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:

Robert D. Engle
Head, Major Department

Myron J Henry
Chairman, Examining Committee

Henry L Parsons
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

LIST OF TABLES

LIST OF DIAGRAMS

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].

<u>Theorem 1.2.2</u>. (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].

<u>Theorem 1.2.3</u>. 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 $f \in C(I)$.

<u>Theorem 1.2.4</u>. (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

$$\|f - \varphi\|_I \geq \|f - Tf\|_I + \gamma \|\varphi - Tf\|_I \tag{1.2}$$

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

$$(1.3) \qquad \|Tg - Tf\|_I \leq \gamma \|g - f\|_I$$

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

$$\|Tg - Tf\|_I \leq \lambda_\Gamma \|g - f\|_I \tag{1.5}$$

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 $Tg \in \Phi$,

$$\gamma_f \|Tg - Tf\| \leq \|f - Tg\|_I - \|f - Tf\|_I$$
$$\leq \|f - g\|_I + \|g - Tg\|_I - \|f - Tf\|_I$$
$$\leq \|f - g\|_I + \|g - Tf\|_I - \|f - Tf\|_I$$
$$\leq 2\|g - f\|_I.$$

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 $a \leq x_0 < x_1 < \ldots < x_n \leq b$, define

(1.6) $K(x_0,\ldots,x_n)$
$$= \sup\{\|\varphi\|_I : \varphi \in \Phi \text{ and } (-1)^i\varphi(x_i) \geq -1,\ i = 0,\ldots,n\}.$$

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

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

Proof. By the Haar condition,

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

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

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

where $D[x_1^k,\dots,x_{i-1}^k,x,x_{i+1}^k,\dots,x_n^k]$ and $D[x_1^k,\dots,x_n^k]$ are given by (1.1) and $\{\varphi_1,\dots,\varphi_n\}$ is a basis for Φ. Since $\varphi_1,\dots,\varphi_n$ 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 $k \to \infty$, $i = 1,\dots,n$, and $D[x_1,\dots,x_n] \neq 0$,

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

$$= y_j$$

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

$j = 1,\ldots,n$. 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 $f - Tf$,

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

$i = 0,\ldots,n$, 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 $i = 0,\ldots,n$,

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

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

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

$i = 0,\ldots,n$. 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

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

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

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

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

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.

Theorem 1.3.7. 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

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

as $k \to \infty$.

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 $\|f_k^*\|_I < 1/k$, $f_k^* \to 0$ 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 $x_0 < x_1 < \dots < x_n$ constitute an alternation set for $f - T_X f$ in X. Then

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

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

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

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

$i = 0, \dots, n$, 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 $i = 0,\dots,n$,

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

Subtracting (1.13) from (1.12), we get

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

Hence,

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

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

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

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

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

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

Let $M = \sup\{\|\varphi\|_I : \varphi \in \Phi \text{ and } \forall i = 0,\dots,n, \exists \xi_i \in [\alpha_i,\beta_i] \ni (-1)^i \varphi(\xi_i) \geq -1\}$. 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 $X \cap [\alpha_i,\beta_i] \neq \emptyset$, $i = 0,\dots,n$.

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

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

By the definition of M,

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

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,

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

and

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

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

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

Let $x_0^k < x_1^k < \dots < x_n^k$ be an alternation set for $S_k = f - T_{X_k} f$ in X_k. We extract subsequences and relabel so that $x_i^k \to x_i \in I$, $i = 0,\dots,n$. We now show that $x_{i-1} < x_i$, $i = 1,\dots,n$. 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,

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

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

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

Then

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

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

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

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

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

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

Hereafter, $\mathscr{P}$ 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 $(\mathscr{P}F)(x,y) = T(g,x)Q(h,y)$. In addition, if $F(x,y) = g(x) + h(y)$ and if Φ and Ψ contain the constant functions, then $(\mathscr{P}F)(x,y) = T(g,x) + Q(h,y)$. 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 $\mathscr{P}$.

Two questions regarding possible invariance properties of $\mathscr{P}$ naturally arise. Is $\mathscr{P}$ 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

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

is it necessarily true that $\mathscr{P}F \equiv \mathscr{P}^*F$? Also, is $\mathscr{P}F$ 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 $\mathscr{P}F \equiv 0$. 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 <u>modulus of continuity</u> [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].

<u>Theorem 2.2.2</u>. (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 $i = 0, \ldots, n$,

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

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 $i = 0$,

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

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

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

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

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

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

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

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

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

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

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

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

Lemma 2.2.4. 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.

<u>Lemma 2.2.5</u>. 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

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

Applying Lemma 2.2.4, we thus have

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

and for $i > 0$,

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

Combining these with (2.4) we get

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

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.

<u>Theorem 2.2.6</u>. If $\frac{\partial^k F}{\partial x^k} \in C(D)$, $n \geq k$, then

$$(2.5)\quad \|F-\mathcal{P}_{n,m}F\|_D \le \left(\frac{\pi}{2}\right)^k \left\|\frac{\partial^k F}{\partial x^k}\right\|_D (3+2n\sqrt{2})/[(n+1)n\ldots(n-k+2)]$$

$$+ (1+n\sqrt{2})\ \omega_y(F,D,\pi/(m+1)).$$

Proof. By Jackson's theorem, for $y \in J$

$$\|F_y - T_n(F_y,\cdot)\|_I \le \left(\frac{\pi}{2}\right)^k \left\|\frac{\partial^k F_y}{\partial x^k}\right\|_I / [(n+1)n\ldots(n-2+k)]$$

Combining this inequality with that of Lemma 2.2.5, Theorem 2.2.6 now follows.

If $k \ge 2$, the first term in the bound of $\|F - \mathcal{P}_{n,m}F\|_D$ in (2.5) converges to zero as $n \to \infty$ and the second term can be forced to tend to zero by choosing m sufficiently large. This is more easily seen in the following corollary of Theorem 2.2.6.

Corollary 2.2.7. If $\frac{\partial^2 F}{\partial x^2}$, $\frac{\partial F}{\partial y} \in C(D)$, $n \ge 2$, then

$$(2.6)\quad \|F - \mathcal{P}_{n,m}F\|_D \le \left(\frac{\pi}{2}\right)^2 \left\|\frac{\partial^2 F}{\partial x^2}\right\|_D (3 + 2n\sqrt{2})/[n(n+1)]$$

$$+ \pi \left\|\frac{\partial F}{\partial y}\right\|_D (1 + n\sqrt{2}) / (m+1).$$

Proof. We need only note that by the mean value theorem,

$$\omega_y(F, D, \pi/(m+1)) \le \left\|\frac{\partial F}{\partial y}\right\|_D \pi/(m+1),$$

and Corollary 2.2.7 now follows from Theorem 2.2.6.

2.3 Continuity of the operator $\mathcal{P}$.

In this section, we prove that the product approximation operator $\mathcal{P}$ defined in (2.1) is continuous with respect to the uniform norm $\|\cdot\|_D$ on D. The continuity of $\mathcal{P}$ is of particular importance in the computation of product approximations. Due to inherent errors made by computing machinery, when we set out to find $\mathcal{P}F$, we actually obtain the product approximation of a function that is uniformly close to F. Continuity of $\mathcal{P}$ would insure that the resulting product approximation is uniformly near $\mathcal{P}F$.

We retain the notations of Section 2.1 and fix $F \in C(D)$. As in Section 2.1, we let

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

be the best approximation of $F_y(x) = F(x,y)$. We define

$$\rho(y) = \|F_y - T(F_y, \cdot)\|_I.$$

The uniform continuity of F on D and Theorem 1.2.6 imply that ρ is continuous on J. The continuity of $\mathcal{P}$ will follow from the next two lemmas.

Lemma 2.3.1. Given $\epsilon > 0$, there is a $\delta = \delta(F,\epsilon) > 0$ such that if $(\alpha_1(y),\ldots,\alpha_n(y)) \in [C(J)]^n$ and

$$\|F_y - \sum_{i=1}^{n} \alpha_i(y)\varphi_i\|_I < \rho(y) + \delta$$

for all $y \in J$, then $\max_{i=1,\ldots,n} |f_i(y) - \alpha_i(y)| < \epsilon$ for all $y \in J$.

Proof. Assume otherwise. Then there is a sequence $\{(\alpha_1^k(\cdot),\ldots,\alpha_n^k(\cdot))\}$ in $[C(J)]^n$ such that

$$(2.7) \qquad \|F_y - \sum_{i=1}^{n} \alpha_i^k(y)\varphi_i\|_I < \rho(y) + 1/k$$

for all $y \in J$ and such that for each k there is a $y_k \in J$ where

(2.8) $$\max_{i=1,\ldots,n} |f_i(y_k) - \alpha_i^k(y_k)| \geq \epsilon.$$

Since J is compact, we may assume after relabeling that $y_k \to y^* \in J$. By (2.7)

$$\left\|\sum_{i=1}^{n} \alpha_i(y_k)\varphi_i\right\|_I \leq \|F\|_D + \|\rho\|_J + 1,$$

and since $\{\varphi_1,\ldots,\varphi_n\}$ is linearly independent, $\{(\alpha_1^k(y_k),\ldots,\alpha_n^k(y_k))\}$ is a bounded sequence. We may further extract a subsequence and relabel so that $\alpha_i^k(y_k) \to \lambda_i$, $i = 1,\ldots,n$. Passing to the limit in (2.7) we have

$$\|F_{y^*} - \sum_{i=1}^{n} \lambda_i\varphi_i\|_I \leq \rho(y).$$

By the uniqueness of best approximations, $\lambda_i = f_i(y^*)$, $i = 1,\ldots,n$. However, passing to the limit in (2.8) gives

$$\max_{i=1,\ldots,n} |f_i(y^*) - \lambda_i| \geq \epsilon.$$

which is a contradiction, and thus Lemma 2.3.1 is proven.

Lemma 2.3.1 is a generalization of Weinstein's Theorem 2.2 in [30] which implies the continuity of the coefficient functions in $T(F_y,\cdot)$ with respect to y. We now apply Lemma 2.3.1 to establish the continuous dependence of the coefficient functions in $T(F_y,\cdot)$ on $F \in C(D)$.

<u>Lemma 2.3.2</u>. Given $\epsilon > 0$, there is a $\delta = \delta(F,\epsilon) > 0$ such that if $G \in C(D)$ and $\|G - F\|_D < \delta$, then $\|g_i - f_i\|_J < \epsilon$, $i = 1,\ldots,n$, where $G_y(x) = G(x,y)$ and $T(G_y,\cdot) = \sum_{i=1}^{n} g_i(y)\varphi_i$.

Proof. Denote

$$\mu(y) = \|G_y - T(G_y,\cdot)\|_I.$$

Note that

$$\begin{aligned}\mu(y) - \rho(y) &= \|G_y - T(G_y,\cdot)\|_I - \|F_y - T(F_y,\cdot)\|_I \\ &\leq \|G_y - T(F_y,\cdot)\|_I - \|F_y - T(F_y,\cdot)\|_I \\ &\leq \|G_y - F_y\|_I \\ &\leq \|G - F\|_D.\end{aligned}$$

Let $\epsilon > 0$ be given. By Lemma 2.3.1, we may select $\delta > 0$ such that $\|F_y - T(G_y,\cdot)\|_I < \rho(y) + \delta$ for all $y \in J$ implies that $\max_{i=1,\ldots,n} |f_i(y) - g_i(y)| < \epsilon$ for all $y \in J$. Now suppose $\|G - F\|_D < \delta/2$. Then $\mu(y) - \rho(y) < \delta/2$ and

$$\begin{aligned}\|F_y - T(G_y,\cdot)\|_I &\leq \|F - G\|_D + \|G_y - T(G_y,\cdot)\|_I \\ &< \delta/2 + \mu(y) \\ &< \rho(y) + \delta .\end{aligned}$$

Hence, $\|g_i - f_i\|_J < \epsilon$, $i = 1,\ldots,n$.

<u>Theorem 2.3.3</u>. The product approximation operator $\mathcal{P}$ is a continuous mapping of $C(D)$ into the approximating space spanned by $\{\psi_j\varphi_i : i = 1,\ldots,n;\ j = 1,\ldots,m\}$.

Proof. Let $F \in C(D)$ and $\epsilon > 0$ be given. Since the Tchebycheff approximation operator Q is continuous at $f_1,\ldots,f_n$, there is a $\sigma > 0$ such that if $\|g_i - f_i\|_J < \sigma$, $i = 1,\ldots,n$, then

$$\|Q(g_i,\cdot) - Q(f_i,\cdot)\|_I < \frac{\epsilon}{n\|\varphi_i\|_I} ,$$

$i = 1,\ldots,n$. By Lemma 2.3.2, we may select $\delta > 0$ so that if $\|G - F\|_D < \delta$, then $\|g_i - f_i\|_J < \sigma$, $i = 1,\ldots,n$, where $G_y(x) = G(x,y)$ and $T(G_y,\cdot) = \sum_{i=1}^n g_i(y)\varphi_i$.

Now if $G \in C(D)$ and $\|G - F\|_D < \delta$, then

$$\|\mathcal{P}G - \mathcal{P}F\|_D \leq \sum_{i=1}^n \|\varphi_i\|_I \|Q(g_i,\cdot) - Q(f_i,\cdot)\|_J$$

$$< \sum_{i=1}^n \|\varphi_i\|_I \cdot \frac{\epsilon}{n\|\varphi_i\|_I}$$

$$= \epsilon.$$

Hence, $\mathcal{P}$ is continuous at F.

2.4 Point-Lipschitz continuity of $\mathcal{P}$.

The question of when the operator $\mathcal{P}$ satisfies the stronger version of point-Lipschitz continuity is now considered. It is shown that if F satisfies an additional condition, then $\mathcal{P}$ satisfies a point-Lipschitz condition at F, and an example is given to show that this condition is essential. We retain the notations of Sections 2.1 and 2.3.

Theorem 2.4.1. If $F \in C(D)$ and $F_y \notin \Phi$ for all $y \in J$, then there is a constant $\lambda_F > 0$ such that

$$\|\mathcal{P}G - \mathcal{P}F\|_D \leq \lambda_F \|G - F\|_D$$

for all $G \in C(D)$.

Proof. We note that $\{F_y : y \in J\}$ is a compact subset of $C(I)$ which does not meet Φ. By Theorem 1.3.7, there is constant $\lambda > 0$ such that

$$(2.9) \qquad \|T(G_y,\cdot) - T(F_y,\cdot)\|_I \leq \lambda \|G_y - F_y\|_I$$

for all $G \in C(D)$. We may apply Theorem 1.2.6 to $f_1,\ldots,f_n$ to obtain constants $\sigma_i > 0$, $i = 1,\ldots,n$, such that

$$(2.10) \qquad \|Q(g,\cdot) - Q(f_i,\cdot)\|_J \leq \sigma_i \|g - f_i\|_J$$

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

Let $G \in C(D)$. Then we may apply (2.9) and (2.10) to get

$$\|\mathcal{P}G - \mathcal{P}F\|_D \leq \sum_{i=1}^{n} \|\varphi_i\|_I \, \|Q(g_i,\cdot) - Q(f_i,\cdot)\|_J$$

$$\leq \sum_{i=1}^{n} \sigma_i \|\varphi_i\|_I \, \|g_i - f_i\|_J.$$

On the finite dimensional space Φ, the two norms $\|\cdot\|_I$ and $\|\alpha_1\varphi_1 + \ldots + \alpha_n\varphi_n\|_* = \max_{i=1,\ldots,n} |\alpha_i|$ are equivalent. Let $K > 0$ be such that $\|\varphi\|_* \leq K\|\varphi\|_I$ for all $\varphi \in \Phi$. For fixed $y \in J$ and $\nu = 1,\ldots,n$,

$$\begin{aligned} |g_\nu(y) - f_\nu(y)| &\leq \max_{i=1,\ldots,n} |g_i(y) - f_i(y)| \\ &\leq K\|T(G_y,\cdot) - T(F_y,\cdot)\|_I \\ &\leq K\lambda\|G_y - F_y\|_I \\ &\leq K\lambda\|G - F\|_D. \end{aligned}$$

Finally, we obtain

$$\|\mathcal{P}_G - \mathcal{P}_F\|_D \leq \left[\lambda K \sum_{i=1}^{n} \sigma_i\|\varphi_i\|_I\right] \|G - F\|_D,$$

and we may take $\lambda_F = \lambda K \sum_{i=1}^{n} \sigma_i\|\varphi_i\|_I$.

We conclude this chapter with an example that shows that the conditions of Theorem 2.4.1 are minimal in a sense. The example uses two fundamental results which we now present. Although a less specialized version holds, the first result is stated and proved for the particular setting in which it will be used.

Lemma 2.4.2. Suppose $h:[a,b] \to [-1,1]$ is continuous and $h(a) = 0$. Then there is a continuous function $\delta:(0,1] \to (a,b]$ such that for each $y \in (0,1]$ if $a \leq x < \delta(y)$, then $|h(x)| < y$.

Proof. Define $\theta(x) = \sup \{|h(t)|:t \in [a,x]\}$. Clearly, θ is monotone increasing. We show that θ is continuous. Let $\epsilon > 0$ be given. There is a $\tau > 0$ such that if $s,t \in [a,b]$ and $|s - t| < \tau$, then $||h(s)| - |h(t)|| \leq |h(s) - h(t)| < \epsilon$. Let $x_1,x_2 \in [a,b]$ such that $|x_1 - x_2| < \tau$. Without loss of generality, $x_1 < x_2$. Then $\theta(x_1) \leq \theta(x_2)$. By the choice of τ, $\sup \{|h(t)|:t \in [x_1,x_2]\} \leq |h(x_1)| + \epsilon$. Thus

$$\begin{aligned}\theta(x_2) &= \max \{\theta(x_1), \sup \{|h(t)|:t \in [x_1,x_2]\}\}\\ &\leq \max \{\theta(x_1), |h(x_1)| + \epsilon\}\\ &\leq \max \{\theta(x_1), \theta(x_1) + \epsilon\}\\ &= \theta(x_1) + \epsilon.\end{aligned}$$

Thus θ is continuous.

Define $\zeta(x) = x - a + \theta(x)$. Note that ζ is a continuous and strictly increasing function of $[a,b]$ onto

$[0,A]$ where $A = b - a + \theta(b)$. Thus $\zeta^{-1}:[0,A] \to [a,b]$ exists and is continuous. Define $\delta:(0,\infty) \to (a,b]$ by

$$\delta(y) = \begin{cases} \zeta^{-1}(y), & 0 < y \leq A, \\ b, & y \geq A. \end{cases}$$

Since $\zeta^{-1}(A) = b$, δ is continuous. If $y > A$ and $a \leq x < \delta(y)$, then

$$|h(x)| \leq \theta(x) \leq x - a + \theta(x) = \zeta(x) \leq \zeta(b) = A < y.$$

If $0 < y \leq A$ and $a < x < \delta(y)$, then

$$\zeta(x) < \zeta(\delta(y)) = \zeta(\zeta^{-1}(y)) = y$$

and

$$|h(x)| \leq \theta(x) \leq \zeta(x) < y.$$

Thus Lemma 2.4.2 is established.

Although the best approximation operator T is known to be nonlinear, the next result asserts a linearity property of T.

<u>Lemma 2.4.3</u>. Let Φ be an n-dimensional Haar subspace of $C(I)$ and for $f \in C(I)$ let $T(f,\cdot)$ denote the best

uniform approximation of f over I from Φ. If $f, g \in C(I)$, $f - T(f,\cdot)$ and $g - T(g,\cdot)$ have a common alternation set, $(f - T(f,\cdot))(g - T(g,\cdot)) \geq 0$ on this alternation set, and $\alpha, \beta \geq 0$, then

$$T(\alpha f+\beta g,\cdot) = \alpha T(f,\cdot) + \beta T(g,\cdot).$$

Proof. Let $x_0 < x_1 < \dots < x_n$ constitute this alternation set. Since

$$\|(\alpha f+\beta g) - (\alpha T(f,\cdot) + \beta T(g,\cdot))\|_I$$
$$\leq \alpha \|f - T(f,\cdot)\|_I + \beta \|g - T(g,\cdot)\|_I$$

and for $i = 0,\dots,n$

$$|(\alpha f(x_i) + \beta g(x_i)) - (\alpha T(f,x_i) + \beta T(g,x_i))|$$
$$= |\alpha(f(x_i)-T(f,x_i)) + \beta(g(x_i)-T(g,x_i))|$$
$$= \alpha |f(x_i) - T(f,x_i)| + \beta |g(x_i) - T(g,x_i)|$$
$$= \alpha \|f - T(f,\cdot)\|_I + \beta \|g - T(g,\cdot)\|_I,$$

we have that

$$|(\alpha f(x_i) + \beta g(x_i)) - (\alpha T(f,x_i) + \beta T(g,x_i))|$$

$$= \|(\alpha f+\beta g) - (\alpha T(f,\cdot)+\beta T(g,\cdot))\|_I,$$

$i = 0,\ldots,n$. Moreover, it is clear that the residual $(\alpha f+\beta g) - (\alpha T(f,\cdot)+\beta T(g,\cdot))$ alternates in sign at the x_i. By the alternation theorem,

$$T(\alpha f+\beta g,\cdot) = \alpha T(f,\cdot) + \beta T(g,\cdot).$$

Example 2.3. We assume that Φ is an n-dimensional Haar subspace of $C(I)$, $n \geq 2$, and Ψ is an m-dimensional Haar subspace of $C(J)$, $J = [0,1]$, which contains the constant functions. We demonstrate a function $F \in C(D)$ for which $F_y \in \Phi$ for just one $y \in J$ and a sequence $\{G^k\}$ in $C(D)$ such that each $G^k \neq F$ and

$$\frac{\|\mathcal{P}G^k - \mathcal{P}F\|_D}{\|G^k - F\|_D} \to \infty$$

as $k \to \infty$.

Since $n \geq 2$, there is a $\varphi \in \Phi$ such that $\varphi(a) = 0$ and $\|\varphi\|_I = 1$. From Lemma 2.4.2, there is a function $\delta:(0,1] \to (a,b]$ such that for each $y \in (0,1]$, if $a \leq x < \delta(y)$, then $|\varphi(x)| < y$. Define $n + 2$ continuous functions $\delta_\nu : (0,1] \to (0,b]$ by

$$\delta_\nu(y) = a + \frac{\nu}{n+2}\,[\delta(y) - a],$$

$\nu = 1,\ldots,n+2$. Note that $\delta_{n+2} \equiv \delta$. Also let $\delta_0 \equiv a$. Moreover, note that $\delta_0 < \delta_1 < \cdots < \delta_{n+2}$. This is illustrated in Diagram 1 for the case n = 2.

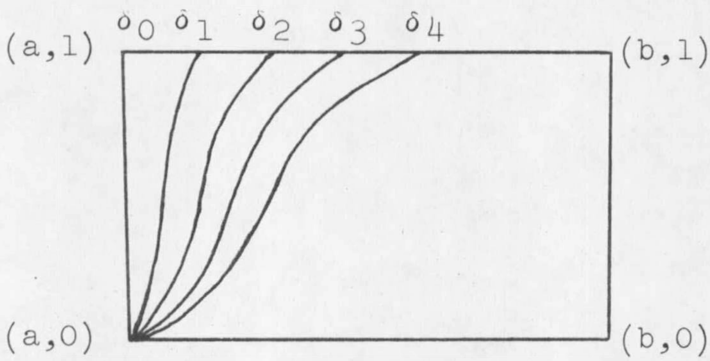

Diagram 1. Sketch of δ_0, δ_1, δ_2, δ_3, δ_4 in the case n = 2.

