JOURNAL OF APPROXIMATION THEORY 9, 20–43 (1973)

Chebyshev Approximation by γ-Polynomials

DIETRICH BRAESS

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

Communicated by G. Meinardus

Received: July 19, 1970

1. INTRODUCTION

Let X be an interval on the real axis, and let $C(X)$ be endowed with a weighted Chebyshev-norm. We consider the approximation of real-valued functions $f(x)$ by γ-polynomials [7]

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

where T is a subset of $\mathbf{R}$ and $\gamma \in C(T \times X)$. Interesting examples of kernels $\gamma(t, x)$ are e^{tx}, $\cosh tx$, x^t, arctg tx, $(1 + tx)^{-1}$ and $(x - t)_+^n$. The interest in γ-polynomials stems from approximation by exponentials and by splines. A uniform theory can be formulated, since the functions

$$\gamma(t_1, x), \gamma(t_2, x), \ldots, \gamma(t_N, x)$$

form a Chebyshev system for distinct t_i with the above mentioned kernels, except for $(x - t)_+^n$. For this kernel, one obtains a weak Chebyshev system [11], and, therefore, the corresponding splines require some special consideration.

For Hobby and Rice [7, 17], as well as for de Boor [2], the existence of best approximations was of main interest. They noticed that one has to consider the closure of the families. If the derivatives $\gamma^{(\mu)} = (\partial^\mu / \partial t^\mu) \gamma$ exist[1] and are continuous in $T \times X$, one has to adjoin the extended γ-polynomials

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

[1] Only derivatives in t are used in this paper.

to the proper γ-polynomials (1.1). However, the uniqueness of the best approximation may be lost through this extension, as is known in the case of approximation by exponentials [3]. In addition, local best approximation may exist [4], which makes the computation of best approximation more difficult [5].

For these reasons, one is interested in characteristics of best approximation. For best approximation in the sense of Chebyshev, we shall draw far-reaching conclusions from the fact, stated by Karlin [10], that Haar's condition implies a generalized Descartes rule. Aside from uniqueness theorems and alternant-criteria, we obtain results about generalized signs. In this way we study the topological structure of the families. In addition, the case $N = 2$ is treated completely. The problem of approximation with positive factors occupies a special position. We shall see that in this nonlinear theory, not only the length of the alternant but also the sign of the error-function yields important information.

2. Sign-Regular and Totally Positive Kernels

Let T and X be subsets of $\mathbf{R}$ and let $\gamma(t, x) \in C(T \times X)$. Karlin [10] considered the determinants

$$\gamma\begin{pmatrix} t_1, t_2, ..., t_r \\ x_1, x_2, ..., x_r \end{pmatrix} = \begin{vmatrix} \gamma(t_1, x_1) & \gamma(t_1, x_2) & \cdots & \gamma(t_1, x_r) \\ \gamma(t_2, x_1) & \gamma(t_2, x_2) & \cdots & \gamma(t_2, x_r) \\ & & \cdots & \\ \gamma(t_r, x_1) & \gamma(t_r, x_2) & \cdots & \gamma(t_r, x_r) \end{vmatrix}. \tag{2.1}$$

Definition 2.1. Let $\gamma(x, t) \in C(T \times X)$. If there exist $\epsilon_1, \epsilon_2, ..., \epsilon_r$, each either $+1$ or -1, such that, for all

$$\begin{matrix} x_1 < x_2 < \cdots < x_p, & x_i \in X \\ t_1 < t_2 < \cdots < t_p, & t_i \in T \end{matrix} \qquad (1 \leqslant p \leqslant r) \tag{2.2}$$

the relation

$$\epsilon_p \cdot \gamma\begin{pmatrix} t_1, t_2, ..., t_p \\ x_1, x_2, ..., x_p \end{pmatrix} > 0 \tag{2.3}$$

holds, then the kernel $\gamma(t, x)$ is called strictly sign-regular of order r (abbreviated: SSR_r). If $\epsilon_1, \epsilon_2, ..., \epsilon_r$ are all positive, $\gamma(t, x)$ is a strictly totally positive kernel of order r (STP_r). If such a property of $\gamma(t, x)$ holds for every r then r is omitted.

For connected sets T and X, being SSR_r is obviously equivalent to $\gamma(t_1, x)$, $\gamma(t_2, x), ..., \gamma(t_r, x)$ forming a Chebyshev system for distinct t_i. The theory can be generalized to the extended γ-polynomials (1.2). In the determinant (2.1) we follow the convention of Karlin: If $t_m, t_{m+1}, \ldots, t_{m+k}$ is a block of coincident arguments, then the $(m+p)$ th row in (2.1) is to be replaced by $\gamma(t_m, x_i), \gamma^{(1)}(t_m, x_i), \ldots, \gamma^{(k)}(t_m, x_i)$, for $p = 1, 2, \ldots, k$, i.e., we replace (2.1) by

$$\gamma^* \begin{pmatrix} t_1, t_2, \ldots, t_r \\ x_1, x_2, \ldots, x_r \end{pmatrix} = \begin{vmatrix} \gamma(t_1, x_1) & \gamma(t_1, x_2) & \cdots & \gamma(t_1, x_r) \\ & \cdots & & \\ \gamma(t_m, x_1) & \gamma(t_m, x_2) & \cdots & \gamma(t_m, x_r) \\ \gamma^{(1)}(t_m, x_1) & \gamma^{(1)}(t_m, x_2) & \cdots & \gamma^{(1)}(t_m, x_r) \\ & \cdots & & \\ \gamma^{(k)}(t_m, x_1) & \gamma^{(k)}(t_m, x_2) & \cdots & \gamma^{(k)}(t_m, x_r) \\ & \cdots & & \end{vmatrix}. \tag{2.4}$$

DEFINITION 2.2. Let $\gamma(t, x)$ be $(r-1)$ times differentiable in t, let $(\partial^{r-1}/\partial t^{r-1})\, \gamma \in C(T \times X)$, and let each $\epsilon_1, \epsilon_2, \ldots, \epsilon_r$ be 1 or -1. Assume that

$$\epsilon_p \cdot \gamma^* \begin{pmatrix} t_1, t_2, \ldots, t_p \\ x_1, x_2, \ldots, x_p \end{pmatrix} > 0 \qquad (1 \leqslant p \leqslant r) \tag{2.5}$$

for all $t_1 \leqslant t_2 \leqslant \cdots \leqslant t_p$ $(t_i \in T)$ and for all $x_1 < x_2 < \cdots < x_r (x_i \in X)$. Then $\gamma(t, x)$ is an extended sign-regular kernel of order r in the t variable ($\mathrm{ESR}_r(t)$). If all of $\epsilon_1, \epsilon_2, \ldots, \epsilon_r$ are $+1$, $\epsilon(t, x)$ is an extended totally positive kernel ($\mathrm{ETP}_r(t)$).

3. GENERALIZED SIGNS

The extended γ-polynomials can be written in the form

$$F[a] = F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \gamma^{(\mu)}(t_\nu, x) \tag{3.1}$$

with $t_1 < t_2 < \cdots < t_l$ and $\alpha_{\nu M_\nu} \neq 0$ for $\nu = 1, 2, \ldots, l$. Here

$$k = k(a) = \sum_{\nu=1}^{l} (1 + M_\nu)$$

is the order of the γ-polynomial. The order coincides with the length of the γ-polynomial $l = l(a)$ if $F(a, x)$ has the special form (1.1). The parameters t_ν are called characteristic numbers of $F(a, x)$, and the set $\{t_\nu, \nu = 1, 2, \ldots, l(a)\}$ is its spectrum. The parameters $\alpha_{\nu\mu}$ are called factors of $F(a, x)$.

For every γ-polynomial of order k, we define recursively a sign-vector $\{s_1, s_2, \ldots, s_k\}$ with k components [4].

