JOURNAL OF APPROXIMATION THEORY 4, 194–203 (1971)

Sets of Best Approximation in Certain Classes of Normed Spaces

BRUCE L. CHALMERS

AND

LE BARON O. FERGUSON

Department of Mathematics, University of California, Riverside, California 92502

Communicated by Oved Shisha

Received January 30, 1970

INTRODUCTION

Let C^rI denote the space of r-times continuously differentiable functions on the interval $I = [b, c]$ of the real line $\mathbf{R}$. The question of uniqueness of best approximation of functions in C^rI by functions in a finite-dimensional subspace, with respect to various norms, has been investigated in several papers. For example, Garkavi [3] examines the problem using the ordinary supremum norm

$$\|f\|_\infty = \sup_{x \in I} |f(x)|,$$

while Moursund [7] and Johnson [5] use the norm

$$\|f\| = \max[\|f\|_\infty, \|f^{(1)}\|_\infty, \ldots, \|f^{(r)}\|_\infty].$$

In this paper we consider uniqueness of best approximation in certain classes of normed spaces which include, e.g., C^rI with the norm

$$\|f\| = \max[|f(a)|, |f^{(1)}(a)|, \ldots, |f^{(r-1)}(a)|, \|f^{(r)}\|_p],$$

$1 \leqslant p \leqslant \infty$, where $\|\cdot\|_p$ denotes the L^p norm and a is a fixed point in I. Applied to these norms, Sections 2–4 yield results in the cases $p = \infty$, $1 < p < \infty$, and $p = 1$, respectively.

In Section 1, the classes of normed spaces to be treated are described, and a representation for continuous linear functionals on these spaces is obtained. In Sections 2–4, various subclasses of these normed spaces are investigated. The results of Section 2 are obtained by realizing the spaces there as spaces of continuous functions on a compact topological space and then applying the Haar condition as generalized by Rubenstein [3, *p.* 94]. The results of Sections 3 and 4 are obtained by using the representation of Section 1 and the Hahn–Banach Theorem. Section 5 is a remark concerning certain other norms.

1. Preliminaries

Let B be any normed linear space with norm $\|\cdot\|_B$. Let S be a vector space, and T a nonzero linear transformation from S into B such that the nullspace $\eta(T)$ has a finite dimension r. Further, let $\{\mathscr{L}^\alpha\}_{\alpha=1}^r$ be a set of r linear functionals on S which are independent on $\eta(T)$.

Example. Let $S = C^rI$, the space of all r-times continuously differentiable functions on $I = [b, c] \subset \mathbf{R}$, $\mathscr{L}^\alpha f = f^{(\alpha-1)}(a)$, $\alpha = 1, 2, \ldots, r$, for some fixed $a \in I$, and $Tf = f^{(r)}$. Then $T(S) = C^0I$ and $\eta(T)$ is the set of all polynomials of degree $\leqslant r - 1$. We can take $B = L^pI$, the space of all Lesbesgue p-th power integrable functions on I, with the usual norm $\|\cdot\|_p$, $1 \leqslant p \leqslant \infty$.

Definition. If $f \in S$, define $\|f\| = \max[\max_\alpha |\mathscr{L}^\alpha f|, \|Tf\|_B]$.

It is easily seen that $\|\cdot\|$ is a norm on S. Indeed, $\|Tf\|_B$ fails, in general, to be a norm on S only because the nullspace of T is not necessarily the zero element alone. The information contained in the numbers $\{\mathscr{L}^\alpha f\}_{\alpha=1}^r$ is, in a sense, the minimum that must be added to obtain a true norm.

