Linear and nonlinear product approximation
by Darrell Patrick Schmidt

A thesis submitted in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in Mathematics
Montana State University

Abstract:
Let(Formula not captured by OCR) and (Formula not captured by OCR) be n- and m-dimensional Haar sets defined on the intervals I = [a,b] and J = [c,d], respectively. The (linear) product approximation of (Formula not captured by OCR)is defined to be (Formula not captured by OCR) (Formula not captured by OCR) where (Formula not captured by OCR) is the best approximation of F y(x) = F(x,y) from Φ = span(Formula not captured by OCR) with respect to the uniform norm ||.||I and (Formula not captured by OCR) is the best uniform approximation of J fi(y) over J from Ψ = span (Formula not captured by OCR) The rational product approximation of F &isin: C(D) is defined in a similar fashion. It is shown that both methods of product approximation possess desirable properties analogous to the classical theory for univariate Tchebycheff approximation. This study contributes to the recent developments in multivariate approximation theory initiated by S. E. Weinstein and M. S. Henry.

As the above considerations involve univariate Tchebycheff approximation some results in this area are also established. In particular, uniform extensions of the classical Freud's theorem are proven for both the generalized polynomial and the rational settings.

LINEAR AND NONLINEAR

PRODUCT APPROXIMATION

by

DARRELL PATRICK SCHMIDT

A thesis submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Mathematics

Approved:

Head, Major Department

Chairman, Examining Committee

Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

June, 1976

ACKNOWLEDGMENT

I especially want to thank my advisor, Dr. Myron S. Henry, for his able guidance, patience, and encouragement throughout my graduate school experience. I have benefited immensely from my association with him. Thanks are also due to Dr. Jackson Henry for his many suggestions in the programming aspects of this study. This acknowledgment would not be complete without expressing appreciation to Mrs. Diane Stovall who typed this manuscript so efficiently.

TABLE OF CONTENTS

CHAPTER PAGE

INTRODUCTION 1

I. TCHEBYCHEFF APPROXIMATION 6

1.1 Introduction 6
1.2 Tchebycheff approximation 7
1.3 The constants of the strong unicity theorem and Freud's theorem 11
1.4 Point-Lipschitz condition on subsets of I 28

II. LINEAR PRODUCT APPROXIMATION 35

2.1 Introduction to product approximations . 35
2.2 Degree of approximation 41
2.3 Continuity of the operator $\mathcal{P}$ 49
2.4 Point-Lipschitz continuity of $\mathcal{P}$ 54

III. RATIONAL APPROXIMATION 70

3.1 Introduction 70
3.2 Relationships involving $\mathbb{R}(n,m)$, $\mathbf{C}$, $\mathbb{R}^*$ and $\mathbf{C}^*$ 74
3.3 Uniform strong unicity theorem and uniform Freud's theorem for $R(C(f),\cdot)$. 85

IV. RATIONAL PRODUCT APPROXIMATION 96

4.1 Introduction 96
4.2 Definedness of $\mathcal{R}F$ 100
4.3 Continuity of $\mathcal{R}$ 106

V. DISCRETE RATIONAL PRODUCT APPROXIMATION AND COMPUTATION 121

5.1 Definitions and notations 121
5.2 Existence of $\mathcal{R}_{X\times Y}(x,y)$ and comparison to $\mathcal{R}(x,y)$ 124

CHAPTER PAGE

V. (Continued)
5.3 Discrete strong unicity theorem 141
5.4 Algorithms 148
5.5 Proof of Theorem 5.4.1 155
5.6 Computation 169

VI. CONCLUSIONS 174

BIBLIOGRAPHY 176

LIST OF TABLES

Table		Page
1	. . . .	171
2	. . . .	172

LIST OF DIAGRAMS

Diagrams		Page
1	. . . .	61
2	. . . .	62
3	. . . .	67

ABSTRACT

Let $\{\varphi_1,\ldots,\varphi_n\}$ and $\{\psi_1,\ldots,\psi_m\}$ be n- and m-dimensional Haar sets defined on the intervals $I = [a,b]$ and $J = [c,d]$, respectively. The (linear) product approximation of $F \in C(D)$, $D = I \times J$, is defined to be $(\mathcal{P}F)(x,y) = \sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij}\psi_j(y)\varphi_i(x)$ where $\sum_{i=1}^{n} f_i(y)\varphi_i(x)$ is the best approximation of $F_y(x) = F(x,y)$ from $\Phi = \text{span}\,\{\varphi_1,\ldots,\varphi_n\}$ with respect to the uniform norm $\|\cdot\|_I$ and $\sum_{j=1}^{m} f_{ij}\psi_j(y)$ is the best uniform approximation of $f_i(y)$ over J from $\Psi = \text{span}\,\{\psi_1,\ldots,\psi_m\}$, $i = 1,\ldots,n$. The rational product approximation of $F \in C(D)$ is defined in a similar fashion. It is shown that both methods of product approximation possess desirable properties analogous to the classical theory for univariate Tchebycheff approximation. This study contributes to the recent developments in multivariate approximation theory initiated by S. E. Weinstein and M. S. Henry.

As the above considerations involve univariate Tchebycheff approximation, some results in this area are also established. In particular, uniform extensions of the classical Freud's theorem are proven for both the generalized polynomial and the rational settings.

INTRODUCTION

The concern of this report is the uniform approximation of continuous real-valued functions defined on a rectangle. The approximating functions are either polynomials or rational functions in two variables. The method of approximation considered is that of product approximation formally defined in the polynomial setting by Weinstein [30] in 1969 and extended to the rational setting by Brown and Henry [3] in 1973. Product approximation is an alternative to best uniform approximation, or Tchebycheff approximation, of functions defined on a rectangle. Reasons for studying an alternative method are the loss of uniqueness of best approximations in the multivariate setting and a resulting lack of a theory corresponding to the classical theory of univariate Tchebycheff approximation. The focus of this investigation is to establish a theory for product approximation that parallels the classical univariate Tchebycheff theory.

Given a rectangle $D = I \times J = [a,b] \times [c,d]$ and Tchebycheff systems $\{\varphi_1,\ldots,\varphi_n\}$ and $\{\psi_1,\ldots,\psi_m\}$ on I and J, respectively, the (polynomial) product approximation of $F \in C(D)$ is the function

$$\mathcal{P}F = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij}\psi_j\varphi_i$$

where $\sum_{i=1}^{n} a_i(y)\varphi_i$ is the best uniform approximation of $F_y(x) = F(x,y)$ over I from $\Phi = \text{span}\,\{\varphi_1,\ldots,\varphi_n\}$ and $\sum_{j=1}^{m} a_{ij}\psi_j$ is the best uniform approximation $a_i(y)$ over J from $\Psi = \text{span}\,\{\psi_1,\ldots,\psi_m\}$. Definedness of $\mathcal{P}F$ depends on the continuity of the coefficient functions $a_i(y)$, $i = 1,\ldots,n$. In the setting described above, these coefficient functions are continuous. The rational product approximation of F over D is analogously defined; however, normality conditions must be placed on F to insure the continuity of the corresponding coefficient functions. A feature of product approximation is that $\mathcal{P}F$ is uniquely defined.

The first consideration is the degree of approximation of F by $\mathcal{P}F$ when $\Phi = \Phi_n$ and $\Psi = \Psi_m$ consist of the algebraic polynomials of degree less than or equal to n and m, respectively. A density theorem due to Weinstein [30] asserts that the uniform error of approximation

$$\|F - \mathcal{P}F\|_D = \sup \{\, |F(x,y) - (\mathcal{P}F)(x,y)| : (x,y) \in D\}$$

can be made arbitrarily small by appropriately choosing n and m sufficiently large. This result, however, does not provide a means of estimating the error of approximation. This is accomplished by establishing bounds on $\|F - \mathcal{P}F\|_D$ similar to those of Jackson's theorem [5] for Tchebycheff approximation.

A consideration that is essential for computation is the continuity of the product approximation operators. Due to intrinsic errors made by computing machinery, computation of $\mathcal{P}F$ yields the product approximation of a function that is uniformly close to F. Continuity of the operator $\mathcal{P}$ will force the product approximation obtained to be uniformly near $\mathcal{P}F$. Continuity theorems for both the polynomial and rational product approximation operators are proven, and conditions which guarantee the stronger point - Lipschitz continuity are determined.

