JOURNAL OF APPROXIMATION THEORY 4, 401–418 (1971)

Uniqueness of Best Approximation by Monotone Polynomials*

R. A. LORENTZ

Gesellschaft für Mathematik und Datenverarbeitung, 5201 Birlinghoven (Schloss) West Germany

Communicated by Oved Shisha

Received February 20, 1970

The purpose of this paper is twofold. First, we prove the uniqueness of a polynomial of best uniform approximation, in a certain class $\mathscr{P}$ of "monotone" polynomials, to a given continuous function. This is the content of Theorem 3.1 which complements the results of Lorentz and Zeller [1]. Secondly, we prove (Theorem 6.1) that a polynomial of best L_1 approximation in the class $\mathscr{P}$, to a given continuous function is also unique. This is the analog of Jackson's theorem for general polynomials. As a preliminary to Theorem 6.1, we give a necessary condition for a polynomial in $\mathscr{P}$ to be a polynomial of best L_1 approximation to an integrable function.

1. PRELIMINARIES

Let $1 \leqslant k_1 < \cdots < k_p$ be positive integers and let $\epsilon_i = \pm 1$, $i = 1,..., p$. For a positive integer n (which will remain fixed throughout the paper), we denote by $\mathscr{P} = \mathscr{P}(k_1 ,..., k_p ; \epsilon_1 ,..., \epsilon_p)$ the set of all polynomials P, of degree not exceeding n, satisfying

$$\epsilon_i P^{(k_i)}(x) \geqslant 0, \qquad a \leqslant x \leqslant b, \quad i = 1,..., p. \tag{1.1}$$

Since (1.1) holds automatically for $k_i > n$, we may as well assume in the following that $k_p \leqslant n$, which we do.

Compactness and convexity arguments show that for each f in $C[a, b]$ there exists at least one polynomial in $\mathscr{P}$ of best uniform approximation. If f is merely known to be integrable, then there is at least one polynomial in $\mathscr{P}$ of best L_1 approximation. Our problem is to prove that if f is continuous,

* This work was partially supported by the Air Force Office of Scientific Research, Contract AF 49(638)-1401.

then there exists in $\mathcal{P}$ only one polynomial of best uniform approximation, and only one polynomial of best L_1 approximation. Simple examples show that if f is discontinuous, there may be in $\mathcal{P}$ more than one polynomial of best L_1 approximation.

In [1], Lorentz and Zeller developed to a considerable extent the theory of the class $\mathcal{P}$. However, they proved their main result, the uniqueness theorem for best uniform approximation, only for the case $p = 1$. We treat here the more intricate general case.

For any pair $P \in \mathcal{P}$, $f \in C[a, b]$, we define the sets

$$A(P, f) = \{x \mid x \in [a, b], |f(x) - P(x)| = \|f - P\|\}, \tag{1.2}$$

$$B_j(P) = \{x \mid x \in [a, b], P^{(k_j)}(x) = 0\}, \qquad j = 1,..., p. \tag{1.3}$$

Here $\|\cdot\|$ is the uniform norm on $[a, b]$.

We shall use the following results from [1]:

LEMMA 1.1. *A polynomial $P \in \mathcal{P}$ is a polynomial of best uniform approximation to $f \in C[a, b]$ if and only if there is no polynomial Q of degree not exceeding n, for which*

$$Q(x)\, \sigma[f(x) - P(x)] < 0, \qquad x \in A(P, f), \tag{1.4}$$

$$\epsilon_i Q^{(k_i)}(x) < 0, \qquad x \in B_i(P), \quad i = 1,..., p, \tag{1.5}$$

where $\sigma(u)$ is the sign of u.

LEMMA 1.2. *For a given $f \in C[a, b]$, there exists among all polynomials in $\mathcal{P}$ of best uniform approximation to f, a "minimal polynomial" P_0, which satisfies*

$$A(P_0, f) \subset A(P, f), \tag{1.6}$$

$$B_i(P_0) \subset B_i(P), \qquad i = 1,..., p, \tag{1.7}$$

$$P_0(x) = P(x), \qquad x \in A(P_0, f) \tag{1.8}$$

for any other polynomial $P \in \mathcal{P}$ of best uniform approximation to f. Moreover, $\deg(P_0) \geqslant \deg(P)$ *for any such P.*

Remark. The last claim is not proved in [1] but is immediate.

An essential tool in the following is the concept of "free" or "poised" matrices and the associated Birkhoff interpolation problem. Let $E = (e_{ij})$ be an $m \times (n + 1)$ matrix. E is called an "incidence" matrix if E has only zeros or ones as entries. That is, for each $1 \leqslant i \leqslant m$, $0 \leqslant j \leqslant n$, e_{ij} is 0 or 1.

We set $e = \{(i, j) \mid e_{ij} = 1\}$. It is generally assumed that an incidence matrix has exactly $n + 1$ nonzero entries. However, for the sake of convenience, we do not make this assumption.

If the number of nonzero entries of some incidence matrix E is $n + 1$, then for any choice of points $x_1 < x_2 \cdots < x_m$ and of real numbers b_{ij}, $(i, j) \in e$, we can associate with E the following Birkhoff interpolation problem (B.I.P.) for a polynomial Q of degree not exceeding n: Determine Q so that

$$Q^{(j)}(x_i) = b_{ij}, \qquad (i, j) \in e.$$

Conversely, consider the problem of determining a polynomial Q of degree $\leqslant n$, satisfying the $n + 1$ conditions:

$$Q^{(\alpha_{ij})}(y_i) = \beta_{ij}, \qquad 0 \leqslant \alpha_{ij} \leqslant n \tag{1.9}$$

where the integers α_{ij}, the reals β_{ij} and the (distinct) reals y_j are given. Let $\lambda_1 < \cdots < \lambda_m$ be the points y_i arranged in increasing order. Then to this problem there corresponds an $m \times (n + 1)$ incidence matrix $E = (e_{ij})$ for which $e_{ij} = 1$ if $Q^{(j)}(\lambda_i)$ appears in one of the conditions (1.9) and $e_{ij} = 0$ otherwise.

