JOURNAL OF APPROXIMATION THEORY 34, 12–36 (1982)

Nonlinear L_p Approximation for $1 < p < 2$

M. AKHLAGHI

Department of Mathematics, Shiraz University, Shiraz, Iran

AND

JERRY M. WOLFE

Department of Mathematics, University of Oregon, Eugene, Oregon 97403

Communicated by E. W. Cheney

Received August 15, 1980

INTRODUCTION

In this paper we shall consider certain questions regarding the uniqueness of best approximations and the continuity properties of the best approximation operator for nonlinear approximation in L_p spaces with $1 < p < 2$. In particular we shall study the following two questions:

(1) Given $L_p(\mu, [a, b])$ with μ a complete regular Borel measure on $[a, b]$ and $1 < p < 2$ and given an approximating set $M \subset L_p$, what is the topological "size" of the set of elements in L_p having unique best approximations in M?

(2) Given $f_0 \in L_p$ and M as in (1), does the best approximation operator P satisfy a Lipschitz condition on some neighborhood of f_0? That is, is there some $k > 0$ and some $\delta > 0$ (depending only on f_0) such that $\|f - f_0\| < \delta$ implies that $\|P(f) - P(f_0)\| \leqslant K \|f - f_0\|$? (If this is the case, we would say that the approximation problem is stable provided that K is not too large.)

For $2 \leqslant p < \infty$ these questions were answered in [1] and [2], respectively, for a large class of approximating families that includes generalized rational functions and the so called Γ-families [3]. Some results on uniqueness and characterization for the case $p = 1$ may be found in [4] and [5]. The techniques used in [1] and [2] relied on the continuity of the second (Frechet) derivative of the error functional with respect to the parameters. As will be seen, for $1 < p < 2$ these second derivatives need not exist much less be continuous.

0021-9045/82/010012-25\$02.00/0

Using a smoothing technique described later, we shall show, for example, that the set of elements having unique best approximations from the set R_m^n of ordinary rational functions contains an open and dense subset of L_p for each $1 < p < 2$. This extends the corresponding result in [1] to those values of p. In contrast to the results of [2], for $1 < p < 2$, we show that even for finite dimensional linear approximation, no local Lipschitz conditions holds (at least when Lebesgue measure is used). We shall show that under certain conditions an inequality of the form $\|P(f) - P(f_0)\|_p \leqslant K\|f - f_0\|_\infty$ obtains.

Approximation Problem

The approximation problem we shall consider is as follows: Let A denote a continuous map from an open subset S of R^N to the normed linear space H. Given $f \in H$ we seek an $x_0 \in S$ such that $\|A(x_0) - f\| = \inf_{x \in S} \|A(x) - f\|$, where $\|\cdot\|$ is the norm on H.

In the cases we consider, H will be $L_p(\mu, [a, b])$, where μ is a complete regular Borel measure on $[a, b]$ and $1 < p < \infty$ and usually $p < 2$. Moreover, we will assume that the map $x \to A''(x, \cdot, \cdot)$ exists and is continuous on S, where $A''(x, \cdot, \cdot)$ denotes the second Frechet derivative of A with respect to x. We will usually shorten $L_p(\mu[a, b])$ to L_p.

I. Characterization and Uniqueness of Best Approximations

For $f \in L_p$ fixed, let $F(x) = \int_a^b |A(x)(t) - f(t)|^p \, d\mu$ $p > 1$. Then finding a best approximation to f from $A(S) = \{A(x) \mid x \in S\}$ is equivalent to finding an $x \in S$ that minimizes F. The chain rule and Lebesgue's dominated convergence theorem easily yield the following formula for the derivative of F in the direction h:

$$F'(x, h) = p \int_a^b |A(x)(t) - f(t)|^{p-1} \operatorname{sgn}(A(x)(t) - f(t)) \, A'(x, h)(t) \, d\mu$$

$$\text{for all} \quad h \in R^N. \tag{1}$$

At a local minimum of F we, of course, have $F'(x, h) = 0$ for all $h \in R^N$. This immediately yields the following basic lemma.

Lemma 1. *Suppose $A(x) \in A(S)$ is a local best approximation to $f \in L_p$, where $p > 1$. Then*

$$\int_a^b |A(x)(t) - f(t)|^{p-1} \operatorname{sgn}(A(x)(t) - f(t)) \, A'(x, h)(t) \, d\mu = 0$$

$$\textit{for all} \quad h \in R^N.$$

If $p \geqslant 2$ one would also have the necessary condition that $F''(x, h, h) \geqslant 0$ for all $h \in R^N$ if x is a local minimum of F. However, for $1 < p < 2$, proceeding *formally* and differentiating (1) under the integral sign we arrive at the expression

$$\begin{aligned} p(p-1) \int_a^b \frac{(A'(x, h)(t))^2}{|E(x, t)|^{2-p}} \, d\mu \\ + p \int_a^b |E(x, t)|^{p-1} \operatorname{sgn}(E(x, t)) \, A''(x, h, h)(t) \, d\mu, \end{aligned} \tag{2}$$

where $E(x, t) = A(x)(t) - f(t)$.

At a local best approximation $E(x, t)$ will have sign changes in $[a, b]$ so that the first of the integrals in (2) may be infinite. (The second one, however, is always finite.) This indicates that $F''(x, h, h)$ may fail to exist. We should note here that the formal differentiation process does *not* yield (2) if $p = 1$ and for this reason the techniques of this paper will not apply to that case. The problem of computing $F''(x, h, h)$ for $p = 1$ is discussed in [4].

Our first task is to prove that $F''(x, h, h)$ exists and is equal to (2) if (2) is finite. This proves to be suprisingly delicate and we require several preparatory lemmas. In addition we shall adopt the following convention.

Convention. Given $f \in L_p$, $1 < p < 2$, $x \in S$, $h \in R^N$, the function $(A'(x, h)(t))^2/|E(x, t)|^{2-p}$ shall be given the value zero whenever $A'(x, h)(t) = 0$ even if $E(x, t) = 0$ at the same t.

LEMMA 2. *Let $\{g_\lambda\}$, $\lambda \in R$, $\lambda \neq 0$ be a family of μ measurable functions on $[a, b]$ that are finite valued μ. a.e. and converge to a μ.a.e. finite function g with $g(t) \geqslant 0$ on $[a, b]$ as $\lambda \to 0$. Let $G_\lambda \equiv \{t \mid g_\lambda(t) < 0\}$ and $G_0 \equiv \{t \mid g(t) = 0\}$. Then if $\mu(G_0) = \alpha$ we have $\overline{\lim}_{\lambda \to 0} \mu(G(\lambda)) \leqslant \alpha$.*

Proof. See [6, p. 21].

LEMMA 3. *For all a, b, and δ real with $b \neq 0$ and $0 < \delta < 1$ the inequality*

$$\frac{||a+b|^\delta - |a|^\delta|}{|b|} \leqslant \frac{1}{|a|^{1-\delta}}$$

is valid.

Proof. See [6, p. 22].

LEMMA 4. *For all a and b and $0 < \delta < 1$ the inequality $||a|^\delta - |b|^\delta| \leqslant |a-b|^\delta$ holds.*

Proof. Elementary. ∎

LEMMA 5. *If* $\{\Phi_\nu\}$ *is a sequence in* L_p *converging to* $\Phi \in L_p$ $1 < p < 2$ *then* $|\Phi_\nu|^{p-1} \to |\Phi|^{p-1}$ *in* L_q, *where* $q = p/(p-1)$.

Proof. Using Lemma 4 we have

$$\int_a^b ||\Phi_\nu|^{p-1} - |\Phi|^{p-1}|^q \, d\mu \leqslant \int_a^b (|\Phi_\nu - \Phi|^{p-1})^q \, d\mu = \int_a^b |\Phi_\nu - \Phi|^p \, d\mu$$
$$\to 0 \quad \text{as} \quad \nu \to \infty. \qquad \blacksquare$$

THEOREM 1. *Let* $f \in L_p(\mu, [a, b])$, $1 < p < 2$, *where* μ *is a complete regular Borel measure. Let* $x \in S$ *and* $h \in R^N$ *be fixed. Then the functional* $F(x) = \int_a^b |A(x)(t) - f(t)|^p \, d\mu$ *is twice (Gateaux) differentiable in the direction h provided that* $\int_a^b A'(x, h)(t)^2/|E(x, t)|^{2-p} \, d\mu < \infty$, *where* $E(x, t) = A(x)(t) - f(t)$ *and where the integrand is defined by our convention at all points where the numerator vanishes. In fact*

$$F''(x, h, h) = p(p-1) \int_a^b \frac{(A'(x, h)(t))^2}{|E(x, t)|^{2-p}} \, d\mu$$
$$+ p \int_a^b |E(x, t)|^{p-1} A''(x, h, h)(t) \operatorname{sgn}(E(x, t)) \, d\mu.$$

Proof. Let $x \in S$ and $h \neq 0$ in R^N be fixed. Since x is fixed for the proof, we will use $E(t)$ instead of $E(x, t)$ to denote $A(x)(t) - f(t)$. Similarly, for each λ, $E_\lambda(t)$ will denote $A(x + \lambda h)(t) - f(t)$. Now, $(1/p)\; F'(x, h) = \int_a^b |E(t)|^{p-1} \operatorname{sgn}(E(t))\, A'(x, h)(t) \, d\mu$. By definition, $F''(x, h, h) = \lim_{\lambda \to 0}((F'(x + \lambda h, h) - F'(x, h))/\lambda)$ and $F'(x + \lambda h, h) - F'(x, h)$ can be written $(1/p)[(F'(x + \lambda h, h) - F'(x, h)] = I_1(\lambda) + I_2(\lambda) + I_3(\lambda)$, where

$$I_1(\lambda) = \int_a^b |E_\lambda(t)|^{p-1} (A'(x + \lambda h, h)(t) - A'(x, h)(t)) \operatorname{sgn}(E_\lambda(t)) \, d\mu,$$

$$I_2(\lambda) = \int_a^b (|E_\lambda(t)|^{p-1} - |E(t)|^{p-1})\, A'(x, h)(t) \operatorname{sgn}(E_\lambda(t)) \, d\mu,$$

and

$$I_3(h) = \int_a^b |E(t)|^{p-1} A'(x, h)(t)(\operatorname{sgn}(E_\lambda(t)) - \operatorname{sgn}(E(t))) \, d\mu.$$

We shall consider these three integrals separately.

