THE UNIFORM APPROXIMATION OF A FUNCTION AND ITS DERIVATIVES BY POLYNOMIALS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

FREDERICK JAMES SCHUURMANN

1967

This is to certify that the

thesis entitled

"The Uniform Approximation of a Function

and its Derivatives by Polynomials"

presented by

Frederick James Schuurmann

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

David Moursund

D. Moursund
Major professor

Date July 10, 1967

O-169

ABSTRACT

THE UNIFORM APPROXIMATION OF A FUNCTION AND
ITS DERIVATIVES BY POLYNOMIALS

by Frederick James Schuurmann

Let X be a compact subset of the real line, and let I be a finite collection of nonnegative integers including 0. The function $f(x)$ and the base functions $\{\phi_i(x)\}_{i=1}^{n}$ are assumed to be in $C^q(X)$ while the weight functions $w_k(x)$ for $k \varepsilon I$ are continuous on X. The problem is to find real scalars $a_1, a_2, \ldots, a_n$ which minimize

$$\max_{(x,k)\varepsilon X\times I} \left| w_k(x) \frac{d^k}{dx^k} \left[f(x) - \sum_{i=1}^{n} a_i \phi_i(x) \right] \right| .$$

The main objective is to provide an efficient method for the computation of a best approximation on a digital computer when X is a closed interval.

First the problem of the existence of a best approximation is discussed. Then the characterization of a best approximation is treated to provide a basis for the computational algorithm as well as to provide a method of determining when an approximation is a best approximation. Results are also obtained concerning the dimension of the space of best approximations.

Next, sufficient conditions for the uniform convergence of the polynomial and its derivatives to the function and its derivatives as the number of base functions increases are given.

It was also found that with the appropriate hypotheses the approximation of a function and its first derivative on an interval by certain classes of base functions is unique.

Finally, two algorithms which use linear programming to find best approximations on finite point sets are presented and convergence theorems given. Several computational examples are also presented.

THE UNIFORM APPROXIMATION OF A FUNCTION AND ITS DERIVATIVES BY POLYNOMIALS

By

Frederick James Schuurmann

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1967

5 47130
12-20-67

ACKNOWLEDGMENTS

The author is deeply indebted to his major professor Dr. David G. Moursund for suggesting this thesis problem and for his patient and helpful guidance during the preparation of this thesis.

The author would also like to thank Michigan State University for making available the computing facilities used in the development of the algorithms and the calculation of the examples in this paper.

TABLE OF CONTENTS

		Page
ACKNOWLEDGMENTS		ii
INTRODUCTION		1
Chapter		
I.	EXISTENCE AND CHARACTERIZATION	3
	1. Introduction	3
	2. Existence	9
	3. Characterization	13
II.	CONVERGENCE CONSIDERATIONS	30
	1. Uniform Convergence of the Polynomial to the Function	30
	2. Continuous Dependence of the Approximating Polynomial on the Function	32
	3. The de la Vallée Poussin Algorithm	38
III.	UNIQUENESS RESULTS	44
	1. An Example of Nonuniqueness	44
	2. A General Uniqueness Theorem	47
	3. The Trigonometric Polynomials	57
	4. A Special Class of Polynomials	58
IV.	COMPUTATIONAL ALGORITHMS FOR A BEST APPROXIMATION	62
	1. Introduction	62
	2. The Continuity of e(T)	65
	3. Two Algorithms	72
	4. Convergence of the Algorithms	76
	5. Computational Procedures	83
	6. Computational Examples	86
BIBLIOGRAPHY		89
INDEX OF NOTATIONS		92

INTRODUCTION

Many advances have been made in the study of approximation theory since the rise of the electronic computer. The computer has stimulated the study of approximation theory because approximations are necessary in the efficient handling of many problems, and the computer provides the means whereby approximations may be computed. Many problems which once were dismissed as impractical because of the difficulties of computation are now solvable using an electronic computer.

A recent bibliography of approximation theory, a survey of recent Russian literature in approximation theory, and a number of papers on approximation theory are given in Garabedian [10]. Another earlier survey of literature in approximation theory is given by Buck [3,4]. Recently, several new books have been published which deal exclusively with approximation theory. Among these are Rice [23] and Cheney [5]. Both of these books contain extensive bibliographies.

The problem considered in this paper is that of approximating a given function f by a polynomial in such a way as to make one or more of the derivatives of the approximating polynomial approximate the corresponding derivatives of the function f in some prescribed manner.

In the first chapter, the existence and characterization of best approximations are treated extensively. Next the uniform convergence of best approximations to f and the continuous dependence of best approximations on f are studied. A theoretical algorithm for the computation of a best approximation is also presented. Then several classes of problems are described which have unique best approximations. Finally the computational problems of finding a best approximation are considered. Two algorithms are given and convergence theorems presented. Several numerical examples are also given.

CHAPTER I

EXISTENCE AND CHARACTERIZATION

1. Introduction

Let X be a compact subset of the real line, and let I be an finite collection of nonnegative integers including 0. The function f(x) and the base functions $\phi_1(x),\phi_2(x),\ldots,\phi_n(x)$ for fixed $n \geq 1$, are assumed to be in $C^q(X)$ where q is the largest integer in I. For each $k\epsilon I$ let $w_k(x)$ be a continuous weight function on X. (We show in Lemma 1.5 that there is no loss of generality in assuming that the weight functions are nonnegative.) The following standard notation will be used:

$$D \equiv \frac{d}{dx}\ , \quad D^k \equiv \frac{d^k}{dx^k}\ , \quad Df(x_o) \equiv \left.\frac{df(x)}{dx}\right|_{x = x_o} .$$

1.1 Definition. Let T be a closed subset of X×I and let $g\epsilon C^q(X)$. Then define $MT[g(x)] = \max_{(x,k)\epsilon T} |w_k(x)D^k g(x)|$. If $T = X \times I$ we will write $M[g(x)]$ for $MT[g(x)]$. (A set $T \subset X \times I$ is called closed if the sets $V_k = \{x:(x,k)\epsilon T\}$ for $k\epsilon I$ are all closed subsets of X.)

1.2 Problem. Find real scalars $a_1,a_2,\ldots,a_n$ such that

$$M[f(x) - \sum_{i=1}^{n} a_i\phi_i(x)]$$

is a minimum. A solution to this problem is called a best approximation to f(x) on X×I with weight functions $\{w_k(x)\}$.

As a notational convenience, points in E^n are represented by $\alpha = (a_1,a_2,\ldots,a_n)$, $\beta = (b_1,b_2,\ldots,b_n)$, etc., while polynomials are represented by

$$P(x,\alpha) = \sum_{i=1}^{n} a_i\phi_i(x).$$

Also, for any closed set $T \subset X\times I$, let

$$e(T) = \min_{\alpha\in E^n} MT[f(x) - P(x,\alpha)].$$

It is shown in Theorem 1.16 that this minimum exists. e(T) will be called the deviation of a best approximation to f(x) on the set T. We define

$$R(T) = \{\alpha\in E^n : MT[f(x) - P(x,\alpha)] = e(T)\}$$

to be the set of all best approximations to f(x) on T. When T is X×I we will write R for R(T) and e for e(T). The norm to be used on the polynomial coefficients which are points of E^n is

$$||\alpha|| = \max_{i=1,2,\ldots,n} |a_i| \quad .$$

Throughout this paper the symbol T, with or without a subscript or superscript, will denote a compact subset of $X \times I$.

1.3 Theorem. $MT[g]$ is defined for all $g \in C^q(X)$ and has the properties:

(a) $0 \leq MT[g] < \infty$

(b) $MT[g] = 0$ if $g \equiv 0$

(c) $MT[tg] = |t|MT[g]$ for any real scalar t

(d) $MT[g + h] \leq MT[g] + MT[h]$.

The first three properties are obvious. The fourth can be proved using the triangle inequality for real numbers. Since we allow the weight functions to be zero, we cannot prove the converse of (b); hence $MT[g]$ is a pseudonorm.

1.4 Corollary. The set R(T) of best approximations is convex.

Proof: If $P(x,\alpha)$ and $P(x,\beta)$ are two best approximations and $0 \leq t \leq 1$, then

$$MT[f(x) - tP(x,\alpha) - (1 - t)P(x,\beta)]$$

$$= MT[t(f(x) - P(x,\alpha)) + (1 - t)(f(x) - P(x,\beta))]$$

$$\leq MT[t(f(x) - P(x,\alpha))] + MT[(1 - t)(f(x) - P(x,\beta))]$$

$$= tMT[f(x) - P(x,\alpha)] + (1 - t)MT[f(x) - P(x,\beta)] = e(T)$$

Observe that the inequality in the above relation is actually an equality for all t. If this were not true, some linear combination of $P(x,\alpha)$ and $P(x,\beta)$ would be a better approximation than either $P(x,\alpha)$ or $P(x,\beta)$. Therefore, any convex combination of two best approximations is a best approximation. Q. E. D.

1.5 Lemma. Let $\{w_k(x)\}$ for $k \varepsilon I$ be continuous weight functions and

$$M_1[g] = \max_{(x,k)\varepsilon X\times I} \left| w_k(x) D^k g(x) \right|, \quad M_2[g] = \max_{(x,k)\varepsilon X\times I} \left| |w_k(x)| D^k g(x) \right|.$$

Then $P(x,\alpha)$ is a best approximation to $f(x)$ in the pseudonorm M_1 if and only if $P(x,\alpha)$ is a best approximation to $f(x)$ in the pseudonorm M_2.

Proof: $P(x,\alpha)$ is a best approximation in $M_1[g]$ if it minimizes

$$\max_{(x,k)\varepsilon X\times I} |w_k(x) D^k[f(x) - P(x,\alpha)]| . \tag{1}$$

In the case of $M_2[g]$, $P(x,\alpha)$ must minimize

$$\max_{(x,k)\varepsilon X\times I} \left| |w_k(x)| |D^k[f(x) - P(x,\alpha)]| \right|. \tag{2}$$

It is evident that for a given α, the quantities (1) and (2) are identical. This implies that an α which minimizes one of them minimizes both, and completes the proof.

Since the set R of best approximations does not change if the absolute value of the weight function is substituted for the weight function, we will henceforth assume that all weight functions are nonnegative.

In order to study some properties of the base functions of the approximating polynomials, we introduce the following matrices.

1.6 Definition. The interpolation matrix of a set Q which is a set of m ordered pairs $\{(x_i,k_i)\}_{i=1}^m$ or ordered triples $\{(x_i,k_i,s_i)\}_{i=1}^m$ is the m×n matrix B whose entry in the ith row and jth column is

$$b_{ij} = D^{k_i}\phi_j(x_i), \; i = 1,2,\ldots,m \; ; \; j = 1,2,\ldots,n.$$

This matrix will be denoted by IM(Q).

The weighted interpolation matrix of a set Q which is a set of m ordered pairs $\{(x_i,k_i)\}_{i=1}^m$ or ordered triples $\{(x_i,k_i,s_i)\}_{i=1}^m$ is the m×n matrix C whose entry in the ith row and jth column is

$$c_{ij} = w_{k_i}(x_i)D^{k_i}\phi_j(x_i), \; i = 1,2,\ldots,m ; \; j = 1,2,\ldots,n.$$

This matrix will be denoted by WIM(Q).

In the above definition, it should be observed that the interpolation matrix, or weighted interpolation matrix of a set of ordered triples does not depend on the third element of the triple, the s_i. In addition, the matrices IM(Q) and WIM(Q) always have n columns if Q is nonempty.

1.7 Assumption. We will assume that the base functions $\phi_1(x), \phi_2(x), \ldots, \phi_n(x)$ are linearly independent in the sense that

$$M[P(x,\alpha)] = 0 \text{ implies } ||\alpha|| = 0. \qquad (3)$$

This is a natural assumption since without it one cannot hope to prove uniqueness of a best approximation; i.e., if $M[P(x,\alpha)] = 0$ and $P(x,\beta)$ is a best approximation to f(x), then $P(x,c\alpha + \beta)$ is also a best approximation for any real c.

The problem of approximating a function f(x) which is equal to some polynomial in the n base functions on $\{(x,k)\varepsilon X\times I: w_k(x) \neq 0\}$ is easily solved and of little interest. These problems will be dropped from further consideration and hence we may assume that $e > 0$ for any particular problem under consideration.

1.8 Lemma. If $MT[P(x,\alpha)] = 0$ implies $||\alpha|| = 0$, then T contains a set Q of n points such that $\det[WIM(Q)] \neq 0$. (Here det denotes the determinant.)

Proof: Assume that the conclusion is false. Let T' be a set of m points from T for which B = WIM(T') has rank $m < n$, where m is the maximum rank of any weighted interpolation matrix in points from T. Let (x_o, k_o) be any point from T and define $T'' = T' \cup \{(x_o,k_o)\}$. Then the rank of $B' = WIM(T'')$ is m and the row b_{m+1} of B' corresponding to (x_o,k_o) can be written as a linear combination of the rows of B. Hence in matrix notation we have $b_{m+1} = \gamma B$ where γ is a row matrix giving the proper multiples of the rows of B which are needed to produce b_{m+1}. Next consider the space of solutions to the problem $B\alpha^t = 0$. This problem has a solution space of dimension $n - m$, and since $n > m$ there exists a nonzero solution α_o. The weighted value of this polynomial $P(x,\alpha_o)$ at (x_o,k_o) is given by $b_{m+1}\alpha_o^t = \gamma B\alpha_o^t = 0$. Hence $MT[P(x,\alpha_o)] = 0$ and $||\alpha_o|| \neq 0$ which contradicts the hypothesis of this lemma. Q. E. D.

2. Existence

1.9 Theorem. $G(\alpha) \equiv MT[f(x) - P(x,\alpha)]$ is a continuous function for $\alpha \varepsilon E^n$.

Proof: Let $K = \max_{i=1,2,\ldots,n} MT[\phi_i(x)]$. If $K = 0$, $G(\alpha)$ is constant for all $\alpha \varepsilon E^n$ and hence is continuous. If $K \neq 0$, let α_o be a point of E^n. Given any $\varepsilon > 0$ let $\delta = \varepsilon/nK$. Then $||\beta - \alpha_o|| < \delta$ implies

$$|G(\beta) - G(\alpha_o)| = |MT[f(x) - P(x,\beta)] - MT[f(x) - P(x,\alpha_o)]|$$

$$\leq |MT[f(x) - P(x,\beta) - f(x) + P(x,\alpha_o)]| = MT[\sum_{i=1}^{n} (a_i - b_i)\phi_i(x)]$$

$$\leq \max_{i=1,2,\ldots,n} |a_i - b_i| \cdot nK = ||\alpha_o - \beta|| \cdot nK < \varepsilon. \text{ Q. E. D.}$$

1.10 <u>Corollary</u>. The set $R(T)$ of best approximations is closed.

<u>Proof</u>: If $R(T)$ is empty, the result is trivial. Otherwise $R(T) = \{\alpha : G(\alpha) = e(T)\}$ and hence $R(T) = G^{-1}[e(T)]$. Since $G(\alpha)$ is continuous and $e(T)$ is a closed subset of the real numbers, $R(T)$ is closed. Q. E. D.

1.11 <u>Lemma</u>. If $\alpha \varepsilon R(T)$ (the set of best approximations on the set T) then $MT[P(x,\alpha)] \leq 2 \cdot MT[f(x)]$.

<u>Proof</u>: $MT[P(x,\alpha)] - MT[f(x)] \leq MT[f(x) - P(x,\alpha)] \leq MT[f(x)]$ where the second inequality holds because $P(x,\alpha) \equiv 0$ cannot be a better approximation than any best approximation. The result follows from these inequalities. Q. E. D.

1.12 <u>Definition</u>. A family of polynomials $\{P(x,\alpha)\}$ for $\alpha \varepsilon S$ is said to be bounded at a set of points $Q = \{(x_i,k_i)\}_{i=1}^{m}$ or a set $Q' = \{(x_i,k_i,s_i)\}_{i=1}^{m}$ if there exists some constant z such that $|D^{k_i}P(x_i,\alpha)| \leq z$ for all $\alpha \varepsilon S$.

1.13 <u>Lemma</u>. Let $\{P(x,\alpha)\}$ for $\alpha \varepsilon S \subset E^n$ be a family of polynomials in n base functions which is bounded by B' at the n points of $T \subset X \times I$. Then if $\det[WIM(T)] \neq 0$, there exists a constant B such that $||\alpha|| \leq B$ for all $\alpha \varepsilon S$.

<u>Proof</u>: Since $\det[WIM(T)] \neq 0$, any polynomial in these n base functions is determined by the values of the polynomial at the points of T. Then if $V(x_i,k_i)$, $i = 1,2,\ldots,n$ gives the values of a polynomial $P(x,\alpha)$ at the points of

T and L_{ij} is the cofactor of c_{ij} from the ith row and jth column of WIM(T), the coefficients of $P(x,\alpha)$ are

$$a_j = [\sum_{i=1}^{n} L_{ij} V(x_i,k_i)]/\det[WIM(T)],\ j = 1,2,\ldots,n.$$

Then if $L = \underset{ij}{\text{Max}}|L_{ij}|$, $E = |\det[WIM(T)]|$ and $|V(x_i,k_i)| \leq B'$ for $i = 1,2,\ldots,n$ we have

$$|a_j| \leq nLB'/E \quad \text{for all } \alpha\varepsilon S.$$

Thus $B = nLB'/E$ is a uniform bound on $||\alpha||$ for $\alpha\varepsilon S$. Q. E. D.

1.14 <u>Theorem</u>. The set R(T) of best approximations on the set T is bounded if $MT[P(x,\alpha)] = 0$ implies $||\alpha|| = 0$.

