



This is to certify that the
thesis entitled
"Approximation From Varisolvent and Unisolvent
Families Whose Members Have Restricted Ranges".

presented by

J. Edward Tornga

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

Gerald D. Taylor
Major professor

Date February 15, 1971

1000

ABSTRACT

APPROXIMATION FROM VARISOLVENT AND UNISOLVENT FAMILIES WHOSE MEMBERS HAVE RESTRICTED RANGES

By

J. Edward Tornga

We consider the questions of existence, characterization, and uniqueness for the following approximating problem. Approximate in the uniform norm a real valued function $f \in C(X)$, where X is a compact set contained in the real closed interval $[a,b]$, from a subset of a certain family of continuous real valued functions defined on $[a,b]$. The subset considered is the subset of the family lying between two curves u and ℓ , where $u > \ell$.

Our family is a varisolvent family in Chapter 1. We also look at the constant error curve difficulty in the characterization theorem for a varisolvent family. In Chapter 2 we consider a family which is unisolvent. Adding a Haar subspace condition to a varisolvent family gives us strong uniqueness and continuity of the best approximation operator theorems in Chapter 3. Finally in Chapter 4 we consider continuous generalized weight function approximation

where local solvency of a varisolvent family is replaced by property A.

Our theorems are the natural extension of the theorems for the same problem considered by G. D. Taylor for a linear family of functions.

APPROXIMATION FROM VARISOLVENT
AND UNISOLVENT FAMILIES WHOSE MEMBERS
HAVE RESTRICTED RANGES

By

J. Edward Tornga

A THESIS

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

DOCTOR OF PHILOSOPHY

Department of Mathematics

1971

ACKNOWLEDGEMENTS

My sincere appreciation goes to my major Professor Gerald D. Taylor for his guidance. I am especially indebted to him for his time, his many suggestions, and his encouragement. I would also like to thank my wife Sondra for her moral support and patience during this time.

TABLE OF CONTENTS

		Page
	INTRODUCTION.....	1
Chapter		
1	APPROXIMATION FROM MEMBERS OF A VARISOLVENT FAMILY HAVING RESTRICTED RANGES.....	4
	Section 1-1: Introduction and Definitions.....	4
	Section 1-2: Existence of Best Approximations.....	7
	Section 1-3: Characterization of Best Approximations....	9
	Section 1-4: The Constant Error Curve Condition in the Characterization Theorem.....	20
	Section 1-5: Uniqueness of Best Approximations.....	30
2	APPROXIMATION FROM MEMBERS OF A UNISOLVENT FAMILY HAVING RESTRICTED RANGES.....	34
	Section 2-1: Introduction and Definitions.....	34
	Section 2-2: Existence, Charac- terization, and Uniqueness of Best Approximations.....	35
3	APPROXIMATION FROM MEMBERS OF A VARISOLVENT FAMILY HAVING RESTRICTED RANGES WITH THE ADDITIONAL HYPOTHESIS OF A HAAR SUBSPACE EXISTING FOR EACH MEMBER OF THE FAMILY.....	41
	Section 3-1: Introduction and Definitions.....	41

		Page
Chapter		
	Section 3-2: Strong Unicity and Continuity of the Best Approximation Operator Theorems.....	51
4	GENERALIZED WEIGHT FUNCTION APPROXIMATION WITH PROPERTY A AND RESTRICTION BETWEEN THE CURVES u AND ℓ WHERE $u > \ell$	59
	Section 4-1: Introduction and Definitions.....	59
	Section 4-2: Existence of Best Approximations.....	63
	Section 4-3: Uniqueness and Characterization of Best Approximations....	66
	BIBLIOGRAPHY.....	77

INTRODUCTION

Beginning with Tchebyshev (see pg. 224-227 of [5] for specific details), the following problem was considered: approximate a continuous function $f(x)$ on $[a,b]$, a,b real finite numbers, from a set of continuous linear functions $\{P(A,x)\}$ on $[a,b]$ where $A = (a_1, \dots, a_n) \in E_n$, Euclidean n space, and the $\{P(A,x)\}$ forms a Tchebyshev set of degree n [†]. That is, for $P(A_1,x) \neq P(A,x)$, both belonging to $\{P(A,x)\}$, $P(A_1,x) - P(A,x)$ can have at most $n - 1$ zeros. By 1918 existence, uniqueness, and characterization of the best approximation to f on $[a,b]$ from $\{P(A,x)\}$ were determined.

In 1949 Motzkin [16] and in 1950 Tornheim [26], extended this type of theory to a non-linear case. Their approximating family was a set of unisolvent functions. That is, a set of continuous functions $\{U(A,x)\}$, $A = (a_1, \dots, a_n) \in E_n$, of degree n such that given points $\{x_i: x_i < x_{i+1}, i = 1, \dots, n\}$ on $[a,b]$ and real numbers $\{y_i\}_{i=1}^n$, there exists a unique $U(A_1,x)$ belonging to $\{U(A,x)\}$ which interpolates the y_i at the x_i . Existence, uniqueness, and characterization of the best approximation

†: all functions considered in this paper are real-valued functions

were discussed.

Since many useful non-linear approximating families, such as the rational and exponential functions were not unisolvent, J. R. Rice [20] in 1961 weakened the hypotheses of a unisolvent family to a locally unisolvent family of continuous functions $\{F(A,x)\}$ having variable degree. Rice called this family a varisolvent family. Exponentials and rationals as well as other non-linear functions have been shown to be varisolvent families (see table 7.1, page 40 of [20]). Questions of existence, uniqueness, and characterization of best approximations from a varisolvent family were examined by Rice.

In 1967 C. Dunham noticed that the proofs of Tornheim's and Rice's characterization theorems for the unisolvent and varisolvent families were incomplete. The unisolvent characterization proof by Tornheim has been completed (see R. Barrar and H. Loeb [1]), whereas the varisolvent characterization proof is only partially completed (see D. Braess [4] and R. Barrar and H. Loeb [1]). By assuming additional conditions on the varisolvent family, however, the characterization theorem has been completed. R. Barrar and H. Loeb [3] gave one such additional condition. They assumed the existence of a Haar subspace for each member of the approximating family in order to complete the proof (see Chapter 3).

G. D. Taylor [22], [23], and [24] among others, has examined the situation of the first paragraph in a

more general setting along with certain restrictions on the approximating family $\{P(A,x)\}$. In particular, the problem was considered when the only acceptable approximations were members of $\{P(A,x)\}$ whose ranges were between two curves. Questions concerning existence, uniqueness, and characterization of best approximations were studied in this restricted setting.

In this paper, we will examine non-linear approximations in the restricted setting of Taylor. Restricted varisolvent families will be considered in Chapter 1. We will also look at the completeness of Rice's characterization theorem when additional hypotheses are assumed. Chapter 2 will consist of approximation by unisolvent families in our restricted setting. A strong uniqueness theorem and a continuity of the best approximation operator theorem will comprise Chapter 3. Our last chapter, Chapter 4, will consider generalized weight function approximation with property A in our restricted setting.

Throughout the paper, we will of course assume that our approximating families are non-empty.

CHAPTER I

APPROXIMATION FROM MEMBERS OF A VARISOLVENT FAMILY HAVING RESTRICTED RANGES

Section 1-1: Introduction and Definitions.

Let E_n represent Euclidean n space. For $A_i \in E_n$, let $A_i = (a_1^i, a_2^i, \dots, a_n^i)$. Let R be the set of real numbers. Let X be a compact subset of R with $X \subset [a, b]$ where a and b are finite and fix P a subset of E_n . Let V be the set of functions $\{F(A, x): Px[a, b] \rightarrow R\}$ where

(1-1) $F(A, x)$ is continuous in the sense that, given

$A_0 \in P$, $x_0 \in [a, b]$, and $\epsilon > 0$, there exists

a $\delta > 0$ such that $A \in P$, $x \in [a, b]$, and

$\|A_0 - A\| + |x - x_0| < \delta$ imply that

$|F(A_0, x_0) - F(A, x)| < \epsilon$ where $\|A_0 - A\| =$

$$\max_{1 \leq i \leq n} |a_i^0 - a_i|$$

(1-2) $F(A, x)$ is locally solvent of degree $m(A)$, i.e.,

suppose we are given a set of points

$S = \{x_j: a \leq x_1 < x_2 \cdots < x_{m(A)} \leq b\}$ and an $\epsilon > 0$.

Then there exists a $\delta = \delta(S, F(A, x), \epsilon) > 0$.

such that

$$|y_j - F(A, x_j)| < \delta \quad j = 1, \dots, m(A)$$

implies that there exists an $F(A_1, x) \in V$

such that both

$$F(A_1, x_j) = y_j \quad j = 1, \dots, m(A)$$

and

$$\|F(A, x) - F(A_1, x)\| < \epsilon$$

hold where

$$\|F(A, x) - F(A_1, x)\| = \max_{x \in [a, b]} |F(A, x) - F(A_1, x)|$$

(1-3) $F(A, x)$ has property Z of degree $m(A)$ on $[a, b]$,

i.e. for any $F(A_*, x) \neq F(A, x)$, $F(A_*, x) \in V$, we

have that $F(A_*, x) - F(A, x)$ has at most $m(A) - 1$

zeros on $[a, b]$ (we assume $m(A) \geq 1$).

We will assume that $m(A)$ is uniformly bounded for all $A \in P$ and that X has more points (in number) than any $m(A)$ for $A \in P$. It should be noted that the degrees of two different members of V may not be the same.

The family of functions V defined above was called a varisolvent family by J. R. Rice [20]. As was mentioned in the introduction to this paper, Rice considered the questions of existence, uniqueness and characterization of best approximations from a family V .

We will examine a generalization of Rice's work. Suppose an additional assumption that only members of V lying between two curves are to be used in the approximating problem. Do we then obtain comparable existence, uniqueness and characterization results?

Let $\mathcal{A}$ be the subset of V where

(1-4) $F(A, x) \in \mathcal{A}$ if $\ell(x) \leq F(A, x) \leq u(x)$ for all $x \in X$

where $u(x)$ and $\ell(x)$ are defined below.

Thus $\mathcal{A}$ is a subset of V bounded between two curves $u(x)$ and $\ell(x)$.

We will consider the existence, characterization, and uniqueness of approximating a given real-valued continuous function f on X in the uniform norm by members of $\mathcal{F}$. That is to say, can we determine an $F(A, x) \in \mathcal{F}$ such that

$$\|F(A, x) - f(x)\| = \inf \{\|F(A, x) - f(x)\| : F(A, x) \in \mathcal{F}\}$$

where again (and for the entire paper),

$$\|F(A, x) - f(x)\| = \max_{x \in X} |F(A, x) - f(x)|$$

First we require a few more definitions. Since we want $\|F(A, x) - f(x)\| = K < \infty$ for some $F(A, x) \in \mathcal{F}$, we must ensure that u does not assume the value $-\infty$ and ℓ the value $+\infty$. Also when u and ℓ assume finite values, we will wish to examine the distance, between u and members of $\mathcal{F}$, and between ℓ and members of $\mathcal{F}$. To ensure the existence of maxima and minima (of $u(x) - F(A, x)$ or $\ell(x) - F(A, x)$, $F(A, x) \in \mathcal{F}$), we require that u and ℓ are continuous on closed subsets of X . Therefore, let u and ℓ be defined on X such that

- (i) ℓ may assume the value $-\infty$, but never $+\infty$.
- (ii) u may assume the value $+\infty$, but never $-\infty$.
- (iii) $X_{-\infty} = \{x : \ell(x) = -\infty\}$ and $X_{+\infty} = \{x : u(x) = +\infty\}$ are open subsets of X .
- (iv) ℓ is continuous on $X - X_{-\infty}$ and u is continuous on $X - X_{+\infty}$.
- (v) $\ell < u$ for all $x \in X$.

Note that (i) $\rightarrow$ (v) imply that

$$\inf [u(x) - \ell(x) : x \in X] = d > 0.$$

The reason for (v) will be discussed in the remark at the end of Section 1-5.

J. R. Rice assumes that $A \neq A_*$ implies that $F(A, x) \neq F(A_*, x)$. Instead of this, we will use R. B. Barrar and H. L. Loeb's notation in [3]. Let N be the maximal value of $m(A)$ for $A \in P$. A function $f \in C[a, b]$ (i.e. continuous on $[a, b]$) will be called a normal point in $C[a, b]$ if it has a best approximation $F(A_*, x)$ which has the property that $m(A_*) = N$. Note that if $F(A, x) \in \mathcal{A}$ and $m(A_*) = N$, then $F(A, x)$ itself is a normal point in $C[a, b]$. We say that A_n is equivalent to A_n' if $F(A_n, x) = F(A_n', x)$. Also the sequence $\{A_n\}$ is equivalent to the sequence $\{A_n'\}$ if A_n is equivalent to A_n' for each n . We will require this concept of equivalence for Theorem 3-2.

The above definitions, unless specifically changed, will apply throughout this paper.

Section 1-2: Existence of Best Approximations.

The existence of a best approximation to f on X from $\mathcal{A}$ is not assured from our definitions about $\mathcal{A}$. We will parallel the discussion on page 9 in [20] by J. R. Rice to obtain one criterion for the existence of a best approximation. On page 6 of [20], the following theorem is given:

Let $\{f_\alpha\}$ be a uniformly bounded infinite set of functions continuous on $[0, 1]$ with property Z of fixed degree. Then $\{f_\alpha\}$ contains a pointwise convergent

subsequence.

Noting that the above theorem holds for our family $\mathcal{F}$, on $[a,b]$ we have

Theorem 1-1: If the limit of every pointwise convergent sequence of members of $\mathcal{F}$ belong to $\mathcal{F}$, then a best approximation to f on X exists from $\mathcal{F}$.

Proof: There always exists a sequence $\{F(A_n, x)\}_{n \in \mathbb{N}} \in \mathcal{F}$ such that

$$\lim_{n \rightarrow \infty} \|F(A_n, x) - f(x)\| = \inf_{F(A, x) \in \mathcal{F}} \|F(A, x) - f(x)\|$$

By taking subsequences, since $m(A)$ is uniformly bounded, there exists a subsequence $\{F(A_s, x)\}$ of $\{F(A_n, x)\}$ with property Z of fixed degree such that

$$\lim_{s \rightarrow \infty} \|F(A_s, x) - f(x)\| = \inf_{F(A, x) \in \mathcal{F}} \|F(A, x) - f(x)\|$$

Applying the theorem of J. R. Rice mentioned in the first paragraph gives a pointwise convergent subsequence $\{F(A_t, x)\}$ of $\{F(A_s, x)\}$ such that

$$\lim_{t \rightarrow \infty} \|F(A_t, x) - f(x)\| = \inf_{F(A, x) \in \mathcal{F}} \|F(A, x) - f(x)\|$$

Therefore, by our hypothesis, the limit of this pointwise convergent sequence belongs to $\mathcal{F}$ and we have a best approximation to f on x from $\mathcal{F}$.

