JOURNAL OF APPROXIMATION THEORY 33, 296–307 (1981)

# On the Local Continuity of the Chebyshev Operator

R. HETTICH

*Universität Trier, FBIV, Postfach 3825, 55 Trier, West Germany*

AND

H. TH. JONGEN

*Technische Hogeschool Twente, Postbus 217, Enschede, The Netherlands*

*Communicated by G. Meinardus*

Received February 20, 1980

## 1. INTRODUCTION

Let $B$ be a compact metric space and $A$ a non empty subset of $C(B)$, $C(B)$ the space of continuous real functions on $B$ normed by

$$\|f\|_\infty = \max_{x\in B} |f(x)|.$$

Then, given $f \in C(B)$, the Chebyshev approximation problem is to find a best approximation $a^f \in A$, satisfying for every $a \in A$,

$$\|f - a^f\|_\infty \leqslant \|f - a\|_\infty .$$

If this inequality holds for all $a \in A \cap U$, $U \subset C(B)$ some neighborhood of $a^f$, $a^f$ is called a locally best approximation. It is well-known that, given $A$, the mapping

$$m: \quad \begin{array}{l} C(B) \to \mathbb{R} \\ f \quad \to m(f) = \inf\limits_{a\in A} \|f-a\|_\infty \end{array}$$

is continuous [5]. Much more problematic is the dependence of a best approximation $a^f$ on $f$. Let $M(f) \subset A$ be the (possibly empty) set of best approximations to $f$. Then one can define the mapping



0021-9045/81/120296–12\$02.00/0

$$M: \begin{array}{ll} C(B) & \to \mathscr{P}(A) \\ f & \to M(f) \end{array},$$

where $\mathscr{P}(A)$ is the set of subsets of $A$. In this general form the problem is treated in [2].

More usual is to consider instead of $M$ the so-called $T$-operator:

DEFINITION 1. Let $D_T \subset C(B)$ be the set of $f$ for which there is exactly one best approximation $a^f \in A$. The $T$-operator is the mapping

$$T: \begin{array}{ll} D_T & \to A \\ f & \to a^f \end{array}.$$

An important property in investigating $T$ is strong unicity:

DEFINITION 2. $a^f \in A$ is a strongly unique best approximation to $f$ if there is a $\gamma > 0$ such that for every $a \in A$

$$\|a - f\|_\infty \geqslant \|a^f - f\|_\infty + \gamma \|a^f - a\|_\infty .$$

If every $f \in C(B)$ has a strongly unique best approximation, then $T$ is said to have the strong unicity property.

The following theorems hold [5].

THEOREM 1 (Freud, cf. [4]). *If $T$ has the strong unicity property, then, for every $f \in C(B)$, there is a $\lambda = \lambda(f) > 0$ such that $\|Tf - Tg\|_\infty \leqslant \lambda \|f - g\|_\infty$ for every $g \in C(B)$. Especially, $T$ is continuous on $D_T = C(B)$.*

THEOREM 2 [15]. *Suppose that $B$ contains at least $n+1$ points. Let $A \subset C(B)$ be a linear Haar-subspace of dimension $n$. Then $T$ has the strong unicity property.*

The following theorem is important in considering continuity at a given $f$.

THEOREM 3. *Let $a^f$ be a strongly unique best approximation to $f$. If there is an $\varepsilon > 0$ such that $\{a \in A \mid \|a - a^f\|_\infty \leqslant \varepsilon\}$ is compact, then there is a neighborhood $U_f$ of $f$ and a constant $\lambda > 0$ such that for every $g \in U_f$ there exists a best approximation $a^g$ and $\|a^f - a^g\|_\infty \leqslant \lambda \|f - g\|_\infty$.*

In order to compute a best approximation numerically, usually a parametrization

$$a: \begin{array}{l} P \to A \\ p \to a(p, \cdot) \end{array}, \qquad P \subset \mathbb{R}^n,$$

is assumed to be given and $p^f \in P$ is to be determined such that $a(p^f, \cdot) = a^f$. Then from a numerical point of view it is an important question whether the appropriately defined function $f \to p^f$ is continuous. Let $\delta(a(p, \cdot))$ be the dimension of the tangent-space $S(a)$ at $a(p, \cdot)$ with respect to the parametrization. Then, if $S(a)$ is a Haar-space and, for all $a' \in A$, $a' - a$ has at most $\delta(a(p, \cdot)) - 1$ zeros in $B$, the normality of $f$—i.e., $\delta(a(p^f, \cdot)) = \max_{p \in P} \delta(a(p, \cdot))$—is sufficient for the assumptions of Theorem 3 to hold [1, 5]. This result can be applied for instance to rational an exponential approximation.

