

POLYNOMIALS OF BEST APPROXIMATION ON A REAL FINITE POINT SET. I⁽¹⁾

BY

T. S. MOTZKIN AND J. L. WALSH

Polynomials $p_n(x)$ of given degree n (≥ 0) of best approximation to a given real function $f(x)$ on a given real finite point set E have various properties in common, especially those relating to oscillation of the difference $f(x) - p_n(x)$; here approximation is measured according to any of the classical deviations (i.e. measures of approximation), such as those of Tchebycheff and least p th powers ($p > 0$). In the present paper we study such properties in a systematic manner.

The totality of extremal polynomials for the classical deviations and unspecified positive weights is identical with the totality of *juxtapolynomials* as defined below, and §§1-3 are devoted to the striking relation between juxtapolynomials and weak oscillation of $f(x) - p_n(x)$ on E . Then §§4-6 consider the properties of boundedness, closure, and connectedness of the set of juxtapolynomials, while §§7-9 study the characterization of juxtapolynomials relative to the special classical norms of $p > 1$, of Tchebycheff, and of $0 < p < 1$. Part II of this paper, to be published separately, contains further developments based on the consideration of polynomials of best approximation and the corresponding weights as points in $(n+1)$ -space and m -space.

Introduction. A *polynomial of degree n* is a function of the form $a_0z^n + a_1z^{n-1} + \dots + a_n$.

We say that the polynomial $q_n(z)$ of degree n is a *closer polynomial* to $f(z)$ on E than $p_n(z)$ (also a polynomial of degree n) provided $q_n(z) \neq p_n(z)$ and provided

$$(0.1) \quad |f(z) - q_n(z)| < |f(z) - p_n(z)| \quad \text{on } E \text{ where } f(z) - p_n(z) \neq 0,$$

$$(0.2) \quad f(z) - q_n(z) = f(z) - p_n(z) \quad \text{on } E \text{ where } f(z) - p_n(z) = 0.$$

If there exists no closer polynomial to $f(z)$ on E than $p_n(x)$ it is called a *juxtapolynomial* to $f(z)$ on E .

If E contains fewer than $n+1$ points, there exist infinitely many polynomials of degree n coinciding with $f(z)$ on E , so this case is henceforth excluded.

In the sequel we shall frequently use the

Received by the editors July 20, 1957.

⁽¹⁾ This research was sponsored in part by the Office of Naval Research, U. S. Navy, and by the Office of Scientific Research, Air Research and Development Command, U. S. Air Force. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Summary published in Proc. Nat. Acad. Sci. USA vol. 43 (1957) pp. 845-846.

REMARK. If $p_n(z)$ is a juxtapolynomial to $f(z)$ on E , then it is also a juxtapolynomial to $f(z)$ on any set containing E .

If E and $f(z)$ are given, there always exists a juxtapolynomial of degree n to $f(z)$, indeed one coinciding with $f(z)$, on a subset of E consisting of $n+1$ points; hence there exists a juxtapolynomial of degree n to $f(z)$ also on E . In this connection, compare Lemma 3 below.

A polynomial of given degree of best approximation for any of the classical deviations with positive weights is a juxtapolynomial; the same is true for any *monotonic* deviation $\delta[\delta_1, \delta_2, \dots, \delta_m]$, $\delta_k = |f(z_k) - p_n(z_k)|$, namely provided δ is positive if $\sum \delta_k > 0$, and provided δ decreases when all the non-vanishing δ_k decrease and the vanishing δ_k remain unchanged. This concept is essentially due to Fejér (1922); compare [4, §3].

If we choose $f(z) \equiv z^{n+1}$, the polynomial $z^{n+1} - p_n(z)$ is an *extremal polynomial* or is a *T-polynomial* if it minimizes a deviation, and if (0.1) and (0.2) are impossible is called an *infrapolynomial*. The concept of infrapolynomial is due to Fekete and von Neumann; the analogous concept of juxtapolynomial is called by Fekete [2] *nearest polynomial*, a term which we reserve for extremal polynomials that minimize a given deviation.

If $p_n(z)$ is a juxtapolynomial to $f(z)$ on E and different from $f(z)$ there, then $p_n(z)$ is a polynomial of best approximation to $f(z)$ on E in the sense of Tchebycheff with weight function $1/|f(z) - p_n(z)|$, for the inequality

$$\max_{z \text{ on } E} \frac{|f(z) - q_n(z)|}{|f(z) - p_n(z)|} < 1 = \max_{z \text{ on } E} \frac{|f(z) - p_n(z)|}{|f(z) - p_n(z)|}$$

is impossible if $q_n(z)$ is a polynomial of degree n . This remark generalizes a remark made by Fekete [1] concerning *T-polynomials*.

The following lemma is not as trivial as it might seem at first sight:

LEMMA 0. Every juxtapolynomial $p_n(x)$ of degree n to a real function $f(x)$ on a real finite point set E containing at least $n+1$ points is also real.

If $p_n(x)$ satisfies the hypothesis, we set $p_n(x) \equiv p'_n(x) + ip''_n(x)$, where $p'_n(x)$ and $p''_n(x)$ are polynomials of degree n which are real for real x . If $p''_n(x) \not\equiv 0$ we show the existence of a closer polynomial $q_n(x) \equiv p'_n(x) + \epsilon r_n(x)$ than $p_n(x)$ to $f(x)$ on E , where the real polynomial $r_n(x)$ of degree n is to be defined, and $\epsilon (> 0)$ also to be determined. The totality E' of points of E in which $p''_n(x)$ vanishes is not greater than n in number. On E' where $p'_n(x) = f(x)$ we set $r_n(x) = 0$; thus (0.2) is satisfied in such points. On E' where $p'_n(x) \neq f(x)$ we set $\text{sig } r_n(x) = \text{sig } [f(x) - p'_n(x)]$, so (0.1) is satisfied in such points if ϵ is sufficiently small; our total conditions on the values of $r_n(x)$ relate to a number of points x not greater than n . On the finite set $E - E'$ we have $|f(x) - p'_n(x)| < |f(x) - p_n(x)|$, so (0.1) follows for ϵ sufficiently small. Then $q_n(x)$ is a closer polynomial to $f(x)$ on E , contrary to the hypothesis that $p_n(x)$ is a juxtapolynomial.

