THE REMEZ EXCHANGE ALGORITHM FOR APPROXIMATION WITH LINEAR RESTRICTIONS

BY

BRUCE L. CHALMERS

ABSTRACT. This paper demonstrates a Remez exchange algorithm applicable to approximation of real-valued continuous functions of a real variable by polynomials of degree smaller than n with various linear restrictions. As special cases are included the notion of restricted derivatives approximation (examples of which are monotone and convex approximation and restricted range approximation) and the notion of approximation with restrictions at poised Birkhoff data (examples of which are bounded coefficients approximation, ϵ-interpolator approximation, and polynomial approximation with restrictions at Hermite and "Ferguson-Atkinson-Sharma" data and pyramid matrix data). Furthermore the exchange procedure is completely simplified in all the cases of approximation with restrictions at poised Birkhoff data. Also results are obtained in the cases of general linear restrictions where the Haar condition prevails. In the other cases (e.g., monotone approximation) the exchange in general requires essentially a matrix inversion, although insight into the exchange is provided and partial alternation results are obtained which lead to simplifications.

1. **Introduction, preliminaries and examples.** This paper demonstrates a Remez exchange algorithm applicable to approximation of real-valued continuous functions of a real variable by polynomials of degree smaller than n with various linear restrictions. In [2] existence and uniqueness results were proved in this setting. As special cases are included the notion of restricted derivatives approximation (examples of which are monotone and convex approximation and restricted range approximation) and the notion of approximation with restrictions at poised Birkhoff data (examples of which are bounded coefficients approximation, ϵ-interpolator approximation, and polynomial approximation with restrictions at Hermite and "Ferguson-Atkinson-Sharma" data and pyramid matrix data).

The key to the uniqueness theory is a generalized Haar condition [2, p. 310]. A new condition called nearly Haar is introduced in the present paper and plays an important role in the development of a Remez exchange algorithm which is

shown to converge to the best approximation in this general setting (Theorems 3.1 and 3.2). Furthermore, the method of exchange can be completely simplified (see §4) in all the cases of approximation with restrictions at uniform rank n (see Definition A) Birkhoff data (e.g., bounded coefficients approximation, restricted range approximation with bounded coefficients and, more generally, approximation with pyramid restrictions) and/or with interpolatory restrictions at poised Birkhoff data. The exchange is as simple as the familiar Tchebycheff exchange in the unrestricted case (which is also the same as in the case of restricted range approximation [13]). Also, results are obtained in the cases of general linear restrictions where the Haar condition prevails (see §4). In the other cases (e.g., monotone approximation) the exchange in general requires essentially a matrix inversion. Even here, however, a complete insight into the exchange and why it does not simplify to the same extent as in the case of bounded coefficients approximation is explained (see §4). Nevertheless partial alternation results are obtained which lead to simplifications in the exchange.

In [5] P. J. Laurent proves the convergence of a Remez exchange algorithm for one-sided (see Note 5 of §3) restricted approximation assuming the usual Haar condition, which is not, however, satisfied, for example, in the case of restricted derivatives approximation (e.g., monotone and convex approximation). In [13] G. D. Taylor and M. J. Winter prove the convergence of a Remez algorithm for two-sided Restricted Range Approximation. By combining these two methods we obtain the convergence of a Remez algorithm for two-sided restricted approximation provided the ordinary Haar condition is satisfied (Theorem 3.1). Further, by use of the nearly Haar condition we can approximate the problem in the more general setting (e.g., restricted derivatives approximation) by a discrete problem in which the Haar condition holds (Theorem 3.2).

Laurent's results do apply to the case of discrete uniform rank n Birkhoff data (see Note 5 of §3) but the exchange still involves a matrix inversion. In §4 (Theorems 2, 3, 6, 7) it is seen, as mentioned above, that in this case (and even if the data is not discrete) the exchange can be completely simplified.

In the following the setting and background of our investigation is described in detail.

Let E denote a compact subset of the real line and let CE denote the real space of real-valued continuous functions on E with the usual supremum norm $\| \cdot \|_E$ or $\| \cdot \|$. Let V^n be an n-dimensional subspace of CE. For some fixed index space A, let $\{L^\alpha\}_{\alpha \in A}$ be a compact space (usual norm) of linear functionals defined on V^n, such that, for each p in V^n, $L^\alpha p$ is a continuous function on A, where A inherits the norm topology of $\{L^\alpha\}_{\alpha \in A}$. Now set

$$V_0^n = \{p \in V^n; l_\alpha \leqslant L^\alpha p \leqslant u_\alpha, \alpha \in A\},$$

where l_α and u_α are extended-real-valued functions on A and $l_\alpha < +\infty$, $u_\alpha > -\infty$, l_α and u_α are finite on closed sets relative to which they are continuous, and $l_\alpha \leqslant u_\alpha$. We assume, moreover, that $l_\alpha = u_\alpha$ implies α is an isolated point of A.

Let e_x represent point evaluation at x in E. Now consider any fixed f in $CE - V_0^n$ with the restriction that if $L^\alpha = e_x$ for some α in A and some x in E, then $l_\alpha \leqslant f(x) \leqslant u_\alpha$. We are concerned then with approximating f by elements of V_0^n.

We showed in [2] that, if V_0^n is not empty and V^n satisfies a generalized Haar condition, then there exists a unique best approximation p in V_0^n to f.

For the treatment in this paper we will need an additional hypothesis on V^n which is satisfied by all the examples in question.

DEFINITION A. If any kth order subset of $\{N_i\}_{i=1}^{n+1}$ has rank k, we will say $\{N_i\}_{i=1}^{n+1}$ has uniform rank k, i.e., $\Sigma_{j=1}^k \alpha_j N_{i_j} = 0$ implies all $\alpha_j = 0$. We will also say the $(n+1)$-tuple $(N_1, N_2, \ldots, N_{n+1})$ has uniform rank k.

DEFINITION B. V^n is called *nearly Haar* on $\Omega = \{L^\alpha\}_{\alpha \in A} \cup \{e_x\}_{x \in E}$ provided the set of n-tuples $(R_1, R_2, \ldots, R_n) \in \Omega^n$, where the R_i are linearly dependent, forms a closed nowhere dense subset of Ω^n (or alternatively the set of $(n+1)$-tuples $(R_1, R_2, \ldots, R_{n+1}) \in \Omega^{n+1}$ which do not have uniform rank n forms a closed nowhere dense subset of Ω^{n+1}).

DEFINITION. V^n is called Haar (on Ω) if any distinct n elements in Ω are linearly independent, i.e. any set of $n+1$ distinct elements in Ω has uniform rank n.

We now give several examples, the first two of which will be seen to be special cases of Example 3. Existence and uniqueness of best approximation in all these examples was shown in [2] as a consequence of the fact that a generalized Haar condition held. We emphasize that the above stated Haar condition emplies the generalized Haar condition and therefore uniqueness of best approximation holds from V_0^n whenever V^n is Haar (on Ω). For bibliographical references to previous work on these examples, see [2]. In the sequel P_{n-1} will represent the space of real polynomials of degree no larger than $n - 1$, i.e. $P_{n-1} = \{c_0 + c_1 x + \cdots + c_{n-1} x^{n-1}\}$.

EXAMPLE 1 (MONOTONE APPROXIMATION). Let $(\epsilon_0, \epsilon_1, \ldots, \epsilon_q)$ be a $(q+1)$-tuple with elements equal to ± 1 and let $\{k_i\}_{i=0}^q$ be a fixed set of $q + 1$ integers satisfying $0 < k_0 < k_1 < \cdots < k_q \leqslant n - 1$. "Monotone Approximation" in $C[a, b]$ by the set $N = \{p \in P_{n-1}; \epsilon_i p^{(k_i)}(x) \geqslant 0$ on $[a, b], i = 0, 1, \ldots, q\}$ has been studied by O. Shisha, G. G. Lorentz and K. L. Zeller, R. A. Lorentz, and J. A. Roulier.

EXAMPLE 2 (RESTRICTED RANGE APPROXIMATION). Approximation in $C[a, b]$ by the set $N = \{p \in P_{n-1}; l(x) \leqslant p(x) \leqslant u(x), x \in [a, b]\}$, where $l(x)$ and $u(x)$ are extended-real-valued functions, upper and lower semicontinuous

respectively, and $l(x) \leqslant f(x) \leqslant u(x)$ on $[a, b]$, has been investigated by G. D. Taylor and others. With the assumption that $l(x) < u(x)$ on $[a, b]$, uniqueness of best approximation holds for $l(x), u(x)$ continuous on $[a, b]$. In [13], a Remez algorithm was shown to converge.

These two examples are seen to be special cases of the following example.

EXAMPLE 3 (RESTRICTED DERIVATIVES APPROXIMATION). Let $\{k_i\}_{i=0}^q$ be a fixed set of $q + 1$ integers satisfying $0 \leqslant k_0 < k_1 < \cdots < k_q \leqslant n - 1$. Consider approximation in $C[a, b]$ by the set $N = \{p \in P_{n-1}; l_i(x) \leqslant p^{(k_i)}(x) \leqslant u_i(x), i = 0, 1, \ldots, q\}$, where $l_i(x) < +\infty$ and $u_i(x) > -\infty$ are finite on closed sets relative to which they are continuous, and $l_i(x) < u_i(x)$ on $[a, b]$, $i = 0, 1, \ldots, q$. Also assume that if $k_0 = 0$, then $l_0(x) \leqslant f(x) \leqslant u_0(x)$, and if $k_i > 0$, then $l_i(x)$ and $u_i(x)$ are either differentiable or identically infinite on $[a, b]$.

Restricted Derivatives Approximation fits into our scheme as follows. First let $E = [a, b]$ and let $A = \bigcup_{i=0}^q \{(x, i); x \in E\}$, endowed with the obvious topology. Let $V_n = P_{n-1}$. Then for each $\alpha = (x, i) \in A$, we have $L^\alpha p = p^{(k_i)}(x)$. Next we set $l_\alpha = l_i(x), u_\alpha = u_i(x)$ for all $\alpha = (x, i)$ in A, and we find that $N = V_0^n = \{p \in V^n; l_\alpha \leqslant L^\alpha p \leqslant u_\alpha, \alpha \in A\}$.

One can check that the remaining conditions hold. Surely $\{L^\alpha = e_x^{k_i}$, where $e_x^{k_i}$ denotes point evaluation at x of the k_ith derivative$\}$ is a compact set in the dual of V^n, $L^\alpha p = p^{(k_i)}(x)$ is a continuous function on A for each p in V^n, and if $L^\alpha = e_x$ for some α in A and some x in E, then $l_\alpha \leqslant f(x) \leqslant u_\alpha$.

The fact that V^n is nearly Haar follows immediately from well-known properties of $(n - 1)$st degree algebraic polynomials (see [4]).

EXAMPLE 4. Let E be any compact subset of the real line and let $V^n \subset CE$. Further consider s elements $L^1, L^2, \ldots, L^s$ in the dual of V^n such that for each choice of r functionals $L^{j_1}, L^{j_2}, \ldots, L^{j_r}, \{L^{j_1}, L^{j_2}, \ldots, L^{j_r}, e_{x_{r+1}}, e_{x_{r+2}}, \ldots, e_{x_n}\}$ is an independent set for every choice of $n - r$ points $\{x_{r+1}, x_{r+2}, \ldots, x_n\}, r \leqslant s$, provided $e_{x_j} \neq L^{j_i}, i = 1, 2, \ldots, r; j = r + 1, \ldots, n$. Let $V_0^n = \{p \in V^n; l_i \leqslant L^i p \leqslant u_i, i = 1, 2, \ldots, s\}$. V^n is Haar on Ω by definition.

This example fits into our scheme where $A = \{1, 2, \ldots, s\}$ with the discrete topology.

Applying Example 4, we have the following three specific examples:

EXAMPLE 5 (BOUNDED COEFFICIENTS APPROXIMATION). Let $\{k_i\}_{i=0}^q$ be a fixed set of $q + 1$ integers satisfying $0 \leqslant k_1 < k_1 < \cdots < k_q \leqslant n - 1$. Approximation in $C[0, 1]$ by the set $N = \{p \in P_{n-1}; l_i \leqslant p^{(k_i)}(0) \leqslant u_i, i = 0, 1, \ldots, q\}$ has been studied by J. A. Roulier and G. D. Taylor where they proved uniqueness of best approximation provided, in the case $k_0 = 0$, that $l_0 \leqslant f(0) \leqslant u_0$. (Letting $L^i p = p^{(k_i)}(0), i = 0, 1, \ldots, q$, we have the linear independence

required in Example 4 as a result of Theorem C in [1].)

EXAMPLE 6 (ϵ-INTERPOLATOR APPROXIMATION). Fix $x_1 < x_2 < \cdots <$ x_k in E with $k \leqslant n$ and let $\epsilon > 0$ be given. Let V^n be an n-dimensional Haar subspace of CE. Then p in V^n is said to ϵ-interpolate f at $x_1, x_2, \ldots, x_k$ provided $|f(x_i) - p(x_i)| \leqslant \epsilon$ for each x_i. G. D. Taylor proved uniqueness of best approximation of f by ϵ-interpolating elements in V^n.

