Chebyshev Approximations of a Function and Its Derivatives

By D. G. Moursund

1. Introduction. This paper considers the problem of the uniform approximation of a function and its first \( r \) derivatives. Several theorems concerning the number and nature of the extremals of a best approximation are obtained. These results are applied to a special case of approximating a function and its first derivative, and a uniqueness theorem is obtained.

2. Statement of the Problem. Let \( X \) be a compact subset of the real line. Let \( n \geq 0 \) and \( r \geq 1 \) be fixed integers. The function \( f(x) \), which is to be approximated, and the base functions \( \phi_0(x), \phi_1(x), \ldots, \phi_n(x) \) are all assumed to have continuous \( r \)th derivatives. Let \( w_0(x), w_1(x), \ldots, w_r(x) \) be positive weight functions, continuous on \( X \).

1. Definition. \( M[g] = \max_{k=0,1,\ldots,r} \| w_k(x)D^kg(x) \| \) where \( \| \| \) denotes the uniform norm on \( X \), and \( g \) has a continuous \( r \)th derivative.

2. Problem. Find real scalars \( a_0, a_1, \ldots, a_n \) such that \( M[\sum a_k \phi_k(x) - f(x)] \) is a minimum.

The function \( f \) may be given by a table or by other means. By the statement \( g = 0 \) we shall mean that \( D^kg(x) = 0 \) for \( k = 0, 1, \ldots, r \) and for all \( x \) in \( X \). The functional \( M \) is a norm, so that

3. \( 0 \leq M[g] < \infty \).

4. \( M[g] = 0 \) iff \( g = 0 \).

5. \( M[tg] = | t | M[g] \) where \( t \) is any real number.

6. \( M[g + h] \leq M[g] + M[h] \).

Points in Euclidean \( n + 1 \) space are represented by \( \alpha = (a_0, a_1, \ldots, a_n) \), \( \beta = (b_0, b_1, \ldots, b_n) \) etc., while polynomials are given by \( P(x, \alpha) = \sum a_k \phi_k(x) \), etc. In addition we let \( e = \inf M[P(x, \alpha) - f] \) where the infimum is taken over all \( \alpha \) in \( n + 1 \) space, and let \( R = \{ \alpha \in E^{n+1} : M[P(x, \alpha) - f] = e \} \). The base functions \( \phi_0, \phi_1, \ldots, \phi_n \) are assumed to be linearly independent in the sense that \( P(x, \alpha) = 0 \) only if \( \alpha = 0 \).

Using the fact that the norm \( M \) is a continuous linear functional on \( E^{n+1} \) it follows that the set \( R \) of best approximations is closed, bounded, convex and non-empty. These are standard results, and proofs may be found in Achieser [1] and Buck [2].

3. Example. The following example was chosen to illustrate the difference between approximating a function and its derivatives, and ordinary Chebyshev approximation. Let \( w_0 = 1, w_1 = 1, r = 1 \) and suppose that \( f \) and \( Df \) are given by Table 7. The problem is to find \( a_0, a_1, \ldots, a_n \) for various \( n \), so that

\[
M[a_0 + a_1 x + \cdots + a_n x^n - f]
\]

is a minimum.

Received July 10, 1963.
First consider the case \( n = 4 \). For notational convenience let \( x_1 = -1 \), \( x_2 = -1/\sqrt{3} \), \( x_3 = 0 \), \( x_4 = 1/\sqrt{3} \) and \( x_5 = 1 \). Since \( M[f] = 1 \), the set \( R \) of best approximations must be contained in the region of Euclidean 5-space whose points \( \alpha \) satisfy the relation \( M[P(x, \alpha) - f] \leq 1 \). This region is defined by the inequalities

(1) \(-1 \leq P(x_i, \alpha) - f(x_i) \leq 1\), \( i = 1, 2, \ldots, 5 \),
(2) \(-1 \leq DP(x_i, \alpha) - Df(x_i) \leq 1\), \( i = 1, 2, \ldots, 5 \).