CLAIM 1.

$$\lim_{\lambda \to 0} (I_1(\lambda)/\lambda) = \int_a^b |E(t)|^{p-1} \operatorname{sgn}(E(t))\, A''(x, h, h)(t) \, d\mu.$$

Proof of Claim 1. Let $h_\lambda(t) = \operatorname{sgn} E_\lambda(t)$, $h_0(t) = \operatorname{sgn}(E(t))$, $\psi_\lambda(t) = (A'(x+\lambda h)(t) - A'(x, h)(t))/\lambda$, and $\psi_0(t) = A''(x, h, h)(t)$. Note that $E_\lambda(t) = E(t) + \Delta_\lambda(t)$, where $\Delta_\lambda(t) = A(x+\lambda h)(t) - A(x)(t) = \lambda(A'(x, h)(t) + O(\lambda)(t))$, where $\|O(\lambda)\|_p/|\lambda| \leqslant M$ for all $0 < |\lambda| \leqslant \lambda_0$, where M is independent of λ. From this it follows that $h_\lambda(t) = h_0(t)$ except perhaps on the set $C(\lambda) = \{t \mid |E(t)| \leqslant |\Delta_\lambda(t)|\}$. Now

$$\left| \int_a^b [|E_\lambda(t)|^{p-1} h_\lambda(t)\, \psi_\lambda(t) - |E(t)|^{p-1} h_0(t)\, \psi_0(t)]\, d\mu \right|$$

$$\leqslant \int_a^b |E_\lambda(t)|^{p-1} |h_\lambda(t)|\, |\psi_\lambda(t) - \psi_0(t)|\, d\mu$$

$$+ \int_a^b |\psi_0(t)|\, |\, |E_\lambda(t)|^{p-1} h_\lambda(t) - |E(t)|^{p-1} h_0(t)|\, d\mu$$

$$\leqslant \int_a^b |E_\lambda(t)|^{p-1} |\psi_\lambda(t) - \psi_0(t)|\, d\mu$$

$$+ \int_a^b |\psi_0(t)|\, |h_\lambda(t)|\, |\, |E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|\, d\mu$$

$$+ \int_a^b |\psi_0(t)|\, |E(t)|^{p-1} |h_\lambda(t) - h_0(t)|\, d\mu$$

$$\equiv J_1(\lambda) + J_2(\lambda) + J_3(\lambda).$$

We have the following, using Hölder's inequality, where $q = p/p - 1$:

$$J_1(\lambda) \leqslant \|E_\lambda\|_p^{p/q} \|\psi_\lambda - \psi_0\|_p,$$

$$J_2(\lambda) \leqslant \|\psi_0\|_p \|\, |E_\lambda|^{p-1} - |E|^{p-1}\|_q,$$

$$J_3(\lambda) \leqslant 2 \int_a^b |\psi_0(t)|\, |E(t)|^{p-1}\, d\mu$$

$$\leqslant 2\, \|\psi_0\|_p \int_{C(\lambda)} |E(t)|^p\, d\mu^{1/q} \leqslant 2\, \|\psi_0\|_p\, \|\Delta_\lambda\|_p^{p/q}.$$

Then Lemma 5 and the existence of $A''(x, ,)$ imply that $J_i(\lambda) \to 0$, $i = 1, 2, 3$, and claim 1 is proved.

CLAIM 2.

$$\lim_{\lambda \to 0} \frac{I_2(\lambda)}{\lambda} = (p-1) \int_a^b \frac{A'(x, h)(t))^2}{|E(t)|^{2-p}}\, d\mu$$

if the integral if finite.

Proof of Claim 2. Assume the integral is finite. Define $g_\lambda(t)$ by $g_\lambda(t) = [|E_\lambda(t)|^{p-1} - |E(t)^{p-1}]\ \mathrm{sgn}(E_\lambda(t))\ A'(x,h)(t)\cdot(1/\lambda)$. As in the proof of claim 1, let $\Delta_\lambda(t) = A(x+\lambda h)(t) - A(x)(t) = \lambda(A'(x_1 h)(t) + O(\lambda)(t))$ so that $E_\lambda(t) = E(t) + \Delta_\lambda(t)$. Let $S(\lambda) = \{t \mid |A'(x,h)(t)| \leqslant 2\,|O(\lambda)(t)|\}$, $B^+(\lambda) = \{t \mid g_\lambda(t) > 0\}$ and $B^-(\lambda) = \{t \mid g_\lambda(t) \leqslant 0\}$.

We first note that $B^-(\lambda) \supset \{t \mid g_0(t) = 0\} \equiv B_0$, where $g_0(t) = (p-1)((A'(x,h)(t))^2/|E(t)|^{2-p})$ so that $\mu(B^-(\lambda)) \geqslant \mu(B_0)$ for all λ. On the other hand, $g_\lambda(t) \to g_0(t)$ a.e. and $g_0(t)$ is finite valued a.e. since the integral is finite and so by Lemma 2, $\overline{\lim}_{\lambda\to 0}\mu(B^-(\lambda)) \leqslant \mu(B_0)$. Thus, $\lim_{\lambda\to 0}\mu(B^-(\lambda)) = \mu(B_0)$ and hence $\mu\{t \mid g_\lambda(t) \leqslant 0 \text{ and } g_0(t) \geqslant 0\} \to 0$ as $\lambda \to 0$. We now write

$$\frac{I_2(\lambda)}{\lambda} = \int_a^b g_\lambda(t)\,d\mu$$

$$= \int_{C(\lambda)} g_\lambda(t)\,d\mu + \int_{\Omega^+(\lambda)} g_\lambda(t)\,d\mu + \int_{\Omega^-(\lambda)} g_\lambda(t)\,d\mu,$$

where

$$\Omega^+(\lambda) = B^+(\lambda) \cap S^c(\lambda) \qquad \text{and} \qquad \Omega^-(\lambda) = B^-(\lambda) \cap S^c(\lambda).$$

$$\left|\int_{S(\lambda)} g_\lambda(t)\,d\mu\right| \leqslant \int_{S(\lambda)} \left|\,|E_\lambda(t)|^{p-1} - |E(t)|^{p-1}\right| \frac{|A'(x,h)(t)|}{|\lambda|}\,d\mu$$

$$\leqslant 2\int_{S(\lambda)} \left|\,|E_\lambda(t)|^{p-1} - |E(t)|^{p-1}\right| \frac{|O(\lambda)(t)|}{|\lambda|}\,d\mu$$

$$\leqslant 2\,\|\,|E_\lambda|^{p-1} - |E|^{p-1}\|_q \frac{\|O(\lambda)\|_p}{|\lambda|}$$

$$\leqslant 2M\,\|\,|E_\lambda|^{p-1} - |E|^{p-1}\|_q$$

and by Lemma 5 this tends to zero as $\lambda \to 0$. To consider the other two integrals, first write $g_\lambda(t)$ in the form

$$g_\lambda(t) = \frac{|E(t) + \Delta_\lambda(t)|^{p-1} - |E(t)|^{p-1}}{\Delta_\lambda(t)}\,\frac{\Delta_\lambda(t)}{\lambda}\,\mathrm{sgn}(E(t) + \Delta_\lambda(t))\,A'(x,h)(t).$$

