Strong Uniqueness

András Kroó and Allan Pinkus

January 18, 2010

Abstract. This is a survey paper on the subject of strong uniqueness in approximation theory.

MSC: 41A52, 41A50, 41A65

Keywords: strong uniqueness, best approximation, uniqueness

0. Foreword

This is a survey paper on the subject of *Strong Uniqueness* in approximation theory. The concept of strong uniqueness was introduced by Newman, Shapiro in 1963. They proved, among other things, that if M is a finite-dimensional real Haar space in $C(B)$, B a compact Hausdorff space, and if m^* is the best approximant to f from M, then there exists a $\gamma > 0$ such that

$$\|f - m\| - \|f - m^*\| \geq \gamma \|m - m^*\| \tag{0.1}$$

Surveys in Approximation Theory
Volume 5, 2010, pp. 1–91.

ISSN 1555-578X

for all $m \in M$, where γ may depend upon f, M and m^*, but is independent of the specific m. An inequality of the above form valid for all $m \in M$ is what we call *classical strong uniqueness*. This property is stronger than the uniqueness of the best approximant (hence the name).

Strong uniqueness has been much studied, mainly during the 1970s, 1980s and 1990s, and there have been over 100 research papers devoted to the subject. We thought this might be a good time to reflect upon what was considered and accomplished in the study of this subject. We hope that you will agree that the resulting theory is not insignificant.

This paper is organized as follows. We have divided the survey into three parts. The first part is concerned with what we termed above classical strong uniqueness, i.e., inequality (0.1). In Section 1, we exactly characterize the optimal strong uniqueness constant (largest possible γ in (0.1)) via the one-sided Gateaux derivative. In Section 2, we consider some general results regarding classical strong uniqueness in the uniform norm. We look in detail at the case where the approximating set is a finite-dimensional subspace, give a characterization of when we have strong uniqueness, and also prove that in this setting the set of functions with a strongly unique best approximant is dense in the set of functions with a unique best approximant. Finally, we provide a general upper bound on the optimal strong uniqueness constant based on projection constants. In Section 3, we consider the relationship between local Lipschitz continuity of the best approximation operator at a point and classical strong uniqueness at the same point. A general result is that the latter always implies the former. We prove that the converse holds when approximating from a finite-dimensional subspace in the uniform norm. In Section 4, we restrict our attention to approximation from finite-dimensional Haar spaces in the uniform norm. We first prove the result of Newman, Shapiro [1963] mentioned above, i.e., that in this case, we always have strong uniqueness. We then consider various properties of the optimal strong uniqueness constant with respect to the function being approximated. We prove, for example, that the optimal strong uniqueness constant is upper semi-continuous but not necessarily continuous, and it is not uniformly bounded below by a positive constant (i.e., bounded away from 0) if the subspace is of dimension at least 2 and the underlying domain is not discrete. We find lower bounds (that sometimes provide equality) for the optimal strong uniqueness constant via specific elements of the approximating subspace. Assuming the subspace is of dimension n and the error function only attains its norm at $n+1$ points (the generic case), we obtain that the optimal strong uniqueness constant is bounded above by $1/n$. We also give another characterization of the optimal strong uniqueness constant and consider the question of when strong uniqueness and uniqueness are equivalent concepts. In Section 5, we look at classical strong uniqueness in the L^1 norm. Among other results, we prove that if ν is a non-atomic positive measure then the set of functions in $L^1(K,\nu)$ that have a strongly unique best approximant from any finite-dimensional subspace is dense in $L^1(K,\nu)$. In addition, under the above assumptions, we show, as in the uniform norm, that a function in $L^1(K,\nu)$ has a strongly unique best approximant if and only if the best approximation operator from the same finite-dimensional subspace is locally Lipschitz continuous. Furthermore, in this case, we get explicit upper and lower bounds for the optimal strong uniqueness constant based on the Lipschitz continuity of the best approximation operator. We also present some results on classical strong uniqueness in the problem of one-sided L^1 approximation. In Section 6, we consider approximation by rational functions of the form

$$R_{m,n} := \{r = p/q : p \in \Pi_m, q \in \Pi_n, q(x) > 0, x \in [a,b]\},$$

where $\Pi_n = \operatorname{span}\{1, x, \ldots, x^n\}$. The main result reported on is that we have the equivalence of strong uniqueness to a function f from $R_{m,n}$, the operator of best rational approximation from $R_{m,n}$ being continuous at f, and the fact that the unique best rational approximant to f from $R_{m,n}$ is not contained in $R_{m-1,n-1}$.

In the second part of this survey, we discuss what we call *non-classical strong uniqueness*. By this, we mean the existence of a nonnegative strictly increasing function ϕ defined on $\mathbb{R}_+$, and a constant $\gamma > 0$ that may depend upon f and M (and thus on m^*), for which

$$\|f - m\| - \|f - m^*\| \geq \gamma\phi(\|m - m^*\|)$$

for all $m \in M$ (a global estimate), or for all $m \in M$ such that

$$\|f - m\| - \|f - m^*\| \leq \sigma$$

for some $\sigma > 0$ (a local estimate). We start in Section 7 with the case of a uniformly convex norm and prove the basic inequality

$$\|f - m\| - \|f - m^*\| \geq \|f - m\| \, \delta\left(\frac{\|m - m^*\|}{\|f - m\|}\right)$$

valid for all $m \in M$ where $m^* \in M$ is a best approximant to f. Here δ is the usual modulus of convexity of the uniformly convex space. We apply this result and consider also variants thereof. In Section 8, we return to a consideration of the uniform norm. The main result we report on therein is that non-classical strong uniqueness holds for $\phi(t) = t^2$ if $M \subset C^2[a, b]$ is a finite-dimensional unicity space (and not necessarily a Haar space) with respect to $C^2[a, b]$. In Section 9, we return to a consideration of the L^1-norm and obtain two local non-classical strong uniqueness estimates. In one case, we prove that if M is a finite-dimensional unicity space for all $f \in C[a, b]$ in the $L^1[a, b]$ norm then we have non-classical strong uniqueness with $\phi(t) = t\omega^{-1}(f; Dt)$ for some constant $D > 0$ that depends only on f and M. Here ω is the standard modulus of continuity. See Theorem 9.2 for a more detailed explanation. A second result concerns the one-sided L^1 approximation problem. We prove that if $M \subset C^1[a, b]$ is a finite-dimensional unicity subspace for $C^1[a, b]$ in the one-sided L^1-norm, then we have non-classical strong uniqueness with $\phi(t) = tH_f^{-1}(Dt)$ for some constant $D > 0$ that depends only on f and M, where H is based on the moduli of continuity of f' and m' for all $m \in M$. In Section 10, we consider strong uniqueness when we approximate complex-valued functions in the uniform norm. The main result therein is non-classical strong uniqueness with $\phi(t) = t^2$. We also discuss conditions under which classical strong uniqueness holds.

The third part of this survey concerns various applications of strong uniqueness results. The main idea behind these applications is that instead of solving a best approximation problem in a given norm, we replace it by considering another norm, close to the original norm, one that leads to a simpler approximation problem. Strong uniqueness is then applied in order to show that the best approximant in this new norm is sufficiently close to the original best approximant. An estimation of the "closeness" is given in Section 11. Typically, the original norm is modified by replacing it by a similar discrete norm (as considered in Section 12), or by introducing a weight function into the norm. In Section 13, we use strong uniqueness results in order to solve approximation problems in the case when the norm is altered by a weight function. This approach is used to derive asymptotic representations for weighted Chebyshev polynomials.

We have tried to make this survey as self-contained as seemed reasonable. As such there are many additional peripheral results included in this paper that, we hope, put the strong uniqueness results into a reasonable context.

On the other hand, while this survey is lengthy, it is far from complete. Certain specific problems and general topics within the area of strong uniqueness have not been surveyed. For example, a lot of effort and many papers have been concerned with the study of $\gamma_n(f)$, the optimal strong uniqueness constant (in (0.1)) when approximating $f \in C[a,b]$ in the uniform norm from $\Pi_n = \text{span}\{1, x, \dots, x^n\}$. Many different questions were asked concerning this important case. Nevertheless, we have not considered this class of problems. Thus, as is noted in Section 4, Poreda [1976] raised the question of describing the asymptotic behaviour of the above sequence $(\gamma_n(f))$ as n tends to ∞ for a given function f. Henry and Roulier [1978] conjectured that

$$\liminf_{n\to\infty} \gamma_n(f) = 0$$

if f is not a polynomial. This conjecture, sometimes called the Poreda conjecture, was proved for various classes of functions over a string of papers. It was finally solved in the positive by Gehlen [1999]. It should be noted that the above lim inf cannot be replaced by the simple limit. There are functions $f \in C[a,b]$ for which

$$\liminf_{n\to\infty} \gamma_n(f) = 0, \qquad \limsup_{n\to\infty} \gamma_n(f) = 1;$$

see Schmidt [1978]. These same $\gamma_n(f)$ were also studied for specific functions f, or specific classes of function; see for example Henry, Huff [1979], Henry, Swetits [1981], Henry, Swetits, Weinstein [1981], Henry, Swetits [1982] and Henry, Swetits, Weinstein [1983]. There are also papers devoted to looking for sets in $C[a,b]$ over which the strong uniqueness constants are uniformly bounded below by a positive constant; see for example Bartelt, Swetits [1983], Marinov [1983], and Bartelt, Swetits [1988], as well as papers devoted to looking at the optimal strong uniqueness constant as a function of the domain; see for example Henry, Roulier [1977], Bartelt, Henry [1980], and Paur, Roulier [1981]. For a study of strong uniqueness when approximating with constraints, leading to non-classical strong uniqueness, see for example Fletcher, Roulier [1979], Schmidt [1979], Chalmers, Metcalf, Taylor [1983] and Kroó, Schmidt [1991]. All the above is with regards to approximation from Π_n. In addition, strong uniqueness when approximating by splines, with either fixed or variable knots, was considered in Nürnberger, Singer [1982], Nürnberger [1982/83], Nürnberger [1994], Sommer, Strauss [1993], Zeilfelder [1999] and Zwick [1987]. Again this is not discussed in what follows.

We have tried to give exact references, and have also endeavored to make the list of references complete. This list contains all references we have found on the subject of strong uniqueness. Note that not all papers on the list of references are referred to in the body of this survey. We apologize for any omissions and would appreciate information with regard to any additional references.

Part I. Classical Strong Uniqueness

1. Introduction

Let X be a normed linear space and M a subset of X. For any given $f \in X$, we let $P_M(f)$ denote the set of best approximants to f from M. That is, $m^* \in P_M(f)$ if $m^* \in M$ and

$$\|f - m^*\| \le \|f - m\|$$

for all $m \in M$. This set $P_M(f)$ may be empty (non-existence), a singleton (unicity) or larger. In this section, we consider conditions on a linear subspace or convex subset M of X under which we obtain *Classical Strong Uniqueness*. By this, we mean the existence of a strictly positive constant $\gamma > 0$ for which

$$\|f - m\| - \|f - m^*\| \geq \gamma \|m - m^*\|$$

for all $m \in M$ where $m^* \in P_M(f)$. The constant $\gamma > 0$ can and will depend upon f, m^* and M, but must be independent of m. Note that the form of this inequality is the best one might expect since, by the triangle inequality, we always have

$$\|f - m\| - \|f - m^*\| \leq \|m - m^*\|.$$

We first exactly characterize the optimal (largest) γ for which such an inequality holds. This will be done via the one-sided Gateaux derivative that is defined as follows. Given $f, g \in X$, set

$$\tau_+(f, g) = \lim_{t \to 0^+} \frac{\|f + tg\| - \|f\|}{t}. \tag{1.1}$$

The value $\tau_+(f, g)$ is termed the one-sided Gateaux derivative of the norm at f in the direction g. The first thing we prove is that this one-sided Gateaux derivative always exists.

Proposition 1.1. *The one-sided Gateaux derivative of the norm at f in the direction g, namely $\tau_+(f, g)$, exists for every $f, g \in X$.*

Proof: Set

$$r(t) := \frac{\|f + tg\| - \|f\|}{t}.$$

We claim that $r(t)$ is a non-decreasing function of t on $(0, \infty)$ and is bounded below thereon. As this is valid then $\tau_+(f, g)$ necessarily exists.

To see that $r(t)$ is bounded below on $(0, \infty)$, note that

$$\|f + tg\| \geq \|f\| - \|tg\| = \|f\| - t\|g\|.$$

Thus for $t > 0$, we have $r(t) \geq -\|g\|$. The non-decreasing property of r can be shown as follows. Let $0 < s < t$. Then

$$t\|f + sg\| = \|tf + tsg\| = \|s(f + tg) + (t - s)f\| \leq s\|f + tg\| + (t - s)\|f\|.$$

Thus

$$t(\|f + sg\| - \|f\|) \leq s(\|f + tg\| - \|f\|),$$

and that implies $r(s) \leq r(t)$. $\square$

Using this functional τ_+, it is now a simple matter to characterize best approximants from linear subspaces. Namely,

Theorem 1.2. *Let M be a linear subspace of X. Then $m^* \in P_M(f)$ if and only if $\tau_+(f - m^*, m) \geq 0$ for all $m \in M$.*

Proof: $(\Rightarrow)$ Assume m^* is a best approximant to f from M. Therefore

$$\|f - m^* + tm\| \geq \|f - m^*\|$$

for every $m \in M$ and $t > 0$, immediately implying that $\tau_+(f - m^*, m) \geq 0$.

$(\Leftarrow)$ Assume $\tau_+(f - m^*, m) \geq 0$ for all $m \in M$. As M is a linear subspace, this implies that $\tau_+(f - m^*, m^* - m) \geq 0$. From the proof of Proposition 1.1, $r(t)$ is a non-decreasing function of t on $(0, \infty)$. Setting $t = 1$ therein with respect to $\tau_+(f - m^*, m^* - m)$, we obtain

$$\|f - m^* + m^* - m\| - \|f - m^*\| \geq \tau_+(f - m^*, m^* - m) \geq 0.$$

Thus $\|f - m\| \geq \|f - m^*\|$ for all $m \in M$. $\square$

If M is a convex subset of X, then the same arguments prove the following.

Corollary 1.3. *Let M be a convex subset of X. Then $m^* \in P_M(f)$ if and only if $\tau_+(f - m^*, m^* - m) \geq 0$ for all $m \in M$.*

If $\tau_+(f - m^*, m^* - m) > 0$ for all $m \in M$, $m \neq m^*$, then we may have a $\gamma > 0$ for which

$$\|f - m\| - \|f - m^*\| \geq \gamma\|m - m^*\|$$

for all $m \in M$. When there exists a $\gamma > 0$ for which

$$\|f - m\| - \|f - m^*\| \geq \gamma\|m - m^*\|$$

for all $m \in M$, then we say that m^* is a *strongly unique* best approximant to f from M. The reason for this terminology is simply that strong uniqueness is stronger than uniqueness. That is, if m^* is a strongly unique best approximant to f from M, then it is certainly a unique best approximant. While the converse does not hold in general, it does and can hold in various settings. This concept was introduced in Newman, Shapiro [1963] with respect to certain specific spaces. We will state and prove their results in later sections.

This next result characterizes the optimal (largest) γ for which such an inequality holds. We prove this result under the assumption that M is a linear subspace. The case where M is only a convex subset follows analogously, and we will subsequently formally state it as a corollary.

Theorem 1.4. *Let M be a subspace of X and $f \in X \backslash M$. Assume $m^* \in P_M(f)$. Set*

$$\gamma(f) := \inf\{\tau_+(f - m^*, m) : m \in M, \|m\| = 1\}.$$

Then $\gamma(f) \geq 0$ and, for all $m \in M$, we have

$$\|f - m\| - \|f - m^*\| \geq \gamma(f)\, \|m - m^*\|.$$

Furthermore, if $\gamma > \gamma(f)$ then there exists an $\widetilde{m} \in M$ for which

$$\|f - \widetilde{m}\| - \|f - m^*\| < \gamma\, \|\widetilde{m} - m^*\|.$$

Proof: As $\tau_+(f - m^*, m^* - m) \geq 0$ for all $m \in M$ (M is a subspace), we have $\gamma(f) \geq 0$ from Theorem 1.2. Assume $\gamma(f) > 0$. From the fact that $\tau_+(f, cg) = c\,\tau_+(f, g)$ for all $c > 0$, it follows that

$$\tau_+(f - m^*, m^* - m) \geq \gamma(f)\|m - m^*\|$$

for all $m \in M$. From the proof of Proposition 1.1,

$$\frac{\|f - m^* + t(m^* - m)\| - \|f - m^*\|}{t} \geq \tau_+(f - m^*, m^* - m)$$

for all $t > 0$. Setting $t = 1$, we obtain

$$\|f - m\| - \|f - m^*\| \geq \tau_+(f - m^*, m^* - m) \geq \gamma(f)\|m - m^*\|$$

for all $m \in M$.

Now assume $\gamma > \gamma(f)$. By definition, there exists an $\overline{m} \in M$, $\|\overline{m}\| = 1$, for which

$$\tau_+(f - m^*, \overline{m}) < \gamma \|\overline{m}\|.$$

Thus for $t > 0$ sufficiently small, we have

$$\frac{\|f - m^* + t\overline{m}\| - \|f - m^*\|}{t} < \gamma \|\overline{m}\|,$$

implying

$$\|f - m^* + t\overline{m}\| - \|f - m^*\| < \gamma \|t\overline{m}\|.$$

Setting $\widetilde{m} := m^* - t\overline{m}$, we obtain

$$\|f - \widetilde{m}\| - \|f - m^*\| < \gamma \|\widetilde{m} - m^*\|. \quad \square$$

For a convex set M, this same reasoning gives:

Theorem 1.5. *Let M be a convex subset of X and $f \in X \backslash M$. Assume $m^* \in P_M(f)$. Set*

$$\gamma(f) := \inf \left\{ \frac{\tau_+(f - m^*, m^* - m)}{\|m^* - m\|} : m \in M, m \neq m^* \right\}.$$

Then $\gamma(f) \geq 0$ and for all $m \in M$, we have

$$\|f - m\| - \|f - m^*\| \geq \gamma(f)\, \|m - m^*\|.$$

Furthermore, if $\gamma > \gamma(f)$ then there exists an $\widetilde{m} \in M$ for which

$$\|f - \widetilde{m}\| - \|f - m^*\| < \gamma\, \|\widetilde{m} - m^*\|.$$

In much of this paper, we consider approximation from linear subspaces and not from convex subsets. However, most of the results obtained can be easily generalized to convex subsets, as above. We leave their exact statements to the interested reader.

Note that if M is a finite-dimensional subspace (or closed convex subset thereof) then a compactness argument implies that the *inf* in the definition of $\gamma(f)$ in both Theorems 1.4 and 1.5 is in fact a *min*. In the former case, we rewrite an essential consequence of Theorem 1.4 as:

Proposition 1.6. *Let M be a finite-dimensional subspace of X and $f \in X \backslash M$. Assume $m^* \in P_M(f)$. Then m^* is a strongly unique best approximant to f from M, i.e., $\gamma(f)$, as defined in Theorem 1.4, is strictly positive, if and only if there does not exist an $m \in M$, $m \neq 0$, for which $\tau_+(f - m^*, m) = 0$.*

For many normed linear spaces, we have that

$$\tau(f, g) = \lim_{t \to 0} \frac{\|f + tg\| - \|f\|}{t}$$

exists for all $f, g \in X$, $f \neq 0$. (Note that this is the two-sided limit that need not in general exist.) If this is the case then, for any linear subspace M, we have that $m^* \in P_M(f)$ if and only if

$$\tau(f - m^*, m) = 0$$

for all $m \in M$. Thus, we never have classical strong uniqueness in such spaces. If the functional $\tau(f, g)$ exists for all $f, g \in X$, $f \neq 0$, then the space X is said to be *smooth*. Strong uniqueness cannot hold in smooth spaces as strong uniqueness implies that the unit ball has corners. The space X being smooth is equivalent to the existence, for each $f \in X$, $f \neq 0$, of a unique continuous linear functional $h \in X^*$ of norm 1 satisfying $h(f) = \|f\|$. This is true, for example, in any Hilbert space and in L^p for every $p \in (1, \infty)$. In all these cases, we must look for other non-classical type strong uniqueness formulae.

The approach that was taken in this section to characterizing best approximants and classical strong uniqueness may be found in Papini [1978] and Wulbert [1971].

2. Classical Strong Uniqueness in the Uniform Norm

In this section, we assume that B is a compact Hausdorff space and consider $C(B)$, the normed linear space of continuous real-valued functions defined on B, with norm

$$\|f\| = \max_{x \in B} |f(x)|.$$

For each $f \in C(B)$, set $A_f = \{x : |f(x)| = \|f\|\}$ and

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

The formula for the one-sided Gateaux derivative $\tau_+(f, g)$ in $C(B)$ is the following.

Theorem 2.1. *For $f, g \in C(B)$, $f \neq 0$, we have*

$$\tau_+(f, g) = \max_{x \in A_f} [\operatorname{sgn} f(x)]\, g(x).$$

Proof: We first prove that the right-hand-side is a lower bound for $\tau_+(f, g)$. To this end, let $x \in A_f$. Thus $|f(x)| = \|f\|$ and $|f(x) + tg(x)| \leq \|f + tg\|$. Therefore

$$\begin{aligned} \tau_+(f, g) &= \lim_{t \to 0^+} \frac{\|f + tg\| - \|f\|}{t} \\ &\geq \lim_{t \to 0^+} \frac{|f(x) + tg(x)| - |f(x)|}{t} \\ &= \lim_{t \to 0^+} \frac{[\operatorname{sgn} f(x)]\,(f(x) + tg(x)) - [\operatorname{sgn} f(x)]\, f(x)}{t} \\ &= \lim_{t \to 0^+} \frac{t[\operatorname{sgn} f(x)]\, g(x)}{t} \\ &= [\operatorname{sgn} f(x)]\, g(x). \end{aligned}$$

This implies that

$$\tau_+(f, g) \geq \max_{x \in A_f} [\operatorname{sgn} f(x)]\, g(x).$$

We prove the converse direction as follows. For each $t > 0$, let x_t satisfy

$$\|f + tg\| = |f(x_t) + tg(x_t)|.$$

As B is compact, there exists an $x^* \in B$ that is a limit point of the x_t as $t \to 0$. Let t_n denote a sequence, decreasing to zero, along which $x_{t_n} := x_n$ converges to x^*. We first claim that we must have $x^* \in A_f$. If not, then

$$\|f + t_n g\| = |f(x_n) + t_n g(x_n)| \to |f(x^*)| < \|f\|,$$

contradicting the continuity of the norm. As $x^* \in A_f$, it therefore follows that for n sufficiently large,

$$\operatorname{sgn}(f(x_n) + t_n g(x_n)) = \operatorname{sgn} f(x^*).$$

Thus

$$\frac{\|f+t_ng\|-\|f\|}{t_n} = \frac{[\operatorname{sgn} f(x^*)][f(x_n)+t_ng(x_n)]-[\operatorname{sgn} f(x^*)]f(x^*)}{t_n}$$
$$= [\operatorname{sgn} f(x^*)]\left[\frac{f(x_n)-f(x^*)}{t_n}\right] + [\operatorname{sgn} f(x^*)]g(x_n).$$

Note that

$$[\operatorname{sgn} f(x^*)]\left[\frac{f(x_n)-f(x^*)}{t_n}\right] \le 0$$

since $t_n > 0$ and $[\operatorname{sgn} f(x^*)]f(x_n) \le \|f\| = |f(x^*)| = [\operatorname{sgn} f(x^*)]f(x^*)$. Take limits on both sides of the above equality. The left-hand-side has the finite limit $\tau_+(f,g)$, the first term on the right-hand-side is nonpositive while

$$\lim_{n\to\infty}[\operatorname{sgn} f(x^*)]\, g(x_n) = [\operatorname{sgn} f(x^*)]\, g(x^*) \le \max_{x\in A_f}\, [\operatorname{sgn} f(x)]\, g(x).$$

Thus

$$\tau_+(f,g) \le \max_{x\in A_f}\, [\operatorname{sgn} f(x)]\, g(x). \quad \square$$

As a consequence of Theorem 2.1, Theorem 1.2 and Theorem 1.4, we have:

Theorem 2.2. *Let M be a linear subspace of $C(B)$. Then $m^* \in P_M(f)$ if and only if*

$$\tau_+(f-m^*,m) = \max_{x\in A_{f-m^*}}\, [\operatorname{sgn}(f-m^*)(x)]\, m(x) \ge 0$$

for all $m \in M$. Furthermore, m^* is a strongly unique best approximant to f from M if and only if

$$\gamma(f) = \inf_{\substack{m\in M\\ \|m\|=1}}\ \max_{x\in A_{f-m^*}}\, [\operatorname{sgn}(f-m^*)(x)]\, m(x) > 0.$$

This characterization of a best approximant is called the Kolmogorov criterion. This characterization of the optimal strong uniqueness constant $\gamma(f)$ first appeared in Bartelt, McLaughlin [1973]. If M is finite-dimensional, then the above *inf* can be replaced by a *min*. This is not true in infinite dimensions. An example illustrating that fact may be found in Bartelt, McLaughlin [1973]; see Example 4.

In the finite-dimensional setting, the above characterization results can be further refined. We firstly state the following well-known result.

Theorem 2.3. *Let M be an n-dimensional subspace of $C(B)$. Given $f \in C(B)$, we have $m^* \in P_M(f)$ if and only if there exist k distinct points $x_1,\ldots,x_k$ in A_{f-m^*} , and strictly positive values $\lambda_1,\ldots,\lambda_k$, $1 \le k \le n+1$, such that*

$$\sum_{i=1}^{k} \lambda_i[\operatorname{sgn}(f-m^*)(x_i)]m(x_i) = 0$$

for all $m \in M$.

(A variation on a proof of Theorem 2.3 is given by the proof of Theorem 10.3.)

We also have the following result that is proved in Brosowski [1983], and was also later proved in Smarzewski [1990], generalizing results from Bartelt [1974] and Nürnberger [1980].

Theorem 2.4. *Let M be an n-dimensional subspace of $C(B)$. Given $f \in C(B)\backslash M$, we have that $m^* \in M$ is the strongly unique best approximant to f from M if and only if there exist k distinct points $x_1, \ldots, x_k$ in A_{f-m^*}, and strictly positive values $\lambda_1, \ldots, \lambda_k$, with $n+1 \le k \le 2n$, such that*

$$\sum_{i=1}^{k} \lambda_i [\operatorname{sgn}(f - m^*)(x_i)] m(x_i) = 0 \tag{2.1}$$

for all $m \in M$, and

$$\dim M|_{\{x_1,\ldots,x_k\}} = n.$$

Proof: We first prove the "easy" direction in this theorem.
($\Leftarrow$). Assume there exist points $\{x_i\}_{i=1}^k$ satisfying the conditions of the theorem. As the $x_1, \ldots, x_k$ are in A_{f-m^*}, we have

$$\gamma(f) \ge \gamma = \min_{\substack{m \in M \\ \|m\|=1}} \max_{i=1,\ldots,k} \operatorname{sgn}(f - m^*)(x_i) m(x_i).$$

We claim that $\gamma > 0$. If this holds then $\gamma(f) > 0$ and m^* is the strongly unique best approximant to f from M. To see that $\gamma > 0$, recall that since M is a finite-dimensional subspace then from a compactness argument the above minimum is always attained. Thus if $\gamma \le 0$, there exists an $\widetilde{m} \in M$, $\|\widetilde{m}\| = 1$, for which

$$\operatorname{sgn}(f - m^*)(x_i)\widetilde{m}(x_i) \le 0$$

for $i = 1, \ldots, k$. Moreover as $\widetilde{m} \in M$, $\widetilde{m} \ne 0$, then from the fact that $\dim M|_{\{x_1,\ldots,x_k\}} = n$, it follows that $\widetilde{m}(x_i) \ne 0$ at at least one of the x_i. As $\lambda_i > 0$, $i = 1, \ldots, k$, this then implies

$$\sum_{i=1}^{k} \lambda_i \operatorname{sgn}(f - m^*)(x_i)\widetilde{m}(x_i) < 0$$

which contradicts (2.1). Thus $\gamma > 0$.
($\Rightarrow$). Assume $m^* \in M$ is a strongly unique best approximant to f from M. Without loss of generality, we will assume $m^* = 0$.

We start with the case where $\dim M = 1$. Let $M = \operatorname{span}\{m_1\}$. Since

$$\gamma(f) = \max_{x \in A_f} [\operatorname{sgn} f(x)](\pm m_1(x)) > 0,$$

there must exist points $x_1, x_2 \in A_f$ for which we have that both $[\operatorname{sgn} f(x_1)]m_1(x_1) > 0$ and $[\operatorname{sgn} f(x_2)]m_1(x_2) < 0$. Thus, there exist $\lambda_1, \lambda_2 > 0$ satisfying

$$\lambda_1[\operatorname{sgn} f(x_1)]m(x_1) + \lambda_2[\operatorname{sgn} f(x_2)]m(x_2) = 0$$

for all $m \in M$, and $\dim M|_{\{x_1,x_2\}} = 1$. This is the desired result.

We now consider the general case where $\dim M = n$. Since

$$\max_{x \in A_f} [\operatorname{sgn} f(x)]m(x) \ge c > 0$$

for all $m \in M$, $\|m\| = 1$, it follows that $\mathbf{0} \in \mathbb{R}^n$ is in the strict interior of the convex hull of

$$E = \{([\operatorname{sgn} f(x)]m_1(x), \ldots, [\operatorname{sgn} f(x)]m_n(x)) : x \in A_f\} \subset \mathbb{R}^n$$

where $\{m_1, \ldots, m_n\}$ is any basis for M. From a generalization of Carathéodory's Theorem, essentially due to Steinitz, see Danzer, Grünbaum, Klee [1963], the vector $\mathbf{0} \in \mathbb{R}^n$ is in the strict interior of the convex hull of some set of at most $2n$ points of E. That is, there exist $x_1, \ldots, x_k \in A_f$, $k \le 2n$, such that $\mathbf{0} \in \mathbb{R}^n$ is in the strict interior of the convex hull of

$$E^* = \{([\operatorname{sgn} f(x_i)]m_1(x_i), \ldots, [\operatorname{sgn} f(x_i)]m_n(x_i)) : i = 1, \ldots, k\}.$$

From the fact that $\mathbf{0} \in \mathbb{R}^n$ is in the strict interior of the convex hull of E^*, it easily follows that there exist $\lambda_1, \ldots, \lambda_k > 0$, $\sum_{i=1}^k \lambda_i = 1$, for which

$$0 = \sum_{i=1}^k \lambda_i [\operatorname{sgn} f(x_i)] m_j(x_i), \qquad j = 1, \ldots, n,$$

implying

$$0 = \sum_{i=1}^k \lambda_i [\operatorname{sgn} f(x_i)] m(x_i),$$

for all $m \in M$, and also that the vectors

$$([\operatorname{sgn} f(x_i)]m_1(x_i), \ldots, [\operatorname{sgn} f(x_i)]m_n(x_i)), \qquad i = 1, \ldots, k,$$

span $\mathbb{R}^n$. Since $\operatorname{sgn} f(x_i) \neq 0$ for each i, this is equivalent to the fact that the vectors

$$(m_1(x_i), \ldots, m_n(x_i)), \qquad i = 1, \ldots, k,$$

span $\mathbb{R}^n$, i.e.,

$$\dim M|_{\{x_1, \ldots, x_k\}} = n.$$

From this fact and

$$0 = \sum_{i=1}^k \lambda_i [\operatorname{sgn} f(x_i)] m(x_i),$$

for all $m \in M$, it follows that $k \ge n + 1$. □

Is the bound on k, namely $n + 1 \le k \le 2n$, the correct bound? The answer is yes. The value $k = n + 1$ is minimal. If $k \le n$ there is always a nontrivial $m \in M$ that vanishes at $x_1, \ldots, x_{k-1}$ and then $\min\{[\operatorname{sgn}(f - m^*)(x_k)](\pm m(x_k))\} \le 0$ implying $\gamma(f) = 0$. In the next section, we see that in the Haar space setting, we can and do have $k = n + 1$. The case $k = 2n$ is necessary if, for example, M is spanned by n functions with disjoint support. To see this, assume $m_1, \ldots, m_n$ is a basis for M where these basis functions have disjoint support. Assume $f \in C(B) \backslash M$ and consider (2.1). In order that $\dim M|_{\{x_1, \ldots, x_k\}} = n$, it is necessary that among the $x_1, \ldots, x_k$ there is at least one point in the support of each m_j. But we must have at least two points in the support of each m_j among the $x_1, \ldots, x_k$ if (2.1) is to hold. Thus $k = 2n$ is necessary for strong uniqueness in this case and, by the above, there do exist $f \in C(B) \backslash M$ with strongly unique best approximants from M.

It might be conjectured that uniqueness and strong uniqueness are equivalent properties in $C(B)$. This, however, is not true, as can be seen from this example taken from Cheney [1966, p. 82].

Example. Let $M = \mathrm{span}\{x\}$ in $C[-1,1]$, and set $f(x) = 1 - x^2$. As is easily verified, the unique best approximant to f from M is $m^*(x) = 0$. On the other hand, m^* is not a strongly unique best approximant to f from M since $A_f = \{0\}$, and $m(0) = 0$, i.e., $\tau_+(1 - x^2, x) = 0$.

Nevertheless, while uniqueness and strong uniqueness are not equivalent properties in $C(B)$, the set of functions with a strongly unique best approximant is dense in the set of functions with a unique best approximant when approximating from a finite-dimensional subspace. This next result and the subsequent Corollary 2.7 are from Nürnberger, Singer [1982]. The proof given here is from Smarzewski [1988].

Theorem 2.5. *Let M be a finite-dimensional subspace of $C(B)$. Then the set of functions with a strongly unique best approximant is dense in the set of functions with a unique best approximant.*

Note that we have an algebraic characterization for a best approximant and an algebraic characterization for a strongly unique best approximant from any finite-dimensional subspace (Theorems 2.2, 2.3 and 2.4). However there is no known algebraic characterization for when we have a unique best approximant. We get around this by proving the next proposition. We start with some notation.

Let us assume, for ease of presentation, that the zero function is a best approximant to f from M. Thus from Theorem 2.2, we have

$$\max_{x \in A_f} [\mathrm{sgn}\, f(x)] m(x) \geq 0$$

for all $m \in M$. Let (k_n) be any strictly decreasing sequence of positive numbers that converges to zero, and assume $k_n < 1/2$ for all n. Set

$$B_n := \{x : |f(x)| > (1 - k_n)\|f\|\}, \qquad n = 2, 3, \ldots\ .$$

Then we have:

Proposition 2.6. *Let M be a finite-dimensional subspace of $C(B)$. If the zero function is the unique best approximant to f from M then for each n, there exists an $a_n > 0$ such that*

$$\max_{x \in B_n} [\mathrm{sgn}\, f(x)] m(x) \geq a_n \|m\|$$

for all $m \in M$, where B_n is as above.

Remark. The above is a necessary, but not a sufficient condition, for the uniqueness of the best approximant. As an example, consider $f(x) = 1 - |x|$ on $[-1,1]$, and $M = \mathrm{span}\{m_1\}$ where $m_1(x) = x$ on $[-1,1]$. Then, as is readily verified, $B_n = [-k_n, k_n]$ and

$$\max_{x \in B_n} [\mathrm{sgn}\, f(x)] m(x) = k_n \|m\|$$

for all $m \in M$. Moreover, while the zero function is a best approximant to f from M, it is not the unique best approximant as we also have $\pm m_1 \in P_M(f)$.

Proof: Assume to the contrary that the desired inequality does not hold for some $n \geq 2$. This inequality trivially holds for $m = 0$. As such, we can consider it over the boundary of the unit ball of M, and since M is a finite-dimensional subspace, we can use compactness to affirm that there exists an $\widetilde{m} \in M$, $\|\widetilde{m}\| = 1$, such that

$$\max_{x \in B_n} [\operatorname{sgn} f(x)]\widetilde{m}(x) \le 0.$$

Choose any $\alpha > 0$ satisfying $\alpha \le k_n \|f\|$. Thus

$$|\alpha \widetilde{m}(x)| < |f(x)|$$

for all $x \in B_n$, and therefore, we also have thereon

$$|f(x) + \alpha \widetilde{m}(x)| \le |f(x)| + \alpha [\operatorname{sgn} f(x)]\widetilde{m}(x) \le |f(x)| \le \|f\|.$$

On the other hand, for $x \in B \backslash B_n$, we have

$$|f(x) + \alpha \widetilde{m}(x)| \le |f(x)| + \alpha |\widetilde{m}(x)| \le (1 - k_n)\|f\| + k_n \|f\| = \|f\|.$$

Thus $-\alpha \widetilde{m} \in P_M(f)$, contradicting the fact that the zero function is the unique best approximant to f from M. $\square$

One immediate consequence of the above proposition is the following.