Consider the following subsystem of inequalities:

(3) \( a_0 - a_1 + a_2 - a_3 + a_4 \geq 0 \) from (1), \( i = 1 \),
(4) \(-a_0 - a_1 - a_2 - a_3 - a_4 \leq 0 \) from (1), \( i = 5 \),
(5) \( a_1 - 2a_2/\sqrt{3} + a_3 - 4a_4/3\sqrt{3} \geq 0 \) from (2), \( i = 2 \),
(6) \( a_1 + 2a_2/\sqrt{3} + a_3 + 4a_4/3\sqrt{3} \leq 0 \) from (2), \( i = 4 \).

From adding (4), (5), (6) and comparing the result with (3) it follows that any solution to the system (1), (2) must satisfy

(7) \( a_0 - a_1 + a_2 - a_3 + a_4 = 0 \),
(8) \(-a_0 - a_1 - a_2 - a_3 - a_4 = 0 \),
(9) \( a_1 - 2a_2/\sqrt{3} + a_3 - 4a_4/3\sqrt{3} = 0 \),
(10) \( a_1 + 2a_2/\sqrt{3} + a_3 + 4a_4/3\sqrt{3} = 0 \).

From (7), (8) we get \( a_0 + a_2 + a_4 = 0 \) and \( a_1 + a_3 = 0 \), while from (9), (10) it follows that \( a_2 + 2a_4/3 = 0 \). Thus polynomials satisfying (1), (2) must be of the form

\[ P(x, \alpha) = s(x - x^3) + t(-\frac{1}{3} - 2x^2/3 + x^4) \]

where \( s \) and \( t \) remain to be determined. Table 8 gives the results thus far obtained.

<table>
<thead>
<tr>
<th>( x )</th>
<th>-1</th>
<th>(-1/\sqrt{3})</th>
<th>0</th>
<th>(1/\sqrt{3})</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P - f )</td>
<td>-1</td>
<td>(-2s/3\sqrt{3} - 4t/9)</td>
<td>(-t/3 + 1)</td>
<td>(2s/3\sqrt{3} - 4t/9)</td>
<td>1</td>
</tr>
<tr>
<td>( D(P - f) )</td>
<td>(-2s - 8t/3)</td>
<td>(-t/3 + 1)</td>
<td>(-1)</td>
<td>(-2s - 8t/3)</td>
<td>(-1)</td>
</tr>
</tbody>
</table>

It is now evident that if \( P(x, \alpha) \) is a best approximation to \( f \), then \( M[P(x, \alpha) - f] = 1 \). Moreover every polynomial of the form \( P(x, \alpha) = s(x - x^3) + \)

License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use
\[ t(-\frac{1}{3} - 2x^2/3 + x^4) \] is a best approximation provided

\[ -1 \leq -2s/3\sqrt{3} - 4t/9 \leq 1, \]
\[ -1 \leq -t/3 + 1 \leq 1, \]
\[ -1 \leq 2s/3\sqrt{3} - 4t/9 \leq 1, \]
\[ -1 \leq -2(3s + 4t)/3 \leq 1, \]
\[ -1 \leq s \leq 1, \]
\[ -1 \leq 2(-3s + 4t)/3 \leq 1. \]

The solution to this system is a triangular region in the \( s, t \) plane defined by \( t \geq 0, -3s + 4t \leq \frac{1}{3}, \) and \( 3s + 4t \leq \frac{1}{3} \). Thus the case \( n = 4 \) provides an example of a two-dimensional solution space.

To solve the case \( n = 3 \) it is merely necessary to note that the coefficient of \( x^4 \) must be zero. Hence all solutions are given by \( P(x, \alpha) = s(x - x^5), -\frac{1}{3} \leq s \leq \frac{1}{3}. \) For the cases \( n = 0, 1, 2 \) it is evident that the only solution is \( s = 0, t = 0, \) so that \( P(x, \alpha) = 0 \) is the unique best approximation.

For \( n > 4 \) the problem can be formulated as a linear programming problem, and solved using standard techniques. Solutions for these cases are given in 10–12., with a summary of the results in Table 9. The values of \( e \) have been rounded to four places, while the coefficients of the polynomials are given to five decimal places.

The case \( n = 6 \) has a solution space which is at least one dimensional. The solution to \( n = 5 \) is a point in the solution space to \( n = 6, \) so these two have been combined in 10.

