LOWER ESTIMATES FOR THE ERROR OF BEST UNIFORM APPROXIMATION

by

Günter Meinardus
G. D. Taylor

STAN-CS-73-376
July 1973

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

SU326 P30-28

Lower Estimates for the Error of Best Uniform Approximation

by

Günter Meinardus and G. D. Taylor[†,‡]

[†]This report was reproduced with funds from an AEC-NSF grant at Stanford University (AEC AT(04-3)-326 PA30). The second author wishes to thank Professor Gene Golub and the Computer Science Department of Stanford University for office space and secretarial assistance. We would like to thank Fran Brumbaugh for typing this manuscript.

[‡]Supported in part by AFSOR 72-2271.

Lower Estimates for the Error of Best Uniform Approximation

by

Günter Meinardus and G. D. Taylor

Abstract

In this paper the lower bounds of de La Vallée Poussin and Remes for the error of best uniform approximation from a linear subspace are generalized to give analogous estimates based on k points, $k = 1,\ldots,n$.

Introduction. In this paper we shall 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 $x_1 < \cdot \cdot < x_{n+2} \le b$ such that the error function $e(x) = f(x) - h(x)$ satisfies

1. $e(x_i) \neq 0$, $i = 1,\ldots,n+1$,

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

Then,

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

Theorem 2. (Remes) Let π_n, denote the set of all algebraic polynomials of degree $\le n-1$ and let $f \in C[a,b]$. Let $h \in \pi_{n-1}$ and suppose that there exist $n+1$ points $a \le x_1 < \cdots < x_{n+1} \le b$ such that the error function $e(x) = f(x) - h(x)$ satisfies

1. $e(x_i) \neq 0$, $i = 1,\ldots,n+1$

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

Then,

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

In what follows we shall generalize these results to give analogous estimates based on k points, $k = 1,\ldots,n$. For the special cases $k = 1,n$ our estimates will simply be the de La Vallée Poussin estimate and the error of approximation on the points $x_1,\ldots,x_{n+1}$, respectively. For the case $k = 2$, we will 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 $a \leq x_1 < x_2 \ldots < x_{n+1} \leq b$. For each k , $1 \leq k \leq n$ and ν, $1 \leq \nu \leq n-k+1$ define $M_{\nu k}$ by $M_{\nu k} = \{x_\nu, x_{\nu+1},\ldots,x_{\nu+k}\}$. Let $V_n = \langle\varphi_1,\ldots,\varphi_n\rangle$ be a fixed Haar subspace of $C[a,b]$ and for each j , $1 \leq j \leq n$, set $V_j = \langle\varphi_1,\ldots,\varphi_j\rangle$ (i.e., V_j is the subspace of $C[a,b]$ spanned by the functions $\varphi_1,\ldots,\varphi_j$). If V_k $(k = 1,\ldots,n)$ 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

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

where $A_j^{\nu k}$ satisfy $\lambda_\nu^{\nu k} > 0$, $\lambda_j^{\nu k} \neq 0$ for $\nu \leq j \leq \nu+k$, $\operatorname{sgn} \lambda_j^{\nu k} = (-1)^{j-\nu}$, $\sum_{j=\nu}^{\nu+k} |\lambda_j^{\nu k}| = 1$ and $\sum_{j=\nu}^{\nu+k} \lambda_j^{\nu k} \varphi_\mu(x_j) = 0$ for $\mu = 1,\ldots k$. 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

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

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

<u>Theorem 3</u>. Fix k, $1 \leq k \leq n$, r, $0 \leq r \leq k$ and ν, $1 \leq \nu \leq n-k+1$, and assume that V_j satisfies the Haar condition for $j = 1,\ldots,r$ 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, $j = \nu, \ldots, \nu+k-r$:

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

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

<u>Proof</u>. This theorem is valid for $r = 0$ by our remarks concerning the properties of Haar subspaces. Thus, we shall assume $r \geq 1$. 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 $\varphi, \varphi_1, \ldots, \varphi_k$ are linearly independent. Thus,