On the set $\Omega^-(\lambda)$, $\mathrm{sgn}\,\Delta_\lambda(t) = \mathrm{sgn}(\lambda)\ \mathrm{sgn}(A'(x,h)(t))$ by definition of $S(\lambda)$. Thus,

$$\int_{\Omega^-(\lambda)} \frac{|\,|E(t) + \Delta_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|}\,\frac{|\Delta_\lambda(t)|}{|\lambda|}\,|A'(x,h)(t)|\,d\mu$$

$$= \int_{\Omega^-(\lambda)} \frac{|\,|E(t) + \Delta_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|}\,|A'(x,h)(t) + O(\lambda)(t)|$$

$$\times |A'(x,h)(t)|\,d\mu$$

$$\leqslant \int_{\Omega^-(\lambda)} \frac{||E(t)+\Delta_\lambda(t)|^{p-1}-|E(t)|^{p-1}|}{|\Delta_\lambda(t)|}(A'(x,h)(t))^2\,d\mu$$

$$+\int_{\Omega^-(\lambda)} ||E(t)+\Delta_\lambda(t)|^{p-1}-|E(t)|^{p-1}|\,\frac{|O(\lambda)(t)\,A'(x,h)(t)|}{|\Delta_\lambda(t)|}\,d\mu$$

$$\leqslant \text{(by Lemma 3)} \int_{\Omega^-(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu + \int_{\Omega^-(\lambda)} ||E_\lambda(t)|^{p-1}$$

$$-|E(t)|^{p-1}|\cdot\frac{|O(\lambda)(t)|}{|\lambda|}\,\frac{|A'(x,h)(t)|}{|A'(x,h)(t)+O(\lambda)(t)|}\,d\mu,$$

where $(|E(t)+\Delta_\lambda(t)|^{p-1}-|E(t)|^{p-1})/\Delta_\lambda(t)$ is defined to be

$$(p-1)\frac{\operatorname{sgn} E(t)}{|E(t)|^{2-p}} \qquad \text{if} \quad \Delta_\lambda(t)=0.$$

But

$$\int_{\Omega^-(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu = \int_{\Omega^-(\lambda)\cap B_0^c} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu$$
$$= \int_{B^-(\lambda)\cap B_0^c} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu$$

and since $\mu(B^-(\lambda)\cap B_0^c)\to 0$ as observed above, then this integral converges to zero also. Also, since $|A'(x,h)(t)/A'(x,h)(t)+O(\lambda)(t)|\leqslant 2$ on $\Omega^-(\lambda)$ we have using Holder's inequality

$$\int_a^b ||E_\lambda(t)|^{p-1}-|E(t)|^{p-1}|\,\frac{|O(\lambda)(t)|}{|\lambda|}\,\frac{|A'(x,h)(t)|}{|A'(x,h)(t)+O(\lambda)(t)|}\,d\mu$$

$$\leqslant 2\,\||E_\lambda|^{p-1}-|E|^{p-1}\|_q\,\frac{\|O(\lambda)\|_p}{|\lambda|}\to 0 \qquad \text{as} \quad \lambda\to 0 \text{ by Lemma 5.}$$

Finally,

$$\int_{\Omega^+(\lambda)} g_\lambda(t)\,d\mu$$

$$= \int_{\Omega^+(\lambda)} \frac{||E_\lambda(t)|^{p-1}-|E(t)|^{p-1}|}{|\lambda|}\,|A'(x,h)(t)|\,d\mu \ (\text{since } g_\lambda \geqslant 0 \text{ on } \Omega^+(\lambda))$$

$$= \int_{\Omega^+(\lambda)} \frac{||E_\lambda(t)|^{p-1}-|E(t)|^{p-1}|}{|\Delta_\lambda(t)|}\,|A'(x,h)(t)|\,\frac{|\Delta_\lambda(t)|}{|\lambda|}\,d\mu$$

$$= \int_{\Omega^+(\lambda)} \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} |A'(x,h)(t)| \, |A'(x,h)(t) + O(\lambda)(t)| \, d\mu$$

$$\leqslant \int_{\Omega^+(\lambda)} \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} (A'(x,h)(t))^2 \, d\mu$$

$$+ \int_{\Omega^+(\lambda)} \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} |O(\lambda)(t)| \, |A'(x,h)(t)| \, d\mu$$

$$\equiv J_1(\lambda) + J_2(\lambda).$$

Since $|\Delta_\lambda(t)| = |\lambda| \, |A'(x,h)(t) + O(\lambda)(t)|$ and since $|A'(x,h)(t)| > 2\,|O(\lambda)(t)|$ on $\Omega^+(\lambda)$, we have

$$\frac{|O(\lambda)(t)| \, |A'(x,h)(t)|}{|\Delta_\lambda(t)|} \leqslant 2 \frac{|O(\lambda)(t)|}{|\lambda|}.$$

Hence,

$$J_2(\lambda) \leqslant 2 \int_a^b ||E_\lambda|^{p-1} - |E|^{p-1}| \frac{|O(\lambda)(t)|}{|\lambda|} \, d\mu$$

$$\leqslant 2 \, \| \, |E_\lambda|^{p-1} - |E|^{p-1} \|_q \cdot \frac{\|O(\lambda)\|_p}{|\lambda|} \to 0 \quad \text{as} \quad \lambda \to 0.$$

Finally, we need to show that

$$\lim_{\lambda \to 0} J_1(\lambda) = (p-1) \int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}} \, d\mu.$$

Consider

$$\int_a^b \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} (A'(x,h)(t))^2 \, d\mu.$$

By Lemma 3 and Lebesgue's dominated convergence theorem this converges to $(p-1) \int_a^b (A'(x,h)^2 (t)/|E(t)|^{2-p}) \, d\mu$. But

$$\int_a^b \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} (A'(x,h)(t))^2 \, d\mu$$

$$= J_1(\lambda) + \int_{\Omega^+(\lambda)^c} \frac{||E_\lambda(t)|^{p-1} - |E(t)|^{p-1}|}{|\Delta_\lambda(t)|} (A'(x,h)(t))^2 \, d\mu$$

$$\equiv J_1(\lambda) + I(\lambda).$$

$$I(\lambda) \leqslant \int_{\Omega^+(\lambda)^c} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu \leqslant \int_{S(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu + \int_{B^-(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu.$$

Now

$$\int_{B^-(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu = \int_{B^-(\lambda)\cap B_0^c} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu \to 0 \qquad \text{(as above)}$$

and

$$\int_{S(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu \to 0 \qquad \text{as} \quad \lambda \to 0$$

since it follows easily that the characteristic function $\chi_\lambda(\ )$ of $S(\lambda)$ is such that $\chi_\lambda \to \chi_0$ a.e., where χ_0 is the characteristic function of B_0 and the integrand (by definition of B_0) vanishes on this set. This finishes the proof of claim 2.

CLAIM 3. $\lim_{\lambda \to 0} I_3(\lambda)/\lambda = 0$. *Let* $e_\lambda(t) = |E(t)|^{p-1} A'(x,t)(t)(\operatorname{sgn} E_\lambda(t) - \operatorname{sgn}(E(t))$. *Then* $e_\lambda(t) = 0$ *except, perhaps, if* $|E(t) \leqslant |\Delta_\lambda(t)|$. *Let* $T(\lambda) = \{t \mid 0 < |E(t)| \leqslant |\Delta_\lambda(t)|\}$ *and let* $C_1(\lambda) = T(\lambda) \cap S(\lambda)$ *and* $C_2(\lambda) = T(\lambda) \cap S^C(\lambda)$, *where* $S(\lambda)$ *is defined as in claim* 2. *Then*

$$\frac{|I_3(\lambda)|}{\lambda} \leqslant \int_{T(\lambda)} \frac{|e_\lambda(t)|}{|\lambda|}\,d\mu = \int_{C_1(\lambda)} \frac{|e_\lambda(t)|}{|\lambda|}\,d\mu + \int_{C_2(\lambda)} \frac{|e_\lambda(t)|}{|\lambda|}\,d\mu.$$

But

$$\int_{C_1(\lambda)} \frac{|e_\lambda(t)|}{|\lambda|}\,d\mu = \int_{C_1(\lambda)} |E(t)|^{p-1} \frac{|A'(x,h)(t)|}{|\lambda|}\,d\mu \leqslant 2 \int_{C_1(\lambda)} |E(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\lambda|}\,d\mu$$

$$\leqslant 2 \int_{C_1(\lambda)} |\Delta_\lambda(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\lambda|}\, d\mu \leqslant 2\, \|\Delta_\lambda\|_p^{p/q}\, \|O(\lambda)\|_p \to 0$$

as $\Delta \to 0$.