Much work has recently been done in the area of computation of rational product approximations [3,4,12,13]. All the algorithms proposed in these papers involve a discretization of one or both of the intervals I and J. A comparison between the discrete and the non-discrete rational product approximations is thus needed. If X and

Y are finite subsets of I and J, respectively, the rational product approximation of F over $X \times Y$ is defined in a fashion similar to that over $D = I \times J$. Questions of definedness arise and are discussed. It is shown that under appropriate normality conditions if X and Y are sufficiently dense in I and J, then the discrete rational product approximation of F over $X \times Y$ is defined and converges uniformly to the rational product approximation of F over D as X and Y fill out their respective intervals.

A shortcoming of the algorithms proposed in the papers [3, 4, 12, 13] is computational inefficiency. This is due to the fact that the rational Remes algorithm or the differential correction algorithm is used several times in the computation of a rational product approximation. The convergence rate of an algorithm of J. Henry [13] is investigated, and conditions under which the algorithm possesses quadratic convergence properties are given. A modification of this algorithm is made to take advantage of these properties.

Since product approximations are derived from univariate Tchebycheff approximations, this report includes some results for Tchebycheff approximations that are used in establishing a theory for product approximation. In particular, several variations of the strong unicity theorem and Freud's theorem [10] are established.

CHAPTER I

TCHEBYCHEFF APPROXIMATION

1.1 Introduction.

Given a normed linear space V, a nonempty subset M of V, and a point x of V, the best approximation problem is that of finding an element m^* of M such that $\|x - m^*\| = \inf\{\|x - m\| : m \in M\}$. In this case m^* is called a best approximation of x from M. The following existence theorem appears in Cheney [5].

Theorem 1.1.1. A finite dimensional linear subspace of a normed linear space contains at least one best approximation to a given point.

In this chapter, we discuss some aspects of the classical theory of Tchebycheff approximation and present a survey of recent work on the continuity of the Tchebycheff approximation operator.

1.2 Tchebycheff approximation.

Let $C(I)$ be the normed linear space of continuous real-valued functions on the non-degenerate interval $I = [a,b]$ with the uniform, or Tchebycheff, norm $\|\cdot\|_I$ given by $\|f\|_I = \sup\{|f(x)| : x \in I\}$. Let $\Phi = \text{span}\{\varphi_1,\ldots,\varphi_n\}$ be an n-dimensional linear subspace of $C(I)$. The problem of finding best approximations of elements of $C(I)$ from Φ is called the Tchebycheff problem. By Theorem 1.1.1, each $f \in C(I)$ possesses at least one best approximation from Φ. Uniqueness of best approximations is not guaranteed without additional conditions on Φ.

Definition 1.2.1. A set $\{\varphi_1,\ldots,\varphi_n\}$ of n continuous functions on I is said to satisfy the Haar condition if no nontrivial linear combination of $\varphi_1,\ldots,\varphi_n$ vanishes at more than $n - 1$ points in I. In this case, the set $\{\varphi_1,\ldots,\varphi_n\}$ is called a Tchebycheff system on I, and the linear span Φ of $\{\varphi_1,\ldots,\varphi_n\}$ is called a Haar subspace of $C(I)$.

The Haar condition is equivalent to the following condition: For each set $\{x_1,\ldots,x_n\}$ of distinct points in I, the determinant

$$(1.1)\quad D[x_1,\ldots,x_n] = \begin{vmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \varphi_1(x_n) & \varphi_2(x_n) & \cdots & \varphi_n(x_n) \end{vmatrix}$$

is non-zero. This characterization of the Haar condition allows us to interpolate with elements of Φ; that is, given n distinct points $x_1,\ldots,x_n$ in I and n real numbers $y_1,\ldots,y_n$, there is a unique $\varphi \in \Phi$ such that $\varphi(x_i) = y_i$, $i = 1,\ldots,n$.

The Tchebycheff problem where Φ is a Haar subspace of C(I) admits the following characterization theorem; see Cheney [5].

Theorem 1.2.2. (Alternation theorem). Let Φ be an n-dimensional Haar subspace of C(I), $f \in C(I)$, and $\varphi \in \Phi$. Then φ is a best approximation of f from Φ if and only if there are n + 1 points $x_0 < x_1 < \cdots < x_n$ in I such that (i.) $|(f-\varphi)(x_i)| = \|f-\varphi\|_I$, $i = 0,\ldots,n$, and

(ii.) $(f-\varphi)(x_i) = -(f-\varphi)(x_{i-1})$, $i = 1,\ldots,n$.

The set $\{x_0,\ldots,x_n\}$ is called an "alternation set" or an "extremal set" for $f - \varphi$.

Unicity of best approximations is a consequence of the alternation theorem; see Cheney [5].

Theorem 1.2.3. If Φ is a Haar subspace of $C(I)$, then each $f \in C(I)$ possesses a unique best approximation from Φ.

Haar [11] has shown that Haar subspaces are the only finite dimensional subspaces of $C(I)$ that admit unique best approximations to all $f \in C(I)$.

Theorem 1.2.4. (Haar unicity theorem). Let Φ be an n-dimensional subspace of $C(I)$. If Φ is not a Haar subspace of $C(I)$, then there is an $f \in C(I)$ that has infinitely many best approximations from Φ.

Of particular importance is the fact that the Haar unicity theorem extends to the setting of $C(X)$ where X is a compact Hausdorff space; that is, every element of $C(X)$ has a unique best approximation from a finite dimensional subspace Φ of $C(X)$ if and only if Φ is a Haar subspace of $C(X)$. This result can be found in Phelps [25].

We conclude this section by stating two additional theorems which play a significant role in the remainder of this work. When Φ is a Haar subspace of $C(I)$, we let Tf denote the unique best approximation of $f \in C(I)$ from Φ. The strong unicity theorem is due to Newman and Shapiro [23].

<u>Theorem 1.2.5</u>. (Strong unicity theorem). Let Φ be an n-dimensional Haar subspace of $C(I)$. For each $f \in C(I)$, there is a constant $\gamma = \gamma_f > 0$ such that

$$(1.2) \qquad \|f - \varphi\|_I \geq \|f - Tf\|_I + \gamma \|\varphi - Tf\|_I$$

for all $\varphi \in \Phi$.

A related theorem establishing the continuity of the operator T was discovered independently by Freud [10] and Maehly and Witzgall [20].

Theorem 1.2.6. (Freud's theorem). Let Φ be an n-dimensional Haar subspace of $C(I)$. Given $f \in C(I)$, there is a constant $\lambda = \lambda_f > 0$ such that

$$\|Tg - Tf\|_I \leq \gamma \|g - f\|_I \tag{1.3}$$

for all $g \in C(I)$.

1.3 The constants of the strong unicity theorem and Freud's theorem.

The constants γ and λ in the strong unicity theorem and Freud's theorem depend on the function f and the approximating space Φ. In addition, if the approximation problem is considered over a closed subset X of I, then γ and λ may depend on the set X. We may write $\gamma = \gamma(f,\Phi,X)$ and $\lambda = \lambda(f,\Phi,X)$. The behavior of these constants has been the subject of several recent papers. Bartelt [2] and Cline [7] show that $\lambda = \lambda(f,\Phi,X)$ may be chosen independent of $f \in C(X)$ if X is finite. Moreover, Cline shows that if X is infinite and the dimension of Φ is greater than one, then $\lambda = \lambda(f,\Phi,X)$ cannot be chosen independent of f over $C(X)$; that is, given $\epsilon > 0$ there exist $f,g \in C(X)$

such that $\|g - f\|_X < \epsilon$ and $\|T_X g - T_X f\|_X = 1$, where $T_X g$ and $T_X f$ denote the best approximations of g and f from Φ over X, respectively. Bartelt also proves that if $\gamma = \gamma(f,\Phi,I)$ is defined as in Cheney [5],

$$(1.4) \qquad \gamma = \inf_{\varphi \in S} \max_{x \in E(f)} (f-Tf)(x) \|f-Tf\|_I^{-1} \varphi(x)$$

if $f \notin \Phi$, $S = \{\varphi \in \Phi : \|\varphi\|_I = 1\}$, and

$E(f) = \{x \in I : |(f - Tf)(x)| = \|f - Tf\|_I\}$, and

$$\gamma = 1$$

if $f \in \Phi$, then γ may be discontinuous off I but is an upper semicontinuous function of $f \in C(I)$.

