BEST APPROXIMATION IN $L_1[T, \Sigma, \mu]$

A Thesis Submitted

to

LAKEHEAD UNIVERSITY

In Partial Fulfillment of the Requirement

for the

DEGREE

of

MASTER OF SCIENCE

by

Patrick J. Mann

1980

TABLE OF CONTENTS

Chapter I: PRELIMINARIES

1. Introduction 1
2. Preliminary Concepts 3
3. Existence 6

Chapter II: CHARACTERIZATION OF BEST APPROXIMATION

1. Introduction 9
2. First Characterization Theorem 9
3. Second Characterization Theorem 14
4. Characterization in L_1 and C_1 18
5. Best One-Sided Approximation 36
6. Approximation with Restricted Range and Other Constraints 43

Chapter III: UNIQUENESS OF BEST APPROXIMATIONS

1. Introduction 46
2. Uniqueness of a Best Approximation 46
3. Chebyshev Sets 50
4. Chebyshev Subspaces in L_1 53
5. Finite Dimensional Chebyshev Subspaces in L_1^R 57
6. Chebyshev Sets in $C_1[T, \mu]$ 59
7. Convex Chebyshev Sets in L_1 or C_1 64

Chapter IV: STRONG UNIQUENESS

1. Introduction 71
2. Characterization of Strong Uniqueness 74
3. Strong Uniqueness in L_1 81
4. Strongly Chebyshev Subspaces 91

Chapter V: CONCLUSION 95

Bibliography 98

ABSTRACT

Let X be a normed linear space and K a convex set in X. Then $\pi \in K$ is a best approximation to $f \in X \sim K$ if $||f - \pi|| = \inf\{||f - k||: k \in K\}$. The existence of such a best approximation is shown if K is compact, or closed and bounded in a finite dimensional space. Two characterizations of best approximation are proved using a geometrical approach involving functionals in the dual space. These are applied to the space $L_1(T, \sum, \mu)$ under the assumption that its dual is equivalent to $L_\infty(T, \sum, \mu)$ to recover results of Kripke and Rivlin, and Singer.

The same approach is used to derive criterion for the uniqueness of best approximations, and then applied to L_1 to obtain, among others, Jackson's classic theorem on approximation to continuous functions from Haar Subspaces. A result of Phelp's on the non-existence of finite dimensional Chebyshev subspaces in non-atomic L_1^R is also shown.

The concept of strong unicity is presented, and investigated by looking at particular supporting cones. A useful characterization is proposed in L_1 and then applied to prove Wulbert's theorem on strongly Chebyshev subspaces in L_1. It is shown that no Haar subspace is strongly Chebyshev in C_1, and an example of an infinite dimensional strongly Chebyshev subspace is given.

Acknowledgement

I would like to thank Dr. Peter Mah, my supervisor, for the help and encouragement he gave me, and also for the inordinate amount of his time that he spent working with me.

Chapter I

PRELIMINARIES

S1. INTRODUCTION

With the advent of powerful computing machines the idea of approximating functions by other functions parametrized by a finite number of variables has become very important. In practice one finds an approximation and leaves it at that, but the aim is to find the best possible approximation from the set of approximating functions. Obviously we need some sort of measure of how good an approximation we have. In this thesis we will concentrate on the use of the L_1 norm for this purpose. To do this we will use some powerful geometric ideas which can be stated and proved in any normed linear space, so we will first define our best approximations in a rather general way.

Let X be some normed linear space, with the norm $||\cdot||$, and let K be some set in X. If $f \in X$ we say that π is a best approximation to f from K if $||f - \pi|| = \inf\{||f - k||: k \in K\}$ and $\pi \in K$.

Our first task will be to investigate the existence of such a best approximation. If K is compact they exist for any choice of f, but in general existence depends on K, where f is, and the type of norm involved.

If we have an element of K, are there ways of testing to see if it is a best approximation? We can uncover such characterizations by considering the existence and properties of various hyperplanes which separate K and f . (We assume f is outside K .) This approach is very intuitive and easy to understand, and most of the major theorems of approximation in the L_1 norm can be recovered. Since these hyperplanes are defined by functionals in the dual space we have restricted ourselves to a measure μ which ensures that the dual of our space $L_1(T, \sum, \mu)$ is equivalent to the space $L_\infty(T, \sum, \mu)$. This is satisfied, for example, if μ is σ-finite.

The next chapter discusses the question of uniqueness. Again using our geometrical concepts we recover results of Cheney and Wulbert, Jackson, and Singer. In particular we characterize sets which allow only unique best approximations (called Chebyshev sets). In L_1 we show that there are no such Chebyshev subspaces when μ is non-atomic, but do better with continuous functions by proving Jackson's famous theorem on Haar subspaces.

We are not quite finished yet, and go on to introduce strongly unique best approximations. The results of Bartelt and McLaughlin are used in L_1 and some theorems derived. One very nice theorem proves that strict inequality in one of our characterization theorems in a finite dimensional subspace is necessary and sufficient for the best approximation to be strongly unique. A not so nice result states that a Haar subspace is not necessarily a strongly Chebyshev subspace, even though it is Chebyshev.

We will start off by giving a few necessary definitions in the next section.

S2. PRELIMINARY CONCEPTS

In general, the notation used is that of Dunford and Schwartz [16].

Many of the definitions will be introduced as needed, and some basic ideas from topology are assumed. These include open, closed sets and the convergence of sequences (see, for instance, [31]).

For convenience, $\sup\{f(x): x \in X\} = \sup f[X]$ and similarly for the infimum.

2.1. Definition. X is a linear space over a scalar field Φ if X is an additive group under the binary operation $+$ on X together with an operation $m: \Phi \times X \to X$ written as $m(a, x) = ax$ satisfying

$$a(x + y) = ax + ay \qquad x, y \in X \qquad a \in \Phi ;$$

$$(a + b)x = ax + bx \qquad x \in X \qquad a, b \in \Phi ;$$

$$a(bx) = (ab)x \qquad x \in X \qquad a, b \in \Phi ;$$

$$1 \cdot x = x \qquad x \in X .$$

The scalar field Φ will always be the set $\mathbb{C}$ of complex numbers, or the set $\mathbb{R}$ of real numbers.

2.2. Definition. X is a normed linear space over a field Φ if X is a linear space and for each $x \in X$ there corresponds a real number $||x||$, called the norm of x, satisfying

$$||x|| = 0 \text{ if and only if } x = 0 ;$$

$$||x|| \geq 0 ;$$

$$||\lambda x|| = |\lambda| \; ||x|| ;$$

$$||x + y|| \leq ||x|| + ||y|| \quad \text{for } x, y \in X .$$

We let $B(y, \varepsilon) = \{x \in X: ||x - y|| < \varepsilon\}$ and $\overline{B}(y, \varepsilon) = \{x \in X, ||x - y|| \leq \varepsilon\}$ be respectively, the open and closed balls of radius $\varepsilon > 0$ centred at $y \in X$. We let $B^*(y, \varepsilon)$ denote the corresponding open ball in the dual space X^*. Unless otherwise mentioned the topology on X is the norm topology which has all such open balls as a basis.

2.3. Definition. M is a linear subspace of a linear space X if $M \subseteq X$ and M is a linear space over the same field of scalars.

2.4. Definition. A set $K \in X$ is convex, if for all $\lambda \in \mathbb{R}$, $0 \leq \lambda \leq 1$, (the scalar field Φ is $\mathbb{R}$ or $\mathbb{C}$) and $x, y \in K$, $\lambda x + (1 - \lambda)y \in K$.

2.5. Definition. Let X and Y be normed linear spaces over the same field Φ, and L a function mapping X into Y. Then L is linear if $L(x + y) = L(x) + L(y)$ and $L(ax) = aL(x)$ for $x, y \in X, a \in \Phi$. L is bounded if there exists a real number $M \geq 0$ such that $||L(x)|| \leq M||x||$ for all $x \in X$. The set of all bounded linear functions is itself a linear space over the same field. If Y is the scalar field then this set is termed the dual space, X^*, of X. Where $\Phi = \mathbb{C}$ we define $\operatorname{Re} L$ as the real part of L, and note that $\operatorname{Re} L$ is itself a member of X^*. X^* is a Banach space under the norm, $||L|| = \sup\{|L(x)|: ||x|| = 1, x \in X\}$. The boundedness of a linear functional is equivalent to its continuity. For a complete discussion of linear functionals and the dual space refer to [16, Chapter II].

2.6. Definition. Let D be a subset of a normed linear space X, and let $f \in X$. Then π is a best approximation from D to f if $\pi \in D$ and

$$||f - \pi|| = \inf\{||f - d||: d \in D\} .$$

The infimum is termed the distance, $\rho(f, D)$, from f to D. The set of all such best approximations is denoted by $P(D, f)$.

2.7. Definition. Let $(T, \sum, \mu)$ be a measure space, and let f and g be complex-valued functions on T. Define an equivalence relation "$\sim$" by $f \sim g$ if $\mu\{t \in T: f(t) \neq g(t)\} = 0$ and let $[f]$ be the equivalence class which f belongs to. Then $L_1(T, \sum, \mu) = \{[f]: \int_T |f| d\mu < \infty\}$, with the norm $||f|| = \int_T |f| d\mu$. If only the real numbers are being considered we write $L_1^R(T, \sum, \mu)$, and we will suppress the $(T, \sum, \mu)$ wherever it is possible without ambiguity. The definition is abused somewhat (harmlessly) by writing $f \in L_1$ and not $[f] \in L_1$.

Let $L_\infty(T, \sum, \mu) = \{|f|: \text{ess sup } f[T] < \infty\}$ with the norm $||f||_\infty = \text{ess sup } f[T]$. The above comments can be repeated, with the added note that, unless otherwise stated, we assume L_1^* is equivalent to L_∞, as is the case when μ is σ-finite.

2.8. Definition. Suppose for the measure space $(T, \sum, \mu)$ a topology is defined on T so that (i) T is a Tychonoff space; (ii) every open set is in $\sum$ and every non-empty open set has positive measure; and (iii) every singleton is of finite measure. We denote by $C_1(T, \sum, \mu)$

(or simply C_1 or $C_1(T)$ when the measure space is clear) to be the subspace of $L_1(T, \sum, \mu)$ consisting of continuous functions with the L_1 norm. The symbols $C_1^R(T, \sum, \mu)$, C_1^R or $C_1^R(T)$ are used when only real-valued functions are considered. We note that C_1 is not complete.

In references to previous theorems we will write, for example, see theorem 2.4 if the theorem is in the same chapter, or theorem III-2.4 if the theorem is in another chapter (chapter III in this example).

S3. EXISTENCE

Best approximations from closed subsets exist in finite dimensional spaces, but in the general case this is not true. For some special spaces (uniformly or strictly convex) better results can be obtained, but the L_1 case satisfies none of these conditions, so they are not considered here. It is not necessary to postulate convexity; any closed bounded subset in a finite dimensional subspace will do.

More generally closure is not sufficient. The set must be compact to ensure the existence of best approximations. In a finite dimensional space closed and bounded imply compactness so we see that compactness implies the previous paragraph's comments.

3.1. Theorem. Let C be a compact set in a metric space X, with metric d. To each point $f \in X \sim C$ there exists a point π in C with $d(f, \pi) = \inf\{d(f, x): x \in C\}$.

Proof. Let $\delta = \inf\{d(f, x): x \in C\}$. By the definition of the infimum, there exists a sequence of points $\{x_n\} \subseteq C$ such that $\lim_n d(f, x_n) = \delta$. Since C is compact there exists a subsequence $\{y_m\}$ of the sequence $\{x_n\}$ which converges to a point, call it π, of C. By the triangle inequality, $d(f, \pi) \leq d(f, y_n) + d(y_n, \pi)$ for all n. Since the left side is independent of n and the right side $\to \delta$ as $n \to \infty$, $d(f, \pi) \leq \delta$. Since $\pi \in C$, $d(f, \pi) \geq \delta$. Therefore $d(f, \pi) = \delta$ and π is the required point. □

In a normed linear space define, as usual, $d(x, y) = ||x - y||$. Then the above theorem can be quickly applied to yield the following corollary.

3.2. Corollary. For every closed set C in a finite dimensional subspace M of a normed linear space X, the set $P(C, f)$, for any $f \in X \sim C$ is non-empty.

Proof. Choose $y \in C$ arbitrarily. Consider the set $H = \{x \in C: ||f - x|| \leq ||f - y||\}$. Since C is closed this set is also closed, and it is certainly bounded. Since M is finite dimensional H must therefore be compact in X [7, p. 10], and by theorem 3.1 there exists a point $\pi \in H$ such that $||f - \pi|| = \inf\{||f - x||: x \in H\}$. In other words $\pi \in P(H, f)$; but $P(H, f) \subseteq P(C, f)$, and therefore $P(C, f)$ is not empty. □

Cheney gives an example [7, p. 21] showing that finite dimensionality cannot be omitted.

In the rest of this thesis it is assumed that best approximations do indeed exist. However, in the actual computation it might be wise to check to see if there is something to compute.

We will consider best approximation from convex sets. For discussion on general best non-linear approximation see Dierieck [13].

Chapter II

CHARACTERIZATION OF BEST APPROXIMATION

S1. INTRODUCTION

There are two important characterization theorems for best approximation to convex sets. These will be proved and then applied in L_1 and C_1 . Also, specializations to linear subspaces of infinite and finite dimension will be considered.

S2. FIRST CHARACTERIZATION THEOREM

The first characterization theorem is due to Deutsch and Maserick [9, thm. 2.5], valid for a normed linear space, and is a consequence of the Hahn-Banach theorem.

2.1. Definition. Let X be a normed linear space over $\mathbb{C}$ with norm $||\cdot||$. A hyperplane $[L, c]$ is a set of the form, for some $L \in X^* \sim \{0\}$ and $c \in \mathbb{R}$,

$$[L, c] = \{x \in X: \operatorname{Re}L(x) = c\} .$$

2.2. Lemma. Let $L \in X^* \sim \{0\}$, $c \in \mathbb{R}$, and $H = [L, c]$. Then for each $x \in X$,

$$\rho(x, H) = \frac{|\operatorname{Re}L(x) - c|}{||L||} . \qquad (2.2.1)$$

<u>Proof</u>. Since $L \in X^*$, for all $y \in H$, $\frac{|\mathrm{ReL}(x) - c|}{||L||} = \frac{|\mathrm{ReL}(x - y)|}{||L||} \leq$ $\leq ||x - y||$. Therefore $\rho(x, H) \geq \frac{|\mathrm{ReL}(x) - c|}{||L||}$.

If $0 < \varepsilon < ||L||$, then since $||\mathrm{ReL}|| = ||L||$, there exists $z \in X$ such that $|\mathrm{ReL}(z)| > (||L|| - \varepsilon)||z||$, and we see that $|\mathrm{ReL}(z)|\left|\frac{\mathrm{ReL}(x) - c}{\mathrm{ReL}(z)}\right| > (||L|| - \varepsilon)\left|\frac{\mathrm{ReL}(x) - c}{\mathrm{ReL}(z)}\right| ||z||$. Let $y = x - \frac{\mathrm{ReL}(x) - c}{\mathrm{ReL}(z)} z$. Then $\mathrm{ReL}(y) = c$ and $z = \frac{\mathrm{ReL}(z)}{\mathrm{ReL}(x) - c}(x - y)$. Thus $|\mathrm{ReL}(x) - c| > (||L|| - \varepsilon)||x - y||$. Since this is true for all $\varepsilon \in (0, ||L||)$, there exists y such that $\mathrm{ReL}(y) = c$ and $|\mathrm{ReL}(x) - c| > ||L||\ ||x - y||$. Therefore $y \in H$ and $\frac{|\mathrm{ReL}(x) - c|}{||L||} \geq ||x - y|| \geq \rho(x, H)$. □

2.3. <u>Definition</u>. A hyperplane $H = [L, c]$ is said to <u>separate</u> two subsets M and N of X if $\sup \mathrm{ReL}[M] \leq c \leq \inf \mathrm{ReL}[N]$. It can be shown that if H separates a point x from a set M then $\rho(x, M) \geq \rho(x, H)$, and any neighbourhood of a point in H contains points x_1 and x_2 such that $\mathrm{ReL}(x_1) < c < \mathrm{ReL}(x_2)$.

The next theorem is the geometric form of the Hahn-Banach theorem. A functional $L \in X^*$ separates M and N if there exists $c \in \mathbb{R}$ such that $[L, c]$ separates M and N.

2.4. <u>Theorem</u>. Let M and N be two disjoint convex subsets of a normed linear space X, and suppose one of them has an interior point. Then there exists $L \in X^*$, $L \neq 0$ which separates M and N.

Proof. [19, p. 63]. □

If L separates M and N then there exists c such that $[L, c]$ separates M and N. But then $\sup \frac{\mathrm{Re}L[M]}{||L||} \geq \frac{c}{||L||} \geq \inf \frac{\mathrm{Re}L[N]}{||L||}$ since $L \neq 0$. So L can be assumed to have norm 1.

Deutsch and Maserick's "Main Separation Principle" can now be shown.

2.5. Theorem. Let K be a convex subset of a normed linear space X, and let $f \in X \sim \bar{K}$. Then there exists $L \in X^*$ with norm 1 such that $\rho(f, K) = \mathrm{Re}L(f) - \sup \mathrm{Re}L[K]$.

Proof. Let $B = B(f, \rho(f, K))$. Then B is an open, convex set containing f disjoint from K. By theorem 2.4 there exists an $L \in X^*$ with $||L|| = 1$ which separates B and K. Hence $\sup \mathrm{Re}L[K] \leq \inf \mathrm{Re}L[B] \leq \mathrm{Re}\, L(f)$. Let $H = [L, \sup \mathrm{Re}L[K]]$. Then lemma 2.2 implies $\rho(f, H) = \mathrm{Re}L(f) - \sup \mathrm{Re}L[K] \leq \rho(f, K)$. If $\rho(f, H) < \rho(f, K)$ then there exists $y \in H \cap B$. But then B is a neighbourhood of y which lies on one side of H. This is a contradiction since any such neighbourhood contains points from both half spaces determined by H. Therefore $\rho(f, K) =$ $= \mathrm{Re}L(f) - \sup \mathrm{Re}L[K]$. □

Geometrically this theorem shows the existence of a hyperplane separating f and K which "just touches" K, and whose distance from f is the same as $\rho(f, K)$. This "just touches" notion can be stated more exactly as follows.

2.6 Definition. A hyperplane $H = [L, c]$ supports a set K at $\pi \in K$ if $\sup \mathrm{ReL}[K] = \mathrm{ReL}(\pi) = c$.

2.7. Corollary. Let K, f, L and H be as in theorem 2.5. If $\pi \in P(K, f)$ then H supports K at π .

Proof. As in theorem 2.5, let $B = B(f, \rho(f, K))$, and let $\overline{B}$ be the closure of B . Since $\pi \in P(K, f)$, $||f - \pi|| = \rho(f, K)$. Hence $\pi \in \overline{B}$. But $\inf \mathrm{ReL}[\overline{B}] = \inf \mathrm{ReL}[B] \geq \sup \mathrm{ReL}[K]$. Therefore

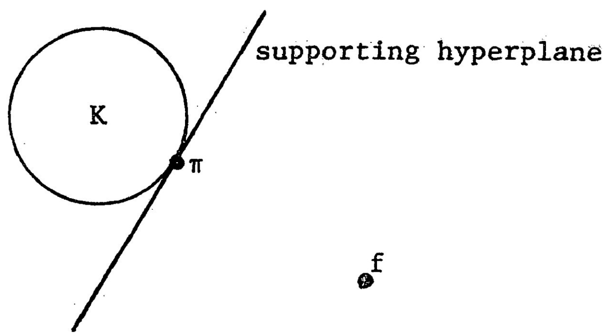
$$\begin{aligned} \mathrm{ReL}(\pi) &\leq \sup \mathrm{ReL}[K] \\ &\leq \inf \mathrm{ReL}[\overline{B}] \\ &\leq \mathrm{ReL}(\pi) . \end{aligned}$$

Thus $\mathrm{ReL}(\pi) = \sup \mathrm{ReL}[K]$. □

Now we are in a position to prove the first characterization theorem. A particular set involved will be referred to a bit later on, so for convenience we give the following definition separate from the theorem.

2.8. Definition. Let $L_\pi = \{L \in X^*: ||L|| = 1 \text{ and } L(f - \pi) = ||f - \pi||\}$, where π is an element of a convex set K and $f \in X \sim \overline{K}$. From the previous work we can see that L_π is not empty, and contains those functions L such that $[L: ||f - \pi||]$ supports the set $\{x \in X: ||x|| = ||f - \pi||\}$ at $f - \pi$.

The first major theorem of this chapter follows. Geometrically the theorem states that π is a best approximation from K to f if and only if there is a supporting hyperplane to K which passes through π, (ie., π is a point of support), and whose distance from f is the same as $\rho(f, K)$.



