neq \varnothing$, then h is a best approximation to F, and $\|E_h\| = e \equiv \|E^*\|$.*

Proof. Let $\xi \in X_0$. Then $\|E_h\| = E_h(\xi) \leqslant e$. But for any $q \in M$ we have $E_q(x) \geqslant E^*(x)$. Hence $\|E_q\| \geqslant e$. Thus $\|E_h\| \leqslant \|E_q\|$ and $\|E_h\| = e$.

THEOREM 3 (Characterization). *Let $h \in M$, and suppose $X_0 = \varnothing$. Then the following three statements are equivalent:*

(a) *h is a best approximation to F;*

(b) *$0 \in \mathscr{H}\{\sigma(x)\hat{x} : x \in X_h\}$, where $\mathscr{H}$ denotes the convex hull and $\hat{x} = (h_1^*(x),..., h_n^*(x))$;*

(c) *there exists an alternation system.*

Proof. (a) $\Rightarrow$ (b). It can be proved by contradiction. Suppose $0 \bar{\in} \mathscr{H}\{\sigma(x)\hat{x} : x \in X_h\}$. Since X_h is clearly compact, we have by the theorem on linear inequalities [2, p. 19] that there exists $q \in M$ such that $\sigma(x)\, q(x) > 0$, $\forall x \in X_h$. We are to prove that there exists $\lambda > 0$ for which $r = h + \lambda q$ satisfies $\|E_r\| < \|E_h\|$.

First, let $\xi \in X_+$. Then $f^-(\xi) > h(\xi)$, $\sigma(\xi) > 0$, $q(\xi) > 0$. By the lower semicontinuity of f^- and the continuity of h and q, $f^-(x) - h(x) \geqslant \varepsilon > 0$ and $q(x) > 0$ are valid in some neighborhood Δ_ξ of ξ. Taking $\lambda_\xi = \varepsilon/\|q\|$, when $0 < \lambda \leqslant \lambda_\xi$, and $x \in \Delta_\xi$ we have

$$f^-(x) - r(x) = f^-(x) - h(x) - \lambda q(x) \geqslant \varepsilon - \lambda_\xi \|q\| = 0.$$

By part (a) of Lemma 2, it follows that

$$E_r(x) < \|E_h\|, \qquad 0 < \lambda \leqslant \lambda_\xi,\ x \in \Delta_\xi. \tag{4}$$

Repeating this argument for $\xi \in X_-$, we can get some neighborhood of ξ and some positive number λ_ξ such that (4) is valid.

Next, let $\xi \bar{\in} X_h$. Then $E_h(\xi) < \|E_h\|$. It is easy to see that $E_r(x)$ is a continuous function of x and λ. Hence there exists a neighborhood Δ_ξ and a positive number λ_ξ such that (4) is valid also.

Since $X_0 = \varnothing$, $\bigcup_{\xi \in X} \Delta_\xi \supset X$. By the finite covering theorem there exists a finite number of neighborhoods $\Delta_\xi, \ldots, \Delta_\eta$ covering X. Taking the minimum of the corresponding positive numbers $\lambda_\xi, \ldots, \lambda_\eta$, denoted by λ, then for such a λ and all $x \in X$ we have $E_r(x) < \|E_h\|$. Hence $\|E_r\| < \|E_h\|$. This contradiction proves the implication (a) $\Rightarrow$ (b).

(b) $\Rightarrow$ (c). This is a standard statement and, for example, it may be found in [2, p. 75].

(c) $\Rightarrow$ (a). It can be shown by Lemma 4.

From Theorem 2 and Theorem 3 we obtain the following complete characterization of best approximations.

THEOREM 4 (Alternation). *$h \in M$ is a best approximation to F if and only if it satisfies one of the two conditions*:

(a) $X_0 \neq \varnothing$;

(b) *There exists an alternation system.*

4. Uniqueness and Other Theorems

THEOREM 5 (de La Vallée Poussin). *If $h \in M$ satisfies $\sigma(x_{i+1}) = -\sigma(x_i)$ $(i = 1, \ldots, n)$ at $n+1$ ordered points* (2), *then*

$$E(F) = \inf_{r \in M} \|E_r\| \geqslant \min_i E_h(x_i).$$

Proof. Suppose on the contrary $E(F) < \min_i E_h(x_i) \leqslant E_h(x_i)$. Let $q \in M$ satisfy $\|E_q\| = E(F)$. By a similar argument made in the proof of Lemma 4 it still follows that $q = h$. But this is not possible.

