SIAM J. MATH. ANAL.
Vol. 45, No. 2, pp. 639–661

ON STABILITY OF THE METRIC PROJECTION OPERATOR*

ANDRÁS KROÓ† AND ALLAN PINKUS‡

Abstract. Let M be a closed linear subspace of a normed linear space X. For a given $f \in X$ denote by $P_M f$ the set of best approximations to f from M. The operator P_M is termed the metric projection onto M. In this paper we are interested in the stability of the metric projection P_M relative to perturbations of the subspace M. We mainly consider the case where $X = L^p$, $p \in [1, \infty]$. We consider a measure of distance $d(M,N)$ between subspaces M and N and estimate $\|P_M f - P_N f\|$ in terms of $d(M,N)$ and $\|f\|$. Typically such an estimate will be of order $d(M,N)^\beta$ with some β which, in general, depends on the geometry of the space X.

Key words. metric projection, strong uniqueness, best approximation in L^p spaces

AMS subject classification. 41A17

DOI. 10.1137/120873534

1. Introduction.

Let X be a normed linear space and consider a subset M of X. For a given $f \in X$ denote by $P_M f$ the set of best approximations to f from M. Thus

$$P_M f := \left\{ m^* : m^* \in M, \|f - m^*\| = \inf_{m \in M} \|f - m\| \right\}.$$

P_M is said to be the metric projection operator onto M. In general, $P_M f$ is a set-valued mapping, i.e., for each f it might be the empty set, a single value in M, or a set in M. In this paper we always assume that $P_M f$ is nonempty and also that $P_M f$ is single-valued, i.e., we have the uniqueness of the best approximation, unless otherwise stated (cf. Corollary 2.6 and 2.7). A natural question that arises with respect to P_M is its stability. The stability of $P_M f$ with respect to small perturbations of f, that is, continuity properties of the metric projection, has been investigated in the literature; see, e.g., Holmes and Kripke [7], Björnestål [3]. In this note we are interested in a different problem, namely: How stable is the metric projection P_M relative to small perturbations of the closed linear subspace M?

Let M and N be closed linear subspaces of X. To address the above question we recall a measure of the distance between the subspaces M and N. There can be different such measures. A measure that will be convenient for our purposes is

$$d(M,N) := \max \left\{ \sup_{\substack{m \in M \\ \|m\|=1}} \inf_{n \in N} \|m - n\|, \sup_{\substack{n \in N \\ \|n\|=1}} \inf_{m \in M} \|n - m\| \right\}. \tag{1.1}$$

This measure is symmetric in M and N, equals 0 if and only if $M = N$, and is a number between 0 and 1. This measure was introduced in Krein and Krasnosel'skii [9] in the Hilbert space setting and was later extended to Banach spaces in Krein, Krasnosel'skii,

*Received by the editors April 13, 2012; accepted for publication (in revised form) January 10, 2013; published electronically March 19, 2013.
http://www.siam.org/journals/sima/45-2/87353.html
†Alfred Rényi Institute of Mathematics, Hungarian Academy of Sciences, and Department of Analysis, Budapest University of Technology and Economics, Budapest, Hungary (kroo@renyi.hu).
This author was supported by OTKA grant T77812.
‡Department of Mathematics, Technion, Haifa 32000, Israel (pinkus@technion.ac.il).

and Milman [10]. A somewhat similar concept was introduced in Gohberg and Markus [5], namely,

$$\widetilde{d}(M,N) := \max\left\{ \sup_{\substack{m\in M\\ \|m\|=1}} \inf_{\substack{n\in N\\ \|n\|=1}} \|m-n\|,\ \sup_{\substack{n\in N\\ \|n\|=1}} \inf_{\substack{m\in M\\ \|m\|=1}} \|n-m\| \right\},$$

i.e., $\widetilde{d}(M,N)$ is the Hausdorff distance between the unit spheres of the linear subspaces M and N. Note that on the one hand $d(M,N) \le \widetilde{d}(M,N)$, while on the other hand it easily follows that

$$\widetilde{d}(M,N) \le \frac{2d(M,N)}{1+d(M,N)} \le 2d(M,N). \tag{1.2}$$

These concepts and some of their basic properties are considered, for example, in Gohberg and Krein [4] and Kato [8]. We will use the following properties of $d(M,N)$ that may be found in these references.

Proposition 1.1. *Let M and N be closed linear subspaces of a Banach space X. If $\dim M > \dim N$, then $d(M,N) = 1$.*

Let $M^\perp$ denote the annihilator of M in the dual space X^*, i.e.,

$$M^\perp := \{f : f \in X^*, f(m) = 0 \text{ all } m \in M\}.$$

Then from standard duality we have the following.

Proposition 1.2. *Let M and N be closed linear subspaces of a Banach space X. Then*

$$d(M,N) = d(M^\perp, N^\perp).$$

The motivation in this paper is different from those of the above-mentioned references. Our goal is to estimate $\|P_M f - P_N f\|$ in terms of $d(M,N)$ and $\|f\|$. Typically such an estimate will be of order $d(M,N)^\beta$ with some $\beta \le 1$ which in general depends on the geometry of the space X. In this paper we mainly consider $X = L^p$ for $p \in [1,\infty]$.

When $X = H$ is a Hilbert space, this theory is well understood. A proof of the fact that

$$\|P_M - P_N\| = d(M,N)$$

may be found in Akhiezer and Glazman [1]. (This text was originally published in Russian in 1950.) For completeness we present a short proof of this result and some extensions.

Let $X = H$ be a Hilbert space with inner product $(\cdot,\cdot)$. Thus P_M and P_N are in fact linear operators and orthogonal projections.

Proposition 1.3. *In any Hilbert space we have*

$$\|P_M - P_N\| = d(M,N) \tag{1.3}$$

for all closed linear subspaces M and N.

Proof. Recall that by (1.1) we defined $d(M,N)$ as

$$d(M,N) = \max\left\{ \sup_{\substack{m\in M\\ \|m\|=1}} \|m - P_N m\|,\ \sup_{\substack{n\in N\\ \|n\|=1}} \|n - P_M n\| \right\}.$$

Consider these suprema in our Hilbert space setting. We have that

$$\sup_{\substack{m\in M\\ \|m\|=1}} \|m - P_N m\| = \sup_{\substack{m\in M\\ \|m\|=1}} \|(I-P_N)m\| = \sup_{\|f\|\le 1} \|(I-P_N)P_M f\| = \|(I-P_N)P_M\|.$$

From symmetry considerations

$$\sup_{\substack{n\in N\\ \|n\|=1}} \|n - P_M n\| = \|(I-P_M)P_N\|.$$

Thus

$$d(M,N) = \max\{\|(I-P_N)P_M\|, \|(I-P_M)P_N\|\}. \tag{1.4}$$

Note that from the self-adjointness of projections we also have

$$\|(I-P_N)P_M\| = \|P_M(I-P_N)\|, \;\; \|(I-P_M)P_N\| = \|P_N(I-P_M)\|. \tag{1.5}$$

Since $P_M - P_N = P_M(I-P_N) - (I-P_M)P_N$ and $P_M(I-P_N) \perp (I-P_M)P_N$, it follows that

$$\|P_M f - P_N f\|^2 = \|(I-P_M)P_N f\|^2 + \|P_M(I-P_N)f\|^2. \tag{1.6}$$

Now using the fact that $P_N^2 = P_N$, we have from (1.4) that for any $f \in H$

$$\begin{aligned} \|(I-P_M)P_N f\|^2 &= \|(I-P_M)P_N^2 f\|^2 \le \|(I-P_M)P_N\|^2 \|P_N f\|^2 \\ &\le d^2(M,N)\|P_N f\|^2. \end{aligned} \tag{1.7}$$

Similarly, we obtain for the second term in (1.6) using that $(I-P_N)^2 = I - P_N$

$$\|P_M(I-P_N)f\|^2 = \|P_M(I-P_N)^2 f\|^2 \le \|P_M(I-P_N)\|^2 \|(I-P_N)f\|^2.$$

Applying (1.5) and (1.4) yields

$$\|P_M(I-P_N)f\|^2 \le d^2(M,N)\|(I-P_N)f\|^2. \tag{1.8}$$

Finally, substituting (1.7) and (1.8) in (1.6), we obtain

$$\|P_M f - P_N f\|^2 \le d^2(M,N)\|P_N f\|^2 + d^2(M,N)\|(I-P_N)f\|^2 = d^2(M,N)\|f\|^2,$$

which verifies that

$$\|P_M - P_N\| \le d(M,N).$$

It remains to show that this upper bound is sharp. By the symmetry of notation (1.1) we can assume without loss of generality that given any $\epsilon > 0$ for some $m \in M, \|m\| = 1$, we have $\|m - P_N m\| \ge d(M,N) - \epsilon$. Then evidently,

$$\|P_M m - P_N m\| = \|m - P_N m\| \ge d(M,N) - \epsilon. \qquad \square$$

What about the exact value of $d(M,N)$ in this Hilbert space setting? As we know (see Proposition 1.1), if M and N are subspaces of different dimensions or different co-dimensions, then $d(M,N) = 1$. We consider the case where M and N are both finite-dimensional subspaces of the same dimension or both subspaces with the same finite co-dimension.

Proposition 1.4. *Let $\{m_1,\ldots,m_r\}$ and $\{n_1,\ldots,n_r\}$ be orthonormal bases for M and N, respectively, with r finite. Let G denote the $r\times r$ matrix $G=((m_i,n_j))_{i,j=1}^r$. Then*

$$d(M,N)=[1-\lambda_r^2]^{1/2},$$

where λ_r is the smallest s-number of the matrix G, i.e., λ_r^2 is the smallest eigenvalue of GG^*.

