JOURNAL OF APPROXIMATION THEORY 16, 150–161 (1976)

Lower Estimates for the Error of Best Uniform Approximation

GÜNTER MEINARDUS

Gesamthochschule Siegen, Huttental-Weidenau, Germany

AND

G. D. TAYLOR*

Department of Mathematics, Colorodo State University, Fort Collins, Colorodo 80523

INTRODUCTION

In this paper, we generalize the lower bounds of de La Vallée Poussin and Remes [2, p. 82] for the error of best uniform approximation from a linear subspace. Precisely, let $C[a, b]$ denote the space of all continuous real valued functions defined on the closed interval $[a, b]$ with norm $\|f\| = \max\{|f(x)| : x \in [a, b]\}$. Then, the above two results are

THEOREM 1 (de La Vallée Poussin). *Let V be an n-dimensional Haar subspace of $C[a, b]$ and let $f \in C[a, b]$. Let $h \in V$ and suppose that there exist $n + 1$ points $a \leqslant x_1 < \cdots < x_{n+1} \leqslant b$ such that the error function $e(x) = f(x) - h(x)$ satisfies*

1. $e(x_i) \neq 0, \qquad i = 1,..., n + 1,$
2. $\operatorname{sgn} e(x_{i+1}) = -\operatorname{sgn} e(x_i), \qquad i = 1,..., n.$

Then,

$$\min_{0 \leqslant i \leqslant n+1} | e(x_i)| \leqslant \rho(f) \equiv \inf_{p \in V} \|f - p\|.$$

THEOREM 2 (Remes). *Let π_{n-1} denote the set of all algebraic polynomials of degree $\leqslant n - 1$ and let $f \in C[a, b]$. Let $h \in \pi_{n-1}$ and suppose that there exist*

* Supported in part by AFSOR 72-2271.

$n+1$ points $a \leqslant x_1 < \cdots < x_{n+1} \leqslant b$ such that the error function $e(x) = f(x) - h(x)$ satisfies

1. $e(x_i) \neq 0, \qquad i = 1,..., n+1,$
2. $\operatorname{sgn} e(x_{i+1}) = -\operatorname{sgn} e(x_i), \qquad i = 1,..., n.$

Then,

$$\min_{1 \leqslant i \leqslant n} \tfrac{1}{2}(|e(x_i)| + |e(x_{i+1})|) \leqslant \rho_n(f).$$

In what follows, we generalize these results to give analogous estimates based on k points, $k = 1,..., n$. For the special cases $k = 1, n$ our estimates will be simply the de La Vallée Poussin estimate and the error of approximation on the points $x_1, ..., x_{n+1}$, respectively. For the case $k = 2$, we have a slight generalization of the Remes estimate in that we do not require the approximants to be algebraic polynomials. Our precise generalization is given in Section 4. In the next two sections, we develop the necessary tools to prove our generalization.

2. Decomposition Theorem

Fix $n+1$ distinct points $a \leqslant x_1 < x_2 \cdots < x_{n+1} \leqslant b$. For each k, $1 \leqslant k \leqslant n$ and ν, $1 \leqslant \nu \leqslant n - k + 1$ define $M_{\nu k}$ by $M_{\nu k} = \{x_\nu, x_{\nu+1}, ..., x_{\nu+k}\}$. Let $V_n = \langle \varphi_1, ..., \varphi_n \rangle$ be a fixed Haar subspace of $C[a, b]$ and for each j, $1 \leqslant j \leqslant n$, set $V_j = \langle \varphi_1, ..., \varphi_j \rangle$ (i.e., V_j is the subspace of $C[a, b]$ spanned by the functions $\varphi_1, ..., \varphi_j$). If V_k ($k = 1,..., n$) satisfies the Haar condition, then, using the standard theory of Haar subspaces [2, p. 19], a linear functional L_ν^k based on $M_{\nu k}$ can be defined by

$$L_\nu^k(f) = \sum_{j=\nu}^{\nu+k} \lambda_j^{\nu k} f(x_j), \qquad f \in C[a, b], \tag{1}$$

where $\lambda_j^{\nu k}$ satisfies $\lambda_\nu^{\nu k} > 0$, $\lambda_j^{\nu k} \neq 0$ for $\nu \leqslant j \leqslant \nu + k$, $\operatorname{sgn} \lambda_j^{\nu k} = (-1)^{j-\nu}$, $\sum_{j=\nu}^{\nu+k} |\lambda_j^{\nu k}| = 1$ and $\sum_{j=\nu}^{\nu+k} \lambda_j^{\nu k} \varphi_\mu(x_j) = 0$ for $\mu = 1,..., k$. The existence and uniqueness subject to $\lambda_\nu^{\nu k} > 0$ and $\sum_{j=\nu}^{\nu+k} |\lambda_j^{\nu k}| = 1$, of such a linear functional is well-known, as well as that

$$|L_\nu^k(f)| = \inf_{h \in V_k} \{\max_{x \in M_{\nu k}} |f(x) - h(x)|\}. \tag{2}$$