We define the function $f_y(x)$ by

$$f_y(x) = \begin{cases} (-1)^{\nu}\left[1 - \dfrac{2(x - \delta_{\nu}(y))}{\delta_{\nu+1}(y) - \delta_{\nu}(y)}\right], \\ \qquad \delta_{\nu}(y) \leq x \leq \delta_{\nu+1}(y),\ \nu = 0,\ldots,n-1, \\ (-1)^{n}\left[1 - \dfrac{x - \delta_{n}(y)}{\delta_{n+1}(y) - \delta_{n}(y)}\right], \\ \qquad \delta_{n}(y) \leq x \leq \delta_{n+1}(y), \\ 0, \qquad \delta_{n+1}(y) \leq x \leq b. \end{cases}$$

A sketch of $f_y(x)$ for $n = 2$ and a fixed value of y is given in Diagram 2.

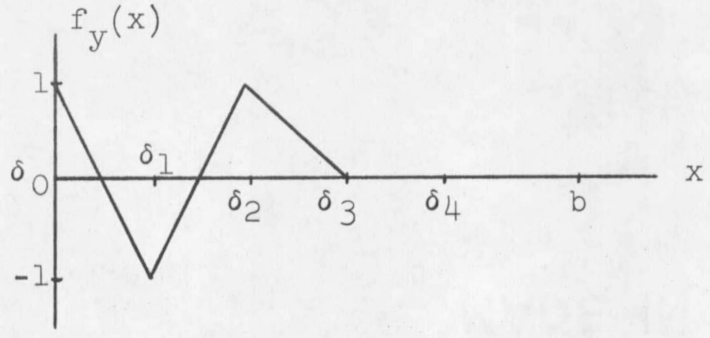

Diagram 2. Sketch of $f_y(x)$ for $n = 2$ and a particular value of y.

By the continuity of each δ_ν, $\{(x,y) \in [a,b] \times (0,1]: \delta_\nu(y) \leq x \leq \delta_{\nu+1}(y)\}$, $\nu = 0,\ldots,n$, and $\{(x,y) \in [a,b] \times (0,1]: \delta_{n+1}(y) \leq x \leq b\}$ are closed subsets of $[a,b] \times (0,1]$. In each of these closed sets, the defining expression for $f_y(x)$ gives a continuous function in x and in y. Moreover, whenever these sets overlap, the two defining expressions involved coincide. Hence, $f_y(x)$ is continuous on $[a,b] \times (0,1]$.

For $y \in (0,1]$, $\nu = 0,\ldots,n$

$$f_y(\delta_\nu(y)) = (-1)^\nu,$$

and by the piecewise linear nature of $f_y(x)$, $\|f_y\|_I = 1$. By the alternation theorem, $T(f_y,\cdot) \equiv 0$.

Define $g_y(x)$ by

$$g_y(x) = \begin{cases} f_y(x) + \varphi(x), & a \leq x \leq \delta_{n+1}(y), \\ \varphi(\delta_{n+1}(y)) \left[1 - \dfrac{x - \delta_{n+1}(y)}{\delta_{n+2}(y) - \delta_{n+1}(y)}\right], & \\ & \delta_{n+1}(y) \leq x \leq \delta_{n+2}(y) \\ 0, & \delta_{n+2}(y) \leq x \leq b. \end{cases}$$

In the same fashion that we showed that $f_y(x)$ is continuous on $[a,b] \times (0,1]$, $g_y(x)$ can be shown to be continuous on $[a,b] \times (0,1]$.

We show that for $0 < y \leq \frac{1}{2}$, φ is the best approximation of g_y on I from Φ. For $\delta_{n+2}(y) \leq x \leq b$,

$$|g_y(x) - \varphi(x)| = |\varphi(x)| \leq 1.$$

For $a \leq x \leq \delta_{n+1}(y)$,

$$|g_y(x) - \varphi(x)| = |f_y(x)| \leq 1.$$

For $\delta_{n+1}(y) \leq x \leq \delta_{n+2}(y)$, the nature of δ yields

$$|g_y(x) - \varphi(x)| \leq |\varphi(\delta_{n+1}(y))| \left[1 - \frac{x - \delta_{n+1}(y)}{\delta_{n+2}(y) - \delta_{n+1}(y)} \right] + |\varphi(x)|$$

$$\leq 2y$$

$$\leq 1.$$

Thus $\|g_y - \varphi\|_I \leq 1$. Now for $\nu = 0,\ldots,n$

$$(g_y-\varphi)(\delta_\nu(y)) = f_y(\delta_\nu(y)) = (-1)^\nu.$$

Thus $\|g_y - \varphi\|_I = 1$ and $T(g_y,\cdot) = \varphi$.

Next we estimate $\|g_y - f_y\|_I$. If $0 \leq x \leq \delta_{n+1}(y)$, then

$$|g_y(x) - f_y(x)| = |\varphi(y)| < y.$$

For $\delta_{n+1}(y) \leq x \leq \delta_{n+2}(y)$,

$$|g_y(x) - f_y(x)| = |g_y(x)| < y.$$

If $\delta_{n+2}(y) \leq x \leq b$, then

$$|g_y(x) - f_y(x)| = 0 < y.$$

Thus $\|g_y - f_y\|_I \leq y$.

We now summarize some observations that can be made from the analysis above:

(i) $f_y(x)$ and $g_y(x)$ are continuous in x and y on $[a,b] \times (0,1]$,

(ii) $\|f_y\|_I = 1$ and $\|g_y\|_I \leq 1 + y$, $y \in (0,1]$,

(iii) $\|g_y - f_y\|_I \leq y$, $y \in (0,1]$,

(iv) $T(f_y,\cdot) \equiv 0$, $y \in (0,1]$,

(v) $T(g_y,\cdot) = \varphi$, $y \in (0, \frac{1}{2}]$,

(vi) $f_y - T(f_y,\cdot)$ and $g_y - T(g_y,\cdot)$ have the same alternation set with the same sign orientation on the alternation set, $y \in (0,\frac{1}{2}]$, and

(vii) $f_y \notin \Phi$, $y \in (0,1]$, since $f_y \not\equiv 0$ and f_y vanishes at at least n points.

We now pick a basis $\{\varphi_1,\ldots,\varphi_n\}$ for Φ in which $\varphi_1 \equiv \varphi$, and let $\{1,\psi_2,\ldots,\psi_m\}$ be a basis for Ψ.

We now define $F \in C(D)$, $D = I \times J = [a,b] \times [0,1]$, by

$$F(x,y) = \begin{cases} yf_y(x), & 0 < y \leq 1, \\ 0, & y = 0. \end{cases}$$

The continuity and boundedness of $f_y(x)$ on $[a,b] \times (0,1]$ imply that $F(x,y)$ is continuous. By property (vii), $F_y \in \Phi$ only for $y = 0$. By property (iv), $T(F_y,\cdot) \equiv 0$ for all $y \in J$. Thus $\mathcal{P}F \equiv 0$ on D.

For each $k = 2,3,\ldots$, pick a continuous function h_k defined on $[1/(k+1), 1/k]$ such that $h_k(1/(k+1)) = h_k(1/k) = 0$, $0 \leq h_k \leq 1/(k+1)$ on $[1/(k+1), 1/k]$, and there exist points $\xi_0 < \xi_1 < \ldots < \xi_m$ in $[1/(k+1), 1/k]$ such that $h_k(\xi_j) = 0$ if j is even and $h_k(\xi_j) = 1/(k+1)$ if j is odd. Diagram 3 gives a candidate for h_k in the case when $m = 3$.

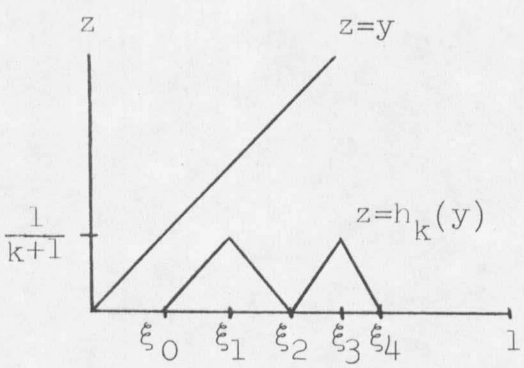

Diagram 3. Illustration of h_k when $m = 3$.

Now for $k = 2,3,\ldots,$ define $G^k \in C(D)$ by

$$G^k(x,y) = \begin{cases} F(x,y), & y \in J - [1/(k+1),\ 1/k] \\ F(x,y) + h_k(y)(g_y(x) - f_y(x)), & \\ & y \in [1/(k+1), 1/k]. \end{cases}$$

We note that $G^k \in C(D)$ and for $y \in J - [1/(k+1), 1/k]$, $T(G^k_y, \cdot) \equiv 0$. For $y \in [1/(k+1), 1/k]$,

$$G^k_y = (y - h_k(y))f_y + h_k(y)g_y.$$

Since $y - h_k(y) \geq 0$ and $h_k(y) \geq 0$, we may apply property (vi) and Lemma 2.4.3, to get

$$T(G_y^k, \cdot) = h_k(y)\varphi_1$$

for $y \in [1/(k+1), 1/k]$. By the oscillating nature of h_k, we see that $\mathcal{P}_{G^k} = \frac{\varphi_1}{2(k+1)}$.

We note that

$$\|\mathcal{P}_{G^k} - \mathcal{P}_F\|_D = \frac{\|\varphi_1\|_I}{2(k+1)} = \frac{1}{2(k+1)} \; .$$

Since

$$(G^k - F)(x,y) = \begin{cases} 0, & y \in J - [1/(k+1), 1/k] \\ h_k(y)(g_k(y) - f_k(y)), & y \in [1/(k+1), 1/k], \end{cases}$$

we have that

$$\|G^k - F\|_D \leq \max_{\frac{1}{k+1} \leq y \leq \frac{1}{k}} h_k(y) \max_{\frac{1}{k+1} \leq y \leq \frac{1}{k}} \|g_y - f_y\|_I$$

$$\leq \frac{1}{(k+1)k} \; .$$

Finally,

$$\frac{\|\mathcal{P}G^k - \mathcal{P}F\|_D}{\|G^k - F\|_D} \geq \frac{k}{2}$$

Thus the conditions of Theorem 2.4.1 are minimal. Finally, we note that the choice of $\varphi_1 = \varphi$ was made for convenience only. Had we used a different basis $\{\varphi_1^*,\ldots,\varphi_n^*\}$ for Φ, we would write $\varphi = \sum_{i=1}^n \alpha_i \varphi_i^*$ and see that $T(G_y^k,\cdot) = \sum_{i=1}^n \alpha_i h_k(y)\varphi_i^*$ for $y \in [1/(k+1),1/k]$. As a result

$$\mathcal{P}G^k = \sum_{i=1}^n \frac{\alpha_i}{2(k+1)} \varphi_i^* = \frac{\varphi}{2(k+1)} .$$

Thus the result of Example 2.3 holds for any choice of basis functions for Φ.

CHAPTER III

RATIONAL APPROXIMATION

3.1 Introduction.

In Chapter III we consider the best approximation problem where the approximating space is a set of rational functions. Let $I = [-1,1]$ and non-negative integers n and m be given. Let $\mathbb{P}_n$ and $\mathbb{Q}_m$ denote the sets of real algebraic polynomials of degree less than or equal to n and m, respectively. The degree of a polynomial P will be denoted by ∂P. We agree that the degree of the zero polynomial is $-\infty$. We now define

$$(3.1) \qquad \mathbb{R}(n,m) = \{R = P/Q : P \in \mathbb{P}_n,\ Q \in \mathbb{Q}_m,\ Q > 0 \text{ on } I\}.$$

We discuss the uniform approximation of functions in $C(I)$ by rational functions in $\mathbb{R}(n,m)$. Although the succeeding results are of interest in themselves, this chapter sets the framework for rational product approximation.

Since a rational function can have several representations as a quotient of two polynomials, we identify a parameter space for $\mathbb{R}(n,m)$. Let $\mathbb{C}$ consist of the zero vector of E_{n+m+1}, $(n+m+1)$-dimensional Euclidean space, and all vectors

$$C = (A;B) = (a_0,\ldots,a_n;\ b_1,\ldots,b_m) \in E_{n+m+1}$$

satisfying

(i) at least one $|a_i| > 0$

(ii) $P(A,x) = a_0 + a_1x + \ldots + a_nx^n$ and $Q(B,x) = 1 + b_1x + \ldots + b_mx^m$ have no common factors except constants, and

(iii) $Q(B,x) > 0$ on I.

For $C = (A;B) \in \mathbf{C}$, let $R(C,x) = P(A,x)/Q(B,x)$ where $P(A,x)$ and $Q(B,x)$ are given above. The mapping

$$(3.2) \qquad C \to R(C,x)$$

is a one-to-one correspondence between $\mathbf{R}(n,m)$ and its parameter space $\mathbf{C}$.

The varisolvent degree of $R(C,\cdot) = P(A,\cdot)/Q(B,\cdot)$ at $C = (A;B) \in \mathbf{C}$, is defined to be

$$(3.3) \quad \mu(C) = \begin{cases} 1 + \max\{n + \partial Q(B,\cdot), m + \partial P(A,\cdot)\}, & R(C,\cdot) \not\equiv 0, \\ 1 + n, & R(C,\cdot) \equiv 0 . \end{cases}$$

We identify

$$(3.4) \quad \mathbf{R}^* = \{R(C,\cdot) \in \mathbf{R}(n,m) : C \in \mathbf{C},\ \mu(C) = 1 + n + m\}$$

and its corresponding parameter space

$$(3.5)\quad \mathbf{C}^* = \{(a_0,\ldots,a_n;b_1,\ldots,b_m) \in \mathbf{C}: a_n \neq 0 \text{ or } b_m \neq 0\}.$$

The following existence and uniqueness theorem appears in Cheney [5].

Theorem 3.1.1. For each $f \in C(I)$ there is a unique $R^* \in \mathbf{R}(n,m)$ such that

$$\|f - R^*\|_I = \inf\{\|f - R\|_I : R \in \mathbf{R}(n,m)\}.$$

For $f \in C(I)$, let $C(f) = (A(f);B(f))$ be the unique coefficient vector in $\mathbf{C}$ such that $R(C(f),\cdot)$ is the best uniform approximation of f from $\mathbf{R}(n,m)$ over I. We say that f is normal (relative to I) if $R(C(f),\cdot) \in \mathbf{R}^*$ or, equivalently, $C(f) \in \mathbf{C}^*$. The concept of normality will play an important role in the theory of rational product approximation.

As in the linear case of Chapter I, best rational approximations are characterized by an alternation condition, but the number of required alternations depends on the varisolvent degree of the approximant, see Rice [28].

Theorem 3.1.2. Let $f \in C(I)$ and $R = R(C,\cdot) \in \mathbf{R}(n,m)$. Then R is the best uniform approximation of f from $\mathbf{R}(n,m)$ over I if and only if there are $1 + \mu(C)$ points, $x_0 < x_1 < \cdots < x_{\mu(C)}$, in I such that

(i) $|(f-R)(x_i)| = \|f-R\|_I$, $i = 0,\ldots,\mu(C)$,

(ii) $(f-R)(x_i) = -(f-R)(x_{i-1})$, $i = 1,\ldots,\mu(C)$.

In this case, $\{x_0,\ldots,x_{\mu(C)}\}$ is called an "alternation set" for $f - R$.

The following rational analog of the strong unicity theorem appears in Cheney [5].

Theorem 3.1.3. Let $f \in C(I)$ be normal. Then there is a constant $\gamma > 0$ such that

$$(3.6) \qquad \|f-R\|_I \geq \|f-R(C(f),\cdot)\|_I + \gamma\|R-R(C(f),\cdot)\|_I$$

for all $R \in \mathbf{R}(n,m)$.

Maehly and Witzgall [21] established a local point-Lipschitz condition for the best rational approximation operator. The following global point-Lipschitz condition appears in Werner [32].

Theorem 3.1.4. Let $f \in C(I)$ be normal. Then there is a constant $\lambda > 0$ such that

$$(3.7) \qquad \| R(C(g),\cdot) - R(C(f),\cdot)\|_I \leq \lambda \| g - f\|_I$$

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

The normality conditions of Theorems 3.1.3 and 3.1.4 are essential. In fact, Maehly and Witzgall [21] show by example that the best rational approximation operator may be discontinuous at non-normal points of $C(I)$.

3.2 Relationships involving $\mathbf{R}(n,m)$, $\mathbf{C}$, $\mathbf{R}^*$, and $\mathbf{C}^*$.

In this section, we establish some topological properties relating $\mathbf{R}(n,m)$ to its parameter space $\mathbf{C}$ and $\mathbf{R}^*$ to its parameter space $\mathbf{C}^*$. We emphasize that $\mathbf{R}(n,m)$ and $\mathbf{R}^*$ carry the topologies inherited from the uniform norm on $C(I)$. The parameter spaces $\mathbf{C}$ and $\mathbf{C}^*$ are topologized by the following metric on E_{n+m+1}. For $C = (A;B)$ $= (a_0,\ldots,a_n;\ b_1,\ldots,b_m) \in E_{n+m+1}$ and $C' = (A';B')$ $= (a_0',\ldots,a_n';\ b_1',\ldots,b_m') \in E_{n+m+1}$

$$(3.8) \qquad \sigma(C,C') = \max_{\substack{i=0,\ldots,n \\ j=1,\ldots,m}} \{ |a_i - a_i'|,\ |b_j - b_j'|\}.$$

The first result we cite is an immediate consequence of Theorem 2 of Rice [27].

<u>Theorem 3.2.1</u>. $\mathbf{R}^*$ is an open subset of $\mathbf{R}(n,m)$.

We proceed to establish that $\mathbf{R}^*$ and $\mathbf{C}^*$ are homeomorphic. This is preceded by three lemmas.

<u>Lemma 3.2.2</u>. If $\{p_k(z)\}$ is a sequence of monic polynomials of degree d with coefficients bounded independent of k, then the zeros of $p_k(z)$ are bounded independent of k.

Proof. Let M be a bound on the coefficients of the $p_k(z)$. We may select $r > 0$ such that $r^d > M(1+r+\ldots+r^{d-1})$. Let $p_k(z) = z^d + \alpha_{d-1}z^{d-1} + \ldots + \alpha_1 z + \alpha_0$. On the circle $|z| = r$, we have

$$\left| \sum_{i=0}^{d-1} \alpha_i z^i \right| \leq M \sum_{i=0}^{d-1} r^i < r^d = |z^d|.$$

By Rouche's theorem, z^d and $p_k(z) = z^d + \sum_{i=0}^{d-1} \alpha_i z^i$ have the same number of zeros in $|z| < r$. Since z^d has d zeros in $|z| < r$, namely 0 with multiplicity d, $p_k(z)$ has all of its d zeros in $|z| < r$. Hence, the zeros of $p_k(z)$ are bounded with respect to k.

Lemma 3.2.3. $\mathbf{C}^*$ is an open subset of E_{n+m+1}.

Proof. Let $C^* = (A^*;B^*) = (a_0^*,\ldots,a_n^*;\ b_1^*,\ldots,b_m^*) \in \mathbf{C}^*$. Then some $|a_i^*| > 0$, $Q^* = Q(B^*,\cdot) > 0$ on I, $P^* = P(A^*,\cdot)$ and Q^* have no common factors except constants, and $a_n^* \neq 0$ or $b_m^* \neq 0$.

It is evident that there is a $\delta_1 > 0$ such that if $C = (A;B) = (a_0,\ldots,a_n;\ b_1,\ldots,b_m) \in E_{n+m+1}$ and $\sigma(C,C^*) < \delta_1$, then some $|a_i| > 0$, $Q(B,\cdot) > 0$ on I, and $a_n \neq 0$ or $b_m \neq 0$.

We must now show that there is a $\delta_2 > 0$ such that if $C = (A;B) \in E_{n+m+1}$ and $\sigma(C,C^*) < \delta_2$, then $P(A,\cdot)$ and $Q(B,\cdot)$ have no common factors except constants. Assume otherwise. Then there is a sequence $C^k = (A^k;B^k) = (a_0^k,\ldots,a_n^k;\ b_1^k,\ldots,b_m^k)$ in E_{n+m+1} such that $\sigma(C^k,C^*) \to 0$ as $k \to \infty$ and $P(A^k,\cdot)$ and $Q(B^k,\cdot)$ have a common non-constant factor and, therefore, have a common complex zero z_k. We assume $a_n^* \neq 0$. The proof for the case in which $b_m^* \neq 0$ is similar. Since $\sigma(C^k,C^*) \to 0$, $a_i^k \to a_i^*$, $b_j^k \to b_j^*$ as $k \to \infty$, $i = 0,\ldots,n$, $j = 1,\ldots,m$. Since $a_n^k \to a_n^*$, $a_n^k \neq 0$ for k sufficiently large. The polynomials

$$N^k = P^k/a_n^k = z^n + \sum_{i=0}^{n-1} (a_i^k/a_n^k) z^i$$

are monic, have fixed degree n, and have bounded coefficients. By Lemma 3.2.2, the zeros of N^k are bounded with respect to k. In particular, the complex numbers z_k are bounded. We now extract a subsequence, relabel, and assume that $z_k \to z^*$ as $k \to \infty$. Since $z_k \to z^*$ and $\sigma(C^k, C^*) \to 0$ as $k \to \infty$,

$$P^*(z^*) = \lim_{k\to\infty} P(A^k, z_k) \text{ and } Q^*(z^*) = \lim_{k\to\infty} Q(B^k, z_k).$$

Thus $P^*(z^*) = Q^*(z^*) = 0$. If z^* is real, then $x - z^*$ is a common factor of P^* and Q^*. If z^* is not real, then $(x-z^*)(x-\overline{z^*}) = x^2 - (z^*+\overline{z^*})x + |z^*|^2$ is a common factor of P^* and Q^*, where $\overline{z^*}$ denotes the complex conjugate of z^*. Either way, we get a contradiction.

Thus $\mathbf{C}^*$ is an open subset of E_{n+m+1}.

<u>Lemma 3.2.4</u>. The mapping (3.2) is a continuous one-to-one map of $\mathbf{C}$ onto $\mathbf{R}(n,m)$.

