JOURNAL OF APPROXIMATION THEORY **12**, 165–181 (1974)

# On the Unicity of Nonlinear Approximation in Smooth Spaces

JERRY M. WOLFE

*Department of Mathematics, University of Oregon, Eugene, Oregon*

*Communicated by E. W. Cheney*

The application of nonlinear approximation theory in strictly convex normed linear spaces presents special problems owing to the fact that best approximations are not necessarily unique [16] and that a complete (and useful) characterization of best approximations is unknown for the commonly used approximation families. In this paper, we shall study several aspects of the unicity problem for a class of nonlinear approximation families in spaces with sufficiently smooth norms. In particular, we will consider the following problems:

(1) When does a given element of the space have a unique best approximation?

(2) How many elements (in a topological sense) have unique best approximations?

In [1] Cheney and Goldstein gave a partial answer to (1) for a class of nonlinear approximating families in a real inner product space. Their result states that if the distance from the point to the approximating set is sufficiently small (a bound is given) then the best approximation is unique. Spiess in his thesis [2] improves their bound and gives several numerical examples. Theorem 1 generalizes these results to the case of a normed linear space with a twice Fréchet-differentiable norm.

Theorems 2 and 3 answer (2) for a class of nonlinear families that include ordinary rational functions (Theorem 4) and the so-called $\Gamma$-families of Hobby and Rice [3] (Theorem 5). The basic result is that under appropriate hypotheses the set of elements having unique best approximations contains an open and dense subset of the underlying space. (A weaker version of this result is proved for the $\Gamma$-families.) In Theorem 6 we show that the number of minima of the functional $\|f - r\|$ where $f \in L_2[0, 1]$ and $r$ runs over the appropriate set of ordinary rational functions is unbounded as we vary $f$. We consider theorems 4, 5, and 6 to be the main results of this paper.

The last portion of the paper is devoted to considering which elements of the approximating family can appear as best approximations to elements other than themselves.

In what follows, $E$ and $H$ will be normed linear spaces, $S$ an open convex subset of $E$, and $A$ a twice Frechet-differentiable map from $S$ to $H$. Thus, elements of $H$ are to be approximated by elements of $A(S) = \{A(s) \mid s \in S\}$. Moreover, $H$ will be assumed to have a twice Fréchet-differentiable norm.

The first and second (Fréchet or Gateaux) derivatives of a transformation $g$ at a point $x$ will be denoted by $g'(x, \cdot)$ and $g''(x, \cdot, \cdot)$ respectively. For convenience the terms Fréchet-derivative and Gateaux-derivative will be shortened to F-derivative and G-derivative. The elementary facts about these derivatives that will be used may be found in [4].

LEMMA 1. *Let $N(g) = \| g \|^r$ for each $g \in H$ with $g \neq 0$ where $r \geqslant 2$, and let $h$ and $k$ be arbitrary in $H$. Then*

(a) *$N$ is twice* F-differentiable.

(b) $N'(tg, h) = t \mid t \mid^{r-2} N'(g, h)$, $N''(tg, h, k) = \mid t \mid^{r-2} N''(g, h, k)$ *and* $\| N'(g, \cdot)\| = r \| g \|^{r-1}$

*where $t$ is any nonzero real number.*

(c) *If $F(x) = N(A(x) - f)$ then $F'(x, k) = N'(A(x) - f, A'(x, k))$ and $F''(x, h, k) = N'(A(x) - f, A''(x, h, k)) + N''(A(x) - f, A'(x, h), A'(x, k))$. Here $f \notin A(S)$ is arbitrary.*

*Proof.* Part (a), the chain rule, and partial differentiation [5, p. 685] imply (c). To prove (a) let $R(x) = x^r$ for each real number $x$ and let $B(g) = \| g \|$ for each $g \in H$. Then $N(g) = R(B(g))$ and by the chain rule $N'(g, h) = r \| g \|^{r-1} B'(g, h)$ from which the relation $\| N'(g, \cdot)\| = r \| g \|^{r-1}$ is clear once we note that $\| B'(g, \cdot)\| = 1$ [4]. Similarly, $N''(g, \cdot, \cdot)$ exists. To finish part (b) we calculate

$$\frac{N(tg + sh) - N(tg)}{s} = \frac{\| tg + sh \|^r - \| tg \|^r}{s}$$
$$= t \mid t \mid^{r-2} \frac{\left\| g + \frac{sh}{t} \right\|^r - \| g \|^r}{(s/t)}$$

which shows that $N'(tg, h) = t \mid t \mid^{r-2} N'(g, h)$ by letting $s \to 0$. Similarly,

$$N''(tg, h, k) = \lim_{s \to 0} \frac{N'(tg + sh, k) - N'(tg, k)}{s}$$
$$= \lim_{s \to 0} \frac{\mid t \mid^{r-2}}{} \frac{N'\left(g + \frac{sh}{t}, k\right) - N'(g, k)}{s/t}$$
$$= \mid t \mid^{r-2} N''(g, h, k).$$

Q.E.D.

*Remark* 1. The point $g = 0$ is exceptional since the norm on $H$ is not G-differentiable there. However, one can verify directly that $N$ is once F-differentiable at 0 and is twice F-differentiable at 0 if $r > 2$. If $r = 2$ $N$ may fail to have two F-derivatives at 0, though in any case it is twice G-differentiable there.

We now consider the problem of determining when a local minimum $z$ of the functional $F(x) = N(A(x) - f)$ is a global minimum. We shall follow the approach of Spiess [2] and consider first the problem restricted to a ray through $z$ in a fixed direction $k$. The global problem is then handled by considering all such rays.

For $z \in S$ and $k \in E$ with $\| k \| = 1$ and $f \notin A(S)$ given, let $\ell_{z,k} = \{x \in S \mid x = z + \lambda k \text{ for some } \lambda \text{ real}\}$, $T_{z,k} = \{x \in \ell_{z,k} \mid \| A(x) - A(z)\| \leqslant 2 \| A(z) - f \|\}$, $\mu_k = \inf_{x \in T_{z,k}} N''(A(x) - f, A'(x, k), A'(x, k))$ and

$$B_k = \sup_{x \in T_{z,k}} \| A''(x, k, k)\|.$$

We shall assume that the quantity $\| A(x) - A(z)\|$ increases monotonically as $x$ moves away from $z$ along $\ell_{z,k}$. More precisely, the function

$$\Phi_k(\lambda) = \operatorname{sgn} \lambda \, \| A(z + \lambda k) - A(z)\|$$

is assumed to be strictly monotone increasing for all values of $\lambda$ such that $z + \lambda k \in S$. Note that this assumption easily implies that $T_{z,k}$ is convex.

