JOURNAL OF APPROXIMATION THEORY **21**, 43–59 (1977)

Best Approximation by Polynomials

STEPHEN D. FISHER*

Department of Mathematics, Northwestern University, Evanston, Illinois 60201

Communicated by Oved Shisha

Received September 5, 1974

INTRODUCTION

This paper gives an account of some old and new results on the topic of best approximation by polynomials and related functions. The starting point is the following theorem, discovered simultaneously by Favard [5] and Achieser and Krein [1], which improves part of the famous 1911 theorem of D. Jackson.

THEOREM 1.1. *Let W_n^* consist of all functions f on $[-\pi, \pi]$ for which $f, f', \dots, f^{(n-1)}$ are continuous and 2π-periodic and $|f^{(n-1)}(x) - f^{(n-1)}(y)| \leq |x - y|$ for all $x, y \in [-\pi, \pi]$. Let T_m be the linear span of $\{1, \cos x, \dots, \cos mx, \sin x, \dots, \sin mx\}$ and set*

$$\beta_{nm} = \sup_{f \in W_n^*} \inf_{s \in T_m} \| f - s \|_{L^\infty(-\pi,\pi)}. \tag{0.1}$$

Then $\beta_{nm} = K_n(m+1)^{-n}$ *where*

$$K_n = (4/\pi) \sum_{j=0}^{\infty} (-1)^{j(n+1)} (2j+1)^{-n-1}. \tag{0.2}$$

Furthermore, there is a solution f_0 of (0.1) *such that* $f_0^{(n)}(x) = \operatorname{sign}(\cos(m+1)x)$ *if n is even,* $f_0^{(n)}(x) = \operatorname{sign}(\sin(m+1)x)$ *if n is odd. If g_0 is any other solution to* (0.1) *then $g_0(x) = \lambda f_0(x - x_0) + c$ where $\lambda = \pm 1$, $x_0 \in [-\pi, \pi]$, and c is a constant.*

In Section 1 we give a proof of Theorem 1.1 which presents some new features. In Section 2 we investigate the analogous problem for approximation by algebraic polynomials of degree m on $[-1, 1]$ where the periodicity requirement on f in the definition of W_n^* is dropped. Theorem 2.1 states that any solution f_0 has the property that $f_0^{(n)}$ assumes only the values 1 and -1 and has exactly $m - n + 1$ sign changes in $(-1, 1)$; that is, f_0 is a perfect spline with exactly $m - n + 1$ knots. When $m = n - 1$, the lowest value of m for which the problem makes any sense, the solution f_0 is a multiple of the

* Research supported by NSF Grant GP-43955.

ISSN 0021-9045

Chebyshev polynomial of degree n. We then show in Theorem 2.2 that Theorem 1.1 can be combined with a result about entire functions of exponential type to give a simple proof of a theorem of S. N. Bernstein on the asymptotic behavior as $m \to \infty$ of the best constant. In Section 3 we apply Theorems 1.1 and 2.2 to give a brief proof of a theorem of M. G. Krein on best approximation on the real line by entire functions of exponential type less than σ. We close this circle of ideas by showing that Krein's theorem easily implies the value of β_{nm} given in Theorem 1.1. In Section 4 we give a very short proof of Babenko's theorem, the "analytic" version of the Favard–Achieser–Krein theorem. Section 5 gives bounds for best approximation when the class of functions is determined by a modulus of continuity condition on the nth derivative.

We make use of several standard notations. If f is a continuous function on the interval I, then $E_m(f; I)$ is the distance from f to the space π_m of algebraic polynomials of degree m or less in the supremum norm

$$E_m(f; I) = \inf_{a_0, \ldots, a_m} \left\| f(x) - \sum_0^m a_j x^j \right\|_{L^\infty(I)}.$$

If g is a continuous, periodic function on $[-\pi, \pi]$, then $E_m^*(g)$ is the distance from g to the space T_m of trigonometric polynomials of degree m or less in the supremum norm

$$E_m^*(g) = \inf_{\substack{a_0, \ldots, a_m \\ b_1, \ldots, b_m}} \left\| g(x) - \left[\frac{a_0}{2} + \sum_1^m (a_k \cos kx + b_k \sin kx) \right] \right\|_{L^\infty[-\pi,\pi]}.$$

We make constant use of the fact that the dual space of $C(I)$ is the space of finite regular Borel measures on I and that the dual space of the continuous periodic functions on $[-\pi, \pi]$ is the space of periodic finite regular Borel measures μ on $[-\pi, \pi]$: $\mu(-\pi) = \mu(\pi)$. We also use the very standard duality relation: if Y is a subspace of a Banach space X and $x_0 \in X$, then

$$\inf\{\| x_0 - x \| : x \in Y\} = \sup\{l(x_0) : \| l \| \leq 1, l \in Y^\perp\}$$

where $Y^\perp$ consists of those elements of the dual space of X which vanish on Y.

1. Approximation by Trigonometric Polynomials

Proof of Theorem 1

We do the proof when n is even; the proof for n odd requires only minor modifications. Let

$$D_n(x) = (1/2\pi)(-1)^{n/2} \sum_{\substack{k=-\infty \\ k \neq 0}}^{\infty} k^{-n} e^{ikx}.$$

For each $f \in W_n{}^*$ we know $f^{(n)}$ exists a.e. and is bounded by one and further

$$f(t) = \int_{-\pi}^{\pi} f^{(n)}(x)\, D_n(t - x)\, dx.$$

Note that $\int_{-\pi}^{\pi} f^{(n)}(x)\, dx = 0$ since $f^{(n-1)}(\pi) = f^{(n-1)}(-\pi)$. Moreover, if h is any function in the unit ball of L^∞ with mean-value 0, then there is a unique function $H \in W_n{}^*$ with $H^{(n)} = h$; H is just the convolution of h and D_n.