The papers of Poreda [26] and Henry and Roulier [14] consider the problem where the approximating space is the set Φ_m of polynomials of degree less than or equal to m. Poreda considers the dependence of $\gamma = \gamma(f,\Phi_m,I)$ on the degree m and demonstrates a function $f \in C(I)$ such that the constant γ becomes unbounded as $m \to \infty$. Henry and Roulier consider the dependence of $\lambda = \lambda(f,\Phi_m, [-\theta,\theta])$ on $\theta \leq 1$, where $f \in C[-1,1]$ is fixed. They demonstrate that λ may be unbounded as $\theta \to 0$ and give conditions that guarantee that λ remains bounded with respect to θ for $0 < \theta \leq \delta$ for some $\delta > 0$.

We propose to determine conditions on a subset Γ of C(I) for which λ can be chosen uniformly over Γ; that is, there is a constant $\lambda_\Gamma > 0$ such that

$$(1.5) \qquad \|Tg - Tf\|_I \leq \lambda_\Gamma \|g - f\|_I$$

for all $f \in \Gamma$ and all $g \in C(I)$. In this section, we give such conditions and present an example which shows that these conditions are essential.

In the remainder of this section, we assume that Φ is a fixed n-dimensional Haar subspace of C(I), where $I = [a,b]$, and T denotes the corresponding Tchebycheff approximation operator. We first note that the constants of the strong unicity theorem and Freud's theorem are reciprocally related. The following lemma appears in Cheney [5].

Lemma 1.3.1. Let $f \in C(I)$. If (1.2) holds for all $\varphi \in \Phi$ with $\gamma = \gamma_f$, then (1.3) holds for all $g \in C(I)$ with $\lambda = 2/\gamma_f$.

Proof. Let $g \in C(I)$. Since $Tg \in \Phi$,

$$\begin{aligned}\gamma_f \|Tg - Tf\| &\leq \|f - Tg\|_I - \|f - Tf\|_I \\ &\leq \|f - g\|_I + \|g - Tg\|_I - \|f - Tf\|_I \\ &\leq \|f - g\|_I + \|g - Tf\|_I - \|f - Tf\|_I \\ &\leq 2\|g - f\|_I.\end{aligned}$$

Thus $\|Tg - Tf\|_I \leq (2/\gamma_f)\|g - f\|_I$.

Before developing sufficient conditions on $\Gamma \subseteq C(I)$ so that (1.5) may hold for all $f \in \Gamma$ and all $g \in C(I)$, we present a characterization of the constants γ and λ different from that of (1.4). For $a \leq x_0 < x_1 < \dots < x_n \leq b$, define

(1.6) $K(x_0,\dots,x_n)$

$$= \sup\{\|\varphi\|_I : \varphi \in \Phi \text{ and } (-1)^i\varphi(x_i) \geq -1,\ i = 0,\dots,n\}.$$

The fact that $K(x_0,\dots,x_n)$ is finite follows from the next two lemmas.

Lemma 1.3.2. Let $a \leq x_1 < \dots < x_n \leq b$ and real numbers $y_1,\dots,y_n$ be given. Suppose that $\{(x_1^k,\dots,x_n^k)\}$ and $\{(y_1^k,\dots,y_n^k)\}$ are two sequences of real numbers such that for each k, $a \leq x_1^k < \ldots < x_n^k \leq b$, and $x_i^k \to x_i$ and $y_i^k \to y_i$ as $k \to \infty$, $i = 1,\ldots,n$. Let φ^0 be the unique element of Φ such that $\varphi^0(x_i) = y_i$, $i = 1,\ldots,n$, and for each k let φ^k be the unique element of Φ such that $\varphi^k(x_i^k) = y_i^k$, $i = 1,\ldots,n$. Then $\varphi^k \to \varphi^0$ uniformly on I as $k \to \infty$.

Proof. By the Haar condition,

$$\|\varphi\|_* = \max_{i=1,\ldots,n} |\varphi(x_i)|$$

defines a norm on Φ. From Cheney [5, p. 78],

$$\varphi^k(x) = \sum_{i=1}^{n} y_i^k D[x_1^k,\ldots,x_{i-1}^k,x,x_{i+1}^k,\ldots,x_n^k]/D[x_1^k,\ldots,x_n^k]$$

where $D[x_1^k,\ldots,x_{i-1}^k,x,x_{i+1}^k,\ldots,x_n^k]$ and $D[x_1^k,\ldots,x_n^k]$ are given by (1.1) and $\{\varphi_1,\ldots,\varphi_n\}$ is a basis for Φ. Since $\varphi_1,\ldots,\varphi_n$ are continuous, the determinant operation is continuous with respect to its entries, $x_i^k \to x_i$ and $y_i^k \to y_i$ as $k \to \infty$, $i = 1,\ldots,n$, and $D[x_1,\ldots,x_n] \neq 0$,

$$\varphi^k(x_j) \to \sum_{i=1}^{n} y_i D[x_1,\ldots,x_{i-1},x_j,x_{i+1},\ldots,x_n]/D[x_1,\ldots,x_n]$$

$$= y_j$$

$$= \varphi^0(x_j),$$

$j = 1,\ldots,n$. Hence, $\|\varphi^k - \varphi^0\|_* \to 0$. Since all norms on the finite dimensional linear space Φ are norm-equivalent, $\|\varphi^k - \varphi^0\|_I \to 0$.

The next lemma is a generalization of an assertion of Rice [28, p. 64].

Lemma 1.3.3. Let $a \leq \alpha_0 \leq \beta_0 < \alpha_1 \leq \beta_1 < \cdots < \alpha_n \leq \beta_n \leq b$ be given. Then the set $\{\varphi \in \Phi : \forall i = 0,\ldots,n, \exists\, x_i \in [\alpha_i,\beta_i] \ni (-1)^i\varphi(x_i) \geq -1\}$ is $\|\cdot\|_I$ - bounded.

Proof. Assume otherwise. Then there are sequences $\{\varphi^k\}$ in Φ and $\{x_i^k\}$ in $[\alpha_i,\beta_i]$, $i = 0,\ldots,n$, such that $(-1)^i\varphi^k(x_i^k) \geq -1$ for all $i = 0,\ldots,n$, and $k = 1,2,\ldots$, and $\|\varphi^k\|_I \to \infty$. We may assume that each $\varphi^k \not\equiv 0$. Since $[\alpha_i,\beta_i]$ is compact, we may extract subsequences and relabel so that $x_i^k \to x_i \in [\alpha_i,\beta_i]$, $i = 0,\ldots,n$.

For each k, there is an index i for which $0 \geq (-1)^i\varphi^k(x_i^k) \geq -1$. Otherwise, φ^k would have n sign changes, and by the Haar condition we would have $\varphi^k \equiv 0$. We may now relabel the sequences and assume there is a fixed index $\nu \in \{0,\ldots,n\}$ such that $0 \geq (-1)^\nu\varphi^k(x_\nu^k) \geq -1$. In particular $|\varphi^k(x_\nu^k)| \leq 1$.

We further pass to subsequences and relabel so that for each index $i \in \{0,\ldots,n\}$, $\varphi^k(x_i^k)$ either converges to some real number η_i or to $\pm\infty$. Let H_1 denote the set of indices i for which $\varphi^k(x_i^k) \to \eta_i$ and H_2 be the set of indices where $\varphi^k(x_i^k) \to \pm\infty$. Since $\nu \in H_1$, $H_1 \neq \emptyset$. For $i \in H_2$, the inequality $(-1)^i \varphi^k(x_i^k) \geq -1$ implies that $(-1)^i \varphi^k(x_i^k) \to +\infty$.

Next we show that H_1 can contain at most n - 1 indices. Since $x_i^k \to x_i$ and $\varphi^k(x_i^k) \to \eta_i$, $i \in H_1$, we would have that φ^k converges uniformly to the unique element of Φ that has values η_i at n points x_i in H_1. This contradicts the assumption $\|\varphi^k\|_I \to \infty$. Thus H_1 contains at most n - 1 indices.

