SMOOTHNESS OF APPROXIMATION

Richard Holmes and Bernard Kripke

INTRODUCTION

Let K be a Chebyshev subset of a Banach space $(X, \| \ \|)$. Then, by definition, K is closed and for every $x \in X$ there exists a unique $P_K(x) \in K$ satisfying the condition

$$\|x - P_K(x)\| = \inf\{\|x - k\|: k \in K\} = \text{dist}(x, K).$$

This map P_K from X to K is called the *best approximation operator* (BAO) *supported by* K; its value $P_K(x)$ at x is the *best approximation to* x *out of* K. It is easy to see that P_K is always a *closed projection*, in the sense that $P_K(x) = x$ whenever $x \in K$, and, if $x_n \to x$ and $P_K(x_n) \to y$, then $y = P_K(x)$. We are interested in the following general question: assuming that K is convex, how does $P_K(x)$ vary as a function of x?

For every closed convex subset of X to be a Chebyshev set, it is necessary and sufficient that X be reflexive and strictly convex (the deep part of this result is due to James; see the remarks in [18, Section 4]). We conjecture that these conditions are also sufficient to guarantee that P_K is always continuous. However, P_K need not be continuous whenever K is a linear Chebyshev subspace of an arbitrary Banach space. This will be shown by Example 4, in which K is a subspace of codimension 2. The weakest condition that is known (to the authors) to imply that P_K is continuous is that K be *approximatively compact*, in other words, that every minimizing sequence in K be compact [5]. In particular, every closed convex subset of a uniformly convex space has this property [6].

Very smooth BAO's are characteristic of Hilbert space. Indeed, a Banach space X of dimension greater than 2 has each of the following three properties if and only if it is a Hilbert space: (a) whenever K is a closed subspace of X, it is a Chebyshev set and P_K is a linear operator ([8]; a stronger result has been established in [19]); (b) whenever K is a closed convex subset of X, it is a Chebyshev set and P_K satisfies the Lipschitz condition $\|P_K(x) - P_K(y)\| \leq \|x - y\|$ [2], [17]; (c) whenever K is a 1-dimensional subspace of X, it is a Chebyshev set and P_K is continuously Gateaux differentiable (Theorem 3). (The case where $\dim X \leq 2$ is special because every Chevyshev subspace of codimension 1 supports a linear BAO (Theorem 3).) Thus a major part of this paper will be devoted to kinds of smooth behavior of P_K that are intermediate between continuity and the uniform smoothness that characterizes Hilbert space. There are two ways in which we can weaken the characteristic properties of Hilbert space: by requiring a weaker property to hold for all closed subspaces or convex sets, or by requiring a strong property to hold, but not uniformly.

An example of the first kind is the *uniform Lipschitz* property of approximation. A reflexive and strictly convex space X has property (UL) if there exists a constant

Received November 3, 1966.
This research was partially supported by the National Science Foundation. Some of the results were originally established in the dissertation of R. Holmes, written at M.I.T. in 1964 under the direction of Professor W. G. Strang.

λ (usually greater than the 1 that is characteristic of Hilbert space) such that for each closed convex subset K of X,

$$\|P_K(x) - P_K(y)\| \leq \lambda \|x - y\| \qquad \text{for all } x, y \in X.$$

Unfortunately, spaces with this property are rare. Not all of them are Hilbert spaces (Example 1), but we conjecture that each has an equivalent Hilbert norm. In particular, if X is an L^p-space $(1 < p < \infty)$, then X is (UL) if and only if $p = 2$ or $\dim X \leq 2$ (Theorem 5). In case X is of *infinite* dimension, one can also obtain this negative result by combining a theorem of Lindenstrauss [12] with the principal result of Murray [15].

We can further weaken the Lipschitz condition by allowing the Lipschitz constant λ to depend on x and K. P_K satisfies a *pointwise Lipschitz condition* at x if there exists a constant $\lambda(x, K)$ such that

$$\|P_K(x) - P_K(y)\| \leq \lambda(x, K) \|x - y\| \qquad \text{for all } y \in X.$$

For example, P_K is known to satisfy this condition when K is a finite-dimensional Chebyshev subspace of $(C(S), \| \ \|_\infty)$, for any compact Hausdorff space S [7], [16]. For the case where K is a finite-dimensional subspace of an L^p-space $(2 < p < \infty)$ we shall show that K has a somewhat stronger property, by proving that P_K is continuously Fréchet differentiable on a certain *proper* subset of X (Theorem 4). In case $\dim X < \infty$, $\sup\{\lambda(x, K): x \in X\} < \infty$, so that the local Lipschitz constant can be chosen to depend only on K, and not on x; but if $\dim X = \infty$, this may fail even if $\dim K = 1$ (Example 2). On the other hand, if $\dim K = \infty$, then P_K need *not* satisfy a pointwise Lipschitz condition for all x (Example 3). Finally, uniform convexity is insufficient to guarantee any of these Lipschitz properties even when $\dim X = 3$ (Example 5).

The smoothness of P_K is related in a very delicate way to the smoothness and rotundity of $\| \ \|$. The approximative properties of K can be altered radically by replacement of $\| \ \|$ with a topologically equivalent norm—this should be borne in mind in connection with our conjecture about (UL) spaces. Indeed, Klee [10] has shown that if K is any Chebyshev hyperplane in a nonreflexive space $(X, \| \ \|)$, then X can be given an equivalent norm under which the only elements of X that have best approximations out of K are those in K itself.

In the first of the two main sections of this paper, we establish results about the smoothness of best approximation in general Banach spaces. These results include a theory of differentiable BAO's in smoothly normed spaces, which is applied in the second section to the study of L^p-spaces, and to the counterexamples.

1. GENERAL RESULTS

Throughout this paper, X will be a real Banach space and X' its dual space. S, the unit sphere of X, consists of all vectors of norm one; S' is the corresponding subset of X'. In either X or X', θ denotes the zero element. For any two sets A and B, $A \setminus B$ denotes the set-theoretic difference $\{x: x \in A, x \notin B\}$. By a "subspace" (respectively, "convex set") we mean a proper, closed, linear (respectively, convex) subset of X. Finally, we denote real n-dimensional Euclidean space by R^n.

We begin with a few elementary propositions about the Lipschitz continuity of BAO's.

PROPOSITION 1. *Let* M *be a Chebyshev subspace of* X. *Then* P_M *is Lipschitzian on* X *if (and only if) it is uniformly continuous on* X. *Consequently,* P_M *is Lipschitzian if (and only if) it is bounded on some parallel set of*

$$\ker P_M = \{x \in X\colon P_M(x) = \theta\}.$$

Proof. The first part follows almost directly from the fact that P_M is homogeneous: $P_M(cx) = c\,P_M(x)$ for all $x \in X$ and $c \in R^1$. Hence, if $\delta(\varepsilon)$ is a modulus of uniform continuity for P_M, then $1/\delta(1)$ is a Lipschitz constant for P_M. To prove the second part: by hypothesis there exist positive constants γ and c such that

$$\sup\{\|P_M(y)\|\colon \operatorname{dist}(y, \ker P_M) \leq \gamma\} \leq c.$$

We observe that for $t \in R^1$, $\operatorname{dist}(ty, \ker P_M) = |t| \operatorname{dist}(y, \ker P_M)$, since P_M is homogeneous. Now choose $\varepsilon > 0$ and set $\delta = \varepsilon\gamma/c$. Suppose that $\|x - y\| < \delta$. Then, since $x - P_M(x) \in \ker P_M$,

$$\operatorname{dist}(y - P_M(x), \ker P_M) \leq \delta, \quad \text{or} \quad \operatorname{dist}(c\varepsilon^{-1}(y - P_M(x)), \ker P_M) \leq \gamma.$$

Thus $\|P_M(c\varepsilon^{-1}(y - P_M(x)))\| \leq c$, whence

$$\|P_M(y) - P_M(x)\| = \|P_M(y - P_M(x))\| \leq \varepsilon. \quad \blacksquare$$

The next two results exhibit the *local* nature of Lipschitz continuity for BAO's. The first says that a locally pointwise Lipschitzian BAO is actually pointwise Lipschitzian, and the second that a BAO is Lipschitzian if it is uniformly locally pointwise Lipschitzian on $S \cap \ker P_M$.

PROPOSITION 2. *Let* M *be a Chebyshev subspace of* X. *Suppose that for some* $x \in X$ *there exist* λ *and* $\delta > 0$ *such that* $\|P_M(x) - P_M(y)\| \leq \lambda\|x - y\|$ *whenever* $\|x - y\| < \delta$. *Then, for all* $y \in X$,

$$\|P_M(x) - P_M(y)\| \leq \max(\lambda, 2 + 4\|x\|/\delta)\,\|x - y\|.$$

Proof. If $\|x - y\| \geq \delta$, then

$$\|x\| \leq (\|x\|/\delta)\,\|x - y\| \quad \text{and} \quad \|y\| \leq \|x\| + \|x - y\| \leq (1 + \|x\|/\delta)\,\|x - y\|.$$

Now, for any $z \in X$,

$$\|P_M(z)\| \leq \|P_M(z) - z\| + \|z\| \leq \|\theta - z\| + \|z\| = 2\|z\|.$$

Therefore $\|P_M(x) - P_M(y)\| \leq 2(\|x\| + \|y\|) \leq 2(1 + 2\|x\|/\delta)\,\|x - y\|$.

COROLLARY. *With* M *as above, assume that there exist* λ, $\delta > 0$ *such that if* $x \in S \cap \ker P_M$ *and* $\|x - y\| < \delta$, *then* $\|P_M(y)\| \leq \lambda\,\|x - y\|$. *Then* P_M *is Lipschitzian.*

Proof. If $x \in M$, then

$$\|P_M(x) - P_M(y)\| = \|x - P_M(y)\| = \|P_M(x - y)\| \leq 2\,\|x - y\|.$$

If $x \notin M$, then, putting $c = \|x - P_M(x)\|$ and $m = P_M(x)$, we find that

$$\|P_M(x) - P_M(y)\| = c\,\|P_M(c^{-1}(x - m)) - P_M(c^{-1}(y - m))\|$$

$$\leq c \max(\lambda, 2 + 4/\delta)\,\|c^{-1}(x - m) - c^{-1}(y - m)\| = \max(\lambda, 2 + 4/\delta)\,\|x - y\|,$$

according to Proposition 2. That is, $\max(\lambda, 2 + 4/\delta) \geq 2$ is a Lipschitz constant for P_M.