Let λ be any (real) periodic measure on $[-\pi, \pi]$ with total variation at most one which is zero on T_m and let $f \in W_n{}^*$. Then by the duality relation we have

$$\begin{aligned}\beta_{nm} &\geqslant \left| \int_{-\pi}^{\pi} f(t)\, d\lambda(t) \right| \\ &= \left| \int_{-\pi}^{\pi} f^{(n)}(x)\, D_\lambda(x)\, dx \right|\end{aligned}$$

where

$$D_\lambda(x) = \int_{-\pi}^{\pi} D_n(t - x)\, d\lambda(t).$$

Now let λ vary over all measures orthogonal to T_m of total variation at most one and let $f^{(n)}$ vary over all functions in the unit ball of L^∞ with mean-value zero. Due to the duality relations we find that

$$\beta_{nm} = \sup_\lambda \{\text{distance in } L^1 \text{ from } D_\lambda \text{ to the constants}\}.$$

Actually, this supremum is a maximum since both $W_n{}^*$ and the unit ball of the space of measures are compact. Now we choose a specific λ; λ consists of $2m + 3$ point masses at the points $-\pi + k\pi/m + 1$, $k = 0, 1,..., 2m + 2$ with weights $1/4m + 4$, $-1/2m + 2$, $1/2m + 2,...$, $-1/2m + 2$, $1/4m + 4$, respectively. For this λ,

$$D_\lambda(x) = (m + 1)^{-n}\, D_n((m + 1)\, x)$$

since λ is orthogonal to $\cos kx$ unless k is a multiple of $m + 1$ in which case the integral has the value 1. Since λ is orthogonat to T_m and has total variation one we have

$$\beta_{nm} \geqslant (m + 1)^{-n}\ [\text{distance of } D_n((m + 1)\, x) \text{ to } \mathbb{R} \text{ in } L^1].$$

Note, however, that the L^1 distance of $D_n((m + 1)\, x)$ to the constants is the same as the distance of $D_n(x)$ to the constants by periodicity and that this number is four times

$$\int_0^{\pi/2} D_n(x)\, dx = (1/\pi) \sum_{k=0}^{\infty} (2k + 1)^{-n-1}\, (-1)^k.$$

Hence,

$$(m+1)^n \beta_{nm} \geqslant (4/\pi) \sum_{k=0}^{\infty} (-1)^k (2k+1)^{-n-1} = K_n .$$

On the other hand, let S be the best L^1 approximation to D_n from T_m; then

$$\| D_n - S \|_1 = (m+1)^{-n} K_n;$$

see [10, p. 114] for the details. If $f \in W_n^*$, then

$$f(x) - s(x) = \int_{-\pi}^{\pi} f^{(n)}(t)[D_n(x-t) - S(x-t)]\, dt$$

where $s(x)$ is some element of T_m. Thus, $\beta_{nm} \leqslant \| D_n - S \|_1$ so that

$$\beta_{nm} = (m+1)^{-n} K_n.$$

Suppose now that $F \in W_n^*$ and that the distance of F from T_m is β_{nm}. Then

$$F(x) - s(x) = \int_{-\pi}^{\pi} F^{(n)}(t)[D_n(x-t) - S(x-t)]\, dt$$

where $s \in T_m$ so that

$$\begin{aligned} \beta_{nm} &\leqslant \| F - s \|_\infty = | F(x_0) - s(x_0) | \\ &\leqslant \| F^{(n)} \|_\infty \| D_n(x_0 - t) - S(x_0 - t) \|_1 \\ &\leqslant \beta_{nm} . \end{aligned}$$

Hence, $F^{(n)}(t)(D_n(x_0 - t) - S(x_0 - t)) \geqslant 0$ a.e. and $|F^{(n)}(t)| = 1$ where $D_n(x_0 - t) - S(x_0 - t) \neq 0$. However, $D_n - S$ changes sign at the points $k\pi/m+1$, $k = 0, \pm 1, \ldots, \pm m$ and only there [10, p. 118], so that F must have the indicated form.

Corollary 1.2. *Let r be a positive integer and let U_r consist of all functions f in W_n^* for which $\hat{f}(k) = (2\pi)^{-1} \int_{-\pi}^{\pi} f(t)\, e^{-ikt}\, dt = 0$ when $|k| \leqslant r$. Then*

$$\max_{f \in U_r} \| f \|_\infty = (r+1)^{-n} K_n .$$

Proof. Let H be a function in U_r which attains the maximum value of the left-hand side. If Q is any element in T_r, then the convolution of $H^{(n)}$ and Q is identically zero. Hence,

$$H(x) = \int_{-\pi}^{\pi} H^{(n)}(t)[D_n(x-t) - Q(x-t)]\, dt$$

so that

$$\begin{aligned} \| H \|_\infty &\leqslant \text{distance of } D_n \text{ to } T_r \text{ in } L^1 \\ &= (r+1)^{-n} K_n \end{aligned}$$

by Theorem 1.1. On the other hand, the function

$$G(x) = (r+1)^{-n} \frac{4}{\pi} \sum_{k=0}^{\infty} (-1)^{k(n+1)} (2k+1)^{-n-1} \cos((2k+1)(r+1)x)$$

is in U_r and $\| G \|_\infty = (r+1)^{-n} K_n$. This completes the proof.

2. Approximation by Algebraic Polynomials

Let I be the closed interval $[-1, 1]$ and let π_m denote the space of algebraic polynomials of degree m or less. Let W_n consist of all functions f on I for which $f, f', \ldots, f^{(n-1)}$ are absolutely continuous and $| f^{(n)} | \leqslant 1$ a.e. Let

$$\alpha_{nm} = \sup_{f \in W_n} E_m(f). \tag{2.1}$$

In this section we prove two theorems. The first describes a property of any solution of (2.1); the second is a simple proof of a theorem of S. N. Bernstein on the asymptotic behavior of α_{nm} as $m \to \infty$.

