JOURNAL OF APPROXIMATION THEORY 24, 273–288 (1978)

A Characterization of the Lagrange Interpolating Projection with Minimal Tchebycheff Norm

THEODORE A. KILGORE

408 Moody Street, League City, Texas 77573

Communicated by E. W. Cheney

Received March 12, 1977

1. INTRODUCTION AND STATEMENT OF THE PROBLEM

Lagrange interpolation is one of the oldest methods for approximating an arbitrary function in $C[a, b]$, the space of continuous functions on the closed interval $[a, b]$ normed with the familiar sup-norm, by means of an element of $\prod_n [a, b]$, the subspace consisting of polynomials of degree n or less. One chooses $n + 1$ points, called nodes, $t_0, t_1, \ldots, t_n$, with $a \leqslant t_0 < t_1 < \cdots < t_n \leqslant b$. Then the ith *Lagrange polynomial* of degree n is defined as

$$y_i(t) = \prod_{\substack{j=0 \\ j \neq i}}^{n} \frac{t - t_j}{t_i - t_j} \qquad (0 \leqslant i \leqslant n).$$

It is easily seen that $y_i(t_j) = 0$ for $i \neq j$, and $y_i(t_i) = 1$; also that $y_0, \ldots, y_n$ form a basis for $\prod_n [a, b]$, and that $\sum_{i=0}^{n} y_i = 1$. The *Lagrange interpolating projection* is the operator L: $C[a, b] \to \prod_n[a, b]$, given by $Lf = \sum_{i=0}^{n} f(t_i)\, y_i$. The equality $\| L \| = \| \sum_{i=0}^{n} | y_i | \|$ is easily established. The expression $\sum_{i=0}^{n} | y_i |$ is called the *Lebesgue function of* L and will be denoted by Λ. For $n \geqslant 2$, $\Lambda(t) \geqslant 1$, and $\Lambda(t) = 1$ only when t is a node. Between t_{i-1} and t_i, for $1 \leqslant i \leqslant n$, Λ is a polynomial, whose analytic continuation we denote by X_i, and Λ (or now X_i) has a unique local maximum $\tau_i \in (t_{i-1}, t_i)$, at which $X_i'(\tau_i) = 0$. Thus we may specifically write $X_i = \sum_{j=0}^{n} y_j \operatorname{sgn} y_j(\tau_i)$. It is also convenient to denote the value $X_i(\tau_i)$ of the local maximum by a single symbol λ_i.

Since the efficiency of approximation by any linear projection P onto a subspace Y of $C[a, b]$ is governed by the inequality

$$\| f - Pf \| \leqslant (1 + \| P \|)\, d(f, Y),$$

it is desirable to minimize $\| L \|$, which, in the case of Lagrange interpolation,

0021-9045/78/0244-0273\$02.00/0

depends exclusively upon strategic placement of the nodes. Now, let $[a, b]$ and $[c, d]$ be any two intervals. The affine transformation of $[a, b]$ to $[c, d]$ induces an isometry between $C[a, b]$ and $C[c, d]$ which carries $\Pi_n [a, b]$ to $\Pi_n[c, d]$. If $P : C [a, b] \to \Pi_n [a, b]$, we may thus obtain a projection $P' : C[c, d] \to \Pi_n[c, d]$ with $\| P' \| \leq \| P \|$, taking c and d to be the extremities of the carrier of P. Therefore without loss of generality we restrict our attention to those nodal configurations where $a = t_0$ and $b = t_n$. (Alternatively, we may fix any two of $t_0 ,..., t_n$ while varying some or all of the others, in order to observe the resulting behavior of $\lambda_1 ,..., \lambda_n$ [7, 14, 27].)

Another simple fact is that if any two nodes are moved toward one another, we have $\| L \| \to \infty$, and, when, for some $i \in \{1,..., n - 1\}$ all nodes save t_i are fixed, λ_i is a strictly increasing function of t_i, and λ_{i+1} is strictly decreasing. In all likelihood, these facts were the ones which in 1931 led Bernstein [5] to conjecture (assuming $a < t_0$ and $t_n < b$) as follows:

> Il parait probable que les plus grand des $n + 2$ maxima de $F(x)$ [here, $\Lambda(t)$], correspondant a tous ces intervalles sera minimé, lorsque tous les maxima seront égaux. Mais je n'ai pu prouver cette affirmation que sous la condition que n croit indéfiniment, et ce n'est que ce dernier cas que nous examinerons actuellement.

We present here three results which among other things completely settle the conjecture of Bernstein.

THEOREM 1. *In order that the Lagrange interpolation from* $C[a, b] \to \Pi_n[a, b]$ *on* $a = t_0 < t_1 < \cdots < t_n = b$ *have minimal norm among all interpolating projections onto* $\prod_n [a, b]$, *it is necessary that* $\lambda_1 = \lambda_2 = \cdots = \lambda_n$.

We denote by C_n the norm of any interpolating projection from $C[a, b]$ onto $\Pi_n[a, b]$ which has minimal norm. In accordance with standard usage, C_n is called the *projection constant for Lagrange interpolation of degree n*. We then have

THEOREM 2. *If* $\lambda_1 = \cdots = \lambda_{n-1} = c$, *for some constant* c, *then* $c > C_{n-1}$. *In particular,* C_n *is a strictly increasing function of* n.

THEOREM 3. *There is a unique configuration of nodes* $a = t_0 < t_1 < \cdots < t_n = b$, *such that Lagrange interpolation on those points yields* $\lambda_1 = \lambda_2 = \cdots = \lambda_n$.

2. Proofs of the Theorems

We begin by noting that the function $(t_0 ,..., t_n) \to (\lambda_1 ,..., \lambda_n)$ is differentiable, and that the derivative is given by the Jacobian $(\partial\lambda_i/\partial t_j)$.

The rank of this matrix is not more than $n - 1$, since the space Π_n is closed under affine transformations of t, as previously mentioned. For the same reason, any system of $n - 1$ rows is equivalent to any other. Thus, in what follows, we fix the nodes t_0 and t_n. If, however, it facilitates comprehension or avoids unpleasant calculations, we fix any convenient pair of nodes for purposes of discussion.

In investigating the rank of the Jacobian matrix, we may employ certain simplifications, the first of which is the formula

$$\frac{\partial \lambda_i}{\partial t_j} = -y_j(\tau_i)\, X_i'(t_j) \qquad (0 \leqslant j \leqslant n,\ 1 \leqslant i \leqslant n).$$

The elegance and symmetry of this expression are enhanced by the fact that τ_i may be treated in the formula as a fixed point rather than as a variable. The second simplification, communicated to me orally by Dr. Dietrich Braess, is to cancel the denominator $\prod_{k\neq j\ k+0}^{n} (t_j - t_k)$ of y_j from each entry in the jth row of our matrix. Then we may divide each entry in the ith column by the nonzero expression $\prod_{k=0}^{n} (\tau_i - t_k)$, resulting at last in the matrix

$$A = \begin{bmatrix} \dfrac{X_1'(t_0)}{t_0 - \tau_1} & \cdots & \dfrac{X_n'(t_0)}{t_0 - \tau_n} \\ \dfrac{X_1'(t_1)}{t_1 - \tau_1} & \cdots & \dfrac{X_n'(t_1)}{t_1 - \tau_n} \\ \cdot & \cdots & \cdot \\ \dfrac{X_1'(t_n)}{t_n - \tau_1} & \cdots & \dfrac{X_n'(t_n)}{t_n - \tau_n} \end{bmatrix}.$$

When the matrix has been reduced to this new form, we may define polynomials $q_1, \ldots, q_n$ by $q_i(t) = X_i'(t)/t - \tau_i$, since τ_i is a root of X_i'.

