JOURNAL OF APPROXIMATION THEORY 36, 294–303 (1982)

A Unified Approach to Differential Characterizations of Local Best Approximations for Exponential Sums and Splines

LUDWIG J. CROMME

Universität Göttingen, Institut für Numerische und Angewandte Mathematik, Lotzestr. 16–18, D-3400 Göttingen, West Germany

Communicated by Carl de Boor

Received September 15, 1981; revised March 4, 1982

1. INTRODUCTION

Hobby and Rice [8] were apparently the first to notice that the concept of *γ-polynomials* reveals interesting common features of such important problems as approximation by exponential sums or by splines with variable knots. Later, de Boor [1] extended and simplified their results, and gave a unified approach to existence problems of best approximations for many nonlinear approximation problems. While the above-mentioned questions require only knowledge of the *topology* of the manifold of exponential sums (or splines), an intrinsic knowledge of the *differential structure* of these manifolds is required to derive necessary or sufficient conditions for a given γ-polynomial to be a (local) best approximation. γ-Polynomials constitute a manifold with boundary where the respective dimensions of the tangent cones vary [5]. The closed L_q unit-balls $(0 < q < 1)$ are well-known examples of manifolds with tangent cones of varying dimension.

To identify local best approximations the local structure of the manifold of γ-polynomials must be known. The necessary information is supplied by the regular and C^1-differentiable parametrization of γ-polynomials investigated in a preceding paper [6]. Our analysis therefore depends heavily on that article, and the reader is assumed to be familiar with the notation in [6].

Let $T \subset \mathbb{R}$ be an open interval ($T = \mathbb{R}$ or $T = (a, b) \subset \mathbb{R}$; this assumption will be relaxed later), $X \subset \mathbb{R}$, $|X| \geqslant 2n + 1$ and X compact. For a $\gamma \in C(T \times X)$ the γ-polynomials Γ_n^γ and $\Gamma_{n,k}^\gamma$ are defined as in [6].

For $1 \leqslant q \leqslant \infty$ let $\|\cdot\|_q$ denote the L_q-norm associated with a Lebesgue-measure on X. $g \in \Gamma_n^\gamma$ is a local best approximation to $f \in C(X)$ if there is a $\|\cdot\|_\infty$-neighborhood $U \subset C(X)$ of g with $\|f - h\|_q \geqslant \|f - g\|_q$ for all $h \in U \cap \Gamma_n^\gamma$.

0021-9045/82/120294–10$02.00/0

Otherwise, the general assumptions and the notation are the same as in [6].

2. Tangential Characterizations

Let γ be a normal $(k+1)$-kernel, $k \leqslant n$, $f \in C(X)$, $g \in \Gamma^{\gamma}_{n,k} \setminus \Gamma^{\gamma}_{n-1,k}$ fixed. g can be written as

$$g(\cdot)=\sum_{i=1}^{l}\sum_{j=1}^{m_i} a_{ij}\Delta_t^{j-1}(t_i,\ldots,t_i)\,\gamma(t,\cdot)$$

with the characteristic numbers $t_1 < t_2 < \cdots < t_l$ and this representation is unique. To g we assign the *tangent cone* T_g (the name tangent cone is justified by Theorem 1):

$$T_g=\left\{\sum_{i=1}^{l}\sum_{j=1}^{m_i^*} \alpha_{ij}\Delta_t^{j-1}(t_i,\ldots,t_i)\,\gamma(t,\cdot)\;\middle|\; \begin{array}{l}\alpha_{ij}\in\mathbb{R},\ m_i^* := \begin{cases}2, & m_i=1\\ m_i+2, & \text{otherwise}\end{cases}\\ \alpha_{im_i^*}\cdot a_{im_i}\geqslant 0,\ \text{for } m_i>1.\end{array}\right\}$$

The sign-restriction on $\alpha_{im_i^*}$ is omitted for $m_i = 1$. That is, for a γ-polynomial with all characteristic numbers distinct, T_g is a $2n$-dimensional linear space. We are now ready to state our main result.