Also

$$\int_{C_2(\lambda)} \frac{|e_\lambda(t)|}{|\lambda|}\, d\mu$$

$$\leqslant 2 \int_{C_2(\lambda)} \frac{|E(t)|^{p-1}}{|\lambda|} |A'(x,h)(t)|\, d\mu$$

$$= 2 \int_{C_2(\lambda)} \frac{|E(t)|^{p-1}}{|\Delta_\lambda(t)|} \frac{|\Delta_\lambda(t)|}{|\lambda|} |A'(x,h)(t)|\, d\mu$$

$$\leqslant 2 \int_{T(\lambda)} \frac{|E(t)|^{p-1}}{|\Delta_\lambda(t)|} (A'(x,h)(t))^2\, d\mu$$

$$+ 2 \int_{C_2(\lambda)} |E(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\Delta_\lambda(t)|} |A'(x,h)(t)|\, d\mu$$

$$\leqslant 2 \int_{T(\lambda)} \frac{|E(t)|^{p-1}}{|E(t)|} (A'(x,h)(t))^2\, d\mu$$

$$+ 2 \int_{C_2(\lambda)} |E(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\lambda|} \frac{|A'(x,h)(t)|\, d\mu}{|A'(x,h)(t) + O(\lambda)(t)|}$$

$$\leqslant 2 \int_{T(\lambda)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu + 4 \int_{C_2(\lambda)} |\Delta_\lambda(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\lambda|}\, d\mu$$

$$\equiv J_1(\lambda) + J_2(\lambda).$$

The inequality

$$\int_{C_2(\lambda)} |\Delta_\lambda(t)|^{p-1} \frac{|O(\lambda)(t)|}{|\lambda|}\, d\mu \leqslant \|\Delta_\lambda\|_p^{p/q} \frac{\|O(\lambda)\|}{|\lambda|}$$

shows that $J_2(\lambda) \to 0$ as $\lambda \to 0$. Let χ_λ denote the characteristic function $T(\lambda)$. Let $t \in [a, b]$ be such that $\chi_\lambda(t) \nrightarrow 0$ as $\lambda \to 0$. Then there exists a sequence $\{\lambda_\nu\} \to 0$ such that $\chi_{\lambda_\nu}(t) = 1$ for all ν. But then $t \in T(\lambda_\nu)$ and so $0 < |E(t)| \leqslant |\Delta_{\lambda_\nu}(t)|$ for all ν. But $|\Delta_{\lambda_\nu}(t)| \to 0$ μ.a.e. (since it converges to zero in L_p) and so $\{t \mid \chi_\lambda(t) \nrightarrow 0\}$ is a set of measure zero. Thus $\chi_\lambda \to 0$ μ.a.e. and hence $J_1(\lambda) \to 0$ as $\lambda \to 0$. This proves claim 3 and completes the proof of Theorem 1. ∎

COROLLARY 1. *Let $f \in L_p$ and suppose $x \in S$ is such that $A(x)$ is a local best approximation to f from $A(S)$. Then for each $h \in R^N$ we have*

$$(1) \quad p \int_a^b |E(t)|^{p-1} \operatorname{sgn}(E(t))\, A'(x, h)(t)\, d\mu = 0,$$

$$(2) \quad p(p-1) \int_a^b \frac{[A'(x, h)(t)]^2}{|E(t)|^{2-p}}\, d\mu + p \int_a^b |E(t)|^{p-1} \operatorname{sgn}(E(t))\, A''(x, h, h)(t)\, d\mu \geqslant 0.$$

Proof. Of course (1) is just Lemma 1. To show that (2) holds, we first note that if $\int_a^b ([A'(x, h)(t)]^2/|E(t)|^{2-p})\, d\mu = \infty$ then $(2) = \infty \geqslant 0$. If $\int_a^b ([A'(x, h)(t)]^2/|E(t)|^{2-p})\, d\mu < \infty$, then by Theorem 1, $F''(x, h, h)$ exists and equals the left hand side of (2) where as before $F(x) = \|A(x) - f\|_p^p$. The function $\Phi(\lambda) = F(x + \lambda h)$ has a local minimum at $\lambda = 0$ and is twice differentiable a $\lambda = 0$ and hence $F''(x, h, h) = \Phi''(0) \geqslant 0$. ∎

For convenience of notation we shall denote the quantity in (2) by $F''(x, h, h)$ even when its value is ∞.

Smoothing Technique

Since it is possible that

$$\int_a^b \frac{[A'(x, h)(t)]^2}{|E(t)|^{2-p}}\, d\mu = +\infty$$

we cannot depend on the continuity or even the existence of $F''(x, ., .)$. To overcome this difficulty we now introduce a perturbation in our problem for which we obtain a continuous second derivative. The following lemma defines this perturbation and establishes formulas for the necessary derivatives. The proof is a simple application of Lebesgues dominated convergence theorem and we therefore omit it.

LEMMA 6. *For $f \in L_p$, $1 < p < 2$, define $F_e(f, x)$ by $F_e(f, x) = \int_a^b (E^2(x, t) + e^2)^{p/2}\, d\mu$, where $E(x, t) = A(x)(t) - f(t)$ and $e \geqslant 0$. Then*

$$\text{(i)} \quad \lim_{e \to 0} F_e(f, x) = F(f, x) = \int_a^b |E(x, t)|^p\, d\mu,$$

(ii) $F_e'(f, x, h) = p \int_a^b \frac{E(x,t)\, A'(x,h)(t)}{(E^2(x,t)+e^2)^{(2-p)/2}}\, d\mu, \qquad h \in R^N$

(iii) $\lim_{e \to 0} F_e'(f, x, h) = F'(f, x, h)$

$$= p \int_a^b |E(x,t)|^{p-1} A'(x,h)(t) \operatorname{sgn}(E(x,t))\, d\mu,$$

(iv) $F_e''(f, x, h, h) = I_1(e) + I_2(e) + I_3(e)$, where

$$I_1(e) = p(p-1) \int_a^b \frac{[A'(x,h)(t)]^2}{(E^2(x,t)+e^2)^{(2-p)/2}}\, d\mu,$$

$$I_2(e) = p \int_a^b \frac{E(x,t)\, A''(x,h,h)(t)}{(E^2(x,t)+e^2)^{(2-p)/2}}\, d\mu,$$

and

$$I_3(e) = p(2-p) \int_a^b \frac{e^2 (A'(x,h)(t))^2}{(E^2(x,t)+e^2)^{(4-p)/2}}\, d\mu.$$

THEOREM 2. *Let the perturbation function $F_e(f, x)$ be as in Lemma* 4. *Then*

$$\lim_{e \to 0} F_e''(f, x, h, h)$$

$$= p(p-1) \int_a^b \frac{(A'(x,h)(t))^2}{|E(x,t)|^{2-p}}\, d\mu$$

$$+ p \int_a^b |E(x,t)|^{p-1} A''(x,h,h)(t) \operatorname{sgn} E(x,t)\, d\mu$$

$$\text{if} \quad \int_a^b \frac{[A'(x,h)(t)]^2}{|E(x,t)|^{2-p}}\, d\mu < \infty.$$

Otherwise $\lim_{e \to 0} F_e''(f, x, h, h) = +\infty$.

Proof. Since x and f are fixed in the proof we shall shorten $F_e(f, x)$, $F_e'(f, x, h)$, $F_e''(f, x, h, h)$ and $E(x, t)$ to $F_e(x)$, $F_e'(x, h)$, $F_e''(x, h, h)$ and $E(t)$, respectively. Let $I_1(e)$, $I_2(e)$, and $I_3(e)$ be defined as in Lemma 4.

CLAIM 1.

$$\lim_{e \to 0} I_1(e) = p(p-1) \int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu.$$

Proof. Since $(E^2(t)+e^2)^{(p-2)/2} \leqslant (E^2(t)+\bar{e}^2)^{(p-2)/2}$ if $\bar{e} < e$ the result follows immediately from the monotone convergence theorem.

CLAIM 2.

$$\lim_{e\to 0} I_2(e) = p\int_a^b |E(t)|^{p-1} A''(x, h, h)(t)\,\mathrm{sgn}(E(t))\,d\mu.$$

Proof.

$$\left|(E^2(t)+e^2)^{(p-2)/2} E(t)\,A''(x,h,h)(t)\right|$$

$$\leqslant \left|(E^2(t))^{(p-2)/2} E(t)\,A''(x,h,h)(t)\right| = |E(t)|^{p-1}\,|A''(x,h,h)(t)|$$

which has a finite integral. Since

$$\lim_{e\to 0}\,(E^2(t)+e^2)^{(p-2)/2} E(t)\,A''(x,h,h)(t)$$

$$= |E(t)|^{p-1}\,A''(x,h,h)(t)\,\mathrm{sgn}(E(t))$$

the Lebesgue dominated convergence theorem applies and the claim is proved.

CLAIM 3. $\delta \equiv \lim_{e\to 0} I_3(e) = 0$ *if*

$$\int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu < \infty.$$

Proof. Let $S(e) = \{t \mid E^2(t) > e\}$ and let $T(e)$ denote the complement of $S(e)$. Also, let $g_e(t) = e^2(A'(x,h)(t))^2(E^2(t)+e^2)^{(p-4)/2}$. Then

$$I_3(e) = p(2-p)\left[\int_{S(e)} g_e(t)\,d\mu + \int_{T(e)} g_e(t)\,d\mu\right]$$

and noting that $g_e(t)$ may be written in the form

$$g_e(t) = \frac{(A'(x,h)(t))^2}{(E^2(t)+e^2)^{(2-p)/2}}\,\frac{e^2}{E^2(t)+e^2}$$

we consider two cases:

(i) On $S(e)$ we have

$$g_e(t) \leqslant \frac{(A'(x,h)(t))^2}{(E^2(t)+e^2)^{(2-p)/2}}\,\frac{e^2}{e+e^2} \leqslant \frac{e(A'(x,h)(t))^2}{(E^2(t)+e^2)^{(2-p)/2}}.$$