2.9. Theorem. (First Characterization of Best Approximation). If K is a closed convex subset of the normed linear space X , and $f \in X \sim K$, then $\pi \in P(K, f)$ if and only if there exists $L \in \mathcal{L}_\pi$ with

$$\text{(2.9.1)} \qquad \operatorname{Re}L(\pi) = \sup \operatorname{Re}L[K] .$$

Proof. Assume $\pi \in P(K, f)$. Theorem 2.5 implies the existence of $L \in X^*$ with $||L|| = 1$ and $\rho(f, K) = \operatorname{Re}L(f) - \sup \operatorname{Re}L[K]$. But $\rho(f, K) = ||f - \pi||$ and corollary 2.7 implies $\operatorname{Re}L(\pi) = \sup \operatorname{Re}L[K]$. Then $\operatorname{Re}L(f) - \operatorname{Re}L(\pi) = \operatorname{Re}L(f - \pi) = ||f - \pi||$.

If there exists such an $L \in \mathcal{L}_\pi$, then

$$||f - \pi|| = \operatorname{Re}L(f) - \operatorname{Re}L(\pi) \le \operatorname{Re}L(f) - \operatorname{Re}L(k)$$

for all $k \in K$. But $\operatorname{Re}L(f - k) \le |L(f - k)| \le ||f - k||$ for all $k \in K$. Therefore $\pi \in P(K, f)$. □

Singer's result [40, p. 18, theorem 1.1] can be recovered from the preceding theorem with the aid of the following lemma.

2.10. Lemma. If M is a linear subspace and $L \in X^*$ with $\sup \operatorname{Re}L[M] < \infty$ then $L(m) = 0$ for all $m \in M$.

Proof. Let $m \in M$. If $L(m) = a \neq 0$ then $L(\lambda m) = \lambda L(m) = \lambda a$ for all scalars λ . Since M is a subspace then $\sup \operatorname{Re}L[M] = \infty$. Thus $L(m) = 0$ for all $m \in M$. □

In theorem 2.9, if K is a closed subspace M , then (2.9.1) can be replaced by

(2.9.2) $$L(m) = 0 \text{ for all } m \in M$$

which is Singer's result. He also derives various other reformulations of this theorem. [40, pp. 19-24]

S3. SECOND CHARACTERIZATION THEOREM

Deutsch and Maserick [10, p. 524, theorem 3.9] have given a very nice proof of this theorem from a result of Singer's for linear subspaces. The proof is based on the following set of lemmas which are also of importance if M is of finite dimension.

3.1. Definition. A non-void subset E of a set A is called an extremal subset of A if $x, y \in E$ whenever $x, y \in A$ and $\alpha x + (1 - \alpha)y \in E$ for some $\alpha \in (0, 1)$. Geometrically, this means that no point of E is an interior point of a line segment whose end points are in $A \sim E$. If $E = \{z\}$ then z is termed an <u>extreme</u> point of A.

The next lemma is based on the Krein-Milman Theorem which is stated here without proof.

3.2. <u>Theorem</u>. (Krein-Milman) If K is a compact subset of a locally convex linear space and E is the set of its extreme points then E is not empty and $\overline{co}(E) = \overline{co}(K)$.

<u>Proof</u>. [16, p. 440, theorem 4]. □

3.3. <u>Theorem</u>. [41] If M is a linear subspace of a normed linear space X, and f is an extreme point of the closed unit ball in M^*, then f has an extension to X which is an extreme point of the closed unit ball in X^*.

<u>Proof</u>. Let $K = \{L \in X^*: ||L|| = 1$ and $L(m) = f(m), m \in M\}$. It can be shown that K is a w*-closed, convex subset of $\overline{B}^*(0, 1)$, which implies K is w*-compact. We will show that K is in fact an extremal subset of $\overline{B}^*(0, 1)$ so that the extreme points of K, which exist by the Krein-Milman theorem, are also extreme points of $\overline{B}^*(0, 1)$. If $g, h \in \overline{B}^*(0, 1)$ with $L = \alpha g + (1 - \alpha)h \in K$ for $\alpha \in (0, 1)$, then $\alpha g(m) + (1 - \alpha)h(m) = f(m)$, for all $m \in M$ and it is easy to verify that $||L|| = ||g|| = ||h|| = 1$. Since f is an extreme point of the unit ball in M^*, $g(m) = f(m) = h(m)$ for all

$m \in M$. Therefore $g, h \in K$, and K is an extremal subset of $\overline{B}^*(0, 1)$. Then any extreme point of K is a required extension. □

The following lemma is proved for the case of real scalars, but the proof follows the same lines for the complex case. Note especially that this lemma immediately gives a characterization of best approximations from finite dimensional subspaces.

3.4. Lemma. Let K be a closed convex subset of M , an n-dimensional subspace of X , where $n < \infty$. Let $f \in X \sim K$. Then $\pi \in P(K, f)$ if and only if there exist m extreme points L_i of $\overline{B}^*(0, 1)$ ($m \leq n + 1$ in the real case and $m \leq 2n + 1$ in the complex case) satisfying

$$(3.4.1) \qquad L_i(f - \pi) = ||f - \pi|| , \quad i = 1, \ldots, m$$

$$(3.4.2) \qquad \operatorname{Re} \sum_{i=1}^{m} \lambda_i L_i(\pi) = \sup \operatorname{Re} \sum_{i=1}^{m} \lambda_i L_i[K], \ \lambda_i > 0, \ \sum\lambda_i = 1 .$$

Proof. Let $Y = \operatorname{span}\{M, f\}$. Then the dimension of Y is at most $n + 1$. By a variant of Carathéodory's theorem given in [40, p. 166] we can write the functional $L \in \mathcal{L}_\pi(Y)$ of theorem 2.9 as $L = \sum_{i=1}^{m} \lambda_i L_i$ where $m \leq n + 1$, $\lambda_i > 0$, $\sum\lambda_i = 1$ and the L_i are extreme points of the unit ball in Y . Extend each L_i as in theorem 3.3. Then (2.9.1) implies (3.4.2), and the extended $L \in \mathcal{L}_\pi(X)$ implies $\sum\lambda_i L_i(f - \pi) = ||f - \pi||$. Since $\sum\lambda_i = 1$, we see that $\sum\lambda_i(L_i(f - \pi) - ||f - \pi||) = 0$, but $L_i(f - \pi) - ||f - \pi|| \leq 0$ for each i and $\lambda_i > 0$, so $L_i(f - \pi) - ||f - \pi|| = 0$ for $i = 1, \ldots, m$, giving (3.4.1). □

The second characterization theorem is due to Garkavi, and the proof given below is due to Deutsch and Maserick [10, p. 524].

3.5. <u>Theorem</u>. (Garkavi: Second Characterization of Best Approximation) Let K be a closed convex subset of a normed linear space X, and let $f \in X \sim K$. Then $\pi \in P(K, f)$ if and only if for each $k \in K$, there exists $L = L_k \in X^*$ such that

(3.5.1) $\quad L$ is an extreme point of B^* ;

(3.5.2) $$\operatorname{Re}L(\pi - k) \geq 0 ;$$

(3.5.3) $$L(f - \pi) = ||f - \pi|| .$$

<u>Proof</u>. Assume $\pi \in K$ and such an L exists for each $k \in K$. Then $||f - \pi|| = \operatorname{Re}L(f) - \operatorname{Re}L(\pi) \leq \operatorname{Re}L(f) - \operatorname{Re}L(k) \leq ||f - k||$. Since k is arbitrary $\pi \in P(K, f)$.

Let $\pi \in P(K, f)$. Since $\mathcal{L}_\pi$ is a w*-closed subset of $B^* \equiv \overline{B}^*(0, 1)$, $\mathcal{L}_\pi$ is w*-compact. Therefore $\mathcal{L}_\pi$ has extreme points by the Krein-Milman theorem. Since $\mathcal{L}_\pi$ is also an extremal subset of B^* these extreme points are extreme points of B^*, and they satisfy (3.5.1) and (3.5.3).

Now suppose $k \in K$ and $\operatorname{Re}L(\pi - k) < 0$ for all extreme points L of $\mathcal{L}_\pi$. Let N be the convex set $\{x \in X: x = \lambda k + (1 - \lambda)\pi, \lambda \in [0,1]\}$. We see that $\pi \in P(N, f)$ and N is in the span of k and π, which is two dimensional. Therefore, by lemma 3.4 there exist $m \leq 5$ (complex case) or $m \leq 3$ (real case) extreme points L_i of B^*, and m positive real numbers λ_i such that $\sum\lambda_i = 1$ and $L_i(f - \pi) = ||f - \pi||$, $i = 1, \ldots, m$ with $\operatorname{Re} \sum_{i=1}^{m} \lambda_i L_i(\pi - w) \geq 0$ for all $w \in N$. This last inequality is a restatement of (3.4.2). Since $\lambda_i > 0$, $i = 1, \ldots, m$ there must exist $w \in N$ and j such that $1 \leq j \leq n$ and $\operatorname{Re}L_j(\pi - w) \geq 0$, or, since $w \in N$, $\lambda \operatorname{Re}L_j(\pi - k) \geq 0$ for $\lambda > 0$ giving the required contradiction. □

S4. CHARACTERIZATION IN L_1 AND C_1

The two characterization theorems just proved can now be applied to the special case of approximation in $L_1(T, \sum, \mu)$. The results are all based on the Riesz Representation theorem, which shows the nature of the equivalence between L_∞ and L_1^*.

4.1. Theorem. (Riesz Representation Theorem) If $(T, \sum, \mu)$ (see I-2.7) is a positive, σ-finite measure space, then there is an isometric isomorphism between $L_1^*(T, \sum, \mu)$ and $L_\infty(T, \sum, \mu)$, where $L \in L_1^*$ and $g \in L_\infty$ are related by

$$L(f) = \int fg d\mu \quad \text{and} \quad ||L|| \equiv ||g||_\infty .$$

Proof. [16, p. 289]. □

This theorem is crucial to the application of the previous approximation theorems to the space L_1. Recall that we have decided that the measure space will not be exotic. The use this theorem will be put to indicates that a good definition for "exotic" would be the breakdown of the Riesz representation theorem.

The approximation theorems also indicate the need for a knowledge of the extreme points of the closed unit ball in L_1^* . Hence we prove the following handy lemma.

4.2. <u>Lemma</u>. For L_1 , L is an extreme point of the closed unit ball B^* in L_1^* if and only if there exists a $g \in L_\infty$ such that $|g| = 1$ almost everywhere and $L(f) = \int fg d\mu$ for all $f \in L_1$.

<u>Proof</u>. Since the map and its inverse of theorem 4.1 preserves the extreme points of the unit balls we need only show that g is an extreme point of the closed unit ball $\overline{B}$ in L_∞ if and only if $|g| = 1$ almost everywhere.

Assume $|g(t)| = 1$ for $t \in A$ where $\mu(T \sim A) = 0$, and that there exist f and h in $\overline{B}$, and $\alpha \in (0, 1)$ such that $g = \alpha f + (1 - \alpha)h$. For all $t \in A$,
$|g(t)| = 1 = |\alpha f(t) + (1 - \alpha)h(t)| \leq \alpha|f(t)| + (1 - \alpha)|h(t)| \leq 1$
since h and f have absolute value at most 1 . The functions f, h and g , therefore, map A into the unit disk of the complex plane. However $|g(t)| = 1$ $\forall t \in A$ implies $g(A)$ is a subset of the boundary of the unit disk. But the boundary of the unit disk is precisely the set of extreme points of the unit disk and so if $g(t) = \lambda f(t) + (1 - \lambda)h(t)$, $\lambda \in [0, 1]$, then $f(t) = h(t) = g(t)$ $\forall t \in A$. Since $\mu(T \sim A) = 0$, g is an extreme point.

If g is an extreme point and $|g(t)| \neq 1$ for $t \in E$ where $\mu(E) > 0$, then, since $||g|| \leq 1$, $|g(t)| < 1$ for all $t \in E$. Recall the function sgn is defined by $\operatorname{sgn} f(t) = \frac{f(t)}{|f(t)|}$ if $f(t) \neq 0$, and $\operatorname{sgn} f(t) = 0$ if $f(t) = 0$. Let $f(t) = g(t) = h(t)$ for $t \in T \sim E$ and let $f(t) = \operatorname{sgn} g(t)$ and $h(t) = (2|g(t)| - 1)\operatorname{sgn} g(t)$ for $t \in E$. Then $||f|| \leq 1$ and $||h|| \leq 1$ and $\frac{1}{2} f(t) + \frac{1}{2} h(t) = g(t)$ for $t \in T \sim E$. If $t \in E$, $\frac{1}{2} f(t) + \frac{1}{2} h(t) = |g(t)| \operatorname{sgn} g(t) = g(t)$. But $h(t) \neq f(t)$ for $t \in E$, and $\mu(E) > 0$, so g is not an extreme point. □

Now the characterization theorems can be applied. The first theorem is a restatement of theorem 2.9, using theorem 4.1. The proof will not be included.

4.3. <u>Theorem</u>. Let K be a closed convex subset of L_1 and $f \in L_1 \sim K$. Then $\pi \in P(K, f)$ if and only if there exists $g \in L_\infty$ such that

(4.3.1) $\quad ||g||_\infty = 1$;

(4.3.2) $\quad \operatorname{Re}\int \pi g d\mu = \sup\{\operatorname{Re}\int kgd\mu : k \in K\}$;

(4.3.3) $\quad \int (f - \pi) g d\mu = \int |f - \pi| d\mu$.

4.4. <u>Corollary</u>. Let K be a closed convex subset of L_1 and let $f \in L_1 \sim K$. Then $\pi \in P(K, f)$ if and only if there exist $g \in L_\infty$ with norm 1 such that

$$(4.4.1) \qquad \mathrm{Re}\int(\pi - k)g\,d\mu \geq 0 \quad \text{for all} \quad k \in K ;$$

$$(4.4.2) \qquad g(t) = \overline{\mathrm{sgn}}(f(t) - \pi(t)) \quad \text{for all} \quad t \notin Z(f - \pi) .$$

(Here, and elsewhere, $Z(h)$ denotes the zero set of h and the bar on $\overline{\mathrm{sgn}}(h(t))$ denote the complex conjugate of $\mathrm{sgn}(h(t))$ so that $\overline{\mathrm{sgn}}(h(t)) = \frac{|h(t)|}{h(t)}$ if $t \notin Z(h)$.)

Proof. (4.4.1) follows directly from (4.3.2). (4.3.3) implies $\int((f - \pi)g - |f - \pi|)d\mu = 0$. But $|g| \leq 1$ almost everywhere by (4.3.1) so $(f - \pi)g - |f - \pi| \leq 0$. Thus $(f - \pi)g = |f - \pi|$, or $g = \overline{\mathrm{sgn}}(f - \pi)$ whenever $f - \pi \neq 0$.

The converse again comes from 4.3, or it can be proven directly as follows: $\int|f - \pi|d\mu = \int(f - \pi)g\,d\mu \leq \mathrm{Re}[\int(f - k)g\,d\mu]$ by (4.4.1) and (4.4.2). But $\mathrm{Re}\int(f - k)g\,d\mu \leq \int|f - k|d\mu$ since $||g||_\infty = 1$; so $\pi \in P(K, f)$. □

4.5. Corollary. Let K be a closed convex subset of L_1 , and $f \in L_1 \sim K$. Then $\pi \in P(K, f)$ if and only if there exists a μ-measurable g defined on $Z(f - \pi)$ with $|g| \leq 1$ almost everywhere and

$$(4.5.1) \qquad \mathrm{Re}\int_{T\sim Z(f-\pi)}(\pi - k)\overline{\mathrm{sgn}}(f - \pi)d\mu \geq \mathrm{Re}\int_{Z(f-\pi)} g(\pi - k)d\mu$$

for all $k \in K$.

Proof. Let $\pi \in P(K, f)$. Then (4.4.1) and (4.4.2) imply

$$\mathrm{Re}\int_{T\sim Z(f-\pi)}(\pi - k)\overline{\mathrm{sgn}}(f - \pi)d\mu + \mathrm{Re}\int_{Z(f-\pi)}(\pi - k)g\,d\mu \geq 0$$

where $||g|| = 1$, which immediately gives (4.5.1).

Conversely assuming g as above exists, we extend it to T by defining $g(t) = \overline{\text{sgn}}(f(t) - \pi(t))$ for $t \notin Z(f - \pi)$. Then the extended g satisfies the requirements of 4.4 and $\pi \in P(K, f)$. □

If K is a subspace of L_1 , then these corollaries can be improved. The following theorem is a standard one proved by many authors, among them Kripke and Rivlin [24, p. 104] who gave a proof by considering the derivative of the norm. The proof given here follows the same lines as Singer's proof.

4.6. Theorem. Let M be a closed linear subspace of L_1 , and $f \in L_1 \sim M$. Then $\pi \in P(M, f)$ if and only if

(4.6.1) $$\left|\int m\,\overline{\text{sgn}}(f - \pi)d\mu\right| \le \int_{Z(f-\pi)} |m|d\mu \quad \text{for all } m \in M .$$

Proof. Assume $\pi \in P(M, f)$. Then by corollary 4.5 there exists a g as given in the corollary satisfying (4.5.1) for all $k \in M$. For each m let $m = \pi - k$. Since M is a linear subspace we see that

$$\text{Re}\int m\,\overline{\text{sgn}}(f - \pi)d\mu \ge -\text{Re}\int_{Z(f-\pi)} gmd\mu \quad \text{for all } m \in M .$$

We can replace m by $-m$ to show that only the equality is allowed. Taking the absolute value of both sides and recalling that $|g| \le 1$ then gives (4.6.1).

For the converse assume the inequality is valid. Following Singer we choose $u \in M$ arbitrarily, and define g by

$$g(t) = \begin{cases} \overline{\text{sgn}}(f(t) - \pi(t)) & \text{if } t \notin Z(f - \pi) ; \\ 0 \text{ if } t \in Z(f - \pi) \text{ and either } \mu(Z(f - \pi)) = 0 \\ \quad \text{or } u = 0 \text{ on } Z(f - \pi) ; \\ \left[\dfrac{-\int_{T \sim Z(f-\pi)} u\, \overline{\text{sgn}}(f - \pi) d\mu}{\int_{Z(f-\pi)} |u| d\mu} \right] \overline{\text{sgn}}\, u(t) & \text{otherwise.} \end{cases}$$

We see that $|g| = 1$ outside $Z(f - \pi)$, and on this set (4.6.1) ensures that $|g| \leq 1$. Therefore $g \in L_\infty$, and if $\mu(Z(f - \pi)) > 0$ and $u \neq 0$ on $Z(f - \pi)$,

$$\int gud\mu = \int_{T \sim Z(f-\pi)} u\, \overline{\text{sgn}}(f - \pi) d\mu - 1 \cdot \int_{T \sim Z(f-\pi)} u\, \overline{\text{sgn}}(f - \pi) d\mu$$
$$= 0 .$$

If $\mu(Z(f - \pi)) = 0$ or $u = 0$ on $Z(f - \pi)$ then (4.6.1) implies the same result straightforwardly. We also see that $\int g(f - \pi) = \int |f - \pi| d\mu = ||f - \pi||$, and therefore,

$$||f - \pi|| = \int g(f - \pi) d\mu - \int gud\mu$$
$$= \int g(f - \pi - u) d\mu \leq \int |f - \pi - u| d\mu ,$$

where the last step follows since $|g| \leq 1$ almost everywhere. Therefore $||f - \pi|| \leq ||f - (\pi + u)||$. Since M is a subspace and u is arbitrary, $\pi \in P(K, f)$. □

It might be expected that if M is finite dimensional lemma 3.4 would produce some interesting results, but unfortunately the only result is a trivial refinement of the previous theorem. If $M = \text{span}\{\phi_1, \ldots, \phi_n\}$ then g can be replaced by $\sum_{i=1}^{n} \lambda_i g_i$ where $\lambda_i \geq 0$, $\sum_{i=1}^{n} \lambda_i = 1$ and $|g_i| = 1$ almost everywhere, $i = 1, \ldots, n$. This does not change (4.6.1) at all.