Theorem 1. *For $q \in \mathbb{R}$, $1 \leqslant q \leqslant \infty$, and γ, g, f, T_g as above we have the following necessary and sufficient conditions for g to be a local best approximation to f:*

(a) Necessary condition. *Let g be a local best approximation to f from $\Gamma^{\gamma}_{n,k}$ with respect to $\|\cdot\|_q$.* Then *there is no $\vartheta \in T_g$ such that there exist positive constants c and ε with*

$$\|f-g-\lambda\vartheta\|_q \leqslant \|f-g\|_q-\lambda c \qquad \forall\lambda\in[0,\varepsilon].$$

(b) Characterization. *Statements* (i) *and* (ii) *are equivalent*:

(i) 0 *is a global strong unique best approximation to $f-g$ from T_g, that is, there exists a constant $c>0$ with*

$$\|f-g-\lambda\vartheta\|_q \geqslant \|f-g\|_q+\lambda c \qquad \forall\vartheta\in T_g,\ \|\vartheta\|_\infty=1,\ \lambda\geqslant 0.$$

(ii) *g is a local strong unique best approximation to f from $\Gamma^{\gamma}_{n,k}$. That is, there is a $\|\cdot\|_\infty$-neighborhood U of g and a constant $k>0$ with*

$$\|f-h\|_q \geqslant \|f-g\|_q + k\,\|h-g\|_q \qquad \forall h\in U\cap\Gamma^{\gamma}_{n,k}.$$

(c) Sufficient condition. *For $1 < q < \infty$ let g be a local best approximation to f with respect to those elements from $\Gamma_{n,k}^{\gamma}$ belonging to the same class of multiplicity as g (the same number l of distinct characteristic numbers with the same multiplicities $m_1,\dots,m_l$, respectively). For each $\vartheta \in T_g$ with $\alpha_{im_i^*} \neq 0$ for at least one $i \in \mathbb{N}$ with $m_i > 1$ let there be a constant $c_\vartheta > 0$ with*

$$\|f - g - \lambda\vartheta\|_q \geqslant \|f - g\|_q + \lambda c_\vartheta \qquad \forall \lambda \geqslant 0.$$

Then *g is a local best approximation to f from $\Gamma_{n,k}^{\gamma}$.*

Remarks. Part (a) of the theorem has applications for all $q \in \mathbb{R}$, $1 \leqslant q \leqslant \infty$. Part (b) is meant to cover the cases $q = 1$ and $q = \infty$. Otherwise $(1 < q < \infty)$, neither (i) nor (ii) can be fulfilled. Part (c) covers differentiable norms $(1 < q < \infty)$. The class of multiplicity of g is an open manifold. Therefore, the usual second order conditions can be used to determine whether g is a local best approximation with respect to that class. The assumption on elements ϑ with $\alpha_{im_i^*} \neq 0$ is weak since it can be shown to be "almost always" valid. Part (c) of the theorem cannot be extended to polyhedral norms.

Proof of Theorem 1. Our proof makes use of the differentiability properties of the parametrization of γ-polynomials discussed in [6]. We restrict ourselves to the special cases $k = n$, $l = 1$ to simplify formulas. The general case can be proven in analogy to the following proof. $n = 1$ is trivial, so let n be greater one.

With

$$M := \left\{ (\tau, \delta_1, \dots, \delta_{n-1})^T \in \mathbb{R}^n \,\middle|\, 0 \leqslant \delta_i \leqslant \left(\frac{i+2}{i}\right)^2 \delta_{i+1}, 1 \leqslant i \leqslant n-2 \right\},$$

$p: \mathbb{R}^n \times M \to \Gamma^{\gamma}$ is a parametrization of Γ_n^{γ}:

$$(a_1, \dots, a_n, \tau, \delta_1, \dots, \delta_n)^T \xrightarrow{p} \sum_{i=1}^{n} a_i \sum_{1 \leqslant j_1 < \cdots < j_i \leqslant n} \Delta_t^{i-1}(t_{j_1}, \dots, t_{j_i})\, \gamma(t, \cdot),$$

$$t_j := \tau + (j-1)\sqrt{\delta_{j-1}} - \sum_{k=j}^{n-1} \sqrt{\delta_k} \qquad (j = 1, 2, \dots, n);$$

p is regular at $p^{-1}(g)$; see [6]. Let us now prove part (a) of Theorem 1 by contradiction: ϑ can be written as

$$\vartheta = \sum_{i=1}^{n+2} \alpha_i \Delta_t^{i-1}(\tau, \dots, \tau)\, \gamma(t, \cdot)$$

if g has the representation

$$g=\sum_{i=1}^{n} a_i \binom{n}{i} \Delta_t^{i-1}(\tau,\dots,\tau)\,\gamma(t,\cdot).$$