JUXTAPOLYNOMIALS

1. **Weak oscillation.** We say that a real function $\phi(x)$ *oscillates weakly* on the real set $E: (x_1, x_2, \dots, x_\lambda)$, where $x_1 < x_2 < \dots < x_\lambda$, provided we have either (1.1) or (1.2):

$$(1.1) \quad \phi(x_1) \geq 0, \quad \phi(x_2) \leq 0, \quad \phi(x_3) \geq 0, \dots,$$

$$(1.2) \quad \phi(x_1) \leq 0, \quad \phi(x_2) \geq 0, \quad \phi(x_3) \leq 0, \dots$$

A pair of adjacent inequalities on $\phi(x)$ here may be said to define a *weak sign change* of $\phi(x)$ on E , and of course define a (strong) sign change if the equality signs are omitted. If the equality signs are omitted in (1.1) and (1.2) we have (strong) oscillation of $\phi(x)$ on E in the usual sense, of which weak oscillation is a limiting case as E remains fixed and the values of $\phi(x)$ change continuously and suitably on E .

A continuous function $\phi(x)$ has a *double zero* at a point x_0 , with $\phi(x_0) = 0$, if it does not change sign there, in the sense that $\phi(x) \geq 0$ or $\phi(x) \leq 0$ throughout some neighborhood of x_0 . With this convention (due to S. Bernstein, 1926) we have

LEMMA 1. *If the function $\phi(x)$ is continuous for $x_1 \leq x \leq x_\lambda$ and oscillates weakly on the set $E: (x_1, x_2, \dots, x_\lambda)$, then there exist zeros y_k of $\phi(x)$ such that*

$$x_1 \leq y_1 \leq x_2 \leq y_2 \leq \dots \leq y_{\lambda-1} \leq x_\lambda.$$

That is to say, a suitable set of $\lambda - 1$ zeros of $\phi(z)$ separates weakly the points of E .

It will be noted that no zero y_k is enumerated as such more than once, and it is a consequence of the following discussion that multiple enumeration as successive zeros can occur only twice, namely for a double zero of $\phi(x)$.

CASE I. $\phi(x_1) = \phi(x_2) = \dots = \phi(x_\lambda) = 0$. Here we need merely set $y_k = x_k$, $k = 1, 2, \dots, \lambda - 1$.

CASE II. $\phi(x_1) > 0$, $\phi(x_2) = \dots = \phi(x_{\lambda-1}) = 0$, $\phi(x_\lambda) < 0$. Here λ is even and either $\phi(x)$ has a sign change in $x_1 < x < x_\lambda$ other than at a point x_k or $\phi(x)$ has at least one double zero at some x_k in $x_1 < x < x_\lambda$; in either case the y_k exist satisfying the conclusion.

CASE III. $\phi(x_1) > 0$, $\phi(x_2) = \dots = \phi(x_{\lambda-1}) = 0$, $\phi(x_\lambda) > 0$. Here λ is odd, and again either $\phi(x)$ has a sign change in $x_1 < x < x_\lambda$ other than at a point x_k or $\phi(x)$ has at least one double zero at some x_k in $x_1 < x < x_\lambda$.

CASE IV. $\phi(x_1) > 0$, $\phi(x_2) = \dots = \phi(x_\lambda) = 0$. We need merely set $x_1 < y_1 = x_2 < y_2 = x_3 < \dots < y_{\lambda-1} = x_\lambda$.

All other cases under Lemma 1 can be handled from the cases already treated, by reversing the signs of the assumed inequalities on $\phi(x)$, by reversing the sense of the inequalities on the x_k and y_k , and by combining the various cases in sequence; thus two cases may be combined if the terminal point of one set is the initial point of the second set, namely a point x_k , for which $\phi(x_k) > 0$ or $\phi(x_k) < 0$.

It may be noticed that we have not needed in Lemma 1 the complete continuity of the function $\phi(x)$; it is sufficient if the algebraic sign of $\phi(x)$ cannot reverse as x moves along an interval unless $\phi(x)$ vanishes somewhere on that interval. This latter requirement is weaker than the (Darboux) property of assuming all intermediate values on any interval on which it assumes two values.

LEMMA 2. *If a polynomial $p_n(x)$ of degree n has at least $n+1$ weak sign changes on a set x_1, x_2, \dots, x_{n+2} , then $p_n(x)$ vanishes identically.*

For a polynomial, a double zero in the sense above defined is a zero of order at least two in the usual sense. Under the hypothesis of Lemma 2, $p_n(x)$ has by Lemma 1 a number of zeros on $x_1 \leq x \leq x_{n+2}$ which is at least $n+1$, so $p_n(x) \equiv 0$.

2. Weak oscillation and juxtapolynomials. The following theorem has already been proved for the case of approximation to the function x^{n+1} [5, Theorem 4.1].

THEOREM 1. *If $p_n(z)$ is a polynomial of degree n of best approximation to the function $f(z)$ on a set E' with deviation $\delta'[f(z) - p_n(z)]$ and also a polynomial of degree n of best approximation to $f(z)$ on a set E'' with deviation $\delta''[f(z) - p_n(z)]$, then $p_n(z)$ is a polynomial of degree n of best approximation to $f(z)$ on the set $E = E' + E''$ with deviation $\delta[f(z) - p_n(z)] \equiv \delta'[f(z) - p_n(z)] + \delta''[f(z) - p_n(z)]$.*

If $q_n(z)$ is an arbitrary polynomial of degree n , we have the extremal properties

$$\begin{aligned}\delta'[f(z) - p_n(z)] &\leq \delta'[f(z) - q_n(z)], \\ \delta''[f(z) - p_n(z)] &\leq \delta''[f(z) - q_n(z)],\end{aligned}$$

and by addition these inequalities imply

$$\delta[f(z) - p_n(z)] \leq \delta[f(z) - q_n(z)];$$

this completes the proof. We remark that under the hypothesis of Theorem 1, any polynomial minimizing δ also minimizes δ' and δ'' .