This matrix, originating from a nodal configuration $t_0, \ldots, t_n$, may now be considered as a matrix in which given polynomials $q_1, \ldots, q_n$ are evaluated at points $t_0, \ldots, t_n$. In other words, we may, in this new formulation of the problem, free the points $t_0, \ldots, t_n$ from their original role as nodes. We note that the degree of q_i, for $1 \leqslant i \leqslant n$, does not exceed $n - 2$.

Proof of Theorem 1. By lemmas 1 through 8 to follow, any $n - 1$ of $q_1, \ldots, q_n$ form a basis for Π_{n-2}, given an arbitrary configuration of the nodes. Hence, any $(n - 1) \times (n - 1)$ square submatrix of A is nonsingular. From this it follows that, given any initial nodal configuration, we can produce a perturbation of some or all of any list of $n - 1$ nodes (leaving any pair of nodes fixed), which causes the decrease of any desired $n - 1$ of $\lambda_1, \ldots, \lambda_n$. Thus we may set $t_0 = a$ and $t_n = b$, and vary $t_1, \ldots, t_{n-1}$ from any initial configuration, decreasing $\max_{1\leqslant i\leqslant n} \lambda_i$ until a situation is reached where

$\lambda_1 = \lambda_2 = \cdots = \lambda_n$. Therefore, if $\max_{1 \leqslant i \leqslant n} \lambda_i = \| A \|$ is minimized, it is necessary that $\lambda_1 = \lambda_2 = \cdots = \lambda_n$. ∎

Proof of Theorem 2. Assume an initial configuration $(t_0', ..., t_{n-1}')$; of the nodes in which $t_0' = -1$ and $t_{n-1}' = 1$. Viewing t_0 and t_{n-1} as the fixed pair of nodes, denote $\lambda_i(t_1', ..., t_{n-2}', t_n')$ by λ_i for $1 \leqslant i \leqslant n$. Then, since $t_0 < \tau_1 < t_1 < \tau_2 < t_2 < \cdots < t_{n-1} < \tau_n < t_n$ always, we may apply Lemma 9 to show that the matrix

$$B = \begin{pmatrix} \dfrac{\partial \lambda_1}{\partial t_1} & \cdots & \dfrac{\partial \lambda_{n-2}}{\partial t_1} \\ \cdot & \cdots & \cdot \\ \dfrac{\partial \lambda_1}{\partial t_{n-2}} & \cdots & \dfrac{\partial \lambda_{n-2}}{\partial t_{n-2}} \end{pmatrix}$$

is globally nonsingular. Thus, we may by the Implicit Function Theorem define $(t_1, ..., t_{n-2})$ as an implicit function ψ of t_n, with the given initial point and in accordance with the relation

$$(\lambda_1 - \lambda_1', ..., \lambda_{n-2} - \lambda_{n-2}') = (0, ..., 0),$$

and *I* claim that the domain of ψ contains $[t_n', \infty)$.

Knowing by the Implicit Function Theorem that ψ is defined on an open neighborhood of t_n', assume that it is defined on the interval $[t_n', t_n'')$. It suffices for our proof to show that ψ may be continued to t_n''.

We first wish to show that $| t_i - t_{i-1} |$ is bounded from 0 for $1 \leqslant i \leqslant n$. If $| t_i - t_{i-1} | \to 0$ for some i as $t_n \to t_n''$, then i must be less than n, and $\lambda_j \to \infty$ for every j for which $| t_j - t_{j-1} | \not\to 0$. But then, since $\lambda_1, ..., \lambda_{n-2}$ remain constant, $t_j \to -1$ for $1 \leqslant j \leqslant n - 2$, and therefore $\lambda_{n-1} \to \infty$ and $\lambda_n \to \infty$. Thus $d\lambda_{n-1}/dt_n > 0$ and $d\lambda_{n-1}/dt_n > 0$ for all $t_n \in (t_n', t_n'')$, or else $d\lambda_i/dt_n = 0$ for $i = 1, 2, ..., n - 2$, and $n - 1$ or n at some point in (t_n', t_n''), implying that at that point an $n - 1 \times n - 1$ submatrix of A is singular, a contradiction. But, by Lemma 10, we see that, when $t_1, ..., t_{n-2}$, and t_n are varied in such a manner that $\lambda_1, ..., \lambda_{n-2}$ remain constant, $d\lambda_{n-1}/dt_n$ and $d\lambda_n/dt_n$ must disagree in sign, a contradiction.

Thus, there is an $\epsilon > 0$ such that $| t_j - t_i | \geqslant \epsilon$ for all i, j distinct as $t_n \to t_n''$, implying that on the compact set

$$\{(t_0, ..., t_n): | t_j - t_i | \geqslant \epsilon, t_0 = -1, t_{n-1} = 1, t_n' \leqslant t_n \leqslant t_n''\}$$

the determinant of B is bounded from 0, and thus all the derivatives dt_i/dt_n are bounded. Thus $\lim_{t_n \to t_n''} t_i = t_n''$ exists, and, by continuity, the condition $\lambda i = \lambda_i$, $1 \leqslant i \leqslant n - 2$, holds at the point $(t_0'', ..., t_n'')$. Therefore ψ may be continued to, and hence beyond, t_n''.

Now, as $t_n \to \infty$, we note that $y_n' \to 0$ and $y_n \to 0$ uniformly on $[t_0, t_{n-1}]$. Moreover, for $t \in [t_0, t_{n-1}]$, and for $k < n$, with $t_0 = -1$ and $t_{n-1} = 1$ we have

$$1 = \lim_{n\to\infty} \left| \frac{1 - t_n}{-1 - t_n} \right| \leqslant \lim_{n\to\infty} \left| \frac{t - t_n}{t_k - t_n} \right| \leqslant \lim_{n\to\infty} \left| \frac{-1 - t_n}{1 - t_n} \right| = 1.$$

Restricting our attention to the interval $[-1, 1]$, we see that the interpolation on $[t_0, t_n]$ degenerates in uniform and smooth fashion as $t_n \to \infty$ to an interpolation on the points $t_0, ..., t_{n-1}$ into the space Π_{n-1}, of degree one less than the original space, the points $t_1, ..., t_{n-2}$ being defined as the appropriate limits via the function ψ. That these limits actually exist is a consequence of the above limit computations, which imply that for $0 \leqslant i \leqslant n-1$, y_i and its partial derivatives and y_i' all converge to the expected limits, while y_n and all its derivatives simply vanish. This matter will receive more detailed scrutiny below.

