JOURNAL OF APPROXIMATION THEORY 11, 16–37 (1974)

Chebyshev Approximation by γ-Polynomials, II

DIETRICH BRAESS

Institut für numerische und instrumentelle Mathematik, Westfälische Wilhelms-Universität, 44 Münster, West Germany

Communicated by E. W. Cheney

Best approximation in the sense of Chebyshev is not always unique for γ-polynomials. In this paper we prove that in the normal case the number of best approximations is finite. A necessary and sufficient condition on alternants of local best approximations is established.

In connection with this criterion we define a new structural integer which lies between the order k and the length l, i.e., between the structural integers dominating the theory up to now.

The main tool of this research is the tangent cone. The cone coincides with the tangent space defined by Meinardus and Schwedt [8] provided that all characteristic numbers are distinct. But the tangent cone makes a tangential characterization in the sense of Wulbert [11] possible even if some characteristic numbers coalesce and if the tangent space suffers a loss of dimension [6]. Since the tangent cone is a convex subset of a Haar space we have even a local strong uniqueness condition.

Our investigations depend heavily on the results in [1]. Therefore we proceed with enumerating formulas and theorems. All references to Formulas (1.1)–(8.7) and Theorems 2.1–8.7 are related to that paper. The logical dependence of the different sections is shown in Fig. 1.

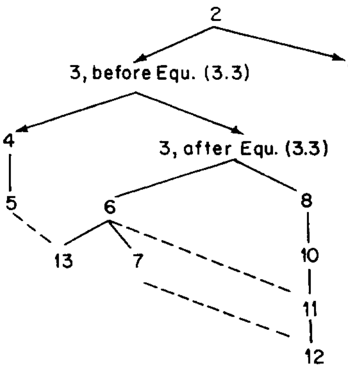

FIGURE 1

INTRODUCTION

Since it is known that sums of exponentials are not varisolvent [2], there are several open problems concerning Chebyshev approximation by these functions. The questions remained unsettled when the investigations were extended to γ-polynomials [1, 5]. Hobby and Rice started the research with the introduction of the proper γ-polynomials of order N,

$$F(a, x) = \sum_{\nu=1}^{N} \alpha_\nu \gamma(t_\nu, x), \qquad \alpha_\nu \in \mathbb{R}, \quad t_\nu \in T, \tag{1.1}$$

where T is a subset of $\mathbb{R}$ and $\gamma \in C(T \times X)$. As has been pointed out already by those authors, it is necessary to close the family of γ-polynomials in order to ensure the existence of best approximations. Assuming the derivatives $\gamma^{(\mu)} = \partial^\mu \gamma / \partial t^\mu$ exist, extended γ-polynomials of the following form are considered:

$$F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \gamma^{(\mu)}(t_\nu, x), \qquad \sum_{\nu=1}^{l} (1 + M_\nu) = k \leqslant N. \tag{1.2}$$

Functions of the form (1.2) containing derivatives may be interpreted as γ-polynomials with coalescing characteristic numbers t_ν.

This extension, however, entailed some serious complications. The most striking drawback was the gap between the sufficient and the necessary conditions on alternants. This gap corresponds to the fact that best approximations are not always unique. Even the question whether the number of best approximations is always finite remained unsettled.

In the present paper we overcome these difficulties. We establish a necessary and sufficient condition by characterizing local best approximations instead of best approximations.

Throughout this paper the family of γ-polynomials V_N is endowed with the relative topology of $C(X)$ which is induced by a (weighted) uniform norm

$$\|f\| = \sup_{x \in X} w(x) \cdot |f(x)|.$$

For this, we restrict our attention to normal families, i.e., extended Descartes families being normal in the sense of Definition 8.1.

There will be 14 constants $c_1, c_2, \ldots, c_{14}$. They are not universal; they will depend on the function $F[a]$, on subspaces, and on sets of indices, but they will be independent of the functions in the neighborhood of $F[a]$ which are compared with $F[a]$.

Finally, we propose the reader glance at the examples in Section 13 in advance. The first example illustrates once more the necessity of the extension mentioned at the beginning. On the other hand for some special functions, it is known that the best approximation is a unique proper γ-polynomial with $\alpha_\nu > 0$ (see Section 5). This was considered by Karlin [7] in some other context.

10. Basic Lemmas

In this section we investigate the neighborhood of those γ-polynomials having a maximally degenerate spectrum,

$$F(x) = \sum_\mu \alpha_\mu \gamma^{(\mu)}(t, x).$$

We establish a map into a certain cone, which in Section 12 will be identified with the tangent cone. Moreover, the distance between the γ-polynomials and their images will be small of order greater than one.

The following lemma will be used repeatedly for the correlation of different estimations.

Lemma 10.1. *Let $h_1, h_2, \ldots, h_m$ be a basis of an m-dimensional subspace of a normed space. There are constants $c_1, c_2 > 0$, such that for every $h = \sum_{i=1}^m \alpha_i h_i$ and for every subset I of the integers from* 1 *to m the following estimations are valid*

$$\Big\| \sum_{i\in I} \alpha_i h_i \Big\| \leqslant c_1 \cdot \| h \|,$$

$$| \alpha_i | \leqslant c_2 \| h \|, \qquad i = 1, 2, \ldots, m.$$

Proof. Estimations of the second kind are often used when the existence of best approximations in finite dimensional subspaces is proved [4]. By setting $c_1 = c_2 \sum_{i=1}^m \| h_i \|$ the first inequality is established. ∎

Throughout the rest of the paper we will assume that $(\partial^{N+2}/\partial t^{N+2})\, \gamma(t, x)$ exists and is continuous. This is no drawback, since the interesting kernels $\gamma(t, x)$ presented in Section 9 are even contained in the class C^∞.

Let F^* be a γ-polynomial with m coalescing characteristic numbers:

$$F^*(x) = \sum_{\mu=1}^m \alpha_\mu \gamma_\mu(\tau, \tau, \ldots, \tau; x), \qquad \alpha_m \neq 0, \quad m \geqslant 1. \tag{10.1}$$

The function is represented by using divided differences. Contrary to the first part of the paper, the number of t-arguments is given as a suffix of γ. This is convenient, since often all arguments coincide or there are arguments repeated and the number is not clear. In particular, we have $\gamma_\mu(\tau, \tau,\ldots, \tau; x) = \gamma^{(\mu-1)}(\tau; x)/(\mu - 1)!$ The γ-polynomials from the neighborhood of F^* are written in the form

$$F(x) = \sum_{\mu=1}^{m} \beta_\mu \gamma_\mu(t_1, t_2,\ldots, t_\mu; x). \tag{10.2}$$

We calculate the derivatives with respect to the parameters β_μ, t_μ and make use of the relation

$$(\partial/\partial t_\nu)\, \gamma_\mu(t_1, t_2,\ldots, t_\mu; x) = \gamma_{\mu+1}(t_1,\ldots, t_\mu, t_\nu; x). \tag{10.3}$$

