JOURNAL OF APPROXIMATION THEORY 37, 29–43 (1983)

A Unified Theory of Strong Uniqueness in Uniform Approximation with Constraints

B. L. CHALMERS

Department of Mathematics, University of California, Riverside, California 92521, U.S.A.

AND

G. D. TAYLOR

Department of Mathematics, Colorado State University Ft. Collins, Colorado 80523, U.S.A.

Communicated by Oved Shisha

Received October 9, 1981; revised April 19, 1982

The previously developed unified theory of constrained uniform approximation from a finite dimensional subspace is extended to treat strong uniqueness and continuity of the best approximation operator.

In this paper the authors extend the unified theory developed in [1, 2] (covering existence, characterization, uniqueness, and computation of best approximations) to treat strong uniqueness and continuity of the best approximation operator.

The usual setting of this theory is as follows. Denote by $C(E)$ the class of all continuous real valued functions defined on E, a compact subset of $[a, b]$ containing at least $n+1$ points, normed with the uniform (Chebyshev) norm, $\|f\| = \max\{|f(t)|: t \in E\}$. Let $V \subset C(E)$ be an n-dimensional subspace of approximants and let V_0 be a nonempty subset of V that is determined by certain linear constraints. Then, given $f \in C(E)$, one says $p \in V_0$ is a best approximation (satisfying the constraints) to f if and only if

$$\|f - p\| = \inf\{\|f - q\|; q \in V_0\}.$$

Examples are monotone approximation and restricted range approximation without equality constraints, both of which are examples of

* Research supported in part by the National Science Foundation, under Grant MCS-80-17056.

0021-9045/83/010029-15$03.00/0

restricted derivatives approximation (RDA). These examples satisfy the generalized Haar and nearly Haar conditions described below. Further examples are bounded coefficients approximation, ε-interpolator approximation, and polynomial approximation with interpolation. All of these are examples where the Haar (and therefore generalized and nearly Haar) condition holds.

The general setting in which all of the above examples lie is best described by a functional analytic approach as follows.

Let A be a compact set of linear functionals in the dual, V^* of V. Note that for each p in V, $\alpha(p)$ is a continuous function on A as α ranges over A. Set

$$V_0 = \{p \in V; l(\alpha) \leqslant \alpha(p) \leqslant u(\alpha), \alpha \in A\},$$

where l and u are extended real-valued functions on A with $l < +\infty$, $u > -\infty$, the set E_l (resp. E_u) on which l (resp. u) is finite is closed, l (resp. u) is continuous on E_l (resp. E_u), and $l(\alpha) \leqslant u(\alpha)$.

Let e_x represent point evaluation at x in E (i.e., $e_x(f) = f(x)$, $\forall f \in C(E)$). Fix f in $C(E) \sim V_0$ with the restriction that if $\alpha = e_x$ for some α in A and some x in E, then $\inf\{\|f - q\|; q \in V_0\} > \max\{l(\alpha) - f(x), f(x) - u(\alpha)\}$; note that this inequality is assured if, for example, $l(\alpha) \leqslant f(x) \leqslant u(\alpha)$. Call such an f admissible. We are concerned then with approximating such admissible f by elements of V_0. One can check that all of the aforementioned examples lie in the above setting.

For example, ordinary monotone approximation, where $V_0 = \{p \in \pi_{n-1}[a, b]; p' \geqslant 0\}$, fits into the scheme as follows. Let e_x^k represent point evaluation of the kth derivative at x in E (i.e., $e_x^k(p) = p^{(k)}(x)$). Set $E = [a, b]$, $V = \pi_{n-1}$, and $A = \{e_x^1; x \in E\}$. Then for each $\alpha = e_x^1 \in A$, $\alpha(p) = p'(x)$, $l(\alpha) \equiv 0$, $u(\alpha) \equiv +\infty$. Then $V_0 = \{p \in V; l(\alpha) \leqslant \alpha(p) \leqslant u(\alpha), \alpha \in A\}$. Note that A is homeomorphic to E with the usual topology. Thus, $\alpha(p) = p'(x)$ is a continuous function on A for each p in V. Moreover, any $f \in C(E) \sim V_0$ is vacuously admissible.

In the following, motivated primarily by [11] and [14] dealing with monotone approximation and by [4] and [6], the authors obtain a unified theory treating the question of strong uniqueness (and the related continuity properties of the best approximation operator) in the general setting described above. Since these results depend on the unified theory previously developed, it is necessary to review a portion of the latter.

By the usual continuity and compactness argument we have the following result.

THEOREM 1 (Existence). *If V_0 is not empty, then there exists a best approximation in V_0 to f.*

DEFINITION 1. If $l(\alpha)=u(\alpha)$ implies α is an isolated point of A, we will say equality condition 1 (EQC1) is satisfied.

All of the preceding examples satisfy EQC1 or can be reformulated so that EQC1 holds (e.g., one-sided approximation with interpolatory restraints). Examples where EQC1 does not hold are restricted range with equality constraints and co-positive approximation.

Note 1. Throughout the remainder of this paper it is to be assumed that EQC1 is satisfied.

In [1] the following definitions generalizing the concept of a Haar subspace are given.

DEFINITION 2. For p in V_0 a set $S=I_1\cup I_3$ with $I_1\subset A$ and $I_3\subset\{e_x; x\in E\}$ is called an *extremal set* for f and p provided

(i) $\alpha(p)=u(\alpha)$ (or $l(\alpha)$), $\alpha\in I_1$,

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

(iii) $e_x\notin I_1$ if $|(f-p)(x)|=\|f-p\|$.

To each $\alpha\in A$ we associate a set (possibly empty) of elements B_α in V^* such that if $p\in V_0$ then $\alpha(p)=l(\alpha)$ (or $u(\alpha)$) implies that for each β in B_α, $\beta(p)=m(\beta)$ (or $n(\beta)$), where $m(\beta)$ (or $n(\beta)$) is some real number depending only on β.

DEFINITION 3. $S'=S\cup I_2$ is called an *augmented extremal set* for f and p if S is an extremal set for f and p and $I_2\subset\bigcup_{\alpha\in I_1}B_\alpha$.