The partial derivatives of p at $a := (a_1,\dots,a_n,\tau,0,\dots,0)^T \in \mathbb{R}^{2n}$ are given by

$$\frac{\partial}{\partial a_i} p = \binom{n}{i} \Delta_t^{i-1}(\tau,\dots,\tau)\,\gamma(t,\cdot), \qquad i=1,2,\dots,n,$$

$$\frac{\partial}{\partial \tau} p = \sum_{i=1}^{n} a_i \binom{n}{i} \cdot i\,\Delta_t^{i}(\tau,\dots,\tau)\,\gamma(t,\cdot),$$

$$\frac{\partial}{\partial \delta_i} p = \sum_{j=2}^{n+2} b_j^{(i)} \Delta_t^{j-1}(\tau,\dots,\tau)\,\gamma(t,\cdot), \qquad i=1,2,\dots,n-1,$$

for certain $b_j^{(i)} \in \mathbb{R}$ with $b_{n+2}^{(i)} \cdot a_n > 0$ $(i=1,2,\dots,n-1)$ [6]. There exist Δa_i, $\Delta\tau$, $\Delta\delta_{n-1} \in \mathbb{R}$, $\Delta\delta_{n-1} \geqslant 0$ such that

$$\sum_{i=1}^{n} \Delta a_i \frac{\partial}{\partial a_i} p(a) + \Delta\tau \frac{\partial}{\partial \tau} p(a) + \Delta\delta_{n-1} \frac{\partial}{\partial \delta_{n-1}} p(a) = \vartheta.$$

With $a_\lambda := (a_1+\lambda\,\Delta a_1,\dots,a_n+\lambda\,\Delta a_n,\tau+\lambda\,\Delta\tau,0,\dots,0,\lambda\,\Delta\delta_{n-1})$ the differentiability of p yields

$$\|p(a_\lambda)-g-\lambda\vartheta\|_q = o(|\lambda|), \qquad \lambda \geqslant 0.$$

This immediality proves (a) and the first part of (b) ("(i) → (ii)"). The other direction of (b) ("(ii) → (i)") is proven in analogy by making use of the differentiability of p.

Only part (c) remains to be proven. With g as above let $h \in \Gamma_{n,k}^{\gamma}$ be an element from a sufficiently small neighborhood of g: $h = p((a_1+\Delta a_1,\dots,a_n+\Delta a_n,\tau+\Delta\tau,\Delta\delta_1,\dots,\Delta\delta_{n-1})^T)$. We have to show that $\|f-h\|_q \geqslant \|f-g\|_q$ holds. This is trivial for $\|f-g\|_q = 0$. Let us therefore assume $\|f-g\|_q > 0$.

Since g is optimal with respect to its class of multiplicity we have for $a' := a_1+\Delta a_1,\dots,a_n+\Delta a_n,\tau+\Delta\tau,0,\dots,0)^T$ the inequality $\|f-p(a')\|_q \geqslant \|f-g\|_q$. Therefore we can restrict our attention to the case $\Delta\delta_i > 0$ for at least one i, $1 \leqslant i \leqslant n-1$. Since both $\|\cdot\|_q$ and p are C^1-functions in the sense of definition 1 in [6], so is the map Q that assigns to every parameter $b \in \mathbb{R}^n \times M$ the approximation error $\|f-p(b)\|_q$. Together with the alleged existence of $c_\vartheta > 0$ this implies $(\partial Q/\partial\delta_i)(b) > 0$ for all b from a neighborhood of $p^{-1}(g)$ and for all i, $1 \leqslant i \leqslant n-1$. Thus, the following chain of inequalities holds, proving that g is indeed a local best approximation:

$$\begin{aligned}\|f-h\|_q &= \|f-h\|_q - \|f-p(a')\|_q + \|f-p(a')\|_q \\ &\geqslant \sum_{i=1}^{n-1} \Delta\delta_i \frac{\partial Q}{\partial \delta_i}(a') + o(\|(\Delta\delta_1,\ldots,\Delta\delta_{n-1})^T\|) + \|f-g\|_q \\ &> \|f-g\|_q\end{aligned}$$

for at least one $\Delta\delta_i$ positive and $\|g-h\|_\infty$ sufficiently small. ∎

3. Applications to Exponential Sums

In this section we apply the general results of the preceding section to Chebyshev ($\|\cdot\|_\infty$)-approximation with exponential sums. Theorem 1 is applicable since $e^{\lambda x}$ is a normal ∞-kernel. The necessary conditions for an exponential sum to be a local best approximation also turn out to be sufficient; we therefore get characterizations of local best approximations. This deduction puts Braess' results (see Lemma 12.2, p. 30 in [3]) in the framework of general γ-polynomials and has the advantage of using *first-order-terms* only.

