JOURNAL OF APPROXIMATION THEORY 13, 375–391 (1975)

On an Approximation Operator of de La Vallée Poussin*

P. D. MORRIS

Department of Mathematics, Pennsylvania State University, University Park, Pennsylvania 16802

K. H. PRICE

Department of Mathematics, Stephen F. Austin College, Nacogdoches, Texas 75961

AND

E. W. CHENEY

Department of Mathematics, University of Texas, Austin, Texas 78712

Communicated by P. L. Butzer

DEDICATED TO PROFESSOR G. G. LORENTZ ON THE OCCASION OF HIS SIXTY-FIFTH BIRTHDAY

1. INTRODUCTION

In 1910, de La Vallée Poussin published in [2] some researches on the following approximation problem: given the continuous function $x(t)$ defined on the interval $-1 \leqslant t \leqslant 1$ and given $n + 1$ distinct points $t_0, \ldots, t_n$ in $[-1, 1]$, it is required to determine that polynomial $p(t)$ of degree $<n$ which best approximates $x(t)$ on the discrete point set $\{t_0, \ldots, t_n\}$. The polynomial p which is sought is therefore expected to produce a minimum value of the following expression:

$$\max_{0 \leqslant i \leqslant n} | x(t_i) - p(t_i)| . \tag{1}$$

If the polynomial p were permitted to be of degree n, then (1) could be made zero by taking p to be the Lagrange interpolating polynomial.

* This research has been supported by the Air Force Office of Scientific Research. Dr. Price was supported by a Faculty Research Grant from Stephen F. Austin State University.

The problem of minimizing (1) did not arise out of unmotivated curiosity. Indeed, the solution of this problem is a crucial step in the systematic determination of the polynomial of best approximation on a continuum. One of de La Vallée Poussin's own theorems states, in fact, that the polynomial of degree $<n$ which best approximates the given function on the interval $[-1, 1]$ is the polynomial of best approximation on an appropriate subset of $[-1, 1]$ consisting of $n + 1$ points.

As is well known, the determination of the polynomial of best approximation on an interval is a nonlinear problem. By this we mean that the operator M which assigns to x its best approximation p is a nonlinear operator on $C[-1, 1]$. In contrast to this, the polynomial which minimizes the expression (1) depends *linearly* on x. In this important characteristic the process resembles Lagrange interpolation: both processes define linear projection operators.

In this essay, we study the general operators of which de La Vallée Poussin's is the prototype. Thus we consider projections onto n-dimensional spaces which utilize $n + 1$ quanta of information concerning the projected element. The information is delivered in the form of values of $n + 1$ linear functionals. We examine families of such projections, and seek out those which are optimal in a certain sense. There exist cases in which de La Vallée Poussin's operator possesses this optimality property.

Some of the preliminary discussion is purely algebraic, and this part is presented first in a general setting. The extremum problems are, however, couched in a space of continuous functions, and our concrete examples all concern the approximating family of algebraic polynomials.

A few conventions about notation and terminology are as follows. The *composition* of two maps is written $A \circ B$. If X is a linear space, if $z_i \in X$ and if $f_i \in X^*$, then the *tensor* notation $L = \sum_{i=1}^{m} f_i \otimes z_i$ denotes a linear operator $L: X \to X$ whose value at x is $Lx = \sum_{i=1}^{m} f_i(x)\, z_i$. A set of functionals $f_1, f_2, \ldots$ defined on a linear space X is *total* over a subspace Y if 0 is the only element of Y having the property $f_1(y) = f_2(y) = \cdots = 0$. Finally, if $Y \subset X$ and $f \in X^*$ then the symbolism $f \perp Y$ signifies that $f(y) = 0$ for all $y \in Y$.

2. Algebraic Properties of Certain Standard Operators

Let X be a normed linear space, and let Y be an n-dimensional subspace in X. Suppose further that a set of $n + 1$ continuous linear functionals on X has been prescribed:

$$\{f_0, f_1, \ldots, f_n\} \subset X^* \tag{2}$$

The problem of de La Vallée Poussin mentioned above has the following interpretation. Using the given set of linear functionals, one defines a semi-norm on X via the equation

$$\Delta(x) = \max_{0 \leqslant i \leqslant n} |f_i(x)| \qquad (x \in X). \tag{3}$$

This leads to a natural approximation problem involving X, Y, and Δ. Namely, for each $x \in X$, one seeks those elements $y \in Y$ for which the deviation $\Delta(x - y)$ is a minimum. In de La Vallée Poussin's original problem, of course, the f_i are "point functionals" defined by $f_i(x) = x(t_i)$, and Y is the subspace of algebraic polynomials of degree $< n$.

In order to give a succinct solution to de La Vallée Poussin's problem, it is convenient to introduce a functional Φ having the following form and properties:

$$\Phi = \sum_{i=0}^{n} \theta_i f_i , \qquad \Phi \perp Y, \qquad \sum_{i=0}^{n} | \theta_i | = 1. \tag{4}$$

The existence of Φ follows from the observation that Y^* is of dimension n, and accordingly, the restricted functionals $f_i \mid Y$ form a linearly dependent set.

THEOREM 1. *Let the set of functionals* (2) *be total over* Y. *Then there exist* $q_0 ,..., q_n$ *in* Y *such that* $f_i(q_j) = \delta_{ij} - \theta_j \operatorname{sgn} \theta_i$. *The operator* $A = \sum_{i=0}^{n} f_i \otimes q_i$ *is a linear projection of* X *onto* Y *and solves de La Vallée Poussin's problem*: $\Delta(x - Ax) \leqslant \Delta(x - y)$ *for all* $x \in X$ *and* $y \in Y$.

Proof. Since $\{f_0 ,..., f_n\}$ is total, a subset of n elements spans Y^*; say $[f_1 ,..., f_n] = Y^*$. Select $q_0 ,..., q_n \in Y$ so that $f_i(q_j) = \delta_{ij} - \theta_j \operatorname{sgn} \theta_i$ $(1 \leqslant i \leqslant n,\ 0 \leqslant j \leqslant n)$. One can verify that this equation is valid then for $i = 0$ also. The verification depends upon calculating $0 = \Phi(q_j) = \theta_0 f_0(q_j) + \theta_j(| \theta_0 | - \delta_{j0})$ and noting that $\theta_0 \neq 0$. (If $\theta_0 = 0$ then $\Phi = \sum_{i=1}^{n} \theta_i f_i \perp Y$.)