<u>Proof</u>: Lemma 1.8 implies that T contains a set $Q = \{(x_i,k_i)\}_{i=1}^{n}$ such that $\det[WIM(Q)] \neq 0$ if $MT[P(x,\alpha)] = 0$ implies $||\alpha|| = 0$. Then letting $W = \underset{k=1,2,\ldots,n}{\text{Min}} |w_{k_i}(x_i)|$ and using Lemma 1.11,

$$2MT[f(x)] \geq MT[P(x,\alpha)] \geq |w_{k_i}(x_i)D^{k_i}P(x_i,\alpha)| \geq W|D^{k_i}P(x_i,\alpha)|$$

for $i = 1,2,\ldots,n$ and $\alpha\varepsilon R(T)$. Since $\det[WIM(Q)] \neq 0$ implies that $W > 0$, it follows that $|D^{k_i}P(x_i,\alpha)| \leq 2MT[f(x)]/W$ for $(x_i,k_i)\varepsilon Q$ and $\alpha\varepsilon R(T)$. Hence Lemma 1.13 implies that R(T) is bounded. Q. E. D.

1.15 <u>Corollary</u>. If $Q = \{(x_i,k_i)\}_{i=1}^{n}$ is a set of points from $X\times I$ such that $\det[WIM(Q)] \neq 0$, then there exists a constant B such that $||\alpha|| < B$ for all $\alpha\varepsilon R(T_1 \cup Q)$ for any $T_1 \subset X\times I$.

<u>Proof</u>: Let $T' = T_1 \cup Q$; then from the proof of Theorem 1.14 we have

$$2M[f(x)] \geq 2MT'[f(x)] \geq W|D^{k_i}P(x_i,\alpha)| \text{ for } i = 1,2,\ldots,n$$

and any $\alpha\varepsilon R(T')$. This implies that

$$\frac{2M[f(x)]}{W} \geq |D^{k_i}P(x_k,\alpha)| \text{ for } i = 1,2,\ldots,n \text{ and } \alpha\varepsilon \bigcup_{T_1 \subset X\times I} R(T_1 \cup Q).$$

Then Lemma 1.13 implies that there exists a constant B such that $||\alpha|| < B$ for all $\alpha\varepsilon R(T_1 \cup Q)$ for any $T_1 \subset X\times I$. Q. E. D.

1.16 <u>Theorem</u>. The set R(T) of best approximations on the set T is nonempty.

<u>Proof</u>:

Case I.--Assume that $MT[P(x,\alpha)] = 0$ implies that $||\alpha|| = 0$. By Theorem 1.14 the set R(T) is bounded. Then $G(\alpha) = MT[f(x) - P(x,\alpha)]$ is a continuous function on a compact set and hence attains its minimum value. Thus a best approximation exists and R(T) is nonempty.

Case II.--If $MT[P(x,\alpha)] = 0$ does not imply that $||\alpha|| = 0$, there is a nontrivial polynomial $P(x,\beta)$ such that $MT[P(x,\beta)] = 0$. This means that

$$w_k(x)D^kP(x,\beta) = w_k(x)\sum_{i=1}^{n} b_i D^k\phi_i(x) = 0 \quad \text{for } (x,k)\varepsilon T$$

where some b_i is nonzero. Therefore, the set of base functions may be reduced in number by at least one without affecting the values that the weighted polynomial can attain on T. i.e., With this reduced set of base functions, we can approximate $f(x)$ on T just as closely as we could using the original set of base functions. This reduction in the number of base functions may be repeated until those remaining satisfy the assumption of Case I. Q. E. D.

3. Characterization

Instead of having one error function as in the approximation of a function by polynomials, we have an error function for each $k \in I$.

1.17 Definition. $L_k(x,\alpha) \equiv w_k(x)D^k[f(x) - P(x,\alpha)]$. The functions $L_k(x,\alpha)$, which are defined for all $(x,k) \in X \times I$ and $\alpha \in E^n$, are called weighted error functions.

Throughout this section on characterization, T will be assumed to be a closed subset of $X \times I$ such that $e(T) > 0$. Such sets T exist because we are assuming that $e = e(X \times I) > 0$.

1.18 Definition. An ordered triple $(x_o,k,s) \in X \times I \times \{-1,1\}$ is called an extremum with respect to the approximation $P(x,\alpha)$ to $f(x)$ on the closed set $T \subset X \times I$ if $(x_o,k) \in T$ and $L_k(x_o,\alpha) = sd$ where $d = MT[f(x) - P(x,\alpha)]$.

1.19 Definition. If $\alpha \in E^n$ and $T \subset X \times I$ is closed, let $C(T,\alpha) = \{(x,k,s) \in X \times I \times \{-1,1\} : (x,k) \in T \text{ and } L_k(x,\alpha) = sd\}$ where $d = MT[f(x) - P(x,\alpha)]$. If $T = X \times I$, let $C(\alpha) = C(T,\alpha)$.

The set $C(T,\alpha)$ is called the set of extrema of the approximation $P(x,\alpha)$ to $f(x)$ in the pseudonorm $MT[g(x)]$.

Since $MT[g(x)]$ gives the maximum absolute value of a finite number of continuous functions on a compact set, every approximation must have at least one extremum. This means that $C(T,\alpha)$ is nonempty for each $\alpha \varepsilon E^n$. The points of $C(T,\alpha)$ are of considerable importance in the computational schemes for solving Problem 1.2 which we discuss in Chapter IV. Here we are interested in the set $C(T,\alpha)$ when $\alpha \varepsilon R(T)$. We shall give certain theorems characterizing the points of $C(T,\alpha)$ in this case.

1.20 Theorem. There exists an $\alpha_o \varepsilon R(T)$ such that for every $\beta \varepsilon R(T)$ and $k \varepsilon I$, $C(T,\alpha_o) \subset C(T,\beta)$.

Proof: We have previously shown that $R(T)$ is closed, convex, and nonempty. Let m be the dimension of $R(T)$. If $m = 0$, $R(T)$ is a single point, the best approximation is unique, and the theorem is true. If $m \geq 1$ the set $R(T)$ has a nonempty interior. Let α_o be in the interior of $R(T)$. It will be shown that this point satisfies the assertion of the theorem.

Let β be any other point in $R(T)$ which is not the same as α_o. Extend the line segment $\overline{\alpha_o \beta}$ beyond α_o to some other point $\alpha_1 \varepsilon R(T)$ so that α_1, α_o and β are on the same line with α_o between α_1 and β. Let $(x_o,k,s) \varepsilon C(T,\alpha_o)$. Then α_o must be a point of the hyperplane

$$w_k(x_o)D^k[f(x_o) - P(x_o,\alpha)] = se(T). \qquad (5)$$

Suppose, for example, that $s = 1$ and that

$$w_k(x_o)D^k[f(x_o) - P(x_o,\beta)] < e(T).$$

Since the three points α_o, α_1 and β are on the line segment $\overline{\alpha_1\beta}$, it follows that α_1 and β must be on opposite sides of the hyperplane in line (5). This means that

$$w_k(x_o)D^k[f(x_o) - P(x_o,\alpha_1)] > e(T)$$

which contradicts the assumption that $\alpha_1 \varepsilon R(T)$. Thus β must be a point of the hyperplane from line (5) since $\beta \varepsilon R(T)$. This same conclusion is also reached if $s = -1$. This implies that every $\beta \varepsilon R(T)$ must satisfy line (5) and hence we conclude that $(x_o,k,s)\varepsilon C(T,\beta)$ for all $\beta \varepsilon R(T)$.
Q. E. D.

We have also proved the following:

1.21 Corollary. If α is an interior point of $R(T)$ then

$$D^kP(x_o,\alpha) = D^kP(x_o,\beta)$$

for every $(x_o,k,s)\varepsilon C(T,\alpha)$ and every $\beta \varepsilon R(T)$.

1.22 Definition. The set $MES(T) = \bigcap_{\beta \varepsilon R(T)} C(T,\beta)$ is called the minimal extremal set (or MES) of the best approximations to $f(x)$ on the set T.

The previous theorem shows that MES(T) is the extremal set of any interior point of R(T). Since $L_k(x,\alpha)$ is continuous in x for fixed k and α, we observe that $C(T,\alpha)$ is closed; hence we have the following:

1.23 Corollary. If T is any closed, nonempty subset of $X \times I$, then MES(T) is compact and nonempty.

So far we do not have any method for determining if a given approximation to f(x) is a best approximation. To alleviate this difficulty, we use the following definition and theorem.

1.24 Definition. A polynomial $P(x,\alpha)$ is said to satisfy Condition A with respect to a set of triples $\{(x_i,k_i,s_i)\}_{i=1}^{m}$ (also called points) where $(x_i,k_i,s_i) \in X \times I \times \{-1,1\}$, if $\operatorname{sgn}[D^{k_i}P(x_i,\alpha)] = -s_i$, $i = 1,2,\ldots,m$.

1.25 Theorem. A polynomial $P(x,\alpha)$ is a best approximation to f(x) on the set T if and only if there is no polynomial $P(x,\beta)$ which satisfies Condition A with respect to $C(T,\alpha)$ the extremal set of the approximation $P(x,\alpha)$.

Proof: Assume that $P(x,\alpha)$ is not a best approximation so that there exists $P(x,\gamma)$ such that $MT[f(x) - P(x,\gamma)] = d' < MT[f(x) - P(x,\alpha)] = d$. Let $P(x,\beta) = P(x,\alpha) - P(x,\gamma)$. Then for any $(x_o,k,s) \in C(T,\alpha)$ we have

$$w_k(x_o)D^k[f(x_o) - P(x_o,\alpha)] = sd.$$

Since $w_k(x_o) > 0$ it follows that

$$D^kP(x_o,\beta) = D^k[f(x_o) - P(x_o,\gamma)] - D^k[f(x_o) - P(x_o,\alpha)]$$

is greater than or less than $s(d' - d)/w_k(x_o)$ as s is -1 or +1 respectively. Therefore, $P(x,\beta)$ satisfies Condition A with respect to $C(T,\alpha)$.

Let $d = MT[f(x) - P(x,\alpha)]$ and assume that there exists a polynomial $P(x,\beta)$ satisfying Condition A with respect to $C(T,\alpha)$. Then let

$$U = \text{Closure } \{(x,k)\epsilon T: \operatorname{sgn} D^kP(x,\beta) = \operatorname{sgn} D^k[f(x) - P(x,\alpha)]\}$$

$$d' = \underset{(x,k)\epsilon U}{\text{Max}} |L_k(x,\alpha)| \qquad r = \underset{(x,k)\epsilon T}{\text{Max}} |w_k(x)|$$

$$s = \underset{(x,k)\epsilon T}{\text{Max}} |D^kP(x,\beta)| \qquad t = \text{Min}\{(d-d')/2rs, d'/rs\}.$$

Then

$$w_k(x)\{D^k[f(x) - P(x,\alpha - \beta t)]\} = L_k(x,\alpha) + w_k(x)D^kP(x,\beta t) \quad (6)$$

and using the quantities defined above we have

$$|w_k(x)D^kP(x,\beta t)| \leq |rtD^kP(x,\beta)| \leq rts =\{\text{Min } (d-d')/2, d'\}.$$

On the set U, the two terms on the right side of (6) are of the same sign and less than d' and (d-d')/2 respectively in absolute value. On T ~ U (where ~ is used to denote the usual set difference) the terms on the right of (6) are of opposite sign and satisfy

$$|L_k(x,\alpha) + w_k(x)D^kP(x,\beta t)| < d$$

since

$$|w_k(x)D^kP(x,\beta t)| \leq d' \quad \text{on } T \sim U.$$

Hence $P(x,\alpha)$ was not a best approximation. Q. E. D.

In order to better understand the characterization of best approximations, we focus our attention on certain subsets of the MES.

1.26 <u>Definition</u>. A set $Q \subset X \times I \times \{-1,1\}$ is called a minimal characterization set (or MCS) if there exists no polynomial satisfying Condition A with respect to Q but for any $Q_0 \subset Q$, $Q_0 \neq Q$ there exists a polynomial which satisfies Condition A with respect to Q_0. (Observe that a MCS has nothing to do with a particular function $f(x)$ being approximated.)

1.27 <u>Definition</u>. If T is a closed subset of $X \times I$, let MCS(T) denote the collection of all minimal characterization sets which are subsets of MES(T).

The above definition makes sense because there may be more than one MCS contained in the set MCS(T) as is shown in the following:

1.28 <u>Example</u>. Let $X = [0,4\pi]$, $I = \{0\}$, $w_0(x) \equiv 1$, $n = 2$, $\phi_1(x) \equiv 1$, $\phi_2(x) = x$, $f(x) = \sin x$. This is an ordinary Chebyshev approximation problem, and we know that the MES is

$C = \{(\pi/2,0,1),\ (3\pi/2,0,-1),\ (5\pi/2,0,1),\ (7\pi/2,0,-1)\}$

Also from the theory of ordinary Chebyshev approximation we know that both

$$D = \{(\pi/2,0,1),\ (3\pi/2,0,-1),\ (5\pi/2,0,1)\}$$

$$E = \{(3\pi/2,0,-1),\ (5\pi/2,0,1),\ (7\pi/2,0,-1)\}$$

are in MCS($X\times I$).

1.29 Definition. Corresponding to each ordered triple (x_o,k,s), define a point

$$(sD^k\phi_1(x_o),sD^k\phi_2(x_o),\ldots,sD^k\phi_n(x_o))$$

in E^n which will be called an n-point (as opposed to the point (x_o,k,s) itself).

The following theorem is a modification and extension of the Lemma of G. F. Voronoi, Remez [22] p. 112.

1.30 Theorem. Let Y be a closed subset of $X\times I\times\{-1,1\}$. Then there exists a polynomial satisfying Condition A with respect to the points of Y if and only if the convex hull of the n-points of Y does not contain the origin of E^n.

Proof: Suppose the convex hull of the n-points of Y does not contain the origin. Then since Y is compact, the convex hull of Y is compact and hence there is a hyperplane $\alpha\cdot Z = a_1z_1 + a_2z_2 + \ldots + a_nz_n = -b$ where $b > 0$ which separates the origin and the convex hull of Y. Thus for each n-point Y_i of Y, the inequality $\alpha \cdot Y_i < 0$ is satisfied. Thus $P(x,\alpha)$ is a polynomial satisfying Condition A with respect to the points of Y.

Assume that there exists a polynomial $P(x,\alpha)$ satisfying Condition A with respect to the points of Y. Then $\alpha \cdot Y_i < 0$ for all n-points Y_i of Y. Then since Y is compact, there exists a constant $b > 0$ such that $\alpha \cdot Y_i < -b$ for all Y_i of Y. Thus the origin is not in the convex hull of the n-points of Y. Q. E. D.

Next we have a theorem which gives an upper bound on the number of points in a MCS. It states that a MCS can have at most n+1 points.

1.31 Theorem. If Y is a closed subset of $X \times I \times \{-1,1\}$ and if the n-points of Y contain the origin in their convex hull, then there is a set of $k \leq n + 1$ n-points of Y whose convex hull contains the origin.

This theorem is an immediate consequence of the Theorem of Carathéodory which is stated below as given in Cheney [6] p. 17.

1.32 Theorem of Carathéodory. Let A be a subset of n-dimensional linear space. Every point of the convex hull of A is expressible as a convex linear combination of n+1 (or fewer) elements of A.

In ordinary Chebyshev approximation by polynomials in a Chebyshev set of base functions a MCS must contain exactly n+1 points. This is because the interpolation matrix in any n distinct points is nonsingular and hence for any n distinct points we have a polynomial satisfying Condition A. In the more general case a MCS may have fewer than n+1 points, and there may be more than one MCS having fewer than n+1 points. Also, there may be two or more minimal characterization sets which do not have the same number of points. These special cases will be shown in the following examples.

1.33 Example. Let $I = \{0,1\}$, $w_0(x) \equiv w_1(x) \equiv 1$, $n = 4$, $\phi_i(x) = x^{i-1}$ for $i = 1,2,3,4$ and $X = \{-1,-1/\sqrt{3},0,1/\sqrt{3},1\}$. The function $f(x)$ is defined on $X \times I$ by the following table.

x	$f(x)$	$f'(x)$
-1	-1	0
$-\frac{1}{\sqrt{3}}$	0	-1
0	1	0
$\frac{1}{\sqrt{3}}$	0	1
1	-1	0

Let $B_1 = \{(-1,0,-1),(-1/\sqrt{3},1,-1),(0,0,1),(1,0,-1)\}$.

$B_2 = \{(-1,0,-1),(1/\sqrt{3},1,1),(0,0,1),(1,0,-1)\}$.

Then

$$B' = \{(-1,1,-1,1),(0,-1,2/\sqrt{3},-1),(1,0,0,0),(-1,-1,-1,-1)\}$$

is the set of n-points of both B_1 and B_2. The origin can be expressed as a convex combination of these points of E^4 by using the multiples

$$\frac{2+\sqrt{3}}{2(4+\sqrt{3})} \quad \frac{\sqrt{3}}{4+\sqrt{3}} \quad \frac{2}{4+\sqrt{3}} \quad \frac{2-\sqrt{3}}{2(4+\sqrt{3})} .$$

Thus using Theorem 1.30 there is no polynomial satisfying Condition A with respect to the points of B_1 or B_2. The nonsignularity of the generalized Vandermonde matrix implies that there is a polynomial satisfying Condition A with respect to any 3 point subset of either B_1 or B_2. Thus B_1 and B_2 are both minimal characterization sets.

