JOURNAL OF APPROXIMATION THEORY 9, 173–191 (1973)

Characterization of Best Approximations by Sums of Exponentials*

DAVID W. KAMMLER

Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901

Communicated by E. W. Cheney

Received September 21, 1971

An exponential sum y can be specified by giving the coefficients $\mathbf{b}$, $\mathbf{c}$ of the corresponding initial value problem $(D^n + c_1D^{n-1} + \cdots + c_n)y = 0$, $D^{j-1}y(0) = b_j$, $j = 1, 2,..., n$. We discuss some of the topological properties of this parametric form, noting that the associated tangential manifold does not suffer a "loss of dimension" when the exponential parameters are allowed to coalesce. Using this representation, we formulate a first order necessary condition which may often be used to characterize a local best L_p-approximation to a given $f \in L_p[0, 1]$, $1 \leqslant p \leqslant \infty$, from the set $V_n(S)$. A sufficient condition is also given, and the use and limitations of the theorems are illustrated by means of several carefully chosen examples.

1. INTRODUCTION

Given $\mathbf{b} = (b_1, b_2, ..., b_n)$ and $\mathbf{c} = (c_1, c_2, ..., c_n)$ from $\mathbf{C}^n$, we define $Y_n(\mathbf{b}, \mathbf{c})$ to be the solution of the initial value problem

$$(D^n + c_1D^{n-1} + \cdots + c_{n-1}D + c_n)\, y(t) = 0, \qquad 0 \leqslant t \leqslant 1, \tag{1}$$

$$D^{j-1}y(0) = b_j, \qquad j = 1, 2,..., n. \tag{2}$$

When y satisfies (1) but does not satisfy any such equation of lower order, we say that y is an exponential sum with order n. We shall let $P_n(\mathbf{c})$ denote the polynomial corresponding to the differential operator in (1) and define

$$\Lambda_n(\mathbf{c}) = \{\lambda \in \mathbf{C}: \lambda \text{ is a root of } P_n(\mathbf{c})\}. \tag{3}$$

* This work was supported in part by the National Science Foundation through the Graduate Fellowship Program and represents a portion of a Ph.D. dissertation directed by Professor Cleve B. Moler and submitted to the Mathematics Department of the University of Michigan.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.

In [6] we have shown that each $f \in L_p[0, 1]$ $(1 \leqslant p \leqslant \infty)$ has a best $\| \|_p$-approximation from the set

$$V_n(S) = \{Y_n(\mathbf{b}, \mathbf{c}): \mathbf{b}, \mathbf{c} \in \mathbf{C}^n, \Lambda_n(\mathbf{c}) \subseteq S\}, \qquad n = 1, 2,...$$

provided that $S \subseteq \mathbf{C}$ is closed. In this paper we shall discuss the parametrization of $V_n(\mathbf{C})$ suggested by (1) and (2) and then formulate necessary conditions and sufficient conditions which may often be used to recognize such a best approximation. By specializing these results we shall obtain the previously known characterization theorems of Rice [8] and Braess [2].

2. The Parametrization $Y_n(\mathbf{b}, \mathbf{c})$

In the formulation of characterization theorems or in the actual construction of best approximations to a given f it is useful to have some means to parametrically represent each $y \in V_n(\mathbf{C})$. For example, we might write y in the form

$$y(t) = \sum_{i=1}^{l} \sum_{j=1}^{n_i} a_{ij} t^{j-1} \exp(\lambda_i^0 t), \qquad n_1 + n_2 + \cdots + n_l = k \leqslant n \tag{4}$$

as is done by Rice [8], in the form

$$y(t) = \sum_{j=1}^{n} d_j \Delta^{j-1}(\lambda_1, \lambda_2, ..., \lambda_j) \exp(\lambda t) \tag{5}$$

(where $\Delta^{j-1}(\lambda_1, \lambda_2, ..., \lambda_j)$ is the $(j-1)$th order divided difference operator with respect to λ at the points $\lambda_1, \lambda_2, ..., \lambda_j$) as is done by Braess [3], in the form

$$y(t) = \mathscr{Y}_n(\mathbf{b}, \boldsymbol{\lambda}; t),$$

where $\mathscr{Y}_n(\mathbf{b}, \boldsymbol{\lambda})$ is the solution of the initial value problem

$$\begin{gathered} \{(D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_n)\}\, y(t) = 0 \\ D^{j-1} y(0) = b_j, \qquad j = 1, 2, ..., n \end{gathered} \tag{6}$$

as is done in [6], or in the form

$$y(t) = Y_n(\mathbf{b}, \mathbf{c}; t),$$

where $Y_n(\mathbf{b}, \mathbf{c})$ is defined by means of the initial value problem of (1) and (2). The conventional form (4) is convenient when the parameters λ_i^0 are held fixed, but it is not at all convenient when these parameters are allowed to coalesce in which case l (and thus the whole form of (4)) is altered. The remaining three forms provide continuous mappings from the euclidean space $\mathbf{C}^{2n}$ onto the $\| \ \|_p$- normed space $V_n(\mathbf{C})$ with $Y_n(\mathbf{b}, \mathbf{c})$ also providing a mapping from $\mathbf{R}^{2n}$ onto the corresponding set $V_n^r(\mathbf{C})$ of real valued exponential sums. When one constrains the exponential parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$ (e.g., by requiring them to be real) the use of either (5) or (6) may be desirable with the form (5) allowing one to cast the problem in a form suitable for the theory of γ-polynomial approximation, cf. Hobby and Rice [5] and de Boor [1]. It will be noted that the choice of the linear parameters $\mathbf{d}$ in (5) is dependent upon the way the corresponding exponential parameters λ_i are ordered into a vector $\boldsymbol{\lambda}$ while this is no longer the case for the linear parameters $\mathbf{b}$ appearing in (6). Moreover, the parametetric form $Y_n(\mathbf{b}, \mathbf{c})$ eliminates this ordering problem altogether by replacing the exponential parameters by the components of $\mathbf{c}$ which are the first n elementary symmetric functions of the λ_i. Hence, $Y_n(\mathbf{b}, \mathbf{c})$ is a natural parametric form to use when some of the λ_i are complex so that no natural ordering is possible.

Given $y \in V_n(\mathbf{C})$ we may uniquely determine $\mathbf{b}$ using (2), but the differential operator of (1) which annihilates y (and, therefore, $\mathbf{c}$) is uniquely determined if and only if y has full order n. Thus, the parametrization $Y_n(\mathbf{b}, \mathbf{c})$ fails to provide a homeomorphism between $\mathbf{C}^{2n}$ and $V_n(\mathbf{C})$. By restricting our consideration to those exponential sums which have full order, however, we obtain the following result.

