PROOF OF THE CONJECTURES OF BERNSTEIN AND ERDOS CONCERNING THE --ETC(U)

JUN 77 C DE BOOR, A PINKUS

UNCLASSIFIED MRC-TSR-1756

END
DATE FILMED
9-77

DOD
MRC Technical Summary Report #1756

PROOF OF THE CONJECTURES OF BERNSTEIN AND ERDÖS CONCERNING THE OPTIMAL NODES FOR POLYNOMIAL INTERPOLATION

Carl de Boor and Allan Pinkus

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

June 1977

(Received May 12, 1977)

Approved for public release
Distribution unlimited

sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

and

National Science Foundation
Washington, D. C.
20550
ABSTRACT

For each \( t \in T := \{ t \in \mathbb{R}^{n-1} : a < t_1 < \cdots < t_{n-1} < b \} \), let \( \Lambda_t(x) \) be the Lebesgue function of the process of polynomial interpolation on \([a,b]\) by polynomials of degree \( \leq n \) at the points

\[
a =: t_0 < t_1 < \cdots < t_{n-1} < t_n := b.
\]

Let \( \lambda_i(t) := \max_{t_{i-1} < x < t_i} \Lambda_t(x) \),

\[ \lambda_i(t) = \max_{t_{i-1} < x < t_i} \Lambda_t(x), \]

\( i = 1, \ldots, n \). Based on work of Kilgore [8], we prove the following conjectures.

(a) Bernstein: \( \| \Lambda_t \|_\infty \) is minimal when \( \lambda_1(t) = \cdots = \lambda_n(t) \).

(b) Erdös: If \( \lambda_i(t) = \lambda^* \), \( i = 1, \ldots, n \), then for all \( s \in T \setminus \{t\} \),

\[ \min_{i} \lambda_i(s) < \lambda^* < \max_{i} \lambda_i(s). \]

Analogous results are proven for trigonometric interpolation.

These results are of interest since \( \| \Lambda_t \|_\infty \) gives the norm of the linear map of polynomial interpolation on the continuous functions and therefore bounds the effect of noisy data on their polynomial interpolant and shows how close the interpolation error is to the best possible error by any method.

AMS (MOS) Subject Classifications: 41A05, 42A12

Key Words: polynomial interpolation, optimal nodes, Lebesgue function, minimal projector.

Work Unit Number 6 (Spline Functions and Approximation Theory)
PROOF OF THE CONJECTURES OF BERNSTEIN AND EHRDÖS CONCERNING THE
OPTIMAL NODES FOR POLYNOMIAL INTERPOLATION

Carl de Boor and Allan Pinkus

Introduction. It is the purpose of this note to complete and extend work of Kilgore [8] on the optimal nodes in polynomial interpolation.

The problem is as follows. Consider the Banach space $C[a,b]$ of continuous functions on the finite interval $[a,b]$, with the usual norm

$$
\|f\| := \max_{a \leq x \leq b} |f(x)| .
$$

Throughout the paper, we take $n$ to be a fixed integer,

$$n \geq 2 .$$

Corresponding to each point $t$ in

$$T := \{ t \in \mathbb{R}^{n-1} : a < t_1 < \ldots < t_{n-1} < b \} ,$$

we construct the linear map $P_t$ of polynomial interpolation in $C[a,b]$ at the $n+1$ points or nodes $a := t_0, t_1, \ldots, t_n := b$. In its Lagrange form,

$$P_t f := \sum_{i=0}^{n} f(t_i) \xi_i$$

with

$$\xi_i(x) := \prod_{j \neq i} \frac{x - t_j}{t_i - t_j} , \quad i \in \{0, n\} .$$

The problem is one of determining optimal nodes, i.e., a point or points $t^* \in T$ for which

$$\|P_{t^*}\| = \inf_{t \in T} \|P_t\| .$$

Here, $\|P_t\| := \sup_{f \in C[a,b]} \|P_t f\|/\|f\|$, as usual. Consideration of this problem is motivated by the fact that $P_t$ is a projector on $C[a,b]$ and its range is $\tau_n$, the subspace of polynomials of degree $\leq n$, which implies that

$$\|f - P_{t^*} f\| \leq (1 + \|P_{t^*}\|) \text{dist}(f, \tau_n) .$$

Sponsored by the United States Army under Contract No. DAA029-75-C-0024 and the National Science Foundation under Grant No. MCS75-17385.
It is well known that $\|p_\ell\|$ can be computed as

$$\|p_\ell\| = \|A_\ell\|,$$

with

$$A_\ell := \sum_{i=0}^{n} |\xi_i|$$

the Lebesgue function of the process. A simple argument shows that $A_\ell(x) \geq 1$ with equality iff $x \in \{t_0, \ldots, t_n\}$. Set

$$\lambda_i(t) := \max_{t_{i-1} \leq x \leq t_i} A_i(x) \text{ for } i \in \{1, n\}.$$

In 1931, S. Bernstein [1] conjectured that $\|p_\ell\|$ is minimal when $A_\ell$ equioscillates, i.e., when $\lambda_1(t) = \lambda_2(t) = \ldots = \lambda_n(t)$. Later, Erdős [7] added to this the conjecture that there is exactly one choice of $t$ for which $A_\ell$ equioscillates and that

$$\min_i \lambda_i(t) \leq \lambda^* := \inf_{\ell \in T} \|p_\ell\| \text{ for every } t \in T.$$

The latter conjecture appears already in Erdős [6] in the form: "min \lambda_i(t) achieves its maximum when $A_\ell$ equioscillates."

Subsequent work on these conjectures and related topics is summarized in Lattmann & Rivlin [11], and in Cheney & Price [4].