If $E = (e_{ij})$ is an $m \times (n + 1)$ incidence matrix and if the corresponding interpolation problem has a unique solution, regardless of the values of the x_i and b_{ij}, then E is said to be "free" (or "poised").

Let

$$m_j = \sum_{i=1}^{m} e_{ij}, \qquad j = 0, 1,..., n.$$

Schoenberg [2] proved that each free incidence matrix E satisfies the Pólya conditions

$$\sum_{j=0}^{k} m_j \geqslant k + 1, \qquad k = 0, 1,..., n.$$

Also useful for us will be the "strong" Pólya conditions for an $m \times (n + 1)$ incidence matrix:

$$\sum_{j=0}^{k} m_j \geqslant k + 2, \qquad k = 0, 1,..., n - 1. \tag{1.10}$$

It is to be noted that we formulate the Pólya and strong Pólya conditions for $m \times (n + 1)$ incidence matrices without requiring the number of nonzero entries to be $n + 1$, although this is not customary.

Atkinson and Sharma [3] gave a sufficient condition for an incidence matrix to be free. To state this condition, we introduce the notions of "maximal" and "supported" sequences.

A "maximal" sequence of the incidence matrix E is a sequence of 1's:

$$e_{ij}, e_{i,j+1}, \dots, e_{i,j+r}$$

which cannot be extended to a longer sequence of 1's of this form. A maximal sequence of E is said to be "supported" if, when written in the above form, there exist integers i_0, i_1, j_0, j_1, with $0 \leqslant j_0 < j$, $0 \leqslant j_1 < j$, and $1 \leqslant i_0 < i < i_1 \leqslant m$, for which $e_{i_0 j_0} = e_{i_1 j_1} = 1$.

If every supported sequence of E has an even number of elements, then E is said to satisfy the Atkinson–Sharma condition. These authors proved that if the $m \times (n+1)$ incidence matrix E (with exactly $n+1$ nonzero entries) satisfies the Atkinson–Sharma condition and the Pólya conditions, then E is free.

2. Intermediate Lemmas

In this section, we establish some lemmas needed for the proof of the uniqueness theorem for the uniform norm. For $f \in C[a, b]$, we define the sets $A(P, f)$, $B_j(P)$ by (1.2) and (1.3) and denote by m, l_j, $j = 1, \dots, p$, the number of elements of these sets, respectively ($m = \infty$ and $l_j = \infty$ not being excluded).

Clearly, if $B_j(P)$ is infinite, then $B_j(P) = [a, b]$. However, $A(P, f)$ can be infinite without being trivial.

By e_j we denote the number of elements of $B_j(P) \cap \{a, b\}$. Then e_j is either zero, one or two.

It would be interesting to characterize completely the cardinality of sets A, B_j that can serve as sets $A(P, f)$, $B_j(P)$ for some pair P, f. In this direction, we have the inequalities (see [1]):

$$2l_j - e_j \leqslant k_{j+1} - k_j, \qquad j = 1, \dots, p-1,$$
$$2l_p - e_p \leqslant n - k_p.$$

We omit the proof (which is not difficult), since these inequalities will not be needed in the sequel.

The following restrictions on the sets B_j are useful.

Lemma 2.1. *For each $P \in \mathscr{P}$,*

$$P^{(k_j+1)}(y) = 0, \qquad y \in B_j(P) \cap (a, b), \quad j = 1, \dots, p. \tag{2.1}$$

If $k_{j+1} = k_j + 1$ *for some* $j = 1,..., p - 1$, *and if* P *is of degree at least* k_j, *then*

$$B_j(P) \subset \{a, b\}. \tag{2.2}$$

Proof. Let $y \in B_j(P) \cap (a, b)$. Then $\epsilon_j P^{(k_j)}(y) = 0$. If $P^{(k_j+1)}(y) \neq 0$, then $\epsilon_j P^{(k_j)}$ would change sign at y. Since $\epsilon_j P^{(k_j)} \geqslant 0$ on $[a, b]$, this is impossible.

If $k_{j+1} = k_j + 1$ for some $j = 1,..., p - 1$ and the degree of P is at least k_j, then $P^{(k_j)}$ must be either monotone increasing or monotone decreasing (and not identically zero), since $\epsilon_{j+1} P^{(k_{j+1})} \geqslant 0$. Since $P^{(k_j)}$ does not change sign on $[a, b]$, it can have a zero only at one of the end-points a or b. This proves (2.2).

It will be convenient to associate with each polynomial $P \in \mathscr{P}$ and each function $f \in C[a, b]$, a certain incidence matrix $E(P, f)$. We assume that $A(P, f)$ is finite and we denote by ν the degree of P. Moreover, we denote by x_i, $i = 1,..., m$, y_{ji}, $k_j \leqslant \nu$; $i = 1,..., l_j$ the elements of $A(P, f)$ and $B_j(P)$, $k_j \leqslant \nu$, respectively. Note that m and l_j are finite for $k_j \leqslant \nu$. Let $E(P, f)$ be the incidence matrix corresponding to the following conditions (for some α_i, β_{ji}, γ_{ji}):

$$Q(x_i) = \alpha_i, \qquad i = 1,..., m, \tag{2.3}$$

$$Q^{(k_j)}(y_{ji}) = \beta_{ji}, \qquad k_j \leqslant \nu, \quad i = 1,..., l_j, \tag{2.4}$$

$$Q^{(k_j+1)}(y_{ji}) = \gamma_{ji}, \qquad a < y_{ji} < b, \quad k_j + 1 \leqslant \nu, \quad i = 1,..., l_j. \tag{2.5}$$

