APPROXIMATION TO SEPARATED FUNCTIONS ON CARTESIAN PRODUCT SPACES

by

Miriam Schapiro Grosof

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in the Belfer Graduate School of Science
Yeshiva University
New York
June 1966

The committee for this doctoral dissertation consisted of:

Harry E. Rauch, Ph.D., Chairman

Donald J. Newman, Ph.D.

Leopold Flatto, Ph.D.

Adam Koranyi, Ph.D.

Introduction

The purpose of this paper is to generalize some of the results contained in Some Theorems on Čebyšev Approximation, by D. J. Newman and H.S. Shapiro [3], and to exhibit the failure of certain other proposed generalizations. This problem was suggested by Professor Donald J. Newman; I wish to acknowledge, with deep gratitude, his innumerable helpful suggestions and constant encouragement.

The work has been carried out, in part, under N.S.F. Summer Fellowships for Graduate Teaching Assistants (1964 and 1965).

Numbers in brackets refer to the bibliography at the end of the paper.

Section 1

In [3], Newman and Shapiro are concerned primarily with uniqueness questions arising from Čebyšev approximation on Cartesian product spaces by ordinary polynomials in $x_1,\dots,x_k$ to functions of form $\sum_{i=1}^{k} F_i(x_i)$.

Definition: A family $\{\varphi^u(x)\}_{u=0,1,\dots}$ of continuous real-valued functions on some compact set X is a Haar sequence* or satisfies the Haar condition if: for any $J \geq 0$, any linear combination $\sum_{u=0}^{J} c_u\varphi^u(x)$ with c_u real and not all zero, has at most J zeroes in X.

Equivalently: $\sum_{u=0}^{J} c_u\varphi^u(x)=0$ for $x=\xi^1,\xi^2,\dots,\xi^{J+1}$ distinct points of X implies

$$c_u = 0 \quad \text{all } u = 0,\dots,J.$$

Approximation by linear combinations of such $\varphi^u(x)$ are of special interest because it is well known (cf. J.R. Rice [5], p.87 ff) that the Haar condition is necessary for the uniqueness of the best approximation even for functions of one variable.

Definition: If $\{\varphi^u(x)\}_{u=0,\dots}$ is a Haar sequence, a Haar polynomial (abbrev. H.p) is any expression of the form $\sum_{u=0}^{J} c_u\varphi^u(x)$. The degree of $\sum_{u=0}^{J} c_u\varphi^u(x)$ is the largest u for which $c_u \neq 0$.

Thus, a H.p. of degree d has at most d distinct zeroes; and if two H.p. of degree $\leq$ d agree at d+1 points, they are identical.

Assume $\{\varphi^u(x)\}_{u=0,\dots}$ is a Haar sequence on X. The proofs of the following Lemmas are immediate, by standard theorems on existence and uniqueness of solutions to systems of linear equations. (Cf. Aitken [1], ch. II).

---

* Cf. Akhiezer [2] p.67 et seq, in which such a family is called a Tchebycheff system with respect to X.

Lemma 1.1 : If $\xi^1,\ldots,\xi^{J+1}$ are distinct values of x, then

$$\begin{vmatrix} \varphi^0(\xi^1) & \varphi^1(\xi^1) & \ldots & \varphi^J(\xi^1) \\ \vdots & \vdots & & \vdots \\ \varphi^0(\xi^{J+1}) & \varphi^1(\xi^{J+1}), & \ldots & \varphi^J(\xi^{J+1}) \end{vmatrix} \neq 0$$

Lemma 1.2 : If $\xi^1,\ldots\xi^{J+1}$ are distinct values of x

and $A_1,\ldots,A_{J+1}$ are real numbers (not necessarily distinct)

Then there exists one and only one H.p. of degree $\leq$ J whose value at each ξ^j is $A_j\varphi^0(\xi^j)$, $j = 1,\ldots,J+1$.

Lemma 1.3 : If $\xi^1,\ldots,\xi^{J+1}$ are distinct values of x

Then there is a unique monic H.p. in x, of degree J, vanishing at ξ^j, $j=1,\ldots,J$.

Proof : The system $\sum_{u=0}^{J} c_u\varphi^u(\xi^j)=0$, $j=1,\ldots,J$ is really

$\sum_{u=0}^{J-1} c_u\varphi^u(\xi^j)= -\varphi^J(\xi^j)$,

which has a unique solution by Lemma 1.1.

Lemma 1.4 : If $\xi^1,\ldots,\xi^J$ are distinct values of x, and $d > J$,

Then there is a unique H.p. $\sum_{u=0}^{d} c_u\varphi^u(x)$ vanishing at $\xi^1,\ldots,\xi^J$, such that $c_d=1$, $c_{d-1}=\ldots=c_J=0$

Proof : Same as for Lemma 1.3.

Related results about the matrices associated with a Haar sequence can be found in Akhiezer [2] p.67 ff.

Suppose now $X_1,\ldots,X_k$ are closed intervals, and that for each $i=1,\ldots,k$, $\{\varphi_i^j(x_i)\}_{j=0,1,\ldots}$ is a Haar sequence on X_i.

Definitions A Haar polynomial (H.p.) in $x_1,\ldots,x_k$ is any finite sum of the form

$$\sum_{\substack{u_i=0 \\ i=1,\ldots,k}}^{d_i} \alpha_{u_1,\ldots,u_k} \varphi_1^{u_1}(x_1)\cdots\varphi_k^{u_k}(x_k)$$ where the α's are real numbers.

The <u>x_i-degree of the H.p.</u> is the largest $\bar{u}_i$ such that

$$\alpha_{u_1,\ldots,u_{i-1},\bar{u}_i,u_{i+1},\ldots,u_k} \neq 0, \text{ for some } u_1,\ldots,\hat{u}_i,\ldots,u_k.$$

The <u>(total) degree of the H.p.</u> is $\max_{\alpha_{u_1,\ldots,u_k}\neq 0} \{u_1 + \ldots + u_k\}$.

The H.p. in $x_1,\ldots,x_k$ of x_i-degree $\leq d_i$ form a vector space of dimension $\prod_{i=1}^{k}(d_i+1)$; moreover the product $\Phi_1(x_1)\cdots\Phi_k(x_k)$, where each $\Phi_i(x_i)$ is a H.p. in x_i, is defined as usual so that any finite sum $\sum_{m=1}^{M}\Phi_1^m(x_1)\ldots\Phi_k^m(x_k)$ is a H.p. in $x_1,\ldots,x_k$. (Note that the product $\varphi_i^u(x_i)\varphi_i^v(x_i)$ is not defined.) Moreover,

<u>Lemma 1.5</u> : Any H.p. $P(x_1,\ldots,x_k)$ can be written in the form

$$\sum_{u=0}^{d_{i_o}} A_u^{i_o}(x_1,\ldots,\hat{x}_{i_o},\ldots,x_k)\varphi_{i_o}^u(x_{i_o})$$

where i_o is any of the $i=1,\ldots,k$; $d_{i_o} = x_{i_o}$-degree of of P; $A_u^{i_o}$ is a H.p. in $x_1,\ldots,\hat{x}_{i_o},\ldots,x_k$.

<u>Proof</u> obvious; same as for ordinary polynomials.

<u>Lemma 1.6</u> : For each $i=1,\ldots,k$ let $\xi_i^1,\ldots,\xi_i^{d_i+1}$ be distinct values of x_i. Then the determinant of order $\prod_{i=1}^{k}(d_i+1)$ whose $((u_1,\ldots,u_k),(\delta_1,\ldots,\delta_k))$-entry* is

* arranged lexicographically : cf. Aitken [1], p.90

$\varphi_1^{u_1}(\xi^{\delta_1})\varphi_2^{u_2}(\xi_2^{\delta_2})\ldots\varphi_k^{u_k}(\xi_k^{\delta_k})$ where $0 \leq u_i \leq d_i$ and

$1 < \delta_i \leq d_i + 1$, $i=1,\ldots,k$, is non-zero.

Proof: Lemma 1.1 and the construction of L.H. Rice [4].

Throughout the preceding there is no requirement that $\varphi^0(x)$ be a constant function, but only that it have no zeroes. Thus, in the case of ordinary polynomials, Lemma 1.2 says a polynomial of degree $d \geq 1$ cannot take on the value A d+1 times.

Clearly, if $\{\varphi^u(x)\}_{u=0,\ldots}$ is a Haar sequence on X, so also is $\left\{\frac{\varphi^u(x)}{\varphi^0(x)}\right\}_{u=0,\ldots}$ and conversely.

Suppose for each $i=1,\ldots,k$ $\{1=\varphi_i^0(x_i),\varphi_i^1(x_i),\varphi_i^2(x_i),\ldots\}$ is a Haar sequence on X_i, and $\overline{\varphi}_i(x_i)$ is a continuous, real-valued function on X_i having no zeroes. Let $P(x_1,\ldots,x_k)$ be a H.p. so

$$P(x_1,\ldots,x_k) = \sum_{\substack{u_i=0 \\ 1 \leq i \leq k}}^{d_i} \alpha_{u_1,\ldots,u_k} \varphi_1^{u_1}(x_1)\ldots\varphi_k^{u_k}(x_k).$$

Define $\overline{P}(x_1,\ldots,x_k) = \sum_{\substack{u_i=0 \\ 1 \leq i \leq k}}^{d_i} \alpha_{u_1,\ldots,u_k} \overline{\varphi}_1^{u_1}(x_1)\cdots\overline{\varphi}_k^{u_k}(x_k)$ where

$\overline{\varphi}_i^u(x_i) = \varphi_i^u(x_i)\cdot\overline{\varphi}_i(x_i)$.

Then $\overline{P}(x_1,\ldots,x_k) = P(x_1,\ldots,x_k)\cdot \prod_{i=1}^{k} \overline{\varphi}_1(x_i)$. For any subset S of $X_1 x\ldots xX_k$, $P(x_1,\ldots,x_k)$ vanishes on S if and only if $\overline{P}(x_1,\ldots,x_k)$ vanishes on S.

