JOURNAL OF APPROXIMATION THEORY 48, 189–200 (1986)

Approximation by Semi-Non-linear Functions*

M. BRANNIGAN[1]

Department of Computer Science, University of Georgia, Athens, Georgia 30602, U.S.A.

Communicated by E. W. Cheney

Received September 24, 1984; revised June 25, 1985

DEDICATED TO THE MEMORY OF GÉZA FREUD

In this paper we discuss the best Chebyshev approximation of continuous real or complex valued functions by a large class of non-linear functions. These functions, which we call semi-non-linear, are non-linear functions of linear functions defined in a general multivariate setting. Full use is made of the concept of H-sets for constrained non-linear approximation, and a characterization of best approximation is stated in terms of H-sets. Unicity of best approximation is discussed, and the special case of approximation by functions of $ax + by + c$ is shown to give uniqueness. This extends a well-known theorem of Collatz (*Z. Angew. Math. Mech.* **36** (1956), 198–211.

INTRODUCTION

The general theory for characterizing best Chebyshev approximation by linear spaces is set out in [1, 2], where use is made most effectively of the concept of H-sets, as originally conceived in [4]. The more general setting of non-linear constrained approximation is set out in [3], where again use is made of H-sets.

Using the theory developed in [3], many classes of non-linear approximating functions can be studied. We consider here a very wide, and useful, class of multivariate non-linear functions, namely those functions formed by taking a non-linear function of a linear function. The resulting set of approximating functions consists of non-linear constrained functions for which the theory of [3] can be readily applied. We shall call these approximating functions semi-non-linear.

To analyze this set of approximating functions we first consider the form of the H-sets, and we show how the H-sets for the non-linear functions relate to the H-sets of the linear space used in the construction of the functions. With this analysis we can then give a characterization of best

* This work was supported in part by National Science Foundation Grant DMS 8319727.

[1] Present address: Information and Graphic Systems, P.O. Box 18885, Atlanta, GA 30326, U.S.A.

0021-9045/86 $3.00

approximation, and show that in this setting best approximations thus characterized are globally best, in contrast to the usual situation in non-linear approximation where only locally best approximations are found. A sufficient condition for uniqueness of the best approximation is given in terms of minimal H-sets.

We consider a particular example of these functions, that being the case when the linear space is spanned by $(1, x, y)$. In [4] uniqueness was proven for the linear problem over strictly convex domains. In [5] uniqueness was shown for the case

$$\frac{1}{L(x, y)},$$

where L are linear functions from the space spanned by $(1, x, y)$. Using the theory developed herein the results of [4, 5] are deducible as special examples of our more general setting.

H-Sets and Characterization

Let $f(t)$ be a strictly monotonic, real-valued, continuously differentiable function of the real variable t on the interval $[a, b]$ with no zero derivatives. Further, let V be an n-dimensional linear subspace of $C(B)$, the space of continuous real-valued functions defined on a compact Hausdorff space B. Let this linear space V have a basis $\{g_1, g_2, ..., g_n\}$ over the same scalar field. Given $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$ we can, therefore, consider the function $F(\alpha, x)$ given by

$$F(\alpha, x) = f(h(\alpha, x)),$$

$$h(\alpha, x) = \sum_{i=1}^{n} \alpha_i g_i(x), \qquad x \in B.$$

That this is non-linear, in general, follows from the generality of the definition of f.

The following analysis also applies if f is complex valued, then we require that the real part of $f(t)$, namely $\text{Re}[f(t)]$, is strictly monotonic and has no zero derivatives. The modification to the theorems is slight in that the real part of the functions in question is used. However, for simplicity we shall assume that our functions are real valued, and leave the reader to derive the theorems for the complex case.

Examples of useful functions. A function used in electrical engineering is

$$\exp(P_n(z)),$$

where $P_n(z)$ is a complex polynomial of degree n. Another often used function is

$$[P_k(x)]^r,$$

where P_k is a polynomial of n variables of degree r. We shall also show how to extend our analysis to functions such as

$$\cos(P_k(x)),$$

with $\cos(t)$ defined on an interval $[0, \pi]$, where the function has zero derivatives at the interval boundaries.