In order to see that A acts like the identity map on Y, it is enough to prove that $f_i \circ (I - A) = \operatorname{sgn} \theta_i \Phi$, because the functionals f_i are total over Y. This equation is easily established by a calculation. Using this equation we complete the proof by writing for any $x \in X$ and $y \in Y$,

$$\begin{aligned} \Delta(x - Ax) &= \max_i |f_i \circ (I - A)\, x | \leqslant | \Phi(x)| \\ &= | \Phi(x - y)| = \Big| \sum \theta_i f_i(x - y) \Big| \\ &\leqslant \max |f_i(x - y)| \sum_j | \theta_j | = \Delta(x - y). \quad \blacksquare \end{aligned}$$

Another important type of projection from X onto Y is the analog of the classical Lagrange interpolator. Select an index i and suppose that the set

$\{f_0, ..., f_{i-1}, f_{i+1}, ..., f_n\}$ is total over Y. (Equivalent to the assumption that $\theta_i \neq 0$.) Then there exist $w_{ij} \in Y$ such that $f_\nu(w_{ij}) = \delta_{\nu j}$ for $0 \leqslant j \leqslant n, j \neq i$, $\nu \neq i$, $0 \leqslant \nu \leqslant n$. Put $w_{ii} = 0$ for convenience and define

$$L_i = \sum_{j=0}^{n} f_j \otimes w_{ij} . \tag{5}$$

One verifies easily the interpolation property $f_j \circ L_i = f_j$ $(i \neq j)$ and the following useful equations:

$$f_i(w_{ij}) = -\theta_j \theta_i^{-1} \qquad (i \neq j,\ \theta_i \neq 0), \tag{6}$$

$$f_i \circ (I - L_i) = \theta_i^{-1} \Phi \qquad (\text{if } \theta_i \neq 0), \tag{7}$$

$$L_j = L_i - \Phi \otimes w_{ij}/\theta_j \qquad (\text{if } \theta_i \theta_j \neq 0). \tag{8}$$

THEOREM 2. *The operator A is a convex linear combination of the operators L_i (only those that exist!). In fact,*

$$A = \sum_{\theta_i \neq 0} | \theta_i | L_i .$$

Consequently, $\| A \| \leqslant \max_{\theta_i \neq 0} \| L_i \|$.

Proof. Let $B = \sum_{\theta_i \neq 0} | \theta_i | L_i$. In order to show that $A = B$ it is enough to prove that $f_i \circ A = f_i \circ B$ for $i = 0, ..., n$. As in the proof of Theorem 1, $f_i \circ A = f_i - (\operatorname{sgn} \theta_i) \Phi$.

$$\begin{aligned} f_i \circ B &= \sum_{\theta_j \neq 0} | \theta_j | f_i \circ L_j = | \theta_i | f_i \circ L_i + \sum_{\substack{\theta_j \neq 0 \\ j \neq i}} | \theta_j | f_i \\ &= | \theta_i | [-\theta_i^{-1} \Phi + f_i] + f_i[1 - | \theta_i |] \\ &= f_i - (\operatorname{sgn} \theta_i) \Phi. \end{aligned}$$

The above calculation must be slightly modified if $\theta_i = 0$. ∎

THEOREM 3. *If P is any projection of the form $\sum_{i=0}^{n} f_i \otimes y_i$ from X onto Y, then P and A are related by the equation*

$$A = P - \Phi \otimes Pq, \qquad q \in X, \qquad f_i(q) = \operatorname{sgn} \theta_i .$$

Proof. The operator $P - A$ must be of the form $\Phi \otimes u$ for some $u \in Y$ (Lemma 2 of [4]). Since $(P - A)\, q = \Phi(q)\, u = u$, it is now only necessary to prove that $Aq = 0$. It suffices to prove $f_i(Aq) = 0$ for each i. By Theorem 1,

$$f_i A q = \sum_j f_j(q) f_i(q_j) = \sum_j (\operatorname{sgn} \theta_j)[\delta_{ij} - \theta_j \operatorname{sgn} \theta_i] = 0. \quad ∎ \tag{10}$$

THEOREM 4. *If all the θ_i are different from zero, then each projection of the form $P = \sum_{i=0}^{n} f_i \otimes y_i$ of X onto Y can be written as an affine linear combination of the interpolating projections L_i; explicitly, $P = \sum_{i=0}^{n} [1 - f_i(y_i)] L_i$.*

Proof. By Theorem 3 of [6], P is an affine linear combination of $L_0, ..., L_n$. Let

$$P = \sum_{i=0}^{n} \lambda_i L_i \qquad \sum_{i=0}^{n} \lambda_i = 1.$$

For each i, let x_i be an element of X such that $f_j(x_i) = \delta_{ij}$ $(0 \leqslant j \leqslant n)$. Then $L_i x_i = 0$. Consequently,

$$f_i(y_i) = f_i(Px_i) = f_i\left(\sum_j \lambda_j L_j x_i\right) = \sum_{j\neq i} \lambda_j f_i L_j x_i = \sum_{j\neq i} \lambda_j f_i x_i$$

$$= \sum_{j\neq i} \lambda_j = 1 - \lambda_i . \quad \blacksquare$$

3. THE SPACE OF CONTINUOUS FUNCTIONS

The considerations of Section II are now specialized to the space $X = C(T)$ of all continuous real-valued functions defined on a compact Hausdorff space T. In X the *norm* is given by $\| x \| = \sup_{t\in T} | x(t) |$. For each $t \in T$, a *point functional* $\hat{t}$ is defined on X by the equation $\hat{t}(x) = x(t)$. Any operator $L: X \to X$ which can be written in the form $L = \sum_{i=1}^{m} \hat{t}_i \otimes x_i$ is said to be *carried* by the point set $\{t_1, ..., t_m\}$. The carrier of an arbitrary operator L is the smallest closed set $S \subset T$ with the property that $Lx = 0$ for all x satisfying $x \mid S = 0$. For a subset S of T it is convenient to define the semi-norm $\| x \|_S = \sup\{| x(s)| : s \in S\}$.

LEMMA 1. *Let L be a linear operator from $C(T)$ to $C(T)$, and let S be its carrier. Then*

$$\| L \| = \sup\{\| Lx \| : \| x \|_S \leqslant 1\} \geqslant \sup\{\| x \| : Lx = x, \| x \|_S \leqslant 1\} \tag{11}$$

If L interpolates on S (i.e., $Lx \mid S = x \mid S$ *for all* x) *then equality occurs in* (11).

Proof. For each $x \in C(T)$ let x' denote another element such that $x \mid S = x' \mid S$ and $\| x \|_S = \| x' \|$. Such an x' is given by the Tietze Extension Theorem. Then

$$\| L \| = \sup \frac{\| Lx \|}{\| x \|} \leqslant \sup \frac{\| Lx \|}{\| x \|_S} = \sup \frac{\| Lx' \|}{\| x' \|} \leqslant \| L \| .$$

Now if L interpolates on S, and if $\| w\|_S \leqslant 1$ then $Lw \mid S = w \mid S$, $L^2 w = Lw$, and $\| Lw \|_S = \| w \|_S \leqslant 1$. Writing $k = \sup\{\| x \| : Lx = x, \| x \|_S \leqslant 1\}$ we have $\| Lw \| \leqslant k$, since Lw is a possible choice for x in the definition of k. Observe that we proved more, namely $\| Lx \| \leqslant k \| x \|_S$ for all x. ∎

In the remainder of this section, we shall assume that an n-dimensional subspace Y has been fixed in X, and that a set of $n + 1$ points $t_0, ..., t_n$ has been fixed in T. The set of functionals $\{\hat{t}_0, ..., \hat{t}_n\}$ is assumed to be total on Y. Our objective (Theorem 5) is a characterization of the minimal projections from X onto Y carried by $\{t_0, ..., t_n\}$.