EXAMPLES. In the case of monotone approximation, $V_0=\{p\in\pi_{n-1}[a,b]; p'\geqslant 0\}$, if $p\in V_0$ and $\alpha(p)=p'(x)=0$ for some $x\in(a,b)$, *then* $\beta(p)=p''(x)=0=m(\beta)$.

Another example of this is found in a combination of monotone and interpolatory constraints. For example, if $V_0=\{p\in\pi_5[0,1]; p'(x)\geqslant 0$ for all $x\in[0,1]$ and $p'''(\frac{1}{2})=0\}$ and S is an extremal set for some f and p that contains α, where $\alpha(q)=q'(\frac{1}{2})$, then the two linear functionals β_1 and β_2, with $\beta_1(q)=q''(\frac{1}{2})$ and $\beta_2(q)=q^{(iv)}(\frac{1}{2})$, adjoined to S will give an augmented extremal set for f and p with $m(\beta_1)=m(\beta_2)=0$.

NOTATION. For f and p fixed, let $S^{\max}$ denote the maximal extremal set for f and p, and let $S^{\max}_{\text{aug}}$ denote the maximal augmented extremal set for f and p.

NOTATION. In the folliwing, p_f will always denote a best approximation in V_0 to f.

DEFINITION 4. *V is generalized Haar with respect to f and p_f*, provided that if $S_{\text{aug}}^{\max}$ for f and p_f has order t, then $S_{\text{aug}}^{\max}$ contains $\min(t, n)$ elements which are linearly independent in V^*. *V is generalized Haar* if V is generalized Haar for all admissible pairs f and p_f (i.e., p_f is a best approximation to f from V_0).

In [1] it is verified that in all the preceding linear examples satisfying EQC1, the generalized Haar condition holds.

The proof of the uniqueness theorem (Theorem 2) below depends on the following partial characterization of best approximations which we will also need below.

LEMMA 1 (Partial characterization). *If V is generalized Haar with respect to f and p_f, then there exists an augmented extremal set for f and p_f of order $n+1$.*

For the proof of Lemma 1 see [1]. An alternate proof, however, under the additional mild assumption (*) below, follows as a corollary to Theorem 4 below (see Note 3). Indeed, the logic of the proof of Lemma 1 given in [1] is similar to the logic of the proof of Theorem 3 (and its corollary Theorem 4) below. We include the proofs of Theorems 2–4 for the sake of completeness and because they are relatively short.

THEOREM 2 (Uniqueness). *If V is generalized Haar, then any best approximation p_f in V_0 to f is unique.*

Proof. If also p^* is a best approximation in V_0 to f then so is $p^{**} = \frac{1}{2}p_f + \frac{1}{2}p^*$, since V_0 is convex. Thus by Lemma 1 there exists a maximal augmented extremal set for f and p^{**} of order $t \geqslant n+1$, say $S_{\text{aug}}^{\max} = I_1 \cup I_3 \cup I_2$ (see Definitions 2 and 3). Then $l(\alpha) \leqslant \alpha(p^{**})$, $\alpha(p_f)$, $\alpha(p^*) \leqslant u(\alpha)$, and $\alpha(p^{**}) = l(\alpha)(u(\alpha))$ therefore implies that $\alpha(p_f) = \alpha(p^*) = l(\alpha)(u(\alpha))$, $\forall \alpha \in I_1$. Hence $\beta(p^*) = \beta(p_f) = m(\beta)(n(\beta))$ $\forall \beta \in I_2$. Finally $|f(x) - p^{**}(x)| = \|f - p^{**}\| = \|f - p_f\| = \|f - p^*\|$ implies $p_f(x) = p^*(x)$, $\forall e_x \in I_3$. But since $t \geqslant n+1$ and some n of the elements of $S_{\text{aug}}^{\max}$ are linearly independent in the dual of V, we have $p_f = p^*$. ∎

Note 2. As an application, uniqueness of best approximation follows in the case of restricted derivatives approximation [1]; also this result was obtained independently and by different methods in [13].

DEFINITION 5. For $p \in V$ define a "signature" function σ on $S^{\max}$ by

$$\sigma(e_x) = \begin{cases} 1 & \text{if } e_x(p) = f(x) - \|f - p\| \\ -1 & \text{if } e_x(p) = f(x) + \|f - p\|, \end{cases}$$

$$\sigma(\alpha) = \quad 1 \qquad \text{if} \quad \alpha(p) = l(\alpha)$$
$$\quad = -1 \qquad \text{if} \quad \alpha(p) = u(\alpha) \neq l(\alpha).$$

That is, σ is 1 at lower extrema and -1 at upper extrema. Note that the admissibility of f insures that σ is well defined (in case e_x is in I_1 or I_3).

Let $S_E = \{\alpha \in A;\ l(\alpha) = u(\alpha)\}$ and recall that this set consists of isolated points of A. We will also need the "0 in the convex hull" criterion for best approximation (which holds in fact in the absence of any Haar condition). For this we must make the additional, very mild, "nonempty"-type assumption

$$\exists\ p_0 \in V_0 \text{ such that } l(\alpha) < \alpha(p_0) < u(\alpha), \qquad \forall\, \alpha \in A \sim S_E. \qquad (*)$$

Define $S^\sigma = \{\sigma(\gamma')\,\gamma';\ \gamma' \in S^{\max} \sim S_E\}$. Set $\bar{V} = \{p \in V;\ \alpha(p) = 0 \text{ for all } a \in E_E\}$ and note that $\dim \bar{V} = n - \dim[S_E]$.