The sets B_1 and B_2 are both subsets of C(0) where 0 is the origin of E^4. Since there is no polynomial satisfying Condition A with respect to B_1 or B_2, there can not be any polynomial which satisfies Condition A with respect to C(0). Hence $\alpha = 0$ is a best approximation, and the sets B_1 and B_2 are both in MCS(X×I) if 0 is an interior point of R.

To show that 0 is interior to R we will determine R. Since 0 is a best approximation, $e = M[f(x) - P(x,\alpha)] = 1$ for $\alpha = 0$ and the following inequalities must be satisfied.

$$-1 \leq f(x_i) - P(x_i,\alpha) \leq 1 \quad i = 1,2,3,4,5 \quad (a)$$

$$-1 \leq f'(x_i) - P'(x_i,\alpha) \leq 1 \quad i = 1,2,3,4,5 \quad (b)$$

$$x_1 = -1,\ x_2 = -1/\sqrt{3},\ x_3 = 0,\ x_4 = 1/\sqrt{3},\ x_5 = 1.$$

In the following computation we will refer to these inequalities by a letter and a number. e.g., (b) 5 refers to line (b) with $i = 5$.

From lines (a) and (b) we have

$$-1 \leq -1 - a_1 + a_2 - a_3 + a_4 \qquad \text{from (a) 1}$$

$$-1 \leq -1 - a_1 - a_2 - a_3 - a_4 \qquad \text{from (a) 5.}$$

Then adding and dividing by 2

$$-1 \leq -1 - a_1 - a_3$$

$$-1 \leq -1 + a_1 \qquad \text{from (a) 3.}$$

Adding these we obtain $a_3 \leq 0$. (7)

$$-1 \leq -1 - a_2 + (2/\sqrt{3})a_3 - a_4 \qquad \text{from (b) 2}$$

$$-1 \leq -1 + a_2 + (2/\sqrt{3})a_3 + a_4 \qquad \text{from (b) 4.}$$

Adding and dividing by 2

$$-1 \leq -1 + (2/\ 3)a_3 \text{ implies } 0 \leq (2/\sqrt{3})a_3$$

and using (7) we know $a_3 = 0$.

Then equation (a)3 implies that $a_1 \geq 0$.

$$-1 \leq -1 - a_1 + a_2 + a_4 \qquad \text{from (a) 1}$$

$$-1 \leq -1 - a_1 - a_2 - a_4 \qquad \text{from (a) 5.}$$

Adding and dividing by 2

$$-1 \leq -1 - a_1 \text{ and hence } a_1 \leq 0 \text{ and thus } a_1 = 0.$$

Equation (b)4 implies that $1 - a_2 - a_4 \leq 1$ and hence $-a_2 - a_4 \leq 0$.

Then (a) 5 implies that $0 \leq -a_2 - a_4$ and hence $a_2 = a_4$.

From the computation using lines (a) and (b) we know that the best approximations are polynomials of the form $a_4(x^3 - x)$. Since this family of polynomials is zero at the points of $B_1 \cup B_2$, the solution space R must be determined by the other points of X×I. Using the equations (b)1 or (b)5 we have $-1 \leq -a_2 + a_4 \leq 1$ or $-1 \leq 2a_4 \leq 1$. Then it can be verified that all of the other inequalities from (a) and (b) are satisfied for these values of a_4. Thus the set R of best approximations is $a_4(x^3 - x)$ for $-1/2 \leq a_4 \leq 1/2$. Therefore, $\alpha = 0$ is interior to R, C(0) is the set MES(X×I), and the sets B_1 and B_2 are in MCS(X×I).

1.34 Example. Let $X = \{-2,-1,0,1,2\}$, $I = \{0,1\}$, $n = 3$, $w_0(x) \equiv w_1(x) \equiv 1$ and $\phi_i(x) = x^{i-1}$ for $i = 1,2,3$. The function f(x) is defined on X×I by the following table:

x	f(x)	f'(x)
-2	1	0
-1	-1	0
0	0	-1
1	1	0
2	-1	0

From ordinary Chebyshev approximation theory we know that

$$B_1 = \{(-2,0,1),(-1,0,-1),(1,0,1),(2,0,-1)\}$$

is a MCS and that $\alpha = 0$ is the unique best approximation with $e = 1$. Now we will show that

$$B_2 = \{(-1,0,-1),(0,1,-1),(1,0,1)\}$$

is also a MCS. We must show that there is no polynomial satisfying Condition A with respect to B_2. Thus we need to verify that the following inequalities are inconsistent.

$$a_1 + a_2 + a_3 < 0 \tag{8}$$

$$a_2 > 0 \tag{9}$$

$$a_1 - a_2 + a_3 < 0 \tag{10}$$

Adding (8) and (10) we have $a_2 < 0$, which contradicts (9). Therefore, B_2 is a MCS. Since both B_1 and B_2 are subsets of $C(0) = MES(X \times I)$, the sets B_1 and B_2 are in $MCS(X \times I)$. Thus the various minimal characterization sets for a best approximation (those in $MCS(X \times I)$) need not have the same number of points.

1.35 Theorem. If T is a finite subset of $X \times I$, the dimension of R(T) is n - k where n is the number of base functions and k is the rank of $B = WIM[MES(T)]$.

Proof: Let $T_o = \{(x,k):(x,k,s) \varepsilon MES(T)\}$ and let $T_1 = T \sim T_o$. If $P(x,\gamma)$ is a best approximation, where γ is in the interior of R(T), (if the best approximation is unique, let $P(x,\gamma)$ be this unique best approximation) then $MT_1[f(x) - P(x,\gamma)] = d < e(T)$. This is because the maximum is taken over points which are not points of MES(T) and because T is finite. Then the continuity of $MT_1[f(x) - P(x,\gamma)]$ in γ implies that there exists an $\varepsilon > 0$ such that $MT_1[f(x) - P(x,\gamma) - P(x,\alpha)] < e(T)$ for $\alpha \varepsilon N(o,\varepsilon)$ a small neighborhood of the origin in E^n.

Next consider the solution space S of $B\alpha^t = 0$. S has dimension n-k where k is the rank of B. We also know that $MT_o[P(x,\alpha)] = 0$ and $MT_o[f(x) - P(x,\gamma) - P(x,\alpha)] = e(T)$ for all $\alpha \varepsilon S$. Then for $\alpha \varepsilon N(o,\varepsilon) \cap S$ it follows that $MT[f(x) - P(x,\gamma) - P(x,\gamma)] = e(T)$ since $T = T_o \cup T_1$. Since $N(o,\varepsilon) \cap S$ has dimension n-k, the dimension of R(T) must be at least n-k. But any two best approximations must be equal at the points of T_o. Then if the dimension of R(T) is greater than n-k, the solution space of $B\alpha^t = 0$ must also be greater than n-k. This is a contradiction to the assumption that k is the rank of B. Q. E. D.

In the following corollaries, T is assumed to be a closed subset of $X \times I$ and not necessarily finite.

1.36 Corollary. If T is any closed subset of $X \times I$, the dimension of R(T) is bounded above by n-k.

1.37 Corollary. If R(T) is not a single point and is a boundary point of R(T), then $C(T,\alpha)$ contains at least one more point than MES(T).

Proof: If not, then the proof of Theorem 1.35 implies that for some $\varepsilon > 0$ there is a neighborhood $N(\alpha,\varepsilon)$ of α such that $N(\alpha,\varepsilon) \cap S \subset R(T)$. This contradicts the assumption that α is on the boundary of R(T).

1.38 Corollary. If the rank of WIM[MES(T)] is equal to n, then the set of best approximations R(T) is a single point and the best approximation is unique.

1.39 Definition. A polynomial $P(\lambda,\alpha)$ is said to have a zero at the point (x_o,k_o) or the point (x_o,k_o,s_o) if $D^{k_o}P(x_o,\alpha) = 0$.

1.40 Lemma. If Q is a MCS of $k \leq n + 1$ points, then the ranks of both IM(Q) and IM(Q_o) are k-1 where Q_o is any k-1 point subset of Q.

Proof: Let m be the rank of IM(Q). This means that the n-points of Q (to be denoted NQ) span a subspace of E^n of dimension m. (The rows of IM(Q) are the n-points of Q except for the factor of -1 if s = -1.) The origin of E^n is in the convex hull of NQ because Q is a MCS. Applying the Theorem of Carathéodory to this m dimensional subspace, the origin may be expressed as a convex combination of some set of not more than m+1 points from NQ. Since Q is a MCS, the origin cannot be expressed as a linear combination of any subset of NQ and hence $k \leq m + 1$. If $k < m + 1$, the rank of IM(Q) is k and there exists a polynomial satisfying Condition A with respect to Q. Thus $k = m + 1$ and the rank of IM(Q) is k-1.

Suppose that the rank of IM(Q_o) is less than k-1 for some k-1 point subset Q_o of Q. This implies that there exists a polynomial P(x) which is zero at the points of Q_o and nonzero at $(x_o,k_o,s_o) = Q \sim Q_o$. Since Q is a MCS there exists a polynomial V(x) which satisfies Condition A with respect to Q_o such that $s_o \neq -\mathrm{sgn}[D^{k_o}V(x_o)]$. Then if $D^{k_o}V(x_o) \neq 0$,

$$V(x) - 2[D^{k_o}V(x_o)]\mathrm{sgn}[D^{k_o}P(x_o)]P(x)/|D^{k_o}P(x_o)|$$

is equal to V(x) at the points of Q_o and is equal to $-D^{k_o}V(x_o)$ at (x_o,k_o,s_o). If $D^{k_o}V(x_o) = 0$, then

$$V(x) - s_o\mathrm{sgn}[D^{k_o}P(x_o)]P(x)/|D^{k_o}P(x_o)|$$

is equal to $V(x)$ at the points of Q_o and is equal to $-s_o$ at (x_o, k_o, s_o). Thus in either case, there exists a polynomial satisfying Condition A with respect to Q. This is a contradiction and implies that the rank of $IM(Q_o)$ is $k-1$. Q. E. D.

1.41 Theorem. If any set in MCS(T) contains $n+1$ points then the best approximation is unique.

Proof: Let Q be a set of $n+1$ points which is in MCS(T). Since we are assuming that $e(T) > 0$ in this section, the ranks of WIM(Q) and IM(Q) are the same. (The weight functions are nonzero on Q.) Since Q is a subset of MES(T) we may apply Corollary 1.38 and Lemma 1.40. Q. E. D.

CHAPTER II

CONVERGENCE CONSIDERATIONS

1. Uniform Convergence of the Polynomial to the Function

In uniform approximation theory one basic property that any set of approximating polynomials should possess is that of the uniform convergence of the polynomials of best approximation to the function being approximated as the number of base functions increases. This is desirable because when approximating a function, a certain maximum allowable deviation is usually specified and the number of base functions is chosen so that this requirement can be satisfied. If the approximating polynomials do not converge uniformly to the function, then for some specified maximum deviation $\varepsilon > 0$ we may use as many base functions as we please, but the deviation e of the best approximation will always be greater than ε. The following example illustrates a typical difficulty which must be avoided.

2.1 Example. Let $f(x) = \sin x + 1$, $I = \{0,1\}$, $X = [-1,1]$, $\phi_i(x) = x^i$ for $i = 1,2,\ldots$ and $w_0(x) \equiv w_1(x) \equiv 1$. Then we know that $DP(x,\alpha)$ can approximate $Df(x)$ as closely as desired using the Weierstrass Approximation Theorem, but the difference between $f(x)$ and $P(x,\alpha)$ is 1 at $x = 0$ for any polynomial in these base functions.

The following theorem gives conditions on the base functions which are sufficient to overcome this difficulty.

2.2 Theorem. Let $\{\phi_i(x)\}_{i=1}^{\infty}$ include the set $\{x^{i-1}\}_{i=1}^{q}$ and let $X = [a,b]$, where $b > a$. Suppose that for any $f(x)\varepsilon C^q(X)$ and any $\varepsilon' > 0$ there exists an n_o and a polynomial $P(x,\beta)$ in n_o base functions such that $|D^q[f(x) - P(x,\beta)]| < \varepsilon'$ on X. Then given any $\varepsilon > 0$ there exists an n and a polynomial $P(x,\alpha) = \sum_{i=1}^{n} a_i\phi_i(x)$ such that $M[f(x) - P(x,\alpha)] < \varepsilon$.

Proof: Let $\varepsilon > 0$ be given and let $\varepsilon' = \varepsilon/E$ where $E = \underset{(x,k)\varepsilon X\times I}{\text{Max}} [|w_k(x)|(b - a)^{q-k}]$. This maximum exists because the functions $w_k(x)$ are continuous on X for $k\varepsilon I$. Obtain a polynomial $P(x,\beta)$ which has $\{x^{i-1}\}_{i=1}^{q}$ among its base functions and for which the coefficients of $\{x^{i-1}\}_{i=1}^{q}$ are zero, such that $|D^q[f(x) - P(x,\beta)]| < \varepsilon'$ on X.

Let $T(x)$ be the Taylor expansion of $P(x,\beta) - f(x)$ about the point $x = a$ with q terms and define $P(x,\alpha) = P(x,\beta) - T(x)$. Then for any $x\varepsilon X$

$$|D^q[f(x) - P(x,\alpha)]| = |D^q[f(x) - P(x,\beta)]| < \varepsilon'.$$

More generally we have

$$|D^k[f(x) - P(x,\alpha)]| < \varepsilon'(b - a)^{q-k} \quad \text{for } k = 0,1,\ldots,q$$

This is true because $D^k[f(a) - P(a,\alpha)] = 0$ for $k = 0,1,\ldots,q-1$. Hence

$$|w_k(x)D^k[f(x) - P(x,\alpha)]| < \max_{x\varepsilon X}|w_k(x)|\varepsilon'(b - a)^{q-k} \leq E\varepsilon' = \varepsilon$$

for all $k\varepsilon I$ and $x\varepsilon X$. Q. E. D.

2. Continuous Dependence of the Approximating Polynomial on the Function

The next theorem is an extension of a theorem by Maehly and Witzgall [14]. It relates the closeness of $P(x,\alpha_f)$ and $P(x,\alpha_g)$, (best approximations to $f(x)$ and $g(x)$ respectively in the pseudonorm M) as a function of the closeness of $f(x)$ and $g(x)$ where $g(x)$ is considered fixed. As expected, under the appropriate hypotheses the best approximations to $f(x)$ and $g(x)$ will be close to each other if $f(x)$ is sufficiently close to $g(x)$.

2.3 Theorem. Let $f(x)$ and $g(x)$ be in $C^q(X)$ with sets of best approximations R_f and R_g respectively on $X\times I$, with respect to the weight functions $\{w_k(x)\}_{k\varepsilon I}$. If $MCS(X\times I)_g$ (the set of minimal characterization sets which are subsets of $MES(X\times I)$ for the function $g(x)$) contains a MCS of n+1 points, then there exists a constant B which depends only on $g(x)$, $\{\phi_i(x)\}_{i=1}^n$ and $\{w_k(x)\}_{k\varepsilon I}$ such that

$$\max_{(x,k)\varepsilon X\times I} |D^k[P(x,\alpha_f) - P(x,\alpha_g)]| \leq B \max_{(x,k)\varepsilon X\times I} |D^k[f(x) - g(x)]|$$

where α_g is the unique best approximation to $g(x)$ and α_f is any best approximation in R_f.

<u>Proof</u>: Let $Q \equiv \{(x_i,k_i,s_i)\}_{i=1}^{n+1}$ be a MCS from $MCS(X\times I)_g$ and let e_f and e_g be the deviations of $P(x,\alpha_f)$ and $P(x,\alpha_g)$ from $f(x)$ and $g(x)$ respectively. (Since it is assumed from the beginning of this paper that e, the deviation of the best approximations to a function, is nonzero for any function under consideration, we will assume that $e_g > 0$.) Also let

$$\delta = \max_{(x,k)\varepsilon X\times I} |D^k[f(x) - g(x)]| \text{ and } W = \max_{(x,k)\varepsilon X\times I} |w_k(x)|.$$

Then

$$e_f \leq M[f(x) - P(x,\alpha_g)] \leq M[f(x) - g(x)]$$

$$+ M[g(x) - P(x,\alpha_g)] \leq W\delta + e_g$$

and hence

$$-W\delta \leq e_g - e_f \tag{1}$$

We also know that

$$s_i w_{k_i}(x_i) D^{k_i}[g(x_i) - P(x_i,\alpha_g)] = e_g \tag{2}$$

$$s_i w_{k_i}(x_i) D^{k_i}[f(x_i) - P(x_i,\alpha_f)] \leq e_f \tag{3}$$

for $i = 1,2,\ldots,n+1$; upon subtracting (3) from (2) and using line (1) we have

$$s_i w_{k_i}(x_i)[D^{k_i}P(x_i,\alpha_f - \alpha_g)] \geq s_i w_{k_i}(x_i)D^{k_i}[f(x_i) - g(x_i)]$$

$$+ e_g - e_f \geq - 2W\delta.$$

Thus letting $\alpha = \alpha_f - \alpha_g$ we see that

$$s_i D^{k_i}P(x_i,\alpha) \geq - \frac{2W\delta}{z} \text{ for } i = 1,2,\ldots n+1 \qquad (4)$$

where $z = \underset{i=1,\ldots,n+1}{\operatorname{Min}} |w_{k_i}(x_i)|$. ($z$ is not zero because $e_g > 0$.)

To complete the proof we shall prove the following lemma.