THEOREM 6 (Uniqueness). *Let $h \in M$ be a best approximation to F. Then h is the unique best approximation to F, if h satisfies one of the following conditions*:

(a) $X_0 = \varnothing$;

(b) *there exists a generalized alternation system*;

(c) $\|E_h\| > e$.

Proof. (a) Since $X_0 = \varnothing$ and h is a best approximation, by Theorem 3, there exists an alternation system. By Lemma 4, h is the unique best approximation to F.

(b) It immediately follows from Lemma 4.

(c) $\|E_h\| > e$ implies $X_0 = \varnothing$, for if $\xi \in X_0$, then $\|E_h\| = E_h(\xi) \leqslant e$ (see the proof of Theorem 2).

THEOREM 7 (Uniqueness). *Let* $h \in M$. *If* $\bar{X}_+ \cap \bar{X}_- = \varnothing$ *and* f^+, $f^- \in C(X)$, *then* h *is the unique best approximation to* F *if and only if there exists a generalized alternation system.*

Proof. The sufficiency follows from Lemma 4. We proceed to the necessity part. Divide the interval $[a, b]$ by the points

$$a = y_0 < y_1 < y_2 < \cdots < y_{\mu-1} < y_\mu = b$$

so that the following two conditions are alternately satisfied on each subinterval $\Delta_i \equiv [y_{i-1}, y_i]$ $(i = 1,..., \mu)$: (1) $\bar{X}_+ \cap \Delta_i \neq \varnothing$, $\bar{X}_- \cap \Delta_i = \varnothing$ and (2) $\bar{X}_- \cap \Delta_i \neq \varnothing$, $\bar{X}_+ \cap \Delta_i = \varnothing$. Thus we need only to prove $\mu > n$, and shall show it by contradiction. Now suppose $\mu \leqslant n$. By the Haar condition there exists $q \in M$ such that $q(x)$ becomes zero and changes signs at the points $y_1,..., y_{\mu-1}$, meanwhile, $q(x) \neq 0$, $\forall x \in (a, b) \backslash \{y_1,..., y_{\mu-1}\}$ [3, p. 30]. Hence by appropriately selecting the sign of q, we have $\sigma(x)\, q(x) > 0$, $\forall x \in \bar{X}_+ \cup \bar{X}_-$.

When $\xi \bar{\in} X_0$, from the proof of (a) $\Rightarrow$ (b) in Theorem 3, it has been shown that there exists some neighborhood Δ_ξ of ξ and some positive number λ_ξ such that (4) is valid, where $r = h + \lambda q$.

When $\xi \in X_0 \cap (\bar{X}_+ \cup \bar{X}_-)$, noting $\bar{X}_+ \cap \bar{X}_- = \varnothing$ and writing $\bar{f} = (f^+ + f^-)/2$, we have $\sigma(\xi)[\bar{f}(\xi) - h(\xi)] > 0$, $\sigma(\xi)\, q(\xi) > 0$. According to continuity there exists some neighborhood Δ_ξ of ξ such that $\sigma(\xi)[\bar{f}(x) - h(x)] \geqslant \varepsilon > 0$ and $\sigma(\xi)\, q(x) > 0$ for all $x \in \Delta_\xi$. Taking $\lambda_\xi = \varepsilon/\|q\|$, we have that $0 < \lambda \leqslant \lambda_\xi$ implies $\sigma(\xi)[\bar{f}(x) - h(x)] \geqslant \sigma(\xi)[\bar{f}(x) - r(x)] \geqslant 0$. By Lemma 2, we get

$$E_r(x) \leqslant \|E_h\|, \qquad 0 < \lambda \leqslant \lambda_\xi, x \in \Delta_\xi. \tag{5}$$

When $\xi \in X_0 \backslash (\bar{X}_+ \cup \bar{X}_-)$, we have $\bar{f}(\xi) < h(\xi) < f^+(\xi)$. Hence there exists some neighborhood Δ_ξ of ξ and some positive number λ_ξ such that $0 < \lambda \leqslant \lambda_\xi$ and $x \in \Delta_\xi$ imply $f^-(x) < r(x) < f^+(x)$. Whence (5) is valid also.