But then

$$0 \leqslant \int_{S(e)} g_e(t)\, d\mu \leqslant e \int \frac{(A'(x,h)(t))^2}{(E^2(t)+e^2)^{(2-p)/2}}\, d\mu$$
$$\leqslant e \int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu \to 0$$

as $e \to 0$ since the integral

$$\int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu < \infty.$$

(ii) Let $\bar{T}(e) = \{t \in T(e) \mid A'(x,h)(t) \neq 0\}$. Then

$$0 \leqslant \int_{T(e)} g_e(t)\, d\mu = \int_{\bar{T}(e)} g_e(t)\, d\mu \leqslant \int_{\bar{T}(e)} \frac{(A'(x,h)(t))^2}{(E^2(t)+e^2)^{(2-p)/2}}\, d\mu$$
$$\leqslant \int_{\bar{T}(e)} \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu.$$

But $\mu(\bar{T}(e)) \to 0$ since $\mu\{t \mid E(t) = 0$ and $A'(x,h)(t) \neq 0\} = 0$ and hence

$$\lim_{e\to 0} \int_{T(e)} g_e(t)\, d\mu = 0.$$

Thus claim 3 is proved. Finally, since $\lim_{e\to 0} I_2(e) < \infty$ and $\delta \geqslant 0$ in any case, if

$$\int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\, d\mu = \infty,$$

then $\lim_{e\to 0} F_e''(x,h,h) = \infty$. ∎

We now need to show that under appropriate hypotheses for each $e > 0$ and $f \in L_p$ there exists at least one $x(e) \in S$ such that $F_e(f, x(e)) = \inf_{x\in S} F_e(f, x)$. We also need to consider what happens to $x(e)$ as $e \to 0$. The following concepts of normality and approximative compactness are crucial to the analysis.

DEFINITION 3. (1) A point $A(x) \in A(S)$ is called normal if A^{-1} exists on a neighborhood of $A(x)$ and is continuous at $A(x)$ and $A'(x, \cdot)$ is one to one. (2) NP will denote the set of points in L_p having at least one normal best approximation from $A(S)$.

DEFINITION 4. Let E be a normal linear space. A subset $M \subset E$ is called approximately compact if for every $x \in E$ and every sequence $\{M_\nu\} \subset M$ with $\|x - M_\nu\| \to \inf_{m \in M} \|x - M\|$, there is a subsequence M_{ν_j} and an $M_0 \in M$ such that $M_{\nu_j} \to M_0$.

Remark. It is easy to see that an approximatively compact subset of a normed linear space always has the property that each element of E has at least one closest point in M.

LEMMA 7. *Let M be an approximatively compact subset of a normed linear space E. Suppose $x \in E$ has $m \in M$ as its unique closest point in M and let $\{x_\nu\}$ be any sequence converging to x and $\{M_\nu\}$ be any corresponding sequence of closest points in M. Then $\|M_\nu - M\| \to 0$.*

Proof. See [7, p. 388].

In the following we shall assume that the set $\overline{A(S)}$ is approximatively compact. In addition we shall assume that each bounded sequence $\{y_k\}$ in $\overline{A(S)}$ has a subsequence $\{y_{k_j}\}$ converging in measure to some limit $y \in \overline{A(S)}$ and that if y is a normal element in $A(S)$, then also $\|y - y_{k_j}\| \to 0$.

Remark. This assumption is satisfied by the standard approximating families such as the rationals and exponentials (see [8] and [3], for example).

The proof of the following "existence" lemma is straightforward but rather lengthy, so we shall omit the proof. It may be found in [6].

LEMMA 8. *Let f be a given element of L_p $1 < p < 2$ and assume that f has a unique best approximation $A(x_0) \in A(S)$ from $\overline{A(S)}$ such that $A(x_0)$ is normal. Then for each $e \geqslant 0$ there exists at least one element $y_e \in \overline{A(S)}$ such that*

$$F_e(f, y_e) = \int_a^b [(y_e(t) - f(t))^2 + e^2]^{p/2} \, d\mu = \inf_{y \in \overline{A(S)}} F_e(f, y)$$

(we shall call y_e best e-approximation to f). Moreover, if $e_\nu \to 0$ and y_ν is a best e_ν-approximation to f, then $\|y_\nu - A(x_0)\|_p \to 0$ and for all ν sufficiently large, $y_\nu = A(x_\nu)$ for some $x_\nu \in S$ where $x_\nu \to x_0$ as $\nu \to \infty$.

LEMMA 9. *Let $f_0 \in L_p$ and $x_0 \in S$ be such that $\inf_{\|h\|>1} F''(f_0, x_0, h, h) = \eta > 0$, where $F''(f_0, x_0, h, h)$ is defined by the quantity (*) of Corollary 1. Then there exist neighborhoods U of f_0 and W of x_0 ($W \subset S$) such that $\inf_{\|h\|=1} \lim_{e \to 0} F_e''(f, x, h, h) \geqslant \eta/2$ for all $(f, x) \in U \times W$.*

Proof. Suppose the lemma is false. Then there exist sequences $\{h_\nu\} \subset R^N$,

$\{e_\nu\} \subset R$, $\{f_\nu\} \subset L_p$, and $\{x_\nu\} \subset S$ such that $\|h_\nu\| = 1$ and $h_\nu \to h \in R^N$, $e_\nu \to 0$, $f_\nu \to f_0$, $x_\nu \to x_0$ for which $F''_{e_\nu}(f_\nu, x_\nu, h_\nu, h_\nu) < \eta/2$ for all ν. But

$$F''_{e_\nu}(f_\nu, x_\nu, h_\nu, h_\nu) \geq p(p-1) \int_a^b \frac{(A'(x_\nu, h_\nu)(t))^2}{(E_\nu^2(t) + e_\nu^2)^{(2-p)/2}} \, d\mu$$

$$+ p \int_a^b \frac{E_\nu(t) A''(x_\nu, h_\nu, h_\nu)(t)}{(E_\nu^2(t) + e_\nu^2)^{(2-p)/2}} \, d\mu \equiv I_\nu + J_\nu,$$

where $E_\nu(t) = A(x_\nu)(t) - f_\nu(t)$.

But the integrand of I_ν is nonnegative and converges μ.a.e. to $((A'(x_0, h)(t))^2/|A(x_0)(t) - f(t)|)^{2-p}$. Hence by Fatou's lemma,

$$\varliminf_\nu I_\nu \geq p(p-1) \int_a^b \frac{(A'(x_0, h)(t))^2}{|A(x_0)(t) - f_0(t)|^{2-p}} \, d\mu.$$

Lebesgue's dominated convergence theorem show that J_ν converges to $p \int_a^b |A(x_0)(t) - f_0(t)|^{p-1} \; A''(x_0, h, h)(t) \; \operatorname{sgn}(A(x_0)(t) - f(t)) \, d\mu$ as $\nu \to \infty$. Thus, $\eta/2 \geq \varliminf_\nu \; F''_{e_\nu}(f_\nu, x_\nu, h_\nu, h_\nu) \geq F''(f_0, x_0, h, h) = \eta > 0$—a contradiction. ∎

Remark. For later purposes we note here that the conclusion of Lemma 9 can be recast in the following form: "There exist neighborhoods U of f_0 and W of x_0 ($W \subset S$) and $e_0 > 0$ such that $F''_e(F, x, h, h) \geq \eta/2$ for all $(f, x) \in U \times W$, $0 < e \leq e_0$, and $h \in R^N$ with $\|h\| = 1$."

We now have the following theorem which is one of the main results of this paper.

THEOREM 3. *Let $f_0 \in L_p$ $(\mu, [a, b])$. $1 < p < 2$, and suppose that f_0 has $A(x_0)$ as its unique best approximation from $\overline{A(S)}$, where $A(x_0)$ is normal and $\overline{A(S)}$ is approximatively compact. Moreover suppose that* $\inf_{\|h\|=1} F''(f_0, x_0, h, h) = \eta > 0$. *Then there is a neighborhood U of f_0 such that each $f \in U$ has a unique best approximation from $A(S)$.*

Proof. From the normality of $A(x_0)$ and Lemma 7 there is a neighborhood $\hat{U}$ of f_0 such that each $f \in \hat{U}$ has at least one best approximation from $A(S)$ (That is, suppose not. Then there is a sequence $\{f_\nu\}$ converging to f_0 such that no best approximations to f_ν from $\overline{A(S)}$ is in $A(S)$. Let y_ν be any best approximation to f_ν from $\overline{A(S)}$. By Lemma 7 $y_\nu \to A(x_0)$ and by normality, $y_\nu = A(x_\nu)$ for all ν sufficiently large for some $x_\nu \in S$—a contradiction. This shows in fact that we may assume that every $f \in \hat{U}$ has all its best approximation from $\overline{A(S)}$ actually in $A(S)$.)