For consistency of notation we write $L_\nu^0(f) = f(x_\nu)$ throughout this paper. Using this notation, we now turn to proving our decomposition theorem.

Theorem 3. *Fix k, $1 \leqslant k \leqslant n$; r, $0 \leqslant r \leqslant k$; and ν, $1 \leqslant \nu \leqslant n - k + 1$ and assume that V_j satisfies the Haar condition for $j = 1,..., r$ and k (if $r = 0$,* *then we only assume this for* $j = k$). *Then, there exists a unique decomposition of the linear functional* L_ν^k *in terms of the linear functionals* L_j^r, $j = \nu, ..., \nu + k - r$

$$L_\nu^k(f) = \sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r(f), \qquad f \in C[a, b], \tag{3}$$

where the real numbers $\lambda_{jr}^{\nu k}$ *are all different from zero,* $\operatorname{sgn} \lambda_{jr}^{\nu k} = (-1)^{j+\nu}$ $j = \nu, ..., \nu + k - r$ *and* $\sum_{j=\nu}^{\nu+k-r} | \lambda_{jr}^{\nu k} | = 1$.

Proof. This theorem is valid for $r = 0$ by our remarks concerning the properties of Haar subspaces. Thus, we assume that $r \geqslant 1$. Since L_ν^k is not the zero linear functional, there exists a function $\varphi \in C[a, b]$ for which $L_\nu^k(\varphi) = 1$. Now, on the point set $M_{\nu k}$ the functions $\varphi, \varphi_1, ..., \varphi_k$ are linearly independent. Thus,

$$f(x) = \alpha\varphi(x) + \sum_{\mu=1}^{k} \alpha_\mu \varphi_\mu(x), \qquad x \in M_{\nu k}, \tag{4}$$

where $\alpha, \alpha_1, ..., \alpha_k$ are unique. We must show, since $L_\nu^k(\varphi) = 1$ and $L_\nu^k(\varphi_\mu) = 0$, $\mu = 1, ..., k$, that there exist numbers $\lambda_{jr}^{\nu k}$, uniquely determined, that satisfy

$$\sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r(\varphi_\mu) = 0, \qquad \mu = 1, ..., k, \qquad \sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r(\varphi) = 1. \tag{5}$$

Since, by definition of L_j^r,

$$\sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r(\varphi_\mu) = 0$$

for $\mu = 1, ..., r$ and every choice of $\lambda_{jr}^{\nu k}$, it is necessary and sufficient to show that the $(k - r + 1) \times (k - r + 1)$ matrix

$$B = \begin{pmatrix} L_\nu^r(\varphi_{r+1}) & \cdots & L_{\nu+k-r}^r(\varphi_{r+1}) \\ \vdots & & \\ L_\nu^r(\varphi_k) & \cdots & L_{\nu+k-r}^r(\varphi_k) \\ L_\nu^r(\varphi) & \cdots & L_{\nu+k-r}^r(\varphi) \end{pmatrix}$$

is nonsingular. To do this, we consider the transposed matrix $B^\top$ and with any fixed vector $b = (b_\nu, ..., b_{\nu+k-r})^\top$, the system of linear equations

$$B^\top a = b, \tag{6}$$

where $a = (\alpha_{r+1}, ..., \alpha_k, \alpha)^\top$ represents a solution (if one exists). Now (6) can be rewritten as

$$L_j^r\left(\alpha\varphi + \sum_{i=r+1}^{k} \alpha_i \varphi_i\right) = b_j, \qquad j = \nu, ..., \nu + k - r. \tag{7}$$

