JOURNAL OF APPROXIMATION THEORY **89**, 380–423 (1997)
ARTICLE NO. AT963059

Best Approximation and Cyclic Variation Diminishing Kernels

Oleg Davydov

Department of Mechanics and Mathematics, Dnepropetrovsk State University, pr. Gagarina 72, Dnepropetrovsk, GSP 320625, Ukraine
E-mail: olda@uni.dnepropetrovsk.ua

and

Allan Pinkus

Department of Mathematics, Technion, IIT, Haifa, Israel
E-mail: pinkus@tx.technion.ac.il

Communicated by Günther Nürnberger

Received November 29, 1995; accepted in revised form May 8, 1996

We study best uniform approximation of periodic functions from

$$\left\{ \int_0^{2\pi} K(x, y)\, h(y)\, dy : |h(y)| \leqslant 1 \right\},$$

where the kernel $K(x, y)$ is strictly cyclic variation diminishing, and related problems including periodic generalized perfect splines. For various approximation problems of this type, we show the uniqueness of the best approximation and characterize the best approximation by extremal properties of the error function. The results are proved by using a characterization of best approximants from quasi-Chebyshev spaces and certain perturbation results. © 1997 Academic Press

1. INTRODUCTION

This paper is about some approximation problems related to cyclic variation diminishing (CVD) kernels. CVD kernels are the periodic analogues of totally positive (TP) kernels. CVD kernels were introduced and discussed in two papers by Schoenberg and coauthors [5, 8] in 1958 and 1959. A more comprehensive consideration is to be found in the book of Karlin [4, Chaps. 5 and 9]. We first define the relevant concepts. We will later return to a general discussion of CVD kernels.

0021-9045/97 $25.00

In what follows $\widetilde{C}$ will denote the set of continuous 2π-periodic functions defined on all $\mathbb{R}$. (The period 2π is chosen for no particular reason.) By $\widetilde{C}^2$ we mean the two-variable functions (kernels) defined on all of $\mathbb{R}^2$ which are continuous and 2π-periodic in each variable.

DEFINITION 1.1. Let $K \in \widetilde{C}^2$. We say that K is a *cyclic variation diminishing kernel of order* $2m-1$ (CVD_{2m-1}) if there exist $\varepsilon_n \in \{-1, 1\}$, $n = 1, ..., m$, such that

$$\varepsilon_n K\begin{pmatrix} x_1, ..., x_{2n-1} \\ y_1, ..., y_{2n-1} \end{pmatrix} = \varepsilon_n \det\{K(x_i, y_j)\}_{i,j=1}^{2n-1} \geqslant 0 \tag{1.1}$$

for all $x_1 < \cdots < x_{2n-1} < x_1 + 2\pi$ and $y_1 < \cdots < y_{2n-1} < y_1 + 2\pi$. We say that the kernel K is *strictly cyclic variation diminishing of order* $2m-1$ (SCVD_{2m-1}) if strict inequality always holds in (1.1). The kernel K is said to be *extended cyclic variation diminishing of order* $2m-1$ (ECVD_{2m-1}) if K is $2m-1$ times continuously differentiable, and the above determinants are strictly positive for all choices of $x_1 \leqslant \cdots \leqslant x_{2n-1} < x_1 + 2\pi$ and $y_1 \leqslant \cdots \leqslant y_{2n-1} < y_1 + 2\pi$, where in case of equal x_i (or y_j) we replace the corresponding rows (columns) by successive derivatives.

We will drop the subscript $2m-1$ from the acronyms CVD, SCVD, or ECVD if we assume that these properties hold for all orders.

Note that the only determinantal conditions imposed are those on the odd-order minors. This "restriction" is a consequence of the periodicity (and a simple rotation of columns or rows). That is, we always have

$$K\begin{pmatrix} x_1, x_2, ..., x_{2n} \\ y_1, y_2, ..., y_{2n} \end{pmatrix} = -K\begin{pmatrix} x_1, x_2, ..., x_{2n} \\ y_2, ..., y_{2n}, y_1 \end{pmatrix},$$

and the "correct" ordering has been maintained. (This is essentially equivalent to the fact that periodic functions have an even number of sign changes (or zeros if the count is done correctly).) Thus (1.1) cannot possibly hold for even-order minors (except in the uninteresting case where the associated determinants are all identically zero). This restriction is a serious drawback and generally weakens the theory. In the standard non-periodic TP case a determinantal inequality of the form (1.1) holds for all orders, and this results in a "stronger" theory. The periodicity is, in a certain sense, a partial compensation.

Essentially equivalent to the CVD, SCVD, and ECVD properties are certain variation diminishing properties; see Karlin [4, Chap. 5, Theorem 6.1]. To explain, let $S_c(f)$ denote the number of sign changes of $f \in \widetilde{C}$ on a period. $\widetilde{Z}_c(f)$ will count the number of zeros of f where nodal zeros (sign changes) are counted once, and nonnodal zeros (zeros which are not sign changes) twice. $Z_c^*(f)$ will, for f sufficiently smooth, denote the number of distinct zeros of f, counting multiplicities. For a vector $\mathbf{c}=(c_1, ..., c_k)$, we let $S_c(\mathbf{c})$ denote the number of (weak) periodic sign changes in the vector $\mathbf{c}$. By this we mean the number of sign changes in any of the sequences

$$c_j, ..., c_k, c_1, ..., c_j,$$

where $c_j \neq 0$, and zero components are discarded. Note that all these values are even numbers (or infinite). We also need the number of sign changes of a 2π-periodic Borel measure μ. We say that such a measure has $2n$ relevant sign changes, denoted by $S_c(\mu)=2n$, if there exist disjoint sets $A_1 < \cdots < A_{2n} < A_1 + 2\pi$, with $\bigcup_{i=1}^{2n} A_i = [a, a+2\pi)$ (some a), such that $(-1)^i \mu$ is a nonnegative measure on A_i and $\mu(A_i) \neq 0$, $i=1, ..., 2n$. If h is a summable 2π-periodic function, then by $S_c(h)$ we mean $S_c(\mu)$, where $d\mu(y)=h(y)\,dy$.

An essentially equivalent definition to the CVD property of the kernel K is that

$$S_c(g) \leqslant S_c(\mu)$$

for all μ as above, where

$$g(x)=\int_0^{2\pi} K(x, y)\,d\mu(y).$$

And similarly, K is SCVD if and only if (up to some minor details)

$$\tilde{Z}_c(g) \leqslant S_c(\mu)$$

for all μ and g as above. Finally, K is ECVD if and only if (up to those minor details again)

$$Z_c^*(g) \leqslant S_c(\mu)$$

for all μ as above, and g sufficiently smooth.

The original two papers which dealt with CVD kernels were [8] by Pólya and Schoenberg and [5] by Mairhuber, Schoenberg and Williamson. Schoenberg had, over the years, developed a theory of totally positive kernels, especially totally positive difference kernels (called Pólya frequency functions). The two papers [5, 8] were the first to consider the periodic versions thereof. As we have already remarked, the even-order minors cannot possibly be of one strict sign and this complicates the theory. (The theory is even today not nearly as complete as the theory of Pólya frequency functions.)

In the first paper [8], Pólya and Schoenberg studied the de la Vallée Poussin means. Let

$$\omega_m(t) = \frac{1}{\binom{2m}{m}} \sum_{\nu=-m}^{m} \binom{2m}{m+\nu} e^{i\nu t}.$$

The transformation

$$V_m(x) = \frac{1}{2\pi} \int_0^{2\pi} \omega_m(x-y)\, f(y)\, dy$$

defines the de la Vallée Poussing means (or V-means) of f. V_m is a trigonometric polynomial of degree at most m, and for every $f \in \tilde{C}$ the V_m uniformly converge to f as $m \to \infty$. The main result of [8] is that the difference kernel $\omega_m(x-y)$ is SCVD_{2m+1}. (We will use the fact, see Karlin [4, Chap. 9, Corollary 3.1], that $\omega_m(x-y)$ is ECVD_{2m+1}.)

In the second paper [5], a more general theory was pursued with regard to CVD difference kernels, i.e., kernels $K \in \tilde{C}^2$ which are CVD and of the form $K(x, y) = k(x-y)$ for some $k \in \tilde{C}$.

The results of both papers, along with numerous generalizations, may be found in Chaps. 5 and 9 of Karlin [4]. In Chap. 9 is studied the many properties of CVD difference kernels. For example, it is shown that if $k(x-y)$ is SCVD_{2m+1} and $k \in \tilde{C}^{(4m)}$, then $k(x-y)$ is in fact ECVD_{2m+1} (see Karlin [4, Chap. 9, Theorem 9.1]). Another result concerning difference kernels which are CVD is that the ε_n in (1.1) are necessarily all equal (see [5, p. 258] or [4, Chap. 5, Theorem 7.1]).

This present paper is, to a large degree, a continuation and extension of [6] to the periodic case. Motivated by work of Sattes [10], the second author considered in [6] approximations (in the uniform norm) to $f \in C[0, 1]$ by functions of the form

$$g(x) = \int_0^1 K(x, y)\, h(y)\, dy,$$

where $|h(y)| \leqslant 1$ and K is strictly totally positive (STP). In addition, numerous related problems were considered such as best approximating by generalized perfect splines with at most n knots, and best approximation from

$$\int_0^1 K(x, y)\, d\mu(y),$$

where $d\mu$ is a nonnegative measure. The results obtained (uniqueness and characterization) were somewhat surprising, considering the fact that the approximating subspace is not finite dimensional (or finite parameter). The results also depended, rather crucially, on an "orientation." It was unclear how one might generalize these results to the periodic case, where there seemed to be no natural "orientation." In [2] the first author, using his results from [1], was able to generalize the main result in [6] to the periodic case. In this paper we review this work (Section 3) and then go on to consider various related problems.

To be more precise, in Section 3 we characterize and prove uniqueness of the best approximation to $f \in \tilde{C}$ from

$$\mathcal{M} = \left\{ \int_0^{2\pi} K(x, y)\, h(y)\, dy : |h(y)| \leqslant 1 \text{ a.e., } y \in [0, 2\pi] \right\},$$

under the assumption that K is SCVD. If $f \notin \mathcal{M}$, then this unique best approximation is necessarily of the form

$$\sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy,$$

for some integer n and some $\xi_1 < \cdots < \xi_{2n} < \xi_{2n+1} = \xi_1 + 2\pi$. We call such functions periodic generalized perfect splines with $2n$ knots. It also exhibits additional properties (see Theorem 3.2).

In Section 4 we restrict our approximating set to a subset of periodic generalized perfect splines with exactly $2n$ knots, where both n and one of the knots is fixed. We characterize and prove the uniqueness of the best approximation to $f \in \tilde{C} \backslash \mathcal{M}$ (Theorem 4.1).

We continue this investigation in Section 5, where we consider approximation from the set of periodic generalized splines with exactly $2n$ knots (but none fixed, Proposition 5.1), and a related problem (Theorem 5.2). Finally, in Section 6, our approximation set is

$$\mathcal{M}_\infty = \left\{ \int_0^{2\pi} K(x, y)\, d\mu(y) : \mu \geqslant 0 \right\}.$$

We prove analogues of some of the results of Sections 3, 4, and 5.

2. PRELIMINARIES

In this section we present various results which will be needed and used in the subsequent analysis. Some of these results may be found in Davydov [1]. However, since that paper is contained in a proceedings in Russian which is probably inaccessible to many readers, we will also present these results with proofs. For variety, the proofs will be somewhat different than those in [1].

An m-dimensional subspace U of continuous functions defined on an interval I is said to be a *Chebyshev* (T-) *space* (or Haar space) if no non-trivial function vanishes at more than $m-1$ distinct points in I. If $u_1, ..., u_m$ is any basis for this space, then it is called a *T-system.* (The terms "space" and "system" are often used interchangeably.) An equivalent definition of a T-system is that

$$U\begin{pmatrix} 1, ..., m \\ x_1, ..., x_m \end{pmatrix} = \det\{u_i(x_j)\}_{i,j=1}^{m} \neq 0$$

for every choice of distinct $x_1, ..., x_m$ in I.

T-spaces have many distinctive properties. One of the more familiar is the characterization (and the uniqueness) of the best approximation to continuous functions in the uniform norm from T-spaces. Since we will deal with 2π-periodic functions, we formulate the result in this setting. We note, for the same reasons as stated in the Introduction, that a T-space in $\tilde{C}$ is necessarily of odd dimension.

THEOREM 2.1. *Let $U_{2m+1} \subset \tilde{C}$ be a T-space of dimension $2m+1$. Let $f \in \tilde{C} \backslash U_{2m+1}$. Then there exists a unique best approximation u^* to f from U_{2m+1}. u^* is characterized by the fact that there exist $2m+2$ points $x_1 < \cdots < x_{2m+2} < x_1 + 2\pi$ and a $\delta \in \{-1, 1\}$ such that*

$$\delta(-1)^i (f - u^*)(x_i) = \|f - u^*\|, \qquad i = 1, ..., 2m+2.$$

We will generally simply say that $f - u^*$ *equioscillates* on $2m+2$ points.

If K is an SCVD kernel, then for every choice of $y_1 < \cdots < y_{2m+1} < y_1 + 2\pi$ (resp., $x_1 < \cdots < x_{2m+1} < x_1 + 2\pi$), the set of functions $K(x, y_1), ..., K(x, y_{2m+1})$ (resp., $K(x_1, y), ..., K(x_{2m+1}, y)$) spans a T-space of dimension $2m+1$. Moreover it will also be necessary that we deal with $2m$ sections of the kernel K, which cannot possibly be a T-system. To this end we present the following definition and result.

DEFINITION 2.1. Let U_{2m} be a $2m$-dimensional subspace of $\tilde{C}$. We say that U_{2m} is a *quasi-Chebyshev* (QT-) space if U_{2m} contains a $(2m-1)$-dimensional T-space and is contained in a $(2m+1)$-dimensional T-space.

Following previous notation, any basis for a QT-space will be called a QT-system. The next result characterizes best approximations from QT-spaces. Note that there is no claim of uniqueness of the best approximation.

