Journal of Approximation Theory **114**, 98–114 (2002)
doi:10.1006/jath.2001.3639, available online at http://www.idealibrary.com on IDEAL®

Best Approximation by the Inverse of a Monotone Polynomial and the Location Problem

Daniel Wulbert

Department of Mathematics 0112, University of California, San Diego, La Jolla, California 92093, U.S.A.
E-mail: dwulbert@ucsd.edu

Communicated by Günther Nürnberger

Received December 15, 2000; accepted in revised form September 12, 2001

Let $\mathscr{L} = \{f \in C[0, 1] : f$ is non-decreasing, $f(0) = 0$ and $f(1) = 1\}$. Let M be a class of monotone polynomials of degree n or less. Then each $f \in \mathscr{L}$ has a unique best uniform (or L_1) approximation from $\{p^{-1}: p \in M \cap \mathscr{L}\}$. The special case for $M = \mathbf{P}_n$ shows that the single-data-point location problem for a one-dimensional domain has a unique solution (uniform or L_1-norm).

Key Words: location problem; best approximations; uniqueness; non-linear; monotone polynomials; approximation of inverses; uniform norm; L_1-norm.

1. INTRODUCTION

As an example of the location problem [1], suppose an unmanned craft were landed at a remote site. To determine its location, x_0, local data, such as the altitude, $a(x_0)$, are assessed by the craft. Space limitations only permit the craft to store an approximation, $p(x)$, to the surface topography, $\{a(x): x \in X\}$. The location problem is to identify the function p, from some class M, whose comparison to the collected data would result in minimizing the error from the geographical position of the landing site.

A particular location problem results by specifying the five undefined concepts above: (i) the domain space X, (ii) the data space, (Y, ρ), (iii) the approximating functions M (with domain X and range Y), (iv) a function D corresponding to a distance between a point in X and a subset of X, and (v) a norm, γ, on the real continuous functions defined on X. Given a function $a: X \to Y$, The location problem is to identify the $q \in M$ that minimize $\{\gamma(D(x, p^{-1}(a(x))) : p \in M\}$ (see the comment following the introduction).

These choices provide a rich collection of theoretical problems and potential applications. However, there were no setting in which the problem

0021-9045/02 $35.00

(as stated above) had been solved. The best results identify properties of the best approximations.

The strongest results are by Berdyshev [2–5] who assumed that $X = [0, 1]$, Y is the real line (i.e., one data point), $M = \mathbf{P}_n - \mathbf{P}_0$, $D(y, T) = \max\{|y-t|: t \in T\}$, γ is the sup-norm, and the real function f to be approximated is continuous. In this setting (even when $n = 1$), best location functions do not always exist. If they do exist, they may not be unique (see Section 10).

For the current work, in addition to Berdyshev's setting, we assume that the approximates, M, are invertible functions contained in $\mathscr{L}$. Then the location problem is equivalent to finding a best approximation to an invertible function $f \in \mathscr{L}$ (where $f = a^{-1}$ in the discussion above) from the non-linear family $M^{-1} = \{p^{-1}: p \in M\}$.

The approximating family here is the monotone polynomials. That is, let $1 = l_1 < l_2 < \cdots < l_\lambda \leqslant n$. Let $s_1 = 1$. For each other i, let s_i be equal either $+1$ or -1. The monotone polynomials that we use are

$$M = \{p : p \in \mathbf{P}_n,\ s_i p^{(l_i)} \geqslant 0,\ \text{ for } i = 1, 2, \ldots, \lambda,\ p(0) = 0,\ p(1) = 1\}.$$

This paper shows that every continuous, f (f in $\mathscr{L}$, resp.), has a unique best uniform (L_1, resp.) approximation from M^{-1}.

Comment on the Description of the Location Problem. If $p \in \mathscr{L}$, then $p^{-1}(a(x))$ is well defined. The general setting is a little more complicated. For dist $\rho(a(x), p(X)) = \min\{\rho(a(x), p(w)) : w \in X\}$, put

$$T_{a,p}(x) = \{y \in X : \rho(a(x), p(y)) = \operatorname{dist} \rho(a(x), p(X))\}.$$

Then replace $p^{-1}(a(x))$ in the discussion above with $T_{a,p}(x)$.

Outline of the Paper. The main results of the paper are in Sections 7. Sections 3, 4, 5, and 6 assemble preliminary results needed for the main theorems. Section 7 contains the proof of uniqueness of best approximations from the inverses of monotone polynomials. Section 8 verifies that best approximations exist, and Section 9 identifies the closure of the inverses of all increasing polynomials. The last section contains remarks on the results, examples, and comparisons to other results in the literature.

2. DEFINITIONS AND NOTATION

The polynomials of degree n or less are $\mathbf{P}_n$, and $\mathbf{P}$ refers to the collection of all polynomials.

The sign of a real number r is written *sgn* r. The closure of a set U is written *cl* U, and the cardinality of U is written *card* U.

The critical set for a function f is $ext(f)=\{x\colon |f(x)|=\|f\|_\infty\}$.

A function $f\in C[0,1]$ is said to have an alternation of length m if there are points $0\leqslant x_1<x_2<\cdots<x_m\leqslant 1$ such that for $i=1,2,\ldots,m-1$, $sgn\, f(x_i)=-sgn\, f(x_{i+1})\neq 0$. If in addition $\{x_i\}_{i=1}^m\subseteq ext(f)$, the alternation is called an extremal alternation.

An n dimensional linear subspace $H\subseteq C[0,1]$ is a Haar space if the only function in H that vanishes at n points is the zero function.

3. PRELIMINARY COMPUTATIONS, $L_1[0,1]$

In this section the term best approximation refers to the L_1-norm.

PROPOSITION 3.1. *For f and p invertible functions in $\mathscr{L}$,*

$$\|f-p^{-1}\|_1=\|f^{-1}-p\|_1.$$

Proof. If $p(a)=A$ and $p(b)=B$, then $\int_A^B p^{-1}(y)\,dy=\int_a^b p^{-1}(p(x))\,p'(x)\,dx=\int_a^b xp'(x)\,dx$. Integration by parts shows that $\int_A^B p^{-1}(y)\,dy=Bb-Aa-\int_a^b p(x)\,dx$.