Moreover, let

$$N = m - 1 + \sum_{k_j \leqslant \nu} l_j + \sum_{k_j+1 \leqslant \nu} (l_j - e_j). \tag{2.6}$$

LEMMA 2.2. *Let* $P \in \mathscr{P}$ *be a polynomial of best uniform approximation to* $f \in C[a, b]$. *Let* $A(P, f)$ *be finite and let* ν *be the degree of* P. *Then the above-defined incidence matrix* $E(P, f)$ *satisfies the Atkinson–Sharma condition and the strong Pólya conditions. Also, if* N *is given by* (2.6), *then*

$$N = m - 1 + \sum_{k_j \leqslant \nu} (2l_j - e_j), \tag{2.7}$$

$$N \geqslant \nu + 1, \tag{2.8}$$

and (2.3–5) *consist of exactly* $N + 1$ *distinct nonoverlapping conditions.*

Proof. We first prove the last claim of the lemma. There will be exactly $N + 1$ distinct conditions (2.3)–(2.5) if we can show that none of the conditions overlap. The only possibility for overlap is if there exists a j_0 with $k_{j_0} < \nu$ and $k_{j_0} + 1 = k_{j_0+1}$ together with points $y_{j_0 i_0} = y_{j_0+1, i_0'} \in (a, b)$.

But according to Lemma 2.1, if $k_{j_0} + 1 = k_{j_0+1}$, then $B_{j_0}(P) \subset \{a, b\}$ and so $y_{j_0 i_0} \notin (a, b)$.

Next we prove (2.7). It is necessary only to prove that $l_j - e_j = 0$ if $k_j + 1 = \nu + 1$ since, then, (2.6) and (2.7) coincide. If $k_j = \nu$, then $P^{(k_j)}$ is a nonzero constant since P is of degree ν. Thus $l_j - e_j = 0$, since $l_j = e_j = 0$.

The last two paragraphs show that the conditions (2.4) and (2.5) come in nonoverlapping pairs, whenever $y_{ji} \in (a, b)$. Thus, any maximal sequence of $E(P, f)$ not lying in the first or last row, either begins in the first column or is of even length. This means that every supported sequence is of even length and, so, $E(P, f)$ satisfies the Atkinson–Sharma condition.

We now show that $E(P, f)$ satisfies the strong Pólya conditions. Since each polynomial P of best approximation deviates from f by $\| f - P \|$ in at least two points, $m \geqslant 2$; hence (1.10) is satisfied for $k = 0$. Assume that (1.10) is not satisfied for some k, $0 < k < \nu$. We shall reach a contradiction. Let $\bar{k}$ be the smallest k for which (1.10) is violated. Consider the incidence matrix $\bar{E}$ consisting of the columns of $E(P, f)$ numbered 0 to $\bar{k}$. By assumption,

$$\sum_{j=0}^{\bar{k}} m_j \leqslant \bar{k} + 1,$$

and (1.10) is satisfied for $0 \leqslant k \leqslant \bar{k} - 1$. This implies that $m_{\bar{k}} = 0$ and that

$$\sum_{j=0}^{\bar{k}} m_j = \bar{k} + 1.$$

Since the $\bar{k}$-th column of $E(P, f)$ has only zeros as entries, no maximal sequence of $E(P, f)$ can cross this column. Consequently, $\bar{E}$ must satisfy the Atkinson–Sharma condition.

We consider the B.I.P. for a polynomial Q of degree not exceeding $\bar{k}$, corresponding to the matrix $\bar{E}$, and with data:

$$\begin{aligned} Q(x_i) &= \alpha_i = -\sigma[f(x_i) - P(x_i)], \qquad i = 1,\ldots, m, \\ Q^{(k_j)}(y_{ji}) &= 0, \qquad k_j \leqslant \bar{k}, \quad i = 1,\ldots, l_j, \\ Q^{(k_j+1)}(y_{ji}) &= 0, \qquad k_j + 1 \leqslant \bar{k}, \quad a < y_{ji} < b, \quad i = 1,\ldots, l_j. \end{aligned} \tag{2.9}$$

Since $\bar{E}$ satisfies the Atkinson–Sharma condition and the strong Pólya conditions, it is free. Hence a Q of degree not exceeding $\bar{k}$ satisfying the Eqs. (2.9) exists. Since Q also satisfies

$$Q^{(k)}(x) = 0, \qquad k > \bar{k},$$

we obtain a contradiction to Lemma 1.1. (We note that (1.4) is satisfied, since m is finite and, hence, $\| f - P \| \neq 0$.)

Finally, we show that (2.8) is satisfied. The proof proceeds similarly. We assume that $N \leqslant \nu$ and reach a contradiction. Since we have proved that $E(P, f)$ satisfies the strong Pólya conditions, we must have $N = \nu$. We consider the B.I.P. (2.9), with $\bar{k}$ replaced by ν. There are exactly $\nu + 1$ conditions. Since $E(P, f)$ satisfies the Atkinson–Sharma condition and the strong Pólya conditions, a solution Q of degree not exceeding ν exists. Since also

$$Q^{(k)}(x) = 0, \qquad k > \nu,$$

Lemma 1.1 is violated. Thus $N \geqslant \nu + 1$.

Remark. By means of the inequality (2.8) and the fact that $E(P, f)$ satisfies the strong Pólya conditions, we can obtain the following inequalities which help to characterize the sets $A(P, f)$ and $B_i(P)$:

$$m \geqslant k_1 + 1,$$

$$m + \sum_{j=1}^{q} (2l_j - e_j) \geqslant k_{q+1} + 1, \qquad q \leqslant n - 1, \quad k_{q+1} - 1 \leqslant \nu,$$

$$m + \sum_{k_j \leqslant \nu} (2l_j - e_j) \geqslant \nu + 2.$$

The proof is not difficult if one keeps in mind the structure of $E(P, f)$. We omit the proof, however, since we do not use these inequalities.