THEOREM 2.1. *Let f be a solution of* (2.1). *Then $f^{(n)}$ assumes only the values* 1 *and* -1 *and has exactly $m - n + 1$ sign changes in* $(-1, 1)$. *If $m = n - 1$, then f is a constant multiple of the* nth *Chebyshev polynomial.*

Proof. The proof closely resembles the proof of the first part of Theorem 1.1.

If $f \in W_n$ and if $f^{(\nu)}(-1) = 0$ for $\nu = 0, \ldots, n-1$, then

$$f(x) = \int_{-1}^{1} f^{(n)}(t)\, \theta(x, t)\, dt$$

where $\theta(x, t) = (x - t)_+^{n-1}/(n-1)!$; that is, $\theta(x, t)$ equals $(x - t)^{n-1}/(n-1)!$ for $-1 \leqslant t \leqslant x$ and 0 for $x \leqslant t \leqslant 1$. Further, if h is in the unit ball of $L^\infty(I)$, then $H(x) = \int_I h(t)\, \theta(x, t)\, dt$ is in W_n.

Let λ be a (real) measure on I which is orthogonal to π_m and has total variation at most one; let $f \in W_n$. Then by the duality relation

$$\begin{aligned} \alpha_{nm} &\geqslant \left| \int_I f(x)\, d\lambda(x) \right| \\ &= \left| \int_I f^{(n)}(t)\, F_\lambda(t)\, dt \right| \end{aligned}$$

where

$$F_\lambda(t) = \int_I \theta(x, t)\, d\lambda(x).$$

As λ runs over the measures of total variation at most one which are orthogonal to π_m and as f runs over W_n we find that

$$\alpha_{nm} = \sup_{\lambda} \|F_\lambda\|_{L^1(I)}. \tag{2.2}$$

The supremum is actually a maximum and equality holds in (2.2) for a measure λ with $m + 2$ points in its support. To see this let H be a function in W_n whose distance to π_m is α_{nm}. Such a function exists since W_n is compact. Let $P \in \pi_m$ be the polynomial of degree m which is closest to H. Then there are points $-1 \leqslant x_0 < x_1 < \cdots < x_{m+1} \leqslant 1$ at which $H(x_k) - P(x_k) = \alpha_{nm}(-1)^k$, $k = 0,..., m+1$. Let λ be a measure supported on $\{x_k : k = 0,..., m+1\}$ of total mass one which is orthogonal to π_m; let λ_k be the weight of λ at x_k, $k = 0,..., m+1$; it is easy to see that $(-1)^k \lambda_k > 0$ and hence $\sum_0^{m+1} (-1)^k \lambda_k = 1$. Then

$$\begin{aligned} \alpha_{nm} &= \sum_{k=0}^{m+1} (H(x_k) - P(x_k))\lambda_k \\ &= \int_I (H - P)\, d\lambda = \int_I H\, d\lambda \\ &= \int_I H^{(n)} F_\lambda\, dx \\ &\leqslant \|F_\lambda\|_1 \leqslant \alpha_{nm}. \end{aligned}$$

Hence, equality holds in (2.2). We have also shown that if $H \in W_n$ is at distance α_{nm} from π_m then there is a measure λ, depending on H, with $m + 2$ points in its support which is orthogonal to π_m and which satisfies

$$\alpha_{nm} = \int_I H^{(n)} F_\lambda.$$

Hence, $H^{(n)}F_\lambda \geqslant 0$ a.e. and $|H^{(n)}| = 1$ where $F_\lambda \neq 0$. However, F_λ is the $(m - n + 1)$st derivative of the function

$$B(x) = \int_I (x - t)_+^m / m!\, d\lambda(t)$$

which is a B-spline and so F_λ has precisely $m - n + 1$ zeros in $(-1, 1)$; see [4, p. 74]. In particular, in the case $m = n - 1$, the smallest value of m for which α_{nm} is finite, we find that $F_\lambda > 0$ on $(-1, 1)$ and so H is a polynomial of degree n; clearly, H must be the nth Chebyshev polynomial suitably normalized and

$$\alpha_{n,n-1} = 2^{-n+1}/n!$$

For emphasis we restate the primary conclusion of Theorem 2.1. *Each solution of* (2.1) *is a perfect spline with exactly* $m - n + 1$ *knots on* $(-1, 1)$.

Comments. (1) It would be most interesting to prove that there is only one solution of (2.1) and to locate its knots. Of course, the solution in the trigonometric case (Theorem 1.1) is also a perfect spline; its knots are regularly spaced at intervals of $\pi/m + 1$. There is a temptation to try the substitution $x = \cos\theta$ to turn the algebraic problem into a trigonometric problem. Of course this will not work since the condition $|f^{(n)}| \leqslant 1$ is not carried into anything useful.

(2) The formula

$$\alpha_{nm} = \max_{f \in W_n} \max_{\lambda} \int_{-1}^{1} f\, d\lambda$$

where λ is the sum of $m + 2$ point masses shows that α_{nm} is the maximum of the $m + 1$st divided difference of f at points $-1 \leqslant x_0 < x_1 < \cdots < x_{m+1} \leqslant 1$ where f is restricted only by the condition that $|f^{(n)}| \leqslant 1$ on $[-1, 1]$.

(3) In the trigonometric case there is a constant C such that $\beta_{nm}(m + 1)^n \leqslant C$ for all choices of m and n. (In fact, $C = \pi/2$ will work). There is no such number for the algebraic case since $\alpha_{n,n-1} = 2^{-n+1}(n!)^{-1}$, and hence $\alpha_{n,n-1}n^n \to \infty$ as $n \to \infty$. However, for n fixed $\alpha_{nm}m^n$ does remain bounded as Theorem 2.2 shows.