DEFINITION 3.1. (a) For the special γ-polynomials having a single characteristic number:

$$F(x) = \sum_{\mu=0}^{M} \alpha_\mu \gamma^{(\mu)}(t, x), \qquad \alpha_M \neq 0, \qquad \operatorname{sign} \alpha_M = \sigma,$$

we set

$$\operatorname{sign} F = \underbrace{\{(-1)^M \cdot \sigma, (-1)^{M-1} \sigma, \ldots, \sigma, -\sigma, \sigma\}}_{M+1}. \tag{3.2}$$

(b) If all characteristic numbers of F_1 are smaller than each of F_2, the following composition rule holds:

$$\operatorname{sign}(F_1 + F_2) = \{\operatorname{sign} F_1, \operatorname{sign} F_2\}.$$

The components of sign F are called the generalized signs of F or the generalized signs of the factors. The number of positive (negative, resp.) signs is denoted $k^+(a)$ $k^-(a)$, resp.). Obviously,

$$k^+(a) + k^-(a) = k(a). \tag{3.3}$$

If $F(x)$ is a proper γ-polynomial (1.1), sign α_j is 0 or s_j $(j = 1, 2, \ldots, k)$, provided $t_1 < t_2 < \cdots < t_N$. To show that Definition 3.1 is plausible for extended γ-polynomials, we consider a limiting process, which will be used in several proofs. For sufficiently differentiable kernels,

$$\sum_{\mu=0}^{M} \beta_\mu \gamma^{(\mu)}(t, x) = \lim_{\substack{t_n \to t \\ t_n \neq t_m \\ n,m=1,\ldots,M+1}} \sum_{\mu=0}^{M} \mu!\, \beta_\mu \gamma(t_1, t_2, \ldots, t_{\mu+1}; x). \tag{3.4}$$

The divided differences of a function $\varphi(t)$ are defined as usual [9]. Equality (3.4) follows from the existence of a mean value τ with

$$\varphi(t_1, t_2, \ldots, t_{\mu+1}) = (1/\mu!)\, \varphi^{(\mu)}(\tau) \tag{3.5}$$

(see, for instance, [9]). Using the formula

$$\varphi(t_1, t_2, \ldots, t_{\mu+1}) = \sum_{n=1}^{\mu+1} \varphi(t_n) \cdot \prod_{\substack{n=1 \\ m\neq n}}^{\mu+1} \frac{1}{t_n - t_m} \tag{3.6}$$

(t_i distinct), we get

$$\sum_{\mu=0}^{M} \beta_\mu \cdot \gamma^{(\mu)}(t, x) = \lim_{\substack{t_n \to t \\ t_n \neq t_m \\ n,m=1,2,\ldots,M+1}} \sum_{n=1}^{M+1} \gamma(t_n, x) \cdot \prod_{m\neq n} \frac{1}{t_n - t_m}$$

$$\times \left\{ M!\, \beta_M + \sum_{\mu=n}^{M} (\mu - 1)!\, \beta_{\mu-1} \prod_{m=\mu+1}^{M+1} (t_n - t_m) \right\}. \tag{3.7}$$

Assume that $\beta_M \neq 0$. If the characteristic numbers are sufficiently close to each other, then the value in the curly brackets takes on the same sign as β_M. We also get

$$\operatorname{sign} \prod_{m \neq n} \frac{1}{t_n - t_m} = (-1)^{p_n}$$

with

$$p_n = \text{number of characteristic numbers } t_m \text{ larger than } t_n .$$

On the right side of (3.7), the coefficients of $\gamma(t_n, x)$ are alternatingly positive and negative, and the coefficient of γ with the largest t_n has the same sign as β_M. This corresponds exactly to Definition 3.1. Setting (see (2.3))

$$\tilde{\epsilon}_1 = \epsilon_1, \quad \tilde{\epsilon}_p = \epsilon_{p-1}/\epsilon_p \; (p > 1),$$

we get a generalized Descartes rule.

LEMMA 3.1. *Let $\gamma(t, x)$ be strictly sign-regular of order r, and let the extended γ-polynomial* (3.1) *satisfy*

$$(-1)^{i-1} F(x_i) > 0, \qquad i = 1, 2,..., r, \tag{3.8}$$

where $x_1 < x_2 < \cdots < x_r$. Then there are at least $r - 1$ sign-changes in the sequence $s_1, s_2,..., s_r$ of generalized signs of F. If the number of changes is $r - 1$, then

$$s_1 = \tilde{\epsilon}_r \cdot \operatorname{sign} F(x_1),$$
$$s_k = \tilde{\epsilon}_r \cdot \operatorname{sign} F(x_r).$$

Proof. If $F(x)$ is a proper γ-polynomial, the statement is a consequence of Theorem 1.2 or a specialization of Theorem 3.1 and of Theorem 1.5 [10, Chapter 5]. In the general case, set $\delta > 0$, and consider the proper γ-polynomial

$$F_\delta(x) = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \cdot \mu! \, \gamma(t_\nu, t_\nu + \delta ... t_\nu + \mu\delta; x).$$

Relations (3.4) and (3.8) yield, for a sufficiently small δ,

$$(-1)^{i-1} F_\delta(x_i) > 0, \qquad i = 1, 2,..., r.$$

Thus, the lemma holds for $F_\delta(x)$. Since the generalized signs of F and of F_δ coincide for sufficiently small δ, the lemma holds for extended polynomials □

If the kernel γ is ESR, a stronger result holds:

THEOREM 3.2. *Let the extended γ-polynomial $F(x)$ of order k satisfy*

$$(-1)^{i-1} F(x_i) \geqslant 0, \qquad i = 1, 2,..., r, \tag{3.9}$$

where $x_1 < x_2 < \cdots < x_r$, *and assume the kernel to be* $\mathrm{ESR}_{\min(k,r)}(t)$. *Then, if* $F \not\equiv 0$, *in the sequence sign F there are at least* $r - 1$ *sign-changes. If the number of sign-changes is* $r - 1$, *we have*

$$s_1 = \tilde{\epsilon}_r , \qquad s_k = \tilde{\epsilon}_r \cdot (-1)^{r-1}. \tag{3.10}$$

Remark. If $F(x)$ is a proper γ-polynomial, Theorem 3.2 can be deduced from the weaker assumption that γ is $\mathrm{SSR}_{\min(k,r)}$.

Proof of Theorem 3.2. We distinguish two cases.

(1) Let $k \geqslant r$. One can choose numbers

$$M_\nu' \leqslant M_\nu , \quad \nu = 1, 2,..., l, \qquad \text{such that} \qquad \sum (1 + M_\nu') = r.$$

Since γ is $\mathrm{ESR}_r(t)$, there exists a γ-polynomial

$$G(x) = \sum_{\nu=1}^{l'} \sum_{\mu=0}^{M_\nu'} \beta_{\nu\mu} \gamma^{(\mu)}(t_\nu , x),$$

which solves the interpolation problem

$$(-1)^{i-1} G(x_i) = +1, \qquad i = 1, 2,..., r.$$

Hence, for every positive δ we have $(-1)^{i-1} [F(x_i) + \delta G(x_i)] > 0$. This, together with Lemma 3.1, prove our conclusion for $\mathrm{sign}(F + \delta G)$. For sufficiently small δ, the latter equals sign F. This completes the proof in Case 1.

(2) Let $k < r$. It follows from Lemma 4.2 of [11, chapter 1] that F vanishes identically, since it has at least k zeros, counting nonnodal zeros twice. □

4. Uniqueness, Sign-Distribution

In the following, let X be a compact interval, and let the space $C(X)$ be endowed with a (weighted) Chebyshev norm:

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

with $w \in C(X)$, $w(x) > 0$ for $x \in X$. Let $V \subset C(X)$. Then $F^* \in V$ is a best approximation to f in V, if

$$\| f - F^* \| = \inf\{\| f - F \| ; F \in V\}.$$

DEFINITION 4.1. If the kernel $\gamma(t, x)$ is SSR_{2N}, then

$$V_N{}^0 = \left\{ \sum_{\nu=1}^{k} \alpha_\nu \gamma(t_\nu, x), \alpha_\nu \in \mathbf{R}, t_\nu \in T, k \leqslant N \right\} \tag{4.1}$$

is called a Descartes family.

Throughout the rest of this paper, T is assumed to be an open set, unless otherwise stated.