Proof. Let $C^* = (A^*;B^*) \in \mathbf{C}$ and $\epsilon > 0$ be given. Since $Q(B^*,x) > 0$ on I, $0 < \|1/Q(B^*,\cdot)\|_I < \infty$. We may select $\delta > 0$ sufficiently small so that if $C = (A;B) \in \mathbf{C}$ and $\sigma(C,C^*) < \delta$, then

$$\|P(A,\cdot) - P(A^*,\cdot)\|_I < \frac{\epsilon}{2\|1/Q(B^*,\cdot)\|_I}$$

$$\|Q(B,\cdot) - Q(B^*,\cdot)\|_I < \frac{1}{2\|1/Q(B^*,\cdot)\|_I}$$

and

$$\|Q(B,\cdot) - Q(B^*,\cdot)\|_I < \frac{\epsilon}{2\|1/Q(B^*,\cdot)\|_I(1+\|R(C^*,\cdot)\|_I)} .$$

If $C = (A;B) \in \mathbf{C}$ and $\sigma(C,C^*) < \delta$, then for $x \in I$

$$Q(B,x) > Q(B^*,x) - \frac{1}{2\|1/Q(B^*,\cdot)\|_I} \geq \frac{1}{2\|1/Q(B^*,\cdot)\|_I} ,$$

and

$$|R(C,x) - R(C^*,x)| = \left| \frac{P(A,x)}{Q(B,x)} - \frac{P(A^*,x)}{Q(B^*,x)} \right|$$

$$= \frac{1}{Q(B,x)Q(B^*,x)} |P(A,x)Q(B^*,x) - P(A^*,x)Q(B,x)|$$

$$\leq \frac{1}{Q(B,x)} | P(A,x) - P(A^*,x) |$$

$$+ \frac{|R(C^*,x)|}{Q(B,x)} \; |Q(B^*,x) - Q(B,x)|$$

$$< 2\|1/Q(B^*,\cdot)\|_I \, \|P(A,\cdot) - P(A^*,\cdot)\|_I$$

$$+ 2\|1/Q(B^*,\cdot)\|_I \, \|R(C^*,\cdot)\|_I \, \|Q(B^*,\cdot) - Q(B,\cdot\, \|_I$$

$$< \epsilon .$$

Hence, $\|R(C,\cdot) - R(C^*,\cdot)\|_I < \epsilon$, and (3.2) is continuous.

Theorem 3.2.5. The restriction of the mapping (3.2) to $\mathbf{C}^*$ is a homeomorphism of $\mathbf{C}^*$ onto $\mathbb{R}^*$.

Proof. By Lemma 3.2.4, the restriction of (3.2) to $\mathbf{C}^*$ is a continuous one-to-one map of $\mathbf{C}^*$ onto $\mathbb{R}^*$. We now demonstrate that this is an open mapping. Let $\ell = n+m+1$ and select ℓ distinct points, $x_1 < x_2 < \dots < x_\ell$, in I. The evaluation mapping

$$(3.9) \qquad R(C,\cdot) \to (R(C,x_1), \dots, R(C,x_\ell))$$

is a continuous map of $\mathbb{R}^*$ onto a subset $\mathcal{Q}$ of E_{n+m+1}. Corollary (3.6) of [21] implies that (3.10) is one-to-one. Composing the maps (3.2) and (3.9), we see that

(3.10) $$C \to (R(C,x_1), \ldots, R(C,x_\ell))$$

is a continuous one-to-one map of the open subset of E_{n+m+1} onto the subset $\mathcal{Q}$ of E_{n+m+1}. We conclude that (3.10) is a homeomorphism of $\mathbf{C}^*$ onto $\mathcal{Q}$ (see [24, p. 137]). Composing the inverse of (3.10)

(3.11) $$(R(C,x_1), \ldots, R(C,x_\ell)) \to C$$

with (3.9), we see that the inverse of the restriction of (3.2) to $\mathbf{C}^*$ is continuous. Thus $\mathbf{C}^*$ and $\mathbf{R}^*$ are homeomorphic under the map (3.2).

The argument which establishes that (3.11) is continuous uses the Brouwer theorem on invariance of open sets [24]. This can also be proven with an application of the inverse function theorem and the fact that if $R \in \mathbf{R}^*$, then $\{1,x,\ldots,x^n, xR(x),\ldots,x^mR(x)\}$ is a Tchebycheff system on I.

Although $\mathbf{C}^*$ and $\mathbf{R}^*$ are homeomorphic, the following example indicates that $\mathbf{C}$ and $\mathbf{R}(n,m)$ are not homeomorphic.

Example 3.2.1. Let $n = 0$, $m = 1$, and $C_\epsilon = (\epsilon;1/2)$.

Then $R(C_\epsilon,x) = \dfrac{\epsilon}{1+\frac{1}{2}x}$. It is evident that

$R(C_\epsilon, x) \to 0 = R((0;0), x)$ uniformly on I as $\epsilon \to 0$, but $C_\epsilon \not\to (0;0)$ as $\epsilon \to 0$. Thus $\mathbb{R}(0,1)$ and its parameter space are not homeomorphic.

In the case that $m \geq 1$, let

$$\mathbb{Q}'_m = \{Q \in \mathbb{Q}_m : Q(0) = 0\}.$$

For $R^* = R(C^*, \cdot) = P(A^*, \cdot)/Q(B^*, \cdot) = P^*/Q^* \in \mathbb{R}(n,m)$, $C^* = (A^*;B^*) \in \mathbf{C}$, define

$$\mathbb{P}_n + R^*\mathbb{Q}'_m = \{P + R^*Q : P \in \mathbb{P}_n \text{ and } Q \in \mathbb{Q}'_m\}.$$

We observe that $\mathbb{P}_n + R^*\mathbb{Q}'_m$ is a linear subspace of $C(I)$ and is spanned by $\{1, x, \ldots, x^n, xR^*(x), \ldots, x^m R^*(x)\}$. The following lemma asserts that the dimension of $\mathbb{P}_n + R^*\mathbb{Q}'_m$ is $n + m + 1$ when $R^* \in \mathbb{R}^*$.

Lemma 3.2.6. If $R^* \in \mathbb{R}^*$, then $\mathbb{P}_n + R^*\mathbb{Q}'_m$ is an $(n+m+1)$-dimensional Haar subspace of $C(I)$.

Proof. We show that $\{1, x, \ldots, x^n, xR^*(x), \ldots, x^m R^*(x)\}$ is a Tchebycheff system on I. Suppose $P + R^*Q$ vanishes at $n + m + 1$ distinct points in I, where $P \in \mathbb{P}_n$ and $Q \in \mathbb{Q}'_m$. Then the polynomial $PQ^* + P^*Q$ has at least $n + m + 1$ zeros in I. Since $\partial(PQ^* + P^*Q) \leq n + m$, $PQ^* + P^*Q \equiv 0$. Evaluation of $PQ^* + QP^*$ at $x = 0$ indicates that $P(0) = 0$. If $Q \not\equiv 0$, then

$$R^*(x) = \frac{P(x)/x}{Q(x)/x},$$

except wherever both the polynomials $P(x)/x$ and $Q(x)/x$ vanish. As a result, $\mu(C^*) < 1 + n + m$, which is false. Thus $Q \equiv 0$. Since $Q^* > 0$ on I and $PQ^* \equiv 0$, we now have that $P \equiv 0$. Thus $\{1,x,\ldots,x^n, xR^*(x),\ldots,x^mR^*(x)\}$ is a Tchebycheff system on I. This insures that $\mathbf{P}_n + R^*\mathbf{Q}'_m$ is an (n+m+1)-dimensional Haar subspace of C(I).

Theorem 3.2.7. If $R^* = R(C^*,\cdot)$ where $C^* = (A^*;B^*) = (a_0^*,\ldots,a_n^*;\, b_1^*,\ldots,b_m^*) \in \mathbf{C}^*$, then there is a constant $\rho(C^*) > 0$ such that

$$\|P-R^*Q\|_I \geq \rho(C^*)\left(\sum_{i=0}^{n} |a_i-a_i^*| + \sum_{j=1}^{m} |b_j-b_j^*|\right)$$

for all $P(x) = \sum_{i=0}^{n} a_ix^i \in \mathbf{P}_n$ and all $Q(x) = 1 + \sum_{j=1}^{m} b_jx^j$.

Proof. Since $C^* \in \mathbf{C}^*$, $R^* \in \mathbf{R}^*$, and Lemma 3.2.6 insures that $\{1,x,\ldots,x^n,xR^*(x),\ldots,x^mR^*(x)\}$ is a basis for $\mathbf{P}_n + R^*\mathbf{Q}'_m$. By the equivalence of all norms on a finite dimensional normed linear space, there is a constant $\rho(C^*) > 0$ such that

$$\|P + R^*Q\|_I \geq \rho(C^*)\left(\sum_{i=0}^{n} |a_i| + \sum_{j=1}^{m} |b_j|\right)$$

for all $P + R^*Q = \sum_{i=0}^{n} a_i x^i + R^*(x)\sum_{j=1}^{m} b_j x^j \in \mathbf{P}_n + R^*\mathbf{Q}'_m$. In particular, we choose

$$(3.12) \qquad \rho(C^*) = \inf\{\|P + R^*Q\|_I : P(x) = \sum_{i=0}^{n} a_i x^i, \; Q(x) = \sum_{j=1}^{m} b_j x^j, \sum_{i=0}^{n} |a_i| + \sum_{j=1}^{m} |b_j| = 1\}.$$

Now suppose $P(x) = \sum_{i=0}^{n} a_i x^i$ and $Q(x) = 1 + \sum_{j=1}^{m} b_j x^j$. Also let $P^* = P(A^*,\cdot)$ and $Q^* = Q(B^*,\cdot)$. Then

$$\begin{aligned}\|P-R^*Q\|_I &= \|P - P^* + R^*Q - R^*Q\|_I \\ &= \|(P-P^*) + R^*(Q^*-Q)\|_I \\ &\geq \rho(C^*)\left(\sum_{i=0}^{n} |a_i - a_i^*| + \sum_{j=1}^{m} |b_j - b_j^*|\right)\end{aligned}$$

since $(Q-Q^*)(0) = 0$. Thus Theorem 3.2.7 is proven.

We complete this section by generalizing Theorem 3.2.7.

Theorem 3.7.8. If $\mathcal{C}$ is a compact subset of $\mathbf{C}^*$, then $\inf_{C \in \mathcal{C}} \rho(C) > 0$ where $\rho(C)$ is defined in (3.12).

Proof. We first verify that $\rho(C)$ is a continuous function of C over $\mathbf{C}^*$. Let $C_1, C_2 \in \mathbf{C}^*$. Since $\mathbf{P}_n + R(C_1,\cdot)\mathbf{Q}'_m$ is finite dimensional, we may choose $P_1(x) = \sum_{i=0}^{n} a_i^1 x^i$ and $Q_1(x) = \sum_{j=1}^{m} b_j^1 x^j$ such that $\sum_{i=0}^{n} |a_i| + \sum_{j=1}^{m} |b_j| = 1$ and $\|P_1 + R(C_1,\cdot)Q_1\|_I = \rho(C_1)$. Thus

$$\begin{aligned}
\rho(C_2) - \rho(C_1) &\leq \|P_1 - R(C_2,\cdot)Q_1\|_I - \|P_1 - R(C_1,\cdot)Q_1\|_I \\
&\leq \|Q_1(R(C_2,\cdot) - R(C_1,\cdot))\|_I \\
&\leq \|Q_1\|_I \, \|R(C_2,\cdot) - R(C_1,\cdot)\|_I \\
&\leq m\|R(C_2,\cdot) - R(C_1,\cdot)\|_I .
\end{aligned}$$

Reversing the roles of C_1 and C_2 in the above argument, we have

$$(3.13) \qquad |\rho(C_2) - \rho(C_1)| \leq m\|R(C_2,\cdot) - R(C_1,\cdot)\|_I .$$

Inequality (3.13) and Theorem 3.2.5 now insure that $\rho(C)$ is a continuous function of C over $\mathbf{C}^*$.

From Theorem 3.2.7, $\rho(C)$ is a positive valued continuous function over the compact set $\mathcal{C}$. This implies that $\inf_{C\in\mathcal{C}} \rho(C)$ is attained by some $C \in \mathcal{C}$, and, hence,

$\inf_{C\in\mathcal{C}} \rho(C) > 0$.

3.3 Uniform strong unicity theorem and uniform Freud's theorem for $R(C(f),\cdot)$.

In this section, we establish theorems for the best rational approximation operator that correspond to Theorems 1.3.6 and 1.3.7 for the best polynomial approximation operator. The rational Freud's theorem is due to Maehly and Witzgall [21]. The particular version stated in Theorem 3.1.4 is due to Werner [32]. As in the case of linear product approximation, the uniform rational Freud's theorem will be applied to the continuity theory for rational product approximation.

We will make use of the following two lemmas which appear on pages 164 and 165 of Cheney [5]. For $R^* \in \mathbf{R}(n,m)$, we let

$$\mathbf{P}_n + R^*\mathbf{Q}_m = \{P + R^*Q : P \in \mathbf{P}_n,\ Q \in \mathbf{Q}_m\}.$$

If $R^* = P^*/Q^*$, then $P^* - R^*Q^* \equiv 0$. Thus

$$\dim (\mathbb{P}_n + R^*\mathbb{Q}_m) \leq 1 + n + m . \tag{3.14}$$

Lemma 3.3.1. Let $f \in C(I)$. If $\mathbb{P}_n + R(C(f),\cdot)\mathbb{Q}_m$ is Haar subspace of $C(I)$, then 0 is the only element φ of $\mathbb{P}_n + R(C(f),\cdot)\mathbb{Q}_m$ such that $\varphi(y)[f(y) - R(C(f),y] \geq 0$ for all $y \in Y = \{y \in I: |f(y) - R(C(f),y)| = \|f - R(C(f),\cdot)\|_I\}$.

Lemma 3.3.2. Let $R^* = P^*/Q^* \in \mathbb{R}(n,m)$ where $Q^* > 0$ on I and $\dim (\mathbb{P}_n + R^*\mathbb{Q}_m) = 1 + n + m$. If $P \in \mathbb{P}_n$, $Q \in \mathbb{Q}_m$, $\|P\|_I + \|Q\|_I = \|P^*\|_I + \|Q^*\|_I$, $P = R^*Q$, and $Q \geq 0$ on I, then $P = P^*$ and $Q = Q^*$.

We first present the uniform strong unicity theorem for rational approximation.

Theorem 3.3.3. Let Γ be a compact subset of $C(I)$ where $\Gamma \cap \mathbb{R}(n,m) = \emptyset$ and f is normal for all $f \in \Gamma$. Then there is a constant $\gamma_\Gamma > 0$ such that

$$\|f - R\|_I \geq \|f - R(C(f),\cdot)\|_I + \gamma_\Gamma \|R - R(C(f),\cdot)\|_I$$

for all $f \in \Gamma$ and all $R \in \mathbb{R}(n,m)$.

Proof. Assume otherwise. Then there are sequences $\{f_k\}$ in Γ and $\{R^k\}$ in $\mathbf{R}(n,m)$ such that each $R^k \neq R(C(f_k),\cdot)$ and

$$(3.15) \qquad \gamma_k = \frac{\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I}{\|R^k - R(C(f_k),\cdot)\|_I} \to 0$$

as $k \to \infty$. The compactness of Γ allows us to extract subsequences and relabel so that $f_k \to f \in \Gamma$ uniformly over I. The normality of f and Theorem 3.1.4 insure that $R(C(f_k),\cdot) \to R(C(f),\cdot)$ uniformly over I.

We renormalize the rational functions by dividing numerator and denominator by the sum of the norms of the numerator and the denominator. Thus we write $R^k = N^k/D^k$, $R(C(f_k),\cdot) = N_{f_k}/D_{f_k}$, and $R(C(f),\cdot) = N_f/D_f$ where

$$(3.16) \quad \|N^k\|_I + \|D^k\|_I = \|N_{f_k}\|_I + \|D_{f_k}\|_I = \|N_f\|_I + \|D_f\|_I = 1$$

and $D^k > 0$, $D_{f_k} > 0$, and $D_f > 0$ over I. Equation (3.16) implies that N^k and D^k are uniformly bounded with respect to k. Thus we may extract subsequences and relabel so that $N^k \to N^* \in \mathbf{P}_n$ and $D^k \to D^* \in \mathbf{Q}_m$ uniformly over I as $k \to \infty$. We observe that $D^* \geq 0$ on I.

We now show that $N^* = N_f$ and $D^* = D_f$. If $\|R^k\|_I$ were unbounded, then $\|f_k - R^k\|_I$ would be unbounded and for an appropriate subsequence

$$\gamma_k = \frac{\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I}{\|R^k - R(C(f_k),\cdot)\|_I}$$

$$\geq \frac{\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I}{\|f_k - R^k\|_I + \|f_k - R(C(f_k),\cdot)\|_I}$$

$$= \frac{1 - \|f_k - R(C(f_k),\cdot)\|_I \,/\, \|f_k - R^k\|_I}{1 + \|f_k - R(C(f_k),\cdot)\|_I \,/\, \|f_k - R^k\|_I}$$

$$\to 1$$

as $k \to \infty$. This would contradict (3.15). Thus $\|R^k\|_I$ is bounded with respect to k. Now let

$$Y = \{y \in I: |f(y) - R(C(f),y)| = \|f - R(C(f),\cdot)\|_I\} ,$$

and for $y \in Y$, let

$$s(y) = \text{sgn}\,[f(y) - R(C(f),y)]$$

where

$$\operatorname{sgn}(t) = \begin{cases} 1, & t > 0, \\ 0, & t = 0, \\ -1, & t < 0. \end{cases}$$

For $y \in Y$,

$$(3.17)\quad \gamma_k \| R^k - R(C(f_k),\cdot)\|_I = \| f_k - R^k \|_I - \| f_k - R(C(f_k),\cdot)\|_I$$

$$\geq s(y)(f_k(y) - R^k(y)) - s(y)(f_k(y) - R(C(f_k),y))$$

$$+ s(y)(f_k(y) - R(C(f_k),y)) - \| f_k - R(C(f_k),\cdot)\|_I .$$

Since $f_k - R(C(f_k),\cdot) \to f - R(C(f),\cdot)$ uniformly over I and $f \notin \mathbf{R}(n,m)$,

$$s(y)(f_k(y) - R(C(f_k),y)) \to s(y)(f(y) - R(C(f),y)) > 0.$$

Thus for sufficiently large k,

$$s(y)(f_k(y) - R(C(f_k),y)) = | f_k(y) - R(C(f_k),y)|.$$

Inequality (3.17) now implies that for $y \in Y$,

$$\gamma_k \| R^k - R(C(f_k),\cdot)\|_I + \| f_k - R(C(f_k),\cdot)\|_I - | f_k(y) - R(C(f_k),y)|$$

$$\geq s(y)(R(C(f_k),y) - R^k(y))$$

$$= \frac{s(y)(R(C(f_k),y)D^k(y) - N^k(y))}{D^k(y)} .$$

Thus

$$(3.18)\qquad s(y)(R(C(f_k),y)D^k(y) - N^k(y))$$
$$\leq D^k(y)\{\gamma_k \| R^k - R(C(f_k),\cdot)\|_I$$
$$+ \| f_k - R(C(f_k),\cdot)\|_I - |f_k(y) - R(C(f_k),y)|\}.$$

Since $|D^k(y)| \leq 1$, $\gamma_k \to 0$, $\| R^k - R(C(f_k),\circ)\|_I$ are bounded with respect to k, and

$$\| f_k - R(C(f_k),\cdot)\|_I - |f_k(y) - R(C(f_k),y)|$$
$$\to \| f - R(C(f),\cdot)\|_I - |f(y) - R(C(f),y)| = 0,$$

we may pass to a limit in inequality (3.18) to get

$$(3.19)\qquad s(y)(R(C(f),y)D^*(y) - N^*(y)) \leq 0$$

for all $y \in Y$.

Since f is normal, Lemma 3.2.6 implies that $\mathbf{P}_n + R(C(f),\cdot)\mathbf{Q}'_m$ is an (n+m+1)-dimensional Haar subspace of $\mathbf{P}_n + R(C(f),\circ)\mathbf{Q}_m$. Inequality (3.14) now insures that $\mathbf{P}_n + R(C(f),\cdot)\mathbf{Q}_m = \mathbf{P}_n + R(C(f),\cdot)\mathbf{Q}'_m$ and is an (n+m+1)-dimensional Haar subspace of C(I). Applying inequality (3.19) and Lemma 3.3.1, we now have that $N^* = R(C(f),\cdot)D^*$.

Equality (3.16) implies that

$\|N^*\|_I + \|D^*\|_I = \|N_f\|_I + \|D_f\|_I = 1$. Lemma 3.3.2 now implies that $D^* = D_f$ and $N^* = N_f$.

We thus have that $N^k \to N_f$ and $D^k \to D_f$ uniformly on I. In addition, since $R(C(f_k),\cdot) \to R(C(f),\cdot)$ uniformly on I as $k \to \infty$, Theorem 3.2.5 implies that $C(f_k) \to C(f)$. Hence $P(A(f_k),\cdot) \to P(A(f),\cdot)$ and $Q(B(f_k),\cdot) \to Q(B(f),\cdot)$ uniformly on I as $k \to \infty$. Thus $\|P(A(f_k),\cdot)\|_I + \|Q(B(f_k),\cdot)\|_I$ $\to \|P(A(f),\cdot)\|_I + \|Q(B(f),\cdot)\|_I \neq 0$. But $D_{f_k} = Q(B(f_k),\cdot)/(\|P(A(f_k),\cdot)\|_I + \|Q(B(f_k),\cdot)\|_I)$ and $D_f = Q(B(f),\cdot)/(\|P(A(f),\cdot)\|_I + \|Q(B(f),\cdot)\|_I)$. Hence, $D_{f_k} \to D_f$ uniformly on I. Since $D_f > 0$ over I, we may select an $\epsilon > 0$, discard an initial segment of the sequences, and assume that $D^k \geq \epsilon$ and $D_{f_k} \geq \epsilon$ on I for all k.

Let $\ell = n + m + 1$. Since $\mu(C(f_k)) = \ell$ for all k, Theorem 3.1.2 implies that there is an alternation set for $f_k - R(C(f_k),\cdot)$ consisting of $\ell + 1$ points:

$$-1 \leq x_0^k < x_1^k < \ldots < x_\ell^k \leq 1.$$

We may pass to a subsequence and relabel so that $x_i^k \to x_i$ as $k \to \infty$, $i = 0,\ldots,\ell$. Since $f \notin \mathbf{R}(n,m)$ we may apply a similar argument to that on page 25 to see that

$$-1 \leq x_0 < x_1 < \ldots < x_\ell \leq 1.$$

For $-1 \leq t_0 < \dots < t_\ell \leq 1$, let $K(t_0,\dots,t_\ell)$ $= \sup\{\|p\|_I$: p is a polynomial of degree less than or equal to ℓ and $(-1)^i p(t_i) \geq -1$, $i = 0,\dots,\ell\}$. By Lemma 1.3.5,

$$K(x_0^k,\dots,x_\ell^k) \to K(x_0,\dots,x_\ell)$$

as $k \to \infty$, and, hence, $K(x_0^k,\dots,x_\ell^k)$ are bounded by some constant $A > 0$ with respect to k.