We now use Theorem 1.1 and some other facts to give a proof of the following theorem of Bernstein [2], proved in 1947; see [11, p. 293] for a proof in the English language.

THEOREM 2.2. $\operatorname{limit}_{m\to\infty} m^n \alpha_{nm} = K_n$ *where* K_n *is the constant given by* (0.2).

Proof. For the first part of the proof it is technically somewhat easier to work on $[-\pi, \pi]$. Let $\bar{\alpha}_{nm}$ be the number analogous to α_{nm} for the interval $[-\pi, \pi]$; then $\bar{\alpha}_{nm} = \pi^n \alpha_{nm}$. We shall show that

$$\limsup_{m\to\infty} m^n \bar{\alpha}_{nm} \geqslant \pi^n K_n .$$

Let h lie in the unit ball of $L_\infty(-\pi, \pi)$; then $H(x) = \int D_n(x - t)\, h(t)\, dt$ is in $W_n(-\pi, \pi)$ and differs from the usual nth indefinite integral of h by an algebraic polynomial of degree $n - 1$. Hence, as in the proof of Theorem 1.1 or Theorem 2.1

$$\bar{\alpha}_{nm} = \max\{\| D_\lambda \|_{L^1}\}$$

where the supremum is taken over all measures λ on $[-\pi, \pi]$ which are orthogonal to π_m and which have total variation at most one.

Let ϵ be a small positive number and let r be the greatest integer in $m/(1 + \epsilon)\pi$. We shall need the following standard fact, which dates back at least to Bernstein in 1912 (See [10, p. 77]).

LEMMA 2.3. *Let $R > 1$ and let E_R be the ellipse $x = (\pi/2)(R + R^{-1})\cos\theta$, $y = (\pi/2)(R - R^{-1})\sin\theta$, $0 \leq \theta \leq 2\pi$. Suppose f is holomorphic on and within E_R and bounded by M on E_R. Then*

$$E_m(f; -\pi, \pi) \leq 2MR^{-m}(R - 1)^{-1}. \tag{2.3}$$

Continuing the proof of Theorem 2.2 we take $f(z) = e^{ikz}$ in Lemma 2.3. The maximum of $|f(z)|$ on E_R is at most $\exp[(k/2)(R - R^{-1})\pi]$. When $0 \leq k \leq r$, this in turn is no more than $\exp[m(R - R^{-1})/2(1 - \epsilon)]$. Choose R so close to 1 that

$$\exp[(R - R^{-1})/2(1 - \epsilon)] < \rho R$$

where $\rho < 1$. Then we have the estimate

$$E_m(e^{ikx}; -\pi, \pi) \leq 2(R - 1)^{-1}\rho^m \tag{2.4}$$

for $0 \leq k \leq r$.

Now let h be any continuous, 2π-periodic function on $[-\pi, \pi]$ which is bounded by 1. Let λ be a measure on $[-\pi, \pi]$ which is orthogonal to π_m and has total variation 1 or less. Then

$$\int_{-\pi}^{\pi} h(t)\, D_\lambda(t)\, dt = \int H(x)\, d\lambda(x)$$

$$= \sum_{k=-r}^{r} |k|^n \hat{h}(k)\, \hat{\lambda}(k)$$

where $\hat{h}(k) = (1/2\pi)\int_{-\pi}^{\pi} h(t)\, e^{-ikt}\, dt$ and $\hat{\lambda}(k) = \int_{-\pi}^{\pi} e^{-ikx}\, d\lambda(x)$. Set

$$H_r(x) = \sum_{|k| \leq r} |k|^n \hat{h}(k)\, \hat{\lambda}(k)\, e^{ikx}.$$

Then

$$\int_{-\pi}^{\pi} h(t)\, D_\lambda(t)\, dt = H_r(0) = \sum_{\substack{k=-r \\ k \neq 0}}^{r} |k|^n \hat{h}(k)\, \hat{\lambda}(k). \tag{2.5}$$

The estimate (2.4) shows that $|\hat{\lambda}(k)| \leq 2(R - 1)^{-1}\rho^m$ if $|k| \leq r$ and thus the second sum in (2.5) is no larger than $4m(R - 1)^{-1}\rho^m$. Furthermore, the nth derivative of H_r (recall n is even) differs from

$$\sum_{-r}^{r} \hat{h}(k)\, \hat{\lambda}(k)\, e^{ikx} = \int_{-\pi}^{\pi} h(x - t)\, d\lambda(t) \tag{2.6}$$

by the term

$$\sum_{|k|>r} \hat{h}(k)\, \hat{\lambda}(k)\, e^{ikx}. \tag{2.7}$$

We know that (2.7) does not exceed $4m(R-1)^{-1}\rho^m$ and (2.6) is clearly bounded by 1. (We extend h to be 2π-periodic on $\mathbb{R}$.) Hence, Corollary 1.2 implies that

$$|H_r(0)| \leqslant \|H_r\|_\infty \leqslant (r+1)^{-n} K_n(1+4m(R-1)^{-1}\rho^m).$$

This implies that

$$\left|\int_{-\pi}^{\pi} h(t)\, D_\lambda(t)\, dt\right|$$
$$\leqslant (r+1)^{-n} K_n(1+4m(R-1)^{-1}\rho^m) + 4m(R-1)^{-1}\rho^m$$
$$\leqslant \pi^n(1+\epsilon)^n m^{-n}K_n + 4m(R-1)^{-1}\rho^m(1+(r+1)^{-n}K_n).$$

Hence, $\bar{\alpha}_{nm} = \sup_{\|h\|_\infty \leqslant 1} |\int_{-\pi}^{\pi} h(t)\, D_\lambda(t)\, dt|$ is also bounded by the same quantity so that

$$m^n\bar{\alpha}_{nm} \leqslant \pi^n(1+\epsilon)^n K_n + 4m^{n+1}\rho^m(R-1)^{-1}(1+(r+1)^{-n}K_n)$$

and this yields $\limsup_{m\to\infty}(m^n\bar{\alpha}_{nm}) \leqslant \pi^n(1+\epsilon)^n K_n$. Since ϵ is arbitrary, we have established

$$\limsup_{m\to\infty}(m^n\bar{\alpha}_{nm}) \leqslant \pi^n K_n\,.$$

To prove that $\liminf_{m\to\infty}(m^n\alpha_{nm}) \geqslant K_n$ we return to the interval $I = [-1, 1]$ and use a few elementary facts about entire functions of exponential type.