As $t_n \to \infty$, $\lambda_i' = \lambda_i$ for $1 \leqslant i \leqslant n-2$, and λ_{n-1} decreases strictly monotonically, since λ_n must increase. Therefore the initial value of $\max_{1\leqslant i\leqslant n} \lambda_i$ is not less than the limiting value of $\max_{1\leqslant i\leqslant n-1} \lambda_i$ as $t_n \to \infty$. Since this is true for arbitrary initial nodes, it is in particular true for any initial configuration of nodes yielding $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-1} = c$, where c is some constant. In the limit, we obtain $\lambda_1 = \lambda_2 = \cdots = \lambda_{n-2} = c$, while $\lambda_{n-1} < c$. Therefore, by Theorem 1, the resulting limit cannot yield $\max_{1\leqslant i\leqslant n-1} \lambda_i = C_{n-1}$, whence $c > C_{n-1}$. In particular, $C_{n-1} < C_n$ for $n \geqslant 3$. It is well known that $C_1 = 1$ and $C_2 = \frac{5}{4}$.

Up to the matter of explaining in detail what happens as $t_n \to \infty$, we have completed the proof of Theorem 2. In order to see that $(t_1, ..., t_{n-2})$ indeed approaches a limit as $t_n \to \infty$, let us consider replacing t_n by $\theta = 1/t_n$. We may define $y_0^{(n-1)}, ..., y_{n-1}^{(n-1)}$ to be the polynomials of degree $n-1$ which interpolate on the points $-1 = t_0, t_1, ..., t_{n-1} = 1$. Then for $t \in \mathbb{R}$ we may define

$$z_i(t) = [y_i^{(n-1)}(t)](\theta t - 1)/(\theta t_i - 1) \qquad \text{for} \quad 0 \leqslant i \leqslant n-1, \quad \text{and}$$

$$z_n(t) = \theta^n \prod_{j=0}^{n-1} \left[\frac{t - t_j}{1 - \theta t_j} \right].$$

It is easily seen that, for $0 < \theta < 1$, the polynomials z_i agree with the polynomials of interpolation on the points $t_0, ..., t_{n-1}$, and t_n, where $t_n = 1/\theta$, and $1 = t_{n-1} < t_n$. If on the other hand $-1 < \theta < 0$, the polynomials z_i agree with the polynomials of interpolation on $t_0, ..., t_{n-1}$, and t_n, where $t_n = 1/\theta$ as before, but $t_n < t_0 = -1$. If $\theta = 0$, then $z_i = y_i^{(n-1)}$ for $0 \leqslant i \leqslant n-1$, whereas $z_n = 0$. Moreover, $z_0, ..., z_n$ are clearly continuous in $(t, t_0, ..., t_{n-1}, \theta)$, so long as $-1 < \theta < 1$ and $-1 = t_0 < t_1 < \cdots < t_{n-1} = 1$; indeed they are analytic, implying that $\tau_1, ..., \tau_{n-1}$ and $\lambda_1, ..., \lambda_{n-1}$ are analytic on the same domain, while τ_n and λ_n are defined and analytic for $0 < \theta < 1$, escaping definition when $\theta = 0$.

Now, when $\theta = 0$, the matrix B is nonsingular by Lemma 8, and nonsingular at all other values of θ by Lemma 9. Moreover, the matrix

$$A' = \begin{pmatrix} \dfrac{\partial \lambda_1}{\partial t_1} & \cdots & \dfrac{\partial \lambda_{n-1}}{\partial t_1} \\ \cdot & \cdots & \cdot \\ \dfrac{\partial \lambda_1}{\partial t_{n-2}} & \cdots & \dfrac{\partial \lambda_{n-1}}{\partial t_{n-2}} \\ \dfrac{\partial \lambda_1}{\partial \theta} & \cdots & \dfrac{\partial \lambda_{n-1}}{\partial \theta} \end{pmatrix},$$

which contains B as a submatrix, is defined and continuous at all values of $t_1, \dots, t_{n-1}$, and θ. It is of course nonsingular only when $\theta \neq 0$.

Therefore, to the function $\psi(t_n)$ defined above there corresponds naturally a function $\bar{\psi}(\theta)$, governed by the same conditions on $\lambda_1, \dots, \lambda_{n-2}$ and agreeing in its outputs when $t_n = 1/\theta \quad 0$. If $\bar{\psi}$ has been defined by means of an initial point $\theta' > 0$, then its domain clearly includes $(0, \theta']$. Since the matrix A' is defined and continuous at all values of θ (in particular when $\theta = 0$), and since the matrix B, a submatrix of A' and hence continuous at 0, is globally nonsingular, we may repeat the arguments used above to demonstrate that, for $1 \leqslant i \leqslant n - 1$, $| t_i - t_{i-1} | \not\to 0$ as $\theta \to 0$, justifying the assertion that $t_1, \dots, t_{n-2}$ converge to well-defined limits as $t_n \to \infty$. ∎

Proof of Theorem 3. Let $n \geqslant 3$. For an inductive proof, assume that whenever interpolation into Π_{n-1} is carried out on nodes $t_0, \dots, t_{n-1}$ with $t_0 = -1$ and $t_{n-1} = 1$, the map $(t_1, \dots, t_{n-2}) \mapsto (\lambda_1, \dots, \lambda_{n-2})$ is a global homeomorphism, and that for each choice of $c \in (C_{n-2}, \infty)$ there exist unique nodes $(t_1, \dots, t_{n-2})$ yielding $\lambda_1 = \cdots = \lambda_{n-2} = c$, and that if $c > Cn - 1$, then $\lambda_{n-1} < C_{n-1}$. One sees easily that, when $n = 3$, these conditions hold trivially, and that for any n these inductive hypotheses imply the uniqueness of nodes yielding $\lambda_1 = \cdots = \lambda_{n-1}$ for interpolation into Π_{n-1}.

Our proof is completed by a further investigation into the properties of the function $\bar{\psi}$ defined in the proof of Theorem 2. We have shown there that the nodes $t_1, \dots, t_{n-2}$ converge to well-defined limits as $\theta \to \infty$. The function $\bar{\psi}$ has been implicitly defined by an initial θ', an intitial set of nodes $t_1', \dots, t_{n-2}'$. and initial values $\lambda_1', \dots, \lambda_{n-2}'$. By our inductive hypothesis, and by continuity, the nodes approached as $\theta \to 0$ must be those unique nodes which yield $\lambda_1 = \lambda_1', \dots, \lambda_{n-2} = \lambda_{n-2}'$ when $\theta = 0$. These observations imply that, given $\theta' \in (0, 1)$, the function $\bar{\psi}$ induced by $\lambda_i = \lambda_i'$ for $1 \leqslant i \leqslant n - 2$ with initial point $(t_1', \dots, t_{n-2}', \theta')$ contains $[0, \theta']$ in its domain. We have further shown that all such functions which may exist have the same value at $\theta = 0$. This however makes it impossible for more than one such function to exist, since the matrix B is nonsingular when $\theta = 0$, implying that a function $\theta \mapsto (t_1, ..., t_{n-2})$ may be defined by initial values of $(t_1, ..., t_{n-2})$ occurring when $\theta = 0$, uniquely determined by $(\lambda_1, ..., \lambda_{n-2})$, according to our inductive hypothesis.