THEOREM 1. *Let the mapping*

$$Y_n : \mathbf{C}^{2n} \to V_n(\mathbf{C}) \qquad \text{by} \qquad (\mathbf{b}, \mathbf{c}) \to Y_n(\mathbf{b}, \mathbf{c}) \tag{7}$$

be defined by means of the initial value problem (1) *and* (2), *and let*

$$\mathbf{F}_{2n} = Y_n^{-1}\{V_{n-1}(\mathbf{C})\}$$

(*with $V_0(\mathbf{C})$ denoting the set whose only element is the zero function in the event $n = 1$.*) *Then $\mathbf{F}_{2n}$ is a nowhere dense closed subset of $\mathbf{C}^{2n}$, and the restriction*

$$Y_n : \mathbf{C}^{2n}\backslash\mathbf{F}_{2n} \to V_n(\mathbf{C})\backslash V_{n-1}(\mathbf{C}) \tag{8}$$

of (7) *to $\mathbf{C}^{2n}\backslash\mathbf{F}_{2n}$ is a continuously differentiable homeomorphism provided $\mathbf{C}^{2n}$ has the usual topology and $V_n(\mathbf{C})$ has the topology induced by any one of the $\| \ \|_p$-norms, $1 \leqslant p \leqslant \infty$.*

Proof. From the definition of (7) in terms of the initial value problem (1) and (2) it is clear that (7) is continuously differentiable and that the restriction (8) is a bijection. Next, for any given $y \in V_n(\mathbf{C})\backslash V_{n-1}(\mathbf{C})$ there exists a best $\| \ \|_p$-approximation, y_0, from $V_{n-1}(\mathbf{C})$ (cf. [6, Theorem 2]) so that $V_n(\mathbf{C})\backslash V_{n-1}(\mathbf{C})$ contains the open ball of radius $\| y - y_0 \|$ which is centered at y. It follows that $V_{n-1}(\mathbf{C})$ is a closed subset of $V_n(\mathbf{C})$ and that $\mathbf{F}_{2n}$ (the inverse image of $V_{n-1}(\mathbf{C})$ under the continuous mapping (7)) is a closed subset of $\mathbf{C}^{2n}$. Moreover, the closed set $\mathbf{F}_{2n}$ can contain no open ball in $\mathbf{C}^{2n}$ (since by slightly perturbing the initial conditions $\mathbf{b}$ one can insure that $Y_n(\mathbf{b}, \mathbf{c})$ has full order n) and, therefore, $\mathbf{F}_{2n}$ is nowhere dense. It remains to be shown that the inverse of (8) is continuous.

Accordingly, let $Y_n(\mathbf{b}, \mathbf{c})$, $Y_n(\mathbf{b}_\nu, \mathbf{c}_\nu)$, $\nu = 1, 2,\ldots$ be chosen in such a manner that each has full order n and that

$$\lim \|Y_n(\mathbf{b}, \mathbf{c}) - Y_n(\mathbf{b}_\nu, \mathbf{c}_\nu)\|_p = 0. \tag{9}$$

We must show that this implies that $\{(\mathbf{b}_\nu, \mathbf{c}_\nu)\}$ converges to $(\mathbf{b}, \mathbf{c})$ in $\mathbf{C}^{2n}$.

As a first step, we shall show that $\{\mathbf{c}_\nu\}$ is bounded. Indeed, if this is not the case, then the corresponding sequence of spectral sets $\{\Lambda_n(\mathbf{c}_\nu)\}$ is unbounded so that (after passing to a subsequence, if necessary) we may obtain the decomposition

$$Y_n(\mathbf{b}_\nu, \mathbf{c}_\nu) = u_\nu + v_\nu + w_\nu, \qquad \nu = 1, 2,\ldots,$$

where $\{v_\nu\}$ is a sequence from $V_{n-1}(\mathbf{C})$ which $\| \ \|_p$- converges to some $v \in V_{n-1}(\mathbf{C})$ and where $\{u_\nu\}$, $\{w_\nu\}$ are U, W-sequences, respectively, from $V_n(\mathbf{C})$, cf. [6, Section 3]. Together with (9) and [6, Lemma 2-iv] this gives

$$\begin{aligned} 0 &= \lim \| u_\nu + v_\nu + w_\nu - Y_n(\mathbf{b}, \mathbf{c})\|_p \\ &= \lim \| u_\nu + v + w_\nu - Y_n(\mathbf{b}, \mathbf{c})\|_p \\ &\geqslant \| v - Y_n(\mathbf{b}, \mathbf{c})\|_p, \end{aligned}$$

which implies that $Y_n(\mathbf{b}, \mathbf{c}) \in V_{n-1}(\mathbf{C})$ contrary to hypothesis. Thus, $\{\mathbf{c}_\nu\}$ is bounded.

Now since $\{\mathbf{c}_\nu\}$ is bounded, the corresponding sequence of spectral sets is also bounded, and together with (9) and [6, Lemma 1] this implies that $\{\mathbf{b}_\nu\}$ is also bounded. Hence, some subsequence of $\{(\mathbf{b}_\nu, \mathbf{c}_\nu)\}$ is convergent. Let $(\mathbf{b}^*, \mathbf{c}^*)$ denote the limit of this subsequence (which we shall continue to call $\{(\mathbf{b}_\nu, \mathbf{c}_\nu)\}$.) By using (9) and the continuity of (7) we find

$$\| Y_n(\mathbf{b}^*, \mathbf{c}^*) - Y_n(\mathbf{b}, \mathbf{c})\|_p = \lim \| Y_n(\mathbf{b}_\nu, \mathbf{c}_\nu) - Y_n(\mathbf{b}, \mathbf{c})\|_p = 0$$

and since (8) is a bijection this implies that $(\mathbf{b}^*, \mathbf{c}^*) = (\mathbf{b}, \mathbf{c})$. $\square$

COROLLARY 1. *The further restriction*

$$Y_n : \mathbf{R}^{2n}\backslash\mathbf{F}_{2n} \to V_n^r(\mathbf{C})\backslash V_{n-1}^r(\mathbf{C}) \tag{10}$$

of (8) *is also a continuously differentiable homeomorphism.*

Unfortunately, there is no "ideal" parametrization for the class of exponential sums, i.e., there is no homeomorphism from $\mathbf{C}^{2n}$ onto $V_n(\mathbf{C})$. Indeed, since $\mathbf{C}^{2n}$ is locally compact while $V_n(\mathbf{C})$ is not, no such homeomorphism can exist [4, p. 238–239]. (To see that $V_n(\mathbf{C})$ with the $\| \ \|_p$- norm is not locally compact, we note that any $\| \ \|_p$- neighborhood of the zero function in $V_n(\mathbf{C})$ contains *a* U-sequence which contains no convergent subsequence.) Nevertheless, as $\mathbf{b}, \mathbf{c}$ vary over suitably small open balls with centers at $\mathbf{b}_0$, $\mathbf{c}_0$, respectively, in $\mathbf{C}^n$ we see from Theorem 1 that $Y_n(\mathbf{b}, \mathbf{c})$ varies over some neighborhood of $y_0 = Y_n(\mathbf{b}_0, \mathbf{c}_0)$ in $V_n(\mathbf{C})$ provided y_0 has full order n. When y_0 does not have full order, however, we cannot explore a neighborhood of y_0 in $V_n(\mathbf{C})$ so simply, and, indeed, we do not have any means of reasonably describing such a neighborhood in terms of any presently known parametric form. (Similar observations have been made by Rice [8, p. 152] and de Boor [1, p. 182].)

Given $\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^* \in \mathbf{C}^n$ we define

$$h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*) = \sum_{i=1}^{n} \left\{ b_i^* \cdot \frac{\partial}{\partial b_i} + c_i^* \cdot \frac{\partial}{\partial c_i} \right\} Y_n(\mathbf{b}, \mathbf{c}), \tag{11}$$

and we refer to the set

$$H_n(\mathbf{b}, \mathbf{c}) = \{h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*) \colon \mathbf{b}^*, \mathbf{c}^*, \in \mathbf{C}^n\} \tag{12}$$

as the tangential manifold of $y = Y_n(\mathbf{b}, \mathbf{c})$ with respect to the parameters $\mathbf{b}$, $\mathbf{c}$. Clearly $H_n(\mathbf{b}, \mathbf{c})$ is a linear space with dimension at most $2n$ which contains the n-dimensional space

$$L_n(\mathbf{c}) = \{h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, 0) \colon \mathbf{b}^* \in \mathbf{C}^n\} \tag{13}$$

of solutions of (1). A useful alternative description of $H_n(\mathbf{b}, \mathbf{c})$ is provided by the following lemma.