It follows that with no loss in generality it can be assumed that $\varphi_i^0(x_i)=1$, each $i=1,\ldots,k$, and that assumption will be made from here on.

The following are direct consequences of Lemma 1.6:

<u>Lemma 1.7</u> : For each $i=1,\ldots,k$ let $\xi_i^1,\ldots,\xi_i^{d_i+1}$ be distinct values of x_i. Let $C_{\delta_1,\ldots,\delta_k}$ $(1 \le \delta_i \le d_i+1)$ be $\prod_{i=1}^{k}(d_i+1)$ numbers not necessarily distinct. Then there exists a unique H.p. $P(x_1,\ldots,x_k)$ of x_i-degree d_i such that

$$P(\xi_1^{\delta_1},\ldots,\xi_k^{\delta_k}) = C_{\delta_1,\ldots\delta_k}.$$

<u>Lemma 1.8</u> : In particular, if all $C_{\delta_1,\ldots,\delta_k}$ in Lemma 1.7 are zero, $P(x_1,\ldots,x_k)$ vanishes term-by-term: all $\alpha_{u_1,\ldots,u_k} = 0$.

<u>Lemma 1.9</u> : If two H.p. in $x_1,\ldots,x_k$, each of which has x_i-degree $\le d_i$ $(i=1,\ldots,k)$, agree on the $\prod_{i=1}^{k}(d_i+1)$ k-tuples of Lemma 1.7, then they are identical.

<u>Lemma 1.10</u> : Let $P(x_1,\ldots,x_k)$ be a H.p. and suppose P to have been represented as in Lemma 1.5, for some fixed i_o. Then $P \equiv 0$ if and only if $A_u^{i_o}(x_1,\ldots,\hat{x}_{i_o},\ldots x_k) \equiv 0$ each $u=0,\ldots,d_{i_o}$.

<u>Proof</u> Induction on k, using Lemma 1.8.

<u>Definition</u> . A continuous real-valued function $F(x_1,\ldots,x_k)$ on $X_1 x\ldots xX_k$ is <u>separated</u> if it can be written $F_1(x_1)+\ldots+F_k(x_k)$ where each $F_i(x_i)$ is continuous on X_i. The function $F_i(x_i)$ is the (i^{th}) <u>separate component</u> of F.

Observe that if $P(x_1,\ldots,x_k)$ is a separated H.p. on $X_1 x\ldots xX_k$, then the i^{th} separate component of P is a H.p. also.

Let N be any non-negative integer. For each $1 \le i \le k$ let there be given a closed interval X_i and two sets of points Σ_i^+ and and Σ_i^- in X_i, which separate each other, such that the total number of points in Σ_i^+ and Σ_i^- together is N+2. Thus, if N is even, $N+2=2r_i$, so Σ_i^+ and Σ_i^- each contain r_i points; whereas, if N is odd, $N+2=2s_i+1$ so one set contains s_i points and the other s_i+1.

Since each family $\{1,\varphi_i^1(x_i),\ldots,\varphi_i^N(x_i)\}$ satisfies the Haar condition on X_i, it follows (cf. Akhiezer [1] p.74 ff) that for any function $F_i(x_i)$ real-valued and continuous on X_i there exists a unique H.p. of degree $\leq$ N of least Čebyšev deviation from $F_i(x_i)$ on X_i. The (strong) extremal signatures for $\{1,\varphi_i^1,\ldots,\varphi_i^N\}$ are precisely of the form $\Sigma_i^+ \cup \Sigma_i^-$.

Let $\Sigma^+=\Sigma_1^+ \times\ldots\times \Sigma_k^+$, $\Sigma^-= \Sigma_1^- \times\ldots\times \Sigma_k^-$. The construction of [3], §2 applies here, so we have

Theorem 1 : For each $1 \leq i \leq k$, let X_i be a closed interval, let $F_i(x_i)$ be a continuous real-valued function on X_i, let $P_i^*(x_i)$ be the H.p. of degree $\leq$ N of least Čeybšev deviation from $F_i(x_i)$ on X_i. Then among all H.p. $P(x_1,\ldots,x_k)$ of degree $\leq$ N there is none whose Čebyšev deviation from $F(x_1,\ldots,x_k)=\sum_{i=1}^{k} F_i(x_i)$ on $X_1\times\ldots\times X_k$ is less than that of $\sum_{i=1}^{k} P_i^*(x_i)$.

That is, if $\Sigma_i^+ \cup \Sigma_i^-$ is an extremal signature for $\{1,\varphi_i^1,\ldots,\varphi_i^N\}$ then $\Sigma^+ \cup \Sigma^-$ is an extremal signature for the set $\{\varphi_1^{u_1}(x_1)\ldots\varphi_k^{u_k}(x_k);\ u_i \geq 0,\ u_1+\ldots+u_k \leq N\}$.

Section 2

We shall now prove

Theorem 2 : If $P(x_1,\ldots,x_k)$ is a H.p. of degree $\leq$ N which vanishes on Σ^+ and on Σ^- then $P \equiv 0$.

There will then follow immediately

Theorem 3 : The H.p. $\sum_{i=1}^{k} P_i^*(x_i)$ of Theorem 1 is the unique H.p. of degree $\leq$ N of least deviation from $\sum_{i=1}^{k} F_i(x_i)$ on $X_1\times\ldots\times X_k$. That is, $\Sigma^+ \cup \Sigma^-$ is a strong extremal signature.

(The terminology of the preceding follows [3]).

The proof of Theorem 2 is based upon several lemmas.

Suppose first that for each $i=1,\ldots,k$ a non-empty set of points S_i is given, call them $\xi_i^1,\xi_i^2,\ldots,\xi_i^{r_i}$ (all distinct). Let $\mathcal{S}_i$ = {all H.p. in x_i vanishing on S_i}. Observe that no non-trivial (i.e., non-zero) H.p. in x_i of degree $< r_i$ belongs to $\mathcal{S}_i$.

Lemma 2.1 : Let $P(x_1,\ldots,x_k)$ be a H.p. and suppose that for any choice of $\xi_1,\ldots,\xi_{k-1}$, $P(\xi_1,\ldots,\xi_{k-1},x_k)$ vanishes at each point of S_k. Then there exists a finite collection $\Phi_k^1(x_k),\Phi_k^2(x_k),\ldots\ldots,\Phi_k^t(x_k)$ of H.p. in $\mathcal{S}_k$, and also H.p. $B^1(x_1,\ldots,x_k)$, $B^2(x_1,\ldots,x_k),\ldots,B^t(x_1,\ldots,x_k)$ of x_k-degree zero, such that $\sum_{\ell=1}^{t} B^\ell(x_1,\ldots,x_k)\ \Phi_k^\ell(x_k) = P(x_1,\ldots,x_k)$.

Proof : (By induction on k.) If k=1, statement is obvious, because we assumed $\varphi_i^0(x_i)=1$. Assume it is true for H.p. in k-1 variables; we will show it is true for k. Let $P(x_1,\ldots,x_k)$ be a H.p. satisfying the hypothesis. Let d_1 be the x_1-degree of P; let $\xi_1^1,\ldots,\xi_1^{d_1+1}$ be distinct values of x_1. $P(\xi_1^j,x_2,\ldots,x_k)$ vanishes at each point of S_k, for every choice of $x_2,\ldots,x_{k-1}$, every $1 \leq j \leq d_1+1$. By the inductive hypothesis,

$$P_j(x_2,\ldots,x_k)=P(\xi_1^j,x_2,\ldots,x_k)=\sum_{\ell=1}^{t_i} B_j^\ell(x_2,\ldots,x_k)\cdot \Phi_{k,j}^\ell(x_k)$$

where $\Phi_{k,j}^\ell \in \mathcal{S}_k$ and B_j^ℓ has x_k-degree zero. Next, let $\Omega_1^j(x_1)$ be the H.p. of degree d_1 which is 1 at ξ_1^j and 0 at $\xi_1^{\bar{j}}$ $(\bar{j} \neq j)$ $j=1,\ldots,d_1+1$ (Lemma 1.7); Ω_1^j obviously has x_k-degree zero. Let $Q(x_1,\ldots,x_k) = \sum_{j=1}^{d_1+1} P_j(x_2,\ldots,x_k)\cdot\Omega_1^j(x_1)$; Q has x_1-degree $\leq d_1$. Q agrees with P for all values of

$x_2,\ldots,x_k$ in each of the d_1+1 values of x_1, hence $P \equiv Q$ by Lemma 1.9.

$\therefore P(x_1,\ldots,x_k)$ has a representation of the desired form.

Let I_k be the set of all H.p. of the form $\sum_{i=1}^{k}(\sum_{\ell_i=1}^{t_i} B_i^{\ell_i}(x_1,\ldots,x_k)\,\Phi_i^{\ell_i}(x_i))$

$$\left.\begin{array}{l}\text{where every } \Phi_i^{\ell_i}(x_i) \in \mathcal{S}_i \\ \text{and } x_i\text{-degree of } B_i^{\ell_i} \text{ is zero}\end{array}\right\} \quad i=1,\ldots,k$$

By Lemmas 1.5 and 2.1 every H.p. in I_k can be written

$\sum_{m=1}^{M} \theta_1^m(x_1)\theta_2^m(x_2)\ldots\theta_k^m(x_k)$, where $\theta_i^m(x_i)$ is a H.p.,

and, for each $m \;\exists\; i_m \ni \theta_{i_m}^m(x_{i_m}) \in \;\; _{i_m}$.

Clearly, every H.p. in I_k vanishes on $S_1 \times\ldots\times S_k$.

<u>Lemma 2.2</u> : The set of all Haar polynomials vanishing on $S_1\times\ldots\times S_k$ is precisely the set I_k.

<u>Proof</u> : In view of the immediately preceding remarks, it will suffice to show:

$P(x_1,\ldots,x_k)$ vanishes on $S_1 \times\ldots\times S_k$ implies $P\in I_k$.

