JOURNAL OF APPROXIMATION THEORY **24**, 119–145 (1978)

Chebyshev Approximation by γ-Polynomials. III. On the Number of Best Approximations

DIETRICH BRAESS

Institut für Mathematik, Ruhr-Universität, D-4630 Bochum, Federal Republic of Germany

Communicated by E. W. Cheney

Received February 28, 1977

Es wird gezeigt, daß bei der Tschebyscheff-Approximation durch γ-Polynome der Ordnung k höchstens $k!$ lokal beste Approximationen auftreten. Damit wird insbesondere das Problem der Mehrdeutigkeit bei der Exponentialapproximation bewältigt. Für die Untersuchung muß das Konzept generischer Eigenschaften in der nichtlinearen Approximationstheorie eingeführt werden, um Morsetheorie anwenden zu können. Die Situation ist völlig anders als bei der L_2-Approximation mit Exponentialsummen, da es dort keine Schranke für die Zahl der lokalen Lösungen geben kann.

INTRODUCTION

In 1967 Hobby and Rice [12] introduced the idea of γ-polynomials. Their concept provided a natural generalization of approximation by sums of exponentials. In the same year it was shown that best uniform approximation is not always unique [4]. Therefore the question arose of whether the number of solutions is always finite and whether there is a finite bound:

$$c_k = \sup_{f \in C(X)} \{\text{number of local best approximations to } f \text{ in } V_k\}. \quad (14.1)$$

Here we have already taken into account that it is mathematically more elegant and gives more insight to consider not only the global solutions but also the local ones.

The problem was settled only for the simplest nontrivial cases [3]. Besides $c_1 = 1$ we have $c_2 = 2$. More recently the rough estimate $c_3 \leqslant 9$ was presented in [9].

In 1973 the author [7] anounced the result that $c_k \leqslant k!$ for all k. Unfortunately, a serious gap was later detected in the proof. It became apparent that it is impossible to neglect certain cases of degeneracy. Only now, 3 years later, we have succeeded in bridging the gap by using the concept of generic properties. When the standard construction for all local solutions is

0021-9045/78/0242–0119$02.00/0

developed, certain exceptional cases are ignored. Nevertheless, the situation is more advantageous than most in topology. Here the final result is extended also to the exceptional functions. With the aid of perturbation techniques it is proved that the bound holds in all cases.

Though the rigorous treatment of the problem must be done in a very abstract setting, the basic idea may be better described from the numerical viewpoint. When one intends to compute a best approximation and applies Newton's method, then a sequence is generated which in general will converge only to a local best approximation. Which one of the (possibly many) local solutions is found, greatly depends on the starting point of the iteration. The following question is the key to the solution: Is it possible to characterize a set of starting points from which all solutions are reached, when applying Newton-like algorithms?

The continuous analog of Newton's method is just what we need when applying critical point theory. To be more specific, we use the introductory part of Morse theory as in [8]. This is possible because the manifolds under consideration are trivial from the homotopical point of view. On the other hand the theory is by no means trivial, because the manifolds are not compact. The lack of compactness is the reason for the complexity of the analysis.

The results show once more that uniform approximation is always something special in nonlinear approximation theory. As was shown by Wolfe [18], one has $c_k = \infty$ even for $k = 1$ when the approximation problem is considered in the L_2-case.

This paper is a continuation of the author's two papers on γ-polynomials [6]*). Therefore, we proceed with enumerating formulas and theorems. All references to Eqs. (1.1)–(13.1) and Theorems 2.1–12.5 refer to those papers.

The standard notation for a γ-polynomial is

$$F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=0}^{m_\nu - 1} \alpha_{\nu\mu} \gamma^{(\mu)}(t_\nu, x), \qquad \alpha_{\nu m_\nu} \neq 0. \tag{14.2}$$

When the family V_N of γ-polynomials of order $\leqslant N$ is considered, the kernel γ is assumed to be extended sign-regular of order $2N$, where the extension is in the t-variable [14]. In particular, this means that the derivatives $\gamma^{(\mu)} = \partial^\mu/\partial t^\mu \gamma$ exist for $\mu \leqslant 2N - 1$.

In the interest of simplicity some additional assumptions are made. They are natural, and have been verified for the interesting families of functions, e.g., for the sums of exponentials they are consequences of Schmidt's compactness results [16]. Specifically we will assume that T, the domain of the characteristic numbers $t_1, t_2, \ldots, t_N$, is an open connected subset of $\mathbb{R}$.

* We want to indicate that part II of the paper contains a serious misprint. Please start reading on the top of page 17 until the end of the third paragraph. Then proceed with page 16 and continue on page 17 with the fourth paragraph.

Then without loss of generality T may be identified with $\mathbb{R}$. Moreover, V_N is assumed to be normal in the sense of Section 8, i.e., for each $F_0 \in V_N \backslash V_{N-1}$ there is a neighborhood $U(F_0)$ such that the spectrum (= set of characteristic numbers) of all γ-polynomials in $U(F_0)$ belongs to a compact subset of T. Finally we assume that each sequence of γ-polynomials which is bounded in $C(X)$ contains a subsequence which converges to a γ-polynomial in the topology of compact convergence on X. Then each closed and bounded set in $V_N \backslash V_{N-1}$ is compact in the norm topology.

The consequences of our theory for the numerical solution of the approximation problem are obvious. In particular the difficulties in the classical algorithms for treating spline functions with free knots [2] may be overcome by the regularization procedure [11] which yields the connection with γ-polynomials.

15. A Partial Uniqueness Result

In Section 12 local best approximations (for short: LBAs) were characterized in terms of alternants. The criteria provide conditions which are both necessary and sufficient. In the framework of critical point theory [8] the characterization theorem (Theorem 12.3) is recognized as a consequence of the fact that $V_N \backslash V_{N-1}$ is a Haar embedded manifold.

The theory of Haar embedded manifolds is central to the present paper. In particular we apply the Nonzero Index Theorem which was derived in [8] with methods from global analysis. To this end we have to modify slightly the parametrization used in Section 12.

Definition 15.1. A subset G of a normed linear space E is called a C^1-manifold (with boundary), if for every $F_0 \in G$ there is a neighborhood $U \subset G$ with the following properties:

(i) There is a closed convex set $C \subset \mathbb{R}^n$ and a homeomorphism $g : W \to U$ with W relatively open in C. (g will be called centered at F_0 if $g^{-1}(F_0) = 0$.)

(ii) g is a Fréchet differentiable map and the derivative $d_a g$ is continuous in a.

(iii) There is a continuous mapping

$$\kappa : U \to d_{a_0} g \left(\bigcup_{\lambda > 0} \lambda C \right)$$

with $F_0 = g(a_0)$, satisfying

$$\begin{aligned} \kappa(F_0) &= 0, \\ \| F - F_0 - \kappa(F) \| &= o(\| \kappa(F) \|). \end{aligned} \tag{15.1}$$

DEFINITION 15.2. $h \in E$ is a tangent ray at F to $G \subset E$, if there is a continuous mapping from [0, 1] to G which sends λ to $F_\lambda \in G$ such that

$$\| F_\lambda - F - \lambda h \| = o(\lambda), \qquad \text{as } \lambda \to 0. \tag{15.2}$$

The set of all tangent rays at F is called the tangent cone and is denoted by $C_F G$.

For C^1-manifolds the tangent cone can be easily calculated since

$$C_F G = d_a g\left(\overline{\bigcup_{\lambda>0} \lambda C}\right) \tag{15.3}$$

if $F = g(a)$. We will verify that the tangent cone for an element $F(a)$ in V_N / V_{N-1} is just the cone $W(a)$ explicitly given in Definition 12.1.

To develop a suitable parameterization for V_N we consider first the neighborhood of a γ-polynomial with only one characteristic number t. Let its multiplicity be $m \geqslant 2$. Define the mapping $g : A \to V_n$, where

$$\begin{aligned} g(\beta_1, ..., \beta_m, t_1, t_2, ..., t_{m-2}, u, v) = \sum_{\mu=1}^{m-2} & \beta_\mu \gamma_\mu(t_1, t_2, ..., t_\mu; x) \\ & + \tfrac{1}{2}\beta_{m-1}[\gamma_{m-1}(t_1, ..., t_{m-2}, u + v^{1/2}; x) \\ & + \gamma_{m-1}(t_1, ..., t_{m-2}, u - v^{1/2}; x)] \\ & + \beta_m \gamma_m(t_1, ..., t_{m-2}, u + v^{1/2}, u - v^{1/2}; x) \end{aligned} \tag{15.4}$$

The domain

$$\begin{aligned} A = \{(\beta_1, \beta_2, ..., \beta_m, t_1, ..., t_{m-2}, u, v) \in \mathbb{R}^{2m}; \\ v \geqslant 0, u - v^{1/2} \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_{m-2} \leqslant u + v^{1/2}\} \end{aligned} \tag{15.5}$$

is a convex set. Indeed, the inequalities are equivalent to

$$\begin{aligned} -v &\leqslant 0, \\ u - v^{1/2} - t_1 &\leqslant 0, \\ t_\mu - t_{\mu+1} &\leqslant 0, \qquad \mu = 1, 2, ..., m-3, \\ t_{m-2} - u - v^{1/2} &\leqslant 0, \end{aligned}$$

and the left-hand sides are convex functions of the arguments.