LEMMA 1. *Let* $\mathbf{b}, \mathbf{c} \in \mathbf{C}^n$ *and assume that the exponential sum* $y = Y_n(\mathbf{b}, \mathbf{c})$ *has order* k. *Let* $\lambda_1, \lambda_2, \ldots, \lambda_k$ *be the essential exponential parameters of* y *so that*

$$Q_n(\mathbf{b}, \mathbf{c}; D) = \begin{cases} 1 & \text{if } k = 0, \\ (D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_k) & \text{if } k \geqslant 1 \end{cases}$$

is the lowest order monic differential operator which annihilates y. *Then the linear space* $H_n(\mathbf{b}, \mathbf{c})$ *given by* (12) *has dimension* $n + k$ *with*

$$H_n(\mathbf{b}, \mathbf{c}) = \{h \colon Q_n(\mathbf{b}, \mathbf{c}; D)\, P_n(\mathbf{c}; D)\, h \equiv 0\} \tag{14}$$

(*where* $P_n(\mathbf{c}; D)$ *is again the differential operator of* (1).)

Proof. If $y = Y_n(\mathbf{b}, \mathbf{c})$ and $h = h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*)$ then by using (1), (2), and (11) we find that h satisfies the initial value problem

$$P_n(\mathbf{c}; D)\, h(t) = -\sum_{i=1}^{n} c_i{}^* D^{n-i} y(t), \qquad 0 \leqslant t \leqslant 1, \tag{15}$$

$$D^{j-1}h(0) = b_j{}^*, \qquad j = 1, 2,..., n \tag{16}$$

and, hence, vanishes identically if and only if the driving term on the right side of (15) and the initial conditions of (16) both vanish. But by hypothesis y has order k and, therefore, the functions $D^{j-1}y$, $j = 1, 2,..., n$ span a linear space of dimension k. Hence, from (15) and (16) we conclude that $H_n(\mathbf{b}, \mathbf{c})$ has dimension $n + k$. Finally, from (15) and the definition of Q_n, we find that $H_n(\mathbf{b}, \mathbf{c})$ is contained within the $n + k$ dimensional kernel of the differential operator $Q_n(\mathbf{b}, \mathbf{c}; D)\, P_n(\mathbf{c}, D)$ so that (14) holds. □

The corresponding tangential manifolds which result when $V_n(\mathbf{C})$ is parametrized by using the roots of $P_n(\mathbf{c})$, e.g. (4), (5), or (6), suffer a "loss of dimension" (cf. [1, p. 182]) when some of these roots coalesce. Indeed, if we let $\mathscr{Y}_n(\mathbf{b}, \boldsymbol{\lambda})$ denote the solution of (6) and if we define

$$g_n(\mathbf{b}, \boldsymbol{\lambda}, \mathbf{b}^*, \boldsymbol{\lambda}^*) = \sum_{i=1}^{n} \left\{ b_i{}^* \cdot \frac{\partial}{\partial b_i} + \lambda_i{}^* \cdot \frac{\partial}{\partial \lambda_i} \right\} \mathscr{Y}_n(\mathbf{b}, \boldsymbol{\lambda}), \tag{17}$$

$$G_n(\mathbf{b}, \boldsymbol{\lambda}) = \{ g_n(\mathbf{b}, \boldsymbol{\lambda}, \mathbf{b}^*, \boldsymbol{\lambda}^*) \colon \mathbf{b}^*, \boldsymbol{\lambda}^* \in \mathbf{C}^n \}$$

then by using arguments analogous to those used in the proof of Lemma 1 we obtain the following alternative description of the linear space $G_n(\mathbf{b}, \boldsymbol{\lambda})$.

LEMMA 2. *Let* $\mathbf{b}, \boldsymbol{\lambda} \in \mathbf{C}^n$, *and assume that the exponential sum* $y = \mathscr{Y}_n(\mathbf{b}, \boldsymbol{\lambda})$ *can be expressed in the form* (4) *with* $a_{in_i} \neq 0$ *for each* $i = 1, 2,..., l$ *and with* $\lambda_1{}^0, \lambda_2{}^0,..., \lambda_l{}^0$ *being distinct numbers selected from among the components of* $\boldsymbol{\lambda}$. *Let*

$$R_n(\mathbf{b}, \boldsymbol{\lambda}; D) = \begin{cases} 1 & \text{if } l = 0, \\ (D - \lambda_1{}^0)(D - \lambda_2{}^0) \cdots (D - \lambda_l{}^0) & \text{if } l \geqslant 1, \end{cases}$$

and let $\sigma_n(\boldsymbol{\lambda}) \in \mathbf{C}^n$ *be defined in such a manner that the components of* $\boldsymbol{\lambda}$ *are the roots of* $P_n(\boldsymbol{\sigma}_n(\boldsymbol{\lambda}))$. *Then the linear space* $G_n(\mathbf{b}, \boldsymbol{\lambda})$ *given by* (17) *has dimension* $n + l$ *with*

$$G_n(\mathbf{b}, \boldsymbol{\lambda}) = \{ g \colon R_n(\mathbf{b}, \boldsymbol{\lambda}; D)\, P_n(\boldsymbol{\sigma}_n(\boldsymbol{\lambda}); D)\, g \equiv 0 \}. \quad \square$$

Using Lemmas 1 and 2 we see that the exponential sum

$$y(t) = t \exp(t) \tag{18}$$

from $V_2(\mathbf{C})$ has the four-dimensional tangential manifold

$$H_2(\mathbf{b}_0, \mathbf{c}_0) = \{h: (D-1)^4 h \equiv 0\} \qquad \mathbf{b}_0 = (0, 1), \qquad \mathbf{c}_0 = (-2, 1)$$

with respect to the parametrization of (1) and (2) while the corresponding tangential manifold

$$G_2(\mathbf{b}_0, \boldsymbol{\lambda}_0) = \{g: (D-1)^3 g \equiv 0\}, \qquad \mathbf{b}_0 = (0, 1), \qquad \lambda_0 = (1, 1),$$

which results from any of the parametric forms (4), (5), or (6) is only three-dimensional.

The use of the full dimensionality of $H_2(\mathbf{b}_0, \mathbf{c}_0)$ may be restricted due to some constraint we wish to impose upon the exponential parameters. For example, suppose we wish to consider exponential sums which lie close to the y of (18) and which continue to have real exponential parameters, λ_1, λ_2. We, thus, wish to find those $\mathbf{b}^*, \mathbf{c}^* \in \mathbf{C}^2$ for which $Y_2(\mathbf{b} + \alpha\mathbf{b}^*, \mathbf{c} + \alpha\mathbf{c}^*)$ lies in $V_n(\mathbf{R})$ for all sufficiently small $\alpha > 0$, or, equivalently, we wish to find $\mathbf{c}^*$ such that the roots of the corresponding auxiliary polynomial

$$P_2(\mathbf{c} + \alpha\mathbf{c}^*; \lambda) = \lambda^2 + (\alpha c_1{}^* - 2)\lambda + (\alpha c_2{}^* + 1)$$

are both real for all small $\alpha > 0$. For this to be the case it is both necessary and sufficient that $c_1{}^*, c_2{}^* \in \mathbf{R}$ with $c_1{}^* + c_2{}^* \leqslant 0$. Thus, by constraining the exponential parameters to lie in the set $S = \mathbf{R}$ we prohibit the full use of the four-dimensional manifold $H_2(\mathbf{b}_0, \mathbf{c}_0)$, but we certainly have more freedom than would be the case if we only made use of the three-dimensional manifold $G_2(\mathbf{b}_0, \boldsymbol{\lambda}_0)$ which results from the alternative parametrization using (4), (5), or (6).