Theorem 1 is needed in our proof of

THEOREM 2. *If $E: (x_1, x_2, \dots, x_\lambda)$ is an arbitrary real set with $x_1 < x_2 < \dots < x_\lambda$, $\lambda \geq n+2$, if $f(x)$ is an arbitrary (real) function defined on E , and if $p_n(x)$ is a polynomial of degree n such that $f(x) - p_n(x)$ has at least $n+1$ consecutive weak sign changes on E , then $p_n(x)$ is a juxtapolynomial of $f(x)$ on E , and indeed is a polynomial of best approximation to $f(x)$ on E in the sense of least first powers.*

Since $f(x) - p_n(x)$ has at least $n+1$ weak sign changes on E , it oscillates weakly on some subset E_0 of E containing precisely $n+2$ points. On E_0 we

may replace $f(x)$ by its equal, a polynomial $P_{n+1}(x)$ of degree $n+1$, so $P_{n+1}(x) - p_n(x)$ oscillates weakly on E_0 and by Lemma 1 the zeros of $P_{n+1}(x) - p_n(x)$ separate weakly the points of E_0 . Hence [4, §8, Corollary] $P_{n+1}(x) - p_n(x)$ is a constant multiple of the polynomial $z^{n+1} + \cdots$ of least norm on E_0 in the sense of least first powers with suitable weights. Thus [4, §2] even if $p_n(x) \equiv f(x)$ on E_0 , $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E_0 in the sense of least first powers with suitable weights.

Every point of E belongs to some subset E_0 of $n+2$ points on which $f(x) - p_n(x)$ oscillates weakly. For if we have one such set E_0 , an arbitrary point X of $E - E_0$ which lies between two points of E_0 may be adjoined to E_0 provided one or the other points of E_0 adjacent to X is deleted from E_0 in such a way as to preserve weak oscillation; an arbitrary point X of $E - E_0$ which does not lie between two points of E_0 may be adjoined to E_0 provided a suitable one of the original endpoints of E_0 is deleted.

By application of Theorem 1 to a number of suitable subsets E_0 of E , where the totality of such subsets contains E , it follows that $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E in the sense of least first powers with suitable positive weights. Consequently $p_n(x)$ is also a juxtapolynomial to $f(x)$ on E . Theorem 2 is established.

The first part of Theorem 2 can be established still more simply. Suppose $f(x) - p_n(x)$ has at least $n+1$ weak sign changes on E ; we assume (0.1) and (0.2) for some polynomial $q_n(x) (\not\equiv p_n(x))$ of degree n and shall reach a contradiction. It follows from (0.1) that $q_n(x) - p_n(x)$ has the same algebraic sign as does $f(x) - p_n(x)$ on E where $f(x) - p_n(x) \neq 0$, and it follows from (0.2) that $q_n(x) - p_n(x) = 0$ on E where $f(x) - p_n(x) = 0$. Consequently $q_n(x) - p_n(x)$ also has at least $n+1$ weak sign changes on E , and by Lemma 2 vanishes identically, contrary to assumption.

3. Weak oscillation and least first powers. Theorem 2 is contained in

THEOREM 3. *Let the real function $f(x)$ be defined on the real finite set E containing at least $n+2$ points. The three classes of polynomials $p_n(x)$ of degree n defined by each of the following properties are identical:*

- (1) $f(x) - p_n(x)$ has at least $n+1$ weak sign changes on E ,
- (2) $p_n(x)$ is a juxtapolynomial to $f(x)$ on E ,
- (3) $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E in the sense of least first powers with some positive weights.

We have shown that (1) implies (2) and (3); (3) obviously implies (2); it remains to prove that (2) implies (1): If $p_n(x)$ (by Lemma 0 necessarily real) is a juxtapolynomial of degree n to $f(x)$ on E , then $f(x) - p_n(x)$ has at least $n+1$ weak sign changes on E .

If the sequence $f(x_1) - p_n(x_1), f(x_2) - p_n(x_2), \cdots$ exhibits no more than n (weak) sign changes on E , there exists a polynomial $q_n(x)$ of degree n which vanishes where $\phi(x) \equiv f(x) - p_n(x)$ vanishes on E , and otherwise on E has

the same algebraic sign as $f(x) - p_n(x)$. To count sign changes here, at a point X of E where $\phi(X) \neq 0$ we write $\phi(X) > 0$ or $\phi(X) < 0$; at successive zeros X_1, X_2, \dots following (and preceding) such a point X we write alternately $\phi(X_1) \leq 0, \phi(X_2) \geq 0$, etc. or $\phi(X_1) \geq 0, \phi(X_2) \leq 0$, etc. In defining $q_n(x)$ we may place a double zero of $q_n(x)$ at a zero x_k of $\phi(x)$ which is flanked by two similar inequalities such as $\phi(x_{k-1}) \geq 0, \phi(x_{k+1}) \geq 0$ at least one of which is strong. Then for suitably chosen $\epsilon (> 0)$, the function $\phi(x) - \epsilon q_n(x)$ vanishes wherever $\phi(x)$ vanishes on E , and at all other points of E is numerically less than $\phi(x)$, so $p_n(x) + \epsilon q_n(x)$ is a closer polynomial to $f(x)$ on E than is $p_n(x)$, contrary to hypothesis. Theorem 3 is established.

For given positive weights, the polynomial of best approximation in the sense of least first powers need not be unique [5, Theorem 1.1].