As in Section 2, Φ is a functional such that $\Phi = \sum_{i=0}^n \theta_i \hat{t}_i$, $\Phi \perp Y$, and $\sum | \theta_i | = 1$. A bonus is the fact that $\| \Phi \| = 1$.

Now let P be a projection of X onto Y carried by $\{t_0, ..., t_n\}$. Then there exist $y_i \in Y$ such that $P = \sum_{i=0}^n \hat{t}_i \otimes y_i$. It is elementary to prove that

$$\| P \| = \| \sum | y_i | \| = \| \Lambda \|,$$

where Λ is the Lebesgue function of P, $\Lambda = \sum | y_i |$. Two further functions associated with P are

$$v(t) = \sum_{i=0}^n \theta_i \operatorname{sgn} y_i(t),$$

$$u(t) = \sum \{| \theta_i | : y_i(t) = 0\} \qquad (= 0 \text{ if all } y_i(t) \neq 0).$$

For any $x \in X$, the *critical* set of x is the set $\operatorname{crit}(x) = \{t : t \in T$ and $| x(t)| = \| x \|\}$.

LEMMA 2. *If* $s \in T$, $z \in X$, $\| z \| \leqslant 1$, $\hat{s}Pz = \| \hat{s} \circ P \|$, *and* $| v(s)| > u(s)$, *then* $\Phi(z)\, v(s) > 0$.

Proof. Since $\hat{s}Pz = \sum z(t_i)\, y_i(s) = \sum | y_i(s)|$, we must have $z(t_i) = \operatorname{sgn} y_i(s)$ whenever $y_i(s) \neq 0$. A quick calculation then shows that $\Phi(z)$ lies between $v(s) - u(s)$ and $v(s) + u(s)$. Since $| v(s)| > u(s)$, $\operatorname{sgn}[v(s) \pm u(s)] = \operatorname{sgn} v(s) = \operatorname{sgn} \Phi(z) \neq 0$. ∎

DEFINITION. Put

$$E = \sup\{\| y \| : y \in Y \text{ and } | y(t_i)| \leqslant 1 \text{ for } 0 \leqslant i \leqslant n\}.$$

By Lemma 1, $\| P \| \geqslant E$ for all projections P from X onto Y carried by $\{t_0, ..., t_n\}$.

LEMMA 3. *If* $\| P \| > E$, *then* $| v | > u$ *on* $\operatorname{crit}(\Lambda)$.

Proof. Suppose that for some $s \in \text{crit}(\Delta)$, $|v(s)| \leqslant u(s)$. Select x_1 and x_2 of norm 1 such that $x_1(t_i) = x_2(t_i) = \text{sgn}\, y_i(s)$ if $y_i(s) \neq 0$, and $x_1(t_i) = -x_2(t_i) = \text{sgn}\, \theta_i$ if $y_i(s) = 0$. Then $\hat{s}Px_1 = \hat{s}Px_2 = \| P \|$. A quick calculation shows that $\Phi(x_1) = v(s) + u(s)$ and $\Phi(x_2) = v(s) - u(s)$. Since $|v(s)| \leqslant u(s)$, there exists $\lambda \in [0, 1]$ such that $\lambda\Phi(x_1) + (1 - \lambda)\,\Phi(x_2) = 0$. Let $w = \lambda x_1 + (1 - \lambda)\, x_2$ so that $\Phi(w) = 0$ and $\hat{s}Pw = \| P \|$. By Theorem 3, $Pw = Aw$. By Theorem 1 and its proof, $\Delta(w - Aw) \leqslant |\Phi(w)| = 0$. Thus w, Aw, and Pw all agree on the set $\{t_0, ..., t_n\}$. Consequently, $|\hat{t}_i Pw| = |w(t_i)| \leqslant 1$, and by the definition of E, $\| Pw \| \leqslant E$. Hence $\| P \| = \hat{s}Pw \leqslant \| Pw \| \leqslant E$, a contradiction. ∎

THEOREM 5. *In order that P (as above) be minimal in the class of all projections from X onto Y carried by $\{t_0, ..., t_n\}$ it is necessary and sufficient that no $y \in Y$ have the property that v sgn $y > u$ on* crit(Δ).

Proof. Suppose that P is not minimal in the class described. By Theorem 1 of [4] there exists $y \in Y$ such that for all $s \in S = \text{crit}(\Delta)$,

$$\sum_{i=0}^{n} \{|y_i(s) - \theta_i y(s)| - |y_i(s)|\} < 0. \tag{12}$$

This implies (but is not equivalent to) the inequality

$$\sum_{y_i(s)\neq 0} \{[y_i(s) - \theta_i y(s)]\, \text{sgn}\, y_i(s) - y_i(s)\, \text{sgn}\, y_i(s)\} + \sum_{y_i(s)=0} |\theta_i y(s)| < 0. \tag{13}$$

An equivalent formulation of (13) is simply $-yv + |y|\, u < 0$ on S. From this it is clear that y has no root on S. Hence one can divide by $|y|$ to obtain $v\, \text{sgn}\, y > u$ on S.

For the converse, assume the existence of $y \in Y$ such that $\| y \| = 1$ and $v\, \text{sgn}\, y > u$ on S. For each $(n + 1)$-tuple $(\sigma_0, ..., \sigma_n)$ satisfying $|\sigma_i| = 1$ (for all i), select an element $z \in X$ such that $\| z \| = 1$ and $z(t_i) = \sigma_i$ (for all i). Denote by Z the set of functions z so selected. Observe that Z is finite. Observe also that for any operator $L: X \to Y$ carried by $\{t_0, ..., t_n\}$ we have $\| L \| = \max\{\| Lz \|: z \in Z\}$.

Since $|v| \geqslant v\, \text{sgn}\, y > u$ on S, we have (by Lemma 2) $\Phi(z)\, v(s) > 0$ and $\Phi(z)\, y(s) > 0$ whenever $z \in Z$ and $\hat{s}Pz = \| P \|$. By the compactness of crit(Pz) and the fact that Z is finite, there exists a positive number δ such that $\Phi(z)\, y(s) > \delta$ whenever $z \in Z$ and $\hat{s}Pz = \| P \|$.