For k=1, assertion is obviously true; assume it is true for (k-1) variables. Let $P(x_1,\ldots,x_k)$ vanish on $S_1\times\ldots\times S_k$; Let $\xi_k^1,\xi_k^2,\ldots,\xi_k^{r_k}$ be the pts. of S_k. For $j=1,\ldots,r_k$, $P(x_1,\ldots,x_{k-1},\xi_k^j)$ vanishes on $S_1\times\ldots\times S_{k-1}$, hence, applying the inductive hypothesis,

$P_j(x_1,\ldots,x_{k-1})=P(x_1,\ldots,x_{k-1},\xi_k^j) \in I_{k-1}$

and has a representation of form

$\sum_{i=1}^{k-1}(\sum_{\ell_{i,j}=1}^{t_{i,j}} B_{i,j}^{\ell_{i,j}}(x_1,\ldots,x_{k-1})\Phi_{i,j}^{\ell_{i,j}}(x_i))$

where each $\Phi_{i,j}^{\ell_{i,j}}(x_i) \in \mathcal{S}_i$, and x_i-degree $B_{i,j}^{\ell_{i,j}}(x_1,\ldots,x_{k-1})$ is zero $\begin{cases} i=1,\ldots,k-1 \\ j=1,\ldots,r_k \end{cases}$.

Now, let $\Omega_k^j(x_k)$ be the x_k-Haar polynomial of degree r_k which is 1 at ξ_k^j and 0 at $\xi_k^{\bar{j}}(\bar{j} \neq j)$, $j = 1,\ldots,r_k$ (as in Lemma 2.1).

Form

$$Q(x_1,\ldots,x_k) = \sum_{j=1}^{r_k} P_j(x_1,\ldots,x_{k-1}) \cdot \Omega_k^j(x_k).$$

$P(\xi_1,\xi_2,\ldots,\xi_{k-1},x_k) - Q(\xi_1,\xi_2,\ldots,\xi_{k-1},x_k)$ vanishes at each point of S_k for <u>every</u> choice $x_1 = \xi_1, x_2 = \xi_2, \ldots, x_{k-1} = \xi_{k-1}$, because it is $P(\xi_1,\ldots,\xi_{k-1},x_k) - \sum_{j=1}^{r_k} P(\xi_1,\ldots,\xi_{k-1},\xi_k^j)\cdot\Omega_k^j(x_k)$, so if $x_k = \xi_k^j$, the expression becomes

$P(\xi_1,\ldots,\xi_{k-1},\xi_k^j) - P(\xi_1,\ldots,\xi_{k-1},\xi_k^j)\cdot 1 = 0.$

Thus, by Lemma 2.1,

$P(x_1,\ldots,x_k) - Q(x_1,\ldots,x_k) = \sum_{\ell_k=1}^{t_k} B_k^{\ell_k}(x_1,\ldots,x_k)\cdot\Phi_k^{\ell_k}(x_k)$ where $\Phi_k^{\ell_k}(x_k) \in \mathcal{S}_k$ and x_k-degree of $B_k^{\ell_k}$ is zero.

Since $B_{i,j}^{\ell_{i,j}}$ has x_1-degree zero, $i=1,\ldots,k-1$, and $\Omega_k^{\ell_k}(x_k)$ has x-degree zero, so does $B_{i,j}^{\ell_{i,j}} \cdot \Omega_k^{\ell_k}$.

$\therefore P(x_1,\ldots,x_k)$ has a representation of the desired form.

<u>Lemma 2.3</u>: Let I_k be the set defined in Lemma 2.2.
If $P(x_1,\ldots,x_k) \in I_k$ and the x_i-degree of P is $< r_i$ for <u>each</u> $i=1,\ldots,k$, then $P \equiv 0$.

Proof: By induction on k.

If k=1, we already know $P(x_1)=0$ for $x_1 = \xi_1^1,\ldots,\xi_1^{r_1}$

implies deg $P \geq r_1$ or $P \equiv 0$.

Assume true for k-1. Suppose $P(x_1,\ldots,x_k) \in I_k$.

Let $\xi_k^1,\ldots,\xi_k^{r_k}$ be the points of S_k.

Then it can readily be seen, from Lemma 1.5, that for each $j=1,\ldots,r_k$, $P_j(x_1,\ldots,x_{k-1}) = P(x_1,\ldots,x_{k-1},\xi_k^j)$ has x_i-degree which is $\leq x_i$-degree $P(x_1,\ldots,x_k) < r_i$ for each $i=1,\ldots,k-1$. ∴ By the inductive assumption, $P_j(x_1,\ldots,x_{k-1}) \equiv 0$, $j=1,\ldots,r_k$. By Lemma 2.1,

$$P(x_1,\ldots,x_k) = \sum_{\ell=1}^{t} B^\ell(x_1,\ldots,x_{k-1})\cdot\Phi_k^\ell(x_k) \quad \begin{cases} \text{where } \Phi_k^\ell \in \mathcal{S}_k, \\ B^\ell \text{ is a H.p. in} \\ x_1,\ldots,x_{k-1} \end{cases}$$

which, by Lemma 1.5, $= \sum_{u=0}^{d} A_u(x_1,\ldots,x_{k-1})\varphi_k^u(x_k)$

$$\begin{cases} \text{where } A_u \text{ is a H.p.} \\ \text{in } x_1,\ldots,x_{k-1} \text{ and} \\ d \leq x_k\text{-degree of } P \\ \quad < r_k \end{cases}$$

Suppose $P \not\equiv 0$. Then by Lemma 1.10 there exists some u_o and some $x_1=\xi_1,\ldots,x_{k-1}=\xi_{k-1}$ ∋ $A_{u_o}(\xi_1,\ldots,\xi_{k-1}) \neq 0$.

∴ $P(\xi_1,\ldots,\xi_{k-1},x_k) = \sum_{u=0}^{d} A_u(\xi_1,\ldots,\xi_{k-1})\,\varphi_k^u(x_k)$ is a Haar polynomial in x_k, not all of whose coefficients are zero, of degree $< r_k$, vanishing on S_k. This contradicts the Haar condition. ∴ $P \not\equiv 0$

Lemma 2.4: Let I_k be the set of Lemma 2.2. Let $P(x_1,\ldots,x_k) \in I_k$. Then there exist, for each $i=1,\ldots,k$, Haar polynomials $\Phi_i^{\ell_i}(x_i) \in \mathcal{S}_i$, $\ell_i=1,\ldots,t_i$, and Haar polynomials $B_i^{\ell_i}(x_1,\ldots,x_k)$ of x_i-degree zero such that $P(x_1,\ldots,x_k) = \sum_{i=1}^{k} (\sum_{\ell_i=1}^{t_i} B_i^{\ell_i} \Phi_i^{\ell_i})$ and

$\deg B_i^{\ell_i} \Phi_i^{\ell_i} \leq \deg P$ for all $i=1,\ldots,k$, all $\ell_i=1,\ldots,t_i$.

Proof : Let $\tilde{I}_k$ be the subset of I_k consisting of Haar polynomials P which admit such a representation; suppose $I_k - \tilde{I}_k$ is not empty.

Let d be the minimal degree of all H.p. in $I_k - \tilde{I}_k$. Since every $r_i \geq 1$, we conclude from Lemma 2.3 that $d \geq 1$. Among the H.p. of degree d in $\tilde{I}_k - I_k$ choose those with a minimal number of terms in the leading form; among these, choose those with a minimal number of terms in the next leading form, etc. Call the H.p. so chosen $Q(x_1,\ldots,x_k)$.

$Q \neq 0$. Therefore, by Lemma 2.3, there is an index i_o and a term $c\,\varphi_1^{u_1}(x_1)\,\varphi_2^{u_2}(x_2)\ldots\varphi_k^{u_k}(x_k)$ for which $u_{i_o} \geq r_{i_o}$.

Note $u_1+\ldots+u_k \leq d=\deg Q$. Let $\Gamma_{i_o}^{u_{i_o}}(x_{i_o})$ be the monic H.p. of degree u_{i_o} whose zeroes include the points $\xi_{i_o}^1\ldots,\xi_{i_o}^{r_{i_o}}$ of S_{i_o}, and whose $r_{i_o},\ldots,u_{i_o}-1$ degree terms are absent; (Lemma 1.4); if $u_{i_o}=r_{i_o}$, Γ_{i_o} is the unique H.p. of Lemma 1.3. Consider

$$Q(x_1,\ldots,x_k)-c\varphi_1^{u_1}(x_1)\varphi_2^{u_2}(x_2)\ldots\Gamma_{i_o}^{u_{i_o}}(x_{i_o})\ldots\varphi_k^{u_k}(x_k) = R(x_1,\ldots,x_k).$$

R is certainly in I_k; it differs from Q in having one less term of degree $u_1+\ldots+u_k$, but it has the same number of terms of higher degree.

Moreover, R is in $I_k - \tilde{I}_k$: Suppose R has a representation $\sum_i \sum_{\ell_i} B_i^{\ell_i} \Phi_i^{\ell_i}$ with $\deg B_i^{\ell_i} \Phi_i^{\ell_i} \leq \deg R \leq \deg Q$ all i, all ℓ_i.

Since $c\,\varphi_1^{u_1}(x_1)\ldots\Gamma_{i_o}^{u_{i_o}}(x_{i_o})\ldots\varphi_k^{u_k}(x_k)$ clearly has such a representation (because $u_1+\ldots+u_k \leq d = \deg Q$), it follows that Q has a representation and so is in $\tilde{I}_k$. This contradicts the earlier assumptions for Q.

$\therefore I_k - \tilde{I}_k$ is empty.

Lemma 2.5 : For any $1 \leq i \leq k$: Let $N \geq r_i$; let $\{\Gamma_i^{\omega_i}(x_i)\}_{\omega_i=r_i,\ldots,N}$ be any set of Haar polynomials in x_i, such that $\Gamma_i^{\omega_i}(x_i)$ is monic, of degree precisely ω_i, and vanishes on S_i. Let $\Phi_i(x_i)$ be any Haar polynomial vanishing on S_i, of degree $\leq N$ (and $\geq r_i$). Then there is a unique $(N-r_i-1)$-tuple of real numbers $(\beta_{r_i},\ldots,\beta_N) \ni$

$$\Phi_i(x_i) = \sum_{\omega_i=r_i}^{N} \beta_{\omega_i}\, \Gamma_i^{\omega_i}(x_i).$$

Proof : The uniqueness follows, as usual, from Lemma 1.2. To establish the existence, observe $\Phi_i(x_i)=\sum_{u=0}^{N} c_u\, \varphi_i^u(x_i)$ where $c_{r_i}, c_{r_{i+1}}, \ldots, c_n$ are not all zero.