The set $\{x\colon (f-p^{-1})(x)\neq 0\}$ is the union of a countable number of disjoint open intervals $\{(a_i,b_i)\}$. Let $A_i=f^{-1}(a_i)=p(a_i)$ and $B_i=f^{-1}(b_i)=p(b_i)$. From the line above,

$$\int_{A_i}^{B_i}|f-p^{-1}|\,(y)\,dy=\int_{a_i}^{b_i}|f^{-1}-p|\,(x)\,dx.$$

Summing over all the i's proves the proposition. ∎

COROLLARY 3.2. *For f in $\mathscr{L}$; p_1, p_2 invertible functions in $\mathscr{L}$, and $0\leqslant\lambda\leqslant 1$,*

$$\|f-(\lambda p_1+(1-\lambda)\,p_2)^{-1}\|_1\leqslant\lambda\,\|f-p_1^{-1}\|_1+(1-\lambda)\,\|f-p_2^{-1}\|_1.$$

Proof. For $\varepsilon>0$ $f_\varepsilon(x)=[f(x)+\varepsilon x]/[1+\varepsilon]$ is an invertible function in $\mathscr{L}$. The corollary follows, since from Proposition 3.1, the result is true for f_ε. ∎

COROLLARY 3.3. *Let f be in $\mathscr{L}$. Let $H\subset\mathbf{P}$ be a finite dimensional Haar space. Let U be an open subset of H consisting of invertible functions. Let $U_0=U\cap\mathscr{L}$. If p^{-1} is a local best approximation to f from $U_0^{-1}=\{p^{-1}\colon p\in U_0\}$, then p^{-1} is the unique global best approximation to f from U_0^{-1}.*

Proof. From Proposition 3.1, 0 is a local best approximation $f^{-1}-p$ from $H_0=\{h\in H : h(0)=0, \text{ and } h(1)=0\}$. Since H_0 is a linear space, p is a global best approximation. Global best L_1-approximations from H_0 are unique (e.g., see the proof of Jackson's Theorem in [6, pp. 219–220]). Hence if q^{-1} also were a best approximation from U_0^{-1}, then $p-q$ would be a best approximation to $f^{-1}-p$. ▮

4. PRELIMINARY COMPUTATIONS, $L_\infty[0, 1]$

LEMMA 4.1. *Let p_1 and p_2 be invertible functions in $\mathscr{L}$. Let $0<y<1$. Put $w=p_1^{-1}(y)$ and $v=p_2^{-1}(y)$. The following are equivalent:*

$$\text{(i) } w<v, \qquad \text{(ii) } p_2(w)<p_1(w), \qquad \textit{and} \qquad \text{(iii) } p_2(v)<p_1(v).$$

Proof. Part (i) implies (ii): $p_2(w)<p_2(v)=p_2(p_2^{-1}(y))=p_1(p_1^{-1}(y))=p_1(w)$. The other implications are proved similarly. ▮

LEMMA 4.2. *Let p_1 and p_2 be continuous increasing functions on $[0, 1]$. If $0\leqslant p_1^{-1}(y)<p_2^{-1}(y)\leqslant 1$, then for $0<\lambda<1$,*

$$p_1^{-1}(y)<(\lambda p_1+(1-\lambda)\, p_2)^{-1}(y)<p_2^{-1}(y).$$

Proof. For $w\in[p_1^{-1}(y), p_2^{-1}(y)]$, the function $p_\lambda(w)=(\lambda p_1+(1-\lambda)\, p_2)(w)$ is increasing. We compute p_λ at the end points of this interval.

$$p_\lambda(p_1^{-1}(y))=\lambda y+(1-\lambda)\, p_2)(p_1^{-1}(y))<\lambda y+(1-\lambda)\, p_2)(p_2^{-1}(y))=y.$$

Similarly $y<p_\lambda(p_2^{-1}(y))$. Hence for some $v\in(p_1^{-1}(y), p_2^{-1}(y))$, $p_\lambda(v)=(\lambda p_1+(1-\lambda)\, p_2)(v)=y$. Therefore, $(\lambda p_1+(1-\lambda)\, p_2)^{-1}(y)=v\in(p_1^{-1}(y), p_2^{-1}(y))$. ▮

PROPOSITION 4.3. *For $i=1, 2$, let $p_i\in\mathscr{L}$ be invertible. Let $f\in C[0, 1]$. If $\|f-p_1^{-1}\|_\infty\leqslant 1$ and $\|f-p_2^{-1}\|_\infty<1$, then for $0<\lambda<1$,*

$$\|f-(\lambda p_1+(1-\lambda)\, p_2)^{-1}\|_\infty<1.$$

Proof. For definiteness, assume throughout this proof that $p_1^{-1}(y)\leqslant p_2^{-1}(y)$. If $p_1^{-1}(y)=p_2^{-1}(y)=x$, then $(\lambda p_1+(1-\lambda)\, p_2)(x)=y$, and

$$(\lambda p_1+(1-\lambda)\, p_2)^{-1}(y)=p_1^{-1}(y)=p_2^{-1}(y).$$

Therefore

$$-1+f(y)<p_1^{-1}(y)=(\lambda p_1+(1-\lambda)\, p_2)^{-1}\,(y)=p_2^{-1}(y)<f(y)+1.$$

Otherwise from Lemma 4.2,

$$-1+f(y)\leqslant p_1^{-1}(y)<(\lambda p_1+(1-\lambda)\, p_2)^{-1}\,(y)<p_2^{-1}(y)\leqslant f(y)+1.$$

In either case,

$$-1+f(y)<(\lambda p_1+(1-\lambda)\, p_2)^{-1}\,(y)<f(y)+1,$$

and

$$\|f-(\lambda p_1+(1-\lambda)\, p_2)^{-1}\|_\infty<1. \quad \blacksquare$$

The next two lemmas are known basic properties of Haar spaces written in the forms needed for direct application in the proof of Proposition 4.6