As Rice mentions (page 9 [20]), if there is convergence by functions in $\mathcal{F}$ to functions not in $\mathcal{F}$, one usually enlarges the family $\mathcal{F}$ to include these functions or some type of equivalence is set up between the limit function (not in $\mathcal{F}$) and a function from $\mathcal{F}$. As examples of this,

see pages 42-44 of [20] for exponential families and pages 77-84 of [20] for rational families.

Section 1-3: Characterization of Best Approximations.

The standard Tchebyshev characterization of a best approximation by the number of alternations of the error curve (given by Rice [20] for the unrestricted varisolvent approximating family) extends to our family $\mathcal{F}$, defining alternations as G. D. Taylor does in [23].

A zero x_0 of $g(x) \in C[a,b]$ is said to be a simple zero if $g(x)$ changes sign at x_0 and a double zero if $g(x)$ does not change sign at x_0 . Property Z limits the number of distinct zeros a member of $\mathcal{F}$ may have. We now wish to limit the number of simple and double zeros a member of $\mathcal{F}$ may have. The following Lemma 1-1 is Lemma 7.1 of [20] in our restricted setting.

Lemma 1-1: Let $F(A,x) \in \mathcal{F}$. Then for $F(A_*,x) \in \mathcal{F}$ such that $F(A_*,x) \neq F(A,x)$, we have that $F(A_*,x) - F(A,x)$ can have at most $m(A) - 1$ zeros, counting a simple zero once and a double zero twice.

Proof: Lemma 1-1 is just Lemma 7.1 of [20] with the added condition of (1-4). Therefore $F(A_*,x)$ and $F(A,x)$, considered as unrestricted varisolvent functions, satisfy Lemma 7.1 which implies that $F(A_*,x) - F(A,x)$ can have at most $m(A) - 1$ zeros, counting a simple zero once and a double zero twice. ■

Fix $f(x) \in C(X)$, and let $F(A,x) \in \mathcal{A}$ be given.

Then we define, following Taylor [23],

$$X_{+1} = \{x \in X: f(x) - F(A,x) = \|f(t) - F(A,t)\|\}$$

$$X_{-1} = \{x \in X: f(x) - F(A,x) = -\|f(t) - F(A,t)\|\}$$

$$X_{+2} = \{x \in X: F(A,x) = \ell(x)\}$$

$$X_{-2} = \{x \in X: F(A,x) = u(x)\}$$

$$X_A = X_{+1} \cup X_{+2} \cup X_{-1} \cup X_{-2}$$

The set X_A is said to be the set of critical points of $f(x) - F(A,x)$ on X . As is done in [23], we will divide our problem into two possibilities. Lemma 1-2 and Example 1-1 below are from Taylor [24] and [22].

Lemma 1-2: $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) \neq \phi$

implies that $F(A,x)$ is a best approximation to f on X from $\mathcal{A}$.

Proof: Case 1: $X_{+1} \cap X_{-1} \neq \phi$ implies

$$F(A,x) \equiv f(x).$$

Case 2: $X_{+1} \cap X_{-2} \neq \phi$ or $X_{-1} \cap X_{+2} \neq \phi$

implies that at some critical point, f is a distance $\|f(x) - F(A,x)\|$ above the curve u or below the curve ℓ .

Therefore any other member of $\mathcal{A}$ must be a greater or equal distance from f at this critical point. Therefore $F(A,x)$ is a best approximation. Since

$X_{+2} \cap X_{-2} = \phi$, the proof is complete.

Best approximations satisfying the hypothesis of Lemma 1-2 need not be characterized by alternations as in the case where $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$. The following example is one illustration of this fact.

Example 1-1: Let $f(x) = x^2$ on $[-1,1] = X$.

Let $\mathcal{F} = \{F(A,x): F(A,x) = ax^2 + bx + c, \text{ where } 0 \geq F(A,x) \geq -1\}$. Any member of $\mathcal{F}$ passing through 0 for $x = -1$ and $x = +1$ is a best approximation.

The example above also shows that for

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) \neq \phi,$$

a best approximation from $\mathcal{F}$ may not be unique.

For the remainder of this section we will consider the case $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$.

Definition: $F(A,x) - f(x)$ is said to alternate K times on X if there exist $k + 1$ critical points $\{x_i\}$ in X where $a \leq x_1 < x_2 \cdots < x_{k+1} \leq b$ and such that

$$x_i \in X_{-1} \cup X_{-2} \text{ implies } x_{i+1} \in X_{+1} \cup X_{+2}$$

or

$$x_i \in X_{+1} \cup X_{+2} \text{ implies } x_{i+1} \in X_{-1} \cup X_{-2}$$

for $i = 1, 2, \dots, k$

These $k + 1$ points are said to form an alternant of length k .

A standard type Tchebyshev characterization is as follows: Let $X_{-2} = X_{+2} = \phi$ in the above definition. Then $g(x)$ is a best approximation to f on X if and only if the error curve $f - g$ alternates n times on X (n depends upon the approximating family). As was mentioned in the introduction to this paper, Tchebyshev and his contemporaries examined a continuous linear family of functions $\{P(A, x)\}$ which formed a Tchebyshev set. The above type of characterization of the best approximation was discovered. In 1949 and 1950 Motzkin and Tornheim noted the same type of characterization theorem for their unisolvent families. In 1961 J. R. Rice exhibited a similar type of characterization theorem for his more general non-linear family, the varisolvent family.

In 1967 C. Dunham [6] noted that the characterization proofs of Tornheim and Rice for unisolvent and varisolvent families were incomplete. They both neglected to consider the possibility of a constant error curve (i.e. $f - g \equiv c$). R. Barrar and H. Loeb [1] showed that the constant error curve could not exist for unisolvent families in Tornheim's proof (this was already known since a characterization proof for unisolvent families, different from Tornheim's existed - see Novodvorskiĭ and Pinsker [18]). Barrar and Loeb in the same paper also showed that for $m(A) \leq 3$, a constant error curve could not exist for varisolvent families.

D. Braess [4] has shown that if in an ϵ neighborhood of the best approximation from V (V is a varisolvent family), all the members of V have the same degree as the best approximation, then the best approximation must alternate. Since the degree is an upper semi-continuous function (see Theorem 2 of Rice [20]), it then follows that best approximations of maximal degree must alternate. At this time, however, it is not known in general whether a constant error curve for the best approximation from a varisolvent family can exist.

By adding additional hypotheses, the possibility of a constant error curve can be eliminated. As was mentioned, a Haar condition can be added (see Chapter 3). By using property A (see Chapter 4) in lieu of local solvency, this difficulty can also be overcome. A third possibility would be to assume that the varisolvent family is extendable to a larger interval (see Corollary 3 of Theorem 1-2). A fourth might be to assume closure of $\mathcal{A}$ under pointwise convergence (see Corollary 4 of Theorem 1-2). An obvious fifth possibility would be to assume that each member of V intersects f at some point of X .

The proof of the next theorem follows the standard approaches used by both Tornheim and Rice in their characterization proofs.

Theorem 1-2: Let $f \in C(X)$. $F(A,x) \in \mathcal{F}$, and assume $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, where X is a compact set contained in $[a,b]$.

- (1) If $F(A,x)$ is a best approximation to f from $\mathcal{F}$ and $F(A,x) - f(x) \neq c$, c a non-zero constant, then $F(A,x) - f(x)$ alternates at least $m(A)$ times on X .
- (2) If $F(A,x) - f(x)$ alternates at least $m(A)$ times on X , then $F(A,x)$ is a best approximation to f from $\mathcal{F}$ on X .

Proof: (of (2)) Assume $F(A,x) - f(x)$ alternates at least $m(A)$ times on X . Assume $F(A,x)$ is not a best approximation to f from $\mathcal{F}$. Then there exists an $F(A_*,x)$ in $\mathcal{F}$ such that

$$\|F(A_*,x) - f(x)\| < \|F(A,x) - f(x)\|$$

Let $\{x_j: j = 1, 2, \dots, m(A)+1, x_j < x_{j+1}, x_j \in X\}$ be an alternant of length $m(A)$ for

$F(A,x) - f(x)$ on X . These must be distinct since $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$. Now at these critical points we have

$$F(A_*,x_j) - F(A,x_j) = (F(A_*,x_j) - f(x_j)) - (F(A,x_j) - f(x_j)).$$

We assert that $F(A_*,x) - F(A,x)$ has at least $m(A)$, zeros on $[a,b]$, counting simple zeros once and double zeros twice, which will contradict Lemma 1-1. We will match a counting zero (a double zero has two counting zeros, a simple zero one) to one and only one interval

(x_j, x_{j+1}) for $j = 1, 2, \dots, m(A) + 1$.

Assume $x_1 \in X_{-1} \cup X_{-2}$ for $F(A, x)$.

Then $F(A_*, x_1) - F(A, x_1) \leq 0$ since $F(A_*, x) \in \mathcal{A}$ is a better approximation to $f(x)$ than $F(A, x)$.

Case 1: $F(A_*, x_1) - F(A, x_1) = 0$. Associate the zero x_1 with the interval (x_1, x_2) .

Case 2: $F(A_*, x_1) - F(A, x_1) < 0$. Two possibilities can occur.

(a) $F(A_*, x_\ell) - F(A, x_\ell) \geq 0$ for some $x_\ell \in (x_1, x_2)$. Since $F(A_*, x) - F(A, x)$ is a continuous function on $[a, b]$, it has a zero on (x_1, x_2) . Associate a zero on (x_1, x_2) with the interval (x_1, x_2) .

(b) $F(A_*, x) - F(A, x) < 0$ for all $x \in (x_1, x_2)$.

Since $x_2 \in X_{+1} \cup X_{+2}$ and $F(A_*, x) \in \mathcal{A}$ is a better approximation than $F(A, x)$, there exists a zero at x_2 . Associate this zero with (x_1, x_2) .

We have associated a zero in $[x_1, x_2]$ with (x_1, x_2) . Now

$x_2 \in X_{+1} \cup X_{+2}$ for $F(A, x)$ and therefore

$$F(A_*, x_2) - F(A, x_2) \geq 0.$$

If $F(A_*, x) - F(A, x)$ changes sign in (x_2, x_3) , associate a zero in (x_2, x_3) with (x_2, x_3) . If not, consider cases.

Case 1: $F(A_*, x) - F(A, x) > 0$ on (x_2, x_3) .

Since $x_3 \in X_{-1} \cup X_{-2}$ for $F(A, x)$, a zero exists

at x_3 . Associate this zero with (x_2, x_3) .

Case 2: $F(A_*, x) - F(A, x) < 0$ on (x_2, x_3) .

For a sufficiently small $\delta > 0$, two possibilities can occur,

(a') $F(A_*, x) - F(A, x) \geq 0$ on $(x_2 - \delta, x_2)$

in which case a zero exists at x_2 which was not associated with (x_1, x_2) . Associate this zero with (x_2, x_3) .

(b') $F(A_*, x) - F(A, x) < 0$ on $(x_2 - \delta, x_2)$

in which case a double zero exists at x_2 that was used at most once on (x_1, x_2) . Associate an unused counting zero at x_2 with (x_2, x_3) .

We have associated a zero in $[x_2, x_3]$ with (x_2, x_3) which was not associated with (x_1, x_2) .

Proceed in the same manner with the remaining intervals (x_j, x_{j+1}) , $j = 3, 4, \dots, m(A)$ as was done for (x_2, x_3) .

For $x_1 \in X_{+1} \cup X_{+2}$, the argument would be similar.

In either case, $F(A_*, x) - F(A, x)$ has at least $m(A)$ zeros on $[a, b]$, counting double zeros twice and simple zeros once. This is the desired contradiction of Lemma 1-1 and (2) of Theorem 1-2 is proved.

(of (1)) Assume $F(A, x)$ is a best approximation to f from $\mathcal{F}$ and $F(A, x) - f(x) \neq c$, c a non-zero constant. Further assume that $F(A, x) - f(x)$ alternates exactly $s < m(A)$ times at the points

$$\{x_j: a \leq x_1 < x_2 < \dots < x_{s+1} \leq b, x_j \in X\}.$$

(Note that it is possible for s to be zero here).

Case 1: Assume that a and b are not critical points for every critical point set of $(s + 1)$ points of $F(A, x) - f(x)$. (Note that a and/or b may not even belong to X).

If a (or b) is not a critical point, select a $\delta_1 > 0$ such that $a + \delta_1$ is less than the first possible critical point (or $(b - \delta_1)$ is greater than the last possible critical point). For concreteness, assume a is not a critical point. Determine a $\delta(\epsilon_0)$ where $\delta_1 > \delta > 0$, such that for some $\epsilon_0 > 0$

(sufficiently small), we have

$$\max \{ \ell(x) - f(x), - \|F(A, x) - f(x)\| \} + \epsilon_0 <$$

$$F(A, x) - f(x) < \min \{ u(x) - f(x), \|F(A), x) - f(x)\| \} - \epsilon_0$$

for all $x \in [a, a + \delta] \cap X$. This is possible by continuity, compactness, and the fact that $u(x) > \ell(x)$ on X .

Now, the facts of the previous sentence and

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$$

imply that we can select the following points

$$a = x_0 < x_{m(A)-s} < x_{m(A)-s+1} < \dots < x_{m(A)} = b$$

which divide X into $s + 1$ subsets so that for $\delta_2 > 0$ sufficiently small,

(a) $F(A, x) - f(x)$ alternates exactly once on any two adjacent subsets, but does not alternate on any

one subset.

(b) If $(X_{+1} \cup X_{+2}) \neq \emptyset$ on a subset, then

$$F(A, x) - f(x) < \min \{u(x) - f(x), \|F(A, x) - f(x)\|\} - \delta_2 \text{ for all } x \in (\text{the subset} \cap X)$$

If $(X_{-1} \cup X_{-2}) \neq \emptyset$ on a subset, then $F(A, x)$

$$- f(x) > \max \{\ell(x) - f(x), - \|F(A, x) - f(x)\|\} + \delta_2$$

for all $x \in (\text{the subset} \cap X)$

(c) $\max \{\ell(x) - f(x), - \|F(A, x) - f(x)\|\} + \delta_2 < F(A, x) - f(x)$

$$< \min \{u(x) - f(x), \|F(A, x) - f(x)\|\} - \delta_2 \text{ for all}$$

points in the set $\{x_{m(A)-s}, x_{m(A)-s+1}, \dots, x_{m(A)}\}$

Choose $m(A) - s - 1$ distinct points

$\{x_j: j = 1, 2, \dots, m(A) - s - 1, x_j < x_{j+1}\}$ in $[a, a + \delta]$.

Let $\epsilon = \min(\frac{\delta_2}{2}, \frac{\epsilon_2}{2}) > 0$ and x' denote a point in X

above where $\|F(A, x) - f(x)\|$ is assumed.