THEOREM 3 (Kolmogorov criterion). *Let $f \in C(E) \sim V_0$, and let $p^* \in V_0$ and S^σ be defined as above, where $S^{\max}$ is the maximal extremal set for f and p^*. Then $p^* \in V_0$ is a best approximation to f iff*

$$\max_{\gamma \in S^\sigma} (-\gamma(p)) \geqslant 0 \qquad \textit{for all} \quad p \in \bar{V}.$$

Proof. p^* is a best approximation to f *iff* $\nexists\, p \in \bar{V}$ such that $p^* + \varepsilon p$ (for sufficiently small $\varepsilon > 0$) strictly improves upon p^* at the extrema $(S^{\max} \sim S_E)$ (consideration can be restricted to these extrema by the usual continuity and compactness argument and $(*)$ insures that the improvement at $A \cap (S^{\max} \sim E)$ is strict without loss since if for instance $\alpha(p^* + \varepsilon p) = u(\alpha)$, then p can be replaced by $(1 - \delta)p + \delta p_0$ for $\delta > 0$ sufficiently small), i.e., *iff* $\nexists\, p$ in $\bar{V}$ such that $\operatorname{sgn} \gamma(p) = \sigma(\gamma)$ for all $\gamma \in S^\sigma$, i.e., *iff* $\forall\, p$ in $\bar{V}$, $\max_{\gamma \in S^\sigma}(-\gamma(p)) \geqslant 0$. ∎

As a corollary of Theorem 3 we obtain the following very useful criterion for best approximation.

THEOREM 4 ("0 in the convex hull"-criterion). *p^* is a best approximation to $f \in C(E) \sim V_0$ iff 0 is in the convex hull of some τ ($\leqslant \dim \bar{V} + 1$) elements of $S^\sigma|_{\bar{V}}$, i.e.,*

$$0 = \sum_{i=1}^{\tau} \lambda_i \gamma_i \ \textit{on}\ \bar{V}, \qquad \textit{where} \quad \gamma_i \in S^\sigma, \lambda_i > 0, i = 1,\dots, \tau. \qquad (1)$$

Proof. Let $\dim \bar{V} = m$ and identitfy $\bar{V}$ with $\mathbb{R}^m$. Then $\bar{V}^*$ can of course be identified with (another copy of) $\mathbb{R}^m$. Then $S^\sigma|_{\bar{V}} \subset \bar{V}^*$ and, for $\gamma \in S^\sigma|_{\bar{V}}$

and $p \in \bar{V}$, $\gamma(p)$ is realized as a "dot" product of two m-vectors. Thus $\max_{\gamma \in S^\sigma}(-\gamma(p)) \geqslant 0$ for all $p \in \bar{V}$ represents the fact that for the set $S^\sigma|_{\bar{V}}$ there is no "direction" $p \in \bar{V}$ for which all vectors in $S^\sigma|_{\bar{V}}$ have a negative component. That is, $S^\sigma|_{\bar{V}}$ cannot lie in a half-space in $\mathbb{R}^m$; hence 0 must lie in the convex hull of (τ vectors in) $S^\sigma|_{\bar{V}}$. The fact that τ can be taken $\leqslant m+1$ is Carathéodory's result. ∎

Note 3. As mentioned above, Lemma 1 now follows as a corollary from Theorem 4 since if $t \leqslant n$ in Definition 4, then all the elements of $S_{\text{aug}}^{\max}$ are independent over V. But from formula (1) we conclude that some $\tau' = \tau + n - \dim[S_E]$ members of $S_{\text{aug}}^{\max}$ (in fact some τ' members of $S^{\max}$) are dependent over V. Hence $t \geqslant n+1$.

Another Haar-type condition which again holds for all the preceding linear examples satisfying EQC1, and which is useful for our strong uniqueness discussion, is formulated in [2].

DEFINITION 6. *V is nearly Haar on $\Omega = A \cup \{e_x; x \in E\}$* provided the set of n-tuples $(R_1, R_2, \ldots, R_n) \in \Omega^n$, where the R_i are linearly dependent, forms a closed nowhere dense subset of Ω^n. (Example: monotone approximation and, more generally, restricted derivatives approximation.) *V is Haar* (on Ω) if any distinct n elements in Ω are linearly independent. (Examples: bounded coefficients approximation, restricted range approximation, approximation with Hermite–Birkhoff interpolatory constraints, and, more generally, restrictions at poised Birkhoff data.)

Note 4. If V is Haar (on Ω) then V is both generalized Haar and nearly Haar.

Note 5. If V is nearly Haar on Ω then $\tau = n+1$ in Theorem 4 almost always.

We are now in a position to prove our strong uniqueness (and continuity of the best approximation operator) results.

NOTATION. If V is generalized Haar, let p_f denote the unique (by Theorem 2) best approximation in V_0 to f.

We extend a definition of Schmidt [14] to our setting.

DEFINITION 7. If V is generalized Haar, we say that *p_f is strongly unique of order r* $(0 < r \leqslant 1)$ if, given $N > 0$, there is a constant $\gamma = \gamma(N, f) > 0$ such that

$$\|f - p\| \geqslant \|f - p_f\| + \gamma \|p - p_f\|^{1/r} \text{ for all } p \in V_0 \text{ satisfying } \|p\| \leqslant N. \quad (2)$$

In the case that $r = 1$, the dependence on N is dropped.

NOTATION. Let $S_1^{\max}$, $S_2^{\max}$, $S_3^{\max}$ denote the I_1, I_2, I_3 subsets, respectively, of $S_{\text{aug}}^{\max}$, the maximal augmented extremal set for f and p_f. Further let S_1' and S_3 be the subsets of $S_1^{\max}$ and $S_3^{\max}$, respectively, whose elements appear in (1) (with $\lambda_i > 0$), and let $S_2 = \{\beta \in B_\alpha; \alpha \in S_1'\}$. Finally let $S_1 = S_1' \cup S_E$ and let $S_{\text{aug}}^{\text{special}} = S_1 \cup S_2 \cup S_3$.

Applying the unified theory reviewed above and following procedures developed by Fletcher and Roulier [11] and Schmidt [14] in their treatment of monotone approximation, we obtain the following theory extending that given in [11] and [14] to our general setting.