To be in accordance with the pertinent literature we denote the characteristic numbers (frequencies) $t_1,\ldots,t_n$ by $\lambda_1,\ldots,\lambda_n$. For Γ_n^γ with $\gamma(\lambda,x)=e^{\lambda x}$ we write E_n. To a real number a the function *signum* assigns the sign of a. This notation can be extended to exponential sums [2]: To an exponential sum with one frequency and $a_m \neq 0$ a sign-vector is assigned by

$$\text{signum}\left(e^{\lambda x}\sum_{j=1}^{m} a_j x^{j-1}\right) := (\sigma(-1)^{m-1},\ldots,-\sigma,\sigma)\in \mathbb{R}^m,\ \sigma := \text{signum}(a_m).$$

If all frequencies of g are less than all frequencies of h, and g and h have been assigned sign-vectors, we set: signum $(g+h) :=$ (signum (g), signum (h)). This way, to each element of $E_n \backslash E_{n-1}$ a sign-vector is assigned. The number of sign changes of a sign-vector is defined in a straightforward manner. For example, $-e^{3x}-e^{4.5x}(2+3x^2)$ has the sign-vector $(--+-)$ with 2 sign changes. The following version of *Descartes' rule* will be helpful: With

$$g(x) := \sum_{i=1}^{l} e^{\lambda_i x}\sum_{j=1}^{m_i} a_{ij}x^{j-1}, \qquad \lambda_1<\cdots<\lambda_l;\ a_{1m_1}\neq 0,\ldots,a_{lm_l}\neq 0$$

let there be $k+1$ points $x_0<x_1<\cdots<x_k\in X$ with $g(x_i)\cdot g(x_{i+1})<0$ ($i=0,1,\ldots,k-1$). Then signum (g) has at least k sign changes.

The proof can easily be given by induction: For the induction step apply the induction hypothesis to $e^{\lambda_1 x}((d/dx)\,e^{-\lambda_1 x}g(x))$; see also [3, p. 24].

For the remainder of this section let $g \in E_n \backslash E_{n-1}$ be fixed. We index the frequencies of g such that

$$g(x) = \sum_{i=1}^{l_1} e^{\lambda_i x} \sum_{j=1}^{m_i} a_{ij} x^{j-1} + \sum_{i=l_1+1}^{l} a_{i1} e^{\lambda_i x}$$

and $m_1, \ldots, m_{l_1} \geqslant 2$; $\lambda_1 < \cdots < \lambda_{l_1}$; that is, the multiple frequencies are $\lambda_1, \ldots, \lambda_{l_1}$. Set $\sigma_i :=$ signum (a_{im_i}) $(i = 1, 2, \ldots, l_1)$ and let k denote the number of sign changes in $((-1)^{m_2+1}\sigma_1, \ldots, (-1)^{m_{l_1}+1}\sigma_{l_1-1}, \sigma_{l_1})$; set $k := 0$ for $l_1 = 0$. Then $M := n + l + k$ is the maximum number of sign changes the sign vector of an element from

$$T_g = \left\{ \sum_{i=1}^{l_1} e^{\lambda_i x} \sum_{j=1}^{m_i+2} \alpha_{ij} x^{j-1} + \sum_{i=l_1+1}^{l} (\alpha_{i1} + \alpha_{i2} x) e^{\lambda_i x} \;\middle|\; \alpha_{ij} \in \mathbb{R}, \alpha_{1m_{i_1}} \cdot a_{1m_{i_2}} \geqslant 0, \ldots, \alpha_{l_1 m_{l_1}} \cdot a_{l_1 m_{l_1}} \geqslant 0 \right\}$$

can have, and if the number of sign changes is maximal the sign vector *must have the sign* σ_{l_1} *on the right.* With these definitions we can state

