

MONOTONE APPROXIMATION BY ALGEBRAIC POLYNOMIALS

BY

G. G. LORENTZ⁽¹⁾ AND K. L. ZELLER⁽²⁾

Abstract. A given real continuous function f on $[a, b]$ is approximated by polynomials P_n of degree n that are subject to certain restrictions. Let $1 \leq k_1 < \dots < k_p \leq n$ be given integers, $\varepsilon_i = \pm 1$, given signs. It is assumed that $P_n^{(k_i)}(x)$ has the sign of ε_i , $i = 1, \dots, p$, $a \leq x \leq b$. Theorems are obtained which describe the polynomials of best approximation, and (for $p = 1$) establish their uniqueness. Relations to Birkhoff interpolation problems are of importance. Another tool are the sets A , where $|f(x) - P_n(x)|$ attains its maximum, and the sets B_i with $P_n^{(k_i)}(x) = 0$. Conditions are discussed which these sets must satisfy for a polynomial P_n of best approximation for f . Numbers of the points of sets A , B_i are studied, the possibility of certain extreme situations established. For example, if $p = 1$, $k_1 = 1$, $n = 2q + 1$, it is possible that $|A| = 3$, $|B| = n$.

1. Introduction. This paper deals with approximation, in the uniform norm, of real continuous functions on an interval $[a, b]$ by polynomials P_n of degree n , $P_n(x) = \sum_{k=0}^n a_k x^k$ (where $a_n = 0$ is permitted).

If f increases on $[a, b]$, its polynomial of degree n of best approximation is not necessarily increasing. For example (Roulier [10]), let $f(x) = x^{4k+1}$ on $[-1, +1]$, $k = 1, 2, \dots$. Then the polynomial of best approximation of degree $4k$ for f is $P_{4k}(x) = x^{4k+1} - 2^{-4k} C_{4k+1}(x)$, where C_n is the Chebyshev polynomial, and it is easy to see that $P'_{4k}(0) = -C'_{4k+1}(0) < 0$. It is therefore an interesting problem to find, for a given function f , its polynomial of best approximation among all increasing polynomials, to find the degree of approximation of f by such polynomials, and so on. We will dispense with the assumption that f itself is increasing, since it is not essential for the results of this paper.

More general is the problem of the approximation of a given continuous function f on $[a, b]$, for a given $k = 1, 2, \dots$, by polynomials P_n of degree n that satisfy $P_n^{(k)}(x) \geq 0$, $a \leq x \leq b$. Still more general is the problem, for given integers $1 \leq k_1$

Received by the editors June 6, 1969 and, in revised form, September 22, 1969.

AMS Subject Classifications. Primary 4140; Secondary 4110.

Key Words and Phrases. Monotone approximation, polynomials of best approximation, admissible sets, minimal polynomials, basic polynomials of Lagrange interpolation, Birkhoff interpolation, free incidence matrices, poised incidence matrices.

⁽¹⁾ Supported by a grant from OSR, U.S. Air Force.

⁽²⁾ A considerable part of the work of the second author has been carried out during his visits to the University of Texas, in 1968 and 1969.

Copyright © 1970, American Mathematical Society

$\dots < k_2 < \dots < k_p$ and signs $\varepsilon_i = \pm 1$, $i = 1, \dots, p$, to approximate f by polynomials P_n satisfying the conditions

$$(1.1) \quad \varepsilon_i P_n^{(k_i)}(x) \geq 0, \quad a \leq x \leq b, \quad i = 1, \dots, p.$$

For a given n , we denote this problem (as well as the class of the polynomials P_n that satisfy (1.1)) by $\mathcal{P} = \mathcal{P}(k_1, \dots, k_p; \varepsilon_1, \dots, \varepsilon_p)$. We write \mathcal{P}_k for the special case $p = 1$, $k_1 = k$, $\varepsilon_1 = +1$. In analogy to the case $k = 1$, all these problems will be called problems of monotone approximation.

In the present paper we deal with those aspects of the problem \mathcal{P} (and, more specifically, of \mathcal{P}_k), which, for ordinary approximation, are answered by the theorems of Chebyshev. In our situation, these questions are decidedly more difficult. §2 of our paper deals with the characterization of the polynomials of best approximation P_n . The points $x_1, \dots, x_\mu; y_1, \dots, y_\lambda$ of Theorem 4, which should be regarded as a generalization of the Chebyshev alternance, are $\mu + \lambda \leq n + 2$ in number (the inequality can occur) and depend on the behavior of P_n and of $P_n^{(k)}$. In §3 we prove the uniqueness of the polynomial of best approximation for the problem \mathcal{P}_k , in §§4–5 we discuss possible sets of points x_i , y_i , and possible values of μ , λ . Our problems are related to different interpolation problems: Lagrange interpolation (§2), Hermite interpolation (§5), but most intimately, to the general Birkhoff interpolation problem [2]. Actually, the solvability of this problem under certain conditions (Atkinson and Sharma [1]) is essential for the proof of uniqueness in §3.

Several authors have dealt with the problems of monotone approximation: Shisha [13], Roulier [10], [11], the present authors [4], [5], [6]. The papers [13], [11], [4], [5] discuss the degree of approximation, while the paper [6] deals with the problem \mathcal{P}_1 and can be regarded as an introduction to the present paper. See also Rice [9].

We return to the general problem \mathcal{P} . If $n \geq k_p$, as we shall always assume, there exist polynomials q that satisfy (1.1) in the strict sense: $\varepsilon_i q^{(k_i)}(x) > 0$ on $[a, b]$, $i = 1, \dots, p$. For the proof we take

$$(1.2) \quad q(x) = A_p(x-a)^{k_p} + A_{p-1}(x-a)^{k_p-1} + \dots + A_1(x-a)^{k_1},$$

where $A_p = \varepsilon_p$, and the A_i have the sign of ε_i , $i = 1, \dots, p-1$. If we select A_{p-1} , A_{p-2} , ... inductively taking them sufficiently large, we will have the required inequalities. Compactness arguments show that a polynomial $P \in \mathcal{P}$ of best approximation for $f \in C[a, b]$ exists; and all P with this property form a compact and convex subset of $C[a, b]$. The following theorem has been given by Roulier [10] for the problem \mathcal{P}_1 :

THEOREM 1. *Let a continuous function f and a polynomial $P \in \mathcal{P}$ be given, with P not identically equal to f . Then P is a polynomial of best approximation for f in the class \mathcal{P} if and only if*

$$(1.3) \quad \max_{x \in A} [f(x) - P(x)]Q(x) \geq 0,$$

where A is the set

$$A = A(f, P) = [x : |f(x) - P(x)| = \|f - P\|],$$

for each polynomial Q of degree n that satisfies the conditions

$$(1.4) \quad \varepsilon_i [P^{(k_i)}(x) - Q^{(k_i)}(x)] \geq 0, \quad a \leq x \leq b, \quad i = 1, \dots, p.$$

We omit the proof, which is similar to the proof of the well-known theorem of Kolmogorov [3, p. 18].

On the set $A = A(f, P)$ we define the function

$$(1.5) \quad \sigma(x) = \text{sign} [f(x) - P(x)],$$

which is continuous on A . Conversely, if a polynomial $P \in \mathcal{P}$, a compact set $A \subset [a, b]$ and a continuous sign function σ on A are given, there exists a function $f \in C[a, b]$ for which $A = A(f, P)$.

For the proof, we take $\delta > 0$ arbitrarily and put $g(x) = \delta \sigma(x)$ on A . The complement of A is a union of countably many intervals (α_k, β_k) . We put $g(x) = 0$ on $[\alpha_k + \delta_k, \beta_k - \delta_k]$, taking the $\delta_k > 0$ so small, that $\delta/\delta_k > \|P'\|$ and $\alpha_k + \delta_k < \beta_k - \delta_k$, $k = 1, 2, \dots$, and extend g linearly onto the intervals $(\alpha_k, \alpha_k + \delta_k)$, $(\beta_k - \delta_k, \beta_k)$. Then $f = P + g$ will have the required properties. The question remains whether one can take f here to be continuously differentiable with the properties $\varepsilon_i f^{(k_i)}(x) \geq 0$, $a \leq x \leq b$.