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

where $\alpha, \alpha_1, \ldots, \alpha_k$ are unique. We must show, since $L_{\nu}^{k}(\varphi) = 1$

and $L_\nu^k(\varphi_\mu) = 0$, $\mu = 1,\ldots k$, that there exist numbers $\lambda_{jr}^{\nu k}$, uniquely determined, which satisfy

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

$$\sum_{j=\nu}^{\nu+k-r} \lambda_{jr}^{\nu k} L_j^r(\varphi) = 1 .$$

Since, by definition of L_j^r ,

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

for $\mu = 1,\ldots,r$ 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

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

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

$$(6) \quad B^T a = b$$

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

$$(7) \quad L_j^r(\alpha\varphi + \sum_{i=r+1}^{k} \alpha_i\varphi_i) = b_j \ , \ j = \nu,\ldots,\nu+k-r \ ,$$

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

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

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

(9) $C\hat{\Psi} = b$

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

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

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

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

'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^T is not singular as it maps R^{k-r+1} onto R^{k-r+1}. From this follow 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 $\operatorname{sgn} \lambda_{jr}^{\nu k} = (-1)^{j+\nu}$, $j = \nu,\ldots,\nu+k-r$. Now if $r = k$, then clearly $\lambda_{\nu k}^{\nu k} = 1$. We shall 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, $0 < r \leq k$ 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 $\lambda_{\nu,r-1}^{\nu r} > 0$, since both $\lambda_\nu^{\nu r}$ and $\lambda_\nu^{\nu,r-1}$ are positive. Likewise, applying this expression to $g \in C[a,b]$ where $g(x_\mu) = \delta_{\nu+r\,\mu}$, gives

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

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

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

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

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

Uniqueness of the representation of L_ν^k in terms of L_j^{r-1} gives

$$\lambda^{\nu k}_{\nu,r-1} = \lambda^{\nu k}_{\nu r}\lambda^{\nu r}_{\nu,r-1} > 0 \ ,$$

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

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 $L^k_\nu(g) \neq 0$. 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

$$g(x_\mu) - h(x_\mu) = (-1)^{\mu+\nu}L^k_\nu(g) \ , \quad \mu = \nu,\ldots,\nu+k \ .$$

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

$$L^r_j(g-h) = \sum_{\mu=j}^{j+r} \lambda^{jr}_\mu (g(x_\mu) - h(x_\mu))$$

$$= L^k_\nu(g)(-1)^\nu \sum_{\mu=j}^{j+r} \lambda^{jr}_\mu (-1)^\mu$$

$$= (-1)^{j+\nu}L^k_\nu(g) \ .$$

Hence,

$$L^k_\nu(g) = (-1)^\nu L^k_\nu(g) \sum_{j=\nu}^{\nu+k-r} \lambda^{\nu k}_{jr} (-1)^j$$

or

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

as desired, completing the proof of the theorem. ∎

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

Lemma?. For each k, $1 \le k \le n$ and ν, $1 \le \nu \le n-k+2$, $L_\nu^{k-1}(\varphi_k) \neq 0$ and $\operatorname{sgn} L_\nu^{k-1}(\varphi_k) = \operatorname{sgn} L_{\nu+1}^{k-1}(\varphi_k)$, $v = 1,\ldots,n-k+1$.

Proof. This is clearly true for k = 1. For $k \ge 2$, $|L_\nu^{k-1}(\varphi_k)|$ equals the minimal deviation in approximating φ_k by V_{k-1} on the point set $M_{v,k-1}$. If this were zero, then there would exist $\tilde{\varphi} \in V_k$, , equal to φ_k at the k points of $M_{\nu,k}$! . Since $\varphi_k \notin V_k$, , the difference would then be a function in V_k having k zeros which is not identically zero, contradicting the Haar condition. To prove that $\operatorname{sgn} L_\nu^{k-1}(\varphi_k) = \operatorname{sgn} L_{\nu+1}^{k-1}(\varphi_k)$, 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 $\operatorname{sgn} L_\nu^{k-1}(\varphi_k) = -\operatorname{sgn} L_{\nu+1}^{k-1}(\varphi_k)$ (some ν) on which $L_\nu^{k-1}(\varphi_k) = 0$ holds. Thus, the above arguments preclude this occurring. ∎