THEOREM 2. (Braess [3], Lemma 2.2) (a) *g is a local best approximation to $f \in C(X)$ iff $f - g$ has an alternant of length $M + 1$ with sign $-\sigma_{l_1}$ on the right.*

(b) *Every local best approximation is also a local strong unique best approximation.*

Proof. Let $f - g$ have an alternant of length $M + 1$ with sign $-\sigma_{l_1}$ on the right; that is, there are points $x_0, \ldots, x_M \in X$ with $(f-g)(x_i) = \|f - g\|_\infty \cdot (-1)^{M-i+1}\sigma_{l_1}$. Then 0 must be a best approximation to $f - g$ from T_g: Otherwise, there would be a $\vartheta \in T_g$ with $\vartheta(x_i) \cdot \vartheta(x_{i+1}) < 0$ $(i = 0, 1, \ldots, M-1)$ and $\vartheta(x_M) = -\sigma_{l_1}$. ϑ may be chosen to be of degree $n + l + l_1$ and not of degree $n + l + l_1 - 1$ (make small perturbations, if necessary). Because of Descartes' rule signum (ϑ) must have at least M sign changes. Since this is the maximum number of sign changes for elements of T_g, ϑ cannot have simple zeros to the right of x_M and therefore signum (ϑ) must have the sign signum $\vartheta(x_M) = -\sigma_{l_1}$ on the right. On the other hand, every element from T_g whose sign vector has a maximal number of sign changes must have the sign σ_{l_1} on the right—a contradiction. This proves that 0 is a best approximation to $f - g$ from T_g. Since T_g is a Haar-cone (the affine hull of T_g is a Haar-space) 0 is the *strong unique* best approximation to $f - g$ from T_g ([4, p. 338]; 5, p. 97]). With Theorem 1(b) we conclude that g is a local best approximation.

Let us now assume g is a local best approximation to f from E_n, but $f-g$ has no alternant of length $M+1$ with sign $-\sigma_{l_1}$ on the right. Then there is a sign $\sigma \in \{\pm\}$ and there are points $x_0 := \inf X < x_1 < \cdots < x_q := \sup X$ with: $x \in (x_i, x_{i+1}) \to (f-g)(x) \cdot \sigma \cdot (-1)^{q-i} < \|f-g\|_\infty$ $(i = 0,..., q-1)$; also $q \leqslant M$ or ($q = M+1$ and $\sigma \cdot \sigma_{l_1} = 1$). Since T_g contains a Haar subspace of dimension $n+l$, q must be greater or equal $n+l+1$.

Let T' be a q-dimensional Haar-subspace of the affine hull of T_g such that T' contains an element of T_g with $q-1$ sign changes and sign σ on the right. There is a $0 \neq \vartheta \in T'$ with $\vartheta(x_i) = 0$ $(i = 1, 2,..., q-1)$ and $\vartheta(x) \cdot (f-g)(x) > 0$ $\forall x \in X: |(f-g)(x)| = \|f-g\|_\infty$. ϑ has $q-1$ sign changes (Descartes' rule) and sign σ on the right and is therefore an element of T_g. This implies $\|f-g-\lambda\vartheta\|_\infty < \|f-g\|_\infty$ for every $\lambda > 0$ which is sufficiently small.

Theorem 1(a) yields a contradiction to the optimality of g. Therefore there must exist an alternant of $f-g$ of length $M+1$ with sign $-\sigma_{l_1}$ on the right. ∎

4. Applications to Splines with Variable Knots

Polynomial splines are piecewise polynomial functions. They are smooth yet flexible and therefore very popular for interpolating or approximating functions. There are no general characterizations of local best approximations known in the spline case because—in contrast to the situation for exponential sums— the tangent cone T_g from Section 2 is not a Haar cone. Nevertheless, the necessary condition of Theorem 1(a) is very important because it can be used as a starting point for the design of *numerical algorithms* which converge at least to stationary points.

In this section we consider a couple of special cases in which necessary or sufficient conditions for optimality can be expressed as conditions on the extremal point set.