Applying the finite covering theorem, we find $\lambda > 0$ such that $\|E_r\| \leqslant \|E_h\|$. We note that $r = h + \lambda q \neq h$. Thus r is another best approximation to F, and this is a contradiction. Consequently we conclude $\mu > n$.

THEOREM 8 (Strong Uniqueness). *Let* $h \in M$ *be a best approximation to* F. *If* $X_0 = \varnothing$, *then there exists a constant* $v > 0$ *depending on* F *such that for any* $q \in M$,

$$\|E_q\|^p \geqslant \|E_h\|^p + v\,\|h - q\|^p. \tag{6}$$

Proof. If $\|E_h\| = 0$, then $f_1 = \cdots = f_m = h$ and $\nu = 1$. Let us assume therefore that $\|E_h\| > 0$. By Theorem 3, $0 \in \mathscr{H}\{\sigma(x)\hat{x} : x \in X_h\}$, viz., there exist points $x_0, ..., x_k \in X_h$ such that

$$0 = \sum_{i=0}^{k} \theta_i \sigma(x_i)\, h_j^*(x_i), \qquad \forall \theta_i > 0, j = 1, ..., n.$$

By Carathéodory's Theorem [2, p. 17] and the Haar condition, $k = n$. Let $r \in M$ and $\|r\| = 1$. We have $\sum_{i=0}^{n} \theta_i \sigma(x_i)\, r(x_i) = 0$. By the Haar condition, the numbers $\sigma(x_i)\, r(x_i)$ are not all zero. Since $\theta_i > 0$, we infer that at least one of $\sigma(x_i)\, r(x_i)$ is positive. Consequently $\max_i \sigma(x_i)\, r(x_i) > 0$. Hence $\nu^* = \min_{\|r\|=1} \max_i \sigma(x_i)\, r(x_i) > 0$, for it is the minimum of a positive continuous function on a compact set. Now let $q \in M$. If $q = h$, the conclusion of the theorem is trivial. Otherwise, $r = (h - q)/\|h - q\|$ is of unit norm. Consequently, there exists an index i for which $\sigma(x_i)\, r(x_i) \geqslant \nu^*$.

On the other hand, if $\sigma(x_i) > 0$, $f^-(x_i) > h(x_i)$. And there clearly exists an index j_i such that $f_{j_i}(x_i) \geqslant f^-(x_i)$. Take $0 < \lambda_{j_i}^* < \lambda_{j_i}$ such that the condition $\hat{f}^-(x_i) \geqslant h(x_i)$ is satisfied by the lower critical function $\hat{f}^-$, corresponding to $f_1, ..., f_m$ with $\hat{\lambda}_1, ..., \hat{\lambda}_m$, where

$$\begin{aligned} \hat{\lambda}_j &= \lambda_j/(1 - \lambda_{j_i}^*), \qquad & j \neq j_i, \\ &= (\lambda_{j_i} - \lambda_{j_i}^*)/(1 - \lambda_{j_i}^*), \qquad & j = j_i. \end{aligned}$$

This is indeed possible. In order to do this, it is sufficient to prove $\underline{\lim}_{\lambda_{j_i}^* \to 0} \hat{f}^-(x_i) \geqslant f^-(x_i)$. By the definition of $\hat{f}^-(x_i)$, we have that

$$\sum \hat{\lambda}_j |f_j(x_i) - \hat{f}^-(x_i)|^p \leqslant \sum \hat{\lambda}_j |f_j(x_i) - c|^p, \qquad \forall c$$

or

$$\begin{aligned} &\sum \lambda_j |f_j(x_i) - \hat{f}^-(x_i)|^p - \lambda_{j_i}^* |f_{j_i}(x_i) - \hat{f}^-(x_i)|^p \\ &\quad \leqslant \sum \lambda_j |f_j(x_i) - c|^p - \lambda_{j_i}^* |f_{j_i}(x_i) - c|^p, \qquad \forall c. \end{aligned}$$

