JOURNAL OF APPROXIMATION THEORY 26, 1–13 (1979)

Best Approximation by Piecewise Polynomials With Variable Knots and Degrees

W. DAHMEN AND K. SCHERER

Institut fur Angewandte Mathematik, Universität Bonn, D5300 Bonn, West Germany

Communicated by Carl de Boor

Received November 15, 1977

1. PRELIMINARIES

Approximation by piecewise polynomials of fixed degree and free knots, where only their total number is prescribed in advance, has been studied by several authors (e.g., [4], [3], [1]) as to quantitative behaviour as the number of knots tends to infinity. This kind of approximation can be generalized by varying also the degrees from knot to knot and only fixing the total number of all parameters. Our attention to this was drawn by H. G. Burchard who suggested the study of this more complex problem, first for very smooth functions. In particular, he raised the question whether for analytic functions the optimal approximation would be given by pure polynomial approximation. In this paper, we give, essentially, a positive answer to this.

We need the following notation:

$$E_n(f; [a, b]) = \inf_{p \in \Pi_{n-1}} \| f - p \|, \; E_n(f) =: E_n(f; [-1, 1]),$$

where $\| \;\; \|$ is the sup-norm and Π_{n-1} the space of polynomials of degree $\leqslant n - 1$. To define spaces of piecewise polynomials, we consider pairs (Δ, Z) where Δ is a partition of $[-1, 1]$ into subintervals $\{I_j\}_{j=1}^{k}$ and $Z = (m_1, ..., m_k)$ a corresponding vector in $Z_+{}^k$, and set

$$P(\Delta, Z) = \{f: f_{|I_j} \in \Pi_{m_j - 1}\}. \tag{1}$$

Since we want to fix only the total number of parameters, we introduce

$$P(k, Z) = \bigcup P(\Delta, Z). \tag{2}$$

The union is over all partitions Δ of $[-1, 1]$ into k subintervals and for a fixed $Z \in$ some $Z(k, m)$ where, for $k \leqslant m$,

$$Z(k, m) = \left\{ Z \in Z_+{}^k : \sum_{i=1}^{k} m_i \leqslant m \right\}.$$

0021-9045/79/050001-13$02.00/0

Finally we set

$$P^m = \bigcup_{k \leqslant m} \bigcup_{Z \in Z(k,m)} P(k, Z). \tag{3}$$

Our aim is to investigate dist (f, P^m).

LEMMA 1. (a) *There exists an element* $s^* \in P^m$ *such that*

$$\|f - s^*\| = \text{dist}\,(f, P^m)$$

The corresponding pair (Δ, Z) *is called an optimal partition.*

(b) *Every optimal partition is balanced, in particular*

$$E_{m_i}(f; I_i) = \text{dist}\,(f, P^m)$$

for the segments I_i *and degrees* m_i *of the optimal pair* (Δ, Z).

Proof. Denote by $y_1, \ldots, y_k$ the right hand endpoints of a partition Δ into k subintervals and define, for $Z \in Z(k, m)$,

$$G(y_1, \ldots, y_k) = \inf_{s \in P(\Delta, Z)} \|f - s\|.$$

(Coalescence of some of the y_i's is admitted and to be interpreted in the sense that the corresponding m_i do not appear). G is a continuous function on the compactum $[-1, 1]^k$ because $E_n(f; [a, b])$ is a continuous function of a, b. Hence G takes its minimum which means that there is $\tilde{s} \in P^m$ such that

$$\|f - \tilde{s}\| = \inf_{s \in P(k,Z)} \|f - s\|.$$

Since the union in (3) is taken over a finite set, assertion (a) follows. Part (b) follows from the continuity of $E_n(f; [a, b])$ in a, b.

We remark that balancedness of a pair (Δ, Z) is not sufficient for being optimal, because this is a property of the partition and the influence of Z has still to be taken care of. This is just why we concentrate in the following sections on classes of smooth functions to obtain more information about the possible Z. Further examples in Section 4 show that smoothness alone cannot characterize entirely the type of optimal partitions.