Haar's condition is very restrictive and actually does not hold for nontrivial $B \subset \mathbb{R}^m$, $m > 1$. Therefore, in this paper, we will proceed in a different way. In Section 2 we show that strong unicity is closely related to a sufficient optimality condition of first order to hold. This implies that in nonlinear approximation strong unicity is very restrictive. Therefore, instead of a first order condition, in Section 3 we assume a second order sufficient condition for $p^f$ to be optimal and show that an appropriately defined $T$-operator is locally continuous. We remark that differentiability of the functions under consideration is required for our investigations.

Concerning the numerical relevance of our results, we note that the assumptions required to ensure continuity imply convergence of a Newton-method generalizing the second algorithm of Remes [9]. Thus, the same assumptions imply convergence and numerical stability as well. This generalizes a similar result for strong unicity and the method of linearization [6]. Note that our assumptions are considerably weaker than that of strong unicity. Naturally, as less as normality and strong unicity, our assumptions in general cannot be verified a priori for a given problem.

## 2. Local Theory of First Order in Parameter Space

In the following we assume that $A$ is parametrized

$$a: \begin{array}{l} P \to A \\ p \to a(p, \cdot) \end{array}, \qquad P \subset \mathbb{R}^n \text{ open,}$$

and that $\bar{f} \in C(B)$ and $\bar{p} \in P$ are fixed.

Definition 3. An element $\bar{p} \in P$ is called a locally best approximation to $\bar{f}$ if there is a neighborhood $\bar{U} \subset P$ of $\bar{p}$ such that for every $p \in \bar{U}$

$$\|a(\bar{p}, \cdot) - \bar{f}\|_\infty \leqslant \|a(p, \cdot) - \bar{f}\|_\infty.$$

If, for some $\bar{U}$, equality implies $p=\bar{p}$, $\bar{p}$ is called locally unique. If there are $\bar{U}$ and $\gamma>0$ such that for every $p\in\bar{U}$

$$\|a(p,\cdot)-\bar{f}\|_\infty \geqslant \|a(\bar{p},\cdot)-\bar{f}\|_\infty+\gamma\,|\bar{p}-p| \tag{2.1}$$

($|\ |$ the Euclidean norm on $\mathbb{R}^n$), $\bar{p}$ is called locally strongly unique.

Let $M_L: C(B)\to\mathscr{P}(P)$, be the mapping for which the image of $f$ is the set of locally best approximations to $f$.

DEFINITION 4. Let $\bar{p}$ be a locally unique best approximation to $\bar{f}$. If there are neighborhoods $U_{\bar{f}}$, $U_{\bar{p}}$ of $\bar{f}$, $\bar{p}$ such that for every $f\in U_{\bar{f}}$ the set $M_L(f)\cap U_{\bar{p}}$ contains exactly one element $p^f$, then the local $T$-operator $t$ is defined by

$$t:\quad \begin{array}{c} U_{\bar{f}}\to U_{\bar{p}} \\ f\to p^f \end{array}.$$

*Remark.* If there are $U_{\bar{p}}\subset\mathbb{R}^n$ and $U_{a(\bar{p},\cdot)}\subset C(B)$ such that $a\colon U_{\bar{p}}\to A\cap U_{a(\bar{p},\cdot)}$ is bijective and if there are $\alpha>0$, $\beta>0$ such that

$$\alpha\,\|a(p,\cdot)-a(\bar{p},\cdot)\|_\infty\leqslant|p-\bar{p}|\leqslant\beta\,\|a(p,\cdot)-a(\bar{p},\cdot)\|_\infty,$$

then $a(\bar{p})$ is locally strongly unique if and only if $\bar{p}$ is and $t$ is continuous in $\bar{f}$ if and only if $T$ is.

Since the case $\bar{f}\in A$ is not very exciting, from now on we assume $\bar{f}\notin A$. Furthermore we assume that $a(p,x)$ has a continuous derivative $D_p a(p,x)$ with respect to $p$. Let

$$\bar{E}=\{x\in B\mid |\bar{f}(x)-a(\bar{p},x)|=\|\bar{f}-a(\bar{p},\cdot)\|_\infty\}. \tag{2.2}$$

LEMMA 1 (cf. [8]). *Let $\bar{\sigma}(x)=\operatorname{sign}(\bar{f}(x)-a(\bar{p},x))$. If the system of linear inequalities*

$$\bar{\sigma}(x)\,D_p a(\bar{p},x)\,\xi\geqslant 0,\qquad x\in\bar{E}, \tag{2.3}$$

*has no solution $\xi\neq 0$, then $\bar{p}$ is a locally unique best approximation.*

We show that the condition of Lemma 1 is even equivalent to $\bar{p}$ being locally strongly unique.