Corollary 2.7. *If M is a subspace of ℓ_∞^m, then every unique best approximant to $f \in \ell_\infty^m$ from M is also a strongly unique best approximant to f from M.*

Proof: For n sufficiently large, we always have, in this case, $B_n = A_f$. Apply Proposition 2.6 and the characterization of strong uniqueness found in Theorem 2.2. $\square$

Proof of Theorem 2.5: If $f \in M$, there is nothing to prove. As such, we assume that $f \in C(B) \backslash M$. Without loss of generality, we assume that the zero function is the unique best approximant to f from M. Let B_n and a_n be as in Proposition 2.6.

By the Tietze-Urysohn Theorem, see e.g., Kuratowski [1966], there exists a $f_n \in C(B)$ such that

$$f_n(x) = \begin{cases} \|f\| \operatorname{sgn} f(x), & x \in B_{n+2} \\ f(x), & x \in B_n \cap \overline{(B \backslash B_{n+2})} \end{cases}$$

and

$$(1 - k_n)\|f\| \le f_n(x) \le \|f\|$$

for $x \in B_n$. If $B_n = B_{n+2}$, then we simply set $f_n(x) = \|f\| \operatorname{sgn} f(x)$ thereon. We extend f_n to all of $C(B)$ by setting

$$f_n(x) = f(x)$$

for $x \in B \backslash B_n$. Note that we do not lose continuity in the case where $B_n = B_{n+2}$ since in that case

$$\{x : (1 - k_{n+2})\|f\| \ge |f(x)| \ge (1 - k_n)\|f\|\} = \emptyset.$$

From this construction, we have

$$|f(x) - f_n(x)| \le k_n \|f\|$$

for all $x \in B$. That is, we have

$$\lim_{n \to \infty} \|f - f_n\| = 0.$$

Now, since $A_{f_n} \supseteq B_{n+2} \supseteq A_f$, it follows from Theorem 2.2 that $0 \in P_M(f_n)$ for all n. Furthermore from Proposition 2.6, we have

$$\min_{\substack{m \in M \\ \|m\|=1}} \max_{x \in A_{f_n}} [\operatorname{sgn} f_n(x)] m(x) \ge \min_{\substack{m \in M \\ \|m\|=1}} \max_{x \in B_{n+2}} [\operatorname{sgn} f(x)] m(x) \ge a_{n+2} > 0.$$

Thus, again applying Theorem 2.2, we see that the zero function is the strongly unique best approximant to f_n from M. $\square$

The above begs the question of when the set of functions with a unique best approximant from a given finite-dimensional subspace M is dense in $C(B)$. A similar question was considered by Garkavi [1964], [1965] who proved that the n-dimensional subspace M of $C(B)$ has the property that every function in $C(B)$ has a unique best approximant from M except for a set of first category in $C(B)$ if and only if on each open subset D of B there identically vanish at most $(n-|D|)_+ = \max\{n-|D|, 0\}$ linearly independent functions of M, where $|D|$ is the number of points in D. Thus, for example, the set of functions with a strongly unique best approximant from M is dense in $C[a,b]$ if no $m \in M$, $m \neq 0$, vanishes identically on an open interval.

Bounds for $\gamma(f)$ are difficult to obtain. The following result, from Grothmann [1989], gives an upper bound for $\gamma(f)$. In general, using it to find explicit bounds is very difficult.

Before stating the result, we recall the definition of a *projection constant* of a subspace L in a normed linear space X. It is given by

$$\lambda(L; X) := \inf\{\|P\| : P : X \to L \text{ is a projection}\}.$$

Proposition 2.8. *Let M be a linear subspace of $C(B)$. Let $f \in C(B)$ and assume $m^* \in M$ is a strongly unique best approximant to f. Then*

$$\gamma(f) \leq \frac{\lambda(M|_{A_{f-m^*}}; C(A_{f-m^*}))}{\lambda(M; C(B))}.$$

Proof: Since m^* is a strongly unique best approximant to f from M, we have that

$$\gamma(f) = \min_{\substack{m\in M\\ \|m\|=1}} \max_{x\in A_{f-m^*}} [\operatorname{sgn}(f-m^*)(x)]m(x) > 0.$$

This also implies that

$$\|m\|_{A_{f-m^*}} = \max_{x\in A_{f-m^*}} |m(x)|$$

is a norm on M since no $m \in M$, $m \neq 0$, can vanish identically on A_{f-m^*}. From the above formula for $\gamma(f)$, we have

$$\begin{aligned}\gamma(f) &\leq \min_{\substack{m\in M\\ \|m\|=1}} \max_{x\in A_{f-m^*}} |m(x)| \\ &= \min_{m\in M} \frac{\|m\|_{A_{f-m^*}}}{\|m\|} \\ &= \left(\max_{m\in M} \frac{\|m\|}{\|m\|_{A_{f-m^*}}}\right)^{-1}.\end{aligned}$$

Let $P : C(A_{f-m^*}) \to M|_{A_{f-m^*}}$ be a projection. Define $\sigma : M|_{A_{f-m^*}} \to M$ and $\phi : C(B) \to C(A_{f-m^*})$ by

$$\sigma\left(m|_{A_{f-m^*}}\right) = m$$

for $m \in M$ and

$$\phi(f) = f|_{A_{f-m^*}}$$

for $f \in C(B)$. The linear operator σ is well-defined since no $m \in M$, $m \neq 0$, can vanish identically on A_{f-m^*}. Set

$$Q(f) := \sigma(P(\phi(f))).$$

Q is a projection from $C(B)$ onto M. Thus

$$\lambda(M; C(B)) \leq \|Q\| \leq \|\sigma\| \|P\| \|\phi\|.$$

Now $\|\phi\| = 1$ and, by the above, $\|\sigma\| \leq 1/\gamma(f)$. Thus

$$\lambda(M; C(B)) \leq \frac{1}{\gamma(f)} \|P\|$$

for all projections $P : C(A_{f-m^*}) \to M|_{A_{f-m^*}}$. Therefore

$$\gamma(f) \leq \frac{\lambda(M|_{A_{f-m^*}}; C(A_{f-m^*}))}{\lambda(M; C(B))}. \quad \square$$

In Section 4, where we assume M is a Haar space, we discuss other somewhat more computable bounds for $\gamma(f)$.

3. Local Lipschitz Continuity and Classical Strong Uniqueness

We start once again with an arbitrary normed linear space X. Assume M is a subset of X, and to $f \in X$ there exists a unique best approximant from M. If, for a given $f \in X$, we have an inequality of the form

$$\|P_M(f) - P_M(g)\| \leq \sigma \|f - g\|$$

valid for all $g \in X$ and any element of $P_M(g)$, then we say that the best approximation operator from M is *locally Lipschitz continuous* at f, and call σ a *local Lipschitz constant*.

One of the "uses" of classical strong uniqueness is that it implies local Lipschitz continuity. Before proving this result, we recall that we always have, for every $f, g \in X$,

$$| \, \|f - P_M(f)\| - \|g - P_M(g)\| \, | \leq \|f - g\|.$$

To verify this, note that, assuming $\|f - P_M(f)\| \geq \|g - P_M(g)\|$, then

$$\|f - P_M(f)\| \leq \|f - P_M(g)\| \leq \|f - g\| + \|g - P_M(g)\|,$$

from which the result follows.

Theorem 3.1. *Assume M is a subset of a normed linear space X, $f \in X$, and for some $\gamma > 0$, we have*

$$\|f - m\| - \|f - P_M(f)\| \geq \gamma \|m - P_M(f)\|$$

for all $m \in M$. Then for each $g \in X$ and any element of $P_M(g)$

$$\|P_M(f) - P_M(g)\| \leq \frac{2}{\gamma} \|f - g\|.$$

Proof: By assumption, and an application of the previous inequality,

$$\gamma \|P_M(f) - P_M(g)\| \leq \|f - P_M(g)\| - \|f - P_M(f)\|$$

$$\leq \|f-g\|+\|g-P_M(g)\|-\|f-P_M(f)\| \leq 2\|f-g\|.$$

Thus

$$\|P_M(f)-P_M(g)\| \leq \frac{2}{\gamma}\|f-g\|. \quad \square$$

In the specific case where M is a finite-dimensional Haar space, see Section 4, local Lipschitz continuity of the best approximation operator was proved in Freud [1958] by a different method of proof. For polynomial approximation on an interval, this result can be found in Kirchberger [1902, p. 18–21] (see also Borel [1905, p. 89–92]).

We have proven that strong uniqueness at f implies local Lipschitz continuity at f. The converse direction, i.e., local Lipschitz continuity implying strong uniqueness, does not necessarily hold. It certainly does not hold in an inner product space where we always have

$$\|P_M(f)-P_M(g)\| \leq \|f-g\|$$

for all f, g if M is a subspace. That is, we have local Lipschitz continuity and do not have strong uniqueness, see the discussion at the end of Section 1. However in $C(B)$ the converse does hold, i.e., local Lipschitz continuity at $f \in C(B)$ does imply strong uniqueness at this same f, assuming M is a finite-dimensional subspace. This next theorem is to be found in Bartelt, Schmidt [1984].

Theorem 3.2. *Let B be a compact Hausdorff space and M a finite-dimensional subspace of $C(B)$. For given $f \in C(B)$, the following are equivalent:*

(I) There exists a $\gamma > 0$ such that

$$\|f-m\|-\|f-P_M(f)\| \geq \gamma\|m-P_M(f)\|$$

for all $m \in M$.

(II) There exists a $\sigma > 0$ such that

$$\|P_M(f)-P_M(g)\| \leq \sigma\|f-g\|$$

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

Proof: From Theorem 3.1, we have that (I) implies (II). It remains to prove the converse direction. We prove the result by contradiction. That is, we assume that (I) does not hold for a given f and will prove that (II) does not hold for this same f. Note that if $f \in M$ then (I) and (II) hold, while if the best approximant to f from M is not unique, then neither (I) nor (II) hold. As such, we assume $f \notin M$ and the best approximant to f from M is unique.

Without loss of generality, we can and will assume that the zero function is the unique best approximant to f, and $\|f\| = 1$. Now, since (I) does not hold, there exists, for each $\varepsilon > 0$, an $m_\varepsilon \in M$ satisfying

$$\|f-m_\varepsilon\| < \|f\|+\varepsilon\|m_\varepsilon\| = 1+\varepsilon\|m_\varepsilon\|.$$

Note that this implies that $m_\varepsilon \neq 0$. We divide most of the proof into a series of four lemmas.

Lemma 3.3. *There exist sequences (δ_n) and (α_n) of positive numbers tending to zero and an $\widetilde{m} \in M$, $\|\widetilde{m}\| = 1$, such that*

(i) $\|f-\alpha_n\widetilde{m}\| \leq 1+\delta_n\alpha_n$,

(ii) $\widetilde{m}(x)\operatorname{sgn} f(x) \geq 0$ *for all* $x \in A_f$.

Proof: Let m_ε be as above. We first claim that for $\varepsilon \leq 1/2$, we necessarily have $\|m_\varepsilon\| \leq 4$. To see this, assume $\|m_\varepsilon\| = C$. Then for $0 < \varepsilon \leq 1/2$, we have

$$\frac{C}{2} \geq \varepsilon\|m_\varepsilon\| \geq \|f - m_\varepsilon\| - 1 \geq \|m_\varepsilon\| - \|f\| - 1 = C - 2,$$

whence $C \leq 4$.

As $\|m_\varepsilon\| \leq 4$, we have

$$\lim_{\varepsilon\to 0^+} \varepsilon\|m_\varepsilon\| = 0.$$

Thus

$$\lim_{\varepsilon\to 0^+} \|f - m_\varepsilon\| = \|f\| = \min_{m\in M} \|f - m\|.$$

From a compactness argument, it therefore follows that

$$\lim_{\varepsilon\to 0^+} \|m_\varepsilon\| = 0,$$

i.e., m_ε tends to the zero function. Set

$$m_\varepsilon := \widetilde{m}_\varepsilon\|m_\varepsilon\|$$

and let $\alpha_\varepsilon = \|m_\varepsilon\|$. Thus $\|\widetilde{m}_\varepsilon\| = 1$ for all ε, and α_ε tends to 0 as ε tends to 0. As the $\widetilde{m}_\varepsilon$ are all functions of norm 1 in a finite-dimensional subspace, on a subsequence $\varepsilon_n \to 0^+$, we have

$$\lim_{n\to\infty} \widetilde{m}_{\varepsilon_n} = \widetilde{m}$$

where $\widetilde{m} \in M$ and $\|\widetilde{m}\| = 1$. For convenience, let $\widetilde{m}_n := \widetilde{m}_{\varepsilon_n}$, $m_n := m_{\varepsilon_n}$ and $\alpha_n := \alpha_{\varepsilon_n}$.

We claim that $\widetilde{m}(x)\operatorname{sgn} f(x) \geq 0$ for all $x \in A_f$. Assume not. There then exists an $x^* \in A_f$ for which

$$\widetilde{m}(x^*)\operatorname{sgn} f(x^*) = -c < 0.$$

Thus for n sufficiently large,

$$\widetilde{m}_n(x^*)\operatorname{sgn} f(x^*) < -\frac{c}{2},$$

and therefore

$$\begin{aligned} 1 + \varepsilon_n\alpha_n &= \|f\| + \varepsilon_n\|m_n\| > \|f - m_n\| \\ &\geq |f(x^*) - m_n(x^*)| = |f(x^*)| + \alpha_n|\widetilde{m}_n(x^*)| > 1 + \alpha_n\frac{c}{2}. \end{aligned}$$

But this cannot possibly hold for n sufficiently large as $\varepsilon_n \to 0$. Thus $\widetilde{m}(x)\operatorname{sgn} f(x) \geq 0$ for all $x \in A_f$.

Let $\beta_n = \|\widetilde{m}_n - \widetilde{m}\|$. Therefore β_n tends to zero as n tends to infinity. Now

$$\begin{aligned} [\|f - \alpha_n\widetilde{m}\| - \|f\|] - [\|f - \alpha_n\widetilde{m}_n\| - \|f\|] &= \|f - \alpha_n\widetilde{m}\| - \|f - \alpha_n\widetilde{m}_n\| \\ &\leq \|\alpha_n\widetilde{m} - \alpha_n\widetilde{m}_n\| = \alpha_n\beta_n. \end{aligned}$$

Thus

$$0 \leq \|f - \alpha_n\widetilde{m}\| - \|f\| \leq \|f - \alpha_n\widetilde{m}_n\| - \|f\| + \alpha_n\beta_n \leq \alpha_n\varepsilon_n + \alpha_n\beta_n = \alpha_n(\varepsilon_n + \beta_n).$$

Set $\delta_n := \varepsilon_n + \beta_n$. Then

$$\lim_{n\to\infty} \delta_n = 0$$

and this proves the lemma. $\square$

Recall that we assumed that $0 \in P_M(f)$. This implies, by Theorem 2.3, the existence of k distinct points $x_1, \ldots, x_k \in A_f$, $1 \le k \le n+1$, and strictly positive values $\lambda_1, \ldots, \lambda_k$ such that

$$\sum_{i=1}^{k} \lambda_i [\operatorname{sgn} f(x_i)] m(x_i) = 0$$

for all $m \in M$.

Lemma 3.4. $\widetilde{m}(x_i) = 0$, *for* $i = 1, \ldots, k$.

Proof: From Lemma 3.3 (ii), we have

$$0 \le \widetilde{m}(x_i) \operatorname{sgn} f(x_i)$$

for all $i = 1, \ldots, k$. Thus if we have strict inequality for any i, then we contradict the fact that

$$\sum_{i=1}^{k} \lambda_i [\operatorname{sgn} f(x_i)] \widetilde{m}(x_i) = 0. \quad \square$$

For each given n, let

$$G_n = \{x : |\widetilde{m}(x)| < \delta_n/2\}.$$

Note that G_n is an open neighborhood of the $\{x_1, \ldots, x_k\}$. We define a function $\phi_n \in C(B)$ as follows. Firstly set $\phi_n(x_i) = \alpha_n \delta_n$, $i = 1, \ldots, k$, and $\phi_n(x) = 0$ for all $x \in B \backslash G_n$. Note that on $\{x_1, \ldots, x_k\} \cup B \backslash G_n$, we have

$$0 \le \phi_n(x) \le |\alpha_n \delta_n - \alpha_n |\widetilde{m}(x)|\,|.$$

This follows from the fact that $\widetilde{m}(x_i) = 0$, $i = 1, \ldots, k$. Extend ϕ_n continuously to all of B so that it continues to satisfy

$$0 \le \phi_n(x) \le |\alpha_n \delta_n - \alpha_n |\widetilde{m}(x)|\,|$$

for all $x \in B$. Now set $g_n(x) := f(x)\,[1 + \phi_n(x)]$.

Lemma 3.5. *We have*

$$\|f - g_n\| = \alpha_n \delta_n.$$

Proof: From the definition of g_n, $f(x) - g_n(x) = -f(x)\phi_n(x)$. Since $\|f\| = 1$, we therefore have

$$\|f - g_n\| \le \|\phi_n\|.$$

On G_n, where ϕ_n need not vanish, we have

$$|\widetilde{m}(x)| < \frac{\delta_n}{2}.$$

Thus for $x \in G_n$

$$0 \le \phi_n(x) \le |\alpha_n \delta_n - \alpha_n |\widetilde{m}(x)|\,| = \alpha_n \delta_n - \alpha_n |\widetilde{m}(x)| \le \alpha_n \delta_n.$$

Therefore $\|f - g_n\| \le \alpha_n\delta_n$. Equality holds since

$$|f(x_i) - g_n(x_i)| = |f(x_i)\phi_n(x_i)| = |\phi_n(x_i)| = \alpha_n\delta_n. \quad \square$$

Lemma 3.6. *We have*

$$\min_{m\in M} \|g_n - m\| = \|g_n - \alpha_n\widetilde{m}\| = 1 + \alpha_n\delta_n.$$

Proof: For each $i = 1, \ldots, k$, since $\widetilde{m}(x_i) = 0$, we have

$$|g_n(x_i) - \alpha_n\widetilde{m}(x_i)| = |g_n(x_i)| = |f(x_i)|\,|1 + \phi_n(x_i)| = 1 + \phi_n(x_i) = 1 + \alpha_n\delta_n.$$

Now for $x \in B\backslash G_n$, since $\phi_n(x) = 0$, we have

$$|g_n(x) - \alpha_n\widetilde{m}(x)| = |f(x) - \alpha_n\widetilde{m}(x)| \le \|f - \alpha_n\widetilde{m}\| \le 1 + \alpha_n\delta_n.$$

The latter inequality is from Lemma 3.3 (i). For $x \in G_n$, since $|f(x)| \le 1$ and $|\widetilde{m}(x)| < \delta_n/2$, we have

$$\begin{aligned} |g_n(x) - \alpha_n\widetilde{m}(x)| &= |f(x) + f(x)\phi_n(x) - \alpha_n\widetilde{m}(x)| \\ &\le 1 + |\alpha_n\delta_n - \alpha_n|\widetilde{m}(x)|\,| + \alpha_n|\widetilde{m}(x)| \\ &\le 1 + \alpha_n\delta_n - \alpha_n|\widetilde{m}(x)| + \alpha_n|\widetilde{m}(x)| \;=\; 1 + \alpha_n\delta_n. \end{aligned}$$

This proves that

$$\|g_n - \alpha_n\widetilde{m}\| = 1 + \alpha_n\delta_n.$$

As

$$g_n(x_i) - \alpha_n\widetilde{m}(x_i) = g_n(x_i) = f(x_i)\,[1 + \alpha_n\delta_n]$$

for $i = 1, \ldots, k$, it follows from Theorem 2.3 characterizing best approximants that $\alpha_n\widetilde{m}$ is a best approximant to g_n from M. $\square$

Proof of Theorem 3.2 (cont'd): If (II) holds, then there exists a $\sigma > 0$ such that

$$\|P_M f - P_M g\| \le \sigma\|f - g\|$$

for all $g \in C(B)$. Recall that we assumed, without loss of generality, that the zero function is the best approximant to f from M and $\|f\| = 1$. Taking $g = g_n$ and applying the above lemmas, we obtain

$$\alpha_n = \|\alpha_n\widetilde{m}\| \le \sigma\|f - g_n\| = \sigma\alpha_n\delta_n.$$

Thus $1 \le \sigma\delta_n$ for all n. But $\lim_{n\to\infty}\delta_n = 0$. This is a contradiction. $\square$

Based on Theorems 2.4 and 3.2, we now have a characterization for local Lipschitz continuity of the best approximation operator in the uniform norm from finite-dimensional subspaces. A weaker sufficient condition can be found in Kovtunec [1984].

4. Strong Uniqueness in Haar Spaces in the Uniform Norm

We start with the definition of a Haar space.

Definition 4.1. *An n-dimensional subspace M of $C(B)$ is said to be a Haar space if no nontrivial $m \in M$ vanishes at more than $n-1$ distinct points of B.*

A *unicity space* is any subspace M of a normed linear space X with the property that each $f \in X$ has a unique best approximant to f from M. The following result was proved by Haar [1918] and is built upon earlier results of Young [1908].

Theorem 4.1. *An n-dimensional subspace M of $C(B)$ is a unicity space if and only if it is a Haar space.*

As a result of this theorem, the term Haar space is often applied to any unicity space in any normed linear space. But here we use the term Haar space as that given in the above definition. There are many equivalent definitions of the Haar space property. Here are two that will prove useful.

1) An n-dimensional subspace M of $C(B)$ is a Haar space if and only if $\dim M|_{\{x_1,\ldots,x_n\}} = n$ for every choice of n distinct points $x_1, \ldots, x_n$ in B.

2) Let $m_1, \ldots, m_n$ be a basis for M. Then M is a Haar space if and only if

$$\det(m_i(x_j))_{i,j=1}^n \neq 0$$

for every choice of n distinct points $x_1, \ldots, x_n$ in B.

For $n \geq 2$, there are rather restrictive conditions on B needed to ensure that $C(B)$ can contain a Haar space of dimension n. Essentially, B must be homeomorphic to a subset of S^1. Exact conditions are the content of Mairhuber's Theorem; see Mairhuber [1956], Sieklucki [1958], Curtis [1959], Schoenberg, Yang [1961], and McCullough, Wulbert [1985].

When we have a Haar space, the characterization result Theorem 2.3 can be further strengthened.

Theorem 4.2. *Let M be an n-dimensional Haar space on $C(B)$. Given $f \in C(B)$, we have that m^* is the best approximant to f from M if and only if there exist $n+1$ distinct points $x_1, \ldots, x_{n+1}$ in A_{f-m^*} and strictly positive values $\lambda_1, \ldots, \lambda_{n+1}$ such that*

$$\sum_{i=1}^{n+1} \lambda_i [\mathrm{sgn}(f-m^*)(x_i)] m(x_i) = 0$$

for all $m \in M$.

Proof: This does not differ much from Theorem 2.3. The difference is in the claim that the k therein must be $n+1$. To see this, let

$$\sigma_i = \lambda_i [\mathrm{sgn}(f-m^*)(x_i)], \qquad i = 1, \ldots, k.$$

From Theorem 2.3, we have

$$\sum_{i=1}^{k} \sigma_i m(x_i) = 0$$

for all $m \in M$. Let $m_1, \ldots, m_n$ be any basis for M. Then

$$\sum_{i=1}^{k} \sigma_i m_j(x_i) = 0, \qquad j = 1, \ldots, n.$$

If $k \leq n$, then this implies that

$$\operatorname{rank}\,(m_j(x_i))_{i=1}^{k}{}_{j=1}^{n} < k$$

since $(\sigma_1, \ldots, \sigma_k) \neq \mathbf{0}$. But this contradicts the fact that

$$\operatorname{rank}\,(m_j(y_i))_{i,j=1}^{n} = n$$

for all distinct $y_1, \ldots, y_n$ in B. Thus $k = n + 1$. $\square$

A Haar space on an interval is generally called a *Chebyshev space* (or Tchebycheff space) and often abbreviated a *T-space*. A basis for a T-space is sometimes called a *T-system*. For a T-space, because of the connectedness of the interval, we have:

Proposition 4.3. *Let $m_1, \ldots, m_n$ be any basis for M. Then M is a T-space if and only if*

$$\varepsilon \det(m_i(x_j))_{i,j=1}^{n} > 0$$

for some fixed $\varepsilon \in \{-1, 1\}$ and all $x_1 < \cdots < x_n$.

For T-spaces, we can further specialize Theorem 4.2 into a final and more geometric form. We have

Theorem 4.4. *Let M be an n-dimensional T-space on $C[a, b]$. Given $f \in C[a, b]$, we have that m^* is the best approximant to f from M if and only if there exist points $a \leq x_1 < \cdots < x_{n+1} \leq b$ and a $\delta \in \{-1, 1\}$ such that*

$$(-1)^i \delta (f - m^*)(x_i) = \|f - m^*\|, \qquad i = 1, \ldots, n + 1.$$

Proof: ($\Leftarrow$). Assume m^* exists satisfying the above. If m^* is not a best approximant to f from M, then there exists an $\widetilde{m} \in M$ for which

$$\|f - \widetilde{m}\| < \|f - m^*\|.$$

Then for each $i \in \{1, \ldots, n + 1\}$, we have

$$\begin{aligned} (-1)^i \delta (f - m^*)(x_i) &= \|f - m^*\| \ > \ \|f - \widetilde{m}\| \\ &\geq \ |(f - \widetilde{m})(x_i)| \ \geq \ (-1)^i \delta (f - \widetilde{m})(x_i). \end{aligned}$$

Thus

$$(-1)^i \delta (\widetilde{m} - m^*)(x_i) > 0, \qquad i = 1, \ldots, n + 1.$$

It therefore follows that $\widetilde{m} - m^*$ has at least one zero in each of the intervals (x_i, x_{i+1}), $i = 1, \ldots, n$. That is, $\widetilde{m} - m^*$ has at least n distinct zeros on $[a, b]$. But $\widetilde{m} - m^* \in M$ where M is an n-dimensional T-space. Thus $\widetilde{m} = m^*$, which contradicts our hypothesis.

($\Rightarrow$). Assume m^* is the best approximant to f from M. Then from Theorem 4.2, we have $n+1$ distinct points $a \le x_1 < \cdots < x_{n+1} \le b$ in A_{f-m^*}, and strictly positive values $\lambda_1, \ldots, \lambda_{n+1}$ such that

$$\sum_{i=1}^{n+1} \lambda_i [\operatorname{sgn}(f-m^*)(x_i)] m(x_i) = 0$$

all $m \in M$.

We wish to prove that

$$[\operatorname{sgn}(f-m^*)(x_i)][\operatorname{sgn}(f-m^*)(x_{i+1})] < 0, \qquad i = 1, \ldots, n.$$

To this end, set

$$\sigma_i = \lambda_i [\operatorname{sgn}(f-m^*)(x_i)], \qquad i = 1, \ldots, n+1.$$

Thus, we wish to prove that

$$\sigma_i \sigma_{i+1} < 0, \qquad i = 1, \ldots, n.$$

Let $m_1, \ldots, m_n$ be any basis for M. From the above, it follows that

$$\sum_{i=1}^{n+1} \sigma_i m_j(x_i) = 0, \qquad j = 1, \ldots, n.$$

Since

$$\operatorname{rank}\,(m_j(x_i))_{i=1}^{n+1}{}_{j=1}^{n} = n,$$

we have

$$\sigma_r = \alpha(-1)^r \det(m_j(x_i))_{\substack{r=1 \\ r \ne i}}^{n+1}{}_{j=1}^{n}$$

for all $r = 1, \ldots, n+1$, where $\alpha \ne 0$. Thus from Proposition 4.3,

$$\sigma_r \sigma_{r+1} < 0, \qquad r = 1, \ldots, n. \ \square$$

The following was first proven in Newman, Shapiro [1963] where they introduced the concept of (classical) strong uniqueness.

Theorem 4.5. *Let M be a finite-dimensional Haar subspace of $C(B)$. Then the unique best approximant to f from M is a strongly unique best approximant to f from M.*

Proof: The proof follows immediately from Theorems 2.4 and 4.2, and the definition of a Haar space that implies that we always have $\dim M|_{\{x_1,\ldots,x_{n+1}\}} = n$. $\square$

Thus, when M is a finite-dimensional Haar space, the functional $\gamma(f)$ is strictly positive on $C(B)$. It also has certain simple properties. For example, it is easy to verify that $\gamma(af-m) = \gamma(f)$ for all $a \in \mathbb{R}$, $a \ne 0$, and $m \in M$. Other properties are less obvious and for good reason.

For example, the functional $\gamma(f)$ is in general not a continuous function of f. This is because A_{f-m^*} is a highly noncontinuous function of f.

Example. Assume $B = [-1,1]$ and $M = \operatorname{span}\{1, x\}$. M is a T-space on $C[-1,1]$. If $A_{f-m^*} = \{-1, 0, 1\}$ with sign of the error $+1, -1, +1$, respectively, then $\gamma(f) = 1/3$. However, if $A_{f-m^*} = \{-1, [-1/2, 1/2], 1\}$ with signs $+1, -1, +1$, respectively, i.e., $f - m^*$ attains its norm positively at ± 1 and negatively on the full segment $[-1/2, 1/2]$, then $\gamma(f) = 3/5$. It is not at all difficult to construct an f and a sequence f_n of functions in $C[-1,1]$, all with the zero function as their best approximants from M, and such that $A_{f_n} = \{-1, 0, 1\}$ for all n, while $A_f = \{-1, [-1/2, 1/2], 1\}$.

While we do not have continuity of $\gamma(f)$, we do have upper semi-continuity when M is a finite-dimensional Haar space. This result is contained in Bartelt [1975] who writes that it is due to Phelps.

Proposition 4.6. *Assume M is a finite-dimensional Haar subspace of $C(B)$. Then the optimal strong unicity constant γ is an upper semi-continuous function on $C(B)$.*

Proof: Assume to the contrary that there exists a sequence of functions f_n in $C(B)$ that tend to $f \in C(B)$ such that

$$\gamma(f_n) \geq \gamma(f) + \varepsilon$$

for some $\varepsilon > 0$ and all n. Let $m^* \in P_M(f)$ and $m_n \in P_M(f_n)$ for each n. For given $m \in M$, it follows from the definition of $\gamma(f_n)$ that

$$\|f_n - m\| \geq \|f_n - m_n\| + \gamma(f_n)\|m - m_n\| \geq \|f_n - m_n\| + (\gamma(f) + \varepsilon)\|m - m_n\|.$$

Since M is a Haar space, it is well-known that the best approximation operator is a continuous operator. That is, from the fact that f_n converges to f, it follows that m_n converges to m^*. Thus it also follows that as $n \to \infty$, we have that $\|f_n - m\|$ tends to $\|f - m\|$, $\|f_n - m_n\|$ tends to $\|f - m^*\|$ and $\|m - m_n\|$ tends to $\|m - m^*\|$. Taking limits in the above inequality, we obtain

$$\|f - m\| \geq \|f - m^*\| + (\gamma(f) + \varepsilon)\|m - m^*\|.$$

Since this holds for every $m \in M$, this contradicts the definition of $\gamma(f)$. □

As noted previously, the reason for the lack of continuity of $\gamma(f)$ is in the fact that A_{f-m^*} is highly noncontinuous. When we have continuity of this set, then we have continuity of the $\gamma(f)$. To explain what we mean, let us first define a distance between sets. In what follows, we assume that B is a compact metric space. For sets $C, D \subseteq B$, we let

$$d(C, D) = \sup_{y \in D} \inf_{x \in C} \rho(x, y)$$

where ρ is a metric on B. There are many different such 'distances'. We will make do with this one. This result is contained in Bartelt [1975]

Proposition 4.7. *Assume M is a finite-dimensional Haar subspace of $C(B)$, where B is a compact metric space. Let f_n be a sequence in $C(B)$ converging uniformly to $f \in C(B)\backslash M$. Let $m_n \in P_M(f_n)$ and $m^* \in P_M(f)$. If*

$$\lim_{n \to \infty} d(A_{f_n - m_n}, A_{f-m^*}) = 0,$$

then

$$\lim_{n \to \infty} \gamma(f_n) = \gamma(f).$$

Proof: From Proposition 4.6, it follows that we need only prove that for any given $\varepsilon > 0$ there exists an N such that for all $n > N$, we have

$$\gamma(f_n) + \varepsilon > \gamma(f).$$

We start by simplifying things somewhat. Set $g_n := f_n - m_n$ and $g := f - m^*$. Then $A_{g_n} = A_{f_n - m_n}$ and $A_g = A_{f-m^*}$, $\gamma(g_n) = \gamma(f_n)$ and $\gamma(g) = \gamma(f)$, and the zero function is a best approximant from M to each of the g_n and g. Since M is a Haar space, we again have that m_n converges to m^*. Thus the g_n converge to g, and $g \notin M$. We will, without loss of generality, assume $\|g_n\| = \|g\| = 1$ for all n. We also use the fact that the unit sphere in M is uniformly equicontinuous on B. That is, given $\varepsilon > 0$, there exists a $\delta > 0$ such that if $\rho(x, y) < \delta$ then

$$|m(x) - m(y)| < \varepsilon$$

for all $m \in M$, $\|m\| = 1$. With these preliminaries, we can now prove the desired result.

By assumption, given $\varepsilon > 0$, there exists a $\delta_1 > 0$ such that if $\rho(x, y) < \delta_1$ then

$$|m(x) - m(y)| < \varepsilon$$

for all $m \in M$, $\|m\| = 1$. Similarly from the uniform continuity of g, there exists a $\delta_2 > 0$ such that if $\rho(x, y) < \delta_2$ then

$$|g(x) - g(y)| < 1/2.$$

By assumption, there exists an N_1 such that for all $n > N_1$, we have

$$\|g_n - g\| < 1/2.$$

And finally, by assumption, given $\delta > 0$, there exists an N_2 such that for all $n > N_2$, we have $d(A_{g_n}, A_g) < \delta$, i.e., for each $x \in A_g$, there exists a $y \in A_{g_n}$ such that $\rho(x, y) < \delta$. For our given $\varepsilon > 0$, we therefore set $\delta = \min\{\delta_1, \delta_2\}$ and $N = \max\{N_1, N_2\}$.

Given any $m \in M$, $\|m\| = 1$, let $x^* \in A_g$ satisfy

$$\max_{x \in A_g} [\operatorname{sgn} g(x)] m(x) = [\operatorname{sgn} g(x^*)] m(x^*).$$

By assumption for all $n > N$, we have

$$d(A_{g_n}, A_g) < \delta.$$

Thus there exists a $y^* \in A_{g_n}$ such that $\rho(x^*, y^*) < \delta$ implying that

$$|m(x^*) - m(y^*)| < \varepsilon.$$

In addition, as $|g(x^*)| = 1$, $n > N$ and $\rho(x^*, y^*) < \delta$, we have

$$\operatorname{sgn} g(x^*) = \operatorname{sgn} g(y^*) = \operatorname{sgn} g_n(y^*).$$

Thus

$$\max_{x \in A_g} [\operatorname{sgn} g(x)] m(x) = [\operatorname{sgn} g(x^*)] m(x^*) < [\operatorname{sgn} g_n(y^*)] m(y^*) + \varepsilon.$$

As this is valid for all $m \in M$, $\|m\| = 1$, it implies that for all $n > N$

$$\gamma(f) = \gamma(g) < \gamma(g_n) + \varepsilon = \gamma(f_n) + \varepsilon. \quad \square$$

Remark. In Propositions 4.6 and 4.7, we did not really use the full Haar space property. We used the fact that the f and f_n all had strongly unique best approximants from M.

It would be desirable if we could bound $\gamma(f)$ from below away from zero for all $f \in C(B)$ and thus dispense with the dependence upon f. This is certainly possible if M is a Haar space of dimension 1. In this case, $M = \operatorname{span}\{m\}$ where $m \in C(B)$ has no zero, and from the definition of $\gamma(f)$, we obtain

$$\gamma(f) \geq \min_{\substack{m \in M \\ \|m\|=1}} |m(x)| = \min_{m \in M} \frac{|m(x)|}{\|m\|} > 0,$$

a lower bound independent of f.

Moreover, if B is a finite set, then we can always uniformly bound $\gamma(f)$ from below away from zero if M is a Haar space. This follows simply from a compactness argument. However if B is not a discrete set and M is a Haar space of dimension n, $n \geq 2$, then $\gamma(f)$ cannot be bounded away from zero. The following is a variation on a result proved in Cline [1973].

Proposition 4.8. *Let B be a compact Hausdorff space that is not discrete. Assume $M \subset C(B)$ is a Haar space of dimension n, $n \geq 2$. Then*

$$\inf_{\substack{f \in C(B) \\ \|f\|=1}} \gamma(f) = 0.$$

Proof: Let $\widetilde{x}$ be any accumulation point of B, i.e., every neighbourhood of $\widetilde{x}$ contains at least one other point (and thus an infinite number of points) of B. As M is of dimension n, $n \geq 2$, there exists an $\widetilde{m} \in M$ satisfying $\widetilde{m}(\widetilde{x}) = 0$ and $\|\widetilde{m}\| = 1$. Given $\varepsilon > 0$, choose $n+1$ arbitrary distinct points $\{x_i\}_{i=1}^{n+1}$ in B where $|\widetilde{m}(x_i)| < \varepsilon$. Such points can all be chosen from a suitable neighbourhood of $\widetilde{x}$. Since M is a Haar space of dimension n, there exist $a_1, \ldots, a_{n+1}$, all nonzero, such that

$$\sum_{i=1}^{n+1} a_i m(x_i) = 0 \tag{4.1}$$

for all $m \in M$. Construct an $f \in C(B)$ satisfying $f(x_i) = \operatorname{sgn} a_i$, $i = 1, \ldots, n+1$, $\|f\| = 1$ and $\|f - \widetilde{m}\| \leq 1 + \varepsilon$. Such a construction is clearly possible since $|(f - \widetilde{m})(x_i)| \leq |f(x_i)| + |\widetilde{m}(x_i)| < 1 + \varepsilon$ for all $i = 1, \ldots, n+1$.