Proceed by induction on $N-r_i$:

If $N-r_i=0$, $N=r_i$ and $c_{r_i} \neq 0$. $\Phi_i(x_i)=\sum_{u=0}^{r_i} c_u\, \varphi_i^u(x_i)$, and $\Phi_i(x_i) - c_{r_i}\Gamma_i^{r_i}(x_i)$ is a Haar polynomial of degree $\leq r_i-1$ which vanishes on S_i, hence is idenically zero. $\therefore \Phi_i(x_i)=c_{r_i}\,\Gamma_i^{r_i}(x_i)$. Next, assume proven for $N-r_i \leq n-1$, and suppose $N=r_i+n$.

Then $\Phi_i(x_i)-c_{r_i+n}\Gamma_i^{r_i+n}(x_i)$ is a Haar polynomial of degree $\le r_i+(n-1)=N-1$, hence by the inductive assumption has a representation $\sum_{\omega_i=r_i}^{N-1}\beta_{\omega_i}\Gamma_i^{\omega_i}(x_i)$.

$\therefore \Phi_i(x_i) = c_{r_i+n}\Gamma_i^{r_i+n}(x_i)+\sum_{\omega_i=r_i}^{N-1}\beta_{\omega_i}\Gamma_i^{\omega_i}(x_i)$, and letting $\beta_N=c_{r_{i+n}}$ we have the desired form.

In particular, we could suppose the $\Gamma_i^{\omega_i}(x_i)$ to be the Haar polynomials of Lemma 1.4.

Combining Lemmas 2.2, 2.4 and 2.5 we have

Corollary 2.6 : Given $P(x_1,\ldots,x_k)$ of degree $\le N$, vanishing on $S_1x\ldots xS_k$, there is a representation

$$\sum_{i=1}^{k}\Big(\sum_{\omega_i=r_i}^{N} A_i^{\omega_i}(x_1,\ldots,\hat{x}_i,\ldots x_k)\,\Gamma^{\omega_i}(x_i)\Big)$$

such that $\deg A_i^{\omega_i}\Gamma_i^{\omega_i}\le N$.

$$\deg \Gamma_i^{\omega_i} = \omega_i$$

$$x_i\text{-}\deg A_i^{\omega_i} = 0.$$

Proof : With the notation of Lemmas 2.4 and 2.5,

$$A_i^{\omega_i} = \sum_{\ell_i=1}^{t_i} B_i^{\ell_i}\beta_{\omega_i}^{\ell_i} \qquad \begin{cases} i=1,\ldots,k \\ r_i\le\omega_i\le N \end{cases}.$$

Now, suppose Σ_i^+ and Σ_i^-, r_i and s_i, Σ^+ and Σ^- are as specified in Section 1. Let $P(x_i,\ldots,x_k)$ be a Haar polynomial of degree $\le N$ which vanishes on Σ^+ and Σ^-. Applying Corollary 2.6, we can write

$$P(x_1,\ldots,x_k)=\sum_{i=1}^{k}\Big(\sum_{\omega_i=\rho_i}^{N} A_i^{\omega_i}(x_1,\ldots,x_k)\Phi_i^{\omega_i}(x_i)\Big)$$

$$=\sum_{i=1}^{k}\Big(\sum_{\omega_i=\sigma_i}^{N} B_i^{\omega_i}(x_1,\ldots,x_k)\Psi_i^{\omega_i}(x_i)\Big)$$

where

$$(*)\left\{\begin{array}{l}\rho_i = \text{cardinality of } \Sigma_i^+,\ \sigma_i = \text{cardinality of } \Sigma_i^- \\ \Phi_i^{\omega_i}(x_i) \text{ vanishes on } \Sigma_i^-,\ \Psi_i^{\omega_i}(x_i) \text{ vanishes on } \Sigma_i^- \\ \text{degree } \Phi^{\omega_i} = \text{degree } \Psi_i^{\omega_i} = \omega_i \text{ precisely} \\ x_i\text{-degree } A_i^{\omega_i} = x_i\text{-degree } B_i^{\omega_i} = 0 \\ \text{degree } A_i^{\omega_i}\ \Phi_i^{\omega_i} \leq N,\ \text{degree } B_i^{\omega_i}\ \Psi_i^{\omega_i} \leq N \end{array}\right\}$$

for each $1 \leq i \leq k$;
each ω_i .

There are two cases, according to the parity of N: For N even, $N+2=2r_i$, $\rho_i=\sigma_i=r_i$, and $r_i \leq \omega_i \leq N$ implies $0 \leq N - \omega_i \leq N - r_i = r_i-2$. $\therefore$ degree $A_i^{\omega_i}$, degree $B_i^{\omega_i} \leq r_i-2$. For N odd, $N+2=2s_i+1$, either $\rho_i=s_i$ and $\sigma_i=s_i+1$, or vice versa. $s_i \leq \omega_i \leq N$ implies $0 \leq N-\omega_i \leq N-s_i=s_i-1$, and $s_i+1 \leq N-\omega_i \leq N-(s_i+1)=s_i-2$ $\therefore$ degree $A_i^{\omega_i} \leq s_i-1$ and degree $B_i^{\omega_i} \leq s_i-2$, or vice versa.

[Lemma 2.3 implies we may suppose $N > 0$: for, if $N = 0$, and $\rho_i, \sigma_i = 1$ then $P \equiv 0$].

We will argue by induction on k. The case k=2 is sufficiently interesting and instructive to warrant a separate exposition.

If k=1, the hypothesis says $P(x_1)$ vanishes on N+2 points, yet is of degree $\leq N$, hence $P \equiv 0$ by the Haar condition.

In order to establish the proposition in case k=2 we first make some general observations.

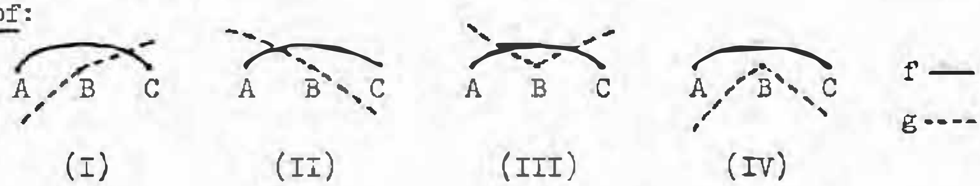
<u>Definition: A function f has an odd zero at ξ</u> if $f(\xi)=0$ and f changes sign at ξ.

<u>A function f has an even zero at ξ</u> if $f(\xi)=0$ and f does not change sign at ξ.

<u>Sublemma 2.7:</u> Given 3 distinct points A,B,C in the real line such that $A < B < C$ and two functions f and g continuous on

$[A,C]$; suppose that $f(A)=g(B)=f(C)=0$, but that neither f nor g has a zero at any other point of $[A,C]$. Then, if B is an odd zero of g, $f-g$ has at least one zero in (A,C); but, if B is an even zero of g, $f-g$ may have two or no zeros in (A,C).

Proof:



W.L.O.G. we may suppose that $f(x) > 0$ for all $A < x < C$. There are four cases, illustrated above:

(I) g changes from negative to positive at B,

$\therefore$ $g(x) < 0$ in $[A,B)$ and $g(x) > 0$ in $(B,C]$.

$\therefore$ $(f-g)(B) > 0$ and $(f-g)(C) < 0$ hence $f-g$ has a zero in $(B,C) \subseteq (A,C)$.

(II) g changes from positive to negative at B: same as (I) mutatis mutandis.

(III) $g(x) > 0$ all $x \in [A,B) \cup (B,C]$.

$(f-g)(A) < 0$ and $(f-g)(C) < 0$, but $(f-g)(B) > 0$.

$\therefore$ $f-g$ has a zero in (A,B) and a zero in (B,C).

(IV) $g(x) < 0$ all $x \in [A,B) \cup (B,C]$. Then $(f-g)(x) > 0$ all $x \in [A,C]$, so $f-g$ has no zeroes in (A,C).

Next, given $A_1 < B_1 < A_2 < \ldots < A_{t-1} < B_{t-1} < B_t$ and functions f and g continuous on $[A_1,A_t]$; suppose f has zeroes precisely at the A_j and g has zeroes precisely at the B_j. From Lemma 2.7 it is easy to see that the number of zeroes of $f-g$ in $[A_1,A_t]$ is $\geq (t-1)-m$ where m is the number of even zeroes of g among $B_1,\ldots,B_{t-1}$. On the other hand, suppose f has zeroes at the A_j

and possibly at some of the B_j (but nowhere else in $[A_1,A_t]$), and g has zeroes at the B_j and possibly at some of the A_j (but nowhere else in $[A_1,A_t]$); then the argument of Lemma 2.7 shows that in any (A_j,A_{j+1}) $j=1,\ldots,t-1$, f-g has at least one, or possibly no or two zeroes, according as g has an odd or even zero at B_j. Therefore the number of zeroes of f-g in $[A_1,A_t]$ is still $\geq (t-1)-m$ as before. Observe finally that if g has more than one zero between A_j and A_{j+1}, then f-g can have no zeroes in (A_j,A_{j+1}) only if g has an even number of such zeroes. That is, in the foregoing, we can replace "g has an odd zero in (A_j,A_{j+1})" by "g has an odd number of zeroes in (A_j,A_{j+1})" and "g has an even zero in (A_j,A_{j+1})" by "g has an even number of zeroes in (A_j,A_{j+1})." Moreover, if $A_1' \leq A_1 < B_1 < A_2' < A_2 < B_2 \ldots < A_{t-1}' \leq A_{t-1} < B_{t-1} < A_t' \leq A_t$, and if f and g are continuous on $[A_1',A_t]$ and if g has no zeroes in any (A_j',A_j), then the number of zeroes of f-g in $[A_1',A_t]$ is $\geq$ number of zeroes of f-g in $\bigcup_{j=1}^{t-1} [A_j,A_{j+1}']$. Hence the number of zeroes of f in $\bigcup_{j=1}^{t} (A_j',A_j)$ does not alter the earlier inequality.