Proof. By definition and since $\{n_1,\ldots,n_r\}$ is an orthonormal basis for N,

$$\max_{\|m\|=1}\min_{n\in N}\|m-n\|^2=\max_{\|m\|=1}\|m-\sum_{j=1}^r(m,n_j)n_j\|^2=\max_{\|m\|=1}\left[1-\sum_{j=1}^r|(m,n_j)|^2\right].$$

Now $m=\sum_{i=1}^r a_im_i$, where $\sum_{i=1}^r|a_i|^2=\|m\|^2$. Set $\mathbf{a}=(a_1,\ldots,a_r)$. Then, continuing the above we get

$$\begin{aligned}
&=1-\min_{\|\mathbf{a}\|=1}\sum_{j=1}^r\left|\sum_{i=1}^r a_i(m_i,n_j)\right|^2\\
&=1-\min_{\|\mathbf{a}\|=1}\sum_{j=1}^r\left(\sum_{i,\ell=1}^r a_i(m_i,n_j)\overline{a}_\ell\overline{(m_\ell,n_j)}\right)\\
&=1-\min_{\|\mathbf{a}\|=1}\sum_{i,\ell=1}^r a_i\overline{a}_\ell\sum_{j=1}^r(m_i,n_j)\,\overline{(m_\ell,n_j)}\\
&=1-\min_{\|\mathbf{a}\|=1}\sum_{i,\ell=1}^r a_i\overline{a}_\ell(GG^*)_{i\ell}=1-\min_{\|\mathbf{a}\|=1}\|G^*\mathbf{a}\|^2.
\end{aligned}$$

It is well known that

$$\min_{\|\mathbf{a}\|=1}\|G^*\mathbf{a}\|^2=\lambda_r^2,$$

where λ_r^2 is the smallest eigenvalue of GG^*.

Note that $d(M,N)$ is actually the maximum of two quantities. But the above expression is symmetric in M and N. Thus

$$d(M,N)=[1-\lambda_r^2]^{1/2}.\qquad\square$$

With regards to the above result, see Golub and Van Loan [6, section 2.6], where they consider subspaces M and N of $\mathbb{R}^n$ of equal dimension, define the distance between these subspaces as $\|P_M-P_N\|$, and obtain Proposition 1.4.

From Proposition 1.2 $d(M,N)=d(M^\perp,N^\perp)$. (In this case, the result follows easily from (1.4).) Thus we also have the following.

Corollary 1.5. *Let $\{m_1,\ldots,m_r\}$ and $\{n_1,\ldots,n_r\}$ be orthonormal systems, r finite. Let M and N be subspaces of co-dimension r defined by*

$$M=\{h:(h,m_i)=0,\quad i=1,\ldots,r\}$$

and

$$N=\{h:(h,n_j)=0,\quad j=1,\ldots,r\}.$$

Let G denote the $r \times r$ matrix $G = ((m_i, n_j))_{i,j=1}^r$. Then

$$d(M, N) = [1 - \lambda_r^2]^{1/2},$$

where λ_r is the smallest s-number of the matrix G.

In section 2 of this paper we outline a general approach to the study of stability of the metric projection P_M based on strong uniqueness. This approach will be shown to provide stability estimates whenever strong uniqueness of some order holds with respect to M or N. However it does not always lead to the best possible rates of stability. In section 3 we shall show that for L^p, $p \geq 2$, somewhat more delicate considerations based on smoothness of the norm lead to sharper estimates for the stability of P_M if nonzero functions of the finite-dimensional subspace M do not vanish on sets of positive measure. We also consider the case where both M and N are subspaces of co-dimension 1. Finally, in section 4 we state some open problems.

Stability of the metric projection is, of course, a highly desirable property. It is, however, not always present. Consider the following simple example.

Example. For $t \in (0, 1)$, let $M_t = \text{span}\{(1, t)\}$ and $N_t = \text{span}\{(1, -t)\}$ be one-dimensional subspaces of $\mathbb{R}^2$ endowed with the uniform norm. It is readily verified that $d(M_t, N_t) = 2t/(1 + t)$. Let $f = (0, 1)$. Then

$$P_{M_t} f = \left(\frac{1}{1+t}, \frac{t}{1+t}\right)$$

while

$$P_{N_t} f = \left(-\frac{1}{1+t}, \frac{t}{1+t}\right).$$

Thus

$$\|P_{M_t} f - P_{N_t} f\|_\infty = \frac{2}{1+t},$$

which does not converge to 0 as t tends to 0. That is, there is no stability. In this example $P_{M_t} f$ and $P_{N_t} f$ tend to two different best approximations to f from the subspace $\text{span}\{(1, 0)\}$.

2. Stability of the metric projection and strong uniqueness. In this section we outline a general approach to the study of the stability of the metric projection based on strong uniqueness type results. Strong uniqueness of best approximations has been extensively investigated over the past 40 years; see the recent survey Kroó and Pinkus [12]. We start with the relevant definitions. We assume, as previously, that M is a closed linear subspace of a normed linear space X, and P_M is its corresponding metric projection. Note that P_M is in general not a linear operator.

DEFINITION 2.1. *The metric projection P_M is said to be strongly unique of order $\alpha > 0$ at M if for each $f \in X$ and every $m \in M$ we have*

$$\gamma_M(f) \|P_M f - m\|^\alpha \leq \|f - m\|^\alpha - \|f - P_M f\|^\alpha \tag{2.1}$$

with some constant $\gamma_M(f) > 0$ depending only on f and M.

Remark. It is readily verified that we must have $\alpha \geq 1$ and $\gamma_M(f) \leq 1$ in the above. Indeed, in (2.1) replace m by tm, divide by t^α, and let $t \to \infty$. This immediately gives $\gamma_M(f) \leq 1$. To prove that $\alpha \geq 1$, choose $f \in X \setminus M$ and set

$m = P_M f - tm^*$, $t > 0$, for any fixed $m^* \in M$, $\|m^*\| = 1$. Then from (2.1) and the triangle inequality we get

$$\gamma_M(f)t^\alpha \le \|f - P_M f + tm^*\|^\alpha - \|f - P_M f\|^\alpha \le (\|f - P_M f\| + t)^\alpha - \|f - P_M f\|^\alpha.$$

By the mean value theorem, this equals

$$\alpha t(\|f - P_M f\| + s)^{\alpha-1}$$

for some $s \in (0, t)$. If $\alpha < 1$, then as $\alpha - 1 < 0$ we have $(\|f - P_M f\| + s)^{\alpha-1} \le \|f - P_M f\|^{\alpha-1}$ and thus

$$\gamma_M(f)t^\alpha \le \alpha t\|f - P_M f\|^{\alpha-1}$$

for all $t > 0$, which is impossible.

Remark. It is also straightforward to verify that whenever the metric projection is strongly unique of order α at M with constant $\gamma_M(f)$ in (2.1), then it is also strongly unique of order β at M for all $\beta \ge \alpha$ with constant $\gamma_M(f)^{\beta/\alpha}$.

The simplest case is where X is a Hilbert space because then

$$\|m - P_M f\|^2 = \|f - m\|^2 - \|f - P_M f\|^2$$

for all $m \in M$, implying that $\alpha = 2$ and $\gamma_M(f) = 1$.

The main result of this section is an estimate for the stability of the metric projection under the assumption of strong uniqueness.

THEOREM 2.1. *Let M be a closed linear subspace of the normed linear space X such that the metric projection P_M satisfies the strong uniqueness condition* (2.1). *Then for any $f \in X$ and every other closed linear subspace $N \subset X$ we have*

$$\|P_M f - P_N f\| \le 10\gamma_M(f)^{-1/\alpha}\|f\|d(M, N)^{1/\alpha}. \tag{2.2}$$

The proof of this theorem is more easily understood if we first prove an ancillary result.

PROPOSITION 2.2. *For $f \in X$ and closed linear subspaces M and N we have*

$$\big|\|f - P_M f\| - \|f - P_N f\|\big| \le 2\,\|f\|\,d(M, N); \tag{2.3}$$

$$\|f - P_M P_N f\| \le \|f - P_M f\| + 4\|f\|\,d(M, N). \tag{2.4}$$

Proof. Assume $\|f - P_N f\| \le \|f - P_M f\|$. By our definition of $d(M, N)$ we have

$$\|P_N f - P_M P_N f\| \le \|P_N f\|\,d(M, N) \le 2\|f\|\,d(M, N), \tag{2.5}$$

since $\|P_N f\| \le 2\|f\|$. Hence

$$\begin{aligned}\|f - P_M f\| &\le \|f - P_M P_N f\| \le \|f - P_N f\| + \|P_N f - P_M P_N f\| \\ &\le \|f - P_N f\| + 2\,\|f\|\,d(M, N)\end{aligned}$$

which verifies (2.3).

To prove (2.4) we write, using (2.3) and (2.5),

$$\begin{aligned}\|f - P_M P_N f\| &\le \|f - P_N f\| + \|P_N f - P_M P_N f\| \\ &\le \|f - P_M f\| + 2\,\|f\|\,d(M, N) + 2\|f\|\,d(M, N),\end{aligned}$$

and this leads to the desired estimate. ◻

We now prove Theorem 2.1.