THEOREM 2.2 (Davydov [1]). *Let $U_{2m} \subset \tilde{C}$ be a QT-space of dimension 2m. Let $f \in \tilde{C} \backslash U_{2m}$. Then $u^* \in U_{2m}$ is a best approximation to f from U_{2m} if and only if there exist 2m points $w_1 < \cdots < w_{2m} < w_1 + 2\pi$, a $\delta \in \{-1, 1\}$, and additional points w'_{2m}, w''_{2m} satisfying*

$$w_{2m-1} < w'_{2m} \leqslant w_{2m} \leqslant w''_{2m} < w_1 + 2\pi$$

such that

$$\text{(a)} \quad \dim U_{2m}|_{\{w_1, ..., w_{2m}\}} < 2m$$

$$\text{(b)} \quad \delta(-1)^i (f - u^*)(w_i) = \|f - u^*\|, \qquad i = 1, ..., 2m-1$$

$$\delta(f - u^*)(w'_{2m}) = \delta(f - u^*)(w''_{2m}) = \|f - u^*\|.$$

We do allow for the possibility that $w'_{2m} = w_{2m} = w''_{2m}$.

Proof. ($\Rightarrow$) Assume that $u^* \in U_{2m}$ is a best approximation to f from U_{2m}. It is known (see, e.g., Rivlin [9, p. 63]) that there exist k distinct points, $1 \leqslant k \leqslant 2m+1$,

$$x_1 < \cdots < x_k < x_1 + 2\pi$$

and real numbers $c_j \neq 0$, $j = 1, ..., k$, such that

$$\text{(i)} \quad \sum_{j=1}^{k} c_j u(x_j) = 0, \qquad \text{all} \quad u \in U_{2m}$$

$$\text{(ii)} \quad (\operatorname{sgn} c_j)(f - u^*)(x_j) = \|f - u^*\|, \qquad j = 1, ..., k.$$

Since U_{2m} contains a T-space of dimension $2m-1$, it follows that

$$S_c(c_1, ..., c_k) \geqslant 2m$$

and thus $k \in \{2m, 2m+1\}$. As a further consequence $\dim U|_{\{x_1, ..., x_k\}} \geqslant 2m-1$. We consider two cases.

(1) $\dim U|_{\{x_1, ..., x_k\}} = 2m$. In this case we must have (from (i)) that $k = 2m+1$. The value $S_c(c_1, ..., c_{2m+1})$ is an even number. As such it must equal $2m$, and $c_j c_{j+1} < 0$, $j = 1, ..., 2m+1$ ($c_{2m+2} = c_1$), for all but one j. Assume without loss of generality that $c_{2m} c_{2m+1} > 0$. Let $u_1, ..., u_{2m}$ by any basis for U_{2m}. Solving for c_j (from (i)), we see that we must have

$$U\begin{pmatrix} 1, ..., 2m \\ x_1, ..., x_{2m-1}, x_{2m} \end{pmatrix} U\begin{pmatrix} 1, ..., 2m \\ x_1, ..., x_{2m-1}, x_{2m+1} \end{pmatrix} < 0.$$

Thus for some $x_{2m} < \tilde{x}_{2m} < x_{2m+1}$, we have

$$U\begin{pmatrix} 1, ..., 2m \\ x_1, ..., x_{2m-1}, \tilde{x}_{2m} \end{pmatrix} = 0.$$

Set $w_i = x_i$, $i = 1, ..., 2m-1$, $w_{2m} = \tilde{x}_{2m}$, $w'_{2m} = x_{2m}$, and $w''_{2m} = x_{2m+1}$. The conditions of the theorem hold.

(2) $\dim U|_{\{x_1, ..., x_k\}} = 2m-1$. In this case we may assume, by a simple argument, that $k = 2m$. Since $S_c(c_1, ..., c_{2m}) = 2m$, the c_j's must alternate in sign. Set $w_i = x_i$, $i = 1, ..., 2m$, and $w'_{2m} = w''_{2m} = w_{2m}$. The conditions of the theorem thus hold.

($\Leftarrow$) Assume that conditions (a) and (b) hold and u^* is not a best approximation to f from U_{2m}. Thus there exists a $\tilde{u} \in U_{2m}$ such that

$$\| f - u^* - \tilde{u} \| < \| f - u^* \|,$$

from which it follows that

$$\delta(-1)^i \, \tilde{u}(w_i) > 0, \qquad i = 1, ..., 2m-1,$$

$$\delta\tilde{u}(w'_{2m}) > 0, \qquad \delta\tilde{u}(w''_{2m}) > 0.$$

Since U_{2m} is contained in a $(2m+1)$-dimensional T-space, $\tilde{u}$ cannot have more than $2m$ distinct zeros. Thus $\tilde{u}$ has no zero in $[w'_{2m}, w''_{2m}]$ and therefore

$$\delta\tilde{u}(w_{2m}) > 0.$$

The function $\tilde{u}$ strictly alternates in sign on the $2m$ points $w_1, ..., w_{2m}$, where

$$\dim U_{2m}|_{\{w_1, ..., w_{2m}\}} < 2m.$$

We prove that this is impossible. For each w_i there exists a v_i in the T-space of dimension $(2m-1)$ contained in U_{2m} which agrees with $\tilde{u}$ at $\{w_1, ..., w_{2m}\} \setminus \{w_i\}$. In addition v_i has at most $2m-2$ zeros. Thus $\tilde{u}$ and v_i have opposite signs at w_i, and $(\tilde{u} - v_i)(w_i) \neq 0$. Renormalizing we have constructed $2m$ functions $z_i = a_i(\tilde{u} - v_i) \in U_{2m}$ satisfying $z_i(w_j) = \delta_{ij}$, $i, j = 1, ..., 2m$. But then

$$\dim U_{2m}|_{\{w_1, ..., w_{2m}\}} = 2m,$$

which is a contradiction. ∎

Remark 2.1. In the above proof of the sufficiency we used the fact that U_{2m} is contained in a $(2m+1)$-dimensional T-space. This same result may be proven by more involved methods, without this assumption.

Remark 2.2. If $u^* \in U_{2m}$ is such that $f-u^*$ equioscillates at $2m+2$ points, then u^* is necessarily the unique best approximation to f from U_{2m}. (This follows from the fact that it is the unique best approximation from the $(2m+1)$-dimensional T-space containing U_{2m}.) Thus there must also exist points for which (a) and (b) hold.

QT-spaces have an additional property which we will find useful. It is the following.

LEMMA 2.3 (Davydov [1]). *Assume that U_{2m} is a QT-space, and*

$$\dim U_{2m}|_{\{w_1, \ldots, w_{2m}\}} < 2m.$$

Then for every choice of

$$y_1 < \cdots < y_{2m} < y_1 + 2\pi$$

satisfying $w_i \leqslant y_i \leqslant w_{i+1}$, $i=1, \ldots, 2m$ ($w_{2m+1} = w_1 + 2\pi$), with $\{w_1, \ldots, w_{2m}\} \neq \{y_1, \ldots, y_{2m}\}$ we have

$$\dim U_{2m}|_{\{y_1, \ldots, y_{2m}\}} = 2m.$$

Proof. Let $u_1, \ldots, u_{2m-1}$ be a basis for the $(2m-1)$-dimensional T-space U_{2m-1} contained in U_{2m} and u_{2m} be such that $u_1, \ldots, u_{2m}$ is a basis for U_{2m}. Since

$$\dim U_{2m}|_{\{w_1, \ldots, w_{2m}\}} < 2m,$$

there exists a non-trivial $v_1 \in U_{2m}$ of the form $v_1 = \sum_{j=1}^{2m} a_j u_j$ which vanishes at the $\{w_i\}$. Furthermore from the T-space property of U_{2m-1} we must have $a_{2m} \neq 0$. Since U_{2m} is contained in a T-space of dimension $2m+1$, the function v_1 must change sign at each of the w_i, and vanish nowhere else.

Similarly if

$$\dim U_{2m}|_{\{y_1, \ldots, y_{2m}\}} < 2m,$$

then there exists a non-trivial $v_2 \in U_{2m}$ of the form $v_2 = \sum_{j=1}^{2m} b_j u_j$ which vanishes at the $\{y_i\}$. From the T-space property of U_{2m-1} we must have $b_{2m} \neq 0$, and since U_{2m} is contained in a T-space of dimension $2m+1$, the function v_2 must change sign at each of the y_i, and vanish nowhere else.

Since $\{w_1, \ldots, w_{2m}\} \neq \{y_1, \ldots, y_{2m}\}$, the function $b_{2m} v_1 - a_{2m} v_2 \in U_{2m-1}$ is not identically zero. However, it has at least $2m$ zeros (where we count zeros which are not sign changes as double zeros in the sense of $\tilde{Z}_c$). This contradicts known properties of T-spaces. ∎

Let us assume that $K \in \tilde{C}^2$ is an SCVD kernel. From Lemma 2.3 it follows that if

$$K\begin{pmatrix} x_1, ..., x_{2m} \\ y_1, ..., y_{2m} \end{pmatrix} = 0$$

for some $x_1 < \cdots < x_{2m} < x_1 + 2\pi$ and $y_1 < \cdots < y_{2m} < y_1 + 2\pi$, then necessarily

$$K\begin{pmatrix} w_1, ..., w_{2m} \\ y_1, ..., y_{2m} \end{pmatrix} \neq 0$$

for every choice of $w_1 < \cdots < w_{2m} < w_1 + 2\pi$ satisfying $x_i \leqslant w_i \leqslant x_{i+1}$, $i = 1, ..., 2m$ $(x_{2m+1} = x_1 + 2\pi)$, with $\{x_1, ..., x_{2m}\} \neq \{w_1, ..., w_{2m}\}$, and thus is of one fixed sign throughout this domain. Let us denote its sign by $\sigma_1(\mathbf{x}, \mathbf{y}) \in \{-1, 1\}$ (to also note its dependence on $\mathbf{x}$ and on $\mathbf{y}$). Similarly

$$K\begin{pmatrix} x_1, ..., x_{2m} \\ z_1, ..., z_{2m} \end{pmatrix} \neq 0$$

for every choice of $z_1 < \cdots < z_{2m} < z_1 + 2\pi$ satisfying $y_i \leqslant z_i \leqslant y_{i+1}$, $i = 1, ..., 2m$ $(y_{2m+1} = y_1 + 2\pi)$, with $\{y_1, ..., y_{2m}\} \neq \{z_1, ..., z_{2m}\}$, and thus is of one fixed sign throughout this domain. Let us denote its sign by $\sigma_2(\mathbf{x}, \mathbf{y}) \in \{-1, 1\}$. There is a relationship between σ_1 and σ_2 which we will use and thus record in this next lemma.

LEMMA 2.4. *Assume that K is an* SCVD *kernel, and*

$$K\begin{pmatrix} x_1, ..., x_{2m} \\ y_1, ..., y_{2m} \end{pmatrix} = 0$$

for some $x_1 < \cdots < x_{2m} < x_1 + 2\pi$ and $y_1 < \cdots < y_{2m} < y_1 + 2\pi$. Let σ_1 and σ_2 be as above. Then

$$\sigma_1(\mathbf{x}, \mathbf{y})\, \sigma_2(\mathbf{x}, \mathbf{y}) = -\varepsilon_m \varepsilon_{m+1},$$

where the ε_n are as defined in Definition 1.1.

Proof. We use a simple form of Sylvester's determinant identity (see Karlin [4, p. 3]) which says that for $x_1 < \cdots < x_{2m+1} < x_1 + 2\pi$ and $y_1 < \cdots < y_{2m+1} < y_1 + 2\pi$, we have

$$K\begin{pmatrix} x_1, ..., x_{2m-1} \\ y_1, ..., y_{2m-1} \end{pmatrix} K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m}, x_{2m+1} \\ y_1, ..., y_{2m-1}, y_{2m}, y_{2m+1} \end{pmatrix}$$

$$= K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m} \\ y_1, ..., y_{2m-1}, y_{2m} \end{pmatrix} K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m+1} \\ y_1, ..., y_{2m-1}, y_{2m+1} \end{pmatrix}$$

$$- K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m+1} \\ y_1, ..., y_{2m-1}, y_{2m} \end{pmatrix} K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m} \\ y_1, ..., y_{2m-1}, y_{2m+1} \end{pmatrix}.$$

By assumption,

$$K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m} \\ y_1, ..., y_{2m-1}, y_{2m} \end{pmatrix} = 0.$$

In addition, we have

$$\varepsilon_m K\begin{pmatrix} x_1, ..., x_{2m-1} \\ y_1, ..., y_{2m-1} \end{pmatrix} > 0,$$

and

$$\varepsilon_{m+1} K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m}, x_{2m+1} \\ y_1, ..., y_{2m-1}, y_{2m}, y_{2m+1} \end{pmatrix} > 0.$$

Finally,

$$\sigma_1(\mathbf{x}, \mathbf{y})\, K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m+1} \\ y_1, ..., y_{2m-1}, y_{2m} \end{pmatrix} > 0,$$

and

$$\sigma_2(\mathbf{x}, \mathbf{y})\, K\begin{pmatrix} x_1, ..., x_{2m-1}, x_{2m} \\ y_1, ..., y_{2m-1}, y_{2m+1} \end{pmatrix} > 0,$$

which proves the lemma. ∎

3. APPROXIMATION FROM $\mathcal{M}$

As previously, we assume that $K \in \tilde{C}^2$ is an SCVD kernel, and set

$$\mathcal{M} = \left\{ g(x) = \int_0^{2\pi} K(x, y)\, h(y)\, dy : |h(y)| \leqslant 1 \text{ a.e., } y \in [0, 2\pi] \right\}.$$

In this section we review the main result from Davydov [2] regarding the best approximation to $f \in \tilde{C}$ from $\mathcal{M}$. To this end we introduce the following definition.

DEFINITION 3.1. A function $g \in \mathcal{M}$ is said to be a *periodic generalized perfect spline* with $2n$ knots if:

(a) $n=0$ and

$$g(x) = \pm \int_0^{2\pi} K(x, y)\, dy;$$

(b) $n \geqslant 1$ and there exist $2n$ points (called *knots*)

$$\xi_1 < \cdots < \xi_{2n} < \xi_1 + 2\pi = \xi_{2n+1}$$

such that

$$g(x) = \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy.$$

This next result and the ideas behind it will be used many times. It is of central importance in determining "orientation" of the best approximation. As such we present it as a separate result.

PROPOSITION 3.1. *Let $n \geqslant 1$, and assume that*

$$g^*(x) = \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy$$

is a best approximation to $f \in \tilde{C} \backslash \mathcal{M}$ from $\mathcal{M}$. Let $\eta \notin \{\xi_1, ..., \xi_{2n}\}$. Then the zero function is a best approximation to $f - g^*$ from

$$\mathcal{A} = \left\{ \sum_{i=1}^{2n} a_i K(x, \xi_i) + bK(x, \eta) : a_i \in \mathbb{R},\, i = 1, ..., 2n,\, \delta b \leqslant 0 \right\},$$

where $\delta = (-1)^{i+1}$ if $\eta \in (\xi_i, \xi_{i+1})$, $i = 1, ..., 2n$.