THEOREM 4. *Inequality* (2.3) *has no solution $\xi\neq 0$ if and only if $\bar{p}$ is a locally strongly unique best approximation.*

*Proof.* First, assume that $\xi_0$, $|\xi_0|=1$, solves (2.3). For $\tau$ sufficiently small, $p(\tau)=\bar{p}+\tau\xi_0\in P$. If $\bar{p}$ is locally strongly unique, then for an arbitrary sequence $\{\tau_i\}$, $\tau_i > 0$, $\lim_{i\to\infty} \tau_i = 0$, for every $i$ sufficiently large there exists an $x_i \in B$ such that

$$|\bar{f}(x_i) - a(p(\tau_i), x_i)| \geqslant \|\bar{f} - a(\bar{p}, \cdot)\|_{\infty} + \gamma\tau_i.$$

Since $B$ is compact we may assume that the sequence $\{x_i\}$ converges, $\lim_{i\to\infty} x_i = x^*$. It is simply proved by contradiction that $x^* \in \bar{E}$. Since $\bar{f} \notin A$, we have $\bar{\sigma}(x^*) \neq 0$. It is sufficient to consider the case $\bar{\sigma}(x^*) = 1$. Then, there is an $i_0$ such that $\bar{f}(x_i) - a(p(\tau_i), x_i) > 0$, $i > i_0$. Therefore, for such $i$

$$\begin{aligned} 0 &< \bar{f}(x_i) - a(\bar{p} + \tau_i \xi_0, x_i) \\ &= \bar{f}(x_i) - a(\bar{p}, x_i) - \tau_i D_p a(\bar{p}, x_i)\, \xi_o + \sigma(\tau_i) \\ &\leqslant \|\bar{f} - a(p, \cdot)\|_{\infty} - \tau_i D_p a(\bar{p}, x_i)\, \xi_0 + \sigma(\tau_i). \end{aligned}$$

Thus, for all $i > i_0$,

$$\gamma\tau_i \leqslant -\tau_i D_p a(\bar{p}, x_i)\, \xi_0 + \sigma(\tau_i).$$

This is seen to be a contradiction to $\gamma > 0$ by observing that (2.3) implies $\lim_{i\to\infty} D_p a(\bar{p}, x_i)\, \xi_0 = D_p a(\bar{p}, x^*)\, \xi_0 \geqslant 0$. Therefore, if (2.3) has a solution other than 0, then $\bar{p}$ is not locally strongly unique.

On the other hand, assume $\bar{p}$ is not locally strongly unique. Then there is a sequence $p^i$, $\lim_{i\to\infty} p^i = \bar{p}$, $p^i = \bar{p} + \tau_i \xi_i \in P$, $|\xi_i| = 1$, $\tau_i > 0$, such that $\|\bar{f} - a(p^i, \cdot)\|_{\infty} = \|\bar{f} - a(\bar{p}, \cdot)\|_{\infty} + \sigma(\tau_i)$. We may assume, that the sequence $\{\xi_i\}$ converges: $\lim_{i\to\infty} \xi_i = \xi_0$, $|\xi_0| = 1$.

Let $x \in \bar{E}$, $\sigma(x) > 0$.Then $\bar{f}(x) - a(p^i, x) \leqslant \bar{f}(x) - a(\bar{p}, x) + \sigma(\tau_i)$ which, by $\bar{f}(x) - a(p^i, x) = \bar{f}(x) - a(\bar{p}, x) - \tau_i D_p a(\bar{p}, x)\, \xi_i + \sigma(\tau_i)$, shows that $D_p a(\bar{p}, x)\, \xi_0 = \sigma(x)\, D_p a(\bar{p}, x)\, \xi_0 \geqslant 0$. The case $\sigma(x) < 0$ is analogous. Consequently $\xi_0$ solves (2.3).

## 3. Some Auxiliary Results

To derive a local theory of second oder some properties of the function spaces on $B$ are required which are given only for special regions $B$.

Definition 5. A nonempty compact subset $B \subset \mathbb{R}^m$ is called a Regular Approximation Region (RAR) if there are functions $h^i \in C^2(\mathbb{R}^m)$, $i = 1,\ldots, \ell$, such that

(i) $B = \{x \mid h^i(x) \leqslant 0, i = 1,\ldots, \ell\}$,

(ii) for every $x \in B$ the gradients $Dh^i(x)$, $i \in L(x) = \{i \mid h^i(x) = 0\}$, are linearly independent.

In the sequel $B$ is assumed to be a RAR with a set of $h^i$, $i = 1,\dots,\ell$, according to Definition 5, chosen once for all. Note that (ii) implies that the interior $B_0$ of $B$ is nonempty and that $B = \operatorname{clos}(B_0)$. Excluded by the definition are for instance $L$-shaped regions or regions with cusps.

By $C^2(B)$ we denote the vector space of real-valued functions, twice continuously differentiable in $B_0$, continuous on $B$, and such that all partial derivatives up to the second order can be extended to functions in $C(B)$. These extensions then are unique and with the norm

$$\|f\|_B = \max_{\substack{|k| \leqslant 2 \\ x \in B}} |f^k(x)|,$$

where $k = (k_1,\dots,k_m) \in (\mathbb{N} \cup \{0\})^m$, $|k| = \sum_{i=1}^m k_i$, and

$$f^k(x) = \frac{\partial^{|k|}}{\partial x_1^{k_1} \cdots \partial x_m^{k_m}} f(x),$$

it is obvious that $C^2(B)$ is a Banach space.