For fixed k,

$$(-1)^{i+\nu}[f_k(x_i^k) - R(C(f_k),x_i^k)] = \|f_k - R(C(f_k),\cdot)\|_I$$

$i = 0,\dots,\ell$, where ν is 0 or 1. Also

$$(-1)^{i+\nu}[f_k(x_i^k) - R^k(x_i^k)] \leq \|f_k - R^k\|_I ,$$

$i = 0,\dots,\ell$. Subtracting we get

$$(-1)^{i+\nu}(R^k(x_i^k) - R(C(f_k),x_i^k)$$

$$\geq - (\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I.$$

Since $R^k - R(C(f_k),\cdot) = (N^k D_{f_k} - D^k N_{f_k})/D^k D_{f_k}$ and $0 < D^k \leq 1$ and $0 < D_{f_k} \leq 1$ on I,

$$(-1)^{i+\nu}(N^k(x_i^k)D_{f_k}(x_i^k) - D^k(x_i^k)N_{f_k}(x_i^k))$$

$$\geq -D^k(x_i^k)D_{f_k}(x_i^k)(\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I)$$

$$\geq -(\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I).$$

The definition of $K(x_0^k,\ldots,x_\ell^k)$ now implies that

$$\|N^kD_{f_k} - D^kN_{f_k}\|I \tag{3.20}$$

$$\leq K(x_0^k,\ldots,x_\ell^k)(\|f_k-R^k\|_I - \|f_k-R(C(f_k),\cdot)\|_I)$$

$$\leq A(\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I).$$

Now for $x \in I$,

$$|N^k(x)D_{f_k}(x) - D^k(x)N_{f_k}(x)|$$

$$= D^k(x)D_{f_k}(x)\ |R^k(x) - R(C(f_k),x)|$$

$$\geq \epsilon^2\ \|R^k - R(C(f_k),\cdot)\|_I\ .$$

Hence,

$$\|N^kD_{f_k} - D^kN_{f_k}\|_I \geq \epsilon^2\ \|R^k - R(C(f_k),\cdot)\|_I\ . \tag{3.21}$$

Combining (3.20) and (3.21) produces the inequality

$$\frac{\|f_k - R^k\|_I - \|f_k - R(C(f_k),\cdot)\|_I}{\|R^k - R(C(f_k),\cdot)\|_I} \geq \frac{\epsilon^2}{A} .$$

This contradicts (3.15), and the proof of Theorem 3.3.3 is finally complete.

We complete Chapter III by establishing a uniform rational Freud's theorem.

<u>Theorem 3.3.4</u>. Let Γ be a compact subset of $C(I)$ where $\Gamma \cap \mathbf{R}(n,m) = \emptyset$ and f is normal for all $f \in \Gamma$. Then there is a constant $\lambda_\Gamma > 0$ such that

$$\|R(C(g),\cdot) - R(C(f),\cdot)\|_I \leq \lambda_\Gamma \|g - f\|_I$$

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

Proof. By Theorem 3.3.3, there is a constant $\gamma_\Gamma > 0$ such that

$$\|f - R\|_I \geq \|f - R(C(f),\cdot)\|_I + \gamma_\Gamma \|R - R(C(f),\cdot)\|_I$$

for all $f \in \Gamma$ and all $R \in \mathbf{R}(n,m)$. Let $f \in \Gamma$ and $g \in C(I)$. Then

$$\gamma_\Gamma \| R(C(g),\cdot)-R(C(f),\cdot)\|_I \leq \| f-R(C(g),\cdot)\|_I - \| f-R(C(f),\cdot)\|_I$$

$$\leq \| f - g\|_I + \| g - R(C(g),\cdot)\|_I$$

$$- \| f - R(C(f),\cdot)\|_I$$

$$\leq \| f - g\|_I + \| g - R(C(f),\cdot)\|_I$$

$$- \| f - R(C(f),\cdot)\|_I$$

$$\leq 2 \| g - f\|_I \ .$$

Thus, choosing $\lambda_\Gamma = 2/\gamma_\Gamma$, Theorem 3.3.4 is proven.

CHAPTER IV

RATIONAL PRODUCT APPROXIMATION

4.1 Introduction.

In Chapter IV, we discuss an extension of product approximation to the setting of multivariate rational approximation. Let $D = I \times J = [-1,1] \times [-1,1]$, and fix the non-negative integers $n, m, r_i, i = 0,\dots,n$, and $s_j, j = 0,\dots,m$. The class $\mathbf{R}_D$ of approximating functions consists of all rational functions

$$R(x,y) = \frac{P(x,y)}{Q(x,y)} = \frac{\sum_{i=0}^{n} \sum_{k=0}^{r_i} a_{ik} y^k x^i}{\sum_{j=0}^{m} \sum_{\ell=0}^{s_j} b_{j\ell} y^\ell x^j}$$

where $Q(x,y) > 0$ on D.

It was indicated in Chapter II that best uniform approximations of $F \in C(D)$ from a finite dimensional linear subspace of $C(D)$ may fail to be unique. In the rational setting, best uniform approximations of $F \in C(D)$ from $\mathbf{R}_D$ may fail to exist. This is illustrated by the following example due to Henry and Weinstein [17].

Example 4.1. Let

$$(4.1) \qquad F(x,y) = \begin{cases} \dfrac{(x+1)^2 + (y+1)^2}{x + y + 2}, & -1 < y \leq 1, \\ x + 1, & y = -1, \end{cases}$$

and let $n = 2$, $m = 1$, $r_i = 2$, $i = 0,1,2$, and $s_j = 1$, $j = 0,1$. For each $\epsilon > 0$, define $R_\epsilon \in \mathbf{R}_D$ by

$$R_\epsilon(x,y) = \frac{(x+1)^2 + (y+1)^2}{x + y + 2 + \epsilon}.$$

Then $R_\epsilon(x,y) \to F(x,y)$ uniformly over D as $\epsilon \to 0$. However, $F \notin \mathbf{R}_D$. Thus F does not have a best uniform approximation from $\mathbf{R}_D$.

The loss of existence or uniqueness of a best approximation and the accompanying computational difficulties have been a deterrent in the development of multivariate rational Tchebycheff approximation. To overcome these problems, Brown and Henry [3] introduced the concept of rational product approximation. The following is a development of rational product approximation which is somewhat different from that of [3]. The motivation for the following treatment comes from the more general concept of composite approximation defined by Brown and Henry [4].

Let $F \in C(D)$, and for each $y \in J$, define $F_y \in C(I)$ by $F_y(x) = F(x,y)$. For $y \in J$, let

$$(4.2) \qquad R(C(F_y),x) = \frac{\sum_{i=0}^{n} a_i^F(y)x^i}{1 + \sum_{j=1}^{m} b_j^F(y)x^j}$$

be the best uniform approximation of F_y from $\mathbb{R}(n,m)$ over I where $C(F_y) = (A(F_y); B(F_y)) = (a_0^F(y),\ldots,a_n^F(y);b_1^F(y),\ldots,b_m^F(y)) \in \mathbf{C}$. The rational function space $\mathbb{R}(n,m)$ is given by (3.1), and the parameter space $\mathbf{C}$ is defined on page 70. If the coefficient functions $a_i^F(y)$, $i = 0,\ldots,n$, and $b_j^F(y)$, $j = 1,\ldots,m$, are continuous on J, then we let

$$(4.3) \qquad T_{r_i}(a_i^F,y) = \sum_{k=0}^{r_i} a_{ik}^F y^k$$

be the best approximation of $a_i^F(y)$ by polynomials of degree less than or equal to r_i, $i = 0,\ldots,n$, in the sense of the uniform norm $\|\cdot\|_J$ and

$$(4.4) \qquad T_{s_j}(b_j^F,y) = \sum_{j=0}^{s_j} b_{j\ell}^F y^\ell$$

be the best uniform approximation of $b_j^F(y)$ over J by polynomials of degree less than or equal to s_j, $j = 1,\ldots,m$. If

(4.5) $$1 + \sum_{j=1}^{m} T_{s_j}(a_j^F, y)\, x^j > 0$$

over D, we define the rational product approximation to be the function $\mathcal{R}F$ given by

(4.6) $$(\mathcal{R}F)(x,y) = \frac{\sum_{i=0}^{n} T_{r_i}(a_i^F, y) x^i}{1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)\, x^j}$$

$$= \frac{\sum_{i=0}^{n} \sum_{k=0}^{r_i} a_{ik}^F\, y^k x^i}{1 + \sum_{j=1}^{m} \sum_{\ell=0}^{s_j} b_{j\ell}^F\, y^\ell x^j} .$$

The present definition of rational product approximation differs from that of Brown and Henry [3] only in choice parameter spaces for $\mathbb{R}(n,m)$. The parameter space used in [3] is an (n+m+2)-dimensional set. As a result, the use of the parameter space $\mathbf{C}$ requires one less best approximation at the second stage (4.3) and (4.4).

The definedness of $\mathcal{R}F$, where $F \in C(D)$, depends both on the continuity of the coefficient vector $C(F_y)$ and on the inequality (4.5) being satisfied for all $(x,y) \in D$. In Section 4.2, we discuss the definedness of $\mathcal{R}F$ and give a survey of variants and extensions of rational product approximation. Section 4.3 consists of a continuity theorem for the operator $\mathcal{R}$ and conditions on F are determined to insure that $\mathcal{R}$ satisfies a local point-Lipschitz condition at F.

4.2 Definedness of $\mathcal{R}F$.

The definedness considerations of Section 4.2 are adapted from composite approximation [4] to the rational setting. We first consider the continuity of the coefficient vector $C(F_y)$. The papers [3] and [17] give examples showing that $C(F_y)$ need not be continuous on J. The following example appears in [17].

Example 4.2. Let $n = 2$, $m = 1$, and

$$F(x,y) = \begin{cases} \dfrac{y + 1 + \frac{1}{2}x}{1 + \frac{1}{2}x}, & -1 \le y < 0, \\ 1, & 0 \le y \le 1. \end{cases}$$

Since $F_y \in \mathbf{R}(2,1)$ for all $y \in J$, $R(C(F_y), \cdot) = F_y$ and

$$C(F_y) = \begin{cases} (y+1,1/2;1/2), & -1 \leq y < 0, \\ (1,0;0), & 0 \leq y \leq 1. \end{cases}$$

We see that the second and third components of $C(F_y)$ are discontinuous at $y = 0$.

We restate the definition of normality from Section 3.1. For $F \in C(D)$ and $y \in J$, we say that F_y is <u>normal</u> (relative to I and $\mathbf{R}(n,m)$) if any of the following equivalent conditions hold:

(i) $R(C(F_y),\cdot) \in \mathbf{R}^*$,

(ii) $C(F_y) \in \mathbf{C}^*$,

where $\mathbf{R}^*$ is defined in (3.4) and $\mathbf{C}^*$ is defined in (3.5). We note that in Example 4.2, F_y is normal for $-1 \leq y < 0$ but is not normal for $0 \leq y \leq 1$. In terms of the varisolvent degree (3.3), this can be restated as

$$\mu(C(F_y)) = \begin{cases} 4, & -1 \leq y < 0, \\ 3, & 0 \leq y \leq 1. \end{cases}$$

We note that the discontinuity of $C(F_y)$ occurs at the point where the varisolvent degree of $R(C(F_y),\cdot)$ changes. A condition which insures the continuity of $C(F_y)$ is that $\mu(C(F_y))$ be constant over J.

Theorem 4.2.1. If $F \in C(D)$ and F_y is normal for all $y \in J$, then $C(F_y)$ is continuous over J.

Proof. The uniform continuity of F over D implies that

$$(4.7) \qquad y \to F_y$$

is a continuous map of J into C(I). Theorem 3.1.4 and the normality of each F_y insures that

$$(4.8) \qquad F_y \to R(C(F_y), \cdot)$$

is a continuous mapping of $\{F_y : y \in J\}$ into $\mathbf{R}^*$. Theorem 3.1.4 asserts that

$$(4.9) \qquad R(C(F_y), \cdot) \to C(F_y)$$

maps $\{R(C(F_y), \cdot) : y \in J\}$ continuously into $\mathbf{C}^*$. Composition of the maps (4.7), (4.8), and (4.9) yields the continuity of $C(F_y)$ over J.

Remark. If $\mu(C(F_y)) = d$ for all $y \in J$ and $d < 1+n+m$, then $C(F_y)$ is continuous over J. In this case, $R(C(F_y), \cdot) \in \mathbf{R}(d-m-1, d-n-1)$ for all $y \in J$, and F_y is normal with respect to the rational function space $\mathbf{R}(d-m-1, d-n-1)$ for all $y \in J$. We may now apply Theorem

4.2.1 to see that $C(F_y)$ is continuous over J whenever $\mu(C(F_y))$ is constant over J.

We next consider the inequality (4.5). We show that if $F \in C(D)$ satisfies the conditions of Theorem 4.2.1, then non-negative integers r_i, $i = 0,\ldots,n$, and s_j, $j = 1,\ldots,m$, can be chosen sufficiently large so that (4.5) is satisfied for all $(x,y) \in D$.

<u>Theorem 4.2.2</u>. Suppose $F \in C(D)$ and F_y is normal for all $y \in J$. Then there are numbers s_j^0, $j = 1,\ldots,m$, such that if $s_j \geq s_j^0$, $j = 1,\ldots,m$, then

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)x^j > 0$$

for all $(x,y) \in D$.

Proof. By Theorem 4.2.1, $b_j(y)$, $j = 1,\ldots,m$, are continuous over J, and as a result,

$$1 + \sum_{j=1}^{m} b_j^F(y)x^j$$

is continuous and positive valued over D. Let

$$\delta = \inf_{(x,y)\in D} \left(1 + \sum_{j=1}^{m} b_j^F(y)x^j\right) > 0.$$

By the Weierstrass approximation theorem [5, p. 66], there are numbers s_j^0, $j = 1,\ldots,m$, such that if $s_j \geq s_j^0$, $j = 1,\ldots,m$, then

$$\| b_j^F - T_{s_j}(b_j^F,\cdot)\|_J < \frac{\delta}{2m} .$$

Now if $s_j \geq s_j^0$, $j = 1,\ldots,m$, and $(x,y) \in D$, then

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^F,y)x^j = 1 + \sum_{j=1}^{m} b_j^F(y)x^j$$

$$- \sum_{j=1}^{m} (b_j^F(y) - T_{s_j}(b_j^F,y))x^j$$

$$\geq 1 + \sum_{j=1}^{m} b_j^F(y)x^j - \sum_{j=1}^{m} \| b_j^F - T_{s_j}(b_j^F,\cdot)\|_J$$

$$> \frac{\delta}{2}$$

$$> 0.$$

The proof of Theorem 4.2.2 is complete.

Thus if $F \in C(D)$ satisfies the normality conditions of Theorem 4.2.1, then $\mathcal{R}F$ can be defined if s_j, $j=1,\ldots,m$, are sufficiently large. Hereafter, we regard the domain $\mathcal{D}$ of $\mathcal{R}$ to be the set of all $F \in C(D)$ such that F_y is normal for all $y \in J$ and (4.5) holds for all $(x,y) \in D$.

We conclude this section by presenting a short survey of alternative considerations for the definedness of $\mathcal{R}$. The paper of Henry and Weinstein [17] overcomes a certain type of discontinuity that may occur in $C(F_y)$. Suppose that $C(F_y)$ is discontinuous only at $y^* \in J$ and we can write

$$C(F_y) = (A_0(y);B_0(y)), \; -1 \leq y < y^*$$

and

$$C(F_y) = (A_1(y);B_1(y)), \; y^* < y \leq 1,$$

where A_0, B_0, A_1, and B_1 are continuous over J and $Q(B_0(y),x) > 0$ and $Q(B_1(y),x) > 0$ over I. Suppose further that

$$\lim_{y \to y^*-} R(C(F_y),x) = \lim_{y \to y^*+} R(C(F_y),x)$$

for $-1 \leq x \leq 1$. Then we can write

$$R(C(F_y),x) = \frac{N_y(x)}{D_y(x)} \qquad (4.10)$$

where $D_y(x) = Q(B_0(y),x)\, Q(B_1(y),x)$ and

$$N_y(x) = \begin{cases} P(A_0(y),x)Q(B_1(y),x), & -1 \le y \le y^*, \\ P(A_1(y),x)Q(B_0(y),x), & y^* < y \le 1 . \end{cases}$$

The coefficient functions in the representation (4.10) are continuous and may be best approximated. The number of coefficient functions to be approximated is thus raised by 2m. J. Henry [12] has developed a modification of the technique of M. Henry and Weinstein [17] and has written a program to compute rational product approximations using his modification.

A further variation of the definition of rational product approximation involves approximating the coefficient functions $a_i^F(y)$, $i = 0,\ldots,n$, and $b_j^F(y)$, $j = 1,\ldots,m$, by rational functions in y. This is done in [3]. The paper [4] discusses composite approximation in which the approximating spaces are varisolvent families of functions (see [27] and [29]).

4.3. Continuity of $\mathcal{R}$.

In this section, we prove that $\mathcal{R}$ maps its domain D continuously into $\mathbb{R}_D$ where both D and $\mathbb{R}_D$ carry the uniform norm topologies. The continuity theorem will follow from a series of lemmas in which we prove that D is an open subset of C(D).

Lemma 4.3.1. Suppose $F \in C(D)$ and F_y is normal for all $y \in J$. Then there is a $\delta = \delta(F) > 0$ such that if $G \in C(D)$ and $\|G - F\|_D < \delta$, then G_y is normal for all $y \in J$.

Proof. Assume this is not the case. Then there is a sequence $\{G^k\}$ in C(D) where $\|G^k - F\|_D \to 0$ as $k \to \infty$ and for each k there is a $y_k \in J$ for which $G^k_{y_k}$ is not normal. Since J is compact, we may assume $y_k \to y^* \in J$. The inequality

$$\|G^k_{y_k} - F_{y^*}\|_I \leq \|G^k - F\|_D + \|F_{y_k} - F_{y^*}\|_I$$

insures that $\|G^k_{y_k} - F_{y^*}\|_I \to 0$. Theorem 3.1.4 and the normality of F_{y^*} guarantee that $R(C(G^k_{y_k}),\cdot) \to R(C(F_{y^*}),\cdot)$ uniformly on I. Since $R(C(F_{y^*}),\cdot) \in \mathbf{R}^*$ and $\mathbf{R}^*$ is an open subset of $\mathbf{R}(n,m)$, a tail of the sequence $\{R(C(G^k_{y_k}),\cdot)\}$ is in $\mathbf{R}^*$. This contradicts the assumption that each $G^k_{y_k}$ is not normal. Thus Lemma 4.3.1 is proven.

For $F \in C(D)$ and $y \in J$, denote

$$(4.11) \qquad \beta_F(y) = \|F_y - R(C(F_y),\cdot)\|_I \, .$$

If F_y is normal for all $y \in J$, then Theorem 3.1.4 implies that $\beta_F(y)$ is continuous over J.

<u>Lemma 4.3.2</u>. Suppose $F \in C(D)$ and F_y is normal for all $y \in J$. Given $\epsilon > 0$ there is a $\delta = \delta(F,\epsilon) > 0$ such that if $C(y) = (A(y);B(y)) \in \mathbf{C}$ for each $y \in J$ and

$$\|F_y - R(C(y),\cdot)\|_I \leq \beta_F(y) + \delta$$

for all $y \in J$, then $\sigma(C(F_y), C(y)) < \epsilon$ for all $y \in J$.

Proof. Assume otherwise. Then there is a sequence $\{C^k(y)\}$, $C^k(y) = (A^k(y);B^k(y)) = (a_0^k(y),\ldots,a_n^k(y);$ $b_1^k(y),\ldots,b_m^k(y)) \in \mathbf{C}$ for all $y \in J$, such that

$$(4.12) \qquad \|F_y - R(C^k(y),\cdot)\|_I \leq \beta_F(y) + 1/k$$

for all $y \in J$, and there is a $y_k \in J$ for which

$$(4.13) \qquad \sigma(C(F_{y_k}), C^k(y_k)) \geq \epsilon .$$

The compactness of J allows us to assume $y_k \to y^* \in J$.

Let $P^k = P(A^k(y_k),\cdot)$, $Q^k = Q(B^k(y_k),\cdot)$, $P^* = P(A(F_{y^*}),\cdot)$, $Q^* = Q(B(F_{y^*}),\cdot)$, $w_k = \|P^k\|_I + \|Q^k\|_I$, and $w^* = \|P^*\|_I + \|Q^*\|_I$. Define $N^k = P^k/w_k$, $D^k = Q^k/w_k$, $N^* = P^*/w^*$, and $D^* = Q^*/w^*$. Since $\|D^k\|_I \leq 1$, by appropriate relabeling we may assume $D^k \to \overline{D} \in \mathbf{Q}_m$. Similarly, $N^k \to \overline{N} \in \mathbf{R}_n$, and $\|\overline{N}\|_I + \|\overline{D}\|_I = 1$. Let $M = \|F\|_D + \|\beta_F\|_J + 1$. From (4.12) we have

$$\|N^k/D^k\|_I = \|P^k/Q^k\|_I \leq M.$$

So, $|N^k(x)| \leq M|D^k(x)|$ for each $x \in I$. Thus

$$|\overline{N}(x)| \leq M|\overline{D}(x)| \tag{4.14}$$

for all $x \in I$. This inequality and $\|\overline{N}\|_I + \|\overline{D}\|_I = 1$ imply $\overline{D} \not\equiv 0$. Thus, using (4.14), we may perform appropriate cancellations to find $N'/D' \in \mathbf{R}(n,m)$ such that

$$\frac{\overline{N}(x)}{\overline{D}(x)} = \frac{N'(x)}{D'(x)} \tag{4.15}$$

where $\overline{D}(x) \neq 0$. Thus at all but the finitely many points where $\overline{D}$ vanishes,

$$\left| F_{y*}(x) - \frac{N'(x)}{D'(x)} \right| = \left| F_{y*}(x) - \frac{\overline{N}(x)}{\overline{D}(x)} \right|$$

$$\leq |F_{y*}(x) - F_{y_k}(x)| + |F_{y_k}(x) - R(C^k(y_k), \cdot)|$$

$$+ \left| \frac{N^k(x)}{D^k(x)} - \frac{\overline{N}(x)}{\overline{D}(x)} \right|$$

$$\leq \| F_{y*} - F_{y_k} \|_I + \| F_{y_k} - R(C^k(y_k), \cdot) \|_I$$

$$+ \left| \frac{N^k(x)}{D^k(x)} - \frac{\overline{N}(x)}{\overline{D}(x)} \right|$$

$$\leq \| F_{y*} - F_{y_k} \|_I + \beta_F(y_k) + 1/k$$

$$+ \left| \frac{N^k(x)}{D^k(x)} - \frac{\overline{N}(x)}{\overline{D}(x)} \right| .$$

Letting $k \to \infty$, we get

$$\left| F_{y*}(x) - \frac{N'(x)}{D'(x)} \right| \leq \beta_F(y^*).$$

By the continuity of $F_{y*} - \frac{N'}{D'}$ over I,

$$\left\| F_{y*} - \frac{N'}{D'} \right\|_I \leq \beta_F(y^*).$$

By the uniqueness of best rational approximations

$$\frac{N'}{D'} = R(C(F_{y*}),\cdot) = R^* = \frac{P^*}{Q^*} = \frac{N^*}{D^*} .$$

These equations, (4.14) and (4.15) imply that $\overline{N} = R^*\overline{D}$ where $\overline{D}(x) \geq 0$ on I and $\|\overline{N}\|_I + \|\overline{D}\|_I = \|N^*\|_I + \|D^*\|_I$. Since F_{y*} is normal, Lemma 2.3.2 implies $\overline{N} = N^*$ and $\overline{D} = D^*$. But $Q^*(0) = Q^k(0) = 1$, $Q^* = w^*D^*$, and $Q^k = w_k D^k$ imply that $D^*(0) = 1/w^*$ and $D^k(0) = 1/w_k$. Since $D^k(0) \to D^*(0)$, $w_k \to w^*$. Thus $P^k \to P^*$ and $Q^k \to Q^*$ uniformly on I. As a consequence, $\sigma(C(F_{y*}), C^k(y_k)) \to 0$. Now (4.13) and the continuity of $C(F_y)$ with respect to y imply $\sigma(C(F_{y*}), C(F_{y*})) \geq \epsilon$ which is false. Hence, Lemma 4.2.2 is proven.