DEFINITION 8. Let f and p_f be given and let $S_{\text{aug}}^{\text{special}} = S_1 \cup S_2 \cup S_3$ be formed from the maximal augmented extremal set for f and p_f. Then define

$$\|g\|' = \max(|g(x)|, |\alpha(g)|), \qquad e_x \in S_3, \alpha \in S_1,$$

$$\|g\|^* = \max(|g(x)|, |\alpha(g)|, |\beta(g)|), \qquad e_x \in S_3, \alpha \in S_1, \beta \in S_2.$$

DEFINITION 9. *V is special generalized Haar with respect to f and p_f* provided that if $S_{\text{aug}}^{\text{special}}$ has order t, then $S_{\text{aug}}^{\text{special}}$ contains $\min(t, n)$ elements which are linearly independent in V^*. *V is special generalized Haar* if V is special generalized Haar for all admissible pairs f and p_f.

The following is a useful and easily applied

Criterion for special generalized Haar. If for every $I_1 \subset A$ there exists $I_2 \subset \bigcup_{\alpha \in I_1} B_\alpha$ such that $V^\downarrow = \{p \in V;\ \gamma(p) = 0\ \forall \gamma \in I_1 \cup I_2\}$ is an ordinary Haar space on E, then V is special generalized Haar. $(\downarrow)$

To check the validity of the above criterion consider $S_{\text{aug}}^{\text{special}} = S_1 \cup S_3 \cup S_2$ and reduce S_2 to S_2' if necessary so that $V^\downarrow = \{p \in V; \gamma(p) = 0\ \forall \gamma \in S_1 \cup S_2'\}$ is an ordinary Haar space on E. But formula (1) provides a dependency on $V^\downarrow$ among the elements of S_3. Hence the order of $S_{\text{aug}}^{\text{special}}$ is $\geqslant n+1$ and $S_{\text{aug}}^{\text{special}}$ contains n independent elements. Thus the condition for V to be special generalized Haar is fulfilled.

Note 6. If V is special generalized Haar, then V is generalized Haar, for then analogously as in Note 3 by use of formula (1), we see that $S_{\text{aug}}^{\text{special}}$ (and hence $S_{\text{aug}}^{\max}$) has order $t \geqslant n+$. Furthermore in all the aforementioned examples, V is easily seen to be special generalized Haar.

EXAMPLE. To see that RDA (restricted derivatives approximation) is special generalized Haar is now even easier than the proof that RDA is generalized Haar given in [1] (which did not make use of formula (1)). Just apply the criterion $(\downarrow)$. For any $I_1 = \{e_x^k\} \subset A$, pick $I_2 \subset \bigcup_{\alpha \in I_1} B_\alpha = \{e_x^{k+1}\}$

such that $I_1 \cup I_2$ form Hermite and supported even block Birkhoff data in an incidence matrix (just as in [1, Theorem 4]). Then $V^{\downarrow}$ is an ordinary Haar space by the well-known Ferguson–Atkinson–Sharma result (see, e.g. [1, Theorem D]).

LEMMA 2. *If V is special generalized Haar, then $\| \|^*$ is a norm on V.*

Proof. Since all the functionals in $S_{\text{aug}}^{\max}$ are linear, $\| \|^*$ is a seminorm. But further, by Definition 9 and Note 6, $S_{\text{aug}}^{\text{special}}$ contains n linearly independent elements of V^*. Hence for $p \in V$, $\|p\|^* = 0$ implies $p \equiv 0$, and $\| \|^*$ is in fact a norm. ∎

THEOREM 5. *Let V be special generalized Haar. Then p_f is strongly unique of order 1 with respect to all p in V_0 satisfying $\alpha(p) = \alpha(p_f)$ for all $\alpha \in S_1$.*

Proof. Since $\| \|^*$ and $\| \|$ are equivalent norms on V, there is a constant $\rho_1 > 0$ such that $\|q\|^* \geqslant \rho_1 \|q\|$ for all $q \in V$. Let $p \in V_0$ with $\alpha(p) = \alpha(p_f)$ for all $\alpha \in S_1$. Then also $\beta(p) = \beta(p_f)$ for all $\beta \in S_2$. Hence $\|p - p_f\|' = \|p - p_f\|^*$. The proof will be complete therefore if we can show

$$\|f - p\| \geqslant \|f - p_f\| + \tau \|p - p_f\|' \tag{3}$$

for some $\tau > 0$, for then $\|f - p\| \geqslant \|f - p_f\| + \tau\rho_1 \|p - p_f\|$ and we can take $\gamma = \tau\rho_1$ and $r = 1$ in (2). To show (3) let e_{x_0} be an arbitrary element in $S_3^{\max}$ and let $\sigma = \sigma(e_{x_0})$. Then $\sigma(f - p)(x_0) = \sigma(f - p_f)(x_0) + \sigma(p_f - p)(x_0) = \|f - p_f\| + \sigma(p_f - p)(x_0)$ and so $\|f - p\| \geqslant \|f - p_f\| + (p_f - p)(x_0)$. Assume for the remainder of the proof that $p \not\equiv p_f$. Then for some $e_{x_0} \in S_3$, $\sigma(p_f - p)(x_0) > 0$. Suppose not; i.e., $\sigma(e_{x_0})(p_f - p)(x_0) \leqslant 0$ $\forall e_{x_0} \in S_3$. Then, from formula (1) of Theorem 4, since $\alpha(p_f - p) = 0$ for all $\alpha \in S_1$ by assumption, we must have $(p_f - p)(x_0) = 0$ $\forall e_{x_0} \in S_3$. But also $\beta(p) = \beta(p_f)$ for all $\beta \in S_2$ and we have therefore by Lemma 2 that $p \equiv p_f$, a contradiction.

Thus $\sup_{e_{x_0} \in S_3} \sigma(e_{x_0})(p_f - p)(x_0) > 0$. Further, the functional $\Gamma(q) = \sup_{e_{x_0} \in S_3} \sigma(e_{x_0}) q(x_0)$ is continuous on the compact set $\mathscr{K} = \{(p_f - p)/\|p_f - p\|' ; \alpha(p) = \alpha(p_f) \ \forall \alpha \in S_1\}$ and hence achieves its infimum. We infer therefore the existence of a $\tau > 0$ such that $\Gamma(q) \geqslant \tau$, $\forall q \in \mathscr{K}$, and conclude that for any admissible p there is a $e_{x_0} \in S_3$ such that $\sigma(e_{x_0})(p_f - p)(x_0) \geqslant \tau \|p_f - p\|'$ and (3) is established. ∎

DEFINITION 10. We say p_f is *semi-strongly* unique of order r $(0 < r \leqslant 1)$ if (2) is valid with $\|p - p_f\|$ replaced by $\|p - p_f\|'$.