Theorem 1. *If $L \in S^*$, the dual of S, then there exist constants $c_1, c_2, \ldots, c_r$ and $\mu \in B^*$ such that $Lf = \sum_{j=1}^r c_j \mathscr{L}^j f + \mu(Tf)$, for all f in S, and*

$$\|L\|_{S^*} = \sum_{j=1}^r |c_j| + \|\mu\|_{B^*}.$$

Proof. If $f \in S_1 = S \cap \eta(\mathscr{L}^1) \cap \cdots \cap \eta(\mathscr{L}^r)$, then $\|f\| = \|Tf\|_B$. So on S_1, $Lf = \mu(Tf)$ for some $\mu \in B^*$ with $\|\mu\|_{T(S_1)^*} = \|\mu\|_{B^*}$ (by using the Hahn–Banach theorem to extend a bounded linear functional from $T(S_1)$ to B while preserving the norm). Now let $e_j \in \eta(T)$, $j = 1, 2, \ldots, r$, so that $\mathscr{L}^i e_j = \delta_{ij}$, $i, j = 1, 2, \ldots, r$. If $f \in S$, then $f - \sum_{i=1}^r (\mathscr{L}^i f)\, e_i \in S_1$, and we have

$$Lf - \sum_{i=1}^r (Le_i)\, \mathscr{L}^i f = L\left[f - \sum_{i=1}^r (\mathscr{L}^i f)\, e_i\right]$$

$$= \mu\left(T\left[f - \sum_{i=1}^r (\mathscr{L}^i f)\, e_i\right]\right) = \mu(Tf).$$

Hence $Lf = \mu(Tf) + \sum_{i=1}^r (Le_i)\, \mathscr{L}^i f = \mu(Tf) + \sum_{i=1}^r c_i \mathscr{L}^i f.$

From this, we have

$$\begin{aligned} | Lf | &\leqslant \| Tf \|_B \| \mu \|_{B^*} + \sum_{i=1}^{r} | c_i | \, | \mathscr{L}^i f | \\ &\leqslant \left(\| \mu \|_{B^*} + \sum_{i=1}^{r} | c_i | \right) \max[\max_{\alpha} | \mathscr{L}^{\alpha} f |, \| Tf \|_B] \\ &= \left(\| \mu \|_{B^*} + \sum_{i=1}^{r} | c_i | \right) \| f \|. \end{aligned}$$

Thus $\| L \|_{S^*} \leqslant \| \mu \|_{B^*} + \sum_{i=1}^{r} | c_i |$.

On the other hand, we can choose an $f \in S_1$ such that $\| Tf \|_B = 1$ and $\mu(Tf) \geqslant \| \mu \|_{B^*} - \epsilon$, since $\| \mu \|_{T(S_1)^*} = \| \mu \|_{B^*}$. Let $f_1 \in \eta(T)$ be such that

$$\mathscr{L}^j f_1 = \overline{\operatorname{sgn} c_j}, \qquad j = 1, 2, \ldots, r.$$

Thus $f + f_1$ is such that

$$\| T(f + f_1) \|_B = 1 \qquad \text{and} \qquad \mathscr{L}^j(f + f_1) = \overline{\operatorname{sgn} c_j}, \quad j = 1, 2, \ldots, r.$$

Hence $\| f + f_1 \| = 1$ and $L(f + f_1) \geqslant \| \mu \|_{B^*} - \epsilon + \sum_{i=1}^{r} | c_i |$. But ϵ is arbitrary. Q.E.D.

If V is a subspace of S, we say that $g \in V$ is a best approximation of an element f of S if

$$\| f - g \| = \inf_{h \in V} \| f - h \| = \rho. \tag{1}$$

It is clear that the set P of best approximations in V of a function in S is convex.

Definition. By the dimension of a convex set P in a finite-dimensional vector space we mean the largest integer k for which there exist points $g_1, g_2, \ldots, g_{k+1}$ in P such that $g_1 - g_{k+1}, g_2 - g_{k+1}, \ldots, g_k - g_{k+1}$ are linearly independent. If P consists of a single point, the dimension of P, dim P, is 0. If P is empty, dim $P = -1$. The maximum dimension of sets $P_V(f)$ of best approximation in V of points f in $S (\supseteq V)$ is called the rank of V. Following [10], we say that V is r-semi-Tchebycheff or r-Tchebycheff if, for all f in S, $-1 \leqslant \dim P_V(f) \leqslant r$ or $0 \leqslant \dim P_V(f) \leqslant r$, respectively. "0-Tchebycheff" is abbreviated to "Tchebycheff."]