LEMMA 4.4. *Let $H\subseteq C[0,1]$ be a Haar space of dimension κ. Let $0\leqslant x_0<x_1<\cdots<x_\kappa\leqslant 1$. If $h\in H$ is such that $(-1)^i\, h(x_i)\geqslant 0$, then $h=0$.*

LEMMA 4.5. *Let $H\subseteq C[0,1]$ be a Haar space of dimension κ. Suppose that $\gamma\leqslant\kappa$ and that $F_1, \ldots, F_\gamma$ are closed sets in $[0,1]$ such that if $v\in F_i$ and $w\in F_{i+1}$, then $v<w$. Then there is an $h\in H$ such that h is positive on F_i for even integers i, and h is negative on F_i for odd integers i.*

PROPOSITION 4.6. *Let $H\subseteq C[0,1]$ be a Haar space of dimension κ. Let U be an open subset of H consisting of invertible functions. Let $U_0=U\cap\mathscr{L}$. Let $f\in C[0,1]$ be such that $f(0)=0$ and $f(1)=1$. The following are equivalent:*

(i) *p^{-1} is a local best approximation to f from $\{p^{-1}\colon p\in U_0\}$,*

(ii) *p^{-1} is the unique global best approximation to f from $\{p^{-1}\colon p\in U_0\}$,*

(iii) *$f-p^{-1}$ has an extremal alternation of length $\kappa-1$ in the open interval $(0,1)$,*

(iv) *$f-p^{-1}$ has 0 as a best approximation from $H_0=\{h\in H: h(0)=0=h(1)\}$.*

Proof. First suppose (iii), that $f-p^{-1}$ has an extremal alternation of length $\kappa-1$ in $(0,1)$. If $p_*^{-1}\in\{p^{-1}\colon p\in U_0\}$ were a global best approximation, then $p^{-1}-p_*^{-1}$ would be alternately non-positive and non-negative on the points of the the union of $\{0,1\}$ with the points of the extremal alternation. By Lemmas 4.1 and 4.7, $p^{-1}=p_*^{-1}$. Thus (iii) implies (ii).

Part (ii) always implies (i).

Now suppose that $f-p^{-1}$ does not have an extremal alternation of length $\kappa-1$ in (0, 1). Then there is a $\gamma \leqslant \kappa-2$ and closed sets $F_1, ..., F_\gamma$ in (0, 1) such that

(1) $v \in F_i$ and $w \in F_{i+1}$, then $v < w$,

(2) $ext(f-p^{-1}) = \bigcup F_i$, and

(3)[a] $(f-p^{-1})(x) = \|f-p^{-1}\|$ if and only if $x \in F_i$ for i even, or

[b] $(f-p^{-1})(x) = \|f-p^{-1}\|$ if and only if $x \in F_i$ for i odd.

We will assume that condition (3)[a] holds. Furthermore, since $(f-p^{-1})(0) = 0 = (f-p^{-1})(1)$, neither 0 nor 1 is in $\bigcup F_i$ (by condition (2)). Therefore, there is $0 < a < b < 1$ such that $[a, b]$ contains each F_i.

Since p^{-1} is order preserving, $p^{-1}(F_1), ..., p^{-1}(F_\gamma)$ are correspondingly ordered closed subsets of $[p^{-1}(a), p^{-1}(b)]$. Since H_0 is a Haar space of dimension $\kappa-2$ on any closed subset of (0, 1), by Lemma 4.5 there is a h in H_0 that is positive on $p^{-1}(F_i)$ for even integers i, and is negative on the $p^{-1}(F_i)$ for i odd. For all positive λ, $p+\lambda h > p$ on $p^{-1}(F_i)$ for even integers i, and $p+\lambda h < p$ on the $p^{-1}(F_i)$ for i odd. For sufficiently small λ, $p+\lambda h$ is invertible. From Lemma 4.1 $(p+\lambda h)^{-1} < p^{-1}$ on F_i for even integers i, and $(p+\lambda h)^{-1} > p^{-1}$ on the F_i for odd indices. Hence, for sufficiently small λ, $(p+\lambda h)^{-1}$ is a better approximation to f then is p^{-1}. This proves that (i) implies (iii).

The equivalence of (iii) and (iv) is classically known. It could also be proven with a simpler version of the argument above that showed the equivalence of (iii) and (ii). ∎

5. PRELIMINARY COMPUTATIONS, MULTIPLICITY OF ZEROS

Section 6 will show that a particular subspace of monotone polynomials is a Haar space. The proof uses a generalized form of Rolle's Theorem which we prove in this section.

Notation. For a function p, let

$Z(p)$ (or Zp) be the zeros of p in $[0, 1]$;

$n(p)$ be the cardinality of $Z(p)$;

$m(p:z) = max\{m$ a non-negative integer: $(x-z)^m$ is a factor of $p\}$;

and

$t(p) = \sum_{z \in Z(p)} m(p:z)$.

LEMMA (Classical Rolle's Theorem). *If* $0 \leqslant a_1 < a_2 < \cdots < a_{n(p)} \leqslant 1$ *are the zeros of* $p \in \mathbf{P}_n$, *then there exists* $C = \{c_1, c_2, ..., c_{n(p)-1}\} \subseteq Z(p')$ *such that:*

(1) $a_i < c_i < a_{i+1}$, *and*

(2) $m(p' : c_i)$ *is an odd integer. Hence for an integer* $l \geqslant 1$,

(3) $t(p^{(l)}) \geqslant t(p) - l$.

Discussion. The objective of this section (Theorem 5.1) is to show that if some higher order derivatives of p have zeros of prescribed multiplicity, then the lower bound, $t(p) - l$, for $t(p^{(l)})$ will be augmented by the sum of the multiplicities of these zeros.

EXAMPLE. If $p(x) = (\frac{2}{3})^6 x - (\frac{1}{3})^6 (x-1) - (x - \frac{1}{3})^6$, then $t(p) = 2$, but $t(p^{(5)}) \geqslant t(p) - 5 + m(p^{(2)} : \frac{1}{3}) = 1$.