This yields

$$\begin{aligned}
(\partial/\partial\beta_\mu)\, F &= \gamma_\mu(t_1,\ldots, t_\mu; x),\\
(\partial/\partial t_\mu)\, F &= \sum_{\rho\geqslant\mu} \beta_\rho \gamma_{\rho+1}(t_1,\ldots, t_\nu, t_\rho; x),\\
(\partial^2/\partial\beta_\mu\, \partial\beta_\nu)\, F &= 0,\\
(\partial^2/\partial\beta_\nu\, \partial t_\mu)\, F &= \begin{cases}\gamma_{\nu+1}(t_1,\ldots, t_\nu, t_\mu; x), & \text{if } \mu \leqslant \nu,\\ 0, & \text{otherwise},\end{cases}\\
(\partial^2/\partial t_\mu^2)\, F &= 2\sum_{\rho\geqslant\mu} \beta_\rho \gamma_{\rho+2}(t_1,\ldots, t_\rho, t_\mu, t_\mu; x),\\
(\partial^2/\partial t_\mu\, \partial t_\nu)\, F &= \sum_{\rho\geqslant\mu} \beta_\rho \gamma_{\rho+2}(t_1,\ldots, t_\rho, t_\mu, t_\nu; x), \qquad \text{if } \mu < \nu.
\end{aligned}$$

Set $\delta_\mu = \beta_\mu - \alpha_\mu$ and $u_\mu = t_\mu - \tau$ for $\mu = 1, 2,\ldots, m$. $F(x)$ is expanded into a Taylor's series in powers of $d = (\delta_1,\ldots, \delta_m, u_1,\ldots, u_m)$ and terms of order one and two are taken into account,

$$\begin{aligned}
F(x) - F^*(x) &= \sum_\mu \delta_\mu\gamma_\mu + \sum_\mu \sum_{\rho\geqslant\mu} u_\mu\alpha_\rho\gamma_{\rho+1} + \sum_\nu \sum_{\mu\leqslant\nu} u_\mu\delta_\nu\gamma_{\nu+1}\\
&\quad + \sum_\mu \sum_{\nu\geqslant\mu} \sum_{\rho\geqslant\nu} u_\mu u_\nu \alpha_\rho \gamma_{\rho+2} + O(d^3)\\
&= \sum_{\rho=1}^{m+2} e_\rho\gamma_\rho + O(d^3).
\end{aligned} \tag{10.5}$$

Here, the coefficients $e_\rho = e_\rho(d)$ are determined by collecting terms with divided differences $\gamma_\rho = \gamma_\rho(\tau, \tau,\ldots, \tau; x)$ of the same order. For convenience, the (dummy) variables $\alpha_0 = \alpha_{-1} = \delta_0 = \delta_{m+1} = 0$ are introduced. Since

$$\sum_\mu \sum_{\nu\geqslant\mu} u_\mu u_\nu = \tfrac{1}{2} \sum_\mu \sum_\nu u_\nu u_\mu + \tfrac{1}{2} \sum_\mu u_\mu^2,$$

we obtain from (10.5)

$$e_\rho = \delta_\rho + (\alpha_{\rho-1} + \delta_{\rho-1}) \sum_{\mu \leqslant \rho-1} u_\mu + \alpha_{\rho-2} \sum_{\mu \leqslant \nu \leqslant \rho-2} u_\mu u_\nu, \quad \rho = 1, 2, \ldots, m+1,$$

$$e_{m+2} = \tfrac{1}{2}\,\alpha_m \left[\left(\sum_\mu u_\mu\right)^2 + \sum_\mu u_\mu{}^2\right]. \tag{10.6}$$

We want to extract from (10.5) a statement which is independent of the special choice of representation. Therefore, we define subsets of γ-polynomials with fixed characteristic numbers by

$$V_N[t_1, \ldots, t_N] = \left\{F \in V_N\,;\, F = \sum_{\mu=1}^{N} \beta_\mu \gamma_\mu(t_1, \ldots, t_\mu\,;\, x);\ \beta_\mu \in \mathbb{R}\right\}. \tag{10.7}$$

Moreover, for each γ-polynomial with coalescing t's represented as in (10.1), the sign of α_m is denoted as leading sign. Recall that this number determines the sign vector introduced in Section 3.

Now we are ready to establish the following lemma.

LEMMA 10.2. *Let $m \geqslant 2$ and V_m be a normal family. Then, given $F^* \in V_m[\tau, \tau, \ldots, \tau] \backslash V_{m-1}$, there exists a constant $c_3 > 0$ and a continuous mapping ϕ: $V_m \backslash V_{m-1} \to V_{m+2}[\tau, \tau, \ldots, \tau]$, such that*

$$\phi(F^*) = 0,$$

and the inequality

$$\| F - F^* - \phi(F)\| \leqslant c_3 \| \phi(F)\|^{3/2} \tag{10.8}$$

holds. Moreover, $\phi(F)$ has the same leading sign as F^* or $\phi(F) \in V_{m+1}$.

Note. If γ is only $(m + 2)$-times differentiable the inequality (10.8) must be replaced by

$$\| F - F^* - \phi(F)\| = o(\| \phi(F)\|). \tag{10.8a}$$

This relation is sufficient for the proofs in the following sections, but we prefer to use (10.8) and to avoid expressions with the Landau-symbols.

Proof of Lemma 10.2. By virtue of Theorem 8.3, a mapping with domain $V_m \backslash V_{m-1}$ is continuous if it is continuous with respect to the parameters of the representation (10.2). Hence, by (10.6) and $\phi_0(F) = \sum_{\rho=1}^{m+2} e_\rho(d)\, \gamma_\rho(\tau, \ldots, \tau; x)$, a continuous mapping ϕ_0: $V_m \backslash V_{m-1} \to V_{m+2}[\tau, \ldots, \tau]$ is defined. As a first step we prove the inequality (10.8) for all F in a neighborhood of F^*, with ϕ replaced by ϕ_0.

For economy, with each (finite) sequence of numbers $\{\beta_\nu\}$, we associate the norm

$$\| \beta \|_n = \sup_{1 \leqslant \nu \leqslant n} | \beta_\nu |.$$

In the particular case when $n = m$, the sufix m will be suppressed. Referring to (10.6), we claim that

$$\| e \|_{m+2} \geqslant c_4 \cdot \max(\| \delta \|, \| u \|)^2,$$

where

$$c_4 = \frac{| \alpha_m |}{8m^2(1 + \| \alpha \|)^2}, \tag{10.9}$$

provided $\| u \| \leqslant 1$, $\| \delta \| \leqslant 1$. Obviously, we have $0 < c_4 < \frac{1}{2} | \alpha_m |$. This implies

$$\begin{aligned} \| e \|_{m+2} \geqslant | e_{m+2} | = \tfrac{1}{2} | \alpha_m | \sum_\mu u_\mu{}^2 \geqslant \tfrac{1}{2} | \alpha_m | \cdot \| u \|^2 \\ = c_4 \cdot \| u \|^2. \end{aligned} \tag{10.10}$$

To prove $\| e \|_{m+2} \geqslant c_4 \| \delta \|^2$ we consider two cases.