In terms of this definition we see that if V is finite-dimensional, then V has rank $\leqslant r$ if and only if V is r-Tchebycheff. In particular, a finite-dimensional subspace V is a space of unique best approximation for any function in S if and only if V is Tchebycheff.

We now state a corollary to Theorem 1 to which we will refer in the sequel.

COROLLARY 1. *For each* $f \in S$, *there exists* $L \in S^*$, $L = \sum_{j=1}^{r} c_j \mathscr{L}^j + \mu T \neq 0$, *such that, whenever* (1) *holds,*

$$c_j \mathscr{L}^j(f-g) = | c_j | \rho, \qquad 1 \leqslant j \leqslant r,$$

$$\mu(T(f-g)) = \| T(f-g)\|_B \| \mu \|_{B^*},$$

and

$$\| T(f-g)\|_B = \rho \qquad \text{if} \quad \mu \neq 0.$$

Proof. From a well-known corollary to the Hahn–Banach Theorem, there exists L in S^* with

$$L(V) = \{0\}, \quad \| L \|_{S^*} = 1, \qquad \text{and} \qquad L(f) = \rho. \tag{2}$$

By Theorem 1, $\| L \|_{S^*} = \sum_{j=1}^{r} | c_j | + \| \mu \|_{B^*}$. Then

$$\begin{aligned} \rho = L(f) = L(f-g) &= \sum_{j=1}^{r} c_j \mathscr{L}^j(f-g) + \mu(T(f-g)) \\ &\leqslant \sum_{j=1}^{r} | c_j | \, | \mathscr{L}^j(f-g)| + \| T(f-g)\|_B \| \mu \|_{B^*} \\ &\leqslant \left(\sum_{j=1}^{r} | c_j | + \| \mu \|_{B^*} \right) \max[\max_\alpha | \mathscr{L}^\alpha(f-g)|, \| T(f-g)\|_B] \\ &= 1 \cdot \| f-g \| = \rho. \end{aligned}$$

Thus, $c_j \mathscr{L}^j(f-g) = | c_j | \rho$, $1 \leqslant j \leqslant r$, $\mu(T(f-g)) = \| T(f-g)\|_B \| \mu \|_{B^*}$, and $\| T(f-g)\|_B = \rho$ if $\mu \neq 0$. Q.E.D.

2. $B = CX$

In this section, let $B = CX$, the space of continuous functions on the compact Hausdorff space X with norm $\| \cdot \|_\infty$.

DEFINITION. A subspace V of S, of dimension n, has property $H(p)$ if, for any linearly independent points $g_1, g_2, \ldots, g_{p+1}$ of V, the number of j's $(1 \leqslant j \leqslant r)$ such that $\mathscr{L}^j g_i = 0$ $(\forall_i)$, plus the number of x's in X such that $(Tg_i)(x) = 0$ $(\forall_i)$, does not exceed $n - p - 1$.

THEOREM 2. *Let V be an n-dimensional subspace of S. For V to be p-Tchebycheff, property $H(p)$ is sufficient. Conversely, if T is onto CX, then property $H(p)$ is necessary for V to be p-Tchebycheff.*

Proof. Consider, first, the case $S = CX$ and $T =$ identity map on CX ($r = 0$). If X is a closed interval, the result is due to Rubenstein [3, p. 94]. If X is arbitrary, the result follows, e.g., by appropriately modifying the proof of Rivlin and Shapiro [8, p. 36] to include the cases $p > 0$.

We will reduce our general case to the above by realizing the normed space S as a subspace of CX', in the case of sufficiency, and as equal to CX', in the case of necessity, for an appropriately chosen compact set X'. Consider the set $X' = \{\mathscr{L}^\alpha\}_{\alpha=1}^r \cup \{\mathscr{L}^x\}_{x\in X}$ of linear functionals on S, where $\mathscr{L}^x f = (Tf)(x)$. Give $\{\mathscr{L}^x\}_{x\in X}$ the topology induced by X in the obvious way and impose on the r elements $\{\mathscr{L}^\alpha\}_{\alpha=1}^r$ the discrete topology. The set X' with the sum topology is compact. It follows that the space S with norm

$$\|f\| = \max[\max_\alpha |\mathscr{L}^\alpha f|, \|Tf\|_{CX}] = \sup[|\mathscr{L}^\alpha f|_{\alpha=1}^r, |(Tf)(x)|_{x\in X}]$$

$$= \sup_{X'} |f(x')|$$

is realized as a subspace of CX', and the sufficiency result follows.