For $j \in H_1$, choose $\psi_j \in \Phi$ such that

$$\psi_j(x_j) = (-1)^j - \eta_j, \text{ and}$$

$$\psi_j(x_i) = 0,\ i \in H_1,\ i \neq j.$$

Since H_1 contains at most n - 1 indices, ψ_j exists (though not uniquely).

For each k, let $\zeta^k = \varphi^k + \sum_{j \in H_1} \psi_j$. By the continuity of each ψ_j and the convergences $x_i^k \to x_i$, and $\varphi^k(x_i^k) \to \eta_i$, $i \in H_1$, and $(-1)^i\varphi^k(x_i^k) \to +\infty$, $i \in H_2$, there is a number M such that

$$|\psi_j(x_i^k) - \psi_j(x_i)| < \frac{1}{n+1}, \; i,j \in H_1,$$

$$|\varphi^k(x_i^k) - \eta_i| < \frac{1}{n+1}, \; i \in H_1, \text{ and}$$

$$(-1)^i\varphi^k(x_i^k) > \sum_{j \in H_1} \|\psi_j\|_I$$

when $n \geq M$.

Let $k \geq M$. If $i \in H_2$, then

$$\begin{aligned}(-1)^i\zeta^k(x_i^k) &= (-1)^i\varphi^k(x_i^k) + \sum_{j \in H_1}(-1)^i\psi_j(x_i^k)\\ &\geq (-1)^i\varphi^k(x_i^k) - \sum_{j \in H_1}\|\psi_j\|_I\\ &> 0\end{aligned}$$

If $i \in H_1$, then

$$\begin{aligned}(-1)^i\zeta^k(x_i^k) &= (-1)^i\varphi^k(x_i^k) + \sum_{j \in H_1}(-1)^i\psi_j(x_i^k)\\ &= (-1)^i\varphi^k(x_i^k) + (-1)^i\psi_i(x_i^k)\\ &\qquad + \sum_{\substack{j \in H_1\\ j \neq i}}(-1)^i\psi_j(x_i^k)\end{aligned}$$

$$= (-1)^i[\varphi^k(x_i^k) - \eta_i]$$
$$+ (-1)^i[\eta_i + \psi_i(x_i)]$$
$$+ (-1)^i[\psi_i(x_i^k - \psi_i(x_i)]$$
$$+ \sum_{\substack{j \in H_1 \\ j \neq 1}} (-1)^i[\psi_j(x_i^k) - \psi_j(x_i)]$$

$$\geq - \frac{1}{n+1} + 1 - \frac{1}{n+1} - \frac{n-2}{n+1}$$
$$= 1 - \frac{n}{n+1}$$
$$> 0.$$

Thus ζ^k has n sign changes, and as a result $\zeta^k \equiv 0$. This implies that $\varphi^k \equiv - \sum_{j \in H_1} \psi_j$ which contradicts the assumption $\|\varphi^k\|_I \to \infty$. This completes the proof of Lemma 1.3.3.

Remark. By applying Lemma 1.3.3. with each $\alpha_i = \beta_i = x_i$, we see that $K(x_0,\ldots,x_n)$ is finite. Moreover, since $(-1)^i\varphi(x_i) \geq -1$ for any $\varphi \in \Phi$ where $\|\varphi\|_I = 1$,

$$1 \leq K(x_0,\ldots,x_n) < \infty$$

whenever $a \leq x_0 < x_1 < \ldots < x_n \leq b$.

The next lemma indicates how the constants γ and λ can be derived from $K(x_0,\ldots,x_n)$.

Lemma 1.3.4. Let $f \in C(I)$, and let $x_0 < x_1 < \ldots < x_n$ be an alternation set for $f - Tf$. Then (1.2) holds for all $\varphi \in \Phi$ with $\gamma = 1/K(x_0,\ldots,x_n)$, and (1.3) holds for all $g \in C(I)$ with $\lambda = 2K(x_0,\ldots,x_n)$.

Proof. Since $x_0 < x_1 < \ldots < x_n$ is an alternation set for $f - Tf$,

(1.7) $$(-1)^{i+\nu}\ (f-Tf)(x_i) = \|f-Tf\|_I,$$

$i = 0,\ldots,n$, where $\nu = 0$ or 1 does not depend on i. Let $\varphi \in \Phi$. If $\varphi = Tf$, then (1.2) clearly holds. Assume $\varphi \neq Tf$. For $i = 0,\ldots,n$,

(1.8) $$-(-1)^{i+\nu}\ (f-\varphi)(x_i) \geq - \|f-\varphi\|_I.$$

Adding (1.7) and (1.8) we get

$$(-1)^{i+\nu}\ (\varphi-Tf)(x_i) \geq - (\|f-\varphi\|_I - \|f-Tf\|_I),$$

$i = 0,\ldots,n$. Since $\varphi \neq f$, $\|f-\varphi\|_I - \|f-Tf\|_I > 0$ and

$$(-1)^i \frac{(-1)^\nu \quad (\varphi - Tf)}{\|f-\varphi\|_I - \|f-Tf\|_I} \geq -1$$

$i = 0,\ldots,n$. From the definition of $K(x_0,\ldots,x_n)$, we obtain

$$\frac{\|\varphi - Tf\|_I}{\|f-\varphi\|_I - \|f-Tf\|_I} \leq K(x_0,\ldots,x_n).$$

Hence,

$$\|f-\varphi\|_I \geq \|f-Tf\|_I + (1/K(x_0,\ldots,x_n)) \|\varphi - Tf\|_I.$$

The second assertion now follows from Lemma 1.3.1.

Lemma 1.3.4 indicates that the constants γ and λ as defined in Lemma 1.3.4 do not depend explicitly on the function f but on an alternation set for f - Tf. In contrast to Bartelt's result [2] that γ may be discontinuous off Φ, we show that $K(x_0,\ldots,x_n)$ is continuous in its domain $\{(x_0,\ldots,x_n) \in I^{n+1} : x_0 < x_1 < \ldots < x_n\}$, and thus γ and λ being appropriately defined depend continuously on the alternation sets. It need be noted, however, that this does not imply that γ and λ may depend continuously on f because alternation sets need not be unique and may not depend continuously on f.

<u>Theorem 1.3.5</u>. The function $K(x_0,\ldots,x_n)$ defined in (1.6) is continuous on $\{(x_0,\ldots,x_n) \in I^{n+1} : x_0 < x_1 < \ldots < x_n\}$.

Proof. Let $a \leq x_0 < x_1 < \ldots < x_n \leq b$ be fixed. There are numbers $a \leq \alpha_0 < \beta_0 < \alpha_1 < \beta_1 < \ldots < \alpha_n < \beta_n \leq b$, such that $x_i \in (\alpha_i,\beta_i)$, $i = 1,\ldots,n-1$. If $x_0 = a$, then $x_0 \in [\alpha_0,\beta_0)$, and if $x_n = b$, then $x_n \in (\alpha_n,\beta_n]$. Otherwise, open intervals may be used for either or both of these indices.

Let $\mathcal{A} = \{\varphi \in \Phi : \forall i = 0,\ldots,n, \exists\, \xi_i \in [\alpha_i,\beta_i] \ni (-1)^i\varphi(\xi_i) \geq -1\}$. Lemma 1.3.3 asserts that $\mathcal{A}$ is uniformly bounded. The uniform boundedness of $\mathcal{A}$ and the finite-dimensionality of Φ imply that $\mathcal{A}$ is equicontinuous.

Let $\epsilon > 0$ be given. We may select $\delta > 0$ such that

(i) if $\varphi \in \mathcal{A}$ and $x,y \in I$ with $|x - y| < \delta$, then $|\varphi(x) - \varphi(y)| < \epsilon$, and

(ii) if $y \in I$ and $|x_i - y| < \delta$, then $y \in [\alpha_i,\beta_i]$, $i = 0,\ldots,n$.