Suppose that $0 = l_0 < l_1 < l_2 < \cdots < l_\lambda \leqslant n$. The main result will be to augment the lower bound, $t(p) - l$, for $t(p^{(l)})$ by the sum of the multiplicities of the zeros of $p^{(l_i)}$. From Rolle's Theorem it is apparent that (i) the parity of the multiplicity of the zeros of the derivatives will be relevant, and (ii) zeros of p and its derivatives that occur at 0 and 1 generally have a different effect on the estimates of $t(p^{(l)})$ then do zeros in (0, 1).

Notation. We set the notation for the general result. For $i = 0, 1, 2, ..., \lambda$ and $j = 1, 2, ..., \beta_i$, let

$$B_i = \{b_{i,j}\}_{j=1}^{\beta_i} = \{b \in Z(p^{(l_i)}) : m(p^{(l_{i-1})} : b) < l_i - l_{i-1};\ \text{if } b \in (0, 1),$$
$$\text{then } m(p^{(l_i)} : b) \geqslant 2\}.$$

If $b_{i,j} \in \bigcup_{i=0}^{\lambda} B_i \cap (0, 1)$, there is a positive even integer, $m_{i,j}$, such that $m(p^{(l_i)} : b_{i,j}) \geqslant m_{i,j}$. If $b_{i,j} = 0$, or if $b_{i,j} = 1$, put $m_{i,j} = m(p^{(l_i)} : b_{i,j})$

Purpose for the notation. The theorem will be that $t(p) - l_\lambda$, the lower bound for $t(p^{(l_\lambda)})$, will be augmented by the sum of the multiplicities of the zeros in $\bigcup_{i=0}^{\lambda} B_i$. However, a point could be a zero of more than one of the derivatives $p^{(l_\lambda)}$. We need to avoid multiple countings of such zeros. For example, the polynomial $p(x) = (\frac{2}{3})^6 x - (\frac{1}{3})^6 (x-1) - (x - \frac{1}{3})^6$, has $m(p^{(2)} : \frac{1}{3}) = 4$ and $m(p^{(4)} : \frac{1}{3}) = 2$. But augmenting the lower bound by the sum of both of these would count the two highest order zeros twice. The assumption above that $m(p^{(l_{i-1})} : b) < l_i - l_{i-1}$ was introduced to prevent redundant counting of zeros.

EXAMPLE. It is possible for a point to intermittently appear in the set of zeros of higher order derivatives of p. As an example, if $0 < a < 1$, and $p(x) = 1 + (x-a)^3 + (x-a)^6$; then $a \in Z(p^{(d)})$ for $d = 1, 2, 4$ and 5, but not 0 or 3.

THEOREM 5.1.

$$t(p^{(l_\lambda)}) \geqslant t(p) - l_\lambda + \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}.$$

Proof. The proof is by induction on λ. We prove the induction step in two parts. The first (Lemma 5.2) estimates $n(p^{(l_i)})$; the second (Lemma 5.3), $t(p^{(l_i)})$. The theorem then is immediate from Lemma 5.3, since for $i = 1, 2, 3, \ldots, \lambda$,

$$t(p^{(l_i)}) \geqslant t(p^{(l_{i-1})}) - (l_i - l_{i-1}) + \sum_{j=1}^{\beta_i} m_{i,j}. \quad \blacksquare$$

LEMMA 5.2. *Let $A_i = \{a \in Zp : m(p, a) = i\}$.*

$$n(p^{(k)}) \geqslant \sum_{i=1}^{k} i\ card(A_i) + (k+1) \sum_{i=k+1}^{n} card(A_i) - k.$$

Proof. The proof is by induction. Since $n(p) = \sum_{i=1}^{n} card(A_i)$, the set C in Rolle's Theorem has cardinality $\sum_{i=1}^{n} card(A_i) - 1$. We conclude that

$$\begin{aligned} n(p') &\geqslant \sum_{i=2}^{n} card(A_i) + \sum_{i=1}^{n} card(A_i) - 1 \\ &= card(A_1) + 2 \sum_{i=2}^{n} card(A_i) - 1. \end{aligned}$$

Now we assume the induction hypothesis that

$$n(p^{(k-1)}) \geqslant \sum_{i=1}^{k-1} i\ card(A_i) + k \sum_{i=k}^{n} card(A_i) - (k-1).$$

From Rolle's Theorem $p^{(k)}$ has $n(p^{(k-1)}) - 1$ "new" zeros which are disjoint from $\bigcup_{i=k+1}^{n} A_i$.

We conclude that

$$\begin{aligned} n(p^{(k)}) &\geqslant \sum_{i=k+1}^{n} card(A_i) + n(p^{(k-1)}) - 1 \\ &= \sum_{i=k+1}^{n} card(A_i) + \sum_{i=1}^{k-1} i\ card(A_i) + k \sum_{i=k}^{n} card(A_i) - (k-1) + 1. \quad \blacksquare \end{aligned}$$

Let $B = \{b_1, b_2, \ldots, b_\beta\} = \{b \in Zp^{(k)} : \text{if } b \in (0, 1), m(p^{(k)} : b) \geqslant 2\} - \{a \in Zp : m(p : a) \geqslant k+1\}$.

Hence, if $b_i \in B \cap (0, 1)$, there is a positive even integer, m_i, such that $m(p^{(k)} : b_i) \geqslant m_i$. If $b_i = 0$, or if $b_i = 1$, let $m_i = m(p^{(k)} : b_i)$.

LEMMA 5.3. *With* $\{m_i\}_{i=1}^{\beta}$ *as defined immediately above,*

$$t(p^{(k)}) \geqslant t(p) - k + \sum_{i=1}^{\beta} m_i.$$

Proof. Again let $A_i = \{a \in Zp : m(p, a) = i\}$. From Rolle's Theorem, $p^{(k)}$ has at least $n(p^{(k-1)}) - 1$ zeros other then the zeros in $\bigcup_{i=k+1}^{n} A_i$. Let $C = \{c_1, c_2, ..., c_{n(p)-1}\}$ denote these zeros.