If $TS = CX$, it is clear that the realization of the space S with norm $\|f\| = \sup_{X'} |f(x')|$ is all of CX', and the necessity result follows. Q.E.D.

EXAMPLE 1. Let $B = C^0 I$ with norm $\|\cdot\|_\infty$, and fix a in $\mathbf{R}$. Let S be the set of all functions f defined on $I \cup \{a\}$ such that $f|_{I\setminus\{a\}}$ is the restriction of a function in $C^0 I$. Let $\mathscr{L}^1 f = f(a)$, and let Tf be the unique continuous extension of $f|_{I\setminus\{a\}}$ to I. Let V be the n-dimensional subspace of S consisting of those elements which coincide on $I\setminus\{a\}$ with $(n-2)$-degree polynomials. By Theorem 2, it follows that V is a Tchebycheff subspace.

EXAMPLE 2. Consider the space $C^1[0, 1]$ with norm

$$\|f\| = \max[|f(a)|, \|f^{(1)}\|_\infty], \qquad a \in [0, 1].$$

Then the space of all linear functions of the form $p(x) = cx + c$ is a 1-dimensional Tchebycheff subspace.

A further example is given in the following corollary.

COROLLARY 2. *Consider $C^r[b, c]$ with norm*

$$\|f\| = \max[|f(a)|, |f^{(1)}(a)|, ..., |f^{(r-1)}(a)|, \|f^{(r)}\|_\infty],$$

where $a \in [b, c]$. Let P_n be the subspace of polynomials of degree n or less. Then if $r \leqslant n$, P_n has rank r. Moreover, if p and q are any two best polynomial approximations to f in $C^r[b, c]$, then $p^{(r)} = q^{(r)}$.

Proof. Here S is the space of the example in Section 1, where $B = L^\infty I$ with norm $\| \cdot \|_\infty$. It is easily checked that the $(n + 1)$-dimensional subspace P_n has property $H(r)$ and so, by Theorem 2, has rank not exceeding r.

We see more, however, from (2) and Corollary 1. If $r > n$, the corollary is trivially true; so assume $r \leqslant n$. If $f \in C^r I$ but $f \notin P_n$, then the L of Corollary 1 cannot be of the form $\sum_{j=1}^{r} c_j \mathscr{L}^j$, where $\mathscr{L}^j f = f^{(j-1)}(a)$, $j = 1, 2, ..., r$. This follows from (2) and the obvious fact that no nontrivial linear combination of the $\{\mathscr{L}^j\}_{j=1}^r$ can vanish identically on P_n. Thus μ of Corollary 1 is $\neq 0$, and, so, $\| f^{(r)} - p^{(r)} \|_\infty = \rho = \inf \| f - h \|$. Suppose $q^{(r)} \neq p^{(r)}$, and $\| f^{(r)} - q^{(r)} \|_\infty = \rho$. We can clearly assume, without loss of generality, that $p^{(\alpha)}(a) = f^{(\alpha)}(a) = q^{(\alpha)}(a)$, $\alpha = 0, 1, ..., r - 1$. But this contradicts the uniqueness of best approximation of $f^{(r)}$ by polynomials in P_{n-r} with respect to $\| \cdot \|_\infty$. Note, also, that P_n has rank exactly r if $r \leqslant n$. Q.E.D.

Note. In [5] and [7] it is shown that the conclusions of Corollary 2 remain valid if we consider, instead, the equivalent norm

$$\| h \| = \max[\| h \|_\infty, \| h^{(1)} \|_\infty, ..., \| h^{(r)} \|_\infty]$$

and make the additional assumption that f be $(r + 1)$-times differentiable.