With each $P \in \mathcal{P}(k_1, \dots, k_p; \varepsilon_1, \dots, \varepsilon_p)$ we associate the sets

$$(1.6) \quad B_i = B_i(P) = [x : P^{(k_i)}(x) = 0], \quad i = 1, \dots, p.$$

For the problem \mathcal{P}_k we have one single set B . A useful remark is that, for $P \in \mathcal{P}$,

$$(1.7) \quad P^{(k_i+1)}(x) = 0, \quad x \in B_i, \quad x \neq a, b.$$

(Otherwise $P^{(k_i)}$ would change sign at x .) A collection of sets B_1, \dots, B_p is possible for the problem \mathcal{P} if there exists a $P \in \mathcal{P}$ with $B_i = B_i(P)$, $i = 1, \dots, p$. This condition restricts the sets B_i considerably. For example, if $p = 2$, $k_1 = 1$, $k_2 = 2$, then B_1 can only be either a one point set, or the interval $[a, b]$.

For the problem \mathcal{P}_k , a set $B \subset [a, b]$, $B \neq [a, b]$, is possible exactly when

$$(1.8) \quad 2l - e \leq n - k,$$

where l is the number of points in B , and e the number of points of B among a, b .

For the proof, let $2l - e > n - k$, then $P^{(k)}$ has more than $n - k$ roots ($l - e$ double roots at the points of B inside (a, b) , and in addition e single roots), so that $P^{(k)} \equiv 0$ and $B = [a, b]$. Conversely, if a set $B \subset [a, b]$ satisfies (1.8) and consists of the points $x_1 < \dots < x_l$ in $[a, b]$, we take $P(x)$ to be the k th indefinite integral of

$$(x - x_1)^2(x - x_2)^2 \cdots (x - x_l)^2.$$

If $x_1 = a$, we replace $(x - x_1)^2$ by $(x - a)$ in this product; if $x_l = b$, we replace $(x - x_l)^2$ by $(b - x)$.

In particular, if B consists just of one point y , then B is possible if and only if $a < y < b$ and $k \leq n - 2$, or if $y = a$ or $y = b$ and $k \leq n - 1$.

2. Characterization of polynomials of best approximation. We wish to improve Theorem 1. For this purpose, we will try to replace condition (1.4) of this theorem, we call it also (a), by a less restrictive condition. If the sets B_i are defined by (1.6), we consider, instead of (a), the conditions

- (a') $\varepsilon_i[P^{(k_i)}(x) - Q^{(k_i)}(x)] > 0, \quad a \leq x \leq b, \quad i = 1, \dots, p,$
 (b) $\varepsilon_i Q^{(k_i)}(x) \leq 0, \quad x \in B_i, \quad i = 1, \dots, p,$
 (b') $\varepsilon_i Q^{(k_i)}(x) < 0, \quad x \in B_i, \quad i = 1, \dots, p.$

Let A , or A' , \dots denote the following statements: There exists a polynomial Q of degree n that satisfies (a), or (a'), \dots , respectively, and violates (1.3), that is, satisfies the condition

$$(2.1) \quad [f(x) - P(x)]Q(x) < 0, \quad x \in A.$$

It is immediately clear that $A' \Rightarrow A$, $B' \Rightarrow B$. Conversely, $A \Rightarrow A'$: If there exists a Q that satisfies (a) and (2.1), then $\bar{Q} = Q - \lambda q$, where q is any of the polynomials (1.2) with $\varepsilon_i q^{(k_i)}(x) > 0$, $a \leq x \leq b$, $i = 1, \dots, p$, and $\lambda > 0$ is sufficiently small, will satisfy (a') and (2.1). Similarly $B \Rightarrow B'$. It is obvious that $A' \Rightarrow B'$. To prove the converse, assume that Q satisfies (2.1) and (b'). With Q , any polynomial $\bar{Q} = \lambda Q$, $\lambda > 0$ satisfies (2.1). For each $i = 1, \dots, p$ we can find an open set $G_i \supset B_i$ with the property $\varepsilon_i Q^{(k_i)}(x) < 0$, $x \in G_i$. Then (a') holds for \bar{Q} if $x \in G_i$. We select $\lambda > 0$ so small that

$$\lambda \max_{i=1, \dots, p} \|Q^{(k_i)}\| < \min_{x \notin G_i; i=1, \dots, p} |P^{(k_i)}(x)|.$$

Then $\varepsilon_i \bar{Q}^{(k_i)}(x) < \varepsilon_i P^{(k_i)}(x)$, $x \notin G_i$, so that (a') holds for \bar{Q} for all x .

This shows that conditions A , A' , B , B' are all equivalent. As a consequence, we have

THEOREM 2. *A polynomial $P \in \mathcal{P}$ is a polynomial of best approximation in the class \mathcal{P} for a given continuous function f if and only if there does not exist a polynomial Q of degree n for which*

$$(2.2) \quad \sigma(x)Q(x) < 0, \quad x \in A,$$

$$(2.3) \quad \varepsilon_i Q^{(k_i)}(x) < 0, \quad x \in B_i, \quad i = 1, \dots, p.$$

As a corollary we see that the property of a given $P \in \mathcal{P}$ to be or not to be a polynomial of best approximation for f depends only on the set $A = A(P, f)$ and the function $\sigma(x)$ given by (1.5), and not on any other property of f .

Our next step is to reduce the number of points x involved in conditions (2.2) and (2.3) to the Chebyshev number $n+2$, and to connect them with certain linear identities, valid for all polynomials Q of degree n .

One can say that conditions (2.2) and (2.3) or (2.4) and (2.5) below express the impossibility of certain Birkhoff interpolation problems (see [2]), where however not the values of the polynomial Q and of its derivatives are prescribed, but the signs of these values.

THEOREM 3. *A polynomial $P \in \mathcal{P}$ is a polynomial of best approximation for a given continuous function f if and only if there exist points $x_j \in A$, $j=1, \dots, \mu$, $y_{ij} \in B_i$, $j=1, \dots, \lambda_i$, $i=1, \dots, p$ with $\mu + \lambda_1 + \dots + \lambda_p \leq n+2$ for which there is no Q that satisfies*

$$(2.4) \quad \sigma(x_j)Q(x_j) < 0, \quad j = 1, \dots, \mu,$$

$$(2.5) \quad \varepsilon_i Q^{(k_i)}(y_{ij}) < 0, \quad j = 1, \dots, \lambda_i, \quad i = 1, \dots, p,$$

or, equivalently, if and only if there exist such points x_j , y_{ij} and constants $b_j > 0$, $b_{ij} > 0$ for which

$$(2.6) \quad \sum_{j=1}^{\mu} b_j \sigma(x_j) Q(x_j) + \sum_{i=1}^p \varepsilon_i \sum_{j=1}^{\lambda_i} b_{ij} Q^{(k_i)}(y_{ij}) = 0$$

holds for all polynomials Q of degree n .