Since $F(A, x)$ is varisolvent of degree $m(A)$, there exists an $F(A_*, x) \in \mathcal{F}$ where

$$(a) \quad F(A_*, x_j) - F(A, x_j) = 0 \quad j = 1, 2, \dots, m(A) - 1$$

$$(b) \quad |F(A_*, x') - f(x')| < |F(A, x') - f(x')|, \text{ and}$$

$$(c) \quad \|F(A, x) - F(A_*, x)\| < \epsilon$$

Now by Lemma 1-1, (a) and (b), $\max |F(A_*, x) - f(x)|$

$< \max |F(A, x) - f(x)|$ for all $x \in ([a + \delta, b] \cap X)$

while (c) implies $\max |F(A_*, x) - f(x)| < \max |F(A, x) - f(x)|$

for all $x \in [a, a + \delta] \cap X$. Therefore

$$\|F(A_*, x) - f(x)\| < \|F(A, x) - f(x)\|$$

(for all $x \in X$) and $F(A_*, x) \in \mathcal{F}$ is a better approximation to f on X than $F(A, x)$. This is a contradiction.

Case 2: Assume both a and b are critical points.

Replace $[a, a + \delta]$ by a similar interval

$$[x_{m(A)-s-\delta}, x_{m(A)-s}]$$

where $\{x_j: j = 0, m(A) - s, \dots, m(A)\}, \delta, \epsilon_0, \delta_2$, and ϵ are defined analogously to the previous construction.

For $m(A) - s - 1$ even, choose $m(A) - s - 1$ points $\{x_j: x_{m(A)-s-\delta} \leq x_1 < x_2 < \dots < x_{m(A)-s-1} < x_{m(A)-s}\}$

and determine an $F(A_*, x) \in \mathcal{F}$ such that

(a') For $x \in \{a, b\}$, $x \in X_{+1} \cup X_{+2}$ implies that

$$F(A_*, x) - f(x) > F(A, x) - f(x), \text{ while } x \in X_{-1} \cup X_{-2} \\ \text{implies } F(A_*, x) - f(x) > F(A, x) - f(x).$$

(b') $\|F(A_*, x) - F(A, x)\| < \epsilon$, and $F(A, x_j) - F(A_*, x_j) = 0$ for $j = 1, 2, \dots, m(A) - 1$.

The evenness of $m(A) - s - 1$ and our construction imply that the two conditions in (a') are in reality only one restriction on our varisolvent function, so that indeed, an $F(A_*, x) \in \mathcal{F}$ does exist. Then

$$\|F(A_*, x) - f(x)\| < \|F(A, x) - f(x)\|$$

which is a contradiction.

For $m(A) - s - 2$ even, choose $m(A) - s - 2$ points

$$\{x_j: x_{m(A)-s-\delta} < x_2 < x_3 \dots < x_{m(A)-s-1} < x_{m(A)-s}\}$$

and determine an $F(A_\star, x) \in \mathcal{A}$ such that (a') and (b') above hold and

$$(c') \quad F(A, x_j) - F(A_\star, x_j) = 0 \quad j = 2, 3, \dots, m(A) - 1.$$

If $F(A, x) - F(A_\star, x)$ has a zero in addition to those in (c'), then (a') implies that this is a double zero, which is impossible by Lemma 1-1. Therefore

$$\|F(A_\star, x) - f(x)\| < \|F(A, x) - f(x)\|$$

which is a contradiction.

Now for all possibilities we have constructed a better approximation to f than $F(A, x)$. This contradicts the fact that $F(A, x)$ can alternate at most $s < m(A)$ times. Therefore (1) of Theorem 1-2 is proven. ■

Section 1-4: The Constant Error Curve Condition in the Characterization Theorem.

It would be very desirable to omit the condition of the constant error curve in (1) of Theorem 1-2. We would then have an if and only if statement for the characterization of best approximations in terms of alternations. This section will consist of corollaries to Theorem 1-2 where the constant error condition is eliminated. One such case with the added hypothesis of a Haar subspace, will be deferred until Chapter 3.

R. Barrar and H. Loeb [1] have shown that for $m(A) \leq 3$, a constant error curve for the best approximation cannot exist. With the addition of our condition on the critical point sets, their reasoning applies for our family $\mathcal{F}$.

Corollary 1: For $m(A) \leq 3$, the constant error curve condition in (1) of Theorem 1-2 can be omitted.

Proof: For concreteness, assume that $f(x) - F(A,x) \equiv c > 0$ (if $c < 0$, a similar argument holds).

Assume $m(A) = 1$. Since $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, $\max_{x \in X} (u(x) - F(A,x)) = s_1 > 0$. Then for $s_1 > \epsilon_1 > 0$, there exists a $\delta_1(\epsilon_1) > 0$ and an $F(A_1, x) \in \mathcal{F}$ such that (i) $\|F(A, x) - F(A_1, x)\| < \epsilon_1$, (ii) $F(A_1, x)$ doesn't intersect $F(A, x)$, and (iii) $F(A_1, a) - F(A, a) = \delta_1 > 0$, since $\mathcal{F}$ is a varisolvent family. But then $F(A_1, x)$ is a better approximation to f from $\mathcal{F}$ than $F(A, x)$ which is a contradiction.

Assume $m(A) = 2$. Again since $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, $\max_{x \in X} (u(x) - F(A, x)) = s_2 > 0$. Then for $s_2 > \epsilon_2 > 0$, there exists a $\delta_2(\epsilon_2) > 0$ and an $F(A_2, x) \in \mathcal{F}$ such that (i) $\|F(A, x) - F(A_2, x)\| < \epsilon_2$, and (ii) $F(A_2, a) - F(A, a) = \delta_2 > 0, F(A_2, b) - F(A, b) = \delta_2 > 0$.

Now, $F(A_2, x) - F(A, x)$ has at most one zero ($m(A) = 2$), but (ii) implies that if a zero occurs, it is a double zero or another simple zero exists. This is impossible, therefore $F(A_2, x) > F(A, x)$ for all $x \in [a, b]$ and by construction, $F(A_2, x)$ is a better approximation to f from $\mathcal{F}$ than $F(A, x)$. This contradicts our hypothesis.

Assume $m(A) = 3$. There exists an $\delta_3 > 0$ constructed as above such that for $s_3 > \epsilon_3 > 0$, there exists by solving a $\delta_3(\epsilon_3) > 0$ and an $F(A_3, x) \in \mathcal{F}$ such that

$$(i) \quad F(A_3, a) = F(A, a), \quad F(A_3, b) = F(A, b),$$

$$(ii) \quad F(A_3, \frac{a+b}{2}) - F(A, \frac{a+b}{2}) = \delta_3, \quad \text{and}$$

(iii) $\|F(A_3, x) - F(A, x)\| < \epsilon_3$. Since $F(A_3, x) - F(A, x)$ can have at most two zeros, and it has those at a and b , by construction, it can have no other zeros and is a non-constant error curve best approximation to f from $\mathcal{F}$.

Applying (1) of Theorem 1-2 says that we can find a better approximation, since $F(A_3, x) - f(x)$ does not alternate enough ($m(A_3) \geq 3$). This concludes the proof. **■**

D. Braess has shown that a constant error curve for a varisolvent family V cannot occur under another condition. His condition is that there be an $\epsilon > 0$ neighborhood of the best approximation $F(A, x) \in V$, such that any member of V lying entirely in this neighborhood

has the same degree as $F(A, x)$. J. R. Rice (pages 5 and 6 of [20]) showed that for $F(A_1, x) \in V$, there existed a $\delta > 0$ neighborhood of $F(A_1, x)$, such that if $F(A, x)$ (belonging to V), was entirely in this neighborhood, then $m(A) \geq m(A_1)$. Braess used both these results to arrive at the following corollary (which is placed in our setting).

Corollary 2: If $F(A, x)$ is the best approximation to f on X from $\mathcal{F}$ such that there exists no member of V (V is the varisolvent family where $\mathcal{F} \subset V$) with degree greater than $m(A)$ lying entirely in some $\epsilon > 0$ neighborhood of $F(A, x)$, then the constant error curve in (1) of Theorem 1-2 can not occur.

One result, noted by Braess, is that if the best approximation has maximal degree, (see Corollary 2 above), a constant error curve can not occur.

The following corollary is due to G. Meinardus and G. D. Taylor (oral communication).

Corollary 3: If for each $F(A, x) \in \mathcal{F}$, there exists an extension $[a_1, b_1]$ of $[a, b]$ (either $-\infty < a_1 < a$ or $\infty > b_1 > b$ - possibly both) and a varisolvent family V' on $[a, b]$ such that $V'|_{[a, b]} = \mathcal{F}$ and for some $F(A_*, x) \in V'$, we have that $F(A_*, x) \equiv F(A, x)$ on $[a, b]$, we can then omit the constant error curve condition in (1) of Theorem 1-2.

Proof: As in the proof of Corollary 1, for concreteness, assume $f(x) - F(A, x) \equiv c > 0$ and $F(A, x)$ is the best approximation to f from $\mathcal{F}$ on X . Then there exists an $s > 0$ such that $\max_{x \in X} (u(x) - F(A, x)) = s$ since

$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$. Using our hypothesis, we then have an $F(A_*, x)$ belonging to V' such that $F(A_*, x) \equiv F(A, x)$ on $[a, b]$. Select $m(A_*) - 1$ distinct points on $[a_1, b_1]$ which are not in $[a, b]$. Now by our solvency condition, there exists an $F(A_1, x) \in V'$ such that $F(A_1, x)$ equals $F(A_*, x)$ at the above $m(A_*) - 1$ points and for some point in $[a, b]$, $F(A_1, x) - F(A_*, x) = s/2$. Now $V'|_{[a, b]} = \mathcal{F}$ implies $F(A_1, x) \in \mathcal{F}$. But by our construction $F(A_1, x) \in \mathcal{F}$ is a better approximation to f than $F(A, x)$. This contradicts our hypothesis. $\blacksquare$

The next corollary to Theorem 1-2 eliminates the constant error curve when $\mathcal{F}$ is closed under pointwise convergence. This does not appear to be that strong a condition, since some form of compactness on $\mathcal{F}$ is usually required to ensure existence of a best approximation (see Section 1-2).

We first require some preliminary results due to C. B. Dunham [7].

Definition: A family G of functions is dense compact on X , a compact space, if every bounded sequence of elements of G has a subsequence converging pointwise on

a dense subset Y of X to an element g of G and for $x \notin Y$, $\liminf_{y \rightarrow x} g(y) \leq g(x) \leq \limsup_{y \rightarrow x} g(y)$ for $y \in Y$.

Note that if $G \subset C(X)$, then the above inequality can be omitted in the definition of dense compactness.

Lemma 1-3: Let G be dense compact. Let $\{f_k\} \in C(X)$ converge uniformly to $f \in C(X)$ and g_k be a best approximation to f_k . Then $\{g_k\}$ has a subsequence $\{g_{k_j}\}$ converging pointwise on a dense subset of X to a best approximation to f .

The theorem mentioned in the first paragraph of Section 1-2 gives us a pointwise convergent sequence for a bounded sequence of functions from $\mathcal{F}$. Therefore if we assume that $\mathcal{F}$ is closed under pointwise convergence, $\mathcal{F}$ will be dense compact.

Corollary 4: If $\mathcal{F}$ is closed under pointwise convergence, and the number of points of X is at least twice the uniform bound of $m(A)$, then the constant error curve condition of (1) in Theorem 1-2 can be omitted.

Proof: For concreteness, assume $f(x) - F(A, x) \equiv c > 0$ (a similar argument holds of $c < 0$). Let $\{x_i: i = 1, 2, \dots, m(A)\}$ and $\{y_i: i = 1, 2, \dots, m(A) - 1\}$ be a set of points in X such that $x_i < y_i < x_{i+1}$. Let $g_n \in C(X)$ be defined as follows,

$$g_n(x) = \begin{cases} f(x) - c/n & \text{for } x = x_i \quad i = 1, 2, \dots, m(A) \\ f(x) & \text{for } x = y_i \quad i = 1, 2, \dots, m(A) - 1 \\ \text{any continuous curve } h(x) \text{ on } X \text{ connecting} \\ g_n(x) \text{ at the } x_i \text{'s and } y_i \text{'s such that} \\ f(x) \geq h(x) \geq f(x) - c/n \text{ on } \{x_i, x_{i+1}\} \text{ for} \\ i = 1, 2, \dots, m(A) - 1. \end{cases}$$

Now since g_n is continuous on X , let $F(A_n, x)$ be the best approximation to g_n on X from $\mathcal{F}$. If

$$\|g_n(x) - F(A_n, x)\| \geq c,$$

then $F(A, x)$ is a best approximation to g_n and by construction $F(A, x)$ is not parallel to g_n . Applying (2) of Theorem 1-2 gives us a better approximation to g_n than $F(A, x)$ (or $F(A_n, x)$). Therefore

$$\|g_n(x) - F(A_n, x)\| < c.$$

We assert that $F(A_n, x)$ is not parallel to g_n . Assume it is, i.e. $g_n(x) - F(A_n, x) \equiv c_n$. We first note that if $c_n = 0$, by construction $F(A_n, x)$ would be a better approximation to f than $F(A, x)$. Therefore $c_n \neq 0$. Next, assume $c_n < 0$. Then $F(A_n, x)$ is above g_n , but $F(A, x)$ is within c of g_n . Therefore since $F(A, x) \neq F(A_n, x)$, (g_n is not parallel to f), $|F(A_n, x) - f(x)| \leq c$. But $F(A_n, x)$ is a

best approximation to f which is not parallel to f on X . Applying (2) of Theorem 1-2 gives us a better approximation to f from $\mathcal{F}$ than $F(A_n, x)$ (or $F(A, x)$). This is not possible, hence $c_n > 0$. We should further note that $|c_n - c| < c/n$, for if not by the construction of g_n and the fact that $g_n(x) - F(A_n, x) \equiv c_n > 0$, $F(A_n, x)$ would be a better approximation to f than $F(A, x)$. Now, $c_n > 0$ implies that $f(x) > F(A_n, x)$ and $g_n(x) > F(A_n, x)$. If $F(A, x)$ and $F(A_n, x)$ do not intersect, then one is above the other and is a best approximation to both f and g_n . This is a contradiction since one function cannot be parallel to both f and g_n . Therefore assume $F(A, x)$ and $F(A_n, x)$ do intersect. Then by construction, at the x_i , $F(A, x) - F(A_n, x) \geq 0$ while at the y_i , $F(A, x) - F(A_n, x) < 0$ since $c_n < c$. Now $F(A, x)$ and $F(A_n, x)$ both being continuous on $[a, b]$ implies that $F(A, x) - F(A_n, x)$ has at least $m(A)$ zeros on $[a, b]$ counting double zeros twice and simple zeros once, because of our construction. This contradicts Lemma 1-1. Therefore the best approximation $F(A_n, x)$ to g_n from $\mathcal{F}$ cannot be parallel to g_n .

Since it is not parallel, Theorem 1-2 says that $F(A_n, x) - g_n(x)$ alternates at least $m(A_n)$ times on X . A further property of our construction is that there exists

a point $x' \in X$ such that $F(A_n, x') - F(A, x') < 0$ (or at x' , $F(A_n, x)$ is below $F(A, x)$). For, if not, $F(A_n, x)$ would be a best approximation to f which is not parallel to f . Again apply (2) of Theorem 1-2 to arrive at a contradiction.