It will be of interest to formulate the

REMARK. If $p_n(x)$ and $q_n(x)$ are different juxtapolynomials of degree n to $f(x)$ on a point set $E: (x_1, x_2, \dots, x_{n+2})$ consisting of $n+2$ points, $x_k < x_{k+1}$, it is not possible to have the inequalities

$$(3.1) \quad f(x_1) - p_n(x_1) \geq 0, \quad f(x_2) - p_n(x_2) \leq 0, \dots,$$

$$(3.2) \quad f(x_1) - q_n(x_1) \leq 0, \quad f(x_2) - q_n(x_2) \geq 0, \dots$$

If these inequalities are valid, subtraction yields

$$q_n(x_1) - p_n(x_1) \geq 0, \quad q_n(x_2) - p_n(x_2) \leq 0, \dots,$$

whence by Lemma 2 we have $q_n(x) \equiv p_n(x)$.

A simple test is available for determining whether (3.1) or (3.2) is characteristic of the set of juxtapolynomials to $f(x)$ on the set E , namely if we have $(-1)^{n+1}\Delta > 0$ or $(-1)^{n+1}\Delta < 0$, where Δ is

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_{n+2} \\ x_1^2 & x_2^2 & \dots & x_{n+2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^n & x_2^n & \dots & x_{n+2}^n \\ y_1 & y_2 & \dots & y_{n+2} \end{vmatrix}$$

and where $y_k = f(x_k)$, then the order is as in (3.1) or (3.2) respectively.

We also have to Theorem 3 the

COROLLARY. If the real finite set E contains more than $n+2$ points, if $p_n(x)$ is a juxtapolynomial of degree n to $f(x)$ on E , then $p_n(x)$ is also a juxtapolynomial to $f(x)$ on some subset E_0 of E consisting of $n+2$ points.

Since $f(x) - p_n(x)$ has at least $n+1$ weak sign changes on E , it has $n+1$ weak sign changes on a subset E_0 of E consisting of $n+2$ points, and consequently (Theorem 3) $p_n(x)$ is a juxtapolynomial to $f(x)$ on E_0 .

The Corollary expresses precisely the fact that if the real set E contains at least $n+2$ points and if $p_n(x)$ is a polynomial of degree n of best approximation to $f(x)$ on E in the sense of least first powers with given weights, then $p_n(x)$ is a polynomial of best approximation to $f(x)$ on some E_0 in the sense of least first powers with suitable weights.

It is clear that the determination of all juxtapolynomials of degree n to a given function $f(z)$ on a set E of $n+1$ points is trivial, for there exists a unique polynomial of degree n coinciding with $f(z)$ on E , and this is the unique juxtapolynomial. However, by the Remark in the Introduction, we may formulate

LEMMA 3. *If $f(z)$ is given on a set E of at least $n+2$ points, any polynomial $p_n(z)$ of degree n which coincides with $f(z)$ on a subset E' of E containing $n+1$ points is a juxtapolynomial to $f(z)$ on E .*

Any such $p_n(z)$ will be called an *interpolation polynomial* to $f(z)$ on E .

TOPOLOGICAL PROPERTIES OF THE SET OF JUXTAPOLYNOMIALS

4. Boundedness. Certain properties of the entire set of juxtapolynomials to $f(x)$ on E are of interest:

THEOREM 4. *Given a real finite set E containing at least $n+2$ points and a function $f(x)$ defined on E , the set of juxtapolynomials $p_n(x)$ of degree n to $f(x)$ on E is bounded on every bounded set.*

Since E is finite, it contains only a finite number of subsets each consisting of $n+2$ points, so by the Corollary to Theorem 3 it is sufficient to treat the case that E consists of precisely $n+2$ points $(x_1, x_2, \dots, x_{n+2})$. Lagrange's interpolation formula expresses $p_n(x_{n+2})$ linearly in terms of $p_n(x_1), \dots, p_n(x_{n+1})$, say $p_n(x_{n+2}) = \sum_{k=1}^{n+1} \beta_k p_n(x_k)$ for an arbitrary polynomial $p_n(x)$ of degree n , where the β_k do not depend on $p_n(x)$. Then we may write for every polynomial

$$(4.1) \quad \sum_{k=1}^{n+2} \gamma_k [f(x_k) - p_n(x_k)] = \gamma, \quad \gamma_k \neq 0,$$

and if (as we now suppose) $f(x)$ is not identically equal to any polynomial of degree n on E , we have $\gamma \neq 0$; if $f(x)$ is equal to a polynomial of degree n on E , of course the set of juxtapolynomials consists of that single polynomial, bounded on every bounded set.

By the remark following the proof of Theorem 3, either (3.1) or (3.2) but not both is characteristic of the juxtapolynomials, so for those polynomials we may write (4.1) in the form

$$(4.2) \quad \sum_{k=1}^{n+2} \alpha_k |f(x_k) - p_n(x_k)| = 1.$$

If $p_n(x)$ is the (interpolation) polynomial of degree n which coincides with

$f(x)$ in all the points of E except x_k , in which $|f(x) - p_n(x)|$ takes the value w_k , then $p_n(x)$ is a juxtapolynomial, and (4.2) becomes $\alpha_k w_k = 1$, whence $\alpha_k > 0$. For an arbitrary juxtapolynomial $p_n(x)$ we then have from (4.2)

$$(4.3) \quad |f(x_k) - p_n(x_k)| \leq 1/\alpha_k, \quad k = 1, 2, \dots, n+2,$$

so the numbers $|p_n(x_k)|$ are bounded for all juxtapolynomials $p_n(x)$, and the theorem follows by Lagrange's interpolation formula.