THEOREM 1. *Let* $F(x) = N(A(x) - f) = \| A(x) - f \|^r$ *for* $x \in S$, *where* $r \geqslant 2$. *Let $z$ and $k$ be as above and suppose there is an open neighborhood $U$ of $z$ such that for all* $x \in U \cap T_{z,k}$, $F(x) > F(z)$ *unless* $x = z$. *Then if*

$$\| A(z) - f \| < \rho_k = 1/3 \, \frac{(\mu_k)^{1/(r-1)}}{r B_k},$$

*$z$ is the unique global minimizer of $F$ on* $\ell_{z,k}$.

*Proof.* Suppose there exists $z_1 \in \ell_{z,k}$ such that $F(z_1) \leqslant F(z)$. Clearly, $z_1 \in T_{z,k}$. By Taylor's Theorem,

$$F(z_1) = F(z) + F'(z, z_1 - z) + 1/2 F''(y, z_1 - z, z_1 - z)$$

for some $y = tz + (1 - t) z_1$ with $0 < t < 1$. Thus,

$$F(z_1) - F(z) = 1/2 F''(y, z_1 - z, z_1 - z)$$

since $F'(z, z_1 - z) = 0$. Therefore,

$$F''(y, z_1 - z, z_1 - z) = W^2 F''(y, k, k) \leqslant 0$$

for some real number $W$, so that

$$\begin{aligned} 0 \geqslant F''(y,k,k) &= N''(A(y)-f, A'(y,k), A'(y,k)) + N'(A(y)-f, A''(y,k,k)) \\ &\geqslant \mu_k - | N'(A(y)-f, A''(y,k,k))| \\ &\geqslant \mu_k - \| N'(A(y)-f, \cdot)\| \cdot \| A''(y,k,k)\| \\ &\geqslant \mu_k - rB_k \| A(y)-f\|^{r-1} \equiv (*). \end{aligned}$$

But $\| A(y)-f\| \leqslant \| A(y)-A(z)\| + \| A(z)-f\| \leqslant 3\| A(z)-f\|$ (Note that $y \in T_{z,k}$).

Hence $(*) \geqslant \mu_k - rB_k 3^{r-1} \| A(z)-f\|^{r-1} > 0$ since $\| A(z)-f\| < \rho_k$. Thus we have a contradiction. Q.E.D.

The following corollary is immediate from Theorem 2.1.

COROLLARY 1. *Suppose $z \in S$ and $U$ is an open neighborhood of $z$ such that $F(x) > F(z)$ for all $x \in U$ unless $x = z$. Assume that for each $k \neq 0$, the function $\Phi_k(\lambda) = \operatorname{sgn} \lambda \cdot \| A(z+\lambda k) - A(z)\|$ is strictly monotone increasing. Then if $\| A(z)-f\| < \rho = \inf_{\|k\|=1} \rho_k$, $z$ is the unique global minimizer of $F$ on $S$.*

EXAMPLE 1. Let $E = S = R$ (the set of real numbers), $H = R^2$ with the Euclidean norm and inner product $[\cdot, \cdot]$. Let $r = 2$ and define $A : S \to R^2$ by $A(x) = (x, x^2)$. Finally, let $f = \{(0, f_2)\}$.

Then $[A(x)-f, A(x)-f]$ has a relative minimum at $z = 0$ if $f_2 < 1/2$. Also, the formulas $A'(x,k) = k(1, 2x)$ and $A''(x,k,k) = k^2(0,2)$ for $k \in R$ are clear. Thus, $\| A'(x,k)\|^2 = (1+4x^2)k^2$ and $\| A''(x,k,k)\|^2 = 4k^4$ so that $B_k = 2$ if $|k| = 1$. We also have that

$$\begin{aligned} \operatorname{sgn} \lambda \| A(z+\lambda k) - A(z)\| &= \operatorname{sgn} \lambda \cdot |\lambda| \cdot |k| \cdot \sqrt{1+\lambda^2 k^2} \\ &= |k| \cdot \lambda \sqrt{1+\lambda^2 k^2} \end{aligned}$$

which is clearly strictly monotone increasing for $k \neq 0$.

$N(g) = [g,g]$ for all $g \in H$ and by direct computation $N'(g,h) = 2[g,h]$ and $N''(g,h,h) = 2[h,h]$. Hence,

$$N''(A(x)-f, A'(x,k), A'(x,k)) = 2(1+4x^2)k^2$$

which implies that $\mu_k = 2$ for $|k| = 1$. Then if

$$\| A(0)-f\| = |f_2| < 1/3\left(\frac{2}{2 \cdot 2}\right) = 1/6$$

$A(0)$ is the unique best approximation to $f$ in $A(S)$.

The following application of Corollary 1 generalizes a result of Cheney and Goldstein [1].

EXAMPLE 2. Let $T$ be a compact Hausdorff space and $m$ a regular Borel measure on $T$. Suppose $\{v_1, ..., v_n\}$ is an independent subset of $C(T)$ (the real valued continuous functions on $T$) with the property that each nonzero $g$ in span $\{v_1, ..., v_n\}$ is such that $m\{t \mid g(t) = 0\} = 0$.

Let $f: R \to R$ be twice differentiable and satisfy $M \geqslant f'(s) \geqslant \alpha > 0$ and $|f''(s)| \leqslant \gamma$ for all $s \in R$ (e.g., $f(s) = s + \arctan(s)$). Let $v(\cdot) = (v_1(\cdot), ..., v_n(\cdot))$ and for $x \in R^n$, let $[v(\cdot), x]$ denote the generalized polynomial $\sum_{i=1}^{n} x_i v_i(\cdot)$. Note that the hypotheses imply that $\|[v(\cdot), x]\|$ is a norm on $R^n$.

Define $A: R^n \to L_p(T, m)$, $p > 2$ by $A(x) = f([v(\cdot), x])$ and let

$$N(h) = \int_T |h|^p \, dm$$

for each $h \in L_p(T, m)$. We then have the formulas $A'(x, k) = f'([v(\cdot), x])[v(\cdot), k]$ and $A''(x, k, k) = f''([v(\cdot), x])\,[v(\cdot), k]^2$ from which the estimate $B \equiv \sup_{\|k\|=1} \|A''(x, k, k)\| \leqslant \gamma K$ is easily obtained where $K = (m(T))^{1/p} \cdot \max_{t \in T} \|v(t)\|^2$ where the usual Euclidean norm is being used for elements of $R^n$.