Using these facts, we can give a recursive scheme for calculating $L_\nu^k(f)$, $f \in C[a,b]$, $7 \le k \le n$, $1 \le v \le n-k+1$. This scheme is displayed in Table 1 where

(10) $L_i^0(f) = f(x_i)$, $i = \nu,\nu+1,\ldots,\nu+k$

(11) $$L_j^m(f) = \frac{L_{j+1}^{m-1}(\varphi_m)\, L_j^{m-1}(f) - L_j^{m-1}(\varphi_m)\, L_{j+1}^{m}(f)}{L_j^{m-1}(\varphi_m) + L_{j+1}^{m-1}(\varphi_m)}, \quad m = 1,\ldots,k;\ j = \nu,\ldots,\nu+k-m.$$

$$
\begin{array}{lllll}
L_{\nu}^{0}(f) & & & & \\
L_{\nu+1}^{0}(f) & L_{\nu}^{1}(f) & & & \\
L_{\nu+2}^{0}(f) & L_{\nu+1}^{1}(f) & L_{\nu}^{2}(f) & & \\
\vdots & \vdots & \vdots & \ddots & \\
L_{\nu+k}^{0}(f) & L_{\nu+k-1}^{1}(f) & L_{\nu+k-2}^{2}(f) & \cdots & L_{\nu}^{k}(f)
\end{array}
$$

Table 1

In the next section, the values $L_j^m(f)$ for fixed m and $j = 1, \ldots, n-m+1$ play a key role in generalizing the Theorems of de La Vallée Poussin and 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 $v = 1,2,\ldots,k$, $r = 0,1,\ldots,\nu-1$ and $j = \nu,\ldots,\nu+k-r$, in addition to the values $L_j^0(f)$, $j = \nu,\ldots,\nu+k$ in order to calculate $L_\nu^k(f)$. Thus, instead of Table 1 we should have possibly written

$$
\begin{array}{llllllll}
 & & & L_{\nu}^{0}(\varphi_1) & L_{\nu}^{0}(f) & & & \\
 & & L_{\nu}^{1}(\varphi_2) & L_{\nu+1}^{0}(\varphi_1) & L_{\nu+1}^{0}(f) & L_{\nu}^{1}(f) & & \\
 & \cdot & L_{\nu+1}^{1}(\varphi_2) & L_{\nu+2}^{0}(\varphi_1) & L_{\nu+2}^{0}(f) & L_{\nu+1}^{1}(f) & & \\
 & \cdot & & \vdots & \vdots & \vdots & & \\
L_{\nu}^{k-1}(\varphi_k) & & & & & & & \\
L_{\nu+1}^{k-1}(\varphi_k) & \cdots\cdots & L_{\nu+k-1}^{1}(\varphi_2) & L_{\nu+k}^{0}(\varphi_1) & L_{\nu+k}^{0}(f) & L_{\nu+k-1}^{1}(f) & \cdots & L_{\nu}^{k}(f)
\end{array}
$$

Table 2

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

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

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

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

Before proceeding to our desired theorem, we would like to relate the above table with the notion of generalized divided differences with respect to a Haar system. In [1] the k-th divided difference of f at $x_j,\dots,x_{j+k}$ with respect to the Haar subspace $V_k = \langle \varphi_1,\dots,\varphi_k \rangle$ is defined by

$$(12) \quad \Delta(f,x_j,\dots,x_{j+k}) \equiv \begin{vmatrix} \varphi_1(x_j) & \varphi_{k-1}(x_j) \dots & f(x_j) \\ \vdots & \vdots & \vdots \\ \varphi_1(x_{j+k}) & \varphi_{k-1}(x_{j+k}) \dots & f(x_{j+k}) \end{vmatrix} \div \begin{vmatrix} \varphi_1(x_j) & \dots & \varphi_k(x_j) \\ \vdots & & \vdots \\ \varphi_1(x_{j+n}) & \dots & \varphi_k(x_{j+n}) \end{vmatrix}$$

Observe that the k-th divided difference (12) is simply a linear functional, Δ , based on the points $x_j,\dots,x_{j+k}$, annihilating $V_k = \langle \varphi_1, \dots +k \rangle$ 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 $V_k = \langle \varphi_1,\dots,\varphi_k \rangle$ is a Haar subspace of $C[a,b]$ for $k = 1,\dots,n$. Because of the uniqueness of Δ it is easily shown that