Substantial progress in answering these conjectures has come only very recently.

Kilgore and Cheney [9] showed the existence of $t \in T$ for which $A_\ell$ equioscillates. This result was considerably strengthened by Kilgore [8] who showed that an optimal Lebesgue function, i.e., a $A_\ell$ for which $\|A_\ell\| = \lambda^*$, must necessarily equioscillate.

In the present paper, which is very much based on Kilgore's analysis, we prove the validity of all of the above conjectures. Explicitly, we prove (Theorem 1) that there is only one $t \in T$ for which $A_\ell$ equioscillates, and we prove (Theorem 2) that

$$\lambda_i(t) \leq \lambda_i(s) \text{ for all } i \in \{1, n\}$$

cannot hold except in the trivial case when $t = s$ from which (1) follows immediately. In addition, we prove analogous results for trigonometric interpolation.
The article is organized as follows. In Section 2, we outline Kilgore's proof of the fact that an optimal Lebesgue function must equioscillate. Section 3 is concerned with the proof of Theorems 1 and 2. In Section 4, we extend these results to the case of trigonometric interpolation. Explicitly, we prove the intuitively obvious fact that trigonometric interpolation on \([0,2\pi]\) at equidistant nodes is optimal.
2. Kilgore's result. In this section, we quickly review the proof of Kilgore's result that an optimal Lebesgue function must equioscillate. This we do for completeness and in order to facilitate its extension to trigonometric interpolation in Section 4. We continue to use the notation introduced in Section 1.

Theorem (Kilgore [8]). If \( \|A_\mathcal{T}^n\| = \lambda^* = \inf_{\mathcal{T}} \|F\| \), then \( A_\mathcal{T}^n \) equioscillates, i.e., then \( \lambda_1(t) = \lambda_2(t) = \ldots = \lambda_n(t) \).

Proof outline. For \( i \in [1,n] \), denote by \( F_i \) the polynomial of degree \( \leq n \) which agrees with \( A_\mathcal{T}^n \) on \( [t_{i-1}, t_i] \). One easily verifies that \( F_i \) is the unique element of \( \mathbb{R}_n \) for which

\[
F_i(t_j) = \begin{cases} (-1)^{i-j} & \text{if } j \in [0,i-1], \\ (-1)^{j-i} & \text{if } j \in [i,n]. \end{cases}
\]

Furthermore, denote by \( \tau_i \) the unique point in \( [t_{i-1}, t_i] \) at which \( A_\mathcal{T}^n \) and \( F_i \) take on the value \( \lambda_i(t) \),

\[
F_i(\tau_i) = \lambda_i(t) = \max_{t_{i-1} \leq x \leq t_i} |F_i(x)| \text{ for all } i \in [1,n].
\]

Kilgore points out that the theorem follows at once if it can be shown that

\[
\text{for each } t \in \mathcal{T}, \text{ for each } k \in [1,n], \text{ and all } \mathcal{A} \text{ close to } \lambda_\mathcal{A} = (\lambda_i(\mathcal{A}))_{i=1}^n \text{ there exists } \mathcal{A}_0 \text{ close to } \mathcal{A} \text{ so that } \lambda_\mathcal{A}_0 = \nu_i \text{ for all } i \neq k.
\]

For, then \( \lambda_i(t) < \|A_\mathcal{T}^n\| \) for some \( k \) implies the existence of \( \mathcal{A}_0 \) (near \( \mathcal{A} \)) for which \( \|A_\mathcal{T}_0^n\| < \|A_\mathcal{T}^n\| \).

Kilgore establishes (2) by showing that

\[
\text{for } t \in \mathcal{T}, \text{ and } k \in [1,n], J_k := \det(\nabla \lambda_i(t)/\nabla t_j)_{i=1,j=1}^{n,n} \neq 0.
\]

His proof of (3) begins with the observation that

\[
\frac{\partial \lambda_i}{\partial t_j} = -F_i'(t_j)S_j(t_i) = \prod_{j \neq k} (\tau_i - t_k) F_i'(t_j) / \prod_{j \neq k} (t_j - t_k)
\]

which shows \( \lambda_i \) to be continuously differentiable on \( \mathcal{T} \) and also shows that (3) is equivalent to
(4) \[ \text{for } t \in T \text{ and } k \in [1,n], \det(q_i(t_j))_{i=1,j=1}^{n} \neq 0, \quad i \neq k \]

with
\[ q_i(x) := F_i'(x)/(x - \tau_i), \quad i \in [1,n]. \]

Since each \( q_i \) is a polynomial of degree \( \leq n - 2 \), (4) is, in turn, equivalent to the linear independence of any \( n - 1 \) of the \( n \) polynomials \( q_1, \ldots, q_n \). For the proof of this linear independence, Kilgore uses eight lemmas. The first five lemmas lead up to the following

**Lemma 6 (of [8]).** On the interval \( [\tau_1, \tau_n] \), the zeros of \( F_1', \ldots, F_n' \) lie in the pattern
\[ i, n, n-1, \ldots, 3, 2, 1, n, n-1, \ldots, 3, 2, 1, n, n-1, \ldots, 3, 2, 1. \]

Here, the number \( i \) denotes a zero of \( F_i' \), and \( i \) denotes the point \( \tau_i \).

It may be instructive for the reader to consider the following alternative argument which obtains Lemma 6 as an immediate corollary to the corresponding result for the zeros of \( F_1', \ldots, F_n' \).