We note moreover that, since $| t_i - t_{i-1} | \not\rightarrow 0$ for $1 \leqslant i \leqslant n - 1$ as $\theta \rightarrow 0$, it is necessary that $\lambda_n \rightarrow \infty$ as $\theta \rightarrow 0$. Thus, since for $\theta > 0$ we can write $t_n = 1/\theta$, we must have $d\lambda_n/d\theta < 0$, else $d\lambda_i/d\theta = (d\lambda_i/dt_n)(dt_n/d\theta) = 0$, for some $\theta > 0$, for $i = 1, 2, ..., n - 2$, and n, a contradiction. By Lemma 10, $d\lambda_{n-1}/d\theta > 0$, and we find on increasing θ from 0 that $(t_1, ..., t_{n-2})$ must follow a now unique path, and only at θ' does λ_{n-1} recover its initial value of λ_{n-1}. There is thus a bijection, clearly continuous, between $\lambda_1, ..., \lambda_{n-1}$ and $t_1, ..., t_{n-2}, t_n$, provided that t_0 and t_{n-1} are fixed. As remarked previously, fixing t_0 and t_{n-1} is equivalent to fixing t_0 and t_n. Thus a global homeomorphism (a diffeomorphism, in fact) exists between $(t_1, ..., t_{n-1})$ and $(\lambda_1, ..., \lambda_{n-1})$, enabling us to carry forward the first part of the induction.

Now, let $c > C_{n-1}$. By Theorem 2, $c > C_{n-2}$ also. We note again that, for $\theta = 0$, there is a unique initial point for the function ψ determined by $\lambda_1 = \cdots = \lambda_{n-2} = c$. As θ increases, we recall that λ_{n-1} increases also, until a point, unique on the graph of ψ, is reached at which $\lambda_1 = \cdots = \lambda_{n-2} = \lambda_{n-1} = c$. That this point is globally unique follows from the established correspondences $(t_1, ..., t_{n-1}) \leftrightarrow (t_1, ..., t_{n-2}, \theta) \leftrightarrow (\lambda_1, ..., \lambda_{n-1})$, and, under the conditions $\lambda_1 = \cdots = \lambda_{n-1} = c$, the correspondence $c \leftrightarrow (t_1, ..., t_{n-2}, \theta)$, thus established is differentiable, whence also $d\lambda_n/dc$ can be computed and shown by Lemma 10 to be negative, implying in combination with Theorem 1 that $\lambda_1 = \cdots = \lambda_n = c$ occurs uniquely and only when $c = C_n$. The second and third parts of the inductive step have been completed, and with them the proof of the theorem. ▮

3. Lemmas on the Lebesgue Function and Related Functions

In a series of lemmas, we establish the necessary properties of $X_1, ..., X_n$, $X_1', ..., X_n'$, and $q_1, ..., q_n$ already referred to, which are used in the proofs of Theorems 1, 2, and 3. One assumption often used in the proofs without explicit statement is that $X_1, ..., X_n$, $X_1', ..., X_n'$, and their roots are all analytic functions of the nodal configuration $t_0, t_1, ..., t_n$, on the domain $t_0 < t_1 < \cdots < t_n$. It is assumed that $n \geqslant 2$ throughout.

Lemma 1. *The polynomials $X_1, ..., X_n$ each have at least $n - 1$ simple roots on $[a, b]$, as do X_1' and X_n'. For each i, $2 \leqslant i \leqslant n - 1$, X_i' has at least $n - 2$ roots on $[a, b]$. Each root of X_i', $1 \leqslant i \leqslant n$, is a local extremum of X_i.*

Proof. This proof is a tedious but straightforward counting argument, based on the fact that

$$\begin{aligned} X_i(t_j) &= (-1)^{j-i} && \text{if } j \geqslant i, \text{ for } 1 \leqslant i \leqslant n \text{ and } 0 \leqslant j \leqslant n, \text{ and} \\ X_i(t_j) &= (-1)^{i-j+1} && \text{if } j < i. \end{aligned}$$

LEMMA 2. *There is no common root for X'_{i-1} and X'_i, $2 \leqslant i \leqslant n$, nor is there a common root for X'_1 and X'_n.*

Proof. We use the easily established identities

$$\begin{aligned} X_{i-1} + X_i &= 2y_{i-1} && \text{for } 2 \leqslant i \leqslant n, \text{ and} \\ X_i + X_n &= 2y_0 + 2y_n && \text{if } n \text{ is odd}, \\ X_1 - X_n &= 2y_0 - 2y_n && \text{if } n \text{ is even}. \end{aligned}$$

Assume that $2 \leqslant i \leqslant n$ and that there is $r \in \mathbb{R}$, $X'_{i-1}(r) = X'_i(r) = 0$. Then also $y'_{i-1}(r) = 0$. Since by Lemma 1 t_{i-1} cannot be a local extremum for X_{i-1} nor for X_i, $r \neq t_{i-1}$. For $j \neq i-1, 0 \leqslant j \leqslant n, y_i(t_j) = 0$. Therefore $y'_i(t_j) \neq 0$. It follows that r is not a node, and that $y_{i-1}(r) \neq 0$. From Lemma 1 it also follows that $X_{i-1}(r) \neq 0$, since all roots of X_{i-1} are simple.

Now since $X'_{i-1}(r) = y'_{i-1}(r) = 0$, it follows that, for any $\alpha \in \mathbb{R}$, $X'_{i-1}(r) + \alpha y'_{i-1}(r) = 0$. Moreover, since for $j \neq i-1, 0 \leqslant j \leqslant n, y_{i-1}(t_j) = 0$, we have $X_{i-1}(t_j) + \alpha y_{i-1}(t_j) = X_{i-1}(t_j)$ when $j \neq i-1$. Thus for any $\alpha \in \mathbb{R}$, $X_{i-1} + \alpha y_{i-1}$ must have at least $n-1$ distinct simple roots. However, since $X_{i-1}(r) \neq 0 \neq y_{i-1}(r)$, there is $\alpha \in \mathbb{R}$ such that $\alpha \neq 0$ and $X_{i-1}(r) + \alpha y_{i-1}(r) = 0$, with the root at r having multiplicity at least two, whence $X_{i-1} + \alpha y_{i+1}$ has at least $n+1$ roots. From this contradiction it follows that X'_{i-1} and X'_i have no common root for $2 \leqslant i \leqslant n$.

For the case of X'_1 and X'_n, we adopt the definition

$$\begin{aligned} P_n &= y_0 - y_n, && n \text{ even} \\ &= y_0 + y_n, && n \text{ odd}, \end{aligned}$$

noting that $P_n(t_j) = 0$ and $P_n(t_j) \neq 0$ for $1 \leqslant j \leqslant n-1$, and any other hypothetical root r of P_n must be simple, whence $P'_n(r) \neq 0$.