The second characterization theorem gives similar results to the preceeding section. They are based, as usual, on the Riesz Representation theorem and the form of the extreme points of the closed unit ball. Note in particular that theorem 3.5 undergoes no change if K is a linear subspace. In fact, Singer's second characterization theorem [40, p. 62] for linear subspaces is the same, so his corollaries can be used here. Some of these corollaries are reproduced below. [40, pp. 63-67]

4.7. Theorem. Let K be a closed convex subset of L_1, and $f \in L_1 \sim K$. Then the following are equivalent.

a) $\pi \in P(K, f)$.

b) For each $k \in K$ there exists $q = q_k \in L_\infty$ with $|q| = 1$ a.e. such that

$$\text{Re}\int (\pi - k)q \, d\mu \geq 0 \tag{4.7.1}$$

$$\int (f - \pi)q \, d\mu = \int |f - \pi| d\mu \tag{4.7.2}$$

c) For each $k \in K$ there exists a μ-measurable $g = g_k$ with $|g| = 1$ a.e. on $Z(f - \pi)$ and

$$(4.7.3) \qquad \mathrm{Re}\int_{T\sim Z(f-\pi)} (\pi - k)\overline{\mathrm{sgn}}(f - \pi)d\mu \geq -\mathrm{Re}\int_{Z(f-\pi)} (\pi - k)g\, d\mu \ .$$

d) For each $k \in K$ there exists a μ-measurable $g = g_k$ with $|g| = 1$ a.e. on $Z(f - \pi)$ such that

$$(4.7.4) \qquad \int|f - \pi|d\mu \leq \mathrm{Re}\int_{T\sim Z(f-\pi)} (f - k)\overline{\mathrm{sgn}}(f - \pi)d\mu + \mathrm{Re}\int_{Z(f - \pi)} (f - k)g\, d\mu \ .$$

<u>Proof</u>. a) $\Leftrightarrow$ b). Apply theorems 3.5, 4.1 and lemma 4.2.

b) $\Rightarrow$ c). Since $|q| = 1$ a.e., (4.7.2) implies $q = \overline{\mathrm{sgn}}(f - \Pi)$ on $T \sim Z(f - \pi)$. Applying this to (4.7.1) results in (4.7.3) as in the proof of 4.4.

c) $\Rightarrow$ b). This proof follows the same lines as the converse in 4.4.

b) $\Rightarrow$ d). (4.7.1) implies

$$\mathrm{Re}\int(f - \pi)q\, d\mu \leq \mathrm{Re}\int(f - k)q\, d\mu \ .$$

Then by (4.7.2) $\int|f - \pi|d\mu \leq \mathrm{Re}\int(f - k)q\, d\mu$ and, as in 4.4, (4.7.2) also implies $q = \overline{\mathrm{sgn}}(f - \pi)$ on $T \sim Z(f - \pi)$. These two observations imply (4.7.4).

d) $\Rightarrow$ b). Let

$$q(t) = \begin{cases} g(t) & t \in Z(f - \pi) \\ \overline{\mathrm{sgn}}(f - \pi) & t \in T \sim Z(f - \pi) \end{cases}$$

Then (4.7.2) follows immediately and

$$\operatorname{Re}\int(\pi - k)q\,d\mu = \operatorname{Re}\int_{T\sim Z(f-\pi)}(f - k)\overline{\operatorname{sgn}}(f - \pi)d\mu + \operatorname{Re}\int_{Z(f-\pi)}(f - k)g\,d\mu - \int|f - \pi|d\mu$$

$$\geq 0$$

where the last inequality is implied by (4.7.4). □

4.8. Corollary. Let K be a closed convex subset of L_1, and $f \in L_1 \sim K$. Then $\pi \in P(K, f)$ if and only if

(4.8.1) $$\operatorname{Re}\int_{T\sim Z(f-\pi)}(\pi - k)\overline{\operatorname{sgn}}(f - \pi)d\mu \geq -\int_{Z(f-\pi)}|\pi - k|d\mu$$

for all $k \in K$.

Proof. If $\pi \in P(K, f)$ then theorem 4.7 part (c) holds. Then for the g in theorem 4.7 (c), $\operatorname{Re}\int_{Z(f-\pi)}(\pi - k)g\,d\mu \leq \left|\int_{Z(f-\pi)}(\pi - k)g\,d\mu\right| \leq$ $\leq \int_{Z(f-\pi)}|\pi - k|d\mu$, and (4.7.3) implies (4.8.1).

For the converse assume (4.8.1), and choose $u \in K$ arbitrarily. Define g by

$$g_u(t) = g(t) = \begin{cases} \overline{\operatorname{sgn}}(\pi - u) & \text{if } t \in Z(f - \pi) \sim Z(\pi - u) \\ 1 & \text{if } t \in Z(f - \pi) \cap Z(\pi - u) . \end{cases}$$

Then $$-\operatorname{Re}\int_{Z(f-\pi)}(\pi - u)g\,d\mu = -\int_{Z(f-\pi)}|\pi - u|d\mu$$

$$\leq \operatorname{Re}\int_{T\sim Z(f-\pi)}(\pi - u)\overline{\operatorname{sgn}}(f - \pi)d\mu .$$

By 4.7(c), $\pi \in P(K, f)$. □

If K is a subspace, then Singer gives another corollary which replaces (4.8.1) by

$$\text{(4.8.2)} \quad \operatorname{Re}\int_{T\sim Z(f-\pi)} m\,\overline{\operatorname{sgn}}(f-\pi)d\mu \geq -\int_{Z(f-\pi)} |m|d\mu \quad \text{for all } m \in K .$$

This can be easily derived from (4.8.1). The above characterizations are, of course, equivalent to corollary 4.4 (or 4.5). In fact, Corollary 4.5 is almost identical to theorem 4.7, part c).

In some restricted cases sharper results can be derived. Consider the real case, $L_1^R[T, \mu]$ where T is an interval $[a, b]$ of the real numbers and that μ is finite, and, as usual, nonatomic. Let the subspace M be equal to span$\{\phi_1, \ldots, \phi_n\}$ where the set $\{\phi_i\}_{i=1}^n$ is linearly independent, i.e., approximation from a finite dimensional subspace.

Also, assume the existence of a set of points x_i which satisfy $a = x_0 < x_1 < \ldots < x_r < x_{r+1} = b$, and

$$\sum_{i=1}^{r+1} (-1)^i \int_{x_{i-1}}^{x_i} \phi_j(t)d\mu(t) = 0 , \; j = 1, \ldots, n .$$

Then the following theorem can be proved; but, unfortunately, the conditions are only sufficient for a best approximation, not necessary. However, Usow [44] has used the theorem to advantage in producing an algorithm. It can also be used when applied to algebraic and trigonometric polynomial approximation.

4.9. <u>Theorem</u>. Let $\{\phi_i, i = 1, \ldots, n\}$ be a set of linearly independent functions in $L_1[T, \mu]$ where T is an interval $[a, b]$ in $\mathbb{R}$ and μ is finite and nonatomic. Let $\{x_i, i = 0, \ldots r + 1\}$ be a set of points, $r \leq n$, such that $a = x_0 < x_1 < \ldots < x_r < x_{r+1} = b$ and

$$(4.9.1) \qquad \sum_{i=1}^{r+1} (-1)^i \int_{x_{i-1}}^{x_i} \phi_j(t)d\mu(t) = 0 , \quad j = 1, \ldots, n .$$

If $\pi \in M$ interpolates f at $\{x_i: i = 1, \ldots, r\}$ and $f - \pi$ changes sign at precisely the points $\{x_i: i = 1, \ldots, n\}$ then $\pi \in P(M, f)$.

<u>Proof</u>. Let $s(x) = \alpha(-1)^i$, $x \in (x_{i-1}, x_i)$, and $s(x_i) = 0$, $i = 0, \ldots, r + 1$, where $\alpha = -\text{sgn}(f(a) - \pi(a))$. Since $f - \pi$ changes sign only at the x_i, $i = 1, \ldots, r$, $\text{sgn}[f(x) - \pi(x)] = s(x)$ for all $x \notin Z(f - \pi)$. Then (4.9.1) implies $\int_a^b s(x)\phi_j(x)d\mu(x) = 0$, $j = 1, \ldots, n$. Choose $m \in M$ arbitrarily. Then $m = \sum_{i=1}^{n} \lambda_i\phi_i$ so that $\int_a^b msd\mu = 0$, and $\int_a^b m\,\text{sgn}(f - \pi)d\mu = \int_a^b m\,s\,d\mu - \int_{Z(f-\pi)} ms\,d\mu \geq$ $\geq 0 - \int_{Z(f-\pi)} |m|d\mu$. Therefore by theorem 4.6, since m is arbitrary, $\pi \in P(M, f)$. □

It is natural then to ask of the existence of such critical points. Hobby and Rice [18] have shown that they do indeed exist in most cases. The theorem is presented here with a very nice proof by Alan Pinkus [34].

4.10. __Theorem.__ [Hobby-Rice] Let $\{\phi_i\}_{i=1}^n$ be real functions in $L_1^R[[0, 1], \mu]$, where μ is finite and nonatomic. Then there exist $\{u_i\}_{i=0}^{r+1}$, $r \le n$ such that $0 = u_0 < u_1 < \ldots < u_r < u_{r+1} = 1$ and

$$(4.10.1) \qquad \sum_{j=1}^{r+1} (-1)^j \int_{u_{j-1}}^{u_j} \phi_i(x)d\mu(x) = 0 \ , \ i = 1, \ldots, n \ .$$

__Proof.__ Recall that a mapping $T: \mathbb{R}^n \to \mathbb{R}^n$ is odd if $T(x) = -T(-x)$. For a set Ω in $\mathbb{R}^{n+1}$ let $\partial\Omega$ be the boundary (the closure without the interior) of Ω and let $C(\partial\Omega, \mathbb{R}^n)$ be the set of all continuous maps from $\partial\Omega$ to $\mathbb{R}^n$. The proof depends on the following version of the Borsuk Antipodality theorem [5, 29]:

Let Ω be a bounded, open, symmetric neighbourhood of 0 in $\mathbb{R}^{n+1}$ and $T \in C(\partial\Omega, \mathbb{R}^n)$, with T odd on $\partial\Omega$. Then there exists $x^* \in \partial\Omega$ for which $T(x^*) = 0$.

For the Hobby-Rice theorem, let $S = \{x = (x_1, \ldots, x_{n+1}): \sum_{i=1}^{n+1} |x_i| = 1\}$, and define $y_0(x) = 0$, $y_j(x) = \sum_{k=1}^{j} |x_k|$, $j = 1, \ldots, n+1$. Let $T: S \to \mathbb{R}^n$ be defined by

$$T_i(x) = \sum_{j=1}^{n+1} (\operatorname{sgn} x_j)\int_{y_{j-1}(x)}^{y_j(x)} \phi_i(t)d\mu(t), \ i = 1, \ldots, n \ .$$

Certainly $\{x: \sum_{i=1}^{n+1} |x_i| < 1\}$ is symmetric and open, and S is its boundary. Since $y_j(x) = y_j(-x)$, $T_i(x) = -T_i(-x)$ and T is odd. It remains to show that T is continuous with respect to x.

The y_j are certainly continuous functions of x, $i = 0, \ldots, n+1$. Consider the integral

$$\int_{y_{j-1}(x)}^{y_j(x)} \phi_i(t)d\mu(t) .$$

We can write this as

$$\int_0^{y_j(x)} \phi_i(t)d\mu(t) - \int_0^{y_{j-1}(x)} \phi_i(t)d\mu(t) .$$

Therefore we must show that the integral $F(x) = \int_0^{y(x)} \phi(t)d\mu(t)$ is a continuous function of x, in which case

$$\int_{y_{j-1}(x)}^{y_j(x)} \phi_i(t)d\mu(t)$$

is the sum of two continuous functions which is itself continuous.

Now, letting $f(x, t) = \phi(t)\psi_{[0,y(x)]}$, where ψ is the characteristic function, $F(x) = \int f(x, t)d\mu(t)$. Choose a point $x \in S$ and any sequence (x_n) converging to x. Since y is continuous, $\lim_{n\to\infty} f(x_n, t) = f(x, t)$ for all t except at $t = y(x)$. But μ is non-atomic so $f(x_n, t)$ converges to $f(x, t)$ almost everywhere since $\mu\{t: t = y(x)\} = 0$. Also $|f(x_n, t)| \leq |\phi(t)|$ and $|f(x, t)| \leq |\phi(t)|$ for all $t \in [0, 1]$. Applying the Lebesgue Dominated Convergence theorem [15, p. 328], $\lim_n F(x_n) = F(x)$. Since this is true for any sequence x_n converging to x, F is a continuous function of x.

Now consider T_i. Sgn x_j is a continuous function of x_j except where x_j changes sign. At this point $x_j = 0$; but then $y_{j-1}(x) = y_j(x)$, so the integral is also zero. Since the integral is continuous the product must approach 0 continuously so

$$\operatorname{sgn}(x_j)\int_{y_{j-1}(x)}^{y_j(x)} \phi_i(t)\,d\mu(t)$$

is continuous in x . Therefore $T \in C(S, \mathbb{R}^n)$ and we can apply Borsuk's theorem to find x^* with

$$\sum_{j=1}^{n+1} (\operatorname{sgn} x_j^*)\int_{y_{j-1}(x^*)}^{y_j(x^*)} \phi_i(t)\,d\mu(t) = 0,\ i = 1, \ldots, n .$$

If $x_j^* = 0$ or $\operatorname{sgn} x_j^* = \operatorname{sgn} x_{j-1}^*$ then the j-th term of the sum can be removed and the $y_j(x^*)$ relabeled to obtain $\{u_j\}_{j=1}^r$ with

$$\sum_{j=1}^{n+1} (-1)^j \int_{u_{j-1}}^{u_j} \phi_i(t)\,d\mu(t) = 0,\ i = 1, \ldots, n . \quad \square$$

4.11. Definition. A set B is a cone if $\lambda b \in B$ for all $\lambda \geq 0$ whenever $b \in B$. If B is also convex then it is called a convex cone.

The next theorem involves an example of a convex cone in C_1^R and a characterization of best approximations from such a cone. The results were first shown by Duffin and Karlovitz [15, p. 672, thm. 7]. First, two lemmas are proved.

4.12. Lemma. If K is a convex cone in a normed linear space X and $L \in X^*$ with $\sup \operatorname{Re} L[K] < \infty$, then $\sup \operatorname{Re} L[K] = 0$.

Proof. If $\sup \operatorname{Re}L[K] > 0$ then there exists $k \in K$ with $\operatorname{Re}L(k) > 0$. By definition 4.11 $\lambda \operatorname{Re}L(k) > 0$ for all $\lambda \geq 0$ so $\sup \operatorname{Re}L[K] = \infty$. Since $0 \in K$, $\sup \operatorname{Re}L[K] \geq 0$, so $\sup \operatorname{Re}L[K] = 0$. □

For the lemma to follow, let M be an n-dimensional subspace of $C_1^R(T)$, and let $K = \{k \in M: k(t) \geq 0$ for all $t \in T\}$. Let $f \in C_1^R(T) \sim K$, and define the set $C = \{\lambda f - k: \lambda \in \mathbb{R}, k \in K\}$. For each $t \in T$, let e_t be the linear functional on $N = \operatorname{span}\{M, f\}$ having the values $e_t(f) = 0$ and $e_t(m) = m(t)$ for all $m \in M$. Let $E = \{e_t: t \in T\}$.

4.13. Lemma. Suppose there exists $m' \in M$ such that $m'(t) > 0$ for all $t \in T$. Then the convex cone, $\operatorname{con}(E)$, generated by E is equal to the polar C°, of C, where

$$C^\circ = \{L \in N^*: L(u) \leq 1 \text{ for all } u \in C\}.$$

Note: For the definitions of some terms to be used in the proof, see [22, pp. 183-184]. By our definition 4.11 we have assumed the vertex is always 0, and the vertex is always in the cone (the cone is pointed).

Proof. By the definition of the convex cone C, C° is the convex cone $\{L \in N^*: L(f) = 0$ and $L(k) \geq 0$ for all $k \in K\}$. We see that C° is w*-closed in fact, closed, and $E \subseteq C^\circ$. Therefore $\overline{\operatorname{con}}(E) \subseteq C^\circ$. If $F \in C^\circ \sim \overline{\operatorname{con}}(E)$ then we can separate F and $\overline{\operatorname{con}}(E)$ by a linear functional $\hat{h}(u) = u(h)$, $u \in N$, for some $h \in N \sim \{0\}$ [16, p. 417, theorem 10, and p. 421, theorem 9]. Therefore there exist real numbers c and $\varepsilon > 0$ such that $F(h) \geq c$ and $L(h) \leq c - \varepsilon$ for all $L \in \overline{\text{con}}(E)$. Since $F \in C^{\circ}$, $F(f) = 0$, and similarly, $L(f) = 0$, $L \in \overline{\text{con}}(E)$. Now $h \in N$ implies $h = \lambda f + m$ for some $m \in M$, and then $F(m) \geq c$, $L(m) \leq c - \varepsilon$ for all $L \in \overline{\text{con}}(E)$. Therefore $L(m) < F(m)$. But $L \in \overline{\text{con}}(E)$ implies $\alpha L \in \overline{\text{con}}(E)$ for all $\alpha \geq 0$. Since $F(m)$ is bounded then $L(m) \leq 0$ for all $L \in \overline{\text{con}}(E)$, or $e_t(m) \leq 0$ for all $t \in T$. But $e_t(m) = m(t)$ so $m(t) \leq 0$ and $-m \in K$. Since $F \in C^{\circ}$ then $F(m) \leq 0$. But $\alpha L(m) < F(m)$ for all $\alpha \geq 0$ and $L \in \overline{\text{con}}(E)$ so we can choose $\alpha = 0$ and $F(m) > 0$, a contradiction.

To complete the proof we will show that $\overline{\text{con}}(E) = \text{con}(E)$. The cone $\text{con}(E)$ is pointed, and it is also a proper cone. If it were not proper then there would exist a t such that $e_t(m') < 0$, a contradiction. Also, since $\text{con}(E)$ is the convex cone generated by E, we see that $\text{con}(E)$ is the cone generated by the convex hull, $\text{co}(E)$, of E. Because of the existence of m', $\text{co}(E)$ does not contain the vertex, 0, of $\text{con}(E)$. We note now that E is in fact closed, as can be shown quite straightforwardly by considering any convergent sequence in E. Since M is finite dimensional and all functions are continuous it can be shown that E is bounded, and therefore compact. By theorem G in [38, p. 78] $\text{co}(E)$ is compact, which in turn implies that $\text{con}(E)$ is closed [22, p. 338]. Therefore $\overline{\text{con}}(E) = \text{con}(E) = C^{\circ}$. □

4.14. Theorem. Let $M = \text{span}\{\phi_1, \ldots, \phi_n\} \subseteq C_1^R$ be such that there exists $m \in M$ with $m(t) > 0$ for all $t \in T$. Let $K = \{m \in M: m(t) \geq 0$ for all $t \in T\}$. Then $\pi \in P(K, f)$ if and only if there exist distinct points $t_1, \ldots, t_s$ and positive numbers $\varepsilon_1, \ldots, \varepsilon_s$ such that $\pi(t_i) = 0$, $i = 1, \ldots, s$ and

$$(4.14.1) \quad \left|\int m \,\text{sgn}(f - \pi)d\mu + \sum_{i=1}^{s} \varepsilon_i m(t_i)\right| \leq \int_{Z(f-\pi)} |m|d\mu \quad \text{for all } m \in M .$$

Proof. Assume $\pi \in P(K, f)$. Since K is a convex cone theorem 4.3 and lemma 4.13 imply the existence of $g \in L_\infty$ such that $\int \pi g d\mu = 0 \geq \int k g d\mu$ for all $k \in K$, $||g||_\infty = 1$, and $g(t) = \text{sgn}(f(t) - \pi(t))$ for all $t \in T \sim Z(f - \pi)$. Let $N = \text{span}\{M, f\}$ and $C = \{\lambda f - k: \lambda \in \mathbb{R}, k \in K\}$. Define $F \in N^*$ by $F(k) = -\int k g d\mu$ for each $k \in M$, $F(f) = 0$, and extending linearly to N. Then $F(u) = \int u g d\mu \leq 0$ for all $u \in C$ and so $F \in C°$. Then lemma 4.13 allows us to write F as a positive combination of points in E. Therefore there exist positive numbers ε_i, $i = 1, \ldots, s$ and points e_{t_i} of E such that $F = \sum_{i=1}^{s} \varepsilon_i e_{t_i}$ where $s < \infty$. Therefore for each $m \in M$, $\int m g d\mu + \sum_{i=1}^{s} \varepsilon_i e_{t_i}(m) = 0$, or

$$\left|\int m \,\text{sgn}(f - \pi)d\mu + \sum_{i=1}^{s} \varepsilon_i m(t_i)\right| = \left|-\int_{Z(f-\pi)} m g d\mu\right|$$

$$\leq \int_{Z(f-\pi)} |m|d\mu .$$

Since $F(\pi) = 0$, $\sum_{i=1}^{s} \varepsilon_i \pi(t_i) = 0$. But $\varepsilon_i > 0$ and $\pi \geq 0$ so $\pi(t_i) = 0$, $i = 1, \ldots, s$.

For the converse, define a linear functional G on M by

$$G(m) = \int m \operatorname{sgn}(f - \pi) d\mu + \sum_{i=1}^{s} \varepsilon_i m(t_i) .$$

If we let $p(m) = \int_{Z(f-\pi)} |m| d\mu$ then it can easily be verified that p is a semi-norm on $C_1^R(T)$. From (4.14.1) we have $|G(m)| \leq p(m)$ for all $m \in M$. By the Hahn-Banach theorem G has an extension to all of $C_1^R(T)$ (which we denote by G again) with $|G(m)| \leq p(m)$ for all $m \in C_1^R(T)$. By the Riesz representation theorem, there exists $g \in L_\infty$ such that

$$G(m) = -\int mg d\mu .$$

Thus for all $m \in M$,

$$\int m \operatorname{sgn}(f - \pi) d\mu + \sum_{i=1}^{s} \varepsilon_i m(t_i) = -\int mg d\mu .$$

In particular, for $k, \pi \in K$,

$$\text{(4.14.2)} \quad \int (\pi - k) \operatorname{sgn}(f - \pi) + \int (\pi - k) g d\mu = -\sum_{i=1}^{s} \varepsilon_i (\pi - k)(t_i)$$

$$= \sum_{i=1}^{s} \varepsilon_i k(t_i)$$

$$\geq 0 .$$

Now define a linear functional H on C_1^R by

$$H(u) = \int u[\operatorname{sgn}(f - \pi) + g] d\mu .$$

Then we have $H(\pi - k) \geq 0$ for all $k \in K$. Since $|G(f - \pi)| \leq p(f - \pi) = 0$ it follows that $\int (f - \pi) g d\mu = 0$. Thus

$$H(f - \pi) = \int (f - \pi) \operatorname{sgn}(f - \pi) d\mu + \int (f - \pi) g d\mu$$

$$= \int |f - \pi| d\mu$$

$$= ||f - \pi|| .$$

Also for any $u \in C_1^R$, we have

$$|H(u)| = |\int (u \operatorname{sgn}(f - \pi) + ug) d\mu|$$

$$\leq |\int u \operatorname{sgn}(f - \pi) d\mu + |\int ug d\mu|$$

$$\leq \int_{T \sim Z(f-\pi)} |u| d\mu + \int_{Z(f-\pi)} |u| d\mu$$

$$= ||u|| ;$$

consequently $||H|| = 1$. By theorem II-2.9 then $\pi \in P(K, f)$. □

S5. <u>BEST ONE-SIDED APPROXIMATION</u>

Here we restrict ourselves to the real case and consider K to be those functions which are all less than or equal to some chosen f. The set K can be further restricted, for instance, by letting K lie in a suitable subspace. Unfortunately our previous theorems turn out to be rather trivial extensions of the basic definition of best approximation, as the following theorem shows. We will content ourselves with sketching two proofs to show the lack of information in the characterization theorems.