From Sublemma 2.7 and the corollary remarks, we conclude

<u>Lemma 2.8:</u> Let Σ_x^+ and Σ_x^- be sets of points which separate each other, entirely contained in some closed bounded real interval X. Let F^+ be a function continuous on X, vanishing on Σ^+

F^- " " " " " " " " Σ^-

(I) If card (Σ_x^+) = card $(\Sigma_x^-) = \tau$, and F^+ has precisely $\tau + \varkappa$ zeroes and F^- has $\leq \tau + \varkappa$ zeroes [counting an even zero as two zeroes and an odd as one] then $F^+ - F^-$ has $\geq (\tau-1)-\varkappa$ zeroes.

(II) If card $(\Sigma_x^+) = \tau + 1$, card $(\Sigma_x^-) = \tau$, F^- has precisely $\tau + \varkappa$ zeroes and F^+ has $\leq \tau + \varkappa$ zeroes, then $F^+ - F^-$ has $\geq ((\tau + 1)-1) - \varkappa = \tau - \varkappa$ zeroes.

(III) If card $(\Sigma_x^+) = \tau + 1$, card $(\Sigma_x^-) = \tau$, F^+ has precisely

$\tau + \varkappa$ zeroes and F^- has $< \tau + \varkappa$ zeroes, i.e., $\tau + M$ where $M \leq \varkappa - 1$, then $F^+ - F^-$ has $\geq ((\tau + 1)-1) - M$ which is $\geq ((\tau + 1)-1) - (\varkappa-1) = \tau-\varkappa+1 > \tau-\varkappa$ zeroes. (In (II) and (III), Σ_x^+ plays the rôle of the A's, and $t=\tau+1$; in (I), Σ_x^- plays the role of the A's, and $t=\tau$.)

We now proceed with the proof of Theorem 2 for k=2. By (*) p.   we have $P(x,y) = S_x^+ + S_y^+ = S_x^- + S_y^-$, where

$$\begin{cases} S_x^+ \text{ vanishes on } \Sigma_x^+ \\ S_y^+ \quad " \quad " \ \Sigma_y^+ \\ S_x^- \text{ vanishes on } \Sigma_x^- \\ S_y^- \quad " \quad " \ \Sigma_y^- \end{cases}$$

Assume not all of these summands vanish identically.

Suppose N even, N = 2r - 2:

For S_x^+ and S_x^- the x-degree $\geq r$ and hence the y-degree $\leq r - 2$.

For S_y^+ and S_y^- the y-degree $\geq r$ and hence the x-degree $\leq r - 2$. But $S_x^+-S_x^- = S_y^--S_y^+$ therefore has x-degree $\leq r - 2$, and so, by Lemma 2.8 (I) at least one of S_x^+, S_x^- has x-degree $\geq r + 1$.

Observe that x-degree of P = x-degree of S_x^+ = x-degree of S_x^- [similarly for y], because no cancellation of terms of degree $\geq r$ can be effected by S_y^+ or S_y^-. $\therefore$ Both S_x^+ and S_x^- have x-degree $\geq r + 1$. In precisely similar fashion, both S_y^+ and S_y^- have y-degree $\geq r + 1$, hence x-degree $\leq r - 3$, so $S_x^+ - S_x^-$ has x-degree $\leq r - 3$.

Suppose it has already been shown that S_x^+ and S_x^- have x-degree $\geq r+m$ [resp.y]. Then S_y^+, S_y^- and $S_y^--S_y^+=S_x^+-S_x^-$ have x-degree $\leq r-m-2$, so by Lemma 2.8 S_x^+ and S_x^- both have x-degree $\geq r+m+1$ [resp.y]. Since this is true for $m \geq 0$, let $m=r-2$ so S_x^+ and S_x^- have x-degree $\geq r+(r-1) > N$. But this contradicts Lemma 2.6. $\therefore S_x^+ = S_x^- \equiv 0$, $S_y^+ = S_y^- \equiv 0$.

Suppose N is odd, so N=2s-1:

One of S_x^+, S_x^- has x-degree $\geq$ s+1, y-degree $\leq$ s-2; the other has x-degree $\geq$ s, y-degree $\leq$ s-1: as before, both have x-degree $\geq$ s+1, y-degree $\leq$ s-2. Likewise, one of S_y^+, S_y^- has y-degree $\geq$ s+1, x-degree $\leq$ s-2; the other has y-degree $\geq$ s, x-degree $\leq$ s-1: $\therefore$ both have y-degree $\geq$ s+1, x-degree $\leq$ s-2. Using Lemma 2.8 (II) or (III) exactly as in the case for N even, we now conclude $S_x^+ = S_y^+ = S_x^- = S_y^- \equiv 0$.

This concludes the special case k=2.

Let k > 2. Assume Theorem 2 has been proved for all Haar polynomials in $\leq$ (k-1) variables. Given $P(x_1,\ldots,x_k)$ written in form (*). Then

$$\sum_{\omega_k=\rho_k}^{N} A_k^{\omega_k} \Phi_k^{\omega_k}(x_k) - \sum_{\omega_k=\sigma_k}^{N} B_k \Psi_k^{\omega_k}$$

$$= \sum_{i=1}^{k-1} \Big(\sum_{\omega_i=\sigma_i}^{N} B_i^{\omega_i} \Psi_i^{\omega_i}(x_i) \Big) - \sum_{i=1}^{k-1} \Big(\sum_{\omega_i=\rho_i}^{N} A_i^{\omega_i} \Phi_i^{\omega_i}(x_i) \Big).$$

For any fixed values $x_1=\xi_1, x_2=\xi_2, \ldots, x_{k-1}=\xi_{k-1}$ the left-hand side is a difference of Haar polynomials in x_k, vanishing on Σ_k^+, Σ_k^- resp.; the right-hand side has x_k-degree $\leq r_k-2$ [s_k-2] if N is even [odd], hence by Lemma 2.8, each sum on the left has x_k-degree $\geq r_k+1$ [s_k+2]. Hence x_i-degree of $A_k^{\omega_k}$ is $\leq N-(r_k+1)=r_k-3$ [$N-(s_k+2)=s_k-3$] each i=1,...,k-1. But now, by a symmetrical argument, it is clear that x_k-degree of $A_i^{\omega_i} \leq r_i-3$ [s_k-3], and likewise for x_k-degree of $B_i^{\omega_i}$. Proceeding as for k=2, we have

$$\sum_{\omega_k=\rho_k}^{N} A_k^{\omega_k}(\xi_1,\ldots,\xi_{k-1})\Phi_k^{\omega_k}(x_k) = \sum_{\omega_k=\sigma_k}^{N} B_k^{\omega_k}(\xi_1,\ldots,\xi_{k-1})\Psi_k^{\omega_k}(x_k) \equiv 0$$

for every $x_1=\xi_1,\ldots,x_{k-1}=\xi_{k-1}$. Hence $A_k^{\omega_k} \equiv 0$, $B_k^{\omega_k} \equiv 0$ and the two sums on the right-hand side above are identically equal. Fix

$x_k = \xi_k$ arbitrarily, and apply the inductive assumption: then the sums with $x_k = \xi_k$ vanish identically. But ξ_k was arbitrary. ∴ The (original) sums on the right-hand side vanish identically. ∴ $P \equiv 0$.

QED Theorem 2

Section 3

Theorem 2 can be regarded as a result about the rank of certain matrices, as follows:

Consider the configuration $\Sigma = \Sigma^+ \cup \Sigma^- = (\Sigma_1^+ \times \ldots \times \Sigma_k^+) \cup (\Sigma_1^- \times \ldots \times \Sigma_k^-)$ as previously defined. We denote by $\gamma_{N,k}$ the number of (lattice)-points in Σ. If N is even, $N+2=2r$; card (Σ_i^+) = card $(\Sigma_i^-) = r$, each $i=1,\ldots,k$, so $\gamma_{N,k}=2r^k$. On the other hand, if N is odd, $N+2=2s+1$; card (Σ_i^+) and card (Σ_i^-) differ by 1, for each $i=1,\ldots,k$, hence one is s and the other $s+1$. Let u = number of i, $1 \leq i \leq k$, for which card $(\Sigma_i^+) = s$. Then Σ^+ consists of $s^u(s+1)^{k-u}$ points, and Σ^- of $(s+1)^u s^{k-u}$ points so $\gamma_{N,k}=s^u(s+1)^{k-u}+(s+1)^u s^{k-u}$. It is easy to see that each choice $u=0,1,\ldots,[k/2]$ produces an essentially different configuration Σ.

Next, a Haar polynomial $P(x_1,\ldots,x_k)$ of degree N in the k variables $x_1,\ldots,x_k$ is of form

$$P(x_1,\ldots,x_k) = \sum_{\substack{u_1+\ldots+u_k=0 \\ u_i \geq 0}}^{N} \alpha_{u_1,\ldots,u_k} \varphi_1^{u_1}(x_1)\ldots\varphi_k^{u_k}(x_k).$$

Lemma 3.1 : P contains as many "monomials" as there are ways to choose non-negative integers $u_1,\ldots,u_k \ni u_1+\ldots+u_k \leq N$. In fact, there are $\binom{N+k}{k}$ such k-tuples $(u_1,\ldots,u_k)$.

Proof: Observe first $\sum_{m=0}^{M} \binom{K+m-1}{m} = \binom{M+K}{K} = \binom{M+K}{K}$, any $M \geq 1$, any $K \geq 1$.