2. APPROXIMATION OF ENTIRE FUNCTIONS

As a first step concerning information about the optimal partition we have

LEMMA 2. *Let* $f \in C^\infty[-1, 1]$. *Then on each subinterval* A *of* $[-1, 1]$ *the restriction of a sequence of optimal partitions* $\{(\Delta_m, Z_m)\}_{m=1}^\infty$ *to* A *must* *contain segments of Δ_m for which the corresponding components of Z_m tend to ∞, as $m \to \infty$ (or f must coincide with some polynomial on A).*

Proof. Suppose the assertion were not true. Then there exists k such that all components of Z_m corresponding to a segment of Δ_m having a point in common with A remain bounded by k as $m \to \infty$.

Now by classical approximation theorems on pure polynomial approximation it is known that $E_m(f)$ is smaller than $O(m^{-\alpha})$ for each $\alpha > 0$. This would imply that we have for A a sequence of partitions consisting of at most m knots and corresponding piecewise polynomials S_m of maximal degree k such that $\|f - S_m\|_A = O(m^{-\alpha})$; for each $\alpha > 0$, in particular $m^k \|f - S_m\|_A \to 0$ for $m \to \infty$. But by a saturation result of Burchard–Hale [4] this implies that f is a polynomial of degree k on A.

Now we consider the following subclass of entire functions

$$G_0 = \left\{ f(z) : f(z) = \sum_{k=0}^{\infty} a_k z^k, \{a_k\}_{k=0}^{\infty} \in \mathscr{L} \right\},$$

where $\mathscr{L}$ is defined by

$$\mathscr{L} = \left\{ \{\alpha_n\}_{n=0}^{\infty} : \text{for each } \epsilon > 0 \text{ exists } n(\epsilon) \text{ such that} \right.$$

$$\left. \left| \frac{\alpha_{n+q}}{\alpha_n} \right| < \epsilon^q \text{ for all } n \geqslant n(\epsilon), q \in \mathbb{N} \right\}.$$

This means that we consider only entire functions with a regular decrease of the coefficients in the Taylor expansion. An example is $\alpha_n = e^{-n^2}$. Note that $\{a_n\} \in \mathscr{L}$ implies $a_n = a_n(f) \neq 0$ for all $n > n_0$ (otherwise f would be a polynomial). One may also assume $a_n \neq 0$ for all $n \in \mathbb{N}$ since otherwise one can consider $\tilde{f} = f + p_0$ where $p_0 \in \Pi_{n_0}$ is an appropriate polynomial such that $a_n(\tilde{f}) \neq 0$ for all $n \in \mathbb{N}$.

One has the following characterization.

LEMMA 3. *A sequence $\{\alpha_n\}$ belongs to $\mathscr{L}$ iff*

$$|\alpha_n| = e^{-t(n)n} \tag{4}$$

where $t(n)$ increases to infinity as $n \to \infty$.

Proof. Clearly each $\{\alpha_n\}$ satisfying (4) belongs to $\mathscr{L}$. Now let $\alpha \in \mathscr{L}$, $\alpha_n = e^{-nt(n)}$ so that

$$\frac{\alpha_{n+1}}{\alpha_n} = \frac{e^{n[t(n)-t(n+1)]}}{e^{t(n+1)}} < \epsilon_n , \tag{5}$$

where $\epsilon_n \to 0$, $n \to \infty$. Assume now that there exists a sequence $\{n_k\}$ such that $t(n_k + 1) \leqslant t(n_k)$ and $t(n_k + 1) < M$ for $k \in \mathbb{N}$. This would lead to a contradiction to (5) since then for k large enough

$$\frac{e^{n_k[t(n_k)-t(n_k+1)]}}{e^{t(n_k+1)}} > \frac{1}{e^M}.$$

Moreover, using Lemma 3, (4) and the definition of $\mathscr{L}$ it is easy to prove

COROLLARY 1. *Let* $\{a_n\} \in \mathscr{L}$, $\epsilon > 0$. Then there exists $n_1(\epsilon)$ such that $|\alpha_n/\alpha_{n-q}| \leqslant \epsilon^q$ *for* $q \leqslant n$, $n \geqslant n_1$