A proof of the following lemma, based on Whitney's Extension Theorem (cf. [14] for instance), is given in [11].

LEMMA 2. *Let $B$ be a RAR. Then, for every $f \in C^2(B)$, there exists an extension $\tilde{f} \in C^2(\mathbb{R}^m)$.*

Let $B$ and $\tilde{B}$ be RAR's such that $B \subset \tilde{B}_0$. The restriction map $\mathscr{R}: C^2(\tilde{B}) \to C^2(B)$, $\mathscr{R}\tilde{f} = \tilde{f}_{|C^2(B)}$, clearly is continuous and consequently $\mathscr{R}^{-1}(\Theta)$, $\Theta \in C^2(B)$ the null-function, is a closed linear subspace of $C^2(\tilde{B})$.

Let $C^2(\tilde{B} \mid B) = C^2(\tilde{B})/\mathscr{R}^{-1}(\Theta)$ be the quotient space with norm

$$|||L||| = \inf_{f \in L} \|f\|_{\tilde{B}}, \qquad L \in C^2(\tilde{B} \mid B).$$

It is well-known (cf. [3]) that $C^2(\tilde{B} \mid B)$ with this norm is a Banach-space. Moreover, the canonical projection $\mathscr{P}: C^2(\tilde{B}) \to C^2(\tilde{B} \mid B)$ is easily seen to be continuous and open, such that the topology given by $|||\ |||$ is the quotient topology with respect to $\mathscr{P}$.

This implies (cf. [12, p. 94ff.]) that a map $\phi: C^2(\tilde{B} \mid B) \to X$, $X$ a topological space, is continuous if and only if $\phi \circ \mathscr{P}: C^2(\tilde{B}) \to X$ is continuous.

Finally the mapping $\mathscr{R}_{\text{ind}}: C^2(\tilde{B} \mid B) \to C^2(B)$ induced by $\mathscr{R}$ is a linear, bijective and continuous map from one real Banach space to another and therefore, by the open mapping theorem, $\mathscr{R}_{\text{ind}}$ is a linear homeomorphism.

Altogether, the following lemma is proved.

LEMMA 3. *Let $B, \tilde{B}$ be RAR's such that $B \subset \tilde{B}_0$, $X$ a topological space, and $\Phi$: $C^2(\tilde{B}) \to X$ a continuous mapping with the property that $\Phi(f) = \Phi(g)$ if $f - g \in \mathscr{R}^{-1}(\Theta)$ (i.e., $\Phi$ depends only on elements from $C^2(\tilde{B} \mid B)$). Then $\Phi$ induces a unique continuous mapping $\Phi_{\text{ind}}: C^2(B) \to X$ such that the following diagram is commutative:*

$$\begin{array}{ccc} C^2(\tilde{B}) & \xrightarrow{\Phi} & X \\ \mathscr{r} \downarrow & & \uparrow \Phi_{\text{ind}} \\ C^2(\tilde{B} \mid B) & \xrightarrow[\mathscr{R}_{\text{ind}}]{} & C^2(B) \end{array}$$

Let $B$ be a RAR. For $|k| \leqslant 2$ we define the mappings

$$T_k: C^2(B) \times B_0 \to \mathbb{R} \qquad \text{by} \quad T_k(f, x) = f^k(x).$$

We remark that $C^2(B) \times B_0$ is an open subset of the Banach space $C^2(B) \times \mathbb{R}^m$ with norm $|||(f, x)||| = \|f\|_B + |x|$.

LEMMA 4. *$T_k$ is continuous for $|k| \leqslant 2$ and continuously Fréchet-differentiable for $|k| \leqslant 1$.*

*Proof.* A trivial calculation gives

$$|T_k(f, x) - T_k(g, y)| \leqslant \|f - g\|_B + |f^k(x) - f^k(y)|.$$

Together with the continuity of $f^k$ this implies the continuity of $T_k$ for each $k$, $|k| \leqslant 2$.

We prove continuous differentiability of $T_0$ only, the proof for $k$, $|k| = 1$, being analogous. Let $(f, x) \in C^2(B) \times B_0$ be fixed and $y \in \mathbb{R}^m$ be such that $x + y \in B_0$. Then

$$\begin{aligned} T_0(f + g, x + y) - T_0(f, x) &= Df(x)\,y + Dg(x)\,y + g(x) + o(|y|) \\ &= Df(x)\,y + g(x) + o(|||(g, y)|||). \qquad (3.1) \end{aligned}$$

The mapping

$$DT_0(f, x): \quad \begin{array}{c} C^2(B) \times \mathbb{R}^m \to \mathbb{R} \\ (g, y) \to Df(x)\,y + g(x) \end{array}$$

is linear and continuous since $T_0$ is continuous. Therefore, (3.1) shows that $DT_0(f, x)$ is the Fréchet-derivative of $T_0$ at $(f, x)$.