We can now define the approximation problem which we address. The norm that we use is the Chebyshev norm, defined by

$$\|\phi\| = \max\{|\phi(x)|: x \in B\}, \qquad \phi \in C(B),$$

and we seek to find that $\alpha \in \mathbb{R}^n$ such that

$$\|G - F(\alpha, \cdot)\|$$

is minimized for a given $G \in C(B)$, the vector α being constrained by

$$a \leqslant \sum_{i=1}^{n} \alpha_i g_i(x) \leqslant b, \qquad \forall x \in B.$$

We shall assume that $\alpha \in \mathbb{R}^n$ is such that there are $x \in B$ for which the above constraints are satisfied as equalities.

It is not difficult in this general constrained approximation setting to include extra constraints. We have in mind such examples as sums of monotonic functions, then assuming that $f_1, f_2, ..., f_k$ are all either monotone increasing or monotone decreasing the function

$$f = \sum_{j=1}^{k} c_j f_j$$

subject to

$$c_j \leqslant 0, \qquad j = 1, ..., k,$$

is also monotonic, and the following theory applies with those extra constraints added. For the purposes of this paper such problems are left to the reader to develop, building on the theory developed herein, and in [3].

We shall denote by W the set of functions $F(\alpha, \cdot)$ with these constraints. The set W is the set of semi-non-linear functions, which is a constrained non-linear set of approximating functions. Hence the general theory of [3] does apply. The subset of $\mathbb{R}^n$, which satisfies the constraints, is the parameter set for the problem, and we shall denote this by P for future reference.

The usefulness of the H-set approach to approximation theory lies in the fact that a study of the approximating set, in this case W, gives us all the properties of the approximation without regard to the specific function to be approximated. Hence, in this vein we now consider what the H-sets of W are. Using the definitions in [3] we obtain, for this particular case, the following definitions.

We shall denote by $f'(t)$ the derivative of $f(t)$ with respect to $t \in \mathbb{R}$. Due to the linearity of $h(\alpha, x)$ the gradient vector of $h(\cdot, x)$ with respect to α is the vector $(g_1(x),..., g_n(x))$. The strictly positive orthant of $\mathbb{R}^s$ consists of all vectors $(c_1,..., c_s)$ with $c_i > 0$, $i = 1,..., s$, which we denote by $\mathbb{R}^s_+$, and $\{x_i\}_1^s$ is the finite set x_i, $i = 1,..., s$. From Definition 1 of [3] we obtain

DEFINITION 1. The set $\{(u_i, t_i), i = 1,..., p\}$, with $u_i \in B$, $t_i \in \mathbb{R}$, $|t_i| = 1$; together with $\{(v_j, s_j), j = 1,..., q\}$, with $v_j \in B$, $s_j \in \mathbb{R}$, $|s_j| = 1$: form an H_1-set with respect to W at $\alpha \in \mathbb{R}^n$ if and only if there exists $\eta \in \mathbb{R}^p_+$, $\mu \in \mathbb{R}^q_+$ such that, $\{u_i\}_1^p \cap \{v_j\}_1^q = \varnothing$ (null set), and

$$\sum_{i=1}^{p} \eta_i t_i f'(h(\alpha, u_i))\, h(\beta, u_i) + \sum_{j=1}^{q} \mu_j s_j h(\beta, v_j) = 0, \qquad \beta \in P,$$

$$\sum_{i=1}^{p} \eta_i = 1.$$

Using the same notation as in Definition 1 we obtain H_2-sets analogous to that in [3].