EXAMPLE 7 (POLYNOMIAL APPROXIMATION WITH INTERPOLATION). Fix $x_1 < x_2 < \cdots < x_k$ in $[a, b]$ with $k \leqslant n$. Let $V^n = P_{n-1}$. Then let $K = \{p \in P_{n-1}; p^{(j)}(x_i) = a_{ij}, 1 \leqslant i \leqslant k \text{ and } 0 \leqslant j \leqslant m_i - 1\}$ where $\{a_{ij}\}$ is a set of real numbers and $\{m_i\}_{i=1}^k$ is a set of positive integers such that $\Sigma_{i=1}^k m_i < n$. Uniqueness of best approximation to $f \in \{f \in C[a, b]; f(x_i) = a_{i0}, 1 \leqslant i \leqslant k\}$ by elements of K was proved in [7]. (Letting $L^{ij}p = p^{(j)}(x_i)$, we have the linear independence required in Example 4 as a result of Theorem C in [2].) Also from Theorem D in [2], it is clear how we can generalize the set K of Example 7. Also in [7], a Remez exchange algorithm was shown to converge in the case of strictly Hermite data.

2. Characterization of best approximation.

DEFINITION. For p in V_0^n a set $S = \{L^\alpha\}_{\alpha \in I_1} \cup \{e_x\}_{x \in I_3}$ of symbols representing distinct elements in the dual of V^n (where $I_1 \subset A$, $I_3 \subset E$) is called an extremal set for f and p if

(i) $L^\alpha p = u_\alpha$ (or l_α), $\alpha \in I_1$,

(ii) $|(f - p)(x)| = \|f - p\|$, $x \in I_3$,

(iii) $e_x \notin \{L^\alpha\}_{\alpha \in I_1}$ if $|(f - p)(x)| = \|f - p\|$.

DEFINITION. Let $S' = \{L^\alpha\}_{\alpha \in I_1} \cup \{e_x\}_{x \in I_3}$ be the extremal set for f and p maximal with respect to I_1 and I_3 (ignore the auxiliary functions M^β in [1]).

Define a "signature" function σ on S' by

$$\sigma(e_x) = 1 \quad \text{if } e_x p = f(x) - \|f - p\|,$$

$$\sigma(e_x) = -1 \quad \text{if } e_x p = \|f - p\| + f(x),$$

$$\sigma(L^\alpha) = 1 \quad \text{if } L^\alpha p = l_\alpha,$$

$$\sigma(L^\alpha) = -1 \quad \text{if } L^\alpha p = u_\alpha.$$

That is, σ is 1 at lower extrema and -1 at upper extrema. Then define $S^\sigma = \{\sigma(R)R; R \in S'\}$.

The following characterization theorem is a restatement of Theorem 3.8.5 in [5] for the present setting (and holds, it should be emphasized, in the absence of any Haar conditions).

THEOREM 1. *p is a best approximation to f if and only if 0 is in the convex hull of some k ($\leqslant n + 1$) members of S^σ; i.e., $0 = \Sigma_{i=1}^{k} \lambda_i N_i$ where $N_i \in S^\sigma$, $\lambda_i > 0$, $i = 1, 2, \ldots, k \leqslant n + 1$.*

EXAMPLES. Examples of this theorem are found in many of the works on constrained approximation (e.g., in the case of Monotone Approximation see [8, p. 5]). See also [3, p. 73].

DEFINITION. Following Laurent [5], we extend the domain of f to $\{\pm L^\alpha\}_{\alpha \in A}$ by setting $L^\alpha f = l_\alpha$ and $(-L^\alpha)f = -u_\alpha$, introduce an auxiliary distance d of f from V_0^n by $d = \inf_{q \in V_0^n} \rho(f - q)$ where $\rho(h) = \sup\{|N(h)|\}$, where N ranges over $\{\pm e_x\}_{x \in E} \cup \{\pm L^\alpha\}_{\alpha \in A}$; i.e., $\rho(h) = \max\{\|h\|_E, \sup_{\alpha \in A}|(\pm L^\alpha)h|\}$.

NOTATION. $N = \pm e_x$ will indicate that $|N(f - p)| = \|f - p\|$. $N = \pm L^\alpha$ will indicate that $N \in S^\sigma$ but $N \neq \pm e_x$.

In the case where $k = n + 1$ and $\{N_i\}_{i=1}^{n+1}$ has uniform rank n, then in Theorem 1, the quantity d is related to the linear combination $l = \Sigma_{i=1}^{n+1} \lambda_i N_i$ (assuming $\Sigma_{N_i = \pm e_x} \lambda_i = 1$) via the following restatement of Theorem 3.9.3 in [5]:

THEOREM 2. *$l(f) = d = \|f - p\|$, where p is the best approximation to f, and p and d are determined by the $n + 1$ linear relations:*

$$N_i p = N_i f - d \quad \text{if } N_i = \pm e_x,$$
$$N_i p = N_i f \quad \text{if } N_i = \pm L^\alpha.$$

3. Convergence of a Remez exchange algorithm.

REMEZ ALGORITHM. Suppose at the νth stage we have the extremel set (of uniform rank n)

$$S_\nu^\sigma = \{N_i^\nu\}_{i=1}^{n+1} = \{\pm L_i^\nu\}_{i=1}^{k_\nu} \cup \{\pm e_{x_i^\nu}\}_{i=k_\nu+1}^{n+1}$$

for f and p_ν, the best approximation to f from

$$V_\nu^n = \{p \in V^n; l_i^\nu \leqslant L_i^\nu p \leqslant u_i^\nu, i = 1, 2, \ldots, k_\nu\}$$

with respect to $\| \ \|_{E^\nu}$ where $E^\nu = \{x_i^\nu\}_{i=k_\nu+1}^{n+1}$. Let $d_\nu = \|f - p_\nu\|_{E^\nu}$.

In order to construct the $(\nu + 1)$st stage, let N_*^ν be a functional yielding the maximum deviation of $f - p_\nu$ on $\{e_x\}_{x \in E} \cup \{\pm L^\alpha\}_{\alpha \in A}$; i.e., N_*^ν yields

$$\max\left\{\max_{N = \pm e_x} N(f - p_\nu) - d_\nu, \max_{N = \pm L^\alpha} N(f - p_\nu)\right\} = \epsilon_\nu;$$

then exchange one of $\{N_i^\nu\}_{i=1}^{n+1}$ for N_*^ν via the Exchange Procedure (see below) to obtain the set

$$S_{\nu+1}^\sigma = \{N_i^{\nu+1}\}_{i=1}^{n+1} = \{\pm L_i^{\nu+1}\}_{i=1}^{k_{\nu+1}} \cup \{\pm e_{x_i^{\nu+1}}\}_{i=k_{\nu+1}+1}^{n+1}.$$

Next calculate $p_{\nu+1}$ (and therefore also $d_{\nu+1}$) according to the prescription of Theorem 2.2.

NOTE 1. By Theorem 2.2 we have that in the above algorithm $p_{\nu+1}$ is the best approximation to f from

$$V_{\nu+1}^n = \{p \in V^n; l_i^{\nu+1} \leqslant L_i^{\nu+1}p \leqslant u_i^{\nu+1}, i = 1, 2, \ldots, k_{\nu+1}\}$$

with respect to $\| \ \|_{E^{\nu+1}}$ where $E^{\nu+1} = \{x_i^{\nu+1}\}_{i=k_{\nu+1}+1}^{n+1}$. (Also $d_{\nu+1} = \|f - p_{\nu+1}\|_{E^{\nu+1}}$.)

NOTE 2. For the 1st stage of the algorithm, let $k_1 = 0$ and $E^1 = \{x_i^1\}_{i=1}^{n+1}$ be a set where p_1 does not interpolate f. Then $d_1 > 0$.

EXCHANGE PROCEDURE (see Laurent [5, p. 147]). Suppose

(1) $0 = \Sigma_{i=1}^{n+1}\lambda_i N_i$, where $\lambda_i > 0$ and $\{N_i\}_{i=1}^{n+1}$ has uniform rank n. Given N_*, we can find an N_j to replace by N_* so that $0 = \Sigma_{i\neq j}\lambda_i^* N_i + \lambda_j^* N_*$ where $\lambda_i^* \geqslant 0$ and $\lambda_j^* > 0$ as follows:

$$N_* - N_{n+1} = \sum_{i=1}^{n} \alpha_i' N_i,$$

whence

(2) $N_* = \Sigma_{i=1}^{n+1}\alpha_i' N_i$ with at least one $\alpha_i' > 0$ (namely $\alpha_{n+1}' = 1$). Then multiply (2) by $\lambda > 0$ so that

(3) $\lambda N_* = \Sigma_{i=1}^{n+1}\alpha_i N_i$ where $\lambda_i \geqslant \alpha_i$ with equality holding for some i. Then subtracting (3) from (1) yields the result.

NOTE 3. Computationally the above exchange involves an $n \times n$ matrix inversion to determine $\{\alpha_i'\}_{i=1}^n$ and a linear search to determine $\lambda^{-1} = \max_i \alpha_i'/\lambda_i$.

NOTE 4. If V^n is Haar on Ω then $\lambda_i = \alpha_i$ for exactly one i.

NOTE 5. The Haar condition imposed in [5, p. 140] in the discussion of constrained approximation precludes both L^α and $-L^\alpha$ from being restraining functionals, since $L^\alpha p = 0$ implies $(-L^\alpha)p = 0$. Thus the imposition of both an upper and lower restriction on the same functional L^α, i.e. $l_\alpha \leqslant L^\alpha p$ and $-u_\alpha \leqslant (-L^\alpha)p$ is not allowable. This theory therefore does not yield as an example the Remez algorithm for ordinary restricted (both above and below) range approximation, obtained in [13]. ([5, p. 149, Proposition 3.9.6] does not hold if both L^α and $-L^\alpha$ occur because then $\lambda L^\alpha + \lambda(-L^\alpha) = 0$. The method of proof in [5] does extend, however, to upper and lower restrictions on L^α if L^α is isolated in the dual.)

The methods of [5] can, however, be combined with the methods of [13] to obtain a Remez algorithm in the general setting with upper and lower restrictions as we see in Theorem 1.

We need the following well-known result.

LEMMA 1. *Let D be a compact metric space with metric ρ. If M is an*

n-dimensional Haar subspace of $C(D)$, $A > 0$ is given and $S = \{(R_1, \ldots, R_n) \in D^n$; $\rho(R_i, R_j) \geq \delta > 0$ for $i \neq j\}$, then there exists $C > 0$ such that if $h \in V^n$ with $|h(R_i)| \leq A$ for some $(R_1, \ldots, R_n) \in S$ then $\max_{R \in D} |h(R)| \leq C$.

NOTATION. If ω belongs to the dual of V^n, then $\|\omega\| = \sup|\omega(h)|$, where h ranges over those functions in V^n of norm 1.

THEOREM 1. *If V^n is Haar on Ω, then the sequence $\{p_\nu\}$ generated by the Remez algorithm converges to the best approximation p.*

PROOF. Since $\|h\|_{E^\nu} \leq \|h\|_E$ and $V^n_\nu \supset V^n_0$, we have $d_\nu = \|f - p_\nu\|_{E^\nu} \leq \|f - p\|_{E^\nu} \leq \|f - p\|_E = d$. By Theorems 2.1 and 2.2 let $l^\nu = \Sigma_{i=1}^{n+1} \lambda_i^\nu N_i^\nu$ where $\lambda_i^\nu > 0$, $\Sigma_{N_i^\nu = \pm e_x} \lambda_i^\nu = 1$, and $l^\nu = 0$ on V^n. Then from Theorem 2.2 and the Exchange Procedure, we have

$$d_{\nu+1} = l^{\nu+1}f = l^{\nu+1}(f - p_\nu) = \sum_{i=1}^{n+1} \lambda_i^{\nu+1} N_i^{\nu+1}(f - p_\nu)$$

$$= \sum_{i \neq j_\nu} \lambda_i^{\nu+1} N_i^\nu(f - p_\nu) + \lambda_{j_\nu}^{\nu+1} N_*^\nu(f - p_\nu)$$

(4)

$$= \sum_{N_i^\nu = \pm e_x; i \neq j_\nu} \lambda_i^{\nu+1} d_\nu + \lambda_{j_\nu}^{j+1} N_*^\nu(f - p_\nu)$$

$$= d_\nu \left[\sum_{N_i^{\nu+1} = \pm e_x} \lambda_i^{\nu+1} \right] + \lambda_{j_\nu}^{\nu+1} [N_*^\nu(f - p_\nu) - \delta^\nu d_\nu]$$

where $\delta^\nu = 1$ if $N_*^\nu = \pm e_x$ and $\delta^\nu = 0$ if $N_*^\nu = \pm L^\alpha$. Hence

$$(5) \qquad d_{\nu+1} = d_\nu + \lambda_{j_\nu}^{\nu+1} \left\{ \max \left[\|f - p_\nu\|_E - d_\nu, \max_{N = \pm L^\alpha} N(f - p_\nu) \right] \right\}.$$

Thus the sequence d_ν is increasing and $d_\nu \leq d$.