It remains to show that the mapping

$$DT_0: C^2(B) \times B_0 \to [C^2(B) \times \mathbb{R}^m]^*$$

is continuous ($*$ denotes the topological dual space). Let $|||\cdot|||^*$ be the norm on $[C^2(B)\times\mathbb{R}^m]^*$ induced by $|||\cdot|||$. Then

$$\begin{aligned}|||DT_0(f,x)-DT_0(\bar{f},\bar{x})|||^* &= \sup_{|||(g,y)|||\leqslant 1} |(DT_0(f,x)-DT_0(\bar{f},\bar{x}))(g,y)|\\ &\leqslant \sup_{|||(g,y)|||\leqslant 1} \{|y|\,|Df(x)-D\bar{f}(\bar{x})|+|g(x)-g(\bar{x})|\}.\end{aligned} \tag{3.2}$$

Since $B_0$ is open, it is no restriction of generality to assume that the segment $[x,\bar{x}]$ is in $B_0$. Then there exists $\Theta\in(0,1)$ such that

$$\begin{aligned}|g(x)-g(\bar{x})| &= |Dg(\bar{x}+\Theta(x-\bar{x}))(x-\bar{x})|\\ &\leqslant |Dg(\bar{x}+\Theta(x-\bar{x}))|\,|x-\bar{x}|.\end{aligned} \tag{3.3}$$

Observing that $|||(g,y)|||\leqslant 1$ implies $|y|\leqslant 1$, (3.2) and (3.3) together with the continuity of $T_k$, $|k|=1$, imply the continuity of $DT_0$ at $(f,x)$.

## 4. Local Theory of Second Order

Consider again the approximation problem of Section 1 but now with $B$ a RAR, $A\subset C^2(B)$, and $\bar{f}\in C^2(B)$. That means from now on, we restrict our considerations to the approximation of twice continuously differentiable functions. The set of approximating functions $A$ is assumed to be locally parametrized in the following sense:

Given $\bar{a}\in A$ there is a $\bar{p}\in\mathbb{R}^n$, an open neighborhood $P$ of $\bar{p}$ and a function $a\in C^2(P\times B)$ such that $a(\bar{p},\cdot)=\bar{a}$ and $a(p,\cdot)\in A$ for $p\in P$. The space $C^2(P\times B)$ is defined analogously to $C^2(B)$ and Lemma 2 holds in the sense that an $h\in C^2(P\times B)$ can be extended to an $\tilde{h}\in C^2(P\times\mathbb{R}^m)$.

In the sequel $\bar{a}\in A$ is fixed and $\bar{p}$, $P$, $a(p,\cdot)$ are given as above. As usual $\|f\|_\infty=\max_{x\in B}|f(x)|$, $f\in C(B)$. We recall that for $\bar{f}\in C^2(B)$, $a\in C^2(P\times B)$ the extensions of all partial derivatives to $\partial B=B\sim B_0$ and $P\times\partial B$ resp. are uniquely determined.

Let the set $\bar{E}$ of extremals of the error function

$$e(\bar{f},\bar{p},x)=\bar{f}(x)-a(\bar{p},x)$$

be given by (2.2). We need some nondegeneracy-assumptions which will be formulated now.

*Assumption* (A). (Cardinality of $\bar{E}$ related to the dimension of the parameter space $P$). There are exactly $r$, $r\leqslant n+1$, points in $\bar{E}$. *Let* $\bar{E}=\{x^1,\dots,x^r\}$.

*Assumption* (B). (Nondegeneracy of the extremal set with respect to B). Suppose (A) holds. Let $\bar{\sigma}_j = \operatorname{sign} e(\bar{f}, \bar{p}, \bar{x}^j)$, $j = 1,..., r$. Then, for $j = 1,..., r$, there are uniquely defined numbers $\bar{w}^{ij} > 0$, $i \in L(\bar{x}^j)$, such that

$$\bar{\sigma}_j D_x e(\bar{f}, \bar{p}, \bar{x}^j) = \sum_{i \in L(\bar{x}^j)} \bar{w}^{ij} Dh^i(\bar{x}^j) \tag{4.1}$$

and such that the quadratic form $\mu' M_j \mu$ is negative definite on the subspace

$$T_j = \{\mu \in \mathbb{R}^m \mid H_j \mu = 0\}, \tag{4.2}$$

where

$$H_j = \begin{pmatrix} \vdots \\ Dh^i(\bar{x}^j) \\ \vdots \end{pmatrix}_{i \in L(\bar{x}^j)} \tag{4.3}$$

an $m \times |L(\bar{x}^j)|$-matrix, and, with $D^2$ denoting the matrix of second derivatives,

$$M_j = \bar{\sigma}_j D_x^2 e(\bar{f}, \bar{p}, \bar{x}^j) - \sum_{i \in L(\bar{x}^j)} \bar{w}^{ij} D^2 h^i(\bar{x}^j). \tag{4.4}$$

*Remark.* The conditions of (B) imply that the $\bar{x}^j$ are locally strict extrema of $e(\bar{f}, \bar{p}, x)$ on $B$.