Following Rice [15], we make

DEFINITION 4.2. A family $V \subset C(X)$ is called varisolvent of degree $m = m(a)$ at $F[a] \in V$, if the following conditions are satisfied:

(1) For all $F[b] \in V$, the difference $F(b, x) - F(a, x)$ has at most $m - 1$ zeros in X or vanishes identically.

(2) For any $x_1 < x_2 < \cdots < x_m$ and any $\epsilon > 0$, there is a $\delta = \delta(a, \epsilon, x_1, ..., x_m)$ such that $| F(a, x_i) - y_i | < \delta$ implies the existence of a function $F[b] \in V$ with

$$F(b, x_i) = y_i, \qquad i = 1, 2, ..., m,$$

and $\| F[b] - F[a] \| \leqslant \epsilon$.

THEOREM 4.1. *Every Descartes family $V_N{}^0$ is varisolvent of degree $N + k(a)$ at $F[a]$.*

Proof. (1) For all $F[b] \in V_N{}^0$, the difference $F[b] - F[a]$ is a γ-polynomial of order $\leqslant N + k(a)$. Thus in this case, (1) of Definition 4.2 is a consequence of γ being SSR_{2N}.

(2) Assume $F(a, x)$ to be of the form (4.1), and let $\alpha_\nu \neq 0$ for $\nu = 1, ..., k$, and $t_1 < t_2 < \cdots < t_k$. Choose numbers $t_\nu \in T$ $(\nu = k + 1, k + 2, ..., N)$ with $t_N > t_{N-1} > \cdots > t_{k+1} > t_k$. Furthermore, set $m = N + k$, and consider the γ-polynomials

$$G(u, x) = \sum_{\nu=1}^{k} u_\nu \gamma(u_{N+\nu}, x) + \sum_{\nu=k+1}^{N} u_\nu \gamma(t_\nu, x) \tag{4.2}$$

corresponding to the vectors $u = (u_1, u_2, ..., u_m)$ of the set

$$U = \{u \in \mathbf{R}^m \mid u_\nu \neq 0, u_{N+\nu} \in T, \nu = 1, 2, ..., k, \\ u_{N+1} < u_{N+2} < \cdots < u_{N+k} < t_{k+1}\}. \tag{4.3}$$

Given distinct points $x_i \in X$, $i = 1, 2, ..., m$, we define the mapping

$$\Phi = (v_1, v_2, ..., v_m) : U \to \mathbf{R}^m$$

by

$$v_i = G(u, x_i), \quad i = 1, 2, ..., m.$$

This mapping is continuous. It is also injective, because the difference of two γ-polynomials (4.2) has at most the order m, and with $m \leqslant 2N$, γ is SSR_m.

Hence, Φ is a homeomorphism. For

$$u_0 = \{\alpha_1, \alpha_2, ..., \alpha_k, 0, 0, ..., 0, t_1, t_2, ..., t_k\},$$

$F(a, x) = G(u_0, x)$. As u_0 is an interior point of U, the vector $v_0 = \{F(a, x_1), ..., F(a, x_m)\}$ is interior to $\Phi(U)$ by Brouwer's theorem on the invariance of domain [8]. In addition, the mapping

$$\begin{aligned} &\Psi: U \to C(X), \\ &\Psi(u) = G(u, x) \end{aligned} \tag{4.4}$$

is continuous.[2] This proves (2) of Definition (4.2). □

DEFINITION 4.3. Given a γ-polynomial $F[a]$ and an $f \in C(X)(f(x) \not\equiv F(a, x))$, we call $\epsilon(a, x) = w(x) \cdot [f(x) - F(a, x)]$ a (weighted) error function. If $x_1 < x_2 < \cdots < x_m$ are extreme points, namely, if

$$|\epsilon(a, x_i)| = \|f - F[a]\|, \qquad i = 1, 2, ..., m, \tag{4.5}$$

and if

$$\epsilon(a, x_i) = -\epsilon(a, x_{i-1}), \qquad i = 2, 3, ..., m, \tag{4.6}$$

then the sequence $(x_1, x_2, ..., x_m)$ is called an alternant of length m; such an alternant has the sign σ on the right (on left, resp.), if the sign of

$$\epsilon(a, x_m) \quad (\text{of } \epsilon(a, x_i), \text{ resp.}) \text{ is } \sigma.$$

Remark. From every alternant of length $m + 1$ one can select one of length m, which has a desired sign on the right or on the left. If the length of an alternant is odd its signs on both ends are equal, and the specification "right" or "left" may be omitted.

THEOREM 4.2. *Let $f \in C(X)$ and let $V_N{}^0$ be a Descartes family.*

(a) *There exists at most one best approximation to f in $V_N{}^0$.*

(b) *$F[a]$ is a best approximation to f in $V_N{}^0$ if and only if an alternant of length $N + k(a) + 1$ exists for $F[a]$.*

Proof. Let $F(a, x) = \sum_{\nu=1}^{k} \alpha_\nu \gamma(t_\nu, x)$ be a best approximation to f. Suppose the error function is a nonzero constant. Then either

$$F(a, x) + \delta \cdot \gamma(t_1, x) \text{ or } F(a, x) - \delta\gamma(t_1, x)$$

[2] This holds, because $\gamma(t, x)$ is uniformly continuous in $T_0 \times X$, where T_0 is any compact subset of T.

would be a better approximation to f, for a sufficiently small $\delta > 0$, since $\gamma(t_1, x)$ has no zero. Thus this case, mentioned by Dunham [6a] and by Barrar and Loeb [1], is excluded, and the results of Rice about varisolvent families [15, 17] yield the conclusions of the theorem. □

The following will be repeatedly used in the sequel. Let us assume that there exists an alternant for $F[a]$ of length r with sign σ on the right

$$\sigma \cdot (-1)^{r-i} \cdot \epsilon(a, x_i) = \|f - F[a]\|, \qquad i = 1, 2,..., r. \tag{4.7}$$

For each F[b] which is at least as good an approximation to f as $F[a]$, we have from (4.7):

$$\sigma \cdot (-1)^{r-1} \cdot \epsilon(b, x_i) \leqslant | \epsilon(b, x_i)| \leqslant \|f - F[a]\|$$
$$= \sigma \cdot (-1)^{r-1} \epsilon(a, x_i), \qquad i = 1, 2,..., r.$$

By substituting $\epsilon(a, x) = w(x) \cdot [f(w) - F(a, x)]$ and the corresponding expression for $\epsilon(b, x)$, and by dividing by $w(x_i)$, we get

$$\sigma \cdot (-1)^{r-1} [F(b, x_i) - F(a, x_i)] \geqslant 0, \qquad i = 1, 2,..., r. \tag{4.8}$$

If $F[b]$ is even a better approximation than $F[a]$, then in (4.8) strict inequalities hold.

THEOREM 4.3. *Let* $F[a]$, resp. $F[b]$, *be the best approximation to* $f \in C(X)$ *in the Descartes family* $V_m{}^0$, resp. $V_n{}^0$. *If* $n \geqslant m$, *then* $F[b]$ *contains at least as many positive, and at least as many negative factors as* $F[a]$.

Remark. The last conclusion holds even for a γ-polynomial $F[b]$ which approximates f at least as good as $F[a]$. It is also true for the generalized signs of extended γ-polynomials which approximate f better than $F[a]$ (cf. Lemma 3.1 for the proof).

Proof of Theorem 4.3. By Theorem 4.2, there is an alternant of length $r = m + k(a) + 1$ for $F[a]$. Assume that $F[b]$ approximates f at least as good as $F[a]$. Only the case $F[a] \neq F[b]$ has to be considered. Then (4.8) holds, and Theorem 3.2 asserts that the factors of the difference $F[b] - F[a]$ change signs at least $m + k(a)$ times, if the terms are reordered according to the size of the characteristic numbers. There are at least $(m + k(a))/2$ positive and at least $(m + k(a))/2$ negative factors in the difference. The positive factors of the difference stem from positive factors of $F[b]$ or negative factors of $F[a]$. Hence, using (3.3), we have

$$k^+(b) + k^-(a) \geqslant (m + k(a))/2 = k^+(a) + k^-(a) + (m - k(a))/2.$$

Since $k(a) \leqslant m$, we obtain

$$k^+(b) \geqslant k^+(a) + (m - k(a))/2 \geqslant k^+(a), \tag{4.9}$$

and in the same way we have $k^-(b) \geqslant k^-(a)$. □

If $m > k(a)$ the last theorem can be sharpened. For the derivation of (4.9), only the fact that γ is $\mathrm{SSR}_{m+k(a)+1}$ was used.