Consider now the sequence of $(n + 1)$-tuples $N^\nu = (N_1^\nu, N_2^\nu, \ldots, N_{n+1}^\nu)$, $\nu = 1, 2, \ldots$. Since $\{\pm e_x\}_{x \in E} \cup \{\pm L^\alpha\}_{\alpha \in A}$ is a compact subset of the dual of V^n and contains all N_j^ν, any subsequence of N^ν has again a subsequence $N^\mu = (N_1^\mu, N_2^\mu, \ldots, N_{n+1}^\mu)$ which converges to $N = (N_1, N_2, \ldots, N_{n+1})$. (Here $\mu = \nu_k$.)

Now all $\{\sigma(N_i)N_i\}_{i=1}^{n+1}$ are distinct; for if not, then by the Haar condition we can find $q \in V^n$ such that

(i) if $N_i = \pm L^\alpha$ then $l_i < L_i q < u_i$ if L_i is not isolated in $\{L^\alpha\}_{\alpha \in A}$, otherwise $l_i = L_i q$, and

(ii) if $N_i = \pm e_x$ then $N_i q = N_i f$.

Thus from the continuity assumptions we would have that there exists a K such that $\mu \geqslant K$ implies $l_i^\mu \leqslant N_i^\mu q \leqslant u_i^\mu$ if $N_i = \pm L^\alpha$ and $|N_i^\mu q - N_i^\mu f| < d_1$ if $N_i = \pm e_x$. Hence q would be a better approximation to f from V^n on E^μ than p_μ since $\|f - p_\mu\|_{E^\mu} = d_\mu \geqslant d_1$.

Hence all $R_i^\mu = o(N_i^\mu)N_i^\mu$ are uniformly separated in the dual space of V^n, i.e., there is an $r > 0$ such that $\|R_i^\mu - R_j^\mu\| > r$ if $i \neq j$. Consider now the associated $(n + 1)$-tuples $\lambda^\mu = (\lambda_1^\mu, \lambda_2^\mu, \ldots, \lambda_{n+1}^\mu)$ and pick $h_\mu \in V^n$ such that $R_i^\mu(h_\mu) = o(N_i^\mu)$ for all $i \in \{1, \ldots, n + 1\} - \{i_0\}$, where i_0 is such that $N_{i_0}^\mu = \pm e_x$. Then by Lemma 1, there is a constant C such that $|N(h_\mu)| \leqslant C$ for all $N \in \Omega$. Hence since $\Sigma_{i=1}^{n+1} \lambda_i^\mu N_i^\mu = 0$ we have $\Sigma_{i \neq i_0} \lambda_i^\mu N_i^\mu(h_\mu) = \lambda_{i_0}^\mu N_{i_0}^\mu(h_\mu)$ where $\lambda_{i_0}^\mu \leqslant 1$. Thus $\Sigma_{i \neq i_0} \lambda_i^\mu \leqslant C$, whence $\Sigma_{i=1}^{n+1} \lambda_i^\mu \leqslant C + 1$. We can therefore assume by taking subsequences if necessary that λ^μ converges to $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{n+1})$. Thus $l^\mu = \lambda^\mu \cdot N^\mu = 0$ on V^n implies that $l = \lambda \cdot N = 0$ on V^n. But $\lambda \neq 0$ since then $d^\mu = l^\mu(f)$ would converge to 0, while $d^\mu \geqslant d_1 > 0$.

Therefore, since $0 = \Sigma_{i=1}^{n+1} \lambda_i N_i$, where $\Sigma \lambda_i > 0$ and $\{o(N_i)N_i\}_{i=1}^{n+1}$ are all distinct, we have that all $\lambda_i > 0$, $i = 1, 2, \ldots, n + 1$, again by the Haar assumption. Thus if $s = \inf \lambda_{j_\mu}^{\mu+1}$, then $s > 0$, and we conclude that $\overline{\lim}_{\mu \to \infty} \|f - p_\mu\|_E - d_\mu \leqslant 0$ and $\overline{\lim}_{\mu \to \infty} \max_{N = \pm L^\alpha} N(f - p_\mu) \leqslant 0$. Now d_μ increases to $d' \leqslant d$ and so $\|p_\mu\|_E \leqslant \|f\| + d'$. Thus p_μ has a subsequence which converges uniformly to p_* which satisfies $\|f - p_*\|_E \leqslant d' \leqslant d$ and $l_\alpha \leqslant L^\alpha p_* \leqslant u_\alpha$ for all $\alpha \in A$. Hence $p_* = p$ the best approximation to f and $d' = d$.

Since therefore any subsequence of the p_ν has again a subsequence which converges uniformly to p, so does p_ν. $\square$

NOTE 6. The argument in the proof of Theorem 1 extends to show that in fact all $R_i^\nu = o(N_i^\nu)N_i^\nu$ are uniformly separated and thus that all λ_i^ν are uniformly bounded, not just for the subsequence $\mu = \nu_k$.

DEFINITION. If Ω_r is a finite subset of Ω such that

$$\sup_{\omega_1 \in \Omega_r} \inf_{\omega_2 \in \Omega_r} \|\omega_1 - \omega_2\| < r,$$

then Ω_r is called a discretization of Ω with mesh size r. If any n arbitrary distinct elements of Ω_r are linearly independent over V^n, then Ω_r is called nonsingular.

NOTE 7. If V^n is nearly Haar, the probability that a random Ω_r is singular is 0. By a modification and generalization of the argument in [6] (see also [3, p. 84 ff]) we obtain the following result in case $V^n = P_{n-1}$.

THEOREM 2. *If* P_{n-1} *is nearly Haar on* Ω *then the sequence* p_ν^r *generated by the Remez algorithm on any nonsingular discretization* Ω_r *converges to the best approximation* p^r. *Furthermore if the sequence* r_m *tends to zero,* p^{r_m} *con-*

verges uniformly to the best approximation p.

PROOF. P_{n-1} is Haar on each Ω_r and so p_ν^r converges to p^r by Theorem 1.

We claim first that the $p_m = p^{r_m}$ are uniformly bounded on E. Let $E_m = \{x; e_x \in \Omega_{r_m}\}$. Then $\|f(x) - p_m(x)\|_{E_m} \leqslant \|f(x)\|_{E_m} \leqslant \|f\| = \|f\|$. Thus $\|p_m(x)\|_{E_m} \leqslant 2\|f\|$ for all m. If $[a, b]$ is the convex hull of E let $I_1, \ldots, I_n$ be n disjoint subintervals of $[a, b]$ with spacing at least ϵ between any two of them. Then since r_m tends to 0, there exists M such that for all $m \geqslant M$ there are points $x_1, \ldots, x_n$ of E_m satisfying $x_j \in I_j$, $j = 1, 2, \ldots, n$. Using the Lagrange interpolating polynomial form we can write

$$p_m(x) = \sum_{k=1}^n p_m(x_k) \prod_{j \neq k} \frac{x - x_j}{x_k - x_j}.$$

Thus

$$|p_m(x)| \leqslant \sum_{k=1}^n |p_m(x_k)| \prod_{j \neq k} \left| \frac{x - x_j}{x_k - x_j} \right| \leqslant 2n\|f\| \left[\frac{b - a}{\epsilon} \right]^{n-1} = R$$

independent of m and x. Hence $\{p_m\}$ is uniformly bounded.

Thus there exists a uniformly convergent subsequence $\{p_{m_k}\}$ with limit p_0. We claim $p_0 \in V_0^n$. If not, then there is an element $L^\beta \in \Omega$ such that $L^\beta p_0 - l_\beta = -\epsilon < 0$ or $L^\beta p_0 - u_\beta = \epsilon > 0$. Suppose the first inequality holds. Then $|L^\beta p_{m_k} - L^\beta p_0| \leqslant \|L^\beta\| \cdot \|p_{m_k} - p_0\|$ which tends to zero. On the other hand, $l_{\alpha_k} \leqslant L^{\alpha_k} p_{m_k}$ for all $L^{\alpha_k} \in \Omega_s$ where $s = r_{m_k}$. Further, $\exists K$ such that $k \geqslant K$ implies the existence of L^{α_k} such that

$$|L^{\alpha_k} p_{m_k} - L^\beta p_0| \leqslant |L^{\alpha_k} p_{m_k} - L^{\alpha_k} p_0| + |L^{\alpha_k} p_0 - L^\beta p_0|$$
$$\leqslant \|L^{\alpha_k}\| \|p_{m_k} - p_0\| + \|L^{\alpha_k} - L^\beta\| \|p_0\| < \epsilon/4$$

while $|l_\beta - l_{\alpha_k}| < \epsilon/4$. Hence

$$L^\beta p_0 - l_\beta = L^\beta p_0 - L^{\alpha_k} p_{m_k} + L^{\alpha_k} p_{m_k} - l_{\alpha_k} + l_{\alpha_k} - l_\beta$$
$$\geqslant - |L^\beta p_0 - L^{\alpha_k} p_{m_k}| - |l_{\alpha_k} - l_\beta| > -\epsilon/2,$$

a contradiction. The case $L^\beta p_0 - u_\beta = \epsilon > 0$ is contradicted analogously. Hence $p_0 \in V_0^n$.

We now show $\|f - p_0\| = \|f - p\|$. Surely $p_0 \in V_0^n$ implies $\|f - p_0\| \geqslant \|f - p\|$. Assume $\|f - p_0\| - \|f - p\| = \epsilon > 0$. Let $y \in E$ where $|f(y) - p_0(y)| = \|f - p_0\|$. Then

$$\|f - p_0\| = |f(y) - p_0(y)|$$

$$\leqslant |f(y) - f(x)| + |f(x) - p_{m_k}(x)| + |p_{m_k}(x) - p_{m_k}(y)|$$

$$+ |p_{m_k}(y) - p_0(y)|$$

where $x \in E_m$ and $|x - y| \leqslant r_{m_k}$. Now

$$|p_{m_k}(x) - p_{m_k}(y)| \leqslant \max_{z \in [a,b]} |p'_{m_k}(z)| \cdot |x - y|$$

$$\leqslant 2(n - 1)^2/(b - a)\|p_{m_k}\|(b - a)$$

by Markov's inequality. But $\|p_{m_k}\|$ is uniformly bounded. Hence since also f is continuous on E and p_{m_k} converges uniformly to p_0, we have that for k large enough

$$\|f - p_0\| \leqslant \max_{x \in E_{m_k}} |f(x) - p_{m_k}(x)| + \epsilon/2.$$

But then $\|f - p_0\| = \|f - p\| + \epsilon$ implies p is a better approximation than p_{m_k} on E_{m_k}, a contradiction. Hence $\|f - p_0\| = \|f - p\|$ which implies $p_0 = p$ by the uniqueness of best approximation. Thus every subsequence of p_m has a subsequence converging to p and so $p_m = p'm$ converges to p. $\square$

NOTE 8. In practice a discretization Ω_r may not be checked beforehand to be nonsingular. And in fact if the elements of Ω_r are chosen in a symmetric fashion which corresponds to a symmetry in the function f, the likelihood that Ω_r is singular increases.

Ω_r may, however, be adjusted as the Remez algorithm proceeds so that it may become nonsingular at least with respect to the sets of elements in the $(n + 1)$-tuples appearing as extremal sets S_r^σ in the algorithm. This may be done as follows.

After the Exchange Procedure, if $S_{\nu+1}^\sigma$ does not have uniform rank n (i.e., if $0 = \Sigma_{i=1}^{n+1}\lambda_i^{\nu+1}N_i^{\nu+1}$ where $\Sigma\lambda_i^{\nu+1} > 0$ but $\lambda_j^{\nu+1} = 0$ for some j), then replace N_*^ν by another element N_{**}^ν in the dual of V^n so that the adjusted $S_{\nu+1}^\sigma$ has uniform rank n, the new Ω_r still has mesh size r. This can be done since V^n is nearly Haar. Then reconstruct the $(\nu + 1)$st stage of the Remez algorithm using the new adjusted Ω_r. The probability is 1 that only a finite number of such adjustments need be made, and if so the proof of Theorem 1 follows exactly as before.

It is easily seen that the result in [6, Theorem 4] proved for convex approximation goes over to the case of Restricted Derivatives Approximation (R.D.A.).

THEOREM 3. *In the case of R.D.A. there exist constants C and D independent of m such that*

$$\|f - p\| - \|f - p^{r_m}\| \leqslant C \cdot r_m^2$$

and

$$\|f - p^{r_m}\| - \|f - p\| \leqslant \omega(f; r_m) + D \cdot r_m$$

where $\omega(f; \sigma)$ is the modulus of continuity function of f.