Now assume that $X'_1(r) = X'_n(r) = 0$. We have seen that $P_n(r) \neq 0$. Also, $X'_1(t_0) \neq 0 \neq X'_n(t_n)$, whence r is not a node, nor is $X_1(r) = 0$, by Lemma 1. Thus there is $\alpha \in \mathbb{R}$, $\alpha \neq 0$, such that $X_1(r) + \alpha P_n(r) = 0$, since $X_1 + (-1)^{n+1} X_n = 2P_n$. The expression $X_1 + \alpha P_n$ takes on the value $(-1)^{j+1}$ at t_j, for $1 \leqslant j \leqslant n-1$. At t_0 the value is $1 + \alpha$, and at t_n the value is $(1+\alpha)(-1)^{n+1}$. If $1 + \alpha = 0$, then $X_1 + \alpha P_n$ vanishes at t_0 and t_n, having thus n roots, with the root at r of multiplicity more than one, implying at least $n+1$ roots. If $1 + \alpha > 0$, then there is a change of sign between t_{n-1} and t_n. If $1 + \alpha < 0$, then there is a change of sign between t_0 and t_1. In either case, the expression has $n-1$ roots at which it changes sign. Since the root at r has multiplicity at least two (or three, if the sign changes at r), we again end with at least $n+1$ roots. Thus X_1' and X_n' have no common root. ∎

LEMMA 3. *All roots of X_1' and X_n' lie on the interval $[\tau_1, \tau_n]$.*

Proof. Consider X_n'. By the proof of Lemma 1, it is clear that all roots of X_n' lie on $(t_0, \tau_n]$. Since $X_1 + (-1)^{n+1} X_n = 2P_n$, we let $t_n \to \infty$ while fixing $t_0, \ldots, t_{n-1}$. This causes P_n' to converge to y_0' uniformly on any compact set. Let t_n be chosen sufficiently large that $P_n' < 0$ on $[t_0, t_1]$. Now if $t_0 < t < \tau_1$, we have $X_1'(t) > 0$, and thus $(-1)^{n+1} X_n'(t) < 0$. By symmetry, the lemma also holds for X_1'. ∎

LEMMA 4. *The roots of X_1' and X_n' interlace as we pass from τ_1 to τ_n.*

Proof. By Lemma 3, all roots of X_1' and X_n' occur on $[\tau_1, \tau_n]$. By Lemmas 1 and 2, X_1' and X_n' each have their full complement of $n-1$ roots on $[\tau_1, \tau_n]$, and no two of these roots coincide. If $1 \leqslant i \leqslant n-2$, X_n' has a root on the interval (t_{i-1}, t_{i+1}) which is a local extremum, since

$$X_n(t_{i-1}) = -X_n(t_i) = X_n(t_{i+1}),$$

and this root is a continuous function of the nodes. Starting at the left of the interval, t_1 may be moved to coincide with the leftmost root of X_n'. Then by the identity

$$X_1' + (-1)^{n+1} X_n' = 2P_n',$$

X_1' is not zero, and by continuity the sign of X_1' at the leftmost root of X_n' is invariant over all nodal configurations. Moving to the interval (t_1, t_3), we may so position t_2 as to coincide with the second root of X_n' from the left. Again $X_1' \neq 0$, and its sign has alternated. Continuing this procedure until we reach (t_{n-3}, t_{n-1}), we then note that, since $X_1(t_{n-2}) = -X_1(t_{n-1}) = X_1(t_n)$, there must occur a root of X_1 on the interval (t_{n-2}, t_n). By Lemma 3, this root must lie to the left of τ_n. ∎

LEMMA 5. *Between τ_1 and the leftmost root of X_{n-1}' on $[\tau_1, \tau_n]$, there is a root of X_n'. The symmetric statement about X_1' also holds.*

Proof. We use the formula $X_{n-1} + X_n = 2y_{n-1}$ to show the first statement. Assuming that n is odd, we have $X_n(t_0) = -X_n(t_1) = X_n(t_2) = 1$, and $X_{n-1}(t_0) = -X_{n-1}(t_1) = X_{n-1}(t_2) = -1$. There is thus a unique minimum of X_n on (t_0, t_2). The node t_1 may be moved to coincide with the leftmost root of X_{n-1} in $[\tau_1, \tau_n]$. Then $X_n'(t_1) = 2y_{n-1}'(t_1) > 0$, and the minimum of X_n must lie to the left of t_1. That point is also a root of X_n' and by

Lemma 3 must lie to the right of τ_1. In case that n is even, the argument is similar. ▮

LEMMA 6. *For all i and j in the range $\{1, 2,..., n\}$, $X_i'(\tau_j) = 0$ if and only if $i = j$. Moreover, if $j \leqslant n - 1$, the polynomial X_i' has exactly one simple root on the interval $[\tau_j, \tau_{j+1}]$.*

Proof. If $n = 2$, then $\tau_1 < \tau_2$, while X_1 and X_2 are linear. Assume inductively that the pattern of roots has been established for all nodal configurations for all values of $n \leqslant N$, and consider any configuration of nodes, $t_0 < t_1 < \cdots < t_N < t_{N+1}$. Let $t_{N+1} \to \infty$. Then on $[t_0, t_N]$ the polynomials $X_1',..., X_N'$ and their respective roots approach uniformly the functions derived from interpolation on $t_1,..., t_N$. Thus by the inductive hypothesis the roots of $X_1',..., X_N'$ which lie on $[\tau_1, \tau_N]$ lie in the desired pattern. Similarly, by allowing $t_0 \to -\infty$ while $t_1,..., t_{N+1}$ are fixed, the roots of $X_2',..., X_{N+1}'$ lying on $[\tau_2, \tau_{N+1}]$ also obey the pattern. Lemmas 3, 4, and 5 complete the inductive step. ▮

Corollary to Lemma 6. The roots of $q_1,..., q_n$ lie in the same locations as those of $X_1',..., X_n'$, save that $\tau_1,..., \tau_n$ are no longer present as roots.

At this point, we adopt a convention. While we now know enough about $q_1,..., q_n$ actually to compute sgn $q_i(\tau_j)$ for all i, j, $1 \leqslant i, j \leqslant n$, it greatly simplifies and clarifies matters, to change the signs of $q_1 \cdots q_n$ in Lemmas 7, 8, and 9 so that $q_i(\tau_1) > 0$, for $1 \leqslant i \leqslant n$. This is possible since $q_i(\tau_1) \neq 0$.

LEMMA 7. *Adopting the above convention, we have, for $2 \leqslant i, j \leqslant n$,* sgn $q_i(\tau_i) =$ sgn $q_1(\tau_i)$, *while* sgn $q_j(\tau_i) = -$sgn $q_1(\tau_i)$ *for $j \neq i$.*

Proof. At τ_1, sgn $q_i(\tau_1) =$ sgn $q_1(\tau_1) = 1$ for $2 \leqslant i \leqslant n$. On the interval $[\tau_{i-1}, \tau_i]$, each of $q_1,..., q_n$ has exactly one root, save that q_{i-1} and q_i have no roots on the interval. ▮

LEMMA 8. *Let $1 \leqslant k \leqslant n$. Then $\{q_1,..., q_n\} \sim \{q_k\}$ is a basis for Π_{n-2}.*

Proof (by contradiction). Assume that there exist $\alpha_1,..., \alpha_n$ with $\alpha_k = 0$, such that not all of $\alpha_1,..., \alpha_n$ are zero, and $Q = \alpha_1 q_1 + \cdots + \alpha_n q_n = 0$. No generality is lost by assuming that $\alpha_1 \geqslant 0$. Also, the case that $k = 1$ is symmetric to the case that $k = n$. Therefore, we assume that $k \neq 1$. We define two subsets $\mathscr{N}$ and $\mathscr{P}$ of $\{2,..., n\}$:

$$\mathscr{N} = \{j \mid 2 \leqslant j \leqslant n \text{ and } \alpha_j < 0\},$$

$$\mathscr{P} = \{j \mid 2 \leqslant j \leqslant n \text{ and } \alpha_j > 0\}.$$

It is evident that, if not all of $\alpha_1,..., \alpha_n$ are zero, and if $\alpha_1 \geqslant 0$, then $\mathscr{N}$ is nonempty, for $Q(\tau_1) = (\alpha_1 q_1 + \cdots + \alpha_n q_n)(\tau_1) = 0$ and $q_i(\tau_i) > 0$ for all i, $1 \leqslant i \leqslant n$, by the convention previously assumed. Therefore some of $\alpha_1, \ldots, \alpha_n$ must be positive and some must be negative.