Suppose the theorem is false. Then there exists a sequence $\{f_\nu\}$ such that $f_\nu \to f_0$ and such that each f_ν has at least two distinct best approximations in

$A(S)$, say, $A(x_\nu)$ and $A(y_\nu)$. By Lemma 7, $\{A(x_\nu)\}$ and $\{A(y_\nu)\}$ both converge to $A(x_0)$ and so by continuity of A^{-1} at $A(x_0)$, $x_\nu \to x_0$ and $y_\nu \to y_0$. By Lemma 9, there are neighborhoods U and W of f_0 and x_0 respectively and a constant $\eta > 0$ such that $\inf_{\|h\|=1} \lim_{e\to 0} F_e''(f, x, h, h) \geqslant \eta/2 > 0$ for all $(f, x) \in U \times W$ and hence $\lim_{e\to 0} F_e''(f, x, h, h) \geqslant \eta/2 > 0$ for all $(f, x) \in U \times W$ and all $\|h\| = 1$. By Taylor's theorem we have

$$F_e(f_\nu, y_\nu) = F_e(f_\nu, x_\nu) + F_e'(f_\nu, x_\nu, h_\nu) + \lambda_\nu^2/2F_e''(f_\nu, z_\nu, h_\nu, h_\nu),$$

where $\lambda_\nu = \|y_\nu - x_\nu\|$, $h_\nu = (y_\nu - x_\nu)/\lambda_\nu$ and $z_\nu = x_\nu + \theta_\nu h_\nu$ for some $\theta_\nu \in (0, 1)$. We may assume that ν is sufficiently large that $f_\nu \in U$, and x_ν, y_ν and z_ν are in W. Thus $(F_e(f_\nu, y_\nu) - F_e(f_\nu, x_\nu) - \lambda_\nu F_e'(f_\nu, x_\nu, h_\nu))/\lambda_\nu^2 = \frac{1}{2}F_e''(f_\nu, z_\nu, h_\nu, h_\nu)$. Now taking the limit on both sides as $e \to 0$ we obtain using Lemma 8, the following inequality: $F(f_\nu, y_\nu) - F(f_\nu, x_\nu) - \lambda_\nu F'(f_\nu, x_\nu, h_\nu) = \lambda_\nu^2/2 \lim_{e\to 0} F_e''(f_\nu, z_\nu, h_\nu, h_\nu)$. Now $F(f_\nu, y_\nu) = F(f_\nu, x_\nu)$ by hypothesis and $F'(f_\nu, x_\nu, h_\nu) = 0$ since x_ν is a local minimizer of $F(f_\nu, \,)$ in S and so we have $0 = (F(f_\nu, y_\nu) - F(f_\nu, x_\nu) - \lambda_\nu F'(f_\nu, x_\nu, h_\nu))/\lambda_\nu^2 = \frac{1}{2} \lim_{e\to 0} F_e''(f_\nu, z_\nu, h_\nu, h_\nu) \geqslant \eta/2 > 0$—a contradiction. ∎

In order to apply Theorem 3 we need to be able to show that there are functions with unique best approximations which also satisfy the second derivative requirements. The following two lemmas establish that the supply of these is abundant. The first of these, Lemma 10, is a standard result which we will not prove. A proof (in the special setting of this paper) may be found in [6].

LEMMA 10. *Let M be a nonempty subset of a strictly convex normed linear space E. Then the of elements having unique closest points in M is a dense subset of the set of elements having at least one closest point in M. In fact if $y \in E$ has $m \in M$ as a closest point then each element of the form $y_\lambda = \lambda y + (1 - \lambda) m \lambda \in (0, 1)$ has m as its unique closest point in M.*

LEMMA 11. *Let $f \in L_p$, $1 < p < 2$, and $x \in S$. Suppose that for each $h \neq 0$,*

$$(1) \quad \int_a^b \frac{(A'(x, h)(t))^2 \, d\mu}{|E(t)|^{2-p}} > 0,$$

$$(2) \quad (p-1) \int_a^b \frac{(A'(x, h)(t))^2}{|E(t)|^{2-p}} \, d\mu + \int_a^b |E(t)|^{p-1} A''(x, h, h)(t) \operatorname{sgn}(E(t)) \, d\mu \geqslant 0.$$

Then

$$(3)\quad \inf_{\|h\|=1} (p-1)\int_a^b \frac{(A'(x,h)(t))^2\,d\mu}{|E_\lambda(t)|^{2-p}} + \int_a^b |E_\lambda(t)|^{p-1} A''(x,h,h)(t)\,\mathrm{sgn}(E_\lambda(t))\,d\mu > 0$$

for each $\lambda \in (0,1)$, *where* $E_\lambda(t) = A(x)(t) - f_\lambda(t)$ *and where* $f_\lambda(t) = \lambda f(t) + (1-\lambda)A(x)(t)$.

Proof. Substituting f_λ for f in (2) we obtain

$$(4)\quad \frac{1}{\lambda^{2-p}}(p-1)\int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu + \lambda^{p-1}\int_a^b |E(t)|^{p-1} A''(x,h,h)(t)\,\mathrm{sgn}(E(t))\,d\mu.$$

But

$$(4) = (2) + (p-1)\frac{1-\lambda^{2-p}}{\lambda^{2-p}}\int_a^b \frac{(A'(x,h)(t))^2}{|E(t)|^{2-p}}\,d\mu - (1-\lambda^{p-1})\int_a^b |E(t)|^{p-1} A''(x,h,h)(t)\,\mathrm{sgn}(E(t)\,d\mu.$$

Clearly, if the second integral in (4) is nonnegative then $(2) \geqslant 0$ implies $(4) > 0$. If the second integral is negative then $-(1-\lambda^{p-1})\int_a^b |E(t)|^{p-1} A''(x,h,t)(t)\,\mathrm{sgn}(E(t))\,d\mu > 0$ and so again $(4) > 0$. Thus for each $h \neq 0$, $(4) > 0$. On the set $T = \{h \in R^N \mid \|h\| = 1$ and $\int_a^b (A'(x,h)(t))^2/|E(t)|^{2-p}\,d\mu = \infty\}$ we have

$$\inf_{h\in T}\left\{(p-1)\int_a^b \frac{(A'(x,h)(t))^2}{|E_\lambda(t)|^{2-p}}\,d\mu + \int_a^b |E_\lambda(t)|^{p-1} A''(x,h,h)(t)\,\mathrm{sgn}(E_\lambda(t))\,d\mu\right\} = +\infty > 1.$$

To finish the proof we note that a simple check shows that $L = \{h \in R^N\colon (1) < \infty\}$ is a subspace of R^N which we may assume is nontrivial.

The map

$$h \to (p-1)\int_a^b \frac{(A'(x,h)(t))^2}{|E_\lambda(t)|^{2-p}}\,d\mu + \int_a^b |E_\lambda(t)|^{p-1} A''(x,h,h)(t)\,\mathrm{sgn}(E_\lambda(t))\,d\mu \equiv \Phi(\lambda, h)$$

is continuous for each fixed $\lambda \in (0, 1)$, L is closed, and $\Phi(\lambda, h) > 0$ so for each $\lambda \in (0, 1)$ we have $\inf_{\|h\|=1} \Phi(\lambda, h) = \delta_\lambda > 0$. Thus, for any h with $\|h\| = 1$, $(3) \geqslant \min\{\delta_\lambda, 1\} > 0$. ∎

Remark. The condition that $\int_a^b (A'(x, h)(t))^2/|E(t)|^{2-p}\, d\mu > 0$ for each $h \neq 0$ is satisfied in the case that x is a normal point (since then $A'(x, \ )$ is a one to one map). For the standard nonlinear families (see [1] and [3] for example) any local best approximation must be normal. Thus for these families at least we see that the set of functions satisfying the hypotheses of Theorem 3 will form a dense subset of those having best approximations from $A(S)$. The main purpose (thus far) of the smoothing technique has been to establish Theorem 3. Having done this the following three results are proved exactly as in [1] and so they will only be stated. Theorems 3, 4 and 5 are extensions of the corresponding results in [1].

LEMMA 12. *Let M be an approximatively compact subset of a strictly convex normed linear space E. Suppose there exists a set $S \subset M$ with the following properties*:

(a) *The subset $T = \{x \in E \backslash M \mid P_m(x) \cap S \neq \varnothing\}$ is dense in $E \backslash M$, where $P_m(x)$ is the subset of best approximations of x from M.*

(b) *For each $x_0 \in T$, $\lambda \in (0, 1)$ and $m_0 \in P_m(x) \cap S$ there is a neighborhood. $V_\lambda(x_0)$ of $\lambda x_0 + (1 - \lambda) m_0$ such that for all $x \in V_\lambda(x_0)$, $P_m(s)$ is a singleton.*

Then the set U of all elements in E having unique best approximations in M contains an open and dense subset of E.

Proof. See [1, p. 172].

THEOREM 4. *Assume $A(S)$ is approximatively compact, that* NP *is a dense subset of $L_p(1 < p)$ and that*

$$\inf_{\|h\|=1} \int_a^b \frac{(A'(x, h)(t))^2}{|E(t)|^{2-p}}\, d\mu > 0$$

whenever $A(x) \in NP$ and $f \neq A(x)$, then the set U of all elements in L_p having a unique best approximations in $A(S)$ contains an open and dense subset of L_p.

DEFINITION 5.

$$R_m^n[a, b] = \left\{ p/q \mid p = \sum_{i=0}^{n} a_i x^i,\ q = \sum_{i=0}^{n} b_i x^i,\ q(x) > 0\ x \in [a, b] \right\}.$$

As on important application of Theorem 4 we have

THEOREM 5. *The set U of functions in $L_p[0, 1]$, $1 < p < \infty$ having unique best approximations in R_m^n contains an open and dense subset of* $L_p[0, 1]$. (Here we are using Lebesgue Measure.)