We now have for each g_n , $n > 1$, a non-parallel best approximation from $\mathcal{F}$, $F(A_n, x)$ which has a point of X below $F(A, x)$. By construction g_n tends uniformly to f as n goes to ∞ . Since we have assumed that $\mathcal{F}$ is dense compact (see the paragraph immediately preceeding Corollary 4), Lemma 1-3 may be applied to give us some subsequence of $\{F(A_n, x)\}_{n=1}^{\infty}$ converging pointwise to a best approximation $F(A_*, x)$ of f from $\mathcal{F}$. Each member of this pointwise convergent (sub) sequence alternates at least once on X and at one critical point is below $F(A, x)$. Therefore a cluster point of critical points of the pointwise convergent sequence exists, which is on $F(A, x)$. Likewise a cluster point of the form $X_{-1} \cup X_{-2}$ exists since each best approximation in the sequence alternates at least once. Therefore since the limiting function $F(A_*, x)$ is a continuous function (it belongs to $\mathcal{F}$ by our closure hypothesis), it must alternate at least once on X . Then $F(A_*, x)$ must alternate $m(A_*)$ times on X or we could apply (2) of Theorem 1-2 to obtain a contradiction. But

$$\|F(A_*, x) - f(x)\| = \|F(A, x) - f(x)\|$$

and $F(A, x)$ is parallel to f . Therefore $F(A_*, x) - F(A, x)$

has at least $m(A_*)$ zeros (counting double zeros twice) on $[a,b]$ which contradicts Lemma 1-1. This contradicts the fact that $F(A,x)$ is parallel to f . ■

Although Corollary 4 is a nice theoretical result, it may not be useful practically. We required in the proof that since each member of the pointwise converging subsequence alternated at least once, the limit function also alternated at least once. If we employed dense compactness only, and our approximating class was not a subset of $C(X)$, it would be possible that convergence on a dense subset would be to the original parallel approximating function $F(A,x)$. This would not give us our desired contradiction. Unfortunately, for exponential and rational families, in order to obtain closure under pointwise convergence, functions are added which may not be continuous, even though they agree with a continuous function on a dense subset of X . Rice [20] in Chapter 8 for exponential families and Chapter 9 for generalized rational families makes an identification in that if a subsequence converges to an element which is a discontinuous best approximation, the closure of the continuous portion of the best approximation then becomes a continuous best approximation. So, although his best approximations are all continuous, his approximating families are enlarged to include discontinuous functions which are pointwise convergent limits of sequences from

the family. Therefore this enlarged family is not really a continuous family, only a dense compact one.

The continuity theorems of the best Tchebyshev operators which I have seen will not overcome this problem. C. B. Dunham's [7] theorem says that the pointwise convergence on a dense subset is actually uniform convergence, while R. H. Barrar and H. Loeb's theorem in Chapter 4 has the additional Haar condition included, which eliminates the possibility of a constant error curve.

The theorem of de la Vallee Poussin applies for $\mathcal{A}$ as well as for the unrestricted variational case (see Rice [20] page 12).

Theorem 1-3: For $F(A, x) \in \mathcal{A}$ if $F(A, x) - f(x) = \phi(x)$ assumes the values $\lambda_i = \phi(x_i)$ at $m(A) + 1$ points $x_i < x_{i+1}$ on X such that $\text{sgn } \lambda_i = -\text{sgn } \lambda_{i+1}$, then

$$\min |\lambda_i| \leq \inf_{F(A_*, x) \in \mathcal{A}} \|F(A_*, x) - f(x)\| \leq \|F(A, x) - f(x)\|$$

Proof: The right hand inequality is trivially true.

Assuming the left is false implies that there exists an $F(A_1, x) \in \mathcal{A}$ such that $F(A, x) - F(A_1, x)$ has at least $m(A)$ zeros on $[a, b]$ which contradicts Lemma 1-1. !

Section 1-5: Uniqueness of Best Approximations.

When $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) \neq \emptyset$, Example 1-1 has shown that the best approximation to f from $\mathcal{A}$ need

not be unique. Therefore we will consider the case when $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$.

Theorem 1-4: If $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$ and the best approximation $F(A,x)$ to f on X from $\mathcal{F}$ does not have a constant error curve, then $F(A,x)$ is the unique best approximation to f from $\mathcal{F}$.

Proof: Changing the first ' $<$ ' in the proof of (2) of Theorem 1-2 to ' $\leq$ ' and repeating this altered proof of (2) of Theorem 1-2 gives us uniqueness of the best approximation. ■

If the best approximation $F(A,x)$ to f on X from $\mathcal{F}$ has a constant error curve where

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi ,$$

any other best approximation must also have a constant error curve. If another best approximation $F(A_*,x)$ were not parallel to f , part (1) of Theorem 1-2 would apply to say that $F(A_*,x) - f(x)$ alternates at least $m(A_*)$ times on X . Then $F(A,x) - F(A_*,x)$ has at least $m(A_*)$ zeros contradicting Lemma 1-1. We therefore have the following corollary to Theorem 1-4.

Corollary 1: If $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$ and the best approximation $F(A,x)$ to f on X from $\mathcal{F}$ has a constant error curve, then at most one other best

approximation $F(A_*, x)$ to f exists and $F(A_*, x) \equiv F(A, x) + c$, where c is a non-zero constant.

Even when $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$ and the best approximation $F(A, x)$ to f on X from $\mathcal{F}$ has a constant error curve, $F(A, x)$ may be unique, since the nature of the functions f , u , and ℓ on X may prohibit another best approximation. The following are some obvious cases where the best approximation, having constant error curve and $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, is unique

- (a) if there exists an $x' \in X$ such that $f(x') \geq u(x')$
or $\ell(x') \geq f(x')$
- (b) if when $u(x) > f(x) > \ell(x)$ for all $x \in X$,
$$\min_{x \in X} \{ \min(u(x) - f(x)), \min(f(x) - \ell(x)) \} < \|F(A, x) - f(x)\|$$

Remark: An open question which I plan to look at later, is whether the previous sections can be generalized to the point where $u(x) = \ell(x)$ at a finite number of points (see G. D. Taylor [24] for the linear case and K. Taylor [25] for rational families). R. Barrar and A. Loeb [2] examined the question of approximating a function f from a varisolvent family which interpolates f at a finite number of points. Extending G. D. Taylor's paper to our setting would give us interpolation and restricted range approximation at the same time.

The non-linear case of $\ell(x) \leq u(x)$ appears, however,

to require much stronger hypotheses. First, a certain amount of differentiability for functions of $\mathcal{F}$ is needed at the points where $u(x) = \ell(x)$ in order to generalize Taylor's paper. Although part (2) of theorem 1-2 can be shown to be true, in part (1) a difficulty occurs in constructing a better approximating function from $\mathcal{F}$ around the points where $u(x) = \ell(x)$. It appears that more than this added differentiability is necessary to obtain (1) of Theorem 1-2. We could assume a uniform bound on a certain order derivative of members of $\mathcal{F}$ where $u(x) = \ell(x)$. This would remove the difficulty around the points $u(x) = \ell(x)$ and allow us to complete part (1) of Theorem 1-2.

CHAPTER 2

APPROXIMATION FROM MEMBERS OF A UNISOLVENT FAMILY HAVING RESTRICTED RANGES

Section 2-1: Introduction and Definitions.

Let U be the set of functions

$\{G(A,x): P \times [a,b] \rightarrow R\}$ which satisfy (2-1) and (2-2) below.

(2-1) $G(A,x)$ is a continuous function on $[a,b]$.

(2-2) Given the set $\{x_i: a \leq x_1 < x_2 < \dots < x_n \leq b\}$

of n distinct points and n arbitrary real numbers $\{y_i\}_{i=1}^n$, then there exists a unique

$G(A_*,x) \in U$ such that $G(A_*,x_i) = y_i$ for

$i = 1, 2, \dots, n$.

U is called a unisolvent family of degree n on $[a,b]$.

As in Chapter 1, we will generalize our problem to where we consider only members of U between our two curves ℓ and u . That is to say, our family of approximating functions $\mathcal{A}$ is a subset of U such that for $G(A,x) \in \mathcal{A}$,

(2-3) $u(x) \geq G(A,x) \geq \ell(x)$.

We will look at existence, characterization, and uniqueness of approximating a given real-valued continuous function f on X in the uniform norm with members of $\mathcal{A}$.

We should first note that each family $\mathcal{G}$ is also a family $\mathcal{F}$ of Chapter 1 (i.e., $G(A,x)$ being unisolvent of degree n implies $G(A,x)$ is varisolvent of degree n). Each member of $\mathcal{G}$ is continuous, satisfies property Z of fixed degree n , and is between the curves u and ℓ . Each member of $\mathcal{G}$ is locally solvent by its solvency property ((2-2)), and Theorem 5, page 460 of Tornheim [26].

We will at times call $\mathcal{G}$ a restricted unisolvent family, even though, $\mathcal{G}$, itself, may not be a unisolvent family.

Section 2-2: Existence, Characterization, and Uniqueness of Best Approximations.

Although we do not have existence of best approximations from a varisolvent family, for $\mathcal{G}$ (our subset of a unisolvent family), best approximations always exist. Tornheim's Theorem 7 [26] gives existence of best approximations for a unisolvent family. Since X is compact and our ℓ and u are sufficiently nice the argument of Theorem 7 applies for our restricted family $\mathcal{G}$ as well. We include it here for completeness.

Theorem 2-1: There exists a best approximation to f on X from $\mathcal{G}$.

Proof: There exists a sequence $\{G(A_n, x)\}$ in $\mathcal{G}$ such that

$$\lim_{n \rightarrow \infty} \|G(A_n, x) - f(x)\| = \inf_{G(A, x) \in \mathcal{G}} \|G(A, x) - f(x)\|.$$

Choose n distinct points $x_1, x_2, \dots, x_n$ in X . By unisolvency, there exists a one to one correspondence between the functions $G(A, x)$ of $\mathcal{B}$ and the set of values $y_1, \dots, y_n$ taken by the functions $G(A, x)$ at $x_1, \dots, x_n$. Let $G(A_n, x_i) = y_{ni}$. Then there is a subsequence $G(A_{n_j}, x_i)$ of $G(A_n, x_i)$ for which $y_{n_j i}$ is convergent for $i = 1, 2, \dots, n$ converging to y_i' . By the nature of u and ℓ , $\ell(x) \leq y_{n_j i} \leq u(x)$ implies $\ell(x) \leq y_i' \leq u(x)$. Let $G(A_*, x_i) = y_i'$. Then by Tornheim's Theorem 5, page 460 in [26], $G(A_{n_j}, x)$ converges uniformly to $G(A_*, x)$ on $X \subset [a, b]$. But again $\ell(x) \leq G(A_{n_j}, x) \leq u(x)$ for all $x \in X$ implies $\ell(x) \leq G(A_*, x) \leq u(x)$ for all $x \in X$. Therefore $G(A_*, x) \in \mathcal{B}$ and is a best approximation to f on X from $\mathcal{B}$.

For the characterization of best approximations from $\mathcal{B}$, we must consider, as in Chapter 1, the placement of critical points. The following lemma is Lemma 1-2 for $\mathcal{B}$, since $\mathcal{B}$ is also a family $\mathcal{F}$ from Chapter 1.

Lemma 2-1: For $G(A, x) \in \mathcal{B}$, if $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) \neq \phi$, then $G(A, x)$ is a best approximation to f on X from $\mathcal{B}$.

Therefore we need only consider the case when

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi.$$

Theorem 1-2 and Corollary 4 to Theorem 1-2 apply for $\mathcal{A}$ (considered as an $\mathcal{F}$ family of Chapter 1). However we can combine these two results for $\mathcal{A}$ if we note that $\mathcal{A}$ is closed under pointwise convergence. Theorem 5 of Tornheim [26] tells us that pointwise convergence gives us uniform convergence and by our selection of X , u , and ℓ , uniform convergence gives us the closure of $\mathcal{A}$. This argument is essentially that of Theorem 2-1 and gives us

Theorem 2-2: Let $G(A,x) \in \mathcal{A}$, $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$ and assume that X contains at least $2n$ points. Then $G(A,x)$ is a best approximation to f on X if and only if $G(A,x) - f(x)$ alternates n times on X .

R. Barrar and H. Loeb [26] proved Theorem 2-2 for an unrestricted unisolvent family omitting the fact that X must contain at least $2n$ points (we of course always assume X contains at least n points). Their proof extends directly to our setting.

Theorem 2-3: Let $G(A,x) \in \mathcal{A}$ and $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$. Then $G(A,x)$ is a best approximation to f on X if and only if $G(A,x) - f(x)$ alternates n times on X .

Proof: We need only show that a constant error curve for the best approximation can not occur, since Theorem 1-2 applies ($\mathcal{A}$ being an $\mathcal{A}$ family). We will show that a constant error curve can not exist for any n (the degree of the unisolvent family U) by induction on n .

For $n = 1$, Corollary 1 of Theorem 1-2 applies and a constant error curve for the best approximation from a restricted unisolvent family of degree 1 can not occur.

Assume that a constant error curve can not occur for restricted unisolvent families of degree $n \leq k - 1$. Let $\mathcal{A}^*$ be a restricted unisolvent family of degree k . Let $G(A_*, x)$ be the best approximation to f on $X \subset [a, b]$ from $\mathcal{A}^*$. Assume without loss of generality that $f(x) - G(A_*, x) \equiv c > 0$ on X . Let

$$\hat{\mathcal{A}} = \{G(A, x) \in \mathcal{A}^* : G(A, x_1) = G(A_*, x_1)\}$$

$$\text{where } x_1 = \max_{x \in X} \{x\}.$$

If $x_1 < b$, the behavior of a member of $\hat{\mathcal{A}}$ or $\mathcal{A}^*$ on $(x_1, b]$ will not affect its norm with f on X . Therefore let

$$\hat{\hat{\mathcal{A}}} = \hat{\mathcal{A}}|_{[a, x_1]}. \quad \text{Then for } 0 < \epsilon < x_1 - a, \hat{\hat{\mathcal{A}}}$$

is a restricted unisolvent family of degree $k - 1$ on $[a, x_1 - \epsilon]$. Let $G_\epsilon(A, x)$ be the best approximation to f on $X \cap [a, x_1 - \epsilon]$ from $\hat{\hat{\mathcal{A}}}$. By induction and Tornheim's results, $f(x) - G_\epsilon(A, x)$ alternates at least

once on $X \cap [a, x_1 - \epsilon]$. If we let

$$\|\phi\|_{\epsilon} = \max_{x \in [a, x_1 - \epsilon]} |\phi(x)|,$$

then for $\epsilon \leq \epsilon_0$,

$$\|f(x) - G_{\epsilon}(A, x)\|_{\epsilon_0} \leq \|f(x) - G_{\epsilon}(A, x)\|_{\epsilon} < c$$

(since $G(A_*, x) \in \hat{\mathcal{H}}$). Hence by a compactness argument using Tornheim's Theorem 5, we may assume

$$\lim_{\epsilon \rightarrow 0} \|G_{\epsilon}(A, x) - \hat{G}(A, x)\|_0 = 0$$

where $\hat{G}(A, x) \in \hat{\mathcal{H}}$. The claim is that $\hat{G}(A, x)$ is also a best approximation to f on X from $\mathcal{H}^*$. If this were false, there would exist an $x_2 \in [a, x_1)$ such that $|f(x_2) - \hat{G}(A, x_2)| > c$. But for small ϵ , $x_2 \in [a, x_1 - \epsilon]$ which implies that

$$|f(x) - G_{\epsilon}(A, x)| < c.$$

Taking the limit as $\epsilon \rightarrow \infty$ gives us a contradiction.