THEOREM 4. *In the case of one-sided R.D.A. (i.e.,*

$$N = \{p \in P_{n-1}; c_i(x) \leqslant \epsilon_i p^{(k_i)}(x), \epsilon_i = \pm 1, i = 0, 1, \ldots, q\}),$$

then the $\lambda_i^\nu m$ occurring in the Remez algorithm are uniformly bounded independent of the discretization Ω_{r_m}.

PROOF. If not, choose a subsequence $\lambda^\nu m = (\lambda_1^\nu m, \ldots, \lambda_{n+1}^\nu m)$ such that $\overline{\lim} \lambda_i^\nu m = \infty$ for some i and choose j_0 so that $\overline{\lim}_{m \to \infty}(\lambda_i^\nu m / \lambda_{j_0}^\nu m) \leqslant 1, 1 \leqslant i \leqslant n + 1$. By taking subsequences if necessary we can replace $\overline{\lim}$ by lim and assume that $N^\nu m = (N_1^\nu m, \ldots, N_{n+1}^\nu m)$ converges to $N = (N_1, \ldots, N_{n+1})$. Then letting $m \to \infty$ in $0 = \Sigma_{i=1}^{n+1}(\lambda_i^\nu m / \lambda_{j_0}^\nu n)N_i^\nu m$ yields $0 = \Sigma_{i=1}^{n+1}\lambda_i'N_i$ where $\lambda_i' = 0$ if $N_i = \pm e_x$ since $\lambda_i^\nu m \leqslant 1$ if $N_i^\nu m = \pm e_x$. Now let l be the smallest superscript in $\{N_i = e_i^l; \lambda_i' \neq 0\}$. Now since for m large at least $l + 1$ of the $N_i^\nu m = \pm e_x$ ($\{N_i^\nu m\}_{i=1}^{n+1}$ has uniform rank n), at least $l + 1$ of the $N_i = \pm e_x$. Hence since we have one-sided approximation, $\Sigma\lambda_i'N_i = \Sigma\lambda_i'e_{x_i}^l = 0$ (where $\epsilon = \pm 1$) on P_l. But if the degree of p is l, then $e_{x_i}^l p = c \neq 0$. Hence $\epsilon c\Sigma\lambda_i' = 0$ is a contradiction. □

4. **Simplifying the exchange.** Suppose $\{R_m\}_{m=1}^{n+1}$ is an arbitrary subset of $\Omega = \{L^\alpha\}_{\alpha \in A} \cup \{e_x\}_{x \in E}$ of uniform rank n and we have a priori knowledge about the signs of the coefficients α_m in

$$(1) \qquad 0 = \sum_{m=1}^{n+1} \alpha_m R_m \quad \text{where} \sum |\alpha_m| > 0.$$

Then the Exchange Procedure can be simplified, since the problem is to replace some $R_{m_0}^\nu$ by R_*^ν so that the signs of the ocefficients in (1) are preserved; i.e., if $R_m^{\nu+1} = R_m^\nu$ ($m \neq m_0$) and $R_{m_0}^{\nu+1} = R_*^\nu$, then in the dependency

$$0 = \sum_{i=1}^{n+1} \alpha_m^{\nu+1}R_m^{\nu+1}\text{sign}(\alpha_m^{\nu+1}) = \text{sign}(\alpha_m^\nu), \quad m = 1, 2, \ldots, n + 1.$$

If, in particular, the signs of the coefficients in (1) are known for every choice of $\{R_m\}_{m=1}^{n+1} \subset \Omega$ then the matrix inversion and linear search of the Exchange Procedure can be eliminated.

NOTATION. We will sometimes refer to sign α_m in (1) as sign R_m.

DEFINITION. Fix $\{x_1, x_2, \ldots, x_k\} \subset [a, b]$. Let L_i^j denote the linear functional on P_{n-1} defined by $L_i^j p = p^{(j)}(x_i)$. Following Schoenberg, let

$$E = (e_{ij})_{i=1,2,\ldots,k}^{j=0,1,\ldots,n-1}$$

be an m-incidence matrix, i.e. each e_{ij} is 0 or 1 and $\Sigma_{(i,j)} e_{ij} = m$. If E is an m-incidence matrix, let $m_j = \Sigma_{i=1}^k e_{ij}$, $j = 0, 1, \ldots, n-1$, and $M_j = \Sigma_{p=0}^j m_p$, $j = 0, 1, \ldots, n-1$. Then E is said to satisfy the Pólya conditions if $M_j \geqslant j + 1$ for $j = 0, 1, \ldots, m-1$, and the strong Pólya conditions if $M_j \geqslant j + 2$ for $j = 0, 1, \ldots, m-2$. The n-incidence matrix E is said to be poised if the set of n linear functionals $\{L_i^j; e_{ij} = 1\}$ is linearly independent on P_{n-1}.

In the following four statements of theorems, suppose that the n-incidence matrix E satisfies the Pólya conditions.

THEOREM A (PÓLYA AND WHITTAKER, SEE [4]). *If $k = 2$, then E is poised.*

THEOREM B (FERGUSON [4, p. 24]). *If $k > 2$ and if $e_{i,j-1} = e_{i,j+p} = 0$, $e_{ij} = \cdots = e_{i,j+p-1} = 1$ implies p is even, then E is poised.*

THEOREM C (SCHOENBERG, SEE [4, p. 25]). *If $x_1 = a$ and $x_k = b$, and if $2 \leqslant i \leqslant k - 1$ and $e_{ij} = 1$ imply $e_{ij'} = 1$ for each $j' \leqslant j$, then E is poised.*

By combining the proofs of Theorems B and C given by Ferguson, we can get the following result, which has also been given by Atkinson and Sharma (see [2]).

THEOREM D. *If $x_1 = a$ and $x_k = b$ and if $2 \leqslant i \leqslant k - 1$ and $e_{i,j-1} = e_{i,j+p} = 0$, $e_{ij} = \cdots = e_{i,j+p-1} = 1$ imply p is even, then E is poised.*

DEFINITION (SEE [4, p. 26]). Let the m-incidence matrix E have k rows. Let f_i be the column index of the first one which appears in row i. E is called a *pyramid matrix* if, for each i, $e_{ij} = 1$ implies $e_{ij'} = 1$ for $f_i \leqslant j' \leqslant j$, and there is some value i ($1 \leqslant i \leqslant k$) so that $f_1 \geqslant f_2 \geqslant \cdots \geqslant f_i$ and $f_i \leqslant f_{i+1} \leqslant \cdots \leqslant f_k$.

As examples, if the transpose matrix E^t is

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(Henceforth in the examples we will often suppress the printing of the last zero rows of E^t.) or

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

then E is a pyramid matrix, while if E^t is

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then E is not a pyramid matrix.

THEOREM E (FERGUSON [4]). *If the n-incidence matrix E is a pyramid matrix satisfying the Pólya conditions, then E is poised with respect to the ordering $x_1 < x_2 < \cdots < x_k$.*

We will determine completely the signs of the coefficients α_i in the dependency $0 = \sum_{m=1}^{n+1} \alpha_m R_m$ where $\{R_m\}_{m=1}^{n+1}$ is any set of derivative evaluations $(e_{x_i}^j)$ whose $(n + 1)$-incidence matrix is a pyramid matrix satisfying the *strong* Pólya conditions. This then will yield a simple alternation scheme completely simplifying the Exchange Procedure in a large class of linearly restricted approximation problems where V^n is Haar on Ω including, for example, Bounded Coefficients Approximation.

NOTATION. The situation of the preceding paragraph will be referred to in the sequel as the pyramid situation, and we will say $R = (R_1, R_2, \ldots, R_{n+1})$ $\in PM_n$. We will also assume $R \in PM_n$ is ordered so that if $R_\mu = e_{x_i}^l$ then $R_{u+1} = e_{x_i}^l$ or $e_{x_{i+1}}^l$ where $j \geqslant l$ if $i \leqslant i_0$ and otherwise $l \geqslant j$.

We first recall the fundamental Tchebycheff result.

THEOREM 1 (TCHEBYCHEFF ALTERNATION). *If M is an m-dimensional Haar subspace of $C([a, b])$ and $\sum_{i=1}^{m+1} \alpha_i e_{x_i} = 0$ on M for $x_i < x_{i+1}$ ($i = 1, 2, \ldots, m$), then $\alpha_i \alpha_{i+1} < 0$ ($i = 1, 2, \ldots, m$).*

LEMMA 1. *In the pyramid situation, $\operatorname{sign}(e_{x_i}^j) \operatorname{sign}(e_{x_{i+1}}^j) = -1$.*

PROOF. (1) must hold on the 1-dimensional space $V^1 = V^n \cap \{p; Rp = 0 \text{ if } R \neq e_{x_{i+k}}^j, \ k = 0, 1\}$. But $V_k^1 = \{p^{(k)}; p \in V^1\}$ is a Haar space on the interval $[x_i, x_{i+1}]$ by Theorem E. Thus, $0 = \alpha_k e_{x_i} + \alpha_{k+1} e_{x_{i+1}}$ on V_k^1, and the result follows from Theorem 1. □

EXAMPLE. Consider E^t (where we replace 1's by .'s).

The signs of the functionals $e^j_{x_i}$ occur above them. Then we see that the alter-

$$\begin{pmatrix} 0 & 0 & \overset{+}{\cdot} & \overset{-}{\cdot} & \overset{+}{\cdot} & \overset{-}{\cdot} & 0 \\ \overset{-}{\cdot} & + & 0 & 0 & 0 & 0 & \overset{+}{} \\ 0 & 0 & 0 & 0 & 0 & 0 & \overset{-}{\cdot} \end{pmatrix}$$

nation pattern of Lemma 2 occurs in the $j = 0$ and $j = 1$ rows. The other sign relationships exhibited in this example will be demonstrated later (see Theorem 3).

From the fact that the determinant of a square matrix is nonzero on an open set we have immediately the following technical lemma. Let Δ denote the dual of V^n.

LEMMA 2. *If $R = (R_1, R_2, \ldots, R_{n+1})$ has uniform rank n then there is an open neighborhood N (in Δ^{n+1}) of R so that if $S = (S_1, S_2, \ldots, S_{n+1}) \in N$ then S has uniform rank n. Moreover, if $0 = \Sigma_{i=1}^{n+1} \alpha_i R_i$ and $0 = \Sigma_{i=1}^{n+1} \beta_i S_i$ and $\alpha_1 \beta_1 > 0$ then $\alpha_i \beta_i > 0$ $(i = 1, 2, \ldots, n + 1)$. In fact if Δ_u^{n+1} denotes the subset of Δ^{n+1} of $(n + 1)$-tuples with uniform rank n, then Δ_u^{n+1} is divided into equivalence classes (components) so that S and T are in the class if and only if there is a continuum between S and T lying in Δ_u^{n+1}.*

The following result is a lemma in the development of the final result on the pyramid situation. It yields the alternation scheme for the important special case of Bounded Coefficients Approximation and the method of proof generalizes to other approximation schemes (see Theorem 5).

THEOREM 2. *Consider the special case of the pyramid situation with $R_m = e^j_{x_i}$ where $0 = x_0 < x_1 < \cdots < x_k$ and $j = 0$ for $1 \leqslant i \leqslant k$. Then $\text{sign}(e_0^q) \cdot \text{sign}(e_0^l) = (-1)^{l-q+1}$ where q and l are adjacent (i.e., no j lies between q and l). Furthermore $\text{sign}(e_{x_i}) \text{sign}(e_{x_{i+1}}) = -1$ and if $e_0 \notin \{R_m\}$ and q is the smallest integer such that $e_0^q \in \{R_m\}$, then $\text{sign}(e_0^q) \text{sign}(e_{x_1}) = (-1)^{q-1}$.*

PROOF. Consider the corresponding Vandermonde matrix

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & & & & \\ 1 & x_k & x_k^2 & \cdots & x_k^{n-1} \\ 1\delta_0 & 0 & 0 & & 0 \\ 0 & 1\delta_1 & 0 & & 0 \\ 0 & 0 & 2\delta_2 & & 0 \\ 0 & 0 & 0 & & (n-1)!\delta_{n-1} \end{pmatrix}$$

where $\delta_s = 1$ if $s \in \{j; e_0^j \in \{R_m\}_{m=1}^{n+1}\} = S_{n-k+1}$, otherwise $\delta_s = 0$. Note that S_{n-k+1} contains $n - k + 1$ elements. Then $\Sigma_{m=1}^{n+1}\alpha_m R_m = 0$ corresponds to the n relations

$$(2) \qquad \sum_{m=1}^{k} \alpha_m x_m^s = -s!\delta_s\alpha_s', \qquad s = 0, 1, \ldots, n-1,$$

where $\{\alpha_s'\}_{\delta_s = 1} = \{\alpha_m\}_{m=k+1}^{n+1}$.