In an effort to deal with situations of this type we shall find it convenient to introduce two new concepts. We define

$$K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*) = \{h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \alpha\mathbf{c}^*): \mathbf{b}^* \in \mathbf{C}^n, \alpha \geqslant 0\} \tag{19}$$

and refer to this subset of $H_n(\mathbf{b}, \mathbf{c})$ as the tangential cone of $Y_n(\mathbf{b}, \mathbf{c})$ at $\mathbf{b}, \mathbf{c}$ in the direction of $\mathbf{c}^*$. We note that

$$\begin{aligned} L_n(\mathbf{c}) &= K_n(\mathbf{b}, \mathbf{c}, \mathbf{0}), \\ H_n(\mathbf{b}, \mathbf{c}) &= \bigcup_{\mathbf{c}^* \in \mathbf{C}^n} K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*), \\ G_n(\mathbf{b}, \mathbf{c}) &= \bigcup_{\boldsymbol{\lambda}^* \in \mathbf{C}^n} K_n(\mathbf{b}, \boldsymbol{\sigma}_n(\boldsymbol{\lambda}), x_n(\boldsymbol{\lambda}, \boldsymbol{\lambda}^*)), \end{aligned} \tag{20}$$

where $\boldsymbol{\sigma}_n$ is defined as in Lemma 2 and

$$x_n(\boldsymbol{\lambda}, \boldsymbol{\lambda}^*) = (d/d\alpha)\, \sigma_n(\boldsymbol{\lambda} + \alpha\boldsymbol{\lambda}^*)|_{\alpha=0}\,.$$

Next, we say that the exponential sum y is accessible through the cone $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ (with respect to $V_n(S)$) provided $y = Y_n(\mathbf{b}, \mathbf{c})$ and there exists some continuously differentiable arc

$$z\colon [0, 1] \to \mathbf{C}^n$$

with

$$\mathbf{z}(\alpha) = \mathbf{c} + \alpha \mathbf{c}^* + o(\alpha) \tag{21}$$

as $\alpha \to 0+$ and with

$$\Lambda_n\,(\mathbf{z}(\alpha)) \subseteq S, \qquad 0 \leqslant \alpha \leqslant 1. \tag{22}$$

For example, when $\Lambda_n(\mathbf{c})$ is contained in the interior of S then $Y_n(\mathbf{b}, \mathbf{c})$ is accessible through each of the cones $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$, $\mathbf{c}^* \in \mathbf{C}^n$. On the other hand, when $S = \mathbf{R}$ we see from the previous discussion that the exponential sum (18) is accessible through the cone $K_2(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ if and only if $\mathbf{b} = (0, 1)$, $\mathbf{c} = (-2, 1)$, and $\mathbf{c}^* \in \mathbf{R}^2$ with $c_1^* + c_2^* \leqslant 0$.

3. First Order Necessary Conditions

Using the notation and concepts just introduced we shall formulate a first order necessary condition which every local best $\|\ \|_p$- approximation to a given f must satisfy. In so doing we first prepare two lemmas.

Lemma 3. *Let $1 \leqslant p \leqslant \infty$ and let $\epsilon, h \in L_p[0, 1]$. Let $\Phi_p[\epsilon, h]$ be defined such that*

$$\Phi_p[\epsilon, h] = \begin{cases} \int_0^1 T[\epsilon, h; t]\, dt & \text{if } p = 1, \\ \|\epsilon\|_p^{1-p} \int_0^1 |\epsilon(t)|^{p-1}\, T[\epsilon, h; t]\, dt & \text{if } 1 < p < \infty, \\ \lim\limits_{\delta\to 0+} \operatorname{ess\,sup}\{T[\epsilon, h; t] \colon |\epsilon(t)| \geqslant \|\epsilon\|_\infty - \delta\} & \text{if } p = \infty, \end{cases} \tag{23}$$

when $\|\epsilon\|_p > 0$ (with $\Phi_p[0, h] = \|h\|_p$) where

$$T[\epsilon, h; t] = \begin{cases} |h(t)| & \text{if } \epsilon(t) = 0, \\ \operatorname{Re}[h(t)\,\overline{\epsilon(t)}/|\epsilon(t)|] & \text{if } \epsilon(t) \neq 0 \end{cases} \tag{24}$$

(with the bar denoting the complex conjugate). Then

$$\|\epsilon + \alpha h\|_p = \|\epsilon\|_p + \alpha \Phi_p[\epsilon, h] + o(\alpha) \tag{25}$$

as α decreases to zero through positive values.

Proof. We may assume $\| \epsilon \|_p , \| h \|_p > 0$. When $1 \leqslant p < \infty$ the Lebesgue dominated convergence theorem [11, p. 89] may be used to show that

$$\lim_{\alpha \to 0+} \alpha^{-1}\{\| \epsilon + \alpha h \|_p^p - \| \epsilon \|_p^p\} = p \| \epsilon \|_p^{p-1} \Phi_p[\epsilon, h]$$

from which (25) follows. When $p = \infty$ we define

$$E_\delta = \{t \in [0, 1]: | \epsilon(t)| \geqslant \| \epsilon \|_\infty - \delta\}, \qquad \delta > 0,$$
$$r(\delta) = \operatorname{ess}_{E_\delta} \sup \operatorname{Re}[\bar{\epsilon} h],$$

noting that $r(\delta)$ monotonically decreases to some limit, $r(0+)$, as $\delta \to 0+$. Setting $\beta = 2\alpha \| h \|_\infty$ we find

$$\begin{aligned}
\| \epsilon \|_\infty^2 + 2\alpha r(0+) - \alpha^2 \| h \|_\infty^2 &= \lim_{\delta \to 0+} \operatorname{ess}_{E_\delta} \sup\{| \epsilon |^2 + 2\alpha \operatorname{Re}[\bar{\epsilon} h] - \alpha^2 \| h \|_\infty^2\} \\
&\leqslant \lim_{\delta \to 0+} \operatorname{ess}_{E_\delta} \sup | \epsilon + \alpha h |^2 \\
&\leqslant \| \epsilon + \alpha h \|_\infty^2 \\
&= \operatorname{ess}_{E_\beta} \sup | \epsilon + \alpha h |^2 \\
&\leqslant \| \epsilon \|_\infty^2 + 2\alpha r(\beta) + \alpha^2 \| h \|_\infty^2
\end{aligned}$$

so that

$$\| \epsilon + \alpha h \|_\infty^2 = \| \epsilon \|_\infty^2 + 2\alpha r(0+) + o(\alpha)$$

from which (25) follows upon taking square roots. □

LEMMA 4. *Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be real numbers, let $\varphi_0 \in C^n(\alpha, \beta)$ be real valued, let*

$$\varphi_1 = (D - \lambda_1) \varphi_0, \qquad \varphi_2 = (D - \lambda_2) \varphi_1, \ldots, \varphi_n = (D - \lambda_n) \varphi_{n-1},$$

and assume that φ_0 has $m > n$ zeros in (α, β). Then φ_ν has at least $m - \nu$ real zeros in (α, β), $\nu = 1, 2, \ldots, n$. In particular, if $y \in V_n(\mathbf{R})$ has n real zeros, then $y \equiv 0$.

Proof. cf. Pólya–Szegö [7, p. 40, #18].