2.4 Lemma. Let $P(x,\alpha)$ satisfy $s_i D^{k_i}P(x_i,\alpha) \geq -2W\delta/z$ for $i = 1,2,\ldots,n+1$, where $Q \equiv \{(x_i,k_i,s_i)\}_{i=1}^{n+1}$ is a MCS of $n+1$ points. Then there is a constant B independent of δ such that

$$\underset{(x,k)\varepsilon X\times I}{\operatorname{Max}} |D^k P(x,\alpha)| \leq B\delta.$$

Proof: Assume that such a constant does not exist. Then for each integer m there exists an α_m such that $P(x,\alpha_m)$ satisfies (4) and $\underset{(x,k)\varepsilon X\times I}{\operatorname{Max}} |D^k P(x,\alpha_m)| > m\delta$. As a consequence of Lemma 1.40 and Lemma 1.13, the sequence $\{P(x,\alpha_m)\}$ is bounded at most $n-1$ points of Q. In addition, for each polynomial in the sequence there must be a point

$(x_i,k_i,s_i)\varepsilon Q$ such that $\mathrm{sgn}[D^{k_i}P(x_i,\alpha_m)] \neq -s_i$. (Otherwise we would have a polynomial satisfying Condition A with respect to Q.) Thus for each α_m there exists a point (x_i^m,k_i,x_i) Q such that

$$0 \leq s_i D^{k_i}P(x_i^m,\alpha_m) \leq 2W\delta/z.$$

Choose a subsequence of $\{P(x,\alpha_m)\}$, also to be denoted $\{P(x,\alpha_m)\}$, so that the sequences $\{D^{k_i}P(x_i,\alpha_m)\}$ converge to λ_i or diverge to $\pm\infty$. This divides Q into two parts, Q_1 and Q_2 where Q_1 is where $\{D^{k_i}P(x_i,\alpha_m)\}$ is bounded and Q_2 is the rest. Thus Q_1 and Q_2 are both nonempty and Q_1 contains no more than n-1 points.

For each (x_j,k_j,s_j) in Q_1 consider the functions $P(x,\beta_j)$ such that

$$D^{k_j}P(x_j,\beta_j) = s_j - \lambda_j$$

$$D^{k_j}P(x_i,\beta_j) = 0 \qquad \text{for } x_i\varepsilon Q_1,\ i \neq j.$$

Then specify $P(x,\beta_j)$ at as many more points to make it unique. (We know that these polynomials exist by Lemma 1.40.) Consider the function

$$P(x,\bar{\alpha}_m) \equiv P(x,\alpha_m) + \sum_{x_i\varepsilon T_1} P(x,\beta_i).$$

For each point of Q_2, $\{D^{k_i}P(x_i,\bar{\alpha}_m)\}$ diverges to $s_i\infty$.

For points of Q_1 we have

$$D^{k_i}P(x_i,\bar{\alpha}_m) = D^{k_i}P(x_i,\alpha_m) + s_i - \lambda_i.$$

Thus for m large enough, $\operatorname{sgn}[D^{k_i}P(x_i,\bar{\alpha}_m)] = s_i$ for $i = 1,2,\ldots,n+1$ and therefore $-P(x,\bar{\alpha}_m)$ satisfies Condition A with respect to Q. This is a contradiction and completes the proof of the lemma and Theorem 2.3.

A more general result which is applicable to all approximation problems under consideration is obtained in the following theorem.

2.5 Theorem. Let M be a pseudonorm over $X \times I$ with weight functions $w_k(x)$ for $k \varepsilon I$. Let $f(x)$ and $g(x)$ be functions with sets of best approximations R_f and R_g respectively. Then given $\varepsilon > 0$ there exists a $\delta > 0$ depending only on ε and $g(x)$ such that if $\alpha_f \varepsilon R_f$, then $M[g(x) - f(x)] < \delta$ implies that

$$\min_{\alpha_g \varepsilon R_g} M[P(x,\alpha_g) - P(x,\alpha_f)] < \varepsilon.$$

Proof: If this theorem is not true, there exists an $\varepsilon_o > 0$ and a sequence of functions $\{f_m(x)\}$ and their corresponding sets of best approximations R_{f_m} from which we obtain a sequence $\{P(x,\alpha_m)\}$ which satisfies the following:

$$\min_{\alpha_g \varepsilon R_g} M[P(x,\alpha_g) - P(x,\alpha_m)] \geq \varepsilon_o$$

for all m. (5)

$$M[g(x) - f_m(x)] < \frac{1}{m}$$

Then using Lemma 1.11 and the triangle inequality we have

$$M[P(x,\alpha_m)] \leq 2\cdot M[f_m(x)] \leq 2(M[g(x)] + 1/m)$$

$$\leq 2M[g(x)] + 2 \text{ for all m.}$$

Since we are assuming that $M[P(x,\alpha)] = 0$ implies $||\alpha|| = 0$, there exists a set T of n points which has a nonsingular interpolation matrix. Lemma 1.13 implies that the coefficients of $\{P(x,\alpha_m)\}$ are uniformly bounded for all m and hence there is a convergent subsequence, which also will be denoted $\{P(x,\alpha_m)\}$, which converges to a polynomial $P(x,\alpha_\infty)$. Then the assumption given in line (5) implies that $P(x,\alpha_\infty)$ is not a best approximation to $g(x)$. Therefore there exists a constant $\varepsilon' > 0$ such that if $\alpha_g \varepsilon R_g$ then

$$M[g(x) - P(x,\alpha_g)] + \varepsilon' = M[g(x) - P(x,\alpha_\infty)]$$

$$\leq M[g(x) - f_m(x)] + M[f_m(x) - P(x,\alpha_m)]$$

$$+ M[P(x,\alpha_m) - P(x,\alpha_\infty)]. \qquad (6)$$

Then determine N so that $m \geq N$ implies that

$$M[g(x) - f_m(x)] \leq \varepsilon'/4 \text{ and } M[P(x,\alpha_m) - P(x,\alpha_\infty)] \leq \varepsilon'/4. \tag{7}$$

Therefore, for $m \geq N$ it follows from (6) and (7) that

$$M[g(x) - P(x,\alpha_g)] + \varepsilon'/2 \leq M[f_m(x) - P(x,\alpha_m)]. \tag{8}$$

But we also have

$$\begin{aligned} M[f_m(x) - P(x,\alpha_g)] &\leq M[f_m(x) - g(x)] + M[g(x) - P(x,\alpha_g)] \\ &\leq \varepsilon'/4 + M[g(x) - P(x,\alpha_g)] \text{ for } m \geq N. \end{aligned} \tag{9}$$

Thus (8) and (9) imply that there exists an m such that

$$\begin{aligned} M[f_m(x) - P(x,\alpha_g)] + \varepsilon'/4 &\leq \varepsilon'/2 + M[g(x) - P(x,\alpha_g)] \\ &\leq M[f_m(x) - P(x,\alpha_m)]. \end{aligned}$$

This is a contradiction since $P(x,\alpha_m)$ is a best approximation to $f_m(x)$. Q. E. D.

3. The de la Vallée Poussin Algorithm

In this section it will be assumed that $X = [a,b]$ where $a < b$. Moreover, the approximation problems under consideration all satisfy Assumption 1.7 and hence from Lemma 1.8 there exists a set Q_o of n points from $X \times I$ for which $\det[WIM(Q_o)] \neq 0$. From the set $X \times I$ choose a countable dense subset Q' which contains the set Q_o. Let $Q'' = \{(x,k)\varepsilon X\times I : w_k(x) = 0\}$ and $Q = Q' \sim Q''$. Since the weight functions are continuous, Q is dense in $X\times I$ except where the weight functions are identically zero on an interval. In addition, since $\det[WIM(Q_o)] \neq 0$, we know that $Q_o \subset Q$.

Let T_m be subsets of Q defined for $m \geq n$ and containing m points such that $Q_o \subset T_k \subset T_{k+1}$ for $k = n, n+1, \ldots$ where the limiting set of points T_m as m approaches ∞ is Q. Also, let δ_m (the spacing or density of the set T_m in Q) be defined as

$$\delta_m = \sup_{(x,k)\varepsilon Q} \{\min_{(x',x')\varepsilon T_m}[|x - x'| + (b-a)|k-k'|]\}. \tag{10}$$

The quantity b-a is used in this definition so that $\delta_m < b-a$ implies that there are at least two points $(x,k)\varepsilon T_m$ with second coordinates equal for each $k\varepsilon I$.

Then one would hope that for $\alpha_m \varepsilon R(T_m)$ and $\alpha^* \varepsilon R$ it would follow that

$$\lim_{m\to\infty} M[f(x) - P(x,\alpha_m)] = M[f(x) - P(x,\alpha^*)] = e$$

More specifically, if $\{(P(x,\alpha_m)\}$ is any sequence with terms from $R(T_m)$, one would hope that any convergent subsequence of $\{P(x,\alpha_m)\}$ would converge to a polynomial in R. These results are proved below, with stronger conclusions being obtained in certain cases.

2.6 Definition. $f(x)$ is said to satisfy a Hölder condition on $[a,b]$ with exponent ρ if there exists a constant λ such that

$$|f(y) - f(x)| \leq \lambda|y - x|^{\rho} \qquad \rho > 0$$

for all x and y in $[a,b]$.

2.7 Remark. If $f'(x)$ satisfies a Hölder condition on $[a,b]$ with exponent $\rho \leq 1$, then $f(x)$ also satisfies a Hölder condition on $[a,b]$ with exponent ρ. This follows from the boundedness of $f'(x)$ on $[a,b]$ and the fact that if $f(x)$ satisfies a Hölder condition with exponent 1, then it also satisfies a Hölder condition with exponent ρ for $0 < \rho < 1$. (See Natanson [21] p. 72.)

2.8 Theorem. Let $D^q f(x)$ and $\{D^q \phi_i(x)\}_{i=1}^{n}$ satisfy a Hölder condition with exponent $\rho \leq 1$. Then there exists a constant K depending on $f(x)$, $\{\phi_i(x)\}_{i=1}^{n}$ and $\{w_k(x)\}_{k\epsilon I}$ such that

$$0 \leq M[f(x) - P(x,\alpha_m)] - e \leq K(\delta_m)^{\rho}$$

for $m \geq n$ and any $\alpha_m \epsilon R(T_m)$. (Here δ_m is defined in line (10).)

<u>Proof</u>: Choose N so that $m \geq N$ implies that $\delta_m < b-a$ and then assume that $m \geq N$ in this proof. Let $(y,k) \varepsilon X \times I$ be such that

$$M[f(x) - P(x,\alpha_m)] = |w_k(y)D^k[f(y) - P(y,\alpha_m)]|$$

for a fixed $\alpha_m \varepsilon R(T_m)$; also, let (z,k) be a point in T_m such that $|y-z| < \delta_m$. Let e and $e(T_m)$ denote the deviations of $P(x,\alpha^*)$ and $P(x,\alpha_m)$ from $f(x)$ on $X \times I$ and T_m respectively where $\alpha^* \varepsilon R$ and $\alpha_m \varepsilon R(T_m)$. Then

$$M[f(x) - P(x,\alpha_m)] = |w_k(y)D^k[f(y) - P(y,\alpha_m)]|$$

$$\leq |w_k(y)D^k[f(y) - f(z)]| + |w_k(y)D^k[f(z) - P(z,\alpha_m)]|$$

$$+ |w_k(y)D^k[P(z,\alpha_m) - P(y,\alpha_m)]|$$

$$\leq WK_1(\delta_m)^\rho + e(T_m) + |w_k(y)D^k[P(z,\alpha_m) - P(y,\alpha_m)]| \tag{11}$$

where $W = \max_{(x,k)\varepsilon X \times I} |w_k(x)|$ and we use Remark 2.10.

Since $Q_o \subset T_m$ for all $m \geq n$, Corollary 1.15 implies that there exists a constant B such that $||\alpha|| < B$ for all $\alpha \varepsilon R(T_m)$ for any $m \geq n$. Therefore, since the base functions satisfy a Hölder condition with exponent ρ, there exists a constant K_2 such that the term on the right in line (11) is less than $WK_2(\delta_m)^\rho$. Thus $M[f(x) - P(x,\alpha_m)] \leq K(\delta_m)^\rho + e$ where $K = W(K_1 + K_2)$. Q. E. D.

2.9 <u>Lemma</u>. Assume that MCS(X×I) contains a set Q_1 of n+1 points and let $P(x,\alpha^*)$ be the unique best approximation to f(x) on X×I. If $P(x,\beta)$ is another approximation to f(x) with deviation $e_\beta = M[f(x) - P(x,\beta)]$ such that $|e_\beta - e| = \varepsilon$ then there exists a constant K" not dependent on ε such that

$$\underset{(x,k)\varepsilon X\times I}{\mathrm{Max}} |D^k[P(x,\beta) - P(x,\alpha^*)]| < K''\varepsilon.$$

<u>Proof</u>: Let $Q_1 \equiv \{(x_i,k_i,s_i)\}_{i=1}^{n+1}$ be the n+1 point set from MCS(X×I). Then for this set of points we have

$$s_i\ w_{k_i}(x_i)D^{k_i}[f(x_i) - P(x_i,\alpha^*)] = e \qquad (12)$$

$$s_i\ w_{k_i}(x_i)D^{k_i}[f(x_i) - P(x_i,\beta)] \leq e_\beta \qquad (13)$$

Subtracting (13) from (12) we have

$$s_i\ w_{k_i}(s_i)D^{k_i}[P(x_i,\beta) - P(x_i,\alpha^*)] \geq e - e_\beta \geq -\varepsilon$$

Then applying the results of Lemma 2.4 we have the existence of the constant K". Q. E. D.

2.10 <u>Corollary</u>. If in addition to the hypothesis of Theorem 2.8 there is a set Q_1 of n+1 points such that $Q_1 \varepsilon MCS(X\times I)$, then there exists a constant K' depending on $\{\phi_i(x)\}_{i=1}^n$, f(x) and $\{w_k(x)\}_{k\varepsilon I}$ such that

$$\max_{(x,k)\in X\times I} |D^k[P(x,\alpha^*) - P(x,\alpha_m)]| \leq K'(\delta_m)^\rho$$

where α_m is any best approximation in $R(T_m)$ and α^* is the unique best approximation in R.

Proof: Let the ε from Lemma 2.19 be equal to $K(\delta_m)^\rho$ from Theorem 2.8. Q. E. D.

2.11 Corollary. If the hypotheses of Theorem 2.8 are satisfied and if $\{P(x,\alpha_m)\}$ is any convergent sequence of polynomials with terms taken from $R(T_m)$, then $\lim_{m\to\infty} P(x,\alpha_m) = P(x,\alpha)$ is a best approximation such that $\alpha\varepsilon R$.

Proof: This follows directly from the continuity of $M[f(x) - P(x,\alpha)]$ in α. Q. E. D.

CHAPTER III

UNIQUENESS RESULTS

In this chapter we shall restrict our attention to the case where X is an interval [a,b] and I is {0,1}. We shall show that under appropriate hypotheses this problem has a unique solution. We begin with an example of nonuniqueness.

1. An Example of Nonuniqueness

In Example 1.33 we saw an approximation problem on a finite point set which did not have a unique solution. The following is an example of nonuniqueness for approximation on an interval.

3.1 Example. Let $X = [0,4]$, $I = \{0,1\}$, $\phi_i(x) = x^{i-1}$ for $i = 1,2,3$, $w_o(x) \equiv 1$ and let

$$w_1(x) = x \ , \ f(x) = x^2/2 - x \ , \ Df(x) = x - 1 \quad \text{for } 0 \le x \le 2$$

$$w_1(x) = 4 - x \ , \ f(x) = -x^2/2 + 3x - 4 \ , \ Df(x) = -x + 3 \quad \text{for } 2 \le x \le 4.$$

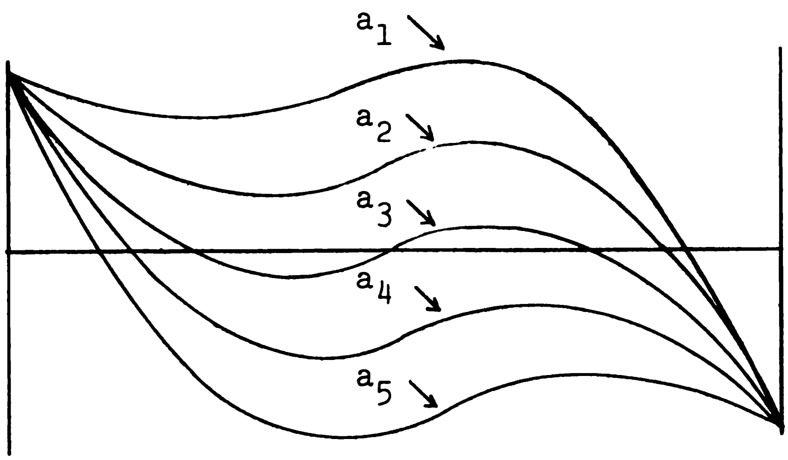
This example was constructed so that $x/2 - 1$ is a best approximation. To verify this, consider the weighted error curves:

$$L_o(x) = w_o(x)[f(x) - x/2+1] = \begin{cases} (x-1)(x-2)/2 & \text{if } 0 \le x \le 2 \\ -(x-2)(x-3)/2 & \text{if } 2 \le x \le 4 \end{cases}$$

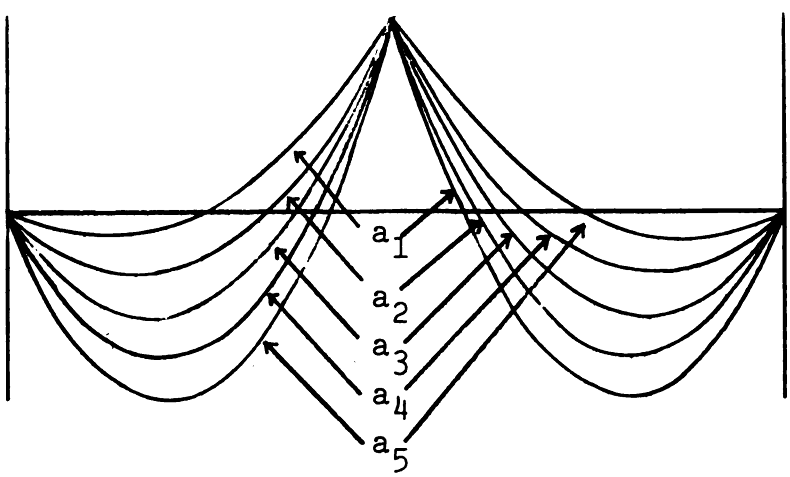
$$L_1(x) = w_1(x)[Df(x) - 1/2] = \begin{cases} x(x-3/2) & \text{if } 0 \le x \le 2 \\ (4-x)(5/2-x) & \text{if } 2 \le x \le 4 \end{cases}$$

It may be verified that the extreme values of $L_o(x)$ on [0,4] are 1 and -1 at x = 0 and x = 4 respectively. The extreme value of $L_1(x)$ is 1 for x = 2. Then the claim is that Q = {(0,0,1),(4,0,-1),(2,1,1)} is in MCS(X×I) and since the deviation of the approximation is not attained at any other point of X×I it follows that Q is equal to MES(X×I). Using Theorem 1.30 it must be shown that the origin is in the convex hull of the n-points of Q which are (1,0,0), (-1,-4,-16) and (0,1,4). Using the multiples 1/6, 1/6 and 2/3 respectively, we have the origin as a convex combination of these points. This shows that there is no polynomial satisfying Condition A with respect to Q. Then using the properties of the generalized Vandermonde matrix, any two point subset Q' of Q does have a polynomial satisfying condition A with Respect to Q'. Thus QϵMCS(X×I), e = 1, and x/2 - 1 is a best approximation.