Proof. In the $n+1$ dimensional space R^{n+1} of points $X = (\xi_1, \dots, \xi_{n+1})$, we associate with each polynomial $Q(x) = a_1 x^n + \dots + a_n x + a_{n+1}$ the linear continuous functional $L(X) = \sum_{i=1}^{n+1} a_i \xi_i$. We consider the points

$$(2.7) \quad \begin{aligned} X &= \sigma(x)(x^n, x^{n-1}, \dots, x, 1), \quad x \in A; \\ Y_i &= \varepsilon_i(n(n-1) \cdots (n-k_i+1)y^{n-k_i}, \dots, k_i!, 0, \dots, 0), \\ &\quad y \in B_i, \quad i = 1, \dots, p. \end{aligned}$$

Let \mathcal{A} be the collection of all these points, then \mathcal{A} is compact in R^{n+1} . The interpolation problem with the given signs is impossible if and only if there is no L for which $L(Z) < 0$ for all $Z \in \mathcal{A}$. This may be replaced by the condition $L(Z) < 0$ for all points Z of the convex hull \mathcal{A}_c of \mathcal{A} . By the theorems of separation of convex sets in R^{n+1} , there is no L with this property, if and only if the origin in R^{n+1} belongs to \mathcal{A}_c . By a theorem of Carathéodory [3, p. 20, Lemma 2], the convex hull \mathcal{A}_c is obtained from \mathcal{A} by taking all convex combinations of all groups of at most $n+2$ points of \mathcal{A} . Thus the interpolation problem is impossible exactly if there exist at most $n+2$ points X_j , Y_{ij} of type (2.7) and numbers $b_j > 0$, $b_{ij} > 0$ such that

$$(2.8) \quad Z_0 = \sum_j b_j X_j + \sum_{ij} b_{ij} Y_{ij} = 0.$$

This is also equivalent to $L(Z_0) = 0$ for all L , that is, to (2.6) for all Q . Since these arguments are invertible, we can also return to the conditions of the form (2.2) and (2.3). This gives the first part of the theorem.

Note that in Theorem 3 we must have

$$(2.9) \quad \mu + (k_1 + 1)\lambda_1 + \dots + (k_p + 1)\lambda_p \geq n+2.$$

For if this sum is $\leq n+1$, then the Hermite interpolation problem is solvable, which assigns arbitrary values to Q at the points x_j and to $Q, Q', \dots, Q^{(k_i)}$ at the points y_{ij} .

As a special case of Theorem 3 we note:

THEOREM 4. *A polynomial P is the polynomial of best approximation for f in the class \mathcal{P}_k if and only if there are points $x_i \in A$, $i = 1, \dots, \mu$, and points*

$$y_j \in B = [x : P^{(k)}(x) = 0], \quad j = 1, \dots, \lambda, \quad \mu + \lambda \leq n + 2,$$

for which there is no Q with the properties

$$(2.10) \quad \sigma(x_i)Q(x_i) < 0, \quad i = 1, \dots, \mu; \quad Q^{(k)}(y_j) < 0, \quad j = 1, \dots, \lambda,$$

or, equivalently, if and only if there exist points x_i , y_j and constants $a_i > 0$, $b_j > 0$ for which

$$(2.11) \quad \sum_{i=1}^{\mu} a_i \sigma(x_i) Q(x_i) + \sum_{j=1}^{\lambda} b_j Q^{(k)}(y_j) = 0$$

holds for all Q of degree n .

EXAMPLE. A polynomial $P_4 \in \mathcal{P}_1$ of best approximation to x^5 on $[-1, +1]$ is $P = cx^3$, where $0 < c < 1$ is so chosen that the negative minimum of $x^5 - cx^3$ on $[0, 1]$ is equal in absolute value to its positive maximum at $x = 1$. In other words, c is the unique positive root of the equation $c - 1 = \frac{2}{5}(\frac{3}{5})^{3/2}c^{5/2}$. For the proof assume that this minimum is taken at $x = a$, $0 < a < 1$. The set A consists of the four points $-1 < -a < a < 1$, B of the point 0, $\sigma(x)$ alternates with $\sigma(-1) = -1$. Since there is a matching identity, of type (2.11), for all polynomials Q of degree 4,

$$-Q(-1) + a^{-3}Q(-a) - a^{-3}Q(a) + Q(1) + 2(a^{-2} - 1)Q'(0) = 0,$$

P is indeed a best polynomial, by Theorem 4.

Restricting ourselves for simplicity to the problem \mathcal{P}_k , we shall show how one can make conditions (2.11) more concrete by the use of the Lagrange polynomials L_j . This works particularly well for small values of λ .

If points $x_1 < \dots < x_\mu$ are given, the basic polynomials of the Lagrange interpolation L_i , $i = 1, \dots, \mu$ are the uniquely defined polynomials of degree $\mu - 1$ with the properties $L_i(x_j) = 0$, $i \neq j$, $L_i(x_i) = 1$. (See [3, p. 36].) Let

$$(2.12) \quad \Pi(x) = (x - x_1) \cdots (x - x_\mu), \quad \Pi_r(x) = x^r \Pi, \quad r = 0, 1, \dots$$

Instead of checking (2.11) for all Q , we can check this condition for all members of a basis of the polynomials of degree n . An example of a basis is given by the polynomials

$$(2.13) \quad L_1, L_2, \dots, L_\mu, \Pi_0, \dots, \Pi_{n-\mu}.$$

(If $\mu = n + 1$, there are no polynomials Π_r .) In fact, each Q can be written in the form $Q = \sum_{i=1}^{\mu} Q(x_i)L_i + R\Pi$, where R is some polynomial of degree $n - \mu$.

THEOREM 5. *A polynomial P is a polynomial of best approximation for f in \mathcal{P}_k if and only if there are $\mu + \lambda \leq n + 2$ points $x_i \in A$, $y_j \in B$ such that:*

- (a) If $\lambda=0$, then $\mu=n+2$ and the signs of $\sigma(x_i)$ alternate (the Chebyshev case).
 (b) If $\lambda \geq 1$, there exist numbers $b_1, \dots, b_\lambda > 0$ for which

$$(2.14) \quad \sum_{j=1}^{\lambda} b_j \Pi_r^{(k)}(y_j) = 0, \quad r = 0, \dots, n-\mu$$

(if $\mu=n+1$, there are no conditions (2.14) and one takes $b_1=1$), and in addition

$$(2.15) \quad \sigma(x_i) \sum_{j=1}^{\lambda} b_j L_i^{(k)}(y_j) < 0, \quad i = 1, \dots, \mu.$$

Proof. (a) If $\lambda=0$, then (2.11) becomes

$$(2.16) \quad \sum_{i=1}^{\mu} a_i \sigma(x_i) Q(x_i) = 0, \quad a_i > 0.$$

If the signs of the $\sigma(x_i)$ do not alternate, there exist at most $\mu-2 \leq n$ points z_k which separate the x_i into at most $n+1$ groups with constant $\sigma(x_i)$ on each group, alternating from group to group. Then $Q(x) = \Pi(x - z_k)$ provides a contradiction to (2.16). This also proves that $\mu=n+2$.

Conversely, if the signs alternate and $\mu=n+2$, we shall select $a_i > 0$ satisfying (2.16). We may assume $\sigma(x_i) = (-1)^i$, and put $a_{n+2} = 1$. Condition (2.16) holds for all Q if it holds for $Q = L_i$, $i = 1, \dots, n+1$, where the L_i are the basic Lagrange polynomials for the knots x_1, \dots, x_{n+1} . For $Q = L_i$, (2.16) becomes

$$a_i = (-1)^{n-i+1} L_i(x_{n+2}),$$

and this is indeed > 0 .

(b) This follows from Theorem 4 by means of the basis (2.13).

A computational way of handling the problem of finding the polynomial of best approximation in \mathcal{P}_k is the following. We first try $\lambda=0$, $\mu=n+2$. If this fails, we try $\lambda=1$, $\mu=n+1$, then $\lambda=1$, $\mu=n$. The case $\lambda=1$, $\mu < n$ is impossible, as we shall see. If this does not work, we try $\lambda=2$, $\mu=n$, then $\lambda=2$, $\mu=n-1$, and so on.

We shall give a rather complete treatment of the case $\lambda=1$. Let $B = \{y\}$.