But by the nature of $\hat{G}(A, x)$,

$$f(x) - \hat{G}(A, x) \neq c_1,$$

c_1 a constant. Corollary 1 of Theorem 1-4 then gives us a contradiction.

Therefore constant error curve best approximations can not occur from restricted unisolvent families. ■

Example 1-1 shows that the best approximation to f

on X from $\mathcal{A}$ need not be unique when

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) \neq \phi.$$

But when $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, we have

Theorem 2-4: $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$ implies

there exists a unique best approximation to f on X from $\mathcal{A}$.

Proof: The proof follows directly from Theorem 2-3 using the standard uniqueness argument. **I**

CHAPTER 3

APPROXIMATION FROM MEMBERS OF A VARISOLVENT FAMILY HAVING RESTRICTED RANGES WITH THE ADDITIONAL HYPOTHESIS OF A HAAR SUBSPACE EXISTING FOR EACH MEMBER OF THE FAMILY

Section 3-1: Introduction and Definitions.

In this chapter we will give $\mathcal{H}$ (of Chapter 1) an additional hypothesis, namely a Haar subspace. This will eliminate the constant error curve condition in the characterization theorem for $\mathcal{H}$ and allow us to give strong unicity and continuity of the best approximation operator theorems. R. Barrar and H. Loeb [3] did this for an unrestricted family. We will generalize their paper to our setting.

We first require some definitions. Let P be an open subset of E_n . Let $\mathcal{G}^*$ be the family of functions

$$\{F(A,x): Px[a,b] \rightarrow R\}$$

where, for $A = (a_1, \dots, a_n) \in P$, each $\frac{\partial F(A,x)}{\partial a_i}$ for

$i = 1, 2, \dots, n$ as well as $F(A,x)$ is continuous in A and x .

Definition: Let $g_1, g_2, \dots, g_n$ be a set of continuous functions on $[a,b]$. Then $\{g_i\}_{i=1}^n$ generates a Haar

subspace of dimension n if $g_1, g_2, \dots, g_n$ forms a linearly independent set of functions and the only linear combination of $g_1, g_2, \dots, g_n$ having n or more zeros is the zero function. (For more information on Haar systems see [5], [9] and [13]). The Haar subspace condition that we will use is

$$(3-1) \quad \text{For } F(A, x) \in \mathcal{G}^*, \left\{ \frac{\partial F(A, x)}{\partial a_i} \right\}_{i=1}^n$$

generates a Haar subspace (Haar system) of dimension $m(A)$.

We further require that for $F(A, x) \in \mathcal{G}^*$, $F(A, x)$ has property Z of degree $m(A)$. Property Z (from Chapter 1) for members of $\mathcal{G}^*$ is

(1-3)' $F(A, x) \in \mathcal{G}^*$ will have property Z of degree $m(A)$ on $[a, b]$, i.e., for any $F(A_*, x) \neq F(A, x)$, $F(A_*, x) \in \mathcal{G}^*$, we have that $F(A_*, x) - F(A, x)$ has at most $m(A) - 1$ zeros on $[a, b]$ ($m(A) \geq 1$).

Again as in the previous chapters, our approximating family $\mathcal{G}^*$ will be the subset of $\mathcal{G}^*$ lying between the two curves ℓ and u (defined in Chapter 1).

Our family $\mathcal{G}^*$ differs from $\mathcal{G}$ of Chapter 1 in that we have replaced local solvency (or 1-2)) by the Haar subspace condition (3-1). That (3-1) is at least as strong a condition as (1-2) will be shown by Lemma 3-1

where we find that members of $\mathcal{F}^*$ satisfy (1-2).

R. Barrar and H. Loeb [3] note that exponential and rational families satisfy this Haar subspace condition (or (3-1)).

The next two lemmas are found in [3]. They apply for $\mathcal{H}^*$. We include their proofs for completeness.

Lemma 3-1: Let $A_* = (a_1^*, \dots, a_n^*)$ and $m(A_*) = q$.

Further, let $x_1, \dots, x_q$ be distinct points in X such that

$$F(A_*, x_i) = c_i \quad i = 1, 2, \dots, q$$

Then for sufficiently small $\epsilon > 0$, there exists a $\delta(\epsilon) > 0$ such that the equations

$$(1) \quad F(A, x_i) = \hat{c}_i \quad i = 1, 2, \dots, q$$

where $|c_i - \hat{c}_i| \leq \delta$ have a solution $A = (a_1, \dots, a_n) \in P$

where $a_i = a_i^*$ for $n \geq i \geq q + 1$ and

$$(2) \quad \|A - A_*\| \leq \epsilon$$

Proof: Let $f_i(a_1, \dots, a_q, c_1, \dots, c_q) \equiv$

$$F(a_1, \dots, a_q, a_{q+1}^*, \dots, a_n^*, x_i) - c_i$$

for $i = 1, 2, \dots, q$. Since the $\frac{\partial F(A_*, x)}{\partial a_i}$ $i = 1, 2, \dots, q$

form a Haar subspace, for $\delta(\epsilon) > 0$ sufficiently small

(or $|c_i - \hat{c}_i| \leq \delta$), we may apply the implicit function

theorem to the f_i system of equations in order to solve

for $a_1, \dots, a_q$ (where $c_i = \hat{c}_i$ for $i = 1, 2, \dots, q$).

The implicit function theorem can be invoked since the Haar subspace implies that the Jacobian of the transformation is non-zero. Therefore there exist a_i , $i = 1, \dots, q$ which solve the system

$$F(a_1, \dots, a_q, a_{q+1}^*, \dots, a_n^*, x_i) - \hat{c}_i = 0 \quad i = 1, \dots, q$$

where $|a_i - a_i^*| \leq \epsilon$ for $i = 1, 2, \dots, q$. Since P is open, ϵ and consequently δ can both be taken small enough to ensure that $A = (a_1, \dots, a_n) \in P$. This proves the lemma. **I**

For the next lemma we should recall the definition of a normal point of $\mathcal{S}^*$. $F(A_*, x) \in \mathcal{S}^*$ is called a normal point if $m(A_*) = N = \text{maximal value of } m(A) \text{ for } A \in P$.

Lemma 3-2: For Lemma 3-1, $F(A_*, x)$ being a normal point implies that there is a unique $F(A, x)$ satisfying (1) of Lemma 3-1.

Proof: Since for each $A \in P$, $m(A) \leq N = m(A_*)$, it follows that if both $F(A, x)$ and $F(A_1, x)$ satisfy (1), they agree at N points. (1-3)' then implies that $F(A, x) \equiv F(A_1, x)$. **I**

We are now ready to state and prove the lemma which will assure that a constant error curve for the best approximation from $\mathcal{S}^*$ can not occur. It is Lemma 3

of [3], modified for our setting. Before we state and prove the lemma, however, we should note that Lemma 3-1 implies that the members of $\mathcal{H}^*$ are locally solvent (or satisfy (1-2)). Therefore the results for $\mathcal{F}^0$ apply for $\mathcal{F}^*$ (since $\mathcal{H}^*$ is a varisolvent family V of Chapter 1).

Lemma 3-3: If $F(A_*, x) \in \mathcal{F}^*$ is a best approximation to f on X from $\mathcal{F}^*$ such that

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi ,$$

then a constant error curve cannot exist.

Proof: Without loss of generality, assume $F(A_*, x) - f(x) \equiv c > 0$ on X . Now a Haar subspace always has a strictly negative function (see [9]). Therefore there exist scalars $\{a_i\}_1^\ell$ such that

$$\sum_{i=1}^{\ell} a_i \frac{\partial F(A_*, x)}{\partial a_i^*} < 0 .$$

Let $A = (a_1, a_2, \dots, a_\ell, 0, 0, \dots) \in E_n$.

Then

$$0 > \sum_{i=1}^{\ell} a_i \frac{\partial F(A_*, x)}{\partial a_i^*} = \left(\sum_{i=1}^{\ell} a_i \frac{\partial F(A_* + tA, x)}{\partial (a_i^* + t a_i)} \right) \Big|_{t=0} = \frac{d F(A_* + tA, x)}{dt} \Big|_{t=0}$$

for all $x \in [a, b]$. By the mean value theorem, for

sufficiently small t , we have that

$$F(A_* + t A, x) < F(A_*, x)$$

for all $x \in [a, b]$. Now $F(A_*, x)$ is continuous with respect to A_* and P is open. Therefore there exists a $t_1 > t > 0$ such that

$$\|F(A_* + t A, x) - f(x)\| < \|F(A_*, x) - f(x)\|.$$

Because $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$, t can also be chosen so small that $F(A_* + t A, x) \in \mathcal{F}^*$ since

$$f(x) < F(A_* + t A, x) < F(A_*, x)$$

for t sufficiently small. ■

Combining Lemma 3-3 and Theorem 1-2, we have

Theorem 3-1: Let $F(A, x) \in \mathcal{F}^*$ satisfy

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi.$$

Then $F(A, x)$ is a best approximation to f on X from $\mathcal{F}^*$ if and only if $F(A, x) - f(x)$ alternates at least $m(A)$ times on X .

Uniqueness of best approximating functions from $\mathcal{F}^*$ for

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$$

follows from Lemma 3-3 and the fact that members of

$\mathcal{F}^*$ are members of $\mathcal{F}$.

We will now reproduce Theorem 2 of [3] in our setting since it will be necessary in the proofs of the next section.

Theorem 3-2: Let $F(A_*, x) \in \mathcal{F}^*$ be a best approximation to f on X where

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$$

and $F(A_*, x)$ is normal. Let the sequence

$$\{F(A_s, x)\}, \quad F(A_s, x) \in \mathcal{F}^*$$

be such that

$$\lim_{s \rightarrow \infty} \|F(A_s, x) - f(x)\| = \|F(A_*, x) - f(x)\|$$

and assume

(3-2) for $\epsilon > 0$, there exists an $M(\epsilon) > 0$ such that

for all $s > M$ and all

$$x \in X, \quad \ell(x) - \epsilon \leq F(A_s, x) \leq u(x) + \epsilon$$

then we can find a sequence $\{A_p'\} \subset P$ such that

$$\lim_{p \rightarrow \infty} \|A_* - A_p'\| = 0,$$

where the sequence $\{A_p'\}$ is equivalent to a subsequence of the $\{A_s\}$, and the last $n - N$ (recall that $m(A_*) = N$) components of each A_p' agree with the corresponding components of A_* .

Proof: Without loss of generality, assume $f(x) \equiv 0$ on X .

Let $\{x_i: a \leq x_1 < x_2 < \dots < x_{N+1}, x_i \in X\}$ be a

critical point set for $F(A_*, x)$. Let $\{F(A_p, x)\}$ be a subsequence of $\{F(A_s, x)\}$ which converges at these critical points. Call the limits at these $N + 1$ points $F(x_j)$. Now by (3-2) and the assumption that $\|F(A_s, x)\|$ tends to $\|F(A_*, x)\|$,

$$(3-3) \quad \max (\ell(x_j), -\|F(A_*, x)\|) \leq F(x_j) \\ \leq \min (u(x_j), \|F(A_*, x)\|),$$

We wish to show that $F(x_j) = F(A_*, x_j)$ for

$j = 1, 2, \dots, N + 1$. If this is true, Lemma 3-1 and Lemma 3-2 imply that a sequence $\{A_p'\}$ equivalent to a subsequence of $\{A_s\}$ can be found such that

$$\lim_{p \rightarrow \infty} \|A_* - A_p'\| = 0$$

and the last $n - N$ components of each A_p' agree with the corresponding components of A_* , proving the theorem.

Assume for concreteness that $x_1 \in X_{-1} \cup X_{-2}$.

Also assume that $F(x_j) \neq F(A_*, x_j)$ for some x_j . Therefore let $F(x_{N+1}) \neq F(A_*, x_{N+1})$ (the method will apply for any other x_j). Let $c = |F(A_*, x_{N+1}) - F(x_{N+1})| > 0$.

By means of a construction we will apply the unrestricted form of Lemma 1-1 and arrive at a contradiction. Let

$$\alpha = \min_{1 \leq j \leq N+1} \{ |F(A_*, x_j) - F(x_j)| : F(A_*, x_j) \neq F(x_j) \}.$$

We will construct a function $F(A, x)$ belonging to $\mathcal{S}^*$

from our local solvency property in the following way.

For $0 < \epsilon < c/2$, there exists a δ sufficiently small where $0 < \delta < \min(\alpha, \epsilon)$ and an $A \in P$ such that (by Lemma 3-1),

$$(3-4) \quad (a) \quad F(A, x_j) = \begin{cases} F(A_*, x_j) + (-1)^{j+1}(\delta/2) & \text{where} \\ & F(A_*, x_j) \neq F(x_j) \\ \\ F(A_*, x_j) + (-1)^j(\delta/2) & \text{where} \\ & F(A_*, x_j) = F(x_j) \end{cases}$$

for $j = 1, 2, \dots, N$

and

$$(b) \quad \|F(A, x) - F(A_*, x)\| \leq \epsilon < c/2.$$

We will consider only p so large that

$$(3-5) \quad |F(A_p, x_j) - F(x_j)| \leq \delta/4 \quad j = 1, 2, \dots, N+1$$

Now we have constructed $F(A, x)$ so that $F(A, x) - F(A_p, x)$

changes sign at the x_j , $j = 1, 2, \dots, N$.

i.e.,

$$(3-6) \quad \text{sgn} (F(A, x_j) - F(A_p, x_j)) = \\ - \text{sgn} (F(A, x_{j-1}) - F(A_p, x_{j-1})) \quad \text{for } j = 2, \dots, N$$

To see this, for concreteness assume that $j = k$ and

$x_k \in X_{+1} \cup X_{+2}$ where

$$F(x_k) = F(A_*, x_k).$$

Then by our construction

$$F(x_k) = F(A_*, x_k) = F(A, x_j) + \delta/2$$

Since

$$|F(x_k) - F(A_p, x_k)| \leq \delta/4$$

we have

$$F(A_p, x_k) > F(A, x_k)$$

There exist two cases for x_{k+1}

Case 1: $F(x_{k+1}) = F(A_*, x_{k+1})$. By (3-4) (a) we have

$$F(A, x_{k+1}) = F(A_*, x_{k+1}) + \delta/2 > F(x_{k+1}) + \delta/4 \quad \text{or}$$

$$F(A_p, x_{k+1}) < F(A, x_{k+1})$$

Case 2: $F(x_{k+1}) \neq F(A_*, x_{k+1})$. By (3-4) (a)

we have

$$F(A_*, x_{k+1}) = F(A, x_{k+1}) + \delta/2 .$$

But $F(A_*, x_{k+1}) - F(x_{k+1}) \geq \alpha$ implies that

$$F(A_*, x_{k+1}) \geq \alpha + F(x_{k+1}) \geq \delta + F(x_{k+1}).$$

Then $F(x_{k+1}) + \delta/2 \leq F(A, x_{k+1})$ or

$$F(A_p, x_{k+1}) < F(A, x_{k+1})$$

Therefore

$$\operatorname{sgn} (F(A, x_k) - F(A_p, x_k)) = - \operatorname{sgn} (F(A, x_{k+1}) - F(A_p, x_{k+1}))$$

when $x_k \in X_{+1} \cup X_{+2}$ and $F(x_k) = F(A_*, x_k)$.