Assumptions (A) and (B) imply that the extremals of the error function $e(f, p, x)$ locally may be considered as continuous functions of $f$ and $p$. More precisely we have:

THEOREM 5. *Assume that* (B) *holds. Then, for* $j = 1,..., r$, *there are neighborhoods* $U_{\bar{f}} \subset C^2(B)$, $U_{\bar{p}} \subset P$, $U_{\bar{x}^j} \subset B$ *of* $\bar{f}$, $\bar{p}$, $\bar{x}^j$ *resp., and continuous functions* $x^j: U_{\bar{f}} \times U_{\bar{p}} \to U_{\bar{x}^j}$, $x^j(\bar{f}, \bar{p}) = \bar{x}^j$, *such that for every pair* $(f, p) \in U_{\bar{f}} \times U_{\bar{p}}$ *the points* $x^j(f, p) \in B$ *are the only local extrema of* $e(f, p, x)$ *in* $\bigcup_{j=1}^r U_{\bar{x}^j}$ *and such that*

$$\|e(f, p, \cdot)\|_\infty = \max_{1 \leqslant j \leqslant r} |e(f, p, x^j(f, p))|.$$

*Proof.* Let $\tilde{B}$ be a RAR such that $B \subset \tilde{B}_0$. Let $\tilde{f}$, $\tilde{a}(p, \cdot) \in C^2(\tilde{B})$ be the extensions of $f$ and $a(p, \cdot)$ according to Lemma 2. For arbitrary but fixed $j \in \{1,..., r\}$ consider the equations

$$\begin{aligned} \bar{\sigma}_j \{D\tilde{f}(x^j) - D_x \tilde{a}(p, x^j)\} - \sum_{i \in L(\bar{x}^j)} w^{ij} Dh^i(x^j) &= 0, \\ h^i(x^j) = 0, \qquad i &\in L(\bar{x}^j). \end{aligned} \tag{4.5}$$

By Lemma 4 the left-hand sides of (4.5) depend continuously Fréchet-differentiable on $\tilde{f}$, $p$, $x^j$, $w^{ij}$on $C^2(\tilde{B}) \times P \times \tilde{B}_0 \times \mathbb{R}^{|L(\bar{x}^j)|}$. From (B) we see that $\tilde{\bar{f}}$, $\bar{p}$, $\bar{x}^j$, $\bar{w}^{ij}$ solve (4.5). Moreover (cf. [7]), (B) implies that the Jacobian in $\tilde{\bar{f}}$, $\bar{p}$, $\bar{x}^j$, $\bar{w}^{ij}$ with respect to $x^j$, $w^{ij}$ is nonsingular.

Therefore the Implicit Function Theorem (cf. [13]) may be applied and yields the existence of neighborhoods $\tilde{U}_{\bar{f}} \subset C^2(\tilde{B})$, $\tilde{U}_{\bar{p}} \subset P$, $\tilde{U}_{\bar{x}^j} \subset B$ and continuously Fréchet-differentiable functions $\tilde{x}^j: \tilde{U}_{\bar{f}} \times \tilde{U}_{\bar{p}} \to \tilde{U}_{\bar{x}^j}$, $\tilde{w}^{ij}: \tilde{U}_{\bar{f}} \times \tilde{U}_{\bar{p}} \to \mathbb{R}_+$ such that for every $(f, p) \in \tilde{U}_{\bar{f}} \times \tilde{U}_p$ the only solution of (4.5) in $\tilde{U}_{\bar{x}^j} \times \mathbb{R}_+^{|L(x^j)|}$ is $\tilde{x}^j(f, p)$, $\tilde{w}^{ij}(f, p)$, $i = 1,..., L(\bar{x}^j)$.

We remark that $\tilde{U}_{\bar{x}^j} \subset B$ is possible due to the equations $h^i(x^j) = 0$, $i \in L(\bar{x}^j)$. From $\tilde{U}_{\bar{x}^j} \subset B$ it follows that $\tilde{x}^j$ only depends on the values of $\tilde{f}$, $\tilde{a}(p, \cdot)$ in $B$. Therefore, Lemma 3 yields the existence of $\hat{U}_{\bar{f}}$ $(= \mathscr{R}_{\text{ind}} \circ \mathscr{P}(\tilde{U}_{\bar{f}}))$, $\hat{U}_{\bar{p}} = \tilde{U}_{\bar{p}}$, $\hat{U}_{\bar{x}^j} = \tilde{U}_{\bar{x}^j} \subset B$ and continuous functions $x^j$ $(= \tilde{x}^j_{\text{ind}})$: $\hat{U}_{\bar{f}} \times \hat{U}_{\bar{p}} \to \hat{U}_{\bar{x}^j}$ such that $x^j(f, p)$ are the only candidates for extrema of $e(f, p, x)$ in $\hat{U}_{\bar{x}^j}$.