### Table

<table>
<thead>
<tr>
<th>( n \leq 4 )</th>
<th>( n = 5 )</th>
<th>( n = 6 )</th>
<th>( n = 7 )</th>
<th>( n = 8 )</th>
<th>( n \geq 9 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e = 1.0 )</td>
<td>( e = .8632 )</td>
<td>( e = .8632 )</td>
<td>( e = .5222 )</td>
<td>( e = .4227 )</td>
<td>( e = 0 )</td>
</tr>
</tbody>
</table>

10. \( n = 5, \) \( n = 6. \) \( P(x) = -.13681 - .86319x + .27362x^2 + 1.81595x^3 - .13681x^4 - 1.08957x^5. \)

<table>
<thead>
<tr>
<th>( x )</th>
<th>(-1)</th>
<th>(-1/\sqrt{3})</th>
<th>(0)</th>
<th>(1/\sqrt{3})</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P - f )</td>
<td>(-.8632)</td>
<td>(.1580)</td>
<td>(.8632)</td>
<td>(-.2796)</td>
<td>(.8632)</td>
</tr>
<tr>
<td>( D(P - f) )</td>
<td>(-.8632)</td>
<td>(-.8632)</td>
<td>(-.8632)</td>
<td>(-.4419)</td>
<td>(-.8632)</td>
</tr>
</tbody>
</table>

11. \( n = 7. \) \( P(x) = -.47789 - .52211x + 3.24209x^2 + 10.49572x^3 - 6.54944x^4 - 22.54194x^5 + 3.37262x^6 + 11.67781x^7. \)

<table>
<thead>
<tr>
<th>( x )</th>
<th>(-1)</th>
<th>(-1/\sqrt{3})</th>
<th>(0)</th>
<th>(1/\sqrt{3})</th>
<th>(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P - f )</td>
<td>(-.5221)</td>
<td>(-.5222)</td>
<td>(.5221)</td>
<td>(.5222)</td>
<td>(-.3031)</td>
</tr>
<tr>
<td>( D(P - f) )</td>
<td>(-.5221)</td>
<td>(-.5220)</td>
<td>(-.5221)</td>
<td>(-.5220)</td>
<td>(.5221)</td>
</tr>
</tbody>
</table>
12. \( n = 8 \). \( P(x) = -1.42265 - 4.2265x + 11.38120x^2 - 31.29829x^3 - 16.74871x^4 + 34.14359x^5 + 8.49038x^6 + 8.10363x^7 + 1.4226x^8 \)

<table>
<thead>
<tr>
<th>( x )</th>
<th>( P - f )</th>
<th>( \frac{P - f}{D(P - f)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-0.4226</td>
<td>-0.4226</td>
</tr>
<tr>
<td>-0.4226</td>
<td>-0.4226</td>
<td>-0.4226</td>
</tr>
<tr>
<td>0</td>
<td>0.4226</td>
<td>0.4226</td>
</tr>
<tr>
<td>0.4226</td>
<td>0.4226</td>
<td>0.4226</td>
</tr>
<tr>
<td>1</td>
<td>0.4226</td>
<td>0.4226</td>
</tr>
</tbody>
</table>

4. **Characterization.** Instead of considering just one error function \( P - f \), as in ordinary Chebyshev approximation, we must consider \( r + 1 \) weighted error functions.

13. **Definition.** \( L_k(x, \alpha) \equiv w_k(x)D^k[P(x, \alpha) - f(x)] \). The functions \( L_k(x, \alpha) \), defined for all \( \alpha \) in Euclidean \( n + 1 \) space and \( k = 0, 1, \ldots, r \) are called **weighted error functions**.

14. **Definition.** Suppose \( M[P(x, \alpha) - f] = d \). The pair \((x_0, k)\), where \( x_0 \in X \) and \( 0 \leq k \leq r \) is an integer, is called an extremal with respect to the approximation \( P(x, \alpha) \) to \( f \) if \( |L_k(x_0, \alpha)| = d \).

Because \( X \) is compact and the functions \( L_k(x, \alpha) \) are continuous, it is evident that every approximation has one or more extremals. We shall now establish some results concerning the extremals of a best approximation.

15. **Definition.** If \( \alpha \in R \), the following notation will be used: \( C_k(\alpha) = \{x \in X : |L_k(x, \alpha)| = e\} \).

16. **Theorem.** There exists an \( \alpha \in R \) such that for every \( \beta \in R \), \( C_k(\alpha) \subseteq C_k(\beta) \) for \( k = 0, 1, \ldots, r \).