If $x_k \in X_{-1} \cup X_{-2}$ or $F(x_k) \neq F(A_*, x_k)$, the argument is similar.

Now (3-6) holds for $j = 2, \dots, N$. But it also holds for $j = N + 1$ by (3-3), (3-4)(b) and (3-5). But $F(A, x)$ and $F(A_p, x)$ both belong to $\mathcal{H}_*$ and by the unrestricted form of Lemma 1-1,

$$F(A, x) - F(A_p, x) \neq 0, (F(A, x_{N+1}) \neq F(A_p, x_{N+1})),$$

can have at most $N - 1$ zeros on $[a, b]$, since

$N = \max m(A)$ for $A \in P$. (3-6) contradicts this and therefore

$$F(A_*, x_{N+1}) = F(x_{N+1})$$

This is the desired result. ■

Section 3-2: Strong Unicity and Continuity of the Best Approximation Operator Theorems.

In this section we will extend Theorems 3 and 4 of [3] to our setting

Theorem 3-3: If $F(A_*, x)$ is a best approximation to f on X from $\mathcal{H}_*$,

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \emptyset,$$

and $F(A_*, x)$ is normal, then there exists an $\alpha > 0$

such that for each $F(A, x) \in \mathcal{F}^*$,

$$\|f(x) - F(A, x)\| \geq \|f(x) - F(A_*, x)\| + \alpha \|F(A, x) - F(A_*, x)\|$$

Proof: Assume $f \notin \mathcal{F}^*$. If Theorem 3-3 is false, a sequence $\{F(A_n, x)\} \subset \mathcal{F}^*$ and a sequence $\{\alpha_n\}$, $\alpha_n > 0$ can be found such that $\lim_{n \rightarrow \infty} \alpha_n = 0$ so that

$$F(A_n, x) \neq F(A_*, x)$$

and such that

$$(3-7) \quad \|f(x) - F(A_n, x)\| = \|f(x) - F(A_*, x)\| + \alpha_n \|F(A_n, x) - F(A_*, x)\|$$

We claim $\{\|F(A_n, x)\|\}$ is bounded.

Consider

$$(3-8) \quad \|F(A_*, x) - F(A_n, x)\| - \|f(x) - F(A_*, x)\| \leq \|f(x) - F(A_*, x)\| + \alpha_n \|F(A_*, x) - F(A_n, x)\|$$

which is true by

$$\begin{aligned} & \|F(A_*, x) - F(A_n, x)\| - \|f(x) - F(A_*, x)\| \leq \\ & \|F(A_*, x) - F(A_n, x)\| - (f(x) - F(A_*, x)) = \\ & \|f(x) - F(A_n, x)\| \end{aligned}$$

and (3-7). Now

$$\|F(A_*, x) - F(A_n, x)\| \neq 0$$

since $F(A_n, x) \neq F(A_*, x)$. Divide both sides of

(3-8) by $\|F(A_*, x) - F(A_n, x)\|$. We have

$$1 - \frac{\|f(x) - F(A_*, x)\|}{\|F(A_*, x) - F(A_n, x)\|} \leq \frac{\|f(x) - F(A_*, x)\|}{\|F(A_*, x) - F(A_n, x)\|} + \alpha_n$$

If $\{\|F(A_n, x)\|\}$ is not bounded, then $1 \leq \alpha_n$ as

$n \rightarrow \infty$, since $\|f(x) - F(A_*, x)\|$ is bounded, which is a contradiction. Therefore assume $\{\|F(A_n, x)\|\}$ is bounded.

Then by (3-7),

$$(3-9) \quad \lim_{n \rightarrow \infty} \|f(x) - F(A_n, x)\| = \|f(x) - F(A_*, x)\|.$$

Apply Theorem 3-2 to (3-9), i.e., there exists a sequence $\{B_k\} \subset P$ converging to A_* where the sequence is equivalent to a subsequence of $\{A_n\}$ and the last $n - N$ components of each B_k agree with the corresponding components of A_* . We should note two things. First $F(B_k, x) \in \mathcal{F}^*$ since $F(A_n, x) \in \mathcal{F}^*$. Second (3-7) remains valid for $\{B_k\}$.

$$\text{Let } \sigma(x) = \begin{cases} -1 & \text{if } x \in X_{-1} \cup X_{-2} \\ 1 & \text{if } x \in X_{+1} \cup X_{+2} \end{cases}$$

By the assumption that $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$,

σ is well defined. Now, we wish to show that

$$(3-10) \quad \alpha_k \|F(B_k, x) - F(A_*, x)\| \geq$$

$$\max_{x \in X_{A_*}} \sigma(x) (F(A_*, x) - F(B_k, x))$$

We have that (3-7) is valid for $\{B_k\}$.

Therefore

$$\alpha_k \|F(B_k, x) - F(A_*, x)\| = \|F(x) - F(B_k, x)\| - \|f(x) - F(A_*, x)\|$$

Now $\alpha_k \|F(B_k, x) - F(A_*, x)\| > 0$ for each k since

$F(A_*, x)$ is the best approximation to f from $\mathcal{F}^*$ and

$F(B_k, x) \in \mathcal{F}^*$. We now consider cases

Case 1: Let $x_1 \in X_{A_*}$ such that $F(A_*, x_1) = \ell(x_1)$

(i.e., $\sigma(x_1) = 1$) . Then since

$$F(B_k, x) \in \mathcal{F}^*, \quad F(B_k, x_1) - F(A_*, x_1) \geq 0$$

$$\text{or } \sigma(x_1) (F(A_*, x_1) - F(B_k, x_1)) \leq 0$$

Case 2: Let $x_1 \in X_{A_*}$ such that $F(A_*, x_1) = u(x_1)$

(i.e., $\sigma(x_1) = -1$) . Then since

$$F(B_k, x) \in \mathcal{F}^*, \quad F(B_k, x_1) - F(A_*, x_1) \leq 0$$

$$\text{or } \sigma(x_1) (F(A_*, x_1) - F(B_k, x_1)) \leq 0$$

Case 3: Let $x_1 \in X_{A_*}$ such that

$$f(x_1) - F(A_*, x_1) = \pm \|f(x) - F(A_*, x)\| .$$

Then

$$\alpha_k \|F(B_k, x) - F(A_*, x)\| =$$

$$\|f(x) - F(B_k, x)\| - \|f(x) - F(A_*, x)\|$$

$$\geq \sigma(x_1) (f(x_1) - F(B_k, x_1)) - \|f(x) - F(A_*, x)\|$$

$$= \sigma(x_1) (f(x_1) - F(B_k, x_1)) - \sigma(x_1) (f(x_1) - F(A_*, x_1))$$

$$= \sigma(x_1) (F(A_*, x_1) - F(B_k, x_1))$$

We therefore have that for all $x_1 \in X_{A_\star}$,

$$\begin{aligned} \alpha_k \|F(B_k, x) - F(A_\star, x)\| \\ \geq \sigma(x_1) (F(A_\star, x_1) - F(B_k, x_1)) \end{aligned}$$

Therefore,

$$\begin{aligned} \alpha_k \|F(B_k, x) - F(A_\star, x)\| \\ \geq \max_{x \in X_{A_\star}} \sigma(x) (F(A_\star, x) - F(B_k, x)) \end{aligned}$$

or (3-10) is shown to be true.

We now wish to show that there exists a $\gamma > 0$ such that for all k ,

$$(3-11) \quad \max_{x \in X_{A_\star}} \{\sigma(x) (F(A_\star, x) - F(B_k, x))\} \geq \gamma \|B_k - A_\star\|$$

Assume (3-11) is false. Then there exists a sequence of positive $\{\gamma_k\}$ such that γ_k tends to 0 and a subsequence of $\{B_k\}$ such that

$$(3-12) \quad \max_{x \in X_{A_\star}} \sigma(x) \frac{(F(A_\star, x) - F(B_k, x))}{\|A_\star - B_k\|} \leq \gamma_k$$

By the mean value theorem for large k ,

$$(3-13) \quad \max_{x \in X_{A_\star}} \sigma(x) \left[\sum_{i=1}^N \frac{\partial F(A_k(x), x)}{\partial a_i} \frac{(a_i^\star - b_{ki})}{\|A_\star - B_k\|} \right] \leq \gamma_k$$

where $B_k = (b_{k1}, \dots, b_{kn})$, $N = m(A_\star)$ and $A_k(x) \in P$ is on the line between B_k and $A_\star$ (k large enough and P

open says that $A_k(x) \in P$). Set $c_k = \frac{A_* - B_k}{\|A_* - B_k\|}$.

Since $\|c_k\| = 1$ for all k , and we are on a compact set, we have a convergent subsequence (which we will not relabel) where this subsequence c_k converges to $c = (c_1, \dots, c_n)$ and $\|c\| = 1$. Using this subsequence in (3-13) and taking limits, we have

$$(3-14) \quad \max_{x \in X_{A_*}} \sigma(x) \sum_{i=1}^N c_i \frac{\partial F(A_*, x)}{\partial a_i} \leq 0$$

Now $\sum_{i=1}^N c_i \frac{\partial F(A_*, x)}{\partial a_i}$ is a non-zero function because

of linear independence and the fact that $\|c\| = 1$.

By (3-14) and Theorem 3-1,

$$\sum_{i=1}^N c_i \frac{\partial F(A_*, x)}{\partial a_i}$$

has at least N zeros which contradicts our Haar subspace hypothesis. Therefore (3-11) is proven.

When we combine (3-10) and (3-11), we have that

$$(3-15) \quad \alpha_k \|F(B_k, x) - F(A_*, x)\| \geq \gamma \|B_k - A_*\|$$

But since $B_k \rightarrow A_*$, by the mean value theorem, there is a $D > 0$ such that for sufficiently large k ,

$$(3-16) \quad \|F(B_k, x) - F(A_*, x)\| \leq D \|B_k - A_*\|$$

From (3-15) and (3-16) we have that

$$\alpha_k \geq \frac{\gamma}{D} > 0$$

which says that $\alpha_k \rightarrow 0$ is impossible. This contradicts our original assumption that the theorem was false. **I**

It is obvious that $F(A, x)$ must belong to $\mathcal{F}^*$ and not just to $\mathcal{H}^*$ in Theorem 3-3 as the following example shows.

Example 3-1: Let $X = [-1, 1]$, $f(x) = x^2$, $u(x) = x^2 + 1/4$, $\ell(x) = -1$ and $\{F(A, x): F(A, x) = ax + b\} = \mathcal{H}^*$.

Then $F(A_*, x) \equiv 1/4$ is the best approximation to $f(x)$ on X from $\mathcal{F}^*$, but for $F(A_1, x) \equiv 1/2 \in \mathcal{H}^*$, we have that

$$\|f(x) - F(A_1, x)\| \neq \|f(x) - F(A_*, x)\| + \gamma \|F(A_*, x) - F(A_1, x)\|$$

for $\gamma > 0$ since $1/2 \neq 3/4 + \gamma \cdot 1/4$ for $\gamma > 0$.

The last theorem of this chapter is a continuity of the best approximation operator theorem. For $\mathcal{H}^*$, the theorem is Theorem 4 of [3]. The proof is based on the strong uniqueness theorem and compactness. Their proof applies in our setting and we have the following theorem.

Theorem 3-4: Let $F(A_*, x)$ be a best approximation to f

on X from $\mathcal{F}^*$ such that $F(A_*, x)$ is normal and

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi .$$

Then,

(1) There exists a $\gamma > 0$ such that $\|f(x) - g(x)\| < \gamma$ implies g has a best approximation on X from $\mathcal{F}^*$, say $F(A_g, x)$

(2) There exists a $\lambda > 0$ such that for all $g(x)$ of (1) above which have a best approximation

$$F(A_g, x), \quad \|F(A_*, x) - F(A_g, x)\| \\ \leq \lambda \|g(x) - f(x)\|$$

Theorem 3-4 is a local continuity theorem. (1) of Theorem 3-4 says that g must be sufficiently close to f to apply the theorem. As an example of a global continuity theorem, see C. Dunham [7]. C. Dunham's theorem, however, assumes that pointwise convergence is uniform convergence (see the paragraphs following Corollary 3 of Theorem 1-2)

CHAPTER 4
GENERALIZED WEIGHT FUNCTION APPROXIMATION
WITH PROPERTY A AND RESTRICTION BETWEEN
THE CURVES u AND ℓ WHERE $u > \ell$

Section 4-1: Introduction and Definitions

D. Moursund [14], D. Moursund and G. D. Taylor [15], G. D. Taylor [22], I. Ninomiya [17], L. Wuytack [27], and C. Dunham [8] among others looked at a generalized weight function for an approximating family of functions. G. D. Taylor [22] examined a generalized weight function for polynomials in our restricted setting, while C. Dunham did likewise except that his approximating family was non-linear in a non-restricted setting. In this chapter we will combine these last two papers and look at a generalized weight function for a class of non-linear functions bounded between the curves u and ℓ .

Our generalized weight function will not be the least restrictive generalized weight function (see [14]). A function $W(x,y)$ mapping $W \times R$ into R will be called a generalized weight function if

- (4-1) (i) $\text{sgn } W(x,y) = \text{sgn } y$
(ii) W is continuous

- (iii) for each $x \in X$, W is a strictly monotone increasing function of y with

$$\lim_{|y| \rightarrow \infty} |W(x, y)| = \infty$$

For our family $\mathcal{F}$ (of Chapter 1), we say that $F(A_*, x) \in \mathcal{F}$ is a best generalized approximation to f with respect to W and $\mathcal{F}$ if

$$\sup_{x \in X} |W[x, f(x) - F(A_*, x)]| \leq \inf_{F(A, x) \in \mathcal{F}} \left(\sup_{x \in X} |W[x, f(x) - F(A, x)]| \right)$$

Our first observation is that if $W(x, y) = y$, our problem is the problem of Chapter 1. Even if $W(x, y) \neq y$ we will show that a form of Chapter 1 applies for our generalized approximation with respect to W and $\mathcal{F}$.

We now require a theorem of D. Moursund and G. D. Taylor [15].