Note that the use of square roots does not spoil the differentiability of the representation.

This follows from

$$\begin{aligned}
&\frac{\partial}{\partial v}\,[\gamma_\mu(\ldots, u + v^{1/2},\ldots; x) + \gamma_\mu(\ldots, u - v^{1/2},\ldots; x)]\\
&\qquad = \gamma_{\mu+2}(\ldots, u + v^{1/2}, u + v^{1/2}, u - v^{1/2},\ldots; x)\\
&\qquad\quad + \gamma_{\mu+2}(\ldots, u + v^{1/2}, u - v^{1/2}, u - v^{1/2},\ldots; x),
\end{aligned} \tag{15.6}$$

$$\begin{aligned}
&\frac{\partial}{\partial v}\,\gamma_\mu(\ldots, u + v^{1/2}, u - v^{1/2},\ldots; x)\\
&\qquad = \gamma_{\mu+2}(\ldots, u + v^{1/2}, u + v^{1/2}, u - v^{1/2}, u - v^{1/2},\ldots; x).
\end{aligned}$$

In the particular case when $m = 2$ and $\gamma(t, x) = e^{tx}$, the representation (15.4) reads (cf. Section 7):

$$g(\beta_1, \beta_2, u, v) = \beta_1 e^{ux} \cosh v^{1/2}x + \beta_2 e^{ux}(\sinh v^{1/2}x)/v^{1/2}.$$

Here only entire functions are involved.

The mapping (15.4) obviously satisfies conditions (i) and (ii) of Definition 15.1. To prove the third condition let a be a parameter corresponding to a γ-polynomial with m coalescing characteristic numbers, i.e.,

$$t_1 = t_2 = \cdots = t_{m-2} = u = \tau, \qquad v = 0.$$

Then the set of tangent vectors to A at a is

$$\begin{aligned}
C_a A &= \overline{\bigcup_{\lambda>0} \lambda(A - a)}\\
&= \{(\delta_1, \delta_2,\ldots, \delta_m, \eta_1,\ldots, \eta_m) \in \mathbb{R}^{2m};\\
&\qquad \eta_m \geqslant 0, \eta_\mu - \eta_{\mu+1} \leqslant 0, \mu = 1, 2,\ldots, m-3\}.
\end{aligned} \tag{15.7}$$

Referring to the calculations in the last part of the proof of Theorem 12.3, we obtain

$$\begin{aligned}
d_a g(C_a A) &= \left\{\sum_{\mu=1}^{m} \tilde{\delta}_\mu \frac{\partial g}{\partial \beta_\mu} + \delta \frac{\partial g}{\partial \mu} + \theta \frac{\partial g}{\partial v}\,;\ \tilde{\delta}_1, \tilde{\delta}_2,\ldots, \tilde{\delta}_m, \delta \in \mathbb{R}, \theta \geqslant 0\right\}\\
&= \left\{\sum_{\mu=1}^{m+2} \delta_\mu \gamma_\mu(\tau,\ldots, \tau; x);\ \delta_\mu \in \mathbb{R}, \delta_{m+2} \cdot \beta_m \geqslant 0\right\}
\end{aligned} \tag{15.8}$$

Now, by Lemma 10.2, a map κ is established with the properties postulated in Definition 15.1. From (15.3) we know that the set in (15.8) is the tangent cone $C_{g(a)} V_m$.

Next we turn to γ-polynomials with more than one distinct characteristic number. Given $F[a] \in V_N \backslash V_{N-1}$ we may apply the parameterization according to (15.4) to each of the l partial sums in the standard form (14.2). From the arguments given in the proof of Theorem 11.2 it follows that a suitable parameterization for an open neighborhood of $F[a]$ is established. Moreover, we recognize that the set $W(a)$ from Definition 12.1 is just the tangent cone $C_{F(a)} V_N$.

The preceding discussion yields an atlas of $V_N \backslash V_{N-1}$ consistent with Definition 15.1. In fact $V_N \backslash V_{N-1}$ is a Haar embedded manifold, because the tangent cones are convex and have the Haar property [8]:

DEFINITION 15.3. Let $v_1, v_2, \ldots, v_n \in C(X)$ and $m \leqslant n$. The convex cone

$$\left\{ u;\ u(x) = \sum_{i=1}^{m} \alpha_i v_i(x);\ \alpha_j \in \mathbb{R} \text{ for } i = 1, 2, \ldots, m,\ \alpha_i > 0 \text{ for } i = m+1, \ldots, n \right\}$$

has the Haar property, if the functions $\{v_i\}_{i \in I}$ span a Haar subspace, whenever

$$\{1, 2, \ldots, m\} \subset I \subset \{1, 2, \ldots, n\}.$$

If each tangent cone to a C^1-manifold G has the Haar property, then G is called a Haar embedded manifold.

DEFINITION 15.4. F is called a critical point to f in G if O is a best approximation to $(f - F)$ in $C_F G$.

Since $V_N \backslash V_{N-1}$ is a Haar embedded manifold, it follows from [8, Theorem 7.1] that F is an LBA to f in $V_N \backslash V_{N-1}$ if and only if F is a critical point. This equivalence has also been derived explicitly for γ-polynomials in Theorem 12.3.

A direct application of the Nonzero Index Theorem from critical point theory is impossible because of the lack of compactness of V_N. It may be applied, however, to get a local result. As usual for any nonnegative real number α put

$$\rho^\alpha = \{F \in V_N;\ \|f - F\| \leqslant \alpha\}. \tag{15.9}$$

THEOREM 15.1. *Let V_N be a normal family and let $f \in C(X)$. Assume that Cp^α is a (connected) component of ρ^α, $\alpha > 0$, which is disjoint from V_{N-1}. Then Cp^α contains exactly one local best approximation.*

Proof. Since Cp^α is bounded, closed, and disjoint from V_{N-1}, it is compact. By applying [8, Satz 7.3] we obtain the theorem. ∎

16. The Maximal Components Associated to Local Best Approximations

Let $F^* \in V_N \backslash V_{N-1}$ be an LBA to f. Denote the component of the level set ρ^α which contains F^* by Cp^α. Further set

$$\beta = \sup\{\alpha; Cp^\alpha \cap V_{n-1} = \varnothing\}.$$

It follows from Theorem 15.1 that F^* is the unique LBA to f in

$$Cp = \bigcup_{\alpha<\beta} Cp^\alpha. \tag{16.1}$$

For this reason we call Cp the associated maximal component. In this section we will prove that the boundary of Cp contains an LBA in V_k, $k \leqslant N-1$. In fact $k = N-1$ holds in most cases (cf. Section 19). From these properties we will obtain a classification of the solutions in V_N in terms of the solutions in V_{N-1}, which finally leads to an enumeration.

Referring to (16.1) we observe that $\beta \leqslant \|f\| = \|f - o\|$, because Cp^α, $\alpha = \|f\|$, contains the γ-polynomials $\lambda \cdot F^*$, $o \leqslant \lambda \leqslant 1$. Furthermore, we claim that $\beta > \|f - F^*\|$. Indeed, from Corollary 12.4 we know that F^* is a strong LBA, i.e., there are numbers $c > 0$, $r > 0$ such that

$$\|f - F\| \geqslant \|f - F^*\| + c\,\|F - F^*\|, \tag{16.2}$$

whenever $F \in V_N$, $\|F - F^*\| < r$. We may assume that r is smaller than the distance of F^* from V_{N-1}. Put $\alpha = \|f - F^*\| + \frac{1}{2}cr$. Hence, ρ^α contains no element of V_N, whose distance from F^* equals r. The component Cp^α of ρ^α containing F^* is disjoint from V_{N-1}.

The next step is the proof of the following:

Assertion 16.1. The closure $\overline{Cp}$ of the maximal component intersects V_{N-1}.

Proof. Suppose to the contrary that the assertion is not true. Then we may construct an extension of Cp which is also disjoint from V_{N-1}. Since $\overline{Cp}$ is bounded and closed, by the normality assumption it is compact. Obviously, we have

$$\|f - F\| = \beta$$

for each F in the boundary $\partial Cp := \overline{Cp} \backslash Cp$. Since $\overline{Cp}$ is connected, ∂Cp contains no strong LBA and therefore no critical point. By [8, Lemma 4.2.], for each $F_0 \in \partial Cp$ there is an open neighborhood $U = U(F_0)$ and a continuous mapping

$$\psi : [0, 1] \times \overline{U} \to V_N$$

such that

$$\begin{aligned} &\psi(0, F) = F, \\ &\|f - \psi(t, F)\| < \|f - \psi(s, F)\|, \qquad F \in U,\ 0 \leqslant s < t \leqslant 1. \end{aligned} \tag{16.3}$$

After reducing U, if necessary, the following properties hold:

(1) $\bar{U} \cap V_{N-1} = \varnothing$,

(2) $\bar{U}$ is compact and connected,

(3) $\|f - \psi(1, F)\| < \beta$ for each $F \subset \bar{U}$.

Note that $F \in Cp$ implies $\psi(t, F) \in Cp$ for $0 < t \leqslant 1$. Since $U \cap Cp$ is not empty, $\psi(1, U)$ intersects Cp. From the connectedness of $\psi(1, \bar{U})$ and (3) we obtain $\psi(1, \bar{U}) \subset Cp$. Consequently,

$$\|f - F\| > \beta \qquad \text{whenever } F \in \overline{U \backslash Cp}. \tag{16.4}$$

Otherwise the orbit $\psi(t, F)$, $0 < t \leqslant 1$, would establish a connecting arc between F and $\psi(1, F)$ which runs below the level β and hence in Cp.