If f is an analytic function

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} C_k T_k(x)$$

represented by its Taylor-series and its Fourier–Tchebycheff-series respectively the following well-known formula holds (cf. Bernstein [2, p. 116])

$$C_{k+1} = \frac{1}{2^k} \sum_{j=0}^{\infty} a_{k+1+2j} \binom{k+1+2j}{j} 2^{-2j}. \tag{6}$$

A further simple result in [2, p. 115] leads to

PROPOSITION 1. $f \in G_0$ *implies*

$$C_{n+1} \sim \frac{1}{2^n} a_{n+1} \sim E_{n+1}(f).$$

Here $p_n \sim q_n$ means that for some fixed constants A, B, we have for all n $|p_n| \leqslant A\,|q_n|$, $|q_n| \leqslant B\,|p_n|$.

We shall make use of this to estimate the best polynomial approximation of f on arbitrary sub-intervals $[\alpha - h, \alpha + h] \subset [-1, 1]$, using the transformation

$$g(\alpha, h; t) = g(t) = f(\alpha + ht),$$

so that $E_{n+1}(g) = E_{n+1}(f; [\alpha - h, \alpha + h])$.

One easily gets

$$g(t) = \sum_{k=0}^{\infty} a_k(\alpha + ht)^k = \sum_{k=0}^{\infty} (A_k(\alpha)\, h^k)\, t^k$$

with

$$A_k(\alpha) = \sum_{j=0}^{\infty} \left(a_{k+j} \binom{k+j}{j} \alpha^j \right).$$

For $\{a_k\} \in \mathscr{L}$ one can find, again following Bernstein's arguments, constants c_1, c_2 independent of α such that

$$c_1 a_k \leqslant A_k(\alpha) \leqslant c_2 a_k, \quad k \in \mathbb{N}, \tag{7}$$

i.e. $A_k(\alpha) \in \mathscr{L}$. By (6) one gets for the Tchebycheff-coefficients

$$C_{k+1}(\alpha, h) = \frac{1}{2^k} \sum_{j=0}^{\infty} A_{k+1+2j}(\alpha)\, h^{k+1+2j} \binom{k+1+2j}{j} \frac{1}{2^{2j}}$$
$$= \frac{h^{k+1}}{2^k} A_{k+1}(\alpha)\left[1 + \frac{A_{k+1+2}(\alpha)}{A_{k+1}(\alpha)} \frac{h^2}{2} \binom{k+3}{1} + \cdots\right]$$

so that we can again conclude (as in (7))

$$C_{k+1}(\alpha, h) \sim \frac{h^{k+1}}{2^k} A_{k+1}(\alpha),\ C_{k+1}(\alpha, h) \in \mathscr{L}$$

and we get by Proposition 1,

PROPOSITION 2.

$$C_{n+1}(\alpha, h) \sim \frac{h^{n+1}}{2^n} A_{n+1}(\alpha) \sim \frac{h^{n+1}}{2^n} a_{n+1}$$
$$\sim E_{n+1}(f; [\alpha - h, \alpha + h]) \sim h^{n+1} E_{n+1}(f) \tag{8}$$

This leads to

THEOREM 1. *$f \in G_0$ implies for m large enough that the optimal partition (Δ_m, Z_m) in* dist (f, P^m) *is unique and realized by pure polynomial approximation, i.e. $\Delta_m = \{-1, 1\}$, $Z_m \in Z(1, m)$ and* dist $(f, P^m) = \|f - s^*\|$ *with $s^* \in \Pi_{m-1}$.*

Proof. Suppose an optimal partition has two neighboring subintervals $I_{h_i} = [\alpha_i - h_i, \alpha_i + h_i]$ with corresponding degrees n_i $(i = 1, 2)$ and $n_1 + n_2 = n$. By Lemma 1 we have $E_{n_1}(f; I_{h_1}) = E_{n_2}(f; I_{h_2})$. We now assume

$$\frac{E_n(f; I_{h_1} \cup I_{h_2})}{E_{n_1}(f; I_{h_1})} = \frac{E_n(f; I_{h_1} \cup I_{h_2})}{E_{n_2}(f; I_{h_2})} \geqslant 1. \tag{9}$$