**Proof.** Let \( m \) be the dimension of the convex set \( R \). If \( m = 0 \), \( R \) is a single point, the best approximation is unique, and the theorem is true. If \( m \geq 1 \) then \( R \) has a non-empty interior. Let \( \alpha \) be an arbitrary point in the interior of the set \( R \). It will be shown that this point satisfies the assertion of the theorem.

Let \( \beta \neq \alpha \) be an arbitrary point of \( R \). Extend the line \( \overline{\alpha \beta} \) to the boundary of \( R \), calling the points of intersection \( \alpha_1 \) and \( \beta_1 \) as indicated by Figure 17. If \( \beta \) is a boundary point of \( R \), then \( \beta_1 = \beta \).

Now consider the functions \( L_k[x, q\alpha_1 + (1 - q)\beta_1] \) where \( 0 < q < 1 \). If \( D^kP(x_0, \alpha_1) \neq D^kP(x_0, \beta_1) \) then one of the following relations holds:

(17) \( L_k(x_0, \alpha_1) < L_k[x_0, q\alpha_1 + (1 - q)\beta_1] < L_k(x_0, \beta_1) \),

(18) \( L_k(x_0, \beta_1) < L_k[x_0, q\alpha_1 + (1 - q)\beta_1] < L_k(x_0, \alpha_1) \).
Hence if \( q \in (0, 1) \) is arbitrary then \( x_0 \in C_k[q\alpha_1 + (1 - q)\beta_1] \) if and only if \( x_0 \in C_k(\alpha_1), x_0 \in C_k(\beta_1) \) and \( D^kP(x_0, \alpha_1) = D^kP(x_0, \beta_1) \).

Since \( \alpha \) is an interior point of the line segment \( \alpha_0\beta_1 \), it follows from the above result that if \( x_0 \in C_k(\alpha) \) then \( x_0 \in C_k(\alpha_1), x_0 \in C_k(\beta_1) \) and \( D^kP(x_0, \alpha_1) = D^kP(x_0, \beta_1) \). This implies, using the result again, that \( x_0 \in C_k[q\alpha_1 + (1 - q)\beta_1] \) for \( 0 \leq q \leq 1 \), which completes the proof.

18. Corollary. Suppose the best approximation is not unique, and let \( \alpha \) be a point in the interior of \( R \). Then if \( P(x, \beta) \) is any best approximation, \( D^k(x_0, \alpha) = D^k(x_0, \beta) \) for every extremal \( (x_0, k) \) of the approximation \( P(x, \alpha) \) to \( f \).

Given an approximation \( P(x, \alpha) \) to \( f \) we would like to have a means of telling if it is a best approximation. In addition, if \( P(x, \alpha) \) is not a best approximation we would like to be able to find a better approximation. The following definition is useful in this endeavor.

19. Definition. Suppose \( M[P(x, \alpha) - f] = d \). A polynomial \( P(x, \beta) \) is said to satisfy Condition A with respect to the approximation \( P(x, \alpha) \) to \( f \), if for each extremal \( (x_0, k) \) of this approximation, \( \text{sgn} D^kP(x_0, \beta) = -\text{sgn} D^k[P(x_0, \alpha) - f(x_0)] \).

20. Theorem. A polynomial \( P(x, \alpha) \) is a best approximation to \( f \) if and only if there is no polynomial \( P(x, \beta) \) which satisfies Condition A with respect to the approximation.

Proof. Suppose first that \( P(x, \alpha) \) is not a best approximation, so that there exists a polynomial \( P(x, \beta) \) such that \( M[P(x, \beta) - f] < M[P(x, \alpha) - f] \). Then the polynomial \( P(x, \beta - \alpha) \) satisfies Condition A with respect to the approximation \( P(x, \alpha) \).

Next suppose that \( P(x, \beta) \) satisfies Condition A with respect to the approximation \( P(x, \alpha) \). Using the compactness of \( X \) and the continuity of the functions \( D^kP(x, \beta) \) and \( L_k(x, \alpha) \), one can establish that for each \( k \) there exists a constant \( T_k > 0 \) such that if \( 0 < t \leq T_k \) then \( ||L_k(x, \alpha + t\beta)|| < M[P(x, \alpha) - f] \). A proof of this result may be found in Remez [3, p. 38]. Hence if \( 0 < t \leq \min \{T_0, T_1, \ldots, T_r\} \) then \( M[P(x, \alpha + t\beta) - f] < M[P(x, \alpha) - f] \).