Banach spaces having the uniform Lipschitz property of approximation (property (UL)) were defined in the introduction, and it was noted that all Hilbert spaces have this property. Clearly, property (UL) is a very strong restriction on the geometry of a Banach space, and it is not obvious that there are *any* non-Hilbert (UL) spaces. We shall return to this problem in the next section, contenting ourselves for the present with the following result, which allows us to replace "convex subset" in the above definition by "line."

PROPOSITION 3. *A reflexive and strictly convex Banach space* X *has property* (UL) *if and only if there exists a constant* $\lambda \geq 1$ *such that, for all* $x, y \in X$,

$$\sup\{\|P_L(x) - P_L(y)\| : L \textit{ is a one-dimensional subspace of } X\} \leq \lambda\,\|x - y\|.$$

Proof. The hypothesis immediately implies that $\|P_L(x) - P_L(y)\| \leq \lambda\,\|x - y\|$ if L is any *line*, that is, any translate of a one-dimensional subspace. Let K be a convex subset of X, and choose $x, y \in X$. Then $L = \{t P_K(x) + (1 - t) P_K(y) : t \text{ real}\}$ is the line through $P_K(x)$ and $P_K(y)$. We see that $\mathrm{dist}(x, L) \leq \mathrm{dist}(x, K)$, and similarly for y. Denote by K' the line segment $\{t P_K(x) + (1 - t) P_K(y) : 0 \leq t \leq 1\}$; then $K' \subset K \cap L$. Since $K' \subset K$, no point in K' can be closer to x (or to y) than $P_K(x)$ (or $P_K(y)$). Thus $P_L(x)$ (respectively, $P_L(y)$) either equals $P_K(x)$ (respectively, $P_K(y)$), or else it is contained in $L \setminus K'$. By the convexity of the function $\mathrm{dist}(\cdot, L)$, $P_L(x)$ and $P_L(y)$ lie on opposite sides of K'. That is,

$$\|P_K(x) - P_K(y)\| \leq \|P_L(x) - P_L(y)\|,$$

and this establishes the proposition.

The next two propositions concern approximation in *smooth* Banach spaces. By definition, X is smooth if at each point in S there exists a *unique* supporting hyperplane of the closed unit ball. By a result of Mazur [13], X is smooth if and only if the Gateaux derivative

$$G(x, y) = \lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{t}$$

exists for all $x, y \in S$. This actually defines G for all $x \neq \theta$ and all y, and for fixed x, $G(x, y)$ is a linear functional of norm 1; that is, $G(x, \cdot) \in S'$ if $x \neq \theta$ [9, Theorem 2.2].

PROPOSITION 4. *Let* X *be a smooth Banach space. Let* $\{M_\alpha\}$ *be an arbitrary family of Chebyshev subspaces of* X, *and suppose that* M, *the closed linear span of* $\{M_\alpha\}$, *is also a Chebyshev subspace of* X. *Then* $\ker P_M = \bigcap_\alpha \ker P_{M_\alpha}$.

Proof. Clearly, $P_M(x) = \theta$ implies $P_{M_\alpha}(x) = \theta$ for each α. Suppose that $P_{M_\alpha}(x) = \theta$ for all α. Then $G(x, m_\alpha) = 0$ for each $m_\alpha \in M_\alpha$. Since each $m \in M$ is the limit of finite linear combinations of suitable m_α and $G(x, \cdot) \in S'$, we conclude that $G(x, m) = 0$. In other words, $G(x, \cdot)$ belongs to $M^\perp$, the annihilator of M

Now, by a theorem of Singer [21, Theorem 1], if M is any subspace of a Banach space X and $x \in X \setminus M$, then θ is a best approximation to x out of M if and only if there exists $f \in S' \cap M^{\perp}$ such that $f(x) = \|x\|$. In the case at hand, we may take $f(\cdot) = G(x, \cdot)$. ■

In the next section, we shall apply the following result to the problem of differentiable best approximation in L^p-spaces.

PROPOSITION 5. *Let* M *be a Chebyshev subspace of a smooth Banach space* X. *Suppose that* $M = M_1 \oplus M_2$ *and that* M_1 *and* M_2 *are both Chebyshev subspaces. For any* $x \in X$, *define* $T_x: X \to M_1$ *by* $T_x(y) = P_{M_1}(x - P_{M_2}(x - y))$. *Then* T_x *has a unique fixed point,* m^* *say, in* M_1. *Further,* $P_M(x) = m^* + P_{M_2}(x - m^*)$.

Proof. $P_M(x)$ may be written uniquely as $m_1 + m_2$, where $m_1 \in M_1$, $m_2 \in M_2$. Then $m_1 = P_M(x) - m_2 = P_M(x - m_2)$, and since $m_1 \in M_1 \subset M$, $m_1 = P_{M_1}(x - m_2)$. Similarly, $m_2 = P_M(x - m_1) = P_{M_2}(x - m_1)$, whence m_1 is a fixed point of T_x in M_1. To show uniqueness, let m be any such fixed point of T_x. Then

$$\theta = P_{M_1}(x - m - P_{M_2}(x - m)),$$

because $m \in M_1$. That is, $x - m - P_{M_2}(x - m) \in \ker P_{M_1}$; but clearly it also belongs to $\ker P_{M_2}$. By Proposition 4, $x - m - P_{M_2}(x - m) \in \ker P_M$; this shows that $P_M(x) = m + P_{M_2}(x - m)$. Thus the fixed point is unique and we can obtain it by projecting $P_M(x)$ along M_2 onto M_1. ■

Most of the remaining results of this section involve a notion of differentiable best approximation. If M is a Chebyshev subspace of a Banach space X, we define the Gateaux derivative of P_M at x in the direction y by

$$P'_M(x, y) = \lim_{t \to 0} \frac{P_M(x + ty) - P_M(x)}{t},$$

whenever the limit exists. We restrict attention to subspaces, in order to avoid the corners and edges of the general convex Chebyshev set where this derivative does not exist (consider, for example, approximation out of a triangle or line segment in 2-space). With each such subspace M we associate an idempotent map $\psi_M: X \setminus M \to S \cap \ker P_M$, defined by

$$\psi_M(x) = \|x - P_M(x)\|^{-1}(x - P_M(x)).$$

We now observe the following facts about the derivative of P_M (they are immediate consequences of the definitions):

(1) $P'_M(x, cy) = c\,P'_M(x, y)$, if either side exists;

(2) if $x \in M$, $P'_M(x, y) = P_M(y)$, for each y; and

(3) if $x \notin M$, and one or the other of $P'_M(x, y)$ and $P'_M(\psi_M(x), y)$ exists, then both exist and are equal.

THEOREM 1. *Let* M *be a Chesbyshev subspace of* X.

(a) *Suppose that for some* $x_0 \in X \setminus M$ *and some* $\delta > 0$, $P'_M(x, y)$ *exists whenever* $\|x - x_0\| < \delta$ *and* $y \in S$. *If*

$$\sup\{\|P'_M(x, y)\| : \|x - x_0\| < \delta,\ y \in S\} \equiv \lambda < \infty,$$

then to every $r < \delta$ *there corresponds a constant* λ_r *such that*

$$\|P_M(x) - P_M(z)\| \le \lambda_r \|x - z\|$$

for all z, *whenever* $\|x - x_0\| < r$. *In particular,* P_M *is pointwise Lipschitzian at* x_0.

(b) *Suppose that* $P'_M(x, y)$ *exists for all* $x = \psi_M(x)$ *and all* $y \in S$, *and that*

$$\sup\{\|P'_M(x, y)\| : x = \psi_M(x),\ y \in S\} \equiv \lambda < \infty.$$

Then P_M *is Lipschitzian.*

Proof of (a). Suppose that $\|x_1 - x_0\| < r < \delta$. Then $\|P'_M(x, y)\| \le \lambda$ if $\|x - x_1\| < \delta_1 \equiv \delta - r$. For any (fixed) $y \in S$ and $g \in S'$, we may define a differentiable function F on the interval $(-\delta_1, \delta_1)$ by $F(t) = g(P_M(x_1 + ty))$. Then

$$|g(P_M(x_1 + ty) - P_M(x_1))| = |F(t) - F(0)| = |F'(s)|\,|t|,$$

where $0 < |s| < |t|$. But

$$|F'(s)| = |g(P'_M(x_1 + ty, y))| \le \|P'_M(x_1 + ty, y)\| \le \lambda.$$

Now, if $\|z - x_1\| < \delta_1$, we set $t = \|z - x_1\|$ and $y = \|z - x_1\|^{-1}(z - x_1)$. Then the argument above shows that $|g(P_M(z) - P_M(x_1))| \le \lambda\,\|z - x_1\|$. By the Hahn-Banach theorem, we may choose g so that

$$|g(P_M(z) - P_M(x_1))| = \|P_M(z) - P_M(x_1)\|.$$

Hence we have shown that P_M is locally pointwise Lipschitzian at x_1, and so we may apply Proposition 2 to obtain our result.

Proof of (b). We first note that $P_M(x + ty)$ has a t-derivative that is uniformly bounded over all x, all real t, and all $y \in S$. This is because

$$\frac{d}{dt} P_M(x + ty) = P'_M(x + ty, y) = P'_M(\psi_M(x + ty), y),$$

and the last member is bounded by hypothesis if $x + ty \notin M$, while in the contrary case $\|P'_M(x + ty, y)\| = \|P_M(y)\| \le 2\,\|y\| = 2$. Now we choose any $u \ne v$ in X, put $y = \|u - v\|^{-1}(u - v)$, and choose any $g \in S'$. Defining F by $F(t) = g(P_M(v + ty))$, we have the inequalities

$$|F'(t)| \le \left\|\frac{d}{dt}\, g \circ P_M(v + ty)\right\| \le \max(2, \lambda).$$

Hence, by the mean-value theorem,

$$|g(P_M(u) - P_M(v))| = |F(\|u - v\|) - F(0)| \le \max(2, \lambda)\,\|u - v\|.$$

We complete the proof by applying the Hahn-Banach theorem to choose g so that $|g(P_M(u) - P_M(v))| = \|P_M(u) - P_M(v)\|$.

The preceding theorem applies in particular to the pleasant situation where P_M is continuously differentiable in Fréchet's sense. Let L(X, M) be the space of all bounded linear operators from X to M with the usual norm topology. We say that P_M is *Fréchet*-C^1 on the open set $X \setminus M$ if there exists a continuous map $U: X \setminus M \to L(X, M)$ such that $P_M'(x, y)$ exists and is given by $P_M'(x, y) = U(x)y$.