<u>Lemma 4.3.3</u>. Suppose $F \in C(D)$ and F_y is normal for all $y \in J$. Given $\epsilon > 0$, there is a $\delta = \delta(F,\epsilon) > 0$ such that whenever $G \in C(D)$ and $\|G - F\|_I < \delta$, $\sigma(C(G_y), C(F_y)) < \epsilon$ for all $y \in J$.

Proof. Let $\epsilon > 0$ be given. From Lemma 4.3.2, select $\delta > 0$ such that $G \in C(D)$ and $\|F_y - R(C(G_y),\cdot)\|_I \leq \beta_F(y) + \delta$ for all $y \in J$ implies $\sigma(C(G_y), C(F_y)) < \epsilon$ for all $y \in J$.

Suppose $G \in C(D)$ and $\|G - F\|_I < \delta/2$. Then for any

$y \in J$,

$$\|F_y - R(C(G_y),\cdot)\|_I \le \|F_y - G_y\|_I + \|G_y - R(C(G_y),\cdot)\|_I$$
$$\le \|F_y - G_y\|_I + \|G_y - R(C(F_y),\cdot)\|_I$$
$$\le 2\|F_y - G_y\|_I + \|F_y - R(C(F_y),\cdot)\|_I$$
$$< \beta_F(y) + \delta .$$

Thus $\sigma(C(G_y),\ C(F_y)) < \epsilon$ for all $y \in J$.

Lemma 4.3.4. Suppose $F \in D$ and let

$$\tau = \inf_{(x,y)\ D} (1 + \sum_{j=1}^{m} T_{s_j}(b_j^F,y)x^j) > 0.$$

Then there is a $\delta = \delta(F) > 0$ such that if $G \in C(D)$ and $\|G - F\|_D < \delta$, then

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^G,y)x^j > \tau/2$$

for all $(x,y) \in D$.

Proof. By the continuity of T_{s_j} at b_j^F (see Theorem 1.2.6), there is a $\gamma > 0$ such that $\|b_j^G - b_j^F\|_J < \gamma$ implies that $\|T_{s_j}(b_j^G,\cdot) - T_{s_j}(b_j^F,\cdot)\|_J < \tau/2m$. Pick $\delta > 0$, via

Lemma 4.3.3, so that $G \in C(D)$ and $\|G-F\|_D < \delta$ implies that $\|b_j^G - b_j^F\|_J < \gamma$, $j = 1,\ldots,m$. For such G,

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^G,y)x^j \geq 1 + \sum_{j=1}^{m} T_{s_j}(b_j^F,y)x^j$$

$$- \sum_{j=1}^{m} \|T_{s_j}(b_j^G,\cdot) - T_{s_j}(b_j^F,\cdot)\|_I$$

$$\geq \tau - \tau/2$$

$$= \tau/2.$$

Remark. As we have taken the domain $\mathcal{D}$ of $\mathcal{R}$ to be the collection of all $F \in C(D)$ such that F_y is normal for all $y \in J$ and (4.5) is satisfied for all $(x,y) \in D$, Lemmas 4.3.1 and 4.3.4 insure that the domain of $\mathcal{R}$ is an open subset of $C(D)$.

We are now in a position to prove that $\mathcal{R}$ maps $\mathcal{D}$ continuously into $\mathbb{R}_D$.

<u>Theorem 4.3.5</u>. The rational product approximation operator $\mathcal{R}$ is continuous over $\mathcal{D}$.

Proof. Let $F \in \mathcal{D}$. In view of the above remark, $\mathcal{R}$ is defined in a neighborhood of F. Let $\{G^k\}$ be a sequence in C(D) where $\|G^k - F\|_D \to 0$. By Lemma 4.3.3, $C(G_y^k) \to C(F_y)$ uniformly over J. By Theorem 1.2.7, $T_{r_i}(a_i^{G^k}, \cdot) \to T_{r_i}(a_i^F, \cdot)$ and $T_{s_j}(b_j^{G^k}, \cdot) \to T_{s_j}(b_j^F, \cdot)$ uniformly over J. Hence,

$$\sum_{i=0}^{n} T_{r_i}(a_j^{G^k}, y)x^i \to \sum_{i=0}^{n} T_{r_i}(a_i^F, y)x^i$$

and

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^{G^k}, y)x^j \to 1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)x^j$$

uniformly over D. Moreover, by Lemma 4.3.4, there is a $\tau > 0$ such that

$$1 + \sum_{j=1}^{m} T_{s_j}(b_j^{G^k}, y)x^j \geq \tau$$

over D for all sufficiently large k. Hence, $\mathcal{R}G^k \to \mathcal{R}F$ uniformly over D. Thus $\mathcal{R}$ is continuous at F.

We conclude this chapter by establishing an analog of Theorem 2.4.1 for rational product approximation. We note that $\mathcal{R}$ satisfies a local point-Lipschitz condition while $\mathcal{P}$ satisfies a global point-Lipschitz condition.

Theorem 4.3.6. Suppose $F \in \mathcal{D}$ and $F_y \notin \mathbb{R}(n,m)$ for all $y \in J$. Then there are constants $\lambda > 0$ and $\delta > 0$ such that if $G \in C(D)$ and $\|G - F\|_D < \delta$, then $\|\mathcal{R}G - \mathcal{R}F\|_D \leq \lambda \|G - F\|_D$.

Proof. We first note that $\mathcal{R}$ is defined in a neighborhood of F. By Lemmas 4.3.1 and 4.3.4, we may choose $\delta_1 > 0$ such that if $G \in C(D)$ and $\|G - F\|_D < \delta_1$, then $G \in \mathcal{D}$ and

$$1 + \sum_{i=1}^{m} T_{s_j}(b_j^G, y)x^j \geq \tau/2$$

for all $(x,y) \in D$ where

$$\tau = \inf_{(x,y)\ D} \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)x^j\right) > 0.$$

Thus if $G \in C(D)$ and $\|G - F\|_D < \delta_1$, then for $(x,y) \in D$

$$|(\mathcal{R}G)(x,y) - (\mathcal{R}F)(x,y)|$$

$$= \Bigg| \left(\sum_{i=0}^{n} T_{s_i}(a_i^G, y)x^i \right) \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)x^j \right) - \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^G, y)x^j \right) \left(\sum_{i=0}^{n} T_{r_i}(a_i^F, y)x^i \right) \Bigg|$$

$$\div \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^G, y)x^j \right) \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^F, y)x^j \right)$$

$$\le 2/\tau^2 \left| \left(\sum_{i=0}^{n} [T_{r_i}(a_i^G,y) - T_{r_i}(a_i^F,y)]x^i \right) \right.$$

$$\cdot \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j^F,y)x^j \right)$$

$$+ \left(\sum_{j=1}^{m} [T_{s_j}(b_j^F,y) - T_{s_j}(b_j^G,y)]x^j \right)$$

$$\left. \cdot \left(\sum_{i=0}^{n} T_{r_i}(a_i^F,y)x^i \right) \right|$$

$$\le 2/\tau^2 \left\{ \left(1 + \sum_{j=1}^{m} \|T_{s_j}(b_j^F,\cdot)\|_J \right) \right.$$

$$\cdot \left(\sum_{i=0}^{n} \|T_{r_i}(a_i^G,\cdot) - T_{r_i}(a_i^F,\cdot)\|_J \right)$$

$$+ \left(\sum_{i=0}^{n} \|T_{r_i}(a_i^F,\cdot)\|_J \right)$$

$$\left. \cdot \left(\sum_{j=1}^{m} \|T_{s_j}(b_j^G,\cdot) - T_{s_j}(b_j^F,\cdot)\|_J \right) \right\}$$

We now apply Theorem 1.2.7 to the operators T_{r_i} at a_i^F, $i = 0,\ldots,n$, and T_{s_j} at b_j^F, $j = 1,\ldots,m$, to get σ_i, $i = 0,\ldots,n$, and λ_j, $j = 1,\ldots,m$, which do not depend on G, such that

(4.16) $\|\mathcal{R}G - \mathcal{R}F\|_D$

$$\le 2/\tau^2 \left\{ K_1 \sum_{i=0}^{n} \sigma_i \|a_i^G - a_i^F\|_J + K_2 \sum_{j=1}^{m} \lambda_j \|b_j^G - b_j^F\|_J \right\}$$

where $K_1 = 1 + \sum_{j=1}^{m} \|T_{s_j}(b_j^F,\cdot)\|_J$ and
$K_2 = \sum_{i=0}^{n} \|T_{r_i}(a_i^F,\cdot)\|_J$.

Now we note that $\{F_y : y \in J\}$ is a compact set of normal functions in $C(I)$ disjoint from $\mathbb{R}(n,m)$. Thus by Theorem 3.3.4, there is a constant $\lambda^* > 0$, independent of G, such that

$$\begin{aligned}\|R(C(G_y),\cdot) - R(C(F_y),\cdot)\|_I &\\ &\le \lambda^* \|G_y - F_y\|_I \\ &\le \lambda^* \|G - F\|_D\end{aligned}$$

for all $y \in J$.

Now $\{R(C(F_y),\cdot) : y \in J\}$ is a compact subset of $\mathbb{R}^*$. By Theorems 3.2.7 and 3.2.8, there is a constant $\rho > 0$, which does not depend on G, such that

$$(4.17)\quad \begin{aligned}&\|P(A(G_y),\cdot) - R(C(F_y),\cdot)Q(B(G_y),\cdot)\|_I \\ &\quad \ge \rho\left(\sum_{i=0}^{n} |a_i^G(y)-a_i^F(y)| + \sum_{j=1}^{m} |b_j^G(y)-b_j^F(y)|\right)\end{aligned}$$

for all $y \in J$. Furthermore, by Lemma 4.3.3, numbers $\delta_2 > 0$ and $K_3 > 0$ may be chosen so that if $G \in C(D)$ and $\|G - F\|_D < \delta_2$, then $\|Q(B(G_y),\cdot)\|_I \le K_3$ for all $y \in J$.

We choose $\delta = \min(\delta_1, \delta_2) > 0$. Suppose $G \in C(D)$ and $\|G - F\|_D < \delta$. Then

(4.18) $$\|P(A(G_y),\cdot) - R(C(F_y),\cdot)Q(B(G_y),\cdot)\|_I$$

$$\leq \|Q(B(G_y),\cdot)\|_I \, \|R(C(G_y),\cdot) - R(C(F_y),\cdot)\|_I$$

$$\leq K_3\lambda^* \|G - F\|_D .$$

Applying (4.17) and (4.18), we obtain

(4.19) $$\|a_i^G - a_i^F\|_J \leq \frac{K_3\lambda^*}{\rho} \|G - F\|_D ,$$

$i = 0,\ldots,n$, and

(4.20) $$\|b_j^G - b_j^F\|_J \leq \frac{K_3\lambda^*}{\rho} \|G - F\|_D ,$$

$j = 1,\ldots,m$. Finally, applying (4.16), (4.19), and (4.20), we obtain

$$\|\mathcal{R}G - \mathcal{R}F\|_D \leq \lambda \, \|G - F\|_D ,$$

whence

$$\lambda = \frac{K_3\lambda^*}{\rho} \left[K_1 \sum_{i=0}^{n} \sigma_i + K_2 \sum_{j=1}^{m} \lambda_i \right].$$

Thus a continuity theory for the rational product approximation operator has been established. We remark that such results are not possible for best rational Tchebycheff approximation in the multivariate setting.

CHAPTER V

DISCRETE RATIONAL PRODUCT APPROXIMATION AND COMPUTATION

5.1. Definitions and notations.

The algorithms to compute rational product approximations of $F \in C(D)$, $D = I \times J = [-1,1] \times [-1,1]$, presented in the papers [4,12,13] involve a discretization of one or both of the intervals I and J. A theory for rational product approximation over discrete arrays of points is thus needed. The case in which only the interval J is discretized is treated by Brown and Henry [4]. We consider the more general concept of rational product approximation of $F \in C(D)$ over $X \times Y$ where X and Y are closed subsets of I and J, respectively.

We first recall the definition of rational product approximation over D. The discretization concern of Chapter V does not involve purturbations of the function $F \in C(D)$. Thus we suppress the notations indicating dependence on F. Let the non-negative integers n, m, r_i, $i = 0,\dots,n$, and s_j, $j = 1,\dots,m$, be given. Associated with n and m is the rational function space

$$\mathbf{R}(n,m) = \{R = P/Q : P \in \mathbf{P}_n,\ Q \in \mathbf{Q}_m,\ Q > 0 \text{ on } I\}$$

where $\mathbf{P}_n$ and $\mathbf{Q}_m$ denote the spaces of polynomials of degree less than or equal to n and m, respectively.

Let $F \in C(D)$ and for each $y \in J$, $F_y \in C(I)$ is given by $F_y(x) = F(x,y)$. In the remainder of this chapter, it is assumed that $F \in \mathcal{D}$. That is to say, F_y is normal with respect to I and $\mathbf{R}(n,m)$ for all $y \in J$ and inequality (4.5) holds for all $(x,y) \in D$. Let

(5.1) $$R(C(y),x) = \frac{\sum_{i=0}^{n} a_i(y)x^i}{1 + \sum_{j=1}^{m} b_j(y)x^j},$$

$C(y) = (A(y);B(y)) = (a_0(y),\ldots,a_n(y);b_1(y),\ldots,b_m(y)) \in \mathbf{C}$, be the best uniform approximation of F_y over I from $\mathbf{R}(n,m)$. Furthermore, we let

(5.2) $$\Delta(y) = \|F_y - R(C(y),\cdot)\|_I$$

$$= \inf\{\|F_y - R\|_I : R \in \mathbf{R}(n,m)\}.$$

Now let

(5.3) $$T_{r_i}(a_i,y) = \sum_{k=0}^{r_i} a_{ik}y^k x^i$$

be the best uniform approximation of a_i over J by polynomials of degree less than or equal to r_i, $i = 0,\ldots,n$, and let

$$(5.4) \qquad T_{s_j}(b_j,y) = \sum_{\ell=0}^{s_j} b_{j\ell} y^{\ell} x^j$$

be the best uniform approximation of b_j over J by polynomials of degree less than or equal to s_j, $j = 1,\ldots,m$. The corresponding rational product approximation of F over D is

$$(5.5) \qquad \mathcal{R}(x,y) = \frac{\sum_{i=0}^{n} T_{r_i}(a_i,y)x^i}{1 + \sum_{j=1}^{m} T_{s_j}(b_j,y)x^j} \; .$$

Suppose X is a closed subset of I. Associated with X is the approximating space

$$(5.6) \qquad \mathbb{R}_X(n,m) = \{P = R/Q : P \in \mathbb{P}_n,\ Q \in \mathbb{Q}_m,\ Q > 0 \text{ on } X\} .$$

We observe that $\mathbb{R}(n,m) \subseteq \mathbb{R}_X(n,m)$. For $y \in J$, let

$$(5.7) \qquad \Delta_X(y) = \inf\{\|F_y - R\|_X : R \in \mathbb{R}_X(n,m)\}.$$

Suppose that each F_y has a unique best uniform approximation from $\mathbb{R}_X(n,m)$ over X which can be expressed as

$$(5.8) \qquad R(C_X(y),x) = \frac{\sum_{i=0}^{n} a_i^X(y)x^i}{1 + \sum_{j=1}^{m} b_j^X(y)x^j} ,$$

$C_X(y) = (a_0^X(y),\ldots,a_n^X(y);\ b_1^X(y),\ldots,b_m^X(y)) \in E_{n+m+1}$, where the coefficient vector $C_X(y)$ is continuous on J. If Y is a closed subset of J, we let $T_{r_i}^Y(a_i^X,y)$ be the best uniform approximation of a_i over Y by polynomials of degree less than or equal to r_i, $i = 0,\ldots,n$, and let $T_{s_j}^Y(b_j^X,y)$ be the best uniform approximation of b_j^X over Y by polynomials of degree less than or equal to s_j, $j = 1,\ldots,m$. The rational product approximation of F over $X \times Y$ is now defined to be the function

$$(5.9)\qquad \mathcal{R}_{X\times Y}(x,y) = \frac{\sum_{i=0}^{n} T_{r_i}^Y(a_i^X,y)x^i}{1 + \sum_{j=1}^{m} T_{s_j}^Y(b_j^X,y)x^j}$$

provided the denominator in (5.9) does not vanish over $X \times Y$.

The definedness of $\mathcal{R}_{X\times Y}(x,y)$ depends on the existence and uniqueness of $R(C_X(y),\cdot)$ and the ability to write it as in (5.8). In Section 5.2, we show that if X is "sufficiently dense" in I, then $R(C_X(y),\cdot)$ exists, is unique, and is contained in $\mathbf{R}(n,m)$. In this case, $\mathcal{R}_{X\times Y}(x,y)$ will be definable.

In Section 5.2, it is also shown that as X and Y converge to I and J, respectively, in the sense of the density measure (1.10), the corresponding rational product approximation $\mathcal{R}_{X\times Y}(x,y)$ converges uniformly to $\mathcal{R}(x,y)$ over D.

In the remaining sections of Chapter V, we discuss algorithms to compute $\mathcal{R}_{X\times Y}$ where X and Y are finite subsets of I and J, respectively. In this case, $\mathcal{R}_{X\times Y}(x,y)$ is called discrete rational product approximation. In particular, an algorithm is presented in Section 5.4 and is contrasted with an algorithm of J. Henry [13]. In Section 5.5, "quadratic convergence" properties of this algorithm are established.

5.2 Existence of $\mathcal{R}_{X\times Y}(x,y)$ and comparison to $\mathcal{R}(x,y)$.

If X is a closed subset of I, the density of X in I is defined as in (1.10)

$$d_I(X) = \sup_{s\in I} \inf_{t\in X} |s - t| .$$

Similarly, if Y is a closed subset of J, the density of Y in J is

$$d_J(Y) = \sup_{s\in J} \inf_{t\in Y} |s - t| .$$

We first state a characterization of best uniform approximations of $f \in C(I)$ over a closed subset X of I from $\mathbb{R}(n,m)$. This result is due to Dunham [8,9]. Although he proves this for the general setting of "alternating Tchebycheff approximation," we state it for the special case of rational approximation.

Lemma 5.2.1. Let X be a closed subset of I. Let $f \in C(I)$, and suppose $R(C,\cdot) \in \mathbb{R}(n,m)$ has varisolvent degree $\mu(C)$. Then $R(C,\cdot)$ is a best uniform approximation of f over X from $\mathbb{R}(n,m)$ if and only if there are $1 + \mu(C)$ points in X

$$x_0 < x_1 < \cdots < x_{\mu(C)}$$

such that $|f(x_i) - R(C,x_i)| = \|f - R(C,\cdot)\|_X$, $i = 0,\ldots,\mu(C)$, and $(f(x_i) - R(C,x_i)) = -(f(x_{i-1}) - R(C,x_{i-1}))$, $i = 1,\ldots,\mu(C)$.

Remark. We note that Lemma 5.2.1 provides a characterization for best approximations over X from $\mathbb{R}(n,m)$ and not from $\mathbb{R}_X(n,m)$.

It is assumed that $F \in \mathcal{D}$. If X is a closed subset of I and $y \in J$, define

$$(5.10)\quad \mathfrak{M}(X,y) = \{R \in \mathbb{R}_X(n,m) : \|F_y - R\|_X < \Delta_X(y) + d_I(X)\}.$$

We show that if X is sufficiently dense in I, then all candidates for $R(C_X(y),\cdot)$ are in $\mathbb{R}(n,m)$.

Lemma 5.2.2. Let $F \in \mathcal{D}$. There are numbers $\epsilon > 0$ and $\delta > 0$, independent of $y \in J$, such that if X is a closed subset of I and $d_I(X) < \delta$, then $Q(x) \geq \epsilon$ on I for all $P/Q \in \mathfrak{M}(X,y)$ where $P \in \mathbb{P}_n$, $Q \in \mathbb{Q}_m$, $Q > 0$ on X, and $\|Q\|_I = 1$.

Proof. Assume that no such $\epsilon > 0$ and $\delta > 0$ exist. Then there is a sequence $\{X_k\}$ of closed subsets of I such that

$$(5.11)\qquad \lim_{k\to\infty} d_I(X_k) = 0$$

and for each k there are points $\xi_k \in I$ and $y_k \in J$ and a rational function $R_k = P_k/Q_k \in \mathfrak{M}(X_k,y_k)$ where $P_k \in \mathbb{P}_n$, $Q_k \in \mathbb{Q}_m$, $Q_k > 0$ on X_k, $\|Q_k\|_I = 1$ and

$$(5.12)\qquad \lim_{k\to\infty} Q_k(\xi_k) = 0.$$

Since $R_k \in \mathfrak{M}(X_k; y_k)$,

$$\begin{aligned}\|F_{y_k} - R_k\|_{X_k} &\le \Delta_{X_k}(y_k) + d_I(X_k) \\ &\le \|F_{y_k} - R(C(y_k),\cdot)\|_{X_k} + d_I(X_k) \\ &\le \|F_{y_k} - R(C(y_k),\cdot)\|_I + 2 \\ &= \Delta(y_k) + 2 \\ &\le \|\Delta\|_J + 2.\end{aligned}$$

Applying the triangle inequality, we get

$$\|R_k\|_{X_k} \le \|F\|_D + \|\Delta\|_J + 2 = M.$$

Thus

$$(5.13) \qquad |P_k(x)| \le M|Q_k(x)|$$

for all $x \in X_k$.

We note that the sequence $\{Q_k\}$ is uniformly bounded over I. We now show that $\{P_k\}$ is uniformly bounded over I. Choose α_i, β_i, $i = 0,\ldots,n+1$, where

$$-1 \le \alpha_0 < \beta_0 < \alpha_1 < \beta_1 < \cdots < \alpha_{n+1} < \beta_{n+1} \le 1.$$

Lemma 1.3.3 ensures that $A = \sup\{\|P\|_I : P \in \mathbf{P}_n$ and

$\forall i = 0,\ldots,n+1$, $\exists x_i \in [\alpha_i,\beta_i] \ni (-1)^i P(x_i) \geq -1\}$ is finite. By (5.11) we may choose k_0 such that if $k > k_0$, then $X_k \cap [\alpha_i,\beta_i] \neq \emptyset$, $i = 0,\ldots,n+1$. Suppose $k > k_0$. For $i = 0,\ldots,n+1$, we select $\zeta_i \in X_k \cap [\alpha_i,\beta_i]$ and apply (5.13) to get $(-1)^i P_k(\zeta_i) \geq -M$. Thus $\|P_k\|_I \leq AM$. Hence $\{P_k\}$ is uniformly bounded by $\max\{AM, \|P_1\|_I,\ldots,\|P_{k_0}\|_I\}$.