From (4.1) and the fact that $f(x_i) = \operatorname{sgn} a_i$, $i = 1, \ldots, n+1$, we have that $0 \in P_M(f)$. Thus from the strong uniqueness property of Haar spaces, we necessarily have

$$\|f - \widetilde{m}\| - \|f\| \geq \gamma(f)\|\widetilde{m}\|$$

implying

$$\varepsilon = (1+\varepsilon) - 1 \geq \|f - \widetilde{m}\| - \|f\| \geq \gamma(f)\|\widetilde{m}\| = \gamma(f),$$

i.e., $\gamma(f) \leq \varepsilon$. Thus

$$\inf_{\substack{f \in C(B) \\ \|f\|=1}} \gamma(f) = 0. \quad \square$$

While there is no uniform strictly positive lower bound on $\gamma(f)$, there are results concerning sets in $C(B)$ over which the strong uniqueness constants are uniformly bounded from below by a positive constant; see for example Bartelt, Swetits [1983], Marinov [1983], and Bartelt, Swetits [1988]. In addition, there are specific bounds on $\gamma(f)$ related to the structure of the A_{f-m^*}. To explain how these latter bounds arise, we start with a construction.

As usual, we assume that M is an n-dimensional Haar space in $C(B)$. Given any $n+1$ distinct point $x_1, \ldots, x_{n+1} \in B$, there exist a_i, all nonzero, such that

$$\sum_{i=1}^{n+1} a_i m(x_i) = 0$$

for all $m \in M$. Normalize the a_i so that

$$\sum_{i=1}^{n+1} |a_i| = 1.$$

As M is a Haar space of dimension n, there also exist $m_j \in M$ satisfying

$$m_j(x_i) = \operatorname{sgn} a_i, \qquad i = 1, \ldots, n+1,\ i \neq j.$$

Thus for each $j \in \{1, \ldots, n+1\}$,

$$0 = \sum_{i=1}^{n+1} a_i m_j(x_i) = \sum_{\substack{i=1 \\ i \neq j}}^{n+1} |a_i| + a_j m_j(x_j) = (1 - |a_j|) + a_j m_j(x_j),$$

and therefore

$$[\operatorname{sgn} a_j] m_j(x_j) = 1 - \frac{1}{|a_j|} < 0.$$

These m_j, $j = 1, \ldots, n+1$, will play a special role in determining and bounding $\gamma(f)$. The m_j were introduced in Cline [1973]. The important property of the m_j is the following.

Proposition 4.9. *Assume M is an n-dimensional Haar subspace of $C(B)$. Given any $n+1$ distinct points $\{x_i\}_{i=1}^{n+1}$ in B, let the $\{a_i\}_{i=1}^{n+1}$ and $\{m_j\}_{j=1}^{n+1}$ be as constructed above. Then*

$$\min_{\substack{m \in M \\ \|m\|=1}} \max_{i=1,\ldots,n+1} [\operatorname{sgn} a_i] m(x_i) = \min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

Proof: We first prove that any n of the above $\{m_j\}_{j=1}^{n+1}$ are linearly independent, i.e., span M. Assume not, and without loss of generality let us assume that $m_2, \ldots, m_{n+1}$ are linearly dependent. Thus

$$\sum_{k=2}^{n+1} b_k m_k = 0$$

for some nontrivial choice of $b_2, \ldots, b_{n+1}$. At x_1, we have $m_k(x_1) = \operatorname{sgn} a_1$ for all $k = 2, \ldots, n+1$. Thus

$$\sum_{k=2}^{n+1} b_k = 0.$$

At x_j, $j \neq 1$, we have

$$\begin{aligned} 0 = [\operatorname{sgn} a_j] \sum_{k=2}^{n+1} b_k m_k(x_j) &= [\operatorname{sgn} a_j] b_j m_j(x_j) + \sum_{\substack{k=2 \\ k \neq j}}^{n+1} b_k \\ &= b_j \left(1 - \frac{1}{|a_j|}\right) + \sum_{\substack{k=2 \\ k \neq j}}^{n+1} b_k = -\frac{b_j}{|a_j|}. \end{aligned}$$

As $|a_j| > 0$, we have $b_j = 0$ for all $j \neq 1$, a contradiction.

Let $m \in M$ be normalized so that

$$\max_{i=1,\ldots,n+1} [\operatorname{sgn} a_i] m(x_i) = 1.$$

Note that the above maximum is always strictly positive for any $m \in M$, $m \neq 0$, since we have $\sum_{i=1}^{n+1} a_i m(x_i) = 0$, and $m \neq 0$ can only vanish at at most $n-1$ distinct points. Assume

$$[\operatorname{sgn} a_r] m(x_r) = 1.$$

As any n of the above $\{m_j\}_{j=1}^{n+1}$ are linearly independent, we have

$$m = \sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i m_i.$$

At x_r, we have

$$1 = [\operatorname{sgn} a_r]m(x_r) = \sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i [\operatorname{sgn} a_r] m_i(x_r) = \sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i.$$

We now prove that $c_i \geq 0$ for all i as above. To this end, note that for $j \in \{1, \ldots, n+1\}\backslash\{r\}$,

$$\begin{aligned} 1 \geq [\operatorname{sgn} a_j]m(x_j) &= [\operatorname{sgn} a_j]\sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i m_i(x_j) = \sum_{\substack{i=1\\ i\neq r,j}}^{n+1} c_i + [\operatorname{sgn} a_j] c_j m_j(x_j) \\ &= (1 - c_j) + c_j(1 - \frac{1}{|a_j|}) = 1 - \frac{c_j}{|a_j|}. \end{aligned}$$

Thus

$$1 \geq 1 - \frac{c_j}{|a_j|}$$

and as $|a_j| > 0$, this implies that $c_j \geq 0$.

Returning to our definition of m, we have, using the fact that $c_i \geq 0$ and $\sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i = 1$,

$$\|m\| = \|\sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i m_i\| \leq \sum_{\substack{i=1\\ i\neq r}}^{n+1} c_i \|m_i\| \leq \max_{i=1,\ldots,n+1} \|m_i\|.$$

That is, for every $m \in M$ satisfying

$$\max_{i=1,\ldots,n+1} [\operatorname{sgn} a_i]m(x_i) = 1,$$

we have

$$\frac{1}{\|m\|} \geq \min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

In addition, we have

$$\max_{i=1,\ldots,n+1} [\operatorname{sgn} a_i]m_j(x_i) = 1$$

for $j = 1, \ldots, n+1$. $\square$

Let $f \in C(B)\backslash M$ and m^* be the best approximant to f from M. Recall that as M is an n-dimensional Haar subspace of $C(B)$, there exist $n+1$ distinct point $x_1, \ldots, x_{n+1} \in A_{f-m^*}$ and strictly positive $\lambda_1, \ldots, \lambda_{n+1}$, with $\sum_{i=1}^{n+1} \lambda_i = 1$, satisfying

$$\sum_{i=1}^{n+1} \lambda_i [\mathrm{sgn}(f - m^*)(x_i)] m(x_i) = 0$$

for all $m \in M$. In what follows, we let the $m_j \in M$, $j = 1, \ldots, n+1$, be as constructed above, with respect to these $x_1, \ldots, x_{n+1}$ and $a_i = \lambda_i [\mathrm{sgn}(f - m^*)(x_i)]$, $i = 1, \ldots, n+1$. From Proposition 4.9, we immediately obtain the following results.

Theorem 4.10. *Assume M is an n-dimensional Haar subspace of $C(B)$. Let $f \in C(B)\backslash M$ and m^* be the best approximant to f from M. Let the m_j, $j = 1, \ldots, n+1$, be as constructed above. Then*

$$\gamma(f) \geq \min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

Furthermore, if $A_{f-m^*} = \{x_1, \ldots, x_{n+1}\}$ then

$$\gamma(f) = \min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

Proof: We have that

$$\begin{aligned} \gamma(f) &= \min_{\substack{m \in M \\ \|m\|=1}} \max_{x \in A_{f-m^*}} \mathrm{sgn}(f - m^*)(x) m(x) \\ &\geq \min_{\substack{m \in M \\ \|m\|=1}} \max_{i=1,\ldots,n+1} \mathrm{sgn}(f - m^*)(x_i) m(x_i). \end{aligned}$$

From Proposition 4.9, we see that this latter quantity equals

$$\min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

This proves the first statement of the theorem. The second statement follows analogously since the inequality is then an equality. □

The first statement in Theorem 4.10 is in Cline [1973]. The second statement follows easily from results in Cline [1973] together with the characterization of Theorem 2.2, as noted in Henry, Roulier [1978]. According to Blatt [1986], the first inequality was partially implicitly arrived at in Freud [1958].

If A_{f-m^*} has exactly $n+1$ points, then from the above construction each of the m_j is “admissible”, i.e., satisfies

$$\max_{x \in A_{f-m^*}} [\mathrm{sgn}(f - m^*)(x)] m_j(x) = 1, \qquad j = 1, \ldots, n+1.$$

However, if A_{f-m^*} has more than $n+1$ points then it may be that some of the above constructed m_j satisfy

$$\max_{x \in A_{f-m^*}} [\mathrm{sgn}(f - m^*)(x)] m_j(x) > 1.$$

This is why equality need not hold in the first part of Theorem 4.10.

Given A_{f-m^*}, it is also not clear, a priori, which choices of $n+1$ points $x_1, \ldots, x_{n+1}$ in A_{f-m^*} can be nade so as to satisfy

$$\sum_{i=1}^{n+1} \lambda_i[\mathrm{sgn}(f-m^*)(x_i)]m(x_i) = 0$$

for all $m \in M$ with $\lambda_i > 0$, $i = 1, \ldots, n+1$ (except of course if A_{f-m^*} contains only $n+1$ points). However if B is a connected interval then the "eligible" $x_1 < \cdots < x_{n+1}$ are exactly those for which $[\mathrm{sgn}(f-m^*)(x_i)][\mathrm{sgn}(f-m^*)(x_{i+1})] < 0$, $i = 1, \ldots, n$, (see Theorem 4.4).

In Bartelt, Henry [1980], there is an example of a function f for which $\gamma(f)$ is strictly greater than $\min_{j=1,\ldots,n+1} 1/\|m_j\|$ when one varies over all possible "eligible" $\{x_i\}_{i=1}^{n+1}$ (see the previous paragraph). It is the following.

Example. Let $M = \Pi_1$ be the polynomials of degree ≤ 1 in $C[0,1]$. Let $f \in C[0,1]$ be piecewise linear with nodes $f(0) = f(1/2) = f(1) = 1$ and $f(1/4) = f(3/4) = -1$. It can be readily verified that $\gamma(f) = 3/5$ while the maximum of $\min_{j=1,\ldots,n+1} 1/\|m_j\|$ where we choose 3 arbitrary alternating points is $1/5$ (attained choosing the points $1/4, 1/2, 3/4$).

A simple consequence of Theorem 4.10 is the following, as was noted in Blatt [1984b].

Corollary 4.11. *Assume M is an n-dimensional Haar subspace of $C(B)$. Let $f \in C(B)\backslash M$ and m^* be the best approximant to f from M. Assume $A_{f-m^*} = \{x_1, \ldots, x_{n+1}\}$. Then*

$$\gamma(f) \leq \frac{1}{n}.$$

Proof: From Theorem 4.10, we have

$$\gamma(f) = \min_{j=1,\ldots,n+1} \frac{1}{\|m_j\|}.$$

From the construction of the m_j

$$|m_j(x_j)| = \frac{1}{\lambda_j} - 1.$$

Since $\lambda_i > 0$ and $\sum_{i=1}^{n+1} \lambda_i = 1$, there exists a λ_j satisfying $\lambda_j \leq 1/(n+1)$. Thus, for this j,

$$\|m_j\| \geq |m_j(x_j)| = \frac{1}{\lambda_j} - 1 \geq n,$$

implying $\gamma(f) \leq 1/n$. □

This bound of $1/n$ is independent of both B and M. There is no reason to assume that it is attainable. However Cline [1973] proved that if $B = [-1,1]$, $M = \Pi_{n-1} = \mathrm{span}\{1, x, \ldots, x^{n-1}\}$, and the $\{x_i\}_{i=1}^{n+1}$ are the extreme points of the Chebyshev polynomial of degree n, then $\gamma(f) = 1/(2n-1)$. So at least in these cases the order $O(n)$ is attainable.

If $M = \Pi_{n-1}$ and $B = [a,b]$, then it is well known that if $f \in C^n[a,b]$ and $f^{(n)}(x)$ is strictly of one sign for all $x \in [a,b]$, then A_{f-m^*} has exactly $n+1$ points. As such, it follows from Corollary 4.11 that the strong unicity constant associated with, for example, e^x necessarily tends to 0 as $n \uparrow \infty$. It has been shown by Gehlen [1999] that if f is not a polynomial, then necessarily $\liminf \gamma_n(f) = 0$, where $\gamma_n(f)$ is the strong unicity constant for f on $[a,b]$ with respect to the polynomial approximating space Π_{n-1}. This was formally conjectured in Henry, Roulier [1978] and is sometimes called the Poreda conjecture; see Poreda [1976]. A lot of work went into this conjecture until it was finally proven by Gehlen. It should be noted that the above lim inf cannot be replaced by the simple limit. There are functions $f \in C[a,b]$ for which

$$\liminf_{n\to\infty} \gamma_n(f) = 0, \qquad \limsup_{n\to\infty} \gamma_n(f) = 1;$$

see Schmidt [1978]. From this example, we also see that $\gamma_n(f)$ is not, in general, a monotone function of n. Nothing is known concerning the behavior of the associated $\gamma_n(f)$ on other dense subspaces.

If A_{f-m^*} contains more than $n+1$ points, then via the $\|m_j\|$, we cannot necessarily calculate the exact value of $\gamma(f)$. It is therefore natural to ask how we can calculate $\gamma(f)$. The following, from Schmidt [1980], characterizes the elements of M that attain the minimum in the formula for the determination of $\gamma(f)$.

Theorem 4.12. *Assume M is an n-dimensional Haar subspace of $C(B)$. Let $f \in C(B)\backslash M$ and m^* be the best approximant to f from M. Assume $\widetilde{m} \in M$, $\|\widetilde{m}\| = 1$, satisfies*

$$\gamma(f) = \max_{x \in A_{f-m^*}} [\operatorname{sgn}(f - m^*)(x)]\widetilde{m}(x).$$

Then given any $x^* \in B$ satisfying

$$|\widetilde{m}(x^*)| = \|\widetilde{m}\|,$$

there exist points $x_1, \ldots, x_n \in A_{f-m^*}$ and strictly positive values $\lambda_1, \ldots, \lambda_{n+1}$ such that

$$[\operatorname{sgn}(f - m^*)(x_i)]\widetilde{m}(x_i) = \gamma(f), \qquad i = 1, \ldots, n,$$

and

$$\sum_{i=1}^{n} \lambda_i [\operatorname{sgn}(f - m^*)(x_i)] m(x_i) - \lambda_{n+1} [\operatorname{sgn} \widetilde{m}(x^*)] m(x^*) = 0$$

for all $m \in M$.

Proof: There necessarily exists an $\widetilde{m} \in M$, $\|\widetilde{m}\| = 1$, satisfying

$$\gamma(f) = \max_{x \in A_{f-m^*}} [\operatorname{sgn}(f - m^*)(x)]\widetilde{m}(x).$$

Set

$$D := \{x : x \in A_{f-m^*}, [\operatorname{sgn}(f - m^*)(x)]\widetilde{m}(x) = \gamma(f)\}$$

and let x^* be as above, i.e., $|\widetilde{m}(x^*)| = \|\widetilde{m}\|$. We first claim that there does not exist an $m \in M$ satisfying

$$[\operatorname{sgn}(f - m^*)(x)] m(x) < 0, \qquad x \in D,$$

and

$$\widetilde{m}(x^*) m(x^*) > 0.$$

For if such an m existed, then for $\varepsilon > 0$ sufficiently small $m_\varepsilon = \widetilde{m} + \varepsilon m$ would satisfy

$$[\operatorname{sgn}(f - m^*)(x)] m_\varepsilon(x) < \gamma(f), \qquad x \in A_{f-m^*}$$

and

$$\|m_\varepsilon\| > \|\widetilde{m}\| = 1,$$

contradicting the fact that

$$\gamma(f) = \min_{\substack{m\in M\\ \|m\|=1}} \max_{x\in A_{f-m^*}} \operatorname{sgn}(f-m^*)(x)m(x) = \max_{x\in A_{f-m^*}} \operatorname{sgn}(f-m^*)(x)\widetilde{m}(x).$$

Let $m_1,\ldots,m_n$ be any basis for M. For each $x \in D$, set

$$M(x) = ([\operatorname{sgn}(f-m^*)(x)]m_1(x),\ldots,[\operatorname{sgn}(f-m^*)(x)]m_n(x)) \in \mathbb{R}^n,$$

and let

$$N(x^*) = (-[\operatorname{sgn}\widetilde{m}(x^*)]m_1(x),\ldots,-[\operatorname{sgn}\widetilde{m}(x^*)]m_n(x)).$$

As M is a linear subspace and there exists no $m \in M$ such that

$$[\operatorname{sgn}(f-m^*)(x)]m(x) < 0, \qquad x \in D$$

and

$$-[\operatorname{sgn}\widetilde{m}(x^*)]m(x^*) < 0,$$

it follows that $\mathbf{0} \in \mathbb{R}^n$ is in the convex hull of $\{M(x)\}_{x\in D} \cup N(x^*)$. From Carathéodory's Theorem, this implies that $\mathbf{0}$ is a convex combination of at most $n+1$ points in this set. If these k points, $1 \le k \le n+1$, do not include $N(x^*)$, then there exist $\lambda_i \ge 0$, $\sum_{i=1}^k \lambda_i = 1$, such that

$$\sum_{i=1}^k \lambda_i[\operatorname{sgn}(f-m^*)(x_i)]m(x_i) = 0$$

for all $m \in M$. Substituting $m = \widetilde{m}$, and since $[\operatorname{sgn}(f-m^*)(x_i)]\widetilde{m}(x_i) = \gamma(f) > 0$ for $i = 1,\ldots,k$, we get a contradiction. Thus there exist $\lambda_i \ge 0$, $\sum_{i=1}^k \lambda_i = 1$, $\lambda_k > 0$, such that

$$\sum_{i=1}^{k-1} \lambda_i[\operatorname{sgn}(f-m^*)(x_i)]m(x_i) - \lambda_k[\operatorname{sgn}\widetilde{m}(x^*)]m(x^*) = 0$$

for all $m \in M$. In fact, since M is a Haar space of dimension n, we must have $k = n+1$ and $\lambda_i > 0$ for all $i = 1,\ldots,n+1$. $\square$

Example. Let us return to the example presented just before Corollary 4.11, namely let $M = \Pi_1$ be the polynomials of degree ≤ 1 in $C[0,1]$. Let $f \in C[0,1]$ be piecewise linear with nodes $f(0) = f(1/2) = f(1) = 1$ and $f(1/4) = f(3/4) = -1$. It can be readily verified that $\gamma(f) = 3/5$ while the maximum of $\min_{j=1,\ldots,n+1} 1/\|m_j\|$ where we choose 3 alternating points is $1/5$ (attained choosing the points $1/4, 1/2, 3/4$). The $\widetilde{m}$ that gives this $\gamma(f) = 3/5$ is obtained by considering the linear function $\widetilde{m}$ satisfying $\widetilde{m}(1/4) = -3/5$ and $\widetilde{m}(1) = 3/5$, i.e., $\widetilde{m}(x) = (8x-5)/5$.

Remark. In the case where B is well-ordered and M is a T-system thereon, then the $x_1 < \cdots < x_n$ as in the statement of Theorem 4.12 are such that either $f - m^*$ strictly alternates thereon, in which case the $\widetilde{m}$ takes its norm at $x^* < x_1$ with $\operatorname{sgn}\widetilde{m}(x^*) = \operatorname{sgn}(f-m^*)(x_1)$ or $\widetilde{m}$ takes its norm at $x^* > x_n$ with $\operatorname{sgn}\widetilde{m}(x^*) = \operatorname{sgn}(f-m^*)(x_n)$, or $f - m^*$ strictly alternates thereon except for having the same sign at x_i and x_{i+1} for some $i \in \{1, \dots, n\}$ and $\widetilde{m}$ takes its norm at $x_i < x^* < x_{i+1}$ with $\operatorname{sgn} \widetilde{m}(x^*) = \operatorname{sgn}(f - m^*)(x_i) = \operatorname{sgn}(f - m^*)(x_{i+1})$.

For which subspaces of $C(B)$ are uniqueness and strong uniqueness equivalent? That is, which subspaces M have the property that if $f \in C(B)$ has a unique best approximant from M, then it is also a strongly unique best approximant? We know that they are equivalent if M is a finite-dimensional Haar space since in this case, we always have both uniqueness and strong uniqueness of the best approximant. In McLaughlin, Somers [1975], it is proved that if B is an interval $[a, b]$, then uniqueness and strong uniqueness for all functions in the space are equivalent if and only if M is a Haar space, i.e., if M is not a Haar space on $C[a, b]$ then there exist functions with a unique best approximant, but not a strongly unique best approximant. The proof is difficult and complicated, and is restricted to the case $B = [a, b]$.

If B is a nonconnected set, there may exist subspaces M of $C(B)$ such that uniqueness and strong uniqueness are equivalent, and yet M is not a Haar space. As an example, consider $M = \operatorname{span}\{(1, 0)\}$ in $\mathbb{R}^2$. Then for $f = (x, y)$, we can have $f \in M$, i.e., $y = 0$, in which case we trivially have both uniqueness and strong uniqueness of the best approximant. On the other hand, if $f \notin M$, i.e., $y \neq 0$, then f never has a unique best approximant. The vectors $c(1, 0)$ are best approximants to (x, y) from M for all $c \in [x - |y|, x + |y|\,]$ as for all such c, we have

$$\|(x, y) - c(1, 0)\| = \max\{|x - c|, |y|\} = |y|.$$

Thus uniqueness and strong uniqueness hold simultaneously or do not hold at all. A more general result is the following

Proposition 4.13. *Assume B is a compact Hausdorff space and M is an n-dimensional subspace of $C(B)$. Assume $B = B_1 \cup \cdots \cup B_k$ where each B_j is both open and closed, $M|_{B_j} = M_j$ is a Haar space of dimension n_j (assuming $n_j \geq 1$) on B_j, and*

$$\sum_{j=1}^{k} n_j = n.$$

That is, we assume M_j (if $n_j \geq 1$) has a basis of functions that vanish off B_j. Then uniqueness and strong uniqueness are equivalent on $C(B)$.

Proof: From the above assumptions, we can write

$$M = M_1 \oplus M_2 \oplus \cdots \oplus M_k$$

and each $m \in M$ has a unique decomposition of the form

$$m = \sum_{j=1}^{k} m_j$$

where $m_j \in M_j$ (and m_j vanishes off B_j). Thus

$$\|f - m\|_B = \|f - \sum_{j=1}^{k} m_j\|_B = \max_{j=1,\dots,k} \|f - m\|_{B_j} = \max_{j=1,\dots,k} \|f - m_j\|_{B_j}.$$

We first claim that f has a unique best approximant from M if and only if

$$\min_{m\in M}\|f-m\|_B = \min_{m_j\in M_j}\|f-m_j\|_{B_j}$$

for each $j\in\{1,\ldots,k\}$ for which $n_j\geq 1$. Assume $m^*\in M$ is a best approximant to f from M, and there exists an $\ell\in\{1,\ldots,k\}$ for which $n_\ell\geq 1$ and

$$\min_{m\in M}\|f-m\|_B > \min_{m_\ell\in M_\ell}\|f-m_\ell\|_{B_\ell}.$$

Let $m_\ell^*\in M_\ell$ be the best approximant to f from M_ℓ on B_ℓ. As $n_\ell\geq 1$, there exists an $\widetilde{m}_\ell\in M_\ell$, $\widetilde{m}_\ell\neq 0$, and $\widetilde{m}_\ell$ vanishes off B_ℓ. Thus

$$\min_{m\in M}\|f-m\|_B > \|f-(m_\ell^*+\varepsilon\widetilde{m}_\ell)\|_{B_\ell}$$

for all ε sufficiently small, implying that $m^*+\varepsilon\widetilde{m}_\ell$ are best approximants to f from M for all ε sufficiently small. Thus, if f has a unique best approximant from M then

$$\min_{m\in M}\|f-m\|_B = \min_{m_j\in M_j}\|f-m_j\|_{B_j}$$

for each $j\in\{1,\ldots,k\}$ with $n_j\geq 1$.

Let us now assume that

$$\min_{m\in M}\|f-m\|_B = \min_{m_j\in M_j}\|f-m_j\|_{B_j}$$

for each $j\in\{1,\ldots,k\}$ for which $n_j\geq 1$. We claim that in this case, we always have strong uniqueness that also, of course, implies uniqueness. Let $m_j^*\in M_j$ be the unique best approximant to f from M_j on B_j (that vanishes off B_j). Set $m^*:=\sum_{j=1}^k m_j^*$. Thus m^* is a best approximant to f from M and

$$\|f-m^*\|_B = \|f-m_j^*\|_{B_j}$$

for all j with $n_j\geq 1$. As M_j is a Haar space (assuming $n_j\geq 1$) on B_j, there exists a $\gamma_j(f)>0$ for which

$$\|f-m_j\|_{B_j} - \|f-m_j^*\|_{B_j} \geq \gamma_j(f)\|m_j-m_j^*\|_{B_j}$$

for all $m_j\in M_j$. Set

$$\gamma(f) := \min_{\substack{j=1,\ldots,k\\ n_j\geq 1}}\gamma_j(f).$$

Let $m\in M$ and $m=\sum_{j=1}^k m_j$ where $m_j\in M_j$. For each $j\in\{1,\ldots,k\}$ with $n_j\geq 1$, we have, since $\|f-m^*\|_B=\|f-m_j^*\|_{B_j}$,

$$\begin{aligned}\|f-m\|_B-\|f-m^*\|_B &\geq \|f-m_j\|_{B_j}-\|f-m_j^*\|_{B_j}\\ &\geq \gamma_j(f)\|m_j-m_j^*\|_{B_j}\\ &\geq \gamma(f)\|m_j-m_j^*\|_{B_j}.\end{aligned}$$

If $n_j=0$, then $m_j=m_j^*=0$ and thus

$$\|f-m\|_B-\|f-m^*\|_B\geq 0=\gamma(f)\|m_j-m_j^*\|_{B_j}.$$

As we now have

$$\|f - m\|_B - \|f - m^*\|_B \geq \gamma(f)\|m_j - m_j^*\|_{B_j}$$

for all $j = 1, \ldots, k$, this implies that

$$\|f - m\|_B - \|f - m^*\|_B \geq \gamma(f)\|m - m^*\|_B. \quad \square$$

It would be interesting to determine exact necessary and sufficient conditions for when uniqueness and strong uniqueness are equivalent on $C(B)$.

Unfortunately, the conditions of Proposition 4.13 are not necessary. For example, consider the space $M = \operatorname{span}\{m_1, m_2\}$ on $B = [0,1] \cup [2,3]$ where

$$m_1(x) = \begin{cases} x, & x \in [0,1] \\ 1, & x \in [2,3] \end{cases}$$

and

$$m_2(x) = \begin{cases} 1, & x \in [0,1] \\ 0, & x \in [2,3] \end{cases}.$$

In this example, uniqueness and strong uniqueness are equivalent, but this example does not satisfy the conditions of Proposition 4.13. Uniqueness and strong uniqueness are equivalent because for uniqueness of the best approximant from M to hold it is necessary that the error function attains its norm alternately at least three times in $[0,1] \cup [2,3]$, and two of these alternants must be in $[0,1]$. Details are left to the reader. But in these cases, we also get strong uniqueness. This latter fact follows from Theorem 2.4.

For $\dim M = 1$, we have the following result. The argument is from McLaughlin, Somers [1975].

Proposition 4.14. *Assume B is a compact Hausdorff space and $M = \operatorname{span}\{m\}$ is a 1-dimensional subspace of $C(B)$. Assume uniqueness and strong uniqueness from M are equivalent on $C(B)$. Then either M is a Haar space on B, i.e., m vanishes nowhere on B, or $B = B_1 \cup B_2$ where B_1, B_2 are both open and closed, and m vanishes nowhere on B_1 but vanishes identically on B_2.*

Proof: To prove the proposition, it suffices to prove the following. If $x_0 \in B$ is such that $m(x_0) = 0$, then m vanishes identically in a neighbourhood of x_0. Assume this is not the case. That is, there exists an $x_0 \in B$ such that $m(x_0) = 0$ and m does not vanish identically in any neighbourhood of x_0. Then, we can choose some x_0 such that $m(x_0) = 0$ and there exists a small, closed, nondegenerate set N such that $x_0 \in \partial N$ and $\varepsilon m(x) > 0$ for all $x \in N \backslash \{x_0\}$ for some $\varepsilon \in \{-1, 1\}$. We assume, taking cm if necessary for some $c \in \mathbb{R}$, that $\varepsilon = 1$ and $\|m\| = 1$. Let (x_n) be a sequence of points in N converging to x_0. Thus $m(x_n) > 0$ for all n, and $m(x_n)$ converges to zero. Let $y \in B \backslash N$ with $m(y) \neq 0$. We define $f \in C(B)$ as follows. First set $f(y) = -\operatorname{sgn} m(y)$ and $f(x_n) = 1 - m^2(x_n)$. Note that $f(x_0) = 1$. Extend f to be in $C(B)$ satisfying $\|f\| = 1$ and $|f(x)| < 1$ for all $x \in B \backslash \{x_0, y\}$.

As $\|f\| = 1$ and $f(x_0) = 1$, it follows that the 0 function is a best approximant to f from M. From the construction of f, we have $A_f = \{x_0, y\}$. A simple calculation gives

$$\gamma(f) = \min_{\pm} \max\{\pm m(x_0), \mp(\operatorname{sgn} m(y))m(y)\} = 0.$$

Thus, we do not have strong uniqueness of the best approximant to f from M. It remains to prove we have uniqueness.

For $\alpha > 0$,

$$|f(y) - \alpha m(y)| = |-\operatorname{sgn} m(y) - \alpha(\operatorname{sgn} m(y))|m(y)|| = 1 + \alpha|m(y)| > 1,$$

implying αm is not a best approximant for $\alpha > 0$. Assume $\alpha < 0$ is such that αm is a best approximant. Thus for each x_n, we have

$$|f(x_n) - \alpha m(x_n)| = |1 - m^2(x_n) - \alpha m(x_n)| \leq 1.$$

This implies that we must have

$$m^2(x_n) + \alpha m(x_n) \geq 0$$

for all n. But $\alpha < 0$ and $m(x_n) > 0$ converges to zero, which is impossible. Thus the 0 function is the unique best approximant to f from M. □

5. Classical Strong Uniqueness in the L^1 Norm

In this section, as the title indicates, we will study strong uniqueness when our approximating norm is L^1. To this end, let us assume that K is a set, Σ a σ-field of subsets of K and ν a positive measure on Σ. By $L^1(K,\nu)$, we mean the set of real-valued functions f that are ν-measurable, and such that $|f|$ is integrable. The norm on $L^1(K,\nu)$ is given by

$$\|f\|_1 = \int_K |f(x)| \, \mathrm{d}\nu(x).$$

Two important sets when approximating in this $L^1(K,\nu)$ norm are the following. For each $f \in L^1(K,\nu)$, we define its zero set as

$$Z(f) = \{x : f(x) = 0\}$$

and also

$$N(f) = K \backslash Z(f).$$

Note that $Z(f)$ is ν-measurable.

To understand when classical strong uniqueness might hold, we first calculate the $\tau_+(f,g)$ of (1.1).

Theorem 5.1. *For given $f, g \in L^1(K,\nu)$, we have*

$$\tau_+(f,g) = \int_{Z(f)} |g(x)| \, \mathrm{d}\nu(x) + \int_K [\operatorname{sgn} f(x)] g(x) \, \mathrm{d}\nu(x).$$

Proof: Let us assume that $f, g \neq 0$. By definition,

$$\tau_+(f,g) = \lim_{t \to 0^+} \frac{\|f + tg\|_1 - \|f\|_1}{t}.$$

For $t > 0$,

$$\begin{aligned}
\frac{\|f+tg\|_1 - \|f\|_1}{t} &= \frac{1}{t}\left[\int_K (|f+tg| - |f|)\,\mathrm{d}\nu\right] \\
&= \frac{1}{t}\left[\int_{Z(f)} t|g|\,\mathrm{d}\nu + \int_{N(f)} (|f+tg| - |f|)\,\mathrm{d}\nu\right] \\
&= \int_{Z(f)} |g|\,\mathrm{d}\nu + \frac{1}{t}\left[\int_{N(f)} (|f+tg| - |f|)\,\mathrm{d}\nu\right].
\end{aligned}$$

On $N(f)$,

$$\left|\frac{|f+tg| - |f|}{t}\right| \le |g|$$

and

$$\begin{aligned}
\frac{|f+tg| - |f|}{t} &= \frac{|f+tg|^2 - |f|^2}{t\,[|f+tg| + |f|]} \\
&= \frac{2fg + t|g|^2}{|f+tg| + |f|}.
\end{aligned}$$

Thus on $N(f)$,

$$\lim_{t\to 0^+} \frac{|f+tg| - |f|}{t} = \frac{2fg}{2|f|} = [\operatorname{sgn} f]g.$$

Applying Lebesgue's Dominated Convergence Theorem, we obtain

$$\tau_+(f,g) = \int_{Z(f)} |g|\,\mathrm{d}\nu + \int_K [\operatorname{sgn} f]g\,\mathrm{d}\nu. \quad \square$$

As a consequence of Theorem 5.1, Theorem 1.2 and Theorem 1.4, we have:

Theorem 5.2. *Let M be a linear subspace of $L^1(K,\nu)$ and $f \in L^1(K,\nu)\backslash M$. Then $m^* \in P_M(f)$ if and only if*

$$\left|\int_K [\operatorname{sgn}(f-m^*)(x)]m(x)\,\mathrm{d}\nu(x)\right| \le \int_{Z(f-m^*)} |m(x)|\,\mathrm{d}\nu(x)$$

for all $m \in M$. Furthermore, $m^* \in M$ is a strongly unique best approximant to f from M if and only if

$$\gamma(f) = \inf_{\substack{m\in M\\ \|m\|_1=1}} \left\{\int_{Z(f-m^*)} |m(x)|\,\mathrm{d}\nu(x) + \int_K [\operatorname{sgn}(f-m^*)(x)]m(x)\,\mathrm{d}\nu(x)\right\} > 0.$$

The first statement in this theorem can be found in James [1947, p. 291].

Recall (see Proposition 1.6) that if M is a finite-dimensional subspace, then the infimum in the definition of $\gamma(f)$ is a minimum, i.e., it is attained. To prove strong uniqueness, it therefore suffices to verify that

$$\int_{Z(f-m^*)} |m(x)|\,\mathrm{d}\nu(x) + \int_K [\operatorname{sgn}(f-m^*)(x)]m(x)\,\mathrm{d}\nu(x) > 0$$

for each $m \in M$, $\|m\| = 1$, and that is equivalent to showing that

$$\left| \int_K [\operatorname{sgn}(f - m^*)(x)] m(x) \, \mathrm{d}\nu(x) \right| < \int_{Z(f-m^*)} |m(x)| \, \mathrm{d}\nu(x)$$

for all $m \in M$, $m \neq 0$.

We consider two scenarios. In the first case, ν is a discrete positive measure with a finite number of points of support, and in the second case ν is a non-atomic positive measure. The two cases radically differ.

Assuming ν is a discrete positive measure with a finite number of points of support, we are effectively considering approximation in the $\ell_1^m(\mathbf{w})$ norm where

$$\ell_1^m(\mathbf{w}) := \{\mathbf{x} = (x_1, \ldots, x_m) : \mathbf{x} \in \mathbb{R}^m, \|\mathbf{x}\|_{\mathbf{w}} = \sum_{i=1}^m |x_i| w_i\}$$

and $\mathbf{w} = (w_1, \ldots, w_m)$ is a fixed strictly positive vector (weight). In this case, it transpires that uniqueness and strong uniqueness are equivalent. It is also true that for a fixed weight $\mathbf{w}$ most subspaces are unicity spaces, i.e., for most subspaces M there is a unique best approximant to each vector $\mathbf{x}$ from M. However for some subspaces, we do not have uniqueness of the best approximant. For example, in ℓ_1^2 with weight $(1,1)$, i.e., norm $\|\mathbf{x}\|_{\mathbf{w}} = |x_1| + |x_2|$, and $M = \operatorname{span}\{(1,1)\}$ then to each and every $\mathbf{x} \notin M$, there is no unique best approximant. On the other hand, all the 1-dimensional subspaces in ℓ_1^2 other than $M = \operatorname{span}\{(1, \pm 1)\}$, are unicity spaces.