II. CONTINUITY PROPERTIES OF THE BEST APPROXIMATION OPERATOR

In this section we shall study the continuity properties of the best L_p-approximation operator for $1 < p < 2$. We shall employ the perturbation technique of the previous section in our analysis. As will be seen, this allows us full use of the implicit function theorem which is the main tool in the analysis. From the analysis of section I, we have the following "easy" result on continuity.

THEOREM 6. *Suppose $f_0 = L_p[\mu, [a, b]]$, $1 < p < \infty$, has $A(x_0)$ as its unique best approx. from $A(S)$. Further assume $A(x_0)$ is normal, $\overline{A(S)}$ is approximatively compact and that*

$$\inf_{\|h\|=1} (p-1) \int_a^b \frac{(A'(x_0, h)(t))^2}{|E(x_0, t)|^{2-p}} d\mu + \int_a^b |E(x_0, t)|^{p-1} A''(x_0, h, h)(t) \operatorname{sgn}(E(x_0, t))\, d\mu > 0.$$

Then the best projection operator P for $A(S)$ is continuous at f_0.

Proof. By Theorem 3, P is well defined on a neighborhood of f_0 and since $\overline{A(S)}$ is approximately compact, and since $A(x_0)$ is normal we have (as in the proof of Theorem 3) that $P(f_v) \to P(f_0)$. From Theorem 5 and its proof, the following is immediate.

COROLLARY 2. *In the case that μ is Lebesgue measure on $[a, b]$ and $A(S) = R_m^n$ then the best approximation operator is continuous on an open and dense subset of $L_p[a, b]$, $1 < p < \infty$.*

In the case $p \geqslant 2$, Wolfe showed in [2] that at a point f_0 as in Theorem 5 above, the operator P is in fact differentiable and hence Lipschitz continuous. Surprisingly this is not necessarily the case if $1 < p < 2$ even with a linear approximating family as will be shown presently. Throughout this section we will assume that $\overline{A(S)}$ is approximatively compact. Moreover it will be necessary to use a more precise and, unfortunately, more cumbersome notation since the function f will now be considered a variable.

As before let $F_e(f, x) = \int_a^b [(A(x)(t) - f(t))^2 + e^2]^{p/2}\, d\mu \equiv$

$\int_a^b (E^2(f,x)(t)+e^2)^{p/2}\,d\mu$ *for each* $f \in L_p$ *and* $x \in S$. *For a given* f, *necessary conditions for* $x \in S$ *to be a local minimum of* $F_e(f,)$ *are given by*

(1) $(1/p)\,F_e'(f,x,h)=0$ *for all* $h \in R^N$,

(2) $(1/p)\,F_e''(f,x,h,h) \geqslant 0$ *for all* $h \in R_N$.

Conditions (1) *and* (2) *may be cast in the following equivalent form*

(1)′ $\psi_e(f,x)=0$,

(2)′ $\langle h, D\psi_{e,x}(f,x)\rangle \geqslant 0$ *for all* $h \in R^N$ *where* $\psi_e(f,x)=(\psi_e^1(f,x),\ldots,\psi_e^N(f,x))^T$ *with*

$$\begin{aligned}\psi_e^j(f,x) &= \frac{1}{p}\frac{\partial F_e(f,x)}{\partial x_j}\\ &= \int_a^b |E^2(f,x)(t)+e^2|^{\frac{p-2}{2}} E(f,x)(t)\,\frac{\partial A}{\partial x_j}(x)(t)\,d\mu \quad j=1,\ldots,N\end{aligned}$$

and where $D\psi_{e,x}(f,x)$ *is the Jacobian matrix of* $\psi_e(f,x)$ *with respect to* x *and* $\langle\ ,\ \rangle$ *is the usual inner product on* R^N. *Let* $D\psi_{e,f}(f,x)(\ )$ *denote the Frechet derivative of* $\psi_e(f,x)$ *with respect to* f. *A simple calculation shows that for each* $g \in L_p$,

$$\begin{aligned}D\psi_{e,f}^j(f,x)(g) = -\int_a^b &\frac{g(t)(\partial A/\partial x_j)(t)}{(E^2(f,x)(t)+e^2)^{2-p/2}}\\ &\times \frac{(p-1)\,E^2(f,x)(t)+e^2}{E^2(f,x)(t)+e^2}\,d\mu, \qquad j=1,\ldots,N\end{aligned}$$

and $D\psi_{e,f}(f,x)(g)=(D\psi_{e,f}(f,x)(g),\ldots,D\psi_{e,f}^N(f,x)(g))^T$. *We now have the following basic result.*

THEOREM 7. *Assume* $f_0 \in L_p$ *and* $x_0 \in S$ *are such that* $A(x_0)$ *is normal and is the unique best approximation to* f_0 *from* $A(S)$ *and satisfies*

$$\begin{aligned}(1)\quad \inf_{\|h\|=1}\ (p-1)\int_a^b &\frac{(A'(x_0,h)(t))^2}{|E(f_0,x_0)(t)|^{2-p}}\,d\mu\\ &+\int_a^b |E(f_0,x_0)(t)|^{p-1}\\ &\times \operatorname{sgn}(E(f_0,x_0)(t))\,A''(x_0,h,h)(t)\,d\mu=\eta>0.\end{aligned}$$

Then there is an $e_0>0$ *such that for each* e *with* $0<e\leqslant e_0$ *there exist neighborhoods* U_e *of* f_0 *and* V_e *of* x_0 *and a map* $x_e\colon U_e \to V_e$ *such that*

(a) $\psi_e(f,x_e(f))=0$ *for all* $f \in U_e$,

(b) $\psi_e(f,x)=0$ *with* $f \in U_e$ *and* $x \in V_e$ *implies that* $x=x_e(f)$,

(c) *$x_e(\,)$ is differentiable on U_e with $x_e'(f)(g) = D\psi_{e,x}^{-1}(f, x_e(f))$ $(D\psi_{e,f}(f, x_e(f))(g))$ for all $g \in L_p$. In fact the map is continuously differentiable.*

Proof. We first note that by Lemma 9 there exists an $e_1 > 0$ and neighborhoods U of f_0 and V of x_0 respectively such that $0 < e \leqslant e_1$, $(f, x) \in U \times V$ and $\|h\| = 1$ imply that $F_e''(f, x, h, h) \geqslant \eta/2 > 0$. Also, by shrinking e_1 further if necessary we may assume that $F_e(f_0, \cdot)$ achieves a unique minimum at some $x_e \in V$. (This is an easy consequence of Lemma 8 and Taylor's theorem.) Thus $\psi_e(f_0, x_e) = 0$ and the condition $F_e''(f_0, x_e, h, h) \geqslant \eta/2 > 0$ for all $\|h\| = 1$ implies that $D\psi_{e,x}^{-1}(f_0, x_e)$ exists since $\langle h, D\psi_{e,x}(f, x)h\rangle = F_e''(f, x, h, h)$ so that $D\psi_{e,x}(f_0, x_e)$ is positive definite. Also the maps $(f, x) \to D\psi_{e,x}(f, x)$ and $(f, x) \to D\psi_{e,f}(f, x)$ are easily seen to be continuous on $U \times V$ in the product topology on $L_p \times R^N$. Thus, we may apply the (generalized) implicit function theorem [9, p. 230] and the result follows. ∎

Now using the differentiability of the map $x_e(\,)$ and the fact that for appropriately small e, the best e-approximation operator P_e is given by $P_e(f) = A(x_e(f))$ is follows that P_e is differentiable with respect to f. But then the generalized mean value theorem will yield that P_e is Lipschitz continuous at f_0. That is, there exists a constant K_e depending on f_0 and a neighborhood W of f_0 such that $f \in W$ implies that $\|P_e(f) - P_e(f_0)\|_p \leqslant K_e \|f - f_0\|_p$. (For the details of this argument see [2].)

It is even possible to show that we may use the same neighborhood W for all e sufficiently small. The question then is what happens as $e \to 0$? We know that $P_e(f) \to P(f)$ and $P_e(f_0) \to P(f_0)$, where P is the unperturbed best approximation operator for $A(S)$. If K_e stayed bounded then the Lipschitz continuity of P could be established. As the next example shows, however, this program will not succeed in general.