THEOREM 6. *Let V be special generalized Haar. Then p_f is semi-strongly unique of order 1 (with respect to all p in V_0).*

Proof. We must show

$$\|f-p\| \geqslant \|f-p_f\| + \gamma \|p-p_f\|' \qquad \text{for some} \quad \gamma > 0. \tag{4}$$

Just as in the proof of Theorem 5, for e_{x_0} an arbitrary member of S_3 and $\sigma = \sigma(e_{x_0})$, $\|f-p\| \geqslant \|f-p_f\| + \sigma(p_f - p)(x_0)$. We claim now that if $\mathscr{K} = \{q = (p_f - p)/\|p_f - p\|'; \|p_f - p\|' \neq 0\}$, then $\inf_{q \in \mathscr{K}} \max_{x_0 \in S_3} \sigma(e_{x_0})\, q(x_0) = \gamma > 0$. Assume not; then there are $q_m \in \mathscr{K}$ such that $\lim_{m \to \infty} \max_{x_0 \in S_3} \sigma(e_{x_0})\, q_m(x_0) \leqslant 0$. But also since $p_m \in V_0$ we have $\sigma(\alpha)\, \alpha(q_m) \leqslant 0$ $\forall \alpha \in S_1$, and we conclude therefore from formula (1) of Theorem 4 that $\lim_{m \to \infty} q_m(x_0) = 0$ $\forall e_{x_0} \in S_3$ and $\lim_{m \to \infty} \alpha(q_m) = 0$ $\forall \alpha \in S_1$. Hence $\lim_{m \to \infty} \|q_m\|' = 0$ while $\forall m$ $\|q_m\|' = 1$, a contradiction. We conclude that for any amissible p there is a $e_{x_0} \in S_3$ such that $\sigma(p_f - p)(x_0) \geqslant \gamma \|p_f - p\|'$ and (4) is established. ∎

The proof of the following lemma proceeds exactly as in the classical case ([7, p. 82] or [11]) and is included for completeness.

LEMMA 3. *Let V be special generalized Haar. Then there exists a positive number K such that for all admissible $g \in C(E)$*

$$\|p_f - p_g\|' \leqslant K \|f - g\|. \tag{5}$$

In fact we may take $K = 2/\gamma$ where γ is from (4).

Proof. By (4) $\|p_f - p\|' \leqslant \gamma^{-1}(\|f-p\| - \|f-p_f\|)$ for any $p \in V_0$. Thus $\|p_f - p_g\| \leqslant \gamma^{-1}(\|f-p_g\| - \|f-p_f\|) \leqslant \gamma^{-1}(\|f-g\| + \|g-p_g\| - \|f-p_f\|) \leqslant \gamma^{-1}(\|f-g\| + \|g-p_f\| - \|f-p_f\|) \leqslant \gamma^{-1}(\|f-g\| + \|g-f\| + \|f-p_f\| - \|f-p_f\|) = (2/\gamma)\|f-g\|$. ∎

We can now establish the continuity of the best approximation operator $B(f) = p_f$.

THEOREM 7. *Let V be special generalized Haar. Then the best approximation operator B is continuous on the admissible functions in $C(E)$.*

Proof. We parrot the proof of Theorem (5.2) in [11]. It suffices to show that if $\{g_m\}_{m=1}^{\infty}$ is a sequence of elements of $C(E)$ satisfying $\lim_{m \to \infty} g_m = f$ uniformly on E, then $\lim_{m \to \infty} B(g_m) = B(f)$. First, by (5),

$$\lim_{m \to \infty} \|B(f) - B(g_m)\|' = 0. \tag{6}$$

Further, $\|B(g_m)\| \leqslant \|B(g_m) - g_m\| + \|g_m\| \leqslant \|p_0 - g_m\| + \|g_m\| \leqslant 1 + 2\|f\| + \|p_0\|$ for m sufficiently large, where $p_0 \in V_0$ is fixed. Thus $\{B(g_m)\}_{m=1}^{\infty}$ is bounded. Now assume that $\lim B(g_m) \neq B(f)$. Then there is an $\varepsilon > 0$ and a subsequence $\{B(g_{m_k})\}_{k=1}^{\infty}$ such that $\|B(g_{m_k}) - B(f)\| \geqslant \varepsilon$,

$k = 1, 2, \ldots$. Furthermore $\{B(g_{m_k})\}_{k=1}^{\infty}$ is bounded. Hence, this sequence has a subsequence which converges, and assume without loss that the sequence itself converges to $q \in V_0$.

We will now show that $q = B(f)$ and thus reach a contradiction to the above assumption. Define $p_k = B(g_{m_k})$ and $p_f = B(f)$. From (6) we have $\lim_{k \to \infty} p_k(x_0) = p_f(x_0)$ $\forall e_{x_0} \in S_3$ and $\lim \alpha(p_k) = \alpha(p_f)$ $\forall \alpha \in S_1$. On the other hand, since $\lim_{k \to \infty} p_k = q$ we have $q(x_0) = p_f(x_0)$ $\forall e_{x_0} \in S_3$ and $\alpha(q) = \alpha(p_f)$ $\forall \alpha \in S_1$. Moreover we also have $\beta(q) = \beta(p_f)$ for all $\beta \in S_2$. Hence by Lemma 2, $q \equiv p_f = B(f)$, a contradiction. ∎

Since if V is Haar (on Ω) then $\|\ \|'$ is equivalent to $\|\ \|$ on V, we obtain the following corollary to Theorem 6 and formula (5).