We may now extract subsequences and relabel so that $P_k \to P \in \mathbf{P}_n$ and $Q_k \to Q \in \mathbf{Q}_m$, uniformly on I, and $\xi_k \to \xi \in I$ and $y_k \to y \in J$.

Let $x \in I$. By (5.11) there is a sequence $\{x_k\}$ where each $x_k \in X_k$ and $\lim_{k\to\infty} x_k = x$. The uniform convergence of P_k to P and the continuity of P imply that $\lim_{k\to\infty} P_k(x_k) = P(x)$. Similarly, $\lim_{k\to\infty} Q_k(x_k) = Q(x)$. Thus we conclude that

$$|P(x)| \leq M|Q(x)| \tag{5.14}$$

for all $x \in I$,

$$\|Q\|_I = 1, \tag{5.15}$$

and

$$Q(x) \geq 0$$

for $x \in I$.

Equation (5.15) implies that $Q \not\equiv 0$ and thus has at most finitely many zeros in I. By (5.14) each zero of Q of a given multiplicity is also a zero of P of at least the same multiplicity. Thus we may perform finitely many cancellations to obtain $R' = P'/Q' \in \mathbf{R}(n,m)$ where $P' \in \mathbf{P}_n$, $Q' \in \mathbf{Q}_m$, and $Q' > 0$ on I such that

$$\frac{P(x)}{Q(x)} = R'(x)$$

for all $x \in I$ where $Q(x) \neq 0$.

For any $x \in I$ where $Q(x) \neq 0$, there is a sequence $\{x_k\}$ such that each $x_k \in X_k$ and $\lim_{k\to\infty} x_k = x$. Then

$$(5.16) \quad |F_y(x)-R'(x)| = \left| F_y(x) - \frac{P(x)}{Q(x)} \right|$$

$$\leq |F_y(x)-F_{y_k}(x_k)| + |F_{y_k}(x_k)-R_k(x_k)| + \left| \frac{P_k(x_k)}{Q_k(x_k)} - \frac{P(x)}{Q(x)} \right|$$

$$\leq |F(x,y) - F(x_k,y_k)| + \|F_{y_k} - R_k\|_{X_k} + \left| \frac{P_k(x_k)}{Q_k(x_k)} - \frac{P(x)}{Q(x)} \right|$$

$$\leq |F(x,y) - F(x_k,y_k)| + \Delta_{X_k}(y_k) + d_I(X_k)$$

$$+ \left| \frac{P_k(x_k)}{Q_k(x_k)} - \frac{P(x)}{Q(x)} \right|$$

$$\leq |F(x,y) - F(x_k,y_k)| + \Delta(y_k) + d_I(X_k)$$

$$+ \left| \frac{P_k(x_k)}{Q_k(x_k)} - \frac{P(x)}{Q(x)} \right| \quad .$$

We pass to the limit in (5.16) to get

$$|F_y(x) - R'(x)| \leq \Delta(y).$$

Since R' is continuous over I,

$$\|F_y - R'\|_I \leq \Delta(y).$$

By Theorem 3.1.1, $R' = R(C(y),\cdot)$. However, $Q(\xi) = \lim_{k\to\infty} Q_k(\xi_k) = 0$. Inequality (5.14) implies that $P(\xi) = 0$. Thus P and Q have at least one common zero. Hence, $\partial P' < n$ and $\partial Q' < m$. This contradicts the normality of F_y. This completes the proof of Lemma 5.2.2.

We now apply Lemma 5.2.2 to demonstrate that if X is sufficiently dense in I, then each F_y has a unique best approximation over X from $\mathbf{R}_X(n,m)$.

<u>Lemma 5.2.3</u>. Let $F \in \mathcal{D}$. There is a $\delta > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then F_y has a unique best uniform approximation over X from $\mathbf{R}_X(n,m)$.

Proof. We choose $\epsilon > 0$ and $\delta > 0$, via Lemma 5.2.2, such that if X is a closed subset of I and $d_I(X) < \delta$, then X contains at least $n + 2$ points and $Q \geq \epsilon$ on I for all $P \in \mathbf{P}_n$, $Q \in \mathbf{Q}_m$ where $P/Q \in \mathfrak{M}(X,y)$, $Q > 0$ on X, and $\|Q\|_I = 1$.

Suppose X is a closed subset of I where $d_I(X) < \delta$ and fix $y \in J$. If $X = I$, the conclusion of Lemma 5.2.3 follows from Theorem 3.1.1. We assume X is a proper subset of I. Since X is closed, $d_I(X) > 0$ and thus there is a sequence $\{R_k\}$ in $\mathfrak{M}(X,y)$ such that $\|F_y - R_k\|_X \to \Delta_X(y)$ as $k \to \infty$. We may write $R_k = P_k/Q_k$ where $P_k \in \mathbf{P}_n$, $Q_k \in \mathbf{Q}_m$, $Q_k > 0$ on X, and $\|Q_k\|_I = 1$. We may apply an argument similar to that in the proof of Lemma 5.2.2 to see that the sequence $\{P_k\}$ is uniformly bounded over I. Thus we pass to a subsequence, relabel, and assume that $P_k \to P \in \mathbf{P}_n$ and $Q_k \to Q \in \mathbf{Q}_m$ uniformly over I. Since $Q_k \geq \epsilon$ on I for each k, $R_k = P_k/Q_k$ converges uniformly to P/Q on I and $Q \geq \epsilon$ on I. Hence, $P/Q \in \mathbf{R}_X(n,m)$ and $\|F_y - R_k\|_X \to \|F_y - P/Q\|_X$. Therefore, $\|F_y - P/Q\|_X = \Delta_X(y)$ and P/Q is a best approximation of F_y over X from $\mathbf{R}_X(n,m)$.

For the uniqueness we now suppose that F_y has two best approximations R and R' over X from $\mathbb{R}_X(n,m)$. We write $R = P/Q$ and $R' = P'/Q'$ in reduced form where $P,P' \in \mathbb{P}_n$ and $Q,Q' \in \mathbb{Q}_m$. Since $d_I(X) < \delta$, $R,R' \in \mathbb{R}(n,m)$, and since $\mathbb{R}(n,m) \subseteq \mathbb{R}_X(n,m)$, R and R' are best approximations of F_y over X from $\mathbb{R}(n,m)$. Suppose the varisolvent degree of R is ℓ. Then $\partial P \leq \ell - m - 1$ and $\partial Q \leq \ell - n - 1$. We now apply Lemma 5.2.1 to find $x_0 < x_1 < \ldots < x_\ell$ in X such that

$$(-1)^{i+\nu}(f(x_i) - R(x_i)) = \|f - R\|_X = \Delta_X(y),$$

$i = 0,\ldots,\ell$, where $\nu \in \{0,1\}$ is fixed. Since $\|f - R'\|_X = \Delta_X(y)$,

$$(-1)^{i+\nu}(f(x_i) - R'(x_i)) \leq \Delta_X(y),$$

$i = 0,\ldots,\ell$. Subtracting these inequalities we get

$$(-1)^{i+\nu}(R'(x_i) - R(x_i)) \geq 0,$$

$i = 0,\ldots,\ell$. Since $Q(x_i)Q'(x_i) > 0$,

$$(-1)^{i+\nu}(P'(x_i)Q(x_i) - P(x_i)Q'(x_i)) \geq 0,$$

$i = 0,\ldots,\ell$. But $P'Q - PQ'$ is a polynomial of degree less than or equal to $\ell - 1$. We may now conclude that

$P'Q - PQ' \equiv 0$ (see Rice [28, p. 61]). Since Q, $Q' > 0$ on X, $P/Q \equiv P'/Q'$ on X. Thus uniqueness of best approximations of F_y over X from $\mathbb{R}_X(n,m)$ is established.

Remark. Lemmas 5.3.2 and 5.3.3 ensure that if X is sufficiently dense in I, then each F_y has a unique best approximation over X from $\mathbb{R}_X(n,m)$ which we write as

$$(5.17) \quad R(C_X(y),x) = \frac{P(A_X(y),x)}{Q(B_X(y),x)} = \frac{\sum_{i=0}^{n} a_i^X(y)x^i}{1 + \sum_{j=1}^{m} b_j^X(y)x^j}$$

where $C_X(y) = (A_X(y);B_X(y)) = (a_0^X(y),\ldots,a_n^X(y);\ b_1^X(y),\ldots,b_m^X(y)) \in \mathbf{C}$ for all $y \in J$.

The next result asserts that as X converges to I with respect to the density measure $R(C_X(y),x)$ converges to $R(C(y),x)$ uniformly with respect to x and y. In the univariate case, this result is due to Werner [32]. The following lemma extends Werner's theorem to the uniform sense in y.

<u>Lemma 5.2.4</u>. Let $F \in \mathcal{D}$. Given $\epsilon > 0$ there is a $\delta > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then $\|R(C_X(y),\cdot) - R(C(y),\cdot)\|_I < \epsilon$ for all $y \in J$.

Proof. Assume otherwise. Then there is a sequence $\{X_k\}$ of closed subsets of I and a sequence $\{y_k\}$ in J such that $\lim_{k\to\infty} d_I(X_k) = 0$ and

$$(5.18)\qquad \| R(C_{X_k}(y_k),\cdot) - R(C(y_k),\cdot)\|_I \geq \epsilon .$$

Since J is compact, we may assume $y_k \to y \in J$ as $k \to \infty$. Moreover, we can apply an argument similar to that of the proof of Lemma 5.2.3 to show that a further relabeling produces $R(C_{X_k}(y_k),\cdot) \to R \in \mathbf{R}(n,m)$ uniformly on I as $k \to \infty$. Since $F_{y_k} - R(C_{X_k}(y_k),\cdot) \to F_y - R$ uniformly on I, $F_y - R$ is continuous on I, $d_I(X_k) \to 0$, and $F_{y_k} - R(C(y_k),\cdot)$ $\to F_y - R(C(y),\cdot)$ uniformly on I,

$$\begin{aligned} \|F_y - R\|_I &= \lim_{k\to\infty} \|F_{y_k} - R(C_{X_k}(y_k),\cdot)\|_{X_k} \\ &\leq \lim_{k\to\infty} \|F_{y_k} - R(C(y_k),\cdot)\|_{X_k} \\ &\leq \lim_{k\to\infty} \|F_{y_k} - R(C(y_k),\cdot)\|_I \\ &= \|F_y - R(C(y),\cdot)\|_I . \end{aligned}$$

By Theorem 3.1.1, $R = R(C(y),\cdot)$ and $R(C_{X_k}(y_k),\cdot) \to R(C(y),\cdot)$ uniformly on I. Now

$$\|R(C_{X_k}(y_k),\cdot) - R(C(y_k),\cdot)\|_I$$
$$\leq \|R(C_{X_k}(y_k),\cdot) - R(C(y),\cdot)\|_I + \|R(C(y),\cdot) - R(C(y_k),\cdot)\|_I.$$

The uniform convergence above implies that $\|R(C_{X_k}(y_k),\cdot) - R(C(y),\cdot)\|_I \to 0$, and Theorem 3.1.4 implies that $\|R(C(y),\cdot) - R(C(y_k),\cdot)\|_I \to 0$. Thus $\|R(C_{X_k}(y_k),\cdot) - R(C(y_k),\cdot)\|_I \to 0$. This contradicts (5.18) and Lemma 5.2.4 is proven.

The following corollary is an immediate consequence of Lemma 5.2.4 and Lemma 4.3.2.

<u>Corollary 5.2.5</u>. Let $F \in \mathcal{D}$. Given $\epsilon > 0$, there is a $\delta > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then $\sigma(C(y), C_X(y)) < \epsilon$ for all $y \in J$.

The following result establishes the continuity of the coefficients in $R(C_X(y),x)$ when X is sufficiently dense in I. We shall delay proving this result until Section 5.3.

<u>Lemma 5.2.6</u>. Let $F \in \mathcal{D}$. There is a $\delta > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then $C_X(y)$ is continuous on the interval J.

Thus if X is sufficiently dense in I, then $C_X(y)$ is continuous over J and the coefficient functions $a_0^X(y),\ldots,a_n^X(y)$, $b_1^X(y),\ldots,b_m^X(y)$ may be best approximated. The series of lemmas now culminates in the first major result of Section 5.2.

<u>Theorem 5.2.7</u>. Let $F \in \mathcal{D}$. Then there is a $\delta > 0$ such that if X and Y are closed subsets of I and J, respectively, and $d_I(X) < \delta$ and $d_J(Y) < \delta$, then $\mathcal{R}_{X\times Y}$ exists.

Proof. We must show that if X is sufficiently dense in I, then $R(C_X(y),\cdot)$ exists for all $y \in J$, $C_X(y) \in \mathbb{C}$ for all $y \in J$, and $C_X(y)$ is continuous over J. This is provided by the preceding series of lemmas. We must also show that the denominator of (5.9) does not vanish over $X \times Y$.

Let

$$\tau = \inf_{(x,y)\in D} \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j,y)x^j\right).$$

Since $F \in \mathcal{D}$, $\tau > 0$. Since $b_j \in C(J)$, $j = 1,\ldots,m$, we apply Theorem 1.4.2 to find numbers $\delta_1 > 0$ and $\lambda > 0$ such that

$$\|T_{s_j}^Y(g,\cdot) - T_{s_j}^Y(b_j,\cdot)\|_J \leq \lambda \|g - b_j\|_Y$$

for all $g \in C(J)$, $j = 1,\ldots,m$, whenever Y is a closed subset of J and $d_J(Y) < \delta_1$. We now apply Theorem 3 on p. 87 of [5], to obtain $\delta_2 > 0$ such that if $d_J(Y) < \delta_2$, then $\| T_{s_j}(b_j,\cdot) - T^Y_{s_j}(b_j,\cdot)\|_J < \tau/4\lambda m$, $j = 1,\ldots,m$.

By Lemmas 5.2.2, 5.2.3, and 5.2.6, there is a $\delta_3 > 0$ such that if X is a closed subset of I and $d_I(X) < \delta_3$, then $R(C_X(y),\cdot)$ exists and is unique for all $y \in J$, $C_X(y) \in \mathbf{C}$ for all $y \in J$, and $C_X(y)$ is continuous over J. By Corollary 5.2.5, there is a $\delta_4 > 0$ $(\delta_4 < \delta_3)$ such that if $d_I(X) < \delta_4$, then $\| b_j - b^X_j\|_J < \tau/2\lambda m$, $j = 1,\ldots,m$.

We choose $\delta = \min(\delta_1,\delta_2,\delta_4) > 0$ and suppose X and Y are closed subsets of I and J, respectively, with $d_I(X) < \delta$ and $d_J(Y) < \delta$. We need only show that the denominator of (5.9) is positive over $X \times Y$. Suppose $(x,y) \in D$. Then

$$1 + \sum_{j=1}^{m} T^Y_{s_j}(b^X_j,y)x^j$$

$$\geq 1 + \sum_{j=1}^{m} T_{s_j}(b_j,y)x^j$$

$$- \sum_{j=1}^{m} [T_{s_j}(b_j,y) - T^Y_{s_j}(b_j,y)]x^j$$

$$- \sum_{j=1}^{m} [T^Y_{s_j}(b_j,y) - T^Y_{s_j}(b^X_j,y)]x^j$$

$$\geq 1 + \sum_{j=1}^{m} T_{s_j}(b_j, y)x^j$$

$$- \sum_{j=1}^{m} \| T_{s_j}(b_j, \cdot) - T^Y_{s_j}(b_j, \cdot)\|_J$$

$$- \sum_{j=1}^{m} \| T^Y_{s_j}(b_j, \cdot) - T^Y_{s_j}(b^X_j, \cdot)\|_J$$

$$> \tau - \frac{\tau}{4} - \frac{\tau}{4}$$

$$> 0 .$$

The denominator of (5.9) is positive over D and thus is positive over $X \times Y$. Hence $\mathcal{R}_{X \times Y}$ is defined.

The final result of this section compares $\mathcal{R}_{X \times Y}(x,y)$ to $\mathcal{R}(x,y)$ as X and Y fill out their respective intervals. We comment that the proof bears out that if X and Y are sufficiently dense in their respective intervals, then the denominator of the representation (5.9) for $\mathcal{R}_{X \times Y}(x,y)$ is positive over all of D. This makes it possible to compare $\mathcal{R}_{X \times Y}(x,y)$ to $\mathcal{R}(x,y)$ over D.

<u>Theorem 5.2.8</u>. Let $F \in \mathcal{D}$. Given $\epsilon > 0$, there is a $\delta > 0$ such that if X and Y are closed subsets of I and J,

respectively, with $d_I(X) < \delta$ and $d_J(Y) < \delta$, then $\mathcal{R}_{X\times Y}$ exists and $\|\mathcal{R} - \mathcal{R}_{X\times Y}\|_D < \epsilon$.

Proof. Let

$$\tau = \inf_{(x,y)\in D} \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j,y)x^j\right) > 0.$$

By the proof of Theorem 5.2.7, we may select $\delta_1 > 0$ such that if $d_I(X) < \delta_1$ and $d_J(Y) < \delta_1$, then

$$1 + \sum_{j=1}^{m} T_{s_j}^{Y}(b_j^X,y)x^j > \tau/2$$

for $(x,y) \in D$. Thus for $(x,y) \in D$,

$$(5.19)\quad |\mathcal{R}(x,y) - \mathcal{R}_{X\times Y}(x,y)|$$

$$\leq (2/\tau^2)\left|\left(\sum_{i=0}^{n} T_{r_i}(a_i,y)x^i\right)\left(1 + \sum_{j=1}^{m} T_{s_j}^{Y}(b_j^X,y)x^j\right)\right.$$

$$\left. - \left(1 + \sum_{j=1}^{m} T_{s_j}(b_j,y)x^j\right)\left(\sum_{i=0}^{n} T_{r_i}^{Y}(a_i^X,y)x^i\right)\right|$$

$$\leq (2/\tau^2)\left[\; K_1 \sum_{j=1}^{m} \|T_{s_j}^{Y}(b_j^X,\cdot) - T_{s_j}(b_j,\cdot)\|_J\right.$$

$$\left. + K_2 \sum_{i=0}^{n} \|T_{r_i}(a_i,\cdot) - T_{r_i}^{Y}(a_i^X,\cdot)\|_J\;\right],$$

where $K_1 = \sum_{i=0}^{n} \|T_{r_i}(a_i,\cdot)\|_J$ and

$K_2 = 1 + \sum_{j=1}^{m} \|T_{s_j}(b_j,\cdot)\|_J$. By Theorem 1.4.2, there are numbers $\delta_2 > 0$ and $\lambda > 0$ such that if $d_J(Y) < \delta_2$, then

$$\|T^Y_{r_i}(g,\cdot) - T^Y_{r_i}(a_i,\cdot)\|_J \leq \lambda \|g - a_i\|_Y$$

and

$$\|T^Y_{s_j}(g,\cdot) - T^Y_{s_j}(b_j,\cdot)\|_J < \lambda \|g - b_j\|_Y ,$$

for all $g \in C(J)$, $i = 0,\ldots,n$, and $j = 1,\ldots,m$. Theorem 3 on p. 87 of [5] yields a $\delta_3 > 0$ such that

$$(5.20) \qquad \|T_{r_i}(a_i,\cdot) - T^Y_{r_i}(a_i,\cdot)\|_J < \frac{\epsilon\tau^2}{8\,K_2(n+1)}$$

$i = 0,\ldots,n$, and

$$(5.21) \qquad \|T_{s_j}(b_j,\cdot) - T^Y_{s_j}(b_j,\cdot)\|_J < \frac{\epsilon\tau^2}{8\,K_1 m} ,$$

$j = 1,\ldots,m$, whenever $d_J(Y) < \delta_3$. Moreover, Corollary 5.2.5 produces a $\delta_4 > 0$ such that if $d_I(X) < \delta_4$, then

$$(5.22) \qquad \|a_i - a^X_i\|_J < \frac{\epsilon\tau^2}{8\,K_2(n+1)\lambda}$$

$i = 0,\ldots,n$, and

$$(5.23) \qquad \| b_j - b_j^X \|_J < \frac{\epsilon\tau^2}{8 K_1 m\lambda},$$

$j = 1,\ldots,m$.

Let $\delta = \min(\delta_1, \delta_2, \delta_3, \delta_4) > 0$. If $d_I(X) < \delta$ and $d_J(Y) < \delta$, then inequalities (5.19), (5.20), (5.21), (5.22), and (5.23) imply that

$$|\mathfrak{R}(x,y) - \mathfrak{R}_{X\times Y}(x,y)| < \epsilon.$$

Thus $\|\mathfrak{R} - \mathfrak{R}_{X\times Y}\|_D < \epsilon$, and Theorem 5.2.8 is proven.

We conclude this section by noting that the two theorems of this section justify the discretization of algorithms to compute rational product approximations. In particular, if X and Y are sufficiently dense in I and J, respectively, then the corresponding discrete rational product approximation exists and is uniformly near the rational product approximation of F over D.

5.3 Discrete strong unicity theorem.

In this section, we establish a generalization of Theorem 3.3.3 to the setting of discrete rational approximation. In this case, the strong unicity constant will be independent of F_y, $y \in J$, and the closed subset X of I over which the approximation is done. This result will be used in Section 5.5 in determining quadratic convergence properties for an algorithm to compute rational product approximations. We also state another variant of the strong unicity theorem. We will use this result to prove Lemma 5.2.6.

Theorem 5.3.1. Let $F \in D$ and suppose $F_y \notin \mathbf{R}(n,m)$ for all $y \in J$. Then there are constants $\delta = \delta(F) > 0$ and $\gamma = \gamma(F) > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then the best uniform approximation $R(C_X(y),\cdot)$ of F_y over X from $\mathbf{R}_X(n,m)$ exists for all $y \in J$ and

(5.24) $\|F_y - R\|_X \geq \|F_y - R(C_X(y),\cdot)\|_X + \gamma\|R - R(C_X(y),\cdot)\|_X$

for all $y \in J$ and all $R \in \mathbf{R}_X(n,m)$.

Proof. By Lemma 5.2.3, there is a $\delta_1 > 0$ such that $R(C_X(y),\cdot)$ exists for all $y \in J$ whenever $d_I(X) < \delta_1$.

Assume no such $\gamma > 0$ and $\delta > 0$ ($\delta \leq \delta_1$) exist. Then there are sequences $\{X_k\}$ of closed subsets of I, $\{y_k\}$ in J, and $\{R^k\}$ where $R^k \in \mathbf{R}_{X_k}(n,m)$ such that

(5.25) $$d_I(X_k) \to 0$$

and

(5.26) $$\frac{\|F_{y_k} - R^k\|_{X_k} - \|F_{y_k} - R(C_{X_k}(y_k),\cdot)\|_{X_k}}{\|R^k - R(C_{X_k}(y_k),\cdot)\|_{X_k}} \to 0$$

as $k \to \infty$. We pass to a subsequence and relabel so that $y_k \to y \in J$. Thus $F_{y_k} \to F_y$ uniformly over I as $k \to \infty$. Theorem 3.1.4 and Lemma 5.2.4 now imply that $R(C_{X_k}(y_k),\cdot) \to R(C(y),\cdot)$ uniformly over I as $k \to \infty$.