The equivalence of uniqueness and strong uniqueness in this setting is a consequence of the following, that is essentially contained in Rivlin [1969, Theorem 3.6]; see also Watson [1980, p. 122] and Angelos, Schmidt [1983].

Theorem 5.3. *Let M be a finite-dimensional subspace of $\ell_1^m(\mathbf{w})$ and $\mathbf{x} \in \mathbb{R}^m$. Then $\mathbf{m}^* \in M$ is the unique best approximant to $\mathbf{x}$ from M if and only if*

$$\left| \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)] m_i w_i \right| < \sum_{i \in Z(\mathbf{x}-\mathbf{m}^*)} |m_i| w_i$$

for all $\mathbf{m} \in M$, $\mathbf{m} \neq \mathbf{0}$. Thus every unique best approximant to $\mathbf{x}$ from M is also a strongly unique best approximant to $\mathbf{x}$ from M.

Proof: Assume $\mathbf{m}^* \in M$ is the unique best approximant to $\mathbf{x}$ from M. Thus, by the characterization of the best approximant, we have

$$\left| \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)] m_i w_i \right| \leq \sum_{i \in Z(\mathbf{x}-\mathbf{m}^*)} |m_i| w_i$$

for all $\mathbf{m} \in M$. Assume $\widetilde{\mathbf{m}} \in M$, $\widetilde{\mathbf{m}} \neq \mathbf{0}$, is such that

$$\left| \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)] \widetilde{m}_i w_i \right| = \sum_{i \in Z(\mathbf{x}-\mathbf{m}^*)} |\widetilde{m}_i| w_i.$$

We shall prove that the uniqueness of the best approximant implies that such a $\widetilde{\mathbf{m}}$ cannot exist. From Theorem 5.2, this implies the equivalence of uniqueness and strong uniqueness.

Since we are in the finite-dimensional subspace $\mathbb{R}^m$,

$$\min\{|x_i - m_i^*| : i \notin Z(\mathbf{x} - \mathbf{m}^*)\} = c > 0.$$

Thus there exists an $\varepsilon > 0$ such that for all t satisfying $|t| < \varepsilon$, we have

$$\operatorname{sgn}(x_i - m_i^*) = \operatorname{sgn}(x_i - m_i^* - t\widetilde{m}_i)$$

for each $i \notin Z(\mathbf{x} - \mathbf{m}^*)$. For each such t,

$$\begin{aligned}
\|\mathbf{x} - \mathbf{m}^* - t\widetilde{\mathbf{m}}\|_{\mathbf{w}} - \|\mathbf{x} - \mathbf{m}^*\|_{\mathbf{w}} &= \sum_{i=1}^m \left[|x_i - m_i^* - t\widetilde{m}_i| - |x_i - m_i^*|\right] w_i \\
&= \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)](x_i - m_i^* - t\widetilde{m}_i - x_i + m_i^*) w_i \\
&\quad + \sum_{i \in Z(\mathbf{x} - \mathbf{m}^*)} |t\widetilde{m}_i| w_i \\
&= -t \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i + |t| \sum_{i \in Z(\mathbf{x} - \mathbf{m}^*)} |\widetilde{m}_i| w_i.
\end{aligned}$$

By assumption,

$$\left|\sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i\right| = \sum_{i \in Z(\mathbf{x} - \mathbf{m}^*)} |\widetilde{m}_i| w_i.$$

Thus

$$\|\mathbf{x} - \mathbf{m}^* - t\widetilde{\mathbf{m}}\|_{\mathbf{w}} - \|\mathbf{x} - \mathbf{m}^*\|_{\mathbf{w}} = -t \sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i + |t| \left|\sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i\right|.$$

If $\sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i = 0$, then for all such t, i.e., $|t| < \varepsilon$, we have

$$\|\mathbf{x} - \mathbf{m}^* - t\widetilde{\mathbf{m}}\|_{\mathbf{w}} = \|\mathbf{x} - \mathbf{m}^*\|_{\mathbf{w}}.$$

If $\delta\left(\sum_{i=1}^m [\operatorname{sgn}(x_i - m_i^*)]\widetilde{m}_i w_i\right) > 0$, $\delta \in \{-1, 1\}$, then

$$\|\mathbf{x} - \mathbf{m}^* - t\widetilde{\mathbf{m}}\|_{\mathbf{w}} = \|\mathbf{x} - \mathbf{m}^*\|_{\mathbf{w}}$$

for all t satisfying $0 \le t\delta < \varepsilon$. In either case, we have nonuniqueness of the best approximant. □

In the second case, we assume that ν is a non-atomic positive measure. In this case, neither strong uniqueness (nor uniqueness) is always present, but it is nevertheless around. This next result can be found in Angelos, Schmidt [1983].

Theorem 5.4. *Let ν be a non-atomic positive measure. Let M be a finite-dimensional subspace of $L^1(K, \nu)$. Then the set of $f \in L^1(K, \nu)$ that have a strongly unique best approximant from M is dense in $L^1(K, \nu)$.*

Before proving this result, we recall another characterization of the best approximant via linear functionals that, since ν is a non-atomic positive measure, has the following form. This result was first proved in Phelps [1966].

Theorem 5.5. *Let ν be a non-atomic positive measure. Let M be a finite-dimensional subspace of $L^1(K,\nu)$ and $f \in L^1(K,\nu)$. Then $m^* \in P_M(f)$ if and only if there exists an $h \in L^\infty(K,\nu)$ satisfying*

$$\text{(i) } |h(x)| = 1, \text{ all } x \in K;$$

$$\text{(ii) } \int_K hm \,\mathrm{d}\nu = 0, \text{ all } m \in M;$$

$$\text{(iii) } \int_K h(f - m^*)\,\mathrm{d}\nu = \|f - m^*\|_1.$$

Proof of Theorem 5.4: Let $f \in L^1(K,\nu)$ and assume that f does not have a strongly unique best approximant from M. For convenience, we translate f by an element of $P_M(f)$ so that we may assume, without loss of generality, that $0 \in P_M(f)$. By Theorem 5.5, there exists an $h \in L^\infty(K,\nu)$ satisfying

$$\text{(i) } |h(x)| = 1, \text{ all } x \in K;$$

$$\text{(ii) } \int_K hm \,\mathrm{d}\nu = 0, \text{ all } m \in M;$$

$$\text{(iii) } \int_K hf \,\mathrm{d}\nu = \|f\|_1.$$

Note that conditions (i) and (iii) imply that $h = \operatorname{sgn} f$ ν-a.e. on $N(f)$.

Given $\varepsilon > 0$, it follows from the fact that ν is a non-atomic positive measure and the absolute continuity of integrals that there exists a $\delta > 0$ such that if $C \subseteq K$ with $\nu(C) < \delta$, then

$$\int_C |f| \, d\nu < \varepsilon.$$

Let $C \subseteq K$, $\nu(C) < \delta$, be such that if $hm \le 0$ on C, then $m = 0$. (Such a C can be found; see Pinkus [1989, p. 22].) Set

$$f_\varepsilon := \begin{cases} f, & \text{off } C \\ 0, & \text{on } C. \end{cases}$$

Then

$$\|f - f_\varepsilon\|_1 = \int_C |f| \,\mathrm{d}\nu < \varepsilon.$$

Since $\operatorname{sgn} f = \operatorname{sgn} f_\varepsilon$ off C, it follows that for the above h, we also have

(iii′) $\int_K hf_\varepsilon \,\mathrm{d}\nu = \|f_\varepsilon\|_1$.

Thus $0 \in P_M(f_\varepsilon)$.

As $h = \operatorname{sgn} f = \operatorname{sgn} f_\varepsilon$ on $N(f_\varepsilon) \subseteq N(f)$, we have from (ii)

$$0 = \int_K hm \,\mathrm{d}\nu = \int_{N(f_\varepsilon)} [\operatorname{sgn} f_\varepsilon] m \,\mathrm{d}\nu + \int_C hm \,\mathrm{d}\nu + \int_{Z(f_\varepsilon)\setminus C} hm \,\mathrm{d}\nu$$

for all $m \in M$. If

$$\left| \int_{N(f_\varepsilon)} [\operatorname{sgn} f_\varepsilon] m \,\mathrm{d}\nu \right| = \int_{Z(f_\varepsilon)} |m| \,\mathrm{d}\nu$$

for some $m \in M$, then we must have

$$\left| \int_C hm \,\mathrm{d}\nu + \int_{Z(f_\varepsilon)\setminus C} hm \,\mathrm{d}\nu \right| = \int_C |m| \,\mathrm{d}\nu + \int_{Z(f_\varepsilon)\setminus C} |m| \,\mathrm{d}\nu$$

implying that hm is of one sign on C. However this in turn implies that $m = 0$. Thus

$$\left| \int_{N(f_\varepsilon)} [\operatorname{sgn} f_\varepsilon] m \,\mathrm{d}\nu \right| < \int_{Z(f_\varepsilon)} |m| \,\mathrm{d}\nu$$

for all $m \in M$, $m \neq 0$, and the zero function is therefore the strongly unique best approximant to f_ε from M. $\square$

In Smarzewski [1988], it is proven, in the general $L^1(K,\nu)$ space without any assumptions on the measure ν, that those elements with a strongly unique best approximant are dense in the set of elements with a unique best approximant. His proof in this case parallels his proof of Theorem 2.5. This proof immediately implies Theorem 5.3. It does not directly imply Theorem 5.4. For ν a non-atomic positive measure, the density in $L^1(K,\nu)$ of the set of functions that have a unique best approximant from M goes back to Havinson, Romanova [1972] and Rozema [1974].

In Theorem 3.1, we proved that strong uniqueness at f implies that the best approximation operator from M is locally Lipschitz continuous at f, i.e., if f has a unique best approximant from M and

$$\|f - m\| - \|f - P_M(f)\| \geq \gamma \|m - P_M(f)\|$$

for all $m \in M$ and some $\gamma > 0$, then for each $g \in X$ and any element of $P_M(g)$, we have

$$\|P_M(f) - P_M(g)\| \leq \frac{2}{\gamma} \|f - g\|.$$

This implies that in the inequality

$$\|P_M(f) - P_M(g)\| \leq \sigma \|f - g\|,$$

we can always take the minimal (optimal) $\sigma(f)$ therein to satisfy

$$\sigma(f) \leq \frac{2}{\gamma(f)}$$

where $\gamma(f)$ is the strong uniqueness constant for f. In Theorem 3.2, we proved that on $C(B)$, assuming M is a finite-dimensional subspace, local Lipschitz continuity at f of the best approximation operator and strong uniqueness at f are equivalent. But other than the above inequality, the relationship between $\sigma(f)$ and $\gamma(f)$ is unclear.

In the case of $L^1(K,\nu)$, we again have that local Lipschitz continuity at f of the best approximation operator from M and strong uniqueness at f are equivalent. Full details may be found in Angelos, Kroó [1986]. We consider here the particular case of where ν is a non-atomic positive measure. In this case, despite the positive density statement of Theorem 5.4, for many f, we do not have strong uniqueness. But when we do have strong uniqueness then we also have

$$\frac{1}{\gamma(f)} \leq \sigma(f).$$

Theorem 5.6. *Let ν be a non-atomic positive measure. Let M be a finite-dimensional subspace of $L^1(K,\nu)$. Then $f \in L^1(K,\nu)$ has a strongly unique best approximant from M if and only if the* *best approximation operator from M is locally Lipschitz continuous at f. Furthermore in this case, we have*

$$\frac{1}{\gamma(f)} \leq \sigma(f) \leq \frac{2}{\gamma(f)}$$

where $\gamma(f)$ and $\sigma(f)$ are as detailed above.

Proof: If f has a strongly unique best approximant from M, then we have from Theorem 3.1 that the best approximation operator from M is locally Lipschitz continuous at f, and $\sigma(f) \leq 2/\gamma(f)$. Assume the converse, i.e., the best approximation operator from M is locally Lipschitz continuous at f. We will prove that f has a strongly unique best approximant from M and $1/\gamma(f) \leq \sigma(f)$.

We assume that $f \notin M$ and the best approximant to f is unique. If either of these assumptions does not hold, then the result follows easily. For ease of exposition, we assume that $0 \in P_M(f)$. Let $m \in M$, $m \neq 0$, and assume

$$\|f - m\|_1 - \|f\|_1 = \delta > 0.$$

Set

$$B := \{x : f(x)(f - m)(x) > 0\}$$

and

$$A := K \backslash (B \cup Z(f)).$$

Since ν is a non-atomic positive measure, Theorem 5.5 is valid. Let $h \in L^\infty(K, \nu)$ be as therein.

Define $g \in L^1(K, \nu)$ by

$$g(x) := \begin{cases} m(x), & x \in A \\ f(x), & x \in B \\ |m(x)|h(x) + m(x), & x \in Z(f). \end{cases}$$

We shall prove that $m \in P_M(g)$ and $\|f - g\|_1 \leq \delta$ from which the results will follow.

We note that as $|h| = 1$ on K then

$$\int_K h(g - m)\,\mathrm{d}\nu = \|g - m\|_1$$

if we can prove that $h(g - m) \geq 0$. Now on A, we have $h(g - m) = 0$. On B, we have $h(g - m) = h(f - m) = |f - m|$, since $h = \operatorname{sgn} f$ on $N(f)$, and $B \subseteq N(f)$ is where $\operatorname{sgn} f = \operatorname{sgn}(f - m)$. On $Z(f)$, we have $h(g - m) = |m|$. Thus $h(g - m) \geq 0$ implying that $m \in P_M(g)$. From Theorem 5.5, it therefore follows that $m \in P_M(g)$.

By the above, we have

$$\begin{aligned}
\delta &= \|f - m\|_1 - \|f\|_1 = \int_K |f - m|\,\mathrm{d}\nu - \int_K |f|\,\mathrm{d}\nu \\
&= \int_{A \cup B} |f - m|\,\mathrm{d}\nu + \int_{Z(f)} |m|\,\mathrm{d}\nu - \int_{A \cup B} |f|\,\mathrm{d}\nu \\
&= \int_{A \cup B} (f - m)\operatorname{sgn}(f - m)\,\mathrm{d}\nu - \int_{A \cup B} (f - m)\operatorname{sgn} f\,\mathrm{d}\nu + \int_{Z(f)} |m|\,\mathrm{d}\nu - \int_{A \cup B} m \operatorname{sgn} f\,\mathrm{d}\nu \\
&= \int_{A \cup B} (f - m)[\operatorname{sgn}(f - m) - \operatorname{sgn} f]\,\mathrm{d}\nu + \int_{Z(f)} |m|\,\mathrm{d}\nu + \int_{Z(f)} hm\,\mathrm{d}\nu,
\end{aligned}$$

since $\int_K hm\,\mathrm{d}\nu = 0$ for all $m \in M$ and $h = \operatorname{sgn} f$ on $A \cup B$. On B, we have $\operatorname{sgn}(f-m) = \operatorname{sgn} f$, while on A, we have $\operatorname{sgn}(f-m) \neq \operatorname{sgn} f$. Thus

$$\delta = 2\int_A |f-m|\,\mathrm{d}\nu + \int_{Z(f)} |m| + hm\,\mathrm{d}\nu.$$

Now

$$\begin{aligned}\|f-g\|_1 &= \int_A |f-m|\,\mathrm{d}\nu + \int_B |f-f|\,\mathrm{d}\nu + \int_{Z(f)} ||m|h+m|\,\mathrm{d}\nu \\ &= \int_A |f-m|\,\mathrm{d}\nu + \int_{Z(f)} |m| + hm\,\mathrm{d}\nu \\ &\leq 2\int_A |f-m|\,\mathrm{d}\nu + \int_{Z(f)} |m| + hm\,\mathrm{d}\nu = \delta.\end{aligned}$$

As the best approximation operator from M is locally Lipschitz continuous at f, then

$$\|P_M(f) - P_M(g)\|_1 \leq \sigma(f)\|f-g\|_1$$

for all $g \in L^1(K,\nu)$. Substituting the above g, where we recall that $0 \in P_M(f)$, $m \in P_M(g)$ and $\|f-g\|_1 \leq \delta = \|f-m\|_1 - \|f\|_1$, we have

$$\|m\|_1 = \|P_M(f) - P_M(g)\|_1 \leq \sigma(f)\|f-g\|_1 \leq \sigma(f)[\|f-m\|_1 - \|f\|_1],$$

i.e.,

$$\|m\|_1 \leq \sigma(f)[\|f-m\|_1 - \|f\|_1]$$

for every $m \in M$. This implies that f has a strongly unique best approximant from M, and

$$\frac{1}{\gamma(f)} \leq \sigma(f). \quad \square$$

We now consider strong uniqueness in the one-sided $L^1(K)$ case. Let K be a compact set in $\mathbb{R}^d$ satisfying $K = \overline{\operatorname{int} K}$. We consider $f \in C(K)$ with norm

$$\|f\|_1 = \int_K |f|\,\mathrm{d}\mu$$

where μ is a non-atomic positive finite measure with the property that every real-valued continuous function is μ-measurable, and such that if $f \in C(K)$ satisfies $\|f\|_1 = 0$ then $f = 0$, i.e., $\|\cdot\|_1$ is truly a norm on $C(K)$. For each $f \in C(K)$, set

$$M(f) := \{m : m \in M, m \leq f\},$$

where M is a finite-dimensional subspace of $C(K)$. We consider the problem

$$\inf_{m\in M(f)} \|f-m\|_1 = \inf_{m\in M(f)} \int_K |f-m|\,\mathrm{d}\mu = \inf_{m\in M(f)} \int_K f-m\,\mathrm{d}\mu,$$

since $f - m \geq 0$ for all $m \in M(f)$. This is equivalent to considering

$$\sup_{m\in M(f)} \int_K m\,\mathrm{d}\mu.$$

We let $P_{M(f)}(f)$ denote the set of one-sided best approximants to f from M. That is,

$$P_{M(f)}(f) := \{m^* : m^* \in M(f), \|f-m^*\|_1 \leq \|f-m\|_1, \text{ all } m \in M(f)\}.$$

The following is one characterization of best one-sided L^1-approximations.

Theorem 5.7. *Let $f \in C(K)$. Assume M is an n-dimensional subspace of $C(K)$ containing a strictly positive function. Then a one-sided best L^1-approximation to f from M exists and $m^* \in P_{M(f)}(f)$ if and only if there exist distinct points $\{x_i\}_{i=1}^k$ in K, $1 \le k \le n$, and positive numbers $\{\lambda_i\}_{i=1}^k$ for which*
(a) $(f - m^*)(x_i) = 0, \quad i = 1, \ldots, k.$
(b) For all $m \in M$

$$\int_K m \, \mathrm{d}\mu = \sum_{i=1}^n \lambda_i m(x_i).$$

We have the following result whose conditions are, unfortunately, difficult to verify and generally rarely hold.

Theorem 5.8. *Let M be an n-dimensional subspace of $C(K)$ and $f \in C(K)$. Assume there exist distinct points $\{x_i\}_{i=1}^k$ in K, and an $m^* \in M(f)$ satisfying*
(a) $(f - m^*)(x_i) = 0, \quad i = 1, \ldots, n.$
(b) If $m \in M$ satisfies $m(x_i) = 0$, $i = 1, \ldots, n$, then $m = 0$.
(c) There exist strictly positive values λ_i, $i = 1, \ldots, n$, such that

$$\int_K m \, \mathrm{d}\mu = \sum_{i=1}^n \lambda_i m(x_i)$$

for all $m \in M$.

In this case, m^* is the unique best approximant to f from $M(f)$ and there exists a $\gamma > 0$ such that

$$\|f - m\|_1 - \|f - m^*\|_1 \ge \gamma \|m - m^*\|_1$$

for all $m \in M(f)$.

Proof: Assume $m \in M(f)$. Then from (c) and (a), we have

$$\int_K m \, \mathrm{d}\mu = \sum_{i=1}^n \lambda_i m(x_i) \le \sum_{i=1}^n \lambda_i f(x_i) = \sum_{i=1}^n \lambda_i m^*(x_i) = \int_K m^* \, \mathrm{d}\mu.$$

If equality holds, then we must have $m(x_i) = f(x_i) = m^*(x_i)$, $i = 1, \ldots, n$. Thus from (b), we have $m = m^*$ and therefore m^* is the unique best approximant to f from $M(f)$.

Let $m_i \in M$ satisfy $m_i(x_j) = \delta_{ij}$, $i, j = 1, \ldots, n$. From (b), it follows that we can construct these m_i. Thus, we can write each $m \in M$ is the form $m = \sum_{i=1}^n m(x_i) m_i$. Set

$$\gamma := \min_{i=1,\ldots,n} \frac{\lambda_i}{\|m_i\|_1} > 0.$$

Thus $\lambda_i \ge \gamma \|m_i\|_1$ for $i = 1, \ldots, n$. Assume $m \in M$ satisfies $m(x_i) \ge 0$, $i = 1, \ldots, n$. Then, using (c), we obtain

$$\int_K m \, \mathrm{d}\mu = \sum_{i=1}^n \lambda_i m(x_i) \ge \gamma \sum_{i=1}^n m(x_i) \|m_i\|_1 \ge \gamma \| \sum_{i=1}^n m(x_i) m_i \|_1 = \gamma \|m\|_1.$$

Now for each $m \in M(f)$, we have

$$\|f-m\|_1 - \|f-m^*\|_1 = \int_K (f-m)-(f-m^*)\,\mathrm{d}\mu = \int_K (m^*-m)\,\mathrm{d}\mu.$$

From (a), and since $m \in M(f)$, we have $(m^*-m)(x_i) = (f-m)(x_i) \ge 0$, $i=1,\ldots,n$. Thus

$$\int_K (m^*-m)\,\mathrm{d}\mu \ge \gamma\|m-m^*\|_1$$

implying

$$\|f-m\|_1 - \|f-m^*\|_1 \ge \gamma\|m-m^*\|_1. \quad \square$$

Both Theorems 5.7 and 5.8 are essentially to be found in Nürnberger [1985]. In Section 9, we discuss non-classical strong uniqueness in the L^1 setting.

6. Strong Uniqueness of Rational Approximation in the Uniform Norm

In this section, we consider strong uniqueness of the best uniform approximant of functions $f \in C[a,b]$ by rational functions of the form

$$R_{m,n} := \{r = p/q : p \in \Pi_m, q \in \Pi_n, q(x) > 0, x \in [a,b]\},$$

where $\Pi_n = \operatorname{span}\{1, x, \ldots, x^n\}$. In contrast with the previous approximation problems, we are here faced with a *nonlinear* set of approximants that makes even the question of existence of a best approximant a nontrivial problem. The general technique, based on bounded compactness of the unit ball in finite-dimensional spaces, is not applicable here. Nevertheless, it turns out that existence and uniqueness of the best rational approximant is still valid for arbitrary $f \in C[a,b]$. However, there is an essential difference from uniform polynomial approximation: the possibility of degeneracy. A rational function $r = p/q \in R_{m,n}$ is said to be *degenerate* if $r = 0$ or $r \in R_{m-1,n-1}$. This degeneracy of rational approximants does not effect the uniqueness of the best approximant, but it does spoil the continuity of the best approximation operator and also the property of strong uniqueness.

To explain all this, let us introduce the following quantity called the *defect* of the irreducible rational function $r = p/q \in R_{m,n}$:

$$d(r) := \begin{cases} \min\{m-\partial p, n-\partial q\}, & r \neq 0; \\ n, & r = 0, \end{cases}$$

where ∂p is the degree of the polynomial p. Clearly, $r = p/q \in R_{m,n}$ is degenerate if and only if $d(r) > 0$. Moreover, $d(r)$ is the greatest number such that $r = p/q \in R_{m-d(r),n-d(r)}$.

First, we are going to address the questions of existence, characterization and uniqueness of the best uniform rational approximant. This next result is due to Walsh [1931].

Theorem 6.1. (Existence of Best Rational Approximation) *Any $f \in C[a,b]$ possesses a best approximant from $R_{m,n}$.*

Proof: Let $r_k = p_k/q_k \in R_{m,n}$ be such that

$$\|f-r_k\| \to E(f) := \inf_{r \in R_{m,n}} ||f-r||, \qquad k \to \infty.$$

Then $\|r_k\| \le M, k \ge 0$, and we can also assume the normalization $\|q_k\| = 1, k \ge 0$, i.e., $\|p_k\| \le M, k \ge 0$. Thus, passing to a subsequence, we may assume that $p_k \to p \in \Pi_m$, $q_k \to q \in \Pi_n$, $k \to \infty$, $q \ge 0$, and $||q|| = 1$. Furthermore, the inequality $|p_k| \le Mq_k$, $k \ge 0$, that clearly holds on $[a,b]$ implies that $|p| \le Mq$, on $[a,b]$. Thus p must vanish at every zero of q in $[a,b]$ to at least the same multiplicity, which yields that $r := p/q \in R_{m,n}$ and

$$r_k(x) \to r(x), \quad x \in [a,b] \setminus Z, \tag{6.1}$$

where Z is the discrete set of zeros of q. Choose now an arbitrary $\varepsilon > 0$. Since $[a,b] \setminus Z$ is dense in $[a,b]$, the continuity of functions involved yields that for some $x \in [a,b] \setminus Z$, we have

$$\|f - r\| \le |f(x) - r(x)| + \varepsilon. \tag{6.2}$$

Since (6.1) holds for this x, we have for k large enough

$$|f(x) - r(x)| \le |f(x) - r_k(x)| + \varepsilon \le ||f - r_k|| + \varepsilon \le E(f) + 2\varepsilon.$$

Combining the two inequalities, we obtain $\|f - r\| \le E(f) + 3\varepsilon$, i.e., $\|f - r\| = E(f)$. Thus r is a best approximant to f. $\square$

The next step will be to derive an analogue of the Alternation Theorem for rational approximation. We start with two auxiliary lemmas.

Lemma 6.2. *Let $r \in R_{m,n}$ be irreducible. Then $\Pi_m + r\Pi_n$ is a Chebyshev (T-) space of dimension $m + n + 1 - d(r)$ on $\mathbb{R}$.*

Proof: Let us first show that the dimension of $\Pi_m + r\Pi_n$ equals $m + n + 1 - d(r)$. If $r = 0$, then $d(r) = n$ and this becomes trivial. So assume $r \ne 0$. Clearly,

$$\dim(\Pi_m + r\Pi_n) = \dim(\Pi_m) + \dim(r\Pi_n) - \dim(\Pi_m \cap r\Pi_n) = m + n + 2 - \dim(\Pi_m \cap r\Pi_n).$$

Furthermore, if $p_1 \in \Pi_m \cap r\Pi_n$ with $r = p/q$ then $p_1 q = pq_1$ for some $q_1 \in \Pi_n$. Since the polynomials p, q are relatively prime, it follows that $p_1 = gp$ and $q_1 = gq$, where g is an arbitrary polynomial of degree at most $\min\{m - \partial p, n - \partial q\} = d(r)$. Thus $\dim(\Pi_m \cap r\Pi_n) = d(r) + 1$ and hence $\dim(\Pi_m + r\Pi_n) = m + n + 1 - d(r)$.

To prove the T-space property, assume some $p_1 + rq_1 \in \Pi_m + r\Pi_n$, $p_1 \in \Pi_m$, $q_1 \in \Pi_n$, $r = p/q$, has $m + n + 1 - d(r)$ distinct zeros on $\mathbb{R}$. Then $p_1 q + pq_1$ has these same zeros. But evidently $\partial(p_1 q + pq_1) \le m + n - d(r)$, implying that $p_1 + rq_1 = 0$. Thus $\Pi_m + r\Pi_n$ is a T-space. $\square$

Lemma 6.3. *Let $f \in C[a,b]$ and assume that $r^* \in R_{m,n}$ is its best approximant, and is irreducible. Then for every $r \in \Pi_m + r^*\Pi_n, r \ne 0$, we have*

$$\max_{x \in A_{f-r^*}} [\operatorname{sgn}(f - r^*)(x)]\, r(x) > 0. \tag{6.3}$$

Proof: Set $r^* =: p^*/q^*$ and consider an arbitrary $r = p/q$, $p \in \Pi_m$, $q \in \Pi_n$, without the assumption that $q > 0$ on the interval $[a,b]$. Then for every $t > 0$, small, we have

$$\frac{p^* - tp}{q^* + tq} \in R_{m,n}$$

and

$$\|f - r^*\| \le \|(f - r^*) + \left(r^* - \frac{p^* - tp}{q^* + tq}\right)\|.$$

Clearly

$$r^* - \frac{p^* - tp}{q^* + tq} = tr_1 + O(t^2)$$

where $r_1 := \frac{p^*q + q^*p}{(q^*)^2} \in \frac{1}{q^*}(\Pi_m + r^*\Pi_n)$. Thus by the last two relations, we have

$$\|f - r^* + tr_1\| - \|f - r^*\| \ge Ct^2$$

for $t > 0$, sufficiently small, and some $C \in \mathbb{R}$. This implies that $\tau_+(f - r^*, r_1) \ge 0$ for every $r_1 \in \frac{1}{q^*}(\Pi_m + r^*\Pi_n)$. Applying Theorem 2.1, we obtain that

$$\max_{x \in A_{f-r^*}} [\operatorname{sgn}(f - r^*)(x)]\, r(x) \ge 0, \qquad r \in \Pi_m + r^*\Pi_n.$$

In particular, the last relation yields by Theorem 2.2 that the zero function is a best approximant to f from $\Pi_m + r^*\Pi_n$. But, by Lemma 6.2, this space is a Chebyshev (T-) space. Thus, in view of Theorems 4.5 and 2.2, the strict inequality (6.3) must hold. □

We can now verify the Alternation Theorem which is the analogue of Theorem 4.4 for rational approximation. It may be found in Achieser [1930].

Theorem 6.4. (Characterization of Best Rational Approximation) *Let $f \in C[a,b]$, and $r^* \in R_{m,n}$ be irreducible. Then the following statements are equivalent:*

(i) r^* is a best approximant to f;

(ii) the error function $f - r^*$ equioscillates on at least $m + n + 2 - d(r^*)$ points in $[a,b]$.

Proof: (i) ⇒ (ii). By Lemma 6.3 and Theorem 2.2, the 0 function is a best approximant to $f - r^*$ from $\Pi_m + r^*\Pi_n$. Since, by Lemma 6.2, this space is a T-space of dimension $m + n + 1 - d(r^*)$, Theorem 4.4 implies that (ii) must hold.

(ii) ⇒ (i). Assume $\|f - r\| < \|f - r^*\|$ for some $r \in R_{m,n}$. Since $f - r^*$ equioscillates on at least $m + n + 2 - d(r^*)$ points in $[a,b]$, it follows that $r - r^*$ has at least $m + n + 1 - d(r^*)$ distinct zeros in $[a,b]$. Setting $r^* =: p^*/q^*$ and $r =: p/q$, we obtain that the polynomial $p^*q - pq^*$ has at least $m + n + 1 - d(r^*)$ distinct zeros in $[a,b]$. But clearly $\partial(p^*q - pq^*) \le \max\{\partial p^* + n, \partial q^* + m\} \le m + n - d(r^*)$, a contradiction. Thus r^* is a best approximant. □

As a byproduct of Theorem 6.4, we obtain that every $f \in C[a,b]$ must have a *unique* best approximant from $R_{m,n}$; see Achieser [1930] and Achieser [1947].

Corollary 6.5. (Uniqueness of Best Rational Approximation) *Every $f \in C[a,b]$ possesses a unique best approximant from $R_{m,n}$.*

Proof: Assume that $r^* = p^*/q^*$ is a best approximant to f and $\|f - r\| = \|f - r^*\|$ for some $r = p/q \in R_{m,n}$. By the same argument as in the proof of (ii) ⇒ (i) in Theorem 6.4, it follows that the polynomial $g := p^*q - pq^*$ of degree $\le m + n - d(r^*)$ has at least $N := m + n + 2 - d(r^*)$ points of weak sign change. That is, for some points $a \le x_1 < \cdots < x_N \le b$, we have $g(x_j)g(x_{j-1}) \le 0$ for every $2 \le j \le N$. Since $\partial g \le N - 2$, the Lagrange Interpolation Formula yields that

$$\sum_{j=1}^{N} g(x_j)\frac{1}{\omega'(x_j)} = 0,$$

where $\omega(x) := (x - x_1)\cdots(x - x_N)$. Since $\omega'(x_j)$ also alternate in sign, we must have $g(x_j) = 0$, $1 \le j \le N$, i.e., $g = 0$, and hence $r^* = r$. □

We now have the necessary prerequisites to verify that strong uniqueness of best rational approximation from $R_{m,n}$ holds if and only if the best approximant is non-degenerate. Moreover, our next theorem shows that the continuity of the best rational approximant operator also fails to hold exactly in the case of degeneracy of the best approximant. This theorem combines results of Maehly, Witzgall [1960], Cheney, Loeb [1964] and Werner [1964] (with regards to continuity of the best approximation operator) and Cheney, Loeb [1964] and Cheney [1966] (with regards to strong uniqueness).

Theorem 6.6. *Let $f \in C[a,b]\backslash R_{m,n}$ and let $r^* = p^*/q^* \in R_{m,n}$ be its best rational approximant. Then the following statements are equivalent:*

(i) $d(r^*) = 0$;

(ii) r^* is a strongly unique best approximant;

(iii) the operator of best rational approximation from $R_{m,n}$ is continuous at f.

Proof: (i) $\Rightarrow$ (ii). For an arbitrary $r \in R_{m,n}$, $r \neq r^*$, set

$$\gamma(r) := \frac{\|f - r\| - \|f - r^*\|}{\|r - r^*\|}.$$

In order to verify strong uniqueness, we need to show that there exists $\gamma > 0$ such that $\gamma(r) \geq \gamma$ for every $r \in R_{m,n}$, $r \neq r^*$. Assume, to the contrary, that there exists a sequence $r_k = p_k/q_k \in R_{m,n}$, $r_k \neq r^*$, such that $\gamma(r_k) \to 0$. We may assume, without loss of generality, that $\|p_k\| + \|q_k\| = 1$, $k \in \mathbb{N}$. Passing, if necessary, to a subsequence, it can also be assumed that for some $p \in \Pi_m$, $q \in \Pi_n$, we have $p_k \to p$, $q_k \to q$, as $k \to \infty$, uniformly on $[a,b]$. In particular, $\|p\| + \|q\| = 1$. Assume, in addition, that $\|p^*\| + \|q^*\| = 1$ (where $r^* = p^*/q^*$).

Consider now an arbitrary $x \in A_{f-r^*}$. Then

$$\begin{aligned} \gamma(r_k)\|r^* - r_k\| &= \|f - r_k\| - \|f - r^*\| \\ &\geq (f - r_k)(x)\,\mathrm{sgn}(f - r^*)(x) - (f - r^*)(x)\,\mathrm{sgn}(f - r^*)(x) \\ &= (r^* - r_k)(x)\,\mathrm{sgn}(f - r^*)(x). \end{aligned}$$

Thus using the fact that $q_k > 0$ and $\|q_k\| \leq 1$, we obtain, for any $x \in A_{f-r^*}$,

$$[\mathrm{sgn}(f - r^*)(x)](q_k r^* - p_k)(x) \leq \gamma(r_k)\|r^* - r_k\| q_k(x) \leq \gamma(r_k)\|r^* - r_k\|. \tag{6.4}$$

Note that since $\gamma(r_k) \to 0$, we must have $\|r_k\| \leq M$, $k \in \mathbb{N}$, for some $M > 0$. (Otherwise $\|r_k\| \to \infty$ for a subsequence of k's that in turn would yield that $\gamma(r_k) \to 1$ for the same subsequence.) This means that the right hand side of (6.4) tends to 0 as $k \to \infty$ yielding

$$[\mathrm{sgn}(f - r^*)(x)](qr^* - p)(x) \leq 0, \qquad x \in A_{f-r^*}.$$

In view of Lemma 6.3, this implies that $qr^* = p$, i.e., $p \in \Pi_m \cap r^*\Pi_n$. On the other hand, using the fact that $d(r^*) = 0$, we have $\dim(\Pi_m \cap r^*\Pi_n) = d(r^*) + 1 = 1$ (see the proof of Lemma 6.2 for details). Since evidently p^* is also an element of $\Pi_m \cap r^*\Pi_n$, we must have $p = cp^*$ for some nonzero constant c. But then in view of $qr^* = p$, the relation $q = cq^*$ also holds. Recalling the conditions $\|p\| + \|q\| = \|p^*\| + \|q^*\| = 1$, and $q, q^* > 0$, we clearly have $c = 1$, i.e., $p = p^*$, $q = q^*$. Since $q^* > 0$ on $[a,b]$, it follows that for some $\delta > 0$ the relation $q_k(x) \geq \delta$ holds for all $x \in [a,b]$, for k sufficiently large.

