

ON THE CONTINUITY OF THE NONLINEAR TSCHEBYSCHEFF OPERATOR

R. B. BARRAR AND H. L. LOEB

An existence theorem and a Lipschitz continuity theorem for uniform nonlinear Tschebyscheff approximation are given. These theorems include as special cases known results on generalized rational functions but also yield new results for exponential families.

The classical theorems on the uniqueness and characterization of the best uniform approximations by polynomials have been extended to nonlinear approximating families in the papers of Motzkin [11], Tornheim [16] and Rice [13]. These papers introduce the important ideas of unisolvent and varisolvent families.

Meinardus and Schwedt [10] have stressed the importance of a gradient function in the theory of nonlinear approximation.

In the present paper, we combine both these concepts and are thus enabled to extend the strong unicity theorem and theorems on the continuity of the Tschebyscheff operator to nonlinear approximating families. The strong unicity theorem in the linear case is due to Newman and Shapiro [12]. The continuity theorem for ordinary rational approximation is proved in Maehly and Witzgall [9]. The theorems for generalized rational approximation appear in Cheney and Loeb [3], and Cheney [2].

Our assumptions appear to cover many of the nonlinear approximating families in current use. Some examples are given in the last section of the paper.

2. Notation. Let P be an open subset of real Euclidean M dimensional space E_M . We consider a family V of real valued functions $F(A, x)$ where $A = (a_1, \dots, a_M)$ belongs to P and x belongs to $[0, 1]$. The functions $F(A, x)$ and $\partial F(A, x)/\partial a_i, i = 1 \dots M$ are assumed to be continuous in A and x .

We further assume that the family V satisfies the following conditions:

(A) To each $A \in P$, the functions $\partial F(A, x)/\partial a_i, i = 1 \dots M$ generate a Haar Subspace $W(A)$ of dimension $d(A)$ where $d(A) \geq 1$. For convenience of notation and without loss of generality, we will assume in the statements and proofs of all lemmas and theorems in this paper that $\partial F(A, x)/\partial a_i, i = 1 \dots d(A)$, generate the Haar Subspace.

(B) For each $A \in P, F(A, x) \not\equiv F(A_1, x)$ implies $F(A, x) - F(A_1, x)$ has at most $d(A) - 1$ zeroes. $d(A)$ is sometimes called the degree

of $F(A, x)$. Note that $d(A)$ is the smallest integer for which (A) and (B) hold.

We use the notation that if $f \in C[0, 1]$, then $\|f\| = \max_{0 \leq x \leq 1} |f(x)|$, and if A, A^* belong to P , then $\|A - A^*\| = \max_{1 \leq i \leq M} |a_i - a_i^*|$.

$F(A, x)$ is said to be a best approximation to $f(x) \in C[0, 1]$ if $\|F(A, x) - f(x)\| \leq \|F(A_1, x) - f(x)\|$ for all $A_1 \in P$. We sometimes suppress the x and write $F(A) - f$ for $F(A, x) - f(x)$.

Instead of conditions (A) and (B), it will be clear to the reader that our results hold under the following conditions (in this connection see Rice [15]):

We consider a family V of real valued continuous functions contained in $C[0, 1]$ such that for each $v \in V$, there exists a largest integer $s = d(v)$ with the following properties.

(A') There exists a C^1 mapping $F_v(a_1, \dots, a_s)$ of an open neighborhood of the origin of real Euclidean s dimensional space into V with $F_v(0, \dots, 0) = v$, such that the s functions $\partial F_v / \partial a_i$ $i = 1, \dots, s$ are continuous and form a Haar system of dimension s .

(B') If $v_1 \in V$ intersects v in more than $s - 1$ points then $v_1 \equiv v$.

Although conditions (A) and (B) introduce slight additional difficulties in our proofs because they include redundant parameters, we have nevertheless formulated our results under these conditions since they arise naturally, for example, in the case of ordinary rational function approximation.

DEFINITION. Let N be the maximal value of $d(A)$ for $A \in P$. A function $f \in C[0, 1]$ will be called a normal point in $C[0, 1]$ if it has a best approximation $F(A^*, x)$ which has the property that $d(A^*) = N$. Note if $F(A, x) \in V$ and $d(A) = N$, then $F(A, x)$ itself is a normal point in $C[0, 1]$.

We say A_n is equivalent to A'_n if $F(A_n, x) \equiv F(A'_n, x)$. Furthermore, the sequences $\{A_n\}$ and $\{A'_n\}$ are said to be equivalent if A'_n is equivalent to A_n for each n .

3. Main results.

LEMMA 1. *If $A^* = (a_1^*, \dots, a_M^*)$, let $d(A^*) = q$. Further, let x_1, \dots, x_q be distinct points in $[0, 1]$ for which $F(A^*, x_i) = c_i$ ($i = 1, \dots, q$). Then for sufficiently small $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that the equations*

$$(1) \quad F(A, x_i) = \hat{c}_i \quad i = 1, \dots, q$$

with $|c_i - \hat{c}_i| \leq \delta$ have a unique solution $A = (a_1, \dots, a_M)$ such that

$a_i = a_i^*$ for $M \geq i \geq q + 1$ and

$$(2) \quad \|A - A^*\| \leq \varepsilon .$$

Proof. This is immediate from the implicit function theorem applied to the functions:

$$f_i(a_1, \dots, a_q, c_1, \dots, c_q) \equiv F(a_1, \dots, a_q, a_{q+1}^*, \dots, a_M^*, x_i) - c_i$$

$$i = 1, \dots, q .$$

LEMMA 2. Under the conditions of the above lemma, if $F(A^*, x)$ is a normal point there is only one $F(A, x)$ in V satisfying (1), i.e., if $F(A, x)$ satisfies (1), an A_1 equivalent to A can be found satisfying (2).