3. The Uniqueness Theorem For The Uniform Norm

By means of Lemmas 2.1 and 2.2, the uniqueness theorem for the uniform norm can now be proved.

THEOREM 3.1. *Let $f \in C[a, b]$ be given. Then among all polynomials of $\mathscr{P}$ there is exactly one which approximates f best in the uniform norm.*

Proof. Let P_0 be a minimal polynomial as described in Lemma 1.2. We suppose that there is more than one polynomial of best approximation and reach a contradiction. Let P be any other polynomial of best approximation. If $\deg(P_0) = \nu$, then $\deg(P) \leqslant \nu$. We shall show that $D = P_0 - P$ is identically zero.

Let $A = A(P_0, f)$, $B_j = B_j(P_0)$ and let x_i, y_{ji}, m, l_j, e_j be the points and numbers associated with A and B_j. Since $\deg(P_0) = \nu$, l_j is finite for

$k_j \leqslant \nu$. We may also assume that m is finite since, otherwise, $D \equiv 0$ obviously. By Lemmas 1.2 and 2.1, D satisfies the conditions

$$\begin{aligned} D(x_i) &= 0, \qquad i = 1,\dots, m, \\ D^{(k_j)}(y_{ji}) &= 0, \qquad k_j \leqslant \nu, \quad i = 1,\dots, l_j, \\ D^{(k_j+1)}(y_{ji}) &= 0, \qquad a < y_{ji} < b, \quad k_j + 1 \leqslant \nu, \quad i = 1,\dots, l_j. \end{aligned} \tag{3.1}$$

The incidence matrix corresponding to these conditions, $E(P_0, f)$, is exactly the $E(P, f)$ of Lemma 2.2. The total number of 1's in this matrix is $N + 1 \geqslant \nu + 2$, where N is given by (2.6). If we add to $E(P_0, f)$ zero columns, numbered $\nu + 1$ through N, in order to have a total of $N + 1$ columns, then the new matrix is free, by Lemma 2.2 and the Atkinson–Sharma theorem. Since D is a polynomial of degree not exceeding ν and, hence, not exceeding N, satisfying (3.1), D must be identically zero. This is the desired contradiction.

4. Preliminaries On The L_1 Norm

Let $f \in C[a, b]$. By compactness and convexity arguments, it follows that there exist polynomials in $\mathscr{P} = \mathscr{P}(k_1, \dots, k_p; \epsilon_1, \dots, \epsilon_p)$ which among all polynomials in $\mathscr{P}$, approximate f best in the L_1 norm. Our problem is to show that there is no more than one such polynomial.

Since the theory of approximation in the L_1 norm has not yet been developed for the class $\mathscr{P}$, we shall develop the necessary parts here. In particular, Lemmas 1.1 and 2.1 must be suitably replaced.

Theorem 4.1. *If $P \in \mathscr{P}$ is a polynomial of best L_1 approximation to $f \in C[a, b]$ and if $f - P \neq 0$ a.e. in $[a, b]$, then*

$$\int_a^b Q(x)\, \sigma[f(x) - P(x)]\, dx \leqslant 0 \tag{4.1}$$

for every polynomial $Q \in \mathscr{P}$ of degree not exceeding n. Moreover,

$$\int_a^b P(x)\, \sigma[f(x) - P(x)]\, dx = 0.$$

Proof. Suppose, to the contrary, that there is a $Q \in \mathscr{P}$ of degree not exceeding n for which

$$\int_a^b Q(x)\, \sigma[f(x) - P(x)]\, dx = \delta > 0. \tag{4.2}$$

Since $\sigma(f - P) \neq 0$ a.e., $\sigma[f - (P + \lambda Q)]$ converges to $\sigma(f - P)$ a.e. Hence,

$$\int_a^b Q\sigma[f - (P + \lambda Q)]\, dx \to \int_a^b Q\sigma(f - P)\, dx$$

as $\lambda \to 0$. We may thus choose a $\lambda > 0$ sufficiently small so that

$$\lambda \int_a^b Q\sigma[f - (P + \lambda Q)]\, dx \geqslant \lambda\delta/2 > 0.$$

Let $\bar{P} = P + \lambda Q$. Then $\bar{P} \in \mathscr{P}$ and is of degree not exceeding n. Also

$$\begin{aligned} \| f - P \|_1 &= \int_a^b | f - P |\, dx \geqslant \int_a^b (f - P)\, \sigma[f - (P + \lambda Q)]\, dx \\ &= \int_a^b | f - (P + \lambda Q)|\, dx + \int_a^b \lambda Q\sigma[f - (P + \lambda Q)]\, dx \\ &> \int_a^b | f - (P + \lambda Q)|\, dx = \| f - \bar{P}\|_1 , \end{aligned}$$

which contradicts our assumption on P.

To prove the last assertion of the theorem, we assume that the left hand side of (4.2), for $Q = P$, equals some $\delta \neq 0$. Since

$$\int_a^b P\sigma[f - (1 + \lambda)\, P]\, dx \to \int_a^b P\sigma[f - P]\, dx$$

as $\lambda \to 0$, we may choose λ ($| \lambda | < 1$) so that

$$\lambda \int_a^b P\sigma[f - (1 + \lambda)\, P]\, dx \geqslant \lambda\delta/2 > 0.$$

Then $(1 + \lambda)\, P \in \mathscr{P}$, yet the same calculation as above shows that $\| f - P \|_1 > \| f - (1 + \lambda)\, P \|_1$, a contradiction.