In view of Proposition 2 we may assume without loss of generality $h_1 + h_2 = 1$ and obtain from (9) by (8)

$$C \frac{2^{n_1} a_n}{2^n h_1^{n_1} a_{n_1}} \geqslant 1; \qquad C \frac{a_n}{2^{n_1} a_{n-n_1}(1-h_1)^{n-n_1}} \geqslant 1$$

and from this

$$h_1 < \left(C \frac{2^{n_1} a_n}{2^n a_{n_1}} \right)^{1/n_1} \equiv R_1(n, n_1), \tag{10a}$$

$$h_1 > 1 - \left(\frac{C a_n}{2^{n_1} a_{n-n_1}} \right)^{1/(n-n_1)} \equiv R_2(n, n_1). \tag{10b}$$

We show then that this leads to a contradiction by proving that

$$Q(n_1) = \frac{R_2(n, n_1)}{R_1(n, n_1)} > 1 \tag{11}$$

for n sufficiently large and all n_1 with $n_2 = n - n_1 < n_1 \leqslant n$.

Now from the definition of the class G_0 and from Corollary 1 it follows that

$$Q(n_1) > \frac{1 - (C^{1/n_1}\epsilon/2)^{n_1/(n-n_1)}}{(C^{1/(n-n_1)}\epsilon/2)^{(n-n_1)/n_1}} \geqslant \frac{1 - \left(\frac{C\epsilon}{2}\right)^{n_1/(n-n_1)}}{\left(\frac{C\epsilon}{2}\right)^{(n-n_1)/n_1}}$$

for any $\epsilon > 0$ provided $n \geqslant n(\epsilon)$, or

$$Q(n_1) > (1 - a^x)\, a^{-1/x} \equiv D(x),\ x = n_1/(n - n_1),\ a = \frac{C\epsilon}{2},$$

for any $0 < a < 1$ provided n large enough.

But it is clear that $\lim_{x\to\infty} D(x) = 1$ and for a small enough

$$D'(x) = [x^{-2} a^{-1/x}(1 - a^x) - a^x a^{-1/x}] \log a < 0$$

for all x, $1 \leqslant x \leqslant n$ (or $n_2 < n_1 \leqslant n$) since the term in brackets is positive for such a. Hence (11) must hold.

We remark that the crucial point in the proof is knowledge of the exact relation between $E_n(f)$ and $E_n(f, I)$ for some subinterval $I \subset [-1, 1]$, furnished by the properties of $\mathscr{L}$ and G_0. But next we show that the smoothness of f can be reduced further while still getting an analogous result.

3. Approximation of Analytic Functions

We consider functions f which are analytic in some region D of the complex plane containing the intervall $[-1, 1]$. $\rho(x)$ denotes the radius of convergence when

$$f(z) = \sum_{n=0}^{\infty} a_n(x)(z-x)^n$$

is evaluated at a point $x \in D$.

We define the functions

$$M_k(x) = \sup_{j \geqslant k} \frac{\rho(x)^j \mid a_j(x)|}{\rho(x)^k \mid a_k(x)|}$$

and

$$M(x) = \sup_k M_k(x).$$

Let

$$A_0 = \{f(z);\ M(x) < \infty;\ x \in [-1, 1]\}.$$

We need

Proposition 3. *Let $f \in A_0$. For any z in the complex disk $D(x)$ with center x and radius $\rho(x)$ there holds (for $a_k(x) \neq 0$)*

$$\left| \frac{a_k(z)}{a_k(x)} \right| \leqslant C \frac{\rho(x)^{k+1}}{[\rho(x) - |z-x|]^{k+1}}$$

with a constant independent of z and k uniformly in $x \in I \subset [-1, 1]$.

Proof. Since $M_k(x)$ and $M(x)$ are lower semi-continuous (cf. [5, p. 39], the set

$$V_n = \{x \in I\colon M(x) > n\}$$

is open for each $n \in \mathbb{N}$. Then at least one V_n cannot be dense in I since otherwise (by Baire's theorem) the intersection of all V_n would contain one point x' in I for which $M(x') = \infty$ contradicting the assumption. Hence there exists a closed (open) subinterval of I the points of which do not belong to a certain V_n, i.e., for which $M(x)$ is uniformly bounded.