Also note that we have defined best approximation from below, but that best approximation from above is the same except for the obvious changes.

5.1. Theorem. Let $f \in L_1^R$, and define, for any convex set $C \subseteq L_1^R$, the set $K = \{k \in C: k(t) \leq f(t)$ for all $t \in T\}$. Then $\pi \in P(K, f)$ if and only if $\int \pi d\mu = \sup\{\int k d\mu : k \in K\}$.

Proof. Since $\int |f - k| d\mu = \int (f - k) d\mu$ for all $k \in K$ this is a trivial theorem. As an alternative proof we can apply corollary 4.8 (assuming $f \notin K$). Recall the inequality (4.8.1) specialized to the real case

$$\int_{T \sim Z(f-\pi)} (\pi - k)\operatorname{sgn}(f - \pi) d\mu \geq -\int_{Z(f-\pi)} |\pi - k| d\mu \quad \text{for all } k \in K .$$

In the right side since the integral is taken over $Z(f - \pi)$ we can replace π by f, then remove the absolute value signs as $f \geq k$, and then put back π. We also note that $\operatorname{sgn}(f - \pi) = 1$ on $T \sim Z(f - \pi)$. We then have

$$\int (\pi - k) d\mu \geq 0 \quad \text{for all } k \in K ,$$

which is what we needed. The converse follows also. □

Therefore we need some different methods to discover more useful results. Bojanic and DeVore [4] and DeVore [12] have investigated this subject and have come up with a characterization for best one-sided approximations. Bojanic and DeVore dealt with polynomial approximation which DeVore then generalized to cases where C (see theorem 5.1) is an n-dimensional Haar subspace M, where $n < \infty$. We will present some of DeVore's results here.

First we will have to define Haar systems. Attention is restricted to an interval $[a, b]$ in $\mathbb{R}$, and μ is some finite, non-atomic measure. The last condition is not absolutely necessary but in this thesis we have already noted that μ will be non-atomic in general. DeVore [12] summarizes the major properties of Haar systems, but for a deeper study Karlin and Studden's text is a good reference. [21, chapter 1] Note that sometimes (DeVore) a Haar system is called a Chebyshev system.

5.2. Definition. A set of functions, $\{\phi_1, \ldots, \phi_n\}$ in $C^R[a, b]$ is a Haar system if for any $m \in \text{span}\{\phi_1, \ldots, \phi_n\}$, $m \neq 0$, m has at most $n - 1$ zeroes in $[a, b]$. Equivalently, every determinant,

$$\begin{vmatrix} \phi_1(x_1) & \cdots & \phi_n(x_1) \\ \vdots & & \vdots \\ \phi_1(x_n) & \cdots & \phi_n(x_n) \end{vmatrix}$$

made from n distinct points $x_1, \ldots, x_n$ in $[a, b]$ is non-zero. If a basis of a subspace is a Haar system, then the subspace is called a Haar subspace.

It seems reasonable that, for a Haar system at least, the points where f and π meet (the interpolation points) will be of importance. The following lemma, due to DeVore, immediately indicates how the L_1 integral may be related to these points.

5.3. Lemma. Let M be a Haar subspace of (finite) dimension n. If $t_1, \ldots, t_n$ are any n distinct points in $[a, b]$, then there exist n real numbers $A_1, \ldots, A_n$ such that for any $m \in M$ we have

$$\int_a^b m\,d\mu = \sum_{i=1}^{n} A_i m(t_i) .$$

Proof. [12, Lemma 4.1]. This is a quadrature formula for M. The formula reflects the fact that the Haar property implies the point evaluations at $t_1, \ldots, t_n$ are linearly independent in the n-dimensional space M^* and so the linear functional defined by the integral is a linear combination of these point evaluations. □

We can now use this quadrature formula to give a characterization theorem for best approximation. To make it a little more general DeVore has introduced the concept of essential zeroes, and allowed f to be any real measurable function. The proof is no more complicated if we assume f to be discontinuous so we will follow DeVore's presentation.

5.4. Definition. Let f be a real function on $[a, b]$. A point $t_0 \in [a, b]$ is an essential zero of f if for every $\varepsilon > 0$ and neighbourhood N of t_0 there exists a point $t \neq t_0$, $t \in N$ such that $|f(t)| < \varepsilon$. This concept allows us to cope with discontinuous functions. Note that any zero of a continuous function is also an essential zero. One useful property says that if f has no essential zeroes and no zeroes on a closed set K in $[a, b]$ then using a simple compactness argument, there exists a positive ε such that $\inf|f[K]| \geq \varepsilon$.

5.5. Theorem. [12, Theorems 4.1 and 5.1] Let M be an n-dimensional Haar subspace of $L_1^R[a, b]$, $f \in L_1^R[a, b]$, and $K = \{k \in M: k(t) \leq f(t) \text{ for all } t \in [a, b]\}$.

(i) If $\pi \in P(K, f)$ and $f - \pi$ has precisely m zeroes, $m \leq n$, and all these zeroes are essential zeroes then these zeroes are nodes of a quadrature formula for M with non-negative coefficients.

(ii) If $\pi \in K$ and t_i, $i = 1, \ldots, m$, are essential zeroes of $f - \pi$ such that they are nodes of a quadrature formula for M with non-negative coefficients then $\pi \in P(K, f)$.

Proof. (i) Let $t_1, \ldots, t_m$ be the essential zeroes, $m \leq n$, of $f - \pi$, and assume $\pi \in P(K, f)$. Let $t_{m+1}, \ldots, t_n$ be any other points so that we have a full complement of n distinct points in $[a, b]$. Apply lemma 5.3 to find real numbers A_i such that

$$\int_a^b m d\mu = \sum_{i=1}^{n} A_i m(t_i)$$

for all $m \in M$.

Assume that $\sum_{i=m+1}^{n} |A_i| > 0$. Now define, for $r > 0$, the function $g \in M$ satisfying

$$g(t_i) = \begin{cases} -\frac{1}{r} & i = 1, \ldots, m \\ r \operatorname{sgn} A_i & i = m+1, \ldots, n \end{cases}$$

Then, by the quadrature formula,

$$\int g d\mu = -\frac{1}{r} \sum_{i=1}^{m} A_i + r \sum_{i=m+1}^{n} |A_i| .$$

Choose r large enough to ensure that this expression is positive. Since g is continuous there exists an open set (in the relative topology) N such that $\{t_1, \ldots, t_m\} \subseteq N$, and $g(t) < 0$ for all $t \in N$. We see that the set $H = [a, b] \sim N$ is compact and contains no essential zeroes or zeroes of $f - \pi$, and therefore there exists $\varepsilon > 0$ such that $f(t) - \pi(t) \geq \varepsilon$ for each $t \in H$. Let $\eta = \varepsilon/\sup\{|g(t)|: t \in [a, b]\}$. Now $f - \pi \geq 0$ so we have $\eta g(t) \leq 0 \leq f(t) - \pi(t)$ for all $t \in N$, and also $\eta g(t) \leq \eta \sup\{|g(t)|: t \in [a, b]\} = \varepsilon \leq f(t) - \pi(t)$, $t \in H$. Therefore $\eta g(t) \leq f(t) - \pi(t)$ for all $t \in [a, b]$, and we have that $\eta g + \pi \in K$. Also, r was chosen to ensure that $\int \eta g d\mu > 0$ (since $\eta > 0$) , and therefore $\eta g + \pi$ is a better approximation to f, a contradiction. Thus $\sum_{i=m+1}^{n} |A_i| = 0$, and all the "extra" A_i vanish.

It remains to show that the rest of the A_i are non-negative. Suppose there exists j, $1 \le j \le m$, such that $A_j < 0$. Then take $g \in M$ to satisfy, for some $r > 0$,

$$g(t_i) = \begin{cases} -1 & \text{if } i \ne j \\ -r & \text{if } i = j . \end{cases}$$

Then, again, we have $\int g d\mu = -\sum_{i \ne j} A_i + r|A_j|$, and can choose r large enought to make this positive. As before it is possible to show that $\eta g + \pi \in K$, and we have a contradiction.

(ii) If t_i, $i = 1, 2, \ldots, m$, are essential zeroes of $f - \pi$ then for every $k \in K$, $\pi(t_i) \ge k(t_i)$. For, suppose $\pi(t_j) - k(t_j) = -\delta < 0$ for some $1 \le j \le m$ and some $k \in K$. By continuity, there is a neighbourhood N of t_j such that $-\frac{3\delta}{2} < \pi(t) - k(t) < -\frac{\delta}{2}$ for every t in N. Since t_j is an essential zero of $f - \pi$ there is a t_0 in N for which $-\frac{\delta}{2} < f(t_0) - \pi(t_0) < \frac{\delta}{2}$. Thus

$$\begin{aligned} f(t_0) - k(t_0) &= [\pi(t_0) - k(t_0)] + [f(t_0) - \pi(t_0)] \\ &< -\frac{\delta}{2} + \frac{\delta}{2} \\ &= 0 , \end{aligned}$$

which contradicts that $k \in K$. Thus $\pi(t_i) \ge k(t_i)$, $i = 1, 2, \ldots, m$ and so

$$\int k d\mu = \sum_{i=1}^{m} A_i k(t_i) \le \sum_{i=1}^{m} A_i \pi(t_i) = \int \pi d\mu$$

for all $k \in K$. So $\pi \in P(K, f)$. □

Of course, this theorem is very specialized. No work has been done where C is more general, say, any finite dimensional subspace. This direction is open to further research.

DeVore goes on to investigate quadrature formulae with positive co-efficients in the case where the set $\{\phi_1, \ldots, \phi_n, f\}$ is also a

Haar system. He shows that such formulae always exist in this rather special case. This type of formula can be used when we consider polynomials, and especially trigonometric polynomials, where some very powerful results on interpolation are available. As an example we quote a theorem of Bojanic and DeVore [4, p. 152].

5.6. Theorem. Suppose that $f \in C_1^R[a, b]$ and that $\frac{d^n f}{dx^n}(t) \geq 0$ for all $t \in (a, b)$. If $n = 2\ell$ then the best approximation from the set of polynomials of degree at most $n - 1$ to f from below is defined as follows

$$\pi(t_i) = f(t_i),\ \pi'(t_i) = f'(t_i),\ i = 1, \ldots, \ell$$

where t_i are the nodes of a Gauss quadrature. Similar formulae but for different quadratures hold if $n = 2\ell + 1$.

S6. APPROXIMATION WITH RESTRICTED RANGE AND OTHER CONSTRAINTS

Lewis [25] has chosen the set K to be

$$K = \{k \in M: \ell(t) \leq k(t) \leq u(t) \text{ for all } t \in [a, b]\}$$

where all functions are continuous on $[a, b]$ and ℓ and u are chosen to bracket the chosen function f, i.e.,

$$\ell(t) \leq f(t) \leq u(t) \text{ for all } t \in [a, b]$$

$$\ell(t) < u(t) \text{ for all } t \in [a, b].$$

He calls this approximation with restricted range and derives characterization theorems for the uniform norm. However, the L_1 case seems to be untouched and is open to investigation.

One can also, for instance, take K to be those functions in a subspace M which interpolate f at some points. Lewis [25] has stated the following theorem, which can be easily proved using corollary 4.8.

6.1. Theorem. (Lewis) Let $f \in L_1^R[a, b]$ and M an n-dimensional subspace of $L_1^R[a, b]$. Define $K = \{k \in M: k(t_i) = f(t_i),\ i = 1, \ldots, m\}$, where $t_1, t_2, \ldots, t_m$ are $m \le n$ points in $[a, b]$. Then $\pi \in P(K, f)$ (in $L_1^R[a, b]$) if and only if

$$\left|\int_a^b \phi_i \operatorname{sgn}(f - \pi)d\mu\right| \le \int_{Z(f-\pi)} |\phi_i| d\mu$$

for $i = 1, \ldots, n - m$ where $\phi_1, \ldots, \phi_{n-m}$ is a basis of $K - \pi$.

It would be interesting to see if this result could be extended further, possibly by the use of some of DeVore's quadrature formulae. Rice [37] has done a lot of work with best approximation from "varisolvent" interpolating functions (with some special limit properties). He gives conditions for best approximations to be interpolating functions, but the converse is unfortunately less well covered. In general best approximations from varisolvent functions in $L_1^R[0, 1]$ are interpolating functions. The varisolvent condition is quite restrictive, as varisolvent functions must satisfy a type of Haar condition and any sequence of varisolvent functions with a limit must approach this limit with a uniform rate of convergence.

We can put other conditions on the co-efficients of the basis functions to construct more convex sets K . Although some work in the uniform norm has been done, little is available for L_1 . One special example of this type, that of spline approximation, has been worked on quite extensively, but the topic has a very wide range and is a bit beyond the scope of this thesis.

Chapter III

UNIQUENESS OF BEST APPROXIMATIONS

S1. INTRODUCTION

After having found a best approximation, it is natural to investigate the existence of another one. This chapter starts by giving criteria for the uniqueness of a best approximation, and then introduces Chebyshev sets, which are sets containing unique best approximations to every point in the whole space. A theorem characterizing Chebyshev sets is given, and then used to prove various examples.

S2. UNIQUENESS OF A BEST APPROXIMATION

In this section a best approximation is assumed to be known, and conditions are given for it to be the only best approximation.

2.1. Theorem. Let K be a convex subset of $L_1(T, \sum, \mu)$. If $\pi \in K$ satisfies

$$(2.1.1) \quad \mathrm{Re}\int_T (\pi - k)\overline{\mathrm{sgn}}(f - \pi)d\mu > -\int_{Z(f-\pi)} |\pi - k|d\mu \quad \text{for all } k \in K \sim \{\pi\}$$

then $\{\pi\} = P(K, f)$.

Proof. By Corollary (II-4.8), $\pi \in P(K, f)$. Choose an arbitrary u in $K \sim \{\pi\}$ and define $g \in L_\infty$ by

$$g(t) = \begin{cases} \overline{\operatorname{sgn}}(f(t) - \pi(t)), & t \notin Z(f - \pi) \\ I\ \overline{\operatorname{sgn}}(\pi(t) - u(t)), & t \in Z(f - \pi) \end{cases}$$

where I is given by the expression

$$I = \frac{\int_{T\sim Z(f-\pi)} (\pi - u)\overline{\operatorname{sgn}}(f - \pi)d\mu}{\int_{Z(f-\pi)} |\pi - u|d\mu} .$$

If the denominator vanishes let $I = 0$.

Recalling the definition of $\overline{\operatorname{sgn}}(\pi - u)$ we see that

$$\text{(2.1.2)} \quad \int g(\pi - u)d\mu = \int_{T\sim Z(f-\pi)} (\pi - u)\overline{\operatorname{sgn}}(f - \pi)d\mu - 1 \cdot \int_{T\sim Z(f-\pi)} (\pi - u)\overline{\operatorname{sgn}}(f - \pi)d\mu = 0 ,$$

and also, $|g| \le 1$ on T , and $|g| < 1$ on $Z(f - \pi)$ by the condition (2.1.1). As well, we note $||f - \pi|| = \int (f - \pi)gd\mu$, which can be written as $\int (f - u)gd\mu$ with the aid of (2.1.2). Then $||f - \pi|| \le \int_{T\sim Z(f-\pi)} |f - u|d\mu + \int_{Z(f-\pi)} (f - u)gd\mu < \int |f - u|d\mu$. Since u is arbitrary in $K \sim \{\pi\}$ we have $||f - \pi|| < ||f - u||$ for all $u \in K \sim \{\pi\}$. Therefore $\{\pi\} = P(K, f)$. □

Theorem 2.1 can, as usual, be applied to subspaces to get a result due to Kripke and Rivlin. If K is a subspace M , then condition (2.1.1) can be replaced by

$$(2.1.3)\qquad \left|\int m\,\overline{\operatorname{sgn}}(f - \pi)d\mu\right| < \int_{Z(f-\pi)} |m|\,d\mu \quad \text{for all} \quad m \in M \sim \{0\}\ .$$

The following rather trivial example shows that this condition is not, in general, a necessary condition. It is important when studying strong uniqueness.

2.2. Example. Let $T = [-1, 1]$ with the standard Lebesgue measure, $f(t) = t$, and $M = \operatorname{span}\{1\}$. Then the best approximation is $\pi(t) = 0$, $\int_{-1}^{1} m(t)\operatorname{sgn}(f(t))dt = 0$, and $\mu(Z(f)) = 0$. Therefore

$$\int^{1} m(t)\operatorname{sgn}(f(t))dt = 0 = \int_{Z(f)} |m(t)|dt \quad \text{for all} \quad m \in M\ .$$

But, π is unique as can be easily checked.

It would be advantageous, then, to have a necessary and sufficient condition for uniqueness. V. N. Nikolsky first noted the following theorem, which is based on the following lemma.

2.3. Lemma. If π_1 and π_2 are two distinct elements of $P(K, f)$, then

$$(2.3.1)\qquad \operatorname{sgn}(f(t) - \pi_1(t)) = \operatorname{sgn}(f(t) - \pi_2(t)),\ t \notin Z(f - \pi_1) \cup Z(f - \pi_2)\ .$$

Proof. Since K is convex, $P(K, f)$ is also a convex set. Then $\frac{\pi_1 + \pi_2}{2} \in P(K, f)$ if π_1 and π_2 are (distinct) elements of $P(K, f)$. Therefore $||f - \pi_1|| = ||f - \pi_2|| = ||f - \left(\frac{\pi_1 + \pi_2}{2}\right)||$, or,

$$\int \left|f - \frac{\pi_1 + \pi_2}{2}\right| d\mu = \int \left(\frac{1}{2}\,|f - \pi_1| + \frac{1}{2}|f - \pi_2|\right) d\mu \quad .$$

By the triangle inequality, $|f - \frac{\pi_1 + \pi_2}{2}| \le \frac{1}{2}|f - \pi_1| + \frac{1}{2}|f - \pi_2|$, which implies $|(f - \pi_1) + (f - \pi_2)| = |f - \pi_1| + |f - \pi_2|$ almost everywhere. If $f - \pi_1 \ne 0$ and $f - \pi_2 \ne 0$ then the equality can only be satisfied if $\mathrm{sgn}(f - \pi_1) = \mathrm{sgn}(f - \pi_2)$. □

2.4. Theorem. Let K be a convex subset of L_1. Then $\{\pi\} = P(K, f)$ if and only if $\pi \in P(K, f)$ and

$$(2.4.1) \quad \mathrm{Re}\int_{T\sim Z(f-k)} (k - \pi)\overline{\mathrm{sgn}}(f - k)d\mu < -\int_{Z(f-k)} |k - \pi| d\mu$$

for all $k \in K \sim \{\pi\}$.

Proof. Assume π is unique but (2.4.1) is not true in that there exists $u \in K$ for which

$$(2.4.2) \quad \mathrm{Re}\int_{T\sim Z(f-u)} (u - \pi)\overline{\mathrm{sgn}}(f - u)d\mu \ge -\int_{Z(f-u)} |u - \pi| d\mu .$$

Now

$$\begin{aligned} ||f - u|| &= \mathrm{Re}\int_{T\sim Z(f-u)} (f - u)\overline{\mathrm{sgn}}(f - u)d\mu \\ &= \mathrm{Re}\int_{T\sim Z(f-u)} (\pi - u)\overline{\mathrm{sgn}}(f - u)d\mu + \mathrm{Re}\int_{T\sim Z(f-u)} (f - \pi)\overline{\mathrm{sgn}}(f - u)d\mu \\ &\le \int_{Z(f-u)} |u - \pi| d\mu + ||f - \pi|| - \int_{Z(f-u)} |f - \pi| d\mu \\ &= ||f - \pi|| \end{aligned}$$

where (2.4.2) has been used to obtain the inequality, and for the last step

$f - \pi = u - \pi$ on $Z(f - u)$. Therefore $u \in P(K, f)$ as well and π is not unique.