Next we observe that ax(x-4) is the only family of polynomials in the base functions under consideration which is zero at the points of Q. Hence using Corollary 1.21, the space of best approximations is $P(x,a) \equiv x/2 - 1 + ax(x-4)$ where a is restricted to some interval containing zero. Then it may be verified that the space of best approximations is $P(x,a)$ for $a\varepsilon[-(2\sqrt{2} - 1)/8,(2\sqrt{2} - 1)/8]$. The weighted error curves are given in the following graphs for various values of a.



$L_o(x) = w_o(x)[f(x) - P(x,a)]$ for $a = a_i$, $i = 1,2,3,4,5$



$L_1(x) = w_1(x)[Df(x) - DP(x,a)]$ for $a = a_i$, $i = 1,2,3,4,5$

$a_1 = (2\sqrt{2} - 1)/8$, $a_2 = a_1/2$, $a_3 = 0$, $a_4 = -a_2$, $a_5 = -a_1$.

2. A General Uniqueness Theorem

We have shown by example that a solution to a problem of approximating a function and its derivative need not be unique. However, certain problems do have unique solutions. Such a unique solution can sometimes be detected using the results of Chapter I. For example if one has a solution to a given approximation problem and it has a MCS of n+1 points, then it follows from Theorem 1.41 that this solution is unique. A much more desirable result is one which can be used to claim uniqueness before a solution is found. Such a result will hold provided that $\{\phi_i(x)\}_{i=1}^{n}$ and $f(x)$ satisfy certain conditions. First we shall indicate conditions on the base functions which will be used in a uniqueness proof.

Throughout this chapter the notation $|S|$ will be used to indicate the cardinality of the finite point set S.

We shall first consider the special base functions $\phi_i(x) = x^{i-1}$ for $i = 1,2,\ldots,n$. We shall use the following properties of these base functions:

(a) $D^2\phi_i(x)$ exists on $[a,b]$ for $i = 1,2,\ldots,n$.

(b) $DP(x,\alpha) = \sum_{i=1}^{n} a_i D\phi_i(x)$ has at most n-2 zeros on $[a,b]$ if $DP(x,\alpha) \not\equiv 0$ on $[a,b]$.

(c) At most one base function is constant on $[a,b]$.

In the following two lemmas certain point sets of n and n-1 points are described which have interpolation matrices of rank n and n-1 respectively. A MES cannot have an interpolation matrix for which the rank is equal to the number of points. If it did, there would exist a polynomial satisfying Condition A with respect to this MES. Thus we are able to establish a lower bound on the number of points in a MES. This lower bound is essential to the proof of uniqueness theorems.

3.2 Remark. If it is desired to interpolate the function 0 at the points of a finite set $Q \subset X \times I$ by a polynomial in n base functions, a homogeneous system of $m = |Q|$ linear equations in n unknowns must be solved. From the theory of linear equations, the dimension of the solution space will be n-r where r is the rank of the interpolation matrix of Q in n base functions.

3.3 Definition. A polynomial $P(x,\alpha)$ has a zero at the point (x_o,k) or the point (x_o,k,s) if $D^kP(x_o,\alpha) = 0$.

3.4 Lemma. Given any set $T \subset [a,b]$ such that $|T| = n-1$, define $Q' = T \times \{1\}$. Then if $n > 1$ the interpolation matrix IM(Q') has rank n-1.

Proof: Assume that IM(Q') has rank $k \leq n - 2$. Consider the problem of interpolating the function 0 at the points of Q'. The dimension of the solution space is $n - k \geq 2$. Since at most one base function is constant on [a,b], there exists a polynomial $P(x,\alpha)$ such that $DP(x,\alpha) \not\equiv 0$ on [a,b] and $DP(x,\alpha)$ has n-1 zeros on [a,b]. This contradicts property (b) and completes the proof. Q. E. D.

3.5 Lemma. Let S,T' and T'' be finite sets such that $T'' \subset \{a,b\}$, $T' \subset S \subset [a,b]$ and $|S| + |T' \cup T''| = n$; also, define $T = T' \cup T''$ and $Q = S\times\{0\} \cup T\times\{1\}$. Then IM(Q), the interpolation matrix of Q, is nonsingular if $n > 2$ or S in nonempty.

Proof: First we observe that $n > 2$ implies that $|S| \geq 1$. Then assume that the interpolation matrix for Q is singular and that $n > 2$. This implies the existence of a nontrivial polynomial $P(x,\alpha)$ with zeros at the points of Q. Since T'' can only contain the end points of $[a,b]$, we may apply Rolle's Theorem to every adjacent pair of points from $S\times\{0\}$, if there are any, to conclude that $DP(x,\alpha)$ has at least $|S| - 1$ zeros on $[a,b]$ in addition to the $|T|$ zeros of $T\times\{1\}$. This implies that $DP(x,\alpha)$ has at least $|S| + |T| - 1 = |Q| - 1 = n - 1$ zeros which contradicts property (b) of the base functions under consideration. Thus $DP(x,\alpha) \equiv 0$ and since $|S| \geq 1$, it follows that $P(x,\alpha) \equiv 0$. Q. E. D.

In Corollary 1.21 we established that any two best approximations must agree at the points of a MES. In the following theorem we find that under the appropriate hypotheses the first derivatives of any two best approximations must also agree at the points of MES which are interior to $X\times I$.

3.6 <u>Theorem</u>. Let (x_o,k,s) be a point from MES(X×I) such that x_o is in the interior of $[a,b]$. Then if $Dw_k(x_o)$ and $D^{k+1}[f(x_o) - P(x_o\alpha)]$ exist for $\alpha\varepsilon E^n$ it follows that

$$D^{k+1}P(x_o,\alpha) = D^{k+1}f(x_o) + seDw_k(x_o)/[w_k(x_o)]^2$$

for all $\alpha\varepsilon R$.

<u>Proof</u>: For any $\alpha\varepsilon R$, $L_k(x_o,\alpha) = w_k(x_o)D^k[f(x_o) - P(x_o,\alpha)]$ has a relative extrema at x_o in the interior of $[a,b]$. Since the derivative of $L_k(x,\alpha)$ exists at x_o it must be zero there. Thus

$$Dw_k(x_o)D^k[f(x_o) - P(x_o,\alpha)] + D^{k+1}[f(x_o) - P(x_o,\alpha)]w_k(x_o) = 0$$

and since we assume that $e > 0$ it follows that $w_k(x_o) \neq 0$ and hence

$$D^{k+1}P(x_o,\alpha) = D^{k+1}[f(x_o)] + Dw_k(x_o)D^k[f(x_o) - P(x_o,\alpha)]/w_k(x_o)$$

$$= D^{k+1}f(x_o) + seDw_k(x_o)/[w_k(x_o)]^2$$

since $L_k(x_o,\alpha) = se$. Q. E. D.

The proof of the following uniqueness theorem follows the pattern of the proof used in the classical case of ordinary uniform approximation theory. If there are two best approximations, it follows from Corollary 1.21 and Theorem 3.6 that the difference of the two is a polynomial with a certain number of zeros. This number is dependent on the number of points in the MES. Under appropriate conditions this number of zeros is sufficient to insure that the difference polynomial is identically zero. This completes the proof of uniqueness. This same theorem is given in Moursund [16,17]. The proof here is considerably simplified, and easily extends to other cases of interest.

3.7 <u>Theorem</u>. Let $X = [a,b]$, $I = \{0,1\}$, and let $Dw_o(x)$, $Dw_1(x)$, and $D^2f(x)$ exist on $[a,b]$. If $\phi_i(x) = x^{i-1}$ for $i = 1,2,\ldots,n$ then one of the following is true:

(a) The best approximation is unique.

(b) The best approximation is unique except for an additive constant, and if $P(x,\beta)$ is any best approximation then $DP(x,\beta)$ is the unique best approximation to $Df(x)$ with weight function $w_1(x)$.

<u>Proof</u>: For notational convenience we define the following sets. $MES(X\times I)$ is the minimal extremal set of points for the best approximations on $X\times I$. Then let

$$Q_o = \{(x,k):(x,k,s)\varepsilon MES(X\times I)\}$$

$$S_o = Q_o \cap (a,b)\times\{0\}$$

$$S_a = Q_o \cap \{(a,0)\}$$

$$S_b = Q_o \cap \{(b,0)\}$$

$$T_o = Q_o \cap (a,b)\times\{1\}$$

$$T_a = Q_o \cap \{(a,1)\}$$

$$T_b = Q_o \cap \{(b,1)\}$$

$$W = \{(x,0):(x,0,s)\varepsilon S_o \text{ and } (x,1,s)\varepsilon T_o\}$$

$$S^* = S_o \sim W$$

The set S^* is the set of interior extrema from S_o which do not have the same x coordinate as the interior extrema from T_o.

Case I.--We first consider the special case when $n \leq 2$ and $S_o \cup S_a \cup S_b$ is empty. If $n = 1$ then MES(X×I) must contain at least one point since every approximation has at least one extremum. If $n = 2$ the set MES(X×I) must contain at least two points because Lemma 3.4 implies that there exists a polynomial satisfying Condition A with respect to any single point. Therefore $|T_o| + |T_a| + |T_b| \geq n$ and using Corollary 1.21 the difference between two best approximations is a nontrivial polynomial $P(x,\alpha)$, which is zero at the points of MES(X×I). Thus $DP(x,\alpha)$ has at least n zeros which implies that $DP(x,\alpha) \equiv 0$ using property (b) of the base functions. Thus the best approximations are unique except for an additive constant.

Case II.--The remaining case is when $n > 2$ or $S_o \cup S_a \cup S_b$ is not empty. We begin by showing that

$$2\,|T_o| + |T_a| + |T_b| + |S^*| + |S_a| + |S_b| = m > n. \quad (1)$$

We shall show that if (1) is not satisfied then there exists a polynomial satisfying Condition A with respect to MES(X×I).

Assume that $m \leq n$ and let $S' \subset X \sim \{x\varepsilon X:(x,k,s)\varepsilon MES(X\times I)\}$ such that $|S'| = n-m$. Also define

$$S = \{x:(x,k)\varepsilon T_o \cup S^* \cup S_a \cup S_b\} \cup S'$$

$$T' = \{x:(x,k)\varepsilon T_o\} \qquad T'' = \{x:(x,k)\varepsilon T_a \cup T_b\}.$$

Since all of the sets used in the definition of S, T' and T" are pairwise disjoint,

$$|S| = |T_o| + |S^*| + |S_a| + |S_b| + |S'|$$

$$|T'| = |T_o| \qquad\qquad |T''| = |T_a| + |T_a|$$

and it follows that the sets S, T' and T" satisfy the hypothesis of Lemma 3.5. This implies that IM(Q) is non-singular (the Q is from Lemma 3.5). Since $Q_o \subset Q$, the rank of $IM(Q_o)$ is equal to $|Q_o|$. Thus there exists a polynomial satisfying Condition A with respect to MES(X×I) and the truth of line (1) is verified by this contradiction.

Every best approximation must satisfy Corollary 1.21. Then if there is more than one best approximation, the difference between two of them is a nontrivial polynomial $P(x,\alpha)$ which is zero at every point of $MES(X\times I)$, or equivalently on Q_o. Moreover, since the hypothesis of Theorem 3.6 are satisfied for any (x_o,k,s) where x_o is interior to $[a,b]$, it follows that $DP(x,\alpha)$ must be zero at all points $(x,k)\varepsilon S^* \cup T_o$. Thus $P(x,\alpha)$ has double zeros at the points of T_o and single zeros at each point of $T_a \cup T_b$, which means that $DP(x,\alpha)$ has at least $2|T_o| + |T_a| + |T_b|$ zeros. (See Definition 3.3.) The polymonial $DP(x,\alpha)$ also has zeros at the points of S^*. Then because of the manner in which S^* was defined we may conclude that $DP(x,\alpha)$ must have at least $m' = 2|T_o| + |T_a| + |T_b| + |S^*|$ zeros. From line (1) we see that $m' + |S_a| + |S_b| > n$, and since $|S_a| + |S_b| \leq 2$ it follows that $m' > n-2$. This contradicts property (b) of the polynomials under consideration and implies that $DP(x,\alpha) \equiv 0$. Therefore, the first derivatives of any two best approximations are equal and the best approximations may differ by a constant. If $S_o \cup S_a \cup S_b$ is not empty, $P(x,\alpha)$ (which is a constant) must have at least one zero) and hence $P(x,\alpha) \equiv 0$ and the best approximation is unique. Q. E. D.

In the next theorem, a class of polynomials is described which has the same three properties that were used to prove uniqueness in Theorem 3.7. Since these three properties are the only properties of the base functions that were used in the proof, Theorem 3.7 and its proof are also valid for this class of base functions.

3.8 Theorem. Let $g(x)$ satisfy the following conditions:

(a) $g(x)$ is strictly monotone on $[a,b]$.

(b) $Dg(x)$ has no zeros on $[a,b]$

(c) $D^2g(x)$ exists on $[a,b]$

Then letting

$$P(x,\alpha) = \sum_{j=1}^{n} a_j[g(x)]^{j-1}$$

it follows that $DP(x,\alpha)$ can have at most $n-2$ zeros on $[a,b]$, where zeros of multiplicity 2 or more are counted as two zeros.

Proof: Assume that $DP(x,\alpha)$ has k zeros of multiplicity 1 denoted by $\{x_i\}_{i=1}^{k}$ and m zeros of multiplicity 2 or more denoted by $\{x_i\}_{i=k+1}^{k+m}$, where $k + 2m > n - 2$. Then

$$0 = DP(x_i,\alpha) = Dg(x_i)\sum_{j=2}^{n} a_j(j-1)[g(x_i)]^{j-2}$$

for $i = 1,2,\ldots,k+m$, and

$$0 = D^2P(x_i,\alpha) = [Dg(x_i)]^2\sum_{j=3}^{n} a_j(j-1)(j-2)[g(x_i)]^{j-3}$$

$$+ D^2g(x_i)\sum_{j=2}^{n} a_j(j-1)[g(x_i)]^{j-2}$$

for $i = k+1,\ldots,k+m$.

Since $Dg(x)$ is nonzero on $[a,b]$ it follows that

$$0 = Q(y_i) \equiv \sum_{i=2}^{n} a_j(j-1)(y_i)^{j-2}$$

for $y_i = g(x_i)$, $i = 1,\ldots,k+m$, and

$$0 = Q_1(y_i) \equiv \sum_{j=3}^{n} a_j(j-1)(j-2)(y_i)^{j-3}$$

for $y_i = g(x_i)$, $i = k+1,\ldots,k+m$.

Moreover, $g(x)$ is strictly monotone on $[a,b]$, and hence the numbers $\{y_i\}_{i=1}^{k+m}$ are all distinct. Thus $Q(y)$, which is a polynomial in y of degree $n-2$, has at least $n-1$ zeros. This is a contradiction and completes the proof.

3.9 Corollary. If $g(x)$ satisfies the hypothesis of Theorem 3.8, Theorem 3.7 is true for the base functions

$$\phi_i(x) = [g(x)]^{i-1} \quad \text{for } i = 1,2,\ldots,n$$

Proof: These base functions obviously satisfy properties (a) and (c) listed at the beginning of this section. Moreover, the proof of Theorem 3.7 did not consider zeros of any derivative higher than the second and hence these base functions also satisfy property (b) as it was used in the proof. Since properties (a), (b), and (c) are the only properties of the base functions used in the proof of Theorem 3.7 and the lemmas leading up to it, the proof of this corollary is complete.