Case 1. Let $\| \delta \| \leqslant 2m(1 + \| \alpha \|) \cdot \| u \|$. Combining this relation with (10.10) yields

$$\| e \|_{m+2} \geqslant \tfrac{1}{2} | \alpha_m | \cdot [\| \delta \|/2m(1 + \| \alpha \|)]^2 = c_4 \cdot \| \delta \|^2.$$

Case 2. Let $\| \delta \| \geqslant 2m(1 + \| \alpha \|) \| u \|$. Choose ρ, such that $| \delta_\rho | = \| \delta \|$. From (10.6) we obtain

$$\begin{aligned} | e_\rho | &\geqslant | \delta_\rho | - | \alpha_{\rho-1} + \delta_{\rho-1} | \sum_\nu | u_\nu | - | \alpha_{\rho-2} | \sum_{\mu,\nu} | u_\mu u_\nu | \\ &\geqslant \| \delta \| - (1 + \| \alpha \|)\, m \, \| u \| - \| \alpha \| \cdot m^2 \| u \|^2. \end{aligned}$$

By virtue of the relation of $\| \delta \|$ and $\| u \|$ we estimate

$$\| e \|_{n+2} \geqslant | e_\rho | \geqslant \| \delta \| - \tfrac{1}{2} \| \delta \| - \tfrac{1}{4} \| \delta \|^2 \geqslant \tfrac{1}{4} \| \delta \| \geqslant c_4 \| \delta \|^2.$$

Since $c_4 \leqslant \frac{1}{4}$ this completes the proof of (10.9). Applying Lemma 10.1 to $V_{m+2}[\tau,..., \tau]$ we obtain a constant c_5, such that $h = \sum_{\rho=1}^{m+2} e_\rho \gamma_\rho$ implies

$$\| e \|_{m+2} \geqslant c_5 \| h \|.$$

Since $\gamma(t, x)$ is assumed to be $m + 2$ times continuously differentiable in t, we obtain from (10.5)

$$\| F - F^* - h \| \leqslant c_6 \cdot \max(\| \delta \|, \| u \|)^3, \tag{10.12}$$

with $h = \phi_0(F)$ for every F in a neighborhood of F^*.

Combining (10.9), (10.11), and (10.12), the statement (10.8) is established by setting $c_3 = c_4^{-3/2} \cdot c_5^{3/2} \cdot c_6$.

Since ϕ_0 is continuous and $\phi_0(F^*) = 0$, there exists a number $r > 0$, such that $\| F - F^* \| < r$ ensures inequality (10.8) and $\| \phi_0(F)\| \leqslant (2c_3)^{-2}$. Hence,

$$\| F - F^* - \phi_0(F)\| \leqslant \tfrac{1}{2} \| \phi_0(F)\|.$$

The triangle inequality implies

$$\tfrac{2}{3} \| F - F^* \| \leqslant \| \phi_0(F)\| \leqslant 2 \| F - F^* \|. \tag{10.12}$$

From (10.6) we know that $\phi_0(F) \neq 0$, provided $F \neq F^*$. Consequently,

$$\phi(F) = \begin{cases} \phi_0(F), & \text{if } \| F - F^* \| \leqslant r, \\ \phi_0(F) \cdot \max \left[1, \dfrac{\| F - F^* \|}{2 \| \phi_0(F)\|}\right], & \text{if } \| F - F \| > r \end{cases}$$

defines a continuous mapping. In particular, if $\| F - F^* \| \geqslant r$ holds, the definition implies $\| \phi(F)\| \geqslant \frac{1}{2} \| F - F^* \|$ and

$$\begin{aligned} \| F - F^* - \phi(F)\| &\leqslant \| F - F^* \| + \| \phi(F)\| \leqslant 3 \| \phi(F)\| \\ &\leqslant 3r^{-1/2} \cdot \| F - F^* \|^{1/2} \cdot \| \phi(F)\| \leqslant 6 \cdot r^{-1/2} \cdot \| \phi(F)\|^{3/2}. \end{aligned}$$

With this (10.8) is established for every $F \in V_m \backslash V_{m-1}$, since c_3 may be replaced by $6r^{-1/2}$, if necessary.

Finally, the statement on the leading sign is an immediate consequence of the fact that e_{m+2} is obtained from α_m by multiplying it with a nonnegative number. ∎

If the domain is restricted to those γ-polynomials with coalescing characteristic numbers, we obtain a mapping which admits a sharper estimation.

LEMMA 10.3. *Let* $m \geqslant 1$ *and* V_m *be a normal family. Then given* $F^* \in V_m[\tau, \tau, \ldots, \tau] \backslash V_{m-1}$ *there exists a constant* $c_7 > 0$ *and a continuous mapping of the subset of* $V_m \backslash V_{m-1}$ *containing the elements whose m characteristic numbers coincide, into* $V_{m+1}[\tau, \tau, \ldots, \tau]$, *such that*

$$\phi(F^*) = 0,$$

and the inequality

$$\| F - F^* - \phi(F)\| \leqslant c_7 \cdot \| \phi(F)\|^2 \tag{10.13}$$

holds.

Note. Instead of (10.13) an estimate of the form

$$\| F - F - \phi(F)\| \leqslant c_8 \| \phi(F)\|^{3/2}$$

may be derived.

Outline of proof. Consider in (10.5) and (10.6) only the terms of first order and specialize to the case where $u_1 = u_2 = \cdots = u_m$. It follows that

$$F(x) - F^*(x) = \sum_{\rho=1}^{m+1} \tilde{e}_\rho \gamma_\rho(\tau, \tau, \ldots, \tau; x) + O(d^2),$$

where

$$\tilde{e}_\rho = \delta_\rho + (\rho - 1)\, \alpha_{\rho-1} u_1\,, \qquad \rho = 1, 2, \ldots, m, \qquad e_{m+1} = m\alpha_m u_1\,. \tag{10.14}$$

By considering the cases where $\| \delta \| \leqslant 2m(1 + \| \alpha \|) \cdot | u_1 |$ and $\| \delta \| \geqslant 2m(1 + \| \alpha \|) \cdot | u_1 |$ it is easily verified that

$$\| \tilde{e} \|_{m+1} \geqslant c_9 \cdot \max(\| \delta \|, | u_1 |),$$

where

$$c_9 = | \alpha_m |/2(1 + \| \alpha \|).$$

The remaining part of the proof proceeds by repeating arguments of the proof for the preceeding lemma. ▮

11. Stationary Points

In this section we will prove that there are only a finite number of best approximations apart from some pathological cases. To this end we will verify that to every best approximation there is a neighborhood without further best approximation. This statement follows from a more general result including local best approximations and those γ-polynomials satisfying the necessary condition on the alternant given in Theorem 6.2(ii). Therefore, the notation of stationary elements is introduced.

Definition 11.1. $F[a] \in V_N$ is a stationary element to f in V_N, if there exists an alternant of length $N + l(a) + 1$ for $F[a]$.

Since the alternant condition in Theorem 6.2(ii) also applies to local best approximations (short l.b.a.), every l.b.a. is a stationary element. Moreover,

$F[a]$ satisfies a local strong uniqueness condition (with respect to f) in V_N, if there is a number $c > 0$ and a neighborhood U of $F[a]$ such that

$$\|f - F[b]\| \geqslant \|f - F[a]\| + c \, \| F[b] - F[a]\|$$

holds for every $F[b] \in V_N \cap U$ [Ref. (10)].