We normalize the rational functions and write $R^k = N^k/D^k$, $R(C_{X_k}(y_k),\cdot) = N_{X_k}/D_{X_k}$, and $R(C(y),\cdot) = N^*/D^*$ where N^k, N_{X_k}, $N^* \in \mathbb{P}_n$, D^k, D_{X_k}, $D^* \in \mathbb{Q}_m$, and

(5.27) $$\|N^k\|_I + \|D^k\|_I + \|N_{X_k}\|_I + \|D_{X_k}\|_I = \|N^*\|_I + \|D^*\|_I = 1.$$

Equation (5.27) and an argument similar to the proof of Theorem 3.3.3 (pp. 88 - 91) allow us to further relabel and assume that $N^k \to N^*$, $D^k \to D^*$, $N_{X_k} \to N^*$, and $D_{X_k} \to D^*$ uniformly over I as $k \to \infty$. Since $R(C(y),\cdot) \in \mathbf{R}(n,m)$, we may select an $\epsilon > 0$ such that $\epsilon \leq D^k(x) \leq 1$ and $\epsilon \leq D_{X_k}(x) \leq 1$ for all $x \in I$ and all sufficiently large k.

Now Corollary 5.2.5 implies that for sufficiently large k, $R(C_{X_k}(y_k),\cdot) \in \mathbf{R}(n,m)$ and is the best approximation of F_y over X from $\mathbf{R}(n,m)$ and has varisolvent degree $\ell = n + m + 1$. By Lemma 5.2.1, there is an alternation set

$$-1 \leq x_0^k < x_1^k < \dots < x_\ell^k \leq 1$$

in X_k for $F_{y_k} - R(C_{X_k}(y_k),\cdot)$. We further pluck out a subsequence and relabel so that $x_i^k \to x_i$, $i = 0,\dots,\ell$. Since $F_y \notin \mathbf{R}(n,m)$, an argument similar to that on p. 33 insures that

$$x_0 < x_1 < \dots < x_\ell .$$

Thus by Theorem 1.3.5, $K(x_0^k,\dots,x_\ell^k) = \sup\{\|P\|_I$: P is a polynomial of degree at most $\ell - 1$ and $(-1)^i P(x_i^k) \geq -1$, $i = 0,\dots,\ell\}$ are bounded by some $A > 0$.

We now fix k. For $i = 0,\dots,\ell$,

$$(-1)^{i+\nu}[F_{y_k}(x_i^k)-R(C_{X_k}(y_k),x_i^k)] = \|F_{y_k}-R(C_{X_k}(y_k),\cdot)\|_{X_k}$$

and

$$(-1)^{i+\nu}[F_{y_k}(x_i^k) - R^k(x_i^k)] \leq \|F_{y_k} - R^k\|_{X_k},$$

where $\nu \in \{0,1\}$ is fixed. Subtracting these inequalities produces the inequality

$$(-1)^{i+\nu}[R^k(x_i^k) - R(C_{X_k}(y_k),x_i^k)]$$
$$\geq -(\|F_{y_k} - R^k\|_{X_k} - \|F_{y_k} - R(C_{X_k}(y_k),\cdot)\|_{X_k}).$$

Since $0 < D^k \leq 1$ and $0 < D_{X_k} \leq 1$,

$$(-1)^{i+\nu}[N^k(x_i^k)D_{X_k}(x_i^k) - D^k(x_i^k)N_{X_k}(x_i^k)]$$
$$\geq -(\|F_{y_k} - R^k\|_{X_k} - \|F_{y_k} - R(C_{X_k}(y_k),\cdot)\|_{X_k}).$$

Since $N^kD_{X_k} - D^kN_{X_k}$ is a polynomial of degree at most $\ell - 1$,

$$\|N^kD_{X_k}-D^kN_{X_k}\|_{X_k} \leq A(\|F_{y_k}-R^k\|_{X_k} - \|F_{y_k}-R(C_{X_k}(y_k),\cdot)\|_{X_k}).$$

Now since

$$\|R^k - R(C_{X_k}(y_k),\cdot)\|_{X_k} \leq \frac{1}{\epsilon^2}\|N^kD_{X_k} - D^kN_{X_k}\|_{X_k},$$

we have

$$\frac{\|F_{y_k} - R^k\|_{X_k} - \|F_{y_k} - R(C_{X_k}(y_k),\cdot)\|_{X_k}}{\|R^k - R(C_{X_k}(y_k),\cdot)\|_{X_k}} \geq \frac{\epsilon^2}{A}.$$

This contradicts (5.26), and the proof of Theorem 5.3.1 is complete.

We next state another discrete strong unicity theorem. This theorem is cited without proof in [1]. The proof is similar to the proof of Theorem 5.3.1 and thus is omitted.

Lemma 5.3.2. Let $f \in C(I)$ and X be a closed subset of I containing at least $n + m + 2$ points. Suppose f has a best uniform approximation R^* over X from $\mathbf{R}(n,m)$ and $R^* \in \mathbf{R}^*$. Then there is a constant $\gamma > 0$ such that

$$\|f - R\|_X \geq \|f - R^*\|_X + \gamma \|R - R^*\|_X$$

for all $R \in \mathbf{R}(n,m)$.

We remark that the conditions of Lemma 5.3.2 do not require X to be sufficiently dense in the interval I. We conclude this section with the proof of Lemma 5.2.6.

Proof of Lemma 5.2.6. By Lemmas 5.2.2 and 5.2.3 and Corollary 5.2.5, we may select a $\delta > 0$ such that if X is a closed subset of I and $d_I(X) < \delta$, then for all $y \in J$,

(i) $R(C_X(y),\cdot)$ exists,

(ii) $R(C_X(y),\cdot) \in \mathbf{R}(n,m)$ and thus is the best approximation of F_y over X from $\mathbf{R}(n,m)$, and

(iii) $R(C_X(y),\cdot) \in \mathbf{R}^*$.

We now apply Lemma 5.3.2 to find $\gamma(y)$, $y \in J$, such that

$$(5.28)\quad \|F_y - R\|_X \geq \|F_y - R(C_X(y),\cdot)\|_X + \gamma(y)\|R - R(C_X(y),\cdot)\|_X$$

for all $R \in \mathbf{R}(n,m)$. Now let $y_0 \in J$. For any $y \in J$, we apply (5.28) to get

$$\|R(C_X(y),\cdot) - R(C_X(y_0),\cdot)\|_X$$

$$\leq \frac{1}{\gamma(y_0)}\,[\|F_{y_0} - R(C_X(y),\cdot)\|_X - \|F_{y_0} - R(C_X(y_0),\cdot)\|_X]$$

$$\leq \frac{1}{\gamma(y_0)}\,[\|F_{y_0} - F_y\|_X + \|F_y - R(C_X(y),\cdot)\|_X$$

$$- \|F_{y_0} - R(C_X(y_0),\cdot)\|_X]$$

$$\leq \frac{1}{\gamma(y_0)}\,[\|F_{y_0} - F_y\|_X + \|F_y - R(C_X(y_0),\cdot)\|_X$$

$$- \|F_{y_0} - R(C_X(y_0),\cdot)\|_X]$$

$$\leq \frac{2}{\gamma(y_0)}\,\|F_y - F_{y_0}\|_X\,.$$

Thus $R(C_X(y),\cdot)$ depends continuously on y. However, each $R(C_X(y),\cdot) \in \mathbf{R}^*$, and by Lemma 3.2.5, $\mathbf{R}^*$ and $\mathbf{C}^*$ are homeomorphic. We may now conclude that $C_X(y)$ depends continuously on y over J. Thus Lemma 5.2.6 is proven.

5.4 Algorithms.

The algorithms to compute polynomial or rational product approximations of $F \in C(D)$ given in the papers [4, 12, 13, 31] follow a specific pattern. A finite subset Y of J is selected, and for each $y \in Y$, the best approximation of F_y is computed over I or some prescribed finite subset X of I. The associated coefficient functions are recovered and are best approximated over Y. The best approximations are computed using the Remes algorithm [5] or the differential correction algorithm [1]. The algorithms of J. Henry [12, 13], however, include an outlet to handle a possible discontinuity of the coefficient functions (see p. 105). In this section, we discuss a version of an algorithm of J. Henry [13] and propose a modification of this algorithm.

We first describe the original differential correction (ODC) algorithm to compute the best uniform approximations of $f \in C(I)$ over a finite subset X of I from $\mathbf{R}_X(n,m)$. The ODC algorithm is due to Cheney and Loeb [6], but convergence of the algorithm was established by Barrodale, Powell, and Roberts [1]. We present the ODC algorithm in connection with computing $R(C_X(y),\cdot)$ where $F \in C(D)$.

<u>ODC Algorithm</u>.

Step 1. Choose $R_0 = P_0/Q_0$ where

$$P_0(x) = p_0^0 + p_1^0 x + \ldots + p_n^0 x^n,$$

$$Q_0(x) = q_0^0 + q_1^0 x + \ldots + q_m^0 x^m > 0 \text{ on } X,$$

and $\max_j |q_j^0| = 1$. Let

$$\Delta_0(y) = \|F_y - R_0\|_X.$$

Step 2. At the k-th step, we have $R_k = P_k/Q_k$ where

$$P_k(x) = p_0^k + p_1^k x + \ldots + p_n^k x^n,$$

$$Q_k(x) = q_0^k + q_1^k x + \ldots + q_m^k x^m > 0 \text{ on } x,$$

and $\max_j |q_j^k| = 1$. We denote

(5.29) $$\Delta_k(y) = \|F_y - R_k\|_X.$$

We choose $R_{k+1} = P_{k+1}/Q_{k+1}$ where

$$P_{k+1}(x) = p_0^{k+1} + p_1^{k+1} x + \ldots + p_n^{k+1} x^n$$

and

$$Q_{k+1}(x) = q_0^{k+1} + q_1^{k+1} x + \ldots + q_m^{k+1} x^m$$

to minimize

$$(5.30) \quad \max_{x\in X} \left\{ \frac{|F_y(x)Q_{k+1}(x)-P_{k+1}(x)| - \Delta_k(y)Q_{k+1}(x)}{Q_k(x)} \right\}$$

subject to the constraint

$$(5.31) \qquad \max_j |q_j^{k+1}| = 1 .$$

Barrodale, Powell, and Roberts have shown that the ODC algorithm sustains itself in that $Q_{k+1} > 0$ on X, $\Delta_k(y) \downarrow \Delta_X(y)$, and if F_y is normal, then the convergence is at least quadratic. We remark that Step 2 of the ODC algorithm is accomplished by a linear program. Step 1 is accomplished by choosing P_0 and Q_0 to minimize

$$(5.32) \qquad \max_{x\in X} |F_y(x)Q_0(x) - P_0(x)|$$

subject to the constraint $q_0 = 1$, or if (5.32) does not generate a permissible Q_0, by choosing $P_0 \equiv 0$ and $Q_0 \equiv 1$. The initialization is used by Kaufman and Taylor [18, 19] in a program of the ODC algorithm. Kaufman and Taylor refer to (5.32) as "one step of Loeb's algorithm."

We now describe a simplification of J. Henry's algorithm [13]. We assume $F \in D$. Thus we do not consider the discontinuity outlet.

Algorithm 1.

Step 1. Choose finite subsets

$$X: x_1 < x_2 < \dots < x_r$$

and

$$Y: y_1 < y_2 < \dots < y_s$$

of I and J, respectively.

Step 2. Compute $R(C_X(y_\nu),\cdot)$, $\nu = 1,\dots,s$, using the ODC algorithm with initialization (5.32). Recover the coefficient vector

$$C_X(y_\nu) = (a_0^X(y_\nu),\dots,a_n^X(y_\nu);\ b_1^X(y_\nu),\dots,b_m^X(y_\nu)).$$

Step 3. Compute $T_{r_i}^Y(a_i^X,\cdot)$, $i = 0,\dots,n$, and $T_{s_j}^Y(b_j^X,\cdot)$, $j = 1,\dots,m$, using the ODC algorithm and initialization (5.32).

We remark that Step 3 of Algorithm 1 can compute rational approximations of the coefficient functions, and thus Algorithm 1 can be used to compute rational product approximations in the sense of Brown and Henry [3].

We now describe the proposed modified algorithm.

Algorithm 2.

Step 1. Choose finite subsets

$$X:\ x_1 < x_2 < \dots < x_r$$

and

$$Y:\ y_1 < y_2 < \dots < y_s$$

of I and J, respectively.

Step 2. Compute $R(C_X(y_0),\cdot)$ using the ODC algorithm with initialization (5.32). Recover the coefficient vector

$$C_X(y_0) = (a_0^X(y_0),\dots,a_n^X(y_0);\ b_1^X(y_0),\dots,b_m^X(y_0)).$$

Step 3. Compute $R(C_X(y_\nu),\cdot)$, $\nu = 2,\dots,s$, using the ODC algorithm with initial guess $R_0 = R(C_X(y_{\nu-1}),\cdot)$. Recover the coefficient vector

$$C_X(y\ ) = (a_0^X(y_\nu),\dots,a_n^X(y_\nu);\ b_1^X(y_\nu),\dots,b_m^X(y_\nu)).$$

Step 4. Compute $T_{r_i}^Y(a_i^X,\cdot)$, $i = 0,\dots,n$, and

$T^Y_{s_j}(b^X_j,\cdot)$ using the ODC algorithm with initialization (5.32).

Algorithms 1 and 2 are multiply-iterative in that the iterative ODC algorithm is applied several times. The shortcomings of all the product approximation algorithms is the time of computation of $R(C_X(y),\cdot)$ for several values of y. The ODC algorithm possesses a quadratic rate of convergence [1], and the motivation for Step 3 of Algorithm 2 is to take advantage of this quadratic rate of convergence. It will be shown that there is a constant A, independent of y and sufficiently dense $X \subseteq I$, such that

$$(5.33) \qquad \Delta_{k+1}(y) - \Delta_X(y) \leq A(\Delta_k(y) - \Delta_X(y))^2$$

for all $y \in J$, where $\Delta_{k+1}(y)$ and $\Delta_k(y)$ are given in (5.29). This implies that for all k

$$\Delta_k(y) - \Delta_X(y) \leq \frac{1}{A}\,[A(\Delta_0(y) - \Delta_X(y))]^{2^k}.$$

The motivation for Step 3 of Algorithm 2 follows from Lemma 5.2.6. We infer that if Y is sufficiently dense in J, then the choice of $R(C_X(y_{\nu-1}),\cdot)$ as the guess in the computation of $R(C_X(y_\nu),\cdot)$ would produce

$$A(\Delta_0(y_\nu) - \Delta_X(y_w)) < 1.$$

This inequality and the geometric growth of the exponent 2^k would greatly enhance the convergence rate of each iteration in Step 3 of Algorithm 2. We remark that although Algorithm 1 satisfies the quadratic convergence property (5.33), there is no guarantee that the quadratic convergence is effective until the inequality

$$A(\Delta_k(y) - \Delta_X(y)) < 1$$

is satisfied.

We now state this quadratic convergence property. Since the proof is rather long and technical, we defer the proof until Section 5.5.

Theorem 5.4.1. Suppose $F \in D$ and $F_y \notin \mathbf{R}(n,m)$ for all $y \in J$. Then there are constants $\delta > 0$ and $A > 0$, independent of y, such that if X is a finite subset of I, then

$$\Delta_{k+1}(y) - \Delta_X(y) \leq A(\Delta_k(y) - \Delta_X(y))^2,$$

$k = 0,1,\ldots,$ for all $y \in J$ where $\Delta_k(y)$ and $\Delta_{k+1}(y)$ are the uniform errors at the k-th and (k+1)-st iterations of the ODC algorithm.

5.5 Proof of Theorem 5.4.1.

We now return to the proof of Theorem 5.4.1. Throughout this section, it is assumed that $F \in D$ and $F_y \notin \mathbf{R}(n,m)$ for all $y \in J$. We recall that $F \in D$ implies that F_y is normal with respect to I and $\mathbf{R}(n,m)$ for all $y \in J$. The proof of this section is patterned after the proof of Barrodale, Powell, and Roberts [1] with adjustments made to insure that the number A of (5.33) may be chosen independently of $y \in J$ and sufficiently dense $X \subseteq I$.

The proof of Theorem 5.4.1 is preceded by three lemmas.

Lemma 5.5.1. Let $F \in D$. There are numbers $\delta > 0$ and $\eta > 0$, independent of $y \in J$, such that if X is a closed subset of I, $d_I(X) < \delta$, and $R(C_X(y),\cdot) = P_y/Q_y$ where $P_y(x) = \sum_{i=0}^{n} p_i(y)x^i$, $Q_y(x) = \sum_{j=0}^{m} q_j(y)x^j$, and $\max_j |q_j(y)| = 1$, then $Q_y(x) \geq \eta$ for all $x \in I$.

Proof. By Lemmas 5.2.2 and 5.2.3, we may choose $\epsilon > 0$ and $\delta > 0$ such that whenever $d_I(X) < \delta$, then

$R(C_X(y),\cdot)$ exists for all $y \in J$ and

$$\frac{Q_y(x)}{\|Q_y\|_I} \geq \epsilon$$

for all $x \in I$. Thus

$$Q_y(x) \geq \epsilon \, \|Q_y\|_I$$

on I. The two norms $\|\cdot\|_I$ and $\left\|\sum_{j=0}^{m} q_j x^j\right\|_* = \max_j |q_j|$ on the finite dimensional linear space $\mathbb{Q}_m$ are equivalent. Thus there is a number $\tau > 0$ which does not depend on y or X such that $\|Q_y\|_I \geq \tau \max_j |q_j(y)| = \tau$. Thus

$$Q_y(x) \geq \epsilon\, \tau$$

on I. The choice of $\eta = \epsilon\tau$ completes the proof of Lemma 5.5.1.

Lemma 5.5.2. Suppose $\{f_k\}$ is a sequence of real-valued functions on I and $f_k \to f$ uniformly over I where $f \in C(I)$. Let $\{X_k\}$ be a sequence of subsets of I where $d_I(X_k) \to 0$ as $k \to \infty$. Then $\|f_k\|_{X_k} \to \|f\|_I$ as $k \to \infty$.

Proof. The uniform convergence implies that $\|f_k - f\|_I \to 0$. The inequality

$$\|f_k\|_{X_k} \leq \|f_k\|_I \leq \|f_k - f\|_I + \|f\|_I$$

now implies that

$$\overline{\lim_{k\to\infty}} \|f_k\|_{X_k} \leq \|f\|_I.$$

Now pick $x^* \in I$ such that $|f(x^*)| = \|f\|_I$. Since $d_I(X_k) \to 0$, we may choose a sequence $\{x_k\}$ where $x_k \in X_k$ and $x_k \to x^*$. Then

$$\begin{aligned}\|f\|_I &= |f(x^*)| \\ &\leq |f(x^*) - f(x_k)| + |f(x_k) - f_k(x_k)| + |f_k(x_k)| \\ &\leq |f(x^*) - f(x_k)| + \|f - f_k\|_I + \|f_k\|_{X_k}.\end{aligned}$$

This inequality, the continuity of f, and the uniform convergence $f_k \to f$ over I ensure that

$$\|f\|_I \leq \underline{\lim_{k\to\infty}} \|f_k\|_{X_k}.$$

We conclude that $\|f\|_I = \lim_{k\to\infty} \|f_k\|_{X_k}$.

Lemma 5.5.3. Let $F \in D$. Then there are numbers $\delta > 0$ and $\theta > 0$, independent of $y \in J$, such that if X is a closed subset of I and $d_I(X) < \delta$, then

$$\|Q - Q_y\|_X \leq \theta \, \|R - R(C_X(y),\cdot)\|_X$$

for all $R = P/Q$ where $P(x) = \sum_{i=0}^{n} \overline{p}_i x^i$, $Q(x) = \sum_{j=0}^{m} \overline{q}_j x^j > 0$ on I, and $\max_j |\overline{q}_j| = 1$, and $R(C_X(y),\cdot) = P_y/Q_y$ is given in Lemma 5.5.1.

Proof. We assume otherwise. Then there are sequences $\{X_k\}$ of closed subsets of I with $d_I(X_k) \to 0$ as $k \to \infty$, $\{y_k\}$ in J, and $\{R_k\}$ where $R_k = P_k/Q_k$, $P_k(x) = \sum_{i=0}^{n} \overline{p}_i^k x^i$, $Q_k(x) = \sum_{j=0}^{m} \overline{q}_j^k x^j > 0$ on I, and $\max_j |\overline{q}_j^k| = 1$ such that

$$\frac{\|Q_k - Q_{y_k}\|_{X_k}}{\|R_k - R(C_{X_k}(y_k),\cdot)\|_{X_k}} \to \infty \tag{5.36}$$

as $k \to \infty$. We simplify our notations by writing $R(C_{X_k}(y_k),\cdot) = R_{y_k} = P_{y_k}/Q_{y_k}$ where $P_{y_k}(x) = \sum_{i=0}^{n} p_i^k x^i$ and $Q_{y_k}(x) = \sum_{j=0}^{m} q_j^k x^j$.

We extract subsequences and relabel so that $y_k \to y \in J$.

We write $R(C(y),\cdot) = R = P/Q$ where $P(x) = \sum_{i=0}^{n} p_i x^i$ and $Q(x) = \sum_{j=0}^{m} q_j x^j$. Corollary 5.2.5 and the triangle inequality insure that $C_{X_k}(y_k) \to C(y)$. From this it can be deduced that $P_{y_k} \to P$ and $Q_{y_k} \to Q$ uniformly over I.

Since $F \in \mathcal{D}$, F_y is normal and, therefore, $R \in \mathbb{R}^*$. As a result, P/Q is a reduced representation for R, and $p_n \neq 0$ or $q_m \neq 0$. We consider the case where $p_n \neq 0$. The case where $q_m \neq 0$ can be handled in a similar manner.

Since $P_{y_k} \to P$ uniformly on I, $p_n^k \to p_n$. Since $p_n \neq 0$, a relabeling allows us to assume that all $p_n^k \neq 0$. Since P/Q is in reduced form $\{Q, xQ, \ldots, x^{n-1}Q, P, xP, \ldots, x^m P\}$ is a linearly independent set. The equivalence of all norms on a finite dimensional space implies that

$$H = \inf \{\|QA + PB\|_I : A(x) = \sum_{i=0}^{n-1} a_i x^i,$$
$$B(x) = \sum_{j=0}^{m} b_j x^j, \max_{i,j} \{|a_i|, |b_j|\} = 1\}$$

is positive. For each k, let

$$H_k = \inf \{\|Q_{y_k} A + P_{y_k} B\|_{X_k} : A(x) = \sum_{i=0}^{n-1} a_i x^i,$$
$$B(x) = \sum_{j=0}^{m} b_j x^j, \max_{i,j} \{|a_i|, |b_j|\} = 1\}.$$

We show that the H_k can be bounded away from zero. Since $\|Q_{y_k} A + P_{y_k} B\|_{X_k}$ is continuous with respect to the vector $(a_0, \ldots, a_{n-1}, b_0, \ldots, b_m)$ over the compact set $\{(a_0, \ldots, a_{n-1}, b_0, \ldots, b_m) : \max_{i,j} \{|a_i|, |b_j|\} = 1\}$. Thus we can find $A_k^*(x) = \sum_{i=0}^{n-1} a_i^k x^i$ and $B_k^*(x) = \sum_{j=0}^{m} b_j^k x^j$ where $\max_{i,j} \{|a_i^k|, |b_j^k|\} = 1$ and

$$H_k = \|Q_{y_k} A_k^* + P_{y_k} B_k^*\|_{X_k} .$$

Since $\|A_k^*\|_I \leq n$ and $\|B_k^*\|_I \leq m + 1$, we may further extract subsequences and relabel so that $A_k^* \to A^*$ and $B_k^* \to B^*$ uniformly on I. Then A^* and B^* are polynomials, and $\partial A^* \leq n - 1$ and $\partial B^* \leq m$. Moreover, the maximum of the absolute values of the coefficients of A^* and B^* is 1. Thus

$$H - H_k \leq \|QA^* + PB^*\|_I - \|Q_{y_k} A_k^* + P_{y_k} B_k^*\|_{X_k} .$$

By Lemma 5.5.2, $\lim_{k\to\infty} \|Q_{y_k} A_k^* + P_{y_k} B_k^*\|_{X_k} = \|QA^* + PB^*\|_I$. Hence, $\underline{\lim}_{k\to\infty} H_k \geq H$. Since $H > 0$, a further relabeling can be done and we assume that $H_k \geq d_1$ for all k where $d_1 > 0$ is independent of k.