Proof. For given $f \in X$, apply the strong uniqueness inequality (2.1) with $m = P_M P_N f$ and use (2.4) to obtain

$$\begin{aligned} \gamma_M(f)\|P_M f - P_M P_N f\|^\alpha &\le \|f - P_M P_N f\|^\alpha - \|f - P_M f\|^\alpha \\ &\le (\|f - P_M f\| + 4\|f\|\, d(M,N))^\alpha - \|f - P_M\|^\alpha. \end{aligned} \tag{2.6}$$

Since $\alpha \ge 1$ and $d(M,N) \le 1$, we have as a simple consequence of the mean value theorem

$$\begin{aligned} &(\|f - P_M f\| + 4\|f\|\, d(M,N))^\alpha - \|f - P_M\|^\alpha \\ &\quad \le 4\alpha\|f\|\, d(M,N)(\|f - P_M f\| + 4\|f\|\, d(M,N))^{\alpha-1} \le 4\alpha\|f\|\, d(M,N)(5\|f\|)^{\alpha-1}. \end{aligned}$$

Thus from (2.6) we obtain

$$\gamma_M(f)\|P_M f - P_M P_N f\|^\alpha \le \frac{4\alpha}{5}\, d(M,N)(5\|f\|)^\alpha. \tag{2.7}$$

Finally, from (2.7) and (2.5) we get

$$\begin{aligned} \|P_M f - P_N f\| &\le \|P_M f - P_M P_N f\| + \|P_N f - P_M P_N f\| \\ &\le 5\left(\frac{4\alpha}{5\gamma_M(f)}\right)^{1/\alpha} \|f\| d(M,N)^{1/\alpha} + 2\|f\| d(M,N) \\ &\le 10\gamma_M(f)^{-1/\alpha}\|f\| d(M,N)^{1/\alpha}, \end{aligned}$$

where we used the fact that $\gamma_M(f) \le 1$, $d(M,N) \le 1$, and $\alpha \ge 1$. ∎

Let us obtain strong uniqueness estimates of the form (2.1) for the spaces L^p, $1 < p < \infty$. From Smarzewski [15] we have the following result.

Proposition 2.3. *Let $X = L^p$, $2 < p < \infty$, and let M be any closed linear subspace therein. Then for any $m \in M$*

$$c_p\|P_M f - m\|_p^p \le \|f - m\|_p^p - \|f - P_M f\|_p^p,$$

where $c_p = (p-1)(1+s)^{2-p}$ and s is the unique positive zero of the function $t^{p-1} - (p-1)t - (p-2)$.

It can be shown that this c_p satisfies

$$2^{2-p} < c_p < (p-1)2^{2-p}.$$

For the spaces L^p, $1 < p \le 2$, we have the following result of Smarzewski [18].

Proposition 2.4. *Let $X = L^p$, $1 < p \le 2$, and let M be a closed linear subspace therein. Then for any $m \in M$*

$$(p-1)\|P_M f - m\|_p^2 \le \|f - m\|_p^2 - \|f - P_M f\|_p^2.$$

The main consequences of Propositions 2.3 and 2.4 are that they specify orders of L^p strong uniqueness, namely, $\alpha = 2$ for $1 < p < 2$ and $\alpha = p$ when $2 < p < \infty$, and the constants $\gamma_M(f)$ are here chosen independent of both f and M. As an application of Theorem 2.1 we therefore have the following.

Corollary 2.5. *Let $f \in L^p$ and M and N be closed linear subspace of L^p, $1 < p < \infty$. Then for $2 \le p < \infty$,*

$$\|P_M f - P_N f\|_p \le 10{c_p}^{-1/p}\|f\|_p d(M,N)^{1/p},$$

where c_p is as in Proposition 2.3. *For* $1 < p \le 2$ *we have*

$$\|P_M f - P_N f\|_p \le 10(p-1)^{-1/2}\|f\|_p d(M,N)^{1/2}.$$

We do not claim that these upper bounds are sharp in the power of $d(M,N)$. (The constants can certainly be improved upon.) The above result provides us with an error of the power $d(M,N)^{1/2}$ in L^2 while we know from section 1 that the correct order is $d(M,N)$. In the next section we provide necessary conditions under which in L^p, $p \ge 2$, the power is always linear in $d(M,N)$. It need not hold in general.

The case $p = \infty$ is quite different. If M is a finite-dimensional real Haar space (that is, one always has uniqueness of the best approximation from M) in $C(K)$, K a compact Hausdorff space, then it is known that strong uniqueness of order $\alpha = 1$ holds. This result, due to Newman and Shapiro [13], was where the concept of strong uniqueness was first introduced. However it is also known that the constant $\gamma_M(f)$ depends upon f and M and is not uniformly bounded from below away from zero as we vary over, say, all f in the unit ball of $C(K)$. Thus we have here the optimal order 1 of strong uniqueness, but we have to pay a price in complications due to $\gamma_M(f)$. If M is a finite-dimensional complex Haar space in $C(K)$, K a compact Hausdorff space, then it is known that uniqueness of order $\alpha = 2$ holds. See Newman and Shapiro [13], Smarzewski [17], and Kroo and Pinkus [12] for fuller details. Thus we obtain the following.

Corollary 2.6. *If M is a finite-dimensional Haar space in $C(K)$, K a compact Hausdorff space, then for any closed linear subspace N in $C(K)$ and any $f \in C(K)$*

$$\|P_M f - P_N f\|_\infty \le c_{M,f}\, d(M,N)^\beta, \tag{2.8}$$

where $\beta = 1$ in the real case, $\beta = 1/2$ in the complex case, and $c_{M,f}$ is a constant depending on f and M.

It should be noted that in Corollary 2.6 we do not require that the second subspace N be a Haar space. The metric projection P_N can be set-valued, and (2.8) holds with any element of best approximation from $P_N f$. Thus, as a by-product, we obtain an upper bound for the diameter, diam $P_N f$, of the set $P_N f$ via the distance between M and N.

Corollary 2.7. *If M is a finite-dimensional Haar space in $C(K)$, K a compact Hausdorff space, then for any subspace N in $C(K)$ and any $f \in C(K)$ we have diam $P_N f \le 2c_{M,f}\, d(M,N)^\beta$ with $\beta = 1$ in the real case, $\beta = 1/2$ in the complex case, and the constant $c_{M,f}$ depending on f and M as in* (2.8).

Finally, let us mention that in the case of L^1-approximation a space M being a Haar space is only possible when we restrict ourselves to approximating functions from the smaller space $C_1(K)$ of continuous functions endowed with the L^1 norm. It is also known that if M is a finite-dimensional Haar space, then strong uniqueness in the sense equivalent to (2.1) holds (see Kroó [11]), but the size of the α in (2.1) will in general depend upon f and M and can be arbitrarily large. On the other hand, from results due to Angelos and Schmidt [2] and Smarzewski [16], it is known that if M is a finite-dimensional subspace of $L^1(K,\mu)$, μ a nonatomic positive measure, then the set of $f \in L^1(K,\mu)$ that have a strongly unique best approximant from M (in the classic sense where $\alpha = 1$) is dense in $L^1(K,\mu)$.

Here is an example showing that in L^1 the stability of the metric projection can be of arbitrary order depending on f.

Example. In $L^1[-1,1]$ let $M := \operatorname{span}\{1\}$ be the set of constant functions, $N := \operatorname{span}\{g\}$, where

$$g(x) = \begin{cases} (x+1)/t, & -1 \le x \le -1+t, \\ 1, & -1+t \le x \le 1, \end{cases}$$

and t is any fixed value in $(0,1)$. $d(M,N)$ can be easily calculated since 1 is the best approximation to g from M, while g is the best approximation to 1 from N. In fact $d(M,N) = t/(4-t)$. Consider the functions $f(x) = |x|^a \operatorname{sgn} x$ for $a \in (0,1)$. Clearly $P_M f = 0$. From the equality

$$-\int_{[-1,t/4]} g + \int_{[t/4,1]} g = 0$$

and the characterization of best L^1-approximation, see Pinkus [14], it follows that $P_N f = (t/4)^a g$ is the best approximation to f from N. Hence

$$||P_M f - P_N f||_1 = \left(\frac{t}{4}\right)^a \frac{4-t}{2} \ge \frac{3}{8} t^a.$$

On the other hand we clearly have $d(M,N) \le t$. This means that the quantity $||P_M f - P_N f||_1$ cannot be bounded from above by $c_{M,f} d(M,N)^\beta$ with a β independent of a (i.e., independent of f) and a constant $c_{M,f}$ independent of N. This is very different from the situation in L^p, $p > 1$, or $C(K)$, where such a β always exists.

3. Stability of the metric projection in L^p, $p > 2$. In this section we show that with a different approach, a better stability estimate for the metric projection in L^p, $p > 2$, can be exhibited. More explicitly, we prove that when a certain (reasonable) property holds for the finite-dimensional subspace M, then for every subspace N of the same dimension we have that $\|P_M f - P_N f\|$ is of order $d(M,N)$. This is a considerable improvement when compared with the order $d(M,N)^{1/p}$ of Corollary 2.5. This refinement also fundamentally uses the fact that for $p \in (2,\infty)$ the L^p norm is twice differentiable.

Let M be an r-dimensional subspace of X. We first give a definition of Gateaux differentiability of P_M at M.

DEFINITION 3.1. *Let M be an r-dimensional linear subspace of X. We say that the metric projection P_M is Gateaux differentiable at M if for any basis $\{m_1, \ldots, m_r\}$ for M and any choice of r arbitrary elements $g_1, \ldots, g_r$ of X, we have that for all $f \in X \setminus M$ the limit*

$$\lim_{t \to 0} \frac{P_{M+tG} f - P_M f}{t}$$

exists, where $t \in \mathbb{R}$ and $M + tG = \operatorname{span}\{m_1 + tg_1, \ldots, m_r + tg_r\}$.