COROLLARY 4.4. *Let the best approximation $F[a]$ to f in the Descartes family $V_n{}^0$ exist and be of order $k < n$. Then each better approximating γ-polynomial has one more positive, and one more negative factor than $F[a]$ does.*

In the next section, a sharper result will be required.

THEOREM 4.5. *Let $F[a]$ be a best approximation to $f \in C(X)$ in the Descartes family $V_N{}^0$. Furthermore, let γ be* SSR_{2N+1}. *If $F[a]$ has an alternant with sign $+\tilde{\epsilon}_{2N+1}(-\tilde{\epsilon}_{2N+1}$,* resp.), *then each better approximating γ-polynomial has at least one more positive (negative,* resp.) *factor than $F[a]$ does.*

Proof. It is only necessary to consider the case of an alternant of length $2N + 1$, because otherwise Corollary 4.4 can be applied. Assume $F[b]$ is a better approximation. From (4.8) and Theorem 3.1 it follows that there are at least $2N$ sign-changes in the difference $F[b] - F[a]$, and if the number of sign-changes is exactly $2N$, then the factor of the term with the highest characteristic number is positive. In all cases, the difference contains at least $N + 1$ positive factors. Hence, as in the proof of Theorem 4.3,

$$k^+(b) + k^+(a) \geqslant N + 1 \geqslant k(a) + 1 = k^+(a) + k^-(a) + 1. \quad \square$$

For the construction of best approximations, as suggested in [19], the following is useful.

THEOREM 4.6. *Let $f \in C(X)$, and assume the best approximations to f in the Descartes families $V_N{}^0$ and V_{N-1}^0 exist and are different. Let t_ν and $t_{\nu+1}$ be two consecutive characteristic numbers of the best approximation in V_{N-1}^0. If the associated factors α_ν and $\alpha_{\nu+1}$ have the same sign (the opposite sign,* resp.), *then the interval $(t_\nu, t_{\nu+1})$ contains an odd* (resp., *even*) *number of characteristic numbers of the best approximation in $V_N{}^0$.*

The theorem is an immediate consequence of the fact that there are exactly $N + k(a_{N-1}) - 1$ sign changes in the difference between the two best approximations.

5. Positive Sums

The existence of a best approximation can be guaranteed only if the family in question is closed. Normally, it is considerably easier to determine the closure of a family by allowing only sums with positive factors, i.e., by considering families

$$V_N^+ = \left\{ F(a, x) \mid F(a, x) = \sum_{\nu=1}^{k} \alpha_\nu \gamma(t_\nu, x), \alpha_\nu \geqslant 0, t_\nu \in T, k \leqslant N \right\}.$$

If, for example, $\gamma(t, x) = e^{tx}$ or $\gamma(t, x) = \operatorname{arctg} tx/\operatorname{arctg} t$, then V_N^+ is closed.

Theorem 5.1. *Let $f \in C(X)$ and let γ be* SSR_{2N}.

(a) *There is at most one best approximation to f in V_N^+.*

(b) *A best approximation $F[a]$ in V_N^+ is also a best approximation in $V_{k(a)}^0$.*

Proof. Let $F[a^*]$ be a best approximation in V_N^+ of order k^*. In case there are several best approximations, choose $F[a^*]$ to be one with a maximal order. Obviously, $F[a^*]$ is a best approximation in the subset

$$\left\{ F(a, x) = \sum_{\nu=1}^{k^*} \alpha_\nu \gamma(t_\nu, x) \mid \alpha_\nu > 0, t_1 < t_2 < \cdots < t_{k^*} \right\}.$$

This set is open in V_{k^*}, and varisolvent with the constant degree $2k^*$.

According to Rice [11], there is an alternant of length $2k^* + 1$, and it follows from Theorem 4.1 that $F[a^*]$ is the unique best approximation in $V_{k^*}^0$. Since $V_{k^*}^+ \subset V_{k^*}^0$, $F[a^*]$ is unique in $V_{k^*}^+$. By our choice of $F[a^*]$, uniqueness is assured even in V_N^+. □

Finally, we obtain an alternant criterion which uses not only the length of the alternant but also the sign of the error function.

Theorem 5.2. *Let $f \in C(X)$ and let γ be* SSR_{2N}. *$F[a]$ is a best approximation in V_N^** iff *one of the following conditions holds*:

(1) *There is an alternant of length $2N + 1$.*

(2) *There is an alternant of length $2k(a) + 1$ with the* sign $-\tilde{\epsilon}_{2k(a)+1}$.

Proof.

(1) Assume $k = k(a) = N$. Then (1) is a necessary and sufficient condition by Theorem 5.1b and Theorem 4.2b. Condition (2) is of no interest, because it is more restrictive than (1).

(2) Assume $k = k(a) < N$. If there is an alternant for $F[a]$ of length $2k + 1$, with the sign $-\tilde{\epsilon}_{2k+1}$, then each better approximating γ-polynomial contains a negative factor, by Theorem 4.5. Hence, $F[a]$ is a best approximation in $V_N{}^+$.

On the other hand, if there is no alternant of length $2k + 1$ with sign $-\tilde{\epsilon}_{2k+1}$, then there is none of length $2k + 2$. It follows from Theorem 4.2b that a better approximation exists in V_{k+1}^0, which, according to Theorem 4.5, must have one positive factor more than $F[a]$ does, and thus, it is contained in $V_{k+1}^+ \subset V_N{}^+$. Hence, $F[a]$ is not a best approximation. □

Since condition (2) does not include N, and in the proof γ needed only be $\mathrm{SSR}_{\min(2N,2k+1)}$, we have the following.

COROLLARY 5.3. *Let $f \in C(X)$ and let γ be* SSR_{2N}.
If the best approximation in $V_N{}^+$ exists and is of order $k < N$, it is also the best approximation in $V_M{}^+$ for all $M > N$.

As a specialization of Theorem 4.6 we have the following separation theorem.

THEOREM 5.4. *Let $f \in C(X)$, and let γ be* SSR_{2N}. *If the best approximations in $V_N{}^+$ and V_{N+1}^+ exist, then either they coincide or their characteristic numbers seperate each other.*

6. EXTENDED DESCARTES FAMILIES

DEFINITION 6.1. Assume that the kernel γ is $\mathrm{ESR}_{2N}(t)$. Then the set of extended γ-polynomials

$$V_N = \left\{ F(a, x) = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \gamma^{(\mu)}(t_\nu, x) \mid \alpha_{\nu\mu} \in \mathbf{R}, t_\nu \in T, k = \sum_{\nu=1}^{l} (1 + M_\nu) \leqslant N \right\} \tag{6.1}$$

is called an extended Descartes family.

Such extended families are studied to enable one to prove existence theorems (cf. [2, 7, 17, 19]). On the other hand, for $N \geqslant 2$, these extended families are neither varisolvent, nor asymptotically convex [14], nor are they suns in the sense of Vlasov [6]. A nonuniqueness result for such families is given in Theorem 8.7. Consequently, we cannot expect alternant conditions here which are both necessary and sufficient.

THEOREM 6.1. *Let $f \in C(X)$ and let V_N be an extended Descartes family.*

(a) *If there is an alternant of length $N + k(a) + 1$ for $F[a]$, then $F[a]$ is the unique best approximation to f in V_N.*

(b) *If $F[a]$ is a best approximation to f in V_N, then there is an alternant of length $N + l(a) + 1$ for $F[a]$.*

Proof. (a) Suppose there is an approximation $F[b]$ at least as good as $F(a)$, with $F(b, x) \not\equiv F(a, x)$. Then there are at least $N + k(a)$ sign-changes in the difference $F[a] - F[b]$, according to Theorem 3.2 and formula (4.8). But this is impossible for a γ-polynomial of order $\leqslant N + k(a)$.