Remark. The above proposition states that $\mathcal{A}$ is contained in the *tangent cone* to $\mathcal{M}$ at g^*.

Proof. Without loss of generality we assume that $\eta \in (\xi_{2n}, \xi_1 + 2\pi)$. Thus $\delta = -1$, and in the definition of $\mathcal{A}$ we have $b \geqslant 0$. Assume that the zero function is not a best approximation to $f-g^*$ from $\mathscr{A}$. Then there exists a

$$v(x)=\sum_{j=1}^{2n} a_jK(x,\xi_j)+bK(x,\eta)\in\mathscr{A}$$

such that

$$\|f-g^*-v\|<\|f-g^*\|.$$

Thus for every $\lambda\in(0,1]$ we have

$$\|f-g^*-\lambda v\|\leqslant\|f-g^*\|-\lambda c,$$

where

$$c=\|f-g^*\|-\|f-g^*-v\|>0.$$

Set $\delta_j=\frac{1}{2}(-1)^j a_j\lambda$, $\lambda>0$, small, $j=1,...,2n$, and $\delta_{2n+1}=\frac{1}{2}b\lambda$. (Thus $\delta_{2n+1}\geqslant 0$.) Let $g(x;\boldsymbol{\xi})=g^*(x)$, and for $\boldsymbol{\delta}=(\delta_1,...,\delta_{2n+1})$ as above, set

$$g(x;\boldsymbol{\xi}+\boldsymbol{\delta};\eta)=\sum_{j=1}^{2n-1}(-1)^{j+1}\int_{\xi_j+\delta_j}^{\xi_{j+1}+\delta_{j+1}}K(x,y)\,dy-\int_{\xi_{2n}+\delta_{2n}}^{\eta}K(x,y)\,dy$$
$$+\int_{\eta}^{\eta+\delta_{2n+1}}K(x,y)\,dy-\int_{\eta+\delta_{2n+1}}^{\xi_1+2\pi+\delta_1}K(x,y)\,dy.$$

Now for $\lambda>0$, small,

$$\begin{aligned}
&g(x;\boldsymbol{\xi}+\boldsymbol{\delta};\eta)-g(x;\boldsymbol{\xi})\\
&\quad=\sum_{j=1}^{2n-1}(-1)^{j+1}\left[\int_{\xi_{j+1}}^{\xi_{j+1}+\delta_{j+1}}K(x,y)\,dy-\int_{\xi_j}^{\xi_j+\delta_j}K(x,y)\,dy\right]\\
&\qquad-\left[\int_{\xi_1}^{\xi_1+\delta_1}K(x,y)\,dy-2\int_{\eta}^{\eta+\delta_{2n+1}}K(x,y)\,dy-\int_{\xi_{2n}}^{\xi_{2n}+\delta_{2n}}K(x,y)\,dy\right]\\
&\quad=2\sum_{j=1}^{2n}(-1)^j\delta_jK(x,\xi_j)+2\delta_{2n+1}K(x,\eta)+o(\boldsymbol{\delta})\\
&\quad=\lambda v(x)+o(\lambda).
\end{aligned}$$

Since $\delta_{2n+1}\geqslant 0$, we have $g(\cdot;\boldsymbol{\xi}+\boldsymbol{\delta};\eta)\in\mathscr{M}$. (If $\delta_{2n+1}<0$, this would not be true.) Thus

$$\begin{aligned}\|f-g^*\| &= \|f-g(\cdot;\boldsymbol{\xi})\| \leqslant \|f-g(\cdot;\boldsymbol{\xi}+\boldsymbol{\delta};\eta)\| \\ &= \|f-(g(\cdot;\boldsymbol{\xi})+\lambda v+o(\lambda))\| \\ &= \|f-g^*-\lambda v\|+o(\lambda) \\ &\leqslant \|f-g^*\|-\lambda c+o(\lambda).\end{aligned}$$

But then for $\lambda>0$, sufficiently small, a contradiction ensues. ∎

We now state and reprove the main result in Davydov [2]. We present it here for completeness, and because we apply a slightly different method of proof.

THEOREM 3.2 (Davydov [2]). *Assume that K is an* SCVD *kernel, and $f\in\tilde{C}\backslash\mathscr{M}$. There exists a unique best approximation g^* to f from $\mathscr{M}$. g^* is a periodic generalized perfect spline with $2n$ knots and is characterized as follows.*

(a) *If $n=0$, then*

$$g^*(x)=\delta\int_0^{2\pi} K(x,y)\,dy$$

for some $\delta\in\{-1,1\}$, and there exists a θ such that

$$\varepsilon_1\delta(f-g^*)(\theta)=\|f-g^*\|.$$

(b) *If $n\geqslant 1$, then*

$$g^*(x)=\sum_{j=1}^{2n}(-1)^{j+1}\int_{\xi_j}^{\xi_{j+1}} K(x,y)\,dy$$

for some $\xi_1<\cdots<\xi_{2n}<\xi_1+2\pi=\xi_{2n+1}$, as above, and one of the following is true:

(b1) *$f-g^*$ equioscillates on $2n+2$ points,*

(b2) *there exist $\theta_1<\cdots<\theta_{2n}<\theta_1+2\pi$ such that*

$$K\begin{pmatrix}\theta_1,\ldots,\theta_{2n}\\ \xi_1,\ldots,\xi_{2n}\end{pmatrix}=0$$

and for some θ'_{2n}, θ''_{2n} satisfying $\theta_{2n-1}<\theta'_{2n}\leqslant\theta_{2n}\leqslant\theta''_{2n}<\theta_1+2\pi$ we have

$$\begin{aligned}(-1)^{i+1}\varepsilon_n\sigma_2(\boldsymbol{\theta},\boldsymbol{\xi})(f-g^*)(\theta_i)&=\|f-g^*\|, \qquad i=1,\ldots,2n-1\\ -\varepsilon_n\sigma_2(\boldsymbol{\theta},\boldsymbol{\xi})(f-g^*)(\theta'_{2n})=-\varepsilon_n\sigma_2(\boldsymbol{\theta},\boldsymbol{\xi})(f-g^*)(\theta''_{2n})&=\|f-g^*\|.\end{aligned} \tag{3.1}$$

Proof. From the compactness of $\mathscr{M}$, we have the existence of a best approximation g^* to f from $\mathscr{M}$. We refer to Glashoff [3] where the method of proof shows that g^* must be a periodic generalized perfect spline (with a finite number of knots). The uniqueness follows from a standard convexity argument, since a strict convex combination of two distinct periodic generalized perfect splines is not a periodic generalized perfect spline.

Sufficiency. We assume that g^* satisfies (a) or (b). If (a) holds then for any $g \in \mathscr{M}$, $g \neq g^*$,

$$\varepsilon_1 \delta g(\theta) = \varepsilon_1 \delta \int_0^{2\pi} K(\theta, y)\, h(y)\, dy < \int_0^{2\pi} |K(\theta, y)|\, dy = \varepsilon_1 \delta g^*(\theta)$$

and thus

$$\|f - g^*\| = \varepsilon_1\, \delta(f - g^*)(\theta) < \varepsilon_1\, \delta(f - g)(\theta) \leqslant \|f - g\|$$

and so g^* is the best approximation to f from $\mathscr{M}$.
If (b) holds, and

$$\|f - g\| < \|f - g^*\|$$

for some $g(x) = \int_0^{2\pi} K(x, y)\, h(y)\, dy \in \mathscr{M}$, then

$$\tilde{Z}_c((f - g^*) - (f - g)) = \tilde{Z}_c(g - g^*) \leqslant S_c(h - h^*) \leqslant 2n, \tag{3.2}$$

where $g^*(x) = \int_0^{2\pi} K(x, y)\, h^*(y)\, dy$. (The right most inequality in (3.2) comes from the form of h^*.) If $f - g^*$ equioscillates on $2n + 2$ points, then

$$2n + 2 \leqslant \tilde{Z}_c((f - g^*) - (f - g))$$

and a contradiction immediately ensues from (3.2). This proves the sufficiency of (b1).

Assume that (b2) holds. Here the "orientation" comes into play. From (3.2) we must have $2n = \tilde{Z}_c((f - g^*) - (f - g)) = S_c(h - h^*)$. From (3.1)

$$(-1)^{i+1}\, \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta_i) > 0, \qquad i = 1, ..., 2n - 1$$

$$-\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta'_{2n}) > 0, \qquad -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta''_{2n}) > 0.$$

Since $(f - g^*) - (f - g) = g^* - g$ cannot, by (3.2), have any additional zeros, we must have

$$-\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta_{2n}) > 0,$$

and thus

$$(-1)^{i+1}\, \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta_i) > 0, \qquad i = 1, ..., 2n.$$

Set

$$u_j(x) = \int_{\xi_j}^{\xi_{j+1}} K(x, y) \mid h(y) - h^*(y) \mid dy, \qquad j = 1, ..., 2n.$$

Since $S_c(h - h^*) = 2n$, the function $h - h^*$ does not identically vanish on $[\xi_j, \xi_{j+1}]$ and thus $u_j \neq 0$. Furthermore, since $g \in \mathscr{M}$, we have $|h(y)| \leqslant |h^*(y)|$ for all y and thus

$$g - g^* = \sum_{j=1}^{2n} (-1)^j u_j.$$

Therefore

$$d_i = (-1)^{i+1} \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}) \sum_{j=1}^{2n} (-1)^j u_j(\theta_i) > 0, \qquad i = 1, ..., 2n. \tag{3.3}$$

Recall that

$$\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}) \, K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ y_1, ..., y_{2n} \end{pmatrix} > 0$$

for every choice of $y_1 < \cdots < y_{2n} < y_1 + 2\pi$ satisfying $\xi_i \leqslant y_i \leqslant \xi_{i+1}$, $i = 1, ..., 2n$ with $\{\xi_1, ..., \xi_{2n}\} \neq \{y_1, ..., y_{2n}\}$. Thus

$$\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}) \; U\begin{pmatrix} 1, ..., 2n \\ \theta_1, ..., \theta_{2n} \end{pmatrix} > 0.$$

A simple matrix computation (solve for the coefficient 1 of u_{2n} in (3.3)) shows that

$$1 = \operatorname{sgn}\left(-\varepsilon_n \sum_{k=1}^{2n} d_k \, U\begin{pmatrix} 1, ..., 2n-1 \\ \theta_1, ..., \hat{\theta}_k, ..., \theta_{2n} \end{pmatrix} \right),$$

where

$$U\begin{pmatrix} 1, ..., 2n-1 \\ \theta_1, ..., \hat{\theta}_k, ..., \theta_{2n} \end{pmatrix} = \det\{u_j(\theta_i)\}_{j=1}^{2n-1} {}_{\substack{i=1 \\ i \neq k}}^{2n}.$$

An additional calculation shows that

$$\operatorname{sgn} U\begin{pmatrix} 1, ..., 2n-1 \\ \theta_1, ..., \hat{\theta}_k, ..., \theta_{2n} \end{pmatrix} = \varepsilon_n$$

for each $k = 1, ..., 2n$. This is a contradiction and the sufficiency is proved.

Necessity. Assume that g^* is a periodic generalized perfect spline with $n=0$ knots. Then

$$g^*(x)=\delta\int_0^{2\pi} K(x,y)\,dy$$

for some $\delta\in\{-1,1\}$. If there is no θ such that

$$\varepsilon_1\delta(f-g^*)(\theta)=\|f-g^*\|,$$

then

$$\|f-\lambda g^*\|<\|f-g^*\|$$

for some $\lambda\in(0,1)$ (near 1), which implies that g^* is not a best approximation to f from $\mathscr{M}$. This proves the necessity in the case $n=0$.

Assume that g^* is a periodic generalized perfect spline with $2n$ ($n\geqslant 1$) knots. From Proposition 3.1 the zero function is a best approximation to $f-g^*$ from

$$\mathscr{A}=\left\{\sum_{i=1}^{2n} a_i K(x,\xi_i)+bK(x,\eta) : a_i\in\mathbb{R},\ i=1,\ldots,2n,\ b\geqslant 0\right\},$$

where $\eta\in(\xi_{2n},\xi_1+2\pi)$. This immediately implies that the zero function is a best approximation to $f-g^*$ from the QT-space

$$U_{2n}=\operatorname{span}\{K(\cdot,\xi_1),\ldots,K(\cdot,\xi_{2n})\}.$$

Thus either (b1) holds (i.e., at least $2n+2$ points of equioscillation) or we have exactly $2n$ points of equioscillation as in the statement of Theorem 2.2. It remains to prove the explicit orientation of the sign of the equioscillations as stated in (b2). Assume that $f-g^*$ equioscillates at exactly $2n$ points. Let $\{\theta_i\}_{i=1}^{2n}$, θ'_{2n}, θ''_{2n} be the associated "equioscillation" and "additional" points.

Let

$$V_{2n+1}=\operatorname{span}\{K(\cdot,\xi_1),\ldots,K(\cdot,\xi_{2n}),K(\cdot,\eta)\}.$$

V_{2n+1} is a $(2n+1)$-dimensional T-space. The zero function is therefore not a best approximation to $f-g^*$ from V_{2n+1} (see Theorem 2.1), but is a best approximation to $f-g^*$ from $\mathscr{A}$. Thus if

$$v^*(x)=\sum_{i=1}^{2n} c_i K(x,\xi_i)+dK(x,\eta)$$

is the best approximation from V_{2n+1}, then

$$\|f-g^*-v^*\| < \|f-g^*\|$$

and $d<0$.