THEOREM 2. *Let the exponential sum $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ be accessible through the tangential cone $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ with respect to $V_n(S)$, let $f \in L_p[0, 1]$, and let $\epsilon = f - y_0$. Then a necessary condition for y_0 to be a local best $\| \ \|_p$-approximation to f from $V_n(S)$ is that y_0 be a best $\| \ \|_p$-approximation to f from $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$, and, in particular, that*

$$\Phi_p[\epsilon, -h] \geqslant 0 \tag{26}$$

for each $h \in K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$.

Proof. Let $h \in K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ be chosen. From the definitions (11) and (19) it follows that for some $\mathbf{b}^* \in \mathbf{C}^n$, $\alpha_0 \geqslant 0$ we have

$$\begin{aligned} h &= h_n(\mathbf{b}, \mathbf{c}, \mathbf{b} + \mathbf{b}^*, \alpha_0\mathbf{c}^*) \\ &= h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \alpha_0\mathbf{c}^*) + h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}, \mathbf{0}) \\ &= h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \alpha_0\mathbf{c}^*) + y_0 \end{aligned}$$

and that (after rescaling $\mathbf{c}^*$, if necessary) we may assume $\alpha_0 = 1$. Moreover, since the local best $\| \ \|_p$-approximation y_0 is accessible through $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ we may find a smooth are $\mathbf{z}$ satisfying (21) and (22) so that for all sufficiently small $\alpha \geqslant 0$ we have

$$\begin{aligned} \| f - y_0 \|_p &\leqslant \| f - Y_n(\mathbf{b} + \alpha\mathbf{b}^*, \mathbf{z}(\alpha))\|_p \\ &= \| f - Y_n(\mathbf{b}, \mathbf{z}(0)) - \alpha h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*)\|_p + o(\alpha) \\ &= \| f - y_0 - \alpha(h - y_0)\|_p + o(\alpha) \\ &= \|(1 - \alpha)(f - y_0) + \alpha(f - h)\|_p + o(\alpha) \\ &\leqslant \| f - y_0 \|_p + \alpha\{\| f - h \|_p - \| f - y_0 \|_p\} + o(\alpha). \end{aligned}$$

Hence, $\| f - y_0 \|_p \leqslant \| f - h \|_p$ so that y_0 is a best $\| \ \|_p$-approximation to f from $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$. Finally, using Lemma 3 and the fact that y_0 is a local best $\| \ \|_p$-approximation to f from $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$ we find

$$\begin{aligned} \| f - y_0 \|_p &\leqslant \| f - y_0 - \alpha h \|_p \\ &= \| f - y_0 \|_p + \alpha\Phi_p[\epsilon, -h] + o(\alpha) \end{aligned}$$

for all sufficiently small $\alpha > 0$ from which (26) follows. □

COROLLARY 1. *Let the exponential sum $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ be a local best $\| \ \|_p$-approximation to f from $V_n(S)$, let $\epsilon = f - y_0$, and assume $\| \epsilon \|_p > 0$. Then for each h in the n-dimensional manifold $L_n(\mathbf{c})$ we have*

$$\left| \int_{\epsilon \neq 0} h(t)\, \overline{\epsilon(t)} / |\, \epsilon(t)|\, dt \right| \leqslant \int_{\epsilon = 0} |\, h(t)|\, dt \qquad \text{if} \quad p = 1, \tag{27a}$$

$$\int_{\epsilon \neq 0} h(t)\, \overline{\epsilon(t)} |\, \epsilon(t)|^{p-2}\, dt = 0 \qquad \text{if} \quad 1 < p < \infty, \tag{27b}$$

$$\inf_{|\theta|=1} \lim_{\delta \to 0+} \operatorname*{ess\,sup}_{|\epsilon| \geqslant \|\epsilon\|_\infty - \delta} \operatorname{Re}\{\theta \cdot h(t)\, \overline{\epsilon(t)} / |\, \epsilon(t)|\} \geqslant 0 \qquad \text{if} \quad p = \infty. \tag{27c}$$

If in addition each element of $\Lambda_n(\mathbf{c})$ is an interior point of S, then (27) *also holds for each h in the $n + k$-dimensional manifold $H_n(\mathbf{b}, \mathbf{c})$ (where $k \leqslant n$ is the order of y_0.)*

Proof. We shall first consider the case where $p = 1$. Since y_0 is accessible through the (degenerate) cone $K_n(\mathbf{b}, \mathbf{c}, \mathbf{0}) = L_n(\mathbf{c})$ the inequality (26) or equivalently the inequality

$$\mathrm{Re} \int_{\epsilon \neq 0} h(t)\, \overline{\epsilon(t)} / | \epsilon(t) |\, dt \leqslant \int_{\epsilon = 0} | h(t) |\, dt \tag{28}$$

must hold for each $h \in L_n(\mathbf{c})$. But since $L_n(\mathbf{c})$ is a linear space, (28) must also hold whenever h is replaced by θh where $\theta \in \mathbf{C}$ with $| \theta | = 1$. By appropriately choosing θ we obtain (27a).

Suppose in addition that $\Lambda_n(\mathbf{c})$ is contained in the interior of S. Then y_0 is accessible through every cone $K_n(\mathbf{b}, \mathbf{c}, \mathbf{c}^*)$, $\mathbf{c}^* \in \mathbf{C}$, and, in view of (20) and the theorem, this implies that (28) must hold for every $h \in H_n(\mathbf{b}, \mathbf{c})$. Since $H_n(\mathbf{b}, \mathbf{c})$ is also a linear space this implies as before that (27a) holds for each $h \in H_n(\mathbf{b}, \mathbf{c})$. Hence, the proof is complete for the case $p = 1$.

Analogous considerations apply in the cases where $1 < p < \infty$ and $p = \infty$. $\square$

COROLLARY 2. *Let* $1 \leqslant p < \infty$, *let* y_0 *be a local best* $\| \; \|_p$*-approximation to* f *from* $V_n(S)$ *with* $\| f - y_0 \|_p > 0$, *and assume that* S *possesses some finite accumulation point, i.e., that* S *contains some convergent sequence* $\lambda_1, \lambda_2, \ldots$ *with* $\lambda_i \neq \lambda_j$ *for* $i \neq j$. *In the case where* $p = 1$ *assume in addition that* $f(t) \neq y_0(t)$ *holds for almost all* t. *Then* y_0 *has full order* n.

Proof. We shall first show that the corollary holds when $n = 1$. Indeed, if the zero function is a local best $\| \; \|_p$-approximation to f from $V_1(S)$ then from Corollary 1 we see that

$$\int_0^1 | f(t) |^{p-2}\, \overline{f(t)}\, y(t)\, dt = 0 \tag{29}$$

holds for each $y \in V_1(S)$. This being the case, (29) also holds whenever y is a linear combination of terms from $V_1(S)$, and in particular whenever y may be expressed in the form

$$y(t) = \exp(\lambda_\nu t) \int_0^t \int_0^{t_1} \cdots \int_0^{t_{n-1}} \exp \left\{ \sum_{j=1}^{m} (\lambda_{\nu+j} - \lambda_{\nu+j-1})\, t_j \right\} dt_1\, dt_2 \cdots dt_m$$

for some $\nu = 1, 2, \ldots$ and some $m = 1, 2, \ldots$. Thus, if λ is the limit of the sequence $\{\lambda_\nu\}$, we see that (29) holds whenever

$$\begin{aligned} y(t) &= \exp(\lambda t) \int_0^t \int_0^{t_1} \cdots \int_0^{t_{m-1}} dt_1\, dt_2 \cdots dt_m \\ &= \exp(\lambda t) \cdot t^m / m! \end{aligned}$$

for some $m = 1, 2, \ldots$. Finally, (29) must also hold for all linear combinations of such terms, i.e., whenever

$$y(t) = \exp(\lambda t) \cdot P(t) \tag{30}$$

for some polynomial P. Using (29) and the Hölder inequality we find

$$\begin{aligned} \|f\|_p^p &= \int_0^1 |f|^{p-2} \cdot \bar{f} \cdot [f - y]\, dt \\ &\leqslant \left\{\int_0^1 |f|^p\, dt\right\}^{(p-1)/p} \cdot \left\{\int_0^1 |f - y|^p\, dt\right\}^{1/p} \\ &= \|f\|_p^{p-1} \cdot \|f - y\|_p \end{aligned}$$

must hold whenever y has the form (30) and, hence, $\|f\|_p = 0$. Thus, the corollary holds when $n = 1$.