Theorem 4-1: If W is continuous and $F(A, x)$ is vari-solvent (of degree $m(A)$), then W is varisolvent (of degree $m(A)$).

We also require a slightly altered definition of alternation. (see [22]).

Definition: For $F(A, x) \in \mathcal{F}$, we define $\text{sgn}^*(f(x) - F(A, x))$ by

$$\text{sgn}^*(f(x_i) - F(A, x_i)) = \begin{cases} -1 & \text{if } F(A, x) = u(x) \\ 1 & \text{if } F(A, x) = \ell(x) \\ \text{sgn}(f(x) - F(A, x)) & \text{otherwise} \end{cases}$$

Definition: For $F(A,x) \in \mathcal{F}$, the error curve

$$W[x, f(x) - F(A,x)]$$

is said to alternate $n + 1$ times on X if there exist

$n + 2$ points $x_1 < x_2 < \dots < x_{n+2}$ in X such that

$$\text{sgn}^*(f(x_i) - F(A,x_i)) = -\text{sgn}^*(f(x_{i+1}) - F(A,x_{i+1})) ,$$

$$i = 1, 2, \dots, n + 1$$

and at least one of the following conditions is satisfied by each x_i :

$$(i) \quad |W[x_i, f(x_i) - F(A,x_i)]| =$$

$$\max_{x \in X} |W[x, f(x) - F(A,x)]|$$

$$(ii) \quad F(A,x_i) = \ell(x_i)$$

$$(iii) \quad F(A,x_i) = u(x_i)$$

As before, the x_i are called critical points. An x_i is a positive critical point if

$$\text{sgn}^*(f(x_i) - F(A,x_i)) = 1$$

and a negative critical point if

$$\text{sgn}^*(f(x_i) - F(A,x_i)) = -1$$

Using Theorem 4-1 and our revised definition of alternation we could develop the results of Chapter 1. Rather than this however, we will examine a perhaps more general problem.

J. R. Rice has shown (page 18 of [20]) that the weakest hypothesis with property Z and continuity which

will ensure a Tchebyshev type of characterization theorem is property A (defined below). The difference between the above conditions and varisolvency is essentially the solvency condition (not considering the constant error curve possibility). On page 22 of [20], Rice gives two examples of families satisfying property A which are not varisolvent. It has not been shown yet whether varisolvency implies property A because of the constant error curve difficulty.

We will consider generalized weight function approximation using a family in our restricted setting, where the family has property A in place of local solvency. First, a theorem comparable to Theorem 4-1 will be given for a family with property A (instead of local solvency). We will then examine the questions of existence, uniqueness and characterization for our (property A) family with respect to W.

Let $\mathcal{H}_1^*$ be the family of functions

$$\{F^*(A,x): P \times [a,b] \rightarrow R\}$$

where each $F^*(A,x) \in \mathcal{H}_1^*$ satisfies (1-1), property Z (or (1-3)), and property A. Our approximating family $\mathcal{F}_1^*$ will be the subset of $\mathcal{H}_1^*$ satisfying (1-4) (i.e., members of $\mathcal{H}_1^*$ lying between the curves u and ℓ).

As for our family $\mathcal{F}_1^*$, $F^*(A,x) \in \mathcal{F}_1^*$ is a best generalized approximation to f with respect to W and $\mathcal{F}_1^*$ if

$$(4-2) \sup_{x \in X} |W[x, f(x) - F^*(a, x)]| \leq$$

$$\inf_{F^*(B, x) \in \mathcal{F}_1^*} (\sup_{x \in X} |W[x, f(x) - F^*(B, x)]|) = e$$

We will use our new definition of alternation for $\mathcal{F}_1^*$.

Thus there remains only the definition of property A.

Definition: $F^*(A, x)$ has property A of degree $m(A)$ if for any integer $m < m(A)$, any sequence $\{x_1, \dots, x_m\}$ with

$$a = x_0 < x_1 < \dots < x_{m+1} = b$$

any sign σ , and any real ϵ with

$$0 < \epsilon < \min \{x_{j+1} - x_j : j = 0, \dots, m\},$$

there exists an $F^*(B, x) \in \mathcal{F}_1^*$ such that

$$\|F^*(A, x) - F^*(B, x)\| < \epsilon$$

$$\begin{aligned} \operatorname{sgn} (F^*(A, x) - F^*(B, x)) &= \sigma \quad \text{for } a \leq x < x_1 - \epsilon \\ &= \sigma (-1)^j \quad \text{for } x_j + \epsilon < x < x_{j+1} - \epsilon \\ &= \sigma (-1)^m \quad \text{for } x_m + \epsilon < x \leq b \end{aligned}$$

For $m = 0$,

$$\operatorname{sgn} (F^*(A, x) - F^*(B, x)) = \sigma$$

Section 4-2: Existence of Best Approximations.

Before we consider the existence question we will give the analog of Theorem 4-1 in our setting (with property A).

Theorem 4-2: If W is continuous and $F^*(A,x) \in \mathcal{H}_1^*$,
then W has property Z and property A of degree $m(A)$.

Proof: (The proof parallels that of Theorem 4-1 in [15]). The proof that W has property Z of degree $m(A)$ can be found in [15] (it doesn't depend on property A).

We now show that W has property A of degree $m(A)$.
Let $k < m(A)$, $a = x_0 < x_1 < \dots < x_{k+1} = b$, and ϵ
such that $0 < \epsilon < \min \{x_{j+1} - x_j : j = 0, \dots, k\}$ be
given as well as any sign σ . Since $F^*(A,x)$ has property
 A of degree $m(A)$, there exists an $\epsilon_1 > 0$ and
 $F^*(B,x) \in \mathcal{H}_1^*$ such that

$$\|F^*(A,x) - F^*(B,x)\| =$$

$$\|[f(x) - F^*(A,x)] - [f(x) - F^*(B,x)]\| < \epsilon_1$$

where $\text{sgn} (F^*(A,x) - F^*(B,x)) = \sigma$ for $a \leq x < x_1 - \epsilon_1$

$$= \sigma (-1)^j \quad \text{for } x_j + \epsilon_1 < x < x_{j+1} - \epsilon_1$$

$$= \sigma (-1)^m \quad \text{for } x_m + \epsilon_1 < x \leq b$$

Now $\epsilon_1 < \epsilon$ can be chosen sufficiently small so that
by the continuity of W ,

$$\|W[x, f(x) - F^*(A,x)] - W[x, f(x) - F^*(B,x)]\| < \epsilon$$

But since $\text{sgn } W(x,y) = \text{sgn } y$,

$$\text{sgn} (W[x, f(x) - F^*(A,x)] - W[x, f(x) - F^*(B,x)]) =$$

$$\sigma \quad \text{for } a \leq x < x_1 - \epsilon$$

$$= \sigma (-1)^j \quad \text{for } x_j + \epsilon < x < x_{j+1} - \epsilon$$

$$= \sigma (-1)^m \quad \text{for } x_m + \epsilon < x \leq b$$

Therefore W has property A of degree $m(A)$. If W had property A of degree $m(A) + 1$, we would have a contradiction of property Z . This concludes the proof. $\blacksquare$

We now consider the existence question. Our discussion will parallel that of Chapter 1. We first note that existence of best approximations is not assured under ordinary approximation, and therefore will not be assured under generalized approximation. An example of the above non-existence of a best approximation is as follows.

Example: Let $P = \{a \mid a \text{ is rational}\}$ $\mathcal{A}_1^* = \{F^*(A, x) = a\}$, and f be defined by $f(x) = \sqrt{2}$. Then a best approximation to f from $\mathcal{A}_1^*$ does not exist. (see page 22 of [20]).

Therefore, as in Chapter 1, in this section we will add the hypothesis that $\mathcal{A}_1^*$ is closed under pointwise convergence. This will allow us to obtain existence of a best approximation.

For the ϵ defined in (4-2), let

$$(4-3) \quad \lim_{j \rightarrow \infty} \sup_{x \in X} |W[x, f(x) - F^*(A_j, x)]| = \epsilon$$

Our existence question will be answered in the affirmative if we can find an $F^*(A_r, x) \in \mathcal{A}_1^*$ such that

$$\sup_{x \in X} |W[x, f(x) - F^*(A_r, x)]| = \epsilon$$

Now, since $m(A)$ is bounded for all A , we can find a subsequence of $\{F^*(A_i, x)\}$ of fixed degree such that (4-3) holds (we will not relabel). Applying Theorem 7.2 of Rice [20] to our subsequence $\{F^*(A_i, x)\}$ satisfying (4-3) gives us a pointwise convergent subsequence satisfying (4-3). But our closure assumption on $\mathcal{F}_1^*$ then gives us that there exists an $F^*(A_s, x) \in \mathcal{F}_1^*$ which is the pointwise limit of the pointwise convergent subsequence satisfying (4.3). We must therefore have that

$$\sup_{x \in X} |W[x, f(x) - F^*(A_s, x)]| =$$

$$\max_{x \in X} |W[x, f(x) - F^*(A_s, x)]| = e < \infty$$

since W is continuous on a compact set. Therefore $F^*(A_s, x)$ is a best approximation to f with respect to W and $\mathcal{F}_1^*$. We have just proved the following theorem.

Theorem 4-3: For a generalized weight function W , there exists a best generalized approximation to f on X with respect to W and $\mathcal{F}_1^*$.

Section 4-3: Uniqueness and Characterization of Best Approximations.

Our first observation is that Example1-1 and Lemma1-2 apply in our setting so that we need consider only the

case where $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$. For the remainder of this section we therefore assume that $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$.

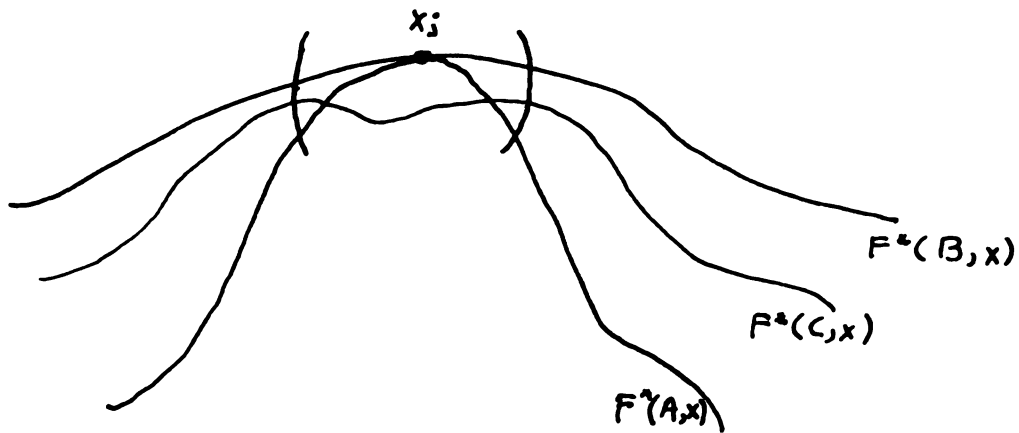
Uniqueness will follow directly from the characterization theorem in the usual manner (see the proof of Theorem 1-4).

For the characterization theorem we require a lemma which C. Dunham [8] gives without proof for the unrestricted setting. For completeness we will give a proof of the lemma in our setting.

Lemma 4-1: For $F^*(A, x) \in \mathcal{F}_1^*$, $F^*(A, x) - F^*(B, x)$ can have at most $m(A) - 1$ zeros, counting double zeros twice, for $F^*(A, x) \neq F^*(B, x) \in \mathcal{F}_1^*$.

Proof: The proof will consist of constructing an $F^*(C, x) \in \mathcal{F}_1^*$ such that $F^*(C, x) - F^*(A, x)$ can have as many distinct zeros as $F^*(B, x) - F^*(A, x)$ has zeros, counting double zeros twice. If x_j is a double zero of $F^*(B, x) - F^*(A, x)$ such that $F^*(B, x) > F^*(A, x)$ in a small neighborhood of x_j (not including x_j of course), $F^*(C, x)$ will be constructed close enough to $F^*(B, x)$ where $F^*(B, x) > F^*(C, x)$ in the neighborhood of x_j . This will ensure two distinct zeros of $F^*(C, x) - F^*(A, x)$

associated only with x_j i.e.,



For $F^*(B,x) < F^*(A,x)$ in a neighborhood of a double zero, $F^*(C,x)$ will be above $F^*(B,x)$ in a manner analogous to the above. Each simple zero of $F^*(B,x) - F^*(A,x)$ will also have a zero of $F^*(C,x) - F^*(A,x)$ associated only with it. We now proceed with the construction.

Let $\{x_i\}_1^k$ be the set of zeros of $F^*(A,x) - F^*(B,x)$.

Let

$$\rho = \min_{\substack{0 \leq i \leq k \\ x_i < x_{i+1}}} \max_{x \in [x_i, x_{i+1}]} |F^*(B,x) - F^*(A,x)| > 0$$

$$\text{where } x_0 = a \text{ and } x_{k+1} = b$$

and

$$\alpha = \min_{\substack{0 \leq i \leq k \\ x_i < x_{i+1}}} |x_{i+1} - x_i| > 0. \text{ Select an}$$

appropriate $\epsilon > 0$ such that $\alpha/2 > \epsilon > 0$ and

$\rho/4 > \epsilon > 0$. We then construct a set of points

$\{x_i'\}_{i=1}^{\ell}$, $\ell \leq k$ in the following way.

Case 1: Let x_1 be a double zero ($a = x_0 = x_1$ or $x_k = x_{k+1} = b$ implies a simple zero). Select an x_1' such that $x_0 + \epsilon < x_1' < x_1 - \epsilon$ (i.e., x_1' is between x_0 and x_1 and an ϵ distance away from both). Proceed to the first simple zero after x_1 (if it exists) and select x_2' as follows, i.e. if x_s is the first simple zero after x_1 , x_2' is selected such that

$$x_{s-1} + \epsilon < x_2' < x_s - \epsilon.$$

Proceed to the next zero, x_{s+1} . Select x_3' where $x_s + \epsilon < x_3' < x_{s+1} - \epsilon$. If x_{s+1} is a simple zero, proceed to x_{s+2} and do the same for x_{s+2} as for x_{s+1} . If x_{s+1} is a double zero, proceed to the next simple zero after x_{s+1} as in selecting x_2' . Continue through all the zeros of $F^*(A, x) - F^*(B, x)$.

Case 2: Let x_1 be a simple zero where $x_1 \neq x_0$. Proceed to x_2 . If x_2 is a double zero, proceed as in Case 1. If x_2 is a simple zero, proceed as if for x_1 in Case 3 below.

Case 3: Let x_1 be a simple zero where $x_1 \neq x_0$. Let x_1' be such that $x_0 + \epsilon < x_1' < x_1 - \epsilon$. Let x_2' be

such that $x_1 + \epsilon < x_2' < x_2 - \epsilon$ (naturally only if x_2 exists). If x_2 is simple proceed and repeat for x_3 . If x_2 is double, proceed as if x_2 is the double zero x_{s+1} in Case 1 above.

Let x_d be the first double zero. Suppose, for concreteness, that near x_j , $F^*(B, x) > F^*(A, x)$.