Case I. $\lambda=1$, $\mu=n+1$. Here there are no Π_r , and with the polynomials L_i for the knots x_1, \dots, x_{n+1} , our conditions become

$$(2.17) \quad \sigma(x_i) L_i^{(k)}(y) < 0, \quad i = 1, \dots, n+1.$$

Case II. $\lambda=1$, $\mu=n$. Here we have to use Π and the Lagrange polynomials for the knots x_1, \dots, x_n . The conditions are

$$(2.18) \quad \Pi^{(k)}(y) = 0, \quad \sigma(x_i) L_i^{(k)}(y) < 0, \quad i = 1, \dots, n.$$

Case III. $\lambda=1$, $\mu < n$ is impossible. Here we have to consider at least Π , Π_1 , and as a part of the condition (2.14) we have $\Pi^{(k)}(y) = \Pi_1^{(k)}(y) = 0$, that is,

$$\Pi^{(k)}(y) = 0, \quad y \Pi^{(k)}(y) + k \Pi^{(k-1)}(y) = 0.$$

It follows that $\Pi^{(k-1)}(y) = 0$, which is impossible, since the zeros of $\Pi^{(k)}$ strictly alternate those of $\Pi^{(k-1)}$.

In case $\lambda=1$, the signs $\sigma(x_i)$ "almost alternate". We shall say that the numbers u_1, u_2, \dots, u_μ form a *semialternance* if none of them vanishes and if they have at least $\mu-2$ changes of sign.

We shall use the following lemma of A. A. Markov:

LEMMA 1. *If L, M are two polynomials of degree n with real distinct roots x_j, y_j that alternate:*

$$(2.19) \quad x_1 \leq y_1 \leq x_2 \leq \dots \leq y_n \quad \text{with } x_j < y_j \text{ for at least one } j,$$

then the roots ξ_j, η_j of the derivatives L', M' are also real, contained in (x_1, y_n) and alternate strictly:

$$(2.20) \quad \xi_1 < \eta_1 < \xi_2 < \dots < \eta_{n-1}.$$

The same applies to the roots of $L^{(k)}, M^{(k)}$, $k < n$.

Since we could not find a proof of this statement with the required strict inequalities in the literature, we supply a simple proof. Let

$$\lambda(x) = \frac{L'}{L}(x) = \sum_{i=1}^n \frac{1}{x-x_i}, \quad \mu(x) = \frac{M'}{M}(x) = \sum_{i=1}^n \frac{1}{x-y_i}.$$

We begin with the following remark. Assume that for some $j=1, \dots, n-1$, $y_j < x_{j+1}$. Then

$$(2.21) \quad \lambda(x) < \mu(x), \quad x \in I = (y_j, x_{j+1}).$$

This follows from the fact that for each $i=1, \dots, n$, $x-x_i \geq x-y_i$, and that both differences are of the same sign. Therefore $1/(x-x_i) \leq 1/(x-y_i)$, with a strict inequality for at least one i .

For a given $j=1, \dots, n-1$, we now prove that $\xi_j < \eta_j$. Without loss of generality we may assume that $y_j < \xi_j$ (otherwise $\xi_j \leq y_j < \eta_j$) and that $\eta_j < x_{j+1}$ (otherwise $\xi_j < x_{j+1} \leq \eta_j$). The interior points ξ_j, η_j of I are the only roots of $\lambda(x)$ and $\mu(x)$, respectively, in this interval. Since $\mu(x) \rightarrow +\infty$ for $x \rightarrow y_j+$, and because of (2.21), we must have $\xi_j < \eta_j$.

If L_1, \dots, L_μ are the basic Lagrange polynomials for the knots $x_1 < \dots < x_\mu$, then it follows from the lemma that the roots $z_1 < z_2 < \dots < z_N$, $N = \mu(\mu-k-1)$ of their derivatives $L_i^{(k)}$ are distributed in (x_1, x_μ) as follows: Smallest is the first root z_1 of $L_\mu^{(k)}$; then comes the first root z_2 of $L_{\mu-1}^{(k)}$, ..., then the first root z_μ of $L_1^{(k)}$, then the second root $z_{\mu+1}$ of $L_\mu^{(k)}$, ..., finally the last root z_N of $L_1^{(k)}$. Also the signs of the derivatives $L_i^{(k)}(x)$ for different positions of x can be simply described. For $x < z_1$, the sign of $L_\mu^{(k)}(x)$ is easily seen to be $(-1)^{\mu-k-1}$; that of $L_{\mu-1}^{(k)}(x)$ is $(-1)^{\mu-k}$; ...; finally that of $L_1^{(k)}(x)$ is $(-1)^k$, so that they alternate. As x moves to the right over the z_k , the signs of the sequence

$$(2.22) \quad L_\mu^{(k)}(x), L_{\mu-1}^{(k)}(x), \dots, L_1^{(k)}(x)$$

change one by one, first the term $L_\mu^{(k)}(x)$, then the second term, and after $L_1^{(k)}(x)$

has changed sign, the first term changes sign again, and so on. As an example, the following table shows the distribution of signs in (2.22) for $\mu=6$, $k=2$:

Interval	L_6''	L_5''	L_4''	L_3''	L_2''	L_1''
$(-\infty, z_1)$	-	+	-	+	-	+
(z_1, z_2)	+	+	-	+	-	+
(z_2, z_3)	+	-	-	+	-	+
(z_3, z_4)	+	-	+	+	-	+
(z_4, z_5)	+	-	+	-	-	+
(z_5, z_6)	+	-	+	-	+	+
(z_6, z_7)	+	-	+	-	+	-
(z_7, z_8)	-	-	+	-	+	-
...
$(z_{18}, +\infty)$	+	-	+	-	+	-

We have therefore:

THEOREM 6. *In case $\lambda=1$ of Theorem 5, the signs $\sigma(x_i)$ form a semialternation.*

For $\lambda=1$, Case II we see that $y \in (x_1, x_\mu)$. In Case I this is not necessarily so. If $x_1 < \dots < x_\mu$ are given, we can select y arbitrarily, different from each of the points z_1, \dots, z_N . Then we select $P \in \mathcal{P}_k$ so that $P^{(k)}$ vanishes only at y , and find the signs $\sigma(x_i)$ by means of (2.17). Then $f \in C$ can be found (see §1) for which P is the polynomial of best approximation in \mathcal{P}_k .

3. Uniqueness of the polynomials of best approximation for the problem \mathcal{P}_k .

As long as we have not proved uniqueness, we must consider the possibility of several polynomials of best approximation for a given problem \mathcal{P} and a given function $f \in C[a, b]$. These polynomials form a compact, convex set \mathcal{B} in C . Among all polynomials of best approximation for f we single out polynomials with the smallest sets $A(f, P)$, $B_i(P)$. We call a polynomial $P_0 \in \mathcal{P}$ of best approximation for f *minimal*, if for any other polynomial $P \in \mathcal{P}$ of best approximation for f , one has

$$(3.1) \quad A(P_0, f) \subset A(P, f), \quad B_i(P_0) \subset B_i(P), \quad i = 1, \dots, p;$$

and if moreover $P(x)$ and $P_0(x)$ coincide on $A(P_0, f)$.