A finite number of such open sets say U_1, U_2,..., U_m cover ∂Cp.

$$U = \bigcup_{j=1}^{m} U_j \supset \partial Cp.$$

(We remark that U is a substitute for a tubular neighborhood of ∂Cp.) The set

$$M = \bar{U} \backslash (U \cup Cp) \tag{16.5}$$

is compact, and the distance function $\|f - F\|$ achieves its minimum at some $F_1 \in M$. From (16.4) we obtain $\|f - F_1\| > \beta$. Since $Cp \cup U$ is connected, this set contains a component of the level set

$$P^{(1/2)(\beta + \|f - F_1\|)}$$

which contradicts the maximality of β. Hence, Cp intersects V_{N-1}. ∎

The investigation of the maximal components requires the handling of sequences $\{F_r\} \subset V_N \backslash V_{N-1}$ which converge to a γ-polynomial $\hat{F}$ with degree $k = k(\hat{F}) \leqslant N - 1$. We will next separate a sequence $\{v_r\}$ from $\{F_r\}$ such that $k(v_r) = k$ and $v_r \to \hat{F}$. Unfortunately, this cannot be done by a simple splitting with the complement satisfying $k(F_r - v_r) = N - k$. This is illustrated by the sequence in V_2 having the elements

$$F_r = 5\gamma(1, x) - 2\gamma(1 + 1/r, x), \qquad r = 1, 2, \ldots,$$

which converges to $3\gamma(1, x) \in V_1$.

The appropriate splitting process will be performed in two steps. Let $\{\lambda_\mu^{(r)}, \mu = 1, 2,..., N\}$ be the spectrum of F_r. We may relabel the characteristic numbers such that after passing to a subsequence the following properties hold with some $j \leqslant N$:

(1) For $\mu \leqslant j$ the limits

$$\lambda_\mu^* = \lim_{r\to\infty} \lambda_\mu^{(r)}$$

exist and belong to spect($\hat{F}$).

(2) For $\mu > j$, the sequence $\{\lambda_\mu^{(r)}\}$ has at most one accumulation point and this one is disjoint from spect($\hat{F}$).

Choose an open subset $T_0 \subset T$ with compact closure, which contains spect($\hat{F}$) but no accumulation point of $\{\lambda_\mu^{(r)}\}$, $\mu > j$. Divide each F_r into two parts

$$F_r = \tilde{v}_r + u_r\,, \qquad \text{spect}(\tilde{v}_r) \subset T_0\,, \text{spect}(u_r) \subset T\backslash T_0\,.$$

We claim that $\{\tilde{v}_r\}$ is bounded. If this is not true, then by passing to a subsequence we have $\| \tilde{v}_r \| \to \infty$. Since $\tilde{v}_r/\| \tilde{v}_r \|$ is a bounded sequence and its spectrum is contained in a compact set, at least one subsequence converges to some $v^* \in V_N$ with $\| v^* \| = 1$ (cf. [3]). Note that $F_r/\| \tilde{v}_r \| \to 0$ implies $u_r/\| \tilde{v}_r \| \to (-v^*)$, which contradicts spect$(u_r) \subset T\backslash T_0$.

Now knowing that $\tilde{v}_r$ is bounded, by the same arguments we get $\tilde{v}_r \to F^*$ and $u_r \to 0$.

For performing the second splitting write

$$\tilde{v}_r = \sum_{\mu=1}^{j} \beta_\mu^{(r)} \gamma_\mu(t_1^{(r)},..., t_\mu^{(r)}; x). \tag{16.6}$$

From $\tilde{v}_r \to \hat{F}$ it follows that

$$\hat{F} = \sum_{\mu=1}^{j} \beta_\mu^* \gamma_\mu(t_1^*, t_2^*,..., t_\mu^* ; x)$$

with $\lim \beta_\mu^{(r)} = \beta_\mu^*$. Assume that the labeling process was performed such that t_1^*, t_2^*,..., t_k^* are the characteristic numbers of $\hat{F}$ with correct multiplicities. Hence,

$$\beta_\mu^* = 0, \qquad \mu > k.$$

Defining v_r to consist of the first k terms of (16.6) and putting $w_r = \tilde{v}_r - v_r$ we have

$$v_r \to \hat{F}, \qquad w_r \to 0.$$

Next, for the lifting of a flow we need an improvement of [8, Lemma 4.2].

LEMMA 16.2. *Let G be a C^1-manifold. Assume that $\hat{F}$ is not a critical point to f_0 in G. Then there are a connected neighborhood U of $\hat{F}$ in G, numbers $c > 0$, $\delta > 0$, and a flow*

$$\psi : [0, 1] \times U \to G$$

such that

$$\| f - \psi(\lambda, F)\| < \| f - F \| - c\lambda, \qquad 0 \leqslant \lambda \leqslant 1,$$

whenever $\| f - f_0 \| < \delta$.

Outline of Proof. Since $\hat{F}$ is not a critical point, by definition we have for an $h \in C_{\hat{F}}G$:

$$\| f_0 - \hat{F} - h \| < \| f_0 - \hat{F} \| . \tag{16.7}$$

The set of elements h satisfying (16.7) is open in $C_{\hat{F}}G$. Referring to (15.3) we may put $h = d_0 g(\alpha b)$ with $b \in C$ and any feasible parameterization g. It is also possible to fix $\alpha = 1$.

Put $c = \frac{1}{8}(\| f_0 - \hat{F} \| - \| f_0 - \hat{F} - h \|)$. By continuity there is a neighborhood C_1 of 0 in C, such that

$$\| f_0 - g(a)\| - \| f_0 - g(a) - d_a g(b - a)\| > 2c,$$

$$\| f_0 - g(a)\| - \| f_0 - g(a) - d_0 g(b - a)\| > 6c,$$

whenever $a \in C_1$. Therefore, $\| f - f_0 \| < c$ implies

$$\| f - g(a)\| - \| f - g(a) - d_a g(b - a)\| > 0,$$

$$\| f - g(a)\| - \| f - g(a) - d_0 g(b - a)\| > 4c.$$

Consequently, with the same arguments as in the proof of Lemma 4.2 in [8] and an appropriate cut-off function χ it follows that with $\psi(\lambda, g(a)) = g(a + \lambda\chi(a) \cdot b)$ a flow with the required properties is established. ∎

Now we are ready to prove:

THEOREM 16.3. *Let V_N be a normal family. Assume that F^* is a local best approximation to $f \in C(X)$ in $V_N \backslash V_{N-1}$. Then the boundary of the (maximal) component Cp assigned to F^* contains an element $\hat{F} \in V_{N-1}$, which is a local best approximation to f in $V_{k(\hat{F})}$.*

Proof. Assume to the contrary that $\hat{F} \in V_{N-1} \cap \overline{Cp}$ is not a critical point in V_k, $k = k(\hat{F})$. Consider a sequence $\{F_r\} \subset Cp$ converging to $\hat{F}$. Put

$$F_r = u_r + v_r + w_r\,,$$

according to the splitting process specified above.

Apply Lemma 16.2 to $\hat{F} \in V_k$ and denote the resulting flow by ψ. Its domain $U \subset V_k$ contains v_r for r sufficiently large. Set

$$v_r(\lambda) = \psi(\lambda, v_r).$$

In particular, the parameters $\xi^r(\lambda) \in \mathbb{R}^{2k}$ for representing $v_r(\lambda)$ in the manifold U have the form

$$\xi^r(\lambda) = \xi^r + \lambda b.$$

The elements of the w_r-sequence may be written as follows:

$$w_r = \sum_{\mu=k+1}^{j} \beta_\mu \gamma_\mu(t_1(\xi^r),\ldots, t_k(\xi^r), t_{k+1}^r,\ldots, t_\mu{}^r; x). \tag{16.8}$$

Moreover, let $w_r(\lambda)$ be the γ-polynomial, which results from a replacement of ξ^r by $\xi^r(\lambda)$ in (16.8). Appealing to (15.6) we estimate

$$\left| \frac{\partial}{\partial \lambda} \gamma_\mu(t_1(\xi(\lambda)),\ldots, t_k(\xi(\lambda)), t_{k+1},\ldots, t_\mu\,; x)\right| \leqslant M, \tag{16.9}$$

with $M < \infty$ for all ξ's in a neighborhood of the parameter vector for $\hat{F}$ and all $t_{k+1}, t_{k+2},\ldots, t_j$ in a neighborhood of spect$(\hat{F})$. Finally, put

$$F_r(\lambda) = v_r(\lambda) + (w_r(\lambda) + u_r)(1 - \lambda), \qquad 0 \leqslant \lambda \leqslant 1.$$

By applying Lemma 16.2 to $(f - u_r - w_r)$ we get

$$\begin{aligned}
&\| f - F_r(\lambda)\| \\
&\quad\leqslant \|(f - u_r - w\,) - v_r(\lambda)\| + (1 - \lambda)\| w\,(\lambda) - w_r \| + \lambda \| u_r + w_r \| \\
&\quad\leqslant \|(f - u_r - w_r) - v_r \| - c \cdot \lambda + \lambda(1 - \lambda)\, M \sum_{\mu>k} | \beta_\mu{}^r | + \lambda \| u_r + w_r \| \\
&\quad\leqslant \| f - F_r \|\,,
\end{aligned}$$

provided that

$$\| u_r + w_r \| \leqslant \frac{c}{2} \qquad \text{and} \qquad \sum_{\mu>k} | \beta_\mu{}^r | \leqslant \frac{c}{2M}\,. \tag{16.10}$$

The conditions (16.10) are satisfied for sufficiently large r. Hence, $F_r(\lambda) \in Cp$ for $0 \leqslant \lambda \leqslant 1$. Since $F_r(1) \in V_k$, this contradicts $Cp^\alpha \cap V_{N-1} = \varnothing$; with $\alpha = \|f - F_r\| < \beta$. ∎

The rest of the paper will be devoted to the question of how many components branch at each LBA in V_k, $k \leqslant N - 1$.