For each $t \in T$, $A(x+h)(t) - A(x)(t) = f([v(t), x+h]) - f([v(t), x]) = f'([v(t), x + \theta_t h])[v(t), h]$ where $0 < \theta_t < 1$ using the mean value theorem. Therefore,

$$\|A(x+h) - A(x)\| = \left[\int_T |f'(v(t), [x + \theta_t h])|^p \cdot |[v(t), h]|^p \, dm\right]^{1/p} \geqslant \alpha\beta \|h\|$$

where $\beta > 0$ is such that $\|[v(\cdot), h]\| \geqslant \beta \|h\|$ for all $h \in R^n$. Moreover, if we define $\psi_k(\lambda) = \|A(x + \lambda h) - A(x)\|^p \equiv N(A(x + \lambda k) - A(x))$ where $k \neq 0$, we have by direct calculation and the mean value theorem that

$$\begin{aligned}
\psi_k'(\lambda) &= N'(A(x + \lambda k) - A(x), A'(x + \lambda k, k)) \\
&= \lambda p \int_T f'([v(t), x + \lambda k]) \cdot f'([v(t), x + \theta_t \lambda k]) \cdot [v(t), k]^2 \\
&\quad \cdot |[v(t), k]|^{p-2} \cdot |f'([v(t), x + \theta_t \lambda k])|^{p-2} \, dm
\end{aligned}$$

(see Lemma 4) which shows that the function $\operatorname{sgn} \lambda \cdot \|A(x + \lambda k) - A(x)\|$ is strictly monotone increasing.

Let $D = \{x : \|A(x) - A(z)\| \leqslant 2\|A(z) - g\|\}$ and $\Omega = \{k : \|k\| = 1\}$ where $z \in R^n$ and $g \in A(R^n)$ are arbitrary. $D$ is clearly closed and is also bounded since if $x \in D$, then $\alpha\beta \|x - z\| \leqslant \|A(x) - A(z)\| \leqslant 2\|A(z) - g\|$.

It is also easily seen that the map $(x, k) \to N''(A(x) - g, A'(x, k), A'(x, k)) = p \cdot (p-1) \int_T | A(x) - g |^{p-2} f'([v(t), x])^2 [v(t), k]^2 \, dm$ (see Lemma 4) is continuous and positive for each $(x, k) \in Dx\Omega$. Hence

$$\mu = \inf_{\|k\|=1} N''(A(x) - f, A'(x, k), A'(x, k)) > 0.$$

Finally, we note that if $z$ is not itself a minimizer of $F(\cdot)$, then $F$ takes on its minimum at some point $z_0$ in the interior of $D$ and we conclude that if $\| A(z_0) - g \| < (\mu/p\gamma K)^{1/p-1} \, 1/3$ (note that this number is less than or equal to the number $\rho_k$ of Theorem 1 for each $k \neq 0$) then $z_0$ is the unique global minimizer of $F$ on $R^n$.

Thus we have an example of a class of nonlinear approximating families in $L_p(T, m)$, $p \geqslant 2$, with the property that if a point is sufficiently close to the approximating set it has a unique closest point in the set.

We now consider the problem of determing the topological size of the set of elements of $H$ possessing more than one best approximation in $A(S)$. We shall need the following standard definition.

DEFINITION 1. A subset $M$ of a normed linear space $F$ is called approximatively compact if for each $f \in F$ and each sequence $\{m_k\} \subset M$ such that $\|f - m_k\| \to \inf_{m \in M} \|f - m\|$ there exists a subsequence $\{m_{k_j}\}$ and an element $m^* \in M$ such that $\| m_{k_j} - m^* \| \to 0$.

LEMMA 2. *Let $M$ be a approximatively compact subset of a normed linear space $H$. Suppose $x \in H$ has $m \in M$ as its unique closest point in $M$ and let $\{x_k\}$ be any sequence converging to $x$ and $\{m_k\}$ any corresponding sequence of closest points in $M$. Then $\| m_k - m \| \to 0$.*

*Proof.* See [6, p. 388].

*Notations and Assumptions.* Unless otherwise stated the following notation and assumptions will be in force for the remainder of this paper. The symbol $E$ shall denote a fixed normed linear space, $S$ an open subset of $E$, and $A$ a twice F-differentiable map from $S$ to $H$ where $H$ is a strictly convex normed linear space with a twice F-differentiable norm. In addition it shall be assumed that $A(S)$ is approximatively compact and that the maps $x \to A''(x, \cdot, \cdot)$ and $g \to N''(g, \cdot, \cdot)$ are continuous on $S$ and $H \sim \{0\}$ respectively where

$$N(g) = \| g \|^r$$

for some $r \geqslant 2$.

THEOREM 2. *Let a map $F$ be defined by $F(f, y) = N(A(y) - f)$ for $f \in H$, $y \in S$, and let $y_0 \in S$ and $f_0 \in H$ be fixed. Assume that $A^{-1}$ exists on a (relatively) open neighborhood of $A(y_0)$ and is continuous at $A(y_0)$ and that*

$$\inf_{\|k\|=1} F''(f_0, y_0, k, k) = \eta > 0$$

*(the differentiation is with respect to $y$). Then if $A(y_0)$ is the unique best approximation to $f_0$ from $A(S)$, there is a neighborhood $V$ of $f_0$ such that $f$ has a unique best approximation in $A(S)$ for each $f \in V$.*

*Proof.* Clearly, $A(S)$ being approximatively compact implies that each $f \in H$ has at least one best approximation in $A(S)$. Suppose the theorem is false. Then there exists a sequence $\{f_n\}$ such that $f_n \to f_0$ and such that each $f_n$ has at least two distinct best approximations in $A(S)$, say $A(y_n)$ and $A(y_n')$. By Lemma 2, $\{A(y_n)\}$ and $\{A(y_n')\}$ converge to $A(y_0)$, and so by continuity of $A^{-1}$ at $y_0$, $y_n \to y_0$ and $y_n' \to y_0$.

Consider the map $(f, y) \to F''(f, y, \cdot, \cdot) = N'(A(y) - f, A''(y, \cdot, \cdot)) + N''(A(y) - f, A'(y, \cdot), A'(y, \cdot))$ from $HXS \to B(E, B(E, R))$ where $B(E, G)$ denotes the set of bounded linear transformations from the normed linear space $E$ to the normed linear space $G$. This map is easily seen to be continuous so that given $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ such that if

$$\| y - y_0 \| + \| f - f_0 \| < \delta(\epsilon),$$

then $| F''(f, y, k, k) - F''(f_0, y_0, k, k)| < \epsilon$ for all $k$ such that $\| k \| = 1$. Let $\epsilon = \eta/2$, $w = \{ y \in S : \| y - y_0 \| < \delta(\epsilon)/2\}$, and

$$U = \{f \in h : \| f - f_0 \| < \delta(\epsilon)/2\}.$$

Then for all $(f, y) \in UXW$, $F''(f, y, k, k) \geqslant \eta/2 > 2$ for each $k$ satisfying $\| k \| = 1$.