Now fix k and let $\alpha = \overline{p}_n^k / p_n$. Define

$$A_k(x) = P_k(x) - \alpha P_{y_k}(x) = \sum_{i=0}^{n-1} a_i x^i ,$$

$a_i = \overline{p}_i^k - \alpha p_i^k$, $i = 0,\ldots,n-1$, and

$$B_k(x) = Q_k(x) - \alpha Q_{y_k}(x) = \sum_{j=0}^{m} b_j x^j ,$$

$b_j = \overline{q}_j^k - \alpha q_j^k$. Then

$$\begin{aligned}
&\| Q_{y_k} A_k - P_{y_k} B_k \|_{X_k} \\
&\qquad = \| Q_{y_k} P_k - \alpha Q_{y_k} P_{y_k} - P_{y_k} Q_k + \alpha Q_{y_k} P_{y_k} \|_{X_k} \\
&\qquad = \| Q_{y_k} P_k - P_{y_k} Q_k \|_{X_k} \\
&\qquad \leq \| Q_{y_k} \|_{X_k} \| Q_k \|_{X_k} \| R_k - R_{y_k} \|_{X_k} \\
&\qquad \leq (m+1)^2 \| R_k - R_{y_k} \|_{X_k} .
\end{aligned}$$

A homogeneity argument yields

$$\| Q_{y_k} A_k - P_{y_k} B_k \|_{X_k} \geq d_1 \max_{i,j} \{ |a_i|, |b_j| \} .$$

Thus

$$|a_i| \leq d_2 \rho ,$$

$i = 0,\ldots,n-1$, and

$$|b_j| \le d_2\rho,$$

$j = 0,\ldots,m$, where $d_2 = (m+1)^2/d_1$ and $\rho = \|R_k - R_{y_k}\|_{X_k}$.

Recall that $b_j = \bar{q}_j^k - \alpha q_j^k$. Choose j where $|\bar{q}_j^k| = 1$. Then

$$1 = |b_j + \alpha q_j^k| \le |b_j| + |\alpha| \le d_2\rho + |\alpha|.$$

Thus $1 - |\alpha| \le d_2\rho$. Now consider j where $|q_j^k| = 1$. We have

$$|\alpha| = |\alpha q_j^k| = |\bar{q}_j^k - b_j| \le 1 + |b_j| \le 1 + d_2\rho.$$

Thus $1 - |\alpha| \ge - d_2\rho$. Thus if $\alpha \ge 0$, then $|1 - \alpha| \le d_2\rho$. In case $\alpha < 0$, we have $1 - \alpha > 0 > -d_2\rho$. For arbitrary $x \in X$,

$$0 < Q_k(x) \le B_k(x) + \alpha Q_{y_k}(x) \le (m+1)d_2\rho + \alpha Q_{y_k}(x).$$

Thus

$$\alpha \ge -(m+1)d_2\rho/Q_{y_k}(x).$$

By Lemma 5.5.1, there is an $\eta > 0$ such that $Q_{y_k}(x) \ge \eta$ for all $x \in X$ and all sufficiently large k. If k is sufficiently large,

$$\alpha \geq -(m+1)d_2\rho/\eta .$$

Thus

$$\begin{aligned} 1 - \alpha &= 1 - |\alpha| - 2\alpha \\ &\leq d_2\rho + 2(m+1)d_2\rho/\eta \\ &= d_2 \, [1 + 2(m+1)/\eta]\rho \\ &= d_3\rho \, . \end{aligned}$$

In any case, $|1 - \alpha| \leq d_3\rho = d_3 \, \| R_k - R_{y_k} \|_{X_k}$.

We now have

$$\begin{aligned} \| Q_k - Q_{y_k} \|_{X_k} &= \| B_k + (1-\alpha) Q_{y_k} \|_{X_k} \\ &\leq \| B_k \|_{X_k} + |1 - \alpha| \, \| Q_{y_k} \|_{X_k} \\ &\leq (m+1)(d_2+d_3) \, \| R_k - R_{y_k} \|_{X_k} \end{aligned}$$

where d_2 and d_3 do not depend on k. This contradicts (5.36), and Lemma 5.5.3 is proven.

We now establish the "uniform quadratic convergence" of the ODC algorithm. With a few minor adjustments, the following proof is that Barrodale, Powell, and Roberts [1]. The constant A in (5.33) is written in terms of the constants of Theorem 5.3.1, Lemma 5.5.1, and Lemma 5.5.3 which have been shown to be independent of $y \in J$ and sufficiently dense $X \subseteq I$.

Proof of Theorem 5.4.1. We choose $\delta > 0$, $\gamma > 0$, $\eta > 0$, and $\theta > 0$, via Theorem 5.3.1, Lemma 5.5.1, and Lemma 5.5.3, such that whenever X is a closed subset of I and $d_I(X) < \delta$, the following properties are satisfied.

(i) For each $y \in J$ and all $R \in \mathbf{R}_X(n,m)$

$$\|F_y - R\|_X \geq \|F_y - R(C_X(y),\cdot)\|_X + \gamma \|R - R(C_X(y),\cdot)\|_X.$$

(ii) For any $y \in J$, if $R(C_X(y),\cdot) = P_y/Q_y$ where $P_y(x) = \sum_{i=0}^{n} p_i(y)x^i$, $Q_y(x) = \sum_{j=0}^{m} q_j(y)x^j$, and $\max_j |q_j(y)| = 1$, then $Q_y(x) \geq \eta$ for all $x \in I$.

(iii) If P_y and Q_y are given as in Property (ii),

$$\|Q - Q_y\|_X \leq \theta \|R - R(C_X(y),\cdot)\|_X$$

for all $R = P/Q$ where $P(x) = \sum_{i=0}^{n} p_i' x^i$,

$Q(x) = \sum_{j=0}^{m} q_j' x^j > 0$ on X, and $\max_j |q_j'| = 1$.

We now fix $y \in J$ and show that the constant A of (5.33) depends only on γ, η, and θ. We let $R_k = P_k/Q_k$ denote the k-th iterate of the ODC algorithm where Q_k is constrained as in Step 2 of the ODC algorithm. For convenience, we denote

$$\Delta = \Delta_X(y) = \|F_y - R(C_X(y),\cdot)\|_X$$

and

$$\Delta_k = \Delta_k(y) = \|F_y - R_k\|_X .$$

From Properties (i) and (iii), we see that

$$\|Q_k - Q_y\|_X \leq \theta \, \|R_k - R(C_X(y),\cdot)\|_X$$
$$\leq \theta(\Delta_k - \Delta)/\gamma .$$

For arbitrary $x \in X$,

$$\frac{Q_y(x)}{Q_k(x)} = \frac{Q_y(x)}{Q_y(x) + [Q_k(x) - Q_y(x)]}$$
$$\geq \frac{Q_y(x)}{Q_y(x) + \theta(\Delta_k - \Delta)/\gamma}$$
$$\geq \frac{\eta}{\eta + \theta(\Delta_k - \Delta)/\gamma}$$

since $s/(s+a)$ is an increasing function of $s > 0$ when $a > 0$ and $Q_y(x) \geq \eta$. Thus

$$\min_{x \in X} \left\{ \frac{Q_y(x)}{Q_k(x)} \right\} \geq \frac{\eta}{\eta + \theta(\Delta_k - \Delta)/\gamma} .$$

Again for any $x \in X$,

$$\frac{Q_k(x)}{Q_{k+1}(x)} = \frac{Q_k(x)}{Q_y(x)} \; \frac{Q_y(x)}{Q_{k+1}(x)}$$

$$\geq \frac{Q_k(x)}{Q_y(x)} \; \frac{\eta}{\eta + \theta(\Delta_{k+1} - \Delta)/\gamma} .$$

Now Theorem 1 of [1] ensures that $\Delta_{k+1} \leq \Delta_k$. Thus

$$\frac{Q_k(x)}{Q_{k+1}(x)} \geq \frac{Q_y(x) + [Q_k(x) - Q_y(x)]}{Q_y(x)} \cdot \frac{\eta}{\eta + \theta(\Delta_k - \Delta)/\gamma}$$

$$\geq \frac{Q_y(x) - \theta(\Delta_k - \Delta)/\gamma}{Q_y(x)} \cdot \frac{\eta}{\eta + \theta(\Delta_k - \Delta)/\gamma}$$

$$\geq \frac{\eta - \theta(\Delta_k - \Delta)/\gamma}{\eta + \theta(\Delta_k - \Delta)/\gamma}$$

since $(s - a)/s$ is an increasing function of $s > 0$ when $a > 0$ and $Q_y(x) \geq \eta$. Hence,

$$\min_{x\in X}\left\{\frac{Q_k(x)}{Q_{k+1}(x)}\right\} \geq \frac{\eta - \theta(\Delta_k-\Delta)/\gamma}{\eta + \theta(\Delta_k-\Delta)/\gamma} .$$

Now by Step 2 of the ODC algorithm

$$\min_{x\in X}\left\{\frac{|F_y Q_{k+1} - P_{k+1}| - \Delta_k Q_{k+1}}{Q_k}\right\}$$

$$\leq \max_{x\in X}\left\{\frac{|F_y Q_y - P_y| - \Delta_k Q_y}{Q_k}\right\}$$

$$= \max_{x\in X}\{[|F_y - R_y| - \Delta_k]\, Q_y/Q_k\}$$

$$\leq \max_{x\in X}\{(\Delta - \Delta_k)\, Q_y/Q_k\}$$

$$= -(\Delta_k - \Delta)\min_{x\in X}\left\{\frac{Q_y}{Q_k}\right\} ,$$

where $R_y = R(C_X(y),\cdot) = P_y/Q_y$. Thus for any $x \in X$, Step 2 of the ODC algorithm implies that

$$|F_y(x) - R_{k+1}(x)| - \Delta_k \leq -(\Delta_k-\Delta)\min_{x\in X}\left\{\frac{Q_y}{Q_k}\right\}\cdot\left\{\frac{Q_k(x)}{Q_{k+1}(x)}\right\}$$

$$\leq -(\Delta_k-\Delta)\min_{x\in X}\left\{\frac{Q_y}{Q_k}\right\}\min_{x\in X}\left\{\frac{Q_k}{Q_{k+1}}\right\} .$$

Thus

$$\Delta_{k+1} - \Delta_k \leq \frac{-\eta^2(\Delta_k-\Delta) + \eta\theta(\Delta_k-\Delta)^2/\gamma}{[\eta + \theta(\Delta_k-\Delta)/\gamma]^2} .$$

Finally, applying this bound we get

$$\Delta_{k+1} - \Delta = \Delta_k - \Delta + \Delta_{k+1} - \Delta_k$$

$$\leq (\Delta_k-\Delta)^2 \left\{ \frac{3\eta\theta/\gamma + \theta^2(\Delta_k-\Delta)/\gamma^2}{[\eta + \theta(\Delta_k-\Delta)/\gamma]^2} \right\}$$

$$\leq \frac{4\theta}{\eta\gamma} (\Delta_k-\Delta)^2 .$$

Thus Theorem 5.4.1 is proven.

5.6 Computation.

In this section, we give the results of several runs of a program to compute rational product approximations of the form

$$(5.37)\qquad \frac{\sum_{i=0}^{n}\left(\sum_{k=0}^{r} a_{ik}y^{k} \Big/ \sum_{\ell=0}^{s} b_{i\ell}y^{\ell}\right)x^{i}}{\sum_{j=0}^{n}\left(\sum_{k=0}^{r} c_{jk}y^{k} \Big/ \sum_{\ell=0}^{s} d_{j\ell}y^{\ell}\right)x^{j}} .$$

We compare the computer processor unit (CPU) time in minutes of Algorithm 1 with that of Algorithm 2. In addition, the uniform error of approximation $\|F - \mathcal{R}\|_{D}$, where $\mathcal{R}$ is given by (5.37), is compared to the error before the coefficient approximation $\max_{y\in J} \|F_y - R(C(y),\cdot)\|_{I}$.

A FORTRAN IV program, RPA, was written to compute rational product approximations using Algorithm 1 or Algorithm 2. The best univariate approximations are computed by a modification of the ODC-package of Kaufman and Taylor [18]. This modification incorporates the use of an optional initial guess for the ODC algorithm in place of the initialization (5.32). The decretization of the intervals $I = J = [-1,1]$ is done by choosing

$X = Y = \{-1.0,-.9,...,.9,1.0\}$. We refer to the approximation of the F_y, $y \in Y$, as Stage 1 of either algorithm; comparisons are made in the computing time for both Stage 1 and the complete algorithm (not including the error check). A continuity check for the coefficient vector $C(y)$ similar to that of J. Henry [12,13] is incorporated into RPA. If $C(y)$ is deemed discontinuous, the program is terminated. We further denote

$$\mathrm{ERRB} = \max_{y \in Y} \| F_y - R(C(y),\cdot)\|_I$$

and

$$\mathrm{ERRA} = \| F - \mathcal{R}\|_D.$$

The uniform error ERRA is computed over the set $X' \times Y'$ where $X' = Y' = \{-1.00,-.99,...,.99,1.00\}$. All examples were executed on the Zerox Sigma 7 computer at Montana State University using double precision arithmetic.

In all the examples, the rational product approximations were the same using Algorithm 1 or Algorithm 2. Table 1 gives the CPU times for Stage 1 and the complete algorithm for both algorithms in a variety of examples. Table 2 gives the corresponding errors ERRB and ERRA. In all cases, the approximation of F_y were done with $n = m = 2$.

TABLE 1

Function	r	s	CPU Stage 1 Alg. 1	CPU Stage 1 Alg. 2	CPU Alg. 1	CPU Alg. 2
e^{x+y}	2	0	.6496	.6134	.7077	.6641
	4	0			.7471	.7164
	2	2			.8780	.8415
e^{-x^2-y}	2	0	.7220	.3586	.7740	.4102
	4	0			.8306	.4598
	2	2			.9612	.6555
$e^{y^2-x^2}$	2	0	.7114	.4399	.7614	.4869
	4	0			.8125	.5386
	2	2			.9477	.7377
$\sin(x+y)$	2	0	.6995	.7367	.7509	.7873
	4	0			.7996	.8374
	2	2			.8724	.9090
$\sin[(x^2+y^2)/5]$	2	0	.5734	.6097	.6236	.6577
	4	0			.6777	.7096
	2	2			.7362	.8391

TABLE 2

Function	r	s	ERRB	ERRA
e^{x+y}	2	0	2.546×10^{-4}	1.228×10^{-1}
	4	0		1.703×10^{-3}
	2	2		4.608×10^{-4}
e^{-x^2-y}	2	0	4.188×10^{-3}	4.802×10^{-2}
	4	0		4.808×10^{-3}
	2	2		4.432×10^{-3}
$e^{y^2-x^2}$	2	0	5.167×10^{-3}	1.098×10^{-1}
	4	0		1.240×10^{-2}
	2	2		8.441×10^{-3}
$\sin(x+y)$	2	0	1.237×10^{-3}	6.877×10^{-2}
	4	0		3.726×10^{-3}
	2	2		1.124×10^{-2}
$\sin[(x^2+y^2)/5]$	2	0	4.445×10^{-5}	3.743×10^{-4}
	4	0		4.867×10^{-5}
	2	2		4.885×10^{-5}

The results of Table 1 indicate that Algorithm 2 is competitive with Algorithm 1 and in some cases the CPU time for Stage 1 of Algorithm 2 is considerably less than that for Algorithm 1. In most cases, however, the CPU times for Stage 1 and the complete Algorithm 1 are nearly the same for both algorithms. For the function $F(x,y) = \sin(x^2 + y)$ a discontinuity in $C(y)$ occured at $y \doteq 1.8$ for both algorithms. Table 2 indicates that in most cases, ERRA is of the same magnitude as ERRB when the coefficient approximation is performed by polynomials of degree 4 or less or by rational functions of "degree (2,2)."

The computational results, although inconclusive, suggest that Algorithm 2 is generally no worse than Algorithm 1 and in some cases is better.

CHAPTER VI

CONCLUSIONS

In this concluding chapter, we summarize the results of this work and indicate some directions for further research in multivariate approximation theory. The papers of J. Henry [12,13] indicate that rational product approximations are computationally competitive with multivariate surface approximation of Kaufman and Taylor [19]. The computation section of Chapter V suggests that some improvements can be made in the existing algorithms for product approximations. Furthermore, it is shown in Chapter V that the rational product approximation algorithms using the ODC algorithm possess a uniform quadratic rate of convergence.

In addition to the computational considerations, this work sets the framework for a theory of product approximations. In the polynomial setting, a degree of approximation theorem is established extending the work of S. E. Weinstein. In both the linear and the rational product approximation settings, continuity of the associated product approximation operators is guaranteed, and under further conditions, these operators satisfy the stronger point-Lipschitz continuity. Due to the loss of uniqueness of best multivariate approximations, results of this type are not possible in the Tchebycheff approximation setting. Finally, in the rational setting, discretization steps in the existing algorithms have been justified.

The computational concerns of continuity and discretization involve purturbations of some ingredients of the approximation problem, namely, of the function to be approximated and the underlying space D. The effect of perturbing the basis functions in computing product approximations has not been studied. A further direction in this area may lie in generalizing the concept of product approximation to abstract settings (for example, tensor products of Banach spaces). Badly needed are a unification and efficiency improvements of methods to compute product approximations and comparisons of methods of multivariate approximation.

BIBLIOGRAPHY

1. Barrodale, J., M. J. D. Powell and F. D. K. Roberts. "The Differential Correction Algorithm for Rational ℓ_∞-Approximation". SIAM J. Numer. Anal. 9(1972):493-504.

2. Bartelt, Martin. "On Lipschitz Conditions, Strong Unicity and a Theorem of A. K. Cline". J. Approximation Theory 14(1975):245-250.

3. Brown, J. A. and M. S. Henry. "Best Rational Product Approximations of Functions". J. Approximation Theory 9(1973):287-294.

4. Brown, J. A. and M. S. Henry, "Best Chebyshev Composite Approximation". SIAM J. Numer. Anal. 12(1975): 336-344.

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

6. Cheney, E. W. and H. L. Loeb. "Two New Algorithms for Rational Approximation". Numer. Math. 3(1961): 72-75.

7. Cline, A. K. "Lipschitz Conditions on Uniform Approximation Operators". J. Approximation Theory. 8(1973):160-172

8. Dunham, Charles B. "Alternating Chebyshev Approximation". Trans. AMS 178(1973):95-109.

9. Dunham, Charles B. "Alternating Chebyshev Approximation with a Non-Continuous Weight Function". Can. Math. Bull. To appear.

10. Freud, G. "Eine Ungleichung für Tschebyscheffsche Approximationspolynome". Acta. Sci. Math. (Szeged) 19(1958):162-164.

11. Haar, A. "Die Minkowskische Geometrie und die Annäherung an Stetige Functionen". Math. Ann. 78(1918):294-311.

12. Henry, J. N. "Computation of Rational Product Approximations". Int. J. of Num. Meth. Engng. To appear.

13. Henry, J. N. "Comparison of Methods for Multivariate Rational Approximation". Submitted for publication.

14. Henry, M. S. and J. A. Roulier. "Uniform Lipschitz Constants on Small Intervals". Submitted for publication.

15. Henry, M. S. and D. Schmidt. "Continuity Theorems for the Product Approximation Operator". Alan G. Law and Badri N. Sahney, editors, Theory of Approximation with Applications. Academic Press, New York. 1976:24-42.

16. Henry, M. S. and D. Schmidt. "Continuity Theorems for the Rational Product Approximation Operator". J. Approximation Theory. To appear.

17. Henry, M. S. and S. E. Weinstein. "Best Rational Product Approximations of Functions II". J. Approximation Theory 12(1974):6-22.

18. Kaufman, E. H. and G. D. Taylor. "An Application of Linear Programming to Rational Approximation". Rocky Mountain J. Math. 4(1974):371-373.

19. Kaufman, E. H. and G. D. Taylor. "Uniform Rational Approximation of Functions of Several Variables". Int. J. Num. Meth. Engng. 9(1975):297-323.

20. Maehly, H. and Ch. Witzgall. "Tschebyscheff-Approximationen in Kleinen Intervallen I: Approximation durch Polynome". Numer. Math. 2(1960):142-150)

21. Maehly, H. and Ch. Witzgall. "Tschebyscheff-Approximationen in Kleinen Intervallen II: Stetigkeitssätze für Gebrochen rationale Approximationen". Numer. Math. 2(1960):293-307.

22. Mairhuber, John C. "On Haar's Theorem Concerning Chebychev Approximation Problems Having Unique Solutions". Proc. Am. Math. Soc. 7(1956): 609-615.

23. Newman, D. J. and Harold S. Shapiro. "Some Theorems on Čhebyšev Approximation". Duke Math. J. 30(1963): 673-682.

24. Newman, M. H. A. Elements of the Topology of Plane Sets of Points. Cambridge University Press, Cambridge. 1954.

25. Phelps, R. R. "Uniqueness of Hahn-Banach Extensions and Unique Best Approximations". Trans. Amer. Math. Soc. 95(1960): 238-255.

26. Poreda, S. J. "Counter Examples for Certain Questions Dealing with the Approximation to Continuous Functions by Polynomials". Submitted for publication.

27. Rice, John R. "Tchebycheff Approximations by Functions Unisolvent of Variable Degree". Trans. Amer. Math. Soc. 99(1961): 298-302.

28. Rice, John R. The Approximation of Functions. Vol. 1. Addison Wesley, Reading, MA. 1964.

29. Rice, John R. The Approximation of Functions. Vol. 2. Addison Wesley, Reading, MA. 1969.

30. Weinstein, Stanley E. "Approximation of Functions of Several Variables: Product Chebychev Approximations I". J. Approximation Theory. 2(1969): 433-447.

31. Weinstein, Stanley E. "Product Approximations of Functions of Several Variables". SIAM J. Numer. Anal. 8(1971): 178-189.

32. Werner, H. "Die Bedeutung der Normalität bei Rationaler Tschebyscheff-Approximation". Computing (Arch. Elektron. Rechnen). 2(1967): 34-52.

DATE	ISSUED TO