3. B IS STRICTLY CONVEX

In this section, let B be any strictly convex normed vector space (e.g., the "L^p-spaces," $1 < p < \infty$), where "strictly convex" is defined as follows:

DEFINITION. B is strictly convex if $\| f + g \| < 2$ whenever $\| f \| = \| g \| = 1$ and $f \neq g$.

THEOREM 3. *Let V be an n-dimensional subspace of S. For V to be p-Tchebycheff it is sufficient that* $\dim[V \cap \eta(T)] \leqslant p$ *and the* $(\mathscr{L}^j)_{j=1}^r$ *are linearly independent in V^*. Conversely, if $p \leqslant n - r$ and $T(S)$ is infinite-dimensional, the above conditions are necessary for V to be p-Tchebycheff.*

Proof. Since V is finite-dimensional, $\exists$ at least one g in V satisfying (1). Thus, if g_1 and g_2 are two best approximations and $\mu \neq 0$, then $\| T(f - g_i)\|_B = \rho$, $i = 1, 2$, and $\mu(T(f - g_1)) = \mu(T(f - g_2)) = \rho \| \mu \|_{B^*}$, by Corollary 1. Hence, if $\mu \neq 0$, $T(f - g_1) = T(f - g_2)$, since B is strictly convex [9, p. 300], and, so, $Tg_1 = Tg_2$.

Now, if V has rank $> p$, then $\exists p + 2$ distinct best approximations

$g_1, g_2, \ldots, g_{p+2}$ such that $(g_{p+2} - g_i)_{i=1}^{p+1}$ are linearly independent. Suppose the $(\mathscr{L}^j)_{j=1}^r$ are linearly independent in V^*. Then it follows from Corollary 1 and (2) that $\mu \neq 0$ and, so, from the above we have $Tg_1 = Tg_2 = \cdots = Tg_{p+2}$. Hence $\dim[V \cap \eta(T)] \geqslant p + 1$.

For the converse, suppose, first, that the $(\mathscr{L}^j)_{j=1}^r$ are linearly dependent on V. Then $\exists p + 1$ linearly independent elements $g_1, g_2, \ldots, g_{p+1}$ in V such that $\mathscr{L}^j g_i = 0$ $(\forall(i,j))$. We may assume that $\sum_{i=1}^{p+1} \| Tg_i \|_B \leqslant 1$. Then, if $\sum_{j=1}^r b_j \mathscr{L}^j = 0$ in V^*, where $\sum_{j=1}^r | b_j | = 1$, choose $f \in \eta(T)$ such that

$$\mathscr{L}^j f = \overline{\operatorname{sgn} b_j}, \qquad j = 1, 2, \ldots, r.$$

If $L = \sum_{j=1}^r b_j \mathscr{L}^j$, then $\| L \|_{S^*} = 1$, and for any $\tilde{g} \in V$, $\| f - \tilde{g} \| \geqslant | L(f - \tilde{g})| = | Lf | = \sum_{j=1}^r | b_j | = 1$. On the other hand, for any $\epsilon_1, \epsilon_2, \ldots, \epsilon_{p+1}(| \epsilon_i | \leqslant 1)$, $| \mathscr{L}^j(f - \sum_{i=1}^{p+1} \epsilon_i g_i)| = | \mathscr{L}^j f | \leqslant 1$ $(1 \leqslant j \leqslant r)$ and $\| T(f - \sum_{i=1}^{p+1} \epsilon_i g_i \|_B \leqslant \sum_{i=1}^{p+1} \| Tg_i \|_B \leqslant 1$. Thus $\{\sum_{i=1}^{p+1} \epsilon_i g_i ; | \epsilon_i | \leqslant 1\}$ forms a $(p + 1)$-dimensional set of best approximations in V to f, and, so, the rank of $V > p$.