Thus, we wish to exhibit a function Ψ in $\langle \varphi_{r+1}, \ldots, \varphi_k, \varphi \rangle$ for which

$$L_j^r(\Psi) = b_j, \qquad j = \nu, \ldots, \nu + k - r \tag{8}$$

is satisfied. Using the representation (1) of each L_j^r, $j = \nu, \ldots, \nu + k - r$, we have that (8) is equivalent to

$$C\hat{\Psi} = b, \tag{9}$$

with $\hat{\Psi} = (\hat{\Psi}(x_\nu), \ldots, \hat{\Psi}(x_{\nu+k}))^\top$ and

$$C = \begin{pmatrix} \lambda_\nu^{\nu r} & \cdots & \lambda_{\nu+r}^{\nu r} & 0 & \cdots & 0 \\ 0 & \cdot & & & \cdot & \\ \vdots & & \cdot & & & \cdot & 0 \\ 0 & 0 & \lambda_{\nu+k-r}^{\nu+k-r,r} & & \cdots & \lambda_{\nu+k}^{\nu+k-r,r} \end{pmatrix}.$$

Since C has maximal rank $k - r + 1$ (as $\lambda_p^{pr} > 0$ for all $p = \nu, \ldots, \nu + k$), the existence of values $\hat{\Psi}(x_p)$, $p = \nu, \ldots, \nu + k$ satisfying (9) is guaranteed. Since $\langle \varphi_1, \ldots, \varphi_k, \varphi \rangle$ forms a basis for $M_{\nu k}$, we can find coefficients, $\alpha, \alpha_1, \ldots, \alpha_k$ so that

$$\tilde{\Psi}(x) = \alpha\varphi(x) + \sum_{\mu=1}^{k} \alpha_\mu \varphi_\mu(x)$$

satisfies $\tilde{\Psi}(x_i) = \hat{\Psi}(x_i)$, $i = \nu, \ldots, \nu + k$. Thus, the function

$$\Psi(x_i) = \alpha\varphi(x) + \sum_{\mu=r+1}^{k} \alpha_\mu \varphi_\mu(x)$$

satisfies (8) as desired and its coefficients are a solution of (6). Hence, by the Fredholm alternative, the matrix $B^\top$ is not singular as it maps R^{k-r+1} onto R^{k-r+1}. From this follows the existence and uniqueness of the numbers $\lambda_{jr}^{\nu k}$.

All that remains to be done is to prove the remaining assertions about the numbers $\lambda_{jr}^{\nu k}$. Let us begin by showing that $\lambda_{jr}^{\nu k} \neq 0$ and that $\operatorname{sgn} \lambda_{jr}^{\nu k} = (-1)^{j+\nu}$, $j = \nu, \ldots, \nu + k - r$. Now, if $r = k$, then clearly, $\lambda_{\nu k}^{\nu k} = 1$. We prove the general result using an induction argument on decreasing r. Thus, let us assume that

$$L_\nu^k = \sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r,$$

for fixed r, $0 < r \leqslant k$, where $\operatorname{sgn} \lambda_{jr}^{\nu k} = (-1)^{j+\nu}$. Consider the relation

$$L_\nu^r = \lambda_{\nu,r-1}^{\nu r} L_\nu^{r-1} + \lambda_{\nu+1,r-1}^{\nu r} L_{\nu+1}^{r-1}.$$

Using the representation (1) of each linear functional of this expression and operating on $\hat{f} \in C[a, b]$, where $\hat{f}(x_\mu) = \delta_{\nu\mu}$, we find that $\lambda_\nu^{\nu r} = \lambda_{\nu,r-1}^{\nu r} \lambda_\nu^{\nu,r-1}$

implying that $\lambda^{\nu r}_{\nu,r-1} > 0$, since both $\lambda^{\nu r}_{\nu}$ and $\lambda^{\nu,r-1}_{\nu}$ are positive. Likewise, applying this expression to $g \in C[a, b]$, where $g(x_\mu) = \delta_{\nu+r,\mu}$, gives

$$\lambda^{\nu r}_{\nu+r} = \lambda^{\nu r}_{\nu+1,r-1}\lambda^{\nu+1,r-1}_{\nu+r}.$$