Since $f_n \to f_0$, $y_n \to y_0$, and $y_n' \to y_0$ we may assume that for all $n$, $(f_n, y_n)$ and $(f_n, y_n')$ lie in $UXW$. Now $F(f_n, y_n) = F(f_n, y_n')$ and so by Taylor's Theorem $0 = F(f_n, y_n) - F(f_n, y_n') = F'(f_n, y_n', y_n - y_n', y_n - y_n') + (1/2) F''(f_n, z, y_n - y_n', y_n - y_n')$ for some $z$ between $y_n$ and $y_n'$. Now, $F'(f_n, y_n', y_n - y_n') = 0$ since $y_n'$ is a local minimizer of $F(f_n, \cdot)$ in $S$, and so

$$0 = \frac{F(f_n, y_n) - F(f_n, y_n')}{\| y_n - y_n' \|^2}$$

$$= \frac{1}{2} F'' \left(f_n, z, \frac{y_n - y_n'}{\| y_n - y_n' \|}, \frac{y_n - y_n'}{\| y_n - y_n' \|}\right) \geqslant \eta/4 > 0$$

since $z \in W$ by convexity and $f_n \in U$. Thus we have a contradiction. Q.E.D.

*Remark* 2. We note here for later use that Theorem 2 is valid even when $S$ is not open in $E$ provided that the point $y_0$ lies in the interior of $S$. Then the condition $N'(A(y_0) - f, A'(y_0, h)) = 0$ for all $h \in E$ is still necessary and the proof is the same.

The following concept of a "normal" element of $A(S)$ is not only useful for the problem at hand, but also plays a basic role in the question of which elements of $A(S)$ can appear as best approximations to elements of $H \sim A(S)$.

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

LEMMA 3. *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 $T = \{x \in E \sim M \mid P_M(x) \cap S \neq \varnothing\}$ is dense in $E \sim M$ where $P_M(x)$ is the subset of best approximations of $x$.*

(b) *For each $x_0 \in T$, $\lambda \in (0, 1)$, and $m_0 \in P_M(x_0) \cap S$ there is a neighborhood $V_\lambda(x_0)$ of $\lambda x_0 + (1 - \lambda) m_0$ such that for each $x \in V_\lambda(x_0)$, $P_M(x)$ 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.* Let $x_0$ be in $T$ and $m_0$ in $P_m(x_0) \cap S$. Using (b) choose for each $\lambda \in (0, 1)$ a neighborhood $V_\lambda(x_0)$ of $x_\lambda \equiv \lambda x_0 + (1 - \lambda) m_0$ with $V_\lambda(x_0) \subset U$. Then let $V_0 = \cup V_\lambda(x_0)$ where the union is taken over all $x_0 \in A$, $m_0 \in P_M(x_0) \cap S$, and $\lambda \in (0, 1)$. Finally, let $V = V_0 \cup M^0$ where $M^0$ denotes the interior of $M$. Clearly $V$ is an open subset of $H$ so it suffices to show that $V$ is dense.

Let $x$ be arbitrary in $E \sim M$. Then there is a sequence $\{x_n\} \subset T$ converging to $x$ by (a). But by definition of $V$, there exists for each positive integer $n$ a $y_n \in V$ such that $\| y_n - x_n \| < 1/n$. Then $y_n \to x$ and so $V$ is dense in $E \sim M$. Similarly, if $x \in M \sim M^0$ then each neighborhood of $x$ intersects $V$. Thus there is a sequence in $V$ converging to $x$ and so $V$ is dense in $E$. Q.E.D.

THEOREM 3. *Assume that $NP$ is a dense subset of $H$ and that*

$$\inf_{\|k\|=1} N''(A(y) - f, A'(y, k), A'(y, k)) > 0$$

*whenever $A(y) \in NP$ and $f \neq A(y)$. Then the set $U$ of all elements in $H$ having unique best approximations in $A(S)$ contains an open and dense subset of $H$.*

*Proof.* We may assume that $H \neq A(S)$. Let $f \in NP \cap (H \sim A(S))$ and let $A(y)$ be any normal best approximation of $f$. By the strict convexity of $H$, each $f_\lambda = \lambda f + (1 - \lambda) A(y)$ where $0 \leqslant \lambda < 1$ has $A(y)$ as its unique best approximation [2, p. 6] and the following conditions hold: $0 = F'(f_\lambda, y, k)$ and $0 \leqslant F''(f_\lambda, y, k, k) = N'(\lambda(A(y) - f), A''(y, k, k)) + N''(\lambda(A(y) - f), A'(y, k), A'(y, k)) = \lambda^{r-2}(\lambda N'(A(y) - f, A''(y, k, k)) + N''(A(y) - f, A'(y, k), A'(y, k)))$.

Since $\inf_{\|k\|=1} N''(A(y) - f, A'(y, k), A'(y, k)) > 0$ and since the above conditions also hold for $\lambda = 1$, it follows that $\inf_{\|k\|=1} F''(f, y, k, k) > 0$ for each $0 < \lambda < 1$. Hence by Theorem 2 there is an open neighborhood $V(f, \lambda, y)$ about each $f_\lambda$ which is contained in $U$. Thus by Lemma 3 the theorem holds. Q.E.D.

LEMMA 4. *Let $(X, m)$ be an arbitrary measure space. For $p \geqslant 2$, let $N(\cdot)$ denote the map $f \to \|f\|_p^p \equiv \int_X |f|^p \, dm$. Then $N$ is twice F-differentiable on $L_p(X, m)$ and the formulas $N'(f, k) = p \int_X f |f|^{p-2} k dm$ and $N''(f, k, h) = p(p-1) \int_X |f|^{p-2} hk dm$ hold, where $f$, $h$, and $k$ are elements of $L_p(X, m)$. Furthermore, the map $f \to N''(f, \cdot, \cdot)$ is continuous everywhere. (See [7] for a proof.)*

Theorem 3 will now be applied to two important types of nonlinear approximating families. The first of these is the set of polynomial rational functions on [0, 1] with fixed degree of numerator and denominator, and the second is the class of $\Gamma$-families whose study was initiated by Hobby and Rice in 1967 [3].

DEFINITION 4.

$$R_m^n[0, 1] \equiv \{ p/q : p = a_0 + \cdots + a_n x^n, q = b_0 + \cdots + b_m x^m,$$

with $q(x) > 0$ for all $x \in [0, 1]\}$. (We shall denote this set more simply by $R_m^n$.)

DEFINITION 5. Let $\mathscr{N} = \{ p/q \in R_m^n : \dim( pQ + qP) = m + n + 1\}$ where $P = \text{span}\{1, x,..., x^n\}$, and $Q = \text{span}\{1, x,..., x^m\}$. Elements of $\mathscr{N}$ are called normal and it is shown in [8] that they comprise the normal elements of $R_m{}^n$ in the sense of Definition 3 using the maps defined below.

*Remark* 3. It is known that $p/q \in \mathscr{N}$ if and only if

$$\min\{n - \partial p, m - \partial q\} = 0$$

and $p$ and $q$ have no common factors where the symbol $\partial$ denotes "degree of" [9].