Furthermore, since inequality (6.3) holds for every r in the finite-dimensional linear space $\Pi_m + r^*\Pi_n$, and because the left hand side of (6.3) is a continuous function of r, we obtain that there exists an $\eta > 0$ such that

$$\max_{x\in A_{f-r^*}} [\mathrm{sgn}(f-r^*)(x)]\, r(x) > \eta\|r\|, \qquad r \in \Pi_m + r^*\Pi_n.$$

Thus, in particular, there exists an $x_k^* \in A_{f-r^*}$ for which

$$[\mathrm{sgn}(f-r^*)(x_k^*)]\,(q_k r^* - p_k)(x_k^*) \ge \eta\|q_k r^* - p_k\|.$$

Using this last inequality together with (6.4) applied with $x = x_k^*$, we arrive at

$$\eta\|q_k r^* - p_k\| \le \gamma(r_k)\|r^* - r_k\| = \gamma(r_k)\|\frac{1}{q_k}(q_k r^* - p_k)\| \le \frac{\gamma(r_k)}{\delta}\|q_k r^* - p_k\|.$$

Since $r_k \ne r^*$, we can divide both sides of this inequality by $\|q_k r^* - p_k\|$ yielding $\eta \le \frac{\gamma(r_k)}{\delta}$. But this contradicts the assumption that $\gamma(r_k) \to 0$. This completes the proof of strong uniqueness.

(ii) $\Rightarrow$ (iii). This implication is trivial since strong uniqueness implies Lipschitz continuity of the best approximation operator; see Theorem 3.1.

(iii) $\Rightarrow$ (i). Assume that to the contrary $d(r^*) > 0$. We separate the proof into two cases.

Case 1. $f - r^*$ has less than $m+n+2$ points of equioscillation. In this case, $f - r^*$ has $m+n+2-d$ points of equioscillation for some $0 < d \le d(r^*)$. Evidently, for any $\varepsilon > 0$, there exists an $r_1 \in R_{m,n}$ such that $\|r^* - r_1\| \le \varepsilon$ and $d(r_1) = 0$. Set $g := f + (r_1 - r^*)$, and denote by $r_2 \in R_{m,n}$ the best approximant to g. Clearly, $r_2 \ne r_1$ since otherwise the error function $g - r_2$ would have less than $m+n+2$ points of equioscillation, in contradiction to the fact that $d(r_1) = 0$ and Theorem 6.4. Thus by the uniqueness of best rational approximation (Corollary 6.5),

$$\|g - r_2\| < \|g - r_1\| = \|f - r^*\|. \tag{6.5}$$

Now, the function $g - r_1 = f - r^*$ has $m+n+2-d$ points of equioscillation, which by (6.5) yields that $r_1 - r_2$ must have at least $m+n+1-d$ zeros in $[a,b]$. Set $r_1 =: p_1/q_1$, $r_2 =: p_2/q_2$. By the above observation that the nonzero polynomial $p_1q_2 - p_2q_1$ has at least $m+n+1-d$ zeros, it follows that $\partial(p_1q_2 - p_2q_1) \ge m+n+1-d$. On the other hand,

$$\partial(p_1q_2 - p_2q_1) \le \max\{m + \partial q_2, n + \partial p_2\} \le m+n-d(r_2).$$

Thus, we obtain $d(r_2) \le d-1$. This implies that $g - r_2$ must have at least $m+n+2-d(r_2) \ge m+n+3-d$ points of equioscillation. Moreover, $\|f-g\| = \|r^* - r_1\| \le \varepsilon$ and hence in view of the continuity of the best approximation operator at f, we have that $g - r_2$ tends uniformly to $f - r^*$ as $\varepsilon \to 0$. Recall that $f - r^*$ has exactly $m+n+2-d$ points of equioscillation, while $g - r_2$ has at least $m+n+3-d$ points of equioscillation. This leads to a contradiction as, in the limit, we cannot lose points of equioscillation.

Case 2. We now assume that $f - r^*$ possesses at least $m+n+2$ points of equioscillation. We also, for simplicity, set $[a,b] = [0,1]$. We first consider the subcase $|(f-r^*)(0)| < \|f - r^*\|$. There then exist $0 < \alpha < 1$ and $\delta > 0$ such that

$$|(f-r^*)(x)| < \alpha\|f - r^*\|, \qquad 0 \le x \le \delta. \tag{6.6}$$

Note that all points of equioscillation of the error curve $f - r^*$ must be in $[\delta, 1]$. Set

$$A := (1-\alpha)\frac{\|f-r^*\|}{\|1/q^*\|}, \quad r_k := \frac{A}{q^*(1+kx)}, \quad r_k^{**} := r^* - r_k.$$

Since r^* is degenerate, it is easy to see that $r_k, r_k^{**} \in R_{m,n}$. Consider now the functions

$$g_k := \frac{A}{q^*}\min\{\frac{1}{1+kx}, \frac{1}{1+k\delta}\}, \quad f_k := f - g_k.$$

Evidently, $g_k = r_k$ if $x \geq \delta$. Furthermore, whenever $0 \leq x \leq \delta$, we have by the choice of A

$$|(r_k - g_k)(x)| = \frac{Ak(\delta - x)}{q^*(1+kx)(1+k\delta)} \leq \frac{Ak\delta}{q^*(1+k\delta)} \leq (1-\alpha)\frac{k\delta\|f-r^*\|}{1+k\delta} \leq (1-\alpha)\|f-r^*\|.$$

Using the last inequality together with (6.6), we obtain for $0 \leq x \leq \delta$

$$|(f_k - r_k^{**})(x)| = |(f-r^*)(x) + (r_k - g_k)(x)| < \alpha\|f-r^*\| + (1-\alpha)\|f-r^*\| = \|f-r^*\|.$$

Recalling that $f_k - r_k^{**} = f - r^*$ whenever $x \geq \delta$ and all points of equioscillation of the error curve $f - r^*$ are in $[\delta, 1]$, it follows that the error curve $f_k - r_k^{**}$ equioscillates at least $m+n+2$ times, i.e., by Theorem 6.4, r_k^{**} is the best approximant to f_k. On the other hand, we clearly have

$$\|f_k - f\| = \|g_k\| \leq \frac{A\|1/q^*\|}{1+k\delta} \to 0, \qquad k \to \infty,$$

while for every k

$$\|r_k^{**} - r^*\| = \|r_k\| \geq \frac{A}{q^*(0)}.$$

Thus the continuity of the operator of best rational approximation fails to hold at f. This verifies the needed statement in the case $|(f-r^*)(0)| < \|f-r^*\|$.

Finally, if $|(f-r^*)(0)| = \|f-r^*\|$, we can introduce a small perturbation to the function f resulting in $f_\varepsilon \in C[0,1]$ such that $\|f - f_\varepsilon\| < \varepsilon$, r^* is still the best approximant to f_ε and $|(f_\varepsilon - r^*)(0)| < \|f_\varepsilon - r^*\|$. By the above argument, the continuity of best rational approximation fails uniformly at f_ε by at least $A/q^*(0)$, and hence it will fail at f as well. □

The ideas in this section have been further generalized to other non-linear families, cf., Barrar, Loeb [1970], Braess [1973], and Braess [1986].

Part II. Non-Classical Strong Uniqueness

7. Uniformly Convex Space

Let X be a uniformly convex space (of dimension at least 2) with norm $\|\cdot\|$. In such a case, we know that we always have uniqueness of the best approximant. But we rarely have strong uniqueness in the classical sense. As such, we look for a different non-classical form of strong uniqueness. This we do in the present context using the modulus of convexity. The *modulus of convexity* on X is defined as follows. For $\varepsilon \in (0, 2]$, we set

$$\delta(\varepsilon) := \inf\{1 - \frac{\|f+g\|}{2} : \|f\|, \|g\| = 1, \|f-g\| \geq \varepsilon\}.$$

We recall that X is uniformly convex if $\delta(\varepsilon) > 0$ for every $\varepsilon > 0$. If X is a uniformly convex space, then the following estimate is valid; see Björnestål [1979].

Theorem 7.1. *Let M be a linear subspace of X. Given $f \in X$ assume that m^* is the best approximant to f from M. Then for all $m \in M$, $m \neq m^*$, we have*

$$\|f-m\| - \|f-m^*\| \geq \|f-m\|\,\delta\left(\frac{\|m-m^*\|}{\|f-m\|}\right).$$

Remark. Note that as $\|m-m^*\| \leq \|f-m\| + \|f-m^*\|$, it follows that

$$\frac{\|m-m^*\|}{\|f-m\|} \leq \frac{\|f-m\| + \|f-m^*\|}{\|f-m\|} = 1 + \frac{\|f-m^*\|}{\|f-m\|} < 2$$

since $\|f-m^*\| < \|f-m\|$. Thus $\delta(\frac{\|m-m^*\|}{\|f-m\|})$ is well-defined.

Remark. δ is a non-decreasing function. If

$$\|f-m\| - \|f-m^*\| \leq \sigma,$$

then, since $\|f-m\| \geq \|f-m^*\|$, we obtain

$$\|f-m\| - \|f-m^*\| \geq \|f-m^*\|\,\delta\left(\frac{\|m-m^*\|}{\|f-m^*\|+\sigma}\right),$$

which is a non-classical strong uniqueness inequality.

Proof: As M is a linear subspace, the above claim is equivalent to verifying

$$\|h-m\| - \|h\| \geq \|h-m\|\,\delta\left(\frac{\|m\|}{\|h-m\|}\right)$$

for any h such that 0 is a best approximant to h from M.

It easily follows from the definition of δ that for any $f, g \in X$ satisfying $\|f\|, \|g\| \leq r$, $r > 0$, we have

$$\left\|\frac{f+g}{2}\right\| \leq r\left[1 - \delta\left(\frac{\|f-g\|}{r}\right)\right].$$

Let, in the above, $f = h$ and $g = h - m$, and $r = \|h-m\| \geq \|h\|$. Substituting, we therefore obtain

$$\left\|\frac{2h-m}{2}\right\| \leq \|h-m\|\left[1 - \delta\left(\frac{\|m\|}{\|h-m\|}\right)\right].$$

Note that as 0 is a best approximant to h from M, it follows that

$$\left\|\frac{2h-m}{2}\right\| = \left\|h - \frac{m}{2}\right\| \geq \|h\|.$$

Thus

$$\|h\| \leq \|h-m\|\left[1 - \delta\left(\frac{\|m\|}{\|h-m\|}\right)\right]$$

that immediately translates into the desired inequality. $\square$

For a similar result, see Wegmann [1975], and also Lin [1989]. From Theorem 7.1, we may obtain the following result that was proven in Smarzewski [1986], Prus, Smarzewski [1987], Smarzewski [1987] by different methods.

Corollary 7.2. *Let M be linear subspace of X. Assume*

$$\delta(\varepsilon) \geq C\varepsilon^q$$

for $\varepsilon \in (0,2]$ and $q \geq 1$. Then for m^*, the best approximant to f from M, we have

$$\|f-m\|^q - \|f-m^*\|^q \geq C\|m-m^*\|^q,$$

for all $m \in M$, $m \neq m^*$. Furthermore, if

$$\|f-m\| - \|f-m^*\| \leq \sigma$$

then we also have

$$\|f-m\| - \|f-m^*\| \geq Ck\|m-m^*\|^q,$$

where $k := \|f-m^*\|/(\|f-m^*\|+\sigma)^q$.

Proof: The first inequality follows immediately from Theorem 7.1. If $q = 1$, there is nothing to prove. For $q > 1$, we have

$$\|f-m\| - \|f-m^*\| \geq \|f-m\|\delta\left(\frac{\|m-m^*\|}{\|f-m\|}\right) \geq \|f-m\|C\left(\frac{\|m-m^*\|}{\|f-m\|}\right)^q.$$

Multiply through by $\|f-m\|^{q-1}$ and use the fact that $\|f-m^*\| \leq \|f-m\|$.

The second inequality follows from the second remark after the statement of Theorem 7.1. □

Remark. Moduli of convexity satisfying

$$\delta(\varepsilon) \geq C\varepsilon^q$$

are called moduli of convexity of power type q. It is known that for any modulus of convexity, we can only have the above holding for $q \geq 2$. (The condition $q \geq 1$ in the above corollary is superfluous.)

It is also known, see Hanner [1956], that if $\delta_p(\varepsilon)$ denotes the modulus of convexity of L^p or ℓ_p, then we have

$$\delta_p(\varepsilon) = \begin{cases} \frac{(p-1)\varepsilon^2}{8} + o(\varepsilon^2), & 1 < p < 2, \\ \frac{\varepsilon^p}{p2^p} + o(\varepsilon^p), & 2 \leq p < \infty. \end{cases}$$

In addition, in Hilbert spaces, we always have $\delta(\varepsilon) = 1-(1-\varepsilon^2/4)^{1/2} = \varepsilon^2/8 + O(\varepsilon^4)$ and Hilbert spaces are the most convex in that $\delta(\varepsilon) \leq 1-(1-\varepsilon^2/4)^{1/2}$ holds for any uniformly convex Banach space. For L_p, we in fact have

$$\delta_p(\varepsilon) \geq c_p\varepsilon^q$$

for all $\varepsilon \in [0,2]$ where $q = \max\{2,p\}$ and

$$c_p := \begin{cases} \frac{p-1}{8}, & 1 < p < 2, \\ \frac{1}{p2^p}, & 2 \leq p < \infty. \end{cases}$$

The following result can be found in Smarzewski [1987]. Assume ϕ is an increasing convex function defined on $\mathbb{R}_+$ satisfying $\phi(0) = 0$. Assume in addition that

$$\delta(\varepsilon) \geq K\phi(\varepsilon)$$

for $\varepsilon \in [0, 2]$, and ϕ is submultiplicative in that there exists a constant $L > 0$ such that for all $t, s \in \mathbb{R}_+$, we have

$$\phi(ts) \leq L\phi(t)\phi(s).$$

Then the following holds.

Proposition 7.3. *Let M be a linear subspace of the uniformly convex space X. Assume ϕ is as above. If m^* is the best approximant to f from M then*

$$\phi(\|f - m\|) - \phi(\|f - m^*\|) \geq KL^{-1}\phi(\|m - m^*\|).$$

for all $m \in M$, $m \neq m^*$.

Proof: Since ϕ is convex and $\phi(0) = 0$, we have for any $\lambda \in [0, 1]$ that

$$\phi(\lambda t) = \phi(\lambda t + (1 - \lambda)0) \leq \lambda\phi(t).$$

Thus

$$\phi(\|f - m^*\|) = \phi\left(\frac{\|f - m^*\|}{\|f - m\|}\|f - m\|\right) \leq \frac{\|f - m^*\|}{\|f - m\|}\phi(\|f - m\|).$$

Therefore, applying Theorem 7.1,

$$\begin{aligned}
\phi(\|f - m\|) - \phi(\|f - m^*\|) &\geq \phi(\|f - m\|) - \frac{\|f - m^*\|}{\|f - m\|}\phi(\|f - m\|) \\
&= \frac{\|f - m\| - \|f - m^*\|}{\|f - m\|}\phi(\|f - m\|) \\
&\geq \delta\left(\frac{\|m - m^*\|}{\|f - m\|}\right)\phi(\|f - m\|) \\
&\geq K\phi\left(\frac{\|m - m^*\|}{\|f - m\|}\right)\phi(\|f - m\|) \\
&\geq KL^{-1}\phi(\|m - m^*\|). \quad \square
\end{aligned}$$

Corollary 7.2 easily follows from Proposition 7.3 by choosing $\phi(t) = t^q$ (where $L = 1$). A different approach is the following which may also be found in Smarzewski [1987].

Theorem 7.4. *Assume there exists a positive constant K and an increasing nonnegative function ϕ defined on $\mathbb{R}_+$ such that*

$$\phi\left(\left\|\frac{g + h}{2}\right\|\right) \leq \frac{1}{2}[\phi(\|g\|) + \phi(\|h\|)] - K\phi(\|g - h\|)$$

holds for all $f, g \in X$. Let M be a linear subspace of X. If m^* is the best approximant to f from M, then

$$\phi(\|f - m\|) - \phi(\|f - m^*\|) \geq 2K\phi(\|m - m^*\|)$$

for all $m \in M$, $m \neq m^*$.

Proof: As m^* is the best approximant to f from M, we have that

$$\left\| \frac{f - m + f - m^*}{2} \right\| = \left\| f - \frac{m + m^*}{2} \right\| \geq \|f - m^*\|.$$

Set $g := f - m^*$ and $h = f - m$. Then, as ϕ is increasing,

$$\phi(\|f - m^*\|) \leq \phi\left(\left\| \frac{f - m + f - m^*}{2} \right\| \right) \leq \frac{1}{2}[\phi(\|f - m^*\|) + \phi(\|f - m\|)] - K\phi(\|m - m^*\|).$$

The desired inequality follows. □

For $p \geq 2$, it is known that for $g, h \in L^p$, we have

$$\|g + h\|_p^p + \|g - h\|_p^p \leq 2^{p-1}(\|g\|_p^p + \|h\|_p^p).$$

This inequality is called Clarkson's Inequality; see Clarkson [1936]. Thus, it follows that we can apply Theorem 7.4 with $\phi(t) = t^p$ and $K = 2^{-p}$ giving

$$\|f - m\|_p^p - \|f - m^*\|_p^p \geq 2^{1-p}\|m - m^*\|_p^p.$$

This improves somewhat the constant $2^{-p}/p$ obtained as a consequence of Corollary 7.2. For a better constant, see Smarzewski [1986]. For $1 < p < 2$, the inequality

$$\left\| \frac{g + h}{2} \right\|_p^2 \leq \frac{1}{2}[\|g\|_p^2 + \|h\|_p^2] - \frac{p(p-1)}{8}\|g - h\|_p^2$$

holds for all $f, g \in X$. This is called Meir's Inequality; see Meir [1984]. Thus, we can take $\phi(t) = t^2$ and $K = p(p-1)/8$ giving

$$\|f - m\|_p^2 - \|f - m^*\|_p^2 \geq \frac{p(p-1)}{4}\|m - m^*\|_p^2.$$

This improves somewhat the constant $(p-1)/8$ obtained as a consequence of Corollary 7.2.

For $p = 2$, both these improved constants are $1/2$. However this is not the optimal constant, as in any Hilbert space, we have

$$\|f - m\|^2 - \|f - m^*\|^2 = \|m - m^*\|^2$$

for all $m \in M$ assuming m^* is a best approximant to f from the linear subspace M.

For other results in this direction, see Angelos, Egger [1984] and Egger, Taylor [1989].

8. The Uniform Norm Revisited

In this section, we return to the uniform norm over an interval $[a, b]$. It transpires that if we restrict our approximation to only "smooth" functions, then a weaker condition than the Haar condition is both necessary and sufficient to insure uniqueness of the best approximant to each "smooth" function if the finite-dimensional subspace M is itself "smooth". The following result, from Garkavi [1959], see also Kroó [1984], characterizes unicity spaces with respect to $C^r[a, b]$ for $r \geq 1$. Note that this characterization is independent of r. The characterization critically uses the fact (independent of $r \geq 1$) that if $x \in A_{f-m^*} \cap (a, b)$ and $f - m^* \in C^r[a, b]$ then we must have $(f - m^*)'(x) = 0$.

Theorem 8.1. *Let r be a fixed positive integer, and let M be an n-dimensional subspace of $C^r[a, b]$. A necessary and sufficient condition for the uniqueness of the best approximant from M to each $f \in C^r[a, b]$ is that the following does* **not** *hold:*
There exist k distinct points $\{x_i\}_{i=1}^k$ in $[a, b]$, $1 \leq k \leq n + 1$, nonzero values $\{\sigma_i\}_{i=1}^k$, and an $\widetilde{m} \in M\backslash\{0\}$ such that
(a) $\sum_{i=1}^k \sigma_i m(x_i) = 0$ all $m \in M$.
(b) $\widetilde{m}(x_i) = 0$, $i = 1, \ldots, k$.
(c) $\widetilde{m}'(x_i) = 0$ if $x_i \in (a, b)$.

In Section 4, we defined a Chebyshev (T-) system as a system of functions $m_1, \ldots, m_n$ defined on $C[a, b]$ such that no nontrivial $m \in M = \text{span}\{m_1, \ldots, m_n\}$ vanishes at more than $n-1$ distinct points of $[a, b]$. We often use the term T-system interchangeably for both $m_1, \ldots, m_n$ and for the space M. We say that $M = \text{span}\{m_1, \ldots, m_n\}$ is an ET_2-system if $M \subset C^1[a, b]$ and no nontrivial $m \in M$ vanishes at more than $n - 1$ distinct points of $[a, b]$, where we count multiplicities up to order 2, i.e., we count x as a double zero if $m(x) = m'(x) = 0$. From simple zero counting, we obtain the following result, that is essentially in Garkavi [1959].

Proposition 8.2. *Let r be a fixed positive integer. Let M be an n-dimensional subspace of $C^r[a, b]$. Assume M contains a T-system of dimension m and is contained in an ET_2-system of dimension ℓ, with $m \geq \ell/2$. Then to each $f \in C^r[a, b]$, there exists a unique best approximant from M.*

Based on the above Theorem 8.1, it is natural to ask whether strong uniqueness always holds in this setting. We first give an example to show that classical strong uniqueness need not hold. We then show that a non-classical strong uniqueness with $\phi(t) = t^2$ holds in this case.

Example. Let $M = \text{span}\{1, x, x^3\}$ in $C[-1, 1]$. Since M contains a T-system of dimension 2 ($\text{span}\{1, x\}$), and is contained in an ET_2-system of dimension 4 ($\text{span}\{1, x, x^2, x^3\}$), it follows from Proposition 8.2 that M is necessarily a unicity space with respect to $C^1[-1, 1]$. Set $f(x) := 2x^2 - 1$. As is easily checked, its best approximant from M is the zero function. Since $A_f = \{-1, 0, 1\}$, we have

$$\gamma(f) = \min_{\substack{m \in M \\ \|m\|=1}} \max\{m(-1), -m(0), m(1)\}.$$

Choosing $m(x) = x - x^3 \in M$, we see that $\gamma(f) = 0$, and therefore classical strong uniqueness does not hold.

Let us consider the above example in further detail. Set $m_a(x) := a(x - x^3)$. For a sufficiently small, we have

$$\|f - m_a\| \geq \|f\| + \gamma\|m_a\|^p$$

for some $\gamma > 0$ if and only if $p \geq 2$. To see this, note that $\|f\| = 1$ and $\|m_a\| = 2|a|/3\sqrt{3}$. Thus the right-hand side of the above is of the form

$$1 + \gamma \left(\frac{2|a|}{3\sqrt{3}} \right)^p.$$

Set

$$g(a) := \|f - m_a\| = \max_{x \in [-1,1]} |(2x^2 - 1) - a(x - x^3)|.$$

For $|a|$ sufficiently small, we have

$$g(a) = \frac{1}{3} + \frac{2}{9}\sqrt{4 + 3a^2} + \frac{8}{27a^2}\left(\sqrt{4 + 3a^2} - 2\right).$$

Expanding $\sqrt{4 + 3a^2}$ in a Taylor series about $a = 0$, we obtain

$$g(a) = 1 + \frac{1}{8}a^2 - ca^4 + \cdots$$

for some $c > 0$. Thus there exists a $\gamma > 0$ such that

$$g(a) \geq 1 + \gamma \left(\frac{2|a|}{3\sqrt{3}} \right)^p$$

for a sufficiently small if and only if $p \geq 2$.

Let $M \subset C^r[a, b]$ be an n-dimensional unicity space with respect to $C^r[a, b]$. The previous example showed that we cannot hope to obtain a classical strong uniqueness theorem. We shall, however, prove the following result due to Kroó [1983a].

Theorem 8.3. *Let r be a positive integer, $r \geq 2$. Let $M \subset C^r[a, b]$ be a finite-dimensional unicity space with respect to $C^r[a, b]$, and let $f \in C^r[a, b]$. Given any positive constant σ, there exists a positive constant γ, dependent on f, M and σ (but independent of specific $m \in M$) such that if m^* is the best approximant to f from M, and if $m \in M$ satisfies*

$$\|f - m\| - \|f - m^*\| \leq \sigma, \tag{8.1}$$

then

$$\|f - m\| - \|f - m^*\| \geq \gamma \|m - m^*\|^2. \tag{8.2}$$

Remark. Theorem 8.3 is a simpler case of a general result in Kroó [1983a]. This more general result also covers the case where $M \subset C^1[a, b]$ is an n-dimensional unicity space with respect to $C^1[a, b]$. The proof thereof is somewhat more cumbersome as the modulus of continuity of f' and m' enter both the statement of the result and the analysis, i.e., a non-classical strong uniqueness result is proven with a more complicated ϕ. We choose to deal with the simpler case as presented here.

Prior to proving this theorem, let us prove an ancillary result concerning local strong uniqueness.

Proposition 8.4. *Let ϕ be any nonnegative increasing function defined on $[0,\infty)$. Assume there exists a $\sigma_1 > 0$ such that if $m \in M$ satisfies*

$$\|f-m\| - \|f-m^*\| \le \sigma_1$$

then

$$\|f-m\| - \|f-m^*\| \ge \gamma_1\phi(\|m-m^*\|)$$

for some $\gamma_1 > 0$. Then given any $\sigma_2 > 0$, there exists a $\gamma_2 > 0$, dependent on γ_1, σ_1, σ_2, ϕ and $\|f-m^*\|$, such that, for every $m \in M$ satisfying

$$\|f-m\| - \|f-m^*\| \le \sigma_2,$$

we have

$$\|f-m\| - \|f-m^*\| \ge \gamma_2\phi(\|m-m^*\|).$$

Proof: We need only consider the case where $\sigma_2 > \sigma_1$. For all $m \in M$ satisfying

$$\|f-m\| - \|f-m^*\| \le \sigma_2,$$

we have

$$\|m-m^*\| \le \|f-m\| + \|f-m^*\| \le 2\|f-m^*\| + \sigma_2.$$

Thus for all such m, we have

$$\phi(\|m-m^*\|) \le \phi(2\|f-m^*\| + \sigma_2).$$

Choose $\gamma^* > 0$ for which

$$\gamma^*\phi(2\|f-m^*\| + \sigma_2) \le \sigma_1$$

and set

$$\gamma_2 = \min\{\gamma_1, \gamma^*\}.$$

Then for $m \in M$ satisfying

$$\|f-m\| - \|f-m^*\| \le \sigma_1,$$

we have

$$\|f-m\| - \|f-m^*\| \ge \gamma_1\phi(\|m-m^*\|) \ge \gamma_2\phi(\|m-m^*\|).$$

For $m \in M$ satisfying

$$\sigma_1 \le \|f-m\| - \|f-m^*\| \le \sigma_2,$$

we have

$$\begin{aligned}\gamma_2\phi(\|m-m^*\|) &\le \gamma^*\phi(\|m-m^*\|) \le \gamma^*\phi(2\|f-m^*\| + \sigma_2)\\ &\le \sigma_1 \le \|f-m\| - \|f-m^*\|. \quad \square\end{aligned}$$

Proof of Theorem 8.3: To prove the theorem, we show that if

$$\|f-m\| - \|f-m^*\| = \delta^2 \tag{8.3}$$

then

$$\sqrt{\gamma}\,\|m - m^*\| \le \delta. \tag{8.4}$$

We shall choose a sufficiently small $\delta > 0$ depending on f and M. This proves the theorem for $\sigma = \delta^2$. We then apply Proposition 8.4 to get the full result (with a different γ).

Recall from Theorem 2.3 that the function $m^* \in M$ is a best approximant to f from the n-dimensional subspace M if and only if there exist k distinct points $x_1, \ldots, x_k \in A_{f-m^*}$, for $1 \le k \le n+1$, and strictly positive values $\lambda_1, \ldots, \lambda_k$ such that

$$\sum_{i=1}^{k} \lambda_i [\mathrm{sgn}(f - m^*)(x_i)] m(x_i) = 0$$

for all $m \in M$.

Furthermore, from Theorem 8.1, the fact that M is a unicity space with respect to $C^r[a,b]$ implies that for the $\{x_i\}_{i=1}^k$ as above, if $m \in M$ satisfies $m(x_i) = 0$, $i = 1, \ldots, k$, and $m'(x_i) = 0$ if $x_i \in (a,b)$, then m is necessarily identically zero. Thus the unicity space property says that the functional $|||\cdot|||$ defined on M by

$$|||m||| = \max_{i=1,\ldots,k} |m(x_i)| + \max_{\substack{i=1,\ldots,k \\ x_i \in (a,b)}} |m'(x_i)|$$

is a norm on M. Since M is finite-dimensional, this norm is equivalent to the uniform norm. Thus, there exists a constant c_1 such that

$$\|m - m^*\| \le c_1 |||m - m^*|||$$

for all $m \in M$. Here and in what follows, the c_j will always denote positive constants that depend only upon f and M. It therefore suffices, in proving (8.4), to prove that, under conditions (8.3), we have

$$|(m - m^*)(x_i)| \le c_2 \delta, \qquad i = 1, \ldots, k, \tag{8.5}$$

and

$$|(m - m^*)'(x_i)| \le c_3 \delta, \qquad \text{if } x_i \in (a,b), \quad i = 1, \ldots, k. \tag{8.6}$$

We assume that (8.3) holds and will prove (8.5) and (8.6). We prove (8.5) and (8.6) in two distinct lemmas. In the first lemma, we prove more than what is formally needed to verify (8.5), but apply this better estimate in the second lemma to prove (8.6).

Lemma 8.5. *Under the above assumptions, (8.5) holds. In fact,* $|(m-m^*)(x_i)| \le c_4 \delta^2$, $i = 1, \ldots, k$.

Proof: We assume $m \in M$. Since $x_i \in A_{f-m^*}$, we have

$$\begin{aligned}
\|f - m\|^2 &\ge |(f - m)(x_i)|^2 = |(f - m^*)(x_i) - (m - m^*)(x_i)|^2 \\
&= \|f - m^*\|^2 - 2\|f - m^*\|\,\mathrm{sgn}[(f - m^*)(x_i)](m - m^*)(x_i) + |(m - m^*)(x_i)|^2.
\end{aligned}$$

From (8.3), we have

$$\|f - m\| - \|f - m^*\| = \delta^2,$$

and therefore, assuming $\delta < 1$,

$$\|f - m\|^2 = (\|f - m^*\| + \delta^2)^2 \le \|f - m^*\|^2 + c_5 \delta^2.$$

Thus
$$-2\|f-m^*\|\,\mathrm{sgn}[(f-m^*)(x_i)](m-m^*)(x_i)+|(m-m^*)(x_i)|^2\le c_5\delta^2$$
implying that
$$\mathrm{sgn}[(f-m^*)(x_i)](m-m^*)(x_i)\ge -c_6\delta^2,\qquad i=1,\ldots,k.$$
Recall that
$$\sum_{i=1}^k \lambda_i[\mathrm{sgn}(f-m^*)(x_i)]m(x_i)=0$$
for all $m\in M$, and normalize the λ_i so that $\sum_{i=1}^k\lambda_i=1$. The λ_i are independent of any specific m. Thus
$$\begin{aligned}-c_6\delta^2 \;\le\; \mathrm{sgn}[(f-m^*)(x_i)](m-m^*)(x_i) \;&=\; -\frac{1}{\lambda_i}\sum_{\substack{j=1\\ j\ne i}}^k \lambda_j[\mathrm{sgn}(f-m^*)(x_j)](m-m^*)(x_j)\\ &=\; \frac{1}{\lambda_i}\sum_{\substack{j=1\\ j\ne i}}^k \lambda_j\left(-[\mathrm{sgn}(f-m^*)(x_j)](m-m^*)(x_j)\right) \;\le\; c_7\delta^2.\end{aligned}$$
We have shown that
$$-c_6\delta^2\le[\mathrm{sgn}(f-m^*)(x_i)](m-m^*)(x_i)\le c_7\delta^2,$$
implying that
$$|(m-m^*)(x_i)|\le c_4\delta^2,\qquad i=1,\ldots,k.\qquad\square$$

Lemma 8.6. *Under the above assumptions, (8.6) holds, i.e.,* $|(m-m^*)'(x_i)|\le c_3\delta$, *if* $x_i\in(a,b)$, $i=1,\ldots,k$.

Proof: Assume $x_i\in(a,b)$. From Lemma 8.5,
$$|(m-m^*)(x_i)|\le c_4\delta^2,$$
where
$$\|f-m\|-\|f-m^*\|=\delta^2.$$

Now for $x_i\in(a,b)$ and $\delta>0$ such that $x_i\pm\delta\in[a,b]$,
$$\begin{aligned}[\mathrm{sgn}(f-m^*)(x_i)](m-m^*)(x_i\pm\delta) &= [\mathrm{sgn}(f-m^*)(x_i)]\left((f-m^*)(x_i\pm\delta)-(f-m)(x_i\pm\delta)\right)\\ &\ge [\mathrm{sgn}(f-m^*)(x_i)](f-m^*)(x_i\pm\delta)-\|f-m\|\\ &= [\mathrm{sgn}(f-m^*)(x_i)](f-m^*)(x_i\pm\delta)-\|f-m^*\|-\delta^2\\ &= [\mathrm{sgn}(f-m^*)(x_i)]\left((f-m^*)(x_i\pm\delta)-(f-m^*)(x_i)\right)-\delta^2.\end{aligned}$$

Set
$$h(y):=(f-m^*)(x_i+y)-(f-m^*)(x_i).$$
As $x_i\in A_{f-m^*}\cap(a,b)$, we have $(f-m^*)'(x_i)=0$. Thus we see that $h(0)=h'(0)=0$. Since $h\in C^r$, $r\ge 2$, in a neighborhood of the origin, it therefore follows that $|h(y)|\le c_8y^2$ for some constant c_8. Thus
$$[\mathrm{sgn}(f-m^*)(x_i)](m-m^*)(x_i\pm\delta)\ge -c_8\delta^2. \tag{8.7}$$

In addition, from Taylor's Theorem,

$$[\operatorname{sgn}(f - m^*)(x_i)](m - m^*)(x_i \pm \delta) \le [\operatorname{sgn}(f - m^*)(x_i)](m - m^*)(x_i)$$

$$\pm[\operatorname{sgn}(f - m^*)(x_i)](m - m^*)'(x_i)\,\delta + c_9\delta^2.$$

Thus, as $|(m - m^*)(x_i)| \le c_4\delta^2$ and using (8.7), we have

$$\begin{aligned}
&\pm\delta[\operatorname{sgn}(f - m^*)(x_i)](m - m^*)'(x_i)\\
&\quad\ge\ [\operatorname{sgn}(f - m^*)(x_i)](m - m^*)(x_i \pm \delta) - [\operatorname{sgn}(f - m^*)(x_i)](m - m^*)(x_i) - c_9\delta^2\\
&\quad\ge\ [\operatorname{sgn}(f - m^*)(x_i)](m - m^*)(x_i \pm \delta) - c_{10}\delta^2 \ge -c_{11}\delta^2.
\end{aligned}$$

We have proven that

$$\pm\delta\,[\operatorname{sgn}(f - m^*)(x_i)](m - m^*)'(x_i) \ge -c_{11}\delta^2,$$

implying

$$|(m - m^*)'(x_i)| \le c_{11}\delta. \ \square$$

Proof of Theorem 8.3 (cont'd): Lemmas 8.5 and 8.6 together prove Theorem 8.3. $\square$

The example prior to Theorem 8.3 shows that, in general, we cannot expect to obtain a power less than 2 if M is not a Haar space. In fact, a result is proved in Kroó [1983a] which shows that this estimate is optimal for any such subspace. We state this here without proof.

Theorem 8.7. *Let r be a positive integer, $r \ge 2$. Let $M \subset C^r[a,b]$ be a n-dimensional unicity space with respect to $C^r[a,b]$, and assume M is not a Haar space, i.e., there exists an $\widetilde{m} \in M$, $\widetilde{m} \ne 0$, with n distinct zeros in $[a,b]$. Then there exists an $f \in C^\infty[a,b]$ with best approximant $m^* \in M$, and a sequence of elements $m_k \in M$, $m_k \to m^*$, such that*

$$\|f - m_k\| - \|f - m^*\| \le \gamma\|m_k - m^*\|^2$$

for some constant $\gamma > 0$ that depends only on f and M and is independent of k.

9. The L^1-Norm Revisited

In this section, we consider the space $L^1[a,b]$ with measure μ. We assume in what follows that μ is a non-atomic positive finite measure on $[a,b]$ with the property that every real-valued continuous function is μ-measurable, and such that if $f \in C[a,b]$ satisfies $\|f\|_1 = 0$ then $f = 0$, i.e., $\|\cdot\|_1$ is truly a norm on $C[a,b]$. We consider finite-dimensional subspaces $M \subset C[a,b]$ that are unicity spaces with respect to all $f \in C[a,b]$ in the $L^1[a,b]$ norm, i.e., for which there exists a unique best approximant from M to each $f \in C[a,b]$ in the $L^1[a,b]$ norm. We restrict our consideration to approximation to continuous functions because if we consider approximation to all functions in $L^1[a,b]$ then there are no finite-dimensional unicity spaces. But there are many unicity spaces in this continuous setting, cf. Pinkus [1989]. One characterization of these unicity spaces is the following which may be found in Cheney, Wulbert [1969, Theorem 24].