Assertion. To each $z \in Z$ there corresponds a positive number $\mu(z)$ with the property that $\| Pz - \lambda\Phi(z)\, y \| < \| P \|$ whenever $0 < \lambda < \mu(z)$. In order to prove this, fix $z \in Z$ and consider first the case when $\| Pz \| < \| P \|$. In this case select $\mu(z)$ so that $0 < \mu(z) < \| P \| - \| Pz \|$. Then $\| Pz - \lambda\Phi(z)\, y \| \leqslant \| Pz \| + \lambda \| \Phi \| \| y \| \leqslant \| Pz \| + \mu(z) < \| P \|$.

Now consider the case $\| Pz \| = \| P \|$. If $\hat{s}Pz = \| P \|$, then by the definition of δ, $\Phi(z)\, y(s) > \delta$. If $\hat{s}Pz = -\| P \|$ then, because $\hat{s}P(-z) = \| P \|$, we have $-\Phi(z)\, y(s) > \delta$. Define the open sets

$$U_1 = \{t \in T\colon \hat{t}Pz > \tfrac{1}{2} \| P \| \text{ and } \Phi(z)\, y(t) > \delta\},$$

$$U_2 = \{t \in T\colon -\hat{t}Pz > \tfrac{1}{2} \| P \| \text{ and } -\Phi(z)\, y(t) > \delta\}.$$

By the preceding remarks, $\operatorname{crit}(Pz) \subset U_1 \cup U_2$. Consequently the set $K = T\backslash(U_1 \cup U_2)$ is a compact set containing no points of $\operatorname{crit}(Pz)$. Hence there is a number α such that $| \hat{t}Pz | < \alpha < \| P \|$ for all $t \in K$. Let $\mu(z)$ be the smaller of $\| P \| - \alpha$ and $\frac{1}{2} \| P \|$. Let $0 < \lambda < \mu(z)$. If $t \in U_1$ then

$$| \lambda\Phi(z)\, y(t)| \leqslant \lambda \leqslant \tfrac{1}{2} \|P\| < \hat{t}Pz$$

$$| \hat{t}Pz - \lambda\Phi(z)\, y(t)| = \hat{t}Pz - \lambda\Phi(z)\, y(t) \leqslant \| P \| - \lambda\delta.$$

If $t \in U_2$ then

$$| \lambda\Phi(z)\, y(t)| \leqslant \tfrac{1}{2} \| P \| < -\hat{t}Pz$$

$$| \hat{t}Pz - \lambda\Phi(z)\, y(t)| = -\hat{t}Pz + \lambda\Phi(z)\, y(t) \leqslant \| P \| - \lambda\delta.$$

If $t \in K$ then

$$| \hat{t}Pz - \lambda\Phi(z)\, y(t)| < \alpha + \lambda < \| P \|.$$

This analysis shows that $\| Pz - \lambda\Phi(z)\, y \| < \| P \|$ as asserted. To complete the proof, let $\lambda = \min\{\mu(z)\colon z \in Z\}$. By the Assertion, $\| Pz - \lambda\Phi(z)\, y \| < \| P \|$ for all $z \in Z$. Since the projection $P - \lambda\Phi \otimes y$ is carried by $\{t_0 ,..., t_n\}$, this implies that $\| P - \lambda\Phi \otimes y \| < \| P \|$. ∎

4. Haar Subspaces on an Interval

We maintain the setting of the previous section, but assume further that T is an interval $[a, b]$, and that Y is a Haar subspace. Recall that the Haar property prohibits each element of $Y\backslash 0$ from having n roots in $[a, b]$. Let P, v, u, and Λ be as described in Section 3.

THEOREM 6. *In order that P be minimal among the projections from X onto Y carried by $\{t_0 ,..., t_n\}$ it is necessary and sufficient that either*

(a) $\| P \| = E$; *or*

(b) *there exist $n + 1$ critical points of Λ, $s_0 < s_1 < \cdots < s_n$, such that* $\operatorname{sgn} v(s_i) = -\operatorname{sgn} v(s_{i-1})$ $(1 \leqslant i \leqslant n)$.

Proof. First suppose that (a) holds. By Lemma 1, every projection $Q: X \to Y$ carried by $\{t_0, ..., t_n\}$ must satisfy $\| Q \| \geq E$. Hence the hypothesis (a) implies that P is minimal.

Next suppose that (b) holds. Then no element $y \in Y$ can have the property $v \operatorname{sgn} y > u$ on $S = \operatorname{crit}(\Lambda)$, for this inequality would require y to have at least n roots. Hence by Theorem 5, P is minimal.

Finally, suppose that P is minimal and that (a) is not true. Then $\| P \| > E$. By Theorem 5, no y in Y satisfies $v \operatorname{sgn} y > u$ on S. By Lemma 3, $| v | > u$ on S. Hence no y in Y can satisfy the inequality $yv > 0$ on S. Since $u \geq 0$, v does not vanish on S. Now we verify that sgn v is continuous on S. If sgn v is discontinuous on S then consider the two sets

$$S_1 = \{t \in S: v(t) > 0\},$$
$$S_2 = \{t \in S: v(t) < 0\}.$$

One of these sets must contain an accumulation point of the other. But this is not possible, for as we now show, S_1 and S_2 are closed. Consider, for example, S_1. A point t belongs to S_1 if and only if $\sum | y_i(t) | = \| P \|$ and $\sum \theta_i \operatorname{sgn} y_i(t) > 0$. Equivalently, there exists a $z \in Z$ (as in the proof of Theorem 5) such that $\hat{t}Pz = \| P \|$ and $\sum z(t_i)\, \theta_i > 0$. Since Z is finite, the set of such t (i.e., S_1) is closed. Now apply the following Lemma to complete the proof. ∎

LEMMA 4. *Let Y be an n-dimensional Haar subspace in $C[a, b]$. Let F be a closed subset of $[a, b]$. Let $\varnothing : F \to \mathbf{R}$ be a function which has no roots and is such that* sgn $\varnothing$ *is continuous. If no $y \in Y$ has the property $\varnothing y \mid F > 0$ then there exist $n + 1$ points $t_0, ..., t_n$ in F such that $t_0 < t_1 < \cdots < t_n$ and $\varnothing(t_{i-1})\, \varnothing(t_i) < 0$ for $i = 1, ..., n$.*