From the first condition we have

$$\begin{aligned} v^*(\theta_i)(f-g^*)(\theta_i) &> 0, \qquad i=1, ..., 2n-1 \\ v^*(\theta'_{2n})(f-g^*)(\theta'_{2n}) &> 0 \\ v^*(\theta''_{2n})(f-g^*)(\theta''_{2n}) &> 0. \end{aligned}$$

Since no $v \in V_{2n+1} \backslash \{0\}$ has more than $2n$ zeros and $v^*(\theta'_{2n})\, v^*(\theta''_{2n}) > 0$, we must have $v^*(\theta_{2n})\, v^*(\theta'_{2n}) > 0$. Therefore v^* alternates in sign on the $\{\theta_i\}_{i=1}^{2n}$. Let $\zeta \in (\theta_{2n}, \theta_1 + 2\pi)$ be such that $v^*(\zeta) = 0$. Solving for d we obtain

$$d = \frac{\sum_{i=1}^{2n} (-1)^{i+1} v^*(\theta_i)\, K\begin{pmatrix} \theta_1, ..., \hat{\theta}_i, ..., \theta_{2n}, \zeta \\ \xi_1, ..., \xi_{2n} \end{pmatrix}}{K\begin{pmatrix} \theta_1, ..., \theta_{2n}, \zeta \\ \xi_1, ..., \xi_{2n}, \eta \end{pmatrix}}.$$

By definition,

$$\operatorname{sgn} K\begin{pmatrix} \theta_1, ..., \theta_{2n}, \zeta \\ \xi_1, ..., \xi_{2n}, \eta \end{pmatrix} = \varepsilon_{n+1}$$

and

$$\operatorname{sgn} K\begin{pmatrix} \theta_1, ..., \hat{\theta}_i, ..., \theta_{2n}, \zeta \\ \xi_1, ..., \xi_{2n} \end{pmatrix} = \sigma_1(\boldsymbol{\theta}, \boldsymbol{\xi})$$

for each $i = 1, ..., 2n$. The $v^*(\theta_i)$ alternate in sign and $d<0$. Thus

$$\operatorname{sgn}(-1)^{i+1} v^*(\theta_i) = -\varepsilon_{n+1} \sigma_1(\boldsymbol{\theta}, \boldsymbol{\xi}).$$

From Lemma 2.4, this implies that

$$\operatorname{sgn}(-1)^{i+1} v^*(\theta_i) = \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}).$$

Thus

$$\begin{aligned} (-1)^{i+1} \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta_i) &= \|f-g^*\|, \qquad i=1, ..., 2n-1 \\ -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta'_{2n}) = -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta''_{2n}) &= \|f-g^*\|, \end{aligned}$$

and the theorem is proved. ∎

The condition $|h(y)| \leqslant 1$ in the definition of $\mathscr{M}$ may be generalized to

$$l(y) \leqslant h(y) \leqslant u(y),$$

where $l, u \in \tilde{C}$ and $l < u$. The same results then hold where h jumps between being equal to l and to u on alternate intervals.

Consider the problem

$$e(\alpha) = \min\{ \|f - \alpha g\| : g \in \mathscr{M} \}.$$

For each $\alpha > 0$ there exists a unique $g_\alpha \in \mathscr{M}$ which attains the above minimum. Assuming that $f \neq \alpha g_\alpha$, the characterization of g_α is given by Theorem 3.2. How does g_α vary with α? (Since g_α is uniquely determined, it may be shown that g_α continuously varies with α.) As α increases the number of knots (and equioscillations) increases. Case (b1) of Theorem 3.2 (where the number of equioscillations is at least two more than the number of knots) occurs exactly at the α for which the number of knots of g_α increases.

Let $\tilde{\alpha}$ be the smallest value for which $f \in \tilde{\alpha}\mathscr{M}$. ($\tilde{\alpha}$ may be infinite.) For each $\alpha \in (0, \tilde{\alpha})$, set

$$g_\alpha(x) = \int_0^{2\pi} K(x, y)\, h_\alpha(y)\, dy.$$

The function h_α is a step function taking on the values ± 1 with $2k(\alpha)$ jumps; i.e., g_α has $2k(\alpha)$ knots.

PROPOSITION 3.3. *On the interval* $(0, \tilde{\alpha})$, *the value* $e(\alpha)$ *is a strictly decreasing function of* α. *Furthermore, if* $0 < \beta < \alpha < \tilde{\alpha}$, *then* $k(\beta) \leqslant k(\alpha)$. *We have* $k(\beta) < k(\alpha)$ *for all* $\alpha \in (\beta, \tilde{\alpha})$ *if* $f - \beta g_\beta$ *equioscillates on at least* $2k(\beta) + 2$ *points.*

Proof. Let $0 < \beta < \alpha < \tilde{\alpha}$. Then $\beta\mathscr{M} \subset \alpha\mathscr{M}$ and $\alpha g_\alpha \in \alpha\mathscr{M} \setminus \beta\mathscr{M}$. Thus from the uniqueness of the best approximation from $\alpha\mathscr{M}$

$$e(\alpha) < e(\beta).$$

Now assume that $f - \beta g_\beta$ equioscillates on $2m$ points. Then

$$\begin{aligned} 2m &\leqslant \tilde{Z}_c((f - \beta g_\beta) - (f - \alpha g_\alpha)) \\ &= \tilde{Z}_c(\alpha g_\alpha - \beta g_\beta) \leqslant S_c(\alpha h_\alpha - \beta h_\beta) = S_c(\alpha h_\alpha) = 2k(\alpha). \end{aligned}$$

Since $m \geqslant k(\beta)$ we obtain $k(\beta) \leqslant k(\alpha)$. If $f - \beta g_\beta$ equioscillates on at least $2k(\beta) + 2$ points, then $k(\beta) + 1 \leqslant k(\alpha)$. ∎

4. A FIXED NUMBER OF KNOTS WITH A FIXED KNOT

Let K be SCVD and for $n=1, 2, \ldots,$ set

$$\mathscr{P}_{2n}^{+}(\xi) = \left\{ \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy : \xi=\xi_1 \leqslant \xi_2 \leqslant \cdots \leqslant \xi_{2n} \leqslant \xi_{2n+1} = \xi_1 + 2\pi \right\}.$$

Note the orientation of sign at $\xi=\xi_1$. Let $f \in \tilde{C}$. In this section we assume that the best approximation to f from $\mathscr{M}$ is not in $\mathscr{P}_{2n}^{+}(\xi)$. We characterize the (unique) best approximation to f from $\mathscr{P}_{2n}^{+}(\xi)$. (Note that this set is not convex.) It follows from a standard compactness argument that a best approximation exists. The following theorem totally characterizes this best approximation.

THEOREM 4.1. *Under the above assumptions there exists a unique best approximation g^+ to f from $\mathscr{P}_{2n}^{+}(\xi)$. g^+ has the form*

$$g^{+}(x) = \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy,$$

where $\xi=\xi_1 < \xi_2 < \cdots < \xi_{2n} < \xi_{2n+1} = \xi_1 + 2\pi$, i.e., $g^+ \in \operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$. It is uniquely characterized by the fact that $f-g^+$ equioscillates on exactly $2n$ points.

Remark. $f-g^+$ cannot possibly equioscillate on $2n$ points which satisfy the conditions (b2) of Theorem 3.2, nor at more than $2n$ points. For it would then be the best approximation to f from $\mathscr{M}$. Note however that no claim is made as to any determinant vanishing which would connect the points of equioscillation and the knots. That is, there is no “orientation” involved in this result. (We could also define $\mathscr{P}_{0}^{+}(\xi)$ (which is independent of ξ and simply contains one function). The same result then holds.) In a totally parallel fashion we can of course define $\mathscr{P}_{2n}^{-}(\xi)$ and obtain the analogous result.

The proof of Theorem 4.1 is technically cumbersome. We divide the proof into two main parts. In the first part we show that if g^+ (a best approximation to f from $\mathscr{P}_{2n}^{+}(\xi)$) is contained in $\operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$, then $f-g^+$ equioscillates on $2n$ points and that this latter condition uniquely characterizes the best approximation from $\mathscr{P}_{2n}^{+}(\xi)$. In the second part we prove that a best approximation must in fact be contained in $\operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$.

PROPOSITION 4.2. *Assume that g^+, a best approximation to f from $\mathscr{P}_{2n}^{+}(\xi)$, is contained in* $\operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$. *Then $f-g^+$ equioscillates on exactly $2n$*

points. Furthermore this latter condition uniquely characterizes the best approximation to f from $\mathscr{P}^+_{2n}(\xi)$.

Proof. For each $\boldsymbol{\xi} = (\xi_1, ..., \xi_{2n})$, $\xi = \xi_1 \leqslant \xi_2 \leqslant \cdots \leqslant \xi_{2n} \leqslant \xi_{2n+1} = \xi_1 + 2\pi$, set

$$g_{\boldsymbol{\xi}}(x) = \int_0^{2\pi} K(x, y)\, h_{\boldsymbol{\xi}}(y)\, dy,$$

where

$$h_{\boldsymbol{\xi}}(y) = (-1)^{j+1}, \qquad \xi_j \leqslant y < \xi_{j+1}, \qquad j = 1, ..., 2n.$$

Obviously $S_c(h_{\boldsymbol{\xi}}) \leqslant 2n$. Moreover a simple argument (see, e.g., Pinkus [7, p. 140]) shows that for any $\boldsymbol{\xi}^1$ and $\boldsymbol{\xi}^2$, as above, we have

$$S_c(h_{\boldsymbol{\xi}^1} - h_{\boldsymbol{\xi}^2}) \leqslant 2n - 2.$$

Now assume that g^+ is a best approximation to f from $\mathscr{P}^+_{2n}(\xi)$, and that $f - g^+$ equioscillates on $2n$ points. Let $g_{\boldsymbol{\xi}^1} \in \mathscr{P}^+_{2n}(\xi)$, $g_{\boldsymbol{\xi}^1} \neq g^+$, satisfy

$$\|f - g_{\boldsymbol{\xi}^1}\| \leqslant \|f - g^+\|.$$

Then

$$2n \leqslant \tilde{Z}_c((f - g^+) - (f - g_{\boldsymbol{\xi}^1})) = \tilde{Z}_c(g_{\boldsymbol{\xi}^1} - g^+)$$

(where we count nonnodal zeros twice). Set

$$g^+(x) = \int_0^{2\pi} K(x, y)\, h^+(y)\, dy.$$

From the SCVD property of K we have

$$\tilde{Z}_c(g_{\boldsymbol{\xi}^1} - g^+) \leqslant S_c(h_{\boldsymbol{\xi}^1} - h^+) \leqslant 2n - 2,$$

which is a contradiction. Thus g^+ is necessarily the unique best approximation to f from $\mathscr{P}^+_{2n}(\xi)$.

Assume that $g^+ \in \operatorname{int} \mathscr{P}^+_{2n}(\xi)$. The perturbation argument given in Proposition 3.1 (without the η and without perturbing $\xi = \xi_1$) proves that the zero function is necessarily a best approximation to $f - g^+$ from

$$\operatorname{span}\{K(\cdot, \xi_2), ..., K(\cdot, \xi_{2n})\}.$$

These $2n - 1$ functions form a T-system and thus $f - g^+$ must equioscillate on at least $2n$ points. Since g^+ is not the best approximation to f from $\mathscr{M}$,

$f - g^+$ cannot equioscillate at more than $2n$ points. This proves the proposition. ∎

It is in proving that any best approximation is necessarily in int $\mathscr{P}_{2n}^+(\xi)$ that we encounter cumbersome technical details. To this end we first prove the result for ECVD_3 kernels. This allows us to consider first derivatives. We then show how to apply a smoothing procedure, using the de la Vallée Poussin means, to obtain the final result.

PROPOSITION 4.3. *Assume that K is* SCVD *and* ECVD_3, *and that the best approximation to f from $\mathscr{M}$ is not in $\mathscr{P}_{2n}^+(\xi)$. If g^+ is a best approximation to f from $\mathscr{P}_{2n}^+(\xi)$, then $g^+ \in \text{int}\, \mathscr{P}_{2n}^+(\xi)$.*

Proof. We assume $g^+ \notin \text{int}\, \mathscr{P}_{2n}^+(\xi)$. Thus

$$g^+(x) = \sum_{j=1}^{2k} (-1)^{j+1} \int_{\eta_j}^{\eta_{j+1}} K(x, y)\, dy,$$

where $\eta_1 < \eta_2 < \cdots < \eta_{2k} < \eta_{2k+1} = \eta_1 + 2\pi$, and $k \leqslant n-1$. For convenience we set $h^+(y) = (-1)^{j+1}$ for $y \in (\eta_j, \eta_{j+1})$, $j = 1, ..., 2k$. Note that ξ may or may not be included among the $\{\eta_j\}_{j=1}^{2k}$. Furthermore, if ξ is included among the $\{\eta_j\}_{j=1}^{2k}$, it may equal an η_s for s odd or s even. (These are different because of the orientation of the jump. We will take $s=1$ or $s=2$.) There are various cases which we will consider.

Case 1. $\xi \notin \{\eta_1, ..., \eta_{2k}\}$.

We first claim that the zero function is a best approximation to $f - g^+$ from

$$\mathscr{A} = \left\{ \sum_{i=1}^{2k} a_i K(x, \eta_i) + bK(x, \xi) : a_i \in \mathbb{R}, \sigma b \leqslant 0 \right\},$$

where $\sigma = \text{sgn}\, h^+(\xi)$.

This result is a direct consequence of Proposition 3.1. The knot ξ here plays the role of η in Proposition 3.1. Note that $g(x; \boldsymbol{\eta} + \boldsymbol{\delta}; \xi) \in \mathscr{P}_{2n}^+(\xi)$, where

$$g(x; \boldsymbol{\eta} + \boldsymbol{\delta}; \xi) = \sum_{j=1}^{2k} (-1)^{j+1} \int_{\eta_j + \delta_j}^{\eta_{j+1} + \delta_{j+1}} K(x, y)\, dy - \sigma \int_{\xi_-}^{\xi_+} K(x, y)\, dy$$

with $\xi_- = \xi$, $\xi_+ = \xi + \delta_{2k+1}$ if $\sigma = -1$, and $\xi_- = \xi - \delta_{2k+1}$, $\xi_+ = \xi$ if $\sigma = 1$.

We now apply the method of proof of Theorem 3.2. Exactly the argument found therein implies that g^+ is a best approximation to f from $\mathscr{M}$, which is a contradiction. We will present much of the argument here, as we shall not do so in the other cases.

Let

$$v(x)=\sum_{i=1}^{2k} c_i K(x, \eta_i)+dK(x, \xi)$$

be the best approximation to $f-g^+$ from the T-space

$$\text{span}\{K(\cdot, \eta_1), ..., K(\cdot, \eta_{2k}), K(\cdot, \xi)\}.$$

Therefore $f-g^+-v$ equioscillates on at least $2k+2$ points. If $\sigma d\leqslant 0$ then $v\in\mathscr{A}$. Thus $v=0$ and $f-g^+$ equioscillates on at least $2k+2$ points. But by Theorem 3.2, this implies that g^+ is a best approximation to f from $\mathscr{M}$, a contradiction. Thus $\sigma d>0$.

The zero function is a best approximation to $f-g^+$ from the QT-space

$$\text{span}\{K(\cdot, \eta_1), ..., K(\cdot, \eta_{2k})\}.$$

This, together with the fact that g^+ is not a best approximation to f from $\mathscr{M}$, implies that $f-g^+$ exhibits exactly $2k$ points of equioscillation as in the statement of Theorem 2.2. We now put this fact together with $\sigma d>0$ (word for word as in the proof of Theorem 3.2) to prove that g^+ satisfies condition (b2) of Theorem 3.2, and thus once again g^+ is a best approximation to f from $\mathscr{M}$. This contradiction implies that g^+ is not of the above form.