(b) We write the best approximation $F[a]$ in the form

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

with $\alpha_\nu = 0$, $\nu = k + 1,..., N$, the numbers t_ν, $\nu = k + 1,..., N$, being distinct and not belonging to the spectrum of $F[a]$. The derivatives

$$\partial F/\partial \alpha_{\nu\mu} = \gamma^{(\mu)}(t_\nu, x), \qquad \nu = 1, 2,..., l, \quad \mu = 0, 1,..., M_\nu,$$

$$\partial F/\partial t_\nu = \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \gamma^{(\mu+1)}(t_\nu, x), \qquad \nu = 1, 2,..., l,$$

$$\partial F/\partial \alpha_\nu = \gamma(t_\nu, x), \qquad \nu = k + 1, k + 2,..., N,$$

form a basis for the space of functions

$$\sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu+1} \beta_{\nu\mu} \gamma^{(\mu)}(t_\nu, x) + \sum_{\nu=k+1}^{N} \beta_\nu \gamma(t_\nu, x).$$

Since γ is $\mathrm{ESR}_{2N}(t)$, this basis is a Haar system of $k + l + (N - k) = N + l$ elements. Thus, V_N satisfies the local Haar condition [13, 14], and the desired conclusion follows from Meinardus and Schwedt's theorem 12 [14]. □

For proper γ-polynomials, $k(a)$ and $l(a)$ coincide; thus, we have the following under the assumptions of Theorem 6.1.

COROLLARY 6.2. *Every best approximation in V_N^0 is the unique best approximation in V_N.*

7. γ-POLYNOMIALS OF ORDER 2

In this section we consider approximation in V_2. Using an improved alternant criterion, we establish that at most two best approximations exist. To this end, we modify Meinardus and Schwedt's Theorem 8 of [14].

Let $A \subset \mathbf{R}^m$, and for every $a \in A$, let $F(a, x)$ be a real function on the real interval X. Assume that for some fixed $a^* \in A$ and some fixed convex non-degenerate cone $\Delta \subset \mathbf{R}^m$, we have $\{a^* + \delta \mid \delta \in \Delta\} \subset A$. Assume that for every $a = (a_1, a_2, \ldots, a_m) \in A$, each $\partial F/\partial a_\nu$ exists and is continuous in $A \times X$. Then

$$F(a^* + \delta, x) - F(a^*, x) = H(x, \delta) + o(\delta), \qquad \delta = (\delta_1, \delta_2, \ldots, \delta_m) \in \Delta,$$

where

$$H(x, \delta) = \sum_{\nu=1}^{m} \delta_\nu v_\nu(x),$$

$$v_\nu(x) = \partial F(a^*, x)/\partial a_\nu, \qquad \nu = 1, 2, \ldots, m.$$

LEMMA 7.1. *If $F[a^*]$ is a best approximation to f in $V = \{F[a] \mid a \in A\}$, then* 0 *is a best approximation to $\epsilon[a^*] = f - F[a^*]$ in $\{H(x, \delta) \mid \delta \in \Delta\}$.* (7.1)

The proof is similar to that of Theorem 8 of [14].

Now we assume the cone Δ to be a half-space.

LEMMA 7.2. *Let $\Delta = \{\delta \mid \delta \in \mathbf{R}^m, \delta_m \geqslant 0\}$ and let $v_1, v_2, \ldots, v_m$ as well as $v_1, v_2, \ldots, v_{m-1}$ satisfy Haar's condition. Furthermore, let the existence of $m - 1$ zeros of $H(x) = H(x, \delta)$ in $x > x_0$ $(x < x_0)$ imply (for a given $\tilde{s} = \pm 1$):*

$$\operatorname{sign} H(x_0, \delta) = -\tilde{s} \cdot \operatorname{sign} \delta_m. \tag{7.2}$$

If $F[a^*]$ is a best approximation to f in $V = \{F[a] \mid a \in A\}$, then there exists an alternant of length m, with the sign $\tilde{s}$ on the left (on the right).

Proof. Since $\{H(x, \delta) \mid \delta \in \Delta\}$ contains a linear Haar subspace of dimension $m - 1$, Lemma 7.1 implies the existence of an alternant $x_1 < x_2 < \cdots < x_m$ for the best approximation $F[a^*]$. Suppose that the error function has at x_1 the sign opposite to $\tilde{s}$, and that there is no alternant of length $m + 1$. Then, in the linear Haar subspace spanned by $v_1, v_2, \ldots, v_m$, there is an element $H = H(x, \delta)$, $\delta \in \mathbf{R}^m$, satisfying $\|f - F[a^*] - H\| < \|f - F[a^*]\|$. This implies

$$(-1)^i \cdot \tilde{s} \cdot H(x_i, \delta) > 0, \; i = 1, 2, \ldots, m.$$

Hence, $H(x, \delta)$ has $m - 1$ zeros, and by considering the sign of $H(x_1, \delta)$, we obtain $\delta_m > 0$, $\delta \in \Delta$, contradicting Lemma 7.1. □

For the sake of clearer presentation, we state the following results only for Descartes families with totally positive kernels, and we omit the geneneralization for sign regular kernels.

THEOREM 7.3. *Let $f \in C(X)$ and let γ be* ETP$_4(t)$. *If*

$$F[a] = \beta_0 \gamma(t_0, x) + \beta_1 \gamma^{(1)}(t_0, x),$$

with $\beta_1 \neq 0$, is a best approximation to $f(x)$ in V_2, then there exists an alternant of length 4 whose sign on the right is opposite to the sign of β_1.

Proof. We write the γ-polynomials of order degree 2 in the form

$$F(a, x) = a_1 \frac{\gamma(a_3 + \sqrt{a_4}, x) + \gamma(a_3 - \sqrt{a_4}, x)}{2} + a_2 \frac{\gamma(a_3 + \sqrt{a_4}, x) - \gamma(a_3 - \sqrt{a_4}, x)}{2 \cdot \sqrt{a_4}}, \quad \text{with} \quad a_4 \geqslant 0, \tag{7.3}$$

where the second quotient should be interpreted as $\gamma^{(1)}(a_3, x)$ for $a_4 = 0$. This occurs when (7.3) describes an extended γ-polynomial. The spectrum consists of the characteristic numbers $a_3 + \sqrt{a_4}$ and $a_3 - \sqrt{a_4}$. For $a_4 = 0$, we have

$$\begin{aligned}
\partial F/\partial a_1 &= \gamma(a_3, x),\\
\partial F/\partial a_2 &= \gamma^{(1)}(a_3, x),\\
\partial F/\partial a_3 &= a_1\gamma^{(1)}(a_3, x) + a_2\gamma^{(2)}(a_3, x),\\
\partial F/\partial a_4 &= \tfrac{1}{2}a_1\gamma^{(2)}(a_3, x) + \tfrac{1}{6}a_2\gamma^{(3)}(a_3, x).
\end{aligned}$$

If the function

$$H(x) = \sum_{\mu=0}^{3} \delta_\mu \gamma^{(\mu)}(a_3, x)$$

has three zeros $\xi_1 < \xi_2 < \xi_3$, then, since γ is ETP$_4(t)$ and by Theorem 3.2, for $x > \xi_3$, H has the same sign as δ_3. Being a linear combination of $\partial F/\partial a_\nu$, $H(x)$ has, for $x > \xi_3$ the same sign as the product $\delta_4 a_2 = \delta_4 \beta_1$. This, by Lemma 7.2, completes the proof. □

The following theorem shows that the alternant criterion is in a certain sense, also sufficient.