What we here term Gateaux differentiability is nonstandard in the sense that the directional differentiability of the metric projection is defined with respect to the subspace M. The Gateaux derivative of the metric projection $P_M f$ with respect to f (not M) was studied in L^p, $p > 2$, see, e.g., Holmes and Kripke [7].

Throughout this section we assume that μ is a positive measure on a set K and $L^p(K,\mu)$ is the usual set of real-valued μ-measurable functions f defined on K for which $|f|^p$ is μ-integrable over K with the standard L^p norm.

DEFINITION 3.2. *Assume M is a finite-dimensional subspace of $L^p(K,\mu)$. We say that M satisfies the Z_μ property if $\mu\{x : m(x) = 0\} = 0$ for every $m \in M$, $m \neq 0$.*

Note that if $\dim M > 1$, then this implies that μ must be a nonatomic measure. We assume that μ is a nonatomic measure in what follows. We first prove the following result characterizing Gateaux differentiable metric projections in $L^p(K,\mu)$.

THEOREM 3.1. *Let M be an r-dimensional subspace of $L^p(K,\mu)$, $p \geq 2$, where μ is a nonatomic measure. If $p = 2$, then P_M is Gateaux differentiable. If $p > 2$, then P_M is Gateaux differentiable at M if and only if M satisfies the Z_μ property.*

We first prove some results that will used be in the proof of Theorem 3.1. Let $m_1, \ldots, m_r$ be a basis for M. As M is finite-dimensional, there exist constants $C_0, c_0 > 0$ such that

$$c_0 \sum_{i=1}^{r} |a_i| \leq \left\| \sum_{i=1}^{r} a_i m_i \right\|_p \leq C_0 \sum_{i=1}^{r} |a_i| \tag{3.1}$$

for all $\{a_i\}_{i=1}^r$. For any given $g_1, \ldots, g_r$, set $G = \text{span}\{g_1, \ldots, g_r\}$ (we do not demand that G be r-dimensional) and let

$$M + tG := \text{span}\{m_1 + tg_1, \ldots, m_r + tg_r\}.$$

LEMMA 3.2. *Let $M = \text{span}\{m_1, \ldots, m_r\}$, $\dim M = r$, and $G = \text{span}\{g_1, \ldots, g_r\}$, $\dim G \leq r$, be subspaces of $L^p(K,\mu)$ with $\|g_i\|_p \leq 1$, $i = 1, \ldots, r$, and c_0 as given in (3.1). If $|t| \leq c_0/2$, then for all $\{a_i\}_{i=1}^r$*

$$c_0/2 \sum_{i=1}^{r} |a_i| \leq \left\| \sum_{i=1}^{r} a_i (m_i + tg_i) \right\|_p .$$

In addition, for each $t_1, t_2 \in [-c_0/2, c_0/2]$, we have

$$d(M + t_2 G, M + t_1 G) \leq \frac{2|t_2 - t_1|}{c_0}.$$

Proof. From the triangle inequality

$$\left\| \sum_{i=1}^{r} a_i (m_i + tg_i) \right\|_p \geq \left\| \sum_{i=1}^{r} a_i m_i \right\|_p - |t| \left\| \sum_{i=1}^{r} a_i g_i \right\|_p .$$

From (3.1), $|t| \leq c_0/2$, and $\|g_i\|_p \leq 1$ for all i, we can continue this as

$$\geq c_0 \sum_{i=1}^{r} |a_i| - \frac{c_0}{2} \sum_{i=1}^{r} |a_i| = \frac{c_0}{2} \sum_{i=1}^{r} |a_i|,$$

which is the first estimate of the lemma.

Now, for any $h = \sum_{i=1}^r a_i(m_i + t_2 g_i) \in M + t_2 G$, $\|h\|_p = 1$, we have by the previous estimate,

$$\begin{aligned} \|h - P_{M+t_1G} h\|_p &\leq \left\| \sum_{i=1}^{r} a_i (m_i + t_2 g_i) - \sum_{i=1}^{r} a_i (m_i + t_1 g_i) \right\|_p \\ &\leq |t_2 - t_1| \sum_{i=1}^{r} |a_i| \leq \frac{2|t_2 - t_1|}{c_0} \|h\|_p = \frac{2|t_2 - t_1|}{c_0}. \end{aligned}$$

From symmetry considerations it therefore follows that $d(M + t_2 G, M + t_1 G) \leq 2|t_2 - t_1|/c_0$. This proves the lemma. $\square$

LEMMA 3.3. *Assume M and G are as in Lemma* 3.2. *Let $P_{M+tG}f := \sum_{i=1}^r a_i(t)$ $(m_i + tg_i)$ denote the best approximation to f from $M + tG$. The $a_i(t)$, $i = 1, \dots, r$, are continuous functions of t on* $[-c_0/2, c_0/2]$.

Proof. First note that $E(t) := \|f - P_{M+tG}f\|_p$ is a continuous function of t. This follows from (2.3) and Lemma 3.2. Assume that $t_k \to t_0$ but $a_j(t_k)$ does not converge to $a_j(t_0)$ for some j. (In fact passing to a subsequence, if necessary, we may assume that it is separated from it.) Since $\|P_{M+tG}f\|_p \le 2\|f\|_p$ we have by Lemma 3.2 that $\{a_i(t_k)\}$ is a bounded sequence in k. Thus, passing if necessary to a subsequence, $a_i(t_k) \to a_i^*$ for every i. This implies that $P_{M+t_kG}f$ converges to some $P^*f \in M + t_0G$. Since $E(t)$ is continuous in t and the best approximation is unique (we have a strictly convex norm) we must have $P^*f = P_{M+t_0G}f$ and $a_j^* = a_j(t_0)$. This is a contradiction, hence the lemma holds. ▯

Proof of Theorem 3.1. Let $f \in L^p(K, \mu) \backslash M$, $p \ge 2$, where $M = \text{span}\{m_1, \dots, m_r\}$ is an r-dimensional subspace therein. As previously, let $P_{M+tG}f = \sum a_i(t)(m_i + tg_i)$ denote the best $L^p(K, \mu)$ approximation to f from $M + tG$. For $i = 1, \dots, r$, set

$$S_i(a_1, \dots, a_r, t) := \int_K \left| f - \sum_{j=1}^r a_j(m_j + tg_j) \right|^{p-1} \text{sgn}\left(f - \sum_{j=1}^r a_j(m_j + tg_j) \right)(m_i + tg_i)\, d\mu.$$

By the well-known L^p-orthogonality it follows that for every t

$$S_i(a_1(t), \dots, a_r(t), t) = 0, \qquad i = 1, \dots, r.$$

Taking into account that $p \ge 2$, straightforward differentiation yields at $t = 0$ and $(a_1, \dots, a_r) = (a_1(0), \dots, a_r(0))$,

$$\frac{\partial S_i}{\partial a_j} = (1 - p) \int_K |f - P_M f|^{p-2} m_i m_j \, d\mu, \qquad i, j = 1, \dots, r, \tag{3.2}$$

and

$$\begin{aligned} \frac{\partial S_i}{\partial t} = & \int_K |f - P_M f|^{p-1} \text{sgn}(f - P_M) g_i \, d\mu \\ & + (1 - p) \int_K |f - P_M f|^{p-2} m_i \left(\sum_{j=1}^r a_j g_j \right) d\mu, \qquad i = 1, \dots, r. \end{aligned}$$

Let

$$J(t) = \det \left(\frac{\partial S_i}{\partial a_j} \right)_{i,j=1}^r$$

at $(a_1, \dots, a_r) = (a_1(t), \dots, a_r(t))$, and let $J_\ell(t)$ be the determinant of the matrix resulting from $J(t)$, where we replace its ℓth column by $\{\frac{\partial S_i}{\partial t}\}_{i=1}^r$. We first claim that $J(0) \neq 0$. Indeed, if $J(0) = 0$, then from (3.2)

$$\sum_{i,j=1}^r b_i b_j \int_K |f - P_M f|^{p-2} m_i m_j \, d\mu = 0$$

for some $(b_1, \dots, b_r) \neq 0$. But this then implies that

$$\int_K |f - P_M f|^{p-2} m^2 \, d\mu = 0,$$

where $m = \sum_{i=1}^r b_i m_i \in M \setminus \{0\}$. If $p = 2$, then we actually get $\int_K m^2\, d\mu = 0$, which is an immediate contradiction. For $p > 2$ we see that this implies that m vanishes μ-a.e. on the set where $f - P_M f$ does not vanish. Since $f \in L^p(K,\mu) \setminus M$ this contradicts our assumption that M satisfies the Z_μ property. Hence $J(0) \neq 0$.

By the implicit function theorem

$$\frac{\partial a_i}{\partial t}(0) = \frac{J_i(0)}{J(0)}, \qquad i = 1, \ldots, r.$$

Now, by Lemma 3.3, the functions $a_i(t)$ are continuous functions of t for t sufficiently small. Thus we obtain

$$\begin{aligned}\lim_{t\to 0} \frac{P_{M+tG} f - P_M f}{t} &= \lim_{t\to 0} \frac{\sum_{i=1}^r a_i(t)(m_i + t g_i) - \sum_{i=1}^r a_i(0) m_i}{t} \\ &= \lim_{t\to 0} \sum_{i=1}^r \left(\frac{a_i(t) - a_i(0)}{t} \right) m_i + \sum_{i=1}^r a_i(t) g_i \\ &= \sum_{i=1}^r \frac{J_i(0)}{J(0)} m_i + a_i(0) g_i,\end{aligned}$$

implying that P_M is in fact Gateaux differentiable at M.