Let $\epsilon > 0$ be given. Let m be a positive integer and define

$$F_m(x) = (4/\pi)\sum_0^\infty (-1)^k (2k+1)^{-n-1}\cos((2k+1)(1+\epsilon)\, mx),$$
$$-\infty < x < \infty$$

and set $F(x) = F_1(x)$. Suppose for each m in a sequence of $m \to \infty$ there is a polynomial p_m of degree m with

$$\|F_m - p_m\|_{L^\infty(I)} \leqslant (1-\delta)\, K_n\,, \qquad \delta > 0,$$

where δ is independent of m. Then a change of variables yields

$$\|F(x) - q_m(x)\|_{L^\infty(-m,m)} \leqslant (1-\delta)\, K_n$$

where $q_m(x) = p_m(x/m)$, $-\infty < x < \infty$. Now

$$q_m^{(k)}(0) = m^{-k}p_m^{(k)}(0), \qquad k = 0,\ldots, m$$

and by a classical inequality of Markov [3. 46, (83)],

$$|p_m^{(k)}(0)| \le m^k \max_{|x| \le 1} |p_m(x)|$$
$$\le 2K_n m^k.$$

Hence,

$$|q_m^{(k)}(0)| \le 2K_n \qquad \text{for } k = 0, \ldots, n.$$

This implies that a subsequence of $\{q_m\}$ converges uniformly on compact subsets of the plane to an entire function G of exponential type 1 or less. Clearly, G satisfies the inequality

$$\|F - G\|_{(-\infty,\infty)} \le (1 - \delta) K_n.$$

However, since F is periodic with period $2\pi(1+\epsilon)^{-1}$ we may assume G has this period and hence G is constant. But no constant is within distance K_n of F. This contradiction shows that

$$\liminf_{m\to\infty} E_m(F_m ; I) \ge K_n$$

However, $(1+\epsilon)^{-n} m^{-n} F_m \in W_n$ so that

$$\liminf_{m\to\infty} (m^n \chi_{nm}) \ge K_n(1+\epsilon)^{-n}.$$

Thus,

$$\liminf_{m\to\infty} m^n \chi_{nm} \ge K_n$$

and the theorem is proved.

3. Approximation on the Line by Entire Functions of Exponential Type

Let E_σ, $\sigma > 0$, be the space of entire functions of exponential type *less than* σ which are bounded on the real axis. Such function f necessarily satisfies the growth condition

$$|f(x + iy)| \le e^{\rho|y|}\left(\sup_{-\infty<t<\infty} |f(t)|\right)$$

for some $\rho < \sigma$. In this section we use Theorems 1.1 and 2.2 to give a simple proof of the following theorem of Krein [9].

THEOREM 3.1. *Let V_n consist of all bounded functions f on $(-\infty, \infty)$ which satisfy $|f^{(n)}| \le 1$ on $(-\infty, \infty)$. Let*

$$\gamma_{n\sigma} = \sup_{f\in V_n} \inf_{G\in E_\sigma} \|f - G\|_{(-\infty,\infty)}.$$

Then $\gamma_{n\sigma} = \sigma^{-n} K_n$ where K_n is the constant given in (0.2).

Proof. Again we take n to be even. Let

$$F_\sigma(x) = \frac{1}{\sigma^n}\frac{4}{\pi}\sum_{k=0}^{\infty}(-1)^k\frac{\cos(2k+1)\,\sigma x}{(2k+1)^{n+1}}, \qquad -\infty < x < \infty.$$

Then $F_\sigma \in V_n$; suppose $G \in E_\sigma$ and $\| G - F_\sigma \| \leqslant K_n(1-\delta)\,\sigma^{-n}$, $\delta > 0$. Let $F(x) = \sigma^n F_\sigma(x/\sigma)$ and $H(z) = \sigma^n G(z/\sigma)$. Then $\| F - H \|_{(-\infty,\infty)} \leqslant (1-\delta)\,K_n$ and H is entire of exponential type less than 1. Since F is 2π-periodic, we may assume H is also and hence H is constant. But the distance from F to the constants is K_n. Thus, the distance from F_σ to E_σ is $K_n\sigma^{-n}$ so that $\gamma_{n\sigma} \geqslant K_n\sigma^{-n}$.