LEMMA 11.1 *Let $f \in C(X)$ and $p \geqslant 1$. Then the set of functions in V_N having an alternant of length $\leqslant p$, is open in V_N.*

Proof. Set $\epsilon(x) = f(x) - F^*(x)$. Suppose that the corresponding alternant has the exact length q, where $q \leqslant p$. Then, by standard arguments the interval X is divided into q subintervals by $q - 1$ points $\xi_1 < \xi_2 < \cdots < \xi_{q-1}$ such that

$$\sigma \cdot (-1)^i \cdot w(x) \cdot \epsilon(x) < \| \epsilon \|, \qquad \text{if} \quad x_i \in [\xi_{i-1}, \xi_i], \qquad i = 1, 2, \ldots, q, \quad (11.1)$$

where $\sigma = +1$ or $\sigma = 1$, and ξ_0, ξ_q denote the end points of X. Hence,

$$0 < \rho = \min_{1 \leqslant i \leqslant q} \min_{\xi_{i-1} \leqslant x \leqslant \xi_i} \{\| \epsilon \| - \sigma \cdot (-1)^i \cdot w(x) \cdot \epsilon(x)\}.$$

Assume that $\| F - F^* \| < \rho/2$. Then (11.1) holds with $\epsilon(x)$ replaced by $f(x) - F(x)$. Hence there is no alternant of length $q + 1$ to $f - F$. ▮

Now we are ready to prove the main result of this section.

THEOREM 11.2. *Let V_N be a normal family and let $F[a] \in V_N \backslash V_{N-1}$ be a stationary element to f in V_N. Then there is a neighborhood of $F[a]$ containing no further stationary element.*

Proof. If $F[a]$ is a stationary element, there is an alternant with exact length $N + l(a) + 1 + p$, where $p \geqslant 0$. By virtue of Lemma 11.1 there is a neighborhood of $F[a]$, which contains no stationary point $F[b] \in V_N$ with $l(b) > l(a) + p$. Therefore, in the remainder of this proof we restrict our attention to elements $F[b] \in V_N$, for which $l(b) \leqslant l(a) + p$ is valid. We establish a local strong uniqueness condition in the restricted set. From this, by arguments similar to the theorem of de la Vallee–Poussin the statement is established.

We write $F[a]$ in the form

$$F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=1}^{m_\nu} \alpha_{\nu\mu} \cdot \gamma_\mu(t_\nu, \ldots, t_\nu; x), \tag{11.2}$$

and the elements of a neighborhood are represented by using the same integer parameters $l = l(a)$ and m_ν, $\nu = 1, 2, \ldots, l$,

$$F(b, x) = \sum_{\nu=1}^{l} \sum_{\mu=1}^{m_\nu} \beta_{\nu\mu} \cdot \gamma_\mu(t_{\nu 1}, \ldots, t_{\nu\mu}; x). \tag{11.3}$$

This representation is possible as long as the characteristic numbers $t_{\nu\mu}$ differ from t_ν by less than $\frac{1}{2} \min | t_\nu - t_{\nu+1} |$. Moreover, the parameters $\beta_{\nu\mu}$, $t_{\nu\mu}$ are continuous functions of $F[b]$. Indeed, for the parameters $t_{\nu\mu}$ the statement is an immediate consequence of Theorem 8.3. The continuity of the $\beta_{\nu\mu}$ follows from the same theorem, since the functions used in the bases of (11.3) and (8.1) are connected by simple continuous transformations. Hence, for $\nu = 1, 2, \ldots, l$ the projections which send $F[b]$ to the partial sums

$$\Psi_\nu\colon F[b] \to \sum_{\mu=1}^{m_\nu} \beta_{\nu\mu} \cdot \gamma_\mu(t_{\nu 1}, \ldots, t_{\nu\mu}; x) \tag{11.4}$$

are defined on a neighborhood of $F[a]$ and are continuous mappings into V_{m_ν}.

Let I be a subset of the integers from 1 to l containing at most p elements. We restrict ourselves to those elements in V_N whose characteristic numbers coalesce in certain partial sums. To be more precise, we assume

$$t_{\nu 1} = t_{\nu 2} = \cdots = t_{\nu m_\nu}, \qquad \text{if} \quad \nu \notin I. \tag{11.5}$$

Now Lemmas 10.2 and 10.3 are applied to the partial sums $\Psi_\nu(F[b])$ for $\nu \in I$ and for $\nu \notin I$, respectively. We obtain l constants $c^{(\nu)}$ depending only on $F[a]$ and the subset I, such that for $F[b]$ there are l functions $h^{(\nu)}$ satisfying

$$\begin{aligned} h^{(\nu)} &\in V_{m_\nu+2}[t_\nu, \ldots, t_\nu], && \text{if} \quad \nu \in I, \\ h^{(\nu)} &\in V_{m_\nu+1}[t_\nu, \ldots, t_\nu], && \text{if} \quad \nu \notin I, \end{aligned} \tag{11.6}$$

and

$$\| \Psi_\nu(F[b]) - \Psi_\nu(F[a]) - h^{(\nu)} \| \leqslant c^{(\nu)} \| h^{(\nu)} \|^{3/2}, \qquad \nu = 1, 2, \ldots, l.$$

By summing up it follows that

$$\| F[b] - F[a] - h \| \leqslant \sum_{\nu=1}^{l} c^{(\nu)} \| h^{(\nu)} \|^{3/2},$$

where $h = \sum_{\nu=1}^{l} h^{(\nu)}$. By virtue of Lemma 10.1 the norms of the terms $h^{(\nu)}$,

($\nu = 1, 2, \ldots, l$) may be estimated by $\| h \|$. Hence, there exists a number $c_{10} > 0$ depending on $F[a]$ and I, such that[1]

$$\| F[b] - F[a] - h \| \leqslant c_{10} \cdot \| h \|^{3/2}. \tag{11.7}$$

Denote the number of elements of I by p_1. By construction we have $p_1 \leqslant p$. There is an alternant of length $N + l + p_1 + 1$ to $\epsilon(x) = f(x) - F(a, x)$. Hence, zero is the best approximation to f in the linear subspace $W = V_{N+l+p_1}[t_1, \ldots, t_1, t_2, \ldots, t_l]$, where each characteristic number is counted with the multiplicity given in (11.6). Since $\gamma(t, x)$ is an extended sign-regular kernel, W satisfies the Haar condition. By virtue of the strong unicity theorem [4, p. 80] there is a constant $c_{11} > 0$ such that

$$\| \epsilon - h \| \geqslant \| \epsilon \| + c_{11} \cdot \| h \|, \qquad h \in W. \tag{11.8}$$

Combining (11.7) and (11.8), we obtain

$$\| f - F[b] \| \geqslant \| f - F[a] \| + c_{11} \| h \| - c_{10} \| h \|^{3/2}. \tag{11.9}$$

Since the mapping which sends $F[b]$ to $h \in W$ is continuous, and zero is the image of $F[a]$, we have

$$\| h \| < \tfrac{1}{4}(c_{11}/c_{10})^2 \tag{11.10}$$

for the image of every element in a neighborhood of $F[a]$. The inequalities (11.9) and (11.10) imply

$$\| f - F[b] \| \geqslant \| f - F[a] \| + \tfrac{1}{2} c_{11} \| h \|. \tag{11.11}$$

To finish the proof, we may assume that the characteristic numbers of $F[b]$ violate the relation

$$t_{\nu 1} = t_{\nu 2} = \cdots = t_{\nu m_\nu}$$

for each $\nu \in I$, since otherwise our argument may be repeated with a reduced set I. Hence, $l(b) \geqslant l(a) + p_1$. Suppose, that $F[b]$ is a stationary element.