The extension to $n = 2, 3, \ldots$ is now immediate. Indeed, if $y_0 \in V_{n-1}(S)$ and $\|f - y_0\|_p > 0$, then the zero function cannot be a local best $\|\ \|_p$-approximation to $f - y_0$ from $V_1(S)$, and, hence, y_0 cannot be a local best $\|\ \|_p$-approximation to f from $V_n(S)$. □

The mild requirement that S have a finite accumulation point cannot be relaxed, e.g. the zero function is the unique best $\|\ \|_2$-approximation to $f \equiv 1$ from $V_1(S)$ when $S = \{\nu \cdot 2\pi i: \nu = \pm 1, \pm 2, \ldots\}$.

Somewhat different formulations and proofs of Corollary 2 for the case of $V_n^r(\mathbf{R})$ are given by Hobby and Rice [5, p. 99] and by de Boor [1, p. 179] with the case $p = 1$ being left undecided. By examing the above proof we see that when $p = 1$ and the zero function is a best $\|\ \|_1$-approximation to f, then we cannot deduce that (29) holds for all $y \in V_1(S)$ without imposing the additional hypothesis that $f \neq 0$ almost everywhere.

EXAMPLE 1. Let us define

$$f(t) = \begin{cases} 1 & \text{if } \ 1/3 \leqslant t \leqslant 2/3, \\ 0 & \text{otherwise} \end{cases}$$

and set

$$\begin{aligned} y(t) &= a \exp(\lambda t), \\ s &= \exp(|\lambda|/3). \end{aligned}$$

We then find

$$\begin{aligned}\|f+y\|_1 &\geqslant \int_0^{1/3} |y|\,dt + \int_{1/3}^{2/3} [f-|y|]\,dt + \int_{2/3}^{1} |y|\,dt \\ &= \|f\|_1 + (1-s+s^2)\int_0^{1/3} |y|\,dt \\ &= \|f\|_1 + [(1-s+s^2)/(1+s+s^2)]\int_0^1 |y|\,dt \\ &\geqslant \|f\|_1 + (1/3)\|y\|_1 \\ &\geqslant \|f\|_1\end{aligned}$$

with equality holding throughout if and only if $a = 0$. Thus, the zero function is the unique best $\|\ \|_1$-approximation to f from $V_1(\mathbf{C})$. □

The following example has been chosen to show that a best $\|\ \|_p$-approximation need not be unique and to show that the necessary conditions of Theorem 2 and Corollaries 1 and 2 are not sufficient conditions to characterize a local best $\|\ \|_p$-approximation.

EXAMPLE 2. We define

$$f(t) = 90t^2 - 90t + 16, \qquad 0 \leqslant t \leqslant 1$$

and first seek a best $\|\ \|_2$-approximation to f from $V_1(\mathbf{R})$. If the exponential sum

$$y(t) = a\exp(\lambda t)$$

is a best $\|\ \|_2$-approximation to f, then from (27b) we find

$$\int_0^1 [f(t) - y(t)]\exp(\lambda t)\,dt = 0$$

or equivalently

$$a = \left\{\int_0^1 f(t)\exp(\lambda t)\,dt\right\} \Big/ \left\{\int_0^1 \exp(2\lambda t)\,dt\right\}. \tag{31}$$

Using (31) and defining

$$\psi(\lambda) = \|f-y\|_2^2,$$

we find

$$\psi(\lambda) = 46 - B(\lambda/2)^2/\{(\lambda/2)^5 \sinh(\lambda/2)\cosh(\lambda/2)\},$$

where

$$B(u) \equiv (45 + 16u^2)\sinh u - 45\,u\cosh u.$$

It can be shown that ψ' vanishes only at the relative maximum of ψ at $\lambda = 0$ and at the absolute minima of ψ at $\lambda = \pm 12.3064...$ so that

$$y_+(t) = 0.00008931... \exp(12.3064...\, t),$$
$$y_-(t) = y_+(1 - t)$$

are the two best $\| \|_2$-approximations to f from $V_1(\mathbf{R})$.

Suppose now that we seek a best $\| \|_2$-approximation to f from the larger class $V_1(\mathbf{C})$. Taking

$$y_0(t) = Y_1(1, 0, t) \equiv 1,$$

we find find that (by the construction of f)

$$\int_0^1 [f(t) - y_0(t)]\, dt = \int_0^1 [f(t) - y_0(t)]\, t\, dt = 0$$

so that y_0 is the unique best $\| \|_2$-approximation to f from the tangential manifold

$$H_1(1, 0) = \{h\colon D^2 h \equiv 0\}.$$

Thus, y_0 satisfies the necessary conditions of Theorem 2 and its accompanying corollaries, but y_0 fails to be a best $\| \|_2$-approximation to f from $V_1(\mathbf{R})$ and, hence, from the larger set $V_1(\mathbf{C})$. Finally, we note that even in the larger class $V_1(\mathbf{C})$ f has no unique best $\| \|_2$-approximation. Indeed, since $f(t) = \overline{f(t)} = f(1 - t)$ we see that

$$y_1(t) = a \exp(\lambda t), \qquad y_2(t) = \bar{a} \exp(\bar{\lambda} t),$$
$$y_3(t) = a \exp[\lambda(1 - t)], \qquad y_4(t) = \bar{a} \exp[\bar{\lambda}(1 - t)]$$

all $\| \|_p$-approximate f equally well, and when $a, \lambda \neq 0$ at least two of these four functions must be distinct. □

In the special case where f is a real valued continuous function, one might attempt to characterize a best $\| \|_\infty$-approximation to f from $V_n^r(S)$ in terms of the number of alternations of the corresponding error curve, cf. [2; 8; 10, Chapter 8]. Such a characterization can be made provided the exponential parameters are real so that Lemma 4 can be utilized.

Corollary 3. *Let the exponential sum $y_0 \in V_n^r(S)$ be expressible in the form* (4) *with $\lambda_i^0 \in \mathbf{R}$ and $a_{in_i} \neq 0$, $i = 1, 2,..., l$, and let $f \in C[0, 1]$ be real valued.*

(a) *If $S \subseteq \mathbf{C}$ and each λ_i^0 is an interior point of S with respect to the* *topology of* $\mathbf{C}$, *then a necessary condition for* y_0 *to be a local best* $\|\ \|_\infty$*-approximation to* f *from* $V_n^r(S)$ *is that the error curve* $\epsilon = f - y_0$ *alternate at least* $n + k$ *times on* $[0, 1]$.

(b) *If* $S \subseteq \mathbf{R}$ *and each* λ_i^0 *is an interior point of* S *with respect to the topology of* $\mathbf{R}$, *then a necessary condition for* y_0 *to be such a local best* $\|\ \|_\infty$*-approximation is that* ϵ *alternate at least* $n + l$ *times on* $[0, 1]$.