Now, as functions of s on $[0, \infty)$, $\{x_m^s\}_{m=2}^k$ form a $(k-1)$st order Tchebycheff system since all x_i are distinct and positive. Thus by the well-known nonsingularity of the Vandermonde matrix, the α_m are uniquely determined (up to a multiplicative constant) by noting that $\delta_s = 0$ for $s \in \{0, 1, \ldots, n-1\} - S_{n-k+1} = T_{k-1}$ where T_{k-1} contains $k - 1$ elements. For a fixed nonzero k-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_k)$, consider $F(s) = \Sigma_{m=1}^k \alpha_m x_m^s$. Then as we have noted $F(s)$ has exactly $(k-1)$ zeros on $[0, \infty)$ and they occur in T_{k-1}. Hence F is of constant sign as s varies between two elements of T_{k-1}. By (2) therefore, $\text{sign}(\alpha_s')$ is constant for all $s \in S_{n-k+1}$ between two consecutive elements of T_{k-1} and changes sign as s passes over each element of T_{k-1}. We conclude that $\text{sign}(e_0^q)\,\text{sign}(e_0^l) = (-1)^{l-q+1}$ where q and l are adjacent in S_{n-k+1}.

Lemma 1 shows that $\text{sign}(e_{x_i})\,\text{sign}(e_{x_{i+1}}) = -1$. If $e_0 \notin \{R_m\}$, by Lemma 2 we can move e_{x_1} continuously to e_0 staying always in the (uniform rank n) pyramid situation to conclude that $\text{sign}(e_0^q)\,\text{sign}(e_{x_1}) = (-1)^{q-1}$ (which is the case if $x_1 = 0$). $\square$

COROLLARY 1. *Theorem 2 provides the alternation pattern for Bounded Coefficients Approximation, thereby completely simplifying the Exchange Procedure in this case.*

EXAMPLE. Recall that

$$V_0^n = \left\{ p = \sum_{i=0}^{n-1} a_i x^i \in P_{n-1}[a, b]; c_i \leqslant a_i \leqslant b_i \quad (i = 0, 1, \ldots, n-1) \right\},$$

where c_i and b_i are given extended real numbers, and $c_0 \leqslant f(0) \leqslant b_0$. Then

$$V_0^n = \{p \in P_{n-1}[a, b]; l_i \leqslant e_0^i p \leqslant u_i \quad (i = 0, 1, \ldots, n-1)\}$$

where $l_i = i!c_i$ and $u_i = i!b_i$. Suppose at stage ν the $n + 1$ functionals forming the extremal set are

$$\overset{+15}{e_0} \quad \overset{+12}{e_0} \quad \overset{+11}{e_0} \quad \overset{+10}{e_0} \quad \overset{-6}{e_0} \quad \overset{-5}{e_0} \quad \overset{-}{e}_{x_1} \quad \overset{+}{e}_{x_2} \cdots \overset{-}{e}_{x_{k-1}} \quad \overset{+}{e}_{x_k}$$

where $x_i < x_{i+1}$ and we write $\overset{\pm}{R}$ for $\pm R$. Thus if $N_*^\nu = \overset{+}{e}_x$, where $x_1 < x < x_2$, N_*^ν replaces $\overset{+}{e}_{x_2}$ while if $N_*^\nu = \overline{e}_x$, N_*^ν replaces $\overline{e}_{x_1}$. If $N_*^\nu = \overline{e}_x$, where $0 < x <$

x_1 then N_*^ν replaces $\bar{e}_{x_1}$ while if $N_*^\nu = \overset{+}{e}_x$, N_*^ν replaces $\bar{e}_0^5$. If $N_*^\nu = \bar{e}_0^7$ then N_*^ν replaces $\overset{+10}{e}_0$, while if $N_*^\nu = \overset{+7}{e}_0^\nu$ then N_*^ν replaces $\bar{e}_0^6$. If $N_*^\nu = \overset{+}{e}_x$ where $x_{n+1} < x$ then N_*^ν replaces $\overset{+}{e}_{x_{n+1}}$, while if $N_*^\nu = \bar{e}_x$ then N_*^ν replaces $\overset{+15}{e}_0$.

By a change of variable $(x \longleftarrow -x)$ in Theorem 2 we obtain the following result.

COROLLARY 2. *Consider* $R \in PM_n$ *with* $R_m = e_{x_i}^j$ *where* $x_1 < \cdots < x_k$ $< x_{k+1} = b$ *and* $j = 0$ *for* $1 \le i \le k$. *Then* $\text{sign}(e_b^q)\,\text{sign}(e_b^l) = -1$ *where* q *and* l *are adjacent (i.e., no j lies between q and l). Furthermore* $\text{sign}(e_{x_i})\,\text{sign}(e_{x_{i+1}}) =$ -1. *If* $e_b \notin \{R_m\}$ *and* q *is the smallest integer such that* $e_b^q \in \{R_m\}$, *then* $\text{sign}(e_b^q)\,\text{sign}(e_{x_k}) = -1$.

EXAMPLE.

$$\overset{+}{e}_{x_1}\,\overset{-}{e}_{x_2}\,\cdots\,\overset{+}{e}_{x_{k-1}}\,\overset{-}{e}_{x_k}\,\overset{+5}{e}_b\,\overset{-6}{e}_b\,\overset{+10}{e}_b\,\overset{-11}{e}_b\,\overset{+12}{e}_b\,\overset{-15}{e}_b$$

Note 1. If $b = 0$, Corollary 2 gives the alternation scheme for Bounded Coefficients Approximation on $[a, 0)$.

The following lemma provides a useful method for determining the signs of the coefficients in (1) in many cases of restraints at Birkhoff data.

LEMMA 3. *Suppose* $R \in PM_n$ *and* $e_{x_i}^j$ *and* $e_{x_{i+1}}^j$ *are entries in R with signs $-$ and $+$ respectively, and such that if $e_{x_r}^l$ follows $e_{x_{i+1}}^j$ in the ordering, then $l > j$. Let R' be the $(n+1)$-tuple coinciding with R except that $e_{x_{i+1}}^j$ is replaced by $e_{x_i}^{j+1}$. Then either* (i) $R' \in PM_n$ *and* $\text{sign}(R_i') = \text{sign}(R_i)$ *if* $R_i \ne$ $e_{x_i}^j$, *or* (ii) R' *less* $e_{x_i}^j$ *belongs to* PM_{n-1} *where each j is replaced by $j - 1$.*

PROOF. That $R' \in PM_n$ or $R' - \{e_{x_i}^j\} \in PM_{n-1}$ is immediate from the "$l > j$" condition. Then write

$$R_1 = \sum_{k=2}^{n=1}{}' \alpha_{l,}(h)R_k + \left(\frac{\alpha(h) - \beta(h)}{h}\right)e_{x_i}^j + \beta(h)\frac{[e_{x_i+h}^j - e_{x_i}^j]}{h}$$

where $0 < h \le x_{i+1} - x_i$ and Σ' denotes $\Sigma_{k=2}^{n+1}$ minus the two terms where $R_k = e_{x_i}^j$ and $e_{x_i+h}^j$. Then $\text{sign}\,\alpha_k(h) = \text{sign}\,\alpha_k(x_{i+1} - x_i)$ and $\text{sign}\,\beta(h) =$ $\text{sign}\,\beta(x_{i+1} - x_i)$, since none of $\alpha_k(h)$ and $\beta(h)$ can have a zero for $0 < h \le$ $x_{i+1} - x_i$. This follows since $R(h) = R$ with $e_{x_{i+1}}^j$ replaced by $e_{x_i+h}^j$ remains in PM_n and thus has uniform rank n. But also

$$R_1 = \sum_{k=2}^{n+1}{}' \alpha_k R_k + \gamma e_{x_i}^j + \beta e_{x_i}^{j+1},$$

with all coefficients nonzero except perhaps γ. Since $(e_{x_i+h}^j - e_{x_i}^j)/h$ approaches $e_{x_i}^{j+1}$ weakly in Δ, the dual of V^n, a finite dimensional space, the convergence

is also norm convergence in Δ. Hence $\alpha_k(h)$ converges to α_k whence sign $\alpha_k =$ sign $\alpha_k(x_{i+1} - x_i)$ and $\beta(h)$ converges to β whence sign $\beta =$ sign $\beta(x_{i+1} - x_i)$. $\square$

EXAMPLE A. In the situation of Theorem 2, by repeated application of Lemma 3 the functionals e_{x_i} may be replaced by derivative evaluation functionals at a single point, say x_1, i.e., we may replace e_{x_i} by $e_{x_1}^{(i-1)}$, $1 \leqslant i \leqslant k$. To illustrate this consider the following example.

$$
\begin{pmatrix} + & - & + & - \\ \vdots & & & \vdots \\ \sigma & & (2) & (1) \\ \vdots & & & \\ & & & \\ \tau & & & \end{pmatrix}
\begin{array}{c} \longrightarrow \\ L.3(1) \\ L.3(2) \end{array}
\begin{pmatrix} + & - \\ \vdots & \vdots \\ \sigma & \\ \vdots & \vdots \\ \tau & \end{pmatrix}
\begin{array}{c} \longrightarrow \\ L.3 \end{array}
\begin{pmatrix} + & - \\ \vdots & \vdots \\ \sigma & \\ \vdots \\ \tau \end{pmatrix}
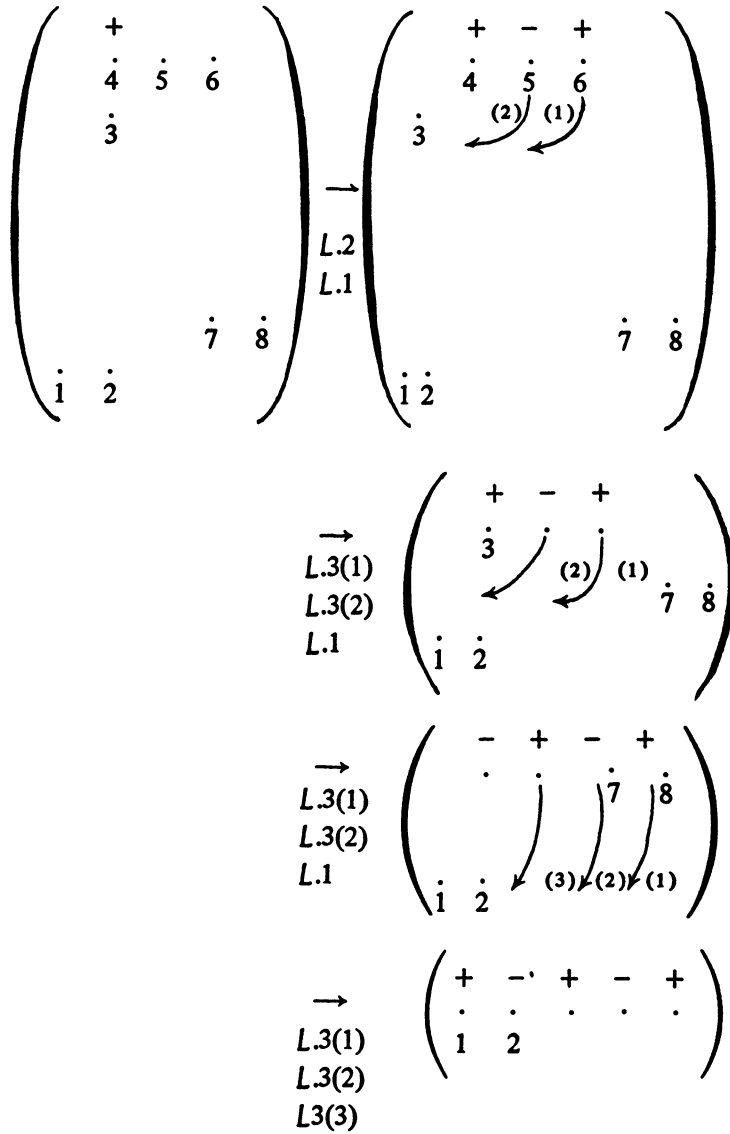$$

Note that the signs $-, +, \sigma, \tau$ of the functionals e_{x_1}, e_0, e_0^1, e_0^3 do not change. Thus as an alternative in the proof of Theorem 2 we can examine the Vandermonde matrix

$$
\begin{pmatrix}
1 & x_1 & x_1^2 & \cdots & x_1^{n+1} \\
0 & 1 & 2x_1 & \cdots & (n-1)x_1^{n-2} \\
0 & 0 & 2 & \cdots & (n-2)(n-1)x_1^{n-3} \\
\vdots & \vdots & \vdots & \cdot & \\
0 & 0 & 0 & \cdot & (n-1)_{k-1}x_1^{n-k} \\
1\delta_0 & 0 & 0 & \cdots & 0 \\
0 & 1\delta_1 & 0 & \cdots & 0 \\
0 & 0 & 2\delta_2 & \cdots & 0 \\
0 & 0 & 0 & \cdots & (n-1)!\delta_{n-1}
\end{pmatrix}
$$

Then $F(s) = \Sigma_{m=1}^k \alpha_m (s)_{m-1} x_1^{s-m+1}$, where $(s)_{m-1} = s(s-1) \cdots (s-m+1)$. Without loss of generality take $x_1 = 1$. Thus $F(s) = \Sigma_{m=1}^k \alpha_m (s)_{m-1} = p_k(s)$ where $p_k(s)$ is a kth degree algebraic polynomial whose only zeros are at the $k-1$ points s where $\delta_s = 0$.