DEFINITION 2. $[\{u_i, t_i, \eta_i(\beta), p\}, \{v_j, s_j, \mu_j(\beta), q\}]$ forms an H_2-set with respect to W at α if and only if $\{u_i\}_1^p \cap \{v_j\}_1^q = \varnothing$, and

$$\sum_{i=1}^{p} \eta_i(\beta)\, t_i(F(\beta, u_i) - F(\alpha, u_i)) + \sum_{j=1}^{q} \mu_j(\beta)\, s_j h(\beta - \alpha, v_j) = 0, \qquad \beta \in P,$$

$$\sum_{i=1}^{p} \eta_i(\beta) = 1,$$

where $\eta(\beta) \in \mathbb{R}^p_+$ and $\mu(\beta) \in \mathbb{R}^q_+$, both depending on β.

H_1- and H_2-sets with respect to W are said to be minimal if no proper subset of the set $\{u_i\}_1^p \cup \{v_j\}_1^q$ can form an H_1- or H_2-set, respectively.

From the nature of our definition of W we obtain

THEOREM 1. *The set* $\{x_l, \lambda_l, e_l, k\}$, *with* $x_l \in B, \lambda_l > 0, e_l \in \mathbb{R}, |e_l| = 1$, $l = 1,..., k$, *is an* H_1-*set with respect to* W *at* α *if and only if*

$$\begin{pmatrix} g_1(x_1) \cdots g_1(x_k) \\ \vdots \qquad \vdots \\ g_n(x_1) \cdots g_n(x_k) \end{pmatrix} \begin{pmatrix} \lambda_1 e_1 \\ \vdots \\ \lambda_k e_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Proof. Suppose we have $\{(u_i, t_i)\}_1^p$ and $\{(v_j, s_j)\}_1^q$ as in Definition 1 forming an H_1-set with respect to W at α. We now put $k = p + q$, and

$$\{x_l\}_1^k = \{u_i\}_1^p \cup \{v_j\}_1^q,$$
$$\{e_l\}_1^k = \{t_i\}_1^p \cup \{s_j\}_1^q.$$

As f is strictly monotonic with no zero derivative, f' has the same sign for all values in the interval $[a, b]$. Also $\alpha \in P$, thus values of $h(\alpha, \cdot)$ lie in this same interval $[a, b]$, hence we can assign $\lambda_l > 0$ to be

$$\{\lambda_l\}_1^k = \{\eta_i f'(h(\alpha, u_i))\}_1^p \cup \{\mu_j\}_1^q \qquad \text{if } f' \text{ is positive,}$$
$$\{\lambda_l\}_1^k = \{-\eta_i f'(h(\alpha, u_i))\}_1^p \cup \{\mu_j\}_1^q \quad \text{if } f' \text{ is negative.}$$

We thus obtain, by substitution in Definition 1

$$\sum_{l=1}^{k} \lambda_l e_l h(\beta, x_l) = 0,$$

which is

$$(\beta_1, ..., \beta_n) \begin{pmatrix} g_1(x_1) \cdots g_1(x_k) \\ \vdots \qquad \vdots \\ g_n(x_1) \cdots g_n(x_k) \end{pmatrix} \begin{pmatrix} \lambda_1 e_1 \\ \vdots \\ \lambda_k e_k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

and the result follows.

Conversely, suppose $\{x_l, \lambda_l, e_l, k\}$ satisfies the matrix relation, then by multiplying by $\beta \in P$ we obtain, as above,

$$\sum_{l=1}^{k} \lambda_l e_l h(\beta, x_l) = 0.$$

This satisfies Definition 1 in the following manner: Put $p = k$, $q = 0$, and define η_i, $i = 1,..., k$ such that for f' positive,

$$\eta_i f'(h(\alpha, x_i)) = \lambda_i, \qquad i = 1,..., k,$$

or for f' negative

$$-\eta_i f'(h(\alpha, x_i)) = \lambda_i, \qquad i = 1,..., k,$$

and in each case

$$\sum_{i=1}^{k} \eta_i = 1.$$

Hence, the theorem is proved.