To see that $\mathscr{P}$ is nonempty, we consider the point τ_k. By Lemma 7, we have $\operatorname{sgn} q_j(\tau_k) = -\operatorname{sgn} q_1(\tau_k)$ for $j \neq 1, k$. If $\mathscr{P}$ is void, this implies $\alpha_1 \geqslant 0$ and $\alpha_j \leqslant 0$ for all $j \geqslant 2$, with $\alpha_j < 0$ for at least one $j \geqslant 2$, implying that $Q(\tau_k) > 0$. However, we have assumed that $Q(t) = 0$ for all t. Thus $\mathscr{P}$ is nonempty.

We now set $N = \alpha_1 q_1 + \sum_{j \in \mathscr{N}} \alpha_j q_j$ and $P = \sum_{j \in \mathscr{P}} \alpha_j a_j$, yielding $Q = N + P$. Investigating $\operatorname{sgn} N(\tau_i)$ for $2 \leqslant i \leqslant n$, we see that $\operatorname{sgn} N(\tau_1) = -1$, since $P(\tau_1) > 0$. Let $\alpha \leqslant i \leqslant n$. Then at τ_i we have two possibilities. If $i \in \mathscr{N}$, then $i \notin \mathscr{P}$. Thus, by Lemma 7, $\operatorname{sgn} q_j(\tau_i) = -\operatorname{sgn} q_1(\tau_i)$ for all $j \in \mathscr{P}$. Therefore, $0 \neq \operatorname{sgn} P(\tau_i) = -\operatorname{sgn} N(\tau_i) = -\operatorname{sgn} q_1(\tau_i)$. On the other hand, if $i \notin \mathscr{N}$, then again by Lemma 7, $\operatorname{sgn} q_j(\tau_i) = -\operatorname{sgn} q_1(\tau_i)$ for all $j \in \mathscr{N}$. Thus $\operatorname{sgn} N(\tau_i) = \operatorname{sgn} q_1(\tau_i)$ for every value of i in $\{2, \ldots, n\}$.

By Lemma 6 and its corollary, q_1 has exactly one root on the interval (τ_i, τ_{i+1}) for $2 \leqslant i \leqslant n - 1$, for a total of $n - 2$ roots, its full complement. By the convention assumed before Lemma 7, we see that $q_1(\tau_1) > 0$, while $\operatorname{sgn} q_1(\tau_i) = (-1)^i$ for $2 \leqslant i \leqslant n$. Above, we have seen that $\operatorname{sgn} N(\tau_1) = -\operatorname{sgn} q_1(\tau_1) = -1$, while $\operatorname{sgn} N(\tau_i) = \operatorname{sgn} q_1(\tau_i)$ for $2 \leqslant i \leqslant n$. Therefore, for $1 \leqslant i \leqslant n$, $\operatorname{sgn} N(\tau_i) = (-1)^i$, obliging $N(\tau_i)$ to have a root in each interval (τ_i, τ_{i+1}) for $1 \leqslant i \leqslant n - 1$, a total of $n - 1$ roots. Since the degree of N is not more than $n - 2$, this is an absurdity which proves that all the α's are zero. ∎

As previously stated, the proof of Lemma 8 completes the proof of Theorem 1.

LEMMA 9. *There is no nontrivial linear combination of* $q_1, \ldots, q_{n-2}$ *which has roots* $t_1, t_2, \ldots, t_{n-2}$, *such that* $\tau_1 < t_1 < \tau_2 < t_2 < \cdots < \tau_{n-2} < t_{n-2} < \tau_{n-1}$.

Proof. Assume that such a nontrivial combination $Q = \alpha_1 q_1 + \cdots + \alpha_n q_n$ exists, with $\alpha_1 \geqslant 0$ and $\alpha_{n-1} = \alpha_n = 0$. Then each root t_i has multiplicity one, and the expression alternates sign at $\tau_1, \tau_2, \ldots, \tau_{n-1}$. We define exactly as in Lemma 8 the sets $\mathscr{N}$ and $\mathscr{P}$ and the polynomials N and P, such that $Q = N + P$. Since we have assumed $\alpha_1 \geqslant 0$, we must in this case check two possibilities: $Q(\tau_1) < 0$ and $Q(\tau_1) > 0$.

If $Q(\tau_1) < 0$, then $N(\tau_1) < 0$, and for $1 \leqslant i \leqslant n - 1$ we have $\operatorname{sgn} Q(\tau_i) = (-1)^i$. We show that $\operatorname{sgn} N(\tau_i) = (-1)^i$ for $2 \leqslant i \leqslant n$, leading to the same contradiction as in the previous lemma.

If $i \notin \mathscr{N}$, then by Lemma 7, $\operatorname{sgn} q_j(\tau_i) = -\operatorname{sgn} q_1(\tau_i)$ for all $j \in \mathscr{N}$. Therefore $\operatorname{sgn} \alpha_j q_j(\tau_i) = \operatorname{sgn} q_1(\tau_i)$ for all $j \in \mathscr{N}$, and $\operatorname{sgn} N(\tau_i) = \operatorname{sgn} q_1(\tau_i) = (-1)^i$. We remark in particular that $n - 1 \notin \mathscr{N}$ and $n \notin \mathscr{N}$, since $\alpha_{n-1} = \alpha_n = 0$ from the outset.

If $i \in \mathcal{N}$, then $i \notin \mathcal{P}$, and again by Lemma 7, sgn $\alpha_j q_j(\tau_i) = -\text{sgn } q_1(\tau_i) = -(-1)^i = -\text{sgn } Q(\tau_i)$, for all $j \in \mathcal{P}$. Therefore, since P has sign opposite to Q or is equal zero if $\mathcal{P} = \varnothing$, and since $Q = N + P$, it is necessary that N and Q agree in sign, and sgn $N(\tau_i) = \text{sgn } Q(\tau_i) = (-1)^i$. Thus for all i, $1 \leqslant i \leqslant n$, sgn $N(\tau_i) = (-1)^i$, and N has $n-1$ roots, while its degree is not more than $n-2$. Therefore, the possibility that Q actually exists and $Q(\tau_1) < 0$ is not viable.

The other possibility is that $Q(\tau_1) > 0$. In that case, we necessarily have $P(\tau_1) > 0$, for, if not, $P(\tau_1) = 0$, whence P is the zero polynomial, and $Q = N$. This cannot occur, since sgn $Q(\tau_i) = (-1)^{i-1}$ for $1 \leqslant i \leqslant n-1$, and (because $n-1, n \notin \mathcal{N}$) sgn $N(\tau_{n-1}) = -\text{sgn } N(\tau_n)$.