Lemmas 1–3 give rise to a technique for solving the alternation problem for an arbitrary pyramid situation $(R \in P)$.

EXAMPLE. Suppose we are given the following, $R = (R_1, \ldots, R_8) \in PM_7$, and we wish to determine all the signs, given that sign$(R_4) = 1$.

$$
\begin{pmatrix}
& + & & & \\
& \dot{4} & \dot{5} & \dot{6} & \\
& \dot{3} & & & \\
& & & & \\
& & & \dot{7} & \dot{8} \\
\dot{1} & \dot{2} & & &
\end{pmatrix}
\xrightarrow[\substack{L.2 \\ L.1}]{}
\begin{pmatrix}
& + & - & + & \\
& \dot{4} & \dot{5} & \dot{6} & \\
& & {}_{(2)}\!\!\nearrow & {}^{(1)} & \\
\dot{3} & & & & \\
& & & & \\
& & & \dot{7} & \dot{8} \\
\dot{1}\,\dot{2} & & &
\end{pmatrix}
$$

$$
\xrightarrow[\substack{L.3(1) \\ L.3(2) \\ L.1}]{}
\begin{pmatrix}
& + & - & + & \\
& \dot{3} & \dot{} & \dot{} & \\
& \nearrow & {}_{(2)}\!\!\nearrow & {}^{(1)} & \dot{7} \; \dot{8} \\
\dot{1} & \dot{2} & & &
\end{pmatrix}
$$

$$
\xrightarrow[\substack{L.3(1) \\ L.3(2) \\ L.1}]{}
\begin{pmatrix}
- & + & - & + \\
\dot{} & \dot{} & \dot{7} & \dot{8} \\
& & & \\
\dot{1} & \dot{2} & {}_{(3)}\;{}_{(2)}\;{}_{(1)}
\end{pmatrix}
$$

$$
\xrightarrow[\substack{L.3(1) \\ L.3(2) \\ L3(3)}]{}
\begin{pmatrix}
+ & - & + & - & + \\
\dot{} & \dot{} & \dot{} & \dot{} & \dot{} \\
\dot{1} & \dot{2} & & &
\end{pmatrix}
$$

Hence R_i $(i = 1, \ldots, 8)$ have signs $+, -, +, +, -, +, -, +$ respectively.

An examination of this example indicates immediately the proof of the following alternation scheme for the general pyramid situation. We refer to the topmost row of functionals at each stage as the determining block.

THEOREM 3. *Consider the general pyramid situation $(R \in PM_n)$. Let $R_k = e^j_{x_t}$ and $R_{k+1} = e^r_{x_s}$ and $\sigma = \text{sign}(R_k)\,\text{sign}(R_{k+1})$. If $r < j$, then $\sigma = (-1)^{j-r+1}$, while if $r \geqslant j$, then $\sigma = -1$.*

PROOF. From Lemma 3 we see that as the kth determining block of functionals in $PM_{(n-k)}$ moves down to form the $(k+1)$st determining block in $PM_{(n-k-1)}$, the kth block loses only the leftmost functional and the other functionals retain the same sign. Thus if new functionals are picked up at the right and on the left to form the $(k+1)$st determining block, ordinary alternation occurs among all the functionals in the new $(k+1)$st determining block. The result follows by induction. $\square$

Note 2. Theorem 2 also follows from this technique of repeated application of Lemma 3, since Theorem 2 is a special case of Theorem 3. The proof given in Theorem 2 generalizes, however, to situations where the "Lemma 3" type proof does not (see Theorem 5).

DEFINITION. $A \geqslant B$ will mean that for $a \in A$ either $a \geqslant b$ for all $b \in B$ or $a \leqslant b$ for all $b \in B$.

COROLLARY 3. *Theorem 3 provides the alternation pattern for approximation from $V_0^n = \{p \in P_{n-1}[a, b]; l_i(x) \leqslant p^{(i)}(x) \leqslant u_i(x)$ on E_i, where $[a, b] \subset E_0$ and $E_i \geqslant \bigcup_{j=0}^{i-1} E_j$ for $i = 0, 1, \ldots, n-1\}$.*

As examples we have restricted range with Bounded Coefficients Approximation and approximation by polynomials which are required to be monotone (or convex) in an interval outside $[a, b]$.

Note. The exchange is made so that the alternation pattern is preserved. The exchange is done exactly as in the case of Bounded Coefficients Approximation on the "left side" of the pyramid while the ordinary alternation pattern holds on the "right side" of the pyramid.

COROLLARY 4. *By Lemma 2, if $R \in PM_n$ then the alternation scheme of Theorem 3 also holds for any S in the same component as R in Δ_u^{n+1}. For example we can replace each $R_m = e_{x_i}^j$ by $S_m = \int_{-\infty}^{\infty} (\ )^{(j)} d\mu_{ji}$ where μ_{ji} is a positive measure such that the convex hull of its support c_{ji} contains x_i and such that none of the c_{ji} intersect.*

THEOREM 4. *If V^n is Haar on $\Omega = \{L^\alpha\}_{\alpha \in A} \cup \{e_x\}_{x \in E}$, where $E = [a, b]$, then the ordinary alternation pattern occurs among the point evaluations. (I.e., if $R = (R_1, \ldots, R_{n+1}) \in \Omega^{n+1}$, all R_i distinct, and $\{e_{x_i}\}_{i=1}^k = \{R_i\}_{i=1}^{n+1} \cap \{e_x\}_{x \in E}$, then $\text{sign}(e_{x_i}) \, \text{sign}(e_{x_{i+1}}) = -1, i = 1, 2, \ldots, k-1.$)*

PROOF. $V^{k-1} = V^n \cap \{p; R_i p = 0$ for all $R_i \neq e_x\}$ is an ordinary Haar space on $[a, b]$ and the result follows from Theorem 1. $\square$

Note 3. If V^n is Haar on Ω and if in the Exchange Procedure $R_*^\nu = e_x$ where $x_1 < x < x_2$ and x_1, x_2 are adjacent in E^ν, then $\text{sign}(e_{x_1}) \, \text{sign}(e_{x_2}) = -1$. The Exchange Procedure is therefore simply that e_x will replace that e_{x_i} such that $\text{sign}(e_x) = \text{sign}(e_{x_i})$.

By generalizing the procedure of Theorem 2, we now treat the case of n arbitrary linear restrictions on P_{n-1}, i.e., $\{L^\alpha\}_{\alpha \in A} = \{L^i\}_{i=0}^{n-1}$ and $V^n = P_{n-1}[a, b]$ is Haar on $\Omega = \{L^i\}_{i=0}^{n-1} \cup \{e_x\}_{x \in [a,b]}$. Without loss of generality let $[a, b] = [0, 1]$.

We wish to determine the signs of α_i in $0 = \Sigma_{i=1}^{n+1} \alpha_i R_i$ if $R = (R_1, R_2, \ldots, R_{n+1}) \in \Omega^{n+1}$ where all R_m are distinct. Suppose without loss of generality that $R_{k+1} = e_0$ (by Lemma 2 e_{x_0} can be moved continuously to e_0 without changing the signs of the α_i), $R_{k+m+1} = L^{jm}$ for $m = 1, \ldots, n - k$, and $R_i = e_{x_i}$ for $i = 1, 2, \ldots, k$ where $x_i < x_{i+1}$. Then since $0 = \Sigma_{i=1}^{n+1} \alpha_i R_i$ must also hold on $P_{n-1} \cap [L^m p = 0, m = 0, 1, \ldots, n - k - 1]$ and since P_{n-1} is Haar on Ω we see from Theorem 4 that the ordinary alternation pattern occurs among the e_{x_i}, i.e. $\text{sign}(e_{x_i}) \, \text{sign}(e_{x_{i+1}}) = -1$.

We now examine the alternation pattern for the L^m. By the same argument as that in Example A we can use Lemma 3 to replace e_{x_i} by e_1^i, $i = 1, \ldots, k$. (Correspondingly replace R by R'.) Then let $v_i \in P_{n-1}$ be dual to $\{L^i\}_{i=0}^{n-1}$, i.e., $L^i v_j = \delta_{ij}$ $(0 \leqslant i, j \leqslant n - 1)$, and consider the Vandermonde-type matrix

$$
\begin{pmatrix}
v_0(1) & v_1(1) & \cdots & v_{n-1}(1) \\
v_0^{(1)}(1) & v_1^{(1)}(1) & \cdots & v_{n-1}^{(1)}(1) \\
\vdots & \vdots & & \vdots \\
v_0^{(k-1)}(1) & v_1^{(k-1)}(1) & \cdots & v_{n-1}^{(k-1)}(1) \\
\delta_0 & 0 & \cdots & 0 \\
0 & \delta_1 & \cdots & 0 \\
\vdots & \vdots & & \vdots \\
0 & 0 & \cdots & \delta_{n-1}
\end{pmatrix}
$$

where $\delta_s = 1$ if $L^s \in \{R_m\}_{m=1}^{n+1}$, otherwise $\delta_s = 0$. Denote $\{s; \delta_s = 1\}$ by S_{n-k+1}. Now extend $v_i(x)$ from $\{0, 1, \ldots, n - 1\}$ to $[0, \infty)$ as follows.

$$v_i(x) = \sum_{j=0}^{n-1} a_{ij} x^j, \quad i = 0, 1, \ldots, n - 1,$$

so define

$$v_s(x) = x^s \sum_{j=-(n-1)}^{n-1} p_j(s) x^j,$$

where $p_j \in P_{n-|j|-1}$ and $p_j(i) = a_{i(j+i)}$, $0 \leqslant i + j \leqslant n - 1$. Let

$$q_m(s) = \sum_{|j| \leqslant n-1} (j+s)_{m-1} p_j(s), \quad m = 1, 2, \ldots, n.$$

$\{q_m\}$ is a set of auxiliary algebraic polynomials depending only on $\{L^i\}^{n-1}$ and the ordering $L^0, L^1, \ldots, L^{n-1}$. (The p_j depend on the $\{v_i\}_{i=0}^{n-1}$ being ordered.)

Then $\Sigma_{m=1}^{n+1} \alpha_m R'_m = 0$ corresponds to the n relations

$$(3) \qquad \sum_{m=1}^{k} \alpha_m v_s^{(m-1)}(1) = -\delta_s \alpha'_s, \quad s = 0, 1, \ldots, n-1,$$

where $\{\alpha'_s\}_{\delta_s=1} = \{\alpha_m\}_{m=k+1}^{n+1}$. Now P_{n-1} is Haar on $[0, 1]$, the $\{v_i\}_{i=0}^{n-1}$ are linearly independent in P_{n-1}, and $\delta_s = 0$ for $s \in \{0, 1, \ldots, n-1\} - S_{n-k+1} = T_{k-1}$ where T_{k-1} contains $k-1$ elements. Thus the α_m are uniquely determined (up to a multiplicative constant). For a fixed nonzero k-tuple $(\alpha_1, \alpha_2, \ldots, \alpha_k)$, consider

$$F_R(s) = \sum_{m=1}^{k} \alpha_m v_s^{(m-1)}(1) = \sum_{m=1}^{k} \alpha_m \sum_j (j+s)_{m-1} p_j(s)$$

$$= \sum_{m=1}^{k} \alpha_m q_m(s).$$

Then $\alpha'_s = F_R(s)$ for $s \in S_{n-k+1}$. Suppose now that $R_*^\nu (\neq e_x)$ is to replace some R_i^ν in the Exchange Procedure. Determine $Z_{R^\nu}(s)$ to be a polynomial $\Sigma_{m=1}^{k} \beta_m q_m(s)$ which vanishes at $s \in T_{k-1} - \{S_0\}$ and takes the value 1 at s_0 where $R_{s_0} = R_*^\nu$. Then consider $Q_\lambda(s) = F_{R^\nu}(s) + \lambda Z_{R^\nu}(s)$, where $|\lambda|$ is the first value larger than 0 so that $Q_\lambda(s) = 0$ for some $s \in S_{n-k+1}$, say s_1. Note that

$$|\lambda| = \min_{S_{n-k+1}} |F_{R^\nu}(s)/Z_{R^\nu}(s)|.$$

Then R_*^ν replaces R_{s_1} and $F_{R^{\nu+1}}(s) = F_{R^\nu}(s) + \lambda Z_{R^\nu}(s)$.

We summarize this result as follows.