COROLLARY 1. *H_1-sets with respect to W are independent of α.*

Proof. As has been shown, H_1-sets are defined completely by a quadruple $\{x_l, \lambda_l, e_l, k\}$ and a matrix relationship. Thus if we have an H_1-set with respect to W at α then Theorem 1 shows that the matrix relation holds. Suppose now that we have a quadruple $\{x_l, \lambda_l, e_l, k\}$ and hence the matrix relation as in Theorem 1. We can choose any $\alpha' \in P$ and use the converse in Theorem 1 to obtain an H_1-set with respect to W at α'. Thus H_1-sets with respect to W are independent of the parameter, and we can drop the reference to α in the statements in H_1-sets.

COROLLARY 2. *H_1-sets with respect to W are H-sets with respect to V, and vice versa.*

Proof. This follows because the definition of H-sets with respect to V is that of the above matrix relationship, see [1, 2].

We can extend the equivalence of H-sets to include H_2-sets in the following manner.

THEOREM 2. *The class of minimal H_1-sets with respect to W is equivalent to the class of minimal H_2-sets with respect to W.*

Proof. Given an H_2-set with respect to W at α. This is an H_1-set with respect to W from Theorem 3 in [3], where it is shown that the class of H_2-sets is included in the class of H_1-sets.

Conversely, suppose $\{x_i, \lambda_i, e_i, k\}$ forms a minimal H_1-set with respect to W. Due to the fact that f is a continuously differentiable function of a real variable we obtain, for any $\alpha, \beta \in P$, using the mean value theorem

$$[F(\beta, x_i) - F(\alpha, x_i)] = [h(\beta, x_i) - h(\alpha, x_i)]\, f'(h(\gamma_i, x_i)),$$

where $\gamma_i \in P$ for $i = 1,..., k$.

Because $\{x_i, \lambda_i, e_i, k\}$ forms an H-set with respect to V (Corollary 2, Theorem 1), then no α, β exists such that (see [1, 2])

$$e_i[h(\beta, x_i) - h(\alpha, x_i)] > 0, \qquad i = 1,..., k.$$

Hence, due to the fact that f' is never zero and f is strictly monotonic, no α, β can exist such that

$$e_i[F(\beta, x_i) - F(\alpha, x_i)] > 0, \qquad i = 1,..., k.$$

(Note, if f' is negative we use $-e_i$ throughout). Hence from Theorem 2 in [3], $\{x_i, \lambda_i, e_i, k\}$ contains an H_2-set with respect to W at α. Due to the minimality of the H-set with respect to V this containment must be equality, as no subset of the $\{x_i\}$ can form an H-set with respect to V. Hence, the theorem follows.

We note here that the dependence on α can be dropped from the statement of an H_2-set because any α, β can be chosen in Theorem 2.

We have thus arrived at the conclusion that all minimal H-sets for the space W are minimal H-sets for the space V, and conversely. These are computable due to the matrix formulation, hence their interest.

Using this, a characterization of best approximation is possible. In what follows extremal points of a function are values $x \in B$ for which the norm of the function is attained.

THEOREM 3. *Consider the approximation of* $G \in C(B)$ *by* W. *If an* $\alpha \in \mathbb{R}^n$ *can be found such that a subset* $\{x_i : i = 1,..., p\}$ *of the extremal points of* $G - F(\alpha, \cdot)$, *together with signs* e_i *such that*

$$e_i[G(x_i) - F(\alpha, x_i)] = \|G - F(\alpha, \cdot)\|, \qquad i = 1,..., n,$$

and $\{y_j : j = 1,..., q\}$ *satisfying either*

$$h(\alpha, y_j) = \begin{cases} b & \text{with sign } -1, \ j = 1,..., q, \text{ or} \\ a & \text{with sign } +1, \ j = 1,..., q, \end{cases}$$

forms a minimal H-set with respect to V, then $f(\alpha, \cdot)$ *is a global best approximation to G by W.*

Conversely, let $F(\alpha, \cdot)$ *be a global best approximation to G by W, then the set of extremal points M of* $G - F(\alpha, \cdot)$, *together with signs* $e(x)$ *such that*

$$e(x)(G(x) - F(\alpha, x)) = \|G - F(\alpha, \cdot)\|, \qquad x \in M,$$

and points $y \in N \subset B$ *such that*

$$h(\alpha, y) = \begin{cases} b & \text{with sign } -1, \text{ or} \\ a & \text{with sign } +1, \end{cases}$$

will contain a minimal H-set with respect to V.