If $2 \leqslant i \leqslant n$, then either $i \notin \mathcal{P}$ or $i \in \mathcal{P}$. If $i \notin \mathcal{P}$, then by Lemma 7, sgn $P(\tau_i) = -\text{sgn } q_1(\tau_i)$. We note that $n-1, n \notin \mathcal{P}$. If $i \in \mathcal{P}$, then $i \notin \mathcal{N}$. Thus sgn $N(\tau_i) = \text{sgn } q_1(\tau_i) = -\text{sgn } Q(\tau_i)$, unless perhaps $\mathcal{N}$ is empty and $N = 0$. Therefore, since $Q = N + P$, and since N and Q disagree in sign, it is necessary that P agree in sign with Q and sgn $P(\tau_i) = \text{sgn } Q(\tau_i) = -\text{sgn } q_1(\tau_i)$. Combining the cases, we see that sgn $P(\tau_i) = (-1)^{i-1}$ for $1 \leqslant i \leqslant n$, resulting in $n-1$ roots for P, which is again a contradiction. Thus, the inequality $Q(\tau_1) > 0$ is also not viable. Our result is therefore implied by exhaustion of all other possibilities. ∎

LEMMA 10 (de Boor and Pinkus [29]). Let J_k denote the expression

$$\det(\partial\lambda_i/\partial t_j)_{i=1, j\neq k, j=1}^{n, n-1}, \qquad \text{for} \quad 1 \leqslant k \leqslant n.$$

Then at any nodal configuration we have $(-1)^k (J_k/J_1) < 0$ for each $k \geqslant 2$ and moreover $\partial\lambda_1/\partial\lambda_k < 0$ for $k \geqslant 2$. Symmetrically, $\partial\lambda_n/\partial\lambda_k < 0$ for $k \leqslant n-1$.

Proof (included for the sake of completeness). These facts follow for all configurations of nodes if they can be shown for a particular configuration of nodes, in view of the fact that $J_1, \ldots, J_n$ do not vanish. To prove the results for a particular configuration $(t_1, \ldots, t_{n-1})$ observe that since $J_1(t_1, \ldots, t_{n-1})/0$, we can find a continuously differentiable function G on some open neighborhood V of the point $(\lambda_i(t_1, \ldots, t_{n-1}))_{i=2}^n$ and an open neighborhood U of $(t_1, \ldots, t_{n-1})$ such that $\lambda_1(s_1, \ldots, s_{n-1}) = G(\lambda_2(s_1, \ldots, s_{n-1}), \ldots, \lambda_n(s_1, \ldots, s_{n-1}))$ for all $(s_1, \ldots, s_{n-1}) \in U$. Also, by Cramer's rule,

$$\partial\lambda_1 = \sum_{k=2}^{n} (-1)^k (J_k/J_1)\, \partial\lambda_k,$$

and therefore

$$\frac{\partial G}{\partial\lambda_k} = \frac{\partial\lambda_1}{\partial\lambda_k} = (-1)^k \frac{J_k}{J_1} \qquad \text{for} \quad 2 \leqslant k \leqslant n.$$

If now for some $k, 2 \leqslant k \leqslant n, (-1)^k J_k/J_1 > 0$, then we could find $(s_1, ..., s_{n-1}) \in U$ such that $\lambda_i(s_1, ..., s_{n-1}) = \lambda_i(t_1, ..., t_{n-1})$ for $2 \leqslant i \leqslant n, i \neq k$, while $\lambda_i(s_1, ..., s_{n-1}) < \lambda_i(t_1, ..., t_{n-1})$ for both $i \neq 1$ and $i = k$. Hence, if $(t_1, ..., t_{n-1})$ were optimal, $(s_1, ..., s_{n-1})$ would also be optimal, contradicting Theorem 1. This proves our results for $(t_1, ..., t_{n-1})$ and thus for all other nodal configurations as well. ▮

From Lemmas 9 and 10, Theorems 2 and 3 follow.

Historical Notes

Since this paper has presented a solution of Bernstein's conjecture already mentioned, it should be appropriate to record something of the history of the problem. Lagrange interpolation is, of course, a very old procedure for approximation, and, when the question arose of estimating the error in such procedures, it has been for some time naively assumed that if the points of interpolation "fill up" the interval upon which interpolation is taking place, the results get "better and better." This was disproved by examples by Meray and Runge around the turn of the century.

In fact, quite the opposite is the case, when one interpolates arbitrary continuous functions on a given interval. Bernstein [23] in 1914 suspected after investigating equally spaced nodes that the minimal norm for interpolation grows logarithmically with the degree of the polynomials used. Faber [24] in 1914 also showed that, given any array of nodes $\{t_0^{(0)}, ..., t_n^{(n)}\}_{n=1}^{\infty}$, there must exist a function f for which $\sum_{i=0}^{n} f(t_i^{(n)}) y_i^{(n)}$ fails to converge to f. Further investigations into such matters were carried on by Hahn [25].

As a consequence of such results, the problem of finding or characterizing nodes which would minimize the norm of the interpolation operator began to arouse interest. Useful results in this area have been hitherto scarce, save for results giving asymptotic bounds on the projection constant C_n. The works of Bernstein[5], Erdös [26], Ehlich and Zeller [13], Luttman and Rivlin [7], and others have shown that a constant c exists, such that

$$2/\Pi \log n - c \leqslant C_n \leqslant 2/\Pi \log n + c, \qquad \text{for all } n.$$

Interpolation on the zeros of the Tchebycheff polynomial of degree n produces results which are asymptotically optimal. However, Luttman and Rivlin in [7] have found better points of interpolation for $n \leqslant 40$. Erdös in [8] has added to Bernstein's conjecture by conjecturing that, for all nodal configurations, if for some i, $\lambda_i < Cn$, then a j exists such that $\lambda_j > C_n$. This conjecture could be affirmatively settled by modifications of the proofs contained here, but it has in fact been resolved by de Boor and Pinkus in [29]. One consequence of Erdös' conjecture is exact information about the quality of interpolation on a given set of nodes through the difference between the greatest and the least of the local maxima of the Lebesgue function. Measured in this manner, interpolation on the Tchebycheff nodes is indeed very good and improves as n grows.

It is of interest to mention that no elegant general method, whether a formula or a special algorithm, has yet been discovered which serves to compute the nodes yielding $\lambda_1 = \lambda_2 = \cdots = \lambda_n$. However, as a direct consequence of the closure of Π_n under affine transformations of the variable t, and of Theorem 3, the desired nodes must be symmetric about the midpoint of the interval $[a, b]$.