Since $\operatorname{sgn} \lambda^{\nu r}_{\nu+1} = (-1)^r$ and $\operatorname{sgn} \lambda^{\nu+1,r-1}_{\nu+r} = (-1)^{r-1}$, it follows that $\operatorname{sgn} \lambda^{\nu r}_{\nu+1,r-1} = -1$. Therefore,

$$\begin{aligned} L_\nu{}^k &= \sum_{j=\nu}^{\nu+k-r} \lambda^{\nu k}_{jr} L_j{}^r \\ &= \lambda^{\nu k}_{\nu r}\lambda^{\nu r}_{\nu,r-1}L_\nu^{r-1} + \sum_{j=\nu+1}^{\nu+k-r} (\lambda^{\nu k}_{j-1,r}\lambda^{j-1,r}_{j,r-1} + \lambda^{\nu k}_{jr}\lambda^{jr}_{j,r-1})\, L_j^{r-1} \\ &\quad + \lambda^{\nu k}_{\nu+k-r,r}\lambda^{\nu+k-r,r}_{\nu+k-r+1,r-1}L^{r-1}_{\nu+k-r+1}\,. \end{aligned}$$

Uniqueness of the representation of $L_\nu{}^k$ in terms of L_j^{r-1} gives

$$\begin{aligned} \lambda^{\nu k}_{\nu,r-1} &= \lambda^{\nu k}_{\nu r}\lambda^{\nu r}_{\nu,r-1} > 0, \\ \operatorname{sgn} \lambda^{\nu k}_{j,r-1} &= \operatorname{sgn}(\lambda^{\nu k}_{j-1,r}\lambda^{j-1,r}_{j,r-1} + \lambda^{\nu k}_{jr}\lambda^{jr}_{j,r-1}) = (-1)^{j+\nu}, \\ &\qquad j = \nu + 1,\dots, \nu + k - r, \end{aligned}$$

and

$$\operatorname{sgn} \lambda^{\nu k}_{\nu+k-r,r-1} = \operatorname{sgn}(\lambda^{\nu k}_{\nu+k-r,r}\lambda^{\nu+k-r,r}_{\nu+k-r+1,r-1}) = (-1)^{k-r+1},$$

which completes the inductive argument. Finally, to show that

$$\sum_{j=\nu}^{\nu+k-r} | \lambda^{\nu k}_{jr} | = 1,$$

take $g \in C[a, b]$ so that $L_\nu{}^k(g) \neq 0$. Let $h \in V_k$ be the best approximation to g on the point set $M_{\nu k}$. From the standard theory of Haar subspaces, we have that

$$g(x_\mu) - h(x_\mu) = (-1)^{\mu+\nu}L_\nu{}^k(g), \qquad \mu = \nu,\dots, \nu + k.$$

Thus, for $\nu \leqslant j \leqslant \nu + k - r$,

$$\begin{aligned} L_j{}^r(g - h) &= \sum_{\mu=j}^{j+r} \lambda^{jr}_\mu(g(x_\mu) - h(x_\mu)) \\ &= L_\nu{}^k(g)(-1)^\nu \sum_{\mu=j}^{j+r} \lambda^{jr}_\mu(-1)^\mu \\ &= (-1)^{j+\nu}\, L_\nu{}^k(g). \end{aligned}$$

Hence,

$$L_\nu^k(g) = (-1)^\nu L_\nu^k(g) \sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k}(-1)^j,$$

or

$$\sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k}(-1)^{j+\nu} = \sum_{j=\nu}^{\nu+k-r} | \lambda_{jr}^{\nu k} | = 1,$$

as desired, completing the proof of the theorem. ■

3. Recursive Computation of the Linear Functionals L_ν^k

In this section, we give a recursive scheme for constructing the values of the linear functional L_ν^k applied to a given function f. To accomplish this, first we must observe that $L_\nu^{k-1}(\varphi_k)$ is never zero and has a constant sign as a function of ν, $1 \leqslant \nu \leqslant n - k + 2$, provided V_k satisfies the Haar condition.

Lemma 1. *For each k, $1 \leqslant k \leqslant n$ and ν, $1 \leqslant \nu \leqslant n - k + 2$, $L_\nu^{k-1}(\varphi_k) \neq 0$ and* sgn $L_\nu^{k-1}(\varphi_k) =$ sgn $L_{\nu+1}^{k-1}(\varphi_k)$, $\nu = 1,..., n - k + 1$.