3. The Trigonometric Polynomials

In this section on trigonometric polynomials, X will be the half open interval $[0,2\pi)$ where the topology of X is the ordinary topology on $[0,2\pi]$ with the points 0 and 2π being identified with each other. The polynomials will be represented by

$$P(x,\alpha) = a_1 + \sum_{i=1}^{k} a_{2i}\cos(ix) + a_{2i+1}\sin(ix).$$

$P(x,\alpha)$ will be called a polynomial of degree k if $|a_{2k}| + |a_{2k+1}| > 0$. It is well known that a trigonometric polynomial of degree $k > 0$ has at most $2k$ zeros counting multiplicity, on the interval X. Thus one can easily show:

3.10 <u>Lemma</u>. Given sets S and T such that $T \subset S \subset X$ and $|S| + |T| = 2k+1$, the interpolation matrix $IM(Q)$ of $Q = S\times\{0\} \cup T\times\{1\}$ in $n = 2k+1$ base functions is nonsingular.

3.11 <u>Theorem</u>. Let $X = [0,2\pi)$, $I = \{0,1\}$, and let $Dw_o(x)$, $Dw_1(x)$, and $D^2f(x)$ exist on X; then if the approximating polynomials are the trigonometric polynomials of degree k, one of the following is true:

(a) The best approximation in unique.

(b) The best approximation is unique except for an additive constant and if $P(x,\beta)$ is any best approximation, then $DP(x,\beta)$ is the unique best approximation to $Df(x)$ with weight function $w_1(x)$.

The proof of this theorem is almost identical to that of Case II of Theorem 3.7. Using the notation introduced there, the sets S_a, S_b, T_a, and T_b are empty (and hence not used) since X has no end points. The two differences between these two proofs are as follows:

(1) Lemma 3.10 is used here instead of Lemma 3.5.

(2) $DP(x,\alpha)$ in n base functions can have at most $n-1 = 2k$ zeros for trigonometric polynomials of degree k instead of the n-2 zeros as assumed in property (b) of section (2).

4. A Special Class of Polynomials

This section will deal with the special case where $\phi_i(x) = x^{i+k}$ for $i = 1,2,\ldots,n$ where $k \geq 0$ is a fixed integer. Let $I = \{0,1\}$ and $X = [a,b]$ where $0 < a < b$. (Results similar to those obtained in this section hold if $a < b < 0$.) The functions $D^2f(x)$, $Dw_o(x)$ and $Dw_1(x)$ will be assumed to exist on $[a,b]$. It should be noted that both $P(x,\alpha)$ and $DP(x,\alpha)$ in these base functions can have at most n-1 zeros in the interval $[a,b]$. This is because $P(x,\alpha)$ is a polynomial of degree n+k with k+1 zeros at $x = 0$ and $DP(x,\alpha)$ is a polynomial of degree n+k-1 with k zeros at $x = 0$. From this property of the base functions under consideration we have the following lemma.

3.12 <u>Lemma</u>. Given sets S and T' such that $T' \subset S \subset [a,b]$ and a set T'' which is empty or is equal to $\{b\}$ where $|S| + |T' \cup T''| = n$, define $T = T' \cup T''$ and $Q = S\times\{0\} \cup T\times\{1\}$. Then $IM(Q)$ is nonsingular.

<u>Proof</u>: Assume that $IM(Q)$ is singular. Then the problem of interpolating 0 at the points of Q has a nontrivial solution $P(x,\alpha)$. If $|S| + |T'| = n$, this polynomial $P(x,\alpha)$ has n zeros in $[a,b]$ counting multiplicity. This implies that $P(x,\alpha) \equiv 0$ and hence $IM(Q)$ must be nonsingular. If $|S| + |T'| < n$ it follows that T'' is not empty. Then applying Rolle's Theorem to $P(x,\alpha)$ on the whole real line we find that $DP(x,\alpha)$ has at least k zeros at $x = 0$ and at least $n-1$ more zeros in $(0,b)$. The zeros counted cannot include the zero at the point b since b is the furthest point from $x = 0$ of all points in $[a,b]$. Thus $DP(x,\alpha)$ has at least $k+(n-1)+1 = k+n$ zeros which means that $DP(x,\alpha) \equiv 0$. Moreover, $P(x,\alpha)$ has a zero at $x = 0$ so $P(x,\alpha) \equiv 0$ and $IM(Q)$ must also be nonsingular in this case. Q. E. D.

3.13 <u>Lemma</u>. Let $t_1,t_2,\ldots,t_m$ be m distinct points in (a,b). Suppose $D^3g(x)$ exists on $[a,b]$ and $g(a) = g(b) = 0$, and $Dg(t_i) = D^2g(t_i) = 0$ for $i = 1,2,\ldots,m$. Then there exists at least $2m+1$ zeros of $Dg(x)$ in (a,b) counting multiplicity.

<u>Proof</u>: If $Dg(t_i) = D^2g(t_i) = D^3g(t_i) = 0$ for some i, the conclusion is true. Therefore, assume that $D^3g(t_i) \neq 0$ for $i = 1,2,\ldots,m$. Then the points t_i for $i = 1,2,\ldots,m$ are relative maximum or minimum points for $Dg(x)$. If the $2m$ zeros which we assume for $Dg(x)$ are the only ones that $Dg(x)$ has in (a,b), it follows that $Dg(x)$ is either nonnegative or nonpositive on $[a,b]$. Hence $g(x)$ is a monotone function on $[a,b]$ and $g(x) \equiv 0$. Here we have a contradiction and the proof is complete.

3.14 <u>Theorem</u>. Under the assumptions of this section and the assumption that $w_1(a) = 0$, the best approximation is always unique.

The assumption that $w_1(a) = 0$ implies that the points $(a,1,1)$ and $(a,1,-1)$ cannot be in the set $MES(X\times I)$. This restriction of the points of $MES(X\times I)$ is necessary in order that the method of proof used in Theorem 3.7 may also be used here.

The proof of this theorem is almost identical to that of Case II of Theorem 3.7. Using the notation introduced there, the set T_a is empty (and hence not used) since $w_1(a) = 0$. The three differences are as follows:

(1) Lemma 3.12 is used here instead of Lemma 3.5.

(2) In counting the zeros that $DP(x,\alpha)$ must have we use Lemma 3.13.

(3) $DP(x,\alpha)$ in n base functions can have at most $n-1$ zeros for the polynomials under

consideration rather than the n-2 zeros as assumed in property (b) of section 2.

Finally, since there is no constant among the base functions, two best approximations cannot differ by a constant. Thus $DP(x,\alpha) \equiv 0$ implies that $P(x,\alpha) \equiv 0$.

CHAPTER IV

COMPUTATIONAL ALGORITHMS FOR A BEST APPROXIMATION

1. Introduction

The most basic tool in the computation of uniform approximations is linear programming, which can be used to find best approximations on finite point sets. The approximation problem under consideration may be stated as a linear programming problem as follows: find a point $\alpha = (a_1, a_2, \ldots, a_n)$ in E^n such that

$$|w_k(x)D^k[f(x) - \sum_{i=1}^{n} a_i\phi_i(x)]| \leq e \quad \text{for } (x,k)\varepsilon X\times I \qquad (1)$$

where e is to be a minimum. This is equivalent to the problem:

maximize -e subject to the constraints

$$\left.\begin{aligned} w_k(x)D^k f(x) &\geq -e + w_k(x)D^k P(x,\alpha) \\ -w_k(x)D^k f(x) &\geq -e - w_k(x)D^k P(x,\alpha) \end{aligned}\right\} \text{for } (x,k)\varepsilon X\times I.$$

Since (1) implies that e is nonnegative, the quantity -e is bounded above. Hence when applied to an approximation problem on a finite set $T \subset X\times I$, the simplex algorithm for solving a linear programming problem will terminate at a best approximation on T. (See Arden [1].) For a comprehensive study of linear programming, see Dantzig [6].

In Chapter II the de la Vallée Poussin Algorithm for finding a best approximation was given. In this algorithm each term of the sequence of polynomials which converges to a best approximation is a best approximation on some finite point set. Therefore, linear programming may be used to generate this sequence. This algorithm is acceptable if we do not demand an approximation which is very close to a best approximation. If a high degree of accuracy is desired, this algorithm is applicable but impracticle because of the large number of constraints that must be included in the linear program.

To eliminate the need for solving linear programming problems with large numbers of constraints, an algorithm based on the characterization of best approximations will be developed. In the case of the uniform approximation of a function by polynomials over a Chebyshev set of base functions, the Remez Algorithm is a very efficient method for generating a sequence of polynomials which converges to a best approximation. At each step of this algorithm a best approximation is computed on n+1 points. This particular linear programming problem can be reduced to solving a particular system of n+1 simultaneous linear equations in n+1 unknowns. Thus at no time in the process does one need to solve a linear programming problem involving a very large number of constraints.

In the more general problem to be considered here, which includes the weighted approximation of a function and its derivatives by polynomials in a general set of base functions, the direct extension of the Remez Algorithm does not necessarily converge to a best approximation. Moreover, the general n+1 point problem is not easily converted into a system of simultaneous linear equations except in certain special cases, (see Moursund and Stroud [18]). Hence we will rely heavily upon the standard linear programming algorithm, and develop an algorithm which uses the fact that the deviation of a best approximation on any compact set $T \subset X \times I$ is determined by any MCS in MES(T). (Recall that any MCS has at most n+1 points where n is the number of base functions.)

In the next section we shall show that with appropriate hypotheses, the deviation of a best approximation on a finite point set is a continuous function of the point set. Hence computationally we may look for a sequence of finite point sets which converge to a finite point set containing a MCS from MES($X \times I$) and then conclude that the deviation on these points sets will approach e. From this sequence of finite point sets we may determine an approximation which is close to some best approximation.

Throughout this chapter X will be a closed interval [a,b] where $a < b$ (except in Example 4.9 where X includes a finite point set in addition to an interval).

2. The Continuity of $e(T)$

The set $X \times I$ is a subset of E^2; therefore, the distance function $p[(x,k),(x',k')] = |x - x'| + |k - k'|$ is a metric on $X \times I$. Using this metric we define the distance from a point (x',k') to a compact set $T \subset X \times I$ to be

$$d'[(x',k'),T] = \min_{(x,k)\varepsilon T} p[(x',k'),(x,k)].$$

In order to measure the distance between two compact sets T and Q, with the requirement that the distance between two sets is zero if and only if they are identical, we define the following distance for compact subsets of $X \times I$.

$$d(T,Q) = \mathrm{Max}\{\max_{(x,k)\varepsilon Q} d'[(x,k),T], \max_{(x,k)\varepsilon T} d'[(x,k),Q]\}$$

It can be shown that this distance function is a metric on the compact subsets of $X \times I$. Moreover, the space of all compact subsets of $X \times I$ is a compact metric space using this metric. (See Michael [15].) Metrics such as this are commonly called Hausdorff metrics on subspaces of a metric space.

In this chapter the symbol T with or without a subscript or superscript will denote a compact subset of $X \times I$. The following definition concerning the convergence of sequences of subsets of $X \times I$ is the usual one for metric spaces.

4.1 Definition. A sequence $\{T_m\}$ converges to a set T_o if for each $\varepsilon > 0$ there exists an integer $N = N(\varepsilon)$ such that $m \geq N$ implies that $d(T_m, T_o) < \varepsilon$.

4.2 Lemma. Let the sequence $\{T_m\}$ converge to T_o and let S be a compact subset of E^n; then for any $\varepsilon > 0$ there exists an integer $N = N(\varepsilon)$ such that $m \geq N$ implies

$$|MT_m[f(x) - P(x,\alpha)] - MT_o[f(x) - P(x,\alpha)]| < \varepsilon$$

for all $\alpha \varepsilon S$.

Proof: The function $w_k(x)D^k[f(x) - P(x,\alpha)]$ is a continuous function of x,k, and α on $X \times I \times E^n$ and hence uniformly continuous on the compact set $X \times I \times S$. Thus given any $\varepsilon > 0$ there exists a $\delta(\varepsilon)$ such that $|x_1 - x_2| + |k_1 - k_2| < \delta$ implies that

$$|w_{k_1}(x_1)D^{k_1}[f(x_1) - P(x_1,\alpha)] - w_{k_2}(x_2)D^{k_2}[f(x_2) - P(x_2,\alpha)]| < \varepsilon$$

for (x_1,k_1) and (x_2,k_2) in $X \times I$ and for all $\alpha \varepsilon S$. Since $\{T_m\}$ converges to T_o there exists an integer $N = N(\delta)$ such that $m \geq N$ implies that $d(T_m,T_o) < \delta$. Therefore, it follows that

$$|MT_m[f(x) - P(x,\alpha)] - MT_o[f(x) - P(x,\alpha)]| < \varepsilon$$

for $m \geq N$ and $\alpha \varepsilon S$. Q. E. D.

4.3 <u>Definition</u>. Let $T \subset X \times I$ be a set of m points; then $Nb(T,\delta) \equiv \{Q: Q \subset X \times I$ has m points and $d(Q,T) \leq \delta\}$ is a neighborhood of T of radius δ.

4.4 <u>Lemma</u>. If $T \subset X \times I$ is a set of n points where $\det[WIM(T)] \neq 0$, there exists a $\delta > 0$ such that $Q \varepsilon Nb(T,\delta)$ implies $\det[WIM(Q)] \neq 0$.

<u>Proof</u>: Since a determinant is a continuous function of its elements and since all of the elements of WIM(Q) are continuous on $X \times I$, the quantity $\det[WIM(Q)]$ is a continuous function of the points of Q. Then if $\det[WIM(T)]$ is nonzero, it must be nonzero in some neighborhood of T. Q. E. D.

4.5 <u>Theorem</u>. Let $\{T_m\}$ be a sequence of compact subsets of $X \times I$ which converges to a compact set $T_o \subset X \times I$. If T_o contains a set of n points T' such that $\det[WIM(T')] \neq 0$ then $\lim_{m \to \infty} e(T_m) = e(T_o)$.

<u>Proof</u>: We know that $e(T_m) - e(T_o) \leq MT_m[f(x) - P(x,\alpha_o)] - e(T_o)$ where $\alpha_o \varepsilon R(T_o)$ the set of best approximations to f(x) on the set T_o. Then Lemma 4.2 implies that for any $\varepsilon > 0$ there exists an integer $N = N(\varepsilon)$ such that $m \geq N$ implies

$$MT_m[f(x) - P(x,\alpha_o)] - MT_o[f(x) - P(x,\alpha_o)] < \varepsilon.$$

Thus for $m \geq N$ we have $e(T_m) - e(T_o) < \varepsilon$.

The set T_o contains a set T' of n points such that $\det[WIM(T')] \neq 0$. Thus by Lemma 4.4 there exists a $\delta_o > 0$ such that $Q \varepsilon Nb(T',\delta_o)$ implies that $\det[WIM(Q)] \neq 0$.

Since $\{T_m\}$ converges to $T_o \supset T'$ there exists an integer $N_1 = N_1(\delta_o)$ such that $m \geq N_1$ implies that there exists a $Q_m \varepsilon Nb(T', \delta_o)$ such that $Q_m \subset T_m$. Then Lemma 1.11 implies that

$$MQ_m[P(x,\alpha_m)] \leq MT_m[P(x,\alpha_m)] \leq 2MT_m[f(x)] \leq 2M[f(x)] \quad (2)$$

where $\alpha_m \varepsilon R(T_m)$, the set of best approximations on the set T_m. Let $B_{ij}[WIM(Q)]$ denote the cofactor of the entry in the ith row and jth column of $WIM(Q)$ and define

$$B = \max_{\substack{Q \varepsilon Nb(T'_o,\delta_o) \\ i = 1,2,\ldots,n \\ j = 1,2,\ldots,n}} |B_{ij}[WIM(Q)]| \qquad E = \min_{Q \varepsilon Nb(T'_o,\delta_o)} |\det[WIM(Q)]|. \qquad (3)$$

Since the coefficients of a polynomial $P(x,\alpha)$ are uniquely determined by the values it has on any point set $Q \equiv \{(x_i,k_i)\}_{i=1}^{n}$ in $Nb(T'_o,\delta_o)$, these coefficients are given by

$$a_j = \frac{\sum_{i=1}^{n} B_{ij}[WIM(Q)] \; w_{k_i}(x_i) \; D^{k_i} \; P(x_i,\alpha)}{\det[WIM(Q)]}, \quad j = 1,2,\ldots,n.$$

Then if $\alpha_m \varepsilon R(T_m)$ for $m \geq N_1$ it follows from lines (2) and (3) that

$$|a_{m_j}| \leq 2nB \cdot M[f(x)]/E, \qquad j = 1,2,\ldots,n$$

where $\alpha_m = (a_{m_1}, a_{m_2}, \ldots, a_{m_n})$. Let S be the set of all $\alpha \varepsilon E^n$ for which $||\alpha|| \leq 2nB \cdot M[f(x)]/E$. It follows then from Lemma 4.2 that for any $\varepsilon > 0$ there exists an integer $N_2 = N_2(\varepsilon)$ such that $m \geq N_2$ implies

$$MT_o[f(x) - P(x,\alpha)] - MT_m[f(x) - P(x,\alpha)] < \varepsilon$$

for all $\alpha \varepsilon S$ (which includes all α_m for $m > N_1$). Thus

$$e(T_o) - e(T_m) \leq MT_o[f(x) - P(x,\alpha_m)]$$

$$- MT_m[f(x) - P(x,\alpha_m)] < \varepsilon$$

for $m \geq N_2$. Q. E. D.

4.6 <u>Corollary</u>. Under the hypotheses of Theorem 4.5 there exists an integer N such that $R(T_m)$ is contained in some compact set S for all $m \geq N$.

4.7 <u>Corollary</u>. Let $\{T_m\}$ converge to a set T'_o. Then $\lim_{m\to\infty} e(T_m) \leq e(T'_o)$ if the limit exists.