COROLLARY. *Suppose that* M *is a Chebyshev subspace of* X *and that* P_M *is Fréchet*-C^1 *on* $X \setminus M$. *Then, for each* $x \in X \setminus M$ *there exist a neighborhood* N_x *of* x, *and a constant* λ_x *such that, for every* y,

$$\|P_M(x') - P_M(y)\| \leq \lambda_x \|x' - y\| \qquad \textit{whenever } x' \in N_x.$$

Further, if $\dim X < \infty$, *then* P_M *is Lipschitzian on* X.

Proof. The function U occurring in the definition of Fréchet-C^1 is bounded on a neighborhood of x, because it is continuous. Therefore, Theorem 1(a) applies. If in addition $\dim X < \infty$, then the set $\{x: x = \psi_M(x)\}$ is compact, and so the function $x \to \|U(x)\|$ is bounded there. Hence P_M is Lipschitzian, by Theorem 1(b).

In connection with the existence of Fréchet-C^1 best approximation, we mention in passing the following result. Its proof is rather long, and we omit it, since we do not use the result in this paper. *Theorem:* Let X be smooth and strictly convex, and suppose that P_M is Fréchet-C^1 on $X \setminus M$ whenever $\dim M = 1$. Then P_M is Fréchet-C^1 on $X \setminus M$ whenever $\dim M < \infty$.

We turn now to the problem of the existence of the Gateaux derivative of P_M for the case where $\dim M < \infty$. In the next theorem we shall provide a solution to this problem for Banach spaces X with a sufficiently smooth norm. To begin with, fix $x \neq \theta$ and assume that for any $y, z \in X$ the function $N(s, t) = \|x + sy + tz\|$ is twice continuously differentiable in a neighborhood of (0, 0). Then we may define a functional $\langle \cdot, \cdot \rangle_x$ on $X \times X$, by the formula $\langle y, z \rangle_x = \frac{\partial^2 N}{\partial s \partial t}(0, 0)$. We claim that $\langle \cdot, \cdot \rangle_x$ is a continuous, symmetric, bilinear form on $X \times X$. For, letting $G(u, v) = \lim\limits_{s \to 0} \frac{\|u + sv\| - \|u\|}{s}$, we note that

$$\langle y, z \rangle_x = \lim_{t \to 0} \frac{G(x + tz, y) - G(x, y)}{t}.$$

We have already observed that $G(u, \cdot) \in S'$ for any $u \neq \theta$. Therefore, the uniform boundedness principle applies, and $\langle y, z \rangle_x$ is a bounded linear functional in y. By the assumed continuity of the second partials of N, $\langle \ , \ \rangle_x$ is symmetric and is therefore a bilinear form continuous in each variable separately. But X is complete and hence is a Baire space. Therefore $\langle \ , \ \rangle_x$ is actually continuous on $X \times X$ (Schaefer [20, Theorem 5.1]). We also note that if $c \neq 0$, then

$$\langle y, z \rangle_{cx} = |c|^{-1} \langle y, z \rangle_x.$$

THEOREM 2. *Let* $M = \operatorname{span}\{m_1, \cdots, m_n\}$ *be a finite-dimensional Chebyshev subspace of* X. *Choose an* $x = \psi_M(x)$ *and a* $y \in S$. *Set*

$$f(s_1, \cdots, s_n, t) = \left\| x + ty - \sum_1^n s_k m_k \right\|,$$

and assume that f *is* C^2 *in a neighborhood of the origin in* R^{n+1}. *Let* P_x *be the matrix of the form* $\langle \cdot, \cdot \rangle_x$ *in the basis vectors* $m_1, \cdots, m_n$; *that is, let* $(P_x)_{ij} = \langle m_i, m_j \rangle_x$. *Let* $q_x(y)$ *be the* n*-vector whose* i*th component is* $\langle m_i, y \rangle_x$. *Then, if* P_x *is invertible,* $P'_M(x, y)$ *exists and is given by the formula*

$$P'_M(x, y) = \sum_{k=1}^n (P_x^{-1} q_x(y))_k m_k.$$

Proof. During this proof we shall write $\bar{s}$ for $(s_1, \cdots, s_n)$ and $\bar{s}m$ for $s_1 m_1 + \cdots + s_n m_n$. For sufficiently small $|t|$, the relation $\bar{s}_0 m = P_M(x + ty)$ holds only if $\frac{\partial f(\bar{s}_0, t)}{\partial s_i} = 0$ for $i = 1, \cdots, n$. Put

$$G(t, \bar{s}) = \left(\frac{\partial f(\bar{s}, t)}{\partial s_1}, \cdots, \frac{\partial f(\bar{s}, t)}{\partial s_n} \right);$$

then G is C^1 in a neighborhood of the origin in R^{n+1}. Observe that $\frac{\partial G(t, \bar{s})}{\partial \bar{s}}$, the Jacobian matrix of G with respect to $\bar{s}$, is by definition of $\langle \, , \, \rangle_x$ precisely the matrix $P_{x+ty-\bar{s}m}$. By our hypotheses, $G(0, \bar{0}) = \bar{0}$, while $\frac{\partial G(0, \bar{0})}{\partial \bar{s}}$ is nonsingular. According to the implicit-function theorem (Dieudonné [4, Section 10.2]), there exist a connected neighborhood U of the origin of R' and a *unique* continuously differentiable mapping $\bar{S}$ of U into R^n such that $\bar{S}(0) = \bar{0}$ and $G(t, \bar{S}(t)) = \bar{0}$ for each $t \in U$. Since P_M is continuous (because M is finite-dimensional and so approximatively compact), and since $P_M(x) = \theta$, it follows that $P_M(x + ty) = \bar{S}(t)m$ for sufficiently small $|t|$. Finally (and again by the implicit-function theorem),

$$P'_M(x, y) = \bar{S}'(0)m = -\left(\frac{\partial G(0, \bar{0})}{\partial \bar{s}} \right)^{-1} \circ \left(\frac{\partial G(0, \bar{0})}{\partial t} \right) m,$$

which is exactly the formula for $P'_M(x, y)$ given in the statement of the theorem. ∎

Remarks on Theorem 2. (a) The rank of the form $\langle \, , \, \rangle_x$ on the subspace M is equal to the rank of the matrix of the form in any basis; therefore, the existence of $P'_M(x, y)$ does not depend on the choice of basis $\{m_1, \cdots, m_n\}$ for M. (b) The formula of the theorem shows that whenever the matrix $P_{\psi_M(x)}$ is invertible, then $P'_M(x, y)$ exists for all $y \in S$, and hence for *all* $y \in X$, and is a bounded linear operator in y; that is, $P'_M(x, \cdot) \in L(X, M)$. (c) For $x = \psi_M(x)$, the existence of P_x^{-1} is equivalent to the positive-definiteness of the form $\langle \, , \, \rangle_x$ on M. Indeed, being the Hessian form of the convex function $f(\cdot, 0)$, $\langle \, , \, \rangle_x$ is certainly positive-*semidefinite* on M. If it is actually positive-*definite* on M, it is an inner product there, and P_x is its Gram matrix. Otherwise, its rank is less than n, and so P_x is singular.

In the next section, the preceding results on differentiability will be applied to several problems on the Lipschitz continuity of best approximation in L^p-spaces. The remainder of this section deals with the *linearity* of best approximation in general Banach spaces. We have already observed that BAO supported by a Chebyshev subspace M is always closed, homogeneous, and idempotent. Hence, if P_M is *additive*, then it is actually a bounded linear projection of X onto M, and in fact $\|P_M\| \leq 2$. The next theorem shows that several obviously necessary conditions for the linearity of P_M are also sufficient.

THEOREM 3. *The following properties of a Chebyshev subspace* M *of a Banach space* X *are equivalent:*

(a) P_M *is linear;*

(b) ker P_M *is a subspace;*

(c) ker P_M *contains a subspace* N *for which* $X = M + N$;

(d) P_M *is quasi-linear; that is, there exists a constant* K *such that*

$$\|P_M(x+y)\| \leq K(\|P_M(x)\| + \|P_M(y)\|);$$

(e) P_M *is continuously Gateaux differentiable.*

Proof. Clearly, (a) implies all the other properties, and (d) ⇒ (b). Assume (c); then each $x \in X$ can be written uniquely as $x = m + n$ $(m \in M,\ n \in N)$. Hence $P_M(x) = m$, and we note that m depends linearly on x; that is, (c) ⇒(a). Note that N must actually be all of ker P_M. Next we show that (b) ⇒ (a). Choose any $x, y \in X$; then $x - P_M(x)$, $y - P_M(y)$, and $x + y - P_M(x+y)$ all belong to ker P_M. Therefore $(x - P_M(x) + y - P_M(y)) \in \ker P_M$, and so

$$P_M(x+y) - P_M(x) - P_M(y) \in M \cap \ker P_M = \{\theta\}.$$

Finally, we show that (e) ⇒ (a). If $\delta > 0$, then

$$P'_M(\delta x, y) = \delta P'_M(x, \delta^{-1} y) = P'_M(x, y).$$

Hence, if P_M is Gateaux-C^1, then

$$P'_M(x, y) = \lim_{\delta \to 0} P'_M(\delta x, y) = P'_M(0, y) = P_M(y).$$

Therefore,

$$P_M(x+y) = P_M(x) + \int_0^1 \frac{d}{dt} P_M(x+ty)\,dt$$

$$= P_M(x) + \int_0^1 P'_M(x+ty, y)\,dt = P_M(x) + P_M(y). \quad \blacksquare$$

As immediate consequences of this theorem we see that P_M is always linear if codim $M = 1$; if $\dim X = 2$ and X is strictly convex, then X has property (UL); and, in the notation of Proposition 4, if every P_{M_α} is linear, then so is P_M. Also, from Theorem 3 and a theorem of Rudin and Smith [19], we deduce that if X is strictly convex and P_M is Gateaux-C^1 whenever M is a one-dimensional subspace of X, then X is a Hilbert space.

Finally, we observe that P_M is linear if and only if the projection $I - P_M$ is linear and *has norm one*. Further, assuming that P_M is linear, we find that $\|P_M\| = 1$ if and only if $(I - P_M)(x)$ is a best approximation to x out of $\ker P_M$. Thus, if P_M is linear and $\ker P_M$ is a Chebyshev subspace, then $\|P_M\| = 1$ if and only if

$$P_{\ker P_M} = I - P_M, \qquad \text{that is,} \qquad \ker P_{\ker P_M} = M;$$

in other words, the subspaces M and $\ker P_M$ are mutually orthogonal in the sense of James [8].