Proof. The system of inequalities $y(t) \operatorname{sgn} \varnothing(t) > 0$, $(y \in Y,\ t \in F)$ is inconsistent. Hence if $u_1, ..., u_n$ forms a basis for Y, then the system of inequalities $\sum_{i=1}^{n} \lambda_i u_i(t) \operatorname{sgn} \varnothing(t) > 0$ $(\lambda_i \in \mathbf{R}, t \in F)$ is inconsistent. It follows (see [10, p. 19]) that the origin of $\mathbf{R}^n$ lies in the convex hull of $\{\operatorname{sgn} \varnothing(t)\, \bar{t}: t \in F\}$ where $\bar{t}$ denotes the n-tuple $(u_1(t), ..., u_n(t))$. By Caratheodory's Theorem [10, p. 17], 0 lies in the convex hull of some $n + 1$ points $\operatorname{sgn} \varnothing(t_i)\, \bar{t}_i$. Arranging these in the order $t_0 < \cdots < t_n$ and applying the Lemma of [10, p. 74], we obtain $\varnothing(t_{i-1})\, \varnothing(t_i) < 0$. ∎

EXAMPLE. If $[a, b] = [-1, 1]$, if $Y = \Pi_2$, and if $\{t_0, ..., t_2\} = \{\pm 1, \pm s\}$ for some $s \in [\frac{1}{2}, 1)$ then the best approximation operator A of Section 2 is minimal, and its Lebesgue function has only 3 critical points. Hence by Theorem 6, $\| A \| = E$. See the examples of Section 6.

THEOREM 7. (*Strong Unicity*). *If P is minimal among the projections carried on $\{t_0, ..., t_n\}$ and if $\| P \| > E$, then P is the unique such minimal projection. Indeed there exists $\alpha > 0$ such that all other projections Q carried by $\{t_0 \ldots t_n\}$ satisfy $\| Q \| = \| P + \Phi \otimes y \| \geq \| P \| + \alpha \| y \|$.*

Proof. By Theorem 6, there exist $s_0, ..., s_n \in \text{crit}(\Lambda)$ such that $v(s_i)\, v(s_{i-1}) < 0$ $(1 \leq i \leq n)$. For each i, select $z_i \in Z$ such that $\hat{s}_i P z_i = \| P \|$. By Lemma 3, $| v(s) | > u(s)$. By Lemma 2, $\Phi(z_i)\, v(s_i) > 0$. Hence $\| Q \| \geq \hat{s}_i Q z_i = s_i P z_i + \Phi(z_i)\, y(s_i) = \| P \| + \Phi(z_i)\, y(s_i)$. Let α be the infimum of $\max_i \Phi(z_i)\, w(s_i)$ as w ranges over the surface of the unit sphere in Y. Since $v(s_i)$ alternates in sign, so does $\Phi(z_i)$. Since $w(s_i)$ cannot, $\alpha > 0$. Thus $\| Q \| \geq \| P \| + \max_i \Phi(z_i)\, y(s_i) \geq \| P \| + \alpha \| y \|$. ∎

THEOREM 8. *If $\| Pz \| < E$ whenever $| z(t_i) | = 1$ and $\Phi(z) \neq 0$, then P is minimal among the projections from X onto Y carried by $\{t_0, ..., t_n\}$ but it is not unique. Indeed, for all y in an ϵ-sphere of Y, $\| P + \Phi \otimes y \| = \| P \| = E$.*

Proof. Since the set Z (defined in the proof of Theorem 5) is finite, the number

$$\epsilon = E - \max\{\| Pz \| : z \in Z, \Phi(z) \neq 0\},$$

is positive. Since $\| P \| = \max_{z \in Z} \| Pz \|$, we conclude (using Eq. (7)) that

$$\begin{aligned} E \leq \| P \| &= \max\{\| Pz \| : z \in Z, \Phi(z) = 0\} \\ &\leq \max\{\| Pz \| : z \in Z, (Pz)(t_i) = z(t_i)\} \\ &= \max\{\| y \| : y \in Y, | y(t_i) | = 1\} \leq E. \end{aligned}$$

Now if $\| y \| \leq \epsilon$ and $\Phi(z) \neq 0$, then

$$\| Pz + \Phi(z)\, y \| \leq \| Pz \| + \| y \| \leq (E - \epsilon) + \epsilon = E.$$

If $\Phi(z) = 0$ then $\| Pz + \Phi(z)\, y \| = \| Pz \| \leq E$. Thus $\| P + \Phi \otimes y \| = \max_{z \in Z} \| Pz + \Phi(z)\, y \| \leq E$. ∎

5. THE CONSTANT LEBESGUE FUNCTIONS

An interesting phenomenon in the study of minimal projections is the occurrence of constant Lebesgue functions. The Fourier projections into spaces of trigonometric polynomials have constant Lebesgue functions, for example. A projection into an n-dimensional space and carried by $n + 1$ points can have a constant Lebesgue function, but it cannot be minimal among the projections with that carrier. This fact, together with various other related matters is discussed in this section.

We assume that $X = C[a, b]$ and that Y is an n-dimensional Haar subspace containing the constants. Without loss of generality, the points t_i are ordered: $a \leqslant t_0 < t_1 < \cdots < t_n \leqslant b$. As a consequence of this, $\theta_i\theta_{i-1} < 0$ for $i = 1,..., n$ [7, *p.* 20]. By changing the signs of all θ_i, we can arrange that $\operatorname{sgn} \theta_i = (-1)^i$. Since $\Phi(1) = 0$, $\sum_{i=0}^n \theta_i = 0$. In order to rule out some trivial special cases, we assume that $n \geqslant 3$.

LEMMA 15. *Let $P = \sum \hat{s}_i \otimes y_i$ be any projection with finite carrier and constant Lebesgue function from $C[a, b]$ into Y. Then the y_i's do not change sign.*

Proof. From the equation $P = \sum \hat{s}_i \otimes y_i$ drop any terms for which $y_i = 0$. The roots of the remaining y_i's form a finite set. Let J be an interval in which no y_i has a root. Then y_i on J has a constant sign, σ_i. On J, we have $\sum \sigma_i y_i = \sum | y_i | = c$. Since $c \in Y$, and J contains as many points as the dimension of Y, we conclude that $\sum \sigma_i y_i = c$ everywhere. Hence $c = \sum \sigma_i y_i \leqslant \sum | y_i | = c$, and $\sigma_i = \operatorname{sgn} y_i$ everywhere, except at roots of of y_i. Thus y_i does not change sign. ∎

LEMMA 16. *Let $P = \sum_{i=0}^n \hat{t}_i \otimes y_i$ be a projection of $C[a, b]$ onto Y. If the y_i do not change sign, say* $\operatorname{sgn} y_i = \sigma_i$ a.e., *then $\sigma_i\sigma_{i-1} < 0$, $i = 1,..., n$.*

Proof. Let $Q = \sum_{i=1}^n \hat{t}_i \otimes w_i$ be the interpolating projection for nodes $t_1,..., t_n$. Then $P = Q + \hat{t}_0 \circ (I - Q) \otimes y_0$ and $y_i = w_i - w_i(t_0)\, y_0$ for $i = 1,..., n$. Fixing $i \geqslant 1$, select s so that $\operatorname{sgn} w_i(s) = (-1)^i \sigma_0$. Since $\operatorname{sgn} w_i(t_0) = -(-1)^i$, we have $\operatorname{sgn} y_i(s) = \operatorname{sgn}[w_i(s) - w_i(t_0)\, y_0] = (-1)^i \sigma_0$. Hence $\sigma_i = (-1)^i \sigma_0$. ∎