Suppose now that $\dim[V \cap \eta(T)] > p$. That is, $\exists p + 1$ linearly independent elements $g_1, g_2, \ldots, g_{p+1}$ in $V \cap \eta(T)$. Let $F = \{j \mid \mathscr{L}^j g_i = 0 \ (\forall_i)\}$. Let k be the number of elements of F. We may also assume $| \mathscr{L}^j g_i | \leqslant 1$ $(\forall(i,j))$. Choose $\{M^j\}_{j=1}^{n-p-k}$ to be linearly independent elements in the dual of some finite-dimensional space $T(W)$ containing $T(V)$. Then the set of linear functionals $\{\mathscr{L}^j\}_{j \in F} \cup \{M^j T\}_{j=1}^{n-p-k}$ is linearly dependent in the dual V^* of V. If not, they would span an $(n - p)$-dimensional subspace of V^* and, thus, there would be a nontrivial linear combination $\sum_{i=1}^{p+1} d_i g_i = g_* \in V^{**} = V$ such that $\mathscr{L}^j g_* = 0$ $(j \in F)$ and $M^j Tg_* = 0$ $(1 \leqslant j \leqslant n - p - k)$ is not true, a contradiction. Thus, $\exists$ scalars b_j, $j \in F$, and c_j $(1 \leqslant j \leqslant n - p - k)$, not all zero, such that $L' = \sum_{j \in F} b_j \mathscr{L}^j + \sum_{j=1}^{n-p-k} c_j M_j T$ is zero on V. If all $c_j = 0$ then the $\mathscr{L}^j$'s are linearly dependent in V^*, which case has already been treated above. Hence we assume that not all $c_j = 0$, which implies that $M_W = \sum_{j=1}^{n-p-k} c_j M_j$, as an element of $T(W)^*$, is $\neq 0$. Let M be a norm-preserving extension of this functional from $T(W)$ to B. Let $a = \| M \|$. Assume, without loss of generality, that $\sum_{j \in F} | b_j | + a = 1$. It follows from Theorem 1 that if $L = \sum_{j \in F} b_j \mathscr{L}^j + MT$, then $\| L \|_{S^*} = 1$. Let h be an element in $T(W)$ with $\| h \|_B \leqslant 1$ and $M_W(h) = a$. (This is possible, since $T(W)$ is finite-dimensional and, therefore, reflexive.) Pick $f' \in T^{-1}\{h\}$. Next, pick $f_1 \in \eta(T)$ such that

$$\mathscr{L}^j f_1 = \overline{\operatorname{sgn} b_j} - \mathscr{L}^j f' \qquad (j \in F),$$

while $\mathscr{L}^j f_1 = -\mathscr{L}^j f'$ $(j \notin F, 1 \leqslant j \leqslant r)$. Let $f = f' + f_1$. Then

$$\mathscr{L}^j f = \overline{\operatorname{sgn} b_j} \quad (j \in F), \qquad \| Tf \|_B = \| Tf' \|_B = \| h \|_B \leqslant 1 \qquad \text{and} \qquad \| f \| \leqslant 1.$$

Furthermore, $M(Tf) = M(h) = M_W(h) = a$. Then, for any $\tilde{g} \in V$, $\|f - \tilde{g}\| \geqslant |L(f - \tilde{g})| = |Lf - L\tilde{g}| = |Lf| = |\sum_{j \in F} b_j \mathscr{L}^j f + M(Tf)| = \sum_{j \in F} |b_j| + a = 1$. On the other hand, for any $\epsilon_1, \epsilon_2, ..., \epsilon_{p+1}$

$$\left(|\epsilon_i| \leqslant \frac{1}{p+1}\right), \qquad \left|\mathscr{L}^j\left(f - \sum_{i=1}^{p+1} \epsilon_i g_i\right)\right| = |\mathscr{L}^j f| \qquad \text{if } j \in F$$

and

$$\left|\mathscr{L}^j\left(f - \sum_{i=1}^{p+1} \epsilon_i g_i\right)\right| = \left|\sum_{i=1}^{p+1} \epsilon_i \mathscr{L}^j g_i\right| \qquad \text{if } j \notin F, \quad 1 \leqslant j \leqslant r.$$

So $|\mathscr{L}^j(f - \sum_{i=1}^{p+1} \epsilon_i g_i)| \leqslant 1$, $1 \leqslant j \leqslant r$. Also $\| T(f - \sum_{i=1}^{p+1} \epsilon_i g_i)\|_B = \| Tf \|_B \leqslant 1$. Thus $\{\sum_{i=1}^{p+1} \epsilon_i g_i ; 0 \leqslant |\epsilon_i| \leqslant (p+1)^{-1}\}$ forms a $(p+1)$-dimensional set of best approximations in V of f and, so, the rank of $V > p$. Q.E.D.