C has the following properties:

(i) *card* $C \geqslant n(p^{(k-1)}) - 1$,

(ii) $C \cap [\bigcup_{i=k+1}^{n} A_i] = \varnothing$, and

(iii) If $c \in C$, then $m(p^{(k)} : c)$ is odd,

$$\begin{aligned}
t(p^{(k)}) \geqslant{} & \sum_{a \in \bigcup_{i=k+1}^{n} A_i} m(p^{(k)} : a) + \sum_{c_i \notin B} m(p^{(k)} : c_i) + \sum_{b_i \notin C} m(p^{(k)} : b_i) \\
& + \sum_{b_i \in C \cap B} m(p^{(k)} : b_i) \\
\geqslant{} & \sum_{i=k+1}^{n} (i-k)\, card(A_i) + card(C - B) + \sum_{b_i \notin C} m_i + \sum_{b_i \in C \cap B} [m_i + 1] \\
\geqslant{} & \sum_{i=k+1}^{n} (i-k)\, card(A_i) + card(C - B) + \sum_{b_i \notin C} m_i + \sum_{b_i \in C \cap B} m_i \\
& + card(C \cap B) \\
\geqslant{} & \sum_{i=k+1}^{n} (i-k)\, card(A_i) + card\, C + \sum_{b_i \in B} m_i \\
\geqslant{} & \sum_{i=k+1}^{n} (i-k)\, card(A_i) + [n(p^{(k-1)}) - 1] + \sum_{i=1}^{\beta} m_i \\
\geqslant{} & \sum_{i=k+1}^{n} (i-k)\, card(A_i) \\
& + \left[\sum_{i=1}^{k-1} i\, card(A_i) + k \sum_{i=k}^{n} card(A_i) - (k-1) - 1\right] + \sum_{i=1}^{\beta} m_i \\
={} & \sum_{i=1}^{k-1} i\, card(A_i) + \sum_{i=k}^{n} i\, card(A_i) - k + \sum_{i=1}^{\beta} m_i \\
={} & t(p) - k + \sum_{i=1}^{\beta} m_i.
\end{aligned}$$

6. THE HAAR SPACE D_U

In the last section we estimated the total multiplicity of a function p using the multiplicities, $m(p, b_{i,j})$, of its zeros at various derivatives $p^{(l_i)}$. In this section we specify, $\{b_{i,j}\}$ and $m_{i,j}$ and look at the space of polynomials whose l_ith derivative have those prescribed multiplicities at the specified points.

DEFINITION OF D_u. Let $u = (\{l_i\}, \{b_{i,j}\}, \{m_{i,j}\})$, where:

$0 = l_0 < l_1 < l_2 < \cdots < l_\lambda \leqslant n$;

$b_{i,j} \in [0, 1]$ for $i = 0, 1, 2, \ldots, \lambda$, and $j = 1, 2, \ldots, \beta_i$;

for all (i, j), $m_{i,j}$ is a positive integer, and for $b_{i,j} \in (0, 1)$, $m_{i,j}$ is a positive even integer;

if $b_{i,j} = b_{\alpha,v}$ for $i < \alpha$, then $m_{i,j} < l_\alpha - l_i$; and

for each $\gamma = 1, 2, \ldots, n$; $\sum_{i=\gamma}^{n} \sum_{j=1}^{\beta_i} m_{i,j} \leqslant n - l_\gamma$.

Put $u = (\{l_i\}, \{b_{i,j}\}, \{m_{i,j}\})$, and

$$D_u = \{p \in \mathbf{P}_n : p^{(l_i+k)}(b_{i,j}) = 0;\ k = 0, 1, \ldots, m_{i,j} - 1\}.$$

Purpose of the Conditions in the Definition. The last condition guarantees that $D_u - \{0\}$ is not empty. Without such an assumption, one might be hypothesizing that a polynomial $p \in D_u$ (which has degree $\leqslant n$) has a l_γth derivative with more than $n - l_\gamma$ zeros (counting multiplicity).

THEOREM 6.1. *D_u is a Haar space of dimension $n + 1 - \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}$.*

Proof. First we show that the dimension is at least as large as stated. We define linear functionals on $\mathbf{P}_n$. Put $L_{i,k,j}(p) = p^{(l_i+k)}(b_{i,j})$. Then

$$D_u = \{L_{i,k,j}^{-1}(0) : i = 0, 1, 2, \ldots, \lambda;\ j = 1, 2, \ldots, \beta_i;\ k = 0, 1, \ldots, (m_{i,j} - 1)\}.$$

Since there are $\sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}$ linear functionals, the dimension of D_u is greater than or equal $n + 1 - \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}$.

Now suppose that $p \in D_u$ has $n + 1 - \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}$ zeros. To complete the proof of the theorem, we need to show that $p = 0$. We have

$$t(p) \geqslant n(p) \geqslant n + 1 - \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}.$$

By Theorem 5.1, for any $0 \leqslant \gamma \leqslant \lambda$,

$$\begin{aligned} t(p^{(l_\gamma)}) &\geqslant t(p) - l_\gamma + \sum_{i=0}^{\gamma} \sum_{j=1}^{\beta_i} m_{i,j} \\ &\geqslant n + 1 - \sum_{i=0}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j} - l_\gamma + \sum_{i=0}^{\gamma} \sum_{j=1}^{\beta_i} m_{i,j} \\ &\geqslant n + 1 - l_\gamma + \sum_{i=\gamma+1}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j}. \end{aligned}$$

When $\gamma = \lambda$, the double sum is zero and we conclude

$$t(p^{(l_\lambda)}) \geqslant n - l_\lambda + 1.$$

Since $p^{(l_\lambda)} \in \mathbf{P}_{n-l_\lambda}$, we deduce that $p^{(l_\lambda)} = 0$. Therefore, $deg\ p \leqslant l_\lambda - 1$.