As an additional remark, we note that (4.3) can be written

$$(4.4) \quad |f(x_k) - p_n(x_k)| \leq w_k;$$

here w_k is the value of the first member of (4.4) when $p_n(x)$ is the interpolation polynomial taking the values $f(x)$ on $E - x_k$. To prove the boundedness of all juxtapolynomials of degree n to a variable (uniformly bounded) function $f(x)$ on a variable set E , it is thus sufficient to consider the special juxtapolynomials which coincide with $f(x)$ in $n+1$ points of E . In particular, *if E_λ is a uniformly bounded variable finite point set containing at least $n+2$ points, if the variable function $f_\lambda(x)$ defined on E_λ is uniformly bounded, and if the distance between consecutive points of E_λ is bounded from zero, then the totality of all juxtapolynomials for all E_λ and for all $f_\lambda(x)$ is uniformly bounded on every bounded set, for all λ .* The condition that the distance between consecutive points of E_λ be bounded from zero is essential to the conclusion, as is shown by the example $n=1$, E_λ containing three points and otherwise merely restricted to the interval $(0, 1)$, and $f_\lambda(x)$ taking on but the two values zero and unity. The special juxtapolynomials coinciding with $f_\lambda(x)$ in two points of E_λ are not uniformly bounded.

5. Closure. Here we prove

THEOREM 5. *If $f(x)$ and a real set $E(x_1, x_2, \dots, x_m)$ containing at least $n+2$ points are given, the set of juxtapolynomials to $f(x)$ on E is closed.*

We need to show that if $f_1(x), f_2(x), \dots$ are a sequence of juxtapolynomials of degree n to $f(x)$ on E which converge (necessarily uniformly on every bounded set [6, Lemma 2]) to a polynomial $P(x)$ of degree n , then $P(x)$ is also a juxtapolynomial of degree n to $f(x)$ on E . Since E is finite, there exists only a finite number of subsets of E each of $n+2$ points. An infinite number of the $f_n(x)$ are juxtapolynomials to $f(x)$ on some fixed subset E_0 of E consisting of $n+2$ points, by the Corollary to Theorem 3, and by the Remark to Theorem 3 only one of the sets of inequalities like (3.1) and (3.2) is pertinent; we except the trivial case $f_n(x) \equiv f(x)$ on E for every n . Such a set of inequalities as $f(X_1) - f_k(X_1) \geq 0, f(X_2) - f_k(X_2) \leq 0, \dots$, valid for every k is also valid for the limit function $P(x)$, which by Theorem 3 is a juxtapolynomial to $f(x)$ on E_0 and on E .

In connection with the closure of the set of juxtapolynomials, also the following theorem is significant:

THEOREM 6. *Let E be an arbitrary point set, let F be a family of approximating functions defined on E , and let $\phi, \phi_1, \phi_2, \phi_3, \dots$ be a sequence of functions each defined on E . Let $\delta_k = \delta_k(\phi_k - g)$ or $\delta_k(\phi - g)$ be a measure of the deviation on E of ϕ_k or ϕ from an arbitrary function g of F , defined for $k=1, 2, \dots$, and let also $\delta = \delta(\phi - g)$ be a measure defined for ϕ and every g of F . Suppose δ has the property*

$$\lim_{k \rightarrow \infty} \delta_k(\phi_k - f_k) = \delta(\phi - f)$$

whenever $\phi_k \rightarrow \phi$ and $f_k \rightarrow f$ uniformly on E , where f and the f_k belong to F . If $\phi_k \rightarrow \phi$ uniformly on E , if f_k and f are functions of least deviation δ_k to ϕ_k and ϕ on E respectively, and if $f_k \rightarrow f$ uniformly on E , then f is a function of F of least deviation δ to ϕ on E .

We have by hypothesis for an arbitrary g in F

$$\delta_k(\phi_k - g) \geq \delta_k(\phi_k - f_k),$$

whence by taking limits

$$\delta(\phi - g) \geq \delta(\phi - f),$$

as we were to prove. In Theorem 6 are included the classical norms for approximation by polynomials of fixed degree n in the sense of least weighted p th powers, $0 < p$, where the function approximated, the exponent p , and the weight function all vary continuously. Deviations $\delta_k(\phi, g)$ and $\delta(\phi, g)$ may also be considered in generalizing Theorem 6, which depend on the functions ϕ and g otherwise than merely on their difference.

6. Connectedness. A further property of the totality of juxtapolynomials is expressed in

THEOREM 7. *Let $E: (x_1, x_2, \dots, x_m)$ be a set of m ($\geq n+2$) real distinct points and let $f(x)$ be given on E . In the space of $n+1$ real dimensions representing a real polynomial of degree n by its coefficients, the set of juxtapolynomials of degree n to $f(x)$ on E is connected.*

If $m = n+2$, any two such juxtapolynomials $p_n(x)$ and $q_n(x)$ can be connected linearly; for by the Remark to Theorem 3 we have $(x_k < x_{k+1})$ either

$$\begin{aligned} f(x_1) - p_n(x_1) &\geq 0, & f(x_2) - p_n(x_2) &\leq 0, \dots, \\ f(x_1) - q_n(x_1) &\geq 0, & f(x_2) - q_n(x_2) &\leq 0, \dots, \end{aligned}$$

or the reverse inequalities. Inequalities in the original sense are then valid also for the function

$$\lambda[f(x) - p_n(x)] + (1 - \lambda)[f(x) - q_n(x)], \quad 0 \leq \lambda \leq 1,$$

which is therefore (Theorem 3) likewise a juxtapolynomial to $f(x)$ on E . Indeed, we have proved that for $m = n+2$ the set of juxtapolynomials is convex.

For arbitrary m , any two juxtapolynomials which oscillate weakly on the same subset E_0 of E containing precisely $n+2$ points can be connected linearly by juxtapolynomials, so we need merely prove that some juxtapolynomial on any subset E_0 of E containing $n+2$ points can be connected to some juxtapolynomial on any other subset E_1 of E also containing $n+2$ points. It is sufficient to prove this fact for such subsets E_0 and E_1 which have $n+1$ points in common. However, any (interpolation) polynomial of degree n coinciding with $f(x)$ in these latter $n+1$ points is a juxtapolynomial on both E_0 and E_1 , by Lemma 3, so Theorem 7 is established.