17. Generic Properties

Three integers have been assigned to each γ-polynomial F: the number of distinct characteristic numbers $\ell = \ell(F)$, the number of characteristic numbers counting multiplicities $k = k(F)$, and a number $L = L(F)$ satisfying

$$\ell \leqslant L \leqslant k. \tag{17.1}$$

By the characterization theorem of Hobby and Rice [12] an alternant of length $N + k + 1$ guarantees that a γ-polynomial in V_N is a best approximation. On the other hand, it was observed in 1967 that the converse is not true. The solutions have not always an alternant of this length. It is only necessary that there is an alternant of length $N + L + 1$. When the first correct necessary condition was derived the gap was even larger, and the smaller bound $N + \ell + 1$ was given.

The difference between the necessary and the sufficient conditions becomes effective only if $\ell < k$, i.e., if at least one characteristic number of the γ-polynomial has a multiplicity greater than one. At first glance one would hope that two of the characteristic numbers will almost never coalesce, because one expects that small perturbations of f remove multiplicities greater than 1 and the gap would not be serious.

The following analysis will show, however, that in most cases small perturbations do not change the multiplicities of the characteristic numbers. In fact, for the solutions to most functions in $C(X)$ the integers ℓ and L coincide. This means that there is even the tendency for the gap to become maximal.

On the other hand certain degeneracies occur very rarely. It is crucial for the development in Section 19 that we can eliminate the exceptional cases.

It may be appropriate at this point to recall the definitions of the integer L. If $F \in V_N{}^0$ put $L = \ell = k$. Otherwise, if $F \in V_N \backslash V_N{}^0$, then write F in a form in which the terms with multiplicity one and greater than one are separated:

$$F = \sum_{\nu=1}^{\ell_1} \sum_{\mu=1}^{m_\nu} \alpha_{\nu\mu} \gamma_\mu(t_\nu, \ldots, t_\nu; x) + \sum_{\nu=\ell_1+1}^{\ell} \alpha_{\nu 1} \gamma(t_\nu, x) \tag{17.2}$$

with

$$t_1 < t_2 < \cdots < t_{\ell_1},\ \alpha_{\nu m_\nu} \neq 0,\ \nu = 1, 2,\ldots, \ell_1 .$$

Similar to (12.2) we put

$$\sigma_\nu = \operatorname{sign} \alpha_{\nu m_\nu},\ \nu = 1, 2,\ldots, \ell_1 ,$$

$$\begin{aligned} r_\nu &= 1, \qquad \text{if } \sigma_\nu\sigma_{\nu+1}(-1)^{m_\nu+1} < 0,\ \nu = 1, 2,\ldots, \ell_1 - 1, \\ &= 0, \qquad \text{otherwise,} \end{aligned} \tag{17.3}$$

$$L = \ell + \sum_{\nu<\ell_1} (1 - r_\nu). \tag{17.4}$$

Observe that each characteristic number is counted with multiplicity one or two when L is evaluated. In particular, we have $L = \ell$ as long as $N \leqslant 3$.

By Theorem 12.3, F is an LBA to f in $V_N \backslash V_{N-1}$ if and only if there is an alternant of length $N + L + 1$ with sign $-\sigma_{\ell_1}$ on the right. We note that $N + L - 1$ is the maximal number of sign changes for the elements in the tangent cone $C_F V_N$.

Definition 17.1. The function $f \in C(X)$ is an unexceptional point in $C(X)$ (with respect to V_N) if the following properties hold for the approximation in V_k, $k \leqslant N$.

(i) For each local best approximation F to f in V_k there is an alternant of exact length $k + \ell(F) + 1$. Consequently, the equality $\ell(F) = L(F)$ holds.

(ii) Each local best approximation to f in V_k has the maximal order k.

(iii) There is a neighborhood U of f in $C(X)$ such that the number of local best approximations in V_k is the same for all $g \in U$.

(iv) For each local best approximation F to f in V_k the end points of the interval X do not belong to the alternant.

At the moment we consider the term "unexceptional point" as a purely formal definition. It will be shown in the sequel that the definition is consistent with the use in differential topology [15, p. 20]. A subset of a topological space is said to be residual, if it can be expressed as a countable intersection of dense, open sets. Its elements may be denoted as unexceptional points. A property is called generic if it holds for the elements of a residual set at least.

For the justification of the definition above we need an improvement of Theorem 15.1. It will be repeatedly used when perturbation techniques are applied.

Theorem 17.1. *Let G be a Haar embedded manifold* (e.g., *let* $G = V_N \backslash$

V_{N-1}), and let $f \in C(X)$. Assume that $F_0 \in G$ is a local best approximation to f. Then there is a neighborhood W of f in $C(X)$ and a neighborhood U of F_0 in G, such that U contains exactly one local best approximation to each $g \in W$ in G.

Proof. Let $F_0 \in G$ be an LBA to f. Since G is a finite-dimensional manifold and local strong uniqueness holds, we know that for some $\beta > \| f - F_0 \|$ the component Cp of ρ^β containing F_0 is compact.
Since we may eventually replace G by a neighborhood of F_0 in G we may assume that ρ^β is already connected. Choose $\delta < \frac{1}{4}(\beta - \| f - F_0 \|)$, $\delta > 0$. Given $g \in C(X)$, $\| f - g \| < \delta$, put

$$\tilde{\rho}^\alpha = \{F \in G; \| g - F \| \leqslant \alpha\}.$$

Observe that $|\| f - F \| - \| g - F \|| < \delta$ implies

$$\rho^{\alpha-\delta} \subset \tilde{\rho}^\alpha \subset \rho^{\alpha+\delta}, \quad \alpha \in \mathbb{R}. \tag{17.5}$$

Let F_1 be a best approximation to g in the compact set ρ^β. From $\| g - F_1 \| \leqslant \| g - F_0 \| < \| f - F_0 \| + \delta$ we conclude that $F_1 \in \tilde{\rho}^{\beta-3\delta} \subset \rho^{\beta-2\delta}$. Hence, F_1 is not a boundary point of ρ^β and is an LBA in G. Assume that F_2 is another LBA to g in $\rho^{\beta-2\delta}$. Since $\rho^{\beta-2\delta}$ is connected and contained in $\tilde{\rho}^{\beta-2\delta}$, both F_1 and F_2 belong to the same component of $\tilde{\rho}^{\beta-\delta}$. This set is compact because it is a closed subset of ρ^β. By the uniqueness theorem for Haar embedded manifolds [8] we have $F_1 = F_2$. ∎

Combining Theorem 12.5 and Theorem 17.1 we obtain the following corollary:

COROLLARY 17.2. *Suppose the conditions of Theorem 17.1 hold. Moreover, define U and W as in the theorem. Then the mapping from $W \subset C(X)$ to $U \subset V_N$ which sends each f to its local best approximation is continuous.*

The corollary generalizes a result of Schmidt. In [17] the continuity of the metric projection was derived for exponentials under additional restrictions on the alternant. An earlier result of Barrar and Loeb [1] applies to the varisolvent subset $V_N{}^0$.
The anounced justification of Definition 17.1 is now established under a hypothesis which will be proved in Section 19. We refer to (14.1) for the definition of the constants c_k.

LEMMA 17.3. *Assume that $c_k < \infty$, $k = 1, 2,..., N$. Then the elements of $C(X)$ being not exceptional with respect to V_N form an open, dense subset of $C(X)$. Moreover, there is an unexceptional f having exactly c_N local best approximations in V_N.*

Proof. For convenience, we introduce the set with the γ-polynomial of order zero $V_0 = \{0\}$, and start the inductive proof at $N = 0$. The set of functions f such that there is only an alternant of length 1 to $f - 0$ is open and dense in $C(X)$, Hence, the obvious extension of the lemma to $N = 0$ holds.

Let $N \geqslant 1$ and assume that the elements of $C(X)$ being not exceptional with respect to V_{N-1} form an open, dense subset $C_1 \subset C(X)$. This and Theorem 6.1(b) imply that no LBA to $f \in C_1$ in V_N has an order less than N. Hence, (ii) is a generic property. (Here and in the sequel, (i) through (iv) refer to Definition 17.1.)

Given a positive integer j, by Theorem 17.1 the subset of elements in $C(X)$ having at least j local solutions in $V_N \backslash V_{N-1}$ is open. Combining this and the assumption $c_N < \infty$ we conclude that the subset $C_2 \subset C(X)$ satisfying condition (iii) is open and dense.