This corollary follows directly from the first part of the proof of Theorem 4.5.

4.8 <u>Corollary</u>. Suppose T' is a set of n points for which $\det[WIM(T')] \neq 0$. If $T' \subset T_m$, for $m = 1,2,\ldots.$, and if $\lim_{m\to\infty} T_m = T_o$, then

$$\lim_{m\to\infty} e(T_m) = e(T_o).$$

The following is an example where e(T) is not a continuous function of T.

4.9 Example. Let $X = \{-1,-1/\sqrt{3},0,1/\sqrt{3}\} \cup [4/5,1]$, $n = 4$, $I = \{0,1\}$, $\phi_i(x) = x^{i-1}$ for $i = 1,2,3,4$ and $w_0(x) \equiv w_1(x) \equiv 1$. The function $f(x)$ is defined by the following table where k is to be determined later.

x	$f(x)$	$Df(x)$
-1	-1	0
$-1/\sqrt{3}$	0	-1
0	1	0
$1/\sqrt{3}$	0	1
$[4/5,1]$	k	0

Also let $T_0 = \{(1,0),(0,0),(-1,0),(1/\sqrt{3},1),(-1/\sqrt{3},1)\}$ and $T_i = \{(1-1/i,0),(0,0),(-1,0),(1/\sqrt{3},1),(-1/\sqrt{3},1)\}$ for $i \geq 5$. Then $\{T_i\}$ converges to T_0.

The value of $e(T_i)$ on any set T_i for $i \geq 5$ can be determined from the following system of equations where we let $t_i = 1 - 1/i$.

$$\begin{pmatrix} 1 & t_i & (t_i)^2 & (t_i)^3 & s_1 \\ 1 & -1 & 1 & -1 & s_2 \\ 1 & 0 & 0 & 0 & s_3 \\ 0 & 1 & 2/\sqrt{3} & 1 & s_4 \\ 0 & 1 & -2/\sqrt{3} & 1 & s_5 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ e_i \end{pmatrix} = \begin{pmatrix} k \\ -1 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

In this system of equations the quantities s_j for $j = 1,\ldots,5$ are to be determined to minimize the absolute value of e_i where s_j may be chosen to be 1,-1, or 0. Using Cramer's rule

$$e_i = \frac{kB_1 - B_2 + B_3 + B_4 - B_5}{\sum_{j=1}^{5} s_j B_j}$$

where B_j is the cofactor of s_j and it is noted that $B_j \neq 0$ for $j = 2,3,4,5$. Then if we choose $s_j = \mathrm{sgn}(B_j)$ it follows that $e(T_i) = |e_i|$ will be a minimum. Moreover, $B_1 = 0$ and hence $e(T_i)$ is independent of k. Therefore we may conclude that $e(T_i) \leq 1$ for all $i \geq 5$ and any k.

Next consider the approximation problem on the set T_o. Since the polynomials do not satisfy the requirement that $MT_o[P(x,\alpha)] = 0$ implies $||\alpha|| = 0$ we may reduce the number of base functions by at least 1. Since $x = x^3$ at the points of T_o we will use the base functions $\{1,x,x^2\}$. Consider the subset $T' = \{(1,0),(-1,0),(1/\sqrt{3},1),(-1/\sqrt{3},1)\}$ of T_o. Then $e(T_o) \geq |e'|$ where e' is given by

$$\begin{pmatrix} 1 & 1 & 1 & s_1 \\ 1 & -1 & 1 & s_2 \\ 0 & 1 & 2/\sqrt{3} & s_3 \\ 0 & 1 & -2/\sqrt{3} & s_4 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ e' \end{pmatrix} = \begin{pmatrix} k \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

Using Cramer's rule and the same notation and arguments as above, we have

$$e' = \frac{kB_1 - B_2 + B_3 + B_4}{\sum_{j=1}^{4} s_j B_j} .$$

Since $B_1 = 4/\sqrt{3}$ we may make $|e'|$ and hence $e(T_o)$ as large as we desire by choosing k properly. Thus $e(T)$ is not continuous at the point set T_o.

3. Two Algorithms

In this section we shall discuss two algorithms which can be implemented on a computer, and which (theoretically) can be used to solve the interval version of the problem of this thesis. Each algorithm involves the determination of a sequence of point sets, and uses linear programming to calculate best approximations to $f(x)$ on the finite point sets.

Since the continuity of $e(T)$ depends on the presence of a set T' of n points which has a weighted interpolation matrix of rank n, we must be able to determine such a set T' for use in the computational algorithm. (Such a set T' exists because of the restrictions on the base functions and Lemma 1.8.) In problems of computational interest such a set may be found quite easily. For example, if the base functions $\{\phi_i(x)\}_{i=1}^{n}$ are a Chebyshev set on $[a,b]$, then any n points from $[a,b]\times\{0\}$ for which $w_k(x)$ is nonzero will have a nonsingular weighted interpolation matrix. Also, from the continuity of det[WIM(T')] in the points of T' we know that if any non-trivial polynomial $P(x,\alpha)$ in the base functions $\{\phi_i(x)\}_{i=1}^{n}$ can have at most a finite number of zeros on X, then in any neighborhood of a set of points from $[a,b]\times\{0\}$ for which $w_k(x)$ is nonzero, there will be a set T' for which WIM(T') is nonsingular. Moreover, it should be noted that such a set T' is independent of the function f(x) being approximated and hence if a number of functions are to be approximated on X×I with the same base functions and weight functions, a set T' may be determined once and used for all functions being approximated.

The convergence of the following algorithms depends on the existence and use of this point set T'.

<u>Algorithm I</u>

Let T' be a set of n points such that det[WIM(T')] $\neq$ 0 and let $\varepsilon_o \geq 0$ be some fixed real number. (In the theorems which follow we shall show that the use of $\varepsilon_o = 0$ will give a convergent computational scheme in certain special problems. In actual computational practice, most problems fall into this class. The use of $\varepsilon_o > 0$ gives a slightly different algorithm which will tend to be computationally slower than the $\varepsilon_o = 0$ algorithm. However, we shall show that the algorithm converges in all cases if $\varepsilon_o > 0$.) Then choose some finite point set T''_o of at least n+1 points from X×I such that $e(T''_o) > 0$. (This set T''_o is usually chosen to be a set of equally spaced points from each set $X\times\{k\}$ for $k\varepsilon I$ which includes the points (a,k) and (b,k) for each $k\varepsilon I$. If the first choice of T''_o has a deviation $e(T''_o) = 0$, then additional points must be chosen from $X\times I$ and included in this set T''_o until the deviation is non-zero. Such sets T''_o such that $e(T''_o) > 0$ exist because we are only considering problems for which any best approximation on $X\times I$ has a deviation $e > 0$.) Next determine a best approximation $P(x,\alpha_o)$ to $f(x)$ on T''_o. (This may be done using linear programming.) Let Q_o be a subset of the extremal point set $C(T''_o,\alpha_o)$ such that Q_o contains some minimal characterization set from $MCS(T''_o)$ and let $T_o = \{(x,k):(x,k,s)\varepsilon Q_o\}$. (The set Q_o and the sets Q_m for any integer $m \geq 1$ given below are easily determined from the linear program by taking the points of $C(T''_o,\alpha_o)$ or $C(T''_m,\alpha_m)$ respectively which correspond to the slack variables which are set equal to zero to obtain a solution of the linear programming problem. The set Q_o or Q_m will usually include all the points of $C(T''_o,\alpha_o)$ or $C(T''_m,\alpha_m)$ respectively.) Then let $t = 0$ and define the sequence of point sets $\{T_m\}$ and the sequence of approximations $\{P(x,\alpha_m)\}$ recursively as follows ($m = 0$ initially). Obtain a point $T'_m = \{(x_m,k_m)\}$ where $|L_k(x,\alpha_m)|$ attains its maximum on $X\times I$. Let

$$T''_{m+1} = \begin{cases} T_m \cup T'_m & \text{if } t = 0 \\ T_m \cup T' \bigcup_{i=m-t}^{m} T'_i & \text{if } t > 0 \end{cases}$$

and find a best approximation $P(x,\alpha_{m+1})$ to $f(x)$ on T''_{m+1} (again we may use linear programming). Let Q_{m+1} be a subset of the extremal point set $C(T_{m+1},\alpha_{m+1})$ such that Q_{m+1} contains some minimal characterization set from $MCS(T''_{m+1})$ and let $T_{m+1} = \{(x,k):(x,k,s)\varepsilon Q_{m+1}\}$. Then if $e(T_{m+1}) \geq e(T_m) + \varepsilon_o$ set $t = 0$ and if $e(T_{m+1}) < e(T_m) + \varepsilon_o$ increase t by 1.

Algorithm II

This algorithm differs from Algorithm I only in the determination of the set T'_m which is defined as follows. For each approximation $P(x,\alpha_m)$ we define the set

$$B = \{(x,k)\varepsilon X\times I: |L_k(x,\alpha_m)| \geq e(T_m)\}$$

The set B consists of at most a countable number of closed intervals (allowing single points as degenerate intervals) from $X\times I$. Then combine the intervals from B into larger sets $J_1,J_2,\ldots,J_j$ so that this collection of sets $\{J_i\}$ satisfies the following requirements. Let J^* be any fixed element in $\{J_i\}$, then

(a) only one of the following is true for all $(x,k)\varepsilon J^*: L_k(x,\alpha_m) \geq e(T_m)$ or $L_k(x,\alpha_m) \leq - e(T_m)$,

(b) $\{k:(x,k)\varepsilon J^*\}$ contains only one integer,

(c) if (x,k) and (y,k) are any two points from J^* such that $x < y$ then there is no $(z,k)\varepsilon B$ such that $x < z < y$ and $L_k(x,\alpha_m) \leq - se(T_m)$ where $s = \mathrm{sgn}[L_k(x,\alpha_m)]$,

(d) if J' is any element from $\{J_i\}$ which is adjacent to J^* then $\text{sgn}[L_k(x',\alpha_m)] = -\text{sgn}[L_k(x^*,\alpha_m)]$ where (x',k) J' and $(x^*,k)\varepsilon J^*$.

Since the function $L_k(x,\alpha_m)$ is continuous and $e(T_m) > 0$ the collection of sets $\{J_i\}$ consists of at most a finite number of subsets of $X\times I$. Then from each set J_i, pick out a point where $|L_k(x,\alpha_m)|$ attains its maximum value on J_i and let T'_m consist of all these points. (In computational practice the set T'_m can usually be defined to be the set of all points (x,k) which are relative maxima of $|L_k(x,\alpha_m)|$ over $X\times I$ for which $|L_k(x,\alpha_m)| \geq e(T_m)$. The only difficulty that this could present is that T'_m might contain an infinite number of points, or T'_m might contain too many points for efficient computation of a best approximation on a finite point set.)

4. Convergence of the Algorithms

The first convergence theorem says that for any approximation problem under consideration in this chapter, the polynomial $P(x,\alpha_m)$ as generated by either of the two algorithms with $\varepsilon_o > 0$ will be close to a best approximation if m is taken large enough. First we observe that the proof of the following lemma follows directly from the compactness of $X\times I$.

4.10 Lemma. Let $\{(x_m,k_m)\}_{m=1}^{\infty}$ be a sequence of points from $X \times I$ and let $Q_m = \bigcup_{i=1}^{m-1} (x_i,k_i)$. Then given any $\varepsilon > 0$ there exists an integer $N = N(\varepsilon)$ such that $m \geq N$ implies that $d[(x_m,k_m),Q_m] < \varepsilon$.

4.11 Theorem. Let $\varepsilon_o > 0$ for Algorithm I be fixed and let T' be a set of n points such that $\det[WIM(T')] \neq 0$. Then given any $\varepsilon > 0$ and any finite point set T''_o for which $e(T''_o) > 0$, there exists an integer N depending on $\varepsilon_o, \varepsilon, T'$ and T''_o such that if Algorithm I is followed starting with T''_o, it follows that

$$M[f(x) - P(x,\alpha_m)] - e \leq \varepsilon \quad \text{for } m \geq N,$$

and

$$e - e(T_m) \leq \varepsilon \quad \text{for } m \geq N.$$

Proof: The sequence $\{e(T_m)\}$ is monotone increasing and bounded above and hence converges to some real number $e' \leq e$. Let m_o be some integer for which $e' - e(T_{m_o}) < \varepsilon_o$. Then it follows that $T' \subset T''_m \subset T''_{m+1}$ for all $m > m_o$. Since $T' \subset T''_m$ for $m > m_o$, Corollary 1.15 implies that there exists a compact set $S \subset E^n$ such that $R(T_m) \subset S$ for all $m > m_o$. Thus the weighted error functions $L_k(x,\alpha)$ defined for all $k \varepsilon I$, are uniformly continuous in α and x for $\alpha \varepsilon S$ and $x \varepsilon X$. Therefore, given any $\varepsilon > 0$ there exists a δ such that $|L_k(x_1,\alpha) - L_k(x_2,\alpha)| < \varepsilon$ for all $\alpha \varepsilon S$, $k \varepsilon I$, and $x_1,x_2 \varepsilon X$ satisfying $|x_1 - x_2| < \delta$.

Then from Lemma 4.10 there exists an integer $N = N(\delta)$ such that

$$d[(x_m,k_m),T''_m] < \delta \quad \text{for } m \geq N.$$

Thus for some $(x',k')\varepsilon T''_m$ we have

$$|L_{k_m}(x_m,\alpha_m)| - |L_{k'}(x',\alpha_m)| \leq |L_{k_m}(x_m,\alpha_m) - L_{k'}(x',\alpha_m)| < \varepsilon.$$

Since $|L_{k'}(x',\alpha_m)| \leq e(T_m) \leq e$ and

$$e \leq M[f(x) - P(x,\alpha_m)] = |L_{k_m}(x_m,\alpha_m)|$$

we have $M[f(x) - P(x,\alpha_m)] - e < \varepsilon$ and $e - e(T_m) < \varepsilon$ for all $m \geq N$. Q. E. D.

The above theorem and proof are easily extended to Algorithm II without any significant changes.

The second convergence theorem is one which is applicable in certain special cases. It states that for certain problems the two algorithms will converge to a best approximation with $\varepsilon_o = 0$ if we start with a point set T''_o for which $e(T''_o)$ is close enough to e. We shall first prove three lemmas to simplify the proof of the theorem.

4.12 Lemma. Let $P(x,\alpha)$ be a best approximation to $f(x)$ on $T \subset X \times I$ where MCS(T) contains a set of n+1 points. If $|L_{k_o}(x_o,\alpha)| > e(T)$ for some $(x_o,k_o)\varepsilon X\times I$, then $e(T \cup \{(x_o,k_o)\}) > e(T)$.

Proof: Assume that $e(T \cup \{(x_o,k_o)\}) = e(T)$. Let α' be any best approximation in $R(T \cup \{(x_o,k_o)\})$; then α' is also in $R(T)$. Theorem 1.41 implies that $R(T)$ contains one unique best approximation and thus that $\alpha = \alpha'$. This contradicts the assumption that $|L_{k_o}(x_o,\alpha)\} > e(T)$. Q. E. D.

4.13 Lemma. Let $\{f_m(x)\}_{i=1}^{\infty}$ be a uniformly convergent sequence of continuous functions on a compact set X which converges to $f(x)$ and for which there exists a sequence $\{x_m\}_{m=1}^{\infty}$ converging to x_o and satisfying $f_m(x_m) = \underset{x\varepsilon X}{\text{Max}}|f_m(x)|$. Then $\underset{x\varepsilon X}{\text{Max}}|f(x)| = |f(x_o)|$.

Proof: For each m and any $x\varepsilon X$, $|f_m(x)| \leq |f_m(x_m)|$. Then using the triangle inequality we have

$$|f_m(x)| \leq |f_m(x_m)| \leq |f(x_o)| + |f_m(x_m) - f_m(x_o)| + |f_m(x_o) - f(x_o)|$$

for any $x\varepsilon X$. Then taking the limit as $m\to\infty$ we have $|f(x)| \leq |f(x_o)|$ for all $x\varepsilon X$. (In taking this limit we must use the fact that a uniformly convergent sequence of continuous functions is equicontinuous. See Goffman [11], p. 106.) Q. E. D.

4.14 Lemma. Let $\lim_{m\to\infty} T_m = T_o$, $\lim_{m\to\infty} e(T_m) = e(T_o)$ and $\lim_{m\to\infty} \beta_m = \beta_o$ where $\beta_m \varepsilon R(T_m)$ for all m. Then $\beta_o \varepsilon R(T_o)$.

Proof: Since $e(T_m) = MT_m[f(x) - P(x,\beta_m)]$ for all m, it follows that

$$MT_m[f(x) - P(x,\beta_o)] \leq e(T_m) + M[P(x,\beta_m) - P(x,\beta_o)]. \quad (4)$$

Taking the limit of both sides of (4) as $m\to\infty$, $MT_o[f(x) - P(x,\beta_o)] \leq e(T_o)$ which implies that $\beta_o \varepsilon R(T_o)$.

4.15 Theorem. Suppose that in a particular problem of approximating a function $f(x)$ on $X\times I$ all of the point sets in $MCS(X\times I)$ contain $n+1$ points. Then there exists a constant $B < e$ such that if T''_o is any finite subset of $X\times I$ for which $e(T''_o) > B$ and if either Algorithm I or Algorithm II is applied with $\varepsilon_o = 0$ beginning with the set T''_o, it follows that each approximation $P(x,\alpha_m)$ will be the unique best approximation on T''_m (and hence also on T_m) and $\{P(x,\alpha_m)\}$ will converge to the best approximation on $X\times I$.