When $0 \leqslant \gamma < \lambda$, we have hypothesized that $\sum_{i=\gamma+1}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j} \leqslant n - l_{\gamma+1}$, and we have

$$t(p^{(l_\gamma)}) \geqslant l_{\gamma+1} - l_\gamma + 1 + 2 \sum_{i=\gamma+1}^{\lambda} \sum_{j=1}^{\beta_i} m_{i,j} \geqslant l_{\gamma+1} - l_\gamma + 1.$$

We apply this conclusion with $\gamma = \lambda - 1$ and get

$$t(p^{(l_{\lambda-1})}) \geqslant l_\lambda - l_{\lambda-1} + 1.$$

Since

$$deg\ p^{(l_{\lambda-1})} \leqslant deg\ p - l_{\lambda-1} \leqslant l_\lambda - 1 - l_{\lambda-1},$$

we have that $p^{(l_{\lambda-1})} = 0$. Continuing the same argument inductively, we conclude that $p = p^{(l_0)} = 0$. ∎

Let $\{l_i\}$ be as before. Let $p \in \mathbf{P}_n$ be specified. Let $\{b_{i,j}(p)\}_{j=1}^{\beta_i}$ be equal

$$\{b \in Z(p^{(l_i)}) : p^{(l_i)} \neq 0;\ m(p^{(l_{i-1})} : b) < l_i - l_{i-1};\ \text{if } b \in (0,1), \\ \text{then } m(p^{(l_i)} : b) \text{ is an even integer}\},$$

and let $u(p) = (\{l_i\}, \{b_{i,j}\}, \{m(p^{(l_i)} : b_{i,j})\})$.

THEOREM 6.2. *$D_{u(p)}$ is a Haar space of dimension $deg\ p - \sum_{i=\gamma}^{\lambda} \sum_{j=1}^{\beta_i} m(p^{(l_i)} : b_{i,j})$.*

Proof. Without loss of generality, we may assume that $deg\, p = n$. The only hypothesis that needs to be verified is that for each $\gamma = 0, 1, 2, ..., n$;

$$\sum_{i=\gamma}^{\lambda} \sum_{j=1}^{\beta_i} m(p^{(l_i)} : b_{i,j}) \leqslant n - l_\gamma.$$

From Theorem 5.1,

$$t(p^{(l_\lambda)}) \geqslant t(p^{(\gamma)}) - (l_\lambda - l_\gamma) + \sum_{i=\gamma}^{\lambda} \sum_{j=1}^{\beta_i} m(p^{(l_i)} : b_{i,j}).$$

So if the condition were not met, there would be a γ for which

$$t(p^{(l_\lambda)}) > t(p^{(\gamma)}) - l_\lambda + n \geqslant n - l_\lambda.$$

But $p^{(l_\lambda)}$ has degree $n - l_\lambda$, and so it cannot have more than $n - l_\lambda$ zero (counting multiplicities). ∎

7. UNIQUENESS OF BEST APPROXIMATIONS

Let the monotone polynomials, M, and their inverses, M^{-1} be as defined in the introduction. For $p \in \mathbf{P}_n$, let $u(p)$ be as defined before Theorem 6.2.

LEMMA 7.1. *If $p \in M$, then $D_{u(p)}$ is a Haar space of dimension $deg\, p + 1 - \sum_{i=1}^{\lambda} \sum_{j=1}^{\beta_i} m(p^{(l_i)}; b_{i,j})$.*

Proof. This is immediate from Theorem 6.2 since for a monotone polynomial, p, all zeros of $p^{(l_\gamma)}$ (for $\gamma > 0$) in (0, 1) have even multiplicity. ∎

THEOREM 7.2. *Let $f \in C[0, 1]$ be such that $f(0) = 0$ and $f(1) = 1$. A local best uniform approximation, p^{-1}, to f from M^{-1} is the unique global best uniform approximation from M^{-1}.*

THEOREM 7.3. *Let $f \in \mathscr{L}$. A local best L_1 approximation, p^{-1}, to f from M^{-1} is the unique global best L_1 approximation from M^{-1}.*

Notation for the Proof. For $p \in M$ put

$$T(p) = \sum_{b \in B} m(p^{(l_i)}; b) = \sum_{i=1}^{\lambda} \sum_{j=1}^{\beta_i} m(p^{(l_i)}; b_{i,j}).$$

For $K=0, 1, \ldots, n$; and $J=0, 1, \ldots, n$ let,

$$M^{-1}_{K,J}=\{p^{-1}\in M^{-1}: deg\ p=n-J,\ \text{ and } T(p)=K\}.$$

Define $(\mu, \nu)<(K, J)$ if $\mu\leqslant K.\ \nu\leqslant J$, and at least one of these inequalities is a strict inequality.

The results preceding these theorems have been arranged so that the proofs for both theorems above are identical. The proof will be by induction on the (K, J) in $M^{-1}_{K,J}$.

Proof. Since $U=\{p\in \mathbf{P}_n : T(p)=0,\ \text{ and } deg\ p=n\}$ is an open subset of $\mathbf{P}_n$, if p^{-1} were in $M^{-1}_{0,0}$, then, from Corollary 3.3 (for the L_1-norm) and Proposition 4.6 (for the uniform norm) p^{-1} would be the unique global best approximation to f from from M^{-1}.

For the induction hypothesis we assume that if $p^{-1}\in M^{-1}_{\mu,\nu}$, for $(\mu, \nu)<(K, J)$, and p^{-1} is a local best approximation to f from M^{-1}, then p^{-1} is the unique global best approximation to f from M^{-1}.

For the induction step, let $p^{-1}\in M^{-1}_{K,J}$, be such that p^{-1} is a local best approximation to f. Let V be a neighborhood of p^{-1} in M^{-1} such that if $v^{-1}\in V$, then $\|f-p^{-1}\|\leqslant\|f-v^{-1}\|$. Now suppose that $q^{-1}\in M^{-1}$ is such that $\|f-p^{-1}\|\geqslant\|f-q^{-1}\|$. Then from Corollary 3.2 (for the L_1-norm) and Lemma 4.2 (for the uniform norm) for all $0\leqslant c\leqslant 1, p_c^{-1}=[(1-c)\,p+cq]^{-1}\in M^{-1}$ is at least as good an approximation to f as is p^{-1}. For small c, $p_c^{-1}\in V$. So for these small c, p_c^{-1} must also be a best local approximation to f.