Remarks. The theorem also holds if it is only assumed that $f \in L_1[a, b]$. The proof is word for word the same.

One may wish to improve the theorem in analogy to the corresponding theorem for unrestricted polynomials. That is, one could try to replace the inequality in (4.1) by equality and also try to prove the converse. However, the example $n = 1$, $\mathscr{P} = \mathscr{P}(1; 1)$, $[a, b] = [0, 1]$ and $f(x) = 1 - x$ proves the first conjecture to be false. In this case, $P(x) = 1/2$ is clearly the polynomial of best approximation. Since $\sigma(f-P)=1$ for $0 \leqslant x < 1/2$ and $\sigma(f-P)=-1$ for $1/2 < x \leqslant 1$, $Q(x)=x$ satisfies

$$\int_0^1 Q\sigma(f-P)\,dx < 0.$$

That the converse is false can also be seen by considering this example. Any polynomial P_1 in $\mathscr{P}$ with $P_1(1/2)=1/2$ satisfies $\sigma(f-P_1)=\sigma(f-P)$. Thus, by Theorem 4.1, (4.1) holds. Yet P_1 is not necessarily a polynomial of best L_1 approximation to f.

Let $B_j = B_j(P)$, for $P \in \mathscr{P}$, be defined as before. The following corollary, which is a slight improvement of Theorem 4.1, proves to be very useful.

COROLLARY 4.2. *If $P \in \mathscr{P}$ is a polynomial of best L_1 approximation to $f \in C[a, b]$ in $\mathscr{P}$, and if $f-P \neq 0$ a.e., then*

$$\int_a^b Q(x)\,\sigma[f(x)-P(x)]\,dx \leqslant 0 \tag{4.3}$$

for all polynomials Q of degree not exceeding n satisfying

$$\epsilon_j Q^{(k_j)}(y) \geqslant 0, \qquad y \in B_j\,, \quad j=1,\ldots,p. \tag{4.4}$$

Proof. Suppose the conclusion is false. That is, suppose that there is a polynomial Q of degree not exceeding n which satisfies (4.4) but for which

$$\int_a^b Q\sigma[f-P]\,dx = \delta > 0.$$

Let P_1 be some polynomial of degree not exceeding n for which

$$\epsilon_j P_1^{(k_j)}(x) > 0, \qquad x \in [a, b], \quad j=1,\ldots,p.$$

That such polynomials exist is shown in [1]. Let $\bar{Q} = Q + \mu P_1$. Clearly, for some $\mu > 0$ sufficiently small, we have

$$\int_a^b \bar{Q}\sigma(f-P)\,dx \geqslant \delta/2 > 0.$$

The calculations in the proof of Theorem 4.1 can be used to show that $\|f-P\|_1 > \|f-(P+\lambda\bar{Q})\|_1$ as soon as $\lambda > 0$ is sufficiently small. We reach a contradiction if we can show that $P+\lambda\bar{Q} \in \mathscr{P}$ for some λ.

Let l_j, e_j, y_{ji} be the numbers and points associated with $B_j(P)$, $j = 1,..., p$, as before. Since Q satisfies (4.4), we know that

$$\epsilon_j \bar{Q}^{(k_j)}(y_{ji}) > 0, \qquad j = 1,..., p, \quad i = 1,..., l_j .$$

Thus, for each j, there is an open neighborhood 0_j of B_j such that

$$\epsilon_j \bar{Q}^{(k_j)}(y_{ji}) > 0, \qquad x \in 0_j .$$

Clearly,

$$\epsilon_j (P + \lambda \bar{Q})^{(k_j)}(x) > 0, \qquad x \in 0_j , \quad j = 1,..., p, \tag{4.5}$$

for all $\lambda > 0$. Since $P^{(k_j)} > 0$ on $[a, b] - 0_j$, we have

$$0 < \beta_j = \min_{x \in [a,b]-0_j} \{P^{(k_j)}(x)\}.$$

If we choose λ so small that

$$\lambda \sup_{x \in [a,b]} | \bar{Q}^{(k_j)}(x)| < \beta_j , \qquad j = 1,..., p,$$

then

$$\epsilon_j (P + \lambda \bar{Q})^{(k_j)}(x) > 0, \qquad x \in [a, b] - 0_j . \tag{4.6}$$

Combining (4.5) and (4.6), we see that $P + \lambda \bar{Q} \in \mathscr{P}$ if λ, μ are chosen as above.

Remark. This corollary remains valid for $f \in L_1[a, b]$.

Instead of $A(P, f)$, the set which is relevant for L_1 approximation is $D(P, f)$, the set of points where $f - P$ changes sign. If $g \in C[a, b]$ is nonzero a.e. in $[a, b]$, we say that g changes sign m times in $[a, b]$ if there exist m points $a < x_1 < \cdots < x_m < b$ for which g is either nonnegative or non-positive in each of the intervals $[a, x_1]$, $[x_1, x_2]$,..., $[x_m, b]$, the signs $\geqslant$, $\leqslant$ alternating from each interval to its immediate neighbor. We say that g changes sign on $\{x_i\}$. $m = 0$ means g is always $\geqslant 0$ or always $\leqslant 0$ in $[a, b]$. If no such m exists, we set $m = \infty$. If $f \in C[a, b]$, $P \in \mathscr{P}$ and $f - P \neq 0$ a.e. in $[a, b]$, $m = m(P, f)$ will henceforth be the number of times $f - P$ changes sign in $[a, b]$ and, if $0 < m < \infty$, $D(P, f) = \{x_1 ,..., x_m\}$ will denote the set of points where $f - P$ changes sign.