THEOREM 7.4. *Let $f \in C(X)$, and let V_2 be an extended Descartes family with a totally positive kernel. Assume that the γ-polynomial*

$$F(a, x) = \alpha_0\gamma(t_0, x) + \alpha_1\gamma^{(1)}(t_0, x),$$

where α_1 positive (negative), satisfies the alternant condition of Theorem 7.3. Then $F[a]$ is the unique best approximation in the subfamily

$$V = \{F \in V_2 \mid \operatorname{sign}(F) = (-, +)\} \qquad (V = \{F \in V_2 \mid \operatorname{sign}(F) = (+, -)\}). \tag{7.4}$$

Proof. We assume $F[b] \in V, F(b, x) \not\equiv F(a, x)$, to be an approximation at least as good as $F[a]$. According to Theorem 3.2 and (4.8),

$$\operatorname{sign}(F[b] - F[a]) = (+, -, +, -).$$

In order to reach a contradiction, we distinguish three cases.

(a) Both characteristic numbers of $F[b]$ are larger than that of $F[a]$. Then the difference has the sign $(+, -, -, +)$.

(b) Both characteristic numbers of $F[b]$ are smaller than that of $F[a]$. Then the difference has the sign $(-, +, +, -)$.

(c) The characteristic number of $F[a]$ lies between those of $F[b]$. Then the difference has the sign $(-, +, -, +)$. □

From Theorem 7.4 and Corollary 6.2 we have the following

COROLLARY 7.5. *Let $f \in C(X)$ and let V_2 be an extended Descartes family with a totaly positive kernel. Then at most two best approximations exist. If two distinct best approximations exist, they have the form*

$$\beta_0^{(i)}\gamma(t_0^{(i)}, x) + \beta_1^{(i)}\gamma^{(1)}(t_0^{(i)}, x), \qquad i = 1, 2.$$

where $\beta_1^{(1)}$ and $\beta_1^{(2)}$ have opposite signs.

For $N = 2$, the theory is now quite complete. One cannot expect sharper results; functions with two best approximations are known for the exponential kernel [3].[3]

8. THE CONNECTED COMPONENTS OF NORMAL DESCARTES FAMILIES

The generalized signs give a certain structure to the Descartes families. We shall see that these signs characterize the connected components of $V_N - V_{N-1}$ under relatively weak conditions. In this section, T may be any locally compact, σ-compact set in $\mathbf{R}$; it need not be open.

First we develop a parameterization of γ-polynomials which describes their topological structure. This is not provided by the representation (1.2), e.g., one cannot see from there that in every neighborhood of $F(a, x) = \gamma^{(1)}(t, x)$ there are functions of the form

$$(1/\delta)(\gamma(t + \delta, x) - \gamma(t, x)).$$

[3] We can conclude from the existence of several best approximations that the extended Descartes families are not suns [6] and that the Kolmogoroff criterion is not a necessary condition. But these properties do hold for the subfamilies of Theorem (7.4) and for V_N^+.

THEOREM 8.1. *Let the kernel $\gamma(t, x)$ be $N - 1$ times continuously differentiable in t, and let $\gamma^{(N-1)}(t, x) \in C(T \times X)$. Then the γ-polynomials of order $\leqslant N$ can be written in the form* (*cf.* (3.4) *for notation*)

$$F(a, x) = \sum_{\mu=1}^{N} \beta_\mu \gamma(t_1, t_2, ..., t_\mu; x), \qquad t_1 \leqslant t_2 \leqslant \cdots \leqslant t_N. \tag{8.1}$$

The characteristic number t_i appears $(m + 1)$ times in (8.1), *if $\gamma^{(m)}(t_i, x)$ appears in the representation* (1.2). *The mapping corresponding to* (8.1):

$$\begin{aligned} &\Phi: A \to C(X), \\ &\Phi(a) = F[a], \end{aligned} \tag{8.2}$$

is continuous in

$$\begin{aligned} A = \{a = (\beta_1, \beta_2, ..., \beta_N, t_1, ..., t_N) \mid \beta_\nu \in \mathbf{R}, t_\nu \in T, \\ \nu = 1, 2, ..., N, \textit{ and } t_1 \leqslant t_2 \leqslant \cdots \leqslant t_N\} \subset \mathbf{R}^{2N}. \end{aligned}$$

Remark. The characteristic numbers are labeled differently in (6.1) and in (8.1); multiplicity is treated differently.

Proof of Theorem (8.1). To prove the possibility of the representation (8.1) it is sufficient to show that $\gamma^{(m)}(t_n, x)$, with any $n > m$, can be expressed as a linear combination of

$$\gamma(t_1, t_2, ..., t_\mu; x), \qquad \mu = 1, 2, ..., n, \tag{8.3}$$

whenever $t_n = t_{n-1} = \cdots = t_{n-m}$. For $n = 1$ this is obvious. We assume it to hold for $n - 1$, and distinguish two cases:

(1) Let $t_1 = t_2 = \cdots = t_n$. Then

$$\gamma^{(m)}(t_n, x) = m! \cdot \gamma(t_1, t_2, ..., t_{m+1}; x)$$

follows directly from (3.5).

(2) Let $t_1 \neq t_n$. Since $m < n - 1$, the inductive hypothesis yields that we can express $\gamma^{(m)}(t_n, x)$ as a linear combination of the $n - 1$ functions $\gamma(t_2, t_3, ..., t_\mu; x)$, $\mu = 2, 3, ..., n$, which in turn, can be expressed by the functions (8.3), using the formula

$$\gamma(t_2, t_3, ..., t_\mu; x) = \gamma(t_1, t_2, ..., t_{\mu-1}; x) + (t_\mu - t_1)(t_1, ..., t_\mu; x).$$

Now we prove that the mapping (8.2) is continuous. Since $\gamma \in C^{(\mu-1)}(T \times X)$, the divided differences $\gamma(t_1, t_2, ..., t_\mu; x)$ are continuous in $T^\mu \times X$). Thus, $(t_1, ..., t_\mu) \to \gamma(t_1, ..., t_\mu; x)$ is a continuous mapping into $C(X)$. □

The mapping (8.2) is injective only for functions in $V_N - V_{N-1}$, i.e., only for γ-polynomials of maximal order. With regard to the inverse mapping, we have the following.

LEMMA 8.2. *Using the representation* (8.1) *let the sequence* $F^\rho = F[a^\rho]$ *converge to* F *belonging to an extended Descartes family* V_N. *If the spectrum of* F^ρ *converges to the spectrum of* F, *then the parameter* a^ρ *converges to an* a, *for which* $F[a] = F$.

Remark. It is sufficient for the sequence F^ρ to converge to F at $2N$ distinct points $x_i \in X$; convergence in the strong topology is not necessary.

Proof of Lemma 8.2. The connection between the values of the functions F^ρ at N points $x_1 < x_2 < \cdots < x_N$ and the parameters $\beta_\mu{}^\rho$:

$$F^\rho(x_i) = \sum_{\mu=1}^{N} \Gamma^\rho_{i\mu} \cdot \beta_\mu{}^\rho, \qquad i = 1, 2, \ldots, N,$$

$$\Gamma^\rho_{i\mu} = \gamma(t_1{}^\rho, t_2{}^\rho, \ldots, t_\mu{}^\rho; x),$$

is given by a converging sequence of matrices Γ^ρ which, by Theorem 8.1, are not singular. As the inverse matrices approach the inverse of the limit matrix, the proposition follows from the convergence of the values $F^\rho(x_i)$. □

The assumptions of Lemma 8.2 hold, if the limit function is a proper γ-polynomial of order $k = N$ and if T is open, since then (4.4) defines a homeomorphism, and thus the convergence of the characteristic numbers follows from the convergence of the γ-polynomials in $2N$ points. For the extended γ-polynomials, we cannot establish convergence of the spectra by such simple arguments. On the other hand, it is possible to prove the results for Descartes families of interest, such as exponential sums [4, 18]. Therefore, we define the following.

DEFINITION 8.1. Let the families V_N be endowed with the topology $\mathscr{F}$. For each $F_0 \in V_N - V_{N-1}$, let there exist a neighborhood $U(F_0)$ such that the characteristic numbers for all $F \in U(F_0)$ belong to a compact subset of T. Then we call V_N normal relative to $\mathscr{F}$. If V_N is normal relative to the norm topology, then V_N is called a normal family.