Theorem 9.1. *Let μ be a non-atomic positive measure, as above. Let M be a finite-dimensional subspace of $C[a,b]$. Then M is a unicity space for all $f \in C[a,b]$ in the $L^1[a,b]$ norm if and only if there does not exist an $h \in L^\infty[a,b]$ and an $\widetilde{m} \in M$, $\widetilde{m} \neq 0$, for which*

$$\text{(i) } |h(x)| = 1, \ \textit{all } x \in [a,b],$$
$$\text{(ii) } \int_a^b hm \, \mathrm{d}\mu = 0, \ \textit{all } m \in M,$$
$$\text{(iii) } h|\widetilde{m}| \textit{ is continuous.}$$

Let $\omega(f;\delta)$ denote the standard modulus of continuity of $f \in C[a,b]$, i.e.,

$$\omega(f;\delta) = \max\{|f(x) - f(y)| : x, y \in [a,b], \ |x-y| \le \delta\}.$$

We recall that $\omega(f;\cdot)$ is a continuous nonnegative non-decreasing function for $\delta \ge 0$ with $\omega(f;0) = 0$. For any such continuous nonnegative non-decreasing function F, let $F^{-1}(x) = \min\{y : F(y) = x\}$ denote its inverse. Note that F^{-1} is nonnegative and increasing, but may have jumps, i.e., it need not be continuous. This next result from Kroó [1981a] provides a non-classical strong uniqueness result in the case where M is a finite-dimensional unicity space in $C[a,b]$ in the $L^1[a,b]$ norm. It is well-known, for example, that Haar spaces satisfy this condition. For Haar spaces, this next result was obtained by Björnestål [1975].

Theorem 9.2. *Let M be a finite-dimensional unicity space in $C[a,b]$ in the $L^1[a,b]$ norm. If $f \in C[a,b]$, m^* is the best approximant to f from M, and $\sigma > 0$ is an arbitrary positive constant, then for any $m \in M$ satisfying $\|f-m\|_1 - \|f-m^*\|_1 \le \sigma$, we have*

$$\|f-m\|_1 - \|f-m^*\|_1 \ge \gamma \|m-m^*\|_1 \omega^{-1}(f; D\|m-m^*\|_1)$$

where $\gamma, D > 0$ are constants that depend only on f, σ and M.

Proof: For convenience, we assume that $m^* = 0$, i.e., replace $f - m^*$ by f. Set $\varepsilon := \|f-m\|_1 - \|f\|_1$. We shall prove that if $\varepsilon \le 1$ then

$$\varepsilon \ge \gamma \|m\|_1 \omega^{-1}(f; D\|m\|_1)$$

for some constants $\gamma, D > 0$ that depend only on f and M. From Proposition 8.4, this suffices to prove the theorem.

As $0 \in P_M(f)$, there exists, by Theorem 5.5, an $h \in L^\infty[a,b]$ satisfying

$$\text{(i) } |h(x)| = 1, \ \text{all } x \in [a,b],$$
$$\text{(ii) } \int_a^b hm \, \mathrm{d}\mu = 0, \ \text{all } m \in M,$$
$$\text{(iii) } \int_a^b hf \, \mathrm{d}\mu = \|f\|_1.$$

From (i) and (iii), it follows that $h(x) = \operatorname{sgn} f(x)$ μ-a.e. on $N(f)$.

For each $x \in [a,b]$, we define the δ-neighborhood of x as

$$O_\delta := (x-\delta, x+\delta) \cap [a,b].$$

Given the above $h \in L^\infty[a,b]$, let $S(f)$ denote the set of $x \in [a,b]$ for which

$$\mu(O_\delta(x) \cap \{x : h(x) = 1\}) > 0$$

and

$$\mu(O_\delta(x) \cap \{x : h(x) = -1\}) > 0$$

for all $\delta > 0$. That is, $S(f)$ is the set of essential sign changes of h. It is easily verified that if $S(f)$ is not empty then it is closed and hence compact. Thus for each fixed $\delta > 0$, we have

$$\min_{\sigma=\pm1} \inf_{x \in S(f)} \mu(O_\delta(x) \cap \{x : h(x) = \sigma\}) = \beta_f(\delta) > 0. \tag{9.1}$$

We have yet to show that $S(f) \neq \emptyset$. This follows immediately from Theorem 9.1. In fact, we have that

$$|||m||| := \max_{x \in S(f)} |m(x)|$$

is a norm on M. We prove this as follows. By the equivalence of norms on finite-dimensional subspaces of $C[a,b]$ it suffices to prove that only the zero function in M vanishes identically on $S(f)$. Assume to the contrary that there exists an $\widetilde{m} \in M$, $\widetilde{m} \neq 0$, that vanishes on $S(f)$. We claim that we may choose h as above so that $h|\widetilde{m}| \in C[a,b]$. To see this, assume $\widetilde{m}$ does not vanish on an interval (c,d). As $(c,d) \cap S(f) = \emptyset$ for each $x \in (c,d)$, there exists a $\delta_x > 0$ such that $h = 1$ or $h = -1$ μ-a.e. on $O_{\delta_x}(x)$. From the connectedness of (c,d), this implies that we may take $h = 1$ or $h = -1$ on all of (c,d). Thus $h|\widetilde{m}|$ is continuous on (c,d), and hence it easily follows that $h|\widetilde{m}| \in C[a,b]$. But, by Theorem 9.1, this contradicts the fact that M is a unicity space for all $f \in C[a,b]$ in the $L^1[a,b]$ norm. No such $\widetilde{m}$ exists and therefore $|||\cdot|||$ is a norm on M.

As $|||\cdot|||$ is a norm on M, it therefore follows that there exists a constant $c_1 > 0$, that depends only on M, satisfying

$$|||m||| \geq c_1 \|m\|_\infty$$

for all $m \in M$. Here and in what follows, the c_j will always denote positive constants that depend only upon f and M.

Returning to the proof of the theorem, we have

$$\begin{aligned}
\varepsilon &= \|f-m\|_1 - \|f\|_1 \\
&= \int_a^b (f-m)\,\mathrm{sgn}(f-m)\,\mathrm{d}\mu - \int_a^b f\,\mathrm{sgn}\,f\,\mathrm{d}\mu \\
&= \int_{N(f)} (f-m)\,\mathrm{sgn}(f-m)\,\mathrm{d}\mu + \int_{Z(f)} (-m)\,\mathrm{sgn}(-m)\,\mathrm{d}\mu - \int_{N(f)} f\,\mathrm{sgn}\,f\,\mathrm{d}\mu \\
&= \int_{N(f)} (f-m)\,[\mathrm{sgn}(f-m) - \mathrm{sgn}\,f]\,\mathrm{d}\mu - \int_{N(f)} m\,\mathrm{sgn}\,f\,\mathrm{d}\mu + \int_{Z(f)} |m|\,\mathrm{d}\mu.
\end{aligned}$$

From (ii), we have

$$\int_a^b hm\,\mathrm{d}\mu = 0$$

for all $m \in M$, while $h = \mathrm{sgn}\,f$ on $N(f)$. Thus

$$\begin{aligned}
\varepsilon &= \int_{N(f)} (f-m)\,[\mathrm{sgn}(f-m) - \mathrm{sgn}\,f]\,\mathrm{d}\mu + \int_{Z(f)} mh\,\mathrm{d}\mu + \int_{Z(f)} |m|\,\mathrm{d}\mu \\
&= 2\int_{\{x : f(f-m)<0\}} |f-m|\,\mathrm{d}\mu + \int_{Z(f)} |m| + mh\,\mathrm{d}\mu.
\end{aligned} \tag{9.2}$$

For $m \in M$, as above, let $x^* \in S(f)$ satisfy

$$|m(x^*)| = |||m||| \geq c_1 \|m\|_\infty$$

and for convenience set $\sigma^* := \operatorname{sgn} m(x^*)$. Thus, there exists a $\delta > 0$ such that for all $x \in O_\delta(x^*)$,

$$\sigma^* m(x) = |m(x)| \geq \frac{c_1}{2} \|m\|_\infty \tag{9.3}$$

and thus by (9.1)

$$\mu(O_\delta(x^*) \cap \{x : h(x) = \sigma^*\}) \geq \beta_f(\delta) = c_2 > 0. \tag{9.4}$$

Let

$$B := O_\delta(x^*) \cap \{x : x \in N(f), \operatorname{sgn} f(x) = \sigma^*\}$$

and

$$Q := O_\delta(x^*) \cap \{x : x \in Z(f), h(x) = \sigma^*\}.$$

By (9.4), we have

$$\mu(B) + \mu(Q) \geq c_2.$$

We now consider three options.

I. Assume $\mu(Q) \geq c_2/2$. Since $Q \subseteq Z(f)$, we have from (9.2) that

$$\begin{aligned} \varepsilon \ &\geq \ \int_{Z(f)} |m| + mh \, \mathrm{d}\mu \geq \int_Q |m| + mh \, \mathrm{d}\mu \ = \ 2 \int_Q |m| \, \mathrm{d}\mu \\ &\geq \ c_1 \|m\|_\infty \mu(Q) \geq c_3 \|m\|_1 \end{aligned} \tag{9.5}$$

due to the equivalence of the $\|\cdot\|_1$ and $\|\cdot\|_\infty$ norms on M. Now if $m \in M$ satisfies $\|f - m\|_1 - \|f\|_1 \leq 1$ then

$$\|m\|_1 \leq \|f - m\|_1 + \|f\|_1 \leq 2\|f\|_1 + 1,$$

i.e., all such m are uniformly bounded in norm, and therefore $\omega^{-1}(f; \|m\|_1)$ is also uniformly bounded above for all such m. Thus (9.5) implies the desired result.

II. Assume $\mu(B) \geq c_2/2$ and $\sigma^* f(x) < (c_1/4)\|m\|_\infty$ for all $x \in O_\delta(x)$. From (9.3) and our assumption, we have that

$$B \subseteq \{x : f(x)(f - m)(x) < 0\}$$

and therefore from (9.2) and (9.3)

$$\begin{aligned} \varepsilon \ &\geq \ 2 \int_{\{x : f(f-m) < 0\}} |f - m| \, \mathrm{d}\mu \ \geq \ 2 \int_B |f - m| \, \mathrm{d}\mu \\ &\geq \ 2 \int_B \frac{c_1}{2} \|m\|_\infty - \frac{c_1}{4} \|m\|_\infty \, \mathrm{d}\mu \ = \ \frac{c_1}{2} \|m\|_\infty \mu(B) \ \geq \ c_4 \|m\|_1 \end{aligned}$$

which, as above, implies the result.

III. Assume $\mu(B) \geq c_2/2$ and there exists an $\widetilde{x} \in O_\delta(x^*)$ for which $\sigma^* f(\widetilde{x}) \geq (c_1/4)\|m\|_\infty$. Since $x^* \in Z(f)$, we have $f(x^*) = 0$. Assume, without loss of generality, that $\widetilde{x} > x^*$. Choose

$x^* \le x_1 < x_2 < \widetilde{x}$ so that $f(x_1) = 0$, $\sigma^* f(x_2) = (c_1/4)\|m\|_\infty$ and $0 < \sigma^* f(x) < (c_1/4)\|m\|_\infty$ on (x_1, x_2). As $f(x_2) - f(x_1) = \sigma^*(c_1/4)\|m\|_\infty$, we have

$$x_2 - x_1 \ge \omega_f^{-1}\left(\frac{c_1}{4}\|m\|_\infty\right). \tag{9.6}$$

Since $(x_1, x_2) \subseteq O_\delta(x^*)$ from (9.3), we see that

$$f(x)(f - m)(x) < 0$$

for $x \in (x_1, x_2)$. Thus from (9.2) and (9.6)

$$\begin{aligned} \varepsilon \;&\ge\; 2\int_{x_1}^{x_2} |f - m| \,\mathrm{d}\mu \;\ge\; 2\int_{x_1}^{x_2} \frac{c_1}{2}\|m\|_\infty - \frac{c_1}{4}\|m\|_\infty \,\mathrm{d}\mu \\ &=\; \frac{c_1}{2}\|m\|_\infty (x_2 - x_1) \;\ge\; \frac{c_1}{2}\|m\|_\infty \omega_f^{-1}\left(\frac{c_1}{4}\|m\|_\infty\right) \;\ge\; \gamma\|m\|_1 \omega_f^{-1}\left(D\|m\|_1\right). \end{aligned}$$

This proves the theorem. □

We now turn to the one-sided L^1-approximation problem. For this problem, unicity spaces are rare when considering approximating all $f \in C[a,b]$ from some finite-dimensional subspace M in $C[a,b]$; see Pinkus [1989]. The situation is different if we consider the case where both the approximating subspace M and the functions to be approximated are in $C^1[a,b]$, i.e., are continuously differentiable. In this case, there are many unicity spaces.

For convenience, we again consider the interval $[a,b]$. The more general case can be found in Pinkus [1989]. For each $f \in C^1[a,b]$, we define $Z_1(f)$ to be the set of zeros of f in $[a,b]$ that only includes interior zeros if the function and its derivative vanish, i.e., $x \in (a,b)$ is in $Z_1(f)$ only if $f(x) = f'(x) = 0$. Note that if $f \in C^1[a,b]$ satisfies $f \ge 0$, then $Z_1(f) = Z(f)$. In what follows, we also always assume that M contains a strictly positive function. With these assumptions, we have the following characterization of unicity subspaces that is from Pinkus, Strauss [1987]; see also Strauss [1982] and Pinkus [1989].

Theorem 9.3. *An n-dimensional subspace M of $C^1[a,b]$ is a unicity space for $C^1[a,b]$ in the one-sided L^1-norm problem if and only if there does not exist an $m^* \in M$, $m^* \ne 0$, points $\{x_i\}_{i=1}^k$ in $Z_1(m^*)$, $1 \le k \le n$, and positive values $\{\lambda_i\}_{i=1}^k$ satisfying*

$$\int_a^b m \,\mathrm{d}\mu = \sum_{i=1}^k \lambda_i m(x_i)$$

for all $m \in M$.

Let $m_1, \ldots, m_n$ be any basis for the n-dimensional subspace M. In what follows, we assume $M \subset C^1[a,b]$. Given an $f \in C^1[a,b]$, we set

$$H_f(\delta) := \omega(f';\delta) + \sum_{i=1}^n \omega(m_i';\delta).$$

With these assumptions and definitions, we can now state the next result which is in Kroó, Sommer, Strauss [1989]. Recall that $M(f) := \{m : m \in M, m \le f\}$.

Theorem 9.4. *Assume $M \subset C^1[a,b]$ is a finite-dimensional unicity subspace for $C^1[a,b]$ in the one-sided L^1-norm problem. Assume $f \in C^1[a,b]$, m^* is the best approximant to f from $M(f)$, and $\sigma > 0$ is an arbitrary positive constant. If $m \in M(f)$ is such that $\|f-m\|_1 - \|f-m^*\|_1 \le \sigma$, then we have*

$$\|f-m\|_1 - \|f-m^*\|_1 \ \ge \ \gamma\|m-m^*\|_1 H_f^{-1}(D\|m-m^*\|_1)$$

for some constants $\gamma, D > 0$ that depend only on f, σ and M.

If, for example, f and the elements of M are such that their derivatives are all in Lip α, $0 < \alpha \le 1$, then $H_f(\delta) = O(\delta^\alpha)$, and from the above, we obtain

$$\|f-m\|_1 - \|f-m^*\|_1 \ \ge \ \gamma\|m-m^*\|_1^{(\alpha+1)/\alpha}$$

for some other constant $\gamma > 0$ that depends only on f and M.

Proof: For convenience, we assume that $m^* = 0$, i.e., replace $f - m^*$ by f. Set

$$\varepsilon := \|f-m\|_1 - \|f\|_1,$$

and assume $0 < \varepsilon \le 1$, as in the proof of Theorem 9.2. Since the zero function is a best one-sided L^1-approximation from the n-dimensional subspace M, we have that $f \ge 0$, and from Theorem 5.7, there exist distinct points $\{x_i\}_{i=1}^k$ in $Z_1(f)$, $1 \le k \le n$, and positive numbers $\{\lambda_i\}_{i=1}^k$ for which

$$\int_a^b m\,\mathrm{d}\mu = \sum_{i=1}^k \lambda_i m(x_i)$$

for all $m \in M$.

Now

$$\varepsilon = \|f-m\|_1 - \|f\|_1 = \int_a^b (f-m)\,d\mu - \int_a^b f\,\mathrm{d}\mu = -\int_a^b m\ \mathrm{d}\mu = -\sum_{i=1}^n \lambda_i m(x_i).$$

As $f(x_i) = 0$ and $f - m \ge 0$, we have $m(x_i) \le 0$ and thus

$$\varepsilon = \sum_{i=1}^k \lambda_i |m(x_i)|.$$

Recall that our aim is to prove that there exist constants $\gamma, D > 0$ that depend only on f and M such that

$$\varepsilon = \sum_{i=1}^k \lambda_i |m(x_i)| \ge \gamma\|m\|_1 H_f^{-1}(D\|m\|_1).$$

As M is a unicity space for $C^1[a,b]$ in the one-sided L^1-norm problem, we have from Theorem 9.3 that there does not exist an $\widetilde{m} \in M$, $\widetilde{m} \ne 0$, points $\{y_i\}_{i=1}^k$ in $Z_1(\widetilde{m})$, $1 \le k \le n$, and positive values $\{\mu_i\}_{i=1}^k$ satisfying

$$\int_a^b m\,\mathrm{d}\mu = \sum_{i=1}^k \mu_i m(y_i)$$

for all $m \in M$. Thus, for the above $\{x_i\}_{i=1}^k$, if $m \in M$ satisfies $m(x_i) = 0$, $i = 1, \ldots, k$, and $m'(x_i) = 0$, $x_i \in (a, b)$, then necessarily $m = 0$. This implies that

$$|||m||| := \sum_{i=1}^{k} |m(x_i)| + \sum_{x_i \in (a,b)} |m'(x_i)|$$

is a norm on M and therefore

$$\|m\|_1 \le c_1 |||m||| \tag{9.7}$$

for all $m \in M$. As usual, the c_j will always denote positive constants that depend only upon f and M.

As $\varepsilon = \sum_{i=1}^n \lambda_i |m(x_i)|$ where $\lambda_i > 0$ for all i, we have

$$|m(x_i)| \le c_2 \varepsilon, \qquad i = 1, \ldots, k. \tag{9.8}$$

To use the norm $|||m|||$, we must also bound the $|m'(x_i)|$ for $x_i \in (a, b)$. To this end, assume $m = \sum_{i=1}^n b_i m_i$. Then

$$\omega(m'; \delta) \le \sum_{i=1}^{n} |b_i|\, \omega(m_i'; \delta) \le (\max_{j=1,\ldots,n} |b_j|) \sum_{i=1}^{n} \omega(m_i'; \delta).$$

As $m_1, \ldots, m_n$ is a basis for M, it follows that $\max_{j=1,\ldots,n} |b_j|$ is also a norm on M. Therefore $\max_{j=1,\ldots,n} |b_j| \le c_3 \|m\|_1$. Furthermore, $\sum_{i=1}^n \omega(m_i'; \delta) \le H_f(\delta)$. Thus

$$\omega(m'; \delta) \le c_3 \|m\|_1 H_f(\delta). \tag{9.9}$$

For m satisfying $\|f - m\|_1 - \|f\|_1 \le 1$, it follows that

$$\|m\|_1 \le \|f - m\|_1 + \|f\|_1 \le 2\|f\|_1 + 1.$$

Thus $\|m\|_1$ is uniformly bounded above and from (9.9), we have

$$\omega(m'; \delta) \le c_4 H_f(\delta). \tag{9.10}$$

Consider $x_i \in (a, b)$. Choose $1 \ge \eta > 0$ such that the $(x_i - \eta, x_i + \eta)$ are disjoint intervals of (a, b). For each $m \in M$, $0 < h < \eta$ and $b = \operatorname{sgn} m'(x_i)$, we have

$$m(x_i + bh) = m(x_i) + bh\, m'(x_i) + E = m(x_i) + h|m'(x_i)| + E. \tag{9.11}$$

An easy estimate and (9.10) shows that

$$|E| \le h\omega(m'; h) \le c_4 h H_f(h). \tag{9.12}$$

Since $f(x_i) = f'(x_i) = 0$, we therefore have

$$|f(x_i + bh)| = |f(x_i + bh) - f(x_i)| \le h\omega(f'; h) \le hH_f(h). \tag{9.13}$$

From $m(x_i+bh)\le f(x_i+bh)$, we have by (9.7), (9.11), (9.12) and (9.13)

$$\begin{aligned}|m'(x_i)| &= \frac{m(x_i+bh)-m(x_i)-E}{h}\\ &\le \frac{1}{h}\left(f(x_i+bh)+c_2\varepsilon-E\right)\\ &\le \frac{1}{h}\left(hH_f(h)+c_2\varepsilon+c_4hH_f(h)\right)\\ &= c_5\left(H_f(h)+\frac{\varepsilon}{h}\right).\end{aligned} \tag{9.14}$$

Using the equivalence of norms, from (9.7), and (9.14), we finally obtain

$$\begin{aligned}\|m\|_1 &\le c_1\left(\sum_{i=1}^{k}|m(x_i)|+\sum_{x_i\in(a,b)}|m'(x_i)|\right)\\ &\le c_1\left(kc_2\varepsilon+kc_5\left(H_f(h)+\frac{\varepsilon}{h}\right)\right)\\ &\le c_6\left(H_f(h)+\frac{\varepsilon}{h}\right)\end{aligned} \tag{9.15}$$

for all $h\in(0,\eta)$.

Let us recall that we wish to bound $\varepsilon=\|f-m\|_1-\|f\|_1$ from below. To this end, we consider two cases.

Case 1: $\|m\|_1\le 2c_6\varepsilon/\eta$. In this case, it follows from the monotonicity of H_f that

$$H_f^{-1}(\|m\|_1)\le H_f^{-1}(2c_6\varepsilon/\eta)\le c_7$$

and therefore

$$\|m\|_1H_f^{-1}(\|m\|_1)\le 2c_6c_7\varepsilon/\eta,$$

i.e.,

$$\gamma\|m\|_1H_f^{-1}(\|m\|_1)\le\varepsilon.$$

Case 2: $\|m\|_1>2c_6\varepsilon/\eta$. Set $h:=2c_6\varepsilon/\|m\|_1<\eta$ in (9.15) to obtain

$$\|m\|_1\le c_6\left(H_f\left(\frac{2c_6\varepsilon}{\|m\|_1}\right)+\frac{\|m\|_1}{2c_6}\right).$$

Thus

$$\|m\|_1\le\frac{\|m\|_1}{2}+c_6H_f\left(\frac{2c_6\varepsilon}{\|m\|_1}\right)$$

implying

$$c_8\|m\|_1\le H_f\left(\frac{2c_6\varepsilon}{\|m\|_1}\right)$$

and therefore

$$H_f^{-1}\left(c_8\|m\|_1\right)\le\frac{2c_6\varepsilon}{\|m\|_1}$$

whence

$$\frac{\|m\|_1}{2c_6}H_f^{-1}\left(c_8\|m\|_1\right)\le\varepsilon.$$

This proves the theorem. □

A converse result is proved in Kroó, Sommer, Strauss [1989] that shows that this estimate is optimal. We state it here without proof.

Theorem 9.5. *Assume $M \subset C^1[a,b]$ is an n-dimensional unicity subspace for $C^1[a,b]$ in the one-sided L^1-norm problem. Then there exists an $f \in C^1[a,b]$ with m^* the best approximant to f from $M(f)$, and a sequence of elements $m_k \in M(f)$, $m_k \to m^*$, such that*

$$\|f - m_k\|_1 - \|f - m^*\|_1 \leq \gamma \|m_k - m^*\|_1 H_f^{-1}(D\|m_k - m^*\|_1)$$

for some constants $\gamma, D > 0$ that depend only on f and M and are independent of k.

10. Strong Uniqueness in Complex Approximation in the Uniform Norm

In this section, we consider the question of strong uniqueness in the space of continuous complex-valued functions endowed with the uniform norm. The principal result here will indicate that classical strong uniqueness fails, in general, in the complex setting in a Haar space. Instead, we shall derive a non-classical strong uniqueness type result of order 2, i.e., prove that for m^* the best approximant to f from M, and any $m \in M$ sufficiently close to m^*, we have

$$\|f - m\| - \|f - m^*\| \geq \gamma \|m - m^*\|^2,$$

where $\gamma > 0$ depends only on f and M. To this end, we shall need to extend some results from Section 2 to the complex setting. Throughout this section, $C(K)$ will denote the space of complex-valued functions, continuous on the compact Hausdorff space K. First, we shall require a formula for the complex τ-functional that is similar to the one provided by Theorem 2.1. We recall that $A_f = \{x : |f(x)| = \|f\|\}$, and if $z = x + \mathrm{i}y$ where $x, y \in \mathbb{R}$, then we set $\Re z = x$ and $\Im z = y$.

Theorem 10.1. *For any functions $f, g \in C(K)$, $f \neq 0$, we have*

$$\tau_+(f,g) = \frac{1}{\|f\|} \max_{x \in A_f} \Re(\overline{f(x)} g(x)).$$

Proof: Note that for any $z_1, z_2 \in \mathbb{C}$ and $t \in \mathbb{R}$, we have

$$|z_1 + t z_2|^2 = |z_1|^2 + 2t \,\Re\, \overline{z_1} z_2 + t^2 |z_2|^2. \tag{10.1}$$

Using this relation, we obtain for every $x \in A_f$, $f \neq 0$,

$$\tau_+(f,g) \geq \lim_{t \to 0^+} \frac{|f(x) + t g(x)|^2 - |f(x)|^2}{t(|f(x) + t g(x)| + |f(x)|)} = \frac{\Re(\overline{f(x)} g(x))}{|f(x)|},$$

i.e.,

$$\tau_+(f,g) \geq \frac{1}{\|f\|} \max_{x \in A_f} \Re(\overline{f(x)} g(x)).$$

This proves the desired lower bound for $\tau_+(f,g)$.

In order to verify the upper bound, choose $t_n \to 0^+$ and let $x_n \in K$ be such that $\|f + t_n g\|$ is attained at x_n. Similar to the argument used in the proof of Theorem 2.1, we can assume, without loss of generality, that $x_n \to x^* \in A_f$. Then using the fact that $|f(x_n)|^2 \leq |f(x^*)|^2$, we obtain

$$\begin{aligned} \tau_+(f,g) &= \lim_{n \to \infty} \frac{|f(x_n) + t_n g(x_n)|^2 - |f(x^*)|^2}{t_n(|f(x_n) + t_n g(x_n)| + |f(x^*)|)} = \lim_{n \to \infty} \frac{|f(x_n)|^2 - |f(x^*)|^2}{t_n(|f(x_n) + t_n g(x_n)| + |f(x^*)|)} \\ &\quad + \lim_{n \to \infty} \frac{2\Re(\overline{f(x_n)} g(x_n))}{|f(x_n) + t_n g(x_n)| + |f(x^*)|} \leq \frac{\Re(\overline{f(x^*)} g(x^*))}{|f(x^*)|}. \end{aligned}$$

(Note that since the limit on the left hand side and the second limit on the right hand side exist, the first limit on the right hand side must exist, as well.) Clearly this provides the needed upper bound. □

Applying the above and the argument of Theorem 1.4, we obtain

Corollary 10.2. *Let $f \in C(K)$ and M be a linear subspace of $C(K)$. Then $m^* \in M$ is a best approximant to f from M if and only if*

$$\max_{x \in A_{f-m^*}} \Re(\overline{(f-m^*)(x)}m(x)) \geq 0, \quad m \in M.$$

Moreover, m^* is the strongly unique best approximant to f from M if and only if

$$\inf_{m \in M, \|m\|=1} \max_{x \in A_{f-m^*}} \Re(\overline{(f-m^*)(x)}m(x)) > 0.$$

The next theorem provides an analogue of Theorem 2.3 in the complex setting. Since this is a key result in this section, we include a proof.

Theorem 10.3. *Let $f \in C(K)$ and M be an n-dimensional subspace of $C(K)$. Then $m^* \in M$ is a best approximant to f from M if and only if there exist points $x_1, \ldots, x_k \in A_{f-m^*}$, and positive numbers $\lambda_1, \ldots, \lambda_k, 1 \leq k \leq 2n+1$, such that for every $m \in M$, we have*

$$\sum_{j=1}^{k} \lambda_j \overline{(f-m^*)(x_j)}m(x_j) = 0. \tag{10.2}$$

Proof: ($\Leftarrow$) We may assume that $\sum_{j=1}^{k} \lambda_j = 1$ and $m^* = 0$. Then using (10.2), we have for every $m \in M$

$$\|f\|^2 = \sum_{j=1}^{k} \lambda_j \overline{f(x_j)} f(x_j) = \sum_{j=1}^{k} \lambda_j \overline{f(x_j)}(f-m)(x_j) \leq \|f\| \cdot \|f-m\|.$$

Dividing both sides of the above inequality by $\|f\|$ yields that $m^* = 0$ is a best approximant to f from M.

($\Rightarrow$) Assume again that $m^* = 0$. Let $M = \text{span}\{m_1, \ldots, m_n\}$. For any $x \in A_f$, consider the vectors

$$\mathbf{u}_x := (\Re(\overline{f(x)}m_1(x)), -\Im(\overline{f(x)}m_1(x)), \ldots, \Re(\overline{f(x)}m_n(x)), -\Im(\overline{f(x)}m_n(x))) \in \mathbb{R}^{2n}.$$

Furthermore, let $D := \{\mathbf{u}_x, x \in A_f\} \subset \mathbb{R}^{2n}$. We claim that $\mathbf{0}$ lies in the convex hull of D. Indeed, if this is not the case, then by the separating hyperplane theorem, there exists a $\mathbf{c} = (c_1, \ldots, c_{2n}) \in \mathbb{R}^{2n}$ such that $(\mathbf{c}, \mathbf{u}_x) < 0$ for every $x \in A_f$. This clearly yields that for some $\widetilde{m} \in M$, we have $\Re(\overline{f(x)}\widetilde{m}(x)) < 0$ for every $x \in A_f$. But in view of Corollary 10.2, this contradicts the assumption that 0 is a best approximant to f from M. Thus $\mathbf{0}$ belongs to the convex hull of $D \subset \mathbb{R}^{2n}$ and, by the Carathéodory Theorem, $\mathbf{0}$ is a convex linear combination of at most $2n+1$ points of D. Evidently, this verifies relations (10.2). $\square$

We are now in a position to prove a strong uniqueness type result for complex Chebyshev approximation. This is the main result of this section and can be found in Newman, Shapiro [1963]. The proof as presented here is from Smarzewski [1989]. As previously, we shall call the n-dimensional subspace M of $C(K)$ a Haar space if every nontrivial element of M has at most n distinct zeros in K.

Theorem 10.4. *Let M be a finite-dimensional Haar space in $C(K)$. Let $f \in C(K)$ and m^* be the best approximant to f from M. Then there exists a $c > 0$ that depends only upon f and M such that*

$$\|f - m\|^2 - \|f - m^*\|^2 \geq c\|m - m^*\|^2, \tag{10.3}$$

for all $m \in M$. In addition, if $m \in M$ satisfies

$$\|f - m\| - \|f - m^*\| \leq \sigma$$

then

$$\|f - m\| - \|f - m^*\| \geq \gamma\|m - m^*\|^2, \tag{10.4}$$

where $\gamma := c/(2\|f - m^*\| + \sigma) > 0$ and thus also depends only upon f, M and σ.

Proof: Inequality (10.4) is an immediate consequence of (10.3) since if $\|f - m\| - \|f - m^*\| \leq \sigma$, then

$$\begin{aligned} c\|m - m^*\|^2 &\leq \|f - m\|^2 - \|f - m^*\|^2 = (\|f - m\| + \|f - m^*\|)(\|f - m\| - \|f - m^*\|) \\ &\leq (2\|f - m^*\| + \sigma)(\|f - m\| - \|f - m^*\|). \end{aligned}$$

It remains to prove (10.3). Assume M is a Haar space of dimension n. We recall from Theorem 10.3 that

$$\sum_{j=1}^{k} \lambda_j \overline{(f - m^*)(x_j)} m(x_j) = 0$$

for all $m \in M$, where $x_1, \ldots, x_k \in A_{f-m^*}$, $\lambda_1, \ldots, \lambda_k > 0$ and, since M is a Haar space, we have $n + 1 \leq k \leq 2n + 1$ and, more importantly,

$$\left(\sum_{j=1}^{k} \lambda_j |m(x_j)|^2 \right)^{1/2} =: |||m|||$$

is a norm on M. Hence there exists a $c > 0$ such that

$$|||m|||^2 \geq c\|m\|^2 \tag{10.5}$$

for all $m \in M$.

For each $m \in M$ and $j \in \{1, \ldots, k\}$, we have

$$\begin{aligned} \|f - m\|^2 &\geq |(f - m)(x_j)|^2 = |(f - m^*)(x_j) + (m - m^*)(x_j)|^2 \\ &= \|f - m^*\|^2 + 2\Re\left(\overline{(f - m^*)(x_j)}(m - m^*)(x_j)\right) + |(m - m^*)(x_j)|^2. \end{aligned}$$

Multiply the above by $\lambda_j > 0$ and sum over j. Assuming, without loss of generality, that $\sum_{j=1}^{k} \lambda_j = 1$, we obtain

$$\|f - m\|^2 \geq \|f - m^*\|^2 + 2\Re\left(\sum_{j=1}^{k} \lambda_j \overline{(f - m^*)(x_j)}(m - m^*)(x_j) \right)$$

$$+\sum_{j=1}^{k} \lambda_j |(m-m^*)(x_j)|^2.$$

Applying (10.2) and (10.5), this gives

$$\|f-m\|^2 - \|f-m^*\|^2 \geq c\|m-m^*\|^2. \quad \square$$

Clearly, estimations (10.3) and (10.4) provide strong uniqueness type results that are weaker than the classical strong uniqueness result that holds in the real case when approximating by elements of Haar spaces. This raises the natural question of whether classical strong uniqueness can also hold for every complex function. Our next proposition, see Gutknecht [1978] for a particular example thereof, shows that this is not the case. In fact, it turns out that, in general, (10.4) provides the best possible estimate.

Proposition 10.5. *Let M be any subspace in $C(K)$ containing the constant function. Assume that there exists an $f \in C(K)$ such that $\Im f \equiv 0$ on K and $m^* \equiv 0$ is the best approximant to f from M. (Such an f will clearly exist, for instance, when M possesses a basis consisting of real functions.) Then for $m := \mathrm{i}b \in M$, $b \in \mathbb{R}$, we have the following converse to (10.4):*

$$\|f-m\| - \|f-m^*\| \leq \frac{\|m-m^*\|^2}{2\|f\|}. \tag{10.6}$$

Proof: Indeed, by the above assumptions, $\|f-m\|^2 = \|f\|^2 + \|m\|^2$ and since $m^* \equiv 0$, the above inequality easily follows. $\square$

From (10.6), it follows that there cannot exist a $\gamma > 0$ such that

$$\|f-m\| - \|f-m^*\| \geq \gamma\|m-m^*\|$$

for all $m \in M$.

Despite the fact that classical strong uniqueness fails, in general, in the complex case, it is possible to give a relatively simple sufficient condition for it to hold. Let $f \in C(K)$ and assume that m^* is its best approximant. We shall say that the set of points $A \subset A_{f-m^*}$ is *extremal* if m^* is a best approximant to f from M also on the set A. Moreover, A is a minimal extremal set if no proper subset of A is extremal. Clearly, any set of points $x_1, \ldots, x_k \in A_{f-m^*}$, $k \in \{1, \ldots, 2n+1\}$, satisfying conditions (10.2) of Theorem 10.3 must be extremal. Hence the cardinality of any minimal extremal set can be at most $2n+1$. This also follows easily from Corollary 10.2.

We now give a sufficient condition for classical strong uniqueness to hold in the complex case. This sufficient condition is to be found in Theorem 2 in Gutknecht [1978].

Theorem 10.6. *Let $f \in C(K)$ and M be an n-dimensional subspace of $C(K)$. Assume that m^* is a best approximant to f from M possessing a minimal extremal set of cardinality $2n+1$. Then m^* is a strongly unique best approximant of f, i.e., for all $m \in M$, we have*

$$\|f-m\| - \|f-m^*\| \geq \gamma\|m-m^*\|,$$

for some $\gamma > 0$.