Proof. Clearly, this is true for $k = 1$. For $k \geqslant 2$, $| L_\nu^{k-1}(\varphi_k)|$ equals the minimal deviation in approximating φ_k by V_{k-1} on the point set $M_{\nu,k-1}$. If this were zero, then there would exist $\tilde{\varphi} \in V_{k-1}$, equal to φ_k at the k points of $M_{\nu,k-1}$. Since $\varphi_k \notin V_{k-1}$, the difference would then be a function in V_k having k zeros that is not identically zero, contradicting the Haar condition. To prove that sgn $L_\nu^{k-1}(\varphi_k) =$ sgn $L_{\nu+1}^{k-1}(\varphi_k)$, one uses the continuous dependence of $L_\nu^{k-1}(\varphi_k)$ on the points to show that a new selection of points could be made in the event sgn $L_\nu^{k-1}(\varphi_k) = -$sgn $L_{\nu+1}^{k-1}(\varphi_k)$ (some ν), on which $L_\nu^{k-1}(\varphi_k) = 0$ holds. Thus, the above arguments preclude that this occurs. ■

Using these facts, we can give a recursive scheme for calculating $L_\nu^k(f)$, $f \in C[a, b]$, $1 \leqslant k \leqslant n$, $1 \leqslant \nu \leqslant n - k + 1$, This scheme is displayed in Table I, where

$$L_i^0(f) = f(x_i), \qquad i = \nu, \nu + 1,..., \nu + k, \tag{10}$$

$$L_j^m(f) = \frac{L_{j+1}^{m-1}(\varphi_m) L_j^{m-1}(f) - L_j^{m-1}(\varphi_m) L_{j+1}^{m}(f)}{L_j^{m-1}(\varphi_m) + L_{j+1}^{m-1}(\varphi_m)},$$

$$m = 1,..., k, \qquad j = \nu,..., \nu + k - m. \tag{11}$$

In the next section, the values $L_j^m(f)$ for fixed m and $j = 1,..., n - m + 1$ play a key role in generalizing the Theorems of de La Vallée Poussin and

TABLE I

$L_\nu^0(f)$				
$L_{\nu+1}^0(f)$	$L_\nu^1(f)$			
$L_{\nu+2}^0(f)$	$L_{\nu+1}^1(f)$	$L_\nu^2(f)$		
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
$L_{\nu+k}^0(f)$	$L_{\nu+k-1}^1(f)$	$L_{\nu+k-2}^2(f)$	$\cdots$	$L_\nu^k(f)$

Remes. With this in mind, we would like to discuss the actual computation of $L_\nu^k(f)$ in some more detail. In an actual computation one must compute and store the values $L_j^r(\varphi_\nu)$ for $\nu = 1, 2,..., k$, $r = 0, 1,..., \nu - 1$ and $j = \nu,..., \nu + k - r$, in addition to the values $L_j^0(f)$, $j = \nu,..., \nu + k$ in order to calculate $L_\nu^k(f)$. Thus, instead of Table I we possibly should have written

TABLE II

			$L_\nu^0(\varphi_1)$	$L_\nu^0(f)$		
		$L_\nu^1(\varphi_2)$	$L_{\nu+1}^0(\varphi_1)$	$L_{\nu+1}^0(f)$	$L_\nu^1(f)$	
	$\cdot$	$L_{\nu+1}^1(\varphi_2)$	$L_{\nu+2}^0(\varphi_1)$	$L_{\nu+2}^0(f)$	$L_{\nu+1}^1(f)$	
$L_\nu^{k-1}(\varphi_k)$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$
$L_{\nu+1}^{k-1}(\varphi_k)$	$\cdots\cdots$	$L_{\nu+k-1}^1(\varphi_2)$	$L_{\nu+k}^0(\varphi_1)$	$L_{\nu+k}^0(f)$	$L_{\nu+k-1}^1(f)$ $\cdots$	$L_\nu^k(f)$

The above procedure can be interpreted in terms of the process of Gaussian elimination. Indeed, consider the following system of linear equations

$$\sum_{\nu=1}^{n} \alpha_\nu \varphi_\nu(x_\mu) + (-1)^\mu \lambda = f(x_\mu), \qquad \mu = 1,..., n + 1,$$

in the unknowns $\alpha_1,..., \alpha_n, \lambda$. If one applies Gaussian elimination (no pivoting) with the constraint that the coefficient of λ is $(-1)^\mu$ in the μth row in each step, then, after $(k - 1)$ steps, the last $n - k + 1$ rows are

$$\sum_{\nu=k}^{n} \alpha_\nu L_\mu^{k-1}(\varphi_\nu) + (-1)^\mu \lambda = L_\nu^{k-1}(f), \qquad \mu = 1,..., n - k + 1.$$