To put splines in the context of γ-polynomials we need the notation of truncated power series: $(x)_+ := \max(0, x)$;

$$\begin{aligned} (x)_+^m &:= ((x)_+)^m, && m > 0 \\ &:= 1, && m = 0, x > 0 \\ &:= 1/2, && m = 0, x = 0 \\ &:= 0, && m = 0, x < 0. \end{aligned}$$

For simplicity let X be a closed interval, $X = [a, b] \subset \mathbb{R}$. Let us examine the necessary and sufficient conditions for the spline s,

$$s(x) := \sum_{i=1}^{l} \sum_{j=1}^{m_l} a_{ij} \Delta_t^{j-1}(t_i,..., t_i)(x-t)_+^m,$$

$a := t_0 < t_1 < \cdots < t_l < t_{l+1} := b$, to be a local best approximation to $f \in C(X)$ with respect to the maximum norm $\|\cdot\|_\infty$.

THEOREM 3. *With s as above, $m \geqslant m_l + 2$ $(i = 1, 2,..., l)$, and $s \in \Gamma_{n,m}^\gamma \backslash \Gamma_{n-1,m}^\gamma$ with $\gamma(t, x) := (x-t)_+^m$ we get:*

(a) Sufficient Condition. *For an $i \in \mathbb{N}$, $0 \leqslant i \leqslant l$, with $\|f-s\|_\infty = \|(f-s)|_{[t_i,t_{i+1}]}\|_\infty$ let $(f-s)|_{[t_i,t_{i+1}]}$ have an alternant of length $m+2$. Then s is a local strong unique best approximation from $\Gamma_{n,m}^\gamma$ to f with respect to the maximum norm.*

(b) Necessary condition. *Let s be a local best approximation from $\Gamma_{n,m}^\gamma$ to f with respect to the maximum norm. Then there is an i, $0 \leqslant i \leqslant l$, such that $\|(f-s)|_{[t_i,t_{i+1}]}\|_\infty = \|f-s\|_\infty$ and at least one of the following three conditions holds:*

(1) *$i \geqslant 1$, $m_i = 1$, and $(f-s)|_{[t_i,t_{i+1}]}$ has an alternant of length 3;*

(2) *$i \geqslant 1$, $m_i \geqslant 2$, and $(f-s)|_{[t_i,t_{i+1}]}$ has an alternant of length m_i^* with sign $(a_{im_i}(-1)^{m_i^\dagger}$ on the left;*

(3) *$i = 0$.*

Proof. Rewriting the tangent cone T_s greatly simplifies the proof:

$$T_s = \left\{ \sum_{i=1}^{l} \sum_{j=1}^{m_i^\dagger} \alpha_{ij} \Delta_t^{j-1}(t_i,..., t_i)(x-t)_+^m \;\middle|\; \begin{array}{l} \alpha_{ij} \in \mathbb{R}, m_i^* := 2, m_i = 1; \\ m_i^* := m_i + 2, m_i \geqslant 2; \alpha_{im_i^*} \cdot a_{im_i} \geqslant 0, \text{ if } m_i \geqslant 2 \; (i = 1,..., l) \end{array} \right\}$$

$$= \left\{ h \in C(X) \;\middle|\; \begin{array}{l} h|_{[t_0,t_1]} \equiv 0, \forall i, 1 \leqslant i \leqslant l \; \exists \alpha_0,..., \alpha_m \in \mathbb{R}: \\ h|_{[t_i,t_{i+1}]}(x) = \sum_{j=0}^{m} \alpha_j (x-t_i)^j \\ h \text{ is } m - m_i^* \text{ times continuously differentiable at } t_i, \text{ and} \\ \alpha_{m+1-m_i^\dagger} \cdot a_{im_i}(-1)^{m_i^\dagger+1} \geqslant 0, \text{ if } m_i \geqslant 2. \end{array} \right\}$$

Part (a) of Theorem 3 follows immediately from this representation of T_s and Theorem 1, (b) (i) → (ii) using the fact that every tangent cone element is a mth degree polynomial on $[t_i, t_{i+1}]$. We prove part (b) of Theorem 3 by contradiction. Let s be a local best approximation such that for no $i \in \mathbb{N}$, $0 \leqslant i \leqslant l$, one of the three given conditions holds.