[1] If (10.8) is replaced by (10.8a), we obtain the weaker relation

$$\| F[b] - F[a] - h \| = o(h).$$

This does not affect the proof, since we may replace the condition (11.10) below by

$$\| F[b] - F[a] - h \| < \tfrac{1}{2} c_{11} \cdot \| h \|.$$

We will derive a contradiction from (11.11). Set $q = N + l(a) + p_1 + 1$. Let $x_1 < x_2 < \cdots < x_q$ be the points of an alternant to $F[b]$. From (11.10) and (11.11) we conclude

$$\begin{aligned}
&(-1)^i \cdot \sigma \cdot w(x_i) \cdot [f(x_i) - F(a, x_i) - h(x_i)] \\
&\quad \geqslant (-1)^i \cdot \sigma \cdot w(x_i) \cdot [f(x_i) - F(b, x_i)] - \| F[b] - F[a] - h \| \\
&\quad \geqslant \| f - F[b] \| - c_{10} \| h \|^{3/2} \\
&\quad \geqslant \| f - F[a] \| + \tfrac{1}{2} \cdot c_{11} \| h \| - c_{10} \| h \|^{3/2} \\
&\quad > \| f - F[a] \| \geqslant (-1)^i \sigma \cdot w(x_i) \cdot [f(x_i) - F(a, x_i)], \qquad i = 1, 2,..., q,
\end{aligned}$$

where $\sigma = +1$ or $\sigma = -1$. Hence,

$$(-1)^i \sigma \cdot h(x_i) > 0, \qquad i = 1, 2,..., q,$$

and $h \in W$ has $q - 1 = N + l(a) + p_1$ zeros, contradicting the Haar condition. Consequently, $F[b]$ is not a stationary element, provided that $F[b]$ lies in the neighborhood of $F[a]$ determined by (11.10) and the characteristic numbers satisfy (11.5).

Since this consideration may be repeated for all sets I with p or less integers, the statement holds for all elements with $l(b) \leqslant l(a) + p$. ∎

From Theorem 11.2 we obtain the following corollary.

COROLLARY 11.3. *Let V_N be a normal family. Then the set of local best approximations and the set of stationary elements are countable.*

Proof. The proof proceeds by induction on N. There is at most one stationary element in $V_1 = V_1^0$. Assume that the statement has been proven for $N - 1$ and consider the stationary elements in V_N. By virtue of Theorem 8.5 $V_N \backslash V_{N-1}$ is σ-compact. With this, it follows from the preceding theorem that the subset of stationary points in $V_N \backslash V_{N-1}$ is countable. On the other hand, those stationary elements, which are already contained in V_{N-1}, are also stationary with respect to V_{N-1}. By the inductive hypothesis they can be counted. This completes the proof because l.b.a.'s are also stationary elements. ∎

We remind the reader that we have more information in the case, when $N = 2$. By virtue of Theorem 7.2 there is at most one l.b.a. in each sign class $V_2(s_1, s_2)$. Combining the methods of Section 7 with the results of Theorem 12.3 we obtain an analogous statement for the sign classes of V_3, apart from $V_3(+-+)$ and $V_3(-+-)$. In the latter sign classes there may be more than one l.b.a. as is shown in Section 13.

In the particular case when we are concerned with exponentials, i.e., with $\gamma(t, x) = e^{tx}$ on $\mathbb{R} \times X$, then we have a sharper result. The subsets [9]

$$\{F \in V_N\,;\, \|f - F\| \leqslant M\}$$

are compact or empty provided that $M < \inf\{\|f - F\|; F \in V_{N-1}\}$. As an immediate consequence we have the following corollary.

COROLLARY 11.4. *Let V_N be the family of exponentials of order $\leqslant N$. Suppose that there is no best approximation to f in V_N with an order $k \leqslant N - 1$.*[2] *Then there are only a finite number of best approximations to f in V_N.*

12. THE CRITERION FOR LOCAL BEST APPROXIMATIONS

According to Theorem 6.2 the elements $F[a]$ in the particular subset $V_N{}^0$ are best approximations (and l.b.a.) if and only if there is an alternant of length $N + k(a) + 1$ for $F[a]$. Then a best approximation $F[a]$ may also be characterized by the fact that zero is the unique best approximation to $f - F[a]$ in the linear tangent space. Following the terminology of Wulbert [11] we have a tangential characterization.

If, on the other hand, $F[a] \in V_N \backslash V_N{}^0$, the tangent space suffers a loss of dimension, as was pointed out by Kammler [6]. This corresponds to the mentioned gap between necessary and sufficient conditions in Theorem 6.2. In this section we will overcome these difficulties by introducing a convex cone which will be denoted as tangent cone. A γ-polynomial will be shown to be an l.b.a., if and only if zero is a best approximation in the tangent cone. Hence, l.b.a.'s are tangentially characterizable (if the terminology of Wulbert [11] is extended in an appropriate way). Moreover, an l.b.a. may be identified by an alternant condition.

We confine ourselves to the study of γ-polynomials of maximal order, since the extension to γ-polynomials with lower order requires more intricate considerations.

Let $F[a] \in V_N \backslash V_{N-1}$. To investigate the neighborhood of $F[a]$ we represent the γ-polynomials almost in the form as in the preceding section, but contrary to (11.2) the partial sums with $m_\nu = 1$ and $m_\nu \geqslant 2$ are separated:

$$F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=1}^{m_\nu} \alpha_{\nu\mu} \gamma_\mu(t_\nu\,, t_\nu\,, \ldots, t_\nu\,; x) + \sum_{\nu=l_1+1}^{l} \alpha_{\nu 1} \gamma(t_\nu\,, x)$$

[2] This additional hypothesis is not necessary when $N \leqslant 3$.

with

$$t_1 < t_2 < \cdots < t_{l_1}, \qquad \alpha_{\nu m_\nu} \neq 0, \quad \nu = 1, 2,..., l. \tag{12.1}$$

In particular, $l_1 = l_1(a)$ denotes the number of characteristic numbers t_ν associated with a multiplicity $m_\nu \geqslant 2$. (The possibility that l_1 or $l - l_1$ is equal zero is not excluded at this stage. As usual, sums are considered as not written if the indexing set is void.)

In addition to $l(a)$ and $k(a)$ some significant parameters are defined in (12.2) for $F[a]$. The most important of them is $L = L(a)$.

$$\begin{aligned}
&\sigma_\nu = \operatorname{sign} \alpha_{\nu m_\nu}, \qquad \nu = 1, 2,..., l_1,\\
&r_\nu = \begin{cases} 1, & \text{if } \sigma_\nu \sigma_{\nu+1} \cdot (-1)^{m_{\nu+1}} < 0, \\ 0, & \text{otherwise,} \end{cases} \qquad \nu = 1, 2,..., l_1 - 1,\\
&r_{l_1} = 0,\\
&K = l + l_1, \qquad L = K - \sum_{\nu=1}^{l_1} r_\nu .
\end{aligned} \tag{12.2}$$

EXAMPLE. Let $\gamma(t, x) = e^{tx}$ and $F(a, x) = (-x + 5)\, e^{-3x} - x^2 e^{2x} + 4e^x$. Then we have $l = 3$, $k = 6$, $r_1 = 1$, $r_2 = 0$, $K = 5$, $L = 4$.

Note that σ_ν refers to the leading sign of the νth partial sum. Obviously, we have $l \leqslant L \leqslant K \leqslant k = N$. This corresponds to our intention to fill the gap of the criterion when $l(a) < k(a)$ holds. We introduce a convex cone $W(a)$ in V_{N+K}.