Proof. That all H-sets involved in this approximation are H-sets with respect to V follows from Theorems 1 and 2. Because, also, the class of

H_1-sets and the class of H_2-sets with respect to W are equivalent, then from Theorems 7, 9, and 10 in [3] Theorem 3 follows.

Due to the fact that the H-sets for the space W are the H-sets for the linear space V then it is not surprising that many of the theorems from linear approximation apply. If we define $\rho(G)$ by

$$\rho(G) = \min\{\|G - F(\alpha, \cdot)\| : \alpha \in P\}$$

then we obtain directly from Theorem 6 in [3]

THEOREM 4. *Let $\{x_i, \lambda_i, e_i, p\}$ together with points $\{y_j\}_1^q$, such that for some $\alpha \in \mathbb{R}^n$*

$$h(\alpha, y_j) = \begin{cases} b & \text{with sign } -1,\ j = 1,\dots, q, \text{ or} \\ a & \text{with sign } +1,\ j = 1,\dots, q, \end{cases}$$

forms a minimal H-set with respect to V, and given $G \in C(B)$ such that

$$e_i[G(x_i) - F(\alpha, x_i)] > 0, \qquad i = 1,\dots, p,$$

then

$$\min\{e_i[G(x_i) - F(\alpha, x_i)] \leqslant \rho(g) \leqslant \|G - F(\alpha, \cdot)\|.$$

We note here that the H-sets which characterize the best approximations consist of extremal points and points where the constraints are attained. We shall call all values of $x \in B$ for which the constraints are attained active points.

For the consideration of uniqueness we have

THEOREM 5. *Suppose $F(\alpha, \cdot)$ and $F(\beta, \cdot)$ are distinct best approximations to $G \in C(B)$ by W, then the extremal and active points of $G - F(\alpha, \cdot)$ and $G - F(\beta, \cdot)$ contain the same minimal H-sets with respect to V.*

Proof. Let $\{x_i, \lambda_i, e_i, k\}$ be a minimal H-set with respect to V such that $\{x_i\}_1^k$ is contained in the set of extremal and active points of $G - F(\alpha, \cdot)$, *and such that*

$$e_i[G(x_i) - F(\alpha, x_i)] = \rho(G), \qquad i = 1,\dots, p$$

and

$$e_i h(\alpha, x_i) = a \text{ or } b, \qquad i = p+1,\dots, k.$$

That such an H-set exists follows from Theorem 3.

To prove the theorem let us assume the contrary. First assume that $\{x_i\}_1^p$ is not contained in the extremal points for $G - F(\beta, \cdot)$ or for some $1 \leqslant j \leqslant p$,

$$e_j[G(x_j) - F(\beta, x_j)] \neq \rho(G),$$

then

$$e_i[G(x_i) - F(\beta, x_i)] \leqslant e_i[G(x_i) - F(\alpha, x_i)], \qquad i = 1,..., p,$$

and strict inequality holds for some i. Hence,

$$0 \leqslant e_i[F(\beta, x_i) - F(\alpha, x_i)], \qquad i = 1,..., p.$$

Due to the differentiability of f we have

$$F(\beta, x_i) - F(\alpha, x_i) = h(\beta - \alpha, x_i)\, f'(h(\gamma_i, x_i)), \qquad \gamma_i \in P,\ i = 1,..., k.$$

Using this relationship and the fact that f is strictly monotonic with no zero derivatives we obtain for $i = 1,..., p$,

$$e_i[F(\beta, x_i) - F(\alpha, x_i)] \geqslant 0,$$

implies

$$e_i h(\beta - \alpha, x_i) \geqslant 0 \qquad \text{for} \quad f' > 0.$$

This same inequality applies for $f' < 0$; in this case the signs $-e_i$ are used throughout (see Theorem 2). From our assumption strict inequality must hold for some i.