Let $\hat{d}$ be any cluster point of $\hat{f}^-(x_i)$ as $\lambda_{j_i}^* \to 0$. Then by the continuity of $E_d(x)$ from the preceding inequality it follows that

$$\sum \lambda_j |f_j(x_i) - \hat{d}|^p \leqslant \sum \lambda_j |f_j(x_i) - c|^p, \qquad \forall c.$$

Namely, $\hat{d} \geqslant f^-(x_i)$ and $\underline{\lim}_{\lambda_{j_i}^* \to 0} \hat{f}^-(x_i) \geqslant f^-(x_i)$. Using part (a) of Lemma 2 from $\hat{f}^-(x_i) \geqslant h(x_i) > q(x_i)$ (for $\sigma(x_i)\, r(x_i) > 0$) it follows that

$$\sum \hat{\lambda}_j |f_j(x_i) - h(x_i)|^p \leqslant \sum \hat{\lambda}_j |f_j(x_i) - q(x_i)|^p$$

or

$$\sum \lambda_j |f_j(x_i) - h(x_i)|^p - \lambda_{j_i}^* |f_{j_i}(x_i) - h(x_i)|^p$$
$$\leqslant \sum \lambda_j |f_j(x_i) - q(x_i)|^p - \lambda_{j_i}^* |f_{j_i}(x_i) - q(x_i)|^p.$$

Similarly, if $\sigma(x_i) < 0$, there exists an index j_i and a positive number $\lambda_{j_i}^*$ such that $0 < \lambda_{j_i}^* < \lambda_{j_i}$ implies that the preceding inequality is valid. Taking $\lambda^* = \min_i \{\lambda_{j_i}^*\} > 0$, we have (here using the inequality $(|u| + |v|)^p \geqslant |u|^p + |v|^p$ [2, p. 11, Prob. 8]) that

$$\begin{aligned} \|E_q\|^p \geqslant E_q^p(x_i) &= \sum \lambda_j |f_j(x_i) - q(x_i)|^p - \lambda_{j_i}^* f_{j_i}(x_i) - q(x_i)|^p \\ &\quad + \lambda_{j_i}^* |f_{j_i}(x_i) - q(x_i)|^p \\ &\geqslant \sum \lambda_j |f_j(x_i) - h(x_i)|^p - \lambda_{j_i}^* f_{j_i}(x_i) - h(x_i)|^p \\ &\quad + \lambda_{j_i}^* |f_{j_i}(x_i) - q(x_i)|^p \\ &\geqslant E_h^p(x_i) + \lambda_{j_i}^* |h(x_i) - q(x_i)|^p \\ &\geqslant \|E_h\|^p + \lambda^* (v^*)^p \|h - q\|^p = \|E_h\|^p + v \|h - q\|^p, \end{aligned}$$

where $v = \lambda^* (v^*)^p$ depends only on F. The Theorem is proved.

Now, for $G \equiv (g_1, ..., g_m) \in \mathscr{F}$, let $\mathscr{E}G$ be any best approximation to G from M. Then we have the following theorem.

THEOREM 9 (Continuity). *If $X_0 = \varnothing$ for $\mathscr{E}F$ and F, then to F there corresponds a number $\lambda > 0$ such that for all $G \in \mathscr{F}$,*

$$\|\mathscr{E}F - \mathscr{E}G\|^p \leqslant \lambda [E(F) + 2\rho(F, G)]^{p-1} \rho(F, G), \tag{7}$$

where $\rho(F, G) = \|\{\sum \lambda_j |f_j - g_j|^p\}^{1/p}\|$.

Proof. Let $h = \mathscr{E}F$ and $q = \mathscr{E}G$ in (6) and the inequality $u^p - v^p \leqslant pu^{p-1}(u - v)$ with $u \geqslant v \geqslant 0$. We have

$$v \|\mathscr{E}F - \mathscr{E}G\|^p \leqslant \|E_{\mathscr{E}G}\|^p - \|E_{\mathscr{E}F}\|^p \leqslant p \|E_{\mathscr{E}G}\|^{p-1} (\|E_{\mathscr{E}G}\| - \|E_{\mathscr{E}F}\|). \tag{8}$$

But

$$\begin{aligned} \|E_{\mathscr{E}G}\| &\leqslant \rho(F, G) + \left\| \left\{ \sum \lambda_j |g_j - \mathscr{E}G|^p \right\}^{1/p} \right\| \\ &\leqslant \rho(F, G) + \left\| \left\{ \sum \lambda_j |g_j - \mathscr{E}F|^p \right\}^{1/p} \right\| \\ &\leqslant 2\rho(F, G) + \|E_{\mathscr{E}F}\|. \end{aligned}$$

Inserting this in (8) and noting $E_{\mathscr{E}F} = E(F)$, we let $\lambda = 2p/v$ and obtain (7).