LEMMA 4.3. *Let $f \in C[a, b]$. If there is more than one polynomial in $\mathscr{P}$ of best L_1 approximation to f, then there exists among them a polynomial P_0,* *called a minimal polynomial, such that* $f - P_0 \neq 0$ a.e. *in* $[a, b]$ *and such that if* $P \in \mathscr{P}$ *is any other polynomial of best* L_1 *approximation to* f, *then*

$$B_j(P_0) \subset B_j(P), \qquad j = 1,..., p.$$

Also $\deg(P_0) \geqslant \deg(P)$ *and, if* $m(P_0, f)$ *is finite, then* $f - P$ *is zero on* $D(P_0, f)$.

Proof. Let F be the set of all polynomials in $\mathscr{P}$ which are of best approximation to f. Let

$$B_i = \bigcap_{P \in F} B_i(P), \qquad i = 1,..., p.$$

It is clear that each B_i is either the entire interval $[a, b]$ or has only a finite number of points. In either case, there exist a finite number of polynomials $P_{i,1},..., P_{i,r_i} \in F$ such that

$$\bigcap_{j=1}^{r_i} B_i(P_{ij}) = B_i, \qquad i = 1,..., p.$$

Let

$$P_1 = \left(\sum_{i=1}^{p} r_i\right)^{-1} \sum_{i=1}^{p} \sum_{j=1}^{r_i} P_{ij}.$$

Clearly, $P_1 \in \mathscr{P}$. By convexity, $P_1 \in F$. Also

$$B_i(P_1) = B_i \subset B_i(P), \qquad P \in F, \quad i = 1,..., p. \tag{4.7}$$

If $\deg(P_1) = \max_{P \in F}[\deg(P)]$, let $P_2 \in F$ be arbitrary, but distinct from P_1. If $\deg(P_1) < \max_{P \in F}[\deg(P)]$, let $P_2 \in F$ be such that $\deg(P_2) = \max_{P \in F}[\deg(P)]$. Since there can be only a countable number of polynomials P for which $f - P = 0$ on a given set of nonzero measure, there is a λ, $0 < \lambda < 1$, for which $P_0 = \lambda P_1 + (1 - \lambda) P_2$ satisfies $f - P_0 \neq 0$ a.e. in $[a, b]$. Moreover, since there is at most one possible choice of λ for which $\deg[\lambda P_1 + (1 - \lambda) P_2] \neq \max_{P \in F}[\deg(P)]$, we may choose λ so that $\deg(P_0) = \max_{P \in F}[\deg(P)]$. Then $P_0 \in F$, and (4.1) together with the positivity of λ imply that

$$B_i(P_0) = B_i \subset B_i(P), \qquad P \in F, \qquad i = 1,..., p.$$

Let $P \in F$. Then

$$\int_a^b |f - P_0| \, dx + \int_a^b |f - P| \, dx = \int_a^b |f - P_0 + f - P| \, dx$$

implies that $\sigma(f - P_0) = \sigma(f - P)$ a.e. in $[a, b]$ and, hence, by continuity, everywhere in $[a, b]$. Here equality is meant in the sense that

$$[\sigma(f - P_0)] \cdot \sigma(f - P)(x) \geqslant 0.$$

The rest of the Lemma follows from this equality.

5. An Intermediate Lemma

We prove now the main lemma needed for the uniqueness theorem. As in the case of the uniform norm, we introduce, for each $f \in C[a, b]$ and each $P \in \mathscr{P}$ of best L_1 approximation to f, an incidence matrix $I(P, f)$. Let ν be the degree of P. Assume that $m(P, f)$ is finite. For $k_j \leqslant \nu$, the sets B_j and, hence, the corresponding numbers e_j, l_j are finite. We define $I(P, f)$ to be the incidence matrix corresponding to the conditions

$$\begin{aligned} Q(x_i) &= \alpha_i, \qquad i = 1,..., m(P, f), \\ Q^{(k_j)}(y_{ji}) &= \beta_{ji}, \qquad k_j \leqslant \nu, \quad i = 1,..., l_j, \\ Q^{(k_j+1)}(y_{ji}) &= \gamma_{ji}, \qquad a < y_{ji} < b, \quad k_j + 1 \leqslant \nu, \quad i = 1,..., l_j, \\ Q(x_0) &= \alpha_0, \end{aligned} \tag{5.1}$$

where $x_0 \in (a, x_1)(x_0 \in (a, b)$ if $m(P, f) = 0)$ and the α_i, β_{ji}, γ_{ji}, α_0 are arbitrary reals. $I(P, f)$ has $\nu + 1$ columns.

We define N by

$$N = m + \sum_{k_j \leqslant \nu} l_j + \sum_{k_j+1 \leqslant \nu} (l_j - e_j), \tag{5.2}$$

where $m = m(P, f)$.

Lemma 5.1. *Let $P \in \mathscr{P}$ be a polynomial of best L_1 approximation to $f \in C[a, b]$ such that $f - P \neq 0$* a.e. *and such that $m = m(P, f)$ is finite. If $I(P, f)$ and N are defined as above, then $I(P, f)$ satisfies the Atkinson–Sharma condition and the strong Pólya conditions. $I(P, f)$ has exactly $N + 1$ nonzero entries, where*

$$N = m + \sum_{k_j \leqslant \nu} (2l_j - e_j). \tag{5.3}$$

Moreover, $N \geqslant \nu + 1$.

Proof. That (5.3) holds and that $I(P, f)$ has exactly $N + 1$ nonzero entries follow exactly as in the proof of Lemma 2.2, since these are properties of the class $\mathscr{P}$.

If we omit the last condition of (5.1), the resulting incidence matrix satisfies the Atkinson–Sharma condition since it has exactly the same form as $E(P, f)$ of Lemma 2.2. Since the addition of the omitted condition can neither create a supported sequence of odd length nor cause a maximal sequence of odd length to be supported, $I(P, f)$ satisfies the Atkinson–Sharma condition.