We add the remark that *for approximation in the sense of least first powers with given weights on a given set E of $m \geq n+2$ points to $f(x)$ by a polynomial of degree n , the set of polynomials of best approximation forms a convex set.* If $p_n(x)$ and $q_n(x)$ are polynomials of best approximation:

$$\mu = \sum \mu_k |f(x_k) - p_n(x_k)| = \sum \mu_k |f(x_k) - q_n(x_k)|,$$

so also is $\lambda p_n(x) + (1-\lambda)q_n(x)$, $0 < \lambda < 1$:

$$\begin{aligned} & \sum \mu_k |f(x_k) - \lambda p_n(x_k) - (1-\lambda)q_n(x_k)| \\ & \leq \lambda \sum \mu_k |f(x_k) - p_n(x_k)| + (1-\lambda) \sum \mu_k |f(x_k) - q_n(x_k)| \\ & \leq \mu; \end{aligned}$$

the strong inequality signs are impossible.

SPECIAL JUXTAPOLYNOMIALS

7. Strong oscillation and least p th powers, $p > 1$. Theorem 3 has a precise analogue for approximation in the sense of least p th powers, $p > 1$:

THEOREM 8. *Let the real function $f(x)$ be defined on the real finite set E containing at least $n+2$ points, and suppose $f(x)$ is not identical on E to any polynomial of degree n . The three classes of polynomials $p_n(x)$ of degree n defined by each of the following properties are identical, and (3) is independent of p (> 1):*

- (1) $f(x) - p_n(x)$ has at least $n+1$ strong sign changes on E ;
- (2) $p_n(x)$ is a proper (i.e. different from $f(x)$) at each point of E) juxtapolynomial to $f(x)$ on some subset E_0 of E containing $n+2$ points;
- (3) $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E in the sense of least p th powers ($p > 1$) with some positive weights.

Our proof of Theorem 8 follows in part the method of proof of Theorem 3.

If $f(x) - p_n(x)$ has at least $n+1$ strong sign changes on E , it has $n+1$ strong sign changes on some subset E_0 of E consisting of $n+2$ points, and $p_n(x)$ is by Theorem 3 a juxtapolynomial to $f(x)$ on E_0 , hence is such a juxtapolynomial on E .

On E_0 we may express $f(x)$ as a polynomial of degree $n+1$ [as in 4, §2], so $f(x) - p_n(x)$ is on E_0 a constant multiple of a polynomial $T_{n+1}(x) \equiv x^{n+1} + \dots$ which has $n+1$ strong sign changes on E_0 . Then $T_{n+1}(x)$ is [5, Theorem 5.2]

an extremal polynomial on E_0 for least p th power norm, and $p_n(x)$ is a polynomial (necessarily unique) of best approximation to $f(x)$ on E_0 in the sense of least p th powers with suitable weights. Every point of E at which $f(x) - p_n(x) \neq 0$ belongs to some subset E_0 of E , as in the proof of Theorem 2, and hence by Theorem 1 all such points of E can be adjoined to E_0 without altering the extremal character of $p_n(x)$. Every point of E at which $f(x) - p_n(x) = 0$ can similarly be adjoined to the new E_0 (see Lemma 4 below), so it follows that (1) implies both (2) and (3).

LEMMA 4. *If $p_n(z)$ is a polynomial of degree n of best approximation to $f(z)$ on a set E containing m ($\geq n+2$) points, in the sense of minimizing the norm*

$$\sum_{k=1}^m \mu_k |f(z_k) - p_n(z_k)|^p, \quad \mu_k > 0, p > 0,$$

and if $f(z_0) - p_n(z_0) = 0$, where z_0 is not in E , then $p_n(z)$ also minimizes the norm

$$\sum_{k=0}^m \mu_k |f(z_k) - p_n(z_k)|^p,$$

where μ_0 (> 0) is arbitrary.

For every polynomial $q_n(z)$ of degree n we have

$$\begin{aligned} \sum_{k=1}^m \mu_k |f(z_k) - q_n(z_k)|^p &\geq \sum_{k=1}^m \mu_k |f(z_k) - p_n(z_k)|^p \\ &= \sum_{k=0}^m \mu_k |f(z_k) - p_n(z_k)|^p, \end{aligned}$$

so we have

$$\sum_{k=0}^m \mu_k |f(z_k) - q_n(z_k)|^p \geq \sum_{k=0}^m \mu_k |f(z_k) - p_n(z_k)|^p.$$

We now prove that (3) implies (1). If $f(x) - p_n(x)$ ($\neq 0$ on E) has fewer than $n+1$ strong sign changes on E , there exists a polynomial $q_n(x)$ of degree n which has the same algebraic sign as $f(x) - p_n(x)$ at every point of E where $f(x) - p_n(x) \neq 0$. We readily compute for small ϵ (> 0)

$$\begin{aligned} &\frac{d}{d\epsilon} \sum \mu_k |f(x_k) - p_n(x_k) - \epsilon q_n(x_k)|^p \\ &= -p \sum_+ \mu_k |f(x_k) - p_n(x_k) - \epsilon q_n(x_k)|^{p-1} q_n(x_k) \\ &\quad + p \sum_- \mu_k |f(x_k) - p_n(x_k) - \epsilon q_n(x_k)|^{p-1} q_n(x_k) \\ &\quad + p \sum_0 \mu_k \epsilon^{p-1} |q_n(x_k)|^p, \end{aligned}$$

where the summations are taken over the subsets of E on which $f(x) - p_n(x)$ is respectively positive, negative, or zero. For $\epsilon = 0$ the second member reduces to the first two terms and is obviously negative, so $p_n(x)$ is not extremal. By Theorem 3 it is seen that (2) implies (1). This completes the proof of Theorem 8.

Under the conditions of Theorem 8, any polynomial $p_n(x)$ satisfying (1), (2), or (3) is also a polynomial of best approximation to $f(x)$ on E_0 for suitable weights in the sense of least p th powers ($p > 1$).