Assume $p > 2$ and M does not satisfy the Z_μ property. We will prove that P_M is not Gateaux differentiable at M. As M does not satisfy the Z_μ property there exists an $\widetilde{m}_1 \in M \setminus \{0\}$ such that $\widetilde{m}_1 = 0$ on a set $D \subset K$ of positive μ measure. We claim that there exists a basis $\{\widetilde{m}_1, \widetilde{m}_2, \ldots, \widetilde{m}_r\}$ for M such that the identically zero function is the best $L^p(K,\mu)$ approximation to $\widetilde{m}_1$ from $\operatorname{span}\{\widetilde{m}_2, \ldots, \widetilde{m}_r\}$ on D^c, the complement of D in K. To see this, let $M = \operatorname{span}\{m_1, \ldots, m_r\}$, where the $\{m_i\}_{i=1}^r$ are an arbitrary basis for M. Consider the equation

$$\int_{D^c} |\widetilde{m}_1|^{p-1} (\operatorname{sgn} \widetilde{m}_1)(a_1 m_1 + a_2 m_2 + \cdots + a_r m_r)\, d\mu = 0,$$

i.e.,

$$\sum_{i=1}^r a_i \int_{D^c} |\widetilde{m}_1|^{p-1} (\operatorname{sgn} \widetilde{m}_1) m_i \, d\mu = 0.$$

This one equation in r variables has $r-1$ linearly independent solutions (with respect to the $\{a_i\}_{i=1}^r$), and thus there are $r-1$ linearly independent elements $\{\widetilde{m}_2, \ldots, \widetilde{m}_r\}$ of M on K satisfying

$$\int_{D^c} |\widetilde{m}_1|^{p-1} (\operatorname{sgn} \widetilde{m}_1) \widetilde{m}_i \, d\mu = 0, \qquad i = 2, \ldots, r. \tag{3.3}$$

Obviously $\widetilde{m}_1 \notin \operatorname{span}\{\widetilde{m}_2, \ldots, \widetilde{m}_r\}$, and therefore $M = \operatorname{span}\{\widetilde{m}_1, \ldots, \widetilde{m}_r\}$. The orthogonality relations (3.3) imply that the identically zero function is the best $L^p(K,\mu)$ approximation to $\widetilde{m}_1$ from $\operatorname{span}\{\widetilde{m}_2, \ldots, \widetilde{m}_r\}$ on D^c.

Let $f \in L^p(K,\mu) \setminus \{0\}$ be such that $f = 0$ on D^c and the identically zero function is the best $L^p(K,\mu)$ approximation to f from M on D. That is, $\operatorname{supp} f \subseteq D$, and

$$\int_D |f|^{p-1} (\operatorname{sgn} f) m \, d\mu = 0, \qquad m \in M.$$

Such a choice of f is possible since μ is a nonatomic measure. (The restriction of $L^p(K,\mu)$ to D is infinite-dimensional.) Note that this also implies that $P_M f = 0$. Set

$g_1 := f$, $g_i := 0$, $i = 2, \ldots, r$, i.e., $M + tG = \operatorname{span}\{\widetilde{m}_1 + tf, \widetilde{m}_2, \ldots, \widetilde{m}_r\}$. Thus

$$\begin{aligned}\|f - P_{M+tG}f\|_p^p &= \inf_{a_j} \left\| f - a_1(\widetilde{m}_1 + tf) - \sum_{j=2}^{r} a_j \widetilde{m}_j \right\|_p^p \\ &= \inf_{a_j} \int_D \left| (1 - a_1 t) f - \sum_{j=2}^{r} a_j \widetilde{m}_j \right|^p d\mu + \int_{D^c} \left| a_1 \widetilde{m}_1 + \sum_{j=2}^{r} a_j \widetilde{m}_j \right|^p d\mu.\end{aligned}$$

Whatever the choice of a_1, it follows from the fact that the identically zero function is the best $L^p(K, \mu)$ approximation to f from M on D and the identically zero function is the best $L^p(K, \mu)$ approximation to $\widetilde{m}_1$ from $\operatorname{span}\{\widetilde{m}_2, \ldots, \widetilde{m}_r\}$ on D^c that

$$\inf_{a_2,\ldots,a_r} \int_D \left| (1 - a_1 t) f - \sum_{j=2}^{r} a_j \widetilde{m}_j \right|^p d\mu = \int_D |(1 - a_1 t) f|^p \, d\mu$$

and

$$\inf_{a_2,\ldots,a_r} \int_{D^c} \left| a_1 \widetilde{m}_1 + \sum_{j=2}^{r} a_j \widetilde{m}_j \right|^p d\mu = \int_{D^c} |a_1 \widetilde{m}_1|^p \, d\mu.$$

Thus

$$\begin{aligned}\|f - P_{M+tG}f\|_p^p &= \inf_{a_1} \int_D |(1 - a_1 t) f|^p \, d\mu + \int_{D^c} |a_1 \widetilde{m}_1|^p \, d\mu \\ &= \inf_{a_1} |1 - a_1 t|^p \int_D |f|^p \, d\mu + |a_1|^p \int_{D^c} |\widetilde{m}_1|^p \, d\mu.\end{aligned}$$

Setting

$$A := \int_D |f|^p \, d\mu, \quad B := \int_{D^c} |\widetilde{m}_1|^p \, d\mu,$$

this reduces to considering the ℓ_2^p minimum problem

$$\inf_{a_1} |1 - a_1 t|^p A + |a_1|^p B.$$

Assume $t > 0$. The standard ℓ_2^p orthogonality implies that the infimum (minimum) is uniquely attained by $a_1^*(t)$ satisfying

$$-|1 - a_1^*(t)t|^{p-1} \operatorname{sgn}(1 - a_1^*(t)t)\, At + |a_1^*(t)|^{p-1} \operatorname{sgn}(a_1^*(t))\, B = 0.$$

As $t > 0$, this implies that $\operatorname{sgn}(1 - a_1^*(t)t)\ \operatorname{sgn}(a_1^*(t)) > 0$. Thus $a_1^*(t), 1 - a_1^*(t)t > 0$, and

$$\frac{a_1^*(t)}{1 - a_1^*(t)t} = \left(\frac{At}{B} \right)^{\frac{1}{p-1}}. \tag{3.4}$$

The right-hand side of (3.4) tends to zero as $t \downarrow 0$, and

$$0 < a_1^*(t) \leq \frac{a_1^*(t)}{1 - a_1^*(t)t}.$$

Thus $a_1^*(t) \to 0$ as $t \downarrow 0$, which implies that

$$\frac{a_1^*(t)}{1 - a_1^*(t)t} \le 2a_1^*(t)$$

for $t > 0$ sufficiently small. We therefore have

$$\left(\frac{At}{B}\right)^{\frac{1}{p-1}} \le 2a_1^*(t) \tag{3.5}$$

for $t > 0$ sufficiently small.

Recall that $P_M f = 0$ and $P_{M+tG} f = a_1^*(t)(\widetilde{m}_1 + tf)$. Thus from (3.5) we have

$$\|P_M f - P_{M+tG} f\|_p = \|a_1^*(t)(\widetilde{m}_1 + tf)\|_p = |a_1^*(t)| \|\widetilde{m}_1 + tf\|_p \ge Ct^{\frac{1}{p-1}} \tag{3.6}$$

for $t > 0$ sufficiently small. Since $p > 2$ this proves that P_M is not Gateaux differentiable at M. □

We now present the main result of this section. We prove that if the Z_μ property holds for a finite-dimensional subspace M in $L^p(K, \mu)$, $p > 2$, then the size of $\|P_M f - P_N f\|_p$ is of order at most $d(M, N)$. Moreover the Z_μ property is necessary to attain this improvement.

THEOREM 3.4. *Consider the space $L^p(K, \mu)$, $p > 2$, where μ is a nonatomic measure. Let M be an r-dimensional subspace of $L^p(K, \mu)$ satisfying the Z_μ property, i.e., $\mu\{x : m(x) = 0\} = 0$ for every $m \in M \setminus \{0\}$. Then for every $f \in L^p(K, \mu)$, $p > 2$, there exists a constant $c_{M,f}$ depending on M and f such that for any r-dimensional subspace N of $L^p(K, \mu)$, $p > 2$, we have*

$$\|P_M f - P_N f\|_p \le c_{M,f} d(M, N).$$

Furthermore, the Z_μ property of M is necessary in order for the above estimate to hold.

Proof. The necessity of the Z_μ property follows by the same argument as in the proof of Theorem 3.1. That is, if M does not satisfy the Z_μ property, then from (3.6) and the result of Lemma 3.2, namely, the estimate

$$d(M, M + tG) \le 2|t|/c_0,$$

where $c_0 > 0$ depends only on M, it follows that we cannot have

$$\|P_M f - P_{M+tG} f\|_p \le c_{M,f} d(M, M + tG),$$

where $c_{M,f}$ is independent of t.