THEOREM 7. *For each $f \in C$ and each problem \mathcal{P} there is a minimal polynomial of best approximation.*

Proof. Let

$$(3.2) \quad A = \bigcap_{P \in \mathcal{B}} A(f, P); \quad B_i = \bigcap_{P \in \mathcal{B}} B_i(P), \quad i = 1, \dots, p.$$

Each of the sets $B_i(P)$ is either identical with $[a, b]$, or finite. There exist therefore

finitely many $P_\nu \in \mathcal{B}$, $\nu=1, \dots, N$ with the property that $B_i = \bigcap_{\nu=1}^N B_i(P_\nu)$, $i=1, \dots, p$. This argument does not apply to the sets A . However, since all sets $A(f, P)$ are compact, there exists a sequence $P_\nu \in \mathcal{B}$, $\nu=N+1, \dots$ for which

$$A = \bigcap_{\nu > N} A(f, P_\nu).$$

Let

$$(3.3) \quad P_0(x) = \sum_{\nu=1}^{\infty} \frac{1}{2^\nu} P_\nu(x).$$

This uniformly convergent series of polynomials of bounded degree can be differentiated termwise, therefore $P_0 \in \mathcal{P}$, and $P_0 \in \mathcal{B}$. Moreover $|f(x) - P_0(x)| = \|f - P_0\|$ can happen for some x only if all differences $|f(x) - P_\nu(x)|$ are equal to $\|f - P_0\| = \|f - P_\nu\|$. Hence $A = A(f, P_0)$, and also $B_i = B_i(P_0)$, $i=1, \dots, p$. Finally if P is another element of \mathcal{B} , then $P(x) = P_0(x)$ on A , for otherwise the polynomial $Q = \frac{1}{2}(P_0 + P)$ would have smaller sets A , B_i than the corresponding sets of P_0 .

Let P be a polynomial of best approximation in the class \mathcal{P}_k for a function f , let $A = A(f, P)$, $B = B(P)$ and let m, l, e denote the numbers of points in the sets A, B and $B \cap \{a, b\}$ (so that $0 \leq e \leq 2$). To prove our main theorem about the uniqueness of the polynomial of best approximation, we shall need some inequalities valid for these numbers and for n, k . One of these is given by (1.8). In addition we have

THEOREM 8. *If P_n is not a polynomial of degree lower than k , that is, if the set B is finite, then*

$$(3.4) \quad m \geq k+2.$$

Proof. Assume that $m \leq k+1$. We put $\Pi(x) = \prod_{j=1}^l (x - y_j)^2$ with the stipulation that the first factor of the product is $x - a$ if $y_1 = a$, and that the last factor is $b - x$ if $y_l = b$. By (1.8), Π is a polynomial of degree $2l - e \leq n - k$. We can find in $[a, b]$ exactly $k+1$ points $z_1 < \dots < z_{k+1}$, among which there are all of the x_i . We define Q by means of a repeated integral

$$(3.5) \quad Q(x) = \int_a^x dt_1 \cdots \int_a^{t_{k-1}} \Pi(t_k) dt_k - R_k(x),$$

where $R_k(x)$, a polynomial of degree k , is so chosen that $Q(z_i) = 0$, $i=1, \dots, k+1$. This can be done by Lagrange interpolation. We now show that the leading coefficient A of R_k is strictly positive. The divided difference of order k , $[z_1, z_2, \dots, z_{k+1}]$ of Q is zero. On the other hand, this difference is a certain repeated integral of the k th derivative of Q [8, p. 16, formula (40)]. The corresponding divided difference of the integral in (3.5) is strictly positive, since its k th derivative, $\Pi(x) > 0$ on $[a, b]$ except for isolated points. The divided difference of R_k is a positive multiple of A . Hence $A > 0$.

Thus Q is a polynomial of degree n , which satisfies $Q(x_i) = 0$, $i=1, \dots, m$, and $Q^{(k)}(x) = \Pi(x) + \text{Const}$, where the constant is negative. Therefore $Q^{(k)}(y_j) < 0$,

$j=1, \dots, l$. If in (2.11) $\lambda > 0$, we have obtained a contradiction with Theorem 4. If, on the other hand, $\lambda = 0$ in (2.11), then by Theorem 5(a), $m \geq \mu = n + 2 \geq k + 2$.

The conclusion of Theorem 8 may be wrong, if the degree of P_n is less than k . If, for example, f is a strictly decreasing function on $[a, b]$, then for each n , its polynomial of best approximation $P_n \in \mathcal{P}_1$ is the constant $\frac{1}{2}[f(a) + f(b)]$. We have $m=2, k=1$.

Later, in Theorem 16, we shall see that $m=3$ is possible for arbitrary large n for the problem \mathcal{P}_1 .

From the inequality (2.9), obtained by Hermite interpolation, it follows that

$$(3.6) \quad m + (k+1)l \geq n+2.$$

A stronger inequality, namely (3.10) below, requires Birkhoff interpolation [2].

This interpolation may be described as follows. (The following terminology is due to Schoenberg [12].) For integers $m=1, 2, \dots, n=1, 2, \dots$ we consider "incidence matrices" $E=(\varepsilon_{ij})$, $i=1, \dots, m, j=0, \dots, n$, with elements ε_{ij} that are 0 or 1. By e we denote the set of pairs (i, j) for which $\varepsilon_{ij}=1$; we assume that the number of elements of e is $n+1$. The matrix E satisfies the *Pólya condition* if

$$(3.7) \quad \sum_{i=1}^m \sum_{j=0}^r \varepsilon_{ij} \geq r+1, \quad r=0, \dots, n.$$

The matrix E is called *free* or *poised*, if for any selection of points $x_1 < x\beta < \dots < x_m$ and real data b_{ij} , $(i, j) \in e$, the Birkhoff interpolation problem

$$(3.8) \quad Q^{(j)}(x_i) = b_{ij}, \quad (i, j) \in e$$

has a solution among all polynomials Q of degree n . If E is free, then it satisfies the Pólya condition [12].

A *maximal sequence* of the matrix E is a sequence of 1's

$$(3.9) \quad \varepsilon_{ij}, \varepsilon_{i,j+1}, \dots, \varepsilon_{i,j+p}$$

which cannot be extended to a longer sequence of 1's. A sequence of 1's (3.9) is said to be *supported*, if there exist pairs $(i_1, j_1), (i_2, j_2) \in e$ such that $i_1 < i < i_2$, and $j_1, j_2 < j$. Atkinson and Sharma [1] proved that a matrix E is free if E satisfies the Pólya condition and if each supported maximal sequence of E has an even number of elements. (For a simple proof of this result see [7].) By means of this theorem we derive

THEOREM 9. *For each polynomial $P_n \in \mathcal{P}_k$ of best approximation, of degree not less than k ,*

$$(3.10) \quad m + 2l - e \geq n + 2.$$

Proof. Assume that (3.10) does not hold, then $m + 2l - e \leq n + 1$. We consider the Birkhoff interpolation problem (for polynomials Q of degree n)

$$(3.11) \quad Q(x_i) = a_{i,0}, \quad Q^{(k)}(y_j) = b_{j,k}, \quad i = 1, \dots, m, \quad j = 1, \dots, l.$$

We add to (3.11) conditions

$$(3.12) \quad Q^{(k-1)}(y_j) = b_{j,k-1}, \quad a < y_j < b$$

with arbitrary data $b_{j,k-1}$, unless $k=1$ and $y_j=x_i$ in which case we take $b_{j,0}=a_{i,0}$. If necessary, we add some conditions of the form

$$(3.13) \quad Q(z_q) = c_q,$$

(with $z_q \neq x_i$, $z_q \neq y_i$) to have a total number of conditions (3.11), (3.12), (3.13) equal to $n+1$. From Theorem 8 it follows that the incidence matrix E , that corresponds to this interpolation problem, satisfies the Pólya condition, and from the theorem of Atkinson and Sharma it follows that E is free. Hence (3.11) has a solution Q for any choice of data. But this contradicts Theorem 2.

We can now prove the main result of this section.

THEOREM 10. *The polynomials of best approximation for the problem \mathcal{P}_k are unique.*

Proof. Let $f \in C[a, b]$, let P be one of its minimal polynomials of best approximation, let P_1 be any other polynomial of best approximation. We shall prove that $D = P - P_1$ is identically zero. Let $A = A(f, P)$, $B = B(P)$. We count the number of zeros of $D^{(k)}$ (taking into account their multiplicity). We know that $D(x)=0$, $x \in A$ and that $D^{(k)}(y)=0$, $y \in B$, also (see (1.7)) that $D^{(k+1)}(y)=0$ if $y \in B$, $y \neq a, b$. If $B=[a, b]$, then $D^{(k)}$ is identically zero. If B is finite, that is, if P is of degree not less than k , we prove the same fact as follows. Let $a \leq z_1 \leq \dots \leq z_p \leq b$ be all zeros of $D^{(k-1)}$ in $[a, b]$. Then by Rolle's theorem, $p \geq m - k + 1$.