It will be shown in Corollary 2 that the elements of $\mathscr{N}$ are precisely the ones that can appear as best approximations to functions not in $R_m^n$.

Let $S = \{y = (a_0, ..., a_n, b_1, ..., b_m) \in R^{m+n+1}$ :

$1 + b_1x + \cdots + b_mx^m > 0$ for all $x \in [0, 1]\}$ and define $A : S \to L_p[0, 1]$ by $A(y) = (a_0 + \cdots + a_nx^n)/(1 + b_1x + \cdots + b_mx^m)$. Then $A(S) = R_m^n$ since if $r = (a_0 + \cdots + a_nx^n)/(b_0 + \cdots + b_mx^m)$ is in $R_m^n$ then $b_0 \neq 0$, so $r$ has a representation with $b_0 = 1$.

THEOREM 4. *The set U of functions in $L_p[0, 1]$, $p \geq 2$, having unique best approximations in $R_m{}^n$ contains an open dense subset.*

*Proof.* $S$ is easily seen to be open and $A(S)$ is weakly closed and hence approximatively compact [10 and 6, p. 368]. Also, each element of $A(S) \sim R_{m-1}^{n-1}$ is normal in the sense of Definition 3. $A$ is twice F-differentiable on $S$ with $A'(y, h) = \Delta/q^2(y)$ and

$$A''(y, h, h) = \frac{2\Delta}{q^2(y)} \frac{q(h) - 1}{q(y)}$$

where $h \in R^{m+n+1}$, $\Delta = p(y)(q(h) - 1) - p(h)\, q(y)$, and where

$$p(u) = u_1 + u_2x + \cdots + u_{n+1}x^n \quad \text{and} \quad q(u) = 1 + u_{n+2}x + \cdots + u_{m+n+1}x^m$$

for $u = (u_1, ..., u_{m+n+1})$. From the second formula it is easily established that the map $y \to A''(y, \cdot, \cdot)$ is continuous.

Suppose $f \neq A(y)$ and $A(y) \in \mathscr{N}$. Then for $k \neq 0$ in $R^{m+n+1}$ we have that the (Lebesgue) measure of $\{x : A'(y, k)(x) = 0\}$ is zero since $A'(y, \cdot)$ is one-to-one so that $A'(y, k)$ is a nonzero rational function on [0, 1]. Thus by the continuity of the map $k \to N''(A(y) - f, A'(y, k), A'(y, k))$ and the compactness of $\{k : \| k \| = 1\}$, $\inf_{\|k\|=1} N''(A(y) - f, A'(y, k), A'(y, k)) > 0$. Also, $NP = L_p[0, 1] \sim R_{m-1}^{n-1}$ as remarked above and so is dense in $H$. Hence the result follows from Theorem 3. Q.E.D.

The $\Gamma$-families of Hobby and Rice can be described as follows: A function $\gamma(t, x)$ from $T \times [0, 1]$ to the real numbers is given, where $T$ is a subset of the reals. For a fixed positive integer $n$, consider the family

$$F = \left\{ f(x) = \sum_{i=1}^{N} a_i\gamma(t_i, x) \text{: } a_i \text{ is real, and } t_i \in T \text{ for all } i \right\}.$$

Then given $g \in L_p[0, 1]$ we seek a best approximation to $g$ from $F$. However, since $F$ is not closed, in general [11, p. 43] it is necessary to consider the closure of $F$. If $T$ is compact (and if $\gamma(t, x)$ satisfies certain conditions given in Theorem 5 below) then it is known that the closure in $L_p[0, 1]$ for any $1 \leqslant p \leqslant \infty$ is given by

$$\bar{F} = \left\{ \sum_{i=1}^{k} \sum_{j=0}^{m_i} a_{ij} \gamma^{(j)}(t_i, x) \colon \sum_{i=1}^{k} (m_i + 1) \leqslant N \text{ and } t_i \in T \text{ for all } i \right\}$$

where $\gamma^{(j)}(t, x)$ denotes $(\partial^j \gamma / \partial t^j)(t, x)$. (See [8] or [9] for example).

Until recently, the parameterization of $\bar{F}$ has presented great difficulties since the natural parameterization of elements of $F$ by the $a_i$'s and $t_i$'s does not extend to $\bar{F}$ in any simple way. However, in [12], Barrar and Loeb have introduced a parameterization of $F$ that can be easily extended to $\bar{F}$. To parameterize $\bar{F}$, define a map $A : D \to C[0, 1]$ by

$$A(c_1, \ldots, c_n, a_1, \ldots, a_n)(x) \equiv A(c, a)(x)$$

$$= \frac{1}{2\pi i} \int_K \frac{c_1 z^{n-1} + \cdots + c_n}{z^n + a_1 z^{n-1} + \cdots + a_n} \gamma(z, x)\, dz$$

where $D = \{(c_1, \ldots, c_n, a_1, \ldots, a_n) \equiv (c, a) \in R^{2n} \mid z^n + a_1 z^{n-1} + \cdots + a_n$ has all its roots in $T\}$ and where $K$ is any contour in $U$ with $T$ in its interior. Now $D$ is a closed subset of $R^{2n}$ and so for differentiation one can extend $A$ to the open set $V = \{(c, a) \in R^{2n} \mid z^n + a_1 z^{n-1} + \cdots + a_n$ has all its roots in the interior of $K\}$. Since $\gamma(z, x)$ is real valued for real $z$ it follows from the residue theorem and Schwarz's principle of reflection that $A(c, a)(\cdot)$ is real valued for each $(c, a) \in V$. It is a lengthy but straightforward exercise (see [14]) to verify that the map $A$ satisfies all the hypotheses of Theorem 3. However, $D$ is not open in $R^{2n}$ so that the usual orthogonality condition for a best approximation (i.e. $N'(A(x) - f, A'(x, h)) = 0$ for all $h \in R^{2n}$) is no longer necessary in all cases. It is clear, however, that the condition is still necessary whenever the best approximation lies in the relative interior of the original family $F$. Thus using Theorem 2 (see Remark 2) and the techniques of Theorem 3 we have the following weaker version of Theorem 4.