Repeating the process we find an open covering by such intervals of I. Since I is compact we can find a finite subcovering so that $M(x)$ is uniformly bounded on I. This gives

$$|a_j(x)| \leqslant C\rho(x)^{k-j} |a_k(x)| \tag{12}$$

uniformly in $j \geqslant k$ and $x \in I$.

Now for $z \in D(x)$:

$$\frac{a_k(z)}{a_k(x)} = \frac{1}{k!} \sum_{j=k}^{\infty} \frac{j!}{(j-k)!} \frac{a_j(x)}{a_k(x)} (z-x)^{j-k}$$

and by (12)

$$\left| \frac{a_k(z)}{a_k(x)} \right| \leqslant C \frac{1}{k!} \sum_{j=k}^{\infty} \frac{j!}{(j-k)!} \left(\frac{|z-x|}{\rho(x)} \right)^{j-k}.$$

Since

$$\frac{1}{k!} \sum_{j=k}^{\infty} \frac{j!}{(j-k)!} a^{j-k} = \frac{1}{k!} \frac{d^k}{da^k} \sum_{j=0}^{\infty} a^j$$

$$= \frac{1}{k!} \frac{d^k}{da^k} \left(\frac{1}{1-a} \right) = \frac{1}{(1-a)^{k+1}}$$

it follows

$$\left| \frac{a_k(z)}{a_k(x)} \right| \leqslant C \frac{1}{\left(1 - \dfrac{|z-x|}{\rho(x)}\right)^{k+1}} = C \frac{\rho(x)^{k+1}}{(\rho(x) - |z-x|)^{k+1}}.$$

Remark. $M(x) < \infty$ is equivalent with $|a_j(x)/a_k(x)| \leqslant C(x)\rho(x)^{k-j}$, $j \geqslant k, \forall k \in \mathbb{N}$ and (12) just means $C(x) \leqslant M$ on I. There are evident examples of analytic functions which have this property as well as those which do not.

The set A_0 corresponds to the above introduced subset G_0 of entire functions since it picks those analytic functions which have a regular decrease in their power series expansion.

THEOREM 2. *Let* $f \in A_0$ *and* $\inf_{x \in [-1,1]} \rho(x) > 0$. Then the number of subintervals of the optimal partitions (Δ_m, Z_m) remains bounded as $m \to \infty$, i.e., one has essentially pure polynomial approximation.

Proof. We assume that $[a, c]$ and $[c, b]$ are two intervals of a balanced optimal partition with degrees n_1 and n_2, respectively, and $n = n_1 + n_2$. We suppose

$$E_n = E_n(f; [a, b]) > E_{n_1} = E_{n_2}, \tag{13}$$

where $E_{n_1} = E_{n_1}(f; [a, c])$ and $E_{n_2} = E_{n_2}(f; [c, b])$. Basic for the following is now the well known equation

$$E_n = 2[(b-a)/4]^n \, |a_n(z)|$$

for some $z \in [a, b]$ where we recall $a_n(z) = f^{(n)}(z)/n!$. With similar equations for E_{n_1} and E_{n_2} (13) implies

$$[(b-a)/4]^n \mid a_n(z) \mid > [(c-a)/4]^{n_1} \mid a_{n_1}(x_1) \mid = [(b-c)/4]^{n_2} \mid a_{n_2}(x_2) \mid$$

Proposition 3 and (13) then yield from this ($\mid b - a \mid$ small enough in comparison to $\inf_x \rho(x)$)

$$\left[\frac{b-a}{c-a}\right]^n \left[\frac{4}{c-a}\right]^{n_1-n} > \frac{1}{C} \rho(x_1)^{n-n_1} \frac{[\rho(x_1) - \mid z - x_1 \mid]^{n+1}}{\rho(x_1)^{n+1}} \tag{14a}$$

and

$$\left[\frac{b-a}{b-c}\right]^n \left[\frac{b-c}{4}\right]^{n_1} > \frac{1}{C} \rho(x_2)^{n_1} \frac{[\rho(x_2) - \mid z - x_2 \mid]^{n+1}}{\rho(x_2)^{n+1}}. \tag{14b}$$