The zeros of $p_c^{(l_i)}$ (counting multiplicity) are the common zeros of $q^{(l_i)}$ and $p^{(l_i)}$ (since both derivatives are always non-positive—or they are always non-negative). Also $deg\ p_c=\max\{deg\ p, deg\ q\}$. Hence if either there were a zero of $p^{(l_i)}$ (counting multiplicity) that were not also a zero of $q^{(l_i)}$, or if $deg\ q$ were greater than $deg\ p$, then p_c^{-1} would be in $M^{-1}_{\mu,\nu}$ for some $(\mu.\nu)<(K, J)$. From the induction hypothesis, p_c^{-1} would be the unique global best approximation to f. Since this would be true for all sufficiently small c, we must have that $p=q$.

We have shown that if q^{-1} were a global best approximation to f, then

(1) all the zeros of $p^{(l_i)}$ (counting multiplicity) are also zeros of $p^{(l_i)}$,

and

(2) $deg\ p\geqslant deg\ q$.

Let

$$u(p)=(\{l_i\}_{i=0}^{\lambda};\ \{b_{i,j}\}_{i=0,\,j=1}^{\lambda,\,\beta_i};\ \{m(p^{(l_i)}; b_{i,j})\}_{i=0,\,j=1}^{\lambda,\,\beta_i})$$

be defined for p as before Lemma 7.1.

Then both q and p are in $D_{u(p)}$, and p is in the open subset of $D_{u(p)}$

$$U = \{d \in D_{u(p)} : T(d) = T(p), \deg d = \deg p\}.$$

From Lemma 7.1, $D_{u(p)}$ is a Haar space. We again use Corollary 3.3 (for the L_1-norm) and Lemma 4.6 (for the uniform norm) to conclude $p = q$. ▮

8. EXISTENCE OF BEST APPROXIMATIONS

LEMMA 8.1. *M^{-1} is compact in any topology defined on $C[0, 1]$ that is weaker then the uniform norm.*

Proof. Since M is a closed and bounded subset of $\mathbf{P}_n$, we only need to show that the map taking p to p^{-1} is continuous on M. Let $p_i \to p_0$, since M is finite dimensional $p_i \to p_0$ uniformly. For $y \in [0, 1]$,

$$y = p_i(p_i^{-1}(y)) = p_0(p_i^{-1}(y)) + [p_i(p_i^{-1}(y)) - p_0(p_i^{-1}(y))].$$

The term in brackets goes to zero uniformly in y. So $p_0(p_i^{-1}(y))$ converges to y (uniformly in y). Since p_0^{-1} is uniformly continuous, applying it to both sides implies that $p_i^{-1}(y)$ converges to $p_0^{-1}(y)$ (uniformly in y). ▮

COROLLARY 8.2. *Every $f \in L_p[0, 1]$ $(1 \leqslant p \leqslant \infty)$ has a best L_p approximation in M^{-1}.*

9. THE UNIFORM CLOSURE OF M_∞^{-1}

DEFINITION. Put

$$M_\infty = \{p\colon p \in \mathbf{P}, p' \geqslant 0, p(0) = 0, p(1) = 1\}, \quad \text{and} \quad M_\infty^{-1} = \{p^{-1}\colon p \in M_\infty\}.$$

THEOREM 9.1. *$cl\, M_\infty^{-1} = \mathscr{L}$.*

Proof. It is apparent that $cl\, M_\infty^{-1} \subseteq \mathscr{L}$. We show the converse by first proving that functions with additional smoothness properties are in $cl\, M_\infty^{-1}$.

First, a C^1-function f in $\mathscr{L}$ is in $cl\, M_\infty^{-1}$ if $f' > 0$ for all $y \in [0, 1]$. Suppose that $(f^{-1})' \geqslant 2\delta > 0$. Let $\varepsilon > 0$. We want to show that there is a $p \in M_\infty$ such that $\|f - p^{-1}\| < \varepsilon$. Let p be a polynomial that approximates f^{-1} in the sense that: (i) $\|(f^{-1})' - p'\| < \delta$ (so $p' > \delta$ on $[0, 1]$), (ii) $\|(f^{-1}) - p\| < \varepsilon\delta$, and (iii) $p(0) = 0$ and $p(1) = 1$.

Let $0 \leqslant y \leqslant 1$, and put $x = f(y)$. From the Mean-Value Theorem there is a ζ such that

$$p(p^{-1}(y)) - p(x) = p'(\zeta)[p^{-1}(y) - x]$$

and

$$p^{-1}(y) = \frac{y - p(x)}{p'(\zeta)} + x.$$

Therefore,

$$|f(y) - p^{-1}(y)| = \left| x - \left[x + \frac{y - p(x)}{p'(\zeta)} \right] \right| \leqslant \frac{|f^{-1}(x) - p(x)|}{|p'(\zeta)|} \leqslant \varepsilon.$$

Second, a C^1-function f in $\mathscr{L}$ is in the closure of the C^1-functions in $\mathscr{L}$ that have a positive derivative. For any $\varepsilon > 0$, let $h(x) = [f(x) + \frac{\varepsilon}{2} x]/[1 + \frac{\varepsilon}{2}]$. Then h is a C^1-function in $\mathscr{L}$ with a positive derivative such that, $\|f - h\| \leqslant \varepsilon$.

Third, $\mathscr{L}$ is in the closure of the C^1-functions in $\mathscr{L}$. To prove this let $K(x)$ be a non-negative, C^1 function on the whole real line such that $\int_{-\infty}^{\infty} K(x)\, dx = 1$. Put $K_\lambda(x) = \lambda K(\lambda x)$.

Extend f to the whole real line by letting $f(x) = 0$ for $x \leqslant 0$ and letting $f(x) = 1$ for $1 \leqslant x$.

Let $(f * K_\lambda)(x) = \int_{-\infty}^{\infty} f(x - y)\, K_\lambda(y)\, dy$.

Since f is continuous on $[0, 1]$, $f * K_\lambda$ are C^1-functions that converges to f uniformly [10, Theorem 3 and Corollary 1].

Furthermore, $f * K_\lambda$ is non-decreasing since if $x_1 < x_2$; then $f(x_1 - y) \leqslant f(x_2 - y)$, $f(x_1 - y)\, K(\frac{y}{\lambda}) \leqslant f(x_2 - y)\, K(\frac{y}{\lambda})$, and $f * K_\lambda(x_1) \leqslant f * K_\lambda(x_2)$.