Proof: Let $\{x_1, \ldots, x_{2n+1}\} \subset A_{f-m^*}$ be a minimal extremal set. Then by Theorem 10.3, there exist corresponding positive numbers $\lambda_1, \ldots, \lambda_{2n+1}$ such that (10.2) holds with $k = 2n+1$. Assume that to the contrary m^* is not a strongly unique best approximant to f from M. Then by Corollary 10.2, there exists an $m_0 \in M$, $m_0 \neq 0$, such that

$$\max_{x \in A_{f-m^*}} \Re(\overline{(f-m^*)(x)} m_0(x)) = 0.$$

Since relations (10.2) hold for this m_0, we obtain that

$$\Re(\overline{(f-m^*)(x_j)} m_0(x_j)) = 0, \quad j = 1, \ldots, 2n+1. \tag{10.7}$$

Let $\{m_1, \ldots, m_n\}$ be a basis for M. For each $j \in \{1, \ldots, 2n+1\}$, consider the vectors

$$\mathbf{u}_j := (\Re(\overline{(f-m^*)(x_j)} m_i(x_j)), -\Im(\overline{(f-m^*)(x_j)} m_i(x_j)),\ i = 1, \ldots, n) \in \mathbb{R}^{2n},$$

and set $B := \{\mathbf{u}_j\}_{j=1}^{2n+1} \subset \mathbb{R}^{2n}$. Then relations (10.7) yield that the set B lies in a $(2n-1)$-dimensional hyperplane H of $\mathbb{R}^{2n}$. Moreover, by the standard argument repeatedly used above, $\mathbf{0}$ belongs to the convex hull of B (otherwise relations (10.2) would fail for some $m \in M$). Now, using the fact that the convex hull of B is of dimension $2n-1$, it follows from the Carathéodory Theorem that $\mathbf{0}$ is a convex linear combination of at most $2n$ points from B. But this in turn means that (10.2) holds for a proper subset of $\{x_1, \ldots, x_{2n+1}\}$ (with some λ's). Clearly this set of at most $2n$ points is extremal too, contradicting the condition of minimality of the extremal set $\{x_1, \ldots, x_{2n+1}\}$. $\square$

Thus if m^* is a best approximant with minimal extremal set consisting of exactly $2n+1$ points then the best approximation is strongly unique in the classical sense. This result is not vacuous. It is easy to construct examples that satisfy the criteria of Theorem 10.6. Our next example, see Gutknecht [1978], shows that we do not necessarily have classical strong uniqueness for a minimal extremal set consisting of fewer than $2n+1$ points.

Example. Let $K := \{|z| = 1 : z \in \mathbb{C}\}$ be the unit circle, $M := \Pi_{n-1}$ the space of polynomials of degree at most $n-1$, and

$$f(z) := \frac{z^n + z^{3n}}{2}.$$

Then it is easy to see that $\|f\| = 1$ and $A_f = \{z_k := \mathrm{e}^{\mathrm{i}\pi k/n} : k = 1, \ldots, 2n\}$, i.e., A_f is the set of all roots of unity of order $2n$. In particular, $f(z_k) = (-1)^k$, $k = 1, \ldots, 2n$. We claim that $m^* = 0$ is the best approximant of f. Indeed, if for some $m \in M$, we had $\|f - m\| < \|f\| = 1$ then evidently

$$(-1)^k \Re m(z_k) > 0, \quad k = 1, \ldots, 2n, \tag{10.8}$$

where $m = \sum_{j=0}^{n-1} d_j z^j$, $d_j = a_j + \mathrm{i} b_j$. Now $\Re m(\mathrm{e}^{\mathrm{i}t}) = \sum_{j=0}^{n-1} (a_j \cos jt - b_j \sin jt)$, i.e., $T_n := \Re m(\mathrm{e}^{\mathrm{i}t})$ is a trigonometric polynomial of degree at most $n-1$. By (10.8), T_n has at least $2n-1$ distinct zeros on the unit circle. Thus, we must have $T_n = 0$, and that contradicts (10.8). Hence $m^* = 0$ is a best approximant to f from M. It is also the unique best approximant since M is a Haar space. Note that by Corollary 10.2 this best approximant is not strongly unique in the classical sense since choosing m to be an imaginary constant, we have

$$\max_{x \in A_{f-m^*}} \Re(\overline{(f-m^*)(x)} m(x)) = \max_{x \in A_f} \Im(\overline{f(x)}) = 0.$$

It remains to show that the extremal set $A_f = \{z_k := \mathrm{e}^{\mathrm{i}\pi k/n} : k = 1, \ldots, 2n\}$ is minimal. Let us delete any point of A_f, say z_1, and assume that the remaining set $\{z_k := \mathrm{e}^{\mathrm{i}\pi k/n} : k = 2, \ldots, 2n\}$ is still extremal. (From symmetry considerations, deleting z_1 does not restrict the generality.) Clearly, we can choose a trigonometric polynomial $T^* = \Re m(\mathrm{e}^{\mathrm{i}t})$, $m \in \Pi_{n-1}$, of degree $n-1$ such that $(-1)^k T^*(k\pi/n) < 0$, $k = 2, \ldots, 2n$. But in view of Corollary 10.2 this means that $m^* = 0$ is no longer a best approximant on this set of $2n-1$ points. Thus the extremal set A_f consisting of $2n$ points is minimal, but the best approximant is not strongly unique in the classical sense.

Hence the condition of Theorem 10.6 is sharp, in a certain sense. Nevertheless, as the next example verifies, this condition is not necessary for classical strong uniqueness to hold.

Example. Let $K := \{|z| = 1 : z \in \mathbb{C}\}$ be the unit circle, $M := \Pi_{n-1}$ be the space of polynomials of degree at most $n-1$, and $f(z) := z^n$. Then $A_f = K$ and as in the previous example the function $m^* = 0$ is the unique best approximant to z^n from M with the set $\{z_k := \mathrm{e}^{\mathrm{i}\pi k/n} : k = 1, \ldots, 2n\}$ being a minimal extremal set of $2n$ points. Let us show that this best approximant is strongly unique in the classical sense. Indeed, if classical strong uniqueness is not valid then by Corollary 10.2 for some nontrivial $m = \sum_{j=0}^{n-1} d_j z^j$, $d_j = a_j + ib_j$, we would have

$$\max_{|z|=1} \Re(\overline{z^n} m(z)) = 0.$$

On the other hand, evidently $T_n := \Re(\overline{z^n} m(z)) = \sum_{j=0}^{n-1} a_j \cos(n-j)t + b_j \sin(n-j)t$ is a nonpositive trigonometric polynomial of degree n such that

$$\int_0^{2\pi} T_n(t)\,\mathrm{d}t = 0.$$

Thus T_n must be identically zero implying that m must be identically zero. By this contradiction, it follows that the best approximation is strongly unique in the classical sense. In fact, the optimal strong uniqueness constant in this case is $1/n$; see Rivlin [1984a].

An exact necessary and sufficient condition for when one has strong uniqueness in the classical sense is to found in Blatt [1984], who credits the result to Brosowski [1983]. These conditions are somewhat technical and are not detailed here.

Part III. Applications of Strong Uniqueness

We shall briefly consider various applications of strong uniqueness type results. The main idea behind these applications is the following: instead of solving the best approximation problem in a given norm, we replace it by considering another norm, close to the original norm, one that leads to a simpler approximation problem. Strong uniqueness is then applied in order to show that the best approximant in this new norm is sufficiently close to the original best approximant. Typically, the original norm is modified by replacing it by a similar discrete norm, or by introducing a weight function into the norm. We first start with some general remarks concerning approximation in nearby norms, and then proceed to a discussion of specific examples.

11. Strong Uniqueness and Approximation in Nearby Norms

Let X be a normed linear space with norm $\|\cdot\|$, and let M a finite-dimensional subspace of X. Assume that the sequence of seminorms $\|\cdot\|_k$, $k \in \mathbb{N}$, approximate the given norm, i.e.,

$\lim_{k\to\infty} \|f\|_k = \|f\|$, for all $f \in X$. We also assume that for a given $f \in X$, its best approximant $P_M(f) \in M$ with respect to the norm $\|\cdot\|$ is unique, and denote by $P_M^{(k)}(f) \in M$ the set of best approximants to f in the seminorms $\|\cdot\|_k$, $k \in \mathbb{N}$. We are interested in estimating the deviation of $P_M^{(k)}(f)$ from $P_M(f)$. This approach was first considered by Kripke [1964] and Peetre [1970].

To estimate the deviation of $P_M^{(k)}(f)$ from $P_M(f)$, we introduce a quantity measuring the deviation of $\|\cdot\|_k$ from $\|\cdot\|$ uniformly on the set $S(M) := \{m \in M : \|m\| = 1\}$, the unit sphere in M, namely

$$\eta_k(M) := \sup_{m\in S(M)} |\, \|m\|_k - 1|.$$

Lemma 11.1. *For any finite-dimensional subspace $M \subset X$, we have*

$$\lim_{k\to\infty} \eta_k(M) = 0.$$

Proof: Assume to the contrary that for a subsequence k_i, there exist elements $u_i \in S(M)$ such that

$$|\, \|u_i\|_{k_i} - 1| \geq \delta > 0, \qquad i \in \mathbb{N}. \tag{11.1}$$

Let $u_i = \Sigma_{j=1}^n a_j^i m_j$, where $m_1, \ldots, m_n$ is a basis for M. Since $S(M) \subset X$ is compact, we may assume, without loss of generality, that u_i converges to u^* as $i \to \infty$ for some $u^* = \Sigma_{j=1}^n a_j^* m_j \in S(M)$, i.e., $\|u^*\| = 1$. Moreover, by the equivalence of norms in finite-dimensional spaces,

$$\lim_{i\to\infty} \max_{1\leq j\leq n} |a_j^i - a_j^*| = 0.$$

In addition,

$$\lim_{i\to\infty} \sum_{j=1}^n \|m_j\|_{k_i} = \sum_{j=1}^n \|m_j\|.$$

Thus

$$\|u^* - u_i\|_{k_i} \leq \max_{1\leq j\leq n} |a_j^i - a_j^*| \sum_{j=1}^n \|m_j\|_{k_i}, \tag{11.2}$$

and the right-hand-side converges to 0 as $i \to \infty$.

Hence, we obtain by (11.2) that

$$\lim_{i\to\infty} |1 - \|u_i\|_{k_i}| \leq \lim_{i\to\infty} (|1 - \|u^*\|_{k_i}| + \|u^* - u_i\|_{k_i}) = |1 - \|u^*\|\,| = 0.$$

But this clearly contradicts (11.1). $\square$

Corollary 11.2. *If $k \in \mathbb{N}$ is sufficiently large so that $\eta_k(M) < 1/2$, then for all $m \in M$, we have*

$$\frac{2}{3}\|m\|_k \leq \|m\| \leq 2\|m\|_k. \tag{11.3}$$

Assume now that for a given $f \in X$, with unique best approximant $P_M(f) \in M$, non-classical strong uniqueness holds. That is, we assume there exists a nonnegative strictly increasing function ϕ (depending only on f and M), defined on $[0, \sigma]$, such that for all $m \in M$ satisfying $\|f - m\| - \|f - P_M(f)\| \leq \sigma$, we have

$$\|f - m\| - \|f - P_M(f)\| \geq \phi(\|m - P_M(f)\|). \tag{11.4}$$

Note that the above relation yields that $\phi(t) \leq t$ on $[0, \sigma]$. Moreover, if $\phi(t) \geq ct$ thereon then classical strong uniqueness holds at f. Using Lemma 11.1 and (11.4), we will estimate the deviation of $P_M^{(k)}(f)$ from $P_M(f)$. In what follows, $\eta_k(f + M)$ will be an abuse of notation for $\eta_k(\text{span}\{f, M\})$.

Proposition 11.3. *Assume that M is a finite dimensional subspace in X and a strong uniqueness type estimate of the form (11.4) holds for a given $f \in X$. Then for any best approximant $P_M^{(k)}(f)$ and any k sufficiently large so that $\eta_k(f+M) < 1/2$, we have*

$$\phi(\|P_M^{(k)}(f) - P_M(f)\|) \le 8\|f\|\eta_k(f+M).$$

In particular,

$$\lim_{k\to\infty} \|P_M^{(k)}(f) - P_M(f)\| = 0.$$

Proof: From (11.3) applied to both $P_M^{(k)}(f)$ and f, and since $\|P_M^{(k)}(f)\|_k \le 2\|f\|_k$, we have

$$\|f - P_M^{(k)}(f)\| \le \|f\| + \|P_M^{(k)}(f)\| \le \|f\| + 2\|P_M^{(k)}(f)\|_k \le \|f\| + 4\|f\|_k \le 7\|f\|.$$

From the definition of $\eta_k(f+M)$ and the above estimate, we have

$$|\,\|f - P_M^{(k)}(f)\| - \|f - P_M^{(k)}(f)\|_k| \le \|f - P_M^{(k)}(f)\|\eta_k(f+M) \le 7\|f\|\eta_k(f+M).$$

Thus

$$\begin{aligned}\|f - P_M^{(k)}(f)\| &\le \|f - P_M^{(k)}(f)\|_k + 7\|f\|\eta_k(f+M) \;\le\; \|f - P_M(f)\|_k + 7\|f\|\eta_k(f+M)\\ &\le \|f - P_M(f)\| + \|f - P_M(f)\|\eta_k(f+M) + 7\|f\|\eta_k(f+M)\\ &\le \|f - P_M(f)\| + 8\|f\|\eta_k(f+M).\end{aligned}$$

Combining now the last estimate with (11.4) immediately yields the desired statement. The limit follows from Lemma 11.1 since

$$\lim_{k\to\infty} \eta_k(f+M) = 0$$

and $\phi(0) = 0$ while $\phi(t) > 0$ for $t > 0$. □

Remark. The above proposition yields the general estimate

$$\|P_M^{(k)}(f) - P_M(f)\| \le \phi^{-1}(c\eta_k(f+M)), \tag{11.5}$$

where, by Lemma 11.1, the quantity on the right hand side tends to 0 as $k \to \infty$. In order to use this estimate to obtain rates of convergence, one needs to find sharp bounds for both η_k and ϕ. In many instances, this leads to sharp estimates for the deviation of $P_M^{(k)}(f)$ from $P_M(f)$ (for example, discrete Chebyshev approximation, Pólya algorithm). However, in some cases, this approach does not yield sharp bounds, even when both η_k and ϕ are precisely determined (this happens, for example, in the case of discrete L_1-approximation). In what follows, we provide a brief summary of these results.

12. Discretization of Norms

It is considerably easier to solve an approximation problem when the L_p-norm is replaced by a discrete L_p-norm. In this section, we shall investigate the size of the error of this discretization technique.

Let us denote by

$$\|f\|_k := \max_{0\le j\le k} |f(j/k)|$$

and

$$\|f\|_{k,p} := \frac{1}{k^{1/p}}(\sum_{j=0}^{k-1} |f(j/k)|^p)^{1/p}, \qquad 1 \le p < \infty$$

the discrete uniform and L_p-norms on $[0,1]$, respectively.

In what follows, we deal with the example $M = \Pi_n$. We first consider the case of the discrete uniform norm.

Theorem 12.1. *Let $f \in C^2[0,1]$ and $M = \Pi_n$. Set $X := C[0,1]$, with the usual uniform norm thereon, and*

$$\|f\|_k := \max_{0\le j\le k} |f(j/k)|.$$

Then there exists a constant C such that

$$\eta_k(f+\Pi_n) \le (C\|f''\| + 4n^4(1+C\|f\|))/(8k^2)$$

for all k, and therefore

$$\|P_M^{(k)}(f) - P_M(f)\| = O(k^{-2}).$$

Proof: We know from preceding results that classical strong uniqueness holds in this case, i.e., we can use estimate (11.5) with $\phi(t) = \gamma_n(f)t$, where $\gamma_n(f)$ is the strong unicity constant for the function f. We now need to estimate $\eta_k(f+\Pi_n)$ for $f \in C^2[0,1]\backslash\Pi_n$. Let $g \in S(f+\Pi_n)$ and assume that $\|g\| = g(x^*) = 1$, $x^* \in [0,1]$. If $x^* = j/k$ for some $j \in \{0,1,\ldots,k\}$, then $\|g\| = \|g\|_k = 1$. Thus, we may assume that $x^* \neq j/k$, $j \in \{0,1,\ldots,k\}$, i.e., in particular $x^* \in (0,1)$ and $g'(x^*) = 0$. Using Taylor's formula, we have that for every $x \in [0,1]$,

$$g(x) = g(x^*) + \frac{g''(\xi)}{2}(x-x^*)^2 \tag{12.1}$$

with some ξ between x and x^*. Obviously, we can choose a $j \in \{0,1,\ldots,k\}$ so that $|x^*-j/k| \le 1/2k$. Setting $x = j/k$ in (12.1) yields

$$1 = \|g\| \le \|g\|_k + \frac{1}{8k^2}\|g''\|.$$

Finally, $g = \alpha f + p_n$ for some $p_n \in \Pi_n$. As $\|g\| = 1$ and $f \notin \Pi_n$, it easily follows that $|\alpha| \le C = [\min_{p\in\Pi_n} \|f-p\|]^{-1}$. Thus $\|p_n\| \le 1 + C\|f\|$. Therefore by the Markov inequality,

$$\|g''\| \le C\|f''\| + 4n^4(1+C\|f\|).$$

From these estimates, we obtain

$$\eta_k(f+\Pi_n) \le (C\|f''\| + 4n^4(1+C\|f\|))/(8k^2).$$

In particular, $\eta_k(f+\Pi_n) = O(1/k^2)$, and hence by (11.5), we have

$$\|P_M^{(k)}(f) - P_M(f)\| = O(k^{-2}). \;\square$$

The above estimate turns out to be sharp, in general.

Let us now turn our attention to the discrete L_p-norms, $1 < p < \infty$.

Theorem 12.2. *Let $f \in C^1[0,1]$ and $M = \Pi_n$. Set $X := L_p[0,1]$, $1 < p < \infty$, and*

$$\|f\|_{k,p} := \frac{1}{k^{1/p}}(\sum_{j=0}^{k-1} |f(j/k)|^p)^{1/p}.$$

Then there exist constants C_p and c_p such that

$$\eta_k(f + \Pi_n) \le pc_p\frac{n^2}{k}(1 + C_p\|f\|_p + C_p\|f'\|_p)$$

for all k, and therefore

$$\|P_M^{(k)}(f) - P_M(f)\|_p = O(k^{-\theta_p}), \qquad 1 < p < \infty,$$

where $\theta_p := \min\{1/2, 1/p\}$.

Proof: Consider $f \in C^1[0,1]$ and let us estimate $\eta_k(f + \Pi_n)$ from above. Setting $\Delta_k := [j/k, (j+1)/k]$, we obtain, by repeated application of Hölder's inequality, for any $g \in C^1[0,1]$,

$$\begin{aligned}
\Big| \|g\|_p^p - \|g\|_{k,p}^p \Big| &= \Big| \int_0^1 |g(x)|^p \, \mathrm{d}x - \frac{1}{k}\sum_{j=0}^{k-1} |g(j/k)|^p \Big| \\
&= \Big| \sum_{j=0}^{k-1} \int_{\Delta_k} (|g(x)|^p - |g(j/k)|^p) \, \mathrm{d}x \Big| \le \sum_{j=0}^{k-1} \int_{\Delta_k} \left(\int_{j/k}^{x} p|g(\xi)|^{p-1} |g^{'}(\xi)| \, \mathrm{d}\xi \right) \mathrm{d}x \\
&\le \frac{p}{k} \sum_{j=0}^{k-1} \left(\int_{\Delta_k} |g(\xi)|^p \, \mathrm{d}\xi \right)^{1-1/p} \left(\int_{\Delta_k} |g'(\xi)|^p \, \mathrm{d}\xi \right)^{1/p} \\
&\le \frac{p}{k} \left(\sum_{j=0}^{k-1} \int_{\Delta_k} |g(\xi)|^p \, \mathrm{d}\xi \right)^{1-1/p} \left(\sum_{j=0}^{k-1} \int_{\Delta_k} |g'(\xi)|^p \, \mathrm{d}\xi \right)^{1/p} = \frac{p}{k}\|g\|_p^{p-1}\|g'\|_p.
\end{aligned}$$

Thus

$$\Big| \|g\|_p^p - \|g\|_{k,p}^p \Big| \le \frac{p}{k}\|g\|_p^{p-1}\|g'\|_p$$

which in turn implies

$$\Big| \|g\|_p - \|g\|_{k,p} \Big| \le \frac{p}{k}\|g'\|_p \ . \tag{12.2}$$

Let now $g := \alpha f + p_n \in S(f + \Pi_n)$. As $\|g\|_p = 1$, it follows, as above, that $|\alpha| \le C_p = [\min_{p\in\Pi_n} \|f - p\|_p]^{-1}$. Thus $\|p_n\|_p \le 1 + C_p\|f\|_p$ and using the L_p-Markov inequality, we obtain $\|p'_n\|_p \le c_p n^2(1 + C_p\|f\|_p)$. Combining this estimate with (12.2) yields

$$\eta_k(f + \Pi_n) \le pc_p\frac{n^2}{k}(1 + C_p\|f\|_p + C_p\|f'\|_p).$$

Thus, we can use the estimate

$$\eta_k(f + \Pi_n) = O(k^{-1}), \qquad 1 \le p < \infty, \tag{12.3}$$

in (11.5). Recall that by Corollary 7.2 and the subsequent Remark in the case of the L_p-norm, $1 < p < \infty$, we have the estimate $\phi^{-1}(t) = O(t^{\theta_p})$ where $\theta_p := \min\{1/2, 1/p\}$. Thus, we obtain by (11.5),

$$\|P_M^{(k)}(f) - P_M(f)\|_p = O(k^{-\theta_p}), \qquad 1 < p < \infty. \quad \square$$

Remark. When $p = 1$, Theorem 9.2 implies that for $f \in C^1[0,1]$, non-classical strong uniqueness holds with $\phi(t) = ct^2$. Thus repeated application of (12.3) and (11.5) yields

$$\|P_M^{(k)}(f) - P_M(f)\|_1 = O(k^{-1/2}).$$

When $p = 1$, it is shown in Kroó [1981b] that if $f \in C^2[0,1]$ and $f - P_M(f)$ has a finite number of zeros then the above estimate can be replaced by the sharp upper bound

$$\|P_M^{(k)}(f) - P_M(f)\|_1 = O(k^{-1}).$$

13. Asymptotic Representation of Weighted Chebyshev Polynomials

In this section, we shall use strong uniqueness results in order to solve approximation problems in the case where the norm is altered by a weight function. This approach will be used to derive asymptotic representations for weighted Chebyshev polynomials. Let $w > 0$ be a positive weight on $[-1, 1]$ and denote by

$$T_{n,p}(x, w) := x^n + q_{n-1}(x), \qquad q_{n-1} \in \Pi_{n-1},$$

the monic polynomial that deviates least from 0 in the weighted L_p-norm on $[-1, 1]$. That is,

$$\|T_{n,p}(x, w)w\|_p = \inf\{\|(x^n + g_{n-1}(x))w\|_p : g_{n-1} \in \Pi_{n-1}\}, \qquad 1 \le p \le \infty.$$

In addition, let $T^*_{n,p}(x, w) := T_{n,p}(x, w)/\|T_{n,p}(x, w)w\|_p$ be the normalized Chebyshev polynomial. The problem of finding asymptotic representations for weighted Chebyshev polynomials has been studied by many authors, but satisfactory solutions were given only in the case $p = 2$; see Bernstein [1930]. In addition, in the case $p = \infty$, Bernstein [1930] gave an asymptotic formula for $\|T_{n,\infty}(x, w)w\|_\infty$, Fekete, Walsh [1954/55] found the n-th root asymptotics of these Chebyshev polynomials, while in a more recent paper, Lubinsky, Saff [1987] gave their asymptotics outside of $[-1, 1]$. In this section, based on results of Kroó, Peherstorfer [2007], [2008], we shall outline a complete solution to this classical problem. This solution will be based on the strong uniqueness estimates derived in the previous sections. Another basic tool consists of the fact that for the specific weight $w = 1/\rho_m$, $\rho_m > 0$, $\rho_m \in \Pi_m$, $m < n$, an explicit formula for the minimal polynomials was already found by Chebyshev (see Akhiezer [1947]).

The case $p = \infty$**.** It is known (see Akhiezer [1947]) that

$$T^*_{n,\infty}(\cos\phi, 1/\rho_m) = \Re\{z^{-n} g_m^2(z)\}, \qquad z = \mathrm{e}^{\mathrm{i}\phi}, \tag{13.1}$$

where $\rho_m \in \Pi_m$ is a polynomial positive on $[-1, 1]$ and $g_m \in \Pi_m$ is the certain polynomial (unique up to a multiplicative constant of modulus 1) all of whose zeros lie in $|z| > 1$ and such that

$$|g_m(\mathrm{e}^{\mathrm{i}\phi})|^2 = \rho_m(\cos\phi), \qquad \phi \in [0, \pi].$$

Moreover, the corresponding L_1-Chebyshev polynomial is given by

$$T^*_{n-1,1}(\cos\phi, 1/\rho_m) = -\Im\{z^{-n} g_m^2(z)\}/\sin\phi, \qquad z = \mathrm{e}^{\mathrm{i}\phi}.$$

In addition, these polynomials satisfy the relation

$$\Re\{z^{-n} g_m^2(z)\}^2 + \Im\{z^{-n} g_m^2(z)\}^2 = \rho_m^2. \tag{13.2}$$

It follows immediately from (13.2) that the zeros of $\Im\{z^{-n} g_m^2(z)\}$ are the $n+1$ equioscillation points of $T^*_{n,\infty}(\cos\phi, 1/\rho_m)$. Moreover, in view of the alternation theorem (Theorem 4.4), (13.2) also implies that

$$T^*_{n-1,\infty}(\cos\phi, \sqrt{1-x^2}/\rho_m) = \Im\{z^{-n} g_m^2(z)\}/\sin\phi. \tag{13.3}$$

So a natural idea is to approximate a positive weight w by reciprocals of positive polynomials $1/\rho_m$ (the degree of approximation can be estimated by the Jackson Theorem), and then apply strong uniqueness type results in order to obtain the required asymptotics. Clearly, we shall need rather precise strong uniqueness type estimates that also take into account the dependence upon n.

We first recall that the positive continuous weight function $w(\cos\phi)$ can be represented in the form

$$w(\cos\phi) = \frac{1}{|\pi(\mathrm{e}^{\mathrm{i}\phi})|^2},$$

where $\pi(z)$, the so called Szegő function of w, is the function nonzero and analytic in $|z| < 1$ given by the formula

$$\pi(z) := \exp\left\{-\frac{1}{4\pi}\int_0^{2\pi} \frac{\mathrm{e}^{\mathrm{i}\phi} + z}{\mathrm{e}^{\mathrm{i}\phi} - z} \log w(\cos\phi)\, \mathrm{d}\phi\right\}.$$

Let us denote by $C^{k+\alpha}[-\pi,\pi]$ the class of 2π-periodic functions whose k-th derivative, $k \in \mathbb{N}$, satisfies the Lip α property. This next result provides an asymptotic formula for the weighted L_∞-Chebyshev polynomials for positive weights $w(\cos\phi) \in C^{2+\alpha}[-\pi,\pi]$.

Theorem 13.1. *Let $w(\cos\phi) \in C^{2+\alpha}[-\pi,\pi]$ with $0 < \alpha < 1$, $w(x) > 0$, $x \in [-1,1]$. Then*

$$T^*_{n,\infty}(\cos\phi, w) = \Re\{\mathrm{e}^{-\mathrm{i}n\phi}(\pi(\mathrm{e}^{\mathrm{i}\phi}))^2\} + O(n^{-\alpha})$$

uniformly for $\phi \in [0,\pi]$.

First, we shall need a lemma estimating the deviation between Chebyshev polynomials corresponding to different weights via the strong unicity constant.

Lemma 13.2. *Let $w_1, w_2 \in C[-1,1]$ be positive weight functions. Then*

$$\|T^*_{n,\infty}(\cdot, w_1)w_1 - T^*_{n,\infty}(\cdot, w_2)w_2\| \le \frac{c}{\gamma_{n-1}(T^*_{n,\infty}(w_1))}\|w_1 - w_2\|,$$

where $\gamma_{n-1}(T^*_{n,\infty}(w_1))$ is the strong unicity constant of $T^*_{n,\infty}(w_1)w_1$ with respect to the Haar space $w_1\Pi_{n-1}$, and $c > 0$ depends only on w_1, w_2.

Proof: We shall denote below by c_j constants depending only on w_1, w_2. Clearly,

$$T^*_{n,\infty}(\cdot, w_1) = a_n x^n - p_1 =: T^*_1, \quad T^*_{n,\infty}(\cdot, w_2) = b_n x^n - p_2 =: T^*_2, \qquad p_1, p_2 \in \Pi_{n-1}.$$

Denote by $T_1 := T_1^*/a_n, T_2 := T_2^*/b_n$ the corresponding monic polynomials. Then using the extremal property of T_1, we obtain

$$\frac{1}{a_n} = \|T_1 w_1\| \leq \|T_2 w_1\| \leq \left\| T_2 w_2 \frac{w_1}{w_2} \right\| \leq \frac{1}{b_n} \left\| \frac{w_1}{w_2} \right\|.$$

Replacing T_1 by T_2, we can obtain an analogous inequality

$$\frac{1}{b_n} \leq \frac{1}{a_n} \left\| \frac{w_2}{w_1} \right\|.$$

Thus combining these estimates, we have

$$1 - c_1 \|w_1 - w_2\| \leq \frac{1}{\left\| \frac{w_1}{w_2} \right\|} \leq \frac{a_n}{b_n} \leq \left\| \frac{w_2}{w_1} \right\| \leq c_2 \|w_1 - w_2\| + 1.$$

It therefore follows that

$$\left| \frac{a_n}{b_n} - 1 \right| \leq c_3 \|w_1 - w_2\|. \tag{13.4}$$

Set now

$$q := p_1 - \frac{a_n}{b_n} p_2 = T_1^* - \frac{a_n}{b_n} T_2^* \in \Pi_{n-1}. \tag{13.5}$$

Since 0 is the best approximant to $w_1 T_1^*$ in $w_1 \Pi_{n-1}$ in the uniform norm, it follows by the classical strong uniqueness inequality (applied to the Haar space $w_1 \Pi_{n-1}$) together with (13.4) and (13.5) that

$$\begin{aligned}
\gamma_{n-1}(T_1^*) \|q w_1\| \ &\leq \ \|(T_1^* - q) w_1\| - \|T_1^* w_1\| \ = \ \frac{a_n}{b_n} \|T_2^* w_1\| - \|T_1^* w_1\| \\
&= \ \frac{a_n}{b_n} \|T_2^* w_1\| - 1 \ \leq \ \frac{a_n}{b_n} \left\| \frac{w_1}{w_2} \right\| - 1 \ \leq \ \frac{a_n}{b_n} (1 + c_4 \|w_1 - w_2\|) - 1 \\
&\leq \ (1 + c_3 \|w_1 - w_2\|)(1 + c_4 \|w_1 - w_2\|) - 1 \ \leq \ c_5 \|w_1 - w_2\|,
\end{aligned} \tag{13.6}$$

where $\gamma_{n-1}(T_1^*)$ is the strong unicity constant of $T_1^* w_1$ with respect to the Haar space $w_1 \Pi_{n-1}$.

Finally, using (13.4) and (13.6), we arrive at

$$\begin{aligned}
\|T_1^* w_1 - T_2^* w_2\| \ &\leq \ \|q w_1\| + \|T_2^* (w_2 - \frac{a_n}{b_n} w_1)\| \\
&\leq \ \|q w_1\| + \|T_2^* (w_2 - w_1)\| + \|T_2^* w_1 \left(1 - \frac{a_n}{b_n} \right) \| \leq \frac{c_6}{\gamma_{n-1}(T_1^*)} \|w_1 - w_2\|,
\end{aligned}$$

which is the required estimate. $\square$

We now need to estimate the strong unicity constant $\gamma_{n-1}(T^*_{n,\infty}(w_1))$ (which is the strong unicity constant of $T^*_{n,\infty}(w_1)w_1$ with respect to the Haar space $w_1\Pi_{n-1}$) in the special case when $w_1 = 1/\rho_m$. Moreover, it is important in deriving asymptotic relations that the dependence of this quantity on n is revealed. We shall use Theorem 4.10 of Section 4 in order to provide a precise bound.

Lemma 13.3. *Let $\rho_m \in \Pi_m$, $m < n$, be such that $0 < A \le \rho_m \le B$ on $[-1, 1]$. Then*

$$\gamma_{n-1}(T^*_{n,\infty}(\cdot, 1/\rho_m)) \ge \frac{c}{n^2}, \tag{13.7}$$

where the constant c depends only on A and B.

Proof: Let $-1 = x_1 < \cdots < x_{n+1} = 1$ be the equioscillation points of $T^*_{n,\infty}(\cdot, 1/\rho_m)/\rho_m$ and $m_k \in \Pi_{n-1}$, $1 \le k \le n+1$, be the polynomials of Theorem 4.10 with respect to the Haar space Π_{n-1}/ρ_m, i.e.,

$$m_k(x_j)/\rho_m(x_j) = \operatorname{sgn} T^*_{n,\infty}(x_j, 1/\rho_m), \qquad 1 \le j \le n+1, \quad j \ne k.$$

By the second statement of Theorem 4.10,

$$\gamma_{n-1}(T^*_{n,\infty}(\cdot, 1/\rho_m)) = \min_{1\le k\le n+1} \frac{1}{\|m_k/\rho_m\|}. \tag{13.8}$$

Set now for any $1 \le k \le n+1$,

$$Q_k(x) := (x - x_k)T^*_{n,\infty}(x, 1/\rho_m) + (1 - x^2)T^*_{n-1,\infty}(x, \sqrt{1-x^2}/\rho_m).$$

It follows by (13.2) and (13.3) that the equioscillation points of $T^*_{n,\infty}(x, 1/\rho_m)/\rho_m$ are the zeros of $(1 - x^2)T^*_{n-1,\infty}(x, \sqrt{1-x^2}/\rho_m)$. Moreover, since $m < n$, the relation (13.2) also implies that the leading coefficients of $T^*_{n,\infty}(x, 1/\rho_m)$ and $T^*_{n-1,\infty}(x, \sqrt{1-x^2}/\rho_m)$ coincide. It therefore follows that $Q_k \in \Pi_n$, $Q_k(x_k) = 0$, $1 \le k \le n+1$, and

$$m_k(x) = \frac{Q_k(x)}{x - x_k}, \qquad 1 \le k \le n+1.$$

Obviously, $\|Q_k\| \le c$, $1 \le k \le n+1$, with a constant depending only on A, B. This easily implies, by the Markov inequality, that $\|m_k\| \le c_1 n^2$, $1 \le k \le n+1$. Applying this inequality together with (13.8) yields (13.7). □

Proof of Theorem 13.1: When $0 < \alpha < 1$, the smoothness condition imposed on the weight w implies that its Szegő function $\pi(z)$ is also $C^{2+\alpha}$ on $|z| = 1$ (see Kroó, Peherstorfer [2008] for details). Therefore there exists a sequence of polynomials $g_m \in \Pi_m$ such that uniformly on $|z| \le 1$

$$|\pi(z) - g_m(z)| \le \frac{c}{m^{2+\alpha}}. \tag{13.9}$$

Since $\pi(z) \ne 0$ on $|z| \le 1$, it follows that for $m \ge m_0$ the function g_m also does not vanish in $|z| \le 1$ and

$$\rho_m(\cos\phi) := |g_m(\mathrm{e}^{\mathrm{i}\phi})|^2 \ge c > 0. \tag{13.10}$$

Furthermore, by Chebyshev's result (see (13.1)),

$$T^*_{n,\infty}(\cos\phi, 1/\rho_m) = \Re\{z^{-n} g_m^2(z)\}, \qquad z = \mathrm{e}^{\mathrm{i}\phi}. \tag{13.11}$$

We are now ready to verify the statement of the theorem. Indeed, since

$$w(\cos\phi) = \frac{1}{|\pi(\mathrm{e}^{\mathrm{i}\phi})|^2},$$

we have by (13.9) and (13.10) that

$$\left| w(x) - \frac{1}{\rho_m(x)} \right| \le \frac{c_1}{m^{2+\alpha}}, \qquad x \in [-1,1]. \tag{13.12}$$

Thus using Lemmas 13.2 and 13.3 together with (13.9), (13.10) and (13.12), see also (13.11), we obtain

$$\|T^*_{n,\infty}(\cos\phi, w) - \Re\{\mathrm{e}^{-in\phi}(\pi(\mathrm{e}^{\mathrm{i}\phi}))^2\}\| \le \|T^*_{n,\infty}(\cos\phi, w) - T^*_{n,\infty}(\cos\phi, 1/\rho_m)\| +$$

$$\|T^*_{n,\infty}(\cos\phi, 1/\rho_m) - \Re\{\mathrm{e}^{-in\phi}(\pi(\mathrm{e}^{\mathrm{i}\phi}))^2\}\| \le \frac{c_2 n^2}{m^{2+\alpha}}, \qquad x = \cos\phi.$$

Putting in the last estimate $m := \lfloor n/2 \rfloor$ yields the statement of the theorem. □

The case $1 \le p < \infty$**.** Similarly to the case $p = \infty$, the following explicit representation for weighted Chebyshev polynomials is given in Akhiezer [1947] for the weight function $\rho_{m,p} := \frac{1}{\sqrt{1-x^2}\rho_m^{p/2}}$, where $\rho_m \in \Pi_m$ is positive on $[-1,1]$,

$$T^*_{n,p}(\cos\phi, \rho_{m,p}) = \lambda_p \Re\{z^{-n} g_m(z)\}, \qquad z = \mathrm{e}^{\mathrm{i}\phi},$$

where λ_p is a constant depending only on p and, as above, g_m is related to ρ_m by

$$|g_m(\mathrm{e}^{\mathrm{i}\phi})|^2 = \rho_m(\cos\phi), \qquad \phi \in [0,\pi].$$

Let now w be a positive weight and set $w_p := \frac{w^{p/2}}{\sqrt{1-x^2}}$, $1 \le p < \infty$.