This theorem provides the basis for a computational scheme similar to the first method of Remez [3, p. 36]. In addition it provides a useful tool for investigating the uniqueness of a best approximation.

We know that if a differentiable function has a relative extremum at a point interior to its domain of definition, then the derivative must be zero at that point. Under proper assumptions this allows us to derive additional conditions which a best approximation must satisfy.

21. Theorem. Let \( (x_0, k) \) be an extremal of the approximation \( P(x, \alpha) \) to \( f \), with \( M[P(x, \alpha) - f] = d \). Suppose that at the point \( x_0 \) both \( D^{k+1}[P(x, \alpha) - f] \) and \( Dw_k(x) \) exist. If for every \( \epsilon > 0 \) there exist points \( x_1, x_2 \) in \( X \) such that \( x_0 - \epsilon < x_1 < x_0 < x_2 < x_0 + \epsilon \) then \( DL_k(x_0, \alpha) = 0 \) at the point \( x_0 \).

Proof. Suppose that \( L_k(x_0, \alpha) = d \). Then

\[
DL_k(x_0, \alpha) = \lim_{x \to x_0} \frac{L_k(x, \alpha) - d}{x - x_0}.
\]

For all \( x \), \( L_k(x, \alpha) - d \leq 0 \). Hence if \( x < x_0 \) then \( (L_k(x, \alpha) - d)/(x - x_0) \geq 0 \) while if \( x > x_0 \) then \( (L_k(x, \alpha) - d)/(x - x_0) \leq 0 \). Since the approach to the limit
may be made from either side of \( x_0 \), through points of \( X \), it follows that the limit is zero. A similar argument holds if \( L_k(x_0, \alpha) = -d \) so the proof is complete.

22. Corollary. Suppose that \( P(x, \alpha) \) is a best approximation to \( f \), with \( \alpha \) in the interior of \( R \), and that the hypotheses of Theorem 21 are satisfied at the point \((x_0, k)\). Then for any other best approximation \( P(x, \beta) \) it follows that

\[
D^{k+1}P(x_0, \alpha) = D^{k+1}P(x_0, \beta).
\]

Proof. From Corollary 18 we know that \((x_0, k)\) is an extremal of the approximation \( P(x, \beta) \) to \( f \). Hence

\[
DL_k(x_0, \alpha) = 0 = DL_k(x_0, \beta).
\]

Using the product rule for differentiation we have

\[
Dw_k(x_0)D^k[P(x_0, \alpha) - f(x_0)] + w_k(x_0)D^{k+1}[P(x_0, \alpha) - f(x_0)] = 0,
\]

\[
Dw_k(x_0)D^k[P(x_0, \beta) - f(x_0)] + w_k(x_0)D^{k+1}[P(x_0, \beta) - f(x_0)] = 0.
\]

Since \( D^k[P(x_0, \alpha) - f(x_0)] = D^k[P(x_0, \beta) - f(x_0)] \) and \( w_k(x_0) \neq 0 \) the result is established.

5. Uniqueness. The previous theorems give us considerable information about the extremals of a best approximation. We shall use these results to establish the following theorem.

Let \( X = [-1, 1] \), \( r = 1, \phi_i = x^i, i = 0, 1, \cdots, n \), and suppose \( w_0(x), w_1(x), \) and \( f''(x) \) exist. The weight functions are assumed to be positive.

23. Theorem. Under the above conditions one of the following assertions is true:

(19) The best approximation is unique.

(20) The best approximation is unique except for an additive constant; moreover, if \( P \) is any best approximation then \( DP \) is the unique best Chebyshev approximation of degree \( n - 1 \) to \( Df \) with weight function \( w_0(x) \).

We shall first discuss the conditions under which the second assertion will be true. Suppose there exists a best approximation \( P(x, \gamma) \), with \( M[P(x, \gamma) - f] = e \), such that \( \| w_0(x)[P(x, \gamma) - f(x)] \| < e \). Then \( \| w_1(x)D[P(x, \gamma) - f(x)] \| = e \). Now suppose that assertion (20) is false, so that there exists a polynomial \( P(x, \alpha) \) such that \( \| w_1(x)D[P(x, \alpha) - f(x)] \| < e \).