We denote by l_1 the number of points of B among the z_i with $a < z_i < b$. Also, let $l(0)$, $l(i)$, $i=1, \dots, p-1$ and $l(p)$ denote the number of points of B in the open intervals (a, z_1) , (z_i, z_{i+1}) , $i=1, \dots, p-1$, and (z_p, b) , respectively.

Among the points a, b there are e points of B . Hence there are at least e zeros of $D^{(k)}$ there. Each point z_i , $a < z_i < b$, which belongs to B , is at least a double zero of $D^{(k)}$. We count here at least $2l_1$ zeros of $D^{(k)}$. Similarly, the intervals (a, z_1) and (z_p, b) contain together at least $2l(0) + 2l(p)$ zeros of $D^{(k)}$. An interval (z_i, z_{i+1}) contains at least $2l(i)$ zeros of $D^{(k)}$. However, between any two adjacent zeros of the polynomial $D^{(k-1)}$ there is an *odd* number of zeros of its derivative. This gives at least $2l(i) + 1$ zeros of $D^{(k)}$ in (z_i, z_{i+1}) .

Altogether, $D^{(k)}$ has at least

$$\begin{aligned} e + 2l_1 + 2l(0) + 2l(p) + \sum_{i=1}^{p-1} [2l(i) + 1] &= e + 2 \left[l_1 + \sum_{i=0}^p l(i) \right] + p - 1 \\ &= e + 2(l - e) + p - 1 \geq m + 2l - e - k \end{aligned}$$

zeros. By Theorem 9, this is $\geq n - k + 2$. Hence $D^{(k)}$ is identically zero in all cases.

It follows that D is a polynomial of degree $k-1$. But D has $m \geq k+2$ zeros, by (3.4). Thus D must vanish identically. This completes the proof.

It is a striking feature of this proof, that in many places of it the required inequalities are much weaker than the complete truth. For Theorem 9, $m \geq k+1$ is sufficient instead of (3.4), and in the last argument the inequalities $m+2l-e \geq n+1$ and $m \geq k$ would suffice in place of (3.10) and (3.4).

REMARK. In the case $k=1$, and when the function f is differentiable, there is a simpler proof of Theorem 10, which does not depend upon Theorem 9 and the Birkhoff interpolation. Let P, P_1 be two polynomials of best approximation, let P be minimal. Let D, A, B be defined as before. We put $A_0 = A \setminus B$, $m_0 = |A_0|$, $e_A = |A_0 \cap \{a, b\}|$, $e_B = |B \cap \{a, b\}|$. Then

$$(3.14) \quad m_0 + 2l \geq n + 2.$$

For otherwise we can, by Hermite interpolation, find a polynomial Q with prescribed values $Q(x)$, $x \in A$, $Q'(y)$, $y \in B$, and the "parasitic" prescribed values $Q(y)$, $y \in A \cap B$. At the points of $B \setminus \{a, b\}$ we have double zeros of D , at the points of $A_0 \setminus \{a, b\}$, $(f-P)'=0$, $(f-P_1)'=0$, hence $D'=(P-P_1)'=0$. Altogether D' has at least

$$2l - e_B + m_0 - e_A \geq m_0 + 2l - 2 \geq n$$

zeros.

4. Possible configurations: sets A, B . We discuss again the problem

$$\mathcal{P}(k_1, \dots, k_p; \varepsilon_1, \dots, \varepsilon_p).$$

If sets A, B_1, \dots, B_p in $[a, b]$ are given, we can ask whether they can be realized as sets $A(f, P), B_1(P), \dots, B_p(P)$ of some function $f \in C[a, b]$ and of its polynomial of best approximation $P \in \mathcal{P}$. We recall the notion of possible sets B_1, \dots, B_p of §1. For simplicity we shall assume that the sets A, B_1, \dots, B_p are finite. We denote the points of A by x_j , $j=1, \dots, m$, the points of B_i by y_{ij} , $j=1, \dots, l_i$, $i=1, \dots, p$. By means of Theorem 3 we obtain:

THEOREM 11. *Sets A, B_i , $i=1, \dots, p$ can be realized in the problem \mathcal{P} if and only if the sets B_1, \dots, B_p are possible and if in addition there is a nontrivial identity*

$$(4.1) \quad \sum_{j=1}^m a_j Q(x_j) + \sum_{i=1}^p \varepsilon_i \sum_{j=1}^{l_i} b_{ij} Q^{(k_i)}(y_{ij}) = 0,$$

with $b_{ij} \geq 0$, and at least one $b_{ij} > 0$, which is valid for all polynomials Q of degree n .

Condition (4.1) can be replaced by the following: There is no polynomial Q of degree n that vanishes at all points x_j , $j=1, \dots, m$, and satisfies

$$(4.2) \quad \varepsilon_i Q^{(k_i)}(y_{ij}) > 0, \quad j=1, \dots, l_i; \quad i=1, \dots, p.$$

Proof. (a) Sufficiency. We define $\sigma(x_j) = \pm 1$, so as to have $a_j \sigma(x_j) \geq 0$, $j=1, \dots, m$.

By §1, there exist $f \in C[a, b]$ and $P \in \mathcal{P}$ for which $A = A(f, P)$, $B_i = B_i(P)$, $i = 1, \dots, p$. Then (4.1) shows that condition (2.6) of Theorem 3 is satisfied.

(b) Necessity. Conversely, if (2.6) holds, we have a nontrivial identity (4.1), perhaps with all $b_{ij} = 0$. In the latter case, since any $n+1$ points $X_i = (x_i^n, x_i^{n-1}, \dots, x_i, 1)$ are linearly independent in the space R^{n+1} , we must have $m = n+2$. Then the X_i span R^{n+1} , hence each point $Y \in R^{n+1}$ is a linear combination of the X_i . Then we have (4.1) with a prescribed $b_{ij} = 1$, all other $b_{ij} = 0$.

Condition (4.1) obviously implies (4.2). Conversely, if (4.2) holds, then in R^{n+1} there is no hyperplane through the points $0, X_1, \dots, X_m$ which leaves all points

$$Y_{ij} = \varepsilon_i(n \cdots (n - k_i + 1)y_i^{n-k_i}, \dots, k_i!, 0, \dots, 0)$$

strictly on one side. This means that the convex hull of the points Y_{ij} contains at least one point of the subspace spanned by the points X_i . But this is identical with (4.1).

EXAMPLES. For $n=2$, $k_1=1$, the sets $A=\{-1, 1\}$, $B=\{0\}$ are realizable, because, for all polynomials Q of degree 2, $Q(-1) - Q(1) + 2Q'(0) = 0$. Likewise, for $n=4$, $k_1=1$, $k_2=3$, the identity

$$Q(-1) - Q(1) + 2Q'(0) + \frac{1}{3}Q'''(0) = 0,$$

valid for all polynomials of degree 4, shows that the sets $A=\{-1, 1\}$, $B_1=B_2=\{0\}$ are realizable.

For the problem \mathcal{P}_k , each finite set B is possible for a sufficiently large n (see (1.8)). As a special case of Theorem 11 we have

THEOREM 12. *Sets $A=\{x_1, \dots, x_m\}$ and $B=\{y\}$, where $n \geq k+2$ if $a < y < b$, and $n \geq k+1$ if $y=a$ or $y=b$ can be realized by an extremal pair $f, P \in \mathcal{P}_k$ if and only if there is an identity*

$$(4.3) \quad \sum_{i=1}^m a_i Q(x_i) = Q^{(k)}(y),$$

valid for all polynomials Q of degree n , or if and only if $Q^{(k)}(y)$ vanishes for all Q with zeros x_1, \dots, x_m .