For brevity we introduce the notation $\gamma = n_1/n$ and ($i = 1, 2$)

$$\rho_i = \rho(x_i),\ e_{n,i} = \left[1 - \frac{\mid z - x_i \mid}{\rho_i}\right]^{(n+1)/n} C^{1/n}.$$

Then taking the n-th root in (14, a, b) we obtain

$$\left[\frac{b-a}{c-a}\right]\left[\frac{4}{c-a}\right]^{\gamma-1} > \rho_1^{1-\gamma} e_{n,1}$$

$$\left[\frac{b-a}{b-c}\right]\left[\frac{b-c}{4}\right]^{\gamma} > \rho_2^{\gamma} e_{n,2}.$$

With the abbreviations $\delta = (b-a)/(c-a)$ and $X = (b-a)/4$ these inequalities are:

$$\delta^\gamma X^{1-\gamma} > \rho_1^{1-\gamma} e_{n,1},$$

$$[\delta/(\delta-1)]^{1-\gamma} X^\gamma > \rho_2^\gamma e_{n,2},$$

and with $v = \rho_1/X$, $w = \rho_2/X$, $x = \gamma/(1-\gamma)$ we have:

$$\delta > v^{1/x} e_{n,1}^{1/\gamma}, \tag{15a}$$

$$\delta < \frac{w^x e_{n,2}^{1/(1-\gamma)}}{w^x e_{n,2}^{1/(1-\gamma)} - 1}. \tag{15b}$$

Now, if $v \geqslant w$, v can be replaced by w in (15a). Since $w^x/(w^x - 1)$ decreases for fixed x and increasing w, (15b) remains true with w replaced by v in case $v \leqslant w$. By analogous arguments with respect to $e_{n,1}$ and $e_{n,2}$ we get the following couple of inequalities

$$\delta > v^{(1-\gamma)/\gamma} e_n^{1/\gamma} = R_1(\gamma, v),$$
$$\delta < \frac{v^{\gamma/(1-\gamma)} e_n^{1/(1-\gamma)}}{v^{\gamma/(1-\gamma)} e_n^{1/(1-\gamma)} - 1} = R_2(\gamma, v), \tag{16}$$

where

$$v = \min\left(\frac{4\rho(x_1)}{b-a}, \frac{4\rho(x_2)}{b-a}\right),$$
$$e_n = \min(e_{n,1}; e_{n,2}).$$

Inequalities (16) are a consequence of our initial assumption (12). So we want to show a contradiction to (16) by proving that

$$R_2(\gamma, v) < R_1(\gamma, v) \tag{17}$$

for n large enough all $0 \leqslant \gamma \leqslant 1$ (actually $0 \leqslant \gamma \leqslant 1/2$ suffices since without loss of generality $n_1 \leqslant n_2$), and $v \geqslant v_0$. The latter condition is satisfied if $\inf_{x\in[-1,1]} 4\rho(x)/(b - a) \geqslant v_0$. This means that for $b - a$ small enough (depending on $\inf \rho(x)$ and the total number n of parameters) pure polynomial approximation on $[a, b]$ is optimal which is the assertion of the theorem.

In order to prove (17) we consider the function

$$H(\gamma, v) = \frac{R_2(\gamma, v)}{R_1(\gamma, v)} = \frac{v^{\gamma/(1-\gamma)} e_n^{1/(1-\gamma)}}{v^{(1-\gamma)/\gamma} e_n^{1/\gamma} [v^{\gamma/(1-\gamma)} e_n^{1/(1-\gamma)} - 1]}$$
$$= \frac{(ve_n)^{(2\gamma-1)/\gamma(1-\gamma)}}{(ve_n)^{1/(1-\gamma)}\, v^{-1} - 1}.$$

One easily verifies for v sufficiently large (such that $ve_n > 1$) $\lim_{\gamma\to 0} H(\gamma, v) = 0$, or more specifically

$$H(\gamma, v) < \tfrac{1}{2}; \qquad 0 < \gamma \leqslant \gamma_0, \qquad v \geqslant v_0. \tag{18}$$