Define a polynomial

\[
P_q(x) = q \int_0^x DP(x, \alpha) \, dx + (1 - q)P(x, \gamma), \quad \text{for} \quad 0 \leq q \leq 1.
\]

Then

\[
\| w_0(x)[P_q(x) - f(x)] \| \leq q \| w_0(x) \left[ \int_0^x DP(x, \alpha) \, dx - f(x) \right] \|
\]

\[
+ (1 - q) \| w_0(x)[P(x, \gamma) - f(x)] \|.
\]

Hence there exists a \( t > 0 \) such that if \( 0 < q \leq t \) then \( \| w_0(x)[P_q(x) - f(x)] \| < e \). However, \( \| w_1(x)D[P_q(x) - f(x)] \| \leq q \| w_1(x)D[P(x, \alpha) - f(x)] \| + (1 - q) \| w_1(x)D[P(x, \gamma) - f] \| < e \). Hence if \( q < t \) then \( P_q(x) \) is a better approximation to \( f \) than is \( P(x, \gamma) \); this is a contradiction. Therefore, if there exists any best approximation \( P(x, \gamma) \) such that \( \| L_0(x, \gamma) \| < e \) then (20) is true.
From the theory of equations we know that if two polynomials of degree \( n \) agree on a set of \( s + t \) points, where \( s > 0 \) and \( t \geq 0 \), if their first derivatives agree on a subset of \( t \) of these points and if \( s + 2t \geq n + 1 \), then the polynomials are identical. In addition if \( s + t \) distinct points are specified, with \( s > 0 \), then it is possible to find a polynomial of degree \( \leq n \) which has arbitrary values on the \( s + t \) points, and whose derivative has arbitrary values on a subset of \( t \) of these points, provided \( s + 2t \leq n + 1 \). For the remainder of the proof we shall assume, for each \( \alpha \in R \), that \( ||L_0(x, \alpha)|| = \epsilon \). Let \( \beta \) be a point in the interior of \( R \). For notational convenience define:

\[
\begin{align*}
S &= C_0(\beta) \\
T &= C_1(\beta) \\
U &= S \cap T \cap (-1, 1).
\end{align*}
\]

Let \( s, t, u \) be the cardinality of \( S, T, U \) respectively.

Using Corollary 18 and Corollary 22 it follows that if \( P(x, \alpha) \) is any best approximation then

\[
\begin{align*}
(25) & \quad P(x, \beta) = P(x, \beta) \quad \forall \ x \in S, \\
(26) & \quad DP(x, \alpha) = DP(x, \beta) \quad \forall \ x \in S \cap (-1, 1), \\
(27) & \quad DP(x, \alpha) = DP(x, \beta) \quad \forall \ x \in T, \\
(28) & \quad D^2P(x, \alpha) = D^2P(x, \beta) \quad \forall \ x \in T \cap (-1, 1).
\end{align*}
\]

If we assume that \( \alpha \) and \( \beta \) are distinct we can now give a lower bound for \( n \), the degree of the approximating polynomials, in terms of \( s, t, u \). We can compute an upper bound for the degree of a polynomial which satisfies Condition A with respect to the approximation \( P(x, \beta) \), in terms of the same \( s, t, u \). We shall show that the former is \( \geq \) the latter, leading to a contradiction of the assumption \( \alpha \neq \beta \), and completing the proof.

Since the points \( +1 \) and \( -1 \) can be in \( S \) and \( T \) independently, it is necessary to consider each possible placement of these points as a separate case. The arguments involved in each case are similar, so that only one example will be given. A summary of all cases is given in Table 24.