2. APPLICATIONS AND EXAMPLES

Let (Ω, S, μ) be a *positive measure space;* that is, let Ω be a set, S a sigma algebra of subsets of Ω, and μ a nonnegative, countably additive set function (a measure) on S. If $1 < p < \infty$, then $L^p = L^p(\Omega, S, \mu)$ is the Banach space of all (equivalence classes of) μ-measurable, real-valued functions f defined on Ω for which $\|f\|^p = \int_\Omega |f(\cdot)|^p d\mu < \infty$. It is well known that for such values of p, L^p is uniformly convex (Clarkson [3]). Hence, as we noted in the Introduction, every convex subset of L^p is an approximatively compact Chebyshev set and so supports a continuous BAO. In the special case where Ω is a finite set of n points, we shall denote the associated L^p-space by $\ell^p(n, w)$. Thus $x \in \ell^p(n, w)$ means that x is an n-tuple $(x_1, \cdots, x_n)$ of real numbers, and

$$\|x\|^p = \int_\Omega |x(\cdot)|^p d\mu = \sum_{i=1}^n w_i |x_i|^p,$$

where the "weight" $w_i > 0$ is the μ-measure of the ith point of Ω.

To apply the results of the preceding section to approximation problems in L^p-spaces, we shall need the following technical lemma.

LEMMA 1. *Let* $X = L^p(\Omega, S, \mu)$ *for some* p $(2 < p < \infty)$.

(a) *If* $x, y, z \in X$ *and* $\|x\| = 1$, *then*

$$\langle y, z\rangle_x \equiv \left.\frac{\partial^2 \|x + sy + tz\|}{\partial s\, \partial t}\right|_{s=t=0}$$

$$= (p - 1)\left\{\int_\Omega yz\,|x|^{p-2} d\mu - p\int_\Omega xy\,|x|^{p-2} d\mu \int_\Omega xz\,|x|^{p-2} d\mu\right\}.$$

If, in addition, $P_Y(x) = \theta$, *where* $Y = \operatorname{span}\{y\}$, *then*

$$\langle y, z\rangle_x = (p - 1)\int_\Omega yz\,|x|^{p-2} d\mu.$$

(b) *For fixed* $y, z \in X$, *the real-valued function* $x \mapsto \int_\Omega yz\,|x|^{p-2} d\mu$ *is continuous on* X.

(c) *For fixed* $y \in X$, *the map* $x \mapsto y\,|x|^{p-2}$ *from* X to $X' = L^q$ $(p^{-1} + q^{-1} = 1)$ *is continuous.*

Proof. The first formula of (a) is obtained by differentiation under the integral sign. This is justified by the theorems in McShane [14, Section 39], together with the Hölder inequality. The second formula follows if we note that if $P_Y(x) = \theta$, then

$$0 = \frac{d}{dt} \int_\Omega |x + ty|^p \, d\mu \bigg|_{t=0} = p \int_\Omega y \,\mathrm{sgn}(x)\, |x|^{p-1} \, d\mu = p \int_\Omega xy\, |x|^{p-2} \, d\mu .$$

We next observe that (b) follows quite directly from (c), so that it will suffice to prove the latter. Put $r = p/(p - 2)$. We can factor the map in question as a product $\phi \circ \psi$, where

$$\psi\colon L^p \to L^r \qquad \text{and} \qquad \phi\colon L^r \to L^q$$

are defined by $\psi(x) = |x|^{p-2}$ and $\phi(x) = yx$. Now ϕ is clearly continuous, since it is linear and Hölder's inequality shows that $\|\phi(x)\|_q \leq \|y\|_p \|x\|_r$. To show that ψ is continuous, it is convenient to consider separately the cases $2 < p \leq 3$ and $3 < p < \infty$. In the first case, the number $s = p - 2$ satisfies the condition $0 < s \leq 1$, and since $\big||A|^s - |B|^s\big| \leq |A - B|^s$ (because the function $A \to A^s$ is subadditive on the positive reals) we obtain the inequality

$$\|\psi(x) - \psi(y)\|_r \leq \|x - y\|_p^{p-2} .$$

Now suppose that $p > 3$, and put $z(\cdot) = \max(|x(\cdot)|, |y(\cdot)|)$. We can apply the mean-value theorem to show that

$$|\psi(x) - \psi(y)| \leq (p - 2) z^{p-3} \big||x| - |y|\big| \leq (p - 2) z^{p-3} |x - y| .$$

If $s = (p - 2)/(p - 3)$, then $\|z\|_p \leq \|\, |x| + |y| \,\|_p \leq \|x\|_p + \|y\|_p$, and Hölder's inequality shows that

$$\|\psi(x) - \psi(y)\|_r^r \leq (p - 2)^r \,\|\, |x - y|^r \,\|_{p-2} \,\|z^{(p-3)r}\|_s$$
$$= (p - 2)^r \|x - y\|_p^r \|z\|_p^{p/s} \leq (p - 2)^r \|x - y\|_p^r (\|x\|_p + \|y\|_p)^{p/s} .$$

This completes the proof of the lemma.

The lemma has immediate application to the differentiability of best approximation in L^p-spaces. For suppose M is a finite-dimensional subspace of $L^p(\Omega, S, \mu)$ $(2 < p < \infty)$, say $M = \mathrm{span}\{m_1, \cdots, m_n\}$. If $x = \psi_M(x)$, then for the matrix P_x of Theorem 2 we have the formula

$$(P_x)_{ij} = \langle m_i, m_j \rangle_x = (p - 1) \int_\Omega m_i m_j |x|^{p-2} \, d\mu .$$

Similarly, the ith component of the vector $q_x(y)$ is $(p - 1) \int_\Omega m_i y\, |x|^{p-2} \, d\mu$. Now q_x is a continuous linear map from X to R^n, and by the lemma, both q_x and P_x depend continuously on x. Therefore, by Theorem 2 and the subsequent remarks, if M has the property that $\int_\Omega m^2 |x|^{p-2} \, d\mu > 0$ whenever $m \in M \setminus \{\theta\}$ and $x = \psi_M(x)$, then P_M is Fréchet-C^1 on the open set $X \setminus M$.

To provide examples of such subspaces M, we let Ω be a compact Hausdorff space, S the σ-algebra of Borel sets in Ω, μ a regular nonatomic Borel measure on S, and $\{m_1, \cdots, m_n\}$ a *Chebyshev system* on Ω (that is, the m_i are continuous on Ω and no nontrivial linear combination of $\{m_i\}$ has more than $n - 1$ zeros in Ω). More generally, of course, it would simply suffice to take any finite-dimensional M in any L^p-space and to ask that no nonzero member of M vanish on a set of positive measure. But the former example does have a close connection with the present work. For it is well known [18, Theorem 3.6] that the span of a (finite) Chebyshev system is a Chebyshev subspace of $C(\Omega)$ in the uniform norm, and further, that the best uniform approximation to an $f \in C(\Omega)$ is obtainable as the uniform limit of the best L^p-approximations to f as $p \to \infty$ (see Buck [1, Theorem 5] or Kripke [11]).

We turn now to our main result on the smoothness of best approximation in L^p-spaces $(p > 2)$. It will be seen from subsequent examples that within the hierarchy of smoothness properties considered in this paper, these results are the best possible.

THEOREM 4. *Let* M *be a finite-dimensional subspace of* $L^p(\Omega, S, \mu)$ *for some* $p > 2$. *If* $x \in L^p \setminus M$, *there exist a neighborhood* N_x *of* x *and a constant* $\lambda_x \geq 1$ *such that*

$$\|P_M(x') - P_M(y)\| \leq \lambda_x \|x' - y\| \quad \textit{for all } x' \in N_x \textit{ and } y \in L^p.$$

In case $\dim L^p < \infty$, *there exists a single Lipschitz constant* λ, *independent of* x, *such that*

$$\|P_M(x) - P_M(y)\| \leq \lambda \|x - y\| \quad \textit{for all } x, y \in L^p.$$

Before proving this theorem, we isolate as a second technical lemma a key portion of the argument, which is valid in general Banach spaces.

LEMMA 2. *Let* M, M_1, M_2 *be Chebyshev subspaces of* X *such that* $M = M_1 \oplus M_2$. *Suppose that* $x \in \ker P_M$, *and that* K_1, K_2, K_3, r *are positive constants such that*

(i) $$\|P_{M_1}(x) - P_{M_1}(z)\| \leq K_1 \|x - z\| \quad \textit{for all } z \in X;$$

(ii) $$\|x' - x\| \leq r \Rightarrow \|P_{M_2}(x') - P_{M_2}(z)\| \leq K_2 \|x' - z\| \textit{ for all } z;$$

(iii) $$\|x' - x\| \leq \frac{3}{2} r,\ m \in M_2, \quad \textit{and}$$

$$\|m\| \leq 3r \Rightarrow \|P_{M_1}(x' + m) - P_{M_1}(x')\| \leq K_3 \|m\|;$$

(iv) $$K_2 K_3 < 1.$$

Then, if $\|y\| \leq (1 - K_2K_3)(K_2 + K_1K_2)^{-1} r$, *we have the inequality*

$$\|P_M(x + y)\| \leq \{K_2 + (1 + K_2)(K_1 + K_2K_3)(1 - K_2K_3)^{-1}\} \|y\|.$$

Proof. Choose a y such that $\|y\| \leq (1 - K_2K_3)(K_2 + K_1K_2)^{-1} r$, and put $x' = x + y$. Conditions (i) and (ii) imply $K_1, K_2 \geq 1$, since we may take $z \in M_i$ and get the relations

$$\|P_{M_i}(x) - P_{M_i}(z)\| = \|P_{M_i}(x - z)\| \leq K_i \|x - z\| \quad (i = 1, 2).$$

Therefore, $\|y\| = \|x' - x\| \leq r$. Next we observe that the map $T_{x'}$ (see Proposition 5) maps the ball

$$B = M_1 \cap \{z: \|z\| \leq (K_1 + K_2 K_3)(K_2 + K_1 K_2)^{-1} r\}$$

into itself. Indeed, suppose that $m \in B$. Then

$$\begin{aligned}
\|T_{x'}(m)\| &= \|T_{x'}(m) - P_{M_1}(x)\| \\
&\leq \|P_{M_1}(x' - P_{M_2}(x' - m)) - P_{M_1}(x')\| + \|P_{M_1}(x') - P_{M_1}(x)\| \\
&\leq K_3 \|P_{M_2}(x' - m)\| + K_1 \|x' - x\| \leq K_2K_3 \|y - m\| + K_1 \|y\| \\
&\leq K_2K_3 \|m\| + (K_1 + K_2K_3) \|y\| \\
&\leq K_2K_3(K_1 + K_2 K_3)(K_2 + K_1K_2)^{-1} r + (K_1 + K_2K_3)(1 - K_2K_3)(K_2 + K_1K_2)^{-1} r \\
&= (K_1 + K_2K_3)(K_2 + K_1K_2)^{-1} r .
\end{aligned}$$