If V_N is a normal family, then each family V_M with $1 \leqslant M \leqslant N$ is normal, too. For, assume F^ρ to be a sequence in V_M and $\lim F^\rho = F \in V_M - V_{M-1}$; choose a γ-polynomial $\tilde{F}$ of order $N - M$, with a spectrum disjoint from that of F. The conclusion follows from $\lim (F^\rho + \tilde{F}) = F + \tilde{F} \in V_N - V_{N-1}$.

THEOREM 8.3. *Let the extended Descartes families V_N be endowed with a topology $\mathscr{F}$ having the following properties*:

(1) *$\mathscr{F}$ is the norm-topology or a weaker one.*

(2) *The convergence of a filter implies the convergence of the functions in at least $2N$ points $x_i \in X$.*

(3) *V_N is normal relative to $\mathscr{F}$.*

Let ϕ be the mapping defined by (8.1) *and* (8.2), *then*

$$\Phi^{-1} : V_N - V_{N-1} \rightarrow A$$

is a homeomorphism.

Proof. We already know that Φ is continuous. Let $\mathscr{G}$ be a filter which converges to $F^0 \in V_N - V_{N-1}$. From $\mathscr{G}$ we can select a sequence F^ρ such that

$$\lim F^\rho(x_i) = F^0(x_i), \qquad i = 1, 2, ..., 2N,$$

holds for $2N$ points $x_i \in X$. By virtue of (3), the spectra of F^ρ are contained in a compact subset of T. Thus, the set of characteristic numbers contains a convergent subsequence, which we call F^ρ, again. By Lemma 8.2, $a^\rho = \Phi^{-1}(F^\rho)$ converges. This consideration can be applied to every subsequence. □

V_N is normal relative to the norm-topology, if V_N is normal relative to a weaker topology. As the other properties stated in Theorem 8.3 hold for the strong topology, we have the following.

COROLLARY 8.4. *Let V_N be an extended Descartes family. Then all topologies with the properties stated in Theorem* 8.3 *are equivalent to the norm-topology in $V_N - V_{N-1}$.*

Therefore, we can restrict ourselves in the following to normal families, although it is often convenient for existence proofs to use weaker topologies [18, 19].

THEOREM 8.5. *Let V_N be a normal Descartes family. Then $V_N - V_{N-1}$ is a locally compact, σ-compact space.*[4]

For proof, the reader may verify that V_{N-1} is closed in the normal family V_N and that $\Phi^{-1}(V_N - V_{N-1})$ is locally compact. Write T as a union of compact sets:

$$T = \bigcup_{m=1}^{\infty} T_m, \qquad T_m \subset T_{m+1}, \qquad T_m \text{ compact}, \tag{8.4}$$

[4] Thus $V_N - V_{N-1}$ is paracompact, i.e., a normal topological space. This motivated Definition 8.1.

and observe that $V_N - V_{N-1}$ is the union of the compact subsets

$$K_m = \{F \in V_N \mid \text{spectrum}(F) \subseteq T_m, \|F\| \leqslant m, \inf\{\|F - G\|, G \in V_{N-1}\} \geqslant 1/m\}. \tag{8.5}$$

Making use of the sign vectors introduced in Section 3, we define the 2^N classes

$$V_N(s) = \{F \in V_N - V_{N-1}, \text{sign}(F) = s\}. \tag{8.6}$$

Obviously,

$$V_N - V_{N-1} = \bigcup_s V_N(s), \quad V_N(s) \cap V_N(s') = \varnothing \qquad \text{for} \quad s \neq s', \tag{8.7}$$

because each γ-polynomial of maximal degree is associated with a unique sign vector with N components.

The following generalizes a result for exponential sums [4].

THEOREM 8.6. *Let the set of parameters T be connected. Then the $2N$ sign classes $V_N(s)$ in normal Descartes families V_N form the connected components of $V_N - V_{N-1}$.*

Proof. The subset of proper γ-polynomials in each sign class $V_N(s)$ is connected, because the representation (1.1) defines a continuous mapping from a convex set in $\mathbf{R}^{2N}$ onto $V_n(s) \cap V_N{}^0$. As was pointed out before Lemma 3.1, every γ-polynomial can be represented as a limit of proper γ-polynomials, the elements of the sequence carrying the same sign s. Hence, the sets $V_N(s)$ are connected.

For the same reason, it is sufficient to show that $V_N(s)$ is the closure of $V_N(s) \cap V_N{}^0$, in order to prove that $V_N(s)$ is closed in $V_N - V_{N-1}$. Let $F[a] = \lim F[a^\rho]$. We replace the derivatives in the normal representation (1.2) by divided differences

$$F[a] = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu} \cdot \mu!\, \gamma(t_\nu, t_\nu, \ldots, t_\nu; x)\, \gamma(t, x).$$

From Theorem 8.3 we know that exactly $(1 + M_\nu)$ characteristic numbers of this sequence converge towards t_ν. By enumerating them in the manner

$$t_{10}^\rho \cdots t_{1M_1}^\rho, \qquad t_{20}^\rho \cdots t_{2M_2}^\rho \cdots t_{lM_l}^\rho,$$

we get $\lim t_{\nu\mu}^\rho = t_\nu$. We then write

$$F[a^\rho] = \sum_{\nu=1}^{l} \sum_{\mu=0}^{M_\nu} \alpha_{\nu\mu}^\rho \cdot \mu!\, \gamma(t_{\nu 0}^\rho, t_{\nu 1}^\rho, \ldots, t_{\nu\mu}^\rho; x).$$

Since γ is $\mathrm{ESR}_{2N}(t)$, $\lim \alpha_{\nu\mu}^{\rho} = \alpha_{\nu\mu}$ is obtained in the same way as in the proof of Lemma 8.2 through convergence of the sequence at N points. By applying the same considerations as in the proof of Lemma 3.1 to each of the l partial sums, it follows that, for large ρ, the sequence belongs to the same sign class as the limit function.

Finally, from (8.7), it follows that

$$V_N(s) = (V_N - V_{N-1}) - \bigcup_{s' \neq s} V_N(s').$$

Hence, the sets $V_N(s)$ are not only closed but also open. (8.7) defines a partition into disjoint connected open and closed sets. □

Theorem 8.6 has important consequences for the numerical construction of best approximations. In most cases, V_N is an existence set, because from every bounded sequence a subsequence may be selected which converges pointwise on a dense subset of X to an element of V_N. If V_N is normal, then $V_N(s) \cup V_{N-1}$ is closed (compare Corollary 8.4), and $V_N(s) \cup V_{N-1}$ is also an existence set for each sign vector s.

If a best approximation in one of these subfamilies is not contained in V_{N-1}, one has "local best approximation." Using proofs analogous to those in [4, Section 11], we obtain local best approximations which may not be global ones, provided we exclude certain degeneracies and consider the standard case. Namely, we assume:

(1) The best approximation in V_{N-1} is a proper γ-polynomial, i.e., it is contained in V_{N-1}^0 and does not vanish identically.

(2) The best approximations in V_N and in V_{N-1} are not identical.

(3) The factors of the best approximation in V_N are not all positive or all negative.

We see that local best approximations may exist even if the (global) best approximation is a proper γ-polynomial and is, thus, unique. In any case, the other minima are extended γ-polynomials in $V_N - V_N{}^0$.

When using iterative processes for the determination of best approximations, we have to see to it that the iterative sequence does not converge towards a minimum other than a best approximation [5].

With the same assumptions on the topological structure we obtain the following nonuniqueness theorem.