DEFINITION 12.1. Let $F[a] \in V_N \backslash V_{N-1}$. Then

$$W(a) = \Big\{ h \in V_{N+K} ; h = \sum_{\nu=1}^{l} \sum_{\mu=1}^{m_\nu^*} \delta_{\nu\mu} \gamma_\mu(t_\nu ,..., t_\nu ; x)$$

$$\delta_{\nu\mu} \in \mathbb{R},\ \sigma_\nu \cdot \delta_{\nu m_\nu^*} \geqslant 0 \text{ for } \nu = 1, 2,..., l_1 \Big\}, \tag{12.3}$$

where

$$m_\nu^* = \begin{cases} m_\nu + 2, & \nu = 1, 2,..., l_1, \\ 2, & \nu > l_1 . \end{cases}$$

is called the tangent cone at $F[a]$ to V_N.

In the remainder of this section we restrict our attention to elements $F[a] \in V_N \backslash V_N^0$, i.e., to γ-polynomials with a degenerate spectrum, since otherwise the tangent cone coincides with the (linear) tangent space. In addition, we will assume that $\gamma(t, x)$ is an extended totally positive kernel, though the main results may be modified so as to include extended sign-regular kernels without major difficulties.

Since the tangent cone contains γ-polynomials, we may consider the associated generalized signs (c.f. Definition 3.1). The constraints of (12.3) imply a bound for the number of sign changes.

LEMMA 12.1. *Let $h \in W(a)$ be a γ-polynomial of order p, and let $(s_1, s_2, \ldots, s_p)$ be the sign vector assigned to h. Then there are at most $N + L - 1$ sign changes in the sequence $s_1, s_2, \ldots, s_p$. If the number of sign changes equals $N + L - 1$, then $s_p = \sigma_{l_1}$ holds.*

The inductive proof is left to the reader. (Observe that the signs assigned to the terms with index $\nu > l_1$ appear in pairs and have no influence on the deficiency of sign changes.)

To prepare the main result we characterize best approximations in the tangent cone.[3]

LEMMA 12.2. *Let γ be an extended totally positive kernel of order $2N$ and let $g \in C(X)$. Assume that $F[a] \in V_N \backslash \{V_N{}^0 \cup V_{N-1}\}$.*

(i) *Zero is the best approximation to g in $W(a)$, if and only if there is an alternant of length $N + L$ with* sign *$-\sigma_{l_1}$ on the right.*

(ii) *If zero is a best approximation to g in $W(a)$, then there is a constant $c_{12} > 0$ such that for each $h \in W(a)$*

$$\| g - h \| \geqslant \| g \| + c_{12} \cdot \| h \|. \tag{12.4}$$

Proof. We may assume $g \not\equiv 0$, since there is nothing to prove for $g \equiv 0$.

Let zero be a best approximation to g in $W(a)$. Assume that there is no alternant of length $N + L$ with sign $-\sigma_{l_1}$ on the right, in particular, there is no alternant of length $N + L + 1$. On the other hand, since $W(a)$ contains a Haar subspace of dimension $N + l$, there is an alternant of exact length $N + l + p$ with $p \geqslant 1$. We select a subset I of p elements from the set of integers $\{\nu; 1 \leqslant \nu \leqslant l_1, r_\nu = 0\}$ by the following procedure. If the alternant has the sign $-\sigma_{l_1}$ on the right, we remove the greatest integer from the given set. The subset which contains the p greatest integers of the (reduced) set is denoted by I. Since zero is not the best approximation in an arbitrary $N + l + p$ dimensional subspace satisfying the Haar condition, there is a function

$$h(x) = \sum_{\nu=1}^{l} \sum_{\mu=1}^{\bar{m}_\nu} \delta_{\nu\mu} \gamma_\mu(t_\nu, t_\nu, \ldots, t_\nu; x),$$

[3] Uniqueness of best Chebyshev approximation may be proved for more general convex sets. Let V be a finite dimensional subspace of $C(X)$ and let $\varphi_1, \varphi_2, \ldots, \varphi_m$ be linear functionals with domain V. Suppose that for any subset I of numbers from 1 to m the subspaces $\{h \in V; \varphi_\nu(h) = 0, \nu \in I\}$ satisfy the Haar condition. Then the best approximation to each $f \in C(X)$ in $W = \{h \in V; \varphi_\nu(h) \geqslant 0, \nu = 1, 2, \ldots, m\}$ is unique.

where

$$\tilde{m}_\nu = \begin{cases} m_\nu + 2, & \text{if } \nu \in I, \\ m_\nu + 1, & \text{otherwise,} \end{cases}$$

such that $\| g - h \| < \| g \|$ holds. By standard arguments $h(x)$ has $N + l + p - 1$ zeros and the generalized signs of h alternate. Moreover, the signs are fixed by the sign of the alternant. This implies sign $\delta_{\nu m_\nu+2} = \sigma_\nu$ for $\nu \in I$ and $h \in W(a)$, contradicting the optimality of the zero function. This proves the existence of an alternant as stated.

To prove the converse, assume that there is an alternant of length $N + L$ with the postulated sign, i.e., there are points $x_1 < x_2 < \cdots < x_{N+L}$ such that

$$(-1)^{N+L-i+1} \sigma_{l_1} \cdot w(x_i) = \| g \|, \qquad i = 1, 2,..., N + L.$$

From Lemma 12.1 and Theorem 3.2 we know that the inequalities

$$g(x_i) \cdot h(x_i) \geqslant 0, \qquad i = 1, 2,..., N + L$$

are violated by each $h \in W(a)$, $h \not\equiv 0$. Hence,

$$\max_{1 \leqslant i \leqslant N+L} \{-g(x_i)\, h(x_i)\} > 0.$$

By compactness arguments we have

$$\inf_{\substack{h \in W(a) \\ \|h\|=1}} \max_{1 \leqslant i \leqslant N+L} \{-w(x_i)\, g(x_i) \cdot h(x_i)\} = c_{12} \cdot \| g \| > 0, \tag{12.6}$$

where c_{12} is a positive constant. From this we obtain for each $h \in W(a)$,

$$\begin{aligned} \| g - h \| &\geqslant \max_{1 \leqslant i \leqslant N+L} w(x_i) \mid g(x_i) - h(x_i) | \\ &= \max_{1 \leqslant i \leqslant N+L} \{\| g \| - (g(x_i)/\| g \|)\, w(x_i) \cdot h(x_i)\} \\ &\geqslant \| g \| + 1/\| g \| \cdot c_{12} \cdot \| g \| \cdot \| h \| = \| g \| + c_{12} \cdot \| h \|. \end{aligned}$$

Hence, zero is a best approximation to g in $W(a)$ and the strong uniqueness condition (ii) holds. ∎

Now we are ready to identify an l.b.a. by an alternant and to establish the tangential characterization.