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

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

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

and

$$(15)\quad \Delta(f,x_\nu,\ldots x_{\nu+k}) \equiv \frac{\Delta(f,x_{\nu+1},\ldots,x_{\nu+k}) - \Delta(f,x_\nu,\ldots,x_{\nu+k-1})}{\Delta(\varphi_{k+1},x_{\nu+1},\ldots,x_{\nu+k}) - \Delta(\varphi_{k+1},x_\nu,\ldots,x_{\nu+k-1})} ,$$

$$\nu = 1,\ldots,n \ , \quad k = 1,\ldots,n-\nu+1 \ ;$$

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

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

1. $L_j^m(e) \neq 0$, $j = 1, \ldots, n-m+1$,

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

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

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

<u>Proof</u>. It is known that $|L_1^n(f)| \leq \rho_n(f)$. Thus,

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

<u>Corollary 1</u>. Suppose $\varphi_1, \ldots, \varphi_n$ 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, $a \leq x_1 < x_2 < \ldots < x_{n+1} < b$, such that the error function $e(x) = f(x) - h(x)$ satisfies

1. $e(x_i) \neq 0$, $i = 1,\dots,n+1$,

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

Then

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

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

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

5. <u>The Polynomial Case</u>. Theorem 4 is even new in the case that $\varphi_\nu(x) = x^{\nu-1}$, $\nu = 1,\dots,n$. Therefore, it may be of interest to briefly outline 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 $a \leq x_\nu < x_{\nu+1} < \dots < x_{\nu+k} \leq b$,

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

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

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, $r \le k$. Then, there exists a unique partial fraction decomposition

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

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

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

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

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

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

Multiplying (16) by $C^k_\nu f(z)$ and integrating, gives Theorem 3 with $r^k_\nu = L^k_\nu$ and $\lambda^{\nu k}_{jr} = \frac{C^k_\nu}{C^r_j} d^{\nu k}_{jr}$.

6. A Numerical Example. Let $X = \{x_i : x_i = \frac{i}{64}, \ i = 0,1,\dots,64\}$, $f(x) = \tan x$, $\varphi_i(x) = x^{i-1}e^x$, $i = 1,\dots,5$. We shall use the above techniques in conjunction 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 $X=(x_i : x_i = \frac{i}{64}, \ i = 0,1,\dots,64\}$. Taking $x_9, x_{18}, x_{27}, x_{36}, x_{45}$ and x_{54} as our initial guess, we find that $h_1(x) = .00277e^x + .96068xe^x - .80272x^2e^x + .37561x^3e^x + .03142x^4e^x$ is the best approximation to f on this set from V with error .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 (see Table 1):

-.002774					
.000140	-.001601				
-.000075	.000111	-.000875			
.000094	-.000084	.000099	-.000509		
-.000280	.000179	-.000131	.000114	-.000315	
.014042	-.006412	.002629	-.001227	.000607	.000452

Table 3

Thus, $.000075 \leq .000084 \leq .000099 \leq 000114 \leq .000315 \leq .000452 \leq \text{dist}(f,V) \leq$

.01402 . Continuing we get after the second exchange that .00045 ≤ .00049 ≤ .00061 ≤ .00069 ≤ .00094 ≤ dist (f,V) ≤ .0027 ; after the third exchange that .0094 ≤ .00094 ≤ .00094 ≤ .00097 ≤ .000978 ≤ .001005 ≤ dist (f,V) ≤ .001250 showing that we now are within .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 .00010059 ≤ .00010059 ≤ .00010066 ≤ .00010069 ≤ .00010087 ≤ dist (f,V) ≤ .00010192 so that we are now within .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. Karlin, S., and W. J. Studden, Tchebycheff Systems with Applications to Analysis and Statistics, Interscience, New York, 1966.

2. Meinardus, G., A roximation of Functions: Theory and Numerical Methods, Springer-Verlag, Berlin, 19 7. -

The University of Erlangen - Nürnberg
Erlangen, Germany

and

Stanford University
Stanford, California