THEOREM 9. *Let $A = \sum_{i=0}^n \hat{t}_i \otimes q_i$ be the projection of best approximation from $C[a, b]$ onto Y. Then the set of projections with constant Lebesgue functions carried by $\{t_0,..., t_n\}$ is $\{P_c : c \leqslant c_1 \text{ or } c \geqslant c_2\}$ where $P_c = A - \Phi \otimes c$,*

$$\| P_c \| = | c |, \qquad c_1 = \min_{i,t} \frac{q_i(t)}{\theta_i}, \qquad c_2 = \max_{i,t} \frac{q_i(t)}{\theta_i}.$$

Proof. We have $P_c = A - \Phi \otimes c = \sum \hat{t}_i \otimes [q_i - \theta_i c]$. If $c \geqslant c_2$ then $c \geqslant \theta_i^{-1} q_i(t)$ for all i and t. If i is even, then $\theta_i > 0$ and $q_i - \theta_i c \leqslant 0$. If i is odd, then $\theta_i < 0$ and $q_i - \theta_i c \geqslant 0$. Hence

$$\Lambda_{P_c} = \sum | q_i - \theta_i c | = - \sum (-1)^i (q_i - \theta_i c) = +c.$$

If $c \leqslant c_1$ then $c \leqslant \theta_i^{-1} q_i(t)$ for all i and t. As before, this implies that

$q_i - \theta_i c \geqslant 0$ when i is even and it implies that $q_i - \theta_i c \leqslant 0$ when i is odd. Hence $\Lambda_{P_c} = \sum | q_i - \theta_i c | = \sum(-1)^i(q_i - \theta_i c) = -c$.

For the converse, suppose that P is a projection which is carried by $\{t_0 \dots t_n\}$ and which has a constant Lebesgue function. Then for an appropriate $u \in Y$,

$$P = A - \Phi \otimes u = \sum \hat{t}_i \otimes [q_i - \theta_i u].$$

By Lemmas 1 and 2, each function $q_i - \theta_i u$ has a constant sign σ_i except at its roots, and $\sigma_i = (-1)^i \sigma_0$. Hence if $\| P \| = c$ then

$$c = \sum | q_i - \theta_i u | = \sum (-1)^i \sigma_0[q_i - \theta_i u] = -\sigma_0 u.$$

Thus u is $+c$ or $-c$. If $u = c$ then $\sigma_i = -(-1)^i$ and $P = P_c$. Furthermore, $0 \leqslant -(-1)^i(q_i - \theta_i c) = -(-1)^i q_i + | \theta_i | c$ whence $c \geqslant (-1)^i q_i / | \theta_i | = q_i/\theta_i$. Thus $c \geqslant c_2$. In the case $u = -c$ we have $\sigma_i = (-1)^i$, so that $0 \leqslant (-1)^i(q_i - \theta_i u) = (-1)^i(q_i + \theta_i c) = (-1)^i q_i + | \theta_i | c$, whence $-c \leqslant q_i/\theta_i$ and $-c \leqslant c_1$. Now $P_{-c} = A - \Phi \otimes (-c) = A - \Phi \otimes u = P$. ∎

THEOREM 10. *Each projection of $C[a, b]$ onto Y which is carried by $\{t_0, \dots, t_n\}$ and has constant Lebesgue function must have norm $\geqslant n$ if n is odd, and norm $\geqslant n - 1$ if n is even.*

Proof. The θ_i with even index sum to $\frac{1}{2}$ because $1 = \sum | \theta_i | = \sum \theta_i = 2(\theta_0 + \theta_2 + \theta_4 + \cdots)$. It follows that the θ_i with odd index sum to $-\frac{1}{2}$. By Theorem 9, the minimum norm for the projections contemplated in Theorem 10 is $c = \min\{c_2, -c_1\}$. Now

$$c_2 = \max_{i,t} \theta_i^{-1} q_i(t) \geqslant \max_i \theta_i^{-1} q_i(t_i) = \max_i \theta_i^{-1}(1 - | \theta_i |)$$
$$= \max_{i \text{ even}} \theta_i^{-1}(1 - \theta_i) = -1 + (\min_{i \text{ even}} \theta_i)^{-1}.$$

If n is even then there are $(n + 2)/2$ θ_i with even index. Hence

$$\frac{n + 2}{2} \min\{\theta_0, \theta_2, \theta_4, \dots\} \leqslant \theta_0 + \theta_2 + \theta_4 + \cdots = \tfrac{1}{2}$$

whence $c_2 \geqslant n + 1$. If n is odd, there are $(n + 1)/2$ θ_i with even index. Hence the same reasoning as before leads to the conclusion that $c_2 \geqslant n$. A similar analysis for $-c_1$ leads us to conclude that $-c_1 \geqslant n - 1$ when n is even and $-c_1 \geqslant n$ if n is odd. ∎

In the next theorem, we use the notation $Z(y)$ to denote the set of roots of a function y.

LEMMA 7. *Let $P = \sum_{i=0}^{n} \hat{t}_i \otimes y_i$ be a projection from $C[a, b]$ onto Y. Assume that*

(i) $\operatorname{sgn} y_i = \sigma_i$ a.e.

(ii) $\sigma_i = (-1)^i \sigma_0$

(iii) $Z(y_i) \cap Z(y_j) \cap \operatorname{crit} \Lambda_P = \Box$ *for* $i \neq j$.

Then P is not minimal in the class of all projections from $C(T)$ onto Y carried by $\{t_0, ..., t_n\}$.

Proof. Put $S = \operatorname{crit} \Lambda_P$. Let $\{s_1, ..., s_m\} = S \cap \bigcup_{i=0}^{m} Z(y_i)$. For each i, s_i is a root of exactly one of the y_j's because of hypothesis (iii). Hence for $i = 1, ..., m$ we can select an open neighborhood U_i of y_i in such a way that $\bar{U}_i \cap \bar{U}_j = \Box$ for $i \neq j$. Let

$$\delta_1 = \min_{1 \leqslant i \leqslant m} \min_{\substack{0 \leqslant j \leqslant n \\ j \neq i}} \min_{t \in \bar{U}_i} | y_j(t)/\theta_j | ,$$

$$\delta_2 = \min_{0 \leqslant i \leqslant n} \min_{s \in S \setminus \bigcup_{\nu=1}^{m} U_\nu} | y_i(s)/\theta_i | .$$

Select $\epsilon > 0$ such that $\epsilon < \min(\delta_1, \delta_2)$. We shall prove that for some $\lambda > 0$, $P - \lambda\Phi \otimes \epsilon\sigma_0$ has norm less than $\| P \|$. By Theorem 1 of [4],it suffices to prove that $\sum | y_i(s) - \theta_i \epsilon \sigma_0 | < \sum | y_i(s)|$ for all $s \in S$.