COROLLARY 1. *Let V be Haar (on Ω). Then p_f is strongly unique (of order 1) with respect to all p in V_0. Moreover, B is Lipschitz-continuous.*

The following definition is introduced in [6].

DEFINITION 11. *p_f is strongly unique with respect to the rate (function) u* ($u \in C[0, \infty)$, u is increasing and $u(0) = 0$) if for each $N > 0$ there exists a constant $\gamma > 0$ such that

$$\|f - p\| \geqslant \|f - p_f\| + \gamma u(\|p - p_f\|) \tag{7}$$

for all $p \in V_0$ satisfying $\|p\| \leqslant N$. We will say that the *rate* (of strong unicity) *is at best u* if (7) cannot be satisfied by any u_1, where $u(t) = o(u_1(t))$, $t \to 0$.

In case $u(t) = t^{1/r}$ for some constant r $(0 < r \leqslant 1)$ then we will according to convention, also say *p_f is strongly unique of order r* (as in Definition 7).

EXAMPLE A ([4]). $E = [a, b]$, $V_0 = \{p \in \pi_{n-1};\ p^{(1)}(x) \geqslant 0,\ a \leqslant x \leqslant b$ and $p^{(2)}(x_0) = \cdots = p^{(2m-1)}(x_0) = 0$ for $x_0 \in (a, b)$ fixed and $n \geqslant 2m + 1\}$. Then (7) holds, where $u(t) = t^{2m}$ and u is best possible.

EXAMPLE B ([6]). $E = [-\alpha, \alpha]$, $V_0 = V \cap \{p; p'(x) \geqslant 0, -\alpha \leqslant x \leqslant \alpha\}$, where $V = [1, x, h'(x), h(x)]$.

(i) If $h(x) = (\operatorname{sgn} x)\,|x|^{2+s}$, $s > 0$, then (7) holds where $u(t) = t^{s+1}$ and u is best possible.

(ii) If $h(x) = xe^{-x^{-2}}$ and α is sufficiently small, then (7) holds where $u(t) = e^{-c_1 x^{-2/3}}$ and the best possible rate function u satisfies $e^{-c_1 x^{-2/3}} \leqslant u(x) \leqslant x^{-2/3} e^{-c_2 x^{-2/3}}$, for some constants $0 < c_2 < c_1$.

THEOREM 8. *Suppose p_f is strongly unique with rate u. Then for each $K > 0$ there is a constant $\gamma > 0$ such that*

$$\|p_g - p_f\| \leqslant u^{-1} \left(\frac{2\|g-f\|}{\gamma}\right) \tag{8}$$

for all $g \in C(E)$ with $\|g\| \leqslant K$. (That is, in particular, the best approximation oprator $B(f) = p_f$ is continuous.)

Proof. If $\|g\| \leqslant K$ then $\|p_g\| \leqslant MK$ for some M independent of g and according to Definition 11 there exists $\gamma > 0$ such that (7) holds where $N = MK$. Hence $\|f - p_g\| \geqslant \|f - p_f\| + \gamma u(\|p_g - p_f\|)$; i.e.,

$$\begin{aligned}
\|p_g - p_f\| &\leqslant u^{-1} \left(\frac{\|f - p_g\| - \|f - p_f\|}{\gamma}\right) \\
&\leqslant u^{-1} \left(\frac{\|f - g\| + \|g - p_g\| - \|f - p_f\|}{\gamma}\right) \\
&\leqslant u^{-1} \left(\frac{\|f - g\| + \|g - p_f\| - \|f - p_f\|}{\gamma}\right) \leqslant u^{-1} \left(\frac{2\|f - g\|}{\gamma}\right). \quad \blacksquare
\end{aligned}$$

COROLLARY 2. *Suppose p_f is strongly unique with rate u, where u is superhomogeneous of degree ρ ($u(ct) \geqslant c^\rho u(t), c > 0$). Then for each $K > 0$ there is a constant $\lambda > 0$ such that*

$$\|B(g) - B(f)\| \leqslant \lambda \|g - f\|^{1/\rho} \tag{8$'$}$$

for all $g \in C(E)$ with $\|g\| \leqslant K$. That is, the best approximation operator is locally Lipschitz-continuous of order $1/\rho$.

Proof. It follows that u^{-1} is subhomogeneous of degree $1/\rho$. Hence we can continue the last line in the proof of Theorem 8: $u^{-1}(2\|f - g\|/\gamma) \leqslant \gamma^{-1/\rho} u^{-1}(1)\, 2^{1/\rho} \|f - g\|^{1/\rho}$ and so (8$'$) holds where $\lambda = \gamma^{-1/\rho} u^{-1}(1)\, 2^{1/\rho}$. $\blacksquare$

EXAMPLES. In Examples A and B(i) above, u is in fact homogeneous of degree $\rho = 2m$ and $\rho = s + 1$, respectively.

THEOREM 9. *Let V be special generalized Haar. Suppose there exists $u_0 \in C[0, \infty)$, with u_0 increasing and $u_0(0) = 0$, and a constant $\tau > 0$ such that*

$$|\alpha(p - p_f)| \geqslant \tau u_0(|\beta(p - p_f)|) \qquad \forall \beta \in B_\alpha, \quad \forall \alpha \in S_1 \tag{9}$$

whenever $p \in V_0$ and $\|p\| \leqslant N$. Then p_f is strongly unique with rate $u(t) = \min(t, u_0(\kappa t))$ for some constant $\kappa > 0$.

Proof. From Definition 8, $\|p-p_f\|' \geqslant \min(\|p-p_f\|^*, \tau u_0(\|p-p_f\|^*))$. But by Lemma 2, $\|p-p_f\|^* \geqslant \kappa \|p-p_f\|$ for some $\kappa > 0$. We thus have that $\|p-p_f\|' \geqslant \tau \min(\kappa/\tau \|p-p_f\|, \quad u_0(\kappa \|p-p_f\|)) \geqslant \tau \min(\|p-p_f\|, u_0(\kappa\|p-p_f\|)$ for $\tau \leqslant \kappa$ and the conclusion follows from formula (4) of Theorem 6. ▮

COROLLARY 3. *Suppose that, in addition to the hypotheses of Theorem* 9, $u_0'(0)$ *exists and* u_0 *is convex. Then* p_f *is strongly unique with rate* $u(t) = u_0(\kappa t)$ *for some constant* $\kappa > 0$.