Now assume $\{\pi\} \neq P(K, f)$. Then there exists $u \in K$ such that $u \in P(K, f) \sim \{\pi\}$, which implies $||f - u|| = ||f - \pi||$. Thus

$$\mathrm{Re}\int_{T\sim Z(f-u)} (f - u)\overline{\mathrm{sgn}}(f - u)d\mu = \int |f - \pi|d\mu$$

or,

$$\begin{aligned}\mathrm{Re}\int_{T\sim Z(f-u)} &(\pi - u)\overline{\mathrm{sgn}}(f - u)d\mu \\ &= \mathrm{Re}\int_{T\sim Z(f-u)} [|f - \pi| - (f - \pi)\overline{\mathrm{sgn}}(f - u)]d\mu \\ &\quad + \int_{Z(f-u)} |f - \pi|d\mu \ .\end{aligned}$$

Since $f - \pi = u - \pi$ on $Z(f - u)$, the last integral is equal to

$$\int_{Z(f-\pi)} |u - \pi|d\mu \ ,$$

and the first integral on the right vanishes since $\overline{\mathrm{sgn}}(f - u) = \overline{\mathrm{sgn}}(f - \pi)$ on $T \sim (Z(f - u) \cup Z(f - \pi))$ by lemma 2.3. Therefore (2.4.1) is contradicted. □

S3. CHEBYSHEV SETS

A problem which has been extensively investigated is the existence and characterization of sets from which every function from the space has a unique best approximation. The theory for uniform approximation is especially elegant, with some very fine and useful results. The

L_1 case is not so nice. Almost all work done is involved with linear subspaces, and, in the case of a non-atomic measure, no such finite dimensional, real subspace exists. However, the situation can be improved slightly when only continuous functions are considered, although most results are negative.

3.1. Definition. A subset K of a normed linear space X is (semi-) Chebyshev in X if every f in X has (at most one) a unique best approximation from K. In general, the existence of a best approximation is assumed, and conditions are given for uniqueness. Usually 'in X' is dropped wherever possible without ambiguity.

The first theorem is a general one due to Deutsch and Maserick [10, p. 525, thm. 4.2]. Geometrically K is Chebyshev if and only if K has no sides parallel to a side of the unit ball in X.

3.2. Theorem. Let K be a closed convex subset of a normed linear space X. Then K is semi-Chebyshev if and only if there does not exist an $L \in X^*$ such that

(3.2.1) $$||L|| = 1 ;$$

(3.2.2) $L(y_i) = ||y_i||$ for two distinct $y_i \in X$, $i = 1, 2$;

(3.2.3) $\operatorname{Re} L(k_i) = \sup \operatorname{Re} L[K]$ for two distinct $k_i \in K$, $i = 1, 2$;

(3.2.4) $$y_1 - y_2 = k_1 - k_2 .$$

Proof. Assume that K is not semi-Chebyshev. Then there exists $f \in X \sim K$ with at least two best approximations k_1 and k_2 from K. Since K is convex, $\pi = \frac{k_1}{2} + \frac{k_2}{2}$ is also a best approximation. Then the characterization theorem 2.2.9 implies the existence of $L \in X^*$ with $||L|| = 1$, $\text{ReL}(\pi) = \sup \text{ReL}[K]$, and $L(f - \pi) = ||f - \pi||$. Let $y_1 = f - k_2$ and $y_2 = f - k_1$. Then $y_1 - y_2 = k_1 - k_2$ and $L(f - \pi) = ||f - \pi|| = \frac{1}{2} L(f - k_1) + \frac{1}{2} L(f - k_2)$. But $L(f - k_i) \le ||f - k_i||$ since $||L|| = 1$ and also $||f - \pi|| = ||f - k_i||$, $i = 1, 2$. Therefore $L(y_1) = L(f - k_2) = ||f - k_2|| = ||y_1||$ and similarly $L(y_2) = ||y_2||$. Also $\frac{1}{2}\text{ReL}(k_1) + \frac{1}{2}\text{ReL}(k_2) = \text{ReL}(\pi) = \sup \text{ReL}[K]$. Then $\text{ReL}(k_1) = \text{ReL}(k_2) = \sup \text{ReL}[K]$ since k_1 and k_2 lie on the same side of the half-spaces determined by L.

Assume now that such an L exists. Let $f = k_1 + y_2 = k_2 + y_1$. Then $L(f - k_i) = ||f - k_i||$, $i = 1, 2$ and $\text{ReL}(k_1) = \text{ReL}(k_2) = \sup \text{ReL}[K]$. Theorem II-2.9 implies k_1 and k_2 are best approximations to f, and K is not semi-Chebyshev. □

3.3. Corollary. If M is a closed subspace of a normed linear space X then M is semi-Chebyshev if and only if there does not exist an $L \in X^*$ with

(3.3.1) $$||L|| = 1 ;$$

(3.3.2) $$L(m) = 0 \text{ for all } m \in M ;$$

(3.3.3) $L(y_i) = ||y_i||$, $i = 1, 2$, for two distinct elements y_1 and $y_2 \in X$ with $y_1 - y_2 \in M$.

Proof. This follows when lemma II-2.10 is applied to theorem 3.2. □

S4. CHEBYSHEV SUBSPACES IN L_1

4.1. Theorem. A closed subspace M is semi-Chebyshev in $L_1[T, \sum, \mu]$ if and only if there does not exist $g \in L_\infty$ with $||g||_\infty = 1$ and two distinct $y_i \in L_1$, $i = 1, 2$, such that

(4.1.1) $$\int gm d\mu = 0 \text{ for all } m \in M ;$$

(4.1.2) $$\int gy_i d\mu = \int |y_i| d\mu , \quad i = 1, 2 ;$$

(4.1.3) $$y_1 - y_2 \in M .$$

Proof. By Riesz' theorem the L of corollary 3.3 can be written $L(h) = \int hg d\mu$ for some $g \in L_\infty$. Since $||L|| = 1$, $||g||_\infty = 1$. (4.1.1), (4.1.2), and (4.1.3) follow immediately from the last three conditions of corollary 3.3. □

4.2. Theorem. A closed subspace M is semi-Chebyshev in L_1 if and only if there does not exist $g \in L_\infty$ with $||g||_\infty = 1$ and distinct points y_1 and y_2 in L_1 such that

(4.2.1) $$g(t) = \overline{\text{sgn}}\, y_i(t) \text{ for all } t \notin Z(y_i) \quad i = 1, 2 ;$$

(4.2.2) $$\int gm d\mu = 0 \text{ for all } m \in M ;$$

(4.2.3) $$y_1 - y_2 \in M .$$

Proof. (4.2.2) and (4.2.3) follow from (4.1.1) and (4.1.3). (4.1.2) implies $\int (gy_i - |y_i|)d\mu = 0$ where $||g||_\infty = 1$ and $i = 1$ or 2 (as it does throughout this proof). Let $y_i = re^{i\theta}$ and $g = se^{i\phi}$ with the standard conventions, and take the real part of the first expression to get $\int r(s\cos(\theta + \phi) - 1)d\mu = 0$. Since $s \le 1$ by the condition on the norm of g, the integrand is not positive, and therefore $r(s\cos(\theta + \phi) - 1) = 0$. Where r does not vanish, $s \ne 0$ and so $\cos(\theta + \phi) = \frac{1}{s}$; but $s \le 1$ so $\cos(\theta + \phi) = 1$ and $s = 1$. Therefore $g = e^{-i\theta} = \overline{\operatorname{sgn}(y_i)}$ wherever $y_i \ne 0$.

For the converse the g of this theorem satisfies the conditions of theorem 4.1, so M is semi-Chebyshev. □

We note that these theorems follow equally well if only the real case is considered.

The following theorem is a very nice one due to Cheney and Wulbert [8].

4.3. Definition. For a subspace M in L_1, a β-set is a set of the form $Z(f)$ where $0 \in P(M, f)$. Recall that f is in fact an equivalence class of functions, so it follows that $Z(f)$ is itself an equivalence class of sets, as is each β-set. This does not affect any of the theorems so we usually ignore it in the notation.

We can now state the theorem.

4.4. Theorem. If M is a linear subspace of L_1, then M is semi-Chebyshev if and only if 0 is the only element of M vanishing on a β-set.

Proof. Assume M is not Chebyshev. Then there exist distinct π_1 and π_2 in $P(M, f)$ for some $f \in L_1 \sim M$, and $\pi = \frac{\pi_1 + \pi_2}{2}$ is also a best approximation. Therefore

$$\int [|f - \pi| - \frac{1}{2}(|f - \pi_1| + |f - \pi_2|)]d\mu = 0$$

and the triangle inequality implies $|f - \pi| = \frac{1}{2}|f - \pi_1| + \frac{1}{2}|f - \pi_2|$ almost everywhere. Then, on the β-set $Z(f - \pi)$, $|f - \pi_1| + |f - \pi_2| = 0$. Therefore $\pi_1 = \pi_2$ on this set and $\pi_1 - \pi_2$ vanishes there.

If the condition is false then there exists $f \in L_1 \sim M$ and $u \in M$ with $Z(f) \subseteq Z(u)$. Let $y_1 = |u| \operatorname{sgn} f$, and let $y_2 = |u|(\operatorname{sgn} f - \frac{1}{2}\operatorname{sgn} u)$. Then $\operatorname{sgn} y_1(t) = \operatorname{sgn} y_2(t) = \operatorname{sgn} f(t)$ for all $t \notin Z(u)$. Let g have the same value as that g guaranteed by corollary II-4.4 applied to subspaces. Since $0 \in P(M, f)$ and M is a subspace it follows from II-(4.4.1) that $\int mgd\mu = 0$ for all $m \in M$, and from II-(4.4.2) that $g(t) = \overline{\operatorname{sgn}} f(t)$ for all $t \notin Z(f)$. Then $g(t) = \overline{\operatorname{sgn}}\, y_1(t) = \overline{\operatorname{sgn}}\, y_2(t)$ outside $Z(u)$. But by definition $Z(u) = Z(y_1) = Z(y_2)$, and also $y_1 - y_2 = |u|(\frac{1}{2}\operatorname{sgn} u) = \frac{1}{2}u$. Therefore $y_1 - y_2 \in M$ and the conditions of theorem 4.2 are satisfied. Thus M cannot be semi-Chebyshev. □

The next theorem gives a necessary condition for a subspace to be Chebyshev in the real case. It is a rather interesting condition of which use is made in the study of strong uniqueness. The proof is a variant of a proof by Cheney and Wulbert.

4.5. Theorem. If M is a Chebyshev subspace of L_1^R and $\{\pi\} = P(M, f)$, then $\mu(Z(f - \pi) \sim Z(m)) > 0$ for all $m \in M \sim \{\pi\}$.

Proof. Since M is a subspace we can assume $\pi = 0$ without loss of generality. Assume there exists $u \in M \sim \{0\}$ for which $\mu(Z(f) \sim Z(u)) = 0$. Let $Z = Z(u) \sim Z(f)$ and define $h \in L_1$ by

$$h(t) = \begin{cases} |u(t)| \operatorname{sgn} f(t), & t \notin Z \\ f(t), & t \in Z . \end{cases}$$

Then $Z(h) = Z(f)$ and $\operatorname{sgn} h = \operatorname{sgn} f$. Since $0 \in P(M, f)$, $\int m \operatorname{sgn} h d\mu = \int m \operatorname{sgn} f d\mu \le \int_{Z(f)} |m| d\mu = \int_{Z(h)} |m| d\mu$ by theorem II-4.6, and by the same theorem $0 \in P(M, h)$. For any $\theta \in (0, 1)$

$$\int |h - \theta u| d\mu = \int_Z h \operatorname{sgn} h d\mu + \int_{T\sim(Z(f)\cup Z(u))} (h - \theta u)\operatorname{sgn}(h - \theta u) d\mu$$

$$+ \int_{Z(f)\sim Z(u)} (h - \theta u)\operatorname{sgn}(h - \theta u) d\mu$$

$$+ \int_{Z(f)\cap Z(u)} (h - \theta u)\operatorname{sgn}(h - \theta u) d\mu .$$

The last integral vanishes, and so does the next to the last by our assumption that $\mu(Z(f) \sim Z(u)) = 0$. On $T \sim (Z(f) \cup Z(u))$, $\operatorname{sgn}(h - \theta u) = \operatorname{sgn} h = \operatorname{sgn} f$ since $\theta \in (0, 1)$. Thus

$$\int |h - \theta u| d\mu = \int_Z |h| d\mu + \int_{T\sim(Z(f)\cup Z(u))} |h| d\mu$$

$$+ \theta \int_{T\sim(Z(f)\cup Z(u))} (-u) \operatorname{sgn} f d\mu$$

$$= \int |h|\,d\mu + \theta \int (-u)\,\mathrm{sgn}\, f\,d\mu$$

$$\le \|h\| + \theta \int_{Z(f)} |u|\,d\mu \; .$$

But $\int_{Z(f)} |u|\,d\mu = \int_{Z} |u|\,d\mu + \int_{Z(f)\sim Z(u)} |u|\,d\mu = 0$. Thus $\|h - \theta u\| \le \|h\|$ for all $\theta \in (0, 1)$, or $\theta u \in P(M, f)$.

Therefore M is not Chebyshev. □

S5. FINITE DIMENSIONAL CHEBYSHEV SUBSPACES IN L_1^R

We will prove there are no finite dimensional Chebyshev subspaces in L_1^R . The proof is due to Phelps and is based on a theorem of Liapounoff.

5.1. Theorem. (Liapounoff) If $\mu_1, \ldots, \mu_n$ are finite, non-atomic measures on a set T and a σ-field $\sum$ of sets in T , then the subset of $\mathbb{R}^n$ consisting of all n-tuples of the form $(\mu_1(B), \ldots, \mu_n(B))$ for B in $\sum$ is closed and convex.

Proof. Lindenstraus [26] has given a very nice short proof of this theorem.

5.2. Lemma. If M is a finite dimensional subspace of L_1^R and there exists an extreme point L of B^* which annihilates M , then M is not Chebyshev.

Proof. Assume such an L exists. Then $\exists\ g \in L_\infty$ such that $|g| = 1$ a.e., $L(f) = \int fg\,d\mu$ and $\int mg\,d\mu = 0$ for all $m \in M$. Let

$M = \text{span}\{\phi_1, \ldots, \phi_n\}$. Let

$$y_1 = g \sum_{i=1}^{n} |\phi_i| \quad \text{and} \quad y_2 = y_1 - \sum_{i=1}^{n} \phi_i .$$

Then $\text{sgn } y_1(t) = \text{sgn } g(t)$ if $t \notin Z(y_1)$ and $\text{sgn } y_2(t) = \text{sgn } g(t)$ if $t \notin Z(y_1)$. Also, $y_1 - y_2 = \sum\phi_i \in M$. Therefore by theorem 4.2, M is not Chebyshev. □

5.3. <u>Theorem</u>. [33, p. 246, thm. 2.5] If $(T, \sum, \mu)$ contains no atoms then $L_1^R(T, \sum, \mu)$ contains no finite dimensional Chebyshev subspaces.

<u>Proof</u>. Let $M = \text{span}\{\phi_1, \ldots, \phi_n\} \subseteq L_1^R$. Write $\phi_i = \phi_i^+ - \phi_i^-$, where ϕ_i^+, ϕ_i^- are the positive and negative parts. Define, for each $B \in \sum$, $\mu_i^+(B) = \int_B \phi_i^+ d\mu$ and $\mu_i^-(B) = \int_B \phi_i^- d\mu$. The μ_i^+, μ_i^- are finite, non-atomic measures on $(T, \sum)$. By Liapounoff's theorem, the subset of $\mathbb{R}^{2n}$ consisting of all 2n-tuples of the form $(\mu_1^+(B), \mu_1^-(B), \ldots, \mu_n^+(B), \mu_n^-(B))$, for $B \in \sum$, is convex. Hence B can be chosen so that $\mu_i^+(B) = \frac{\mu_i^+(T)}{2}$ and $\mu_i^-(B) = \frac{\mu_i^-(T)}{2}$. Let $g = 1$ on B and -1 on $T \sim B$. Then $\int g\phi_i d\mu = 0$, $i = 1, \ldots, n$ and the linear functional L defined by g is an extreme point of B^*. Therefore, by the lemma, M is not Chebyshev. □

Theorem 5.3 cannot in general be extended to infinite dimensional subspaces as the following simple example of Phelps shows.

5.4. Example. Choose a B contained in T with $\mu(T \sim B) > 0$ and $\mu(B) > 0$, and let $M = \{m \in L_1^R: m(t) = 0$ whenever $t \in B\}$. If $f \in T \sim M$, let $\pi(t) = 0$ on B and $\pi(t) = f(t)$ on $T \sim B$. Then for any $m \in M \sim \{\pi\}$

$$||f - \pi|| = \int_B |f|d\mu + 0 < \int_B |f|d\mu + \int_{T\sim B} |f - m|d\mu = ||f - m|| .$$

Thus π is the unique best approximation and M is Chebyshev. □

While no finite dimensional Chebyshev subspaces in L_1 exists when the measure is non-atomic, we can obtain better results in $C_1[T, \mu]$.

S6. CHEBYSHEV SETS IN $C_1[T, \mu]$

As usual, $C_1[T, \mu]$ is that subset of $L_1[T, \mu]$ consisting of continuous functions. The measure μ is considered to be non-atomic, and T to be Hausdorff, completely regular. Usually T will be an interval and μ the Lebesgue measure. Cheney and Wulbert have, in particular, done much work on $C_1[T, \mu]$ in their paper, and one of their results is reproduced here. Jackson's famous theorem on Chebyshev subspaces will also be shown, together with an example which contradicts the converse of Jackson's theorem. We assume that the sets in question all have best approximations to any $f \in C_1$ so that existence is not considered, and that L_1^* is equivalent to L_∞.

The first two theorems are analogues of theorems 4.1 and 4.2, and the proofs are identical so they are stated here without proof.

6.1. <u>Theorem</u>. A closed M is a semi-Chebyshev subspace in $C_1[T, \sum, \mu]$ if and only if there does not exist $g \in L_\infty$ with $||g||_\infty = 1$ such that

(6.1.1) $\int gm d\mu = 0$ for all $m \in M$;

(6.1.2) $\int gy_i d = \int |y_i| d\mu$ for two distinct $y_i \in C_1$;

(6.1.3) $y_1 - y_2 \in M$.

6.2. <u>Theorem</u>. A closed subspace M is a semi-Chebyshev subspace in C_1 if and only if there does not exist $g \in L$ with $||g||_\infty = 1$ and distinct points y_1 and y_2 in C_1 such that

(6.2.1) $g(t) = \overline{\text{sgn}}\, y_1(t) \quad t \notin Z(y_1)$;

(6.2.2) $g(t) = \overline{\text{sgn}}\, y_2(t) \quad t \notin Z(y_2)$;

(6.2.3) $\int gm d\mu = 0$ for all $m \in M$;

(6.2.4) $y_1 - y_2 \in M$.

Again we remark that these theorems are valid, with the same proofs, in the real case.

The next theorem is due to Cheney and Wulbert [8, theorem 22]. It is the analogue for continuous functions of theorem 4.4, and involves γ-sets which are just β-sets in C_1 . Again the theorem works just as well if we consider the real case.

6.3. Definition. A γ-set is a set of the form $Z(f)$ for which $0 \in P(K, f)$ for some $f \in C_1[T, \mu]$.

6.4. Theorem. If M is a subspace of C_1 , then M is semi-Chebyshev in C_1 if and only if 0 is the only element of M vanishing on a γ-set in C_1 .

Proof. The proof is exactly the same as that of theorem 4.4, except one must note that the functions involved are all continuous. In particular we note that the functions y_1 and y_2 in the proof of theorem 4.4 are continuous because $Z(f) \subseteq Z(u)$. □

We will now prove a famous theorem of Jackson's [20]. It has been proved in many ways by various authors, for example, see [7], [35]. Cheney and Wulbert used the previous theorem to give a proof, but we prefer to use the characterization theorem 6.2.

Recall the definition of Haar subspaces, II-5.2.

6.5. Theorem. If M is a Haar subspace in $C_1^R[a, b]$, then M is Chebyshev in $C_1^R[a, b]$.

Proof. Let $M = \text{span}\{\phi_1, \ldots, \phi_n\}$ be a Haar subspace which is not Chebyshev. By theorem 6.2 there exists $g \in L_\infty$, with $||g||_\infty = 1$, and distinct y_1 and y_2 such that (6.2.1), (6.2.2), (6.2.3), and (6.2.4) hold.

Let $t_1 < t_2 < \ldots < t_m$ be all the points in $[a, b]$ where both y_1 and y_2 are zero. Since $y_1 - y_2 \in M$, $y_1 - y_2$ can have at most $n - 1$ zeroes, and thus $m < n$. Let $t_0 = a$, and $t_{m+1} = b$. If

g changes sign at some $t \in (t_i, t_{i+1})$ for some i, $0 \le i \le m$, then, since y_1 and y_2 are continuous, $y_1(t) = y_2(t) = 0$. But then $t = t_j$ for some j, so g cannot change sign in any interval (t_i, t_{i+1}), $0 \le i \le m$. If there are less than n such intervals subdivide and relabel the endpoints by $a = t_0 < t_1 < \ldots < t_n = b$. Then g is of the same sign, and in fact, constant on each interval. Let $\alpha_i = g_i(t)$ for $t \in (t_i, t_{i+1})$, $i = 0, \ldots, n-1$. Then we can write

$$\int_a^b \phi_j g d\mu = 0 = \sum_{i=0}^{n-1} \alpha_i \int_{t_i}^{t_{i+1}} \phi_j d\mu, \; j = 1, \ldots, n .$$

Let $f_i(\phi_j) = \int_{t_i}^{t_{i+1}} \phi_j d\mu$, $i = 0, \ldots, n-1$; $j = 1, \ldots, n$.