Proof: First we shall show that there exists a constant $B < e$ such that if T is a finite subset of $X\times I$ and $e(T) > B$, then every set in $MCS(T)$ contains $n+1$ points. If this were not so, then for each integer $m \geq 1$ there exists a set of $k_m < n+1$ points T_m such that $e(T_m) \geq e - 1/m$. Then $\{T_m\}$ has a convergent subsequence, also denoted $\{T_m\}$, converging to a set T_o. But Corollary 4.7 implies that $\lim_{m\to\infty} e(T_m) \leq e(T_o)$ and hence that $e = e(T_o)$ since $\lim_{m\to\infty} e(T_m) = e$. Since $e(T_o) = e$ it follows that $R \subset R(T_o)$ and since the sets in $MCS(X\times I)$ contain $n+1$ points, the set R contains one unique best approximation α. Thus

$$MES(T_o) \subset C(T_o,\alpha) \subset C(\alpha) = MES(X\times I)$$

and hence any MCS in MCS(T_o) is also in MCS(X×I). But the sets in MCS(T_o) can consist of at most n points since T_o contains at most n points. This contradicts the assumption that the sets from MCS(X×I) all contain n+1 points.

The set T' from the algorithms need not be considered in this proof since $\varepsilon_o = 0$ and hence t = 0 throughout the algorithm. Assume that $e(T''_o) > B$. Since $e(T_m) \geq e(T''_o)$ for all m, the sets in MCS(T''_o) all consist of n+1 points and hence by Theorem 1.41 each α_m is the unique best approximation on the set T_m for $m \geq 0$. From the three sequences $\{P(x,\alpha_m)\}$, $\{T_m\}$, and $\{T'_m\}$ which are generated by the algorithm, choose three corresponding subsequences also denoted $\{P(x,\alpha_m)\}$ $\{T_m\}$, and $\{T'_m\}$ which converge to $P(x,\alpha_\infty)$, T_∞, and T'_∞ respectively and for which $MT_m[f(x) - P(x,\alpha_m)] = e(T_m)$ and $MT'_m[f(x) = P(x,\alpha_m)] = M[f(x) - P(x,\alpha_m)]$. (The convergent sequence $\{P(x,\alpha_m)\}$ exists by Corollary 4.6.) The sequence $\{e(T_m)\}$ is monotone increasing and bounded above which means that $\lim_{m\to\infty} e(T_m)$ exists. Corollary 4.7 implies that $e(T_\infty) \geq e(T_m) > B$ for all m and hence that the set MCS(T_∞) contains a MCS, denoted Q*, of n+1 points. Let $T^* = \{(x,k):(x,k,s)\varepsilon Q^*\}$. Then Lemma 1.40 implies that the rank of WIM(T*) = WIM(Q*) has rank n. Thus $T^* \subset T_\infty$ and hence the hypotheses of Theorem 4.5 are satisfied which means that $\lim_{m\to\infty} e(T_m) = e(T_\infty)$. Then Lemma 4.14 implies that $P(x,\alpha_\infty)$ is the unique best approximation to f(x) on T_∞.

Assume that $P(x,\alpha_\infty)$ is not a best approximation of $f(x)$ on $X\times I$. From Lemma 4.13 we know that T'_∞ contains an absolute maximum point (x',k') of $|L_k(x,\alpha_\infty)|$ over $X\times I$. Since $P(x,\alpha_\infty)$ is not a best approximation, $|L_{k'}(x',\alpha_\infty)| > e(T_\infty)$ and hence by Lemma 4.12 it follows that $e(T_\infty \cup T'_\infty) > e(T_\infty)$. Moreover, $\lim_{m\to\infty} e(T_m \cup T'_m) = e(T_\infty \cup T'_\infty)$ by Theorem 4.5 since $T^* \subset T_\infty$ and the rank of $WIM(T^*)$ is n. Thus for m large enough, $e(T_m \cup T'_m) > e(T_\infty)$ which is impossible. Thus $e(T'_\infty \cup T_\infty) = e(T_\infty) = e$ and $P(x,\alpha_\infty)$ is the unique best approximation to $f(x)$ on $X\times I$. Thus the original sequence $P(x,\alpha_m)$ generated by the algorithm converges to $P(x,\alpha_\infty)$.

In general we do not know if the hypotheses of the previous theorem are satisfied before solving the problem. However, in computational practice Algorithm II with $\varepsilon_o = 0$ has proved to be an efficient method of computing best approximations even when $e(T''_o)$ is not close to e as required by Theorem 4.14. This is true because for most problems of computational interest the quantity $e(T''_m)$ becomes larger than the number B from Theorem 4.5 for quite small m and hence the theorem may be applicable from this point on.

In problems where $MCS(X\times I)$ contains a set of less than n+1 points or where some points of a set in $MCS(X\times I)$ are very close together, the convergence may be quite slow and a high degree of accuracy obtainable only with multiple precision arithmetic.

5. Computational Procedures

In the actual computational use of Algorithm II using linear programming to find best approximations to $f(x)$ on finite point sets, certain serious difficulties were encountered. The most serious difficulty was that of the extreme inaccuracy of the linear programming solution to certain approximation problems on finite point sets as computed by the simplex algorithm. It was also found that the number of eliminations of the simplex algorithm used in finding best approximations was quite large. However, there is a simple procedure which has been very successful in eliminating both of these difficulties without changing any of the theory behind the application of linear programming to the solution of the approximation problem under consideration.

Instead of solving the system of inequalities of line (1) page 62 on a finite point set T''_m while minimizing e, the following procedure is used. Using the notation of the algorithms and defining $P(x,\alpha_{-m}) \equiv 0$, find a best approximation $P(x,\alpha^*)$ to $f(x) - P(x,\alpha_{m-1})$ on the set T''_m and set $\alpha_m = \alpha_{m-1} + \alpha^*$. This means that the linear programming problem corresponding to line (1) page 62 is modified so that for a fixed α_{m-1} we find a point $\alpha^* = (a_1,a_2,\ldots,a_n)$ in E^n such that

$$|W_k(x)D^k[f(x) - P(x,\alpha_{m-1}) - P(x,\alpha^*)]| \leq e$$

for $(x,k) \varepsilon T''_m$ where e is to be a minimum. Then α_m is defined to be $\alpha_{m-1} + \alpha^*$.

The following example shows how this procedure affects the speed and accuracy of the simplex algorithm solution to specific linear programming problems. (The first method will be called Method I and the later modification will be called Method II.)

4.16 Example. Let $I = \{0,1,2,3,4\}$, $X = [-1,1]$, $n = 9$ $\phi_i(x) = x^{i-1}$ for $i = 1,2,\ldots,9$, $f(x) = \sin(x)$, $W_0(x) \equiv 1$

$$W_1(x) = .1(1 - x^2) \qquad W_3(x) = .001(1 - x^{10})$$

$$W_2(x) = .01(1 - x^6) \qquad W_4(x) \equiv .0001$$

The set T''_0 was chosen to be ten equally spaced points from each set $X\times\{k\}$ for $k \varepsilon I$; thus T''_0 consists of 50 points from $X\times I$. Using single precision arithmetic (ten decimal places of accuracy) on the CDC 3600, several terms of the sequence $\{P(x,\alpha_m)\}$ were generated using Algorithm II with $\varepsilon_0 = 0$. The following table gives a comparison of Method I and Method II and indicates the advantages of Method II in both speed and accuracy. The columns marked "Eliminations" indicates the number of eliminations used in the simplex algorithm in the solution of the linear program.

Method I		Method II		
$e(T''_i)$	Eliminations	$e(T''_i)$	Eliminations	i
7.05269×10^{-8}	36	7.05269×10^{-8}	36	0
8.98299×10^{-8}	49	8.98362×10^{-8}	22	1
1.32349×10^{-23}	59	8.98948×10^{-8}	28	2

Letting α_2 and α'_2 denote the third approximation to f(x) as given by Method I and Method II respectively, it was found that

$$M[f(x) - P(x,\alpha_2)] = 3.53572\times10^{-5}$$

$$M[f(x) - P(x,\alpha'_2)] = 8.99017\times10^{-8}$$

The third linear program which used Method I and required 59 eliminations has a computed deviation $e(T''_2)$ which is essentially zero and hence gives no indication concerning the true deviation of a best approximation on the set T''_2. Moreover the set Q_2 as determined by the linear program did not contain a set from $MCS(T''_2)$ and hence the algorithm could not be continued. (It should be noted that the set T''_1 was the same in both Method I and Method II. The set T''_2 was not the same for both methods because the computed solutions from $R(T''_1)$ were not the same for both methods.)

There is one other difficulty which could occur in theory but which has not been observed in actual computational practice. The sets $T_i = \{(x,k):(x,k,s)\varepsilon Q_i\}$ could have an increasing number of points as i becomes large and hence the whole purpose of solving the problem on a small number of points would be defeated. If this would ever become a problem, the set T_i could be replaced by a subset T^* of no more than n+1 points by determining T^* so that $e(T_i) = e(T^*)$. This can be done by taking various n+1 point subsets of T_i and calculating the deviation on these subsets.

6. Computational Examples

Many approximations were computed on the CDC 3600 using single precision arithmetic (ten decimal places of accuracy) and Algorithm II with $\varepsilon_o = 0$. The following examples are given to illustrate some of the unusual things that can occur in the computation of best approximations.

4.17 Example. This example is the computer solution to the approximation problem given in Example 3.1 which does not have a unique best approximation. It was found that the best approximation obtained by the algorithm depended on the initial point set T''_o. In each case a best approximation was determined in two steps of the algorithm. Thus $P(x,\alpha_1)$ is a best approximation in each case. This occurred because the set

$$\{(x,k):(x,k,s)\varepsilon MES(X\times I)\}$$

was included in the set T''_1 and thus a best approximation could be computed exactly. The various sets T''_0 and best approximations $P(x,\alpha_1)$ are given in Table I.

4.18 <u>Example</u>. Let $X = [-1,1]$, $I = \{0,1\}$, $n = 5$, $\phi_i(x) = x^{i-1}$ for $i = 1,2,\ldots,5$, $W_0(x) \equiv 1$, $W_1(x) = .15(1 - x^6)$, $f(x) = \cos(x)$, $Df(x) = -\sin(x)$. The initial set T''_0 was chosen to be 6 equally spaced points from $X\times\{0\}$ including the end points, and 6 equally spaced points from $X\times\{1\}$, including the end points. In Table II the successive x values of the point sets T_i are given in the columns headed (1), (2),...,(6) for values of $i = 0,1,\ldots,8$. The values which are followed by an asterisk are points (x,1) from $X\times\{1\}$ while those without an asterisk are from $X\times\{0\}$. The column headed "Elim." indicates the number of eliminations that were needed to find a best approximation on the set T''_i using the simplex algorithm and Method II. The set T_1 contained one additional point, namely (0.,0), which is not shown in the table.

It should be noted that the extremal points in columns (4) and (5) are approaching each other as the algorithm proceeds. It has been observed that when this occurs, the rate of convergence of the algorithm is slower than for similar problems where the extremal points do not come together. In situations where points come together it may become necessary to take ε_0 to be positive in the algorithms.

TABLE I

α_1	T''_o
(-1., 0, .125)	{(1.,0), (3.,0), (0.,1), (2.,1), (4.,1)}
(-1., -.3, .2)	{(1.,0), (2.,0), (3.,0), (0.,1), (1.,1), (4.,1)}
(-1., .375, .03125)	{(1.,0), (2.,0), (3.,0), (4.,0), (0.,1), (1.,1)}

TABLE II

i	Elim.	(1)	(2)	(3)	(4)	(5)	(6)	$e(T''_i)$	$M[f(x) - P(x,\alpha_i)]$
0	18	-.200000*	-.600000*	.200000*	.600000*	.200000	1.000000	3.2432582×10^{-5}	6.6870722×10^{-5}
1	15	-.257418*	-.697757*	.257518*	.697758*	.200000	1.000000	4.3076237×10^{-5}	8.6049520×10^{-5}
2	19	-.482257	-.697757*	.000000	.697758*	.394370	1.000000	4.4500309×10^{-5}	5.6240242×10^{-5}
3	13	-1.000000	-.706685*	.000000	.702647*	.493133	1.000000	4.5301110×10^{-5}	4.5553097×10^{-5}
4	12	-1.000000	-.706685*	.000000	.702674*	.712539*	1.000000	4.5332309×10^{-5}	4.5375497×10^{-5}
5	10	-1.000000	-.706685*	.000000	.702674*	.707559*	1.000000	4.5347917×10^{-5}	4.5358426×10^{-5}
6	8	-1.000000	-.706685*	-.000029	.705123*	.707559*	1.000000	4.5347962×10^{-5}	4.5350584×10^{-5}
7	7	-1.000000	-.706685*	-.000029	.706342*	.707559*	1.000000	4.5347986×10^{-5}	4.5348638×10^{-5}
8	6	-1.000000	-.706949*	-.000029	.706949*	.707559*	1.000000	4.5348146×10^{-5}	4.5348308×10^{-5}

BIBLIOGRAPHY

1. Dean N. Arden, The solution of linear programming problems, Mathematical Methods for Digital Computers, eds. Anthony Ralston and Herbert S. Wilf, Wiley, 1960, pp. 263-279.

2. N. I. Achieser, Theory of Approximation, Ungar, New York, 1965.

3. R. C. Buck, Linear spaces and approximation theory, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, pp. 11-23.

4. R. C. Buck, Survey of recent Russian literature on approximation, On Numerical Approximation, ed. Rudolph Langer, Univ. of Wis. Press, 1959, pp. 341-359.

5. E. W. Cheney, Introduction to Approximation Theory, McGraw-Hill, New York, 1966.

6. G. B. Dantzig, Linear Programming and Extensions, Princeton University Press, Princeton, N. J., 1963.

7. P. J. Davis, Interpolation and Approximation, Blaisdell, New York, 1963.

8. P. J. Davis, Approximation theory in the first two decades of electronic computers, Approximation of Functions, ed. H. L. Garabedian, Elsevier, Amsterdam, 1965, pp. 152-163.

9. H. H. Denman, Minimax polynomial approximation, Math of Comp., 20 (1966), pp. 257-265.

10. H. L. Garabedian, Approximation of Functions, Elsevier, Amsterdam, 1965.

11. C. Goffman, Real Functions, Holt, Rinehart and Winston, 1963.

12. G. G. Lorentz, Russian literature on approximation in 1958-1964, Approximation of Functions, ed. H. L. Garabedian, Elsevier, Amsterdam, 1965, pp. 191-215.

13. G. G. Lorentz, Approximation of Functions, Holt, Rinehart and Winston, New York, 1966.

14. H. J. Maehly, and Ch. Witzgall, Tchebysheff-Approximation in Kleinen Intervallen I, Approximation durch Polynome, Numer. Math., 2 (1960), pp. 142-150.

15. E. Michael, Topologies on spaces of subsets, Trans. Amer. Math. Soc., 71 (1951), pp. 151-182.

16. D. G. Moursund, Optimal approximation of functions: Chebyshev-type approximations, Ph.D. Thesis, University of Wisconsin, 1963.

17. D. G. Moursund, Chebyshev approximation of a function and its derivatives, Math. Comp., 18 (1964), pp. 382-389.

18. D. G. Moursund and A. H. Stroud, The best Čebyšev approximation to a function and its derivative on n + 2 points, J. SIAM Numer. Anal., 2 (1965), pp. 15-23.

19. D. G. Moursund, Some computational aspects of the uniform approximation of a function and its derivative, J. SIAM Numer. Anal., 2 (1965), pp. 464-472.

20. D. G. Moursund, Chebyshev solution of n + 1 linear equations in n unknowns, J. ACM 12 (1965), pp. 383-387.

21. I. P. Natanson, Constructive Theory of Functions, United States Atomic Energy Commission, Washington, D. C., 1961.

22. E. Ya. Remez, General Computational Methods of Chebyshev Approximation. The Problem with Linear and Real Parameters, United States Atomic Energy Commission, Washington, D. C., 1962.

23. J. R. Rice, The Approximation of Functions, Vol. I: Linear Theory, Addison-Wesley, Reading, Mass., 1964.

24. T. J. Rivlin and H. S. Shapiro, A unified approach to certain problems of approximation and minimization, J. Soc. Indust. Appl. Math., 9 (1941), pp. 670-699.

25. O. Shisha, The Chebyshev polynomial of best approximation to a given function on an interval, Math. of Comp., 20 (1966), pp. 266-271.

26. A. F. Timan, Simultaneous approximation of functions and their derivatives on the whole real axis, Amer. Math. Soc. Translations, Series 2, 44 (1965), pp. 1-11.

INDEX OF NOTATIONS

	Page
X	3
I	3
f(x)	3
$\{\phi_i(x)\}_{i=1}^{n}$	3
n	3
q	3
$w_k(x)$	3
D	3
MT[ ]	3
M[ ]	3
α, β, γ	4
$P(x,\alpha)$	4
e(T)	4
e	4
R(T)	4
R	4
$\|\| \ \|\|$	4
T, T', T_i, T_m	5, 13, 65
IM(Q), IM(T)	7
WIM(Q), WIM(T)	7
det	8

Page

$L_k(x, \alpha)$ 13

$C(T, \alpha)$ 13

$C(\alpha)$ 13

(x,k,s) 13

Extrema, Extremum. 13

Minimal Extremal Set. 15

MES(T) 15

MES 15

Condition A. 16

MCS 18

Minimal Characterization Set 18

MCS(T) 18

n-point 19

δ_m. 39

$|S|$ 47

$d(T,T')$ 65

$Nb(T,\delta)$ 67