For \( r \in [1,n] \setminus \{i\} \), \( F_i \) changes sign on \( (t_{r-1}, t_r) \), hence must have a zero there. Since \( F_i \) cannot have more than \( n \) zeros, these zeros must all be simple and \( F_i \) has no other zeros in \( [a,b] \). Let \( c^{(i)}_1, \ldots, c^{(i)}_{n-1} \) denote these zeros, in increasing order. Then
\[ c_r^{(i)} \in \begin{cases} (t_{r-1}, t_r), & \text{for } r < i, \\ (t_r, t_{r+1}), & \text{for } r > i. \end{cases} \]

If \( F_i \) has an additional zero, we denote it by \( c_0^{(i)} \) or by \( c_n^{(i)} \) depending on whether it is less than \( a \) or greater than \( b \), respectively.

**Lemma 1.** For \( i < j \), the zeros of \( F_i \) and \( F_j \) strictly interlace. More precisely,
\[ c_r^{(j)} < c_r^{(i)} \quad \text{for all applicable } r \text{ in } [0,n]. \]

**Proof.** The function \( G := F_i - (-1)^{j-i} F_j \) satisfies
\[ G(t_k) = \begin{cases} 0 & \text{for } k \in [0, i - 1] \cup [j, n], \\ 2(-1)^{k-i} & \text{for } k \in [i, j - 1]. \end{cases} \]
Thus, $G_1$ has at least $i + n + 1 - j$ zeros outside $[t_1, t_{j-1}]$, and $j - 1 - i$ zeros in $(t_1, t_{j-1})$. Since $G_1$ is a polynomial of degree $\leq n$, it cannot have any additional zeros and all these zeros must be simple. But, since $G_1(t_1) = 2 > 0$, this shows that $(-1)^{i-r}G_1 > 0$ on $(t_{r-1}, t_r)$ for all $r < i$ and so shows that

\[(5a)\quad t_{r-1} < \sigma_r^{(j)} < \sigma_r^{(1)} < t_r \quad \text{for } r \in [1, i - 1] \]

and also

\[(5b)\quad \sigma_0^{(j)} < \sigma_0^{(1)} < t_0 \quad \text{if these exist.} \]

We have trivially

\[(5c)\quad \begin{align*}
    t_{i-1} &< \sigma_i^{(j)} < t_i, \\
    t_{j-1} &< \sigma_{j-1}^{(1)} < t_j.
\end{align*} \]

Also, $G_1(t_{j-1}) = 2(-1)^{j-1-i}$, hence $(-1)^{r-i}G_1 > 0$ on $(t_r, t_{r+1})$ for $r \geq j$, and therefore

\[(5d)\quad t_r < \sigma_r^{(j)} < \sigma_r^{(1)} < t_{r+1} \quad \text{for } r \in [j, n - 1] \]

and also

\[(5e)\quad t_n < \sigma_n^{(j)} < \sigma_n^{(1)} \quad \text{if these exist.} \]

Finally, the function $G_2 := F_1 + (-1)^jF_j$ satisfies

\[
    G_2(t_k) = \begin{cases} 
        2(-1)^{k-1-i} & \text{for } k \in [0, i - 1], \\
        0 & \text{for } k \in [i, j - 1], \\
        2(-1)^{k-i} & \text{for } k \in [j, n]. 
    \end{cases}
\]

$G_2$ has at least the $j - 1$ zeros $t_1, \ldots, t_{j-1}$ in $[t_{i-1}, t_j]$ and has at least $i + n - j$ zeros outside $[t_{i-1}, t_j]$, giving a total of at least $n - 1$ zeros. Since $G_2(t_{j-1})G_2(t_j) = 4(-1)^{j-1}$, the number of zeros of $G_2$ in $[t_{i-1}, t_j]$ must be of parity $j - 1$. Therefore, since $G_2$ is of degree $\leq n$, it follows that $G_2$ has no other zeros in $[t_{i-1}, t_j]$. This proves that $(-1)^{r-j}G_2 > 0$ on $(t_{r-1}, t_r)$ for $r \in [i, j]$ and so shows that

\[(5f)\quad t_{r-1} < \sigma_{r-1}^{(1)} < \sigma_r^{(j)} < t_r \quad \text{for } r \in [i + 1, j - 1] \]

Concatenation of $(5a-f)$ proves Lemma 1.
Figure 1. Schematic drawing of $F_i$ (solid), $F_j$ (dashed) and $-F_j$ (dotted) for $n = 6$, $i = 3$, $j = 5$. The graphs of $F_i$ and $(-1)^{j-i}F_j$ cross at the $n$ points indicated by $\mathbb{O}$, those of $F_i$ and $-(-1)^{j-i}F_j$ cross at the $n-1$ points indicated by $\mathbb{O}$.

Corollary. The zeros of $F_1, \ldots, F_n$ on $(-\infty, \infty)$ lie in the pattern

$$
0^{(1)}, \ldots, 0^{(1)}, 0^{(n)}, \ldots, 0^{(1)}, 0^{(n)}, \ldots, 0^{(n)}
$$

where $I$ and $J$ are certain integers with $1 \leq I < J \leq n$.

Proof. The corollary is an immediate consequence of Lemma 1 and the additional fact that $0^{(1)}_0$ and $0^{(n)}_n$ necessarily exist.

Since $G_i$ is of degree $n$ for any $i$ and $j$, it follows that $I$ equals $J - 1$ or $J - 2$.

Let now $\gamma^{(1)}_x$ denote the zero of $F_i'$ which lies between $0^{(1)}_{x-1}$ and $0^{(1)}_x$. Since the zeros of $F_i$ and $F_j$ interlace for $i \neq j$, V. A. Markov's well known result [12]
implies that the zeros of $F'_1$ and $F'_j$ interlace, and interlace in the same manner.