On the other and, let $f \in V_n$, $\epsilon > 0$, and let I_m be the interval $I_m = [-m/\sigma(1-\epsilon), m/\sigma(1-\epsilon)]$. Let p_m be the best approximation to f on I_m from π_m, let $g_m(x) = f(mx/\sigma(1-\epsilon))$ and $q_m(x) = p_m(mx/\sigma(1-\epsilon))$. Then

$$\begin{aligned}\| f - p_m \|_{I_m} &= \| g_m - q_m \|_I \\ &= E_m(g_m; I) \\ &\leqslant \alpha_{nm} m^n/\sigma^n(1-\epsilon)^n\end{aligned}$$

so that by Theorem 2.2

$$\limsup_{m\to\infty} E_m(f; I_m) \leqslant K_n/\sigma^n(1-\epsilon)^n.$$

Hence, $\| q_m \|_I \leqslant C$ for all m so that

$$\begin{aligned}| p_m^{(k)}(0)| &= \sigma^k(1-\epsilon)^k\, m^{-k}\, | q_m^{(k)}(0)| \\ &\leqslant C\sigma^k(1-\epsilon)^k\end{aligned}$$

by Markov's inequality. Hence, some subsequence of $\{p_m\}$ converges uniformly on compact subsets of the plane to an entire function G of exponential type less than σ which must satisfy

$$\| f - G \|_{(-\infty,\infty)} \leqslant K_n/\sigma^n(1-\epsilon)^n.$$

Hence,

$$\inf_{G\in E_\sigma} \| f - G \|_{(-\infty,\infty)} \leqslant K_n/\sigma^n(1-\epsilon)^n$$

for each $f \in V_n$ and each $\epsilon > 0$, so that

$$\gamma_{n\sigma} \leqslant K_n/\sigma^n$$

and this completes the proof.

To complete the circle of ideas in Theorems 1.1, 2.2, and 3.1 we show that Theorem 3.1 easily implies the value of the constant β_{nm} in Theorem 1.1. Let

$f \in W_n^*$ and extend f periodically to the line. Then, because a 2π-periodic function in E_{m+1} is a trigonometric polynomial of degree m or less, we have

$$\inf_{G \in E_{m+1}} \| f - G \|_{L^\infty(-\infty,\infty)} = \inf_{T \in T_m} \| f - T \|_{L^\infty(-\pi,\pi)} .$$

But the left-hand side does not exceed $K_n(m+1)^{-n}$ by Theorem 3.1. Hence, $\beta_{nm} \leqslant K_n(m+1)^{-n}$. On the other hand, the function

$$F_m(x) = (m+1)^{-n} \frac{4}{\pi} \sum_0^\infty (\cos(2k+1)(m+1)x)(-1)^{k(n+1)}/(2k+1)^{n+1}$$

lies in W_n^* and has $2m+3$ alternation on $[-\pi, \pi]$ and hence the best approximation to F_m from T_m is zero; thus

$$\beta_{nm} \geqslant \| F_m \|_\infty \geqslant | F_m(0) | = K_n/(m+1)^n .$$

4. An Analytic Version of the Favard–Achieser–Krein Theorem

If we view Theorem 1.1 on the unit circle, then it says that a continuous function whose nth derivative is never larger than 1 can be approximated by a sum of the form

$$s(x) = \sum_{-m}^{m} c_k e^{ikx}$$

with an error of no more that $K_n(m+1)^{-n}$. The "analytic" version of this theorem would be to approximate a function whose negative Fourier coefficients are zero by a sum of the form

$$p(x) = \sum_0^m c_k e^{ikx} ;$$

that is, approximate an analytic function on the unit circle by a polynomial of degree m in the complex variable z. This is the content of the theorem of Babenko [10, p. 126] which is somewhat more general.

THEOREM 4.1. *Let $R \geqslant 1$ and let $A_n(R)$ consist of all analytic functions f on $|z| < R$ which satisfy $|f^{(n)}(z)| \leqslant 1$ for $|z| < R$. Then for $m \geqslant n-1$,*

$$\sup_{f \in A_n(R)} \inf_{p \in \pi_m} \| f(z) - p(z) \|_{|z|=1} = \frac{(m-n+1)!}{(m+1)!} R^{-(m-n+1)} . \tag{4.1}$$

If f is a solution of (4.1) then $f(z) = Cz^{m+1} + p(z)$ for an appropriate constant C and polynomial p of degree $n-1$.

Proof. Assume $R > 1$; the case $R = 1$ follows by taking limits. Let $f \in A_n(R)$; then

$$f(z) = \sum_0^\infty a_j z^j$$

so that

$$f^{(n)}(z) = \sum_0^\infty \{(j+n)!/j!\}\, a_{j+n} z^j, \qquad |z| < R.$$

Let $c_k = k!/(k+n)!$, $k = 0, 1, \ldots$, and set

$$G(\theta) = R^{n-m-1} e^{i(m+1)\theta} \left\{ c_{m-n+1} + 2 \sum_{k=1}^\infty R^{-k} c_{m-n+1+k} \cos k\theta \right\}.$$

Then it is straightforward to check that

$$(1/2\pi) \int_0^{2\pi} f^{(n)}(Re^{i\theta})\, e^{in\theta} G(t-\theta)\, d\theta = f(e^{it}) + p_f(e^{it})$$

where p_f is a polynomial of degree m which depends on f. Hence

$$\begin{aligned} \inf_{q \in \pi_m} \| f - q \| &\leqslant \| f + p_f \| \\ &\leqslant \| f^{(n)}(Re^{i\theta}) \|_\infty \, \| G \|_1 . \end{aligned}$$

However, the term in the brackets in the formula for G is nonnegative since $\{c_k\}$ is nonnegative with nonnegative first and second differences. Hence, $\sup_{f \in A_n(R)} \inf_{q \in \pi_m} \| f - q \| \leqslant R^{n-m-1} c_{m-n+1}$.

On the other hand, the function $F(z) = R^{n-m-1} c_{m-n+1} z^{m+1}$ lies in $A_n(R)$ and is at distance $R^{n-m+1} c_{m-n+1}$ from π_m. This proves the first part of the theorem.