If $s \in S \setminus \bigcup_{\nu=1}^{m} U_\nu$ then $| y_i(s)/\theta_i | \geqslant \delta_2 > \epsilon$. Hence $\sum | y_i(s) - \theta_i \epsilon \sigma_0 | = \sum \{| y_i(s)| - \epsilon |\theta_i |\} = \Lambda(s) - \epsilon$.

If $s \in S \cap \bigcup_{\nu=1}^{m} U_\nu$ then $s \in U_j$ for some j. If $i \neq j$ then $| y_i(s)/\theta_i | \geqslant \delta_1 > \epsilon$. Put $I = \{i\colon | y_i(s)| \geqslant | \theta_i \sigma_0 \epsilon |\}$ and $J = \{i\colon | y_i(s)| < | \theta_i \sigma_0 \epsilon |\}$. Then

$$\begin{aligned}
\sum | y_i(s) - \theta_i \epsilon \sigma_0 | &= \sum_I \{| y_i(s)| - | \theta_i | \epsilon\} + \sum_J \{\epsilon | \theta_i | - | y_i(s)|\} \\
&\leqslant \sum_I | y_i(s)| - \Big\{\Lambda(s) - \sum_I | y_i(s)|\Big\} - \epsilon \Big(1 - \sum_J | \theta_i |\Big) \\
&\quad + \epsilon | \theta_j | \leqslant \Lambda(s) - \epsilon + \epsilon \sum_J | \theta_i | + \epsilon | \theta_j | \\
&\leqslant \Lambda(s) - \epsilon + 2\epsilon | \theta_j | < \Lambda(s).
\end{aligned}$$

Here we used the inclusion $J \subset \{j\}$, and the fact that $| \theta_i | < \frac{1}{2}$. The latter follows from the hypothesis $n \geqslant 3$. ∎

THEOREM 11. *A projection with constant Lebesgue function carried by $n+1$ points from $C[a, b]$ onto Y cannot be minimal in the class of all projections carried by those $n+1$ points.*

Proof. Suppose on the contrary that $P = \sum_0^n \hat{t}_i \otimes y_i$ is a projection with constant Lebesgue function which is minimal in the class of all projections carried by $\{t_0, \ldots, t_n\}$. Let $A = \sum_0^n \hat{t}_i \otimes q_i$ be the "best approximation" operator for these $n+1$ points. Then for some c, $P = A + \Phi \otimes c$, and $y_i = q_i + \theta_i c$. Assume $a \leqslant t_0 < \cdots < t_n \leqslant b$. By Lemma 7, two y_i's have a common root. Suppose s is a root of y_ν and y_μ.

Now $s \notin \{t_0, \ldots, t_n\}$ because

$$y_i(t_j) = \delta_{ij} - (-1)^j \theta_i + \theta_i c = \delta_{ij} + \theta_i[c - (-1)^j].$$

Since $|c| = \|P\| > 1$ by Theorem 9 of [6], $y_i(t_j)$ can vanish only in the case $i = j$.

Define $w = \theta_\mu y_\nu - \theta_\nu y_\mu$. For all j, $w(t_j) = \theta_\mu y_\nu(t_j) - \theta_\nu y_\mu(t_j) = \theta_\mu[\delta_{\nu j} - (-1)^j \theta_\nu + \theta_\nu c] - \theta_\nu[\delta_{\mu j} - (-1)^j \theta_\mu + \theta_\mu c] = \theta_\mu \delta_{\nu j} - \theta_\nu \delta_{\mu j}$. If $j = \nu$, then $w(t_j) = \theta_\mu \neq 0$. Hence $w \in Y \backslash 0$. Trivially, $w(s) = 0$. If $j \in \{0, \ldots, n\} \backslash \{\mu, \nu\}$ then $w(t_j) = 0$. Hence w has n distinct roots, contradicting the Haar property of Y. ∎

Remark. That $n \geqslant 3$ is a necessary hypothesis can be seen from the example of first degree polynomials in $C[a, b]$. Interpolation at the endpoints of the interval is a projection of norm 1 having a constant Lebesgue function.

6. PROJECTIONS ONTO POLYNOMIAL SUBSPACES

The considerations of the preceding sections are specialized further to the space $C[-1, 1]$ and its subspace Π_{n-1} consisting of polynomials of degree $\leqslant n - 1$. Our choice of interval is a matter of convenience, not necessity. And we use the subspace Π_{n-1} rather than Π_n because we want the dimension to be n, as it was in the previous parts of the paper.

THEOREM 12. *For any set N of $n+1$ nodes $t_0 < \cdots < t_n$ in $[-1, 1]$, let $A(N)$ be the projection of best approximation on these nodes, projecting $C[-1, 1]$ onto Π_{n-1}. Among the minimal such projections (varying the nodes but holding n fixed) there is one for which the endpoints of the interval are nodes.*

Proof. Let $A(N)$ be a minimal such projection. If $t_0 > -1$ or $t_n < +1$ then select β and γ so that the map $t \rightarrow t' = \beta t + \gamma$ carries t_0 into -1 and t_n into $+1$. We shall prove that $\|A(N')\| \leqslant \|A(N)\|$, where $N' = (t_0', \ldots, t_n')$.

The operator A is given by the formulas

$$A = \sum_{i=0}^{n} \hat{t}_i \otimes q_i , \qquad q_i = z_i - \theta_i r,$$

$$z_i(t) = \alpha_i \prod_{\substack{j=0 \\ j\neq i}}^{n} (t - t_j),$$

$$\alpha_i = \prod_{\substack{j=0 \\ j\neq i}}^{n} (t_i - t_j)^{-1},$$

$$\theta_i = (-1)^n \alpha_i \Big/ \sum_{j=0}^{n} | \alpha_j | ,$$

$$r = \sum_{i=0}^{n} (-1)^i z_i .$$

These functions and coefficients depend on N, and we note that $q_i(N', t') = q_i(N, t)$. Hence

$$\| A(N')\| = \sup_{-1\leqslant t'\leqslant 1} \sum | q_i(N', t')| = \sup_{t_0\leqslant t\leqslant t_n} \sum | q_i(N, t)|$$

$$\leqslant \sup_{-1\leqslant t\leqslant 1} \sum | q_i(N, t)| = \| A(N)\| . \quad \blacksquare$$