THEOREM 8.7. *Let V_N be a normal Descartes family with $N \geqslant 2$. If the subsets $V_N(s) \cup V_{N-1}$ are existence sets, then there exist at least two best approximations to some $f \in C(X)$ in V_N.*

Proof. Let $F_0 \in V_{N-1}^+ - V_{N-2}^+$. Construct an $f_0 \in C(X) - V_N$ such that $f_0(x) - F_0(x)$ has an alternant of exact length $2N - 1$ and sign $-\tilde{\epsilon}_{2N-1}$. Then, by Theorem 4.2, F_0 is not optimal to f_0 in V_N, and, by Theorem 4.5, the best approximation F_1 is not contained in $V_N{}^+$. Let $s_1 = \operatorname{sign} F_1$. It follows from Theorem 12 in [4] that $\inf\{\|f_0 - F\|; F \in V_N(s)\} < \|f_0 - F_0\|$ whenever s has exactly one negative coefficient. Since the number of connected components is finite, we may select an $s_2 \neq s_1$ such that, with the best approximation $F_2 \in V_N(s_2)$, the inequality $\|f_0 - F\| < \|f_0 - F_2\|$ implies $F \in V_N(s_1)$ or $F \notin V_N$. Observe that to $f_t = f_0 + t(F_2 - f_0)$ the function F_2 is a better approximation than every γ-polynomial in V_{N-1}, if $0 \leqslant t \leqslant 1$. The functions

$$\rho_i(t) = \inf\{\|f_t - F\|; F \in V_N(s_i) \cup V_{N-1}\}, \qquad i = 1, 2,$$

are continuous. From

$$\rho_1(0) \leqslant \rho_2(0), \qquad \rho_1(1) > \rho_2(1) = 0$$

it follows that $\rho_1(t_0) = \rho_2(t_0) < \inf\{\|f_t - F\|; F \in V_{N-1}\}$ for some $t_0 \in [0, 1)$. Hence, $f = f_0 + t_0(F_2 - f_0)$ has two different best approximations in V_N, one contained in $V_N(s_1)$, the other being $F_2 \in V_N(s_2)$. □

The proof is constructive. Observe that f is closer to V_N than f_0 is. Hence in any neighborhood of V_N there are functions f with two best approximations.

Finally, we empharise that we did not even settle the question whether the number of (local) best approximations is always finite. This problem will be treated in a forthcoming paper.

9. Examples of Sign-Regular Kernels

Example 1. $\gamma(t, x) = e^{tx}$, $T = X = (-\infty, +\infty)$. This kernel is ETP [10, Chapter 3, Section 1]. The γ-polynomials in V_N which are bounded in $[a, b] \subset X$ are compact in the topology of compact convergence in (a, b) [4]. All V_N are existence sets, and so are the subfamilies $V_N(s) \cup V_{N-1}$.

Example 2. $\gamma(t, x) = \cosh tx$, $T = X = (0, \infty)$.[5] Each extended γ-polynomial of order m for this kernel is a sum of exponentials (γ-polynomials with kernel e^{tx}) of order $2m$, and therefore has at most $2m - 1$ zeros in $(-\infty, +\infty)$. There are at most $m - 1$ zeros in $(0, \infty)$, because the γ-polynomials are even functions in x. This implies that γ is ESR. The usual considerations on behavior for $x \to \infty$ establish that γ is ETP. In order to get an existence set, it is necessary to use the similar kernel $\gamma(t, x) = \cosh xt^{1/2}$ which is ETP(t) in $T = X = [0, \infty)$. Moreover, we emphasize that approxi-

[5] The kernel is not sign-regular for $X = T = (-\infty, +\infty)$, as conjectured in [7].

mation by γ-polynomials with this kernel is not equivalent to approximation of even functions by exponentials of twice the order.

EXAMPLE 3. $\gamma(t, x) = (1 + tx)^{-1}$, $T = (-1, +1)$, $X = [-1, +1]$. The extended γ-polynomials of order m can be represented as quotients of two algebraic polynomials with a numerator of degree $m - 1$. Hence property ESR holds. The topological structure is similar to that of the exponential case. Via the transformation $t \to t^{-1}$ one gets the similar kernel $\gamma(t, x) = (t + x)^{-1}$.

EXAMPLE 4. $\gamma(t, x) = x^t$, $T = X = (0, \infty)$. By means of the transformation $x \to e^x$ the results of the exponential case can be applied here.

EXAMPLE 5. $\gamma(t, x) = \operatorname{arctg} tx$, $T = X = (0, \infty)$. For any extended γ-polynomial F of order m, the derivative $(d/dx) F$ is a rational function and has at most $m - 1$ zeros in $(0, \infty)$, as can be seen easily. Since $F(0) = 0$, also F has at most $m - 1$ zeros in $(0, \infty)$. Hence, γ is ESR(t). But this kernel does not generate existence sets.

EXAMPLE 6. $\gamma(t, x) = \sin tx$, $T = (0, \tau)$, $X = (0, \pi/2\tau)$, $\tau > 0$. Meinardus proved that γ is SSR[13a]. We claim that γ is even ESR. For an inductive proof, consider $(d/dx)(\sin^2 t_1x \cdot (d/dx)(F(x)/\sin t_1x))$ and apply Rolle's theorem twice.

EXAMPLE 7. $\gamma(t, x) = \cos tx$, $T = [0, \tau)$, $X = [0, \pi/2\tau]$, $\tau > 0$. γ being ESR is established as in the preceding example.

The kernels in Examples 1–5 generate normal families.

ACKNOWLEDGMENTS

I would like to thank Professor G. Meinardus for suggesting this topic. My gratitude also goes to Professor H. Werner for our stimulating discussions.

REFERENCES

1. R. BARRAR AND H. LOEB, On N-parameter and unisolvent families. *J. Approximation Theory* **1** (1968), 180–181.
2. C. DE BOOR, On the approximation by γ-polynomials, *in* "Approximations with Special Emphasis on Spline Functions" (I. J. Schoenberg, Ed.) New York/London, 1969.
3. D. BRAESS, Über die Approximation mit Exponentialsummen. *Computing* **2** (1967), 309–321.

4. D. BRAESS, Über die Vorzeichenstruktur der Exponentialsummen, *J. Approximation Theory* **3** (1970), 101–113.
5. D. BRAESS, Die Konstruktion der Tschebyscheff-Approximierenden bei der Anpassung mit Exponentialsummen, *J. Approximation Theory*, **3** (1970), 261–273.
6. B. BROSOWSKI, "Nichtlineare Tschebyscheff-Approximation," BI Hochschultaschenbuch 808, Bibliographisches Institut, Mannheim, 1968.
6a. C. DUNHAM, Necessity of Alternation, *Canad. Math. Bull.* **11** (1968), 743–744.
7. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rat. Mech. Anal.* **27** (1967), 91–106.
8. W. HUREWICZ AND H. WALLMANN, "Dimension Theory," Princeton University Press, Princeton, NJ, 1948.
9. E. ISAACSON AND H. B. KELLER, "Analysis of Numerical Methods," John Wiley, and Sons, New York/London/Sydney, 1966.
10. S. KARLIN, "Total Positivity I," Stanford University Press, Stanford, 1968.
11. S. KARLIN AND W. STUDDEN, "Tschebyscheff Systems: With Applications in Analysis and Statistics," Interscience Publishers, New York/London/Sydney, 1966.
12. G. KÖTHE, "Topologische lineare Räume," Springer, Berlin/Göttingen/Heidelberg, 1960.
13. G. MEINARDUS, "Approximation of Functions: Theory and Numerical Methods," Springer, Berlin/Heidelberg/New York, 1967.
13a. G. MEINARDUS, Über ein Problem von L. Collatz, *Computing* **8** (1971), 250–254.
14. G. MEINARDUS UND D. SCHWEDT, Nichtlineare Approximation, *Arch .Rat. Mech. Anal.* **17** (1964), 297–326.
15. J. R. RICE, Tschebyscheff approximations by functions unisolvent of variable degree, *Trans. Amer. Math. Soc.* **99** (1961), 298–302.
16. J. R. RICE, Chebyshev-approximation by exponentials, *SIAM J. Appl. Math.* **10** (1962), 149–161.
17. J. R. RICE, "The Approximation of Functions II," Addison Wesley, Reading, MA, 1969.
18. E. SCHMIDT, Zur Kompaktheit bei Exponentialsummen, *J. Approximation Theory* **3** (1970), 445–454.
19. H. WERNER, Tschebyscheff-Approximation with Sums of Exponentials, in "Approximation Theory" (A. Talbot, Ed.) Academic Press, New York/London, 1970.