Proof. Since $u_0'(0)$ exists and u_0 is convex, it is clear that we may choose τ and κ in the proof of Theorem 9 so small that $u_0(\kappa t) < t$ in the interval $[0, N + \|p_f\|]$. ▮

Examples of the application of Theorem 9 (and Corollary 3) are the cases of Theorems 10 and 11 below, where $u_0(t) = t^2$ and $u_0(t) = t^{2+2r}$, respectively. A wide range of additional examples is provided by [6], where the space $V = [1, x, h'(x), h(x)]$, $h \in C^2$, is a Haar space in some neighborhood $(-\alpha, \alpha)$ of the origin and $V_0 = V \cap [p; p'(x) \geqslant 0]$. Several additional assumptions are made on $h(x)$ including $h'(x)/h''(x)$ is asymptotic (as $x \to 0^+$) to $\varphi(x)$, where $\varphi \in C'[0, \infty)$, $\varphi(0) = 0$, $\varphi'(x) > 0$ for $x > 0$. It is easy to see ([6]) that V is special generalized Haar (on Ω). Then, if ψ is any positive continuous function asymptotic to $(h''/h''' - h'/h'')/\varphi$, it is shown that formula (9) (and Corollary 3) holds with $u_0(t) = t\psi(\varphi^{-1}(ct))\, h''(\varphi^{-1}(ct))$ for some constant $c > 0$. (It is shown further, by expanding on the technique developed in [4], that the rate $u_0(t)$ is best possible.) For certain illustrative choices of h see Example B above.

DEFINITION 12. *Generalized restricted derivatives approximation* (*GRDA*) extends ordinary RDA (see, e.g. [1] or [13]) by extending the restraining functionals from $\alpha = e_x^i$ to $\alpha = L_x^i = \sum_{j=0}^{n-1} a_{ij}(x)\, e_x^j$, where the a_{ij} are continuously differentiable on $E = [a, b]$. (Note that at each x, L_x^i represents an arbitrary linear functional on $V = \pi_{n-1}$ the space of $(n-1)$-st degree polynomials on E.) Recall that the restraints are all inequality restraints for RDA as they are therefore for GRDA (i.e., $V_0 = \{p \in \pi_{n-1}; \forall x \in [a, b], l_i(x) \leqslant L_x^i(p) \leqslant u_i(x), l_i(x) < u_i(x), i = 0,..., m\}$). Note that for $x \in (a, b)$, $B_\alpha = \{\beta_\alpha\}$, where $\beta_\alpha = L_x^{i+1} = e_x^1 \circ L_x^i = \sum_{j=0}^{n-1} a_{ij}'(x)\, e_x^j + \sum_{j=0}^{n-1} a_{ij}(x)\, e_x^{j+1}$ and $m(\beta_\alpha)$ (or $n(\beta_\alpha)) = l_i'(x)$ (or $u_i'(x)$). Finally we assume that V is special generalized Haar, as is the case of RDA.

Thus we have all the previous theory at our disposal. Finally, we show that the result of [14] extends to GRDA, provided a_{ij}, l_i, $u_i \in C^2(E)$, $0 \leqslant i \leqslant m$, $0 \leqslant j \leqslant n-1$, whereupon we write $GRDA \in C^2$.

THEOREM 10. *In the case of GRDA $\in C^2$, for each $K > 0$, p_f is strongly unique of order $\alpha = \frac{1}{2}$ with respect to all p in V_0 such that $\|p\| \leqslant K$. Furthermore, $\alpha = \frac{1}{2}$ is in general best possible.*

Proof. We will show (9) holds where $u_0(t) = ct^2$ for some positive constant c and the conclusion will then follow from Corollary 3. We claim that for each of the finite number of pairs $\alpha = L^i_{x_0} \in S_1$ and $\beta_\alpha = L^{i+1}_{x_0} \in S_2$ there exists $c > 0$ for which $|L^i_{x_0}(p_f - p)| \geqslant c\,|L^{i+1}_{x_0}(p_f - p)|^2$ $\forall p \in V_0$ satisfying $\|p\| \leqslant N$. If this is not the case, suppose without loss that α is a lower extremal and then corresponding to each integer $\nu > 0$ there exists $q_\nu \in V_0$ with $\|q_\nu\| \leqslant N$ for which $|L^i_{x_0}(q_\nu - p_f)| < (1/\nu)\,|L^{i+1}_{x_0}(q_\nu - p_f)|^2$, where $L^i_{x_0}(p_f) = l_i(x_0)$ and $L^{i+1}_{x_0}(p_f) = l'_i(x_0)$. Now we may assume that q_ν converges uniformly to $q \in V_0$. Clearly $L^i_{x_0}(q - p_f) = 0$. We can write $L^i_x(q_\nu) - l_i(x) = L^i_{x_0}(q_\nu) - l_i(x_0) + [L^{i+1}_{x_0}(q_\nu) - l'_i(x_0)](x - x_0) + s_\nu(x)(x - x_0)^2 = \beta_\nu + \alpha_\nu(x - x_0) + s_\nu(x)(x - x_0)^2$, where $\beta_\nu \to 0$, $\alpha_\nu \to 0$ ($L^{i+1}_{x_0}(q) = l'_i(x_0)$ since $q \in V_0$ and $L^i_{x_0}(q) = l_i(x_0)$), $|s_\nu(x)| \leqslant N_1$ for all $x \in [a, b]$ and some N_1 independent of ν (GRDA $\in C^2$ implies $[L^i_x(q_\nu)]''$ converges uniformly to $[L^i_x(q)]''$ and $s_\nu(x) = (1/2)(d^2/dy^2)[L^i_y(q_\nu) - l_i(y)]|_{y=\xi_x}$, where ξ_x is a point between x_0 and x), and $L^i_x(q_\nu) \geqslant l_i(x)$ $\forall x \in [a, b]$. Thus $0 \leqslant \beta_\nu + \alpha_\nu(x - x_0) + N_1(x - x_0)^2$ for $x \in [a, b]$. For ν sufficiently large (so that $x \in (a, b)$), set $(x - x_0) = -\alpha_\nu/2N_1$. This gives $\alpha_\nu^2 \leqslant 4N_1\beta_\nu$; i.e., there exists a constant $K_1 = 4N_1$ independent of ν (sufficiently large) such that $|L^{i+1}_{x_0}(q_\nu - p_f)|^2 \leqslant K_1\,|L^i_{x_0}(q_\nu - p_f)|$, which is our desired contradiction.