Before proceeding to our desired theorem, we wish to relate the above table with the notion of generalized divided differences with respect to a Haar system. In [1], the kth divided difference of f at $x_j, \ldots, x_{j+k}$ with respect to the Haar subspaces $V_k = \langle \varphi_1, \ldots, \varphi_k \rangle$ and $V_{k+1} = \langle \varphi_1, \ldots, \varphi_k, \varphi_{k+1} \rangle$ is defined by

$$\Delta(f, x_j, \ldots, x_{j+k}) = \begin{vmatrix} \varphi_1(x_j) & \cdots & \varphi_k(x_j) & f(x_j) \\ \vdots & & \vdots & \vdots \\ \varphi_1(x_{j+k}) & \cdots & \varphi_k(x_{j+k}) & f(x_{j+k}) \end{vmatrix} \div \begin{vmatrix} \varphi_1(x_j) & \cdots & \varphi_{k+1}(x_j) \\ \vdots & & \vdots \\ \varphi_1(x_{j+k}) & \cdots & \varphi_{k+1}(x_{j+k}) \end{vmatrix}. \tag{12}$$

Observe that the kth divided difference (12) is simply a linear functional, Δ, based on the points $x_j, \ldots, x_{j+k}$, annihilating $V_k = \langle \varphi_1, \ldots, \varphi_k \rangle$ and normalized by the requirement that $\Delta(\varphi_{k+1}) = 1$. The assumption that V_{k+1} is a Haar subspace implies that Δ is uniquely determined.

Now, suppose that $V_k = \langle \varphi_1, \ldots, \varphi_k \rangle$ is a Haar subspace of $C[a, b]$ for $k = 1, \ldots, n$. Because of the uniqueness of Δ it is easily shown that

$$\Delta(f, x_\nu, \ldots, x_{\nu+k}) = \frac{L_\nu^k(f)}{L_\nu^k(\varphi_{k+1})} \tag{13}$$

for $k = 1, 2, \ldots, n - 1$. In particular, with the formulas

$$\Delta(f, x_\nu) = \frac{f(x_\nu)}{\varphi_1(x_\nu)}, \qquad \nu = 1, \ldots, n + 1 \tag{14}$$

and

$$\Delta(f, x_\nu, \ldots, x_{\nu+k}) = \frac{\Delta(f, x_{\nu+1}, \ldots, x_{\nu+k}) - \Delta(f, x_\nu, \ldots, x_{\nu+k-1})}{\Delta(\varphi_{k+1}, x_{\nu+1}, \ldots, x_{\nu+k}) - \Delta(\varphi_{k+1}, x_\nu, \ldots, x_{\nu+k-1})},$$
$$\nu = 1, \ldots, n, \qquad k = 1, \ldots, n - \nu + 1; \tag{15}$$

one can construct a generalized divided difference table with respect to given points and a given Markoff system in precisely the same manner that the standard divided difference table is constructed. For the special case that $\varphi_i(x) = x^i$, this is the standard divided difference table and in this case, one has that $\Delta(\varphi_{k+1}, x_\nu, \ldots, x_{\nu+k-1}) = x_\nu + \cdots + x_{\nu+k-1}$, so that it is not necessary to calculate the differences occurring in the denomination of (15). This, incidentally, reduces the operation count of multiplications and divisions from $O(n^3)$ for the general case to $O(n^2)$ for this special case. In a future paper, we intend to discuss the use of these general divided differences for interpolation.

4. Main Theorem

Now, we turn to proving the desired lower estimate. This shall be done using the decomposition theorem on L_1^n,

$$L_1^n(f) = \sum_{j=1}^{n-m+1} \lambda_{jm}^{1n} L_j^m(f),$$

where m is a fixed integer satisfying $0 \leqslant m \leqslant n$. In order that the results of Theorem 3 apply, it is only necessary to assume that $V_r = \langle \varphi_1, ..., \varphi_r \rangle$ is a Haar subspace of $C[a, b]$ for $r = 1, ..., m$ and n.

Theorem 4. *Let $f \in C[a, b]$, $h \in V_n$ and suppose that V_r is a Haar subspace of $C[a, b]$ for $r = 1, ..., m$ and n, where $0 \leqslant m < n$. If there exists a set of $n + 1$ points, $a \leqslant x_1 < x_2 < \cdots < x_{n+1} \leqslant b$, such that the error function $e(x) = f(x) - h(x)$ satisfies*