Proof. Assuming that y_0 is a local best $\|\ \|_\infty$-approximation to f from $V_n^r(S)$ and that each λ_i^0 is an interiror point of S with respect to the topology of $\mathbf{C}$, we see from Corollary 1 that y_0 must be a best $\|\ \|_\infty$-approximation to f from the $n + k$-dimensional real linear space $H_n^r(\mathbf{b}, \mathbf{c})$ whenever $\mathbf{b}, \mathbf{c} \in \mathbf{R}^n$ are chosen in such a manner that $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ with $\Lambda(\mathbf{c}) \subset \mathbf{R}$. But since $\Lambda(\mathbf{c}) \subset \mathbf{R}$, we see from Lemmas 1 and 4 that any basis of the $n + k$-dimensional real linear space $H_n^r(\mathbf{b}, \mathbf{c})$ forms a Haar system, and since y_0 is a best $\|\ \|_\infty$-approximation to f from $H_n^r(\mathbf{b}, \mathbf{c})$ it follows (cf. [9, Chapter 3]) that $f - y_0$ alternates at least $n + k$ times on $[0, 1]$.

More generally, if we can only assume that each λ_i^0 is an interior point of $S \subseteq \mathbf{R}$ with respect to the topology of $\mathbf{R}$, then arguments analogous to those used in the proof of Corollary 1 can be used to show that y_0 must be a best $\|\ \|_\infty$-approximation to f from the $n + l$-dimensional real linear space $G_n^r(\mathbf{b}, \boldsymbol{\lambda})$ with the components of $\boldsymbol{\lambda}$ being interior points of S. Again any basis of $G_n^r(\mathbf{b}, \boldsymbol{\lambda})$ forms a Haar system so that $f - y_0$ must alternate $n + l$ times on $[0, 1]$. $\square$

By constraining the exponential parameters to be real, one loses $k - l$ degrees of freedom in the tangential manifold and this is reflected in the weakened alternation requirement of Corollary 3 which was first pointed out by Braess [2, p. 313]. That this condition cannot be strengthened to a sufficient condition (e.g. as once suggested by Rice [8, p. 158]) may be seen from the following generalization of an example originally suggested by Braess [2, p. 315].

EXAMPLE 3. Let

$$f(t) = \sin(\pi t + \beta), \qquad 0 \leqslant t \leqslant 1$$

(where β is a real parameter which we shall later specialize), let the kth order exponential sum y be a local best $\|\ \|_\infty$-approximation to f from $V_n^r(\mathbf{R})$, and let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_k$ be the essential exponential parameters of y with exactly l of these being distinct. Then if

$$\begin{aligned} \Phi &= \{D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_k)\}(f - y) \\ &= \{(D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_k)\} f, \end{aligned}$$

we see that Φ is a nontrivial linear combination of $\sin(\pi t + \beta)$ and $\cos(\pi t + \beta)$ so that Φ has at most one zero in (0, 1). Using Lemma 4, we conclude that the error curve $\epsilon = f - y_0$ has at most $k + 1$ zeros in (0, 1) so that ϵ alternates at most $k + 1$ times on [0, 1]. On the other hand, from Corollary 3 above we see that ϵ must alternate at least $n + l$ times. Hence, from the inequalities $0 \leqslant l \leqslant k \leqslant n$ and $n + l \leqslant k + 1$ we conclude that either $n = 1$ and $k = l = 0$ (i.e., $y \equiv 0$) or else $k = n$ and $l = 1$ so that y may be parametrized in the form

$$y(\mathbf{a}, \lambda; t) = \{a_0 + a_1 t + \cdots + a_{n-1} t^{n-1}\} \exp(\lambda t),$$

where $\lambda, a_0, \ldots, a_{n-1}$ are real and $a_{n-1} \neq 0$. Thus, for all $n \geqslant 2$ we have a maximum degeneracy in the sense that all of the exponential parameters coalesce. Moreover, since the error curve alternates exactly $n + 1$ times with $k = n$ and $l = 1$ for $n \geqslant 2$ we see from Corollary 3 that no local best $\| \ \|_\infty$-approximation to f from the larger class $V_n^r(\mathbf{C})$ is found in $V_n^r(\mathbf{R})$.

In order to determine λ, $\mathbf{a}$ we may proceed as follows. For each real λ we may use numerical methods to uniquely find the linear parameters $\mathbf{a}(\lambda)$ which make the error curve $f - y(\mathbf{a}(\lambda), \lambda)$ alternate at least n times. With $\mathbf{a}(\lambda)$ thus determined, we compute

$$\psi(\lambda) = \| f - y(\mathbf{a}(\lambda), \lambda) \|_\infty$$

and by numerically locating the local minima of ψ find each local best $\| \ \|_\infty$-approximation to f from $V_n^r(\mathbf{R})$.

When $\beta = 0$ we find that

$$\begin{aligned}
y_1(t) &= 0.5000\ldots \\
y_2^+(t) &= \{0.2762\ldots - 0.2839\ldots t\} \exp(+3.5716\ldots\, t) \\
y_2^-(t) &= \{-0.2762\ldots + 10.1017\ldots\, t\} \exp(-3.5716\ldots\, t) \\
y_3^+(t) &= \{0.2072\ldots - 0.4358\ldots\, t + 0.2290\ldots\, t^2\} \exp(+6.0790\ldots\, t) \\
y_3^0(t) &= -0.0280\ldots + 4.000\ldots\, t - 4.0000\ldots\, t^2 \\
y_3^-(t) &= \{0.2072\ldots - 9.7123\ldots\, t + 100.0094\ldots\, t^2\} \exp(-6.0790\ldots\, t) \\
&\ \ \vdots
\end{aligned}$$

are the local best $\| \ \|_\infty$-approximations to f from $V_1^r(\mathbf{R})$, $V_2^r(\mathbf{R})$, $V_3^r(\mathbf{R})$,... (with $\| \epsilon \|_\infty = | y(0) |$ in each case.) We thus see from this simple example that a best $\| \ \|_\infty$-approximation to a given f from $V_n(\mathbf{R})$ need not be unique and need not alternate $2n$ times. Moreover, there may be a number of local best $\| \ \|_\infty$-approximations which are not best $\| \ \|_\infty$-approximations. $\square$

4. A Sufficient Condition

It has been observed that the first order necessary conditions of the previous section fail to be sufficient to characterize a best $\| \ \|_p$-approximation. When y_0 has full order, we may obtain a sufficient condition by introducing a mild hypothesis on the $o(\alpha)$ term of Lemma 3.

THEOREM 3. *Let $f \in L_p[0, 1]$ with $1 \leqslant p \leqslant \infty$, let $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ be chosen from $V_n(S)$, let $\epsilon = f - y_0$, and assume that y_0 has full order n. Let*

$$\mu(\alpha) = \sup\{\| d^2 Y_n(\mathbf{b} + \alpha\mathbf{b}^*, \mathbf{c} + \alpha\mathbf{c}^*)/d\alpha^2 \|_\infty : 0 \leqslant | b_i^* |, | c_i^* | \leqslant 1\} \tag{32}$$

be defined for $\alpha > 0$ and assume that

$$\liminf_{\alpha \to 0+} (\| \epsilon + \alpha h \|_p - \| \epsilon \|_p)/\alpha^2 > (1/2) \limsup_{\alpha \to 0+} \mu(\alpha) \tag{33}$$

holds for each $h \in H_n(\mathbf{b}, \mathbf{c})$ with $\| h \|_p > 0$. Then y_0 is a local best $\| \ \|_p$-approximation to f from $V_n(S)$.