Suppose that $(y_0,\ldots,y_n) \in I^{n+1}$ where $y_0 < y_1 < \ldots < y_n$ and $\max_{i=0,\ldots,n} |y_i - x_i| < \delta$. We may pick $\varphi_x \in \Phi$ such that $(-1)^i \varphi_x(x_i) \geq -1$, $i = 0,\ldots,n$, and $\|\varphi_x\|_I = K(x_0,\ldots,x_n)$. If i is even, then

$$(-1)^i \varphi_x(y_i) = \varphi_x(y_i) \geq \varphi_x(x_i) - \epsilon \geq -(1+\epsilon).$$

Thus $(-1)^i \varphi_x(y_i)/(1+\epsilon) \geq -1$. If i is odd, then

$$(-1)^i \varphi_x(x_i) = -\varphi_x(x_i) \geq -1.$$

So $\varphi_x(x_i) \leq 1$, and

$$\varphi_x(y_i) \leq \varphi_x(x_i) + \epsilon \leq 1 + \epsilon.$$

Thus $(-1)^i \varphi_x(y_i)/(1+\epsilon) \geq -1$. Hence,

$$\begin{aligned} K(y_0,\ldots,y_n) &\geq \|\varphi_x/(1+\epsilon)\|_I \\ &= K(x_0,\ldots,x_n)/(1+\epsilon). \end{aligned}$$

We may interchange the roles of x and y in the above argument and obtain

$$\frac{K(x_0,\ldots,x_n)}{1+\epsilon} \leq K(y_0,\ldots,y_n) \leq (1+\epsilon)K(x_0,\ldots,x_n),$$

establishing the continuity of $K(x_0,\ldots,x_n)$.

We now state and prove a uniform strong unicity theorem.

Theorem 1.3.6. Let Γ be a compact subset of $C(I)$ where $\Gamma \cap \Phi = \emptyset$. Then there is a constant $\gamma_\Gamma > 0$ such that

$$\|f-\varphi\|_I \geq \|f-Tf\|_I + \gamma_\Gamma\|\varphi-Tf\|_I$$

for all $f \in \Gamma$ and all $\varphi \in \Phi$.

Proof. Assume otherwise. Then there is a sequence $\{f_k\}$ in Γ and a sequence $\{\varphi^k\}$ in Φ such that each $\varphi^k \neq Tf_k$ and

$$(1.9) \qquad \frac{\|f_k-\varphi^k\|_I - \|f_k-Tf_k\|_I}{\|\varphi^k-Tf_k\|_I} \to 0$$

as $k \to \infty$.

Since Γ is compact we may assume after relabeling that $f_k \to f \in \Gamma$. Let $S = f - Tf$ and $S_k = f_k - Tf_k$ for all k. Since $\Gamma \cap \Phi = \emptyset$, $\|S\|_I > 0$.

For each k, let

$$a \leq x_0^k < x_1^k < \ldots < x_n^k \leq b$$

be an alternation set for S_k. We further extract subsequences and relabel so that $x_i^k \to x_i$, $i = 0,\ldots,n$. Then $x_0 \leq x_1 \leq \cdots \leq x_n$.

We show that $x_{i-1} < x_i$, $i = 1,\ldots,n$. Suppose $x_{i-1} = x_i$ for some i. We note that $f_k \to f$ implies $S_k \to S$ (see Theorem 1.2.6). Since $S_k \to S$, $x_{i-1}^k \to x_{i-1} = x_i$, $x_i^k \to x_i$, and S is continuous, k may be chosen large enough so that

$$|S(x_{i-1}^k) - S(x_i^k)| < \|S\|_I$$

and

$$\|S-S_k\|_I < \frac{1}{4}\|S\|_I.$$

Thus

$$\begin{aligned}\|S\|_I &> |S(x_{i-1}^k) - S(x_i^k)| \\ &\geq |S_k(x_{i-1}^k) - S_k(x_i^k)| - |S(x_{i-1}^k) - S_k(x_{i-1}^k)| \\ &\qquad - |S(x_i^k) - S_k(x_i^k)| \\ &\geq 2\|S_k\|_I - 2\|S - S_k\|_I.\end{aligned}$$

But

$\|S_k\|_I \geq \|S\|_I - \|S - S_k\|_I$. Then

$$\begin{aligned}\|S\|_I &> 2\|S\|_I - 4\|S - S_k\|_I \\ &> 2\|S\|_I - 4 \cdot \frac{1}{4}\|S\|_I \\ &= \|S\|_I,\end{aligned}$$

which is false. Thus $x_0 < x_1 < \dots < x_n$.

We may now apply Theorem 1.3.5 to see that

$$K(x_0^k, \dots, x_n^k) \to K(x_0, \dots, x_n)$$

as $k \to \infty$. Thus $K(x_0^k, \dots, x_n^k) \leq A$ for some $A > 0$. By Lemma 1.3.4,

$$\frac{\|f_k - \varphi^k\|_I - \|f_k - Tf_k\|_I}{\|\varphi^k - Tf_k\|_I} \geq \frac{1}{K(x_0^k, \dots, x_n^k)} \geq \frac{1}{A} > 0.$$

This contradicts (1.9) and the proof of Theorem 1.3.6 is complete.

A corresponding uniform Freud's theorem now follows immediately from Lemma 1.3.1 and Theorem 1.3.6.

<u>Theorem 1.3.7</u>. Let Γ be a compact subset of $C(I)$ where $\Gamma \cap \Phi = \emptyset$. Then there is a constant $\lambda_\Gamma > 0$ such that

$$\|Tg - Tf\|_I \leq \lambda_\Gamma \|g - f\|_I$$

for all $f \in \Gamma$ and all $g \in C(I)$.

We conclude this section with an example that shows that the condition $\Gamma \cap \Phi = \emptyset$ cannot be dropped from Theorem 1.3.7.

Example 1.3.1. In the case where $n \geq 2$, we demonstrate a compact subset Γ of $C(I)$ which meets Φ, a sequence $\{f_k^*\}$ in Γ, and a sequence $\{g_k^*\}$ in $C(I)$ such that each $g_k^* \neq f_k^*$ and

$$\frac{\| Tg_k^* - Tf_k^*\|_I}{\| g_k^* - f_k^*\|_I} \rightarrow \infty$$

as $k \rightarrow \infty$.

From Theorem 4 of Cline [7], for each $k = 1,2,\ldots,$ there exist $f_k, g_k \in C(I)$ such that $\| g_k - f_k\|_I < 1/k$ and $\| Tg_k - Tf_k\|_I = 1$. Let $\alpha_k = k(1 + \| f_k\|_I)$, $f_k^* = f_k/\alpha_k$, and $g_k^* = g_k/\alpha_k$. Since $\| f_k^*\|_I < 1/k$, $f_k^* \rightarrow 0$ uniformly on I. As a result, $\Gamma = \{0\} \cup \{f_1^*, f_2^*, \ldots\}$ is a compact subset of $C(I)$. We now apply the homogeneity of the operator T (see Cheney [5, p. 83]) to get

$$\frac{\| Tg_k^* - Tf_k^*\|_I}{\| g_k^* - f_k^*\|_I} = \frac{\| T(g_k/\alpha_k) - T(f_k/\alpha_k)\|_I}{\| (g_k/\alpha_k) - (f_k/\alpha_k)\|_I}$$

$$= \frac{\| Tg_k - Tf_k \|_I / \alpha_k}{\| g_k - f_k \|_I / \alpha_k}$$

$$\geq \quad k \to \infty$$

as $k \to \infty$. Thus λ cannot be chosen uniformly over Γ.

1.4 Point - Lipschitz condition on subsets of I.

In the terminology of the introductory paragraph of Section 1.3, we consider the dependence of $\lambda = \lambda(f,\Phi,X)$ on the closed subset X of I over which the best approximation problem is considered.

We assume that Φ is an n-dimensional Haar subspace of C(I). If X is a closed subset of I containing at least n + 1 points, we let $T_X f$ denote the best approximation of $f \in C(X)$ from Φ in the sense of the norm $\|\cdot\|_X$ defined by $\|f\|_X = \sup \{ |f(x)| : x \in X \}$. For a treatment of Tchebycheff approximation on closed subsets of I, see Cheney [5].

The principle result of this section is that if $f \in C(I)$ is fixed, then the operators $T_X f$ satisfy a point-Lipschitz condition at f with a constant independent of sufficiently dense closed subsets X of I. This result will be used later in Chapter V.

The <u>density</u> of a closed subset X of I is defined to be

$$(1.10) \qquad d(X) = \sup_{y \in I} \inf_{x \in X} |y - x|.$$

We first establish a point-Lipschitz condition for the operator T_X.