If $M=0$, sum on left reduces to $\binom{K-1}{0} = 1$, which is equal to

$\binom{K}{K}$ on the right. Assume true for M-1, so

$$\sum_{m=0}^{M-1}\binom{K+m-1}{m} = \binom{M-1+K}{K}; \text{ but then}$$

$$\binom{K+M-1}{M} + \binom{M-1+K}{K} = \frac{\lfloor K+M-1}{\lfloor M \lfloor K-1} + \frac{\lfloor M-1+K}{\lfloor K \lfloor M-1} = \frac{\lfloor K+M}{\lfloor M \lfloor K} = \binom{M+K}{K}.$$

Next, there are $\binom{k+n-1}{n}$ ways to choose non-negative integers $u_1,\ldots,u_k \ni u_1+\ldots+u_k=n$. For, if k=1, there is evidently only one way to choose u_1, and indeed $\binom{1+n-1}{n}=1$. Assume $k > 1$ and that for any ν, there are $\binom{k-1+\nu-1}{\nu}$ ways to choose $u_1,\ldots,u_{k-1} \ni u_1+\ldots+u_{k-1}\nu$. But for each $0 \leq \nu \leq n$, the choice $u_k=n-\nu$ produces a set $u_1,\ldots,u_k \ni u_1+\ldots+u_k=n$. Hence, there are in all $\sum_{\nu=0}^{n}\binom{k-1+\nu-1}{\nu} = \binom{k-1+n}{n}$ ways to choose $u_1,\ldots,u_k \ni u_1+\ldots+u_k=n$. A second use of the initial observation gives the desired result, as

$$\sum_{n=0}^{N}\binom{k-1+n}{n} = \binom{k+N}{k}.$$

(Another, "nifty", proof is due to D. Berkowitz: choosing non-negative integers $u_1,\ldots,u_k \ni u_1+\ldots u_k \leq N$, is equivalent to filling k places out of N+k, in such a manner that between the $(i-1)^{st}$ filled place and the ith filled place [or to the left of the 1st filled place], u_i empty places should intervene. Clearly there are $\binom{N+k}{k}$ ways to do this.)

To say P vanishes on Σ is to say

$$\sum_{\substack{u_1+\ldots+u_k=0 \\ u_1 \\ u_k\geq 0}}^{N} \alpha_{u_1,\ldots,u_k}\varphi_1^{u_1}(\xi_1)\ldots\varphi_k^{u_k}(\xi_k)=0 \text{ for every } (\xi_1,\ldots,\xi_k)\in\Sigma.$$

By Theorem 2, this implies every $\alpha_{u_1,\ldots,u_k} = 0$. That is, the system of $\gamma_{N,k}$ homogeneous equations in the $\binom{N+k}{k}$ "unknowns" $\alpha_{u_1,\ldots,u_k}$ has only the solution $(0,\ldots,0)$.

<u>Lemma 3.2</u> : $\binom{N+k}{k} \leq \gamma_{N,k}$ for all $k \geq 2$, all $N \geq 0$.

<u>Proof</u> : If $N=0$, assertion is clearly trivial.

If $N=1$, then $s=1$, and we must show $\binom{1+k}{k} = 1+k \leq 2^u+2^{k-u}$, any $0 \leq u \leq k$, any $k \geq 2$. It would suffice, by the elementary calculus, to show $1+k \leq 2^{\frac{1}{2}k+1}$, for $k \geq 2$. However, the function $2^{x+1}-(2x+1)$ is non-negative and has a non-negative first derivative for $x \geq 1$, so we are done.

Suppose now that $N > 1$ and proceed by induction on k.

If $k=2$, and N is even, $\binom{N+2}{2} = \binom{2r}{2} = r(2r-1) < 2r^2 = \gamma_{N,2}$; but, if N is odd, $\binom{N+2}{2} = \binom{2s+1}{2} = s(2s+1) < \gamma_{N,2}$ which is $s^2+(s+1)^2$ or $2s(s+1)$.

Assume $k > 2$ and that the result has been established for $k-1$.

$\binom{N+k}{k} \div \binom{N+k-1}{k-1} = \frac{N+k}{k} \leq [N/2] + 1$, because $N \leq k\cdot[N/2]$ as soon as $N > 1$, $k > 2$. r and s are each $[N/2] + 1$.

For N even, then, $\binom{N+k}{k} \leq r\cdot\binom{N+k-1}{k-1} \leq r\cdot\gamma_{N,k-1} \leq r\cdot 2r^{k-1} = 2r^k$.

For N odd, $\binom{N+k}{k} \leq s\cdot\binom{N+k-1}{k-1} \leq s\cdot\gamma_{N,k-1}$,

so $\binom{N+k}{k} \leq s\cdot\min\limits_{0 \leq u \leq k-1} \{s^u(s+1)^{k-1-u}+s^{k-1-u}(s+1)^u\}$,

which is clearly $\leq \min\limits_{0 \leq u \leq k} \{s^u(s+1)^{k-u} + s^{k-u}(s+1)^u\}$,

thus $\binom{N+k}{k} \leq \gamma_{N,k}$.

Hence the assertion is valid for all k.

From this it follows, since the system must have maximal possible rank, that its rank is $\binom{N+k}{k}$. Moreover, there must exist a sub-lattice $\tilde{\Sigma}$ of $\binom{N+k}{k}$ points, such that the equations $P(\xi_1, \ldots, \xi_k) = 0$, $(\xi_1,\ldots,\xi_k) \in \tilde{\Sigma}$, form an $\binom{N+k}{k}$-square system with non-zero determinant.

Section 4

Theorem 1 can be regarded as saying: a best (Čebyšev) approximation of degree $\leq N$ to a separated function in 2 variables is the separated Haar polynomial which is the sum of the respective best approximations of degree $\leq N$ to the separate components. Theorem 3 says: this Haar polynomial is the unique best approximation of degree $\leq N$.

Certain other attempts to generalize the results of the original paper have led to counterexamples, even when k=2.

Consider approximation on $[0,1]$ by (ordinary) polynomials in the L^p norm, where $\|f\| = \{\int_0^1 |f(x)|^p dx\}^{1/p}$. To say f(x) is unimprovable in the L^p norm by any polynomial of degree $\leq N$, is to say $\|f-\lambda x^u\| \geq \|f\|$ all real λ, all $u=0,\ldots,N$. That is, 0 is the best approximation of degree $\leq N$.

Similarly, the L^p norm on the Cartesian product $[0,1] \times [0,1]$ is given by

$$\|F\| = \{\iint_{[0,1]x[0,1]} |F(x,y)|^p dx\, dy\}^{1/p}$$

and it is easy to see that to say F is unimprovable by a polynomial of degree $\leq N$, means $\|F - \lambda \cdot x^u y^v\| \geq \|F\|$ all real λ, all $u \geq 0$, $v \geq 0 \ni u+v \leq N$.

We will show Theorem 1 does not hold for p=4, k=2, N=0.

Definition: $f \perp g$ (f orthogonal to g) in $L^p(X)$ if $\|f-\lambda g\| \geq \|f\|$ all real λ.

Assert $f \perp g$ in L^4 if and only if $\int f^3 g = 0$:

$$\|f-\lambda g\|^4 = \int(f-\lambda g)^4 = \int f^4 - 4\lambda\int f^3 g + \int(6\lambda^2 f^2 g^2 - 4\lambda^3 fg^3 + \lambda^4 g^4),$$

$$\therefore \int(f-\lambda g)^4 - \int f^4 = -4\lambda\int f^3 g + \lambda^2\int[2(fg)^2 + f^2(2f-\lambda g)^2].$$

The second integral on the right is always non-negative; so if $\int f^3 g \neq 0$, λ can be so chosen that the whole right-hand side is negative, whereas if $\int f^3 g = 0$, the right side is non-negative.

The assertion follows from the fact that $\|a\| \geq \|b\|$ if and only if $\|a\|^4 \geq \|b\|^4$.

It will suffice to exhibit a function $F(x)$, unimprovable by a constant, such that $F(x) + F(y)$ <u>can</u> be improved by a constant, i.e. the best approximation of degree 0 in $[0,1] \times [0,1]$ is <u>not</u> the sum 0+0 of the best approximations to each separate component. That is, $\int_0^1 (F(x))^3 dx = \int_0^1 (F(y))^3 dy = 0$, but $\iint_{[0,1]\times[0,1]} [F(x) + F(y)]^3 dx\, dy \neq 0$.

Observe

$$\iint [F(x)+F(y)]^3 dy\, dy = \iint [(F(x))^3 + 3(F(x))^2 F(y) + 3F(x)(F(y))^2 + (F(y))^3] dx\, dy$$

$$= \int_0^1 [(F(x))^3 + 3(F(x))^2 \cdot \int_0^1 F(y)dy + 3\cdot F(x) \cdot \int_0^1 (F(y))^2 dy + \int_0^1 (F(y))^3 dy] dx$$

$$= 2\{\int_0^1 (F(x))^3 dx + 3 \cdot \int_0^1 (F(x))^2 dx \cdot \int_0^1 F(y)dy\}.$$

$\therefore$ Suffices to exhibit $F(x) \ni \int_0^1 (F(x))^3 dx = 0$ but $\int_0^1 F(x)dx \neq 0$, $\int_0^1 (F(x))^2 dx \neq 0$. Namely, $F(x) = x\,[1-\frac{3}{2}x^2]^{1/3}$:

$$\int_0^1 x\,[1-\tfrac{3}{2}x^2]^{1/3} dx = -\frac{1}{3} \cdot \frac{4}{3}\,[1-\tfrac{3}{2}x^2]^{4/3} \Big|_0^1 = -\frac{1}{4}\,\{[1-\tfrac{3}{2}] - 1\} \neq 0;$$

$\int_0^1 x^2[1-\frac{3}{2}x^2]^{2/3} dx \neq 0$ because the integrand is positive except at $x=0$ or $x=\sqrt{\frac{2}{3}}$; $\int_0^1 x^3(1-\frac{3}{2}x^2)\, dx = \frac{x^4}{4} - \frac{3}{2}\frac{x^6}{6} \Big|_0^1 = 0$.

A more striking counterexample to Theorem 1 is provided by the following: We claim that there exists $F(x) \in L^4[-1,1]$ such that $F(x) + F(y)$ is unimprovable by any quadratic of the form $P(x) + Q(y)$, but is improvable by a multiple of xy. This means $F(x) + F(y)$ is orthogonal to $1, x, x^2, y, y^2$ but <u>not</u> to xy.