1. $L_j^m(e) \neq 0, \quad j = 1, ..., n - m + 1,$
2. $\operatorname{sgn} L_j^m(e) = -\operatorname{sgn} L_{j+1}^m(e), \quad j = 1, ..., n - m,$

where the linear functionals L_j^m, $j = 1, ..., n - m + 1$ are based on the points $x_j, ..., x_{j+m}$. Then

$$\min_{1 \leqslant j \leqslant n-m+1} | L_j^m(e)| \leqslant \rho_n(f) \equiv \inf_{p \in V_n} \|f - p\|.$$

Proof. It is known that $| L_1^n(f)| \leqslant \rho_n(f)$. Thus,

$$\begin{aligned} \rho_n(f) &\geqslant | L_1^n(f)| = | L_1^n(f - h)|, \\ &= \left| \sum_{j=1}^{n-m+1} \lambda_{jm}^{1n} L_j^m(e) \right|, \\ &= \sum_{j=1}^{n-m+1} | \lambda_{jm}^{1n} | \, | L_j^m(e)|, \\ &\geqslant \min_{1 \leqslant j \leqslant n-m+1} | L_j^m(e)|. \quad \blacksquare \end{aligned}$$

Corollary 1. *Suppose $\varphi_1, ..., \varphi_n$ form a Markoff system in $C[a, b]$, $f \in C[a, b]$ and $h \in V_n$. If there exists a set of $n + 1$ points, $a \leqslant x_1 < x_2 < \cdots < x_{n+1} \leqslant b$, such that the error function $e(x) = f(x) - h(x)$ satisfies*

1. $e(x_i) \neq 0, \quad i = 1, ..., n + 1,$
2. $\operatorname{sgn} e(x_i) = -\operatorname{sgn} e(x_{i+1}), \quad i = 1, ..., n.$

Then

$$\min_{1 \leq j \leq n+1} |e(x_j)| \leq \min_{1 \leq j \leq n} |L_j^1(e)| \leq \cdots \leq |L_n^1(e)| \leq \rho_n(f).$$

This is easily proved with repeated applications of the decomposition theorem.

Observe that for the special case of $\varphi_\nu(x) = x^{\nu-1}$, $\nu = 1,..., n$, and $m = 1$, Theorem 4 is precisely the Remes estimate. Also, Theorem 4 is weaker than the de La Vallée Poussin estimate for $\rho_n(f)$ ($m = 0$ case), since one only need assume that V_n is a Haar subspace for this result.

5. The Polynomial Case

Theorem 4 is even new in the case that $\varphi_\nu(x) = x^{\nu-1}$, $\nu = 1,..., n$. Therefore, it may be of interest to outline briefly a second proof of the decomposition theorem for this case. This proof uses Cauchy's integral formula and is the method first used in this study.

Thus, let A be a region in the complex plane containing the closed interval $[a, b]$. Let f be holomorphic in A and real on $[a, b]$ and let C be a simple closed rectifiable path in A containing $[a, b]$ in its interior. Integrating in the positive direction, set

$$\Gamma_\nu^k(f) = \frac{C_\nu^k}{2\pi i} \int_C \frac{f(z)\,dz}{\omega_{\nu k}(z)},$$

where $a \leq x_\nu < x_{\nu+1} < \cdots < x_{\nu+k} \leq b$,

$$C_\nu^k = \left(\sum_{j=\nu}^{\nu+k} \frac{(-1)^j}{\omega'_{\nu k}(x_j)} \right)^{-1} (-1)^\nu,$$

$$\omega_{\nu k}(z) = (z - x_\nu) \cdots (z - x_{\nu+k}).$$

Clearly, Γ_ν^k is a linear functional on $A[a, b]$, the linear space of functions holomorphic in A and real on $[a, b]$, which annihilates π_{n-1}. Using the residue theorem, one gets that

$$\Gamma_\nu^k(f) = C_\nu^k \sum_{j=\nu}^{\nu+k} \frac{f(x_j)}{\omega'_{\nu k}(x_j)}.$$

This relation can be considered to be a continuation of Γ_ν^k to $C[a, b]$.