Hence,

$$h_\lambda(x) = \frac{f * K_\lambda(x) - f * K_\lambda(0)}{f * K_\lambda(1) - f * K_\lambda(0)}.$$

are C^1-function in $\mathscr{L}$ that converge uniformly to f as $\lambda \to \infty$. ∎

10. COMMENTS AND COMPARISONS

L_1 Approximation from M^{-1}. For any continuous function f we proved that if $\|f - p_i^{-1}\|_\infty \leqslant 1$, for $i = 1$ and 2, and $p_i \in M$, then $\|f - (\frac{1}{2} p_1 + \frac{1}{2} p_2)^{-1}\|_\infty \leqslant 1$. We showed this was also true for the L_1-norm if f were an increasing function in $\mathscr{L}$. If f is decreasing this implication is not necessarily true. With an argument paralleling that in Section 3 one can show:

PROPOSITION 10.1. *Let $f(1-x)$, $p_1(x)$, and $p_2(x)$ be increasing functions in $\mathscr{L}$. Let $0 \leqslant \lambda \leqslant 1$.*

$$\|f-(\lambda p_1+(1-\lambda)\, p_2)^{-1}\|_1 \geqslant \lambda\, \|f-p_1^{-1}\|_1+(1-\lambda)\, \|f-p_2^{-1}\|_1.$$

Proof. Show first that for such functions $\|f-p_i^{-1}\|=1-\|f^{-1}-p\|$. ∎

EXAMPLE. The inequality of the proposition can be a strict inequality. For example, when $f^{-1}(x)=1-x$, $p_1(x)=x$, $p_2(x)=x^2$, and $\lambda=\frac{1}{2}$. This follows since on the interval $((\sqrt{5}-1)/2, \frac{1}{2})$ we have that $(1-x)-x<0$ and $(1-x)-x^2>0$. Hence on that interval, $|(1-x)-x|+|(1-x)-x^2| >|2(1-x)-x+-x^2|$. Consequently, $\|(1-x)-x\|_1+\|(1-x)-x^2\|_1 > \|2(1-x)-x-x^2\|_1$. Hence, with this f, p_1, and p_2, $1-\|f^{-1}-\lambda p_1-(1-\lambda)\, p_2\|_1 > \lambda+(1-\lambda)-\|\lambda f^{-1}-\lambda p_1\|_1-\|(1-\lambda)\, f-(1-\lambda)\, p_2\|_1$.

Monotone Approximation. Approximation from the convex set of monotone polynomials is quite well understood. Best approximations, in both the sup-norm and the L_1-norm, are unique [7–8]. The proof here for approximation from the inverses of monotone polynomials involves showing that the error function has 0 as a best approximation from certain classes of monotone polynomials. Some of the lemmas used here are variations of results in monotone polynomial approximation.

The condition $\sum_{i=\gamma}^{n} \sum_{j=1}^{\beta_i} m_{i,j} \leqslant n-l_\gamma$ first introduced in Section 6, has been called the Polya condition.

The Location Problem. V.I. Berdyshev did not assume growth restrictions (e.g., increasing) on either f or the approximates (other than being non-constant). Hence there could be many "locations" where the actual altitude (or the approximating altitude) is attained. That gives rise to both non-existence and to non-uniqueness of best location functions.

EXAMPLE (Non-existence). Let $U(x)=1-\sqrt{1-|x|}$ be defined on $[-1, 1]$. Consider approximating U from $\mathbf{P}_1-\mathbf{P}_0$, the non-constant linear polynomials. First, we observe that the minimum of the location error is less than or equal one. Let $p_k=1+kx$. As $k\to\infty$ the location error for each $x\in[-1, 1]$ goes to one or less. Second, if there were a $p\in\mathbf{P}_1-\mathbf{P}_0$ that produced a location error less then or equal one, then since $U(-1)=1=U(1)$ it is necessary that $p(0)=1$. So $p=1+\alpha x$ for some α. But since $U'(x)\to\infty$ as x approaches the end points of the interval $[-1, 1]$, there will be an interval of points near one of the end points which have location errors exceeding one.

EXAMPLE (Non-unique). Let $V(x)=|x|$ be defined on $[-1, 1]$. Again consider approximating V from $\mathbf{P}_1-\mathbf{P}_0$. Then $1+\alpha x$ is a best approximating location function for every α such that. $|\alpha| \geqslant 1$.

REFERENCES

1. I. N. Beloglazov, G. I. Dzhanzhgava, and G. P. Chigin, "Foundations of Navigation over Geophysical Fields," Nauka, Moscow, 1985. [In Russian]
2. V. I. Berdyshev, Approximation of functions and navigation problem, *in* "Approximation and Optimization," Proc. of International Conference on Approximation and Optimization (Romania)-ICAOR Cluj-Napoca, Vol. I, pp. 177–180, July 29–August 1, 1996.
3. V. I. Berdyshev, Differentiation of location error, *Proc. Steklov Inst. Math.* **219** (1997), 66–72; translated from *Trudy Mat. Inst. Steklov.* **219** (1997), 74–79.
4. V. I. Berdyshev, Polynomial approximation connected with navigation over geophysical fields, *Russian Acad. Sci. Dokl. Math.* **46** (1993), 157–160; translated from *Proc. Akad. Nauk Dokl.* **325** (1992).
5. V. I. Berdyshev, Polynomial approximation and geophysical fields navigation problem, *Russian J. Numer. Anal. Math. Modelling* **8** (1993), 83–100.
6. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
7. G. G. Lorentz and K. L. Zeller, Monotone approximation by algebraic polynomials, *Trans. Amer. Math. Soc.* **149** (1970), 1–18.
8. R. A. Lorentz, Uniqueness of best approximation by monotone polynomials, *J. Approx. Theory* **4** (1971), 401–418.
9. W. Rudin, "Principles of Mathematical Analysis," 3rd ed., McGraw–Hill, New York, 1976.
10. H. S. Shapiro, "Smoothing and Approximation of Functions," Van Nostrand Reinhold Math. Ser., Vol. 24, Van Nostrand–Reinhold, New York, 1969.