Now we claim that $T_{x'}$ is actually a strict contraction of B. To see this, choose m and $\overline{m}$ in B. Then, since

$$\begin{aligned}
\|x - (x' - P_{M_2}(x' - \overline{m}))\| &= \|-y + P_{M_2}(x' - \overline{m})\| \leq \|y\| + K_2 \|y - \overline{m}\| \\
&\leq (1 + K_2) \|y\| + K_2 \|\overline{m}\| \leq [(1 + K_2)(1 - K_2K_3) + K_2(K_2 + K_2K_3)](K_2 + K_1K_2)^{-1} r \\
&= [1 + (1 - K_2K_3)(K_2 + K_1K_2)^{-1}] r < 3r/2,
\end{aligned}$$

and since

$$\|P_{M_2}(x' - m) - P_{M_2}(x' - \overline{m})\| \leq K_2 \|m - \overline{m}\| \leq 2(K_1 + K_2K_3)(1 + K_1)^{-1} r < 3r,$$

we have the inequalities

$$\|T_{x'}(m) - T_{x'}(\overline{m})\| \leq K_3 \|P_{M_2}(x' - m) - P_{M_2}(x' - \overline{m})\| \leq K_2K_3 \|m - \overline{m}\| .$$

Thus, by the contraction-mapping principle, the unique fixed point m^* of $T_{x'}$ lies in B, and

$$\|m^*\| = \|T_{x'}(m^*)\| \leq K_2K_3 \|m^*\| + (K_1 + K_2K_3) \|y\| ,$$

that is, $\|m^*\| \leq (K_1 + K_2K_3)(1 - K_2K_3)^{-1} \|y\|$. Also,

$$\|P_{M_2}(x' - m^*)\| \leq K_2 \|x' - x - m^*\| \leq K_2(\|y\| + \|m^*\|).$$

Thus, by the preceding argument and Proposition 5, we finally obtain the relations

$$\begin{aligned}
\|P_M(x + y)\| &= \|m^* + P_{M_2}(x' - m^*)\| \leq (1 + K_2) \|m^*\| + K_2 \|y\| \\
&\leq \{(1 + K_2)(K_1 + K_2K_3)(1 - K_2K_3)^{-1} + K_2\} \|y\| . \quad \blacksquare
\end{aligned}$$

Proof of Theorem 4. The proof is by induction on dim M; the theorem is trivial if dim M = 0. Suppose that $n \geq 1$ and that the theorem is true for subspaces of dimension less than n. It will be sufficient to prove the theorem under the assumption that $x = \psi_M(x)$; for if there exist $r > 0$ and $\lambda_x \geq 1$ such that

$$\|P_M(x') - P_M(y)\| \leq \lambda_x \|x' - y\| \qquad \text{whenever } \|x' - \psi_M(x)\| < r,$$

then the inequality holds whenever $\|x' - x\| < r \cdot \mathrm{dist}(x, M)$. If $y \in L^p$, we define $Z(y) = \{\omega \in \Omega: y(\omega) = 0\}$. This of course really defines $Z(y)$ as an equivalence class of measurable sets, and we shall therefore identify two measurable sets whose symmetric difference has μ-measure 0. Let $\{m_1, \cdots, m_n\}$ be a basis for M. Without loss of generality, we may assume that $\mu\left(\bigcap_{i=1}^n Z(m_i)\right) = 0$, for otherwise we simply replace Ω by $\Omega \setminus \bigcap_{i=1}^n Z(m_i)$ and restrict all functions to this latter set. Now let M_2 be the subspace of M consisting of the elements that vanish μ-a.e. on $\Omega \setminus Z(x)$, and let M_1 be any subspace of M that is complementary to M_2. The set M_2 cannot be all of M, for if it were, $m_1, \cdots, m_n$ would all vanish μ-a.e. on $\Omega \setminus Z(x)$. By our assumption about $\bigcap_{i=1}^n Z(m_i)$, we could then conclude that $\mu(\Omega \setminus Z(x)) = 0$, contrary to the hypothesis that $\|x\| = 1$. It follows that $\dim M_2 < n$, and hence, by the induction hypothesis, that there exist a neighborhood U of x and a constant K_2 such that

$$x' \in U,\ y \in L^p \Rightarrow \|P_{M_2}(x') - P_{M_2}(y)\| \leq K_2 \|x' - y\|.$$

We next consider approximation out of the subspace M_1. Introducing the form $\langle \cdot, \cdot \rangle_x$, we see by Lemma 1 that

$$\langle y, z \rangle_x = (p - 1) \int_\Omega yz\, |x|^{p-2} d\mu$$

if either $y \in M$ or $z \in M$. By choice of M_1, this form is positive definite on M_1, and so P_{M_1} is Fréchet-C^1 on a neighborhood of x (Theorem 2, Lemma 1). By the corollary to Theorem 1, there then exist a neighborhood V of x and a constant K_1 such that

$$x' \in V \text{ and } y \in L^p \Rightarrow \|P_{M_1}(x') - P_{M_1}(y)\| \leq K_1 \|x' - y\|;$$

we may assume that $V \subset U$. For any such x', let $R_{x'}$ denote the restriction of $P'_{M_1}(x', \cdot)$ to M_2. Now the operator R_x is the zero operator on M_2, because $\langle m, y \rangle_x = 0$ if $m \in M_1$ and $y \in M_2$ (recall the formula of Theorem 2). Since P_{M_1} is Fréchet-C^1 near x, we can find an $\varepsilon > 0$ so small that

$$\|x' - x\| \leq 3\varepsilon \Rightarrow x' \in V \text{ and } \|R_{x'}\| \leq K_3 = K_2/2.$$

If $\|x' - x\| \leq \varepsilon$ and $m \in M_2 \cap \{z: \|z\| \leq \varepsilon\}$, then by the mean-value inequality for Fréchet derivatives $\|P_{M_1}(x' + m) - P_{M_1}(x')\| \leq K_3 \|m\|$.

Let r be a positive number less than $\varepsilon/6$. We can easily verify that if $w \in \ker P_M$ and $\|w - x\| < r$, then Lemma 2 applies uniformly to all such w. The conclusion is that

$$\|P_M(w+y)\| \le \{(2K_1+1)(K_2+1)+K_2\}\,\|y\|$$

provided that $w \in \ker P_M$, $\|w - x\| < r$, and $\|y\| \le r/2K_2(1+K_1)$. We now set

$$\lambda_x = \max\{(2K_1+1)(K_2+1)+K_2,\ 2(1+4(1+r^{-1})K_2(1+K_1))\},$$

and we observe that by Proposition 2, $\|P_M(w+y)\| \le \lambda_x\|y\|$ for all $y \in L^p$ and all $w \in \ker P_M$ with $\|w - x\| < r$. Finally, we may take

$$N_x = \{z: \|z - x\| < r/2\} \cap \{z: \|P_M(z)\| < r/2\}.$$

This is a neighborhood of x, since $P_M(x) = \theta$ and P_M is continuous, and we see that

$$\|P_M(x'+y) - P_M(x')\| \le \lambda_x\|y\| \qquad \text{whenever } x' \in N_x \text{ and } y \in L^p.$$

This completes the proof for L^p-spaces of infinite dimension. In the finite-dimensional case, we can go further, for here we have observed that every $x \in X \setminus M$ is contained in a neighborhood N_x on which P_M satisfies a uniform Lipschitz condition with constant λ_x. Since the set $S \cap \ker P_M$ is compact, it can be covered by finitely many sets of the form $S \cap \ker P_M \cap N_x$. If we let λ be the maximum of the corresponding numbers λ_x, we can apply the corollary to Proposition 2 to conclude that P_M is Lipschitzian on L^p. ■

We next show that in spite of the preceding theorem, the L^p-spaces $(1 < p < \infty)$ do *not* have property (UL) except in trivial special cases.

THEOREM 5. *If* $1 < p < \infty$, *the space* $L^p = L^p(\Omega, S, \mu)$ *does not have property* (UL) *unless* $p = 2$ *or* $\dim L^p \le 2$.

Proof. The two exceptional cases have already been discussed. In all other cases, the underlying measure space must contain three disjoint sets of finite positive μ-measure. The characteristic functions of these sets then span a subspace of L^p isometric with some $\ell^p(3, w)$, for a suitable w. Thus it suffices to prove the theorem for the case $L^p(\Omega, S, \mu) = \ell^p(3, w)$. From now on, p and w are fixed, and $\int y\,d\mu$ means $\sum_{i=1}^3 w_i y_i$ for any $y \in \ell^p(3, w)$. Suppose that m is a nonzero vector in $\ell^p(3, w)$ and that $M = \operatorname{span}\{m\}$. If $x = \psi_M(x)$ and $y \in S$, then we know, by Theorem 2 and Lemma 1(a), that

$$P_M'(x, y) = P_x^{-1} q_x(y)\,m = \left\{\int my\,|x|^{p-2}\,d\mu \Big/ \int m^2\,|x|^{p-2}\,d\mu\right\} m,$$

provided that $\int m^2\,|x|^{p-2}\,d\mu \ne 0$. If we define

$$Q(m, x) = \max\left\{\int my\,|x|^{p-2}\,d\mu: y \in S\right\},$$

then $Q(m, x) = \|m\,|x|^{p-2}\|_q$, where $\|\cdot\|_q$ is the norm in $\ell^q(3, w) = (\ell^p(3, w))'$. We shall show that

$$\sup\left\{\frac{Q(m, x)}{\int m^2\,|x|^{p-2}\,d\mu}: \|m\| \ge 1,\ x = \psi_M(x),\ \int m^2\,|x|^{p-2}\,d\mu > 0\right\} = \infty,$$

and this will prove that $\ell^p(3, w)$ does not have property (UL). For there exists $y' = y'(m, x) \in S$ such that $\int my' |x|^{p-2} d\mu = Q(m, x)$. Hence, if λ is a Lipschitz constant for P_M, then certainly

$$\lambda \geq \| P_M(x + ty') - P_M(x) \| / |t| \qquad \text{for each } t \neq 0,$$

and therefore, by Theorem 2, $\lambda \geq Q(m, x) \Big/ \int m^2 |x|^{p-2} d\mu$, since $\|m\| \geq 1$.