Consider the case of \( -1 \in S, +1 \in S, -1 \in T, +1 \in T \). Then

\[
\begin{align*}
(29) & \quad DP(x, \alpha) = DP(x, \beta) \quad \text{for the} \quad s - 2 \quad \text{points} \quad x \in S \cap (-1, 1), \\
(30) & \quad DP(x, \alpha) = DP(x, \beta) \quad \text{for the} \quad t \quad \text{points} \quad x \in T, \\
(31) & \quad D^2P(x, \alpha) = D^2P(x, \beta) \quad \text{for the} \quad t - 2 \quad \text{points} \quad x \in T \cap (-1, 1).
\end{align*}
\]

Since \( U = S \cap T \cap (-1, 1) \), all of the relations (29) are identical to those given in (30). Hence the polynomials \( DP(x, \alpha) \) and \( DP(x, \beta) \), which are of degree \( n - 1 \), agree on a set of \( (s - 2 - u) + t \) points, and their derivatives agree on a subset of \( t - 2 \) of these points. Hence if \( n - 1 \leq (s - 2 - u) + t + (t - 2) - 1 \) then \( DP(x, \alpha) \equiv DP(x, \beta) \) and \( \alpha = \beta \). That is, the assumption \( \alpha \neq \beta \) forces the conclusion \( n \geq s + 2t - u - 3 \).

A polynomial \( P(x, \gamma) \) will satisfy Condition A with respect to \( P(x, \beta) \) if
\[ \text{sgn } P(x, \gamma) = -\text{sgn } L_0(x, \beta) \quad \text{for } x \in S, \]
\[ \text{sgn } DP(x, \gamma) = -\text{sgn } L_1(x, \beta) \quad \text{for } x \in T. \]

In the case we are considering there are \( s \) points in \( S \), \( t \) points in \( T \) and \( t - u - 2 \) points which are in \( T \) but not in \( S \). Thus a polynomial which is to have arbitrary values on \( S \cup T \), and whose derivative is to have arbitrary values on \( T \), must satisfy a total of \( s + (t - u - 2) + t \) conditions. It follows that there is a polynomial \( P(x, \gamma) \) of degree \( \leq s + 2t - u - 3 \) which satisfies Condition A. This completes the proof in this case.

24. Table.

<table>
<thead>
<tr>
<th>Case</th>
<th>( \alpha = \beta ) unless ( n \geq )</th>
<th>Condition A is satisfied by a polynomial of degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-1 \in S, +1 \in S, -1 \in T, +1 \in T)</td>
<td>( s + 2t - u - 3 )</td>
<td>( s + 2t - u - 3 )</td>
</tr>
<tr>
<td>(-1 \in S, -1 \in T, +1 \in T)</td>
<td>( s + 2t - u - 2 )</td>
<td>( s + 2t - u - 2 )</td>
</tr>
<tr>
<td>(+1 \in S, -1 \in T, +1 \in T)</td>
<td>( s + 2t - u - 2 )</td>
<td>( s + 2t - u - 2 )</td>
</tr>
<tr>
<td>(-1 \in S, +1 \in S, -1 \in T)</td>
<td>( s + 2t - u - 1 )</td>
<td>( s + 2t - u - 1 )</td>
</tr>
<tr>
<td>(-1 \in S, +1 \in S, +1 \in T)</td>
<td>( s + 2t - u )</td>
<td>( s + 2t - u )</td>
</tr>
<tr>
<td>(+1 \in S, +1 \in T)</td>
<td>( s + 2t - u - 1 )</td>
<td>( s + 2t - u - 1 )</td>
</tr>
<tr>
<td>(-1 \in S, +1 \in T)</td>
<td>( s + 2t - u )</td>
<td>( s + 2t - u )</td>
</tr>
<tr>
<td>(+1 \in S)</td>
<td>( s + 2t - u - 1 )</td>
<td>( s + 2t - u - 1 )</td>
</tr>
</tbody>
</table>

5. Remarks. The theorem just proved is, of course, true if the interval \([-1, 1]\) is replaced by any finite interval. A valuable application of this theorem is the case of \( \omega_0(x) = 1 \) and \( \omega_1(x) = k \), a positive constant. In this case we can readily determine which of (19), (20) characterizes the solution to the approximation problem. Let \( P(x) \) be such that \( DP(x) \) is the unique best Chebyshev approximation of degree \( n - 1 \) to \( Df(x) \). Pick the constant coefficient of \( P \) so that \( \| P - f \| \) is as small as possible. Then if \( \| P - f \| \geq k \| D(P - f) \| \) it follows that the best approximation will be unique; otherwise, (20) characterizes the solution. The computational aspects of this problem are quite interesting. A computational scheme will be presented in a later paper.

Michigan State University
East Lansing, Michigan