Example. Consider the example in Section 1, with $B = L^p I$, $1 < p < \infty$, and $a \neq 0$. Then the n-dimensional subspace of all polynomials of the form $a_r x^{r-s} + a_{r+1} x^{r-s+1} + \cdots + a_{r+n-1} x^{r-s+n-1}$ has rank not exceeding max$[0, s]$ in S and, in particular, is a Tchebycheff subspace if s is a nonpositive integer.

Note. The strict convexity of B is not needed in the proof of necessity in Theorem 3.

4. $B = L^1 X$

In this section, let $B = L^1 I$, the space of all Lebesgue-integrable functions h on $I = [b, c] \subset \mathbf{R}$, with $\| h \|_1 = \int |h| = \int_I |h| \, dx$, where dx denotes ordinary Lesbesgue measure. We also assume in this section that $T(S) \subseteq C_{\mathbf{R}} I$, the space of real-valued continuous functions on I.

For a proof of the following lemma see, e.g., [1, p. 219].

Lemma. *Let f and h be elements of $C_{\mathbf{R}} I$. If f has at most a finite number of zeros and if $\int h \operatorname{sgn} f \neq 0$, then for some λ, $\int |f - \lambda h| < \int |f|$.*

In the case $r = 0$ and T is the identity map, the following theorem reduces to the well-known Jackson theorem and the proof reduces to that given in [1, p. 219].

Theorem 4. *Let V be an n-dimensional subspace of S, on which the $(\mathscr{L}^j)_{j=1}^r$ are linearly independent. Then V is Tchebycheff if V has property $H(0)$.*

Proof. Since V is finite-dimensional, $\exists$ at least one g in V, satisfying (1). By Corollary 1, we obtain

$$c_j \mathscr{L}^j(f-g) = \rho \mid c_j \mid, \qquad 1 \leqslant j \leqslant r, \tag{3}$$

and

$$\int_I \mid T(f-g) \mid dx = \rho, \tag{4}$$

since $\mu \neq 0$ by (2), and the $(\mathscr{L}^j)_{j=1}^r$ are linearly independent on V.

Suppose g_1 and g_2 are best approximations of f in V. Let $F = \{j \mid c_j \neq 0\}$. Then (3) shows that there exist constants a_j $(j \in F)$ such that $\mathscr{L}^j g_i = a_j$ $(i = 1, 2)$. Let k be the number of elements in F. Now, since the set of best approximations is convex, $g = \frac{1}{2}(g_1 + g_2)$ is also a best approximation. Hence, by (4), $\int_I (\mid T(f-g)\mid - \frac{1}{2} \mid T(f-g_1)\mid - \frac{1}{2} \mid T(f-g_2)\mid) \, dx = 0$, and, since the integrand is nonpositive and continuous, it is identically zero. Suppose $T(f-g)$ has $n-k$ zeros. Then $T(f-g_1)$ and $T(f-g_2)$ have these $n-k$ zeros in common and, thus, $T(g_1 - g_2)$ has $n-k$ zeros. Since also $\mathscr{L}^j(g_1 - g_2) = 0$ $(j \in F)$ and V has property $H(0)$, we conclude that $g_1 = g_2$.