<u>Lemma 1.4.1</u>. Let X be a closed subset of I containing at least n + 1 points and let $f \in C(X)$. Let $x_0 < x_1 < \ldots < x_n$ constitute an alternation set for $f - T_X f$ in X. Then

$$(1.11) \qquad \| T_X g - T_X f \|_I \leq 2K(x_0, \ldots, x_n) \| g - f \|_X$$

for all $g \in C(X)$ where $K(x_0, \ldots, x_n)$ is defined in (1.6).

Proof. Since $x_0 < x_1 < \ldots < x_n$ is an alternation set for $f - T_X f$ in X,

$$(1.12) \qquad (-1)^{i+\nu} (f - Tf)(x_i) = \| f - T_X f \|_X ,$$

$i = 0, \ldots, n$, where $\nu = 0$ or 1 does not depend on i.

Let $g \in C(I)$. If $g = f$, then (1.11) is trivial. Assume $g \neq f$. For $i = 0,\ldots,n$,

$$(1.13) \quad (-1)^{i+\nu}(g-Tg)(x_i) \leq \|g - T_X g\|_X \leq \|g - T_X f\|_X \leq \|g - f\|_X + \|f - T_X f\|_X .$$

Subtracting (1.13) from (1.12), we get

$$(-1)^{i+\nu}(T_X g - T_X f)(x_i) - (-1)^{i+\nu}(g-f)(x_i) \geq -\|g-f\|_X.$$

Hence,

$$(-1)^{i+\nu}(T_X g - T_X f)(x_i) \geq (-1)^{i+\nu}(g-f)(x_i) - \|g - f\|_X \geq -2\|g - f\|_X.$$

So $(-1)^i(-1)^\nu(T_X g - T_X f)(x_i)/2\|g - f\|_X \geq -1$, $i = 0,\ldots,n$. Hence, $\|T_X g - T_X f\|_I \leq 2K(x_0,\ldots,x_n)\|g - f\|_X$.

Theorem 1.4.2. Let $f \in C(I)$. There exist numbers $\delta > 0$ and $\lambda > 0$ such that if X is a closed subset of I and $d(X) < \delta$, then

$$\|T_X g - T_X f\|_I \leq \lambda\|g - f\|_X$$

for all $g \in C(X)$.

Proof. We consider two cases.

In the first case suppose $f \in \Phi$. Then $T_X f = f$ for any closed subset X of I. If X is a closed subset of I containing at least $n + 1$ points and $g \in C(X)$, then

$$\begin{aligned} \|T_X g - T_X f\|_X &= \|T_X g - f\|_X \\ &\leq \|T_X g - g\|_X + \|g - f\|_X \\ &\leq \|f - g\|_X + \|g - f\|_X \\ &= 2\|g - f\|_X. \end{aligned}$$

Now pick numbers

$$a \leq \alpha_0 < \beta_0 < \alpha_1 < \beta_1 < \cdots < \alpha_n < \beta_n \leq b.$$

Let $M = \sup\{\|\varphi\|_I : \varphi \in \Phi$ and $\forall i = 0,\ldots,n, \exists \xi_i \in [\alpha_i,\beta_i] \ni (-1)^i\varphi(\xi_i) \geq -1\}$. By Lemma 1.3.3, M is finite. We now pick $\delta > 0$ small enough that if $X \subseteq I$ and $d(X) < \delta$, then $X \cap [\alpha_i,\beta_i] \neq \emptyset$, $i = 0,\ldots,n$.

Let X be a closed subset of I with $d(X) < \delta$. Pick $x_i \in X \cap [\alpha_i,\beta_i]$, $i = 0,\ldots,n$. For $g \in C(X)$ and $i = 0,\ldots,n$,

$$(-1)^i(T_X g - T_X f)(x_i) \geq -\|T_X g - T_X f\|_X.$$

By the definition of M,

$$\| T_X g - T_X f \|_I \le M \| T_X g - T_X f \|_X \le 2M \| g - f \|_X .$$

In the second case, we assume $f \notin \Phi$. Suppose no such combination of $\lambda > 0$ and $\delta > 0$ exists. Then there is a sequence $\{X_k\}$ of closed subsets of I and a sequence $\{g_k\}$ where each $g_k \in C(X_k)$ such that $g_k \not\equiv f$ on X_k, $d(X_k) \to 0$, and

$$(1.14) \qquad \frac{\| T_{X_k} g_k - T_{X_k} f \|_I}{\| g_k - f \|_{X_k}} \to \infty$$

as $k \to \infty$.

Let $S = f - Tf$ and $S_k = f - T_{X_k} f$. From Cheney [5, p. 87], since $d(X_k) \to 0$ we have that $S_k \to S$ uniformly on I. This convergence, the continuity of S, the assumption $d(X_k) \to 0$, and the inequalities,

$$\| S_k \|_{X_k} \le \| S_k \|_I \le \| S \|_I + \| S_k - S \|_I$$

and

$$\| S_k \|_{X_k} \ge \| S \|_{X_k} - \| S_k - S \|_{X_k}$$

$$\ge \| S \|_I - (\| S \|_I - \| S \|_{X_k}) - \| S_k - S \|_I ,$$

imply that $\|S_k\|_{X_k} \to \|S\|_I$.

Let $x_0^k < x_1^k < \ldots < x_n^k$ be an alternation set for $S_k = f - T_{X_k} f$ in X_k. We extract subsequences and relabel so that $x_i^k \to x_i \in I$, $i = 0,\ldots,n$. We now show that $x_{i-1} < x_i$, $i = 1,\ldots,n$. Assume that some $x_{i-1} = x_i$. Since $\|S\|_I > 0$ and S is continuous on I, we may choose k large enough so that,

$$|S(x_{i-1}^k) - S(x_i^k)| < \|S\|_I,$$

$$\|S_k - S\|_I < \frac{1}{4}\|S\|_I, \text{ and}$$

$$\|S_k\|_{X_k} \geq \frac{3}{4}\|S\|_I.$$

Then

$$\begin{aligned}\|S\|_I &> |S(x_{i-1}^k) - S(x_i^k)| \\ &\geq |S_k(x_{i-1}^k) - S_k(x_i^k)| - |S_k(x_{i-1}^k) - S(x_{i-1}^k)| \\ &\qquad - |S_k(x_i^k) - S(x_i^k)| \\ &\geq 2\|S_k\|_{X_k} - 2\|S_k - S\|_I \\ &> 2 \cdot \frac{3}{4}\|S\|_I - 2 \cdot \frac{1}{4}\|S\|_I \\ &= \|S\|_I,\end{aligned}$$

which is false. Thus $x_0 < x_1 < \dots < x_n$.

Lemma 1.3.4 insures that $K(x_0^k,\dots,x_n^k) \to K(x_0,\dots,x_n)$, and thus there is a number $A > 0$ such that $K(x_0^k,\dots,x_n^k) \leq A$ for all k. From Lemma 1.4.1,

$$\frac{\| T_{X_k} g_k - T_{X_k} f \|_I}{\| g_k - f \|_{X_k}} \leq 2K(x_0^k,\dots,x_n^k) \leq 2A.$$

This contradicts (1.14), and Theorem 1.4.2 is now established.

CHAPTER II

LINEAR PRODUCT APPROXIMATION

2.1 Introduction to product approximations.

The concern of Chapter II is the approximation of continuous functions defined on a rectangle. Extension of univariate Tchebycheff approximation theory to the multivariate setting has been confronted with severe difficulties. This is evidenced by the following theorem of Mairhuber [22].

Theorem 2.1.1. (Mairhuber). A compact subset X of m-dimensional Euclidean space containing more than n points, $n \geq 2$, serves as the domain of definition for a Tchebycheff system $\{\varphi_1, \ldots, \varphi_n\}$ on X if and only if X is homeomorphic to a closed subset of the circumference of a circle.

A ramification of Mairhuber's result in conjunction with the Haar unicity theorem is that unless X is essentially a circle or a closed subspace of an interval, uniqueness of best uniform approximations of all $f \in C(X)$ from a finite dimensional subspace of C(X) cannot be guaranteed. Uniqueness of best approximations is often needed to facilitate an algorithm used to find the approximation. Moreover, the loss of uniqueness poses difficulties in establishing results corresponding to the classical theory of one-dimensional approximation.