The other possibility contrary to the theorem is that for some $p + 1 \leqslant j \leqslant k$,

$$e_i h(\beta, x_i) \neq a \quad \text{or} \quad b.$$

Then again, noting the sign value of e_i in the characterization theorem, we obtain

$$e_i h(\beta - \alpha, x_i) \geqslant 0, \qquad i = p + 1,..., k,$$

with strict inequality for some i. This contradicts the assumption that $\{x_i, \lambda_i, e_i, k\}$ is a minimal H-set with respect to V (see [1, 2]); hence, the result follows.

COROLLARY 3. *If $\{x_i, \lambda_i, e_i, k\}$ is a minimal H-set with respect to V contained in the set of extremal and active points of $G - F(\alpha, \cdot)$, then $F(\alpha, \cdot)$ is a unique best approximation to G by W if $k = n + 1$.*

Proof. Suppose $F(\beta, \cdot)$ is a further best approximation to G by W, then from Theorem 5

$$F(\beta, x_i) - F(\alpha, x_i) = 0, \qquad i = 1,..., p.$$

Using the relations obtained therein we thus have

$$h(\beta - \alpha, x_i) = 0, \qquad i = 1,..., p,$$

because f' is never zero. Similarly for the active points the same equality holds. These linear homogeneous equations can be written as

$$(\beta_1 - \alpha_1,..., \beta_n - \alpha_n) \begin{pmatrix} g_1(x_1) \cdots g_1(x_k) \\ \vdots \quad\quad \vdots \\ g_n(x_1) \cdots g_n(x_k) \end{pmatrix} = (0,..., 0).$$

As the H-set is minimal with respect to V then this matrix of coefficients has rank $= k - 1$. Thus the only solution to this system is $\alpha - \beta = 0$ and uniqueness follows.

We note here that in the complex valued problem $k = 2n + 1$ is needed for uniqueness.

Our initial requirement that f be strictly monotonic with no zero derivative on $[a, b]$ is used throughout the paper to obtain the results of the theorems. If this requirement is forfeit then there are H-sets for W which are not H-sets with respect to V, hence they are not automatically computable. Some relaxation of this requirement can be made for functions such as $f(t) = \cos(t)$, defined on the interval $[0, \pi]$, where there are zero derivatives at the end points. In such cases every value $x \in B$ and $\alpha \in P$ such that $h(\alpha, x) = a$ or b is an H_1-set with respect to W at α, we need only put $p = 1$, $q = 0$ in Definition 1. However, because for $\beta \neq \alpha$ $h(\beta, x)$ does not necessarily equal a or b then we do not have an H_1-set with respect to W at β. Also x cannot form an H_2-set with respect to W at α nor form an H-set with respect to V.

Notwithstanding these difficulties we have only introduced one other type of H-set, hence little or no complication. Some change has to be made to Theorem 3 in the form of the following rider: "if the only H-sets found from the extremal and active points are H-sets consisting of single elements where $f' = 0$ then the best approximation is achieved but may only be locally best."

Theorems 4 and 5 remain true for such H-sets.

LINEAR TWO-SPACE

We now consider the special case when the linear space V is spanned by $(1, x, y)$, that is, the functions h are of the form $\alpha_1 + \alpha_2 x + \alpha_3 y$, $\alpha \in \mathbb{R}^3$. That approximation by V for strictly convex B and differentiable approximant is unique follows from [4], see also [6, Theorem 25]. For strictly convex B and differentiable approximant uniqueness of approximation was shown in [5] by functions of the form

$$\frac{1}{\alpha_1 + \alpha_2 x + \alpha_3 y}.$$

These examples are particular cases of our general theory when $f(t) = t$ and $f(t) = 1/t$, respectively. We now show that uniqueness follows in our more general setting of semi-non-linear functions of V.