THEOREM 5. *Consider the case where $P_{n-1}[a, b]$ is Haar on $\{L^i\}_{i=1}^{n-1} \cup \{e_x\}_{x \in [a,b]}$. Determine the auxiliary polynomials $\{q_m(s)\}_{m=1}^{n}$ according to the procedure given above. Suppose at the νth stage of the Remez algorithm we have $R^\nu = (R_1^\nu, R_2^\nu, \ldots, R_{n+1}^\nu)$ and we wish to replace some R_i^ν by R_*^ν. Suppose the point evaluations in $\{R_i^\nu\}$ have masses at $x_1 < x_2 < \cdots < x_k$.*

There are four options:

(i) If $R_* = e_x$ where $x_i < x < x_{i+1}$ or $x < x_1$ and $\text{sign}(e_x) = \text{sign}(e_{x_1})$, or $x > x_k$ and $\text{sign}(e_x) = \text{sign}(e_{x_k})$, exchange R_*^ν for some $R_i = e_x$ in the usual way since by Theorem 4 the ordinary alternation pattern occurs among the e_{x_i}, i.e., $\text{sign}(e_{x_i}) \text{sign}(e_{x_{i+1}}) = -1$.

(ii) If $R_\nu^* = e_x$ where $x < x_1$ and $\text{sign}(e_x)\,\text{sign}(e_{x_1}) < 0$ or where $x > x_k$ and $\text{sign}(e_x)\,\text{sign}(e_{x_k}) < 0$, then use the ordinary Exchange Procedure. (R_ν^* will replace $R_i \neq e_x$.)

(iii) Otherwise determine $F_{R^\nu}(s)$ and $Z_{R^\nu}(s)$ according to the procedure given above, i.e.,

$$F_{R^\nu}(s) = \sum_{m=1}^{k} \alpha_m q_m(s) = 0 \quad \text{on } T_{k-1}^\nu = \{s; \delta_s = 0\},$$

$$Z_{R^\nu}(s) = \sum_{m=1}^{k} \beta_m q_m(s) = \begin{cases} 0 & \text{on } T_{k-1}^\nu - \{s_0\}, \\ 1 & \text{if } s = s_0. \end{cases}$$

Let s_1 yield $\min_{S_{n-k+1}} |F_{R^\nu}(s)/Z_{R^\nu}(s)|$ and let $Q(s) = F_{R^\nu}(s) + \lambda Z_{R^\nu}(s)$, where $\lambda = -F_{R^\nu}(s_1)/Z_{R^\nu}(s_1)$.

If $\text{sign } Q(s) = \text{sign } F_{R^\nu}(s)$ on S_{n-k+1}^ν, then R_{s_1} is exchanged for R_ν^*. (Then also $F_{R^{\nu+1}}(s) = Q(s)$ and there is eliminated the need to recalculate $F_{R^{\nu+1}}(s)$ if option (iii) is taken at the next $(\nu + 2)$nd step.)

If $\text{sign } Q(s) \neq \text{sign } F_{R^\nu}(s)$ on S_{n-k+1}^ν, take option (iv).

(iv) $R_*^\nu \neq e_x$ must replace either e_{x_1} or e_{x_k}, and so use the ordinary Exchange Procedure to determine which.

Note 4. In option (iii) to calculate $(\beta_1, \beta_2, \ldots, \beta_k)$ requires a $k \times k$ matrix inversion ($2 \leqslant k \leqslant n$). Compare this with the $n \times n$ matrix inversions required in the ordinary Exchange Procedure. The larger the number of restriction extremals, the greater the saving in computation.

Note 5. Since in the siutation of Theorem 5, there are a fixed finite number (n) of restrictions, a table can be formed (using Theorem 5) listing, once and for all, all the (roughly $2^n = \Sigma_{k=0}^n \binom{n}{k}$) possibilities. From this table the rule of exchange can be determined at each stage of the Remez algorithm.

COROLLARY 5. *If $\{q_m(s)\}_{m=1}^n$ is a Markov system on $[0, n-1]$ (i.e., $\{q_m(s)\}_{m=1}^k$ is a kth order Tchebycheff system, $k = 1, 2, \ldots, n$), then an alternation scheme exists identical to that in Theorem 2. That is, if*

$$R = (L^{i_1}, L^{i_2}, \ldots, L^{i_{n-k}}, e_{x_0}, e_{x_1}, \ldots, e_{x_k}),$$

where the L^{i_j} (and e_{x_i}) are ordered, then

$$(\text{sign } L^{i_q})\,\text{sign}(L^{i_{q+1}}) = (-1)^{i_q - i_{q+1} + 1},$$

$$\text{sign}(L^{i_{n-k}})\,\text{sign}(e_{x_0}) = (-1)^{i_{n-k} - 1}$$

and

$$\text{sign}(e_{x_i})\,\text{sign}(e_{x_{i+1}}) = -1.$$

Also then the Exchange Procedure is completely simplified.

PROOF. The proof follows completely analogously as in Theorem 2. □

EXAMPLE. In the case of Theorem 2 (Bounded Coefficients Approxima-tion), where the e_x^j are ordered as usual, we see that $v_s(x) = x^s$, i.e., $p_0(s) \equiv 1$ while $p_j(s) \equiv 0$, $0 < |j| \leqslant n - 1$, whence $q_m(s) = (s)_{m-1}$, and $\{q_m(s)\}_{m=1}^n$ is a Markov system on $[0, n - 1]$.

DEFINITION. $\{M_i\}_{i=1}^m = \{e_{x_i}^{j_i}\}_{i=1}^m$ is called a set of poised Birkhoff data if $M = \{M_i\}_{i=1}^m \cup \{e_{x_\mu}\}_{\mu=m+1}^n$ is linearly independent in the dual of P_{n-1} for all choices of distinct $e_{x_\mu} \notin \{M_i\}_{i=1}^m$. We will say that P_{n-1} is Haar on $[a, b]$ relative to $\{M_i\}_{i=1}^m$.

EXAMPLE. Let the incidence matrix for M satisfy the conditions of The-orem D. Included here is the case of Hermite data and "Ferguson-Atkinson-Sharma" data.

Consider now approximation problems where $V^n = P_{n-1}[a, b]$ and in-cluded in $\Omega = \{M_i\}_{i=1}^m \cup \{e_x\}_{x \in [a,b]}$ is a set $\{M_i\}_{i=1}^m$ of poised Birkhoff inter-polatory data. That is, $V_0^n = \{p \in V^n; M_i p = m_i \ (i = 1, \ldots, m)\}$, and $M_i f = m_i$ if $M_i = e_x$. We assume without loss of generality that all $m_i = 0$, for we can approximate $f - p_0$ where $M_i p_0 = m_i \ (i = 1, \ldots, m)$. Thus we approximate from $V_0^{n-m} = P_{n-1} \cap \{p; M_i p = 0 \ (i = 1, \ldots, n)\}$.

THEOREM 6. *Suppose P_{n-1} is Haar on $[a, b]$ relative to $\{M_i\}_{i=1}^m$. Then we approximate from $V_0^{n-m} = P_{n-1} \cap \{M_i p = 0 \ (i = 1, \ldots, m)\}$. If $R = (e_{x_1}, e_{x_2}, \ldots, e_{x_{n-m+1}})$ is an $(n - m + 1)$-tuple of extremals where $x_1 < x_2 \leqslant \cdots < x_{n-m+1}$, then alternation occurs as follows: $\text{sign}(e_{x_i}) \ \text{sign}(e_{x_{i+1}}) = (-1)^{q+1}$ where q is the number of Hermite data in $\{M_i\}_{i=1}^m$ at all x where $x_i < x < x_{i+1}$.*

PROOF. Consider $\Sigma_{i=1}^{n-m+1} \alpha_i e_{x_i} = 0$. Suppose for some j, $\text{sign}(e_{x_j}) \cdot \text{sign}(e_{x_{j+1}}) = (-1)^q$. Then let $p \in V_0^{n-m}$ such that $p \neq 0$ and $p(x_i) = 0$ for $i \in \{1, 2, \ldots, n - m + 1\} - \{j, j + 1\}$. Then p has no other zeros in $[x_j, x_{j+1}]$ except those at the Hermite data. So if $x \in (x_j, x_{j+1})$ and $\{e_x, e_x^1, \ldots, e_x^{q_x}\} \subset \{M_i\}_{i=1}^m$ then $p(x) = p'(x) = \cdots = p^{(q_x)}(x) = 0$ implies p changes sign q_x times at x. Hence p changes sign q times in (x_j, x_{j+1}). Thus $\text{sign} \ p(x_j) \cdot \text{sign} \ p(x_{j+1}) = (-1)^q$. Then $\alpha_j \ \text{sign} \ p(x_j) + \alpha_{j+1} p(x_{j+1}) \neq 0$, a contradiction. □

Note. P_{n-1} is Haar on $[a, b]$ relative to $\{M_i\}_{i=1}^m$ implies that $V_0^{n-m} = P_{n-1} \cap \{M_i p = 0 \ (i = 1, \ldots, m)\}$ is Haar in each $E_\delta = [a, b] - N_\delta$ where N_δ is an open neighborhood of total length δ containing each x such that $M_i = e_x$. So since the best approximation p to f from V_0^n exists and is unique (see [2]) and since $p(x) = f(x)$ for $M_i = e_x$, we see that the extremal set of the best ap-proximation to f from V_0^{n-m} lies outside some N_{δ_0}. Hence the Remez algo-rithm can be applied where $\Omega = \{e_x\}_{x \in E_{\delta_0}}$ and δ_0 is sufficiently small.

COROLLARY 6. *Theorem 6 provides the alternation pattern for approximation with interpolation at poised Birkhoff data, thereby completely simplifying the Exchange Procedure.*

Note. As in [7], Theorem 6 can be extended from P_{n-1} to extended Tchebycheff systems. Also other restraining conditions $\{L^\alpha\}_{\alpha \in A}$ can be added with the hypothesis that V_0^{n-m} be Haar on $\Omega = \{L^\alpha\}_{\alpha \in A} \cup \{e_x\}_{x \in E_{\delta_0}}$. Then, as in Theorem 4, the conclusion of Theorem 6 applies to the point evaluation functionals.

EXAMPLE. An example is Polynomial Approximation with Interpolation (at strictly Hermite data). (See Example 7 of §1 and [7].)

We can, as a special case in the above note, add in poised Birkhoff interpolating restraints $\{M_i\}_{i=1}^m$ to pyramid type restraints and determine the alternation pattern by combining the results of Theorems 3 and 6 in the obvious way.

THEOREM 7. *Let V_0^n be as in Corollary 3 and let $\{M_i^j\}_{i=1}^{m_j}$ be poised Birkhoff data on each E_j, $j = 0, 1, \ldots, n-1$. Let $\Sigma m_j = m$. Then the Remez algorithm can be applied to $V_0^{n-m} = V_0^n \cap \{p; M_i^j p = 0\}$ with the following alternation scheme. Let $R \in PM_{n-m}$. Let $R_k = e_{x_i}^t$ and $R_{k+1} = e_{x_s}^r$ and $\sigma = \text{sign}(R_k) \, \text{sign}(R_{k+1})$.*

   (i) *If $r < t$, then $\sigma = (-1)^{t-r+1}$.*

   (ii) *If $r = t$, then $\sigma = (-1)^{q+1}$ where q is the number of Hermite data in $\{M_i^j\}$ at all x where $x_i < x < x_s$.*

   (iii) *If $r > t$, then $\sigma = -1$.*

EXAMPLE. As an example we have restricted range approximation with Hermite interpolation.

Finally we wish to consider the case where V^n is not Haar on Ω and investigate what alternation properties hold and what simplifications can be made in the Exchange Procedure. We will confine our attention to Restricted Derivatives Approximation (R.D.A.).

From Theorem 2.1 we know that p is a best approximation to f if and only if

(4)
$$0 = \sum_{i=1}^r \lambda_i \sigma_i e_{x_i} + \sum_{i=r+1}^{r+s+1} \lambda_i \sigma_i e_{x_i}^{j_i}$$

for some $T = \{e_{x_i}^{k_i}\}_{i=1}^{r+s+1} \subset S'$ where $j_i > 0$, $\sigma_i = \text{sign}(e_{x_i}^{k_i})$, all $\lambda_i > 0$ and $r + s \leqslant n + 1$.

THEOREM 8. *In the case of R.D.A., suppose that if $x_i \in (a, b)$ then $e_{x_i}^{j_i}$ and $e_{x_i}^{j_i+1}$ do not both belong simultaneously to T. Let t equal the number of $e_{x_i}^{j_i}$ where $x_i \in (a, b)$. Then if $r + s + t = n + 1$, we have alternation occurring among the e_{x_i} in (4); i.e., if $x_1 < x_2 < \cdots < x_r$, then $\sigma_i \sigma_{i+1} = -1$, $i = 1, 2,$*

$\ldots, r - 1$. *In fact, the alternation pattern among $\{e^{j_i}_{x_i}\}_{x_i = a \text{ or } b} \cup \{e_{x_i}\}^{r}_{i=1}$ is that of Theorem 3 (the pyramid case).*

PROOF.

$$\{e^{j_i}_{x_i}\}_{x_i = a \text{ or } b} \cup \{e^{j_i}_{x_i}, e^{j_i+1}_{x_i}\}_{x_i \in (a,b)} \cup \{e_{x_\mu}\}^{r-1}_{j_i=1} = \{R^i\}^{n}_{i=1}$$

gives rise to a poised incidence matrix by Theorem D. Thus if $p \in P_{n-1}$ and $R^i p = 0$ $(i = 1, 2, \ldots, n)$, then $p \equiv 0$. Suppose alternation among the e_{x_i} does not occur, say $\sigma_j \sigma_{j+1} = +1$.