EXAMPLE. Select $s \in (0, 1)$ and let $(t_0 , t_1 , t_2 , t_3) = (-1, -s, s, 1)$. Consider the subspace Π_2 of quadratic polynomials in $C[-1, 1]$. For this example we shall give the values of θ_i, the "best approximation operator", and its norm:

$$\theta_0 = s/\alpha, \qquad \alpha = 2(1 + s),$$
$$\theta_1 = -1/\alpha,$$
$$\theta_2 = 1/\alpha,$$
$$\theta_3 = -s/\alpha,$$
$$q_0 = (t^2 - t + st - s^2)/\beta, \qquad \beta = 2(1 - s^2),$$
$$q_1 = (-t^2 - t + st + 1)/\beta,$$
$$q_2 = (-t^2 + t - st + 1)/\beta,$$
$$q_3 = (t^2 + t - st - s^2)/\beta,$$
$$A_s = \sum \hat{t}_i \otimes q_i .$$

$$\| A_s \| = \begin{cases} \dfrac{s + 2}{s + 1} , & 0 < s \leqslant \tfrac{1}{2}, \\[2ex] \dfrac{1 + s^2}{1 - s^2} , & \tfrac{1}{2} \leqslant s < 1. \end{cases}$$

The values of θ_i can be checked by verifying $\sum |\theta_i| = 1$ and $\sum \theta_i \hat{t}_i \in Y^{\perp}$. The values of q_i can be checked by verifying $Ay = y$ for $y \in Y$ and $\sum (-1)^i q_i = 0$. The calculation of $\| A_s \|$ is more tedious. The function $\sum | q_i |$ is even, and piecewise quadratic. An analysis of the roots of the functions q_i is thus called for. The function q_0 has two roots, r_1 and r_2, which satisfy $r_1 < 0 < r_2 < 1$. The function q_1 has one root in the interval $(-\infty, 1)$ and another, r_3, in the interval $(-r_1, r_2)$. Since $q_3(t) = q_0(-t)$ and $q_2(t) = q_1(-t)$, all the roots of Πq_i in $[0, 1]$ are (in order) $-r_1, r_3, r_2$. On $[0, -r_1]$, the Lebesgue function Λ is $-q_0 + q_1 + q_2 - q_3$ and has as its maximum value $(1 + s^2)/(1 - s^2)$. On $[-r_1, r_3]$, $\Lambda = -q_0 + q_1 + q_2 + q_3$. The global maximum of the function is $(5 - s^2)/(4 - 4s^2)$, and this value does not exceed $\max\{\Lambda(0), \Lambda(1)\}$. On the interval $[r_3, r_2]$, $\Lambda = -q_0 - q_1 + q_2 + q_3$, and its maximum value is at r_2. Finally, on the interval $[r_2, 1]$, $\Lambda = q_0 - q_1 + q_2 + q_3$, and its maximum is $(s + 2)/(s + 1)$. The formula for $\| A_s \|$ now depends on the assertion that $(1 + s^2)/(1 - s^2) \geq (s + 2)/(s + 1)$ iff $\frac{1}{2} \leq s < 1$.

THEOREM 13. *The norm of A_s is a minimum only if $s = \frac{1}{2}$, and the value is then $\frac{5}{3}$. For each $s \in [\frac{1}{2}, 1)$, A_s is a minimal in the class of all projections from $C[-1, 1]$ onto Π_2 carried by $\{-1, -s, s, 1\}$. For each $s \in (0, \frac{1}{2})$, A_s is not minimal in this class.*

Proof. The first assertion follows at once from the formula given for $\| A_s \|$. The second and third assertions involve applications of Theorem 5. One needs to know that

$$\operatorname{crit} \sum | q_i | = \begin{cases} \{-1, 1\}, & 0 < s < \frac{1}{2}, \\ \{-1, 0, 1\}, & s = \frac{1}{2}, \\ \{0\}, & \frac{1}{2} < s < 1. \end{cases}$$

If $\frac{1}{2} \leq s < 1$ then v (as defined before Theorem 5) has a root at 0. Hence the sufficient condition for minimality is fulfilled. If $0 < s < \frac{1}{2}$, the function $y(t) = t$ has the property that $v \operatorname{sgn} y > u$ on the critical set, $\{-1, 1\}$. ∎

The proof of the following theorem is too long to be included here.

THEOREM 14. *If $0 < s < \frac{1}{2}$ then the minimal projection of $C[-1, 1]$ onto Π_2 carried by $\{-1, -s, s, 1\}$ is the operator $P_s = A_s - \lambda \Phi \otimes e$, where A_s is as described above,*

$$e(t) = t, \qquad \text{and} \qquad \lambda = s^{-2}(1 - s)^{-1}[s^2 - s - 2 + 2(1 + s - s^3 - s^4)^{1/2}].$$

The norm of P_s is $(2 + s - \lambda)(1 + s)^{-1}$ and the critical points of its Lebesgue function are ± 1 and $\pm\frac{1}{2}(1 + \lambda s)(1 - s)$. If $0 < s < \frac{1}{3}$, then P_s is the unique such minimal projection.

REFERENCES

1. C. DE LA VALLÉE POUSSIN, Sur la méthode de l'approximation minimum, Société Scientifique de Bruxelles, Annales, 2nd partie, *Mémoires* **35** (1911), 1–16.
2. C. DE LA VALLÉE POUSSIN, Sur les polynômes d'approximation et la représentation approchée d'un angle, *Acad. Roy. Belg. Bull. Cl. Sci.* **12** (1910), 808–844.
3. C. DE LA VALLÉE POUSSIN, "Leçons sur l'Approximation des Fonctions d'une Variable Réelle," Paris, Gauthier-Villars, 1919. Reprinted 1952.
4. E. W. CHENEY, Projections with finite carriers, Proceedings of a Conference at Oberwolfach, June 1971, *ISNM* **16** (1972), 19–32.
5. T. S. MOTZKIN AND A. SHARMA, Next-to-interpolatory approximation on sets with multiplicities, *Canad. J. Math.* **18** (1966), 1196–1211.
6. P. D. MORRIS AND E. W. CHENEY, On the existence and characterization of minimal projections, *J. fur die Reine und Angewandte Mathematik* **270** (1974), 61–76.
7. G. MEINARDUS, "Approximation of Functions: Theory and Numerical Methods." Springer, New York, 1967.
8. T. S. MOTZKIN AND A. SHARMA, A sequence of linear polynomial operators and their approximation-theoretic properties, *J. Approximation Theory* **5** (1972), 176–198.
9. S. PASZKOWSKI, Polynomes et Series de Tchebycheff, Vol. II, pp. 1–46, Université Scientifique et Médicale de Grenoble, Institute de Mathématique Appliquées, Seminaire d'Analyse Numérique, 1970–1971.
10. E. W. CHENEY, "Introduction to Approximation Theory," McGraw Hill, New York, 1966.
11. O. SHISHA, Best Approximations on some finite sets, *J. Math. Anal. Appl.* **21** (1968), 347–355.