THEOREM 5. *Let $T$ be a compact subinterval of the real line $R$ and $\gamma(z, x)$ a function on $T \times [0, 1]$ to $R$ satisfying*:

(1) *For some region $U$ of the complex plane containing $T$ the function $\gamma(z, x)$ is defined and analytic in $z$ for each fixed $x \in [0, 1]$ and real valued for real $z$.*

(2) *Each function $\gamma^{(j)}(z, x) \equiv (\partial^j \gamma / \partial z^j)(z, x)$ $j = 0, 1, \ldots, n - 1$ is jointly continuous in $z$ and $x$ on $U \times [0, 1]$.*

(3) *If any function of the form*

$$\sum_{i=1}^{p}\sum_{j=0}^{m_i} a_{ij}\gamma^{(j)}(t_i, x) + \sum_{i=p+1}^{q}\sum_{j=0}^{m_i} [a_{ij}\gamma^{(j)}(\lambda_i, x) + \overline{a_{ij}}\gamma^{(j)}(\overline{\lambda_i}, x)]$$

*is zero for all x, where* $\sum_{i=1}^{p}(m_i + 1) + 2\sum_{i=p+1}^{q}(m_i + 1) \leqslant 2n$, $t_i \in T$, *and*

$$\operatorname{Im}(\lambda_i) \neq 0, \qquad \text{then} \qquad \sum_{i=1}^{q}\sum_{j=0}^{m_i} |a_{ij}| = 0.$$

*Let W denote the set of functions possessing a best approximation in* $\hat{F} = \{\sum_{i=1}^{N} a_i\gamma(t_i, x) : a_i$ *is real and* $t_i \in T^0$ *for all* $i = 1,..., N\}$. *Then the set of functions in W having a unique best approximation in* $\hat{F}$ *contains a subset that is at once open and dense in W in the relative topology and open in* $L_p[0, 1]$ *for* $2 \leqslant p < \infty$.

*Remark* 4. For the choice $\gamma(t, x) = e^{tx}$ all the assumptions are obvious but (3). For a proof of (3) in this case see [10] and [11, p. 45]. Also using the results of Barrar and Loeb in [13] it is easy to show that Theorem 5 also applies to the exponential family for the choice $T = (-\infty, \infty)$.

We shall now consider the question of how many best approximations an element may have. For simplicity we shall restrict our attention to $R_m{}^n$ considered as a subset of $H = L_2[0, 1]$. The proof of the following lemma is quite elementary and has thus been omitted.

LEMMA 5. *Let* $L_0$ *and* $M_0$ *be closed subspaces of the Hilbert space H and let u and v be arbitrary in H. Then there exist* $x \in u + L_0$ *and* $y \in v + M_0$ *such that* $\| x - y \| = dist(u + L_0, v + M_0)$. *Moreover,* $[x - y, z] = 0$ *for any z of the form* $s - y$ *with* $s \in v + M_0$ *or* $t - x$ *with* $t \in u + L_0$.

LEMMA 6. *Let* $M_i$ $i = 1, 2,...$ *be a sequence of finite dimensional subspaces of the Hilbert space H such that* (1) $M_i \cap (\sum_{j=1}^{i-1} M_j) = (0)$ *for all* $i = 2, 3,...$ *Let* $r_i \in M_i$ *be given for* $i = 1, 2,...$ *where* $r_i \neq 0$. *Then for each n,*

$$L_n \equiv \bigcap_{i=1}^{n} (r_i + M_i{}^{\perp})$$

*is nonvoid.*

*Proof.* The proof is by induction. For $n = 1$, there is nothing to do, so assume $L_n = \bigcap_{i=1}^{n}(r_i + M_i{}^{\perp})$ is nonvoid where $n \geqslant 1$. This set is a linear manifold in $H$ and in fact it is simple to check that $L_n = f + M_1{}^{\perp} \cap \cdots \cap M_n{}^{\perp}$ where $f$ is any element of $L_n$. By Lemma 3, there exist $x \in S \equiv r_{n+1} + M_{n+1}^{\perp}$ and $y \in L_n$ such that $\| x - y \| = \operatorname{dist}(S, L_n)$ and $x - y$ is orthogonal to everything of the form $z - y$ where $z \in L_n$ and $w - x$ where $w \in S$. But $\{z - y \mid z \in L_n\} = M_1^\perp \cap \cdots \cap M_n^\perp$ and $\{w - x \mid w \in S\} = M_{n+1}^\perp$. Thus, $x - y \in (M_{n+1}^\perp)^\perp = M_{n+1}$ and $x - y \in (M_1^\perp \cap \cdots \cap M_n^\perp)^\perp = M_1 + \cdots + M_n$ so that $x - y \in M_{n+1} \cap (M_1 + \cdots + M_n) = (0)$. Thus $x = y$ and $L_{n+1} = \bigcap_{i=1}^{n+1} (r_i + M_i^\perp)$ is nonvoid. Q.E.D.

THEOREM 6. *Let $H = L_2[0, 1]$ and let $r_v = p_v/q_v \subset R_m^n$, $v = 1, 2, \ldots$, be such that $p_v$ and $q_v$ have no common factors,* $\deg q_v = m$, $\deg p_v \leqslant n < m$ *and $q_j$ and $q_k$ have no common factors unless $j = k$. Then for each $v = 1, 2, \ldots$, there is an $f_v \in H$ such that $r_1, \ldots, r_v$ are all local minima of the functional $\|f_v - \cdot\|^2$ defined on $R_m^n$.*

*Proof.* By Remark 3, each $r_v$ is normal so that it is sufficient to show that for some $f_v$ $x_1, \ldots, x_v$ are local minima of the functional $N(x) = \| A(x) - f_v \|^2$ where $A(\cdot)$ is the parameter mapping introduced earlier and $x_1, \ldots, x_v$ are the unique parameters with $A(x_i) = r_i$. Now $x$ is a local minimum of $N(x)$ if

(1) $(1/2)\, N'(x, h) = [A(x) - f, A'(x, h)] = 0$ for all $h \in R^{m+n+1}$

(2) $(1/2)\, N''(x, h, h) = [A'(x, h), A'(x, h)] + [A(x) - f, A''(x, h, h)] > 0$

for all $h$ with $\| h \| = 1$. Using the calculations in Theorem 4 and [9] we note that for each $v$, $A'(x_v, R^{m+n+1}) = P_{m+n}/q_v^2 = \{ p/q_v^2 \mid p$ is a polynomial of degree $\leqslant m + n\}$ and $A''(x_v, h, h) \in P_{2m+n}/q_v^3$. Let $M_v = P_{2m+n}/q_v^3$ $v = 1, 2, \ldots$.

*Claim.* If (3) $p_i/q_i^3 = \sum_{j=1}^{i-1} p_j/q_j^3$ $p_\ell \in P_{2m+n}$ $\ell = 1, \ldots, i$, then $p_i \equiv 0$.