To circumvent the loss of unicity, Weinstein [30] devised the concept of product approximation. A description of product approximations and the product approximation operator follows.

Let D denote the rectangle $I \times J = [a,b] \times [c,d]$, and let $F \in C(D)$. For each $y \in J$, let $F_y \in C(I)$ be given by $F_y(x) = F(x,y)$. Let $\{\varphi_1,\ldots,\varphi_n\}$ be a Tchebycheff system on I with linear span Φ. For each $y \in J$, let

$$T(F_y,\cdot) = \sum_{i=1}^{n} f_i(y)\varphi_i$$

be the unique best uniform approximation of F_y on I from Φ. Weinstein [30] has proven that the coefficient functions f_i, $i = 1,\ldots,n$, are continuous on J. Now suppose that $\Psi = \text{span}\,\{\psi_1,\ldots,\psi_m\}$ is an m-dimensional Haar subspace of $C(J)$, and for $i = 1,\ldots,n$, let

$$Q(f_i,\cdot) = \sum_{j=1}^{m} f_{ij}\psi_j$$

be the unique best approximation of f_i on J from Ψ in the sense of the uniform norm $\|\cdot\|_J$. The product approximation of F on D is defined as

$$(2.1)\quad \mathcal{P}F = \sum_{i=1}^{n} Q(f_i,\cdot)\varphi_i = \sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij}\psi_j\varphi_i.$$

Hereafter, $\mathcal{P}$ defined in (2.1) will be called the product approximation operator.

The name "product approximation" is probably derived from the fact that if $F(x,y) = g(x)h(y)$, then $(\mathcal{P}F)(x,y) = T(g,x)Q(h,y)$. In addition, if $F(x,y) = g(x) + h(y)$ and if Φ and Ψ contain the constant functions, then $(\mathcal{P}F)(x,y) = T(g,x) + Q(h,y)$. Several recent papers [3,4,12,13,17] discuss extensions of product approximation to the settings of rational functions and varisolvent families of functions. The results of these papers will be discussed in Chapters IV and V. In this chapter, we consider the degree of approximation of $F \in C(D)$ by product approximations using algebraic polynomials and the continuity of the operator $\mathcal{P}$.

Two questions regarding possible invariance properties of $\mathcal{P}$ naturally arise. Is $\mathcal{P}$ independent of the order in which the variables x and y are prescribed? That is, if

$$Q(F^x,\cdot) = \sum_{j=1}^{m} f_j^*(x)\psi_j$$

where $F^x(y) = F(x,y)$ and

$$\mathcal{P}^*F = \sum_{j=1}^{m} T(f_j^*,\cdot)\psi_j = \sum_{j=1}^{m} \sum_{i=1}^{n} f_{ij}^*\varphi_i\psi_j,$$

is it necessarily true that $\mathcal{P}F \equiv \mathcal{P}^*F$? Also, is $\mathcal{P}F$ independent of the basis functions for the approximating space Φ? The first question is answered in the negative by the following example due to Weinstein [30].

Example 2.1. Let $D = I \times J = [-1,1] \times [-1,1]$, $\Phi = \text{span}\{\varphi_1\}$ where $\varphi_1 \equiv 1$, $\Psi = \text{span}\{\psi_1\}$ where $\psi_1 \equiv 1$, and

$$F(x,y) = \begin{cases} (1-2x^2)y, & -1 \le y \le 0, \\ 2xy, & 0 \le y \le 1. \end{cases}$$

For each $y \in J$, $T(F_y,\cdot) \equiv 0$, and we see that $\mathcal{P}F \equiv 0$. For $x \in I$, $Q(F^x,y) = f_1^*(x)$, where

$$f_1^*(x) = \begin{cases} x^2 + x - \frac{1}{2}, & -1 \le x \le -1/\sqrt{2}, \\ x, & -1/\sqrt{2} \le x \le (1 - \sqrt{3})/2, \\ x^2 - \frac{1}{2}, & (1 - \sqrt{3})/2 \le x \le 0, \\ x^2 + x - \frac{1}{2}, & 0 \le x \le 1/\sqrt{2}, \\ x, & 1/\sqrt{2} \le x \le 1. \end{cases}$$

Best approximation of f_1^* gives $\mathfrak{p}^*F \equiv (2 - \sqrt{2})/4$, and $\mathfrak{p}F \neq \mathfrak{p}^*F$.

The next example answers the second query in the negative.

Example 2.2. Let $D = I \times J = [-1,1] \times [-1,1]$, $\Phi = \text{span}\,\{1,x,x^2\}$, $\Psi = \text{span}\,\{1,y\}$, and

$$F(x,y) = \begin{cases} -4y(1+y), & -1 \le y \le 0, \\ 8x^2y(1-y), & 0 \le y \le 1. \end{cases}$$

Then $T(F_y,x) = f_0(y) + f_2(y)x^2$, where

$$f_0(y) = \begin{cases} -4y(1+y), & -1 \le y \le 0, \\ 0, & 0 \le y \le 1, \end{cases}$$

and

$$f_2(y) = \begin{cases} 0, & -1 \leq y \leq 0 \\ 8y(1-y), & 0 \leq y \leq 1. \end{cases}$$

Now $Q(f_0,y) = \frac{1}{2}$ and $Q(f_2,y) = 1$, and we have

$$(\mathcal{P}F)(x,y) = \frac{1}{2} + x^2.$$

We can write $\Phi = \text{span}\{p_0,p_1,p_2\}$ where $p_0(x) = 1$, $p_1(x) = x$, and $p_2(x) = 2x^2 - 1$ are the first three Tchebycheff polynomials. Then

$$\begin{aligned} T(F_y,x) &= f_0(y) + f_2(y)x^2 \\ &= [f_0(y) + f_2(y)/2]p_0(x) + [f_2(y)/2]p_2(x) \\ &= g_0(y)p_0(x) + g_2(y)p_2(x) \end{aligned}$$

where

$$g_0(y) = \begin{cases} -4y(1+y), & -1 \leq y \leq 0, \\ 4y(1-y), & 0 \leq y \leq 1, \end{cases}$$

and

$$g_2(y) = \begin{cases} 0, & -1 \leq y \leq 0, \\ 4y(1-y), & 0 \leq y \leq 1. \end{cases}$$

Now $Q(g_0,y) = Q(g_2,y) = \frac{1}{2}$. If we let $\mathcal{P}^*F$ denote the product approximation of F with respect to the Tchebycheff polynomials, then

$$(\mathcal{P}^*F)(x,y) = (p_0(x) + p_2(x))/2 = x^2,$$

and thus $\mathcal{P}F \neq \mathcal{P}^*F$.

2.2 Degree of approximation.

In this section, we establish a theorem for product approximation similar to that of Jackson's theorem for Tchebycheff approximation. We fix $D = I \times J = [-1,1] \times [-1,1]$, and let the approximating spaces Φ_n and Ψ_m be the (n+1)- and (m+1)-dimensional spaces of polynomials of degree less than or equal to n and m, respectively. As example 2.2 indicates, we must specify the basis functions for Φ_n. We choose $\{p_0,\ldots,p_n\}$ as the basis for Φ_n where p_i denotes the i-th degree Tchebycheff polynomial. Moreover, let $\{q_0,\ldots,q_m\}$ be any basis for Ψ_m. For $F \in C(D)$ and $y \in J$,

$$T_n(f_y,\cdot) = \sum_{i=0}^{n} f_i(y)p_i \tag{2.2}$$

will denote the best approximation of F_y on I from Φ_n, and

$$Q_m(f_i,\cdot) = \sum_{j=0}^{m} f_{ij}\, q_j$$

will denote the best approximation of f_i on J from Ψ_m. The corresponding product approximation is denoted by $\mathcal{P}_{n,m}F$.

The following density theorem is a special case of a theorem of Weinstein [30].

Theorem 2.2.1. Given $F \in C(D)$ and $\epsilon > 0$, there is an $N(\epsilon)$ such that for each $n > N(\epsilon)$ there is an $M(\epsilon,n)$ such that if $n > N(\epsilon)$ and $m > M(\epsilon,n)$, then $\|F - \mathcal{P}_{n,m}F\|_D < \epsilon$.