Therefore, the corollary implies

**Lemma 2.** The zeros of $F'_1, \ldots, F'_n$ lie in the pattern

$$t_1, t_2, \ldots, t_m, t_{m+1}, \ldots, t_{m+n-1}, t_n$$

where $I$ and $J$ are certain integers with $1 \leq I < J \leq n$.

Lemma 6 of [8] follows from this since $t_i = t_i$, all $i$.

The proof of (4) is now finished as follows. Recall that $q_i$ is a polynomial of degree $\leq n-2$ which vanishes at the zeros of $F'_i$ except for $t_i$. We may assume $q_i(t_i) > 0$, all $i$. Lemma 6 then implies that

$$\text{sgn } q_i(t_j) = (-1)^{j+1} \text{ for } i, j \in [2, n], i \neq j,$$

$$\text{sgn } q_i(t_i) = (-1)^j \text{ for } i \in [2, n],$$

$$\text{sgn } q_i(t_j) = (-1)^j \text{ for } j \in [2, n].$$

Assume now that $\sum a_k q_k = 0$ for some $a \neq 0$ with $a_1 > 0$. Then the set

$N := \{k \in [2, n] : a_k < 0\}$

is not empty since $q_k(t_i) > 0$ for all $k$. Set $P := [2, n] \setminus N$ and consider the function

$$f := a_1 q_1 + \sum_{k \in P} a_k q_k = -\sum_{k \in P} a_k q_k.$$

We have

$$(-1)^j f(t_j) = \sum_{k \in P} a_k (-1)^{j+1} q_k(t_j) > 0 \text{ for } j \not\in P$$

while

$$(-1)^j f(t_j) = a_1 (-1)^j q_1(t_j) + \sum_{k \in N} (-a_k) (-1)^{j+1} q_k(t_j) > 0 \text{ for } j \in P.$$

This shows the polynomial $f$ of degree $\leq n-2$ to have $n-1$ weak sign changes, and therefore $f = 0$ and so, in particular, $P = \emptyset$. Hence $a_k < 0$ for all $k \in [2, n]$. But since $q_k(t_i) > 0$ for all $k$, it then also follows that $a_1 > 0$.

In summary, $\sum a_k q_k = 0$ for some $a \neq 0$ implies that $a_1 a_k < 0$ for all $k \in [2, n]$. In particular, then $a_k \neq 0$ for all $k \in [1, n]$, and (4) follows.
3. Uniqueness. The central result of this article is the following theorem.

Theorem 1. The map \( T \to \mathbb{R}^{n-1} : t \mapsto (\lambda_{i+1}(t) - \lambda_i(t))_{i=1}^{n-1} \) is a homeomorphism of \( T \) onto \( \mathbb{R}^{n-1} \).

In particular, there is exactly one \( t \in T \) with \( \Gamma(t) = 0 \), i.e., exactly one \( t \) for which \( A \) equioscillates. Since Kilgore proved that \( \Gamma \) maps every optimal \( t \) to the point \( 0 \in \mathbb{R}^{n-1} \), Theorem 1 implies at once the validity of Bernstein's conjecture.

Corollary. If \( A \) equioscillates, then \( ||P_s|| < ||P_t|| \) for all \( s \neq t \).

We shall use the following two lemmas in the proof of Theorem 1.

Lemma 3. The map \( \Gamma \) is a local homeomorphism.

Proof. It suffices to show that

\[
\text{for all } t \in T, \det(\delta(\lambda_{i+1} - \lambda_i)/\delta t)_{i=1}^{n-1} \neq 0.
\]

Expanding this determinant by rows, one obtains

\[
det(\delta(\lambda_{i+1} - \lambda_i)/\delta t)_{i=1}^{n-1} = \sum_{k=1}^{n} (-1)^{k+1} J_k
\]

where we use again the abbreviation

\[
J_k := \det(\delta(\lambda_i)/\delta t)_{i=1}^{n}, \quad k \in [1,n].
\]

Hence, it suffices to show that

\[
(6) \quad \text{for some } \epsilon \in \{-1,1\} \text{ and all } t \in T, \ k \in [1,n], \ \epsilon(-1)^k J_k(t) > 0.
\]

But, since \( J_k \) is a continuous function of \( t \) and never vanishes on \( T \) by Kilgore's result, and \( T \) is connected, (6) is proven once we show that, for some \( t \in T, \)

\[
(7) \quad (-1)^k J_k(t)/J_1(t) < 0 \quad \text{for } k \in [2,n].
\]

This we could prove by observing that the last part of the argument for Kilgore's Theorem as we gave it in the preceding section gives precise information about the signs of the \((n-1)\)-minors of the matrix \( (q_i(t_j)) \) which is easily translated into the required information about the sign of \( J_k/J_1 \), all \( k \). But the following argument is more direct and establishes that
a fact which we need again later.

To prove (7) for some \( t \), observe that, since \( J_1(t) \neq 0 \), we can find a continuously
differentiable function \( G \) on some open neighborhood \( V \) of the point \( (\lambda_1(t))^n \) and an
open neighborhood \( U \) of \( t \) so that
\[
\lambda_1(s) = G(\lambda_2(s), \ldots, \lambda_n(s)) \quad \text{for all } s \in U.
\]

Also, by Cramer's rule,
\[
\frac{\partial\lambda_1}{\partial x_k} = \sum_{k=2}^{n} (-1)^{k} \frac{J_{k}/J_{1}}{\lambda_1} \frac{\lambda_k}{\lambda_{(k)}},
\]
and therefore
\[
\frac{\partial G}{\partial \lambda_k} = \frac{\partial\lambda_1}{\partial \lambda_k} = (-1)^{k} \frac{J_{k}/J_{1}}{\lambda_1} \quad \text{for } k \in [2, n].
\]

If now, for some \( k \in [2, n] \), \((-1)^{k} \frac{J_{k}/J_{1}}{\lambda_1} > 0 \), then we could find \( s \in U \) such that
\[
\lambda_1(s) = \lambda_1(t) \quad \text{for } i \in [2, n] \setminus \{k\}
\]
while
\[
\lambda_1(s) < \lambda_1(t) \quad \text{for both } i = 1 \quad \text{and } i = k,
\]

hence, for an optimal \( t \), \( s \) would also be optimal, yet \( \lambda_1 \) would not equioscillate,
contradicting Kilgore's result. This proves (7) for an optimal \( t \) and so proves (8)
and Lemma 3.