Furthermore we choose v_0 so large that

$$H\left(\frac{1}{2}, v\right) = \frac{1}{ve_n - 1} < 1, \; v \geqslant v_0. \tag{19}$$

Finally elementary differentiation yields

$$\frac{\partial}{\partial\gamma} H(\gamma, v) = \frac{\log(ve_n)(ve_n)^{(2\gamma-1)/\gamma(1-\gamma)}\,[\gamma^{-2}[(ve_n)^{1/1-\gamma}\,v^{-1} - 1] - (1-\gamma)^{-2}]}{|(ve_n)^{1/(1-\gamma)}\,v^{-1} - 1\,|^2}$$

Now for v sufficiently large

$$v^{\gamma/(1-\gamma)} e_n^{1/(1-\gamma)} - 1 - \gamma^2/(1-\gamma)^2 > 0,\ \gamma \in [\gamma_0\,,\tfrac{1}{2}]$$

so that $(\partial/\partial\gamma)H(\gamma, v) > 0$ on $[\gamma_0\,, \frac{1}{2}]$ and $v \geqslant v_0$ provided v_0 sufficiently large. This together with (18) and (19) establishes (17).

4. Final Considerations

The assumptions in the definitions of the subclasses G_0 and A_0 of entire and analytic functions, respectively, allowed us to estimate $E_n(f; I_1)$ by $E_n(f; I_2)$ for intervals $I_2 \subset I_1$. It is not known how far these for the proofs of Theorems 1 and 2 essentially needed hypotheses can be relaxed.

Let us add some further remarks concerning the question when polynomial approximation is the unique optimal approximation on some interval $[a, b]$. First, it is clear that smoothness alone cannot be sufficient, e.g., the condition

$$E_n(f;\ [a, b]) < E_{n-1}(f;\ [a, b])$$

is necessary. A somewhat stronger motivation for considering only subsets of classes of smooth functions is given by the following simple proposition which provides a sort of converse to the above theorems for the subset

$$V = \left\{ f \in C[-1, 1] : E_n(f; I_1) \sim E_n(f; I_2) \left| \frac{I_1}{I_2} \right|^n \text{ for } I_1 \subset I_2 \subset [-1, 1] \right\}.$$

Proposition 4. *Let $f \in V$. If there is some interval $[a, b] \subseteq [-1, 1]$ such that (for m sufficiently large) polynomial approximation is best (to achieve* dist (f, P^m)), *then f is analytic on $[a, b]$.*

Proof. We compare $E_n(f;\ [a, b]) = \psi(n)$ with $E_{n_1}(f; I_1)$ where we take $n_1 = [n/2]$ and $I_1 = [a, (a + b)/2]$. By definition of V, $E_{n_1}(f; I_1) < CE_{n_1}(f;\ [a, b])\, 2^{-n_1}$, so that

$$E_{n_1}(f; I_1) \leqslant C \frac{\psi(n/2)}{(\sqrt{2})^n}\,.$$

Now, if f were not analytic on some region containing $[a, b]$, then, for any $\beta > 1$, $\lim_{n\to\infty} \beta^n \psi(n) = \infty$. Hence, for sufficiently large n

$$\begin{aligned} E_{n_1}(f; I_1) &\leqslant C\psi(n/2)(\beta/\sqrt{2})^n \psi(n) \\ &\leqslant C \|f\|_\infty (\beta/\sqrt{2})^n E_n(f; [a, b]) \\ &< E_n(f; [a, b]), \end{aligned}$$

which is a contradiction to the assumption.

The assumptions on f given in the definition of V are similar to those of Theorems 1 and 2. So the question arises whether for the class V polynomial approximation is best if and only if f is analytic. We were not able to answer this. Instead we conclude with an example showing that without additional assumptions the type or the uniqueness of the optimal partitions for dist (f, P^m) need not be related to the smoothness of f at all.

The example may be described as follows:

Let $g(0) > 1$ be odd and define $g(k + 1) = (g(k))^3$ for $k = 0, 1, 2,...$. Let $\{\epsilon_k\}$ be a monotone sequence which tends to zero and $\Delta_k = \epsilon_{k-1} - \epsilon_k$. Then the function

$$f(x) := \sum_{k=1}^{\infty} \Delta_k T_{g(k)}(x),$$

where $T_k(x) = \cos k \arccos x$ is well defined and belongs to $C[-1, 1]$.