Case I ($2 < p < \infty$). Let $\varepsilon > 0$, and choose c so that $c^{p-1} = w_2/w_1$. Let $m = (\varepsilon^p, \varepsilon, 1)$ and $x = r(c, -\varepsilon, 0)$, where $r = r(\varepsilon)$ is chosen so that $x \in S$. Since $\int m \operatorname{sgn}(x) |x|^{p-1} d\mu = 0$, we see that $x = \psi_M(x)$. Now

$$\| m |x|^{p-2} \|_q = (w_1 \varepsilon^{pq} (rc)^{q(p-2)} + w_2 \varepsilon^q (r\varepsilon)^{q(p-2)})^{1/q} = f(\varepsilon) \varepsilon^{p-1},$$

where $f(\varepsilon) \to (w_2 r^{p-q})^{1/q}$ as $\varepsilon \to 0+$. Here we have used the relations $pq = p + q$, $p/q = p - 1$, and $q(p - 2) = p - q$. Similarly, we find that

$$\int m^2 |x|^{p-2} d\mu = w_1 \varepsilon^{2p} (rc)^{p-2} + w_2 \varepsilon^2 (r\varepsilon)^{p-2} = g(\varepsilon) \varepsilon^p,$$

where $g(\varepsilon)$ also approaches $w_2 r^{p-2}$ as $\varepsilon \to 0+$. Therefore

$$Q(m, x) \Big/ \int m^2 |x|^{p-2} d\mu = [f(\varepsilon)/g(\varepsilon)] \varepsilon^{p-1-p} \to +\infty \text{ as } \varepsilon \to 0+.$$

Case II ($1 < p < 2$). We have not discussed the case $p < 2$ before; but this causes no problem in the present situation, because of the simple structure of the spaces $\ell^p(3, w)$. In fact, we can clearly extend Lemma 1(a) to the case $1 < p < 2$ by restricting attention to points x with nonvanishing components. Theorem 2 will then apply, and as in Case I it will suffice to show that $Q(m, x) \Big/ \int m^2 |x|^{p-2} d\mu$ is unbounded. To this end, we let

$$m = (1 - \varepsilon^{1+t(p-1)}, -1, \varepsilon) \qquad \text{and} \qquad x = r(1, b, \varepsilon^t).$$

Here $\varepsilon > 0$, $t > 2/(2 - p)$, and $b = b(\varepsilon)$ is chosen so that $x \in S$. We then verify that

$$\int m^2 |x|^{p-2} d\mu = r^{p-2} \varepsilon^{2+t(p-2)} (w_3 + o(1)) \qquad \text{as } \varepsilon \to 0+,$$

and that

$$\| m |x|^{p-2} \|_q = r^{p-2} \varepsilon^{1+t(p-2)} (w_3 + o(1))^{1/q} \qquad \text{as } \varepsilon \to 0+$$

(here we have used the relations $2 + t(p - 2) < 0$ and $1 + t(p - 1) > 0$). Thus

$$Q(m, x) \Big/ \int m^2 |x|^{p-2} d\mu = \varepsilon^{-1} (w_3 + o(1))^{(1/q)-1} \to +\infty \qquad \text{as } \varepsilon \to 0+. \quad \blacksquare$$

We have already observed in the Introduction that the negative result of Theorem 5 is new only for the case $\dim L^p < \infty$. Indeed, it is meaningless to inquire about property (UL) for infinite-dimensional L^p spaces, since not every BAO on such spaces is even Lipschitzian. For it has recently been shown by Lindenstrauss [12, Section 3, Corollary 2] that if a subspace M of a reflexive Banach space X is the range of a Lipschitzian (or even uniformly continuous) projection on X, then M is actually the range of a bounded *linear* projection on X, that is, M is complemented in X. But by Murray's well-known result [15], not every subspace of infinite-dimensional L^p is complemented. In Example 2 below, we strengthen this result by showing that even a BAO supported by a *line* in ℓ^p may fail to be Lipschitzian.

Our final theorem concerns the Gateaux differentiability of BAO's supported by finite-dimensional subspaces of L^p spaces $(p > 2)$. Theorem 2 is not a complete answer to the existence of such derivatives, because the formula established there is valid only if a certain matrix P_x is nonsingular. As a simple example, consider $\ell^4(4, w)$, where each w_i is 1, and where the subspace M is

$$\text{span}\{(1, 0, 0, 0), (0, 1, 1, -1)\}.$$

We see that $P_M(x) = \theta$ is equivalent to $x_1 = 0$ and $x_2^3 + x_3^3 = x_4^3$; in particular, P_M is not linear. Suppose that $x = \psi_M(x)$. Then the matrix

$$P_x = \begin{bmatrix} 0 & 0 \\ 0 & \sum_{i=2}^{4} x_i^2 \end{bmatrix},$$

and so Theorem 2 is totally inapplicable to the problem of the differentiability of P_M.

THEOREM 6. *Let* M *be a finite-dimensional subspace of* $L^p = L^p(\Omega, S, \mu)$, *where* $2 < p < \infty$. *Then* $P'_M(x, y)$ *exists for each* $x \in L^p \setminus M$ *and each* $y \in L^p$.

Proof. As usual, it suffices to assume that $x = \psi_M(x)$ and that $y \in S$. Let $\{m_1, \cdots, m_n\}$ be a basis for M. As in the proof of Theorem 4, we may assume that $\mu\left(\bigcap_{i=1}^n Z(m_i)\right) = 0$. We then choose the complementary subspaces M_1 and M_2 of M as in that proof. We claim that $P'_M(x, y) = P'_{M_1}(x, y) + P_{M_2}(y - P'_{M_1}(x, y))$. For the rest of this proof we shall write P_i in place of P_{M_i} $(i = 1, 2)$. Now, if t is any real number, we know by Proposition 5 that the map T_{x+ty} has a unique fixed point in M_1, say m(t). That is, $m(t) = P_1[x + ty - P_2(x + ty - m(t))]$. We make the following observations about P_1, P_2, and m(t).

(i) By the uniqueness of m(t) for each t, $m(0) = \theta$;

(ii) $\|m(t)/t\|$ is bounded as $t \to 0$; in fact, we noted in the proof of Proposition 5 that if π is the projection of M onto M_1 along M_2, then $m(t) = \pi \circ P_M(x + ty)$; since π is linear and continuous and P_M is Lipschitzian at x (Theorem 4), it follows that m is Lipschitzian at 0;

(iii) $P_1(x + w) = P'_1(x, w) + R_x(w)$, where the remainder R_x has the property that for each $\varepsilon > 0$ there is a $\delta > 0$ such that $\|R_x(w)\| \leq \varepsilon \|w\|$, provided that $\|w\| < \delta$; here w is any element of L^p, and the result is a consequence of the fact that P_1 is Fréchet-C^1 near x;

(iv) $P'_1(x, y) = \theta$ if $y \in M_2$, by the formula for $P'_1(x, y)$ and the choice of M_2;

(v) $P_2(x + w) = P_2(w)$ for each w, by the choice of M_2.

In view of these facts, we see that

$$m(t) = P_1'[x, ty - P_2(x + ty - m(t))] + R_x[ty - P_2(x + ty - m(t))]$$

$$= t P_1'(x, y) + R_x[ty - P_2(ty - m(t))].$$

Let $\varepsilon > 0$. Since $\|y - P_2(y - m(t)/t)\|$ is bounded for small t, say by A, we have the inequality $\|m(t)/t - P_1'(x, y)\| \leq A\varepsilon$, provided that t is sufficiently small; that is, $m'(0) = P_1'(x, y)$. Finally,

$$\frac{P_M(x + ty) - P_M(x)}{t} = \frac{P_M(x + ty)}{t} = \frac{m(t)}{t} + \frac{P_2(x + ty - m(t))}{t}$$

$$= \frac{m(t)}{t} + P_2\left(y - \frac{m(t)}{t}\right) \rightarrow m'(0) + P_2(y - m'(0)) \quad \text{as } t \rightarrow 0\,;$$

that is, $P_M'(x, y) = P_1'(x, y) + P_2(y - P_1'(x, y))$. ■

Remarks on Theorem 6. (a) $P_M'(x, y)$ is not always linear in y; hence, in particular, we cannot prove in general that P_M has a Fréchet derivative at x or that $P_M'(x, y)$ depends continuously on x. As a simple example, let $m = (m_1, m_2, m_3)$ be a nonzero vector in $\ell^p(3, w)$, where $p > 2$ and $w_i = 1$ for all i. Let $\overline{M} = \text{span}\{m\}$, and let $M = \text{span}\{(m_1, m_2, m_3, 0)\}$ in $\ell^p(4, w)$. Since $\ell^p(3, w)$ is not a Hilbert space, we can choose m so that $P_{\overline{M}}$ is not linear. Let $x = (0, 0, 0, 1) \in \ell^p(4, w)$. Then $x = \psi_M(x)$. If ρ denotes the restriction map from $\ell^p(4, w)$ to $\ell^p(3, w)$ that drops the last component, then $P_M(x + ty) = t P_{\overline{M}}(\rho(y))$, whence $P_M'(x, y) = P_{\overline{M}}(\rho(y))$. However, the map $y \mapsto P_{\overline{M}}(\rho(y))$ is not linear. (b) P_M is not generally Gateaux differentiable if M is an infinite-dimensional subspace of an L^p-space with $2 < p < \infty$. This will emerge as a consequence of Example 3 below.

The remainder of this section consists of the examples to which we have referred throughout the paper. The first example exhibits a method of constructing finite-dimensional non-Hilbert spaces with property (UL), thus proving that property (UL) is not sufficiently strong to characterize the geometry of Hilbert spaces (for 2-dimensional Banach spaces, this has already been observed as a consequence of Theorem 3). We believe, however, that only spaces having equivalent Hilbert norms can have property (UL).