**Lemma 4.** The map \( \Gamma \) takes \( \mathcal{T} \) into \( \mathcal{X}^{n-1} \). Explicitly, if \( t \cdot x \in \mathcal{T} \) with
\( \Delta s_i = 0 \) for some \( i \in [0, n - 1] \), then \( \|\Gamma(t)\| \to \infty \).

**Proof.** Since \( a \Delta s_j = b - a \neq 0 \), there exists \( i \) such that \( \Delta s_i = 0 \) while either
\( \Delta s_{i-1} \) or \( \Delta s_{i+1} \) is not zero. Assume without loss that \( \Delta s_i = 0 \) and \( \Delta s_{i-1} \neq 0 \). Now
pick \( \hat{t} := (t_{i-1} + t_i)/2 \) and let \( x \) be an arbitrary point in \( (t_i, t_{i+1}) \). Then
\[
|\hat{t} - t_r|/|x - t_r| = \begin{cases} \frac{(t - t_{i-1})/(t_{i+1} - t_{i-1})}{2 \Delta t_{i-1}/\Delta t_i} & \text{for } r < i - 1, \\
\frac{1}{2} \Delta t_{i-1}/\Delta t_i & \text{for } r = i, i + 1, \\
1 & \text{for } r > i + 1.
\end{cases}
\]
Therefore, for all \( j \in [0,n] \),
\[
\frac{L_j(t)}{L_j(x)} = \prod_{r \notin j} \frac{t - t_r}{x - t_r} \geq \frac{1}{\Delta t_j} \prod_{r < i} \frac{t - t_{j-1}}{t_{i+1} - t_{j-1}} = \]
as \( \Delta t_j \to 0 \) and \( \Delta t_{i-1} \to \Delta t_{i-1} \neq 0 \). This shows that
\[
limit_{t \to x} \frac{L(t)}{L(x)} = \text{ for every } x \in (t_i, t_{i+1}).
\]
Therefore \( \lim_{t \to x} \frac{L_i(t)}{L_{i+1}(t)} = \infty \), and so \( \lim_{t \to x} \frac{L_{i+1}(t)}{L_i(t)} = -\infty \) since \( L_{i+1} > 1 \).
This proves that \( \lim_{t \to x} \| f(t) \| = 1 \) and so proves the lemma.

Theorem 1 is an immediate consequence of Lemmas 3 and 4 and of the following

Theorem (see, e.g., [2], [11]). A local homeomorphism \( f \) of \( \mathbb{R}^m \) to \( \mathbb{R}^m \) with
\[
\lim \| x \| \to \infty \| f(x) \| = \infty
\]
is a homeomorphism of \( \mathbb{R}^m \) onto \( \mathbb{R}^m \).

In a certain sense, this theorem is trivial since it is a special case of well known
facts regarding covering maps: The function \( f \) is a covering map for \( \mathbb{R}^m \) and so,
since \( \mathbb{R}^m \) is connected and simply connected, \( f \) is a universal covering map, therefore equivalent to any other universal covering map for \( \mathbb{R}^m \), in particular, \( f \) is equivalent
to the identity on \( \mathbb{R}^m \) (see, e.g., [11]; pp. 80-81). But, for completeness, we now
give an outline of a direct proof for the theorem.

The range of \( f \) is open, since \( f \) is locally 1-1 hence an open map. The range
of \( f \) is also closed since \( \lim f(x_r) = a \) implies that the sequence \( (f(x_r)) \) is
bounded, therefore, since \( f \) maps to \( \mathbb{R}^m \) by assumption, \( (x_r) \) is bounded,
hence can be assumed to converge to some \( x \) for which then \( f(x) = a \). This shows that
the range of \( f \) is \( \mathbb{R}^m \).

To show that \( f \) is 1-1, assume that \( f(x) = f(y) \) for some \( x, y \in \mathbb{R}^m \). The
function \( h : I \times I \to \mathbb{R}^m : (s,t) \mapsto (1 - t)h_0(s) + tf(x) \) with \( h_0 : I \to \mathbb{R}^m : s \mapsto f(sx + (1 - s)y) \)
and \( I := [0,1] \) is then a continuous map for which \( h(z) = f(x) \) for all \( z \) in the set
\[
B := ((0) \times I) \cup (I \times (1)) \cup ((1) \times I).
\]
But now, the assumptions on \( f \) allow one to "lift" the map \( h \), i.e., to show the existence
of a continuous map \( g : I \times I \to \mathbb{R}^m \) so that \( fg = h \) and \( g(0,0) = y \), therefore

-11-
\begin{align*}
g(s,0) &= sx + (1 - s)y \quad \text{for all } s \in I. \quad \text{This implies that both } x \text{ and } y \text{ belong to the connected set } g(I) \text{ on which } f \text{ is constantly equal to } f(x), \quad \text{and the fact that } f \text{ is locally } 1 - 1 \text{ now implies that } x = y.