Assume now that $T(f-g)$ has fewer than $n-k$ zeros. Choose points $b = x_0 < x_1 < \cdots < x_{n-k} = c$, including the zeros of $T(f-g)$. Consider $\int h \operatorname{sgn}(T(f-g)) = \sum_{i=1}^{n-k} b_i \varphi_i(h)$, where, for every i, $\mid b_i \mid = 1$ and $\varphi_i(h) = \int_{x_{i-1}}^{x_i} h$. Then $L = \sum_{i=1}^{n-k} b_i \varphi_i T$ vanishes on $V_1 = V \cap (\bigcap_{j \in F} \eta(\mathscr{L}^j))$. For otherwise, by the lemma, for some $h \in V_1$ and λ, $\int \mid T(f-g-\lambda h)\mid = \int \mid T(f-g) - \lambda Th \mid < \int \mid T(f-g)\mid$. Hence, also, for some $\epsilon > 0$, $\int \mid T(f-g-\lambda h - h')\mid < \int \mid T(f-g)\mid$ if $h' \in V$ and $\| h' \| < \epsilon$. Furthermore, since the $(\mathscr{L}^j)_{j \in F}$ are linearly independent on V, we could pick an $h' \in V$ such that $\| h' \| < \epsilon$ and $\mid \mathscr{L}^j(f-g-\lambda h - h')\mid < \rho, j \in F$. Then $g + \lambda h + h'$ would be a better approximation of f than g, which is impossible. There exists, therefore, a nonzero $p \in V_1$ such that $\varphi_i(Tp) = 0$, $i = 1, 2,..., n-k$. But $0 = \varphi_i(Tp) = \int_{x_{i-1}}^{x_i} (Tp)(x) \, dx$ implies that Tp has at least one zero in (x_{i-1}, x_i), $i = 1, 2,..., n-k$. Thus, by property $H(0)$, $p = 0$, a contradiction.

Q.E.D.

Examples. Examples 1 and 2 of Section 2 remain valid if we replace $C^r I$ by $C_{\mathbf{R}}{}^r I$ and $\| \cdot \|_\infty$ by $\| \cdot \|_1$ in Example 2, and let $B = L^1 I$ in Example 1.

5. Remark on Another Norm for S

In the case $B = L^p X$, $1 \leqslant p < \infty$, consider on S the new norm $\|f\|_* = (\sum_\alpha \mid \mathscr{L}^\alpha f \mid^p + \| Tf \|_B^p)^{1/p}$. Then, analogously to the proof of Theorem 2, S can be realized as a subspace of $L^p X^*$, where X^* is the space described in the proof of Theorem 2. Thus, in particular, if $1 < p < \infty$, every closed subspace of S is a Tchebycheff subspace. If $p = 1$, then property $H(0)$ is sufficient for the n-dimensional space V to be Tchebycheff.

References

1. E. W. CHENEY, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
2. P. J. DAVIS, "Interpolation and Approximation," Ginn Blaisdell, Waltham, Mass. 1963.
3. A. L. GARKAVI, On the dimensionality of polytopes of best approximation for differentiable functions, *Izv. Akad. Nauk. SSSR Ser. Mat.* **23** (1959), 93–114 (in Russian).
4. YASUHIKO, IKEBE, A characterization of Haar subspaces in $C[a, b]$, *Proc. Japan Acad.* **44** (1968), 219–220.
5. LEE JOHNSON, Unicity in approximation of a function and its derivatives, *Math. Comp.* **22** (1968), 873–875.
6. GÜNTER MEINARDUS, "Approximation of Functions: Theory and Numerical Methods," (English transl.), Springer-Verlag, New York, 1967.
7. D. G. MOURSUND, Chebyshev approximation of a function and its derivatives, *Math. Comp.* **18** (1964), 382–389.
8. T. J. RIVLIN AND H. S. SHAPIRO, Some uniqueness problems in approximation theory, *Comm. Pure Appl. Math.* **13** (1960), 35–47.
9. W. W. ROGOSINSKI AND H. S. SHAPIRO, On certain extremum problems for analytic functions, *Acta Math.* **90**, 287–318.
10. IVAN SINGER, "Best Approximation in Normed Vector Spaces by Elements of Vector Subspaces," (in Rumanian), Editura Academiei Republicii Socialiste România, Bucharest, 1967.