Assume that F^* is an LBA to $F \in C_1$ and that the length of the alternant does not exceed $N + \ell(F^*) + 1$. For each F in a sufficiently small neighborhood U of F^* in V_N we have $\ell(F) \geqslant \ell(F^*)$. Moreover, for each g in a sufficiently small neighborhood of $f \in C(X)$ the length of the alternant of $g - F$ cannot exceed $N + \ell(F^*) + 1$. From Corollary 17.2 it follows that $\ell(F) = \ell(F^*)$, if $F \in U$ is an LBA to g. Consequently, (i) holds for an open set in $C_1 \cap C_2$.

To prove the density let F^* be an arbitrary LBA to f in $V_N \backslash V_{N-1}$. In particular, $L(F^*) > \ell(F^*)$ is admitted. Write F^* in the form (17.2). Given $\delta > 0$ put

$$F_1 = \sum_{\substack{\nu=1 \\ (r_\nu = 0)}}^{\ell_1 - 1} \left\{ \sum_{\mu=1}^{m_\nu - 1} \alpha_{\nu\mu} \gamma_\mu(t_\nu, \ldots, t_\nu ; x) + \alpha_{\nu m_\nu} \gamma_{m_\nu}(t_\nu, \ldots, t_\nu, t_\nu + \delta ; x) \right\}$$
$$+ \text{ all other terms of } F^* \text{ unchanged.} \tag{17.6}$$

Note that F_1 has been constructed from F^* by separating characteristic numbers such that for F_1 all parameters r_ν in (17.3) equal one. It follows that $\ell(F_1) = L(F_1) = L(F^*)$, and F_1 is a local solution to $g = f + (F_1 - F^*)$. Moreover, by slightly modifying g in the neighborhood of extremal points the number of points of the alternant is reduced to $N + \ell(F_1) + 1$ ones (if necessary).

What happens with the other LBAs when we perform the perturbation process above? From the discussion at the beginning of the proof we know that sufficiently small perturbations do not spoil the relation between the length of the alternant and $N + \ell(F) + 1$. Hence, by a finite number of repetitions we obtain a function such that property (i) holds. Since δ may be chosen arbitrarily small, this proves the density as stated in the lemma.

Finally, if a point of the alternant is an end point of the interval, obviously it can be shifted into the interior by a small perturbation.

Since all constructions may be started from an element with a maximal number of local solutions, the proof is complete. ∎

Now we want to illustrate by an example what was indicated at the beginning of this section. Recall that $c_1 = 1$, $c_2 = 2$. Let $\gamma(t, x) = e^{tx}$, $T = \mathbb{R}$. Then $f(x) = \cos x$ has two best approximations in V_2 when $X = [-1, +1]$ is the approximation interval [4]. For each solution the spectrum collapses to a single point which is counted with multiplicity two. From the preceding lemma and its proof it follows that any g in an open neighborhood of f has two local solutions in V_2 and for each one $k = 2$, $l = 1$.

18. The Standard Construction of Local Best Approximations

In this section a construction for local best approximations in V_N and their maximal components is presented. Each constructed component is characterized by a local solution in V_{N-1} and an additional characteristic number. The basic idea may be found in [5]. However, contrary to the algorithm in [5] it is applied to all local solutions in V_{N-1} and not only to a global one. As we see in the next section the construction yields all local optima in V_N, whenever f is not exceptional.

In order to motivate the subsequent arguments let us begin with a few intuitive and not quite rigorous remarks. Since the domain T is diffeomorphic to $\mathbb{R}$, for simplicity we may assume $T = \mathbb{R}$. The mapping

$$\begin{aligned} \phi &: \mathbb{R}^2 \to V_1 \\ \phi(\alpha, t) &= \alpha \cdot \gamma(t, x) \end{aligned} \tag{18.1}$$

is one–one only if in 2-space the straight line $\{(0, t);\, t \in \mathbb{R}\}$ is contracted to a point. The zero function is the reason that V_1 is not a manifold. Now we may ask whether the singularity cannot be eliminated in the opposite way: does it make sense to blow up the singular point to a one-dimensional set by not identifying the elements $0 \cdot \gamma(t_1, x)$ and $0 \cdot \gamma(t_2, x)$ for $t_1 \neq t_2$? In the same manner we might blow up a γ-polynomial F of order $N - 1$ in V_N by writing it in the form $F + 0 \cdot \gamma(t, x)$. The following analysis gives a positive answer to the question above and indicates which t's are essentially inequivalent.

Let $t_1, t_2, ..., t_p$ be p not necessarily disjoint numbers in T, $p \leqslant N$. Then

$$V_N(t_1, t_2, ..., t_p) = \text{closure}\, \{F \in V_N;\, \{t_1, t_2, ..., t_p\} \subset \text{spect}(F)\} \tag{18.2}$$

denotes the subset of V_N with p characteristic numbers fixed [10]. There is no serious confusion with the notation for the sign classes $V_N(s_1, s_2, ..., s_N)$

because the letters s and t never occur at the same time in connection with the term V_N. Further if $t_1 = t_2 = \cdots = t_m$, $m \leqslant p$, then the abbreviation

$$V_N(m \times t_1, t_{m+1}, \ldots, t_p) = V_N(\underbrace{t_1, \ldots, t_1}_{m \text{ times}}, t_{m+1}, \ldots, t_p) \tag{18.3}$$

is used. In this section we refer only to the special situation with one single number fixed at $t_1 = \tau$:

$$V_N(\tau) = \{F \in V_N\ ;\ \tau \in \operatorname{spect}(F) \text{ or } F \in V_{N-1}\}.$$

Given a local best approximation $\hat{F}$ in V_{N-1}, $k(\hat{F}) = N - 1$, choose a parameterization $g : A \to V_{N-1}$ for a neighborhood of $\hat{F}$ in the Haar embedded manifold V_{N-1}. Moreover, let $\tau \in T$ be disjoint from $\operatorname{spect}(\hat{F})$. We may assume, after reducing the domain A, if necessary, that τ does not belong to the spectrum of any γ-polynomial in $g(A)$. Then by

$$\begin{aligned} &A \times \mathbb{R} \to V_N(\tau), \\ &(a, \alpha) \to g(a) + \alpha \cdot \gamma(\tau, x), \end{aligned} \tag{18.4}$$

a parameterization for a neighborhood U_τ of $\hat{F} = \hat{F} + 0 \cdot \gamma(\tau, x)$ in $V_N(\tau)$ is defined. Therefore, the tangent cone at $F + \alpha\gamma(\tau, x) \in U_\tau$ is easily computed:

$$C_{F+\alpha\gamma(\tau,x)} V_N(\tau) = C_F V_{N-1} + \{\delta \cdot \gamma(\tau, x);\ \delta \in \mathbb{R}\}. \tag{18.5}$$

Hence, U_τ is a Haar embedded manifold. In particular, the elements of the form $F + 0 \cdot \gamma(\tau, x)$ are not degenerate in $V_N(\tau)$ although contained in V_{N-1}.

It follows from (18.5) that $C_F V_N(\tau)$ contains a Haar subspace with dimension $(N - 1) + \ell(\hat{F}) + 1$. As a consequence we have:

ASSERTION 18.1. *Let $\hat{F}$ be a local best approximation to f in V_{N-1}. Moreover, assume that f is not exceptional. If $\tau \notin \operatorname{spect}(\hat{F})$, then $\hat{F}$ is not a critical point in $V_N(\tau)$.*

We remark that the genericity assumption in Assertion 18.1 cannot be abandoned. To verify this, assume that $L(\hat{F}) > \ell(\hat{F})$. Referring to (17.3) we have $r_\nu = 0$ for at least one $\nu < \ell_1$. If τ is chosen such that

$$t_\nu < \tau < t_{\nu+1},$$

then $\hat{F}$ is a critical point in $V_N(\tau)$. This may be verified by arguments as in the proof of Lemma 12.2.

Now we introduce the standard construction.

Construction. Let $\hat{F}$ be a local best approximation to f in $V_{N-1} \setminus V_{N-2}$ and $\tau \notin \operatorname{spect}(\hat{F})$. Let $h \in C_{\hat{F}+0}V_N(\tau)$ satisfy

$$\| f - \hat{F} - h \| < \| f - \hat{F} \| . \tag{18.6}$$

By Definition 15.2 there is a curve $\{F_\lambda, 0 \leqslant \lambda \leqslant 1\} \subset V_N(\tau)$ such that $\| F_\lambda - \hat{F} - \lambda h \| = O(\lambda)$, if $\lambda \to 0$. For sufficiently small $\lambda > 0$ we have

$$\| f - F_\lambda \| < \| f - \hat{F} \| . \tag{18.7}$$

There are two possibilities: The component of the level set $\{F \in V_N ; \|f - F\| < \|f - \hat{F}\|\}$ containing the curve $\{F_\lambda ; \lambda > 0, \lambda \text{ sufficiently small}\}$, is disjoint from V_{N-1}. Then a maximal component has been constructed. If on the other hand the component intersects V_{N-1}, it may be discarded. (The corresponding components may be constructed by starting from another LBA in V_{N-1} lying at a lower level.