\text{This proves the theorem, except for the technical part of } "\text{lifting} \ " h. \quad \text{But this can be proved, e.g., as is Lemma 3 of } [13; \ p. \ 71] \text{ after one has proved, as in the proof of Theorem 2 below, that curves can be lifted uniquely.}

\text{Acknowledgement. \ We are grateful to M. G. Crandall for pointing out to us the above theorem and for joining us in the construction of a proof.}

\text{We now prove Erdos' conjecture that, for every } t \in T,
\quad \lambda_1 \leq \min \lambda_i(t), \quad \max \lambda_i(t).

\text{Theorem 2. If } \lambda_i(s) = \lambda_i(t) \quad \text{for } i = 1, \ldots, n, \text{ then } s = t.

\text{Proof. If } \lambda_i(s) \neq \lambda_i(t) \quad \text{for all } i, \text{ then } s \neq t \text{ by Theorem 1. Hence assume that } \lambda_k(s) < \lambda_k(t) \quad \text{for some } k. \quad \text{This leads to a contradiction as follows.}

\text{The map } f: T \rightarrow R^{n-1}: r \mapsto \lambda(r) := (\lambda_i(r))_{i=1}^n \text{ is a local homeomorphism since } \det f'(r) = J_1(r) \neq 0 \quad \text{for all } r \in T. \quad \text{We can therefore } "\text{lift} \ " \text{any continuous curve } h:[0,1] \rightarrow R^{n-1} \text{ to a curve in } T \text{ as long as } \lambda_i \text{ stays bounded } "\text{along} \ " h. \quad \text{Specifically, let } h:[0,1] \rightarrow R^{n-1}: a \mapsto (1 - a)\lambda(s) + a\lambda(t).

\text{Since } f \text{ is locally } 1 - 1, \text{ there exists, for each } a \in [0,1], \text{ at most one continuous function } g_a:[0,a] \rightarrow T \text{ so that } g_a(0) = s \quad \text{and } f \circ g_a = h \text{ on } [0,a]. \quad \text{Let } A \text{ be the set of such } a. \quad \text{Then } A \text{ is not empty since it contains } 0. \quad \text{Further, } A \text{ is open since, for every } a \in (0,1), \text{ some neighborhood } V \text{ of } g_a(a) \text{ is mapped } 1 - 1 \text{ onto a ball around } h(a) \text{ by } f, \text{ hence } g_a \text{ can be extended continuously to the interval } [0,a] \cup h^{-1}(V) \text{ which contains } a \text{ in its interior. Finally, } A \text{ is closed. To see this, it is sufficient to prove that } [0,a] \subseteq A \text{ implies } a \in A, \text{ which can be done as follows. Since } [0,\hat{a}] \subseteq A, \text{ } g:[0,\hat{a}] \rightarrow T; a \mapsto g_a(a) \text{ defines a continuous map with } g(0) = s \text{ and } f \circ g = h \text{ on } [0,\hat{a}]. \text{ We claim that } g(a) \text{ converges to some point in } T \text{ as } a \rightarrow \hat{a}.

\text{Indeed, for } i \in [2,n], \lambda_i(g(a)) \text{ increases toward } h_i(\hat{a}) = (1 - \hat{a})\lambda_i(s) + \hat{a}\lambda_i(t) \quad \text{as } a \rightarrow \hat{a}, \quad \text{therefore, by (8) in the proof of Lemma 3, } \lambda_1(g(a)) \text{ decreases monotonely as}
\( a \to \hat{a}, \) hence must have a limit since it is bounded below (by 1, for instance). This shows that \( \lim_{a \to \hat{a}} \gamma(g(a)) \) exists in \( \mathbb{R}^{n-1} \), hence \( g(a) \) converges to some point \( \hat{r} \in T \), by Theorem 1. But then, the definition \( q(\hat{a}) := \hat{r} \) provides a continuous extension of \( g \) to \([0,\hat{a}]\) with \( f g(\hat{a}) = h(\hat{a}) \), hence \( \hat{a} \in A \).

This shows that \( A = [0,1] \), hence there exists \( g:[0,1] \to T \) continuous so that \( g(0) = \xi \) and \( f \circ g = h \). Therefore, with \( \tau := g(1) \), we have \( \lambda_i(\tau) = \lambda_i(\tau) \) for all \( i \in [2,n] \), while \( \lambda_1(\tau) < \lambda_1(\tau) < \lambda_1(\tau) \). But, since \( \lambda_k(\tau) < \lambda_k(\tau) \) for some \( k \), it follows that actually

\[
\lambda_1(\tau) < \lambda_1(\tau),
\]
either because \( k = 1 \), or else because \( \lambda_k \) strictly increases along the curve \( g \), therefore \( \lambda_1 \) must strictly decrease along that curve, by (8) in the proof of Lemma 3.

Consider now the curve

\[ h:[0,\hat{a}] \to \mathbb{R}^{n-1} : a \mapsto (\lambda_i(\tau) - a)^n. \]

By the preceding argument, there exists \( \hat{a} > 0 \) and a continuous function \( g:[0,\hat{a}] \to T \) so that \( f \circ g(a) = (\lambda_i(\tau) - a)^n \) for all \( a < \hat{a} \), while \( \lambda_i(g(a)) \) strictly increases from \( \lambda_1(\tau) \) at \( a = 0 \) to \( \tau \) at \( a = \hat{a} \). This implies that

\[
(\lambda_{i+1} - \lambda_i)(g(a)) = (\lambda_{i+1} - \lambda_i)(\tau) \quad \text{for all } i \in [2,n-1],
\]
while \( (\lambda_2 - \lambda_1)(g(a)) = (\lambda_2 - \lambda_1)(\tau) \) decreases from its value \( (\lambda_2(\tau) - \lambda_1(\tau)) \) at \( a = 0 \) to \( \tau \). But since \( \lambda_1(\tau) < \lambda_1(\tau) \), there exists therefore \( a \) so that

\[
(\lambda_2 - \lambda_1)(g(a)) = (\lambda_2 - \lambda_1)(\tau). \quad \text{But then } \Gamma(\tau) = \Gamma(g(a)) \text{ while } g(a) \neq \tau \text{ since, e.g., } \lambda_2(g(a)) < \lambda_2(\tau). \]

This contradiction to Theorem 1 finishes the proof of Theorem 2.