Consider $\int_{-1}^1 [\int_{-1}^1 (F(x) + F(y))^3 \begin{Bmatrix} 1 & y^2 \\ x & y \\ x^2 & xy \end{Bmatrix} dy] dx$: we seek $F(x) \in L^4[-1,1]$ such that

$$\begin{cases} 2\int_{-1}^{1} F^3(x)dx + 3\int_{-1}^{1}F^2(x)dx\cdot\int_{-1}^{1}F(x)dx = 0 \\ 2\int_{-1}^{1}xF^3(x)dx + 3\int_{-1}^{1}xF^2(x)dx\cdot\int_{-1}^{1}F(x)dx+3\int_{-1}^{1}x\cdot F(x)dx\cdot\int_{-1}^{1}F^2(x)dx=0 \\ 2\int_{-1}^{1}x^2F^3(x)dx+3\int_{-1}^{1}F(x)dx\cdot\int_{-1}^{1}F(x)dx+3\int_{-1}^{1}x^2F(x)dx\cdot\int_{-1}^{1}F^2(x)dx+\frac{2}{3}\int_{-1}^{1}F^3(x)dx=0 \\ 3\int_{-1}^{1}xF^2(x)dx\cdot\int_{-1}^{1}xF(x)dx \neq 0. \end{cases}$$

It would certainly suffice to show that there exists $F(x) \in L^4[-1,1]$ such that

$$\begin{cases} F(x) = 0 \text{ on } [-1,0] \\ \int_0^1 F(x)dx = 0 \\ \int_0^1 F^2(x)dx = 1 \\ \int_0^1 (2F^3(x) + 3F(x))\begin{Bmatrix}1\\x\\x^2\end{Bmatrix}dx = 0 \\ \text{but } \int_0^1 xF(x)dx \neq 0 . \end{cases}$$

Suppose no such $F(x)$ exists. Then we would have

$$(*)\begin{cases} F(x)\in L^4[0,1] \\ \int_0^1 F(x)dx = 0 \\ \int_0^1 F^2(x)dx = 1 \\ \int_0^1 (2F^3(x) + 3F(x))\begin{Bmatrix}1\\x\\x^2\end{Bmatrix}dx = 0 \end{cases} \Rightarrow \int_0^1 xF(x)dx = 0 .$$

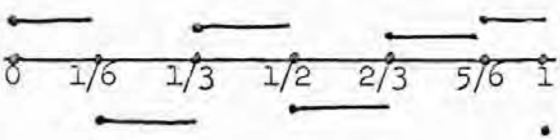
Let $F(x)$ be a function satisfying conditions (*). Then for any $G(x)\in L^4[0,1]$ and any δ

$$\begin{cases} \int_0^1 [F(x) + \delta G(x)]dx = 0 \\ \int_0^1 [F(x) + \delta G(x)]^2dx = 1 \\ \int_0^1 \{2[F(x) + \delta G(x)]^3+3[F(x)+\delta G(x)]\} \begin{Bmatrix}1\\x\\x^2\end{Bmatrix} dx = 0 \end{cases} \Rightarrow \int_0^1 x[F(x)+\delta G(x)]dx=0.$$

Since δ can be chosen arbitrarily small, this means (N.A.S.C.)

$$(**)\qquad \left.\begin{cases} G(x) \in L^4[0,1] \\ \int_0^1 G(x)dx = 0 \\ \int_0^1 F(x)G(x)dx = 0 \\ \int_0^1 (6F^2(x)+3)\cdot G(x) \begin{matrix}1\\ x_2\\ x^2\end{matrix} dx = 0 \end{cases}\right\} \Rightarrow \int_0^1 xG(x)dx=0$$

This set of equations says: whenever $\sqrt[3]{G}$ is orthogonal to $1,F$, $2F^2+1$, $x(2F^2+1)$, $x^2(2F^2+1)$, then $\sqrt[3]{G}$ is orthogonal to x also. i.e. $F(x)$ is such that x is in the linear subspace of $L^4[0,1]$ spanned by these 5 functions. But this implies $F(x)$ satisfies an equation $A(x)F^2+B(x)\cdot F+C(x)=0$, where A,B,C are polynomials in x of degree ≤ 2. $\therefore F(x)$ is continuous on $[0,1]$, except possibly at 2 points (because it is a quadratic surd function of x). Likewise, $F+\delta G$ must be a quadratic surd function of x, and hence continuous, except possibly at 2 pts. for <u>every</u> G satisfying conditions (**). However, given <u>any</u> $F(x)$ satisfying (*), there exist functions $G(x)$ satisfying (**) which fail to be continuous at 5 points, namely



$$G(x)=\begin{cases} a_i & \frac{i-1}{6} \leq x < \frac{i}{6} \quad i=1,\dots,6 \\ a_6 & \text{at } x=1 \end{cases}$$

where
$$\begin{cases} a_1+a_2+a_3+a_4+a_5+a_6=0 \\ \sum_{i=1}^{6} a_i \int_{\frac{i-1}{6}}^{\frac{i}{6}} F(x)dx = 0 \\ \sum_{i=1}^{6} a_i \int_{\frac{i-1}{6}}^{\frac{i}{6}} (2F^2+1)dx = 0 \\ \sum_{i=1}^{6} a_i \int_{\frac{i-1}{6}}^{\frac{i}{6}} x(2F^2+1)dx = 0 \\ \sum_{i=1}^{6} a_i \int_{\frac{i-1}{6}}^{\frac{i}{6}} (2F^2+1)dx = 0 \end{cases}$$

There are 5 homogeneous linear equations in the 6 unknowns $a_1, \dots, a_6$, so there always exist solutions not all zero. But then $F+\delta G$ is <u>not</u> continuous, or rather, fails of continuity at more than 2 points. This contradiction shows that the implication following (*) is not valid.

Consider now weighted Čebyšev norms.

If $f(x)$ is continuous on $[0,1]$ and $\rho(x) \geq 0$ is continuous on $[0,1]$, <u>define</u> $\|f\|_\rho = \sup_{x\epsilon[0,1]} \rho(x)|f(x)|$; likewise; $\|g\|_\sigma$ for functions $g(y)$ with weight $\sigma(y)$. Then the "product norm" can be defined by $\|F\|_{\rho,\sigma} = \text{sum}_{x,y\epsilon[0,1]} \rho(x)\sigma(y)|F(x,y)|$.

We will show that Theorem 1 fails even for $N = 0$, namely, we shall exhibit functions $f(x)$ and $g(y)$, unimprovable by a constant with weights x and y respectively, such that $f(x)+g(y)$ <u>is</u> improvable by a constant, with weight xy. Observe that to say $f(x)$ is unimprovable by a constant, with weight x, is to say $\max_{x\epsilon[0,1]} x|f(x)-c| \geq \max_{x\epsilon[0,1]} x|f(x)|$ for any c. It follows that $\{\alpha | xf(x)=\alpha\}$ is an interval $[-\lambda,\lambda]$, i.e. the range of $xf(x)$ is symmetric. Otherwise, suppose W.L.O.G. $\max_{x\epsilon[0,1]} xf(x) > -\min_{x\epsilon[0,1]} xf(x)$; let $c_o = \frac{1}{2}[\max + \min]$, then $\max_{x\epsilon[0,1]} |xf(x)-c_o x| < \max_{x\epsilon[0,1]} |xf(x)|$. (It is easy to see that this is also a sufficient condition.) Likewise, in order that $f(x) + g(y)$ be unimprovable by a constant, with weight xy, it is necessary that $xy(f(x) + g(y))$ have symmetric range.

So now, choose $f(x) = x-a$ and $g(y) = y-a$ with $a=2\sqrt{2}-2$, then $x(x-a)$ and $y(y-a)$ have symmetric range as x,y run through $[0,1]$. However, $xy(x-a) + xy(y-a)$ does <u>not</u> have symmetric range: by elementary calculus this function achieves its max or min at points for which $x=y=t$, so $\max_{x,y\epsilon[0,1]} xy(x-a)+xy(y-a)= \max_{t\epsilon[0,1]} 2t^2(t-a)$, and likewise for min. But $2t^2(t-a)$ has maximum value $2=2a$ at $t=1$, and minimum value $-\frac{8}{27}a^3$ at $t=2/3a$. $2-2a \neq \frac{8}{27}a^3$ for $a=2\sqrt{2}-2$. Thus range is not symmetric.

$\therefore$ $f(x)+g(y)$ is not unimprovable, so the best approximation is not the sum of the separate best approximations.

In the even simpler case in which only the weight x is involved, consider the function $x((x-a)+(y-\frac{1}{2}))$ with $a = 2\sqrt{2}-2$. By the usual elementary calculations, $\text{range}_{x\epsilon[0,1]}\ x(x-a)=[2\sqrt{2}-3,\ 3-2\sqrt{2}]$, $\text{range}_{y\epsilon[0,1]}\ (y-\frac{1}{2}) = [-\frac{1}{2},\frac{1}{2}]$, but $x((x-a)+(y-\frac{1}{2}))$ has maximum value --in the unit square---of $\frac{3}{2}-a$ at $x=1$, $y=1$ and minimum value of $-\frac{1}{4}(a+\frac{1}{2})^2$ at $x=\frac{1}{2}(a+\frac{1}{2}),y=0$. However, $\frac{1}{4}(a+\frac{1}{2})^2 \neq \frac{3}{2}-a$, so the function is improvable.

Further investigations into the weighted norms are continuing.

We turn now to a question which, although it does not involve a direct generalization of Theorem 1, is nonetheless closely related in spirit. Theorems 1 and 3 say that if a function is of separated form then the (unique) best approximation of degree N is also of separated form. Consider functions on the unit square, of form $f_0(y)+xf_1(y)+x^2f_2(y)+\ldots+x^nf_n(y)$, where each $f_i(y)$ is continuous on $[0,1]$, and Čebyšev approximation by (ordinary) polynomials in x and y. We ask whether there is a best approximation of degree $N > n$ whose degree in x is $\leq n$, i.e of form $p_0(y)+x\,p_1(y) + x^2p_2(y) +\ldots+ x^np_n(y)$ where $p_i(y)$ is a polynomial in y of degree $\leq N-i$. Observe first

$$\max_{x,y}|h_0(y)+xh_1(y)| = \max\{\max_y|h_0(y)|,\max_y|h_0(y) + h_1(y)|\}$$

because h_0+xh_1 is linear in x for each fixed y.