LEMMA 18.2. *Let f be an unexceptional point with respect to V_{N-1}. Then the component constructed depends only on the choice of $\hat{F}$ and on the interval of $\mathbb{R}\backslash\operatorname{spect}(\hat{F})$ containing τ.*

Proof. Given $\tau \notin \operatorname{spect}(\hat{F})$ the standard construction yields a curve in $V_N(\tau)$ such that its elements

$$F_\lambda = \alpha_\lambda \cdot \gamma(\tau, x) + v_\lambda , \qquad 0 < \lambda < 1,$$

satisfy the relation

$$\| f - F_\lambda \| < \| f - \hat{F} \| \tag{18.8}$$

for sufficiently small λ, say for $0 < \lambda \leqslant 1$. Note that

$$s = \operatorname{sign} \alpha_\lambda$$

is independent of λ. By Theorem 17.1 a neighborhood U of $\hat{F}$ in V_{N-1} and a $\delta > 0$ exist such that there is a unique best approximation to g in U provided that

$$\| f - g \| < \delta.$$

Hence, the level set

$$\{v \in U ; \| g - v \| \leqslant \| f - \hat{F} \|\}$$

is connected or void. By specifying $g = f - \alpha_\lambda \cdot \gamma(\tau, x), 0 < \lambda \leqslant 1,$ we conclude that the construction leads to a component which is independent of the choice of the tangent vector h satisfying 18.6 and independent of the curve $\{F_\lambda\}$.

Once again, let $g \in C(X)$ satisfy $\|f - g\| < \delta$. Denote its best approximation in U by F. After reducing δ, if necessary, and recalling Corollary 17.3 we know that the length of the alternant and its sign are the same for $f - \hat{F}$ and $g - F$. Hence, there is a tangent vector

$$h = \beta\gamma(\tau, x) + h_1 \tag{18.9}$$

in the $(N - 1) + \ell(F) + 1$-dimensional linear subspace of the cone $C_{F+0} V_N(\tau)$ satisfying

$$\| g - F - h \| < \| g - F \| . \tag{18.10}$$

The crucial point to be observed here is that $s_0 = \operatorname{sign} \beta$ for β in (18.9) is independent of g. Indeed, by standard arguments it follows from (18.10) that $h(x)$ has $N - 1 + \ell(F)$ zeros. Since h is a γ-polynomial of order $N + \ell(F)$, by Theorem 3.2 its generalized signs are fixed.

Consequently, there is a better approximation in U to the function $g - \lambda\beta \gamma(\tau, x)$, λ sufficiently small, than we have for g. In particular, we may choose $g = f - \alpha\gamma(\tau, x)$ and denote the unique best approximation to this element by $F_{\alpha,\tau}$. It follows that

$$\|f - \alpha\gamma(\tau, x) - F_{\alpha,\tau}\| < \|f - \beta\gamma(\tau, x) - F_{\beta,\tau}\| \quad \text{for } 0 \leqslant |\beta| < |\alpha| , \tag{18.11}$$

as long as $s_0 \cdot \alpha > s_0 \cdot \beta \geqslant 0$ and $\| \alpha \cdot \gamma(\tau, x)\| < \delta$.

Now we compare the standard construction for two distinct numbers τ_1 and τ_2, $\tau_1 < \tau_2$, assuming that

$$[\tau_1 , \tau_2] \cap \operatorname{spect}(\hat{F}) = \varnothing .$$

We may assume that the neighborhood U has been chosen so small such that $[\tau_1 , \tau_2] \cap \operatorname{spect}(F) = \varnothing$ whenever $F \in U$.

Consider the continuous curve

$$t \to \alpha\gamma(t, x) + F_{\alpha,t} \in V_N , \qquad \tau_1 \leqslant t \leqslant \tau_2 , \tag{18.12}$$

where

$$\alpha = s_0 \frac{\delta}{2} \min \left\{ \frac{1}{\| \gamma(t, x)\|} ; \tau_1 < t < \tau_2 \right\} .$$

It follows from (18.11) that the γ-polynomials of the curve (18.12) are better approximations than $\hat{F} = \hat{F} + 0 \cdot \gamma(t, x)$. The endpoints of the curve are contained in the same component of $\{F \in V_N ; \|f - F\| < \|f - \hat{F}\|\}$. The construction with τ_1 and τ_2 leads to the same result. ∎

19. Completeness of the Standard Construction

In this section we prove the following theorem which has the main result of this paper as an immediate consequence.

THEOREM 19.1. *Let V_N be a normal family and let $f \in C(X)$ be an unexceptional element with respect to V_{N-1}. Then each local best approximation to f in V_N may be constructed via its maximal component by applying the standard construction to a local best approximation in V_{N-1} which has the maximal order $k = N - 1$.*

Given an LBA in V_N we know from Theorem 16.3 that the boundary of the associated maximal component Cp contains a local solution $\hat{F}$ in V_k, $k = k(\hat{F}) \leqslant N - 1$. Put $SP := \operatorname{spect}(\hat{F})$. Let $\{F_r\}$ be a sequence in Cp converging to $\hat{F}$ and let

$$F_r = u_r + v_r + w_r, \qquad r = 1, 2, 3, \ldots \tag{19.1}$$

be the splitting defined in Section 16. To prove the theorem we distinguish 5 cases:

Case 1. $k = N - 1$ and u_r is a γ-polynomial of order 1 such that $\operatorname{spect}(u_r)$ is disjoint from an open neighborhood T_0 of SP.

Since V_N is assumed to be normal, the spectrum of each γ-polynomial in some open neighborhood U of $\hat{F}$ in V_{N-1} is contained in T_0. Fix δ as in the proof of Lemma 18.2. For a sufficiently large r we have $\| u_r \| < \delta$. Write $u_r = \alpha_r \cdot \gamma(\tau_r, x)$. It follows that the standard construction with the choice $\tau = \tau_r$ yields the component of the level set containing F_r. This completes the proof for Case 1.

Case 2. $k = N - 1$ but we do not have Case 1.

After passing to a subsequence, if necessary, we know that m characterestic numbers of F_r come close to some $\tau \in SP$, with $m - 1$ being the multiplicity of τ in SP, To complete the proof of Case 2 we will study the approximation problem with m characteristic numbers fixed close to τ.

ASSERTION 19.2. *Let τ be a characteristic number of a local best approximation $\hat{F}$ to f in V_{N-1} and let $m - 1$ be its multiplicity. Given $\eta > 0$ there is a $\delta > 0$ such that a local best approximation F_1 exists to f in $V_N(t_1, t_2, \ldots, t_m)$ satisfying*

$$\| F_1 - \hat{F} \| < \eta$$

provided that

$$| t_i - \tau | \leqslant \delta, \qquad i = 1, 2, \ldots, m. \tag{19.2}$$

Proof. Referring to Definition 12.1 we observe that $C_{\hat{F}} V_N(m \times \tau) \subset C_{\hat{F}} V_{N-1}$. Hence, $\hat{F}$ is a critical point in $V_N(m \times \tau)$. Since a neighborhood of $\hat{F}$ in $V_N(m \times \tau)$ is a Haar embedded manifold, $\hat{F}$ is a strong LBA, i.e., for some $c > 0$ and $\eta_1 > 0$ we have

$$\| f - F \| \geqslant \| f - \hat{F} \| + c \| F - \hat{F} \|,$$

whenever $F \in U := \{F \in V_N(m \times \tau); \| F - \hat{F} \| < \eta_1\}$. Note that $c \leqslant 1$. Since we may reduce η if necessary we assume $2\eta < \eta_1$. Given $t_1, t_2, \ldots, t_m \in T$ define a one-one mapping $\phi : U \to V_N(t_1, t_2, \ldots, t_m)$ by the following procedure: The terms $\sum_{\mu=1}^{m} \beta_\mu \gamma_\mu(\tau, \ldots, \tau; x)$ in the representation of a γ-polynomial in $U \subset V_N(m \times \tau)$ are replaced by $\sum_{\mu=1}^{m} \beta_\mu \gamma_\mu(t_1, t_2, \ldots, t_\mu; x)$. Thereby the factors β_μ are kept. If δ is sufficiently small and (19.2) holds, then

$$\| \phi(F) - F \| < \tfrac{1}{3} c\eta, \qquad \text{for } F \in U \tag{19.3}$$

We claim that the best approximation to f in the compact set $\phi(\bar{U})$ satisfies the statement of the theorem. Indeed we have

$$\| f - \phi(\hat{F}) \| < \| f - \hat{F} \| + \tfrac{1}{3} c\eta.$$

Moreover, $\hat{F} \in \phi(\bar{U})$ and $\| F - \hat{F} \| \geqslant \eta$ imply $\| \phi^{-1}(F) - \hat{F} \| \geqslant \eta - \frac{1}{3} c\eta \geqslant \frac{2}{3}\eta$ and

$$\begin{aligned} \| f - F \| &\geqslant \| f - \phi^{-1}(F) \| - \| F - \phi^{-1}(F) \| \\ &> \| f - \hat{F} \| + c \| \hat{F} - \phi^{-1}(F) \| - \tfrac{1}{3} c\eta \\ &\geqslant \| f - \phi(\hat{F}) \|. \end{aligned}$$

Hence, the optimum is attained at a point of $\phi(U)$ and not on its boundary. Since $\phi(U)$ is open in $V_N(t_1, t_2, \ldots, t_m)$ the proof of Assertion 19.2 is complete. ∎

If η in Assertion 19.2 is chosen sufficiently small, the open set

$$\{F \in V_N(t_1, \ldots, t_m); \| F - \hat{F} \| < 4\eta\}$$

is a Haar embedded manifold. With the same arguments as in the proof of Theorem 17.1 it follows that the LBA given in the assertion is unique. Uniqueness and local compactness imply that the map which sends the m-tuple $t_1, t_2, \ldots, t_m$ to the LBA in $V_N(t_1, t_2, \ldots, t_m)$, is continuous.