THEOREM 12.3. *Let V_N be a normal family with an extended totally positive kernel. If $F(a) \in V_N \backslash \{V_N{}^0 \cup V_{N-1}\}$ and $f \in C(X)$, the following properties are equivalent.*

(i) *$F[a]$ is a local best approximation to f in V_N.*

(ii) *There is an alternant to $f - F(a)$ of length $N + L(a)$ with sign $-\sigma_{l_1}$ on the right.*

(iii) *Zero is the best approximation to $f - F(a)$ in $W(a)$.*

Proof. (iii) $\Rightarrow$ (i). Set $I = \{1, 2,..., l_1\}$. From the proof of Theorem 11.2 we know that there is a continuous mapping Φ from a neighborhood of $F(a)$ in V_N into V_{N+K}. By virtue of the statement on the leading signs in Lemma 10.2, the range of Φ is contained in $W(a)$, i.e., we have

$$\Phi(F[b]) \in W(a),$$
$$\| F[b] - F[a] - \Phi(F[b])\| \leqslant c_{10} \| \Phi(F[b])\|^{3/2}.$$

Set $h = \Phi(F[b])$. If $\| h \| \leqslant (2c_{10})^{-2}$, the triangle inequality implies

$$\| h \| \geqslant \tfrac{1}{2} \| F[b] - F[a]\|. \tag{12.7}$$

By virtue of the preceding lemma, a strong uniqueness condition holds in $W(a)$. Let c_{12} be the attributed constant. Hence,

$$\begin{aligned} \|f - F[b]\| &\geqslant \|f - F[a] + h\| - c_{10} \| h \|^{3/2} \\ &\geqslant \|f - F[a]\| + c_{12} \| h \| - c_{10} \| h \|^{3/2} \\ &\geqslant \|f - F[a]\| + \tfrac{1}{2} c_{12} \| h \| \\ &\geqslant \|f - F[a]\|, \end{aligned} \tag{12.8}$$

provided that $\| h \| < \frac{1}{4}(c_{12}/c_{10})^2$. Consequently, $F[a]$ is an l.b.a.

(i) $\Rightarrow$ (ii). Assume that $F[a]$ is an l.b.a. The partial sums (11.4) will be represented in an appropriate form. Let $m \geqslant 2$ and $u \geqslant 0$. The functions in $V_m[t, t,..., t, t + u^{1/2}, t - u^{1/2}]$ may be written in the form

$$\begin{aligned} F(x) = &\sum_{\mu=1}^{m-2} \beta_\mu \gamma_\mu(t,..., t; x) \\ &+ \tfrac{1}{2}\beta_{m-1}[\gamma_{m-1}(t,..., t, t + u^{1/2}; x) + \gamma_{m-1}(t,..., t, t - u^{1/2}; x)] \\ &+ \beta_m \gamma_m(t,..., t, t + u^{1/2}, t - u^{1/2}; x). \end{aligned} \tag{12.9}$$

This is obvious if $u = 0$ holds and is easily verified for $u > 0$. Observe that a special situation is given in (7.3). The derivatives at $u = 0$ are

$$\partial/\partial\beta_\mu F = \gamma_\mu, \qquad \mu = 1, 2,..., m,$$
$$\partial/\partial t\, F = \sum_{\mu=1}^{m} \mu \cdot \beta_\mu \gamma_{\mu+1},$$
$$\partial/\partial u\, F = \beta_{m-1}\gamma_{m+1} + \beta_m \gamma_{m+2}.$$

All functions of the form

$$\sum_{\rho=1}^{m+2} e_\rho \gamma_\rho(t, t, \ldots, t; x)$$

with e_{m+2} restricted by $e_{m+2} \cdot \beta_m \geqslant 0$ may be written as linear combinations of derivatives

$$\sum_{\mu=1}^{m} \delta_\mu \frac{\partial F}{\partial \beta_\mu} + \delta \frac{\partial F}{\partial t} + \theta \frac{\partial F}{\partial u}$$

with $\theta \geqslant 0$. A simple comparison of coefficients establishes that θ, δ, δ_m, $\delta_{m-1}, \ldots, \delta_1$ can be determined successively such that the wanted function is generated.

To continue the proof we return to the γ-polynomials $F[b]$ lying in the neighborhood of $F[a]$. Starting from the representation (1.2), in addition to $\alpha_{\nu\mu}$ and t_ν, the parameters u_ν are introduced in terms with $m_\nu \geqslant 2$. This means that the partial sums are written in the form given by (12.9). It follows from the preceding discussion that all elements of the tangent cone $W(a)$ may be generated as linear combinations of derivatives with respect to these parameters taking sign restrictions into account. Set

$$M^* = \{x \in X; w(x) \cdot |f(x) - F(a, x)| = \|f - F[a]\|\}.$$

By virtue of Lemma 7.1 which may be applied to l.b.a.'s as well, we obtain

$$\min_{x \in M^*} (f(x) - F(a, x)) \cdot h(x) \leqslant 0$$

for every $h \in W(a)$. Hence, from the Kolmogorov criterion we know that zero is optimal to $f - F[a]$ in $W(a)$.

(ii) $\Rightarrow$ (iii) is already established in the preceding lemma. ▮

The proof of Theorem 12.3, in particular inequalities (12.7) and (12.8), yield a local strong uniqueness condition. This condition also holds for γ-polynomials in $V_N^0 \backslash V_{N-1}^0$ as may be proved with Lemma 9 in [10].

COROLLARY 12.4. *Let V_N be a normal family and $f \in C(X)$. Then every local best approximation in $V_N \backslash V_{N-1}$ satisfies a local strong uniqueness condition.*

The local strong uniqueness condition implies that the l.b.a. changes only slightly if the function f is altered slightly. This can be seen more precisely from the following theorem.

THEOREM 12.5. *Let V_N be a normal family and $f \in C(X)$. Assume that $F[a] \in V_N \backslash V_{N-1}$ is a local best approximation to f. Then there is a constant c_{13} and a neighborhood $\mathcal{U}$ of f in $C(X)$, such that for each $g \in \mathcal{U}$ there is a local best approximation $F[b]$ in V_N satisfying*

$$\| F[b] - F[a] \| \leqslant c_{13} \cdot \| g - f \|.$$

Proof. By virtue of Corollary 12.4 there are constants r, $c_{14} > 0$ such that

$$\| f - F[b] \| \geqslant \| f - F[a] \| + c_{14} \| F[b] - F[a] \| \tag{12.11}$$

provided $\| F[b] - F[a] \| \leqslant r$. We may assume (after reducing r if necessary) that the set

$$V = \{F[b] \in V_N \,;\, \| F[b] - F[a] \| \leqslant r\}$$

does not intersect V_{N-1}. Since V_N is a normal family, V is compact. Assume that g satisfies $\| g - f \| < \frac{1}{2} \cdot c_{14} \cdot r$. Let $F[b^*]$ be a best approximation to g in V. Thus,

$$\| g - F[b^*] \| \leqslant \| g - F[a] \| \leqslant \| g - f \| + \| f - F[a] \|. \tag{12.12}$$

On the other hand, it follows from (12.11) that

$$\begin{aligned} \| g - F[b^*] \| &\geqslant \| f - F[b^*] \| - \| g - f \| \\ &\geqslant \| f - F[a] \| + c_{14} \| F[b^*] - F[a] \| - \| g - f \|. \end{aligned}$$

Combining the last inequalities we obtain

$$\| F[b^*] - F[a] \| \leqslant 2/c_{14} \cdot \| g - f \|.$$

Hence, $F[b^*]$ does not lie on the boundary of V and is an l.b.a. in V_N. ∎

Note that Theorem 12.5 does not state that there is a unique l.b.a.

13. EXAMPLES FOR UNIQUENESS AND NONUNIQUENESS

In this section two examples are discussed. One of them illustrates many pathological features while the other shows an extremely good behavior. Both examples refer to approximation by exponentials.