However, it may be impossible to choose the set E_0 so that $p_n(x)$ is also extremal ($p > 1$) to $f(x)$ on E_0 with the given weights. For instance, suppose $p = 2$ and E a set of $n + 3$ points on which $f(x) - p_n(x)$ has $n + 2$ strong sign changes; there exist weights μ_k so that $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E in the sense of least squares. Consequently we have the orthogonality of $f(x) - p_n(x)$ on E to unity:

$$\sum_1^{n+3} \mu_k [f(x_k) - p_n(x_k)] = 0.$$

There exists no subset E_0 of $n + 2$ points on which the corresponding equation is valid with these same weights μ_k , so there exists no subset E_0 of $n + 2$ points on which $p_n(x)$ is of best approximation to $f(x)$ with the given weights.

In the exceptional case $f(x) \equiv p_n(x)$ on E , of course $p_n(x)$ is a unique polynomial of best approximation to $f(x)$ on E in the sense of least p th power with arbitrary positive weights, even for arbitrary p (> 0).

For given positive weights, the polynomial of best approximation in the sense of least p th powers ($p > 1$) is unique; if two different approximating polynomials yield the same deviation, half their sum yields a smaller deviation.

8. Tchebycheff approximation. We now prove that Theorem 8 is valid also for the case of approximation with positive weights in the sense of Tchebycheff (denoted by $p = \infty$):

THEOREM 9. *Theorem 8 is valid for the case of Tchebycheff approximation, $p = \infty$.*

That (1) implies (2) and (2) implies (1) follows from Theorem 3; assuming (1) or (2) the difference $f(x) - p_n(x)$ has $n + 1$ strong sign changes on some subset E_0 of E containing $n + 2$ points. If $f(x)$ is represented on E_0 by a polynomial of degree $n + 1$, it follows that $f(x) - p_n(x)$ is on E_0 a constant multiple of a polynomial $T_{n+1}(x) \equiv x^{n+1} + \dots$ which has $n + 1$ strong sign changes on E_0 , so [4, Theorem 5; 3, §1] or by Theorem 3 and a remark made in the Introduction, $T_{n+1}(x)$ is a polynomial of least Tchebycheff norm on E_0 with suitable positive weights, and $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E_0 : $(x_1, x_2, \dots, x_{n+2})$ in the sense of Tchebycheff with suitable

positive weights μ_k respectively. Moreover, on E_0 both $\mu_k |T_{n+1}(x_k)|$ and $\mu_k |f(x_k) - p_n(x_k)|$ are [3, §1] constant.

We adjoin to E_0 all points (x_{n+3}, \dots, x_m) of $E - E_0$, and assign weights μ_k to these new points in such a way that $\mu_k |f(x_k) - p_n(x_k)| < \mu_1 |f(x_1) - p_n(x_1)|$, $k > n+2$. Then the measure of approximation of $p_n(x)$ to $f(x)$ is the same on E and E_0 , and $p_n(x)$ is a polynomial of best approximation in the sense of Tchebycheff to $f(x)$ also on E , so (1) and (2) each implies (3).

If (3) is satisfied with positive weights, suppose $f(x) - p_n(x)$ to have no more than n strong sign changes on E ; there exists a polynomial $q_n(x)$ of degree n which has the same sign as does $f(x) - p_n(x)$ in all points of E where $f(x) - p_n(x) \neq 0$. For sufficiently small ϵ (> 0), the Tchebycheff norm with the given weights of $f(x) - p_n(x) - \epsilon q_n(x)$ is less than that of $f(x) - p_n(x)$, so $p_n(x)$ is not extremal.

Theorem 9 is now established; it will be noticed that under the hypothesis of (1) or (2), *the polynomial $p_n(x)$ is of best approximation in the sense of Tchebycheff to $f(x)$ on E_0 with the chosen weights μ_k , and on E_0 the values $\mu_k |f(x_k) - p_n(x_k)|$ are independent of k .*

The following conclusion is in contrast to the situation for values of p , $0 < p < \infty$, where a polynomial of best approximation on E in the sense of least p th powers is likewise a polynomial of best approximation on some E_0 in the sense of least p th powers, but perhaps only with suitably modified weights.

With the hypothesis (3) it is also true that *$p_n(x)$ is a polynomial of best approximation to $f(x)$ on some subset of $n+2$ points of E on which $f(x) - p_n(x)$ has $n+1$ strong sign changes and on which $\mu_k |f(x_k) - p_n(x_k)|$ is a constant with the given μ_k , namely on a subset of E on which $\mu_k |f(x_k) - p_n(x_k)|$ attains its maximum.* Otherwise there exists a polynomial $q_n(x)$ of degree n which has the same algebraic sign as does $f(x) - p_n(x)$ on all points of E where

$$\mu_k |f(x_k) - p_n(x_k)|$$

attains its maximum; we reach a contradiction as before. This consequence of (3) is classical [Kirchberger (1902), Borel (1905)] in the case that all μ_k are unity.

Conversely, *for fixed weights and some polynomial $p_n(x)$ of degree n , if on some set E_0 of $n+2$ points the values of $\mu_k [f(x_k) - p_n(x_k)]$ are of the same and maximum modulus and of alternating signs, then $p_n(x)$ is extremal on E_0 ; there exists no polynomial $q_n(x)$ of degree n which has the same algebraic sign as $f(x) - p_n(x)$ at all points of E_0 , so for no $q_n(x)$ can we have $|f(x) - p_n(x) - q_n(x)| < |f(x) - p_n(x)|$ on E_0 .*

The uniqueness of the polynomial of degree n of best Tchebycheff approximation to $f(x)$ on E with given positive weights follows from the fact that if two such polynomials $p_n(x)$ and $q_n(x)$ exist, then $r_n(x) \equiv [p_n(x) + q_n(x)]/2$ is also such a polynomial; since $\mu_j [f(x_j) - r_n(x_j)]$ takes the value

$$[\max \mu_k |f(x_k) - p_n(x_k)|] = [\max \mu_k |f(x_k) - q_n(x_k)|]$$

at least $n+2$ times on E but with alternating signs, we must have $r_n(x) = p_n(x) = q_n(x)$ in at least $n+2$ points of E , whence $r_n(x) \equiv p_n(x) \equiv q_n(x)$.