Case 2. $\xi\in\{\eta_1, ..., \eta_{2k}\}$ and $k\leqslant n-2$.

We first claim that the zero function is a best approximation to $f-g^+$ from

$$\mathscr{A}=\left\{\sum_{i=1}^{2k} a_i K(x, \eta_i)+bK(x, \zeta) : a_i\in\mathbb{R}, \delta b\leqslant 0\right\},$$

where $\delta=\text{sgn}\, h^+(\zeta)$, and ζ is an arbitrary knot.

We prove this using the argument to be found in Proposition 3.1. We can perturb all the knots exactly as in the proof of Proposition 3.1, and $g(x; \boldsymbol{\eta}+\boldsymbol{\delta}; \zeta)$ will not leave the class $\mathscr{P}^+_{2n}(\xi)$. We consider $g(x; \boldsymbol{\eta}+\boldsymbol{\delta}; \zeta)$ as having the $2k+4\leqslant 2n$ knots $\eta_i+\delta_i$, $i=1, ..., 2k$, ζ, $\zeta+\delta_{2k+1}$, ξ, and ξ (i.e., two knots at the point ξ!).

We now apply the exact same argument as found in Case 1 (and in the proof of Theorem 3.2) which proves that g^+ is a best approximation to f from $\mathscr{M}$. This contradiction again implies that g^+ is not of the above form.

Case 3. $\xi=\eta_1$ and $k=n-1$.

A perturbation argument, as in Cases 1 or 2 (or as in Proposition 3.1), implies that the zero function is a best approximation to $f-g^+$ from

$$\text{span}\{K(\cdot, \eta_1), ..., K(\cdot, \eta_{2n-2})\}.$$

From Theorem 2.2 there exist points $\theta_1 < \cdots < \theta_{2n-2} < \theta_1 + 2\pi$ for which

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n-2} \\ \eta_1, ..., \eta_{2n-2} \end{pmatrix} = 0, \tag{4.1}$$

and for some θ'_{2n-2}, θ''_{2n-2} satisfying $\theta_{2n-3} < \theta'_{2n-2} \leqslant \theta_{2n-2} \leqslant \theta''_{2n-2} < \theta_1 + 2\pi$, we have

$$(-1)^i \, \varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})(f-g^+)(\theta_i) = \|f-g^+\|, \qquad i=1, ..., 2n-3$$

$$\varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})(f-g^+)(\theta'_{2n-2}) = \varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})(f-g^+)(\theta''_{2n-2}) = \|f-g^+\|.$$

(If the sign were reversed, then g^+ would be a best approximation to f from $\mathscr{M}$.)

As such there also exist $\{t'_i\}_{i=1}^{2n-2}$ and $\{t''_i\}_{i=1}^{2n-2}$ satisfying $t'_i \leqslant \theta_i \leqslant t''_i < t'_{i+1}$, $i=1, ..., 2n-2$ $(t'_{2n-1} = t'_1 + 2\pi)$, for which

$$\begin{aligned} &(-1)^i \, \varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})(f-g^+)(t'_i) \\ &\qquad = (-1)^i \, \varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})(f-g^+)(t''_i) = \|f-g^+\|, \end{aligned}$$

and

$$|(f-g^+)(x)| < \|f-g^+\|, \qquad x \in (t''_i, t'_{i+1}), \qquad i=1, ..., 2n-2.$$

In each interval (t''_i, t'_{i+1}), $i=1, ..., 2n-2$, we choose a point $\tau_i \in (t''_i, t'_{i+1})$, and consider the function

$$u(y) = K\begin{pmatrix} \tau_1, ..., \tau_{2n-2} \\ y, \eta_2, ..., \eta_{2n-2} \end{pmatrix}.$$

Since u is periodic, it must have another zero ζ apart from the $\eta_2, ..., \eta_{2n-2}$ (the case of a double zero at one of these points can be avoided by a small perturbation of τ_1). From (4.1) and Lemma 2.4, we must have $\zeta \neq \eta_1$. For convenience, we assume that $\zeta \in (\eta_1, \eta_2)$.

We may also keep $\xi = \eta_1$ fixed, perturb $\eta_2, ..., \eta_{2n-2}$, and add two knots near ζ. It then follows (exactly as in Proposition 3.1) that the zero function is a best approximation to $f-g^+$ from

$$\mathscr{A} = \left\{ \sum_{i=2}^{2n-2} a_i K(x, \eta_i) + bK(x, \zeta) : a_i \in \mathbb{R}, b \leqslant 0 \right\}.$$

We claim that the zero function is not a best approximation to $f-g^+$ from

$$\operatorname{span}\{K(\cdot, \zeta), K(\cdot, \eta_2), ..., K(\cdot, \eta_{2n-2})\}.$$

This follows from Theorem 2.2. If the zero function is a best approximation there exist values $w_1 < \cdots < w_{2n-2} < w_1 + 2\pi$ which are essentially points of equioscillation, and for which

$$K\begin{pmatrix} w_1, ..., w_{2n-2} \\ \zeta, \eta_2, ..., \eta_{2n-2} \end{pmatrix} = 0.$$

But by the choice of the τ_i and equioscillation pattern of $f - g^+$, the $\{w_i\}_{i=1}^{2n-2}$ must strictly interlace the $\{\tau_i\}_{i=1}^{2n-2}$. A contradiction ensues from

$$K\begin{pmatrix} \tau_1, ..., \tau_{2n-2} \\ \zeta, \eta_2, ..., \eta_{2n-2} \end{pmatrix} = 0,$$

and Lemma 2.4.

As such there exists a

$$p(x) = dK(x, \zeta) + \sum_{i=2}^{2n-2} c_i K(x, \eta_i)$$

for which

$$\|f - g^+ - p\| < \|f - g^+\|,$$

and thus

$$(-1)^i \varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})\, p(\theta_i) > 0, \qquad i = 1, ..., 2n-2.$$

If $p \in \mathscr{A}$, then we contradict the fact that the zero function is a best approximation to $f - g^+$ from $\mathscr{A}$.

Solving for d we see that

$$\operatorname{sgn} d = \operatorname{sgn} \frac{-\varepsilon_{n-1} \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta})\, K\begin{pmatrix} \theta_2, ..., \theta_{2n-2} \\ \eta_2, ..., \eta_{2n-2} \end{pmatrix}}{K\begin{pmatrix} \theta_1, ..., \theta_{2n-2} \\ \zeta, \eta_2, ..., \eta_{2n-2} \end{pmatrix}}.$$

Now,

$$\operatorname{sgn} K\begin{pmatrix} \theta_2, ..., \theta_{2n-2} \\ \eta_2, ..., \eta_{2n-2} \end{pmatrix} = \varepsilon_{n-1}$$

(the signs are the same if we delete any θ_i rather than θ_1), and

$$\operatorname{sgn} K\begin{pmatrix} \theta_1, ..., \theta_{2n-2} \\ \zeta, \eta_2, ..., \eta_{2n-2} \end{pmatrix} = \sigma_2(\boldsymbol{\theta}, \boldsymbol{\eta}).$$

This implies that $d<0$ and thus $p \in \mathscr{A}$, which is a contradiction. g^+ is not of the above form.

Case 4. $\xi = \eta_2$ and $k = n-1$.

We assume that g^+ has the form

$$g^+(x) = \sum_{j=1}^{2n-2} (-1)^{j+1} \int_{\eta_j}^{\eta_{j+1}} K(x, y)\, dy,$$

with $\xi = \eta_2$. We are now limited in our perturbation. (The two extra knots must be used to alter the orientation at ξ.) It is here that we make use of the ECVD_3 property of K.

We first claim that the zero function is a best approximation to $f - g^+$ from

$$\mathscr{A} = \left\{ \sum_{i=1}^{2n-2} a_i K(x, \eta_i) + bK'_y(x, \eta_2) : a_i \in \mathbb{R},\, b \geqslant 0 \right\}.$$

The proof of this fact parallels the proof of Proposition 3.1. Assume that the zero function is not a best approximation to $f - g^+$ from $\mathscr{A}$. There then exists a

$$v(x) = \sum_{j=1}^{2n-2} a_j K(x, \eta_j) + bK'_y(x, \eta_2) \in \mathscr{A}$$

such that

$$\|f - g^+ - v\| < \|f - g^+\|.$$

Thus for every $\lambda \in [0, 1]$ we have

$$\|f - g^+ - \lambda v\| \leqslant \|f - g^+\| - \lambda c,$$

where

$$c = \|f - g^+\| - \|f - g^+ - v\| > 0.$$

Set $\delta_j = \frac{1}{2}(-1)^j a_j \lambda$, $\lambda > 0$, small, $j = 1, ..., 2n-2$, $\delta_{2n-1} = \delta_1$, and $\delta'_2 = \sqrt{b\lambda/2}$, $\delta''_2 = \delta_2 + \delta'_2$. Let $g(x; \boldsymbol{\eta}) = g^+(x)$, and for $\boldsymbol{\delta} = (\delta_1, ..., \delta_{2n-2}, \delta'_2)$ as above, set

$$g(x; \boldsymbol{\eta} + \boldsymbol{\delta}) = \sum_{j=3}^{2n-2} (-1)^{j+1} \int_{\eta_j + \delta_j}^{\eta_{j+1} + \delta_{j+1}} K(x, y)\, dy + \int_{\eta_1 + \delta_1}^{\eta_2 - \delta'_2} K(x, y)\, dy$$
$$- \int_{\eta_2 - \delta'_2}^{\eta_2} K(x, y)\, dy + \int_{\eta_2}^{\eta_2 + \delta''_2} K(x, y)\, dy - \int_{\eta_2 + \delta''_2}^{\eta_3 + \delta_3} K(x, y)\, dy.$$

Now for $\lambda > 0$, small,

$$\begin{aligned}
&g(x; \boldsymbol{\eta}+\boldsymbol{\delta}) - g(x; \boldsymbol{\eta}) \\
&\quad = \sum_{j=3}^{2n-2} (-1)^{j+1} \left[\int_{\eta_{j+1}}^{\eta_{j+1}+\delta_{j+1}} K(x, y)\, dy - \int_{\eta_j}^{\eta_j+\delta_j} K(x, y)\, dy \right] \\
&\qquad - \int_{\eta_1}^{\eta_1+\delta_1} K(x, y)\, dy - \int_{\eta_3}^{\eta_3+\delta_3} K(x, y)\, dy \\
&\qquad + 2 \int_{\eta_2}^{\eta_2+\delta_2''} K(x, y)\, dy - 2 \int_{\eta_2-\delta_2'}^{\eta_2} K(x, y)\, dy \\
&\quad = 2 \sum_{\substack{j=1 \\ j \neq 2}}^{2n-2} (-1)^j \int_{\eta_j}^{\eta_j+\delta_j} K(x, y)\, dy + 2 \int_{\eta_2+\delta_2'}^{\eta_2+\delta_2''} K(x, y)\, dy \\
&\qquad + 2 \left[\int_{\eta_2}^{\eta_2+\delta_2'} K(x, y)\, dy - \int_{\eta_2-\delta_2'}^{\eta_2} K(x, y)\, dy \right] \\
&\quad = \left[2 \sum_{j=1}^{2n-2} (-1)^j \delta_j K(x, \eta_j) + o(\boldsymbol{\delta}) \right] + [2(\delta_2')^2 K_y'(x, \eta_2) + o((\delta_2')^2)] \\
&\quad = \lambda v(x) + o(\lambda).
\end{aligned}$$

If $b = 0$, then $\delta_2' = 0$ and $\delta_2'' = \delta_2$, so that $g(\cdot\,; \boldsymbol{\eta}+\boldsymbol{\delta}) \in \mathscr{P}_{2n-2} \subset \mathscr{P}_{2n}^+(\xi)$. If $b > 0$, then $\delta_2', \delta_2'' > 0$ for $\lambda > 0$ sufficiently small, and $g(\cdot\,; \boldsymbol{\eta}+\boldsymbol{\delta}) \in \mathscr{P}_{2n}^+(\xi)$. Thus

$$\begin{aligned}
\|f - g^+\| = \|f - g(\cdot\,; \boldsymbol{\eta})\| &\leqslant \|f - g(\cdot\,; \boldsymbol{\eta}+\boldsymbol{\delta})\| \\
&= \|f - (g(\cdot\,; \boldsymbol{\eta}) + \lambda v + o(\lambda))\| \\
&= \|f - g^+ - \lambda v\| + o(\lambda) \\
&\leqslant \|f - g^+\| - \lambda c + o(\lambda).
\end{aligned}$$

But then for $\lambda > 0$, sufficiently small, a contradiction ensues.

We may now apply the argument found in the proof of Case 1 and in Theorem 3.2. Two things which should be noted are that

$$\operatorname{span}\{K(\cdot, \eta_1), ..., K(\cdot, \eta_{2n-2}), K_y'(\cdot, \eta_2)\}$$

is a T-space, and the determinant

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n-2}, \zeta \\ \eta_1, \eta_2, \eta_2, \eta_3, \eta_4, ..., \eta_{2n-2} \end{pmatrix},$$

for any $\zeta \in (\theta_{2n-2}, \theta_1 + 2\pi)$ is of sign ε_n. For these two properties to hold we need the ECVD_3 property. ∎

We have proved Theorem 4.1 in the case where K is ECVD_3. It remains to remove this extraneous restriction.

PROPOSITION 4.4. *Assume that K is* SCVD, *and that the best approximation to f from $\mathscr{M}$ is not in $\mathscr{P}_{2n}^{+}(\xi)$. If g^{+} is a best approximation to f from $\mathscr{P}_{2n}^{+}(\xi)$, then $g^{+} \in \operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$.*

Proof. We recall that the de la Vallée Poussin kernel

$$\omega_m(t) = \frac{1}{\binom{2m}{m}} \sum_{v=-m}^{m} \binom{2m}{m+v} e^{ivt},$$

is ECVD_{2m+1}. In addition for each function $f \in \tilde{C}$, the transformation

$$V_m(x) = \frac{1}{2\pi} \int_0^{2\pi} \omega_m(x-y)\, f(y)\, dy$$

defines the de la Vallée Poussin means (or V-means) of f. This V_m is a trigonometric polynomial of degree at most m, and uniformly converges to f as $m \to \infty$.

For K, which is SCVD, let

$$K_m(x, y) = \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} \omega_m(x-t)\, K(t, s)\, \omega_m(y-s)\, dt\, ds.$$

It follows from the basic composition formula (see Karlin [4, p. 17]) that K_m is ECVD_{2m+1}. Furthermore, K_m converges uniformly to K as $m \to \infty$.