Finally, we can apply to our problem the polynomials Π_r , $r=0, \dots, n-m$ of (2.12). They constitute a basis for all polynomials that vanish at the points x_1, \dots, x_m . From conditions (4.1) and (4.2) of Theorem 11 we obtain:

THEOREM 13. *Sets $A=\{x_1, \dots, x_m\}$ and $B=\{y_1, \dots, y_l\}$ are realizable for the problem \mathcal{P}_k if and only if either $m \geq n+1$, or else $m \leq n$ and there exist $b_j \geq 0$, not all zero, for which*

$$(4.4) \quad \sum_{j=1}^l b_j \Pi_r^{(k)}(y_j) = 0, \quad r = 0, \dots, n-m.$$

5. Possible configurations: numbers m, l for the problem \mathcal{P}_k . Let n and the problem \mathcal{P}_k , $k \leq n$ be given. If m, l are the numbers of elements of two finite realizable sets A, B for the problem \mathcal{P}_k , what is the range of all possible pairs m, l ? We are far from being able to answer this question, and shall restrict ourselves to a few examples. Of interest seem to be, in the first place, cases with small values of m , with large values of l and with small values of $m+l$. In Theorem 14, we show that $m+l=n$ is possible for an arbitrary large n . In Theorem 15, l is approximately $\frac{1}{2}n$, and $m+l=n+1$ for the problem \mathcal{P}_1 . Theorem 16 is, in some sense, inverse to Theorem 9. In particular, Theorem 16 shows that $m=3, l=n-1$ is possible for $k=1$.

THEOREM 14. *Let $n=2p$, let k be odd and $k < n-1$. Then there exist sets A, B , realizable for the problem \mathcal{P}_k , and consisting of $m=n-2$ and $l=2$ points, respectively.*

Proof. Let the basic interval be $[-1, +1]$. We take $0 < x_1 < \dots < x_{p-1} \leq 1$ arbitrarily. The set A consists of $n-2$ points $\pm x_i, i=1, \dots, p-1$. Let $\Pi(x) = \prod_{i=1}^{p-1} (x^2 - x_i^2)$, and let y be a zero of the even polynomial $(x\Pi)^{(k)}$ in the interval $(0, 1)$. We take $y_1 = -y_2 = y$. We have to check that these sets A, B satisfy the conditions (4.4) of Theorem 13. With $b_1 = b_2 = 1$, the conditions become

$$(5.1) \quad \Pi_r^{(k)}(-y) + \Pi_r^{(k)}(y) = 0, \quad r = 0, 1, 2.$$

For $r=1$ this holds because of the choice of y , for $r=0$ and $r=2$ —because $\Pi^{(k)}$ and $(x^2\Pi)^{(k)}$ are odd polynomials.

THEOREM 15. *For $m=4, 5, \dots$ and arbitrary points $x_1 < \dots < x_m$ in $[a, b]$, there are points y_1, \dots, y_{m-2} in $[a, b]$ so that the sets $A = \{x_1, \dots, x_m\}$ and $B = \{y_1, \dots, y_{m-2}\}$ are sets $A(f, P), B(P)$ for an extremal pair f, P of the problem \mathcal{P}_1 , with $n=2m-3$.*

Proof. Let $\Pi(x) = \prod_{i=1}^m (x - x_i)$, and let $y_1 < \dots < y_{m-1}$ be the zeros of Π' in $[a, b]$. The first $m-2$ of these points are the required y_j . We prove this by means of the condition (4.2) of Theorem 11. Each polynomial Q that vanishes at the points x_i , has the form $Q = \Pi R$, where R is a polynomial of degree $m-3$. We have

$$(5.2) \quad Q'(y_j) = \Pi(y_j)R'(y_j), \quad j = 1, \dots, m-1.$$

We have to prove that not all values $Q'(y_j), j=1, \dots, m-2$ are >0 . Otherwise, since the signs of the $\Pi(y_j)$ alternate, also the signs $R'(y_j), j=1, \dots, m-2$ would alternate. Then R' , a polynomial of degree $m-4$, would have $m-3$ zeros, giving $R'(x) \equiv 0$ and $Q'(y_j) = 0, j=1, \dots, m-1$, a contradiction.

THEOREM 16. *All values of m and l that satisfy $m+2l=n+2, m \geq 3$ are possible for the problem \mathcal{P}_1 .*

The proof will depend on lemmas about the basic polynomials of Hermite interpolation, which are of interest in themselves.

LEMMA 2. Let numbers $\alpha < \beta$ and an integer $l = 1, 2, \dots$ be given. Let P be a polynomial whose zeros are real and lie to the left of α , and assume that $P(x) > 0$ for $x \geq \alpha$. There exist l points

$$(5.3) \quad \alpha < y_1 < \dots < y_l < \beta$$

for which the polynomials

$$(5.4) \quad G_k(x) = P(x)(x - \alpha)(x - \beta) \prod_{j \neq k} (x - y_j)^2, \quad k = 1, \dots, l$$

satisfy

$$(5.5) \quad G'_k(y_k) = 0, \quad k = 1, \dots, l.$$

Proof. We write $\alpha = y_0$, $\beta = y_{l+1}$, and study the polynomials (5.4) for an arbitrary choice of the points y_j , $j = 1, \dots, l$, with $y_0 \leq y_1 \leq \dots \leq y_l \leq y_{l+1}$. We define z_k , $k = 1, \dots, l$ in the following way. If $y_{k-1} = y_{k+1}$, we put $z_k = y_{k-1} = y_{k+1}$, and have $G'_k(z_k) = 0$. If $y_{k-1} < y_{k+1}$, there exists, by Rolle's theorem, a unique z_k , $y_{k-1} < z_k < y_{k+1}$, for which $G'_k(z_k) = 0$. First of all we prove:

$$(5.6) \quad z_k \leq z_{k+1}, \quad k = 1, \dots, l.$$

First let $z_{k+1} = y_{k+1}$, then (5.6) follows from the definition of z_k .

In the general case, let

$$H(x) = P(x)(x - \alpha)(x - \beta) \prod_{j \neq k, k+1} (x - y_j)^2.$$

Differentiating,

$$(5.7) \quad \begin{aligned} G'_k(z_{k+1}) &= H'(z_{k+1})(z_{k+1} - y_{k+1})^2 + 2H(z_{k+1})(z_{k+1} - y_{k+1}), \\ G'_{k+1}(z_{k+1}) &= H'(z_{k+1})(z_{k+1} - y_k)^2 + 2H(z_{k+1})(z_{k+1} - y_k) = 0, \end{aligned}$$

and eliminating $H'(z_{k+1})$,

$$(5.8) \quad G'_k(z_{k+1})(z_{k+1} - y_k) = 2H(z_{k+1})(z_{k+1} - y_{k+1})(y_{k+1} - y_k).$$

Let $z_{k+1} \neq y_{k+1}$, this implies that also $z_{k+1} \neq y_k$, $z_{k+1} \neq y_{k+2}$. We may assume that $z_{k+1} < y_{k+1}$, for otherwise (5.6) is obvious. Since $H(z_{k+1}) < 0$, we derive from (5.8) that $G'_k(z_{k+1}) \geq 0$. Now $G_k(x)$ is negative on (α, β) , and strictly decreasing for $y_{k-1} < x < z_k$, strictly increasing for $z_k < x < y_{k+1}$. Hence we have $z_k \leq z_{k+1}$.