In the exceptional case $f(x) \equiv p(x)$ on E , naturally $p_n(x)$ is the unique polynomial of degree n of best approximation to $f(x)$ on E in the sense of Tchebycheff with arbitrary positive weights.

The present writers have previously [5, Theorem 5.3] proved the particular case relating to infrapolynomials of Theorem 3, but not of Theorems 8 and 9:

THEOREM 10. *Let E be a real finite set containing at least $n+1$ points. The three classes of polynomials $T_n(x) \equiv x^n + \dots$ defined by each of the following properties are identical, and (3) is independent of p , $1 < p \leq \infty$:*

- (1) $T_n(x)$ has at least n strong sign changes on E ;
- (2) $T_n(x)$ is a proper infrapolynomial on some subset E_0 of E containing precisely $n+1$ points;
- (3) $T_n(x)$ is a polynomial of least p th power norm ($1 < p \leq \infty$) on E with some positive weights.

Theorem 10 is a special case of Theorems 8 ($p < \infty$) and 9 ($p = \infty$) where the numbers n are to be adjusted; $f(x)$ in Theorems 8 and 9 is x^n in Theorem 10, and the approximating polynomial in Theorem 10 is $x^n - T_n(x)$, of degree $n-1$.

9. Least p th powers, $0 < p < 1$. We have an analogous theorem which treats approximation in the sense of least p th powers, $0 < p < 1$, precisely as do Theorems 3 and 8-9 for the cases $p = 1$ and $1 < p \leq \infty$:

THEOREM 11. *Let the real function $f(x)$ be defined on the real finite set E containing at least $n+2$ points. The three classes (necessarily finite) of polynomials $p_n(x)$ of degree n defined by each of the following properties are identical, and (3) is independent of p ($0 < p < 1$):*

- (1) $f(x) - p_n(x)$ vanishes in at least $n+1$ points of E ;
- (2) $p_n(x)$ is a juxtapolynomial to $f(x)$ on E equal to $f(x)$ in at least $n+1$ points of E ;
- (3) $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E in the sense of least p th powers ($0 < p < 1$) with some positive weights.

That (1) implies (2) is a consequence of Theorem 2, and that (2) implies (1) is obvious. To prove that (1) implies (3), we choose E_0 , a subset of E containing precisely $n+2$ points, at $n+1$ points of which $f(x) = p_n(x)$. On E_0 we replace $f(x)$ by its equal, a polynomial $P_{n+1}(x)$ of degree $n+1$ [as in 4, §2], so $f(x) - p_n(x)$ is on E_0 a constant multiple of a polynomial $T_{n+1}(x) \equiv x^{n+1} + \dots$ which has $n+1$ zeros on E_0 . Then $T_{n+1}(x)$ is [5, Theorem 1.3] an extremal polynomial on E_0 for least p th power norm with suitable weights,

and $p_n(x)$ is a polynomial of best approximation to $f(x)$ on the set E_0 in the sense of least p th powers, with those same weights. Every point of E other than the $n+1$ zeros assumed to lie in the original E_0 , forms together with those $n+1$ points a possible set E_0 , and hence by Theorem 1 all such points of E not in the original E_0 can be successively adjoined to the original E_0 without altering the extremal character of $p_n(x)$.

We have now shown that (1) implies (3). The converse is immediately available [4, Theorem 6], a result prompted by a remark made to the writers by Professor A. Dvoretzky. The totality of polynomials satisfying (1) is clearly finite. Theorem 11 is established.

Of course under the conditions of Theorem 11 (compare the comments on Theorems 3 and 8) the polynomial $p_n(x)$ is a juxtapolynomial to $f(x)$ on every subset E_0 of E containing at least $n+2$ points and containing $n+1$ points on which $f(x) \equiv p_n(x)$, and $p_n(x)$ is a polynomial of best approximation to $f(x)$ on E_0 in the sense of least p th powers with suitably chosen weights.

In Theorem 11 the polynomial of best approximation need not be unique [5, Theorem 1.3].

Theorem 11 has immediate application to T -polynomials:

THEOREM 12. *Let E be a real finite set containing at least $n+1$ points. The three classes of polynomials $T_n(x) \equiv x^n + \dots$ defined by each of the following properties are identical, and (3) is independent of p , $0 < p < 1$:*

- (1) $T_n(x)$ has n zeros on E ;
- (2) $T_n(x)$ is an *infrapolynomial* on E vanishing in at least n points of E ;
- (3) $T_n(x)$ is a polynomial $x^n + \dots$ of least p th power norm ($0 < p < 1$) on E with suitable positive weights.

BIBLIOGRAPHY

1. Michael Fekete, *On the structure of extremal polynomials*, Proc. Nat. Acad. Sci. U.S.A. vol. 37 (1951) pp. 95-103.
2. ———, *On the structure of polynomials of least deviation*, Bull. Res. Council Israel vol. 5A (1955) pp. 11-19.
3. T. S. Motzkin and J. L. Walsh, *On the derivative of a polynomial and Tchebycheff approximation*, Proc. Amer. Math. Soc. vol. 4 (1953) pp. 76-87.
4. ———, *Least p th power polynomials on a real finite point set*, Trans. Amer. Math. Soc. vol. 78 (1955) pp. 67-81.
5. ———, *The least p th power polynomials on a finite point set*, Trans. Amer. Math. Soc. vol. 83 (1956) pp. 371-396.
6. ———, *Underpolynomials and infrapolynomials*, Illinois J. Math. vol. 1 (1957) pp. 406-426.

UNIVERSITY OF CALIFORNIA,
LOS ANGELES, CALIF.
HARVARD UNIVERSITY,
CAMBRIDGE, MASS.