With m_j defined as before, we want to prove that

$$\sum_{j=0}^{k} m_j \geqslant k + 2, \qquad k = 0, 1, \ldots, \nu - 1. \tag{5.4}$$

Clearly $m \geqslant k_1$, since there always exists a polynomial R of degree m which alternates sign at the x_i, $i = 1, \ldots, m$ and for which

$$\int_a^b R\sigma(f - P)\, dx > 0.$$

If $m \leqslant k_1 - 1$, then $R \in \mathscr{P}$, in contradiction to Theorem 4.1. Since $m_0 = m + 1$, (5.4) is satisfied for $k = 0, 1, \ldots, k_1 - 1$.

Now we assume that (5.4) is false and that $\bar{k}$ is the smallest integer k for which (5.4) is violated. Then necessarily $k_1 \leqslant \bar{k} \leqslant \nu - 1$ and

$$\sum_{j=0}^{\bar{k}} m_j = \bar{k} + 1.$$

Moreover, (5.4) is satisfied for $0 \leqslant k \leqslant \bar{k} - 1$, and $m_{\bar{k}} = 0$.

Let $\bar{I}(P, f)$ be the incidence matrix consisting of the columns of $I(P, f)$ numbered 0 through $\bar{k}$. Since $m_{\bar{k}} = 0$, no maximal sequence of $I(P, f)$ can cross this column. Hence $\bar{I}(P, f)$ satisfies the Atkinson–Sharma condition. Since (5.4) is satisfied for $0 \leqslant k \leqslant \bar{k} - 1$, $\bar{I}(P, f)$ also satisfies the strong Pólya conditions. Since $\bar{I}(P, f)$ has $\bar{k} + 1$ columns and exactly $\bar{k} + 1$ nonzero entries, we may pose the following B.I.P. for polynomials Q of degree not exceeding $\bar{k}$:

$$\begin{aligned} Q(x_i) &= 0, \qquad i = 1, \ldots, m, \\ Q^{(k_j)}(y_{ji}) &= 0, \qquad k_j \leqslant \bar{k}, \quad i = 1, \ldots, l_j, \\ Q^{(k_j+1)}(y_{ji}) &= 0, \qquad a < y_{ji} < b, \quad k_j + 1 \leqslant \bar{k}, \quad i = 1, \ldots, l_j, \\ Q(x_0) &= \sigma_0, \end{aligned} \tag{5.5}$$

where σ_0 is the sign of $f - P$ in (a, x_1). By the Atkinson–Sharma theorem, this problem has a unique nonzero solution Q. We shall show that Q violates Corollary 4.2.

To do this, we first prove that Q has only the zeros $x_1, ..., x_m$ in (a, b) and that these zeros are all of odd multiplicity. The proofs of both of these claims are based on the same idea. If either of them did not hold, then Q would be the unique solution of a homogeneous B.I.P. and, hence, identically zero. In view of (5.5), this is impossible.

In fact, suppose that $Q(\bar{x}) = 0$ for some $\bar{x} \in (a, b)$, $\bar{x} \neq x_i$, $i = 1, ..., m$. Then Q satisfies (5.5) with the last condition replaced by $Q(\bar{x}) = 0$. Let I_1 be the incidence matrix corresponding to these conditions. From previous arguments it is clear that I_1 is free since we have replaced x_0 by $\bar{x}$. So Q is identically zero.

Suppose now that Q has a zero of even multiplicity at one of the x_i, say x_r. Taking into account that the second and third conditions of (5.5) always come in pairs (since $m_{\bar{k}} = 0$), there must be integers $s \geqslant 1$ for which $Q^{(s)}(x_r) = 0$ and for which this equality does not appear in (5.5). We let t be the smallest such s. Then, necessarily, $Q^{(k)}(x_r) = 0$ for $0 \leqslant k \leqslant t$. Since Q is not identically zero, $t \leqslant \bar{k} - 1$. We consider now a new incidence matrix. Let I_2 be the incidence matrix corresponding to (5.5), with the last condition replaced by $Q^{(t)}(x_r) = 0$.

I_2 satisfies the Atkinson–Sharma condition since the entry corresponding to the new condition does not create a supported sequence of odd length, and does not cause a previously unsupported sequence to be supported. We cannot prove that I_2 satifies the strong Pólya conditions. However, we can prove that it satisfies the Pólya conditions, and this suffices.

If $\overline{m}_j$ is the sum of the entries of the column numbered j of I_2, then

$$\sum_{j=0}^{k} \overline{m}_j \geqslant k + 1, \qquad k = 0, 1, ..., \bar{k} - 1, \tag{5.6}$$

since $\bar{I}(P, f)$ satisfies the strong Pólya conditions. Since the total number of nonzero entries of I_2 is $\bar{k} + 1$, (5.6) also holds for $k = \bar{k}$ and, so, I_2 is free. By the Atkinson–Sharma theorem, Q is identically zero.

We see that Q has only the zeros $x_1, ..., x_m$ in (a, b) and that these are of odd multiplicity. It follows that Q alternates sign at the x_i. Since $\sigma(Q)$ agrees with $\sigma(f - P)$ in (a, x_1), the signs agree a.e. and, hence,

$$\int_a^b Q\sigma(f - P)\, dx > 0.$$

Since Q is a polynomial of degree not exceeding $\bar{k}$, $Q^{(k)} \equiv 0$ for $k \geqslant \bar{k} + 1$. This, together with the second equality of (5.5) and the above inequality, violate Corollary 4.2.

This proves that $I(P, f)$ satisfies the strong Pólya conditions.