THEOREM 10. *Let $m=\infty$. Let h_N be a best approximation to $(f_1,\dots,f_N)$ with weights $\lambda_1,\dots,\lambda_N$ from M, and h be one to F from M. If $X_0=\varnothing$ for h and F, then*

$$\|h_N-h\|\to 0, \qquad N\to\infty.$$

Proof. Evidently

$$\sum_{j=1}^{N}\lambda_j|f_j-h_N|^p=\sum_{j=1}^{N}\lambda_j|f_j-h_N|^p+\sum_{j=N+1}^{\infty}\lambda_j|h_N-h_N|^p,$$

so that h_N is also a best approximation to $F_N=(f_1,\dots,f_N,h_N,h_N,\dots)$ with $\lambda_1,\dots,\lambda_N,\lambda_{N+1},\lambda_{N+2},\dots$. Then by Theorem 9, we have that

$$\|h_N-h\|^p\leqslant\lambda[E(F)+2\rho(F,F_N)]^{p-1}\rho(F,F_N), \tag{9}$$

where

$$\rho(F,F_N)=\left\|\left\{\sum_{j=N+1}^{\infty}\lambda_j|f_j-h_N|^p\right\}^{1/p}\right\|$$
$$\leqslant\left\|\left\{\sum_{j=N+1}^{\infty}\lambda_j|f_j|^p\right\}^{1/p}\right\|+\left(\sum_{j=N+1}^{\infty}\lambda_j\right)^{1/p}\|h_N\|.$$

Now, let us estimate $\|h_N\|$. By the definition of h_N, we get

$$\left\|\left\{\sum_{j=1}^{N}\lambda_j|f_j-h_N|^p\right\}^{1/p}\right\|\leqslant\left\|\left\{\sum_{j=1}^{N}\lambda_j|f_j|^p\right\}^{1/p}\right\|,$$

whence

$$\left\|\left\{\sum_{j=1}^{N}\lambda_j|h_N|^p\right\}^{1/p}\right\|\leqslant 2\left\|\left\{\sum_{j=1}^{N}\lambda_j|f_j|^p\right\}^{1/p}\right\|.$$

Thus

$$\|h_N\|\leqslant\left[2\Big/\left(\sum_{j=1}^{N}\lambda_j\right)^{1/p}\right]\left\|\left\{\sum_{j=1}^{N}\lambda_j|f_j|^p\right\}^{1/p}\right\|.$$

Consequently

$$\rho(F,F_N)\leqslant\left\|\left\{\sum_{j=N+1}^{\infty}\lambda_j|f_j|^p\right\}^{1/p}\right\|$$
$$+2\left[\left(\sum_{j=N+1}^{\infty}\lambda_j\right)\Big/\left(\sum_{j=1}^{N}\lambda_j\right)\right]^{1/p}\left\|\left\{\sum_{j=1}^{N}\lambda_j|f_j|^p\right\}^{1/p}\right\|\to 0, \qquad N\to\infty.$$

From (9) it follows that $\|h_N-h\|\to 0$, $N\to\infty$.

ACKNOWLEDGMENT

I am indebted to Professor Hsü Hsien-yü of Peking University for his enthusiastic guidance and help in this work.

REFERENCES

1. W. H. LING, On simultaneous Chebyshev approximation in the 'sum' norm, *Proc. Amer. Math. Soc.* **48**: 1(1975), 185–188.
2. E. W. CHENEY, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
3. S. KARLIN AND W. J. STUDDEN, "Tchebycheff Systems: With Applications in Analysis and Statistics," Interscience, New York, 1966.
4. G. H. HARDY, J. E. LITTLEWOOD, AND G. PÓLYA, "Inequalities," Cambridge Univ. Press, London/New York, 1952.