Let $M = \text{span}\{m_1, \dots, m_r\}$ be an r-dimensional subspace of $L^p(K, \mu)$ with $\|m_i\|_p = 1$, $i = 1, \dots, r$, satisfying the Z_μ property. Let $G = \text{span}\{g_1, \dots, g_r\}$ be such that $\|g_i\|_p \le 1$, $i = 1, \dots, r$. We use the previous notation for $M + tG$ and P_{M+tG}, and for $f \in L^p(K, \mu)$ we set

$$P_{M+tG} f = \sum_{i=1}^r a_i(t)(m_i + tg_i).$$

Assume $f \in L^p(K, \mu) \setminus M$. As in the proof of Theorem 3.1, for $i = 1, \dots, r$, set

$$S_i(a_1, \dots, a_r, t) := \int_K \left| f - \sum_{j=1}^r a_j(m_j + tg_j) \right|^{p-1} \text{sgn}\left(f - \sum_{j=1}^r a_j(m_j + tg_j) \right)(m_i + tg_i)\, d\mu.$$

Thus, for every t,

$$S_i(a_1(t),\dots,a_r(t),t)=0, \qquad i=1,\dots,r.$$

Furthermore, for $(a_1,\dots,a_r)=(a_1(t),\dots,a_r(t))$

$$y_{ij}(t):=\frac{\partial S_i}{\partial a_j}=(1-p)\int_K |f-P_{M+tG}f|^{p-2}(m_i+tg_i)(m_j+tg_j)\,d\mu, \qquad i,j=1,\dots,r,$$

and

$$\begin{aligned}\frac{\partial S_i}{\partial t}&=\int_K |f-P_{M+tG}f|^{p-1}\operatorname{sgn}(f-P_{M+tG}f)g_i\,d\mu\\ &\quad+(1-p)\int_K |f-P_{M+tG}f|^{p-2}(m_i+tg_i)\left(\sum_{j=1}^r a_jg_j\right)d\mu, \qquad i=1,\dots,r.\end{aligned}$$

Set $J(t)=\det(y_{ij}(t))_{i,j=1}^r$, and let $J_\ell(t)$ denote the determinant of the matrix obtained from $J(t)$ on replacing its ℓth column by $\{\frac{\partial S_i}{\partial t}\}_{i=1}^r$. Analogously to the proof of Theorem 3.1, it follows that the Z_μ property of M implies that $J(0)\neq 0$.

We claim that there exists a $t_0(M,f)$ such that $J(t)\neq 0$ for all t satisfying $|t|\le t_0(M,f)$, where $t_0(M,f)$ depends upon M and f. From Lemma 3.3 $J(t)$ is a continuous function of t and $J(0)$, as shown in the proof of Theorem 3.1, is a nonzero quantity depending only on M and f. We will show that

$$|J(0)-J(t)|\le c|t|^{\frac{\gamma}{p}},$$

where $\gamma:=\min\{1,p-2\}$ and c depends only on M and f.

In what follows, all constants c_k will depend upon, at most, M and f. We first get a bound on $|y_{ij}(t)|$. From twice applying Hölder's inequality we have

(3.7)
$$\begin{aligned}|y_{ij}(t)|&\le (p-1)\int_K |f-P_{M+tG}f|^{p-2}|m_i+tg_i|\,|m_j+tg_j|\,d\mu\\ &\le (p-1)\left(\int_K |f-P_{M+tG}f|^p d\mu\right)^{(p-2)/p}\left(\int_K |m_i+tg_i|^{p/2}|m_j+tg_j|^{p/2}d\mu\right)^{2/p}\\ &\le (p-1)\|f-P_{M+tG}f\|_p^{p-2}\|m_i+tg_i\|_p\,\|m_j+tg_j\|_p\\ &\le 4(p-1)\|f\|_p^{p-2}\end{aligned}$$

since $\|f-P_{M+tG}f\|_p\le\|f\|_p$ and $\|m_i+tg_i\|_p\le 2$ when $|t|\le 1$. Because $J(t)=\det(y_{i,j}(t))$ we therefore have for each $i,j\in\{1,\dots,r\}$ that

$$\left|\frac{\partial J}{\partial y_{i,j}}\right|\le c_1, \qquad |t|\le 1.$$

It remains to consider the dependence of $y_{i,j}(t)$ on t. Note that by Corollary 2.5 of the previous section we have for $p>2$

$$\|P_{M+tG}f-P_Mf\|_p\le c_p\|f\|_p d(M+tG,M)^{\frac{1}{p}}$$

for some constant c_p depending only upon p. From Lemma 3.2, we have

$$d(M+tG,M)\le 2|t|/c_0,$$

and therefore

$$\|P_{M+tG}f - P_M f\|_p \le c_2|t|^{\frac{1}{p}}. \tag{3.8}$$

Since

$$y_{i,j}(t) = (1-p)\int_K |f - P_{M+tG}f|^{p-2}(m_i + tg_i)(m_j + tg_j)\,d\mu, \qquad i,j = 1,\dots,r,$$

it is easy to see that

$$y_{i,j}(t) - y_{i,j}(0) = (1-p)\int_K (|f - P_{M+tG}f|^{p-2} - |f - P_M f|^{p-2})m_i m_j\,d\mu + O(t), \tag{3.9}$$

where the $O(t)$ term can be bounded above by $3(p-1)t\|f\|_p^{p-2}$, just as in the estimate (3.7). We estimate the above integrand as follows. For $2 < p \le 3$ we have

$$\big||f - P_{M+tG}f|^{p-2} - |f - P_M f|^{p-2}\big| \le |P_{M+tG}f - P_M f|^{p-2},$$

and thus by the above, twice applying Hölder's inequality, and the fact that $\|m_i\|_p = 1$ we obtain

$$\begin{aligned}&\left|\int_K (|f - P_{M+tG}f|^{p-2} - |f - P_M f|^{p-2})m_i m_j\,d\mu\right| \\ &\quad\le \int_K |P_{M+tG}f - P_M f|^{p-2} m_i m_j\,d\mu \le \|P_{M+tG}f - P_M f\|_p^{p-2}.\end{aligned}$$

For $p > 3$, we have by the mean value theorem

$$\begin{aligned}&\big||f - P_{M+tG}f|^{p-2} - |f - P_M f|^{p-2}\big| \\ &\quad\le (p-2)(|f - P_{M+tG}f| + |f - P_M f|)^{p-3}|P_M f - P_{M+tG}f|.\end{aligned}$$

Applying Hölder's inequality we obtain

$$\begin{aligned}&\int_K (|f - P_{M+tG}f| + |f - P_M f|)^{p-3}|P_M f - P_{M+tG}f|\,|m_i m_j|\,d\mu \\ &\quad\le \left(\int_K (|f - P_{M+tG}f| + |f - P_M f|)^{\frac{(p-3)p}{p-1}}|m_i m_j|^{\frac{p}{p-1}}\,d\mu\right)^{\frac{p-1}{p}} \|P_M f - P_{M+tG}f\|_p.\end{aligned}$$

Continuing we see that

$$\begin{aligned}&\left(\int_K (|f - P_{M+tG}f| + |f - P_M f|)^{\frac{(p-3)p}{p-1}}|m_i m_j|^{\frac{p}{p-1}}\,d\mu\right)^{\frac{p-1}{p}} \\ &\quad\le \left(\int_K (|f - P_{M+tG}f| + |f - P_M f|)^p\,d\mu\right)^{\frac{p-3}{p}} \left(\int_K |m_i m_j|^{\frac{p}{2}}\,d\mu\right)^{\frac{2}{p-1}}.\end{aligned}$$

Again applying Hölder's inequality and the fact that $\|m_i\|_p = 1$ we get

$$\left(\int_K |m_i m_j|^{\frac{p}{2}}\,d\mu\right)^{\frac{2}{p-1}} \le 1.$$

Furthermore, by the triangle inequality

$$\left(\int_K (|f - P_{M+tG}f| + |f - P_M f|)^p \, d\mu\right)^{\frac{p-3}{p}} \le (\|f - P_{M+tG}f\|_p + \|f - P_M f\|_p)^{p-3} \le (2\|f\|_p)^{p-3}.$$

Combining these estimates together with (3.8) and (3.9) we arrive at

$$|y_{i,j}(t) - y_{i,j}(0)| \le c_3 |t|^{\frac{\gamma}{p}},$$

where $\gamma := \min\{1, p-2\}$ and c_3 depends on M and f.

This last estimate and (3.7) yields

$$|J(0) - J(t)| \le c_4 |t|^{\frac{\gamma}{p}}, \qquad |t| \le 1,$$

and therefore for a suitable choice of $t_0(M,f)$ we have

$$|J(t)| \ge |J(0)|/2 > 0, \qquad |t| \le t_0(M,f).$$

Now, by the implicit function theorem (and using that $|\frac{\partial S_i}{\partial t}| \le c_5$ by repeated application of Hölder's inequality),

$$\left|\frac{\partial a_j}{\partial t}(t)\right| = \left|\frac{J_j(t)}{J(t)}\right| \le c_6, \qquad |t| \le t_0(M,f), \qquad j = 1, \dots, r. \tag{3.10}$$

Assume, without loss of generality, that $t_0(M,f) \le c_0/2$. From Lemma 3.2

$$\left\|\sum_{i=1}^r a_i(t) g_i\right\|_p \le \sum_{i=1}^r |a_i(t)| \le \frac{2}{c_0}\|P_{M+tG}f\|_p \le \frac{4}{c_0}\|f\|_p, \qquad |t| \le t_0(M,f). \tag{3.11}$$

Thus we obtain, using (3.10) and (3.11),

$$\begin{aligned} \|P_{M+tG}f - P_M f\|_p &= \left\|\sum_{i=1}^r a_i(t)(m_i + t g_i) - \sum_{i=1}^r a_i(0) m_i\right\|_p \\ &\le c_7|t| + \left\|\sum_{i=1}^r (a_i(t) - a_i(0)) m_i\right\|_p \le c_8|t|, \qquad |t| \le t_0(M,f). \end{aligned} \tag{3.12}$$

Recall that these estimates hold for any set of $\{g_i\}_{i=1}^r$ satisfying $\|g_i\|_p \le 1$, $i = 1, \dots, r$, in $L^p(K,\mu)$, $p > 2$.