**Corollary.** For all \( k \in [1,n] \), the map \( \Gamma_k : T \to \mathbb{R}^{n-1} : \tau \mapsto (\lambda_i(\tau))_{i \neq k} \) is (globally) one-one.

**Proof.** If \( \Gamma_k(\tau) = \Gamma_k(\xi) \), then either \( \lambda_i(\tau) \leq \lambda_i(\xi) \) for all \( i \) or else \( \lambda_i(\tau) > \lambda_i(\xi) \) for all \( i \), hence \( \tau = \xi \) by Theorem 2.

We note that Theorem 2 provides another proof of the characterization of the optimal node vector \( \tau \) as the unique point in \( T \) for which \( \lambda \) equioscillates. Theorem 2
also shows that the optimal node vector is of no practical importance. For Brutman [3] has recently shown that, with

\[ t_i = (a + b + (a - b) \cos \frac{2i + 1}{2n + 2} \pi \cos \frac{\pi}{2n + 2}) / 2, \quad i \in [0, n] \]

the zeros of the Chebyshev polynomial of degree \( n + 1 \), adjusted to the interval \([a, b]\) in such a way that the first and the last zero fall on the end points of the interval,

\[ \max \lambda_i (t) - \min \lambda_i (t) \leq 0.5. \]

Numerical evidence strongly indicates that even

\[ \max \lambda_i (t) - \min \lambda_i (t) < 0.1925 \]

which would mean that the easily constructed node vector (9) produces an interpolation operator \( P_t \) whose norm is within 0.2 of the best possible value for all \( n \).
4. Trigonometric interpolation. In this section, we carry over the analysis of Sections 2 and 3 to the case of interpolation by trigonometric polynomials, i.e., by elements of

$$T_n := \text{span}(1, \cos x, \sin x, \ldots, \cos nx, \sin nx),$$
on \([0,2\pi]).$$

Because of the periodicity, the problem is altered slightly. Corresponding to each point \(t\) in

$$T := \{t \in \mathbb{R}^{2n} : 0 < t_1 < t_2 < \ldots < t_{2n} < 2\pi\},$$

we construct the linear map \(P_T\) of trigonometric interpolation in \(C(0,2\pi)\) at the \(2n+1\) points \(0 = t_0 < \ldots < t_{2n} < 2\pi\). In its Lagrange form,

$$P_T f = \sum_{i=0}^{2n} f(t_i) L_i,$$

with

$$L_i(x) := \frac{\prod_{k=0}^{2n} S(x - t_k)}{S(t_i - t_k)}, \text{ all } i \in [0,2n].$$

Here, we use the abbreviation

$$S(x) := \sin(x/2).$$

We have again \(\|P_T\| = \|A_T\|\) where \(A_T := \sum_i |L_i|\). Set

$$\lambda_i(t) := \max_{t_{i-1/2} \leq x \leq t_i} A_T(x), \text{ for all } i \in [1,2n+1],$$

with \(t_{2n+1} := 2\pi\).

Theorem 3. We have \(\|P_T\| = \lambda^* := \inf_{t \in T} \|P_T\|\) exactly when \(t = t^* := ((1/(2n+1)))_{2n}\), in which case \(A_T\) equioscillates. Furthermore, for any \(t \in T\setminus\{t^*\}\),

$$\min_i \lambda_i(t) < \lambda^* < \max_i \lambda_i(t).$$

Proof. We begin with a proof of the claim that

$$(10) \quad \det(\lambda_i(t)/\lambda_j(t))_{i=1}^{2n+1} = 0 \quad \text{for all } t \in T, k \in [1,2n+1].$$

Let \(P_T\) be the unique trigonometric polynomial of degree \(n\) which agrees with \(A_T\) on
\[ [t_{i-1}, t_i], \text{ for } i \in [1, 2n + 1]. \text{ Thus,} \]
\[
F_i'(t_j) = \begin{cases} 
(-1)^{i-1-j} & \text{for } j \in [0, i - 1], \\
(-1)^{j-1} & \text{for } j \in [i, 2n + 1].
\end{cases}
\]

Let \( \tau_i \) denote the unique point in \([t_{i-1}, t_i]\) at which \( \lambda_i \), and hence \( F_i \), takes on the value \( \lambda_i(t) \). Now
\[
3\lambda_i/2t_j = -F_i'(t_j) \frac{S(t_i - t_j)}{S(t_j - \tau_i)} = \frac{F_i'(t_j)}{S(t_i - t_j)} \sum_{k=0}^{2n} \frac{S(t_j - t_k)}{S(t_j - \tau_i)}
\]
which shows that \( \lambda_i \) is a continuously differentiable function on \( T \) and also shows that (10) is equivalent to
\[
(11) \quad \det(q_i(t_j))_{i=1}^{2n+1} = 0 \quad \text{for all } t \in T, k \in [1, 2n + 1],
\]
where
\[
q_i(x) = F_i'(x)/S(x - \tau_i), \quad i \in [1, 2n + 1].
\]

For the proof of (11), we make use of the following result corresponding to Lemma 6 of [6]. Denote by \( t_1^{(i)}, \ldots, t_{2n}^{(i)} \) the zeros of \( F_i' \) in \([0, 2\pi]\), necessarily all simple, in order.