Set

$$\mathscr{P}_{2n}^{+}(\xi; m) = \left\{ \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K_m(x, y)\, dy \right\}.$$

Let f be as above. It may be shown that for m sufficiently large, the best approximation to f from $\mathscr{P}_{2n}^{+}(\xi; m)$, which we will denote by $g^{+}(\cdot\,; m)$, is not the best approximation to f from the associated $\mathscr{M}_m$. As such it follows from Propositions 4.2 and 4.3 that $g^{+}(\cdot\,; m) \in \operatorname{int} \mathscr{P}_{2n}^{+}(\xi; m)$, it is unique, and $f - g^{+}(\cdot\,; m)$ equioscillates on exactly $2n$ points. Let $m \to \infty$. Then $g^{+}(\cdot\,; m)$ converges uniformly to some $g^{+} \in \mathscr{P}_{2n}^{+}(\xi)$. While it is possible that in the limit $g^{+} \notin \operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$, it nevertheless follows that $f - g^{+}$ exhibits at least $2n$ points of equioscillation. If $g^{+} \notin \operatorname{int} \mathscr{P}_{2n}^{+}(\xi)$ (or $f - g^{+}$ exhibits more than $2n$ points of equioscillation), then g^{+} is the best approximation to f

from $\mathcal{M}$. This contradiction implies that $g^+ \in \operatorname{int} \mathcal{P}_{2n}^+(\xi)$. From Proposition 4.2 g^+ is the unique best approximation to f from $\mathcal{P}_{2n}^+(\xi)$. ∎

Propositions 4.2, 4.3, and 4.4 together prove Theorem 4.1.

5. A FIXED NUMBER OF KNOTS

The best approximation from the set $\mathcal{P}_{2n}^+(\xi)$ ($\mathcal{P}_{2n}^-(\xi)$) is unique and may be easily characterized. Furthermore, if the best approximation is not a best approximation from $\mathcal{M}$, then this characterization is simple and has no "orientation" component.

For $n = 1, 2, ...,$ set

$$\mathcal{P}_{2n} = \bigcup_{\xi} \mathcal{P}_{2n}^+(\xi).$$

That is, $\mathcal{P}_{2n}$ is the set of *periodic generalized perfect splines* with at most $2n$ knots. For $n = 0$,

$$\mathcal{P}_0 = \left\{ \pm \int_0^{2\pi} K(x, y)\, dy \right\}.$$

Is the best approximation to $f \in \tilde{C}$ from $\mathcal{P}_{2n}$ unique and can it be easily characterized? The answer to both questions is no. We present a necessary condition for a best approximation from $\mathcal{P}_{2n}$ (stronger than Theorem 4.1), but also show that this condition is not sufficient. Furthermore we construct a function with many best approximations from $\mathcal{P}_{2n}$. (Note that $\mathcal{P}_{2n}$ is compact, and thus there always exists a best approximation.) In what follows we will take $n \geqslant 1$. The case $\mathcal{P}_0$ is not at all difficult, but is somewhat different.

PROPOSITION 5.1. *Let $f \in \tilde{C}$ and assume that the best approximation to f from $\mathcal{M}$ is not in $\mathcal{P}_{2n}$. If g^* is a best approximation to f from $\mathcal{P}_{2n}$ then*

$$g^*(x) = \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j}^{\xi_{j+1}} K(x, y)\, dy$$

for some $\xi_1 < \cdots < \xi_{2n} < \xi_{2n+1} = \xi_1 + 2\pi$, i.e., $g^* \in \operatorname{int} \mathcal{P}_{2n}$, and there exist $\theta_1 < \cdots < \theta_{2n} < \theta_1 + 2\pi$ such that

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, ..., \xi_{2n} \end{pmatrix} = 0$$

and for some θ'_{2n}, θ''_{2n} *satisfying* $\theta_{2n-1} < \theta'_{2n} \leqslant \theta_{2n} \leqslant \theta''_{2n} < \theta_1 + 2\pi$ *we have*

$$(-1)^i \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta_i) = \|f-g^*\|, \qquad i=1, ..., 2n-1$$

$$\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta'_{2n}) = \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g^*)(\theta''_{2n}) = \|f-g^*\|.$$

Proof. It follows from Theorem 4.1 that $g^* \in \operatorname{int} \mathscr{P}_{2n}$ (and that $f-g^*$ exhibits exactly $2n$ points of equioscillation). The perturbation technique found in Proposition 3.1 implies that the zero function is a best approximation to $f-g^*$ from

$$\operatorname{span}\{K(\cdot, \xi_1), ..., K(\cdot, \xi_{2n})\}.$$

Since this is a QT-space of dimension $2n$, and g^* is not the best approximation to f from $\mathscr{M}$, the remaining statement of the theorem follows. (See also the proof of Case 3 in Proposition 4.3.) ∎

The necessary conditions of Proposition 5.1 are not, in general, sufficient and the best approximation from $\mathscr{P}_{2n}$ is not necessarily unique. We construct an example which exhibits these traits.

Let $k \in \tilde{C}$, $k > 0$, be such that $K(x, y) = k(x-y)$ is SCVD. (In this case $\varepsilon_n = 1$ for all n.) For each $\alpha \in [0, \pi/n)$, set

$$h_{\alpha, n}(y) = (-1)^j, \qquad \alpha + \frac{j\pi}{n} \leqslant y < \alpha + \frac{(j+1)\,\pi}{n},$$

$j = 0, 1, ..., 2n-1$. Let

$$g_\alpha(x) = \int_0^{2\pi} k(x-y)\, h_{\alpha, n}(y)\, dy.$$

Since $g_\alpha(x+\pi/n) = -g_\alpha(x)$ there exists a $\beta \in [0, \pi/n)$ such that g_α alternately attains its norm at the $2n$ points $\beta + \alpha + i\pi/n$, $i = 0, 1, ..., 2n-1$. (We can and will assume that g_α attains its norm only at these $2n$ points.)

It is a known fact, see Pinkus [7, p. 174], that each g_α is a function of minimum norm in $\mathscr{P}_{2n}$. Thus each g_α is a best approximation to $f=0$ from $\mathscr{P}_{2n}$ and uniqueness does not hold.

Note that from the necessary conditions of Proposition 5.1, we have that

$$K\begin{pmatrix} \beta, \beta + \frac{\pi}{n}, ..., \beta + \frac{(2n-1)\,\pi}{n} \\ 0, \frac{\pi}{n}, ..., \frac{(2n-1)\,\pi}{n} \end{pmatrix} = 0$$

and (since $\varepsilon_n = 1$ for all n)

$$g_\alpha\left(\beta + \alpha + \frac{i\pi}{n}\right) = (-1)^i \sigma_2(\beta) \|g_\alpha\|, \qquad i = 0, ..., 2n-1,$$

where the $\sigma_2(\beta)$ is the sign of the appropriate determinant.

Consider g_0 which equioscillates at $\beta + i\pi/n$, $i = 0, 1, ..., 2n-1$. For any f such that $f(\beta + i\pi/n) = 0$, $i = 0, 1, ..., 2n-1$, and

$$\|f - g_0\| = \|g_0\|$$

we have the necessary conditions of Proposition 5.1 holding. Consider g_α for any $\alpha \in (0, \pi/n)$. It is not difficult to see that we may construct $f \in \tilde{C}$, with $f(\beta + i\pi/n) = 0$, $i = 0, 1, ..., 2n-1$, and further satisfying

$$\|f - g_\alpha\| < \|g_\alpha\| = \|g_0\| = \|f - g_0\|.$$

This shows that the necessary condition of Proposition 5.1 is not sufficient.

We now consider a different problem connected with $\mathscr{P}_{2n}$. For $f \in \tilde{C}$, set

$$E_{2n}(\alpha) = \min\{\|f - \alpha g\| : g \in \mathscr{P}_{2n}\}.$$

(Compare this value with the $e(\alpha)$ of Section 3.) We prove the following result.

THEOREM 5.2. *There exists an* $\alpha^* \in [0, \infty)$ *with the following properties*:

(1) *On the interval* $(0, \alpha^*]$ *the value* $E_{2n}(\alpha)$ *is strictly decreasing and* $E_{2n}(\alpha) = e(\alpha)$.

(2) *On the interval* (α^*, ∞) *the value* $E_{2n}(\alpha)$ *is strictly increasing and* $E_{2n}(\alpha) > e(\alpha)$.

(3) *For* $\alpha = \alpha^*$ *there exists a unique* $g_{\alpha^*} \in \mathscr{P}_{2n}$ *which attains the minimum in the above.*

(4) g_{α^*} *is uniquely characterized by the property that* $f - \alpha^* g_{\alpha^*}$ *equioscillates on at least* $2n+2$ *points.*

Proof. For each α, let $g_\alpha \in \mathscr{P}_{2n}$ be such that

$$E_{2n}(\alpha) = \|f - \alpha g_\alpha\|.$$

We know from the results of this section that the g_α is not necessarily uniquely defined.

(1) From Proposition 3.3 it follows that if αg_α is the best approximation to f from $\alpha\mathscr{M}$, i.e., $E_{2n}(\alpha) = e(\alpha)$, then for all $\beta < \alpha$ the function βg_β is

also the best approximation to f from $\beta\mathcal{M}$. Thus $E_{2n}(\alpha)=e(\alpha)$ on some interval $(0, \alpha^*]$. Now $\alpha^*<\infty$. To see this note that $e(\alpha)$ is a non-increasing function while, since

$$\min_{g\in\mathcal{P}_{2n}} \|g\|>0,$$

it follows that $E_{2n}(\alpha)\to\infty$ as $\alpha\to\infty$.

(2) This is the more technically difficult proof in this theorem. We prove it by a perturbation argument.

Let $\alpha>\alpha^*$. From Proposition 5.1, we have that g_α (which is not necessarily uniquely defined) has the form

$$g_\alpha(x)=\sum_{j=1}^{2n}(-1)^{j+1}\int_{\xi_j}^{\xi_{j+1}} K(x, y)\,dy$$

for some $\xi_1<\cdots<\xi_{2n}<\xi_{2n+1}=\xi_1+2\pi$. Furthermore there exist $\theta_1<\cdots<\theta_{2n}<\theta_1+2\pi$ such that

$$K\begin{pmatrix}\theta_1, ..., \theta_{2n}\\ \xi_1, ..., \xi_{2n}\end{pmatrix}=0$$

and for some θ'_{2n}, θ''_{2n} satisfying $\theta_{2n-1}<\theta'_{2n}\leqslant\theta_{2n}\leqslant{}''_{2n}<\theta_1+2\pi$ we have

$$(-1)^i\,\varepsilon_n\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g_\alpha)(\theta_i)=\|f-g_\alpha\|, \qquad i=1, ..., 2n-1$$

$$\varepsilon_n\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g_\alpha)(\theta'_{2n})=\varepsilon_n\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f-g_\alpha)(\theta''_{2n})=\|f-g_\alpha\|.$$

We first note that the functions

$$\{K(\cdot, \xi_1), ..., K(\cdot, \xi_{2n}), g_\alpha(\cdot)\}$$

form a T-system of dimension $2n+1$. In addition, a simple calculation shows that if we evaluate these functions at $2n+1$ consecutive points, then the sign of the associated determinant is $-\varepsilon_{n+1}$.

Since $f-\alpha g_\alpha$ equioscillates on exactly $2n$ points, there exists a

$$v(x)=\sum_{j=1}^{2n} a_j K(x, \xi_j)+d g_\alpha(x)$$

such that

$$\|f-\alpha(g_\alpha+v)\|<\|f-\alpha g_\alpha\|.$$

Thus for every $\lambda \in [0, 1]$

$$\|f-\alpha(g_\alpha+\lambda v)\| \leqslant \|f-\alpha g_\alpha\| - \lambda c,$$

where $c = \|f-\alpha g_\alpha\| - \|f-\alpha(g_\alpha+v)\|$.

From the above it follows that

$$(-1)^i \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})\, v(\theta_i) > 0, \qquad i=1, ..., 2n.$$

These inequalities (together with the sign of the associated determinant of the $(2n+1)$-dimensional T-system) imply that $d<0$, as was the case in the proof of Theorem 3.2.

Let $\delta_j = \frac{1}{2}(-1)^j a_j \lambda$, $j=1, ..., 2n$, $\delta_{2n+1} = \delta_1$, and

$$g(x; \boldsymbol{\xi}+\boldsymbol{\delta}) = \sum_{j=1}^{2n} (-1)^{j+1} \int_{\xi_j+\delta_j}^{\xi_{j+1}+\delta_{j+1}} K(x, y)\, dy.$$

From Proposition 3.1, we have that

$$g(x; \boldsymbol{\xi}+\boldsymbol{\delta}) - g_\alpha(x) = \lambda \sum_{j=1}^{2n} a_j K(x, \xi_j) + o(\lambda).$$

Thus

$$\begin{aligned}(1+\lambda d)\, g(x; \boldsymbol{\xi}+\boldsymbol{\delta}) - g_\alpha(x) &= (1+\lambda d)(g(x; \boldsymbol{\xi}+\boldsymbol{\delta}) - g_\alpha(x)) + \lambda\, d g_\alpha(x)\\ &= \lambda v(x) + o(\lambda).\end{aligned}$$

Set $\alpha_\lambda = \alpha(1+\lambda d)$. Since $d<0$ we have $\alpha_\lambda < \alpha$. Now

$$\begin{aligned}E(\alpha_\lambda) &\leqslant \|f-\alpha_\lambda g(\cdot; \boldsymbol{\xi}+\boldsymbol{\delta})\| = \|f-\alpha(g_\alpha+\lambda v)+o(\lambda)\|\\ &\leqslant \|f-\alpha(g_\alpha+\lambda v)\| + o(\lambda) \leqslant E(\alpha) - \lambda c + o(\lambda).\end{aligned}$$

For $\lambda>0$, sufficiently small, $\lambda c - o(\lambda) > 0$ and thus $E(\alpha_\lambda) < E(\alpha)$. This implies that $E(\alpha)$ is strictly increasing on (α^*, ∞).

(3) and (4). Let αm_α be the best approximation to f from $\alpha\mathscr{M}$, i.e.,

$$e(\alpha) = \|f-\alpha m_\alpha\|,$$

as in Proposition 3.3. For $\alpha \leqslant \alpha^*$ this m_α and the g_α which appears in $E_{2n}(\alpha)$ are identical. However, for $\alpha > \alpha^*$ this m_α is not in $\mathscr{P}_{2n}$, i.e., it has more than $2n$ knots, and $f-\alpha m_\alpha$ exhibits at least $2n+2$ points of equioscillation. As such, from continuity considerations ($\lim_{\alpha \downarrow \alpha^*} m_\alpha = m_{\alpha^*} = g_{\alpha^*}$), $g_{\alpha^*} \in \mathscr{P}_{2n}$ and $f-\alpha^* g_{\alpha^*}$ exhibits at least $2n+2$ points of equioscillation.