The first reference to the problem solved here is the article [5] of Bernstein. In the same work, Bernstein noted that $C_2 = \frac{5}{4}$. This last result, and analogous computations for $n = 3$, also appear in Tureckii [1] and in Neuman [14]. The conjecture of Bernstein is mentioned again in Erdös [8], along with the refinement already mentioned. On the assumption that Bernstein's conjecture was true, Hayes and Powell [2] computed the optimal nodes for $n \leqslant 15$. The present contributor began his work on the problem as a dissertation topic [27]. The problem was not solved in the dissertation, although many of the techniques used in the solution were developed. The main achievement of the dissertation was a proof, by application of fixed point theory, that a nodal configuration exists which equalizes the local maxima of the Lebesgue function for interpolation into any Haar subspace of $C[a, b]$ which contains the constant functions. Most of these have appeared in the article [28] of Kilgore and Cheney. The present paper and that of de Boor and Pinkus [29] were independently conceived as natural extensions of my note [30]. Their paper draws upon their greater experience to simplify the proofs of [30], given here in full, and provides a noninductive proof of Theorem 3 and the conjecture of Erdös. In addition, their paper extends the methods developed here to study the question whether the equally spaced nodes optimize interpolation of 2Π-periodic functions by means of trigonometric polynomials. The intersection between the two articles, aside from Theorem 1 and its proof, lies in Lemma 10, for which I must remain indebted.

The fact that $(t_0, \ldots, t_n) \mapsto (\lambda_1, \ldots, \lambda_n)$ is a differentiable function has, as far as I know, never been recorded as a theorem. Related results, such as the differentiability of $(t_0, \ldots, t_n) \mapsto (\lambda_1 - \lambda_2, \ldots, \lambda_{n-1} - \lambda_n)$, which was demonstrated in an early version of [28], have been known for some time. The formula $\partial\lambda_i/\partial t_j = -y_j(\tau_i)\, X_i(t_j)$ has its ancestry in the excellent paper [16] of Morris and Cheney, where the expression on the right occurs in the more general setting of interpolation into an arbitrary differentiable Haar subspace and is not called a partial derivative. The question of exactly to what class of subspaces the methods of this paper can be applied remains very much open. Many of the tools used here are available in a wide variety of situations, as [27] and [28] have already demonstrated.

Several references appearing in the bibliography have not been cited here.

They have handled apparently related unsolved problems or have formulated the problem under discussion here in other ways which might possibly lead to another method of solution.

ACKNOWLEDGMENTS

This article has benefited greatly from the exhaustive and patient criticism which Dr. Carl de Boor has bestowed upon it. I would also like to thank Dr. Gary Hamrick for discussing with me global applications of the Implicit Function Theorem. Lastly, I extend my appreciation to Dr. Hashem Hibshi, chairman of the Department of Mathematics at the College of Education, King Abdulaziz University at Mecca, for his sympathetic consideration of my desire to complete this project.

REFERENCES

1. A. H. TURECKII, Theory of Interpolation in Problem Form, (Russian) Minsk, 1968, MR 41, #5840.
2. W. E. HAYES AND M. J. D. POWELL, unpublished data, Atomic Energy Research Establishment, Harwell, England, 1969.
3. J. R. CANNON AND K. MILLER, Bounds for the ratio between the fit of interpolation and the best fit, unpublished manuscript, 1964.
4. J. W. ARMSTRONG, Point systems for Lagrange interpolation, *Duke Math. J.* **21** (1954), 511–516.
5. S. BERNSTEIN, Sur la limitation des valeurs d'un polynôme $P_n(x)$ de degré n sur tout un segment par ses valeurs en $n + 1$ points du segment, *Bull. Acad. Sci. URSS Leningrad* (1931), 1025–1050. *Zbl.* **4** (1932), 6.
6. M. J. D. POWELL, On the maximum errors of polynomial approximations defined by interpolation and by least squares criteria, *Computer J.* **9** (1967), 404–407.
7. F. W. LUTTMAN AND T. J. RIVLIN, Some numerical experiments in the theory of polynomial interpolation, *IBM J. Res. Develop.* **9** (1965), 187–191.
8. P. ERDÖS, Some remarks on polynomials, *Bull. Amer. Math. Soc.* **53** (1947), 1169–1176.
9. T. J. RIVLIN, The Lebesgue constants for polynomial interpolation, IBM Research Report RC 4165, December 1972.
10. P. ERDÖS, On some convergence properties of the interpolation polynomials, *Ann. of Math.* **44** (1943), 330–337.
11. P. ERDÖS, Problems and results on the convergence and divergence properties of the Lagrange interpolation polynomials and some extremal problems, *Mathematica (Cluj)* **10** (1968), 65–73.
12. P. ERDÖS AND P. TURAN, On the role of the Lebesgue functions in the theory of the Lagrange interpolation, *Acta Math. Acad. Sci. Hungar.* **6** (1955), 47–66.
13. H. EHLICH AND K. ZELLER, Auswertung der Normen von Interpolations-operatoren, *Math. Ann.* **164** (1966), 105–112.
14. E. NEUMAN, Projections in uniform polynomial approximations, *Zastos. Mat.* **14** (1974), 99–125.
15. E. W. CHENEY AND K. H. PRICE, Minimal projections, "Proceedings of Symposium Held at Lancaster, July, 1969" (A. Talbot, Ed.), pp. 261–289, Academic Press, London, 1970.
16. P. D. MORRIS AND E. W. CHENEY, On the existence and characterization of minimal projections, *J. Reine Angew. Math.* **270** (1974), 61–76.

17. J. MYCIELSKI AND S. PASZKOWSKI, A Generalization of Chebyshev Polynomials, *Bull. Acad. Polon. Sci. Ser. Sci. Math. Astronom. Phys.* **8** (1960), 433–438.
18. V. S. VIDENSKII, A class of interpolation polynomials with nonfixed nodes, *Dokl.* **162** (1965), 647–640.
19. C. DAVIS, Mapping properties of some cebysev systems, *Dokl. Akad. Nauk SSSR* **175** (1967), 840–843; *Soviet Math. Dokl.* **8** (1967).
20. C. DAVIS, Extrema of a polynomial, advanced problem 4714, *Amer. Math. Monthly* **64** (1957), 679–680.
21. C. H. FITZGERALD AND L. L. SCHUMAKER, A differential equation approach to interpolation at extremal points, *J. Analyse Math.* **22** (1969), 117–134.
22. K. H. PRICE AND E. W. CHENEY, Extremal properties of approximation operators, Center for Numerical Analysis Report #54, University of Texas at Austin, August 1972.
23. S. N. BERNSTEIN, Quelques remarques sur l'interpolation, *Math. Ann.* **79** (1918), 1–12.
24. G. FABER, Uber die interpolatorische Darstellung stetiger Funktionen, *Jber. Deutsch. Math. Verein.* **23** (1914), 190–210.
25. H. HAHN, Uber das Interpolations Problem, *Math. Z.* **1** (1918), 115–142.
26. P. ERDÖS, Problems and results in the theory of interpolation, II. *Acta Math. Acad. Sci. Hungar.* **12** (1961), 235–244.
27. T. A. KILGORE, Optimization of the Lagrange Interpolation Process," Ph.D. dissertation, University of Texas at Austin, August 1974.
28. T. A. KILGORE AND E. W. CHENEY, A theorem on interpolation in Haar subspaces, *Aequationes Math.* **14** (1976), 391–400.
29. C. DE BOOR AND A. PINKUS, Proof of the conjectures of Bernstein and Erdös concerning the optimal nodes for polynomial interpolation, *J. Approximation Theory* **24** (1978), 289–303.
30. T. A. KILGORE, Optimization of the norm of the Lagrange interpolation operator, *Bull. Amer. Math. Soc.* **83** (1977), 1069–1071