*Proof.* Using (3) we get (4) $p_i (\prod_{k=1}^{i-1} q_k^3) \equiv q_i^3 (\sum_{j=1}^{i-1} p_j \prod_{k=1,\, k\neq j}^{i-1} q_k^3)$. Then $q_i^3$ divides the left-hand side of (4). But $q_i^3$ is relatively prime to $\prod_{k=1}^{i-1} q_k^3$ so we must have that $q_i^3$ divides $p_i$. But $\deg(q_i^3) = 3m > 2m + n \geqslant \deg p_i$ which implies that $p_i \equiv 0$. Thus, $(M_i \cap \sum_{j=1}^{i-1} M_j) = (0)$ and by Lemma 6, for each $N$, $L_N = \bigcap_{i=1}^{N} (r_i + M_i^\perp)$ is nonvoid. But if $f \in L_N$ then since

$$P_{2m+n}/q_i^3 \supset P_{m+n}/q_i^2$$

we have that (1) and (2) are satisfied so that $\| A(\cdot) - f \|^2$ has $x_1, \ldots, x_N$ as isolated local minima. Q.E.D.

*Remark* 5. A natural question now is whether or not for some $f \in L_2[0, 1]$ the function $\|f - \cdot\|^2$ has infinitely many minima in $R_m^n$. It is known that this is impossible in the case $m = 1$ [2] but the general case is still an open question to the best of our knowledge.

For the remainder of this paper we shall consider the reverse problem of approximation theory. That is, if $H$ is a normed linear space and $M$ a subset of $H$, then given an element $m \in M$, does there exist a point $p \neq m$ such that $m$ is a closest point to $p$ in $M$?

In what follows, the setting will be the same as for Theorem 3. That is, $H$ is strictly convex with a twice F-differentiable norm, $A : S \to H$ is twice F-differentiable, $A(S)$ is approximatively compact and the maps $x \to A''(x, \cdot, \cdot)$ and $g \to N''(g, \cdot, \cdot)$ are continuous.

THEOREM 7. *Let $x \in S$ be given and suppose that*

(a) *$A(x)$ is normal.*

(b) $A'(x, \cdot)$ has closed range in $H$ and is not onto.

(c) *The map $g \to N'(g, \cdot)$ is onto $H^*$.*

(d) $\inf_{\|k\|=1} N''(A(x) - f, A'(x, k), A'(x, k)) > 0$ *whenever* $f \neq A(x)$.

*Then there exists $f \neq A(x)$ such that $A(x)$ is the unique best approximation of $f$ in $A(S)$.*

*Remark* 6. Using Lemma 4, it can easily be shown that hypothesis (c) holds for any $L_p$ space with $\infty > p > 1$.

*Proof of Theorem* 7. By (c) and (b), pick $g \in H$ such that $\| N'(g, \cdot)\| = 1$ and $N'(g, h) = 0$ for every $h = A'(x, k)$, $k \in E$. Let $f = A(x) - g$. Then $f \neq A(x)$ and satisfies the condition $F'(f, x, k) = N'(A(x) - f, A'(x, k)) = 0$ for each $k \in E$. Letting $f_\lambda = \lambda f + (1 - \lambda) A(x)$ for each $\lambda \in [0, 1]$, and proceeding exactly as in the proof of Theorem 3, we have that $F'(f_\lambda, x, k) = 0$ for each $k \in E$ and $\inf_{\|k\|=1} F''(f, x, k, k) > 0$ for $\lambda$ sufficiently small by (d). Thus, for $\lambda$ sufficiently small and positive, $x$ is a relative minimum of the functional $F(f_\lambda, \cdot)$ defined on $S$. By the continuity of $A^{-1}$ on a relative neighborhood of $A(x)$, $A(x)$ is a relative minimum of the functional $N(\cdot - f)$ defined on $A(S)$. Thus for perhaps still smaller $\lambda$, $A(x)$ is the unique best approximation to $f$ in $A(S)$. See [9]. Q.E.D.

For the following result we shall only assume that $H$ has a once F-differentiable norm and that the map $A$ is once F-differentiable on $S$. $N$ will be defined by $N(g) = \| q \|^r$ where $r > 1$ and $g \in H$.

THEOREM 8. *Suppose $p \in A(S)$ satisfies the condition that*

$$\text{span} \bigcup_{x \in A^{-1}(p)} A'(x, E)$$

*is dense in $H$. Then $p$ cannot be a best approximation to any $f \notin A(S)$.*

*Proof.* Suppose $p$ were a best approximation to $f \notin A(S)$. Then for each $x \in S$ satisfying $A(x) = p$, $x$ is a global minimum of $F(f, \cdot) = N(A(\cdot) - f)$. Thus the necessary condition $F'(f, x, k) = N'(A(x) - f, A'(x, k)) = 0$ for all $k \in E$ holds for any such $x$. But then $N'(p - f, g) = 0$ for every $g \in \bigcup_{x \in A^{-1}A(p)} A'(x, E)$ and hence for every $g$ in the linear span of this set. But then, by denseness, $N'(p - f, \cdot) = 0$. However, $\| N'(p - f, \cdot)\| = r \| p - f \|^{r-1}$ which implies $p = f$—a contradiction. Q.E.D.

We will now apply Theorem 8 to a generalized rational family that includes $R_m^n$ as an example and to the $\Gamma$-families defined earlier. Alternate proofs for Theorems 9 and 10 may be found in [9] and [11, p. 45] respectively.

LEMMA 7. *Let $X$ be a compact Hausdorff space and suppose $\rho \in C(X)$ is one to one with $\| \rho \|_\infty \leqslant 1$. Then* $\text{span}\{1/(1 - \lambda\rho) | \lambda \in (-1, 1)\}$ *is uniformly dense in $C(X)$.*

*Proof.* Let $S$ be any compact subset of $[-1, 1]$. Using a result of Achieser [15, p. 254] it is simple to show that span $\{[1/(1 - \lambda x)] | \lambda | < 1\}$ is dense in $C([-1, 1])$ and hence also in $C(S)$ (using Tietzes' Extension Theorem). Since $X$ is compact and $[-1, 1]$ is Hausdorff it follows that $\rho^{-1}$ is continuous on $\rho(X) \equiv S$ given the relative topology. Thus for every $g \in C(X)$, $g \cdot \rho^{-1}$ is continuous on $S$. Then given $\epsilon > 0$, pick $n$, $\alpha_1, ..., \alpha_n$ and $\lambda_1, ..., \lambda_n$ with $|\lambda_i| < 1$ so that $\max_{s=\rho(x)\in S} | g \cdot \rho^{-1}(s) - \sum_{i=1}^n \alpha_i/(1 - \lambda_i\rho)| < \epsilon$. Then $\max_{x\in X} | g(x) - \sum_{i=1}^n \alpha_i/(1 - \lambda_i\rho(x)) | < \epsilon$ and we have finished. Q.E.D.

*Remark* 7. The hypotheses of Lemma 7 imply that $X$ is homeomorphic to a subset of $[-1, 1]$. Thus the possible domains in Theorem 9 below are implicitly limited to such compact sets $X$.