As a consequence we have

ASSERTION 19.3. *Let the conditions of Assertion* 19.2 *hold. Moreover, assume that* $f - \hat{F}$ *has an alternant of length only* $(N - 1) + \ell(\hat{F}) + 1$. *Then* *there is a $\delta > 0$ with the following property: Let $\tau - \delta \leqslant \underline{t}_i \leqslant \bar{t}_i \leqslant \tau + \delta$, $i = 1, 2,..., m$, and F be a local best approximation to f in*

$$W = \bigcup \{V_N(t_1, t_2, ..., t_m); \underline{t}_i \leqslant t_i \leqslant \bar{t}_i, i = 1, 2,..., m\}.$$

If F is sufficiently close to $\hat{F}$, then each characteristic number t_i of F coincides either with $\underline{t}_i$ or with $\bar{t}_i$, $i = 1, 2,..., m$.

Proof. Let F be an LBA to f in W having the characteristic numbers $t_i \in [\underline{t}_i, \bar{t}_i]$, $i = 1, 2,..., m$. Assume that

$$\underline{t}_j < t_j < \bar{t}_j$$

for some j, $1 \leqslant j \leqslant m$. Then F is a local best approximation in $V_N(t_1, ..., t_{j-1}, t_{j+1}, ..., t_m)$. This means that t_i, $1 \leqslant i \leqslant m$, $i \neq j$ are considered fixed while t_j is a free parameter. From the characterization theorem [10, Satz 3.2] it follows that the length of the alternant must be at least $N + \ell(\hat{F}) + 1$. On the other hand, if $\| F - \hat{F} \|$ is sufficiently small, the length of an alternant of $(f - F)$ cannot exceed the length of the alternant of $(f - \hat{F})$. This contradicts the hypothesis. Hence, Assertion 19.3 is proved. ∎

We remark that the existence of an LBA in W is easily verified under the hypothesis that δ is sufficiently small. Let U and ϕ be defined as in the proof of Assertion 19.2, where ϕ depends on $t_1, t_2, ..., t_m$. The best approximation in the compact set

$$\bigcup \{\phi(\bar{U}); \underline{t}_i \leqslant t_i \leqslant \bar{t}_i, i = 1, 2,..., m\}$$

is a local solution in W.

Now we consider a special situation to which Assertion 19.3 can be applied.

Fix δ as in Assertion 19.3. From the discussion in the preceding section we know that there is a unique best approximation to $f - \alpha\gamma(\tau + \delta, x)$ in a neighborhood U of $\hat{F}$ in V_{N-1} provided that $|\alpha|$ is sufficiently small, for instance $|\alpha| < c_1$. Observe that $\hat{F}$ is a critical point to f in $V_N((m - 1) \times \tau, \tau + \delta)$. Hence, $\hat{F}$ is locally optimal in this set. The continuity arguments after Assertion 19.2 shows that there is a unique local solution $F_1 + \alpha\gamma(\tau + \delta, x)$, $F_1 \in U$, $|\alpha| < c_1$, in the set $V_N(t_1, ..., t_{m-1}, \tau + \delta)$ whenever

$$\tau - \delta_2 \leqslant t_i \leqslant \tau + \delta_2, \qquad i = 1, 2,..., m - 1. \tag{19.4}$$

Here δ_2 is supposed to be chosen sufficiently small, $0 < \delta_2 < \delta$. After reducing δ_2 once more, if necessary, we may assume that the same is true when $\tau + \delta$ is replaced by $\tau - \delta$ throughout. Finally put

$$W = \bigcup \{V_N(t_1, t_2, ..., t_m); \tau - \delta \leqslant t_m \leqslant \tau + \delta, \tau - \delta_2 \leqslant t_i \leqslant \tau + \delta_2, \\ i = 1, 2,..., m - 1\}.$$

At this point we recall the sequence $\{F_r\}$ in Cp which was specified when defining Case 2. By construction, $F_r \in W$ holds for sufficiently large r. Hence, $Cp \cap W$ is not empty and there is an LBA F to f in $Cp \cap W$ which is close to $\hat{F}$. By virtue of Assertion 19.3 F has the form $F_1 + \alpha\gamma(t, x)$ with $t = \tau + \delta$ or $t = \tau - \delta$ and $F_1 \in V_{N-1}$. Consequently, Cp can be generated by the standard construction with the additional characteristic number $\tau + \delta$ or $\tau - \delta$. Hence, Case 2 is reduced to Case 1.

We remark that a slight modification of the argument above will be very useful in the next section. Consider the set W as above but change the restriction for t_m

$$\tau - \delta_2 \leqslant t_i \leqslant \tau + \delta_2, \qquad i = 1, 2, \ldots, m - 1$$
$$\tau - \delta_2 \leqslant t_m \leqslant \tau + \delta.$$

Consider the standard construction with the additional characteristic number $\tau - \delta_2$. We conclude that either the constructed component contains an element whose spectrum contains m times $(\tau - \delta_2)$ or the construction with $\tau + \delta$ leads to the same component.

The remaining cases can be dispatched more briefly.

Case 3. $k \leqslant N - 2$ and there is a $\tau \in T\backslash\text{spect}(\hat{F})$ which is an accumulation point of $\{\text{spect}(F_r), r = 1, 2, \ldots,\}$.

Since $(f - \hat{F})$ has an alternant of length $k + \ell + 1$ only and $\tau \notin SP$, the γ-polynomial $\hat{F}$ is not a critical point in $V_{k+1}(\tau)$. By Lemma 16.2 a flow ψ ψ may be defined on a neighborhood of $\hat{F}$ in $V_{k+1}(\tau)$. Similarly, the flow operator may be extended to $V_{k+1}(t)$ for any $t \in [\tau - \delta, \tau + \delta]$ if δ is sufficiently small. Now in the same way as in the proof of Theorem 16.3 a path in Cp is constructed which connects an F_r with an element in $V_{k+1}(t_r)$, $t_r \in \text{spect}(F_r)$, such that all elements of the path are as good approximations as F_r is. This is a contradiction, thus Case 3 is impossible.

Case 4. $k \leqslant N - 2$ and $w_r \neq 0$ for infinitely many r.

After passing to a subsequence, if necessary, we know that m or more characteristic numbers of F_r tend to some $\tau \in SP$. With the technique used when treating Case 2 it is possible to construct a sequence in Cp such that $\tau + \delta$ or $\tau - \delta$ is an accumulation point of the spectra. Hence, this case is reduced to Case 3 and therefore it is also impossible.

Case 5. $k \leqslant N - 2$ but we have neither Case 3 nor Case 4.

After passing to a subsequence, if necessary, we know that $N - k$ characteristic numbers of F_r tend to $+\infty$ or $-\infty$.

ASSERTION 19.4. *Let U be an open neighborhood of $\hat{F}$ in V_k having a compact closure. For any $t_1, t_2, \ldots, t_{N-k} \in T$ let*

$$W = \{v + u \in V_N\,;\, v \in U,\, u \in V_{N-k}(t_1, t_2, \ldots, t_{N-k})\}.$$

If f is not exceptional with respect to V_k, then there is an $\eta > \|f - \hat{F}\|$ such that the level set $\{F \in W; \|f - F\| \leqslant \eta\}$ is compact, whenever $t_1, t_2, \ldots, t_{N-k}$ are sufficiently large.

Proof. Note that $\hat{F} \in W$. By standard arguments we obtain $\overline{W}$ when we replace the condition $v \in U$ by $v \in \overline{U}$. Obviously, for each $\eta > \|f - \hat{F}\|$ the elements below the level η form a compact subset of $\overline{W}$. We will specify η such that the level set is already contained in W.

Let $u + v \in \overline{W}$ and $\|f - u - v\| \leqslant \eta$. Then we have

$$\|u\| \leqslant \|f - \hat{F}\| + \|\hat{F} - v\| + \|f - u - v\|. \tag{19.5}$$

Hence, u is bounded by a constant independent of $t_1, t_2, \ldots, t_{N-k}$. Let $x_0 < x_1 < \cdots < x_{k+\ell}$ be an alternant of $f - \hat{F}$. Since f is not exceptional, x_0 and $x_{k+\ell}$ are interior points of the interval. From the normality assumption and (19.5) we conclude

$$|u(x)| < (\eta - \|f - \hat{F}\|), \qquad x_0 \leqslant x \leqslant x_{k+\ell}$$

provided that $t_1, t_2, \ldots, t_{N-k}$ are sufficiently large. Hence,

$$\begin{aligned} |f(x_i) - v(x_i)| &\leqslant \|f - v - u\| + |u(x_i)| \\ &\leqslant \|f - \hat{F}\| + 2(\eta - \|f - \hat{F}\|), \qquad i = 0, 1, \ldots, k + \ell. \end{aligned}$$

Since $\hat{F}$ is a strong local best approximation to f in U, it follows that $\|v - \hat{F}\|$ is small and $v \in U$, whenever $\eta - \|f - \hat{F}\|$ is sufficiently small. ∎

Let U be an open neighborhood of $\hat{F}$ in V_k. Fix η as in the assertion above. Now recall the sequence $\{F_r\} \subset Cp$ which was specified when defining Case 5. Referring to (19.1) for sufficiently large r we have

$$\|f - v_r\| < \eta,$$

and the characteristic numbers of u_r are sufficiently large. Consequently,

$$\|f - v_r - \lambda u_r\| = \|\lambda(f - F_r) + (1 - \lambda)(f - v_r)\| < \eta, \qquad 0 \leqslant \lambda \leqslant 1.$$

Hence, there is a continuous arc between F_r and v_r, which runs below the level η. Moreover, a continuous arc from v_r to $\hat{F}$ exists in U on which the distance to f does not exceed η. Finally, $\hat{F}$ is not a critical point in $V_{k+1}(\tau)$ with $\tau \in \operatorname{spect}(u_r)$. Hence, there is a continuous curve from $\hat{F}$ to a γ-polynomial $\tilde{F} \in V_{k+1}(\tau)$, such that the distance to f does not exceed $\|f - \hat{F}\| < \eta$. With this we have established that F_r and $\tilde{F}$ belong to the same component of a level set in the Haar embedded manifold W. From the general theory [8, Satz 4.1] it follows that there is a continuous curve between F_r and $\tilde{F}$ such that the distance function $F \to \|f - F\|$ attains its maximum at an endpoint.