As in [14], the "best possible" statement results from the elegant counterexample of Fletcher and Roulier [10], since ordinary monotone approximation is a special case of GRDA $\in C^2$.

DEFINITION 13. s-Augmented GRDA will denote GRDA together with even blocks (of length $\leqslant 2s$) of isolated "interpolatory" derivative side conditions given (i.e., $L^{i+k}_{x_0} = e^k \circ L^i_{x_0}$, $k = 2, 3, \ldots, 2s, 2s+1$ have specified values) and V is assumed to be special generalized Haar (as is easily checked to be the case (modify Example A appropriately) in the case of s-augmented RDA).

THEOREM 11. *In the case of s-augmented GRDA $\in C^2$, for each $K > 0$, p_f is strongly unique of order $r = 1/(2 + 2s)$ with respect to all p in V_0 such that $\|p\| \leqslant K$. Furthermore, $r = 1/(2 + 2s)$ is best possible.*

Proof. The proof follows by combining the techniques and notation of Theorem 10 and [4]. ∎

The following corollary states the fact that the continuity of Theorem 7 is a local Lipschitz continuity of order $1/(2s + 2)$, and follows immediately from Theorem 11 and Corollary 2 (formula (8′)).

COROLLARY 4. *In the case of s-augmented GRDA* $\in C^2$ *for each* $K > 0$ *there is a constant* $\lambda > 0$ *such that*

$$\|B(g) - B(f)\| \leqslant \lambda \|g - f\|^{1/(2s+2)}$$

for all $f, g \in C(E)$ *with* $\|g\| \leqslant K$.

Although the order $1/(2s+2)$ is best possible for strong uniqueness (Theorem 11), it is not known whether the order $1/(2s+2)$ is best possible for the Lipschitz condition above.

THEOREM 12. *If V is nearly Haar then* p_f *is strongly unique (of order* 1) *except when* $[S_1^{\max} \cup S_3^{\max}]$ *forms a closed nowhere dense subset of* Ω^n.

Note 7. If V is nearly Haar, then examples like the Fletcher–Roulier example [11] and that of [4] (used for the "best possible" part of Theorem 11) must arise from the relatively rare situations where no n elements of $S_1^{\max}$ and $S_3^{\max}$ are independent in V^*.

Note 8. In this paper the error between f and p is measured by $\|f - p\| = \sup_{x \in E} e(f(x), p(x))$, where $e(f, p) = |f - p|$. The results are valid, however, if we replace $e(f, p)$ by any error function $e(f, p)$ co-monotone with $|f - p|$ as a function of p [12]. In particular, the theory is valid with $e(f, p) = |1/f - 1/p|$, $f > 0$, $p > 0$, i.e., for uniform reciprocal approximation (see [5]).

Note 9. For strong uniqueness results in L^p-spaces, $1 \leqslant p < \infty$, see [10]. Other pertinent references are [8, 9 (*Added in proof.* See also [15])].

ACKNOWLEDGMENTS

The authors would like to thank the referee for his very careful and helpful reading of the manuscript.

Note added in proof. See also [15].

REFERENCES

1. B. L. CHALMERS, A unified approach to uniform real approximation by polynomials with linear restrictions, *Trans. Amer. Math. Soc.* **166** (1972), 309–316.
2. B. L. CHALMERS, The Remez exchange algorithm for approximation with linear restrictions, *Trans. Amer. Math. Soc.* **223** (1976), 103–131.
3. B. L. CHALMERS AND G. D. TAYLOR, Uniform approximation with constraints, *Jber. Dt. Math-Verein.* **81** (1979), 49–86.

4. B. L. CHALMERS AND G. D. TAYLOR, On the existence of strong unicity of arbitrarily small order, *in* "Approximation III" (E. Cheney, Ed.), pp. 293–298, Proceedings, Conference, Austin, Texas, January, 1980.
5. B. L. CHALMERS, E. H. KAUFMAN, JR., D. J. LEEMING, AND G. D. TAYLOR, Uniform reciprocal approximation subject to linear constraints, preprint.
6. B. L. CHALMERS, F. T. METCALF, AND G. D. TAYLOR, Strong unicity of arbitrary rate, *J. Approx. Theory*, in press.
7. E. W. CHENEY, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
8. L. CROMME, Strong uniqueness—a far reaching criterion for the convergence analysis of iterative procedures, *Numer. Math.* **29** (1978), 179–193.
9. C. B. DUNHAM, The usefulness of strong uniqueness in Chebyshev approximation, preprint.
10. J. ANGELOS AND A. EGGER, Strong uniqueness in L^p, preprint.
11. Y. FLETCHER AND J. A. ROULIER, A counterexample to strong unicity in monotone approximation, *J. Approx. Theory* **27** (1979), 19–33.
12. M. A. GRIESEL, "Uniform Approximation With Respect to a General Error Function," Ph. D. thesis, University of California, Riverside, 1971.
13. J. A. ROULIER AND G. D. TAYLOR, Approximation by polynomials with restricted ranges on their derivatives, *J. Approx. Theory* **5** (1972), 216–227.
14. D. SCHMIDT, Strong unicity and Lipschitz conditions of order 1/2 for monotone approximation, *J. Approx. Theory* **27** (1979), 346–354.
15. A. G. EGGER AND G. D. TAYLOR, Strong uniqueness in convex L^p approximation, preprint.