We define an element h of T_s in the following way: $h|_{[t_0,t_1]} :\equiv 0$; suppose

$h|_{[t_0,t_i]}$ is already defined and $\|h|_{[t_0,t_i]}\|_\infty$ is sufficiently small. Determine $\beta_0,\dots,\beta_m \in \mathbb{R}$ such that

$$\frac{d^k}{dx^k}\bigg|_{x=t_i}\left(\sum_{j=0}^{m-m_i^*} \beta_j(x-t_i)^j\right) = \lim_{x \uparrow t_i} \frac{d^k}{dx^k} h(x), \qquad k=0,1,\dots,m-m_i^*,$$

and

$$(x-t_i)^{m-m_i^*+1} \sum_{j=0}^{m_i^*+1} \beta_{m-m_i^*+1+j}(x-t_i)^j \begin{cases} >0 & \forall x \in (t_i, t_{i+1}]: (f-s)(x) = \|f-s\|_\infty \\ <0 & \forall x \in (t_i, t_{i+1}]: (f-s)(x) = -\|f-s\|_\infty. \end{cases}$$

For a sufficiently small positive constant c set

$$h|_{[t_i,t_{i+1}]}(x) := \sum_{j=0}^{m-m_i^*} \beta_j(x-t_i)^j + c \sum_{j=m-m_i^*+1}^{m} \beta_j(x-t_i)^j.$$

If m_i is greater one $\beta_{m-m_i^*+1}$ can moreover be chosen to have the same sign as $a_{im_i}(-1)^{m_i^*+1}$. Thus, h is defined on $[a,b]$ and an element of T_s. h was constructed such that for small positive λ

$$\|f-s-\lambda h\|_\infty < \|f-s\|_\infty$$

holds. In view of Theorem 1(a), this is a contradiction to the assumption that s is a local best approximation to f. This proves Theorem 3(b). ∎

Remarks. In contrast to the case of exponential sums, splines allow no first-order characterization of local best approximations. For applications additional constraints may be imposed on the approximating class of spline functions. Also, the interval limits a and b are often taken as fixed knots. These and other cases can be handled in analogy to the above analysis with only minor adjustments.

In some cases the assumptions on the differentiability of the kernel ($m \geqslant m_i + 2$) exclude interesting cases such as the coalescing of knots for cubic splines. Since $d^m(x-t)_+^m/dx^m$ is discontinuous only for $x = t$ and even then still bounded, the assumptions of Theorem 3 can indeed be weakened:

COROLLARY. *The assumption $m \geqslant m_i + 2$ in Theorem 3 can be weakened to $m \geqslant m_i + 1$ if the knots are not extremal points, that is,*

$$|(f-s)(t_i)| < \|f-s\|_\infty \qquad \textit{for all } i,\ 1 \leqslant i \leqslant l, \textit{ with } m = m_i + 1.$$

Remark. The sufficient condition for optimality given in Theorem 3(a) is rather stringent since we have only m_i^* free parameters per interval $[t_i, t_{i+1}]$, but require an alternant of length $m + 2$ ($\geqslant m_i^* + 2$). Indeed, with the exception of comparatively few cases an alternant of length $m_i^* + 1$ is sufficient to guarantee local optimality.

References

1. C. de Boor, On the approximation by γ-polynomials, *in* "Approximation with Special Emphasis on Spline Functions" (I. J. Schoenberg Ed.), Academic Press, New York/London, 1969.
2. D. Braess, Über die Vorzeichenstruktur der Exponentialsummen, *J. Approx. Theory* **3** (1970), 101–113.
3. D. Braess, Chebyshev approximation by γ-polynomials, I, II, III, *J. Approx. Theory* **9** (1973), 20–43; **11** (1974), 16–37; **24** (1978), 119–145.
4. D. Braess, Kritische Punkte bei der nichtlinearen Tschebyscheff-Approximation, *Math. Z.* **132** (1973), 327–341.
5. L. J. Cromme, "Approximation auf Mannigfaltigkeiten mit Spitzen-Theorie und numerische Methoden," Habilitationsschrift, Göttingen, 1979.
6. L. J. Cromme, Regular C^1-parametrizations for exponential sums and splines, *J. Approx. Theory* **35** (1982), 30–44.
7. L. J. Cromme, Algorithms for approximation by exponential sums and splines, in preparation.
8. C. R. Hobby and J. R. Rice, Approximation from a curve of functions, *Arch. Rat. Mech. Anal.* **24** (1967), 91–106.
9. P. Zencke, Zur Parametrisierung reeller Exponentialsummen durch lineare Differentialgleichungen mit Parameterrestriktionen, *in* "Numerische Methoden der Approximationstheorie" (L. Collatz, G. Meinardus, and H. Werner Eds.), Band 4, Birkhäuser, Basel, 1978.