EXAMPLE 1. *Examples of finite-dimensional non-Hilbert* (UL) *spaces.* To begin, let H be any Hilbert space with norm σ. Let ρ be a bounded seminorm on H, so that $\rho(\cdot) \leq k\sigma(\cdot)$ for some $k > 0$. We assume that $\rho(u + tv)$ is twice differentiable at $t = 0$ for all $u, v \in H$; this second derivative is then nonnegative, since $\rho(u + tv)$ is convex in t. We define a new norm η on H by $\eta(\cdot) = \sigma(\cdot) + \rho(\cdot)$. Since σ is strictly convex, η is also strictly convex, and thus all convex subsets of H are η-Chebyshev sets. We consider approximation in the η-norm; hence if M is a subspace of H and $x \in H$, then $x = \psi_M(x)$ means $\eta(x) = 1$ and $\eta(x) \leq \eta(x - m)$, for all $m \in M$. We now claim that

$$(*) \qquad \inf\left\{\left.\frac{d^2\eta(x - tm)}{dt^2}\right|_{t=0} : \sigma(m) = 1,\ x = \psi_{\text{span}\{m\}}(x)\right\} > 0.$$

To prove (*), it clearly suffices to show that $d^2\sigma(x - tm)/dt^2|_{t=0}$ is bounded below for all such m and x. Pick such an m, and let $M = \text{span}\{m\}$. By direct computation, the σ-derivative is $(1/\sigma(x))(\sigma(m)^2 - (x, m)^2/\sigma(x)^2)$, where $(\cdot, \cdot)$ is the original inner product on H. But $\sigma(m) = 1$ and $1 = \eta(x) \geq \sigma(x)$, whence this derivative is at least $(1 - (x, m)^2/\sigma(x)^2)$. We now prove that $(x, m)^2 \leq \sigma(x)^2 - (k+1)^{-2}$ (if we grant this, it follows that the derivative in question is at least $(k+1)^{-2}$). To estimate $(x, m)^2$ we first note that the σ-distance from x to M is at least $(k+1)^{-1}$. For suppose that there is $z \in M$ with $\sigma(x - z) < (k+1)^{-1}$. Then $\eta(x - z) < 1$, since $\eta(\cdot) \leq (k+1)\sigma(\cdot)$; but this contradicts $x = \psi_M(x)$. Now let Q denote the *orthogonal* projection of H on M; that is, let Q be the σ-BAO supported by M. Then

$$(x, m)^2 = (Qx, m)^2 \leq \sigma(Qx)^2 = \sigma(x)^2 - \sigma(x - Qx)^2$$
$$= \sigma(x)^2 - \sigma\text{-dist}(x, M)^2 \leq \sigma(x)^2 - (k+1)^{-2}.$$

This completes the proof of (*).

To obtain our examples, we choose H to be R^n, σ to be the usual euclidean norm on R^n, and ρ to be the norm of $\ell^p(n, w)$, where $2 < p < \infty$ and w is an arbitrary n-tuple of positive weights. Because all norms on R^n are equivalent, there exist positive constants c and k for which $c\sigma(\cdot) \leq \rho(\cdot) \leq k\sigma(\cdot)$. We now set $\eta(\cdot) = \rho(\cdot) + \sigma(\cdot)$, and we claim that (R^n, η) has property (UL). Indeed if M is the subspace spanned by a σ-unit vector m, then (*) together with the differentiability of η allows us to deduce from Theorem 2 that $P'_M(x, y)$ exists if $x = \psi_M(x)$ and $\eta(y) = 1$. In fact,

$$P'_M(x, y) = P_x^{-1} q_x(y) m = \langle m, y \rangle_x (\langle m, m \rangle_x)^{-1} m$$
$$= \left(\frac{\partial^2 \eta(x + sm + ty)}{\partial s \partial t} \bigg|_{s=t=0} \right) \left(\frac{d^2 \eta(x - sm)}{ds^2} \bigg|_{s=0} \right)^{-1} m.$$

We now show that $\sup\{ |P_x^{-1} q_x(y)| : \sigma(m) = 1 \text{ and } M = \text{span}\{m\},\ x = \psi_M(x),\ \eta(y) = 1\} < \infty$. If we grant this, then the uniform Lipschitz property of (R^n, η) is a consequence of Theorem 1(b) and Proposition 3. But the content of (*) is exactly the assertion that the set of numbers $\{\langle m, m \rangle_x\}$ is uniformly bounded below over all such x and m. Therefore it suffices to verify that the numbers $q_x(y) = \langle m, y \rangle_x$ are uniformly bounded above. But this is quite straightforward: we simply perform the indicated differentiation and make several applications of the Schwarz-and-Hölder inequalities. In making these estimates, we use the following relations: if $\eta(z) = 1$, then

$$\sigma(z), \rho(z) \leq 1, \qquad 1/\sigma(z) \leq k + 1, \qquad 1/\rho(z) \leq 1 + c^{-1};$$

also, $\rho(m) \leq k$, since $\sigma(m) = 1$.

Remarks on Example 1. (a) The $\ell^p(n, w)$-norm as the perturbing seminorm ρ was a convenient choice, because of its differentiability and the ease with which its derivatives could be bounded. Although other choices for ρ will work, some care must be exercised in making the choice. For example, it can be shown that if we perturb the euclidean norm σ on R^n by adding the seminorm $|(\cdot, x_0)|$, where x_0 is a fixed nonzero vector in R^n, then the resultant space does not have property (UL). (b) Our results in this example, together with the negative results for L^p-spaces $(p \neq 2)$ as presented in Theorem 5, suggest that a Banach space X has property (UL) only if its unit sphere S is sufficiently "curved". Indeed, the essential part of the argument in Example 1 is the bounding away from zero of a measure of the curvature, namely (*). On the other hand, for $p \neq 2$ the L^p-spaces do not have this property, and, in fact, their spheres are "flattened" at certain points.

EXAMPLE 2. *A non-Lipschitzian* BAO *supported by a one-dimensional subspace of* ℓ^p $(p > 2)$. By ℓ^p we mean $L^p(\Omega, S, \mu)$, where $\Omega = \{1, 2, \cdots\}$, S is the family of all subsets of Ω, and $\mu(\{n\}) = 1$ for $n = 1, 2, \cdots$. Define $m \in \ell^p$ by $m = \left(1, 1, \frac{1}{2}, \frac{1}{2}, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^2, \cdots\right)$, and let $M = \text{span}\ \{m\}$. In the notation of the proof of Theorem 5, we shall show that

$$\sup \left\{ Q(m, x) \Big/ \int m^2\, |x|^{p-2}\, d\mu : x = \psi_M(x) \right\} = +\infty,$$

which will prove that P_M is not Lipschitzian. Let

$$x_i = 2^{-1/p}(0, 0, \cdots, 0, -1, 1, 0, 0, \cdots),$$

where the nonzero elements occur in the (2i - 1)st and (2i)th places; then $x_i = \psi_M(x_i)$. But

$$Q(m, x_i) \Big/ \int m^2\, |x_i|^{p-2}\, d\mu = \left(\sum_{j=1}^{\infty} |m_j|^q\, |x_{ij}|^{q(p-2)} \right)^{1/q} \left(\sum_{j=1}^{\infty} m_j^2\, |x_{ij}|^{p-2} \right)^{-1}$$

$$= (2 \cdot 2^q \cdot 2^{-iq} \cdot 2^{q(2-p)/p})^{1/q} (2 \cdot 2^2 \cdot 2^{-2i} \cdot 2^{(2-p)/p})^{-1} = 2^{(1/q)+i-2},$$

which is unbounded as $i \to \infty$.

EXAMPLE 3. *An infinite-dimensional subspace of* ℓ^p *that supports a nondifferentiable* BAO *without the pointwise Lipschitz property.* The space ℓ^p is the same as in the previous example. From Theorem 5 we know that if $1 < p < \infty$ $(p \neq 2)$, then for each $k > 0$ there exists a line L in $\ell^p(3)$ spanned by a unit vector m, an $x = \psi_L(x)$, and a unit vector y such that $P'_L(x, y)$ exists and equals cm with $c > k$. Here $\ell^p(3)$ means $\ell^p(3, w)$ with $w_i = 1$ for all i. Now, for $n = 1, 2, \cdots$, let X_n be the subspace of ℓ^p consisting of all $z = (z_1, z_2, \cdots)$ such that $z_i = 0$ if $i \neq 3n - 2, 3n - 1, 3n$, and let $\pi_n: \ell^p \to \ell^p(3)$ be defined, for $n = 1, 2, \cdots$, as the operator whose value at $z \in \ell^p$ is $(z_{3n-2}, z_{3n-1}, z_{3n})$. By the results of Theorem 5, we can choose three sequences $\{m_n\}_1^\infty$, $\{x_n\}_1^\infty$, $\{y_n\}_1^\infty$ of unit vectors, all in $\ell^p(3)$, and a sequence $\{t_n\}_1^\infty$ of reals with the following properties:

(i) $\lim_{n \to \infty} t_n = 0$;

(ii) $P_{L_n}(x_n) = \theta$, where L_n is the line in $\ell^p(3)$ spanned by m_n;

(iii) $\|P_{L_n}(x_n + t_n y_n)\| \geq 4^n |t_n|$.

Now let M be the subspace of ℓ^p consisting of all vectors z such that $\pi_n(z) \in L_n$ for $n = 1, 2, \cdots$, and let x and y be the unique elements of ℓ^p such that $\pi_n(x) = 2^{-n} x_n$ and $\pi_n(y) = 2^{-n} y_n$. Then $\|x\| \leq 1$, $\|y\| \leq 1$, and moreover, it is easy to see that

$$P_M(z) = m \Leftrightarrow \pi_n(m) = P_{L_n} \circ \pi_n(z) \text{ for every } n.$$

Thus we find in particular that $P_M(x) = \theta$, while

$$\|P_M(x + t_n y)\| \geq \|\pi_n \circ P_M(x + t_n y)\| = \|P_{L_n} \circ \pi_n(x + t_n y)\|$$

$$= \|P_{L_n}(2^{-n}(x_n + t_n y_n))\| = 2^{-n} \|P_{L_n}(x_n + t_n y_n)\| \geq 2^{-n} 4^n |t_n|$$

$$= 2^n |t_n| \geq 2^n \|(x + t_n y) - x\|.$$

Thus the restriction of P_M to the line $\{x + ty: t \text{ real}\}$ fails to be Lipschitzian, and, of course, $P'_M(x, y)$ does not exist.