Therefore, we have constructed a path from F_r to $\tilde{F} \in V_{k+1} \subset V_{N-1}$ which does not leave Cp contradicting $Cp \cap V_{N-1} = \varnothing$. Hence, Case 5 is also impossible.

This completes the proof of Theorem 19.1. ▮

As a consequence of the above arguments we obtain the main result of this paper.

THEOREM 19.5. *Let V_N be a normal family. Moreover, assume that T is an open interval on the real line. Then to each $f \in C(X)$ there are at most $N!$ local best approximations in V_N.*

Proof. By Theorem 4.3 and Corollary 7.5 the statement is true for $N = 1$ and $N = 2$.

Assume that the theorem has already been proved for $1, 2, \ldots,$ and $N - 1$, $N \geqslant 3$. Given $f \in C(X)$, by Lemma 17.3 there is a function $f_1 \in C(X)$ which has at least as many LBAs in V_N as f and which is not exceptional with respect to V_{N-1}. By Theorem 19.1 all LBAs to f_1 in V_N may be generated by the standard construction. It follows from Lemma 18.2 that there are at most $N \cdot c_{N-1} \leqslant N!$ different LBAs to f_1 in V_N. This concludes the inductive proof. ▮

20. γ-POLYNOMIALS OF ORDER 3

The bounds for the number of LBAs given in Theorem 19.5 are not sharp for $N \geqslant 3$. Here we will establish the optimal bound for V_3.

THEOREM 20.1. *Let V_3 be a normal family. Moreover, assume that T is an open interval on the real line. Then to each $f \in C(X)$ there are at most 3 local best approximations in V_3.*

Examples of functions with three LBAs are given in [9]. Therefore, three is indeed the optimal bound.

Proof of Theorem 20.1. It is sufficient to consider functions f which are not exceptional with respect to V_2. In the interests of simplicity we have confined ourselves to extended totally positive kernels γ. With some cost in simplicity γ could equally well have been extended sign regular. One best approximation in V_2 is denoted as $\hat{F}$. We distinguish three cases.

Case 1. $\hat{F} \in V_2^+$ or $(-\hat{F}) \in V_2^+$.

Then $\hat{F}$ is the unique LBA to f in V_2. Since at most three maximal components in V_3 branch from $\hat{F}$, the proof is complete.

Case 2. $\hat{F} \in V_2^0 \backslash V_2^+$ and there is another LBA $\tilde{F}$ in V_2.

At first we will prove that only two maximal components branch from $\hat{F}$. Write

$$\hat{F}(x) = \alpha_1 \gamma(t_1, x) + \alpha_2 \gamma(t_2, x), \qquad \alpha_1 > 0, \alpha_2 < 0 \tag{20.1}$$

After replacing $f(t)$ by $-f(-t)$ if necessary, we assume that $f - \hat{F}$ has a positive alternant. From Theorem 4.5 it follows that each better approximation has at least two positive (generalized) factors. Hence, if we perform the standard construction with $\tau = t_1 + \delta$ or $t_1 - \delta$, then the terms with characteristic numbers in the neighborhood of t_1 have positive factors. Therefore, they will not coalesce. From the remark after discussing Case 2 in Section 19 we conclude that the standard construction with $\tau = t_1 + \delta$ and $\tau = t_1 - \delta$ yield the same component. This proves that no more than two components branch from $\hat{F}$.

Let $\tilde{F}$ be the second LBA to f in V_2. Since the characteristic numbers of $\tilde{F}$ coincide, there are only two possibilities to choose an additional characteristic number τ. What happens, if we put $\tau = t_1 \in \operatorname{spect}(\hat{F})$? By the characterization theorem [10, Satz 3.2] $\hat{F}$ is the unique LBA to f in the restricted set $V_3(t_1)$. Hence, all level sets in $V_3(t_1)$ are connected and $\hat{F}$ must be contained in the level set just constructed. It may be discarded. Consequently, we obtain only one LBA in V_3 when starting with $\tilde{F}$.

Case 3. $\hat{F} \in V_2 \backslash V_2^0$ and there is another LBA $\tilde{F}$ in V_2.

Since the characteristic numbers of $\hat{F}$ and of F, respectively coalesce, from each of them at most two components in V_3 branch, Consequently, we are ready, if only one constructed component is relevant when starting with $\hat{F}$ or $\tilde{F}$. From the remark after Case 2 in Section 19 we conclude that otherwise each level set contains γ-polynomials for which all 3 characteristic numbers coalesce. It follows that we may start the iteration process with the local solutions in $\bigcup_{t \in T} V_3(3 \times t)$. But this set contains at most three local solutions [9, Theorem 5.2].

Thus the proof of Theorem 20.1 is complete. ∎

We remark that the arguments in the proof above may also be used for an improvement of Theorem 19.5. At most $N - 1$ components will branch from any $F \in V_{N-1} \backslash V_{N-1}^+$. If on the other hand $F \in V_{N-1}^+$, then F is unique. This leads to the recursion relation

$$c_N \leqslant (N-1)\, c_{N-1}, \qquad N \geqslant 3.$$

Consequently,

$$c_k = k, \qquad k = 1, 2, 3$$

$$c_k \leqslant \tfrac{3}{2}(k-1)!, \qquad k \geqslant 4.$$

ACKNOWLEDGMENT

The author wishes to express his gratitude to E. W. Cheney for his invitation to the University of Texas, because during his stay in Austin he learned of the concept of generic properties, which is central for this paper.

REFERENCES

1. R. B. BARRAR AND H. L. LOEB, On the continuity of the nonlinear Tschebyscheff operator, *Pacific J. Math.* **32** (1970), 593–601.
2. A. BAUMANN, Die Konstruktion bester Tschebyscheff-Approximationen mit Splines bei freien Knoten, Schriftenreihe des Rechenzentrums, Universität Münster, No. 11 (1974).
3. 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.
4. D. BRAESS, Über Approximation mit Exponentialsummen, *Computing* **2** (1967), 309–321.
5. D. BRAESS, Die Konstruktion der Tschebyscheff-Approximierenden bei der Anpassung mit Exponentialsummen, *J. Approximation Theory* **3** (1970), 261–273.
6. D. BRAESS, Chebyshev approximation by γ-polynomials, I, *J. Approximation Theory* **9** (1973), 20–43; part II, **11** (1974), 16–37.
7. D. BRAESS, Global analysis and Chebyshev approximation by exponentials, *in* "Approximation Theory" (G. G. Lorentz, ed.,), pp. 277–282, Academic Press, New York/London, 1973.
8. D. BRAESS, Kritische Punkte bei der nichtlinearen Tschebyscheff-Approximation, *Math. Z.* **132** (1973), 327–341.
9. D. BRAESS, On the number of best approximations in certain nonlinear families of functions. *Aequationes Math.* **12** (1975), 184–199.
10. J. BRINK-SPALINK, Tschebyscheff-Approximation durch γ-Polynome mit teilweise fixierten Frequenzen. *J. Approximation Theory* **15** (1975), 60–77.
11. J. BRINK-SPALINK, Eine Regularisierung der Splines mit freien Knoten, Schriftenreihe des Rechenzentrums, Universität Münster, No. 23 (1977).
12. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rational Mech. Anal.* **27** (1967), 91–106.
13. D. W. KAMMLER, Existence of best approximations by sums of exponentials. *J. Approximation Theory* **9** (1973), 173–191.
14. S. KARLIN, "Total Positivity," Vol. I, Stanford University Press, Stanford, 1968.
15. Z. NITECKI, "Differentiable Dynamics," MIT Press, Cambridge/London, 1971.
16. E. SCHMIDT, Zur Kompaktheit bei Exponentialsummen, *J. Approximation Theory* **3** (1970), 445–454.
17. E. SCHMIDT, Stetigkeitsaussagen bei der Tschebyscheff-Approximation mit Exponentialsummen, *Math. Z.* **113** (1970), 159–170.
18. J. M. WOLFE, On the unicity of nonlinear approximation in smooth spaces, *J. Approximation Theory* **12** (1974), 165–181.