Furthermore for $g(k) \leqslant n < g(k + 1)$

$$E_{n+1}(f) = \Big\| \sum_{j=k+1}^{\infty} \Delta_j T_{g(j)} \Big\|_\infty = \epsilon_k . \tag{20}$$

Now, $R_k(x) = \sum_{j=k+1}^{\infty} \Delta_j T_{g(j)}(x)$ has $g(k + 1) + 1$ alternating extrema. For the pure polynomial approximation $E_{g(k)+1}(f)$ $g(k) + 1$ parameters are needed. So, in order to realize this error or even a better one by real piecewise polynomials in $P_{g(k)+1}$, there are at most $g(k) + 1$ intervals I_j. The definition of $g(k)$ implies that there is at least one interval I_j which contains more than $g(k) + 1$ extrema of $R_k(x)$, i.e., for $i \leqslant g(k)$, by (20),

$$\max_i E_i(f; I_j) \geqslant E_{g(k)+1}(f; I_j) = E_{g(k)+1}(f; [-1, 1]) = \epsilon_k ;$$

hence

$$\text{dist}\,(f, P^{g(k)+1}) = E_{g(k)+1}(f; [-1, 1]) = \epsilon_k , \qquad k \in \mathbb{N}. \tag{21}$$

Now, consider dist $(f, P^{g(k+1)})$. Let $\Delta_{g(k+1)}$ define intervals $I_1, \ldots, I_j$ so that the corresponding polynomials have degree $l_k^1, \ldots, l_k^j$, i.e.,

$$l_k^1 + \cdots + l_k^j + j = g(k+1)$$

Then dist $(f, P^{g(k+1)})$ can be smaller than $E_{g(k+1)}[f]$ only if I_1 contains at most $l_k^1 + 1$ extrema of $R_k(x)$. The same has to be true for $I_2, \ldots, I_{j-1}$. Hence I_j must contain at least

$$\begin{aligned} g(k+1) + 1 - l_k^1 - \cdots - l_k^{j-1} - (j-1) \\ = g(k+1) + 2 - (l_k^1 + \cdots + l_k^{j-1} + j) \end{aligned}$$

alternating extrema of $R_k(x)$. But since the corresponding polynomial has at most degree $l_k^j = g(k+1) - (l_k^1 + \cdots + l_k^{j-1} + j)$ the error $E_{g(k+1)}(f)$ is not improved and

$$\text{dist } (f, P^{g(k+1)}) = E_{g(k+1)}[f] = \epsilon_k .$$

This yields

$$\text{dist } (f, P^n) \geqslant \text{dist } (f, P^{g(k+1)}) = E_{g(k+1)}(f) = E_{g(k)+1}(f) = \epsilon_k$$

for all $g(k) < n \leqslant g(k+1)$, which together with (21) shows that pure polynomial approximation is optimal. This is independent from the choice of $\{\epsilon_k\}$ and hence from the smoothness of f. So f may be analytic or even entire for rapidly decreasing $\{\epsilon_k\}$, or in contrast for slowly decreasing $\{\epsilon_k\}$ even not differentiable. Note that $E_n(f)$ does not decrease so regularily as is affirmed by the hypothesis in Theorems 1, 2.

References

1. J. Bergh and J. Peetre, On the spaces $V_p(0 < p < \infty)$. Bolletino U. M. I. (4) **10** (1974), 632–648.
2. S. Bernstein and C. de la Vallée-Poussin, "L'approximation" (reprinted), New York, 1970.
3. Yu. Brudnyi, Spline approximation and functions of bounded variation, *Soviet Math. Dokl.* **15**, No. 2 (1974), 518–521.
4. H. G. Burchard and D. F. Hale, Piecewise polynomial approximation on optimal meshes, *J. Approximation Theory* **14** (1975), 128–147.
5. W. Rudin, "Real and Complex Analysis," 2nd ed., McGraw-Hill, New York, 1974.