Now, let N be any fixed r-dimensional subspace of $L^p(K,\mu)$, $p > 2$. As above, consider a basis $\{m_i\}_{i=1}^r$ for M, where $\|m_i\|_p = 1$, $i = 1, \dots, r$. For each $m_i \in M$ let $n_i \in N$ be such that $\|m_i - n_i\|_p \le d(M,N)$, $i = 1, \dots, r$. We assume for the moment that $d(M,N) < t_0(M,f)$. Set $t = d(M,N)$ and

$$g_i := \frac{n_i - m_i}{d(M,N)}, \qquad i = 1, \dots, r.$$

Thus $\|g_i\|_p \le 1$ for all i, and

$$M + tG = \operatorname{span}\{m_1 + t g_1, \dots, m_r + t g_r\} = \operatorname{span}\{n_1, \dots, n_r\}.$$

From Lemma 3.2, the $\{m_i + tg_i\}_{i=1}^r$ are linearly independent, and thus

$$N = \operatorname{span}\{n_1, \ldots, n_r\}$$

and $P_{M+tG}f = P_N f$. From (3.12) we therefore obtain

$$\|P_N f - P_M f\|_p \le c_8|t| = c_8 d(M,N),$$

assuming $d(M,N) < t_0(M,f)$. This is the required estimate. Moreover the estimate holds trivially if $d(M,N) \ge c$ for any fixed $c > 0$. This completes the proof when $f \in L^p(K,\mu) \setminus M$. If $f \in M$, then clearly from the definition of $d(M,N)$ we have

$$\|P_M f - P_N f\|_p = \|f - P_N f\|_p \le \|f\|_p d(M,N).$$

The proof of Theorem 3.4 is complete. □

Remark. From a reading of the proof of Theorem 3.1 it follows that if $f \in L^p(K,\mu) \setminus M$, $p > 2$, is such that there does not exist an $m \in M$, $m \ne 0$, that vanishes μ-a.e. on the set where $f - P_M f$ does not vanish, then $J(0) \ne 0$ therein. This is sufficient to imply from the proof of Theorem 3.4 that there exists a constant $c_{M,f}$, depending on M and f, such that for any r-dimensional subspace N of $L^p(K,\mu)$, $p > 2$, we have

$$\|P_M f - P_N f\|_p \le c_{M,f} d(M,N).$$

Here we do not need to assume that μ is a nonatomic measure.

In this next example we consider ℓ_2^p, $p > 2$. If M is a one-dimensional subspace spanned by m, then for every f such that $m(f - P_M f)$ is not the zero element we have $\|P_M f - P_N f\|_p \le c\, d(M,N)$. On the other hand, if $m(f - P_M f)$ vanishes identically, then $\|P_M f - P_N f\|_p$ is of exact order $d(M,N)^{1/(p-1)}$.

Example. Let $X = \mathbb{R}^2$ endowed with the ℓ_2^p norm, $p > 2$. If $M = \operatorname{span}\{(a,b)\}$, where both a and b are nonzero, then from above the remark, it follows that the rate of stability to every f is at most linear. As such, let us assume without loss of generality that $M = \operatorname{span}\{(1,0)\}$. For $t \in (0,1)$, set $N_t = \operatorname{span}\{(1,t)\}$. From symmetry considerations

$$\max_{\substack{m\in M\\ \|m\|_p=1}} \min_{n\in N_t} \|m - n_t\|_p = \min_\alpha \|(1,0) - \alpha(1,t)\|_p = \min_\alpha \left(|1-\alpha|^p + |\alpha t|^p\right)^{1/p}.$$

The minimum is attained when

$$\alpha^* = \frac{1}{1 + t^{p/(p-1)}} = \frac{1}{1+t^q},$$

where $1/p + 1/q = 1$, and

$$\left(|1-\alpha^*|^p + |\alpha^* t|^p\right)^{1/p} = \frac{t}{(1+t^q)^{1/q}}.$$

On the other hand, also from symmetry considerations, we have

$$\max_{\substack{n\in N_t\\ \|n\|_p=1}} \min_{m\in M} \|n - m\|_p = \min_\alpha \frac{\|(1,t) - \alpha(1,0)\|_p}{\|(1,t)\|_p} = \frac{\|(0,t)\|_p}{\|(1,t)\|_p} = \frac{t}{(1+t^p)^{1/p}}.$$

Since $q < 2 < p$ we therefore have

$$d(M, N_t) = \frac{t}{(1+t^p)^{1/p}}.$$

Let $f = (0,1)$. Obviously $P_M f = 0$. A calculation gives

$$P_{N_t} f = \left(\frac{t^{q/p}}{1+t^q}, \frac{t^q}{1+t^q} \right),$$

and therefore

$$\|P_M f - P_{N_t} f\|_p = \frac{t^{1/(p-1)}(1+t^p)^{1/p}}{1+t^q}.$$

Thus, since $p \geq 2$, we have

$$\|P_M f - P_{N_t} f\|_p \leq C_p \left[d(M, N_t) \right]^{1/(p-1)}$$

for $t > 0$ small, and this is the optimal possible power. By symmetry, when considering N_t for $t \in (-1, 0)$, the same estimate holds except that we replace t by $|t|$. Note that this is better than the power $1/p$ guaranteed by Corollary 2.5 but worse than linearity.

A more detailed calculation shows that for $f = (a, 1)$, any $a \in \mathbb{R}$, we have $P_M f = (a, 0)$ and

$$P_{N_t} f = \left(\frac{a + |t|^{1/(p-1)} \operatorname{sgn} t}{1+|t|^q}, \frac{at + |t|^{p/(p-1)}}{1+|t|^q} \right).$$

From this it follows that $\|P_M f - P_{N_t} f\|_p$ is of order $t^{1/(p-1)}$ as t tends to 0, i.e., of order $d(M, N_t)^{1/(p-1)}$. (For $f \in M$ we always have $\|P_M f - P_{N_t} f\|_p \leq \|f\|_p d(M, N_t)$.) Thus we see that for this subspace M of ℓ_2^p, $p > 2$, the sharp rate of stability for every f not in M is of order $1/(p-1)$.

Note that when $M = \operatorname{span}\{(1,0)\}$ and $f = (a, b)$ we have $f - P_M f = (0, b)$, and thus there exists an $m \in M$, $m \neq 0$, that vanishes on the set where $f - P_M f$ does not vanish.

Why the power $1/(p-1)$ in the above example? One explanation may be in the fact that M is a subspace of co-dimension 1.

Let M and N be closed linear subspaces of co-dimension 1 in X. Set

$$M = \{m : h(m) = 0\},$$
$$N = \{n : g(n) = 0\},$$

where $h, g \in X^*$. Assume $\|h\|_{X^*} = \|g\|_{X^*} = 1$ and also that there exist $r, s \in X$, $\|r\|_X = \|s\|_X = 1$ for which $h(r) = g(s) = 1$. That is, we assume that h and g attain their norms on r and s, respectively. In this case we have

$$P_M f = f - h(f) r$$

and

$$P_N f = f - g(f) s.$$

This follows from the fact that $h(P_M f) = 0$ and thus $P_M f \in M$, and

$$\|f - P_M f\|_X = |h(f)| = |h(f - m)| \le \|f - m\|_X$$

for all $m \in M$.

Now

$$M^\perp = \operatorname{span}\{h\},$$
$$N^\perp = \operatorname{span}\{g\},$$

and from Proposition 1.2

$$d(M, N) = d(M^\perp, N^\perp) = \max\{\min_\alpha \|h - \alpha g\|_{X^*}, \min_\beta \|g - \beta h\|_{X^*}\}.$$

Note that

$$d(M, N) \le \min\{\|h - g\|_{X^*}, \|h + g\|_{X^*}\}.$$

Furthermore, essentially by definition,

$$\widetilde{d}(M^\perp, N^\perp) = \min\{\|h - g\|_{X^*}, \|h + g\|_{X^*}\}.$$

Thus, from (1.2)

$$\min\{\|h - g\|_{X^*}, \|h + g\|_{X^*}\} \le 2d(M, N).$$

We can always replace h by $-h$ and thus r by $-r$. So let us assume that $\|h - g\|_{X^*} \le 2d(M, N)$. Then

$$\begin{aligned}\|P_M f - P_N f\|_X &= \|h(f)r - g(f)s\|_X = \|h(f)r - g(f)r + g(f)r - g(f)s\|_X \\ &\le \|h(f)r - g(f)r\|_X + \|g(f)r - g(f)s\|_X \\ &\le |h(f) - g(f)| + |g(f)|\, \|r - s\|_X.\end{aligned}$$

From the previous inequality

$$|h(f) - g(f)| \le \|h - g\|_{X^*}\, \|f\|_X \le 2\|f\|_X d(M, N),$$

and we also have

$$|g(f)| \le \|f\|_X.$$

Thus

$$\|P_M f - P_N f\|_X \le 2\|f\|_X d(M, N) + \|f\|_X\, \|r - s\|_X. \tag{3.13}$$

It remains to estimate $\|r - s\|_X$. That is, if $\|h\|_{X^*} = \|g\|_{X^*} = \|r\|_X = \|s\|_X = 1$ are such that $h(r) = g(s) = 1$, and $\|h - g\|_{X^*} \le 2d(M, N)$, can we get a good estimate for $\|r - s\|_X$?

We claim that the following results hold in L^p, $1 < p < \infty$.

Theorem 3.5. *Assume M and N are closed linear subspaces of $L^p(K, \mu)$ of co-dimension 1. If $p \in (2, \infty)$, then*

$$\|P_M f - P_N f\|_p \le 4\|f\|_p d(M, N)^{1/(p-1)}.$$

If $p \in (1,2)$, *then*

$$\|P_M f - P_N f\|_p \le c_p \|f\|_p d(M,N),$$

where c_p *is a constant dependent only upon* p.