**Lemma 5.** The zeros of \( F_1', \ldots, F_{2n+1}' \) lie in the pattern
\[
0 < t_1^{(i)} < t_2^{(i-1)} < \ldots < t_{2n}^{(i)} < \tau_{1}^{(i+1)} < \ldots < \tau_{2n-1}^{(i+1)} < \tau_{2n}^{(i+1)} < \ldots < \tau_{2n}^{(i+1)} < 2\pi
\]
for a certain \( i \in [1, 2n] \). Note that \( \tau_1^{(i)} = \tau_i \) and \( \tau_{k+1}^{(i)} = \tau_k \) for \( k \in [2, 2n + 1] \).

The proof of Lemma 5 follows exactly the same lines as the one given in Section 2 for Lemma 6 of [8] (including the use of the trigonometric analog of V. A. Markov's result), except that matters are a little easier since both \( F_i \) and \( F_i' \) have exactly \( 2n \) zeros in \([0, 2\pi]\), for all \( i \).

In order to use Lemma 5 in a proof of (11) much as Kilgore used Lemma 6 of [8] in his proof of (4), we must first show that
\[
0 < s_1 < \ldots < s_{2n} < 2\pi \quad \text{and} \quad \sum_{i=1}^{2n+1} a_{i} q_i(s_j) = 0 \quad \text{for all } j \in [1, 2n]
\]
implies that
\[
\sum_{i=1}^{2n+1} a_{i} q_i = 0.
\]
For this, observe that \( F_i'(x) = \text{const} \sum_{k=1}^{2n} S(x - r_k^{(i)}) \), therefore
\[
q_i(x) = \text{const} \sum_{k=1}^{2n} S(x - r_k^{(i)}) \text{ for all } i \in [1, 2n+1].
\]

Here, \( k \neq i-1 \) is meant to read \( k \neq 2n \) in case \( i = 1 \). This shows that \( q_i \) is not
2\( \pi \)-periodic, but 4\( \pi \)-periodic, and odd about 2\( \pi \), i.e., \( q_i(x + 2\pi) = -q_i(x) \), all \( x \).
Furthermore, the function \( p_i(x) := q_i(2x) \), all \( x \), is in
\[
\mathbb{T}_{2n-1} = \text{span}\{1, \cos x, \sin x, \ldots, \cos(2n-1)x, \sin(2n-1)x\}.
\]
Therefore, the hypotheses of (12) imply that the element \( \Sigma a_i p_i \) of \( \mathbb{T}_{2n-1} \) vanishes at
the 4\( n \) distinct points \( \hat{s}_1, \ldots, \hat{s}_{4n} \) with
\[
\hat{s}_j := \begin{cases} 
\frac{\pi}{2} & \text{for } j \in [1, 2n], \\
\frac{\pi}{2} + \pi & \text{for } j \in [2n+1, 4n].
\end{cases}
\]
and so \( \Sigma a_i p_i = 0 \), proving (12).

The proof of (11) proceeds now as the proof of (4) in Section 2, and, with (10) thus
established, the reasoning in the proofs of Theorems 1 and 2 in Section 3 applies directly
to finish the proof of Theorem 3.

We note in passing that Ehlich & Zeller [5] have proved a formula for \( \lambda^* \) in the
trigonometric case,
\[
(13) \quad \lambda^* = \left( 1 + 2 \sum_{k=0}^{n-1} \left( \frac{\sin \left( \frac{(2k+1)\pi}{2n+1} \right)}{\sin \left( \frac{(2n+1)\pi}{2(2n+1)} \right)} \right)^{-1} \right) / (2n+1).
\]

Finally, the above analysis applies without essential change to the case when we
also fix \( t_{2n} \) at some point \( b < 2\pi \) and consider the optimal choice of \( t_1 < \ldots < t_{2n-1} \)
in (0, b) for trigonometric interpolation.

5. Postscript. After completion of this work in March, we received word from
Theodore Kilgore that he had succeeded in proving Bernstein's conjecture. His proof
proceeds along different lines.
REFERENCES


prove the following conjectures.

(a) **Bernstein**: \( \|A_t\|_\infty \) is minimal when \( \lambda_1(t) = \cdots = \lambda_n(t) \).

(b) **Erdős**: If \( \lambda_i(t) = \lambda^* \), \( i = 1, \ldots, n \), then for all \( s \in T \setminus \{t\} \),

\[
\min_{i} \lambda_i(s) < \lambda^* < \max_{i} \lambda_i(s).
\]

Analogous results are proven for trigonometric interpolation.

These results are of interest since \( \|A_t\|_\infty \) gives the norm of the linear map of polynomial interpolation on the continuous functions and therefore bounds the effect of noisy data on their polynomial interpolant and shows how close the interpolation error is to the best possible error by any method.
PROOF OF THE CONJECTURES OF BERNSTEIN AND ERDOS CONCERNING THE OPTIMAL NODES FOR POLYNOMIAL INTERPOLATION.

Carl de Boor and Allan Pinkus

Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

For each \( t \in T := \{ t \in \mathbb{R}^{n-1} : a < t_1 < \ldots < t_{n-1} < b \} \), let \( \Lambda_t(x) \) be the Lebesgue function of the process of polynomial interpolation on \([a,b] \) by polynomials of degree \( \leq n \) at the points \( a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b \). Let \( \lambda_i(t) := \max_{t_{i-1} < x < t_i} \Lambda_t(x), \ i = 1, \ldots, n \). Based on work of Kilgore [8], we...