We now consider the subset of S of the l -dimensional space R^l that consists of all points (y_1, \dots, y_l) satisfying $\alpha \leq y_1 \leq \dots \leq y_l \leq \beta$. The set of S is compact and convex. For given P , α , β , l we consider the map of S into itself defined by

$$(5.9) \quad (y_1, \dots, y_l) \rightarrow (z_1, \dots, z_l).$$

The map is continuous; to show this, we consider z_k as a function of the point $(y_1, \dots, y_l) \in S$. Let $(\bar{y}_1, \dots, \bar{y}_l)$ be a fixed point of S with the corresponding \bar{z}_k and the polynomial \bar{G}_k . For a given $\delta > 0$, we shall find a neighborhood U of $(\bar{y}_1, \dots, \bar{y}_l)$ in S , so that for all points (y_1, \dots, y_l) of U , $|z_k - \bar{z}_k| < \delta$. If $\bar{y}_{k-1} = \bar{y}_{k+1}$,

this follows from $y_{k-1} \leq z_k \leq y_{k+1}$. In the case $\bar{y}_{k-1} < \bar{z}_k < \bar{y}_{k+1}$, we can assume that δ satisfies $\bar{y}_{k-1} < \bar{z}_k - 2\delta < \bar{z}_k + 2\delta < \bar{y}_{k+1}$. In the interval $I = (\bar{z}_k - \delta, \bar{z}_k + \delta)$, \bar{G}'_k has the simple zero \bar{z}_k . Hence, for some $\rho > 0$, \bar{G}'_k takes values $> \rho$ and values $< -\rho$ on I . Let U be so small that for $(y_1, \dots, y_l) \in U$, $|y_i - \bar{y}_i| < \delta$, $i = k, k+1$, and that $|G'_k(x) - \bar{G}'_k(x)| < \rho$, $\alpha \leq x \leq \beta$. Then the interval (y_k, y_{k+1}) contains I , and G'_k changes sign on I , so that $z_k \in I$ and $|z_k - \bar{z}_k| < \delta$.

It follows that the map (5.9) has a fixed point: $z_k = y_k$, $k = 1, \dots, l$. For this fixed point we have the strict inequalities

$$(5.10) \quad \alpha < y_1 < \dots < y_l < \beta.$$

Otherwise there is a k , $1 \leq k \leq l$ for which $y_{k-1} = y_k < y_{k+1}$. Then z_k satisfies $y_{k-1} < z_k < y_{k+1}$, a contradiction, in view of $z_k = y_k$.

Let $x_1, \dots, x_m, y_1, \dots, y_l$ be distinct points of the interval $[a, b]$, let $m+2l = n+2$, we consider the basic polynomials F_i, G_j, H_j of the Hermite interpolation of degree $n+1$, with knots x_i, y_j . These are polynomial of degree (at most) $n+1$, defined by the relations

$$(5.11) \quad \begin{aligned} F_i(x_k) &= \delta_{ik}, & F_i(y_j) &= F'_i(y_j) = 0, \\ G_j(x_i) &= 0, & G_j(y_h) &= \delta_{jh}, & G'_j(y_h) &= 0, \\ H_j(x_i) &= 0, & H_j(y_h) &= 0, & H'_j(y_h) &= \delta_{jh}, \end{aligned}$$

$i, k = 1, \dots, m; \quad j, h = 1, \dots, l.$

LEMMA 3. If points $x_1 < \dots < x_{m-1} < x_m$ are given, $m \geq 2$, there exist $y_j, j = 1, \dots, l$

$$(5.12) \quad x_{m-1} < y_1 < \dots < y_l < x_m,$$

for which the Hermite polynomials (5.11) have the properties:

$$(5.13) \quad \text{The polynomials } G_j, \quad j = 1, \dots, l \text{ are of degree } n.$$

$$(5.14) \quad \text{The leading coefficients } a_j \text{ of } H_j, \quad j = 1, \dots, l \text{ satisfy } a_j < 0.$$

Proof. We select y_1, \dots, y_l according to Lemma 2, taking $\alpha = x_{m-1}$, $\beta = x_m$, $P(x) = \prod_{i=1}^{m-2} (x - x_i)$. The polynomials F_i, H_j are chosen in the standard fashion. The G_k are defined by

$$(5.15) \quad G_k(x) = b_k \prod_{i=1}^m (x - x_i) \prod_{j \neq k} (x - y_j)^2, \quad k = 1, \dots, l$$

where the b_k are taken to satisfy $G_k(y_k) = 1$. Then condition (5.13) will be satisfied: the degree of each G_k is $m+2l-2 = n$.

The polynomials H_j are of degree $n+1$; they do not vanish identically, hence H_j has simple zeros at $x_i, i = 1, \dots, m, y_j$ and double zeros at $y_k, k \neq j$. At y_j , H_j changes sign from $-$ to $+$, then remains ≥ 0 until x_m , and at x_m changes sign again. Hence $H_j(x) < 0$ for large x , and $a_j < 0$.

Proof of Theorem 16. We take $x_1 < \dots < x_m$ arbitrarily in $[a, b]$, then $y_j, j = 1,$

\dots, l , $x_{m-1} < y_1 < \dots < y_l < x_m$ according to Lemma 3. The set $B = \{y_1, \dots, y_l\}$ is possible for \mathcal{P}_1 , since the condition (1.8) is satisfied in virtue of $m \geq 3$. To check the condition (4.2) of Theorem 11 we show that there does not exist a polynomial Q of degree n that satisfies $Q(x_i) = 0$, $i = 1, \dots, m$, $Q'(y_j) > 0$, $j = 1, \dots, l$.

Each such Q is also a polynomial of degree $n+1$, and hence has a representation by the Hermite interpolation formula:

$$(5.16) \quad Q(x) = \sum_{i=1}^m Q(x_i)F_i(x) + \sum_{j=1}^l Q(y_j)G_j(x) + \sum_{j=1}^l Q'(y_j)H_j(x).$$

The first sum is 0, the second is a polynomial of degree n . The last sum has as its highest term

$$\sum_{j=1}^l a_j Q'(y_j) x^{n+1},$$

and this is not zero by (5.14). Hence a polynomial Q of this type cannot exist.

The method of proof of Theorem 16 can be used for other purposes. For $m=2$ one can give, by means of our considerations, an existence proof of the Gauss integration formulas, which is free of the theory of orthogonal polynomials.

REFERENCES

1. K. Atkinson and A. Sharma, *A partial characterization of poised Hermite-Birkhoff interpolation problems*, SIAM J. Numer. Anal. **6** (1969), 230–235.
2. G. D. Birkhoff, *General mean value and remainder theorems with applications to mechanical differentiation and integration*, Trans. Amer. Math. Soc. **7** (1906), 107–136.
3. G. G. Lorentz, *Approximation of functions*, Holt, New York, 1966. MR **35** #4642.
- 4, 5. G. G. Lorentz and K. L. Zeller, *Degree of approximation by monotone polynomials*. I, II, J. Approximation Theory **1** (1968), 501–504; **2** (1969), 265–269. MR **39** #699.
6. ———, *Gleichmässige Approximation durch monotone Polynome*, Math. Z. **109** (1969), 87–91.
7. ———, *Birkhoff interpolation*, SIAM J. Numer. Anal. (to appear).
8. N. E. Nörlund, *Vorlesungen über Differenzenrechnung*, Springer-Verlag, Berlin, 1924.
9. J. R. Rice, *Approximation with convex constraints*, J. SIAM **11** (1963), 15–32. MR **28** #2387.
10. J. A. Roulier, *Monotone and weighted approximation*, Dissertation, Syracuse University, Syracuse, N. Y., 1968. MR **38** #4875.
11. ———, *Monotone approximation of certain classes of functions*, J. Approximation Theory **1** (1968), 319–324.
12. I. J. Schoenberg, *On Hermite-Birkhoff interpolation*, J. Math. Anal. Appl. **16** (1966), 538–543. MR **34** #3160.
13. O. Shisha, *Monotone approximation*, Pacific J. Math. **15** (1965), 667–671. MR **32** #2802.

UNIVERSITY OF TEXAS,

AUSTIN, TEXAS 78712

UNIVERSITY OF TÜBINGEN,

FEDERAL REPUBLIC OF GERMANY