Then

$$\sum_{i=0}^{n-1} \alpha_i f_i(\phi_j) = 0, \; j = 1, \ldots, n ,$$

and, since all $\alpha_i \ne 0$, we must have $\det(f_i(\phi_j)) = 0$. Therefore there exists a non-trivial set $\{c_1, \ldots, c_n\}$ such that $\sum_{i=1}^{n} c_j f_i(\phi_j) = 0$, $i = 0, \ldots, n-1$, or,

$$\int_{t_i}^{t_{i+1}} \sum_{j=1}^{n} c_j \phi_j d\mu = 0 , \; i = 0, \ldots, n-1 .$$

Since the ϕ_j are continuous, $\sum c_j \phi_j$ must have at least one zero in each interval, or at least n zeroes in total, contradicting the Haar condition. □

Micchelli [30] has extended this theorem to the case where M is a "weakly Chebyshev" subspace. Here, if the ϕ_i span M, then the determinant $\det(\phi_i(x_j))$ must be non-negative for all sets of

x_j rather than strictly positive as in a Haar subspace. Unfortunately he has to restrict f to a special cone so the theorem is not as general as one would like.

Unfortunately Jackson's theorem does not give a necessary condition, in contrast to the uniform case, as the following example shows.

6.6. Example. M may be a Chebyshev subspace but not a Haar subspace in $C_1^R[0, 1]$.

Let M = linear span $\{t\}$. Now M is obviously not Haar since $m(0) = 0$ for all $m \in M$. It is sufficient to show that 0 is the only element in M which vanishes on a γ-set. If this is not true then there exists a γ-set $Z(f)$ for some $f \in C_1^R$ such that m vanishes on $Z(f)$ for some $m \in M \sim \{0\}$. But $Z(m) = \{0\}$, therefore $Z(f) = \{0\}$, or, $f(t)$ can be assumed to be positive for all $t \in (0, 1]$. Then $\mu(Z(f)) = 0$, and, since $0 \in P(M, f)$

$$\int_0^1 m(t)\operatorname{sgn} f(t)dt = c\int_0^1 t\, dt = \frac{c}{2} = 0$$

for all $m \in M$, $m(t) = ct$, which is a contradiction. Therefore M is Chebyshev.

Could we apply theorem 4.5 to $C_1^R[a, b]$? The following example shows that the answer is no.

6.7. Example. Theorem 4.5 is not true in $C_1^R[a, b]$.

Let M be the subspace of all constant functions on $[-1, 1]$ with the standard Lebesgue measure. Then M is a Haar subspace and hence Chebyshev in $C_1^R[-1, 1]$. Choose $f(t) = t$. Then $\{0\} = P(M, f)$, but $Z(f) = 0$, so $\mu(Z(f) \sim Z(m)) = 0$ for all $m \in M$.

S7. CONVEX CHEBYSHEV SETS IN L_1 OR C_1

We can apply theorem 3.2 directly to the L_1 or C_1 case to obtain the first two theorems of this section. Then we can move on to the special case of the convex cone considered in theorem II-4.14, and some other cases of constrained approximation.

7.1. Theorem. K is a semi-Chebyshev convex set in L_1 (respectively C_1) if and only if there does not exist a $g \in L_\infty$, distinct points y_1 and y_2 in L_1 (C_1), and distinct points k_1 and k_2 in K such that $\|g\|_\infty = 1$ and

(7.1.1) $$\int g y_i \, d\mu = \int |y_i| \, d\mu \qquad i = 1, 2 ;$$

(7.1.2) $$\mathrm{Re}\int g k_i \, d\mu \geq \mathrm{Re}\int g k \, d\mu \qquad i = 1, 2 \text{ for all } k \in K ;$$

(7.1.3) $$y_1 - y_2 = k_1 - k_2 .$$

Proof. These three conditions are immediate consequences of theorem 3.2. □

If K is a convex cone then lemma II-4.12 allows us to replace (7.1.2) by

(7.1.4) $\quad \mathrm{Re}\int gk_i d\mu = 0 \geq \mathrm{Re}\int kg d\mu, \; i = 1, 2 \quad \text{for all} \quad k \in K$.

As usual (7.1.1) can also be improved upon to yield the following theorem.

7.2. Theorem. K is a semi-Chebyshev convex cone in L_1 (respectively C_1) if and only if there does not exist $g \in L_\infty$, distinct points $y_1, y_2 \in L_1(C_1)$, and distinct points $k_1, k_2 \in K$ with $||g||_\infty = 1$ and

(7.2.1) $\quad g(t) = \overline{\mathrm{sgn}}\; y_i(t), \; t \notin Z(y_i) \quad i = 1, 2$;

(7.2.2) $\quad \mathrm{Re}\int gk_i d\mu = 0 \geq \mathrm{Re}\int kg d\mu, \; i = 1, 2, \quad \text{for all} \quad k \in K$;

(7.2.3) $\quad y_1 - y_2 = k_1 - k_2$.

Proof. This follows from theorem 7.1. □

Both theorems 7.1 and 7.2 work equally well in the real case.

We next turn to the example of theorem 4.14, that of best positive approximation. Lewis [25] uses this theorem to show that a Chebyshev set is the cone formed from the non-negative elements of an extended Haar system (of order 2). Instead we will use the previous theorem specialized to the real case, where it is also valid. First, a definition.

7.3. Definition. Let $\{\phi_1, \ldots, \phi_n\}$ be a set of differentiable functions in $C_1^R[a, b]$. If every non-zero m in the span of these functions has at most $n - 1$ zeroes in $[a, b]$, counting as two those zeroes where the derivative, m', is also zero, then $\{\phi_1, \ldots, \phi_n\}$ is an extended Haar system of order 2. If a subspace M has such a basis then it is an extended Haar subspace of order 2. For more details see [21].

7.4. Theorem. (Lewis [25]) Let M be an extended Haar subspace of order 2, and $K = \{k \in M: k(t) \geq 0$ for all $t \in [a, b]\}$. Then K is a Chebyshev cone in $C_1^R[a, b]$.

Proof. Assume K is not Chebyshev. Then there exist $g \in L_\infty$, distinct points $y_1, y_2 \in C_1^R$, and distinct points $k_1, k_2 \in K$ where $\|g\|_\infty = 1$ satisfying (7.2.1), (7.2.2) and (7.2.3) (where all quantities are real). For any $\lambda \in (0, 1)$ let $k_\lambda = \lambda k_1 + (1 - \lambda)k_2$ and $y_\lambda = \lambda y_1 + (1 - \lambda)y_2$. We now apply the same chain of reasoning as presented in theorem 6.5 to find $m \leq n$ points such that y_1, y_2, and y_λ change sign only at these points and have the same sign in the intervals. Let $t_1 < t_2 < \ldots < t_m$ be all points in $[a, b]$ such that y_1, y_2, and y_λ ($y_\lambda = 0$ whenever y_1 and y_2 vanish) are all zero. Since $y_1 - y_2 \in M$ by (7.2.3), and M is Haar, $m < n$.

Let $t_0 = a$, and $t_{m+1} = b$. If g changes sign at t in an interval (t_i, t_{i+1}) for some i, $0 \leq i \leq m$, then, since the

y_i are continuous, $y_1(t) = y_2(t) = 0$, and $y_\lambda(t) = 0$. But this is a contradiction, as t must be one of the t_j. Therefore g has the same sign on each interval, and, since $\lambda \in (0, 1)$, y_1, y_2, and y_λ will have the required properties. We note that this is a restatement of lemma 2.3, as $f = k_1 + y_2 = k_2 + y_1$ has two best approximations k_1 and k_2 from K.

Now let $Z(k_\lambda) = \{t_1, \ldots, t_s\}$ (not necessarily related to the previous set). Since M is an extended Haar subspace, $s \le n$. Let e be the number of endpoints a, b in $Z(k_\lambda)$. If y_λ has less than $n + e - 2s$ sign changes in $[a, b] \sim Z(k_\lambda)$ we can find $m \in M$ such that m has the same sign as y_λ and $m(t_i) = 0$ for $i = 1, \ldots, s$. ([21, p. 30], this is taken directly from the proof by Lewis). Since M is linear we can assume $|m| \le k_\lambda$, which implies $k_\lambda + m \in K$. But, $\int (k_\lambda + m) g\,d\mu = 0 + \int |m|\,d\mu$ by definition of g. This is positive and contradicts (7.2.2). Therefore y_λ has at least $n + e - 2s$ sign changes in $[a, b] \sim Z(k_\lambda)$, and so do y_1 and y_2 (since they change sign at the same places). Then $k_1 - k_2 = y_1 - y_2 = 0$ at these points.

However, since all elements of K are non-negative we have that for all $t \in Z(k_\lambda) \sim [a, b]$, $k_1(t) = k_2(t) = k_1'(t) = k_2'(t) = 0$. Then $k_1 - k_2$ has too many zeroes, and we have the required contradiction. □

Lewis has also given a very similar theorem for the case of best approximation from a set of interpolating functions. The proof uses theorem II-6.1 to find a point $k_1 - k_2$ which has too many zeroes in a manner analogous to the last proof, so we will just state the theorem here.

7.5. Theorem. Let f be in $C_1^R[a, b]$, and also let f be differentiable on (a, b). Assume $M = \text{span}\{\phi_1, \ldots, \phi_n\}$ is an extended Haar subspace of order 2, and $K = \{k \in M: k(t_i) = f(t_i), i = 1, \ldots, m\}$ where the t_i are $m < n$ points in $[a, b]$. Then f has a unique best approximation from K.

Proof. [25, theorem 5.4.] □

Lewis gave an example to show that the differentiability of f is necessary.

Can we find a condition on M so that best one-sided approximation from M gives us a Chebyshev set? The answer is, not quite. We will also need the differentiability condition on f. It would be nice if we could use our characterization, theorem 2.1, to show the uniqueness of such best approximations. However, as in section II-5, this theorem just gives the trivial refinement of the definition, that

$$\int \pi d\mu > \int k d\mu \quad \text{for all} \quad k \in K \sim \{\pi\} .$$

Therefore we will have to use some new methods; those of DeVore again. The theorem is due to DeVore and is based on the following lemma which he proves. Recall that the support, $C(\mu)$, of a measure μ is the complement of the union of all open sets of measure 0. Denote by $|C(\mu)|$ the number of points in $C(\mu)$, and by $|Z(f)|$ the number of essential zeroes of f (see definition II-5.4), where in both cases each point in (a, b) is counted twice.

7.6. <u>Lemma</u>. Let $M = \text{span}\{\phi_1, \ldots, \phi_n\}$ be a Haar subspace in $C_1^R[a, b]$, and let $f \in L_1^R[a, b]$. Define $K = \{k \in M: k(t) \leq f(t)$ for all $t \in [a, b]$ If $\pi \in P(K, f)$ then $|Z(f - \pi)| \geq \min\{|C(\mu)| , n\}$.

<u>Proof</u>. [12, pp. 16-17] Note that this has connections with theorem 4.5. □

In this lemma, as in the next theorem, it is only necessary that μ be a Borel measure. The condition of the next theorem, that $|C(\mu)| \geq n$, is not really very restrictive, and most of our "nice" measures easily satisfy this.

7.7. <u>Theorem</u>. (DeVore [12, theorem 3.3]) Let f and K be as in Lemma 7.6, but require that M be an extended Haar subspace of order 2. If f is differentiable on (a, b) , and $|C(\mu)| \geq n$ then the best one sided approximation to f from K is unique.

<u>Proof</u>. Because of continuity, essential zeroes are the same as ordinary ones. Assume $\{\pi_1, \pi_2\} \subseteq P(K, f)$. Then $\pi = \dfrac{\pi_1 + \pi_2}{2}$ is also in $P(K, f)$. Let $t_0 \in Z(f - \pi)$. Since π_1 and π_2 are elements of K we must have $\pi_1(t_0) = \pi_2(t_0) = f(t_0)$, and if

$t_0 \in (a, b)$ we also see that $(f - \pi)'(t_0) = (f - \pi_2)'(t_0) = 0$, or $\pi_1'(t_0) = \pi_2'(t_0) = f'(t_0)$.

Let r be the number of points of $Z(f - \pi)$, and s be the number of points of $Z(f - \pi) \sim \{a, b\}$. By definition, $|Z(f - \pi)| = r + s$, and $r + s \geq n$ by lemma 7.6. Therefore $\pi_1 - \pi_2$ has too many zeroes, and $\pi_1 = \pi_2$. □

We will close our discussion of Chebyshev sets here. The theorems only give some examples of convex Chebyshev sets, but these are quite complete and it is difficult to see where the conditions could be relaxed. It is unfortunate that we need extended Haar systems to guarantee uniqueness, and, as usual, the difficult properties of the L_1 norm have ensured that we have only sufficient conditions, and not some good necessary ones.

Chapter IV

STRONG UNIQUENESS

S1. INTRODUCTION

Strong uniqueness is a concept deriving from the behaviour of elements of the approximating set near the best (unique) approximation. We will start immediately with the definition, and then discuss the idea.

1.1. Definition. Let K be a convex subset of a normed linear space X, and $f \in X \sim K$. Then π is a strongly unique element of best approximation from K to f if there exists a real number $r > 0$ such that

(1.1.1) $$||f - k|| \geq ||f - \pi|| + r||\pi - k|| \quad \text{for all} \quad k \in K .$$

In this case we will write $\pi \in P_r(K, f)$. The existence of r will be implicitly assumed by such a statement.

First we note that the convexity of K is not strictly needed, but as we have restricted ourselves to such sets, it was included.

The inequality (1.1.1) says that if k moves around in K away from π then the approximation of f worsens with the rate of the distance from π. The concept is related to the question of smoothness of the ball in X. Recall that an element x is a point of smoothness of the closed ball of radius $||x||$ (centred at the origin) if there exists one and only exactly one hyperplane supporting the ball at x . If $f - \pi$ is such a point then π is not strongly unique. If $f - \pi$ is not a smooth point then the hyperplanes supporting the ball form a cone which we will use to characterize strongly unique approximations.

These ideas, and the presentation following, are due to papers by Bartelt and McLaughlin [3], and Wulbert [45]. We will present two examples used by Bartelt and McLaughlin to illustrate strong uniqueness from subspaces.

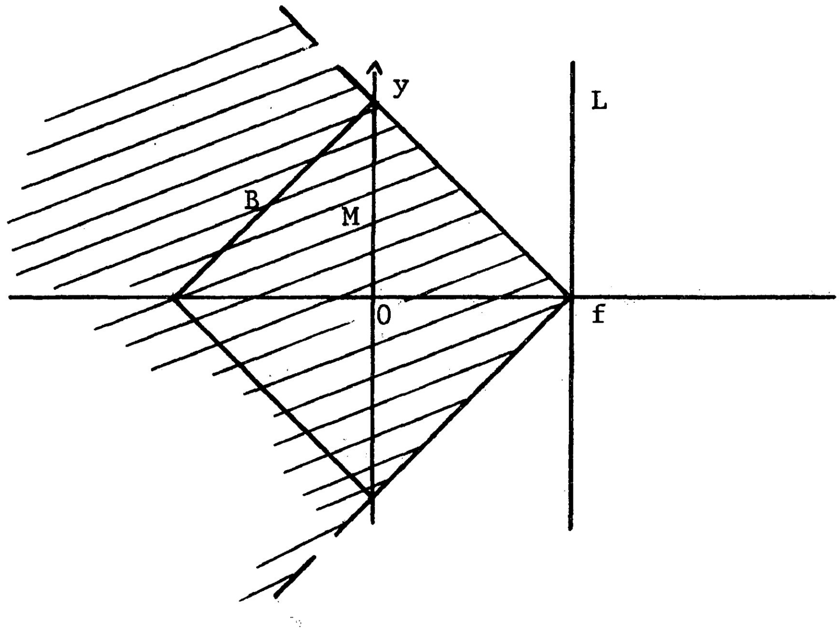
Take $X = \mathbb{R}^2$, with the ℓ_1 norm [figure 1.2] and the ℓ_2 norm [figure 1.3]. Let M be the hyperplane (line) $x = 0$, and f the point $(1, 0)$. In both cases $\{0\} = P(M, f)$, and L is a hyperplane supporting the ball, B , of radius $||f - \pi|| = ||f|| = 1$. With the ℓ_2 norm L is unique. The shaded area is the cone defined by the supporting hyperplanes of the ball. Since L is unique in the ℓ_2 case $f - \pi$ is a smooth point of B , and $\pi = 0$ is not strongly unique. In ℓ_1 , 0 is strongly unique.

These ideas can be formulated more precisely, and lead to our first characterization theorems. Recall the definition of L_π ;

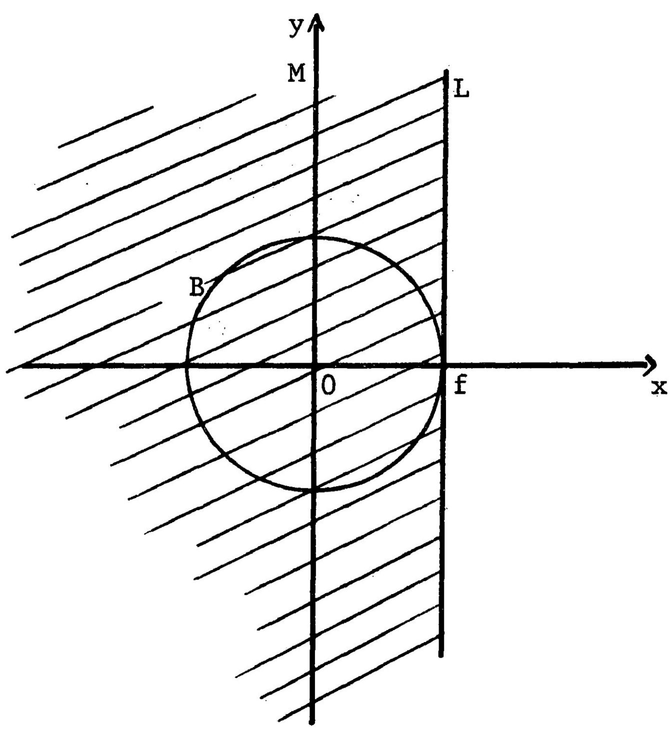
$$L_\pi = \{L \in X^*: L(f - \pi) = ||f - \pi||, ||L|| = 1\} .$$

This is the set of linear functionals which support the ball of radius $||f - \pi||$, centred at the origin, at $f - \pi$. Let $K_\pi = \{x \in X: \mathrm{Re}L(x) \leq ||f - \pi|| \text{ for all } L \in L_\pi\}$. This set is the cone generated by the hyperplanes $[L, ||f - \pi||]$ which support

(1.2)



(1.3)



the ball described above. If $\mathcal{L}_\pi$ consists of only one functional, then $f - \pi$ is a point of smoothness of this ball. If $\mathcal{L}_\pi$ contains an L such that $\operatorname{Re}L(\pi) = \sup \operatorname{Re}L[K]$ then the characterization theorem, II-2.9, implies $\pi \in P(K, f)$.

In a subspace M we will define a special functional on span$\{M, f\}$ by $L_\pi(m + af) = a||f - \pi||$ for any $m \in M$ and constant a , and extend it to X by means of the Hahn-Banach theorem (with the same norm). Certainly $||L_\pi|| \geq 1$. If $||L_\pi|| = 1$, then $L_\pi \in \mathcal{L}_\pi$; but $L_\pi(m) = 0$ for all $m \in M$ so L satisfies the requirements of theorem II-2.9 with (2.9.2), and therefore $\pi \in P(M, f)$.

We can now move on to consider the set K_π and prove the first strong uniqueness theorems.

S2. CHARACTERIZATION OF STRONG UNIQUENESS

The first theorem shows strong uniqueness implies uniqueness and $r \leq 1$ always.

2.1. Theorem. If $\pi \in P_r(K, f)$ then $\{\pi\} = P(K, f)$ and $r \leq 1$.

Proof. $||f - \pi|| \leq ||f - k|| - r||k - \pi|| < ||f - k||$ for all $k \in K \sim \{\pi\}$ by (1.1.1). Therefore π is unique. Since $||f - k|| \leq ||f - \pi|| + ||\pi - k||$, $r \leq 1$. □

If $\pi \in P_r(K, f)$, the next theorem proves all points on the ray from π towards f also have π as a strongly unique best approximation. It can also be shown that if $\pi \in P_r(K, f)$ then $0 \in P_r(K - \pi, f - \pi)$ so in fact it can be assumed that $0 \in P_r(K, f)$ and $||f|| = 1$. This will make some proofs easier.