The remainder of the proof is by standard arguments.

*Remark.* As a consequence of Theorem 2, if assumption (B) holds, the problem locally can be reduced to a discrete one with discretization points $x^j(f, p)$ depending on $f$ and $p$ (cf. [10]).

Finally, to investigate the dependence of $p$ and $f$ we need:

*Assumption* (C) (Nondegeneracy in the parameter space). Assumption (B) holds. Moreover, there are uniquely defined $\bar{u}^j > 0$, $\sum_{j=1}^r \bar{u}^j = 1$ such that

$$\sum_{j=1}^{r} \bar{u}^j \bar{\sigma}_j D_p a(\bar{p}, \bar{x}^j) = 0 \tag{4.6}$$

and such that for every $\xi \in K \sim \{0\}$,

$$K = \{\xi \mid D_p a(\bar{p}, \bar{x}^j)\, \xi = 0, j = 1,..., r\} \tag{4.7}$$

we have

$$\xi^T \left[ \sum_{j=1}^{r} \bar{u}^j \bar{\sigma}_j D_p^2 a(\bar{p}, \bar{x}^j) \right] \xi - \sum_{j=1}^{r} \bar{u}^j \bar{\sigma}_j \bar{\mu}_j^T M_j \bar{\mu}_j > 0, \tag{4.8}$$

where $\bar{\mu}_j$ is uniquely determined from

$$\begin{pmatrix} M_j & H_j^T \\ H_j & 0 \end{pmatrix} \begin{pmatrix} \bar{\mu}_j \\ z \end{pmatrix} = \begin{pmatrix} -D_{xp}^2 a(\bar{p}, \bar{x}^j)\, \xi \\ 0 \end{pmatrix}, \tag{4.9}$$

with $D_{xp}^2 = D_p[D_x^T]$, and $H_j$, $M_j$ given by (4.3), (4.4).

Note that the unicity of $\bar{u}^j$ implies that every set of at most $r - 1$ of the vectors $D_p a(\bar{p}, \bar{x}^j)$, $j = 1,..., r$, is linearly independent.

*Remark.* Assumption (C) is a sufficient condition for $\bar{p}$ to be a locally unique best approximation (cf. [7]).

The following theorem is the main result of this paper:

THEOREM 6. *Assume that* (C) *holds. Then there are neighborhoods* $U_{\bar{f}} \subset C^2(B)$, $U_{\bar{p}} \subset P$, $U_{\bar{x}^j} \subset B$, $U_{\bar{u}^j} \subset \mathbb{R}_+$, $j = 1,..., r$, $U_{\bar{w}^{ij}} \subset \mathbb{R}_+$, $i \in L(\bar{x}^j)$, $j = 1,..., r$, *of* $\bar{f}$, $\bar{p}$, $\bar{x}^j$, $\bar{u}^j$, $\bar{w}^{ij}$ *resp. and continuous functions* $p: U_{\bar{f}} \to U_{\bar{p}}$, $x^j: U_{\bar{f}} \to U_{\bar{x}^j}$, $u^j: U_{\bar{f}} \to U_{\bar{u}^j}$ *and* $w^{ij}: U_{\bar{f}} \to U_{\bar{w}^{ij}}$ *with* $p(\bar{f}) = \bar{p}$, $x^j(\bar{f}) = \bar{x}^j$, $u^j(\bar{f}) = \bar{u}^j$, $w^{ij}(\bar{f}) = \bar{w}^{ij}$ *and such that for every* $f \in U_{\bar{f}}$ $p(f)$ *is a locally unique best approximation to* $f$ *and Assumption* (C) *holds with* $x^j(f)$, $u^j(f)$, $w^{ij}(f)$.

*Proof.* The proof of Theorem 6 follows the same line as that of Theorem 5. Instead of (4.5), now the following system is considered:

$$
\begin{aligned}
\sum_{j=1}^{r} u^j \bar{\sigma}_j D_p a(p, x^j) &= 0, \\
\sum_{j=1}^{r} u^j - 1 &= 0, \\
\bar{\sigma}_j |f(x^j) - a(p, x^j)| - d &= 0, \qquad j = 1,..., r, \\
\bar{\sigma}_j [Df(x^j) - D_x a(p, x^j)] - \sum_{i \in I(\bar{x}^j)} w^{ij} Dh^i(x^j) &= 0, \qquad j = 1,..., r, \\
h^i(x^j) = 0, \quad i \in L(\bar{x}^j), & \qquad j = 1,..., r.
\end{aligned} \tag{4.10}
$$

Assumption (C) shows that $\bar{f}$, $\bar{p}$, $\bar{x}^j$, $\bar{u}^j$, $\bar{w}^{ij}$ is a solution of (4.10). Moreover (cf. [9]), (C) implies that the Jacobian of (4.10) with respect to $p$, $x^j$, $u^j$, $w^{ij}$ in $\bar{f}$, $\bar{p}$, $\bar{x}^j$, $\bar{u}^j$, $\bar{w}^{ij}$ is nonsingular. Therefore, the Implicit Function Theorem may be applied in the same way as in the proof of Theorem 5.