If n=0, the given function is of form $f_0(y)$. Let $p_0(y)$ be the best Cebysev approximation to $f_0(y)$ of degree N. Let $q_0(y)+xq_1(y)+\ldots+x^N q_N(y)$ be a polynomial of degree N. Then

$$\max_{x,y} |(f_0-q_0)-xq_1-\ldots-x^N q_N| \geq \max\{\max_y|f_0(y)-q_0(y)|,\max_y|f_0(y)-q_0(y)-\ldots-q_N(y)|\}$$

$$\geq \max_y|f_0(y)-p_0(y)|,$$

because $q_0(y)$ and $q_0(y)+\ldots+q_N(y)$ were both among the candidates from amongst which $p_0(y)$ was chosen.

Hence $q\,(y)+\ldots+x^N q_N(y)$ does not approximate $f_0(y)$ better than $p_0(y)$.

If n=1, the given function is of the form $f_0(y) + xf_1(y)$: call it $F(x,y)$. Let $p_0(y)$ and $p_1(y)$ be those polynomials of degree N and N-1 respectively, for which $\max\{\max_y |f_0(y) - p_0(y)|, \max_y |f_0(y) + f_1(y) - p_0(y) - p_1(y)|\}$ is a minimum. We assert $P(x,y) = p_0(y) + xp_1(y)$ is the best approximation of degree N with this form. For, let $Q_1(x,y)=q_0(y)+xq_1(y)$ be of degree N or less, then

$$\max_{x,y} |(f_0(y)-q_0(y))+x(f_1(y)-q_1(y))| = \max\{\max_y|f_0-q_0|,\max_y|f_0+f_1-q\ -q_1|\}$$

$$\geq \max\{\max_y|f_0-p_0|, \max_y|f_0+f_1-p_0-p_1|\}$$

by construction

$$= \max_{x,y} |(f_0(y)-p_0(y))+x(f_1(y)-p_1(y))|.$$

Moreover, let $Q_N(x,y)=q_0(y)+xq_1(y)+\ldots+x^N q_N(y)$ be <u>any</u> polynomial of degree N. We will show $\|F-Q_N\| \geq \|F-P\|$, so that P is the best approximation to F, of degree N. It suffices to show

$$\max_{x,y}|f_0(y)-q_0(y) +x(f_1(y)-q_1(y))-x^2 q_2(y)-\ldots-x^N q_N(y)|$$

$$\geq \max_{x,y}|(f_0(y)-p_0(y))+x(f_1(y)-p_1(y))|$$

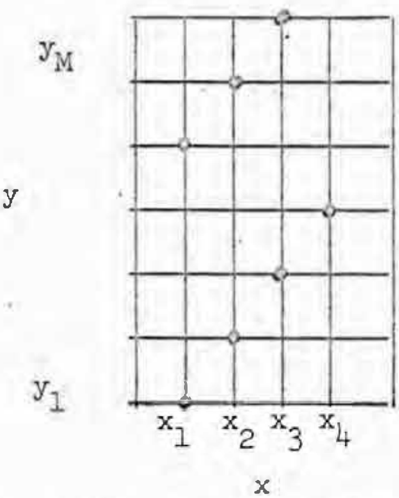
$$= \max\{\max_y|f_0(y)-p_0(y)|,\max_y|f_0(y)+f_1(y)-p_0(y)-p_1(y)|\}.$$

The left-hand expression is $\geq \max\{\max_y|f_0-q_0|, \max_y|f_0+f_1-q_0-(q_1+\ldots+q_N)|\}$, but q_0 and $q_1+\ldots+q_N$ were among the candidates from amongst which p_0 and p_1 were chosen, hence the desired result follows.

Let $n=2$, $N=3$. We shall exhibit a function $F(x,y)=f_0(y)+xf_1(y)+x^2f_2(y)$, unimprovable by any polynomial $p_0(y)+xp_1(y)+x^2p_2(y)$ of degree 3, but improvable by a certain $a_0+a_1x+a^2+a_3x^3$, a_i constants, $a_3 \neq 0$.

By way of preliminary observation, recall from the elementary calculus that given $0 < x_0 < 1$, there always exists a quadratic in x which attains its maximum [or minimum] of 1 [or -1] at $x=x_0$, and its minimum [or maximum] of 0 at x=0 or x=1 according as $x_0 \geq \frac{1}{2}$ or $x_0 \leq \frac{1}{2}$.

Let $M \geq 4$ and $0=y_1 < y_2 < \ldots < y_M=1$ be fixed values of y, all to be determined later. Let $0 \leq x_1 < x_2 < x_3 < x_4 \leq 1$; consider the vertical lines L_i: $x=x_i$ i=1,2.3,4. The points (x_1,y_1), (x_3,y_2), (x_3,y_3), (x_4,y_4), $(x_1,y_5),\ldots,(x_i,y_{4k+i}),\ldots$constitute a finite set which meets any horizontal $y=y_j$ in exactly one point.



To each horizontal $y=y_j$ with $j\equiv1$ mod 4 associate the parabola $\pi_j(x)=\alpha_{j0}+\alpha_{j1}x+\alpha_{j2}x^2$ which attains its minimum value of +1 at x=x, and its maximum value of 0 at x=0 or x=1. Likewise to each horizontal $y=y_j$ with $j\equiv3$ mod 4 associate the parabola $\pi_j(x)$ whose minimum is -1 at $x=x_3$, and minimum of 0 at $x=0$ or $x=1$. And, to each $y=y_j$, $j\equiv 2 \bmod 4$, the parabola $\pi_j(x)$ with max +1 at $x=x_2$ and minimum 0 at $x=0$ or $x=1$, and likewise for $y=y_j, j\equiv 0 \bmod 4$. We can interpolate a surface $F(x,y)$ on the unit square, as follows:

$$\begin{cases} F(x,y_j) = \pi_j(x) & \text{any } x \\ F(x,y)=F(x,y_j)+\left(\dfrac{y-y_j}{y_{j+1}-y_j}\right)[F(x,y_{j+1})-F(x,y_j)] \end{cases}$$

$$\text{for } y_j \leq y \leq y_{j+1}, j=1,\ldots,M-1.$$

$F(x,y)$ is continuous, $|F(x,y)| \leq 1$, and the $\max\limits_{x,y} |F(x,y)|=1$ is taken on <u>only</u> at the distinguished points. Also note $F(x,y_j) \leq 0$, $j\equiv 1,3 \bmod 4$; $F(x,y_j) \geq 0$ $j\equiv 0,2 \bmod 4$. Evidently $F(x,y)$ is of the form $f_0(y) + xf_1(y) + x^2 f_2(y)$.

Let $\bar{x}_1,\bar{x}_2,\bar{x}_3$ be such that $x_1<\bar{x}_1<x_2<\bar{x}_2<x_3<\bar{x}_3<x_4$; let $G(x)=(x-\bar{x}_1)(x-\bar{x}_2)(x-\bar{x}_3)$. $G(x)$ is negative at x_1 and x_3, positive at x_2 and x_4. There exists $\delta > 0$ sufficiently small that $\max\limits_x|\delta G(x)|<\epsilon<\frac{1}{2}$, so $\max\limits_{x,y}|F(x,y) - \delta G(x)| \leq 1-\epsilon$, since $\delta G(x)$ and $F(x,y)$ agree in sign on the distinguished vertical lines. $\therefore F(x,y)$ is indeed improvable by a cubic in x.

It remains to show that no $P(x,y)=p_0(y)+xp_1(y)+x^2p_2(y)$ of degree 3 improves $F(x,y)$. Suppose, on the contrary, $P(x,y)$ is such an improving polynomial. Then P must be <u>negative</u> at the distinguished points on L_1 and L_3 and positive at those on L_2 and L_4. For any $x=\bar{x}$, $P(\bar{x},y)$ is of degree ≤ 3, hence there are at most 2 intervals in which it is positive and at most 2 in which it is negative (otherwise there would be ≥ 4 zeros). We may suppose that M and $y_1,\ldots,y_M$ are so chosen that $y_{j+4}-y_j<\zeta$ where $8\zeta<1$; (it would suffice to choose $y_{j+1}-y_j<\frac{1}{32}$). Then $P(x_1,y)$ is <u>positive</u> on at most a set of measure 2ζ in L_1, i.e. between 2 pairs of distinguished points; likewise, $P(x_3,y)$ is positive on at most a set of measure 2ζ in L_3, $P(x_2,y)$ is negative on at most a set of measure 2ζ in L_2, and $P(x_4,y)$ is negative on at most a set of measure 2ζ in L_4. Since $8\zeta < 1$ there exists a value $y = \bar{y}$ such that $P(x_1,\bar{y}) < 0$, $P(x_2,\bar{y}) > 0$, $P(x_3,\bar{y}) < 0$, $P(x_4,\bar{y}) > 0$, but $P(x,\bar{y})$ is a parabola so this is impossible.

We conclude F is unimprovable by any such $P(x,y)$.

Bibliography

[1] Aitken, A.C. — Determinants and Matrices
Oliver and Boyd, Ltd. London 1958

[2] Akhiezer, N.I. — Theory of Approximations
Fred. Ungar New York 1956

[3] Newman, D.J. and Shapiro, H.S. — Some Theorems on Čebyšev Approximation
Duke Math. Jl., vol.30, No. 4 (Dec. 1963)
p.673-684

[4] Rice, Lepine Hall — Adjoint and Inverse Determinants and Matrices
Journal of Math. and Physics, vol. 5 (1925)
p. 55-64

[5] Rice, John R. — The Approximation of Functions, vol. I
Addison-Welsey 1963