2.2. Theorem. If $\pi \in P_r(K, f)$ then $\pi \in P_r(K, \lambda f + (1 - \lambda)\pi)$ for all $\lambda \geq 0$.

Proof. Let $h = \lambda f + (1 - \lambda)\pi$. If $\lambda \geq 1$ then for any $k \in K$, $\frac{k}{\lambda} + (1 - \frac{1}{\lambda})\pi \in K$. Therefore, for any $k \in K$,

$$\begin{aligned} ||h - k|| &= \lambda||f - (\tfrac{k}{\lambda} + (1 - \tfrac{1}{\lambda})\pi)|| \\ &\geq \lambda||f - \pi|| + r\lambda||\pi - (\tfrac{k}{\lambda} + (1 - \tfrac{1}{\lambda})\pi)|| \\ &= ||\lambda f + (1 - \lambda)\pi - \pi|| + r||\pi - k|| \\ &= ||h - \pi|| + r||\pi - k|| \end{aligned}$$

where the strong uniqueness of π has been used to obtain the inequality. Thus $\pi \in P_r(K, h)$. If $0 < \lambda < 1$ then $||h - \pi|| + ||f - h|| = ||f - \pi||$. For any $k \in K$ we can write

$$\begin{aligned} ||h - k|| &\geq ||f - k|| - ||f - h|| \\ &\geq [||f - \pi|| - ||f - h||] + r||\pi - k|| \\ &= ||h - \pi|| + r||\pi - k|| . \end{aligned}$$

So $\pi \in P_r(K, h)$. If $\lambda = 0$ the result is trivial. □

2.3. <u>Theorem</u>. [3, p. 257, prop. 1] Let M be a subspace of X and assume $\pi \in P_r(M, f)$. Then, for all $m \in M$ and all constants a , (real or complex), $a\pi + m \in P_r(M, af + m)$.

<u>Proof</u>. If $M = \{0\}$ the theorem is trivial. Assume $M \neq \{0\}$. Since $r \leq 1$, if $a = 0$, again, the result is trivial. If $a \neq 0$, then for all $u \in M$, there exists $r > 0$ such that

$$||(af + m) - u|| = |a| \, ||f - \frac{1}{a}(u - m)||$$

$$\geq |a| \, ||f - \pi|| + |a| r ||\pi - \frac{1}{a}(u - m)||$$

$$= ||(af + m) - (a\pi + m)|| + r||(a\pi + m) - u||$$

which implies $a\pi + m = P_r(M, af + m)$ for any $m \in M$. □

The next two theorems, for subspaces, are originally due to Wulbert [45, p. 352, lem. 1]. See also Bartelt and McLaughlin [3, p. 258, Theorem 1].

2.4. <u>Theorem</u>. Let K be any subset of X . If there exists $\pi \in K$ and $r > 0$ such that $\sup\{\mathrm{Re}L(\pi - k): L \in L_\pi\} \geq r||\pi - k||$ for all $k \in K$, then $\pi \in P_r(K, f)$.

<u>Proof</u>. For $k \in K$, $||f - k|| = \sup\{|L(f - k)|: ||L|| = 1\} \geq$ $\geq \sup\{|L(f - k)| : L \in L_\pi\}$. Now $|L(f - k)| \geq \mathrm{Re}L(f - k) =$ $= \mathrm{Re}L(f - \pi) + \mathrm{Re}L(\pi - k) = ||f - \pi|| + \mathrm{Re}L(\pi - k)$ for all $L \in L_\pi$. Therefore,

$$||f - k|| \geq ||f - \pi|| + \sup\{\text{Re}L(\pi - k): L \in \mathcal{L}_\pi\}$$

$$\geq ||f - \pi|| + r||\pi - k|| .$$

Since k is arbitrary, $\pi \in P_r(K, f)$. □

2.5. <u>Theorem</u>. Let K be any convex subset of X . If $\pi \in P_r(K, f)$ then $\sup\{\text{Re}L(\pi - k): L \in \mathcal{L}_\pi\} \geq r||\pi - k||$ for all $k \in K$.

<u>Proof</u>. Let $k \in K$ be arbitrary. Since $\pi \in P_r(K, f)$, for $0 < t < 1$, we have

$$||f - \pi + t(\pi - k)|| = ||f - ((1 - t)\pi + tk)||$$

$$\geq ||f - \pi|| + rt||\pi - k|| .$$

Consequently $\dfrac{||f - \pi + t(\pi - k)|| - ||f - \pi||}{t} \geq r||\pi - k||$. Let

$$c = \lim_{t \to 0^+} \frac{||f - \pi + t(\pi - k)|| - ||f - \pi||}{t} ,$$

which exists by lemma 1 on page 445 in [16]. Moreover, $c \geq r||\pi - k||$. By theorem 5 on page 447 in [16] there exists a linear functional ϕ_0 such that $||\phi_0|| = 1$, $\phi_0(f - \pi) = ||f - \pi||$ and $\phi_0(\pi - k) = c$. Thus we have $\sup\{\phi(\pi - k): \phi \in \mathcal{L}_\pi\} \geq \phi_0(\pi - k) = c \geq r||\pi - k||$. □

2.6. <u>Corollary</u>. Let K be a convex subset of X . Then $\pi \in P_r(K, f)$ if and only if there exists r and $\pi \in K$ with $\sup\{\text{Re}L(\pi - k): L \in \mathcal{L}_\pi\} \geq r||\pi - k||$ for all $k \in K$.

Proof. From theorems 2.4 and 2.5. □

The next corollary is based on the following lemma.

2.7. Lemma. Let A be an open convex subset of a linear topological space X. If a convex function f, defined on A, is bounded above on a neighbourhood of a point $a \in A$, then f is continuous at every point in A.

Proof. See [19, p. 82]. □

2.8. Corollary. Let M be a finite dimensional subspace. Then $\pi \in P_r(M, f)$ if and only if $\pi \in M$ and $\sup\{\mathrm{Re}L(m): L \in L_\pi\} > 0$ for all $m \in M \sim \{0\}$.

Proof. If $\pi \in P_r(M, f)$ then corollary 2.6 immediately shows $\sup\{\mathrm{Re}L(m): L \in L_\pi\} > 0$ for all $m \in M \sim \{0\}$.

Conversely let $p(m) = \sup\{\mathrm{Re}L(m): L \in L_\pi\}$. Then it is easy to verify that p, defined on M, satisfy the following:

(i) $p(m) \geq 0$ with equality hold exactly when $m = 0$;

(ii) $p(\lambda m) = \lambda p(m)$ if $\lambda \geq 0$ and $m \in M$;

(iii) $p(m + n) \leq p(m) + p(n)$ for all $m, n \in M$.

Consequently p is a convex function defined on M. Moreover, p is continuous. To show this, it suffices, by lemma 2.7, to prove p is bounded above on the unit ball $B(M)$. But for any $m \in L_\pi$

and $L \in \mathcal{L}_\pi$, $\mathrm{Re}L(m) \le |L(m)| \le ||L||\ ||m|| \le 1$ and so $p(m) \le 1$ for all $m \in B(M)$.

Now since M is finite dimensional, its unit sphere $S(M) = \{m \in M: ||m|| = 1\}$ is compact and so p attains its minimum $r > 0$ on $S(M)$. Now if $m \in M \sim \{0\}$ is arbitrary then $p\left(\frac{m}{||m||}\right) = \frac{1}{||m||} p(m) \ge r$ and so $p(m) \ge r||m||$ for all $m \in M$. By corollary 2.6, $\pi \in P_r(M, f)$. □

Bartelt and McLaughlin give two more characterizations of strong uniqueness. The first will be stated without proof since it will not be used again.

2.9. <u>Theorem</u>. Let M be a subspace of X. Then the set $K_\pi \cap M$ is bounded for some $\pi \in M$ if and only if $\pi \in P_r(M, f)$.

<u>Proof</u>. See Bartelt and McLaughlin [3, p. 259, theorem 2] for the proof. □

2.10. <u>Theorem</u>. [3, p. 260, theorem 3] Let M be a subspace of X. If $\pi \in P_r(M, f)$ then the set $A_\pi = \{x \in \mathrm{span}\{M, f\}: \mathrm{Re}L_\pi(x) = ||f - \pi||\} \cap K_\pi$ consists exactly of those elements of the form $x = (1 + ia)(f - \pi)$ where $a \in \mathbb{R}$.

<u>Proof</u>. Recall L_π is defined by $L_\pi(m + af) = a||f - \pi||$. Certainly if x is of the required form then $x \in A_\pi$.

Assume $x \in A_\pi$, and $x = b(f - \pi) + m$ for some $m \in M$ and scalar b. Then $\text{ReL}_\pi(x) = (\text{Re}b)||f - \pi||$. By the definition of A_π then $\text{Re}b = 1$ and x is of the form $(1 + ia)(f - \pi) + m$ for some $m \in M$. Now $x \in K_\pi$ implies $\text{ReL}(x) \le ||f - \pi||$ if $L \in L_\pi$. But for $L \in L_\pi$, $\text{ReL}(x) = ||f - \pi|| + \text{ReL}(m)$. Therefore $\text{ReL}(m) \le 0$ for all $L \in L_\pi$. By corollary 2.6 $m = 0$ and $x = (1 + ia)(f - \pi)$. □

2.11. Theorem. [3, p. 261, theorem 4] If M is a finite dimensional subspace of X, and if there exists $\pi \in M$ such that A_π consists exactly of those elements of the form $x = (1 + ia)(f - \pi)$, $a \in \mathbb{R}$, then $\pi \in P_r(M, f)$.

Proof. Assume $\pi \notin P_r(M, f)$. Then corollary 2.8 implies there exists $m \in M \sim \{0\}$ such that $\sup\{\text{ReL}(m): L \in L_\pi\} \le 0$. Let $x = m + (1 + ia)(f - \pi)$. Then $\text{ReL}_\pi(x) = ||f - \pi||$ and for all $L \in L_\pi$, $\text{ReL}(x) \le 0 + ||f - \pi||$ which implies $x \in A_\pi$ for all a, giving a contradiction. □

Bartelt and McLaughlin give an example [3, p. 261] to show that the finite dimensionality of M is in general a necessary condition.

The following theorem indicates the usefulness of strong uniqueness. Let T be the operator $Tf = P(K, f)$. If f has a strongly unique best approximate from a Chebyshev set K then Tf satisfies the following Lipschitz condition at f, which guarantees continuity at f.

2.12. Theorem. [6, p. 82] Let K be a Chebyshev set in X, and $f_0 \in X \sim K$ such that there exists $\pi \in K$ with $\pi \in P_r(K, f_0)$. Then there exists $\lambda > 0$ such that

$$||Tf_0 - Tf|| \leq \lambda ||f_0 - f|| \qquad \forall f \in X .$$

Proof. Since K is Chebyshev, T is a single-valued mapping. By definition, $Tf_0 = \pi$, and for any f, $Tf \in K$. Therefore $r||Tf_0 - Tf|| \leq ||f_0 - Tf|| - ||f_0 - Tf_0|| \leq ||f_0 - f|| + ||f - Tf|| - - ||f_0 - Tf_0||$. But $||f - Tf|| \leq ||f - Tf_0||$ since $Tf \in P(K, f)$ and $||f - Tf_0|| \leq ||f - f_0|| + ||f_0 - Tf_0||$. Then $r||Tf_0 - Tf|| \leq 2||f_0 - f||$. Choose $\lambda = \frac{2}{r}$ to finish the proof. □

S3. STRONG UNIQUENESS IN L_1

As usual whenever $L_1(T, \sum, \mu)$ is considered, μ is assumed to be non-atomic, and L_1^* equivalent to L_∞. The first theorem is just a restatement of theorem 2.4 applied to L_1. Assume now that $X = L_1(T, \sum, \mu)$ and K and M are convex subsets and subspaces, respectively, of $L_1(T, \sum, \mu)$.

3.1. Theorem. For a convex set K, $\pi \in P_r(K, f)$ if and only if there exist $r > 0$ and $\pi \in K$ with

$$\sup\{\mathrm{Re}[\int (\pi - k)\overline{\mathrm{sgn}}(f - \pi)d\mu + \int_{Z(f-\pi)} (\pi - k)g\,d\mu] : |g| \leq 1 \text{ a.e.}\} \geq$$

$$\geq r\int |\pi - k|d\mu \quad \text{for all } k \in K .$$

Proof. A straightforward application of the Riesz Representation theorem and corollary 2.6 gives this theorem since L_π is equivalent to the set $\{g \in L_\infty : \|g\|_\infty = 1$ and $g(t) = \overline{\mathrm{sgn}}(f(t) - \pi(t))$ for all $t \in T \sim Z(f - \pi)\}$. □

3.2. Theorem. $\pi \in P_r(M, f)$ if and only if there exists $r > 0$ such that $\sup\{\mathrm{Re}[\int m\ \mathrm{sgn}(f - \pi)d\mu + \int_{Z(f-\pi)} mg d\mu] : |g| \le 1 \text{ a.e.}\} \ge$ $\ge r\int |m|\, d\mu$ for all $m \in M$.

Proof. Since for subspaces M, $\pi - M = M$ if $\pi \in M$, this theorem follows immediately from theorem 3.1. □

The next two theorems show immediately the correspondence with characterizations of best approximation. Compare especially theorems II-4.6 and III-2.1 (for subspaces).

3.3. Theorem. If $\pi \in P_r(M, f)$ then

(3.3.1) $$\left|\int m\ \overline{\mathrm{sgn}}(f - \pi)d\mu\right| < \int_{Z(f-\pi)} |m|\, d\mu \quad \text{for all } m \in M \sim \{0\} .$$

Proof. By theorem 3.1, for an arbitrary $m \in M \sim \{0\}$, there exist μ-measurable g defined on $Z(f - \pi)$ with $|g| \le 1$ a.e. such that $\mathrm{Re}[\int m\ \overline{\mathrm{sgn}}(f - \pi)d\mu + \int_{Z(f-\pi)} mg d\mu] > 0$, or

$$\mathrm{Re}\int (-m)\overline{\mathrm{sgn}}(f - \pi)d\mu < \mathrm{Re}\int_{Z(f-\pi)} mg d\mu$$

$$\le \int_{Z(f-\pi)} |m|\, d\mu .$$

Since m is arbitrary in $M \sim \{0\}$, (3.3.1) is obtained. □

3.4. Theorem. Let M be a finite dimensional subspace. If

(3.4.1) $$\left|\int m\,\overline{\mathrm{sgn}}(f-\pi)\,d\mu\right| < \int_{Z(f-\pi)} |m|\,d\mu \quad \text{for all} \quad m \in M \sim \{0\}$$

then $\pi \in P_r(M, f)$.

Proof. Assume (3.4.1) holds. Choose $m \in M \sim \{0\}$ arbitrarily. Then (3.4.1) implies $\mathrm{Re}\left[\int m\,\overline{\mathrm{sgn}}(f-\pi)\,d\mu + \int_{Z(f-\pi)} |m|\,d\mu\right] > 0$. Choose $g = \overline{\mathrm{sgn}}\, m$ on $Z(f-\pi)$, and $g = \overline{\mathrm{sgn}}(f-\pi)$ outside $Z(f-\pi)$. Then L , defined by $L(h) = \int gh\,d\mu$, is in $\mathcal{L}_\pi$. Therefore, for all $m \in M \sim \{0\}$, $\sup\{\mathrm{Re}L(m) : L \in \mathcal{L}_\pi\} > 0$ and by corollary 2.8, $\pi \in P_r(M, f)$. □

3.5. Corollary. Let M be a finite dimensional subspace. $\pi \in P_r(M, f)$ if and only if

(3.5.1) $$\left|\int m\,\overline{\mathrm{sgn}}(f-\pi)\,d\mu\right| < \int_{Z(f-\pi)} |m|\,d\mu \quad \text{for all} \quad m \in M \sim \{0\} .$$

Proof. This is directly obtained from theorems 3.3 and 3.4. □

Theorems 2.10 and 2.11 can also be utilized to characterize strong uniqueness in L_1 . The following theorem does so, and is then used to give an alternate proof of theorem 3.4.

Looking back at 2.10 (or 2.11), we see that the set A_π was of some importance, and so we would like to have its counterpart in L_1 . The set F_π of the next definition is such a set, and its relation with A_π will be shown in the proof of the following theorem.

3.6. Definition. Let F_π be that set of elements m in the subspace M for which there exists an $a \in \mathbb{R}$ such that

$$(3.6.1) \quad \int_{Z(f-\pi)} |m + \pi + iaf|\,d\mu \le -\mathrm{Re}\int (m + \pi + iaf)\overline{\mathrm{sgn}}(f - \pi)\,d\mu \ .$$

3.7. Theorem. Let M be a finite dimensional subspace. Then $\pi \in P_r(M, f)$ if and only if F_π consists only of elements of the form $m = -(1 + id)\pi$ for some $d \in \mathbb{R}$.

Proof. Since M is finite dimensional theorems 2.10 and 2.11 combine to give a necessary and sufficient condition for the strong uniqueness of π. This condition, that A_π consists exactly of elements of the form $x = (1 + ia)(f - \pi)$, is related to the composition of F_π as follows:

Now $x \in A_\pi$ implies $x \in \mathrm{span}\{M, f\}$, so we need only consider those elements of K_π of the form $x = m + bf$, for some $m \in M$, and b a scalar. $x \in K_\pi$ if and only if

$$\mathrm{Re}\left[\int x\,\overline{\mathrm{sgn}}(f - \pi)\,d\mu + \int_{Z(f-\pi)} xg\,d\mu\right] \le \int |f - \pi|\,d\mu$$

for all measurable g such that $|g| \le 1$ almost everywhere on $Z(f - \pi)$. (See theorem 3.1. This is an application of the Riescz Representation theorem). Since this is true for all such g we can write equivalently

$$(3.7.1) \quad \mathrm{Re}\left[\int x\,\overline{\mathrm{sgn}}(f - \pi)\,d\mu + \int_{Z(f-\pi)} |x|\,d\mu\right] \le \int (f - \pi)\overline{\mathrm{sgn}}(f - \pi)\,d\mu \ .$$

Last of all, if $x \in A_\pi$, x must satisfy $\mathrm{Re}L_\pi(x) = ||f - \pi||$. By the definition of L_π, $\mathrm{Re}\, b = 1$, and therefore x must be of the form $m + (1 + ia)f$, $a \in \mathbb{R}$.

All these arguments work in reverse, so we can collect them all by stating

$$A_\pi = \{x = m + (1 + ia)f: a \in \mathbb{R},\ \mathrm{Re}\int_{Z(f-\pi)} |x|\,d\mu \le \mathrm{Re}\int (f - \pi - x)\overline{\mathrm{sgn}}(f - \pi)\,d\mu\} .$$

Note that (3.7.1) was rewritten slightly to get the form of the inequality in the statement of A_π . By replacing x by its explicit form, and noting that on $Z(f - \pi)$, $f = \pi$, we see that

$$A_\pi = \{m + (1 + ia)f: a \in \mathbb{R},\ \ (3.6.1) \text{ holds}\} .$$

Theorems 2.10 and 2.11 imply $\pi \in P_r(M, f)$ if and only if A consists exactly of elements of the form $(1 + id)(f - \pi) = -(1 + id)\pi + (1 + id)f$, $d \in \mathbb{R}$. We can trivially restate this as $F_\pi = \{m: (3.6.1) \text{ holds for some } a \in \mathbb{R}\}$ consists exactly of elements of the form $-(1 + id)\pi$ for $d \in \mathbb{R}$. □

3.8. <u>Corollary</u>. Let M be finite dimensional. If (3.4.1) holds then $\pi \in P_r(M, f)$.

<u>Proof</u>. If $\pi \notin P_r(M, f)$ then there exists $m \in F_\pi$ with $m \ne -(1 + ia)\pi$ for any $a \in \mathbb{R}$. Then $m + (1 + ia)\pi \ne 0$ for any $a \in \mathbb{R}$ and

$$\mathrm{Re}\int(-\pi - m - iaf)\overline{\mathrm{sgn}}(f - \pi)d\mu < \int_{Z(f-\pi)} |m + (1 + ia)\pi| d\mu$$
$$+ \mathrm{Re}\int ia(\pi - f)\overline{\mathrm{sgn}}(f - \pi)d\mu$$
$$= \int_{Z(f-\pi)} |m + \pi + iaf| d\mu - \mathrm{Re}(ia||f - \pi||)$$
$$= \int_{Z(f-\pi)} |m + \pi + iaf| d\mu \ .$$

This implies $m \notin F_\pi$, a contradiction. □

3.9. Theorem. Let M be finite dimensional. $\pi \in P_r(M, f)$ if and only if there exists measurable g defined on $Z(f - \pi)$ with $|g| \leq 1$ almost everywhere and

(3.9.1) $$\mathrm{Re}[\int m\ \overline{\mathrm{sgn}}(f - \pi)d\mu + \int_{Z(f-\pi)} mgd\mu] = 0 \quad \text{for all } m \in M ;$$

(3.9.2) $$\int_{Z(f-\pi)} |m| d\mu - \mathrm{Re}\int_{Z(f-\pi)} mgd\mu > 0 \quad \text{for all } m \in M \sim \{0\} .$$

Remark. (3.9.1) is the characterization for best approximation. (3.9.2) gives the strong uniqueness.

Proof. Assume there exists a g satisfying (3.9.1) and (3.9.2). Then

$$\int_{Z(f-\pi)} |m| d\mu + \mathrm{Re}\int m\ \overline{\mathrm{sgn}}(f - \pi) = \int_{Z(f-\pi)} |m| d\mu - \mathrm{Re}\int_{Z(f-\pi)} mgd\mu +$$
$$+ \mathrm{Re}[\int m\ \overline{\mathrm{sgn}}(f - \pi)d\mu + \int_{Z(f-\pi)} mgd\mu] > 0$$

for all $m \in M \sim \{0\}$, and corollary 3.5 implies $\pi \in P_r(M, f)$.