Proof. For each $\alpha > 0$ we define

$$\varphi(\alpha) = \alpha^{-2} \cdot \min\{\| \epsilon + \alpha h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*)\|_p - \| \epsilon \|_p : (\mathbf{b}^*, \mathbf{c}^*) \in \mathbf{K}\}, \tag{34}$$

where

$$\mathbf{K} = \{\mathbf{z} \in \mathbf{C}^{2n} : \max(| z_1 |, | z_2 |, \ldots, | z_{2n} |) = 1\}.$$

The set $\mathbf{K}$ is compact, and this together with the continuity of h_n and the hypothesis (33) implies that

$$\varphi(\alpha) > (1/2)\, \mu(\alpha) \tag{35}$$

holds for all sufficiently small $\alpha > 0$. This being the case we may use the mean value theorem together with (32), (34), and (35) to see that for every $(\mathbf{b}^*, \mathbf{c}^*) \in \mathbf{K}$ and for sufficiently small $\alpha > 0$ we have

$$\begin{aligned} &\| f - Y_n(\mathbf{b} - \alpha\mathbf{b}^*, \mathbf{c} - \alpha\mathbf{c}^*)\|_p \\ &\quad \geqslant \| f - Y_n(\mathbf{b}, \mathbf{c}) + \alpha h_n(\mathbf{b}, \mathbf{c}, \mathbf{b}^*, \mathbf{c}^*)\|_p - \mu(\alpha)\, \alpha^2/2 \\ &\quad \geqslant \| \epsilon \|_p + \alpha^2[\varphi(\alpha) - (1/2)\, \mu(\alpha)] \\ &\quad > \| \epsilon \|_p . \end{aligned}$$

In view of the homeomorphism established by Theorem 1, this shows that y_0 is a local best $\| \ \|_p$-approximation to f from $V_n(\mathbf{C})$ and, therefore, from $V_n(S)$ as well. □

When $p = \infty$ the crutial hypothesis (33) may be simplified as follows.

COROLLARY 1. *Let $f \in C[0, 1]$, let $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ be chosen from $V_n(S)$, let $\epsilon = f - y_0$, and assume that y_0 has full order n. Assume further that*

$$\min\{\operatorname{Re} \overline{\epsilon(t)}\, h(t) \colon |\epsilon(t)| = \|\epsilon\|_\infty\} < 0 \tag{36}$$

holds for every $h \in H_n(\mathbf{b}, \mathbf{c})$ with $\|h\|_\infty > 0$ (i.e., *assume that there is strict inequality in the necessary condition of* (27c).) *Then y_0 is a local best $\|\ \|_\infty$-approximation to f from $V_n(S)$.*

Proof. Let $h \in H_n(\mathbf{b}, \mathbf{c})$ be chosen with $\|h\|_\infty > 0$. Using Lemma 3 and the continuity of f together with (36) we find

$$\begin{aligned}
&\liminf_{\alpha \to 0+} \alpha^{-2}[\|\epsilon - \alpha h\|_\infty - \|\epsilon\|_\infty] \\
&\quad = \liminf_{\alpha \to 0+} \alpha^{-2}[-\alpha \min\{\operatorname{Re} \overline{\epsilon(t)}\, h(t) \colon |\epsilon(t)| = \|\epsilon\|_\infty\} + o(\alpha)] \\
&\quad = +\infty
\end{aligned}$$

so that (33) holds. Thus, from the theorem we conclude that y_0 is a local best $\|\ \|_\infty$-approximation to f from $V_n(S)$. □

COROLLARY 2. *Let $f \in [0, 1]$ be real valued, let the* nth *order exponential sum $y_0 = Y_n(\mathbf{b}, \mathbf{c})$ be chosen from $V_n{}^r(S \cap \mathbf{R})$, and assume that $\Lambda_n(\mathbf{c})$ is contained in the interior of S. Then a necessary and sufficient condition for y_0 to be a local best $\|\ \|_\infty$-approximation to f from $V_n{}^r(S)$ is that the error curve $\epsilon = f - y_0$ alternate at least $2n$ times on* $[0, 1]$.

Proof. If ϵ alternates at least $2n$ times on $[0, 1]$ and if $h \in H_n{}^r(\mathbf{b}, \mathbf{c})$ is chosen such that (36) fails to hold, then h has at least $2n$ zeros in $[0, 1]$ and, therefore, (since $h \in V_{2n}(\mathbf{R})$) we conclude from Lemma 4 that $\|h\|_\infty = 0$. Hence, the sufficiency follows from Corollary 1 above. The necessity follows from Corollary 3 of Theorem 2. □

Using Lemma 4 it can be shown (cf. [2, p. 313]) that a sufficient condition for the kth order exponential sum y_0 to be a best $\|\ \|_\infty$-approximation to the continuous real valued function f from $V_n{}^r(\mathbf{R})$ is that the error curve $\epsilon = f - y_0$ alternate at least $n + k$ times on $[0, 1]$, and in view of Corollary 2 above one might possibly hope to obtain a similar sufficient condition which would apply in the larger class $V_n{}^r(\mathbf{C})$. Unfortunately, the possession of an alternant of any fixed length (no matter how long) is not a sufficient condition to characterize a best $\|\ \|_\infty$-approximation to f from $V_n{}^r(\mathbf{C})$ when $n \geqslant 2$. (We note that this contradicts Rice [9, p. 62–63] where it is incorrectly stated that $V_n{}^r(\mathbf{C})$ is varisolvent under a suitable parametrization.) This is easily seen from the following.

EXAMPLE 4. We define

$$f(t) = \cos(N\pi t), \qquad 0 \leqslant t \leqslant 1,$$

and seek a best $\| \|_\infty$-approximation to f from $V_2{}^r(\mathbf{C})$. Since $f \in V_2{}^r(\mathbf{C})$, f is its own best $\| \|_\infty$-approximation. We note, however, that if

$$y(t) = a \cdot \cos(N\pi t)$$

for some real $a \neq 1$, then the error curve $\epsilon = f - y$ alternates exactly N times on [0, 1]. Clearly, y fails to be even a local best $\| \|_\infty$-approximation to f from $V_2{}^r(\mathbf{C})$. □

REFERENCES

1. C. DE BOOR, On the Approximation by γ-polynomials, *in* "Approximations with Special Emphasis on Spline Functions" (I. J. Schoenberg, Ed.), pp. 157–183, Academic Press, New York, 1969.
2. D. BRAESS, Approximation mit Exponentialsummen, *Computing* **2** (1967), 309–321.
3. D. BRAESS, Über die Vorzeichenstruktur der Exponentialsummen, *J. Approximation Theory* **3** (1970), 101–113.
4. J. DUGUNDJI, "Topology," Allyn and Bacon, Boston, MA, 1966.
5. C. R. HOBBY AND J. R. RICE, Approximation from a curve of functions, *Arch. Rat. Mech. Anal.* **27** (1967), 91–106.
6. D. W. KAMMLER, Existence of best approximations by sums of exponentials, *J. Approximation Theory* **9** (1973), 78–90.
7. G. POLYA AND G. SZEGO, "Aufgaben und Lehrsatz aus der Analysis," Band II, Springer-Verlag, Berlin, 1954.
8. J. R. RICE, Chebyshev approximation by exponentials, *J. Soc. Indust. Appl. Math.* **10** (1962), 149–161.
9. J. R. RICE, "The Approximation of Functions," Vol. I, Addison-Wesley, Reading, MA, 1964.
10. J. R. RICE, "The Approximation of Functions," Vol. II, Addison-Wesley, Reading, MA, 1969.
11. H. L. ROYDEN, "Real Analysis," Macmillan, New York, 1968.