Select σ in property A below so that

$$\text{sgn} (F^*(B, x_d) - F^*(C, x_d)) = + 1$$

(If no double zero exists, property Z proves the lemma).

Now, for the points $\{x_i'\}_{i=1}^{\ell}$ and our $\epsilon > 0$, there

exists (by property A) an $F^*(C, x)$ such that

$$\begin{aligned} \text{sgn} (F^*(B, x) - F^*(C, x)) &= \sigma \quad \text{for } a \leq x < x_1' - \epsilon \\ &= \sigma (-1)^j \quad \text{for } x_j' + \epsilon < x < x_{j+1}' - \epsilon \\ &= \sigma (-1)^{\ell} \quad \text{for } x_{\ell} + \epsilon < x \leq b \end{aligned}$$

where $\|F^*(B, x) - F^*(C, x)\| < \epsilon$.

Now by our construction, $F^*(C, x) - F^*(A, x)$ has as many simple zeros as $F^*(B, x) - F^*(A, x)$ has zeros, counting double zeros twice. Applying property Z proves the lemma. 1

The statement and proof of the characterization theorem which now follows parallels that of Taylor [22] for the linear case.

Theorem 4-4: For a generalized weight function

W and $F^*(A, x) \in \mathcal{F}_1^*$ where

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi ,$$

$F^*(A, x)$ is a best generalized approximation to f with respect to W and $\mathcal{F}_1^*$ if and only if

$$W[x, f(x) - F^*(A, x)]$$

alternates at least $m(A)$ times . (Remember that $\mathcal{F}_1^* \neq \phi$ by assumption).

Proof: Assume $W[x, f(x) - F^*(A, x)]$ alternates at least $m(A)$ times on X . Assume for concreteness that it alternates exactly $m(A)$ times. Assume there exists an $F^*(B, x) \in \mathcal{F}_1^*$ such that

$$\max_{x \in X} |W[x, f(x) - F^*(B, x)]| <$$

$$\max_{x \in X} |W[x, f(x) - F^*(A, x)]| = \rho^*$$

Let $\{x_i: x_i < x_{i+1}\}$ be a critical point set of $m(A) + 1$ points in X for $W[x, f(x) - F^*(A, x)]$. Without loss of generality, assume $\text{sgn}^* (f(x_i) - F^*(A, x_i)) = (-1)^i$ for $i = 1, 2, \dots, m(A) + 1$.

Therefore for i odd, $W[x_i, f(x_i) - F^*(A, x_i)] = -\rho^*$ or $F^*(A, x_i) = u(x_i)$, and for i even,

$$W[x_i, f(x_i) - F^*(A, x_i)] = \rho^*$$

or $F^*(A, x_i) = \ell(x_i)$ since $(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi$.

By the monotone property of W (when needed), we have that $(-1)^{i+1} (F^*(A, x_i) - F^*(B, x_i)) \geq 0$ for $i = 1, 2, \dots, m(A)+1$.

Therefore $F^*(A, x) - F^*(B, x)$ has at least $m(A)$ zeros, which contradicts Lemma 4-1. Therefore if

$$W[x, f(x) - F^*(A, x)]$$

alternates $m(A)$ times on X , $F^*(A, x)$ is a best generalized approximation to f with respect to W and $\mathcal{F}_1^*$.

Assume $F^*(A, x) \in \mathcal{F}_1^*$ is a best generalized approximation to f with respect to W and $\mathcal{F}_1^*$ which alternates $k < m(A)$ times on X . Let $y_1 < y_2 \dots < y_{k+1}$ be a critical point set of X for $W[x, f(x) - F^*(A, x)]$. Fix i , $1 \leq i \leq k$ and suppose y_i is a positive critical point. Let $y' = \sup \{y: y \in X \text{ with } y_i \leq y \leq y_{i+1} \text{ and } y \text{ is a positive critical point}\}$ and let

$$y'' = \inf \{y: y \in X \text{ with } y_i \leq y \leq y_{i+1}$$

and y is a negative critical point}. Since

$$(X_{+1} \cup X_{+2}) \cap (X_{-1} \cup X_{-2}) = \phi,$$

X is compact, and all functions involved are continuous, we have that y' is a positive critical point, y'' a negative critical point, $y' < y''$, and $y \in (y', y'')$ implies y is not a critical point. If $X \cap (y', y'') \neq \phi$, pick an x_i in this intersection. Then at x_i ,

$$(i) \quad \rho = |W[x_i, f(x_i) - F^*(A, x_i)]| < \rho^*$$

$$\text{where } \rho^* - \rho = s > 0$$

$$(ii) \quad \ell(x_i) + s_i < F^*(A, x_i) < u(x_i) - s_i, \text{ where}$$

$$s > s_i > 0$$

Now by continuity, there exists an $\epsilon_i > 0$ such that on

$$[x_i - \epsilon_i, x_i + \epsilon_i], \ell(x) + s_i/2 < F^*(A, x) < u(x) - s_i/2$$

$$\text{and } |W[x, f(x) - F^*(A, x)]| < \rho^* - s_i/2 \text{ for all}$$

$$x \in X \cap [x_i - \epsilon_i, x_i + \epsilon_i].$$

$$\text{If } X \cap (y', y'') = \emptyset, \text{ let } x_i = \frac{(y' + y'')}{2}.$$

Then there exists an $\epsilon_i > 0$, since $y' < y''$ such that

$$X \cap [x_i - \epsilon_i, x_i + \epsilon_i] = \emptyset. \text{ Construct } x_i \text{ and } \epsilon_i \text{ in}$$

a similar manner for each $i = 1, 2, \dots, k$, letting

$$x_0 = \inf \{y: y \in X\} \text{ and } x_{k+1} = \sup \{y: y \in X\}.$$

$$\text{Let } \epsilon^k = \min_{1 \leq i \leq k} \{\epsilon_i\} > 0 \text{ and } s^k = \min_{1 \leq i \leq k} \{s_i/2\} > 0. \text{ Then}$$

$$\text{for } x_p \in \bigcup_{i=1}^k (X \cap [x_i - \epsilon^k, x_i + \epsilon^k]), \text{ we have}$$

$$\text{that } |W[x_p, f(x_p) - F^*(A, x_p)]| < \rho^* - s^k,$$

$$\text{and } \ell(x_p) + s^k < F^*(A, x_p) < u(x_p) - s^k. \text{ Now, by}$$

$$\text{construction } \{X \cap [x_i, x_{i+1}]\}_{i=0}^k \text{ is a set of } k+1$$

intervals in X each of which contains no alternations

and such that $W[x, f(x) - F^*(A, x)]$ alternates exactly once on any two intervals. Each $X \cap [x_i, x_{i+1}]$, $i = 0, \dots, k$ contains a critical point, namely y_{i+1} and no critical point of an opposite sign. Without loss of generality assume y_1 is a positive critical point. Then there exists a $\delta_1 > 0$ such that

$W[x, f(x) - F^*(A, x)] > -\rho^* + \delta_1$ and $F^*(A, x) < u(x) - \delta_1$ for all $x \in X \cap [x_0, x_1]$, since X is compact and all functions involved are continuous and $u(x) > \ell(x)$.

Now since W is a continuous and strictly monotone function of $f(x) - F^*(B, x)$ for $F^*(B, x) \in \mathcal{F}_1^*$, there exists an $0 < \epsilon_1 < \min(s^k, \delta_1/2)$ so small that if $|F^*(A, x) - F^*(A_1, x)| < \epsilon$ and $f(x) - F^*(A, x) >$

$f(x) - F^*(A_1, x)$ for all $x \in X \cap [x_0, x_1]$, then

$$(4-4) \quad |W[x, f(x) - F^*(A, x)] - W[x, f(x) - F^*(A_1, x)]| \leq s^k/3,$$

$$W[x, f(x) - F^*(A_1, x)] > -\rho^* + s^k/3, \text{ and}$$

$$F^*(A_1, x) \leq u(x) - \epsilon_1 \text{ for all } x \in X \cap [x_0, x_1].$$

Repeat this process, i.e., on $X \cap [x_1, x_2]$, there exists

a $\delta_2 > 0$ such that $F^*(A, x) > \ell(x) + \delta_2$ and

$$W[x, f(x) - F^*(A, x)] < \rho^* - \delta_2 \text{ for all } x \in X \cap [x_1, x_2].$$

Then there exists an $0 < \epsilon_2 < \min(s^k, \delta_2/2)$ so small

so that if $|F^*(A, x) - F^*(A_2, x)| < \epsilon_2$ and

$f(x) - F^*(A, x) < f(x) - F^*(A_2, x)$ for all $x \in X \cap [x_1, x_2]$,

we have

$$(4-5) \quad |W[x, f(x) - F^*(A, x)] - W[x, f(x) - F^*(A_2, x)]| \leq s^k/3, \quad$$

$$W[x, f(x) - F^*(A_2, x)] < \rho^* - s^k/3, \quad \text{and}$$

$$F^*(A_2, x) \geq \ell(x) - \epsilon_2 \quad \text{for all } x \in X \cap [x_1, x_2].$$

Continue with $X \cap [x_2, x_3]$ etc.

$$\text{Let } \epsilon^\ell = \min \{\epsilon_1/2, \epsilon_2/2, \dots, \epsilon_k/2\} > 0.$$

Let $\epsilon = \min(\epsilon^\ell, \epsilon^k)$. Now apply property

A to $\{x_1, \dots, x_k\}$, $k < m(A)$ with $\epsilon > 0$ and $\sigma = -1$.

Therefore there exists an $F^*(A_s, x) \in \mathcal{A}_1^*$ such that

$$\|F^*(A, x) - F^*(A_s, x)\| < \epsilon$$

and

$$\text{sgn}(F^*(A, x) - F^*(A_s, x)) = -1 \text{ on } x_0 \leq x < x_1 - \epsilon$$

$$= (-1)^{j+1} \text{ on } x_j + \epsilon < x < x_{j+1} - \epsilon$$

$$= (-1)^{k+1} \text{ on } x_k + \epsilon < x \leq x_{k+1}$$

By construction, property A and strict monotonicity give us that $F^*(A_s, x)$ is a better approximation on the

intervals $X \cap [x_i + \epsilon, x_{i+1} - \epsilon]$ $i = 0, 1, \dots, k$, and

(4-4) and (4-5) along with our selected $\epsilon > 0$ give us

that $|W[x, f(x) - F^*(A_s, x)]| < \rho^*$ and

$$l(x) \leq F^*(A_s, x) \leq u(x) \quad \text{on } X \cap [x_i - \epsilon, x_i + \epsilon]$$

for $i = 1, 2, \dots, k$. Therefore

$$\max_{x \in X} |W[x, f(x) - F^*(A_s, x)]|$$

$$< \max_{x \in X} |W[x, f(x) - F^*(A, x)]| = \rho^*$$

which is a contradiction. Therefore the best generalized approximation $F^*(A, x)$ must alternate at least $m(A)$ times and the theorem is proven.

Uniqueness follows directly from the first part of the proof of Theorem 4-4 where the first ' $<$ ' is replaced by ' $\leq$ ' .

Theorem 4-5: The best approximation of Theorem 4-4 is unique.

Remark: It should be noted that a generalized weight function can be thought of as a transformation.

E. H. Kaufman, Jr. and G. G. Belford in [10] and [11] have examined unisolvent and varisolvent families from this approach.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. R. Barrar and H. Loeb, On N-Parameter and Unisolvent Families, J. Approx. Theory 1 (1968), 180-181.
2. ———, Best Non-linear Approximation with Interpolation, Arch. Rat. Mech. Anal. 33 (1969), 231-237.
3. ———, On the Continuity of the Non-linear Tschebyscheff Operator, Pacific Journal (to appear).
4. D. Braess, On Varisolvency and Alternation.
5. E. W. Cheney, Introduction to Approximation Theory, McGraw-Hill, New York, 1966.
6. C. Dunham, Necessity of Alteration, Canadian Math. Bull. 10 (1968), 743-744.
7. ———, Existence and Continuity of the Chebyshev Operator, Siam Review 10 (1968), 444-446.
8. ———, Chebyshev Approximation with Respect to a Weight Function, J. Approx. Theory 2 (1969), 223-232.
9. S. Karlin and W. J. Studden, Tchebyscheff Systems: With Applications in Analysis and Statistics, Interscience, New York, 1966.
10. E. H. Kaufman, Jr. and G. G. Belford, Transformations of Families of Approximating Functions, J. Approx. Theory (to appear).
11. ———, A Generalization of the Varisolvency and Unisolvency Properties.
12. H. L. Loeb, D. G. Moursund, L. L. Schumaker, and G. D. Taylor, Uniform Generalized Weight Function Polynomial Approximation with Interpolation, Siam J. Numer. Anal. 6 (1969), 284-293.

13. G. Meinardus, Approximation von Funktionen und ihre numerische Behandlung, Springer-Verlag, Berlin, 1964.
14. D. G. Moursund, Chebyshev Approximation Using a Generalized Weight Function, Siam J. Numer. Anal. 3 (1966), 435-450.
15. D. G. Moursund and G. D. Taylor, Uniform Rational Approximation Using a Generalized Weight Function, Siam J. Numer. Anal. 5 (1968), 882-889.
16. T. S. Motzkin, Approximation by Curves of a Unisolvent Family, Bull. Amer. Math. Soc. 55 (1949), 789-793.
17. I. Ninomiya, Generalized Rational Chebyshev Approximation.
18. E. Novodvorskii and I. Pinsker, On a Process of Equalization of Maxima, Usp. Mat. Nauk 6 (1951), 174-181.
19. J. R. Rice, The Approximation of Functions, Vol. 1: Linear Theory, Addison-Wesley, Reading, Mass.
20. ———, The Approximation of Functions, Vol. 2: Advanced Topics, Addison-Wesley, Reading, Mass., 1969.
21. L. L. Schumaker and G. D. Taylor, On Approximation by Polynomials Having Restricted Ranges II, Siam J. Numer. Anal. 6 (1969), 31-36.
22. G. D. Taylor, On Approximation by Polynomials Having Restricted Ranges, Siam J. Numer. Anal. 5 (1968), 258-268.
23. ———, Approximation by Functions Having Restricted Ranges III, J. Math. Anal. Applic. 27 (1969), 241-248.
24. ———, Approximation by Functions Having Restricted Ranges: Equality Case, Num. Math. 14 (1969), 71-78.
25. K. Taylor, Contributions to the Theory of Restricted Polynomial and Rational Approximation, Doctoral thesis, Michigan State University, 1970.

26. L. Tornheim, On N-Parameter Families of Functions and Associated Convex Functions, Trans. Amer. Math. Soc. 69 (1950), 457-467.
27. L. Wuytack, The Existence of a Solution in Constrained Rational Approximation Problems, SIMON STEVIN 43 (1969-1970), 83-99.
28. ———, Some Remarks on a Paper of D. G. Moursund, Siam J. Numer. Anal. 7 (1970), 233-237.
29. ———, Kolmogoroff's Criterion for Constrained Rational Approximation.

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03143 1467