We claim that for any $g \in \mathscr{P}_{2n}$ and $\alpha \geqslant 0$, $\alpha g \neq \alpha^* g_{\alpha^*}$, we have

$$\|f - \alpha^* g_{\alpha^*}\| < \|f - \alpha g\|.$$

Assume not. Then

$$2n+2 \leqslant \tilde{Z}_c((f-\alpha^* g_{\alpha^*}) - (f - \alpha g)) = \tilde{Z}_c(\alpha g - \alpha^* g_{\alpha^*}) \leqslant S_c(\alpha h - \alpha^* h_{\alpha^*}) \leqslant 2n,$$

which is a contradiction. ∎

6. NONNEGATIVE MEASURES

Let $\mathscr{B}$ denote the set of finite Borel measures on $[0, 2\pi)$, and set

$$\mathscr{M}_\infty = \left\{ g(x) = \int_0^{2\pi} K(x, y)\, d\mu(y) : \mu \in \mathscr{B}, \mu \geqslant 0 \right\},$$

where by $\mu \geqslant 0$ we mean that μ is a nonnegative measure. In addition, we set

$$\mathscr{Q}_n = \left\{ \sum_{i=1}^n a_i K(x, \xi_i) : a_i \geqslant 0, \xi_1 \leqslant \cdots \leqslant \xi_n \leqslant \xi_1 + 2\pi \right\},$$

and for any $\xi \in [0, 2\pi)$

$$\mathscr{Q}_n(\xi) = \left\{ \sum_{i=1}^n a_i K(x, \xi_i) \in \mathscr{Q}_n : \xi_1 = \xi \right\}.$$

It will not suffice, in this section, to only assume that K is SCVD. We need slightly more. We assume throughout this section that $K \in \tilde{C}^2$ is continuously differentiable in y, and for each positive integer m there exists an $\varepsilon_m \in \{-1, 1\}$ such that

$$\varepsilon_m K\begin{pmatrix} x_1, ..., x_{2m-1} \\ y_1, ..., y_{2m-1} \end{pmatrix} > 0$$

for all $x_1 < \cdots < x_{2m-1} < x_1 + 2\pi$ and $y_1 \leqslant \cdots \leqslant y_{2m-1} < y_1 + 2\pi$, where at most two consecutive y_j's are permitted to be equal. If $y_j = y_{j+1}$, we replace column $j+1$ by $\{\partial K(x_i, y_j)/\partial y\}_{i=1}^{2m-1}$.

We will prove three main results, paralleling those obtained for $\mathscr{M}$, $\mathscr{P}_n$, and $\mathscr{P}_{2n}^+(\xi)$. We characterize the unique best approximation to $f \in \tilde{C}$ from $\mathscr{M}_\infty$. We give necessary (but not sufficient) conditions for best approximations to f from $\mathscr{Q}_n$, and we characterize the unique best approximation to

f from $\mathcal{Q}_n(\xi)$. Note that $\mathcal{M}_\infty$, $\mathcal{Q}_n$, and $\mathcal{Q}_n(\xi)$ are existence sets. This can be shown in an analogous way to the proof of Lemma 5.2 in Pinkus [6].

An essential tool in proving these results is the following perturbation result.

PROPOSITION 6.1. *Let $n \geqslant 1$, and $g^*(x) = \sum_{i=1}^n a_i K(x, \xi_i)$ be a best approximation to $f \in \tilde{C}$ from $\mathcal{M}_\infty$, where $a_i > 0$, $i = 1, ..., n$, and $\xi_1 < \cdots < \xi_n < \xi_1 + 2\pi$. Then for any $\eta \notin \{\xi_1, ..., \xi_n\}$, the zero function is a best approximation to $f - g^*$ from*

$$\mathcal{A} = \left\{ \sum_{i=1}^n b_i K(x, \xi_i) + c_i \frac{\partial K(x, \xi_i)}{\partial y} + dK(x, \eta) : b_i, c_i \in \mathbb{R}, i = 1, ..., n, d \geqslant 0 \right\}.$$

Proof. Assume not. Let

$$v(x) = \sum_{i=1}^n b_i K(x, \xi_i) + c_i \frac{\partial K(x, \xi_i)}{\partial y} + dK(x, \eta),$$

$d \geqslant 0$, satisfy

$$\|f - g^* - v\| < \|f - g^*\|.$$

Then for each $\lambda \in [0, 1]$

$$\|f - g^* - \lambda v\| \leqslant \|f - g^*\| - \lambda c,$$

where $c = \|f - g^*\| - \|f - g^* - v\|$.

For $\lambda > 0$, small, we set

$$g_\lambda(x) = \sum_{i=1}^n (a_i + \lambda b_i) K(x, \xi_i + \delta_i) + \lambda\, dK(x, \eta)$$

where $\delta_i = \lambda c_i / (a_i + \lambda b_i)$. Since $a_i > 0$, we have $\delta_i \sim \lambda$ for all $i = 1, ..., n$; i.e., they have the same order as $\lambda \downarrow 0$. Now

$$\begin{aligned}
g_\lambda(x) - g^*(x) &= \sum_{i=1}^n (a_i + \lambda b_i) K(x, \xi_i + \delta_i) + \lambda\, dK(x, \eta) - \sum_{i=1}^n a_i K(x, \xi_i) \\
&= \sum_{i=1}^n \lambda b_i K(x, \xi_i) + \sum_{i=1}^n (a_i + \lambda b_i)[K(x, \xi_i + \delta_i) - K(x, \xi_i)] + \lambda\, dK(x, \eta) \\
&= \lambda \sum_{i=1}^n b_i K(x, \xi_i) + \lambda \sum_{i=1}^n c_i \left[\frac{K(x, \xi_i + \delta_i) - K(x, \xi_i)}{\delta_i} \right] + \lambda\, dK(x, \eta) \\
&= \lambda v(x) + o(\lambda).
\end{aligned}$$

Since g^* is a best approximation to f from $\mathscr{M}_\infty$,

$$\|f-g^*\| \leqslant \|f-g_\lambda\| = \|f-(g^*+\lambda v+o(\lambda))\| \leqslant \|f-g^*-\lambda v\| + o(\lambda)$$
$$\leqslant \|f-g^*\| - \lambda c + o(\lambda).$$

For $\lambda > 0$, small, $\lambda c - o(\lambda) > 0$, and a contradiction ensues. ∎

We will need the following analogue of Lemmas 2.3 and 2.4. We present this result without proof as it is a variant on these results. However, it does need and use the previously assumed "extended" SCVD property of K.

LEMMA 6.2. *Let* $\theta_1 < \cdots < \theta_{2n} < \theta_{2n+1} = \theta_1 + 2\pi$, *and* $\xi_1 < \cdots < \xi_n < \xi_{n+1} = \xi_1 + 2\pi$. *Assume that*

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, \xi_1, ..., \xi_n, \xi_n \end{pmatrix} = 0.$$

Then

(a) *For all* $\eta_i \in (\xi_i, \xi_{i+1})$, $i = 1, ..., n$,

$$\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})\, K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, \eta_1, ..., \xi_n, \eta_n \end{pmatrix} > 0$$

for some $\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}) \in \{-1, 1\}$.

(b) *For every choice of* $\zeta_1 < \cdots < \zeta_{2n} < \zeta_1 + 2\pi$ *satisfying* $\theta_i \leqslant \zeta_i \leqslant \theta_{i+1}$, $i = 1, ..., 2n$, $\{\theta_1, ..., \theta_{2n}\} \neq \{\zeta_1, ..., \zeta_{2n}\}$,

$$\sigma_1(\boldsymbol{\theta}, \boldsymbol{\xi})\, K\begin{pmatrix} \zeta_1, ..., \zeta_{2n} \\ \xi_1, \xi_1, ..., \xi_n, \xi_n \end{pmatrix} > 0$$

for some $\sigma_1(\boldsymbol{\theta}, \boldsymbol{\xi}) \in \{-1, 1\}$.

(c) $\sigma_1(\boldsymbol{\theta}, \boldsymbol{\xi})\, \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi}) = -\varepsilon_n \varepsilon_{n+1}$.

We can now state and prove the theorem concerning best approximation from $\mathscr{M}_\infty$.

THEOREM 6.3. *Assume that* K *is as above, and* $f \in \tilde{C} \backslash \mathscr{M}_\infty$. *There exists a unique best approximation* g^* *to* f *from* $\mathscr{M}_\infty$. *Either* $g^* = 0$ *or for some* $n \geqslant 1$, g^* *has the form*

$$g^*(x) = \sum_{i=1}^{n} a_i K(x, \xi_i),$$

where the $a_i > 0$, $i = 1, ..., n$, *and* $\xi_1 < \cdots < \xi_n < \xi_1 + 2\pi$. g^* *is uniquely characterized as follows.*

(a) $g^* = 0$ *if and only if there exists a* θ *such that*

$$f(\theta) = -\varepsilon_1 \|f\|.$$

(b) *If* $n \geqslant 1$, *then one of the following holds*:

(b1) $f - g^*$ *equioscillates on* $2n+2$ *points,*

(b2) *there exist* $\theta_1 < \cdots < \theta_{2n} < \theta_1 + 2\pi$ *such that*

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, \xi_1, ..., \xi_n, \xi_n \end{pmatrix} = 0$$

and for some θ'_{2n}, θ''_{2n} *satisfying* $\theta_{2n-1} < \theta'_{2n} \leqslant \theta_{2n} \leqslant \theta''_{2n} < \theta_1 + 2\pi$ *we have*

$$\begin{gathered} (-1)^{i+1} \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta_i) = \|f - g^*\|, \qquad i = 1, ..., 2n-1 \\ -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta'_{2n}) = -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta''_{2n}) = \|f - g^*\|. \end{gathered} \tag{6.1}$$

Proof. We refer the reader to Lemmas 5.2–5.5 of Pinkus [6] for a proof of the fact that any best approximation to $f \in \tilde{C} \backslash \mathscr{M}_\infty$ from $\mathscr{M}_\infty$ is necessarily of the form $g^* = 0$ or

$$g^*(x) = \sum_{i=1}^{n} a_i K(x, \xi_i),$$

where the $a_i > 0$, $i = 1, ..., n$, and $\xi_1 < \cdots < \xi_n < \xi_1 + 2\pi$. The necessity, sufficiency, and uniqueness of these conditions in the case $g^* = 0$ is easily checked and is left to the reader.

Sufficiency and Uniqueness. Assume that $\|f - g\| \leqslant \|f - g^*\|$ for some $g \in \mathscr{M}_\infty$ of the form

$$g(x) = \int_0^{2\pi} K(x, y)\, d\mu(y)$$

for some $\mu \in \mathscr{B}$, $\mu \geqslant 0$. Set

$$g^*(x) = \sum_{i=1}^{n} a_i K(x, \xi_i) = \int_0^{2\pi} K(x, y)\, d\mu^*(y)$$

(i.e., $d\mu^* = \sum_{i=1}^{n} a_i \delta_{\xi_i}$). Thus

$$\tilde{Z}_c((f - g^*) - (f - g)) = \tilde{Z}_c(g - g^*) \leqslant S_c(\mu - \mu^*).$$

From the form of $\mu - \mu^*$, we know that $S_c(\mu - \mu^*) \leqslant 2n$. If (b1) holds, and $f - g^*$ equioscillates on $2n+2$ points, then

$$2n + 2 \leqslant \tilde{Z}_c((f - g^*) - (f - g))$$

and a contradiction ensues.

Assume that (b2) holds. Then

$$2n = \tilde{Z}_c(g - g^*) = S_c(\mu - \mu^*).$$

Now

$$(g - g^*)(x) = \int_0^{2\pi} K(x, y)\, d(\mu - \mu^*)(y) = \sum_{i=1}^{2n} c_i u_i(x),$$

where $u_{2i-1}(x) = K(x, \xi_i)$, $i = 1, ..., n$, and

$$u_{2i}(x) = \int_{\xi_i +}^{\xi_{i+1} -} K(x, y)\, d\mu(y), \qquad i = 1, ..., n,$$

Note that $c_{2i} = 1$, $i = 1, ..., n$. Since $S_c(\mu - \mu^*) = 2n$, none of the u_i (u_{2i}) vanish identically, and the $\{u_i\}_{i=1}^{2n}$ are a QT-system. From (6.1) we have

$$\begin{aligned} &(-1)^{i+1}\, \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\delta})(g - g^*)(\theta_i) \geqslant 0, \qquad i = 1, ..., 2n-1 \\ &\quad -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta'_{2n}) \geqslant 0, \qquad -\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta''_{2n}) \geqslant 0. \end{aligned} \tag{6.2}$$

Since the $\{u_i\}_{i=1}^{2n}$ are a QT-system, we also have

$$-\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta_{2n}) \geqslant 0.$$

Therefore

$$(-1)^{i+1}\, \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(g - g^*)(\theta_i) \geqslant 0, \qquad i = 1, ..., 2n. \tag{6.3}$$

Now

$$U\begin{pmatrix} 1, ..., 2n \\ \theta_1, ..., \theta_{2n} \end{pmatrix} = \int_{\xi_1 +}^{\xi_2 -} \cdots \int_{\xi_n +}^{(\xi_1 + 2\pi) -} K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, \eta_1, ..., \xi_n, \eta_n \end{pmatrix} d\mu(\eta_n) \cdots d\mu(\eta_1)$$

and thus

$$\sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})\, U\begin{pmatrix} 1, ..., 2n \\ \theta_1, ..., \theta_{2n} \end{pmatrix} > 0.$$

As such $(g-g^*)(\theta_i) \neq 0$ for some $i \in \{1, ..., 2n\}$. Solving for $c_{2n}=1$ in the equations (6.3) we obtain

$$1 = c_{2n} = \frac{\sum_{k=1}^{2n} (g-g^*)(\theta_k)(-1)^k \, U\begin{pmatrix} 1, ..., 2n-1 \\ \theta_1, ..., \hat{\theta}_k, ..., \theta_{2n} \end{pmatrix}}{U\begin{pmatrix} 1, ..., 2n \\ \theta_1, ..., \theta_{2n} \end{pmatrix}}.$$

A calculation similar to that done above shows that

$$\operatorname{sgn} U\begin{pmatrix} 1, ..., 2n-1 \\ \theta_1, ..., \hat{\theta}_k, ..., \theta_{2n} \end{pmatrix} = \varepsilon_n$$

for all $k=1, ..., 2n$. The right-hand side of the above equation therefore has sign -1. This is a contradiction and proves the sufficiency and the uniqueness.