Minimal H-sets in this setting are the minimal H-sets with respect to V, and they are the well-known forms given by (a), (b), and (c) in Fig. 1 (see [4]), where $\circ$, $\times$ are the points of the H-sets with $\circ$ having sign $+1$, and $\times$ sign -1, or vice versa. We can then show

THEOREM 6. *Let B be compact in $\mathbb{R}^2$, and $G \in C(B)$ have first partial derivatives on B. Then the best approximation $F(\alpha, \cdot)$ to G by W can only be non-unique if all the H-sets with respect to V, formed from the extremal and active points of of $G - F(\alpha, \cdot)$, are of type* (c), *and each point is a boundary point of B.*

Proof. Let N be the set of extremal and active points of $G - F(\alpha, \cdot)$. From Theorem 3 N contains points which form minimal H-sets with respect to V. If a minimal H-set with respect to V of type (a) or (b) occurs in N then uniqueness follows from the corollary of Theorem 5. If such H-sets do not occur then all the H-sets formed from N must be of type (c).

Suppose now that (x_1, y_1) belongs to an H-set of type (c) and $(x_1, y_1) \in N$, also suppose that (x_1, y_1) is an interior point of B, and that $F(\beta, \cdot)$ is a further best approximation to G by W. From Theorem 5 this H-set is an H-set formed from the extremal and active points of $G - F(\beta, \cdot)$. Also from Theorem 5 the functions $G - F(\alpha, \cdot)$ and $G - F(\beta, \cdot)$ have a

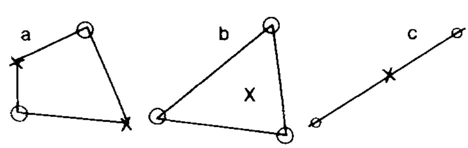

FIGURE 1

maximum or minimum at that point, even when (x_1, y_1) is an active point due to the continuity of G and F. As (x_1, y_1) is an interior point of B we must have zero partial derivatives for each error function, thus

$$\frac{\partial G}{\partial x}(x_1, y_1) = \alpha_2 f'(\alpha_1 + \alpha_2 x_1 + \alpha_3 y_1)$$

$$= \beta_2 f'(\beta_1 + \beta_2 x_1 + \beta_3 y_1),$$

$$\frac{\partial G}{\partial y}(x_1, y_1) = \alpha_3 f'(\alpha_1 + \alpha_2 x_1 + \alpha_3 y_1)$$

$$= \beta_3 f'(\beta_1 + \beta_2 x_1 + \beta_3 y_1).$$

Also from Theorem 5

$$f(\alpha_1 + \alpha_2 x_1 + \alpha_3 y_1) = f(\beta_1 + \beta_2 x_1 + \beta_3 y_1),$$

and hence from the monotonicity of f

$$f'(\alpha_1 + \alpha_2 x_1 + \alpha_3 y_1) = f'(\beta_1 + \beta_2 x_1 + \beta_3 y_1)$$

and

$$\alpha_1 + \alpha_2 x_1 + \alpha_3 y_1 = \beta_1 + \beta_2 x_1 + \beta_3 y_1.$$

From these equations we obtain the fact that $\alpha = \beta$, and hence uniqueness follows in this case. Hence, non-uniqueness will only occur if all the points of (c) type H-sets lie on the boundary of B, and the result follows.

References

1. M. Brannigan, H-sets in linear approximation, *J. Approx. Theory* **20** (1977), 153–161.
2. M. Brannigan, Uniform linear approximation in normed linear spaces, *Numer. Funct. Anal. Optim.* **2** (1980), 79–91.
3. M. Brannigan, Uniform non-linear constrained approximation in normed linear spaces. *Numer. Funct. Anal. Optim.* **6** (1983), 341–361.
4. L. Collatz, Approximation von Funktionen bei einer und bei mehreren unabhängigen Veränderlichen, *Z. Angew. Math. Mech.* **36** (1956), 198–211.
5. C. B. Dunham, Chebyshev approximation by first degree rationals on 2-space, *J. Approx. Theory* **16** (1976), 338–340.
6. G. Meinardus, "Approximation of Functions," Springer-Verlag, New York/Berlin, 1967.