If $f \in A_n(R)$ has maximum distance $d = R^{n-m-1} c_{m-n+1}$ from π_m, then

$$\begin{aligned} d \leqslant \| f + p_f \|_\infty &= | f(e^{it_0}) + p_f(e^{it_0}) | \\ &= \left| \frac{1}{2\pi} \int_0^{2\pi} f^{(n)}(Re^{i\theta})\, e^{in\theta} G(t_0 - \theta)\, d\theta \right| \\ &\leqslant \frac{1}{2\pi} \int_0^{2\pi} | f^{(n)}(Re^{i\theta})\, e^{in\theta} G(t_0 - \theta) |\, d\theta \\ &\leqslant \| f^{(n)}(Re^{i\theta}) \|_\infty \, \| G(t_0 - \theta) \|_1 \\ &\leqslant d. \end{aligned}$$

Hence, $f^{(n)}(\mathrm{Re}^{i\theta})\, e^{in\theta} G(t_0 - \theta)$ has constant argument and $| f^{(n)}(Re^{i\theta}) | = 1$ where $G(t_0 - \theta) \neq 0$. Since G cannot vanish on a set of positive measure, we find that

$$f^{(n)}(Re^{i\theta}) = \lambda e^{i(m-n+1)\theta}$$

where $|\lambda| = 1$. Hence, $f(z) = \lambda R^{n-m-1} c_{m-n+1} z^{m+1} + p(z)$ where $p \in \pi_{n-1}$.

5. Other Moduli of Continuity

Theorem 1.1 may be viewed as establishing the distance from T_m to the set of those functions f for which $f^{(n-1)}$ satisfies the Lipschitz condition $|f^{(n-1)}(x) - f^{(n-1)}(y)| \leqslant |x - y|$ for all x, y in $[-\pi, \pi]$. With that in mind we can ask for the distance from T_m to the set of functions f for which $f^{(n-1)}$ has some other modulus of continuity. We look at this question when the modulus of continuity ω is *concave*; we do not get the precise distance (for this see [8, Section 5]) but do establish upper and lower bounds which are not too far apart. The techniques are elementary but we do make use of a theorem of Korneičuk on the precise value of the constant for the lowest-order case. For simplicity in exposition we impose the modulus of continuity condition on the nth derivative.

Definition. Let $\omega(h)$ be a continuous, concave positive, increasing function on $[0,2\pi]$ with $\omega(0) = 0$ and $\omega(h_1 + h_2) \leqslant \omega(h_1) + \omega(h_2)$. We define $\Lambda_{n\omega}^*$ to be all those functions f for which $f, f', \ldots, f^{(n)}$ are continuous 2π-periodic functions on $[-\pi, \pi]$ and for which $|f^{(n)}(x + h) - f^{(n)}(x)| \leqslant \omega(h)$, $x \in [-\pi, \pi]$, $0 \leqslant h \leqslant 2\pi$.

Theorem 5.1. *Let ω be a concave modulus of continuity and let $\beta_{nm}(\omega) = \sup\{E_m^*(f): f \in \Lambda_{n\omega}^*\}$. Then*

$$(1/\pi)\,\omega(\pi/m+1)\,K_{n+1}(m+1)^{-n} \leqslant \beta_{nm}(\omega) \leqslant \tfrac{1}{2}\omega(\pi/m+1)\,K_n(m+1)^{-n} \tag{5.1}$$

where K_n, K_{n+1} are the constants given by (0.2) *for n and $n+1$, respectively.*

Proof. The theorem of Korneičuk [6] asserts

$$\beta_{0m}(\omega) = \sup\{E_m^*(f) : f \in \Lambda_{0\omega}^*\} = \tfrac{1}{2}\omega(\pi/m+1). \tag{5.2}$$

Also see [9, p. 123]. First we establish the upper bound in (5.1). Let S be the best L^1 approximation to D_n from T_m, let λ be any periodic measure which is orthogonal to T_m and has total variation 1 or less, let $G \in \Lambda_{n\omega}^*$ with $G^{(n)} = g$. For an appropriate choice of λ and G we have

$$\begin{aligned}
\beta_{nm}(\omega) &= \sup\{E_m^*(f): f \in \Lambda_{n\omega}^*\} \\
&= \int_{-\pi}^{\pi} G(t)\, d\lambda(t) \\
&= \int_{-\pi}^{\pi} \left(\int_{-\pi}^{\pi} g(t-x)(D_n(x) - S(x))\, dx\right) d\lambda(t) \\
&= \int_{-\pi}^{\pi} \left(\int_{-\pi}^{\pi} g(t-x)\, d\lambda(t)\right)(D_n(x) - S(x))\, dx.
\end{aligned}$$

Now the function

$$u(x) = \int_{-\pi}^{\pi} g(t - x)\, d\lambda(t)$$

is no larger than $\frac{1}{2}\omega(\pi/m + 1)$ by (5.2). Hence, applying Hölder's inequality we find that

$$\beta_{nm}(\omega) \leqslant \tfrac{1}{2}\omega(\pi/m + 1)\, K_n(m + 1)^{-n}$$

since $\| D_n - S \|_1 = K_n(m + 1)^{-n}$. This gives the upper bound.

To establish the lower bound we choose a particular g and a particular λ. First assume n is even. Let g be the even function of period $2\pi/m + 1$ for which

$$\begin{aligned} g(x) &= \tfrac{1}{2}\omega(\pi/m + 1 - 2x) && 0 \leqslant x \leqslant \pi/2(m + 1), \\ &= -\tfrac{1}{2}\omega(2x - \pi/m + 1) && \pi/2(m + 1) \leqslant x \leqslant \pi/m + 1, \end{aligned}$$

g is in $\Lambda_{0\omega}^{*}$; see [10, p. 45]. Let λ be the measure with $2m + 3$ point masses at the points $-\pi + k\pi/m + 1$, $k = 0, 1,\ldots, 2m + 2$, with weights $1/4m + 4$, $-1/2m + 2$, $1/2m + 2,\ldots, -1/2m + 2$, $1/4m + 4$, respectively. Then λ is orthogonal to $\sin kx$ for all k and to $\cos kx$ if k is not a multiple of $m + 1$; the integral of λ against $\cos p(m + 1)x$ is 1 for all integers p. Since g is $2\pi/m + 1$ periodic and even,