Necessity. From Proposition 6.1, the zero function is a best approximation to $f-g^*$ from

$$\mathscr{A} = \left\{ \sum_{i=1}^{n} b_i K(x, \xi_i) + c_i \frac{\partial K(x, \xi_i)}{\partial y} + dK(x, \eta) : b_i, c_i \in \mathbb{R}, i=1, ..., n, d \geqslant 0 \right\}.$$

This immediately implies that the zero function is a best approximation to $f-g^*$ from the QT-space

$$U_{2n} = \operatorname{span} \left\{ K(\cdot, \xi_1), \frac{\partial K(\cdot, \xi_1)}{\partial y}, ..., K(\cdot, \xi_n), \frac{\partial K(\cdot, \xi_n)}{\partial y} \right\}.$$

Thus either (b1) holds (i.e., at least $2n+2$ points of equioscillation) or we have exactly $2n$ points of equioscillation as in the statement of Theorem 2.2. The proof of the explicit orientation of the sign of the equioscillation as stated in (b2) follows the proof of the analogous result in Theorem 3.2. ∎

Recall that

$$\mathscr{Q}_n = \left\{ \sum_{i=1}^{n} a_i K(x, \xi_i) : a_i \geqslant 0 \right\}.$$

We now prove the analogue of Proposition 5.1 for $\mathscr{Q}_n$.

PROPOSITION 6.4. *Let $f \in \tilde{C}$ and assume that the best approximation to f from $\mathscr{M}_\infty$ is not in $\mathscr{Q}_n$. If g^* is a best approximation to f from $\mathscr{Q}_n$ then*

$$g^*(x) = \sum_{i=1}^{n} a_i K(x, \xi_i)$$

for some $a_i > 0$, $i = 1, ..., n$, *and* $\xi_1 < \cdots < \xi_n < \xi_1 + 2\pi$, *i.e.*, $g^* \in \operatorname{int} \mathcal{Q}_n$, *and there exist* $\theta_1 < \cdots < \theta_{2n} < \theta_1 + 2\pi$ *such that*

$$K\begin{pmatrix} \theta_1, ..., \theta_{2n} \\ \xi_1, \xi_1, ..., \xi_n, \xi_n \end{pmatrix} = 0$$

and for some θ'_{2n}, θ''_{2n} *satisfying* $\theta_{2n-1} < \theta'_{2n} \leqslant \theta_{2n} \leqslant \theta''_{2n} < \theta_1 + 2\pi$ *we have*

$$(-1)^i \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta_i) = \|f - g^*\|, \qquad i = 1, ..., 2n-1$$

$$\varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta'_{2n}) = \varepsilon_n \sigma_2(\boldsymbol{\theta}, \boldsymbol{\xi})(f - g^*)(\theta''_{2n}) = \|f - g^*\|.$$

Proof. The proof is an immediate consequence of Proposition 6.1, and the method of proof in Theorem 6.3. If

$$g^*(x) = \sum_{i=1}^{k} a_i K(x, \xi_i)$$

with $k < n$, then Proposition 6.1 holds since the perturbed $g_\lambda \in \mathcal{Q}_{k+1} \subseteq \mathcal{Q}_n$. We then apply the method of proof of necessity in Theorem 6.3 to prove that g^* is a best approximation to f from $\mathcal{M}_\infty$. From this contradiction we obtain $k = n$.

We now apply the proof of Proposition 6.1 where we set $d = 0$. In this case $g_\lambda \in \mathcal{Q}_n$, so the perturbation is admissible. It follows that the zero function is a best approximation to $f - g^*$ from the $2n$-dimensional QT-space

$$U_{2n} = \operatorname{span}\left\{K(\cdot, \xi_1), \frac{\partial K(\cdot, \xi_1)}{\partial y}, ..., K(\cdot, \xi_n), \frac{\partial K(\cdot, \xi_n)}{\partial y}\right\}.$$

From Theorems 2.2 and 6.3, and since g^* is not a best approximation to f from $\mathcal{M}_\infty$, we see that the desired property must hold. ∎

As in the case of Proposition 5.1 these necessary conditions are not, in general, sufficient. In addition, the best approximation from $\mathcal{Q}_n$ is not necessarily unique. We mention that for $K(x, y) = k(x - y)$ satisfying the "extended" SCVD properties, there exists a non-zero constanct c such that

$$c \sum_{i=1}^{2m} k\left(x - \frac{i\pi}{m}\right)$$

is a best approximation to $f(x) = 1$ from $\mathcal{Q}_{2m}$. But then any translate of this function is also a best approximation, and so there is no uniqueness. Paralleling the analysis in Section 5, one can use this example to construct an $f \in \tilde{C}$ for which the necessary conditions of Proposition 6.4 are not sufficient.

The analogue of Theorem 4.1 for

$$\mathcal{Q}_n(\xi)=\left\{\sum_{i=1}^{n} a_i K(x,\xi_i)\in\mathcal{Q}_n : \xi_1=\xi\right\}.$$

is the following result.

PROPOSITION 6.5. *Assume that the unique best approximation to* $f\in\tilde{C}$ *from* $\mathcal{M}_\infty$ *is not in* $\mathcal{Q}_n(\xi)$. *Then there exists a unique best approximation* g^+ *to* f *from* $\mathcal{Q}_n(\xi)$. g^+ *has the form*

$$g^+(x)=\sum_{i=1}^{n} a_i K(x,\xi_i),$$

where $a_i>0$, $i=1,\ldots,n$, *and* $\xi=\xi_1<\cdots<\xi_n<\xi_1+2\pi$, *i.e.,* $g^+\in\operatorname{int}\mathcal{Q}_n(\xi)$. *It is uniquely characterized by the fact that* $f-g^+$ *equioscillates on* $2n$ *points.*

Proof. Let

$$g^+(x)=\sum_{i=1}^{n} a_i K(x,\xi_i)\in\mathcal{Q}_n(\xi)$$

and assume that $f-g^+$ equioscillates on $2n$ points. Let $g\neq g^+$,

$$g(x)=\sum_{i=1}^{n} b_i K(x,\eta_i)\in\mathcal{Q}_n(\xi).$$

If

$$\|f-g\|\leqslant\|f-g^+\|,$$

then

$$2n\leqslant\tilde{Z}_c((f-g^+)-(f-g))=\tilde{Z}_c(g-g^+).$$

Now

$$(g-g^+)(x)=\sum_{i=1}^{m} c_i K(x,\zeta_i),$$

where $m\leqslant 2n-1$. (Do not double count the ξ_1.) As such $g-g^+$ is contained in a T-space of dimension $2n-1$, but has at least $2n$ zeros, with nonnodal zeros being counted twice. This is a contradiction. Thus g^+ is the unique best approximation to f from $\mathcal{Q}_n(\xi)$.

Now assume that g^+ is a best approximation to f from $\mathcal{Q}_n(\xi)$. To prove that g^+ satisfies the desired conditions, we apply the method used in the proof of Theorem 4.1.

If $g^+ \in \operatorname{int} \mathcal{Q}_n(\xi)$, then the zero function is necessarily a best approximation to $f-g^+$ from

$$\operatorname{span}\left\{K(\cdot, \xi_1), K(\cdot, \xi_2), \frac{\partial K(\cdot, \xi_2)}{\partial y}, ..., K(\cdot, \xi_n), \frac{\partial K(\cdot, \xi_n)}{\partial y}\right\},$$

which is a $(2n-1)$-dimensional T-space. Hence $f-g^+$ exhibits $2n$ points of equioscillation.

Assume that $g^+ \notin \operatorname{int} \mathcal{Q}_n(\xi)$. Thus

$$g^+(x) = \sum_{i=1}^{k} a_i K(x, \eta_i),$$

where $a_i > 0$, $i = 1, ..., k$, $\eta_1 < \cdots < \eta_k < \eta_1 + 2\pi$, and $k < n$. We consider three cases, paralleling the first three cases of Proposition 4.3.

Case 1. $\xi \notin \{\eta_1, ..., \eta_k\}$.

It can be shown exactly as in the proof of Proposition 6.1 that the zero function is a best approximation to $f-g^+$ from

$$\mathcal{A} = \left\{\sum_{i=1}^{k} b_i K(x, \eta_i) + c_i \frac{\partial K(x, \eta_i)}{\partial y} + dK(x, \xi) : b_i, c_i \in \mathbb{R}, i = 1, ..., k, d \geqslant 0\right\}.$$

The argument found in the proof of Theorem 6.3 shows that g^+ is a best approximation to f from $\mathcal{M}_\infty$, which is a contradiction.

Case 2. $\xi \in \{\eta_1, ..., \eta_k\}$, $k \leqslant n-2$.

Suppose that $\xi = \eta_1$. We first claim that the zero function is a best approximation to $f-g^+$ from

$$\mathcal{A} = \left\{\sum_{i=1}^{k} b_i K(x, \eta_i) + c_i \frac{\partial K(x, \eta_i)}{\partial y} + dK(x, \zeta) : b_i, c_i \in \mathbb{R}, i = 1, ..., k, d \geqslant 0\right\},$$

for any $\zeta \notin \{\eta_1, ..., \eta_k\}$. This fact can be shown in the same way as in the proof of Proposition 6.1.

Again, the same analysis as in the proof of Theorem 6.3 shows that g^+ is a best approximation to f from $\mathcal{M}_\infty$, which is a contradiction.

Case 3. $\xi \in \{\eta_1, ..., \eta_k\}$, $k = n-1$.

We shall not give the details of this case, as it is lengthy. It entirely parallels Case 3 of Proposition 4.3, and the above ideas. ▮

We end this paper with a final result concerning the g^+.

Proposition 6.6. *Assume that the best approximation to $f \in \tilde{C}$ from $\mathcal{M}_\infty$ is not in $\mathcal{Q}_{n+1}(\xi)$. Let g_n^+ and g_{n+1}^+ denote the unique best approximations to f from $\mathcal{Q}_n(\xi)$ and $\mathcal{Q}_{n+1}(\xi)$, respectively. If g_n^+ has the form*

$$g_n^+(x) = \sum_{i=1}^{n} a_i K(x, \xi_i),$$

where $a_i > 0$, $i = 1, ..., n$, and $\xi = \xi_1 < \cdots < \xi_n < \xi_1 + 2\pi$, while g_{n+1}^+ has the form

$$g_{n+1}^+(x) = \sum_{i=1}^{n+1} b_i K(x, \eta_i),$$

where $b_i > 0$, $i = 1, ..., n+1$, and $\xi = \eta_1 < \cdots < \eta_{n+1} < \eta_1 + 2\pi$, then

$$\eta_i < \xi_i < \eta_{i+1}, \qquad i = 2, ..., n.$$

Proof. We have

$$2n \leqslant \tilde{Z}_c((f - g_n^+) - (f - g_{n+1}^+)) = \tilde{Z}_c(g_{n+1}^+ - g_n^+)$$
$$= \tilde{Z}_c\left(\sum_{i=1}^{n+1} b_i K(x, \eta_i) - \sum_{i=1}^{n} a_i K(x, \xi_i)\right).$$

Note that $\eta_1 = \xi_1 = \xi$. We have a non-trivial linear combination of $2n$ functions, which form a QT-system, and which vanish at $2n$ distinct points (since $\|f - g_n^+\| > \|f - g_{n+1}^+\|$). The coefficients are uniquely determined, up to multiplication by a constant (since their span contains a T-space of dimension $2n - 1$).

Assume

$$\sum_{i=1}^{2n} c_i K(\theta_j, \zeta_i) = 0, \qquad j = 1, ..., 2n,$$

for some $\theta_1 < \cdots \theta_{2n} < \theta_1 + 2\pi$ and $\zeta_1 < \cdots \zeta_{2n} < \zeta_1 + 2\pi$ (and the c_i not all zero). The c_i are proportional to

$$(-1)^i K\begin{pmatrix} \theta_1, ..., \hat{\theta}_i, ..., \theta_{2n} \\ \zeta_1, ..., \zeta_{2n-1} \end{pmatrix}, \qquad i = 1, ..., 2n.$$

As such the c_i alternate in sign. Moreover, the $\{a_i\}$ and $\{b_i\}$ are all positive. Thus we must have $a_1 > b_1$, and

$$\eta_i < \xi_i < \eta_{i+1}, \qquad i = 2, ..., n. \quad \blacksquare$$

ACKNOWLEDGMENTS

The first author was supported in part by Grant U92000 of the International Science Foundation. He also thanks the Department of Mathematics at the Technion for its kind hospitality and support during his visit in the spring of 1995.

REFERENCES

1. O. V. Davydov, Notes on alternation theorems for periodic functions, *in* "Theory of Approximation of Functions and Summation of Series" (V. P. Motornyi, Ed.), pp. 21–25, Dnepropetrovsk Gos. Univ., Dnepropetrovsk, 1989. [in Russian]
2. O. V. Davydov, Characterization of the best uniform approximation of periodic functions by convex classes defined by strictly CVD kernels, *in* "Approximation Theory VIII," Vol. 1, "Approximation and Interpolation" (C. K. Chui and L. L. Schumaker, Eds.), pp. 177–184, World Scientific, Singapore, 1995.
3. K. Glashoff, Restricted approximation by strongly sign-regular kernels: The finite bang-bang principle, *J. Approx. Theory* **29** (1980), 212–217.
4. S. Karlin, "Total Positivity," Stanford Univ. Press, Stanford, CA, 1968.
5. J. C. Mairhuber, I. J. Schoenberg, and R. E. Williamson, On variation diminishing transformations of the circle, *Rend. Circ. Mat. Palermo* **8** (1959), 241–270.
6. A. Pinkus, Best approximations by smooth functions, *J. Approx. Theory* **33** (1981), 147–178.
7. A. Pinkus, "*n*-Widths in Approximation Theory," Springer-Verlag, Berlin, 1985.
8. G. Pólya and I. J. Schoenberg, Remarks on de la Vallée Poussin means and convex conformal maps of the circle, *Pacific J. Math.* **8** (1958), 295–334.
9. T. J. Rivlin, "The Chebyshev Polynomials," Wiley–Interscience, New York, 1974.
10. U. Sattes, Best Chebyshev approximation by smooth functions, *in* "Quantitative Approximation" (R. A. DeVore and K. Scherer, Eds.), pp. 279–289, Academic Press, New York, 1980.