Then choose $p \not\equiv 0$ satisfying $R^i p = 0$ for all $i \neq j, j + 1$. But then from (4) $0 = \lambda_j \sigma_j p(x_j) + \lambda_{j+1} \sigma_{j+1} p(x_{j+1})$ which contradicts the fact that p cannot have another zero in $[x_j, x_{j+1}]$. The analogous argument shows that in fact the alternation pattern among all $\{e^{j_i}_{x_i}\}_{x_i = a \text{ or } b} \cup \{e_{x_i}\}^{n}_{i=1}$ is that of Theorem 3 (the pyramid case). □

EXAMPLE. In [6, Example 7] we have that in the case of Monotone Approximation where $f(x) = x^7$, on $[-1, 1]$ and $n = 7$, the extremal set is $\{\bar{e}_{-1}, \overset{+}{e}_{-.8}, \bar{e}_{.8}, \overset{+}{e}_1, \overset{+}{e}^1_{-.4}, \overset{+}{e}^1_{+.4}\}$.

EXAMPLE. In [6, Example 2] we have that in the case of convex approximation where $f(x) = |x|$ on $[-1, 1]$ and $n = 6$, the extremal set is $\{\bar{e}_{-1}, \overset{+}{e}_{-.4}, \bar{e}_0, \overset{+}{e}_{.4}, \bar{e}_1, \overset{+}{e}^2_{-1}, \overset{+}{e}^2_1\}$.

Note. Theorem 6 can be strengthened by removing the requirement that $e^{j_i}_{x_i}$ and $e^{j_i+1}_{x_i}$ not both belong to T and requiring instead that after adding some auxiliary functionals $e^{j_i+1}_{x_i}$ to T to form even blocks (see Theorem D) the number of elements in the augmented set T equals $n + 1$.

In the case of Theorem 6 note that the auxiliary extremals are the same auxiliary functionals occurring in the uniqueness results [2]. In fact we have the following result.

COROLLARY 7. *In the situation of Theorem 6, p is a best approximation to f from $V^n_0 \cap \{\sigma_i p^{j_i+1}(x_i) = c'_i$ if $x_i \in (a, b), i = r + 1, \ldots, r + s + 1\}$ and p is a best approximation to f from $K = \{p \in P_{n-1}; c_i \leqslant \sigma_i p^{j_i}(x_i), i = r + 1, \ldots, r + s + 1$ and $c'_i = \sigma_i p^{j_i+1}(x_i)$ if $x_i \in (a, b), i = r + 1, \ldots, r + s + 1\}$.*

PROOF. If $c_i(x)$ is a lower restraining function for $\sigma_i p^{j_i}(x)$ then $\sigma_i p^{j_i}(x) - c_i(x) = 0$ at x_i implies that $\sigma_i p^{j_i+1}(x) - c'_i(x) = 0$. The result follows from Theorem 2.

THEOREM 9. *If in (4) there are only t $e^{j_i}_{x_i}$ such that $x_i \in (x_1, x_r)$ where $t = 0$ or 1, then*

(i) *if $t = 0$, then $r + s = n + 1$ and the alternation pattern coincides with that in Theorem 3 (i.e., the pyramid situation);*

(ii) *if $t = 1$, then $n \leqslant r + s \leqslant n + 1$.*

If $n = r + s$, the alternation pattern is again that of Theorem 3 among all the functionals excluding $e^{j_i}_{x_i}$ which itself appears with a positive coefficient in (4).

If $n + 1 = r + s$, the alternation pattern is the same as in (ii) with exactly one exception.

PROOF. In case (i), the functionals are of pyramid type. In case (ii) if $r + s < n$, then adding $e^{j_i+1}_{x_i}$ would lead to a poised system making the dependency (4) impossible. Thus $r + s \geqslant n$. If $r + s = n$, the conclusion follows from Theorem 8. If $r + s = n + 1$, then exchange one of the functions, say $e^k_{x_i}$ in (4) with $\pm e^{j_i+1}_{x_i}$ by the Exchange Procedure so that $e^{j_i}_{x_i}$ remains with positive coefficient (and all the other functionals have the same sign as in (4)). One of either $+e^{j_i+1}_{x_i}$ or $-e^{j_i+1}_{x_i}$ will accomplish this, for if both replaced $e^{j_i}_{x_i}$, then we would have two essentially different dependencies involving $e^{j_i+1}_{x_i}$ and the other functionals. But this is an impossibility, for the other functionals less $e^{j_i}_{x_i}$ have rank n and $e^{j_i+1}_{x_i}$ can therefore be written in terms of them in exactly one way. Then $T \cup \{e^{j_i}_{x_i}\} - \{e^k_{x_i}\}$ satisfies condition (i) of Theorem 9 by Theorem D. Thus, ignoring $e^k_{x_i}$, we see that the alternation pattern is the same as in (ii). □

EXAMPLES. Theorems 8 and 9 reveal the alternation patterns for all the twelve examples in [6]. For example in [6, Example 5], we have that in the case of convex approximation where $f(x) = \ln 2.1 - \ln(x + 1.1)$ on $[-1, 1]$ and $n = 6$, the extremal set is $\{\bar{e}_{-1}, \bar{e}_{-.9}, \bar{e}_{-.6}, \bar{e}^2_0, \bar{e}_{.6}, \bar{e}_1, \bar{e}^2_1\}$ and in accordance with Theorem 9, if either $\bar{e}_{-.6}$ or $\bar{e}_{.6}$ is suppressed, the remaining functionals (ignoring $\bar{e}^2_0$) alternate as in the pyramid case. ($\pm e^3_0$ would replace either $\bar{e}_{-.6}$ or $\bar{e}_{.6}$ in the Exchange Procedure.)

Another interpretation of the alternation scheme in Theorem 9 would be that if $t = 1$ and $n + 1 = r + s$, then $\bar{e}^{j_i}_{x_i}$ acts like the functional which would make the functionals of pyramid type. In the example mentioned just above, for instance, $\bar{e}^2_0$ acts like $\bar{e}_x$ where $-.6 < x < .6$ (e.g. take $x = 0$).

In all the examples in [6] where $t = 1$ and $n + 1 = r + s$, $e^{j_i}_{x_i}$ acts like $\pm e_{x_i}$ in the above sense. If this were always the case we could always predict which functional $e^{j_i}_{x_i}$ would replace in the Exchange Procedure as long as $t = 0$ and we could simplify the exchange. This need not be the case, however. Consider the following extremal configuration during Monotone Approximation by P_3 where the cubic p_3 (unique up to scalar multiple) is drawn through points symbolized by 1, 2, 5, where c denotes the point between 2 and 5 where $p'_3(c) = 0$ and $\bar{e}$ denotes an extremal to exchange.

$$\bar{1} \quad \overset{+}{2} \quad \bar{3} \quad \bar{5}$$
$$\overset{+}{4}$$
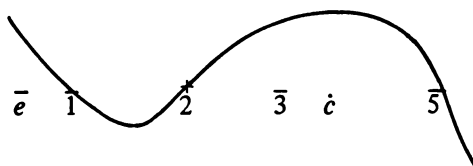
Note that $3 < c$, since $\alpha_3 p_3(3) + \alpha_4 p'_3(4) = 0$ where $\alpha_3\alpha_4 < 0$. Draw a new cubic q_3 through e instead of 1 and let $q'_3(c') = 0$. We can always assume that 3 was close enough to c so that switching from p_3 to q_3 caused c' to be on the
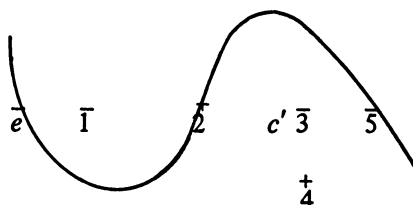
opposite side of 3 as was c. Hence $c' < 3$ and so $\bar{e}$ clearly cannot replace $\bar{1}$ but must instead replace either $\bar{5}$ or $\bar{3}$ (by Theorem 9 there is alternation among the e_x with only one exception). Thus we end up with the following configuration.

$$- \quad - \quad + \quad - \quad \text{or} \quad - \quad - \quad + \quad -$$
$$+ \qquad\qquad\qquad\qquad +$$

Thus even if $t = 0$ it cannot be predicted which functional is replaced by $e^{j_i}_{x_i}$ from the alternation scheme. And also if $t = 1$ and $\bar{e}_x$ is to replace some $\bar{e}_{x_i}$ it cannot in general be determined which one without going through the Exchange Procedure. There are, however, some cases when the rule of exchange is obvious. For example in the configuration

$$\bar{1} \quad \overset{+}{2} \quad \bar{3} \quad \bar{5}$$
$$\overset{+}{4}$$

if $\overset{+}{e}$ occurs between $\bar{3}$ and $\bar{5}$, then surely $\overset{+}{e}$ will replace $\overset{+}{4}$.

We conclude by pointing out that for all the approximation problems discussed in this paper if there is exactly one extremal $(n + 1)$-tuple $N = (N_1, N_2, \ldots, N_{n+1})$ with distinct entries for f and p then, in the Remez algorithm $N^\nu = (N^\nu_1, N^\nu_2, \ldots, N^\nu_{n+1})$ converges to N. We conclude, therefore, by noting the following result.

THEOREM 10. *For any approximation problem discussed in this paper, if there is exactly one extremal $(n + 1)$-tuple $N = (N_1, N_2, \ldots, N_{n+1})$ with distinct entries for f and p the best approximation, then in the Remez algorithm,*

$N^\nu = (N_1^\nu, N_2^\nu, \ldots, N_{n+1}^\nu)$ *converges to* N. *Thus there exists* K *such that* $\nu \geqslant$ K *implies that, in the Exchange Procedure,* N_*^ν *replaces that* $N_{i_0}^\nu$ *yielding* $\min_{1 \leqslant i \leqslant n+1} \|N_i^\nu - N_*^\nu\|$.

Note. The situation of Theorem 10 is, of course, the usual one, and thus if K can be determined, the Exchange Procedure after the Kth stage is greatly simplified.

Note. Computational examples will be carried out on the computer and will appear later.

The author wishes to thank Professor G. D. Taylor for his most helpful encouragement with respect to this paper.

REFERENCES

1. C. Carasso, *Convergence de l'algorithme de Rémès*, J. Approximation Theory 11 (1974), 149–158. MR 50 #11723.

2. B. L. Chalmers, *A unified approach to uniform real approximation by polynomials with linear restrictions*, Trans. Amer. Math. Soc. 166 (1972), 309–316. MR 45 #4030.

3. E. W. Cheney, *Introduction to approximation theory*, McGraw-Hill, New York, 1966. MR 36 #5568.

4. David Ferguson, *The question of uniqueness for G. D. Birkhoff interpolation problems*, J. Approximation Theory 2 (1969), 1–28. MR 40 #599.

5. P.-J. Laurent, *Approximation et optimisation*, Hermann, Paris, 1972.

6. J. T. Lewis, *Approximation with convex constraints*, Univ. of Rhode Island, Tech. Report #11, August 1970.

7. H. L. Loeb, D. G. Moursund, L. L. Schumaker and G. D. Taylor, *Uniform generalized weight function polynomial approximation with interpolation*, SIAM J. Numer. Anal. 6 (1969), 284–293. MR 40 #3679.

8. G. G. Lorentz and K. L. Zeller, *Monotone approximation by algebraic polynomials*, Trans. Amer. Math. Soc. 149 (1970), 1–18. MR 44 #3060.

9. R. A. Lorentz, *Uniqueness of best approximation by monotone polynomials*, J. Approximation Theory 4 (1971), 401–418. MR 45 #779.

10. J. A. Roulier, *Polynomials of best approximation which are monotone*, J. Approximation Theory 9 (1973), 212–217. MR 50 #5279.

11. J. A. Roulier and G. D. Taylor, *Uniform approximation by polynomials having bounded coefficients*, Abh. Math. Sem. Univ. Hamburg 36 (1971), 126–135. MR 49 #953.

12. O. Shisha, *Monotone approximation*, Pacific J. Math. 15 (1965), 667–671. MR 32 #2802.

13. G. D. Taylor and M. J. Winter, *Calculation of best restricted approximations*, SIAM J. Numer. Anal. 7 (1970), 248–255. MR 42 #3978.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, RIVERSIDE, CALIFORNIA 92502