For the following result let $X$ be a compact Hausdorff space, $\mu$ a regular Borel measure on $X$, $\{g_1, ..., g_n\}$ and $\{h_1, ..., h_m\}$ linearly independent subsets of $C(X)$, $Q = \text{span}\{h_1, ..., h_m\}$, $P = \text{span}\{g_1, ..., g_n\}$, and $R^+ = \{p/q : p \in P, q \in Q$, and $q(x) > 0$ for all $x \in X\}$. Let $S = \{(a_1, ..., a_n, b_1, ..., b_m) : b_1h_1(x) + \cdots + b_mh_m(x) > 0$ for all $x \in X\}$ and define $A : S \to L_t(X, \mu)$, $t > 1$, by $A(a_1, ..., a_n, b_1, ..., b_m) = (a_1g_1 + \cdots + a_ng_n)/(b_1h_1 + \cdots + b_mh_m)$. Letting $s = (a_1, ..., a_n, b_1, ..., b_m)$, $p = a_1g_1 + \cdots + a_ng_n$, and $q = b_1h_1 + \cdots + b_mh_m$ we then have by a simple calculation that $A'(s, R^{n+m}) = (pQ + qP)/q^2$.

THEOREM 9. *Assume that there exists $p_0 \in P$ such that $p_0(x) > 0$ for all $x \in X$. Suppose that $r \in R^+$ is given and let $T_r = \{q \in Q : q(x) > 0$ for all $x \in X$ and there exists $p \in P$ such that $p/q = r\mu$ almost everywhere$\}$. Then if $T_r$ contains elements $q_1$ and $q_2$ such that $q_2/q_1$ is one-to-one, $r$ is not the best approximation to any element $f \in L_t(X, \mu)$ except itself for $t > 1$.*

*Proof.* Let $\rho = q_2/q_1$. We may assume $\| \rho \| \infty < 1$. Then for every $|\lambda| < 1$, $q_1 - \lambda q_2 \in T_r$ since if $p_1 \in P$ and $p_2 \in P$ are such that

$$p_1/q_1 = p_2/q_2 = r,$$

then $p_\lambda/q_\lambda = (p_1 - \lambda p_2)/(q_1 - \lambda q_2)$ is equal to $r$ also. Let $x_\lambda \in S$ be such that $A(x_\lambda) = p_\lambda/q_\lambda$. Since $A'(x_\lambda, R^{n+m}) = p_\lambda Q + q_\lambda P/q_\lambda^2$ we have that

$$(p_0/q_1) \cdot (1 - \lambda\rho)^{-1} = p_0/(q_1 - \lambda q_2) \in A'(x, R^{m+n})$$

for each $|\lambda| < 1$. By Lemma 7 and the fact that $p_0(x)/q_1(x) > 0$ for all $x \in X$, span $\{(p_0/q_1) \cdot (1 - \lambda\rho)^{-1} : |\lambda| < 1\}$ is uniformly dense in $C(X)$ and hence dense in $L_t(X, m)$ for $t > 1$. Thus the result follows from Theorem 8.
Q.E.D.

COROLLARY 2. *A nonnormal element of $R_m^n[0, 1]$ cannot be the best approximation to any $f \in L_t[0, 1]$ other than itself for $t > 1$.*

*Proof.* If $r \in R_m^n[0, 1]$ is not normal then by Remark 2, there exist relatively prime polynomials (or they may be constants) $p$ and $q$ such that $\partial p \leqslant n - 1$, $\partial q \leqslant m - 1$, $r = p/q$, and $q(x) > 0$ for all $x$. Then $r[p(1 + x)]/[q(1 + x)]$ and so $T_r$ contains the elements $q$ and $q \cdot (1 + x)$ whose quotient $\rho = 1/(1 + x)$ is clearly one-to-one on $[0, 1]$ with $\|\rho\|_\infty \leqslant 1$.
Q.E.D.

We also have the following application of Theorem 8.

THEOREM 10. *Consider the family $\bar{F} = \{\sum_{i=1}^k \sum_{j=0}^{m_i} a_{ij}\gamma^{(j)}(t_i, x) | t_i \in T$ and $\sum_{i=1}^k (m_i + 1) \leqslant N\}$ where $T$ is a compact subset of the real line, $N$ a fixed positive integer and $\gamma(t, x)$ satisfies the hypotheses of Theorem 5. In addition, assume that* span $\{\gamma(t, x) | t \in T\}$ *is uniformly dense in $C[0, 1]$. Then if an element $f \in \bar{F}$ is not normal (using the parameter map $A$ of Theorem 5) it cannot be the best approximation to any element of $L_p[0, 1]$ other than itself for $p > 1$.*

Again we shall not present a complete proof here, but mention that if $f$ is not normal then one discovers by direct calculation that

$$\bigcup_{x \in A^{-1}(f)} A'(x, R^{2n}) \supset \{\gamma(t, x) \mid t \in T\}$$

from which the result is obvious.

## REFERENCES


1. E. W. CHENEY AND A. A. GOLDSTEIN, A Note on Nonlinear Approximation Theory, "Num. Math, Differentialgleichungen — Approximations Theorie," pp. 251–255. Birhäuser, Basel, 1968.
2. JURGEN SPIESS, "Best Approximation in Strictly Convex Spaces," Thesis, University of Hamburg, 1969.
3. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rational Mech. Anal.* **24** (1967), 91–106.
4. M. M. VAINBERG, "Variational Methods for the Study of Nonlinear Operators," Holden, Day, San Francisco, 1964.
5. L. W. KANTOROVICH AND G. P. ARILOV, "Functional Analysis in Normed Linear Spaces," Macmillan C., New York, 1964.

6. I. SINGER, "Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces," Springer-Verlag, New York, 1970.
7. J. KURZWEIL, On approximation in real Banach spaces, *Stud. Math.* **14** (1954), 214–231.
8. E. W. CHENEY AND H. L. LOEB, On the continuity of rational approximation operators, *Arch. Rat. Mech. Anal.* **21**, 391–401.
9. E. W. CHENEY AND A. A. GOLDSTEIN, Mean-square approximation by generalized rational functions, *Math. Zeitschr.* **95** (1967), pp. 232–241.
10. J. WOLFE, $L_p$ Rational approximation, to appear.
11. JOHN R. RICE, "The Approximation of Functions," Vol. II, Addison-Wesley, Reading, Mass., 1969.
12. R. B. BARRAR AND H. L. LOEB, Extended varisolvence, to appear.
13. R. B. BARRAR AND H. L. LOEB, Nonlinear $L_p$ approximation, *J. Math. Anal. Appl.* **40** (1972), 427–435.
14. J. WOLFE, "The Existence, Uniqueness, and Computation of Best Approximations," Thesis, University of Washington, 1972.
15. N. I. ACHIESER, "Theory of Approximation," Fredreick Ungar, New York, 1956.
16. N. V. EFIMOV AND S. B. STECHKIN, Approximative compactness and Chebychev sets, *Dokl. Akad. Nauk. SSSR* **140** (1961), 522–524.