$$\int_{-\pi}^{\pi} g(t - x)\, d\lambda(t) = g(x).$$

Hence, if $G^{(n)} = g$, then $G \in \Lambda_{n\omega}^{*}$ and

$$\begin{aligned} \beta_{nm}(\omega) &\geqslant \int_{-\pi}^{\pi} G(t)\, d\lambda(t) \\ &= \int_{-\pi}^{\pi} g(x)\, D_n(x)\, dx \\ &= (m + 1)^{-n} \int_{-\pi}^{\pi} g(x)\, D_n((m + 1)\, x)\, dx \end{aligned}$$

since g is $2\pi/m + 1$ periodic. Continuing

$$\begin{aligned} \beta_{nm}(\omega) &\geqslant (m + 1)^{-n} \int_{-\pi}^{\pi} g(x/m + 1)\, D_n(x)\, dx \\ &= 2(m + 1)^{-n} \int_{0}^{\pi} g(x/m + 1)\, D_n(x)\, dx. \end{aligned}$$

Now

$$\begin{aligned} &\int_{0}^{\pi} g(x/m + 1)\, D_n(x)\, dx \\ &\quad = \int_{0}^{\pi/2} g(x/m + 1)[D_n(x) - D_n(\pi - x)]\, dx \\ &\quad = (2/\pi) \int_{0}^{\pi/2} g(x/m + 1)\left(\sum_{0}^{\infty} (2k + 1)^{-n} \cos(2k + 1)\, x\right) dx. \end{aligned}$$

Since the sum in the last integral is nonnegative on $[0, \pi/2]$ we may use the inequality which is valid for $0 \leqslant x \leqslant \pi/2$ because of the concavity of ω

$$\begin{aligned} 2g(x/m+1) &= \omega(\pi/m+1 - 2x/m+1) \\ &= \omega((\pi/m+1)(1 - 2x/\pi)) \\ &\geqslant (1 - 2x/\pi)\,\omega(\pi/m+1). \end{aligned}$$

Hence,

$$\begin{aligned} \beta_{nm}(\omega) &\geqslant \frac{2}{\pi}\,\omega(\pi/m+1) \sum_0^\infty (2k+1)^{-n} \int_0^{\pi/2} (1 - 2x/\pi)\cos(2k+1)x\,dx \\ &= \frac{4}{\pi^2}\,\omega(\pi/m+1) \sum_0^\infty (2k+1)^{-n-2} \\ &= \frac{1}{\pi}\,\omega(\pi/m+1)\,K_{n+1}. \end{aligned}$$

A similar computation gives the same lower bound when n is odd.

COROLLARY 5.2. *Let* $\omega(h) = h^\alpha$, $0 < \alpha \leqslant 1$. *Then*

$$\frac{1}{\pi}\,\pi^\alpha K_{n+1}(m+1)^{-n-\alpha} \leqslant \sup\{E_m{}^*(f) : f^{(n)} \in \mathrm{Lip}_1(\alpha)\} \leqslant \tfrac{1}{2}\pi^\alpha K_n(m+1)^{-n-\alpha}.$$

Comments. (1) When $\alpha = 1$ the left inequality above is actually an equality.

(2) For $n = 1$ the corollary yields the estimates

$$\tfrac{1}{8}\pi^{1+\alpha}(m+1)^{-1-\alpha} \leqslant \beta_{1m}(h^\alpha) \leqslant \tfrac{1}{4}\pi^{1+\alpha}(m+1)^{-1-\alpha}$$

and for $n = 2$ the estimates

$$(1/24)\,\pi^{2+\alpha}(m+1)^{-2-\alpha} \leqslant \beta_{2m}(h^\alpha) \leqslant (1/16)\,\pi^{2+\alpha}(m+1)^{-2-\alpha}.$$

These compare with the exact values

$$\tfrac{1}{4}(1+\alpha)^{-1}\,\pi^{1+\alpha}(m+1)^{-1-\alpha}$$

and

$$\tfrac{1}{8}(2+\alpha)^{-1}\,\pi^{2+\alpha}(m+1)^{-2-\alpha}$$

found by Korneičuk; [5, 6], respectively.

REFERENCES

1. N. I. ACHIESER AND M. G. KREIN, On the best approximation of periodic functions, *Dokl. Akad. Nauk SSSR* **15** (1937), 107–112.
2. S. N. BERNSTEIN, "Collected Works," Vol. II, Akad. Nauk SSSR, Moscow, 1954.

3. W. A. MARKOV, "Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle," Gauthier–Villars, Paris.
4. H. B. CURRY AND I. J. SCHOENBERG, On Polya frequency functions IV; *J. Analyse Math.* **17** (1966), 71–108.
5. J. FAVARD, Sur les meilleures procédés d'approximation de certaines classes des fonctions par des polynômes trigonométriques, *Bull. Sci. Math.* **61** (1937), 109–224; 243–256.
6. N. P. KORNEIČUK, The best uniform approximation of differentiable functions, *Dokl. Akad. Nauk* **141** (1961), 304–307.
7. N. P. KORNEIČUK, The exact value of the best approximation and of widths of some classes of functions, *Dokl. Akad. Nauk* **150** (1963), 1218–1220.
8. N. P. KORNEIČUK, On methods of investigating extremal problems in the theory of best approximation, *Russian Math. Surveys* **29**(3) (1974), 7–43.
9. M. G. KREIN, On the approximation of continuous differentiable functions on the whole real axis, *Dokl. Akad. Nauk* **18** (1938).
10. G. G. LORENTZ, "Approximation of Functions," Holt, Rinehart & Winston, New York, 1966.
11. A. F. TIMAN, "Theory of Approximation of Functions of a Real Variable," translated by J. Berry, Pergamon Press, Oxford, England, 1963.