The principle result of this section gives a bound on $\|F - \mathcal{P}_{n,m}F\|_D$ which indicates the dependence of $M(\epsilon,n)$ on n. We will make use of the orthogonality properties of the Tchebycheff polynomials:

$$(2.3)\quad \int_{-1}^{1} p_i(x)p_j(x)(1-x^2)^{-1/2}dx = \begin{cases} \pi, & i = j = 0, \\ \pi/2, & i = j > 0, \\ 0, & i \neq j, \end{cases}$$

and we note that $\|p_i\|_I = \sup\{|p_i(x)| : x \in I\} = 1$. The error estimates will be given in terms of the modulus of continuity of a function. For $f \in C(I)$, the modulus of continuity [5] of f is

$$\omega(f,I,\delta) = \sup\{|f(x_1) - f(x_2)| : x_1,x_2 \in I, |x_1 - x_2| \leq \delta\}.$$

For $F \in C(D)$, we define

$$\omega_y(F,D,\delta) = \sup\{|F(x,y_1) - F(x,y_2)| : (x,y_1),(x,y_2) \in D, |y_1 - y_2| \leq \delta\}.$$

Uniform continuity of F on D implies that

$$\lim_{\delta \to 0} \omega_y(F,D,\delta) = 0.$$

We now state the classical Jackson's theorem [5, p. 147].

Theorem 2.2.2. (Jackson). (i.) If $f \in C(I)$, then

$$\|f - T_n(f,\cdot)\|_I \leq \omega(f,I,\pi/(n+1)).$$

(ii.) If $f^{(k)} \in C(I)$, $n \geq k$, then

$$\|f - T_n(f,\cdot)\|_I \leq (\pi/2)^k \|f^{(k)}\|_I / [(n+1)n\ldots(n-k+2)].$$

The product approximation analog of Jackson's theorem will follow from the next three lemmas.

Lemma 2.2.3. If $F \in C(D)$ and $T_n(F_y,\cdot)$ is given by (2.2), then for y_1, $y_2 \in J$ and $i = 0,\ldots,n$,

$$|f_i(y_1) - f_i(y_2)| \le \begin{cases} \| T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I, & i = 0, \\ \sqrt{2}\|T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I, & i > 0. \end{cases}$$

Proof. Using (2.3),

$$\int_{-1}^{1}[T_n(F_{y_1},x) - T_n(F_{y_2},x)]p_i(x)(1-x^2)^{-1/2}dx$$

$$= [f_i(y_1) - f_i(y_2)] \int_{-1}^{1}[p_i(x)]^2(1-x^2)^{-1/2}dx$$

$$= \begin{cases} \pi[f_i(y_1) - f_i(y_2)], & i = 0, \\ (\pi/2)[f_i(y_1) - f_i(y_2)], & i > 0. \end{cases}$$

For $i = 0$,

$$|f_0(y_1) - f_0(y_2)|$$

$$= \frac{1}{\pi}\left| \int_{-1}^{1}[T_n(F_{y_1},x) - T_n(F_{y_2},x)](1-x^2)^{-1/2}dx \right|$$

$$\le \frac{1}{\pi} \| T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I \int_{-1}^{1}(1-x^2)^{-1/2}dx$$

$$= \| T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I .$$

For $i > 0$, we apply the Cauchy-Schwarz inequality to get

$$|f_i(y_1) - f_i(y_2)|$$

$$= \frac{2}{\pi}\left|\int_{-1}^{1}[T_n(F_{y_1},x) - T_n(F_{y_2},x)]p_i(x)(1-x_2)^{-1/2}\,dx\right|$$

$$\leq \frac{2}{\pi}\left\{\int_{-1}^{1}[T_n(F_{y_1},x) - T_n(F_{y_2},x)]^2(1-x^2)^{-1/2}dx\right\}^{1/2}$$

$$\cdot \left\{\int_{-1}^{1}[p_i(x)]^2(1-x^2)^{-1/2}dx\right\}^{1/2}$$

$$= \sqrt{\frac{2}{\pi}}\left\{\int_{-1}^{1}[T_n(F_{y_1},x) - T_n(F_{y_2},x)]^2(1-x^2)^{-1/2}dx\right\}^{1/2}$$

$$= \sqrt{\frac{2}{\pi}}\,\|T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I\left\{\int_{-1}^{1}(1-x^2)^{-1/2}dx\right\}^{1/2}$$

$$= \sqrt{2}\,\|T_n(F_{y_1},\cdot) - T_n(F_{y_2},\cdot)\|_I.$$

<u>Lemma 2.2.4</u>. If $F \in C(D)$ and f_i, $i = 0,\ldots,n$, are given by (2.2), then for $i = 0,\ldots,n$,

$$\omega(f_i,J,\delta) \leq \begin{cases} \omega_y(F,D,\delta) + 2\max\limits_{y\in J}\|F_y - T_n(F_y,\cdot)\|_I, & i = 0, \\ \sqrt{2}(\omega_y(F,D,\delta) + 2\max\limits_{y\in J}\|F_y - T_n(F_y,\cdot)\|)_I, & i > 0. \end{cases}$$

Proof. For y_1, $y_2 \in J$ where $|y_1 - y_2| \leq \delta$,

$$\|T_n(y_1,\cdot) - T_n(y_2,\cdot)\|_I$$

$$\leq \|F_{y_1} - F_{y_2}\|_I + \|F_{y_1} - T_n(F_{y_1},\cdot)\|_I + \|F_{y_2} - T_n(F_{y_2},\cdot)\|_I$$

$$\leq \omega_y(F,D,\delta) + 2 \max_{y\in J} \|F_y - T_n(F_y,\cdot)\|_I .$$

This inequality and Lemma 2.2.3 now complete the proof of Lemma 2.2.4.

Lemma 2.2.5. If $F \in C(D)$, then

$$\|F - \mathcal{P}_{n,m}F\|_D \leq (3 + 2n\sqrt{2}) \max_{y\in J} \|F_y - T_n(F_y,\cdot)\|_I$$
$$+ (1+n\sqrt{2})\omega_y(F,D,\pi/(m+1)).$$

Proof. For $(x,y) \in D$,

$$(2.4)\quad |F(x,y) - (\mathcal{P}_{n,m}F)(x,y)| \leq |F_y(x) - T_n(F_y,x)|$$
$$+ |T_n(F_y,x) - (\mathcal{P}_{n,m}F)(x,y)|$$

$$= |F_y(x) - T_n(F_y,x)| + \left| \sum_{i=0}^{n} [f_i(y) - Q_m(f_i,y)]p_i(x) \right|$$

$$\leq \|F_y - T_n(F_y,\cdot)\|_I + \sum_{i=0}^{n} \|f_i - Q_m(f_i,\cdot)\|_J \, \|p_i\|_I$$

$$= \|F_y - T_n(F_y,\cdot)\|_I + \sum_{i=0}^{n} \|f_i - Q_m(f_i,\cdot)\|_J .$$

By Jackson's theorem

$$\|f_i - Q_m(f_i,\cdot)\|_J \le \omega(f_i,J,\pi/(m+1)).$$

Applying Lemma 2.2.4, we thus have

$$\|f_0-Q_m(f_0,\cdot)\|_J \le \omega_y(F,D,\pi/(m+1)) + 2 \max_{y\in J} \|F_y - T_n(F_y,\cdot)\|_I$$

and for $i > 0$,

$$\|f_i-Q_m(f_i,\cdot)\|_I \le \sqrt{2}(\omega_y(F,D,\pi/(m+1)) + 2\max_{y\in J}\|F_y-T_n(F_y,\cdot)\|_I).$$

Combining these with (2.4) we get

$$\|F - \mathcal{P}_{n,m}F\|_D \le (3+2n\sqrt{2}) \max_{y\in J} \|F_y - T_n(F_y,\cdot)\|_I$$
$$+ (1+n\sqrt{2})\, \omega_y(F,D,\pi/(m+1)).$$

Thus the proof of Lemma 2.2.5 is complete.

Under appropriate smoothness conditions on F, the next theorem and its immediate corollary indicate how the $N(\epsilon)$ and $M(\epsilon,n)$ of Theorem 2.2.1 may be chosen.

Theorem 2.2.6. If $\frac{\partial^k F}{\partial x^k} \in C(D)$, $n \ge k$, then