EXAMPLE 13.1. Let us consider the approximation of $f(x) = \cos \pi/2 \cdot x$ in the interval $[-1, +1]$ by exponentials. As was pointed out by Kammler [6], $f(x) - F(a, x)$ has at most $k(a) + 1$ zeros and, therefore, for $N \geqslant 2$, each stationary element in V_N has the maximal order and a maximally degenerate spectrum, i.e., we have $k(a) = N$ and $l(a) = 1$. Hence, stationary elements can be written in the form

$$F(a, x) = e^{tx} \cdot p(x), \tag{13.1}$$

where $p(x)$ is a polynomial of exact degree $N - 1$.

At first we claim that every function in a suitable neighborhood of f in $C(X)$ has a best approximation $F[b]$ in V_N with $l(b) = 1$. Suppose to the contrary that there is a sequence $\{f_n\}$ which converges to $\cos \pi/2 \cdot x$, such that f_n has a best approximation $F[a_n]$ in V_N satisfying $l(a_n) \geqslant 2$. The sequence $\{F[a_n]\}$ is a minimal sequence with respect to $f(x) = \cos \pi/2 \cdot x$. By virtue of Korollar 1 and Satz 4 in [9] a subsequence of $\{F[a_n]\}$ converges uniformly to a best approximation $F[a^*]$ of $\cos \pi/2 \cdot x$. From this, by the same argument as that used in the proof of Lemma 11.1 we conclude that there is only an alternant of length $N + 2$ to $f_n - F[a_n]$ for sufficiently large n. This contradicts $l(a_n) \geqslant 2$.

Consequently, the set of functions in $C(X)$ having a best approximation in $V_N{}^0$ is not dense in $C(X)$, although $V_N{}^0$ is dense in V_N. We emphasize that for this reason we must not restrict our attention to the approximation in $V_N{}^0$, though the restriction would help to avoid many difficulties.

Next we prove that there are at least two best approximations in V_N, provided N is even.[4] Let $F(a^*, x) = e^{\lambda x} \cdot p(x)$ be a best approximation. Since the degree of the polynomial $p(x)$ is odd, $F(a^{**}, x) = F(a^*, -x)$ is a different γ-polynomial. From $f(x) = f(-x)$ we conclude that $F(a^{**})$ is another best approximation.

Finally we verify that one sign class of V_3 contains at least two l.b.a.'s to $\cos \pi/2 \cdot x$ in V_3. It is sufficient to show that there is an l.b.a. $F[a_1] \in V_3(+--+)$ and an l.b.a. $F[a_2] \in V_3(-+-)$. Since the approximation problem with fixed characteristic numbers has a unique solution, at least one of them has a characteristic number $t \neq 0$. Actually, this is true for $F[a_1]$. Hence, $F(a_3, x) = F(a_1, -x)$ defines another l.b.a.

To prove the statement we start from the best approximation $F(\hat{a}, x) = e^{\tau x}(\alpha_1 + \alpha_2 x)$ with $\alpha_2 > 0$ to $f(x)$ in V_2. Choose $t_2 < \tau$ and consider the exponentials of the form $(\beta_1 + \beta_2 x)\, e^{t_1 x} + \beta_3 e^{t_2 x}$ with $t_1 > t_2$. Since the derivatives $\partial F/\partial t_1$, $\partial F/\partial \beta_\nu$, $\nu = 1, 2, 3$ span a Haar subspace and the alternant to $f - F[\hat{a}]$ has length 4, the γ-polynomial $F[\hat{a}]$ is not an l.b.a. in this subset and there is a better approximation $F[b_0]$ with parameter $\beta_2 > 0$. Hence, $F[b_0] \in V_3(s_1, -+)$, where s_1 will be specified later. Denote the best approximation to f in $V_3(s_1, -+) \cup V_2$, which is an existence set, by $F[a_1]$. By

[4] The reader may prove by the same arguments that $f(x) = \sin \pi/2 \cdot x$ has at least two best approximations in V_N, provided N is odd and $N \geqslant 3$.

construction we have $F[a_1] \notin V_2$ and $F[a_1]$ is an l.b.a. in V_3. From (13.1) we conclude $s_1 = +$. If we repeat this procedure with $t_2 > \tau$ we obtain an l.b.a. in $V_3(-+-)$.

After the presentation of the example with bad behavior we discuss a function with the opposite features.

EXAMPLE 13.2. Let us consider the approximation of $f(x) = (1 + x)^{-1}$ in the interval $[0, 1]$. By multiplying $f(x) - F(a, x)$ with $(1 + x)$ we get a γ-polynomial of order $k(a) + l(a) + 1$ and we conclude that the difference has at most $k(a) + l(a)$ zeros. Hence, each stationary element in V_N has the degree $k(a) = N$. We prove by induction that the best approximation to f in V_N belongs to $V_N^+ \backslash V_{N-1}^+$.

This is obvious for $N = 1$. Let us assume that it is true for N. Since $f(x) - F(a, x)$ has $2N$ zeros, we have $(1 + x)[f(x) - F(a, x)] = 1 - (1 + x) F(a, x) \in V_{2N+1}(+, -, +, \ldots, -+)$. By virtue of Theorem 4.5 the best approximation V_{N+1} satisfies $k^+ = N + 1$.

In addition, the best approximation to $(1 + x)^{-1}$ in V_N is unique because we have uniqueness in V_N^+. As was proven in [3] by some simple arguments, the spectra of the best approximations are not bounded when N tends to infinity.

Note added in proof. In Section 12 the characterization of local best approximations was performed by a direct approach. Referring to the author's recent article: "Kritische Punkte bei der nichtlinearen Tschebyscheff-Approximation," *Math. Z.* **132** (1973), 327–341, this approach may be understood in the more general framework of critical point theory. Then the introduction of the tangent cone is also motivated from the point of view of differential topology.

ACKNOWLEDGMENT

The author wishes to thank Professor W. Jäger for many stimulating discussions.

REFERENCES

1. D. BRAESS, Chebyshev approximation by γ-polynomials, *J. Approximation Theory* **9** (1973), 20–43.
2. D. BRAESS, Über die Approximation mit Exponentialsummen, *Computing* **2** (1967), 309–321.
3. D. BRAESS, Die Konstruktion der Tschebyscheff-Approximierenden bei der Anpassung mit Exponentialsummen, *J. Approximation Theory* **3** (1970), 261–273.
4. E. W. CHENEY, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
5. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rat. Mech. Anal.* **27** (1967), 91–106.

6. D. KAMMLER, Characterization of best approximations by sums of exponentials, *J. Approximation Theory* **9** (1973), 173–191.
7. S. KARLIN, On a class of best nonlinear approximation problems, *Bull. Amer. Math. Soc.* **78** (1972), 43–49.
8. G. MEINARDUS AND D. SCHWEDT, Nichtlineare Approximation, *Arch. Rat. Mech. Anal.* **17** (1964), 297–326.
9. E. SCHMIDT, Zur Kompaktheit bei Exponentialsummen, *J. Approximation Theory* **3** (1970), 445–454.
10. D. WULBERT, Uniqueness and differential characterization of approximations from manifolds of functions, *Amer. J. Math.* **93** (1971), 350–366.
11. D. WULBERT, Nonlinear approximation with tangential characterization, *Amer. J. Math.* **93** (1971), 718–730.