To complete the analogy wih the local theory of first order, we define:

DEFINITION 6. $\bar{p}$ is said to be a locally strongly unique best approximation of second order to $\bar{f}$, if there exist a neighborhood $U_{\bar{p}} \subset P$ of $\bar{p}$ and a $\gamma > 0$ such that for every $p \in U_{\bar{p}}$

$$\|a(p, \cdot) - \bar{f}\|_\infty \geqslant \|a(\bar{p}, \cdot) - \bar{f}\|_\infty + \gamma |\bar{p} - p|^2. \tag{4.11}$$

THEOREM 7. *Suppose that* (C) *holds. Then* $\bar{p}$ *is a locally strongly unique best approximation of second order to* $\bar{f}$.

*Proof.* In [9] it has been shown that (C) implies that for $\xi \in \mathbb{R}^n$, $|\xi| = 1$

$$\|\bar{f} - a(\bar{p} + t\xi, \cdot)\|_\infty = \|\bar{f} - a(\bar{p}, \cdot)\|_\infty + \max\{t\bar{\sigma}_j D_p a(\bar{p}, \bar{x}^j)\,\xi$$
$$+ \frac{t^2}{2}\,[\underbrace{\xi^T \bar{\sigma}_j D_p^2 a(\bar{p}, \bar{x}^j)\,\xi - \bar{\sigma}_j \bar{\mu}_j^T M_j \bar{\mu}_j}_{=\xi^T A_j \xi}]\} + o(t^2).$$

Observing that by (C) the matrix $\sum_{j=1}^{r} \bar{u}^j A_j$ is positive definite on $\{\xi \mid D_p a(\bar{p}, \bar{x}^j)\,\xi = 0 \mid j = 1,..., r\}$, (4.11) is easily established.

## References


1. R. B. Barrar and H. L. Loeb, On the continuity of the nonlinear Tschebyscheff operator, *Pacific J. Math.* **32** (1970), 587–601.
2. B. Brosowski, Zur stetigen Abhängigkeit der Menge der Minimalpunkte bei gewissen Minimierungsaufgaben, in "Approximation Theory, Bonn 1976" (R. Schaback and K. Scherer, Eds.), pp. 63–72, Lecture Notes in Mathematics No. 556, Springer-Verlag Berlin/Heidelberg/New York, 1976.
3. A. L. Brown and A. Page, "Elements of Functional Analysis," Van Nostrand–Reinhold, London, 1970.
4. E. W. Cheney, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
5. L. Collatz and W. Krabs, "Approximationstheorie," Teubner, Stuttgart, 1973.
6. L. Cromme, Eine Klasse von Verfahren zur Ermittlung bester nichtlinearer Tschebyscheff-Approximationen, *Numer. Math.* **25** (1976), 447–459.
7. R. Hettich, Kriterien zweiter Ordnung für lokal beste Approximationen, *Numer. Math.* **22**(1974), 409-417.
8. R. Hettich, Kriterien erster und zweiter Ordnung für lokal beste Approximationen bei Problemen mit Nebenbedingungen, *Numer. Math.* **25** (1975), 109–122.
9. R. Hettich, A Newton method for nonlinear Chebyshev approximation, *in* "Approximation Theory, Bonn 1976" (R. Schaback and K. Scherer, Eds.), pp. 222–236, Lecture Notes in Mathematics No. 556, Springer-Verlag, Berlin/New York, 1976.
10. R. Hettich and W. van Honstede, On quadratically convergent methods for semi-infinite-programming, *in* "Semi-Infinite-Programming" R. Hettich (Ed.), pp. 97–111, Lecture Notes in Control and Information Sciences, Springer-Verlag, Berlin/Heidelberg/New York, 1979.
11. R. Hettich and H. Th. Jongen, A Note on the Banach Space $C^k(\bar{\Omega})$, Twente University of Technology, Enschede, The Netherlands, Memorandum TW 267, 1979.
12. J. L. Kelley, "General Topology," Van Nostrand–Reinhold, New York, 1955.
13. S. Lang, "Analysis II," Addison–Wesley, Reading, Mass., 1966.
14. B. Malgrange, "Ideals of Differentiable Functions," Oxford Univ. Press, Oxford, 1966.
15. D. J. Newman and H. S. Shapiro, Some theorems on Chebyshev approximation, *Duke Math. J.* **30** (1963), 673–682.