EXAMPLE. Let μ be Lebesgue measure on $[-1, 1]$, $p = 3/2$, $f(t) = t$ and consider approximating f by constant functions. That is, let $S = R$ and let $A(x)(t) = x$ for each $x \in R$, $t \in [-1, 1]$. Finally, let $g(t) = 1/\sqrt{|t|}$. First, it is clear that the unique best approximation is $x = 0$ and this is true for each of the perturbed norms also. Thus, for all $e \geqslant 0$, $x_e(f_0) = 0$. Also, for general f and x we have

$$D\psi_{e,x}(f, x) = \int_{-1}^{1} [(f(t) - x)^2 + e^2]^{-1/4} \frac{3/2(f(t) - x)^2 + e^2}{(f(t) - x)^2 + e^2} dt,$$

$$D\psi_{e,f}(f, x)(g) = \int_{-1}^{1} [(f(t) - x)^2 + e^2]^{-1/4} \frac{1}{\sqrt{|t|}} \times \frac{(1/2)(f(t) - x)^2 + e^2}{(f(t) - x)^2 + e^2} dt.$$

In particular, for λ arbitrary but sufficiently small that $x_e'(f_0 + \lambda g)$ exists we have

$$x_e'(f_0 + \lambda g)(g)$$

$$= \frac{\int_{-1}^{1} \frac{1}{|t|^{1/2}} \frac{1}{[(t+\lambda|t|^{1/2} - x_{e,\lambda})^2 + e^2]^{1/4}} \frac{(1/2)(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2}{(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2}\, dt}{\int_{-1}^{1} \frac{1}{[(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2]^{1/}} \frac{3/2(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2}{(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2}\, dt}$$

where $x_{e,\lambda} \equiv x_e(f_0 + \lambda g)$.

CLAIM. $\lim_{e,\lambda \to 0} x_e'(f_0 + \lambda g)(g) = \infty$ (where $e > 0$ though this is not really necessary). Since

$$\frac{1}{2} \leqslant \frac{(1/2)\, a^2 + b^2}{a^2 + b^2} \leqslant 1$$

and $1 \leqslant (3/2a^2 + b^2)/(a^2 + b^2) \leqslant 3/2$ if $|a| \cdot |b| > 0$ we have

$$x_e'(f_0 + \lambda g)(g) \geqslant \frac{1}{3} \frac{\int_{-1}^{1} (|t|^{1/2})^{-1} ([t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2]^{1/2})^{-1}\, dt}{\int_{-1}^{1} ([(t+\lambda|t|^{-1/2} - x_{e,\lambda})^2 + e^2]^{1/2})^{-1}\, dt}$$

Also since $x_{e,\lambda} \to x(f_0)$ as $e, \lambda \to 0$ it is sufficient to show that

$$\lim_{e,\lambda,x \to 0} \frac{\int_{-1}^{1} |t|^{-1/2}\, \Phi(e, \lambda, x)(t)\, dt}{\int_{-1}^{1} \Phi(e, \lambda, x)(t)\, dt} = \infty,$$

where $\Phi(e, \lambda, x)(t) = [(t + \lambda|t|^{-1/2} - x)^2 + e^2]^{-1/4}$. We require the following lemma.

LEMMA 13. *Let ψ be a positive even integrable function on $[-1, 1]$ that is continuous except, perhaps, at $t = 0$ and that satisfies $\psi'(t) < 0$ on $(0, 1]$. Let Φ be a positive continuous function on $[-1, 1]$. Define a function g on $[0, 1]$ by*

$$g(\delta) = \frac{\int_{J_\delta} \psi(t)\, \Phi(t)\, dt}{\int_{J_\delta} \Phi(t)\, dt}, \qquad \text{where} \quad J_\delta = [-1, -\delta] \cup [\delta, 1].$$

Then

$$g(\delta) < g(0) \qquad \text{for all} \quad 0 < \delta \leqslant 1.$$

Proof. Since Φ and $\Phi\psi$ are continuous on J_δ, for $\delta > 0$, $g'(\delta)$ exists and a straightforward calculation shows that

$$g'(\delta) = \frac{[\Phi(\delta) + \Phi(-\delta)][\int_{J_\delta} \psi(t)\,\Phi(t) - \psi(\delta)\int_{J_\delta} \Phi(t)\,dt|}{(\int_{J_\delta} \Phi(t)\,dt)^2} < 0$$

since $\psi(\delta) > \psi(t)$ for all $t \in J_\delta$, $t \neq \pm\delta$. ∎

We can now prove the claim above. Let $\{\lambda_n\}$, $\{x_n\}$ and $\{e_n\}$ be arbitrary with $\lambda_n, x_n \to 0$ and $e_n \downarrow 0$. Let Φ_n denote $\Phi(e_n, \lambda_n, x_n)$ and let $\psi = g = |t|^{-1/2}$. Then by Lemma 12 we have for each n and each $1 \geqslant \delta > 0$ that

$$\frac{\int_{-1}^{1} \psi(t)\,\Phi_n(t)\,dt}{\int_{-1}^{1} \Phi_n(t)\,dt} \geqslant \frac{\int_{J_\delta} \psi(t)\,\Phi_n(t)\,dt}{\int_{J_\delta} \Phi_n(t)\,dt}.$$

Fix δ. Then

$$\frac{\int_{J_\delta} \psi(t)\,\Phi_n(t)\,dt}{\int_{J_\delta} \Phi_n(t)\,dt} \to \frac{\int_{J_\delta} (1/|t|)\,dt}{\int_{J_\delta} (1/|t|)\,\frac{1}{2}\,dt} = \frac{-2\log(\delta)}{4(1-\delta^{1/2})}$$

since $\Phi_n(t) \to |t|^{-1/2}$ uniformly on J_δ. Thus

$$\varliminf_{n\to\infty} \frac{\int_{-1}^{1} \psi(t)\,\Phi_n(t)}{\int_{-1}^{1} \Phi_n(t)\,dt}\,dt \geqslant \frac{-2\log(\delta)}{4(1-\delta^{1/2})}$$

But $-2\log(\delta)/4(1-\delta^{1/2}) \to +\infty$ as $\delta \to 0$ so

$$\lim_{n\to\infty} \frac{\int_{-1}^{1} \psi(t)\,\Phi_n(t)\,dt}{\int_{-1}^{1} \Phi_n(t)\,dt} = +\infty. \quad ∎$$

Finally, we may use the claim to show that the map $f \to x(f)$ is not Lipschitz continuous at f_0. (Note that in this example, $x(f) = A(x(f)) = P(f)$.) To do this, suppose there were a constant K and a $\delta > 0$ such that if $\|f - f_0\| < \delta$ then $|x(f) - x(f_0)| \leqslant K\|f - f_0\|$. Then in particular, $|x(f_0 + \lambda g) - x(f_0)| \leqslant K\|g\|$ for all λ sufficiently small and positive.

But since $|x_e(f_0 + \lambda g) - x_e(f_0)|/|\lambda| = |x_e'(f_0 + \lambda^* g)(g)|$ for some $0 < \lambda^* < \lambda$ and since by the claim $\lim_{e,\sigma\to 0} x_e'(f_0 + \sigma g)(g) = +\infty$, then for all e and σ sufficiently small and positive, say, $|x_e'(f_0 + \sigma g)(g)| > 2K\|g\|$. Thus for all sufficiently small and positive e and λ, $|x_e(f_0 + \lambda g) - x_e(f_0)|/\lambda > 2K\|g\|$. But then

$$\begin{aligned} K\|g\| &\geqslant \frac{|x(f_0 + \lambda g) - x(f_0)|}{\lambda} \\ &= \lim_{e\to 0} \frac{|x_e(f_0 + \lambda g) - x_e(f_0)|}{\lambda} > 2K\|g\| \end{aligned}$$

for all λ sufficiently small and positive so we have a contradiction. Thus, P is not Lipschitz continuous at f_0.

Remark. The above example can be generalized to other values of p and other choices of f, etc. On the other hand, using the inequality, $\|x_e'(f,\ )\| \leqslant \|D\psi_{e,x}^{-1}(f, x_e(f))\| \, \|D\psi_{e,f}(f, x_e(f))\|$ it is not difficult to show that if

$$\int_a^b \frac{d\mu}{|E(f_0, x(f_0))(t)|^{2-p}} < \infty,$$

where $A(x(f_0))$ is the unique best approximation to f_0 then an inequality of the form $\|p(f)-p(f_0)\|_p \leqslant K\,\|f-f_0\|_\infty$ is valid. Thus if we are dealing with a discrete set and using counting measure and the error curve does not vanish at any of the data points. The best approximation operator will be Lipschitz continuous. For a detailed analysis of the continuity properties of the best approximation operator see [10].

References

1. J. M. Wolfe, On the unicity of nonlinear approximation in smooth spaces, *J. Approx. Theory* **12** (1974), 165–181.
2. J. M. Wolfe, Differentiability of nonlinear best approximation operators in real inner product spaces, *J. Approx. Theory* **16** (1976), 341–346.
3. J. R. Rice, "The Approximation of Functions," Vol. II, Addison–Wesley, Reading, Mass., 1969.
4. J. M. Wolfe, Nonlinear L_1 approximation of smooth functions, *J. Approx. Theory* **17** (1976), 166–176.
5. J. H. Freilich, Generalized rational approximation in interpolating subspaces of L_1, *J. Approx. Theory* **9** (1973), 380–394.
6. M. R. Aklaghi, "Some Uniqueness Problems in Nonlinear L_p-Approximation," Thesis, University of Oregon, 1978.
7. I. Singer, "Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces," Springer-Verlag, New York, 1970.
8. J. M. Wolfe, Rational L_p approximation, *J. Approx. Theory* **12** (1974), 1–5.
9. L. H. Loomis, and S. Sternberg, "Advanced Calculus," Addison–Wesley, Reading, Mass., 1968.
10. B. O. Bjornestal, Local Lipschitz continuity of the metric projection operator, preprint.