If $\pi \in P_r(M, f)$ then (3.9.1) holds for some g by the characterization theorem and

$$\int_{Z(f-\pi)} |m| - \mathrm{Re}\int_{Z(f-\pi)} mg\,d\mu = \int_{Z(f-\pi)} |m|\,d\mu + \mathrm{Re}\int m\,\overline{\mathrm{sgn}}(f-\pi)\,d\mu$$
$$- \mathrm{Re}\left[\int_{Z(f-\pi)} mg\,d\mu + \int m\,\overline{\mathrm{sgn}}(f-\pi)\,d\mu\right]$$
$$> 0$$

for all $m \in M \sim \{0\}$ where (3.9.1) and corollary 3.5 have been used for the last step. □

3.10. Corollary. $\pi \in P_r(M, f)$, where M is finite dimensional, if and only if the following hold.

(3.10.1) $\mu[Z(f-\pi) \sim Z(m)] > 0$ for all $m \in M \sim \{0\}$.

(3.10.2) There exists a $g \in L_\infty$ of norm 1, such that II-(4.4.1) and II-(4.4.2) hold.

(3.10.3) For any $m \in M \sim \{0\}$, $g \neq \overline{\mathrm{sgn}}(m)$ almost everywhere on $Z(f-\pi) \sim Z(m)$.

Note. We can apply corollary II-4.4 to see immediately that $\pi \in P(M, f)$. Therefore this corollary is a means of testing an already known best approximation for strong unicity.

Proof. (3.10.1) and (3.10.3) imply there exists a g (our g restricted to $Z(f-\pi)$) for which $|g| \le 1$ e.e. and (3.9.2) is true. Since M is a subspace we can rewrite II-(4.4.1) as

$$\mathrm{Re}\int mg\,d\mu = 0 \quad \text{for all} \quad m \in M .$$

Using II-(4.4.2) we get (3.9.1), thereby proving the sufficiency of the conditions.

If there exists a g as in theorem 3.9, then we can expand it to all of T by defining $g = \overline{\mathrm{sgn}}(f - \pi)$ on $T \sim Z(f - \pi)$. Then this g satisfies II-(4.4.1) and II-(4.4.2) by (3.9.1). (3.9.2) can hold only if

$$\int_{Z(f-\pi)} |m|\,d\pi \neq 0$$

for all $m \in M$, which gives us (3.10.1), and also

$$\int_{Z(f-\pi)} |m|\,d\mu \neq \mathrm{Re}\int_{Z(f-\pi)} mg\,d\mu$$

for all $m \in M \sim \{0\}$. Since $|g| \leq 1$ a.e. this is equivalent to (3.10.3), and the necessity is proved. □

It might be thought that this corollary is a bit overly restrictive, that possibly (3.10.3) could be derived from the previous conditions. Certainly this would be extremely fine if it was true, but unfortunately the following example shows that (3.10.3) is needed.

3.11. <u>Example</u>. Let T be the interval $[0, 6]$, and μ the standard Lebesgue measure. Let a typical element in M be of the form, for some $a \in \mathbb{R}$,

$$m(t) = \begin{cases} \frac{a}{2}\, t & 0 \le t \le 2 \\ a & 2 \le t \le 4 \\ \frac{a}{2}\,(6 - t) & 4 \le t \le 6 \end{cases}$$

Then take as f the function

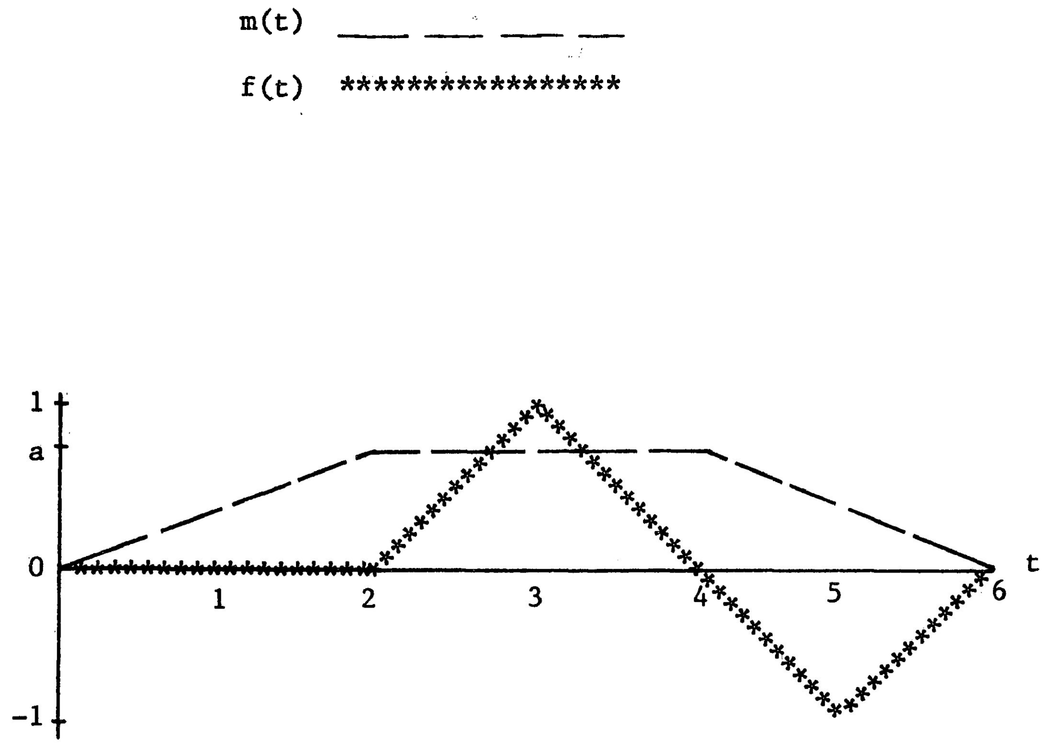
$$f(t) = \begin{cases} 0 & 0 \le t \le 2 \\ (t - 2) & 2 \le t \le 3 \\ (4 - t) & 3 \le t \le 5 \\ (t + 6) & 5 \le t \le 6 \end{cases}$$

(see graph 3.11).

Then we can show that 0 is the unique best approximation to f. Therefore $Z(f - \pi) = Z(f) = [0, 2]$, and, for $m \neq 0$, $Z(m) = \{0, 6\}$. We have $\mu[Z(f) \sim Z(m)] > 0$, as needed. A choice of $g(t) = 1$ for $t \in [0, 2]$, 1 for $t \in (2, 4)$, and -1 for $t \in (4, 6)$ satisfies all the conditions of corollary 3.10, so 0 is not strongly unique. (This can easily be verified.) This is also an example of a unique, but not strongly unique, best approximation, from a finite dimensional subspace.

In this discussion the quantity $\mu(Z(f - \pi))$ has been of some importance. It is related to the concept of smoothness considered in the beginning of this chapter. In L_1, f is a smooth point of the ball of radius $||f||$ if and only if $f(t) \neq 0$ almost everywhere (see [22, p. 350]). Bartelt first showed that for π to be a strongly unique best approximation we must have $\mu(Z(f - \pi)) > 0$

Graph 3.11.



[2, p. 8, theorem 6]. This follows quickly from theorem 3.3. Therefore $f - \pi$ cannot be a smooth point of the ball of radius $||f - \pi||$.

Finally we remark that all of the previous theorems hold in C_1 or the real case, as is easily shown. The proofs are identical.

S4. STRONGLY CHEBYSHEV SUBSPACES

4.1. Definition. A set K is strongly Chebyshev in X if every $f \in X \sim K$ has a strongly unique best approximation from K .

We immediately note that strongly Chebyshev subspaces are Chebyshev, and therefore by the theorem (III-5.3) of Phelps, there are no finite-dimensional strongly Chebyshev subspaces in L_1^R (when the measure is non-atomic). An example will be presented to show the existence of infinite dimensional strongly Chebyshev subspaces. Accordingly, the first theorem is not of much use, but it is an interesting application of some theorems in the last section. (However in the case when the measure is atomic the situation is different.)

4.2. Theorem. [45, p. 354, example 5] If M is a finite-dimensional Chebyshev subspace in $L_1^R(T, \mu)$ then M is strongly Chebyshev in $L_1^R(T, \mu)$.

<u>Proof</u>. Choose $f \in L_1^R \sim M$ arbitrarily. We can assume without loss of generality that $\{0\} = P(M, f)$. Then, by corollary II-4.5 there exists $g \in L_\infty^R$ with $||g||_\infty = 1$ and $g = \operatorname{sgn} f$ on $T \sim Z(f)$ such that $\int mg d\mu = 0$ for all $m \in M$ satisfying (3.9.1). We see that $\{t \in T: |g(t)| < 1\} \subseteq Z(f)$, so we can define $f' \in L_1^R$ by

$$f'(t) = \begin{cases} f(t) & \text{if } f(t) \neq 0 \text{ or } |g(t)| \neq 1 \\ g(t) & \text{if } f(t) = 0 \text{ and } |g(t)| = 1 . \end{cases}$$

It can be shown that $0 \in P(M, f')$ and $Z(f') = \{t \in T: |g(t)| < 1\}$. Since M is Chebyshev, theorem III-4.5 implies $\mu(Z(f') \sim Z(m)) > 0$ for all $m \in M \sim \{0\}$. Therefore

$$\begin{aligned} \int_{Z(f)} |m| d\mu &= \int_{Z(f) \sim Z(f')} |m| d\mu + \int_{Z(f')} |m| d\mu \\ &> \int_{Z(f) \sim Z(f')} mg d\mu + \int_{Z(f')} mg d\mu \\ &= \int_{Z(f)} mg d\mu \end{aligned}$$

and (3.9.2) is satisfied; so by theorem 3.9 0 is a strongly unique best approximation to f . Since f is arbitrary M is strongly Chebyshev. □

4.3. <u>Example</u>. A strongly Chebyshev subspace of infinite dimension in L_1 .

Consider example II-5.4. In that example

$$||f - m|| = \int_B |f| d\mu + \int_{T\sim B} |f - m| d\mu \ , \ ||f - \pi|| = \int |f| d\mu$$

and

$$||\pi - m|| = \int_{T\sim B} |f - m| d\mu \ .$$

Choose $r = 1$. Then $||f - \pi|| + r||\pi - m|| = ||f - m||$ for all $m \in M$ so π is a strongly unique best approximation to f . Since f was arbitrary the subspace is strongly Chebyshev.

Since there do not exist any finite dimensional Chebyshev spaces in L_1 where the measure is non-atomic, Wulbert's theorem is not very useful in this thesis.

Since Haar subspaces are Chebyshev in $C_1^R[a, b]$ it might be conjectured that they are strongly Chebyshev. This is trivially not true in the case of polynomials. As the next theorem shows, no Haar subspace can be strongly Chebyshev in $C_1^R[a, b]$. It is based on the following lemma.

4.4. Lemma. If $\{\phi_i, i = 1, \ldots, n\}$ is a Haar system on an interval $[a, b]$ then there exists ϕ_{n+1} such that $\{\phi_i; i = 1, \ldots, n + 1\}$ is also a Haar system.

Proof. The proof is rather long and complicated and is the subject of a paper by R. A. Zalik. See [46, p. 72, theorem 1]. □

4.5. Theorem. If M is a Haar subspace in $C_1^R[a, b]$ then M is not strongly Chebyshev.

<u>Proof</u>. Let $M = \{\phi_1 , \ldots, \phi_n\}$ and choose f to be the function ϕ_{n+1} guaranteed by lemma 4.3 and find the best approximation π to f from M. By the assumption π exists and is unique. Since $\text{span}\{M, f\}$ is a Haar subspace $f - \pi$ can have at most n zeroes, which implies $\mu(Z(f - \pi) = 0$ if the measure is non-atomic. Therefore (3.10.1) is violated and π is not strongly unique, and hence M is not strongly Chebyshev. □

Are there any finite dimensional subspaces which are strongly Chebyshev in C_1^R ? This is still an open question. On the other hand the Mazur density theorem states that the set of smooth points of the closed ball in any separable B-space is a dense (in fact, a residual) subset of its boundary [15, p. 171], which may suggest that one can always find a point f whose smoothness precludes the existence of a strongly unique best approximation.

Chapter V

CONCLUSIONS

To the theorist the subject of best L_1-approximation from linear subspaces seems tolerably complete, although there are some holes which we will mention shortly. To the more practical investigator there are two omissions; how do we find a best approximation in specific cases, and what is the precision of the approximation?

Neither of these questions have been addressed in any detail. These probably form a subject matter which deserve another survey. Algorithms to determine best approximations have been produced by Barrodale and Young [1], Deutsch, McCabe and Phillips [11], and Usow [44], to mention a few.

Precision is involved with $\rho(f, K)$. Can we put lower bounds, or upper bounds on this quantity? Some fine work has been done with uniform approximation involving the study of H-sets [13], but this has not been carried over to L_1 approximation.

The L_1 norm seems to exult in such behaviour, allowing little of the theory of uniform approximation to be carried over. We have obtained characterization theorems which are useful, but they have none of the elegance of the simple alternation theorems of the uniform norm. If one uses, for instance, II-(4.6.1) to test an element then for each $m \in M$ two integrations must be carried out. The original definition, $\int |f - \pi| d\mu \le \int |f - m| d\mu$, implies only one integration must be completed, so here the characterization theorem has only succeeded in complicating matters.

Theorem II-4.9 uses an alternation condition which looks pretty, especially with algebraic and trigonometric polynomial approximation where the nodal points are known. Unfortunately the condition is sufficient, but not necessary.

The same comments can be applied to uniqueness. Theorem III-2.1 is useful in that a test for best approximation using II-(4.8.1) also tests the criterion III-(2.1.1) for uniqueness. However, this criterion is not a necessary one. A necessary and sufficient condition is presented in theorem III-2.4, but here the range of integration as well as the integrand is changed for each element, making III-(2.4.1) rather cumbersome.

That beautiful result in the continuous uniform case, that subspaces are Chebyshev if and only if they are Haar, is not repeated in C_1^R . Jackson's theorem (III-6.6) is half the result, but the converse is not true. There are no useful sufficient conditions for Chebyshev sets in L_1 or C_1 . Some interesting properties are evident, however. For instance theorem III-4.5 is a useful test.

Very little work has been done studying strong uniqueness in L_1 . Wulbert's result is the most interesting, but not very pertinent to the non-atomic case, and it does not carry over to C_1 .

The existence of strongly Chebyshev subspaces was considered but the results are mostly negative. In the uniform case all Haar subspaces are strongly Chebyshev in $C[a, b]$, but this result is not true in C_1 . In fact, no Haar subspace is strongly Chebyshev. We conjectured that no finite dimensional subspace in C_1 or C_1^R is strongly Chebyshev, but this has yet to be proved.

We can see that there is lots of scope for more work, and hope that this thesis has indicated some of the directions in which future investigations can go.

Bibliography

[1] I. Barrodale and A. Young, "Algorithms for best L_1 and L_∞ linear approximation on a discrete set", Numerische Mathematik 8 (1966), pp. 295-306.

[2] W. M. Bartelt, "Strongly unique best approximates to a function on a set, and a finite set there of", Pacific J. Math. 53 (1974), pp. 1-9.

[3] M. W. Bartelt and H. W. McLaughlin, "Characterizations of strong unicity in approximation theory", J. Approx. Theory 9 (1973), pp. 225-266.

[4] R. Bojanic and R. DeVore, "On polynomials of best one-sided approximation", L'Enseignement Math. 12 (1966), pp. 139-164.

[5] K. Borsuk, "Drei Sütze über die n-dimensionale Euklidische Sphäre", Fund. Math. 20 (1933), pp. 177-190.

[6] M. P. Carrol and H. W. McLaughlin, "On L_1 approximation of discontinuous functions", J. Approx. Theory 8 (1973), pp. 129-132.

[7] E. W. Cheney, Introduction to Approximation Theory, McGraw-Hill, New York, 1966.

[8] E. W. Cheney and D. E. Wulbert, "The existence and unicity of best approximations", Math. Scand. 24 (1969), pp. 113-140.

[9] F. R. Deutsch, "On uniform approximation with interpolatory constraints", J. Math. Anal. Appl. 24 (1968), pp. 62-79.

[10] F. R. Deutsch and P. H. Maserick, "Applications of the Hahn-Banach theorem in approximation theory", SIAM Review 9, no. 3, (July, 1967), pp. 516-530.

[11] F. Deutsch, J. H. McCabe, and G. M. Phillips, "Some algorithms for computing best approximations from convex cones", SIAM J. Numer. Anal. 12, No. 3 (June, 1975), pp. 390-403.

[12] R. DeVore, "One-sided approximation of functions", J. Approx. Theory 1 (1968), pp. 11-25.

[13] C. Dierieck, "Characterization for best nonlinear approximations: A geometrical interpretation", J. Approx. Theory 14 (1975), pp. 163-187.

[14] ___________, "Some remarks on H-sets in linear approximation theory", J. Approx. Theory 2 (1977), pp. 188-204.

[15] R. J. Duffin and L. A. Karlovitz, "Formulation of linear programs in analysis I: Approximation theory", SIAM J. Appl. Math. 16, no. 4, (1968), pp. 662-675.

[16] N. Dunford and J. Schwartz, Linear Operators, Vol. 1, Interscience Publishers Inc., New York, 1966.

[17] E. Hewitt and K. Stromberg, Real and Abstract Analysis. Springer-Verlag, New York, 1965.

[18] C. R. Hobby and J. R. Rice, "A moment problem in L_1 approximation", Proc. AMS 16 (1965), pp. 665-710.

[19] R. B. Holmes, Geometric Functional Analysis and its Applications. Springer-Verlag, New York, 1975.

[20] D. Jackson, "A general class of problems in approximation", Amer. J. of Math. 46 (1924), pp. 215-234.

[21] S. Karlin and W. J. Studden, Tchebycheff Systems: With Applications in Analysis and Statistics. Interscience, New York, 1966.

[22] G. Köthe, Topological Vector Spaces. trans. O. J. H. Garling, Springer-Verlag, New York, 1969.

[23] W. Krabs, "Duality in nonlinear approximation", J. Approx. Theory, 2 (1969), pp. 136-151.

[24] B. R. Kripke and T. J. Rivlin, "Approximation in the metric of $L'(X, \mu)$", Trans. AMS 119 (1965), pp. 101-122.

[25] J. T. Lewis, "Approximation with convex constraints", SIAM Review 15, no. 1 (Jan., 1973), pp. 193-217.

[26] J. Lindenstrauss, "A short proof of Liapounoff's convexity theorem", J. Math. Mech. 15 (1966), pp. 971-972.

[27] G. G. Lorentz, ed. Approximation Theory, Proc. Sympos. Univ. Texas, Academic Press, 1973.

[28] G. G. Lorentz, C. K. Chui, and L. L. Schumaker, editors, Approximation Theory II: Proc. of an International Symposium, University of Texas at Austin, Academic Press, New York, 1976.

[29] M. G. Meyerson and A. H. Wright, "A new and constructive proof of the Borsuk-Ulam theorem", Proc. AMS 73 (1979), pp. 134-136.

[30] C. A. Micchelli, "Best L_1 approximation by weak Chebyshev systems and the uniqueness of interpolating perfect splines", J. Approx. Theory 19 (1977), pp. 1-14.

[31] J. R. Munkres, Topology, a First Course, Prentice-Hall, New Jersey, 1975.

[32] D. D. Pence, "Best approximation by splines satisfying generalized convexity constraints", Univ. Wisconsin-Madison Math. Research Center, Tech. Report no. 1794, 1977.

[33] R. R. Phelps, "Uniqueness of Hahn-Banach extensions and unique best approximation", Trans. AMS 94 (1960), pp. 238-255.

[34] A Pinkus, "A simple proof of the Hobby-Rice theorem", Proc. AMS, 60 (1976), pp. 82-84.

[35] V. Pták, "On approximation of continuous functions in the metric $\int_a^b |x(t)|dt$", Czechoslovak Math J. 8 (1958), pp. 267-273.

[36] J. R. Rice, "Best approximations and interpolating functions", Trans. AMS 101 (1961), pp. 477-498.

[37] J. R. Rice, The Approximation of Functions, Vol. I., Addison-Wesley, Reading, Mass., 1964.

[38] A. W. Roberts and D. E. Varberg, Convex Functions, Academic Press, New York, 1973.

[39] H. L. Royden, Real Analysis, MacMillan, New York, 1968.

[40] I. Singer, Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces, Springer-Verlag, New York, 1970.

[41] _________, "On the extension of continuous linear functionals and best approximations in normed linear spaces", Math. Ann. 159 (1965), pp. 344-355.

[42] _________, The Theory of Best Approximation and Functional Analysis, SIAM Regional Conference Ser. Appl. Math. 13, SIAM, Philadelphia, 1974.

[43] A. Talbot, ed., Approximation Theory, Proc. Symp. Lancaster, July, 1969. Academic Press, 1970.

[44] K. H. Usow, "On L_1 approximation I: Computation for continuous functions and continuous dependence", SIAM J. Numer. Anal. 4, no. 1 (1967), pp. 70-78.

[45] D. Wulbert, "Uniqueness and differential characterization of approximations from manifolds of functions", Amer. J. Math., 93 (1971), pp. 350-366.

[46] R. A. Zalik, "Existence of Chebyshev extensions", J. Math. Anal. and Appl. 51 (1975), pp. 68-75.