The next theorem provides an asymptotic representation for L_p-Chebyshev polynomials when $1 < p < \infty$.

Theorem 13.4. *Let* $1 < p < \infty$*. Then for any positive weight* w *such that* $w \in C^\alpha[-\pi,\pi]$*,* $0 < \alpha < 1$*,*

$$T^*_{n,p}(\cos\phi, w_p) = \lambda_p \Re\{z^{-n}\pi(z)\} + O(n^{-\alpha\theta_p}), \qquad z = \mathrm{e}^{\mathrm{i}\phi},$$

where $\pi(z)$ *is the Szegő function of* w*,* $\theta_p := \min\{1/2, 1/p\}$*, and the* $O(\cdot)$ *is taken with respect to the* L_p*-norm.*

Since the main idea in the proof of Theorem 13.4 is similar to the proof of Theorem 13.1, we shall just briefly outline the proof. First, it is shown that, similarly to Lemma 13.2, under suitable conditions on the weight functions $w_1, w_2 > A > 0$, we have

$$\|T^*_{n,p}(\cos\phi, w_1) - T^*_{n,p}(\cos\phi, w_2)\|_p \le c_{A,p}\|w_1 - w_2\|_p^{\theta_p}.$$

It should be noted that instead of classical strong uniqueness, we use here the fact that in L_p-approximation with $1 < p < \infty$, non-classical strong uniqueness with $\phi^{-1}(t) = O(t^{\theta_p})$ is satisfied, and the constants involved depend only on p.

Finally, similar asymptotic relations can also be given when $p = 1$. In this case, non-classical strong uniqueness with $\phi^{-1}(t) = O(t^{1/2})$ takes place (this follows from Theorem 9.2 applied in the case of Lip 1 functions). In addition, similarly to Lemma 13.3, one has to study how the constants involved depend on n. See Kroó, Peherstorfer [2008] for details.

References

N. I. Achieser [1930], On extremal properties of certain rational functions, *Doklady Akad. Nauk*, 495–499.

N. I. Achieser [1947], *Lectures on the Theory of Approximation*, OGIZ, Moscow-Leningrad, (in Russian) 1947, reprinted in English as *Theory of Approximation*, Frederick Ungar Publishing Co., New York, 1956, and reissued by Dover Publications in 1992.

J. Angelos and A. Egger [1984], Strong uniqueness in L^p spaces, *J. Approx. Theory* **42**, 14–26.

J. R. Angelos, M. S. Henry, E. H. Kaufman, Jr., and T. D. Lenker [1985], Local Lipschitz constants, *J. Approx. Theory* **43**, 53–63.

J. R. Angelos, M. S. Henry, E. H. Kaufman, A. Kroó and T. D. Lenker [1986a], Local Lipschitz and strong unicity constants for certain nonlinear families, in *Approximation Theory V*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 239–242, Academic Press, New York.

J. R. Angelos, M. S. Henry, E. H. Kaufman, A. Kroó and T. D. Lenker [1986b], Local and global Lipschitz constants, *J. Approx. Theory* **46**, 137–156.

J. R. Angelos, M. S. Henry, E. H. Kaufman and T. D. Lenker [1986], Extended Lipschitz constants, in *Approximation Theory V*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 243–246, Academic Press, New York.

J. R. Angelos, M. S. Henry, E. H. Kaufman, Jr., and T. D. Lenker [1988], Bounds for extended local Lipschitz constants, in *Constructive Theory of Functions* (Varna, 1987), 17–26, Publ. House Bulgar. Acad. Sci., Sofia.

J. R. Angelos, M. S. Henry, E. H. Kaufman, Jr., T. D. Lenker and A. Kroó [1989], Local Lipschitz and strong unicity constants for certain nonlinear families, *J. Approx. Theory* **58**, 164–183.

J. Angelos, E. Kaufman, Jr., T. Lenker and M. S. Henry [1991], Bounds for extended Lipschitz constants, *Acta Math. Hungar.* **58**, 81–93.

J. R. Angelos and A. Kroó [1986], The equivalence of the moduli of continuity of the best approximation operator and of strong unicity in L^1, *J. Approx. Theory* **46**, 129–136.

J. Angelos and D. Schmidt [1983], Strong uniqueness in $L^1(X, \Sigma, \mu)$, in *Approximation Theory IV*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 297–302, Academic Press, New York.

J. R. Angelos and D. Schmidt [1988], The prevalence of strong uniqueness in L^1, *Acta Math. Hungar.* **52**, 83–90.

R. B. Barrar and H. L. Loeb [1970], On the continuity of the nonlinear Tschebyscheff operator, *Pacific J. Math.* **32**, 593–601.

R. B. Barrar and H. L. Loeb [1986], The strong uniqueness theorem for monosplines, *J. Approx. Theory* **46**, 157–169.

M. W. Bartelt [1974], Strongly unique best approximates to a function on a set, and a finite subset thereof, *Pacific J. Math.* **53**, 1–9.

M. Bartelt [1975], On Lipschitz conditions, strong unicity and a theorem of A. K. Cline, *J. Approx. Theory* **14**, 245–250.

M. Bartelt [2001], Hausdorff strong unicity in vector-valued Chebyshev approximation on finite sets, in *Trends in Approximation Theory* (Nashville, TN, 2000), K. Kopotun, T. Lyche, M. Neamtu (eds.), 31–38, Innov. Appl. Math., Vanderbilt Univ. Press, Nashville, TN.

M. W. Bartlet and M. S. Henry [1980], Continuity of the strong uniqueness constant on $C[X]$ for changing X, *J. Approx. Theory* **28**, 87–95.

M. Bartelt, E. H. Kaufman, Jr., and J. Swetits [1990], Uniform Lipschitz constants in Chebyshev polynomial approximation, *J. Approx. Theory* **62**, 23–38.

M. W. Bartlet, A. Kroó and J. J. Swetits [1989], Local Lipschitz constants: characterization and uniformity, in *Approximation Theory VI, Volume I*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 65–68, Academic Press, New York.

M. Bartelt and W. Li [1995a], Error estimates and Lipschitz constants for best approximation in continuous function spaces, *Comput. Math. Appl.* **30**, 255–268.

M. Bartelt and W. Li [1995b], Haar theory in vector-valued continuous function spaces, in *Approximation Theory VIII, Volume I*, (College Station, TX, 1995), C. K. Chui, L. L. Schumaker (eds.), 39–46, Ser. Approx. Decompos., 6, World Sci. Publ., NJ, 1995.

M. W. Bartlet and H. W. McLaughlin [1973], Characterizations of strong unicity in approximation theory, *J. Approx. Theory* **9**, 255–266.

M. W. Bartelt and D. Schmidt [1980], On strong unicity and a conjecture of Henry and Roulier, in *Approximation Theory III*, E. W. Cheney, Ed., 187–191, Academic Press, New York.

M. W. Bartelt and D. Schmidt [1981], On Poreda's problem on the strong unicity constants, *J. Approx. Theory* **33**, 69–79.

M. W. Bartelt and D. Schmidt [1984], Lipschitz conditions, strong uniqueness, and almost Chebyshev subspaces of $C(X)$, *J. Approx. Theory* **40**, 202–215.

M. Bartelt and J. Swetits [1983], Uniform strong unicity constants for subsets of $C[a,b]$, in *Approximation Theory IV*, (College Station, Tex., 1983), C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 329–334, Academic Press, New York.

M. Bartelt and J. Swetits [1986], Uniform Lipschitz constants and almost alternation sets, in *Approximation Theory V*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 247–250, Academic Press, New York.

M. Bartelt and J. Swetits [1988], Uniform strong unicity constants for subsets of $C(X)$, *J. Approx. Theory* **55**, 304–317.

M. W. Bartelt and J. J. Swetits [1991a], Local Lipschitz constants and Kolushov polynomials, *Acta Math. Hungar.* **57**, 259–263.

M. W. Bartelt and J. J. Swetits [1991b], New classes of local Lipschitz constants for the best approximation operator on finite sets, in *Approximation Theory* (Kecskemét, 1990), 69–83, Colloq. Math. Soc. János Bolyai, **58**, North-Holland, Amsterdam.

M. W. Bartelt and J. J. Swetits [1993], The strong derivative of the best approximation operator, *Numer. Funct. Anal. Optim.* **14**, 229–248.

M. W. Bartelt and J. J. Swetits [1995], Continuity properties of Lipschitz constants for the best approximation operator, in *Approximation and Optimization in the Caribbean, II* (Havana, 1993), 51–62, Approx. Optim., **8**, Lang, Frankfurt am Main.

M. W. Bartelt and J. J. Swetits [2002], Uniform strong unicity of order 2 for generalized Haar sets, in *Approximation Theory X, Abstract and Classical Analysis* (St. Louis, MO, 2001), C. K. Chui, L. L. Schumaker, J. Stöckler, Eds., 23–30, Vanderbilt Univ. Press, Nashville, TN.

M. Bartelt and J. Swetits [2007], Lipschitz continuity and Gateaux differentiability of the best approximation operator in vector-valued Chebyshev approximation, *J. Approx. Theory* **148**, 177–193.

M. Bartelt and J. Swetits [2008], Lipschitz continuity of the best approximation operator in vector-valued Chebyshev approximation *J. Approx. Theory* **152**, 161–166.

S. N. Bernstein [1930], Sur les polynomes orthogonaux relatifs á un segment fini, *J. Math.* **9**, 127–177.

B. O. Björnestål [1975], Continuity of the metric projection operator II, TRITA-MAT-1975-12, Preprint Series of the Royal Institute of Technology, Stockholm, Sweden.

B. O. Björnestål [1979], Local Lipschitz continuity of the metric projection operator, in *Approximation Theory* (Papers, VIth Semester, Stefan Banach Internat. Math. Center, Warsaw, 1975), pp. 43–53, Banach Center Publ., **4**, PWN, Warsaw.

H. P. Blatt [1984a], On strong uniqueness in linear complex Chebyshev approximation, *J. Approx. Theory* **41**, 159–169.

H. P. Blatt [1984b], Exchange algorithms, error estimations and strong unicity in convex programming and Chebyshev approximation, in *Approximation Theory and Spline Functions*, St. John's, Nfld, 1983, NATO Adv. Sci. Inst. Sec. C Math. Phys. Sci., **136**, 23–63, Reidel, Dordrecht.

H. P. Blatt [1985], Characterization of strong unicity in semi-infinite optimization by chain of references, in *Parametric Optimization and Approximation* (Oberwolfach, 1983) ISNM **72**, 36–46, Birkhäuser.

H. P. Blatt [1986], Lipschitz continuity and strong unicity in G. Freud's work, *J. Approx. Theory* **46**, 25–31.

H. P. Blatt [1990], On the distribution of the points of Chebyshev alternation with applications to strong unicity constants, *Acta Math. Hungar.* **55**, 75–82.

E. Borel [1905], *Leçons sur les Fonctions de Variables Réelles*, Gauthier-Villars, Paris.

A. P. Bosznay [1988], A remark concerning strong uniqueness of approximation, *Studia Sci. Math. Hungar.* **23**, 85–87.

D. Braess [1973], Kritische Punkte bei der nichtlinearen Tschebyscheff-Approximation, *Math. Z.* **132**, 327–341.

D. Braess [1986], *Nonlinear Approximation Theory*, Springer-Verlag, Berlin Heidelberg.

B. Brosowski [1983], A refinement of the Kolmogorov-criterion, in *Constructive Function Theory '81* (Varna, 1981), 241–247, Bulgar. Acad. Sci., Sofia.

B. Brosowski and C. Guerreiro [1986], An extension of strong uniqueness to rational approximation, *J. Approx. Theory* **46**, 345–373.

B. Brosowski and C. Guerreiro [1987], On the uniqueness and strong uniqueness of best rational Chebyshev-approximation, *Approx. Theory Appl.* **3**, 49–70.

B. L. Chalmers, F. T. Metcalf and G. D. Taylor [1983], Strong unicity of arbitrary rate, *J. Approx. Theory* **37**, 238–245.

B. L. Chalmers and G. D. Taylor [1980], On the existence of strong unicity of arbitrarily small order, in *Approximation Theory III*, E. W. Cheney, Ed., 293–298, Academic Press, New York.

B. L. Chalmers and G. D. Taylor [1983], A unified theory of strong uniqueness in uniform approximation with constraints, *J. Approx. Theory* **37**, 29–43.

K. Y. Chan, Y. M. Chen, M. C. Liu and S. M. Ng [1982], An example on strong unicity constants in trigonometric approximation, *Proc. Amer. Math. Soc.* **84**, 79–84.

E. W. Cheney [1965], Approximation by generalized rational functions, in *Approximation of Functions*, H. L. Garabedian, Ed., 101–110, Elsevier, Amsterdam.

E. W. Cheney [1966], *Introduction to Approximation Theory*, McGraw-Hill Book Co., New York-Toronto.

E. W. Cheney and H. L. Loeb [1964], Generalized rational approximation, *J. SIAM Numer. Anal. Ser. B* **1**, 11–25.

E. W. Cheney and D. E. Wulbert [1969], The existence and unicity of best approximations, *Math. Scand.* **24**, 113–140.

J. A. Clarkson [1936], Uniformly convex spaces, *Trans. Amer. Math. Soc.* **40**, 396–414.

A. K. Cline [1973], Lipschitz conditions on uniform approximation operators, *J. Approx. Theory* **8**, 160–172.

L. Cromme [1977/78], Strong uniqueness: a far reaching criterion for the convergence analysis of iterative procedures, *Numer. Math.* **29**, 179–193.

L. Danzer, B. Grünbaum and V. Klee [1963], Helly's theorem and its relatives, in *Convexity, Proc. Symposia in Pure Mathematics, Volume VII*, 101–180, AMS.

F. Deutsch and W. Li [1991], Strong uniqueness, Lipschitz continuity, and continuous selections for metric projections in L_1, *J. Approx. Theory* **66**, 198–224.

F. Deutsch and S. Mabizela [1996], Best interpolatory approximation in normed linear spaces, *J. Approx. Theory* **85**, 250–268.

C. B. Dunham [1980], A uniform constant of strong uniqueness on an interval, *J. Approx. Theory* **28**, 207–211.

C. B. Dunham [1989], Local strong uniqueness for nonlinear approximation, *Approx. Theory Appl.* **5**, 43–45.

C. B. Dunham [1995], Uniform local strong uniqueness on finite subsets, *Aequationes Math.* **49**, 295–299.

A. G. Egger and G. D. Taylor [1983], Strong uniqueness in convex L^p approximation, in *Approximation Theory IV*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 451–456, Academic Press, New York.

A. Egger and G. D. Taylor [1989a], A survey of local and directional local strong uniqueness, in *Approximation Theory VI*, Vol. I, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 239–242, Academic Press, Boston, MA.

A. G. Egger and G. D. Taylor [1989b], Local strong uniqueness, *J. Approx. Theory* **58**, 267–280.

N. H. Eggert and J. R. Lund [1984], Examples of functions whose sequence of strong unicity constants is unbounded, *J. Approx. Theory* **41**, 244–252.

M. Fang [1990], The Chebyshev theory of a variation of L_p $(1 < p < \infty)$ approximation, *J. Approx. Theory* **62**, 94–109.

M. Fekete and J. L. Walsh [1954/55], On the asymptotic behaviour of polynomials with extremal properties, and of their zeros, *J. d'Analyse Math.* **4**, 49–87.

T. Fischer [1989], Strong unicity and alternation in linear approximation and a continuous alternator, in *Approximation Theory VI*, Vol. I, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 255–258, Academic Press, New York.

T. Fischer [1990], Strong unicity in normed linear spaces, *Numer. Funct. Anal. Optim.* **11**, 255–266.

Y. Fletcher and J. A. Roulier [1979], A counterexample to strong unicity in monotone approximation, *J. Approx. Theory* **27**, 19–33.

G. Freud [1958], Eine Ungleichung für Tschebyscheffsche Approximationspolynome, *Acta Sci. Math. Szeged* **19**, 162–164.

A. L. Garkavi [1959], Dimensionality of polyhedra of best approximation for differentiable functions, *Izv. Akad. Nauk SSSR Ser. Mat.* **23**, 93–114. (Russian)

A. L. Garkavi [1964], On Čebyšev and almost Čebyšev subspaces, *Izv. Akad. Nauk SSSR Ser. Mat.* **28**, 799–818; in English translation in *Amer. Math. Soc. Transl.* **96**, (1970) 153–175.

A. L. Garkavi [1965], Almost Čebyšev subspaces of continuous functions, *Izv. Vysš. Učebn. Zaved. Matematika* **1965**, 36–44; in English translation in *Amer. Math. Soc. Transl.* **96**, (1970) 177–187.

W. Gehlen [1999], On a conjecture concerning strong unicity constants, *J. Approx. Theory* **101**, 221–239.

W. Gehlen [2000], Unboundedness of the Lipschitz constants of best polynomial approximation, *J. Approx. Theory* **106**, 110–142.

R. Grothmann [1988a], On the real CF-method for polynomial approximation and strong unicity constants, *J. Approx. Theory* **55**, 86–103.

R. Grothmann [1988b], Local uniqueness in nonuniqueness spaces, *Approx. Theory Appl.* **4**, 35–39.

R. Grothmann [1989], A note on strong uniqueness constants, *J. Approx. Theory* **58**, 358–360.

R. Grothmann [1999], On the problem of Poreda, in *Computational Methods and Function Theory 1997* (Nicosia), N. Papamichael, St. Ruscheweyh and E. B. Saff, Eds., 267–273, Ser. Approx. Decompos., **11**, World Sci. Publ., NJ.

M. Gutknecht [1978], Non-strong uniqueness in real and complex Chebyshev approximation, *J. Approx. Theory* **23**, 204–213.

A. Haar [1918], Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, *Math. Ann.* **78**, 294–311.

O. Hanner [1956], On the uniform convexity of L^p and ℓ^p, *Ark. Mat.* **3**, 239–244.

S. Ja. Havinson and Z. S. Romanova [1972], Approximation properties of finite-dimensional subspaces in L_1, *Mat. Sb.* **89**, 3–15, in English translation in *Math. USSR Sb.* **18**, 1–14.

M. S. Henry [1987], Lipschitz and strong unicity constants, in *A. Haar Memorial Conference*, Vol. I, (Budapest, 1985), 423–444, Colloq. Math. Soc. János Bolyai, **49**, North-Holland, Amsterdam.

M. S. Henry and L. R. Huff [1979], On the behavior of the strong unicity constant for changing dimension, *J. Approx. Theory* **27**, 278–290.

M. S. Henry, E. H. Kaufman, Jr., and T. D. Lenker [1983], Lipschitz constants for small perturbations, in *Approximation Theory IV*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 515–520, Academic Press, New York.

M. S. Henry, E. H. Kaufman, Jr., and T. D. Lenker [1983], Lipschitz constants on sets with small cardinality, in *Approximation Theory IV*, C. K. Chui, L. L. Schumaker, J. D. Ward, Eds., 521–526, Academic Press, New York.

M. S. Henry and J. A. Roulier [1977], Uniform Lipschitz constants on small intervals, *J. Approx. Theory* **21**, 224–235.

M. S. Henry and J. A. Roulier [1978], Lipschitz and strong unicity constants for changing dimension, *J. Approx. Theory* **22**, 85–94.

M. S. Henry, D. P. Schmidt and J. J. Swetits [1981], Uniform strong unicity for rational approximation, *J. Approx. Theory* **33**, 131–146.

M. S. Henry and J. J. Swetits [1980a], Growth rates for strong unicity constants, in *Approximation Theory III*, E. W. Cheney, Ed., 501–505, Academic Press, New York.

M. S. Henry and J. J. Swetits [1980b], Lebesgue and strong unicity constants, in *Approximation Theory III*, E. W. Cheney, Ed., 507–512, Academic Press, New York.

M. S. Henry and J. J. Swetits [1981], Precise orders of strong unicity constants for a class of rational functions, *J. Approx. Theory* **32**, 292–305.

M. S. Henry and J. J. Swetits [1982], Lebesgue and strong unicity constants for Zolotareff polynomials, *Rocky Mountain J. Math.* **12**, 547–556.

M. S. Henry and J. J. Swetits [1984], Limits of strong unicity constants for certain C^∞ functions, *Acta. Math. Hungar.* **43**, 309–323.

M. S. Henry, J. J. Swetits, and S. E. Weinstein [1980], Lebesgue and strong unicity constants, in *Approximation Theory III*, E. W. Cheney, Ed., 507–512, Academic Press, New York.

M. S. Henry, J. J. Swetits, and S. E. Weinstein [1981], Orders of strong unicity constants, *J. Approx. Theory* **31**, 175–187.

M. S. Henry, J. J. Swetits, and S. E. Weinstein [1983], On extremal sets and strong unicity constants for certain C^∞ functions, *J. Approx. Theory* **37**, 155–174.

R. Holmes and B. Kripke [1968], Smoothness of approximation, *Mich. Math. J.* **15**, 225–248.

R. Huotari and S. Sahab [1994], Strong unicity versus modulus of convexity, *Bull. Austral. Math. Soc.* **49**, 305–310.

R. C. James [1947], Orthogonality and linear functionals in normed linear spaces, *Trans. Amer. Math. Soc.* **61**, 265–292.

K. Jittorntrum and M. R. Osborne [1980], Strong uniqueness and second order convergence in nonlinear discrete approximation, *Numer. Math.* **34**, 439–455.

P. Kirchberger, [1902] *Über Tchebychefsche Annäherungsmethoden*, Dissertation. Univ. Göttingen.

A. V. Kolushov [1981], Differentiability of the operator of best approximation, *Math. Notes Acad. Sci. USSR* **29**, 295–306.

V. V. Kovtunec [1984], The Lipschitz property of the operator of best approximation of complex-valued functions, in *The Theory of Approximation of Functions and its Applications*, 80–86, **137**, Akad. Nauk Ukrain. SSR, Inst. Mat., Kiev.

B. Kripke [1964], Best approximation with respect to nearby norms, *Numer. Math.* **6**, 103–105.

A. Kroó [1977a], The continuity of best approximations, *Acta. Math. Hungar.* **30**, 175–188.

A. Kroó [1977b], Differential properties of the operator of best approximation, *Acta. Math. Hungar.* **30**, 319–331.

A. Kroó [1978], On the continuity of best approximation in the space of integrable functions, *Acta. Math. Hungar.* **32**, 331–348.

A. Kroó [1980], The Lipschitz constant of the operator of best approximation, *Acta. Math. Hungar.* **35**, 279–292.

A. Kroó [1981a], On strong unicity of L_1-approximation, *Proc. Amer. Math. Soc.* **83**, 725–729.

A. Kroó [1981b], Best L_1-approximation on finite point sets: rate of convergence, *J. Approx. Theory* **33**, 340–352.

A. Kroó [1981/82], On strong unicity of best approximation in $C(R)$, *Numer. Funct. Anal. Optim.* **4**, 437–443.

A. Kroó [1983a], On the strong unicity of best Chebyshev approximation of differentiable functions, *Proc. Amer. Math. Soc.* **89**, 611–617.

A. Kroó [1983b], On unicity and strong unicity of best approximation in the L_1-norm, *Constructive Function Theory '81*, (Varna, 1981), 396–399, Bulgar. Acad. Sci., Sofia.

A. Kroó [1984], On the unicity of best Chebyshev approximation of differentiable functions, *Acta Sci. Math. (Szeged)* **47**, 377–389.

A. Kroó and F. Peherstorfer [2007], Asymptotic representation of L_p-minimal polynomials, $1 < p < \infty$, *Constr. Approx.* **25**, 29–39.

A. Kroó and F. Peherstorfer [2008], Asymptotic representation of weighted L_∞- and L_1-minimal polynomials, *Math. Proc. Camb. Phil. Soc.* **144**, 241–254.

A. Kroó and D. Schmidt [1991], A Haar-type theory of best uniform approximation with constraints, *Acta Math. Hung.* **58**, 351–374.

A. Kroó, M. Sommer, and H. Strauss [1989], On strong uniqueness in one-sided L^1-approximation of differentiable functions, *Proc. Amer. Math. Soc.* **106**, 1011–1016.

K. Kuratowski [1966], *Topology*, Academic Press, New York.

P.-J. Laurent and D. Pai, On simultaneous approximation, *Numer. Funct. Anal. Optim.* **19**, 1045–1064.

C. Li [2003], On best uniform restricted range approximation in complex-valued continuous function spaces, *J. Approx. Theory* **120**, 71–84.

C. Li and G. A. Watson [1994], On approximation using a peak norm, *J. Approx. Theory* **77**, 266–275.

C. Li and G. A. Watson [1997], Strong uniqueness in restricted rational approximation, *J. Approx. Theory* **89**, 96–113.

C. Li and G. A. Watson [1999], On approximation using a generalized peak norm, *Commun. Appl. Anal.* **3**, 357–371.

W. Li [1989], Strong uniqueness and Lipschitz continuity of metric projections: a generalization of the classical Haar theory, *J. Approx. Theory* **56**, 164–184.

P. K. Lin [1989], Strongly unique best approximation in uniformly convex Banach spaces, *J. Approx. Theory* **56**, 101–107.

H. L. Loeb [1966], Approximation by generalized rationals, *J. SIAM Numer. Anal.* **3**, 34–55.

D. S. Lubinsky and E. B. Saff [1987], Strong asymptotics for L_p-minimal polynomials, $1 < p < \infty$, in *Approximation Theory, Tampa*, E. B. Saff, Ed., LNM **1287**, 83–104.

Z. Ma [1991], Some problems on a variation of L_1 approximation, in *Progress in Approximation Theory*, P. Nevai, A. Pinkus, Eds., 667–692, Academic Press, Boston.

Z. W. Ma and Y. G. Shi [1990], A variation of rational L_1 approximation, *J. Approx. Theory* **62**, 262–273.

H. Maehly and Ch. Witzgall [1960], Tschebyscheff-Approximationen in kleinen Intervallen. I. Approximation durch Polynome, *Numer. Math.* **2**, 142–150.

P. F. Mah [1984], Strong uniqueness in nonlinear approximation, *J. Approx. Theory* **41**, 91–99.

L. K. Malozëmova [1990], On the strong uniqueness constant in a discrete problem of best approximation in the mean, *Vestnik Leningrad Univ. Mat. Mekh. Astronom.* **125**, 26–30; in English translation in *Vestnik Leningrad Univ. Math.* **23**, 33–37.

A. V. Marinov [1983], Strong uniqueness constants for best uniform approximations on compacta, *Mat. Zametki* **34**, 31–46; in English translation in *Math. Notes* **34**, 499–507.

O. M. Martinov [2002], Constants of strong unicity of minimal projections onto some two-dimensional subspaces of $l_\infty^{(4)}$, *J. Approx. Theory* **118**, 175–187.

H. W. McLaughlin and K. B. Somers [1975], Another characterization of Haar subspaces, *J. Approx. Theory* **14**, 93–102.

A. Meir [1984], On the uniform convexity of L^p spaces, $1 < p \le 2$, *Illinois J. Math.* **28**, 420–424.

V. N. Nikolskii [1982], Boundary sets in the strong sense and strong uniqueness of elements of best approximation, *Application of Functional Analysis in Approximation Theory*, 126–130, Kalinin. Gos. Univ., Kalinin, (in Russian).

D. J. Newman and H. S. Shapiro [1963], Some theorems on Cebysev approximation, *Duke Math. J.* **30**, 673–682. Abstract appeared in Notices in 1962 **9**, 143.

G. Nürnberger [1979], Unicity and strong unicity in approximation theory, *J. Approx. Theory* **26**, 54–70.

G. Nürnberger [1980], Strong uniqueness of best approximations and weak Chebyshev systems, in *Quantitative Approximation* (Proc. Internat. Sympos., Bonn, 1979), pp. 255–266, Academic Press, New York-London.

G. Nürnberger [1982], A local version of Haar's theorem in approximation theory, *Numer. Funct. Anal. Optim.* **5**, 21–46.

G. Nürnberger [1982/83], Strong unicity constants for spline functions, *Numer. Funct. Anal. Optim.* **5**, 319–347.

G. Nürnberger [1983], Strong unicity constants for finite-dimensional subspaces, in *Approximation Theory IV*, C. K.Chui, L. L. Schumaker, J. D. Ward, Eds., 643–648, Academic Press, New York.

G. Nürnberger [1984], Strong unicity of best approximations: a numerical aspect, *Numer. Funct. Anal. Optim.* **6** (1983), 399–421.

G. Nürnberger [1985a], Unicity in one-sided L_1-approximation and quadrature formulae, *J. Approx. Theory* **45**, 271–279.

G. Nürnberger [1985b], Global unicity in semi-infinite optimization, *Numer. Funct. Anal. Optim.* **8**, 173–191.

G. Nürnberger [1987], Strong unicity constants in Chebyshev approximation, in *Numerical Methods of Approximation Theory*, Vol. 8 (Oberwolfach, 1986), 144–154, ISNM, **81**, Birkhäuser, Basel.

G. Nürnberger [1994], Strong unicity in nonlinear approximation and free knot splines, *Constr. Approx.* **10**, 285–299.

G. Nürnberger and I. Singer [1982], Uniqueness and strong uniqueness of best approximations by spline subspaces and other subspaces, *J. Math. Anal. Appl.* **90**, 171–184.

P. L. Papini [1978], Approximation and strong approximation in normed spaces via tangent functionals, *J. Approx. Theory* **22**, 111–118.

S. H. Park [1989], Uniform Hausdorff strong uniqueness, *J. Approx. Theory* **58**, 78–89.

S. O. Paur and J. A. Roulier [1980], Uniform Lipschitz and strong unicity constants on subintervals, in *Approximation Theory III*, E. W. Cheney, Ed., 715–720, Academic Press, New York.

S. O. Paur and J. A. Roulier [1981], Continuity and strong unicity of the best approximation operator on subintervals, *J. Approx. Theory* **32**, 247–255.

J. Peetre [1970], Approximation of norms, *J. Approx. Theory* **3**, 243–260.

R. R. Phelps [1966], Cebysev subspaces of finite dimension in L_1, *Proc. Amer. Math. Soc.* **17**, 646–652.

A. M. Pinkus [1989], *On L^1-Approximation*, Cambridge University Press, Cambridge.

A. Pinkus and H. Strauss [1987], One-sided L^1-approximation to differentiable functions, *Approx. Theory Appl.* **3**, 81–96.

S. J. Poreda [1976], Counterexamples in best approximation, *Proc. Amer. Math. Soc.* **56**, 167–171.

B. Prus and R. Smarzewski [1987], Strongly unique best approximation and centers in uniformly convex spaces, *J. Math. Anal. Appl.* **121**, 10–21.

T. J. Rivlin [1969], *An Introduction to the Approximation of Functions*, Blaisdell, Waltham, Mass.

T. J. Rivlin [1984a], The strong uniqueness constant in complex approximation, in *Rational Approximation and Interpolation*, Tampa, 1983, P. R. Graves-Morris, E. B. Saff, R. S. Varga, Eds., 145–149, LNM **1105**, Springer.

T. J. Rivlin [1984b], The best strong uniqueness constant for a multivariate Chebyshev polynomial, *J. Approx. Theory* **41**, 56–63.

E. Rozema [1974], Almost Chebyshev subspaces of $L^1(\mu; E)$, *Pacific J. Math.* **53**, 585–604.

R. Schaback [1978], On alternation numbers in nonlinear Chebyshev approximation, *J. Approx. Theory* **23**, 379–391.

D. P. Schmidt [1978], On an unboundedness conjecture for strong unicity constants, *J. Approx. Theory* **24**, 216–223.

D. Schmidt [1979], Strong unicity and Lipschitz conditions of order $\frac{1}{2}$ for monotone approximation, *J. Approx. Theory* **27**, 346–354.

D. Schmidt [1980a], A characterization of strong unicity constants, in *Approximation Theory III* , E. W. Cheney, Ed., 805–810, Academic Press, New York.

D. Schmidt [1980b], Strong uniqueness for Chebyshev approximation by reciprocals of polynomials on $[0, \infty]$, *J. Approx. Theory* **30**, 277–283.

N. Ph. Seif and G. D. Taylor [1982], Copositive rational approximation, *J. Approx. Theory* **35**, 225–242.

Y. G. Shi [1981], Weighted simultaneous Chebyshev approximation, *J. Approx. Theory* **32**, 306–315.

R. Smarzewski [1986a], Strongly unique best approximation in Banach spaces, *J. Approx. Theory* **46**, 184–194.

R. Smarzewski [1986b], Strongly unique minimization of functionals in Banach spaces with applications to theory of approximation and fixed points, *J. Math. Anal. Appl.* **115**, 155–172.

R. Smarzewski [1987a], Strong unicity in nonlinear approximation, in *Rational Approximation and Applications in Mathematics and Physics*, (Lancut, 1985), 331–350, LNM **1237**, Springer, Berlin.

R. Smarzewski [1987b], Strongly unique best approximation in Banach spaces II, *J. Approx. Theory* **51**, 202–217.

R. Smarzewski [1988a], Strong uniqueness of best approximations in an abstract L^1 space, *J. Math. Anal. Appl.* **136**, 347–351.

R. Smarzewski [1988b], Strong unicity of best approximations in $L_\infty(S, \Sigma,)$, *Proc. Amer. Math. Soc.* **103**, 113–116.

R. Smarzewski [1989], Strong unicity of order 2 in $C(T)$, *J. Approx. Theory* **56**, 306–315.

R. Smarzewski [1990], Finite extremal characterization of strong uniqueness in normed spaces, *J. Approx. Theory* **62**, 213–222.

M. Sommer and H. Strauss [1993], Order of strong uniqueness in best L_∞-approximation by spline spaces, *Acta Math. Hungar.* **61**, 259–280.

H. Strauss [1982], Unicity of best one-sided L_1-approximations, *Numer. Math.* **40**, 229–243.

H. Strauss [1992], Uniform reciprocal approximation subject to coefficient constraints, *Approx. Theory Appl.* **8**, 89–102.

J. Sudolski and A. Wójcik [1987], Another generalization of strong unicity, *Univ. Iagel. Acta Math.* **26**, 43–51.

J. Sudolski and A. P. Wójcik [1990], Some remarks on strong uniqueness of best approximation, *Approx. Theory Appl.* **6**, 44–78.

S. Tanimoto [1998], On best simultaneous approximation, *Math. Japon.* **48**, 275–279.

J. L. Walsh [1931], The existence of rational functions of best approximation, *Trans. Amer. Math. Soc.* **33**, 668–689.

G. A. Watson [1980], *Approximation Theory and Numerical Methods*, J. Wiley and Sons, Chichester.

R. Wegmann [1975], Bounds for nearly best approximations, *Proc. Amer. Math. Soc.* **52**, 252–256.

H. Werner [1964], On the rational Tschebyscheff operator, *Math. Z.* **86**, 317–326.

A. Wójcik [1981], Characterization of strong unicity by tangent cones, in *Approximation and Function Spaces* (Gdansk, 1979), 854–866, North-Holland, Amsterdam-New York.

D. Wulbert [1971], Uniqueness and differential characterizations of approximation from manifolds of functions, *Amer. J. Math.* **93**, 350–366.

S. S. Xu [1995], A note on a strong uniqueness theorem of Strauss, *Approx. Theory Appl. (N.S.)* **11**, 1–5.

C. Yang [1993], Uniqueness of best approximation with coefficient constraints, *Rocky Mountain J. Math.* **23**, 1123–1132.

W. S. Yang and C. Li [1994], Strong unicity for monotone approximation by reciprocals of polynomials, *J. Approx. Theory* **78**, 19–29.

J. W. Young [1907], General theory of approximation by functions involving a given number of arbitrary parameters, *Trans. Amer. Math. Soc.* **8**, 331–344.

F. Zeilfelder [1999], Strong unicity of best uniform approximations from periodic spline spaces, *J. Approx. Theory* **99**, 1–29.

D. Zwick [1987], Strong uniqueness of best spline approximation for a class of piecewise n-convex functions, *Numer. Funct. Anal. Optim.* **9**, 371–379.

András Kroó
Alfred Rényi Institute of Mathematics
Hungarian Academy of Sciences
Budapest, Hungary
and
Budapest University of Technology and Economics
Department of Analysis
Budapest, Hungary
`kroo@renyi.hu`

Allan Pinkus
Department of Mathematics
Technion
Haifa, Israel
`pinkus@tx.technion.ac.il`