Proof. In both cases we will estimate $\|r-s\|_p$ and apply inequality (3.13).

Assume $p \in (2,\infty)$ and $1/p+1/q=1$. Thus $h \in L^q(K,\mu)$ and $r = |h|^{1/(p-1)} \operatorname{sgn} h$. Similarly $s = |g|^{1/(p-1)} \operatorname{sgn} g$. We are given that $\|h-g\|_q \le 2d(M,N)$. Note that

$$\big||a|^\ell \operatorname{sgn} a - |b|^\ell \operatorname{sgn} b\big| \le 2^{1-\ell}|a-b|^\ell$$

for all $a, b \in \mathbb{R}$, and $\ell \in (0,1)$. Now

$$\|r-s\|_p = \left(\int_K |r-s|^p \, d\mu\right)^{1/p} = \left(\int_K \big|\,|h|^{1/(p-1)} \operatorname{sgn} h - |g|^{1/(p-1)} \operatorname{sgn} g\,\big|^p \, d\mu\right)^{1/p}.$$

Since $p > 2$, $1/(p-1) < 1$, and thus we can continue the above as

$$\begin{aligned}&\le 2^{(p-2)/(p-1)} \left(\int_K |h-g|^{p/(p-1)} \, d\mu\right)^{1/p} = 2^{(p-2)/(p-1)} \left(\int_K |h-g|^q \, d\mu\right)^{1/p}\\ &\le 2^{(p-2)/(p-1)} \left(2d(M,N)\right)^{q/p} = 2d(M,N)^{1/(p-1)}.\end{aligned}$$

We have proved that

$$\|P_M f - P_N f\|_p \le 2\|f\|_p [d(M,N) + d(M,N)^{1/(p-1)}] \le 4\|f\|_p d(M,N)^{1/(p-1)},$$

which is the desired result.

Let us now assume that $p \in (1,2)$. As above $r = |h|^{1/(p-1)} \operatorname{sgn} h$, $s = |g|^{1/(p-1)} \operatorname{sgn} g$, and we are given that $\|h-g\|_q \le 2d(M,N)$. Now for $\ell \in (1,\infty)$ consider

$$\big||a|^\ell \operatorname{sgn} a - |b|^\ell \operatorname{sgn} b\big|.$$

If $\operatorname{sgn} a = \operatorname{sgn} b$, then by the mean value theorem

$$\big||a|^\ell \operatorname{sgn} a - |b|^\ell \operatorname{sgn} b\big| = \big||a|^\ell - |b|^\ell\big| \le \ell(|a|+|b|)^{\ell-1}|a-b|. \tag{3.14}$$

If $\operatorname{sgn} a = -\operatorname{sgn} b$, then

$$\begin{aligned}\big||a|^\ell \operatorname{sgn} a - |b|^\ell \operatorname{sgn} b\big| &= |a|^\ell + |b|^\ell\\ &\le (|a|+|b|)^\ell = (|a|+|b|)^{\ell-1}(|a|+|b|) \le \ell(|a|+|b|)^{\ell-1}|a-b|.\end{aligned}$$

Thus (3.14) is valid in both cases.

Now

$$\|r-s\|_p = \left(\int_K |r-s|^p \, d\mu\right)^{1/p} = \left(\int_K \big|\,|h|^{1/(p-1)} \operatorname{sgn} h - |g|^{1/(p-1)} \operatorname{sgn} g\,\big|^p \, d\mu\right)^{1/p}.$$

Since $1 < p < 2$, $1/(p-1) > 1$, and thus we can apply (3.14) to obtain

$$\le \left(\frac{1}{p-1}\right) \left(\int_K [|h|+|g|]^{(2-p)p/(p-1)} \, |h-g|^p \, d\mu\right)^{1/p}.$$

Applying Hölder gives

$$\leq \left(\frac{1}{p-1}\right)\left(\left(\int_K [|h|+|g|]^{p/(p-1)}\ d\mu\right)^{2-p}\left(\int_K |h-g|^q\, d\mu\right)^{p/q}\right)^{1/p}$$
$$= \left(\frac{1}{p-1}\right)\left(\int_K [|h|+|g|]^q\ d\mu\right)^{(2-p)/p}\|h-g\|_q.$$

By the triangle inequality and since $\|h\|_q = \|g\|_q = 1$,

$$\left(\int_K [|h|+|g|]^q\ d\mu\right)^{(2-p)/p} \leq (\|h\|_q + \|g\|_q)^{(2-p)q/p} = 2^{(2-p)/(p-1)}.$$

Thus

$$\|r-s\|_p \leq a_p d(M,N)$$

for some a_p dependent upon p, implying that

$$\|P_M f - P_N f\|_p \leq c_p \|f\|_p d(M,N),$$

where c_p is a constant dependent upon p. ∎

4. Some open questions. This paper leaves open a number of questions. Clearly the Hilbert space situation is the most satisfactory since we have a sharp linear bound for the stability of the orthogonal projections in this case. What is unclear is how to exactly calculate $d(M,N)$ if both M and N are of infinite dimension and infinite co-dimension.

In L^p, $2 < p < \infty$, linear stability of the metric projection onto finite-dimensional subspaces holds if and only if the subspaces satisfy the Z_μ property. However, assuming the Z_μ property, the constant $c_{M,f}$ exhibits a dependence on f. It would be of interest to determine whether (or when) $c_{M,f} = c_M\|f\|_p$. If the Z_μ property fails (which includes the case of atomic measures) or the subspace is infinite-dimensional, then we have by Corollary 2.5 an upper bound of the order $d(M,N)^{1/p}$. Furthermore relation (3.6) yields a lower bound of the form $cd(M,N)^{1/(p-1)}$ which can be attained, as is illustrated by a previous example. And in the case of subspaces of co-dimension 1 we have an upper bound of this order. There is a gap between these upper and lower bounds, which raises the question of finding asymptotically sharp upper bounds in L^p, $2 < p < \infty$, when the Z_μ property fails for finite-dimensional M or when the subspace M is infinite-dimensional.

If $1 < p < 2$, then Corollary 2.5 provides an upper bound of order $d(M,N)^{1/2}$. This is undoubtedly not optimal. What is the best order in this case? Is it possible that we always have linear stability as exhibited in Theorem 3.5?

REFERENCES

[1] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space*, Vol. I, Frederick Ungar, New York, 1961.

[2] J. Angelos and D. Schmidt, *Strong uniqueness in $L^1(X,\Sigma,\mu)$*, in Approximation Theory IV, C. K. Chui, L. L. Schumaker, and J. D. Ward, eds., Academic Press, New York, 1983. pp. 297–302.

[3] B. O. Björnestål, *Local Lipschitz continuity of the metric projection operator*, in Approximation Theory, Banach Center Publ. 4, PWN, Warsaw, 1979, pp. 43–53.

[4] I. C. GOHBERG AND M. G. KREIN, *The basic propositions on defect numbers, root numbers and indices of linear operators*, Amer. Math. Soc. Transl. Ser. 2, 13 (1960), pp. 185–264, (in English); Uspekhi Mat. Nauk (N.S.), 12 (1957), pp. 43–118 (in Russian).
[5] I. C. GOHBERG AND A. S. MARKUS, *Two theorems on the gap between subspaces of a Banach space*, Uspekhi Mat. Nauk, 14 (1959), pp. 135–140 (in Russian).
[6] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., The John Hopkins University Press, Baltimore, MD, 1996.
[7] R. HOLMES AND B. KRIPKE, *Smoothness of approximation*, Michigan Math. J., 15 (1968), pp. 225–248.
[8] T. KATO, *Perturbation Theory for Linear Operators*, Springer-Verlag, New York, 1980.
[9] M. G. KREIN AND M. A. KRASNOSEL'SKII, *Fundamental theorems on the extension of Hermitian operators and certain of their applications to the theory of orthogonal polynomials and the problem of moments*, Uspekhi Mat. Nauk (N. S.), 2 (1947), pp. 60–106 (in Russian).
[10] M. G. KREIN, M. A. KRASNOSEL'SKII, AND D. P. MILMAN, *On the defect numbers of linear operators in a Banach space and on some geometrical questions*, Sb. Trudov Inst. Mat. Akad. Nuajk Ukrain. SSR, 11 (1948), pp. 97–112 (in Russian).
[11] A. KROÓ, *On strong unicity of L_1-approximation*, Proc. Amer. Math. Soc., 83 (1981), pp. 725–729.
[12] A. KROÓ AND A. PINKUS, *Strong uniqueness*, Surv. Approx. Theory, 5 (2010), pp. 1–91.
[13] D. J. NEWMAN AND H. S. SHAPIRO, *Some theorems on Cebysev approximation*, Duke Math. J., 30 (1963), pp. 673–682.
[14] A. M. PINKUS, *On L^1-Approximation*, Cambridge University Press, Cambridge, 1989.
[15] R. SMARZEWSKI, *Strongly unique minimization of functionals in Banach spaces with applications to theory of approximation and fixed points*, J. Math. Anal. Appl., 115 (1986), pp. 155–172.
[16] R. SMARZEWSKI, *Strong uniqueness of best approximations in an abstract L_1 space*, J. Math. Anal. Appl., 136 (1988), pp. 347–351.
[17] R. SMARZEWSKI, *Strong unicity of order* 2 *in* $C(T)$, J. Approx. Theory, 56 (1989), pp. 306–315.
[18] R. SMARZEWSKI, *On an inequality of Bynum and Drew*, J. Math. Anal. Appl., 150 (1990), pp. 146–150.