Finally, we prove that $N \geqslant \nu + 1$. This follows immediately from the previous proof if we take $\bar{k} = \nu$. Under the assumption $N \leqslant \nu$, we would obtain, in contradiction to Corollary 4.2, a polynomial Q of degree not exceeding $\bar{k} = \nu$ which satisfies (4.3) and (4.4).

6. The Uniqueness Theorem For L_1

By means of the lemmas of the previous section, the uniqueness theorem for the L_1 norm can now be proved.

THEOREM 6.1. *Let $f \in C[a, b]$. Then among all polynomials in $\mathscr{P}$ there is exactly one which is of best L_1 approximation to f.*

Proof. We must show that there is no more than one such polynomial. We assume that there is more than one. In accordance with Lemma 4.3, we let P_0 be a minimal polynomial and set $\nu = \deg(P_0)$. Henceforth, we let $D = D(P_0, f)$, $B_j = B_j(P_0)$ and let x_i, y_{ji}, m, l_j, e_j be the corresponding points and numbers for $k_j \leqslant \nu$.

First we consider the possibility $m = \infty$. We claim that there then exist an infinite number of points x in (a, b) such that $f - P_0$ takes on values of opposite signs in every neighborhood of x. Moreover, this in turn implies that every polynomial in $\mathscr{P}$ of best L_1 approximation to f is identical to P_0.

To prove the first claim, we assume that there are only a finite number $a < z_1 < \cdots < z_s < b$ of such points. (Our method takes care, also, of the possibility that there are no such points at all). Since $m = \infty$, $f - P_0$ does not have a constant sign in each of the intervals $(a, z_1), \ldots, (z_s, b)$. Thus, for some j, there are points y_1, y_2 $(z_j < y_1 < y_2 < z_{j+1})$ at which $f - P_0$ takes on values of opposite signs. Assume that $\sigma[(f - P_0)(y_1)] = -1$ and let

$$t = \sup\{x \mid x \in (y_1, y_2), (f - P_0)(x) < 0\}.$$

Then $t < y_2$, $(f - P_0)(x) \geqslant 0$ for $x \in (t, y_2)$ and there is a sequence $t_1, t_2, \ldots$ converging to t, for which $(f - P_0)(t_i) < 0$. Thus, $f - P_0$ takes on values of opposite signs in each neighborhood of t. We reached a contradiction since $z_1, \ldots, z_s$ were assumed to be the only such points.

To prove the second claim, we use the equality

$$\sigma(f - P_0) = \sigma(f - P)$$

in the proof of Lemma 4.3, where $P \in \mathscr{P}$ is any polynomial of best L_1 approximation to f. By the continuity of $f - P$ and by the above equality, $(f - P)(t) = 0$ for each point t of the above type. Since there are an infinite number of such points and since also $f - P_0 = 0$ at such points, P and P_0 agree at an infinite number of points. Hence P is identical with P_0, contradicting the assumption that there is more than one polynomial in $\mathscr{P}$ of best L_1 approximation to f.

We may thus assume that m is finite. Let $P \in \mathscr{P}$ be any polynomial ($\neq P_0$) of best L_1 approximation to f. Then, by Lemma 4.3, $\deg(P) \leqslant \nu$, $B_j \subset B_j(P)$ and $P = P_0$ on D. Then $S = P_0 - P$ satisfies:

$$\begin{aligned} S(x_i) &= 0, \qquad i = 1,..., m, \\ S^{(k_j)}(y_{ji}) &= 0, \qquad k_j \leqslant \nu, \quad i = 1,..., l_j, \\ S^{(k_j+1)}(y_{ji}) &= 0, \qquad y_{ji} \in (a, b), \quad k_j \leqslant \nu, \qquad i = 1,..., l_j. \end{aligned} \tag{6.1}$$

The last equality holds due to Lemma 2.1.

Let N be as in (5.2). Let E be the incidence matrix with N columns corresponding to the conditions (6.1). From Lemma 5.1 it is easy to conclude that E has exactly N nonzero entries, that E satisfies the Atkinson–Sharma condition, that E satisfies the Pólya conditions and that $N \geqslant \nu + 1$. Thus we have a B.I.P. whose incidence matrix is free. It follows that any polynomial of degree not exceeding $N - 1$ which satisfies (6.1) must be identically zero. Since $\deg(S) \leqslant \nu \leqslant N - 1$, S is identically zero and, so, P and P_0 coincide. Thus there is exactly one polynomial in $\mathscr{P}$ which is of best L_1 approximation to f.

Remark. This proof is valid not only for the L_1 norm but also for any $L_1(\rho)$ norm, where $\rho > 0$ a.e. in $[a, b]$. Thus, the following theorem holds.

THEOREM 6.2. *Let $f \in C[a, b]$. Let $\rho \in L_1[a, b]$ and be > 0 a.e. in $[a, b]$. Then among all polynomials in $\mathscr{P}$ there is exactly one which approximates f best in the $L_1(\rho)$ norm. That is, there is exactly one which minimizes*

$$\int_a^b | f - P | \rho \, dx.$$

Proof. Replace dx by $\rho\, dx$ in the proof of Theorem 6.1.

Remark. A similar proof can be carried out for more general Banach spaces of functions. This is a topic to which the author plans to return.

ACKNOWLEDGMENT

The author extends thanks to G. G. Lorentz for suggestions which helped him to substantially shorten some of the proofs.

References

1. G. G. Lorentz and K. L. Zeller, Monotone approximation by algebraic polynomials, *Trans. Amer. Math. Soc.* **149** (1970), 1–18.
2. I. J. Schoenberg, On Hermite-Birkhoff interpolation. *J. Math. Anal. Appl.* **16** (1966), 538–543.
3. K. Atkinson and A. Sharma, A partial characterization of poised Hermite–Birkhoff interpolation problems. *SIAM J. Numer. Anal.* **6** (1969), 230–235.