To prove the decomposition theorem for functions in $A[a, b]$, one must prove first a somewhat unusual partial fraction decomposition. Namely,

Lemma 2. *Let r be a nonnegative integer, $r \leq k$. Then, there exists a unique partial fraction decomposition*

$$\frac{1}{\omega_{\nu k}(z)} = \sum_{j=\nu}^{\nu+k-r} \frac{d_{jr}^{\nu k}}{\omega_{jr}(z)}, \tag{16}$$

where the (real) numbers $d_{jr}^{\nu k}$ are all different from zero and

$$\operatorname{sgn} d_{jr}^{\nu k} = (-1)^{j+\nu+r+k}, \qquad j = \nu,\dots, \nu + k - r. \tag{17}$$

Proof. Multiplying (16) by $\omega_{\nu k}(z)$ and comparing the coefficients of the powers of z leads to an inhomogeneous system of $k - r + 1$ linear equations for the $k - r + 1$ unknowns $d_{jr}^{\nu k}$. The corresponding homogeneous system is equivalent to the decomposition of the zero function. It is easily seen that this system has only the trivial solution. Therefore, the numbers $d_{jr}^{\nu k}$ are uniquely determined. For $r = k - 1$ we have

$$\frac{1}{\omega_{\nu k}(z)} = \frac{d_{\nu,k-1}^{\nu k}}{\omega_{\nu,k-1}(z)} + \frac{d_{\nu+1,k-1}^{\nu k}}{\omega_{\nu+1,k}(z)}.$$

Thus, $d_{\nu,k-1}^{\nu k} < 0$ and $d_{\nu+1,k-1}^{\nu k} > 0$, which corresponds to (17). Induction completes the argument.

Multiplying (16) by $C_\nu^k f(z)$ and integrating gives Theorem 3 with

$$\Gamma_\nu^k = L_\nu^k \qquad \text{and} \qquad \lambda_{jr}^{\nu k} = \frac{C_\nu^k}{C_j^r} d_{jr}^{\nu k}.$$

6. A Numerical Example

Let $X = \{x_i : x_i = i/64,\ i = 0, 1,\dots, 64\}$, $f(x) = \tan x$, $\varphi_i(x) = x^{i-1}e^x$, $i = 1,\dots, 5$. We use the above techniques in conjuction with Remes multiple exchange for finding the best approximation to $f(x) = \tan x$ from $V = \langle e^x, xe^x,\dots, x^4e^x\rangle$ on $X = \{x_i : x_i = i/64,\ i = 0, 1,\dots, 64\}$. Taking x_9, x_{18}, x_{27}, x_{36}, x_{45}, and x_{54} as our initial guess, we find that

$$h_1(x) = 0.00277e^x + 0.96068xe^x - 0.80272x^2e^x + 0.37561x^3e^x + 0.03142x^4e^x$$

is the best approximation to f on this set from V with error 0.000074. Performing the multiple exchange gives new extreme points x_0, x_{14}, x_{26}, x_{39}, x_{50}, x_{64}, where $|f(x_{64}) - h_1(x_{64})| = \|f - h_1\|$. Applying our lower estimates to $f - h_1$ at these points, gives the table (use Table I).

TABLE III

−0.002774					
0.000140	−0.001601				
−0.000075	0.000111	−0.000875			
0.000094	−0.000084	0.000099	−0.000509		
−0.000280	0.000179	−0.000131	0.000114	−0.000315	
0.014042	−0.006412	0.002629	−0.001227	0.000607	0.000452

Thus, $0.000075 \leqslant 0.000084 \leqslant 0.000099 \leqslant 000114 \leqslant 0.000315 \leqslant 0.000452 \leqslant \text{dist}(f, V) \leqslant 0.01402$. Continuing, we get after the second exchange, that $0.00045 \leqslant 0.00049 \leqslant 0.00061 \leqslant 0.00066 \leqslant 0.00069 \leqslant 0.00094 \leqslant \text{dist}(f, V) \leqslant 0.0027$; after the third exchange, that $0.0094 \leqslant 0.00094 \leqslant 0.00094 \leqslant 0.00095 \leqslant 0.000978 \leqslant 0.001005 \leqslant \text{dist}(f, V) \leqslant 0.001250$, showing that we now are within 0.000245 of the error of approximation with h_3 (a relative error of less than 21 %). At the end of the fourth exchange, we find that $0.00010059 \leqslant 0.00010059 \leqslant 0.00010066 \leqslant 0.00010076 \leqslant 0.00010087 \leqslant 0.00010091 \leqslant \text{dist}(f, V) \leqslant 0.00010192$, so that we are now within 0.000001 of the error of approximation with h_4 (a relative error of less than 1 %). The Remes algorithm terminated after the fifth exchange.

References

1. S. Karlin and W. J. Studden, "Tchebycheff Systems with Applications to Analysis and Statistics," Interscience, New York, 1966.
2. G. Meinardus, "Approximation of Functions: Theory and Numerical Methods," Springer-Verlag, Berlin, 1967.