EXAMPLE 4. *A Chebyshev subspace of codimension* 2 *that supports a discontinuous* BAO. This example also exploits the results of Theorem 5. We begin by showing that in each $\ell^p(3)$ $(1 < p < \infty)$ we can find vectors m_p, x_p, y_p such that

(i) $\sup\{\|m_p\|_p, \|x_p\|_p, \|y_p\|_p: 1 \leq p < \infty\} < \infty$;

(ii) if L_p is the line in $\ell^p(3)$ spanned by m_p, then $P_{L_p}(x_p) = \theta$ for all p;

(iii) $\|y_p\|_p$ decreases to 0 as p increases;

(iv) $\|P_{L_p}(x_p + y_p)\|_p \geq 1$

(because we now work with different values of p, we distinguish the norm in $\ell^p(3)$ with the subscript p). To prove this, we first take $x_p = (0, 1, 1)$ for all p. We consider the function $E(\varepsilon, p) = \varepsilon[(1 + 2\varepsilon)^{p-1} - (1 + \varepsilon)^{p-1}]$. Clearly, E is strictly increasing in each variable,

$$\lim_{\varepsilon \to \infty} E(\varepsilon, p) = \lim_{p \to \infty} E(\varepsilon, p) = +\infty \qquad \text{for each } \varepsilon > 0,\ p \geq 1,$$

while $E(0, p) = 0$. Thus for each p there exists a unique $\varepsilon(p) > 0$ such that $E(\varepsilon(p), p) = 1$, and $\varepsilon(p)$ decreases to 0 as $p \to \infty$. Now we set $y_p = (0, 3\varepsilon(p), 0)$ and $m_p = (1, \varepsilon(p), -\varepsilon(p))$, and we claim that conditions (i) to (iv) are satisfied. Clearly (i) and (iii) hold, and (ii) follows readily. Finally, (iv) is certainly satisfied if $P_{L_p}(x_p + y_p) = m_p$, which is equivalent to $P_{L_p}(x_p + y_p - m_p) = \theta$. But

$$\int m_p \operatorname{sgn}(x_p + y_p - m_p)\, |x_p + y_p - m_p|^{p-1}\, d\mu = -1 + E(\varepsilon(p), p) = 0;$$

hence (iv) is also satisfied.

Now let X denote the linear space of all bounded functions defined on the set of positive integers; any such function is identified with a bounded sequence of real numbers. We define an operator $\pi_n: X \to R^3$ by

$$\pi_n(z) = (z_{3n-2}, z_{3n-1}, z_{3n}) \qquad \text{if } z = (z_1, z_2, \cdots) \in X \text{ and } n = 1, 2, \cdots.$$

We now norm X by

$$\|z\| = \sup\{\|\pi_n(z)\|_n: n = 1, 2, \cdots\} + \left[\sum_{n=1}^{\infty} (n^{-1}\|\pi_n(z)\|_n)^2\right]^{1/2}.$$

The space $(X, \|\ \|)$ is a strictly convex Banach space. Let M be the subspace of $(X, \|\ \|)$ consisting of vectors z such that $\pi_n(z) \in L_n$ for $n = 1, 2, \cdots$. Then M is a Chebyshev subspace, and $\pi_n \circ P_M = P_{L_n} \circ \pi_n$. Now let x and y be the unique vectors in X such that $\pi_n(x) = x_n$ and $\pi_n(y) = \|y_n\|_n^{-1} y_n$ for every n. Then $P_M(x) = \theta$, while

$$\begin{aligned}\|P_M(x + \|y_n\|_n y)\| &\geq \|\pi_n \circ P_M(x + \|y_n\|_n y)\|_n \\ &= \|P_{L_n} \circ \pi_n(x + \|y_n\|_n y)\|_n = \|P_{L_n}(x_n + y_n)\|_n \geq 1,\end{aligned}$$

for $n > 1$. But $x + \|y_n\|_n y \to x$ along the line $\{x + ty: t \text{ real}\}$ as $n \to \infty$. Thus the restriction of P_M to this line is discontinuous.

Finally, let Y be the subspace of X spanned by x, y, and M. Then M has codimension 2 in Y and P_M is discontinuous on a line in Y. We claim that $(Y, \| \|)$ is topologically isomorphic with ℓ^∞, that is, with the space X normed by $\|z\|_\infty = \sup\{|z_n|: n = 1, 2, \cdots\}$. To see this, observe that each element z of Y can be written uniquely as $a(z)x + b(z)y + c(z)$, where a and b are continuous linear functionals on Y, and where c is a projection in $L(Y, M)$. It is easy to see that $(M, \| \|)$ is topologically isomorphic with ℓ^∞. If the map $\phi: Y \to \ell^\infty$ is defined by $\phi(z) = (a(z), b(z), (c(z))_1, (c(z))_2, \cdots)$, then ϕ is clearly one-to-one, onto, and continuous; hence ϕ^{-1} is continuous.

A conjecture. The space Y of the preceding example is strictly convex, but not reflexive. It seems unlikely to us that a convex subset of a space that is both reflexive and strictly convex can support a discontinuous BAO.

EXAMPLE 5. *A line in a 3-dimensional, uniformly convex space that supports a* BAO *without the pointwise Lipschitz property.* The computations involved in this example are completely elementary, but the details are tedious. Hence we shall simply sketch the salient points. Our Banach space is (R^3, η), where the norm η has the form $\eta(\cdot) = \rho(\cdot) + \|\cdot\|_5$ (the norm ρ remains to be defined). We represent a point $p \in R^3$ as $p = (x, y, z)$. Consider the semicircles

$$C_1: x = \sqrt{1 - y^2},\ |y| \leq 1,\ z = 1 \quad \text{and} \quad C_2: x = -\sqrt{1 - y^2},\ |y| \leq 1,\ z = -1.$$

Let K be their (compact) convex hull. Then ρ is the norm whose closed unit ball is K. Next we develop a formula for computing the ρ-norm of certain points: if R is the region

$$\{(x, y, z): x < 0,\ y \leq 0,\ z \geq -1,\ x \geq y + z,\ x^2 + y^2 \geq z^2\}$$

and $p = (x, y, z) \in R^3$, then $\rho(p) = -(y + x^2(y + z)^{-1})$. Since η is strictly convex, the z-axis is an η-Chebyshev set, and we let P denote its BAO. We observe that $P((0, 1, 0)) = \theta$ $(= (0, 0, 0))$, because θ is both an $\ell^5(3)$ - best approximation and a ρ-best approximation to $(0, 1, 0)$. We claim that P satisfies no Lipschitz condition at $(0, 1, 0)$. To see this, we first choose an $\bar{x}$ $(0 < \bar{x} < 1)$, and we show that

$$\rho((\bar{x}, 1, 0) - (0, 0, z)) \geq p(\bar{x}, 1, 0) \quad \text{if } z \geq 0.$$

Suppose $-\frac{1}{2}\sqrt{\bar{x}} \leq z < 0$. Then the formula for the ρ-norm shows that

$$\eta((\bar{x}, 1, -z)) = \eta((-\bar{x}, -1, z)) = (1 - \bar{x}^2)(z - 1)^{-1} + (\bar{x}^5 + 1 + |z|^5)^{1/5}.$$

Since $d\eta((\bar{x}, 1, -z))/dz \geq 3\bar{x}^2/16 > 0$, it follows that $\bar{z} < -\sqrt{\bar{x}/2}$ if

$$P((\bar{x}, 1, 0)) = (0, 0, \bar{z}).$$

Therefore, $\eta(P(\bar{x}, 1, 0) - P(0, 1, 0)) = \eta((0, 0, \bar{z})) > \sqrt{\bar{x}/2}\,\eta(0, 0, 1)$. But $((\bar{x}, 1, 0) - (0, 1, 0)) = \bar{x}\eta(0, 0, 1)$.

REFERENCES

1. R. C. Buck, *Linear spaces and approximation theory.* On Numerical Approximation, Proceedings of a Symposium, Madison, April 21-23, 1958, pp. 11-23. Edited by R. E. Langer. Publication no. 1 of the Mathematics Research Center, U.S. Army, The University of Wisconsin. The University of Wisconsin Press, Madison, 1959.

2. W. Cheney and A. A. Goldstein, *Proximity maps for convex sets.* Proc. Amer. Math. Soc. 10 (1959), 448-450.

3. J. A. Clarkson, *Uniformly convex spaces.* Trans. Amer. Math. Soc. 40 (1936), 396-414.

4. J. Dieudonné, *Foundations of modern analysis.* Academic Press, New York, 1960.

5. N. V. Efimov and S. B. Stečkin, *Approximative compactness and Chebyshev sets.* Dokl. Akad. Nauk SSSR 140 (1961), 522-524.

6. K. Fan and I. Glicksberg, *Some geometric properties of the spheres in a normed linear space.* Duke Math. J. 25 (1958), 553-568.

7. G. Freud, *Eine Ungleichung für Tschebyscheffsche Approximationspolynome.* Acta Sci. Math. Szeged 19 (1958), 162-164.

8. R. C. James, *Inner products in normed linear spaces.* Bull. Amer. Math. Soc. 53 (1947), 559-566.

9. ———, *Linear functionals as differentials of a norm.* Math. Mag. 24 (1951), 237-244.

10. V. L. Klee, *Some characterizations of reflexivity.* Rev. Ci. (Lima) 52 (1950), 15-23.

11. B. R. Kripke, *Best approximation with respect to nearby norms.* Numer. Math. 6 (1964), 103-105.

12. J. Lindenstrauss, *On nonlinear projections in Banach spaces.* Michigan Math. J. 11 (1964), 263-287.

13. S. Mazur, *Über schwache Konvergenz in den Räumen* (L^p). Studia Math. 4 (1933), 128-133.

14. E. J. McShane, *Integration.* Princeton University Press, Princeton, N.J., 1944.

15. F. J. Murray, *On complementary manifolds and projections in spaces* L_p *and* ℓ_p. Trans. Amer. Math. Soc. 41 (1937), 138-152.

16. D. J. Newmann and H. S. Shapiro, *Some theorems on Čebyšev approximation.* Duke Math. J. 30 (1963), 673-681.

17. R. R. Phelps, *Convex sets and nearest points.* Proc. Amer. Math. Soc. 8 (1957), 790-797.

18. ———, *Uniqueness of Hahn-Banach extensions and unique best approximation.* Trans. Amer. Math. Society 95 (1960), 238-255.

19. W. Rudin and K. T. Smith, *Linearity of best approximation: a characterization of ellipsoids.* Nederl. Akad. Wetensch. Proc. Ser. A 64 = Indag. Math. 23 (1961), 97-103.

20. H. H. Schaefer, *Topological vector spaces*. Macmillan, New York, 1966.

21. I. Singer, *Charactérization des éléments de meilleure approximation dans un espace de Banach quelconque*. Acta Sci. Math. Szeged 17 (1956), 181-189.

School of Engineering, AFIT, Wright-Patterson AFB, Ohio 45433
and
University of California, Berkeley 94720