Proof. Since for each $A \in P, d(A) \leq N = d(A^*)$ it follows that if both $F(A, x)$ and $F(A_1, x)$ satisfy (1) then $F(A, x) = F(A_1, x)$.

LEMMA 3. If $F(A^*, x)$ is a best approximation to $g(x)$ then $F(A^*, x) - g(x)$ must have at least one zero.

Proof. Assume that the lemma is false, and that

$$F(A^*, x) - g(x) > 0$$

for all $x \in [0, 1]$. It is well known that in a Haar subspace there is always a strictly negative function. Hence there is an $A \in E_M$ such that $(d/dt)F(A^* + tA, x)|_{t=0} < 0$ for all $x \in [0, 1]$. Then by the mean value theorem, for small positive $t, \|F(A^* + tA) - g\| < \|F(A^*) - g\|$. This contradicts the fact that $F(A^*, x)$ is the best approximation to $g(x)$, and the lemma follows.

From Rice's general investigations on varisolvent families [13] it follows that: (see also Rice [14])

THEOREM 1. (1) The function $F(A^*, x)$ is a best approximation to $g(x)$ with respect to V if and only if there is a sequence of $d(A^*) + 1$ points $\{x_i\}$ where $0 \leq x_i < x_{i+1} \leq 1$ such that

$$|F(A^*, x_i) - g(x_i)| = \|F(A^*) - g\|$$

$$F(A^*, x_i) - g(x_i) = -(F(A^*, x_{i+1}) - g(x_{i+1}))$$

(2) Each $f \in C[0, 1]$ has at most one best approximation from V .

Proof. We show how Rice's work applies in the present circumstances. First it follows from Lemma 1 that the set V of functions

is solvent in the sense of Rice. Secondly Rice assumes if $A \neq A^*$ then $F(A, x)$ and $F(A^*, x)$ can intersect at most in $d(A^*) - 1$ points, while in this paper we assume this when $F(A^*, x) \not\equiv F(A, x)$. With this change of definition the reasoning used by Rice still holds. Finally Dunham [4] has recently pointed out that the proof in [13] neglects the possibility that a best approximation has a nonzero constant error curve. However our Lemma 3 rules out this possibility in the family V . Thus the result follows.

THEOREM 2. (Also see Dunham [6].) *Let $F(A^*, x)$ be the best approximation to $g(x)$ and assume $F(A^*, x)$ is normal. Then for any sequence $\{F(A_s, x)\}$ such that $\lim_{s \rightarrow \infty} \|F(A_s) - g\| = \|F(A^*) - g\|$, we can find a sequence $\{A'_p\} \in P$ such that $\lim_{p \rightarrow \infty} \|A^* - A'_p\| = 0$, where the sequence $\{A'_p\}$ is equivalent to a subsequence of the $\{A_s\}$, and the last $M - N$ components of each A'_p agree with the corresponding components of A^* .*

Proof. Without loss of generality we can assume $g(x) \equiv 0$ in the following discussion. Let $0 \leq x_1 < x_2 < \dots < x_{N+1} \leq 1$ be a sequence of $N + 1$ critical points for $F(A^*, x)$, i.e.,

$$(4) \quad \begin{aligned} F(A^*, x_i) &= (-1)^{i+1} F(A^*, x_{i+1}) & i = 1, \dots, N \\ \|F(A^*)\| &= |F(A^*, x_1)|. \end{aligned}$$

For definiteness assume $F(A^*, x_1) = \|F(A^*)\|$.

Now let $\{F(A_p, x)\}$ be a subsequence of $\{F(A_s, x)\}$ that converges at the $N + 1$ points, x_1, \dots, x_{N+1} . We call the limits of these $N + 1$ points $F(x_j), j = 1, \dots, N + 1$. Since $\|F(A_p, x)\| \geq |F(A_p, x_j)|$ it follows that

$$(5) \quad \max_{1 \leq j \leq N+1} |F(x_j)| \leq \|F(A^*, x)\| = |F(A^*, x_k)|.$$

From (4) and (5) it follows that

$$(6) \quad \begin{aligned} (-1)^{j+1} F(A^*, x_j) \geq F(x_j) \geq (-1)^j F(A^*, x_j) \\ j = 1, \dots, N + 1. \end{aligned}$$

We wish to show

$$(7) \quad F(x_j) = F(A^*, x_j) \quad j = 1, \dots, N.$$

It will then follow easily from Lemma 1 and 2 that some sequence $\{A'_p\}$, which is equivalent to a subsequence of $\{A_s\}$, can be found so that $\lim_{p \rightarrow \infty} \|A^* - A'_p\| = 0$ and the last $M - N$ components of each A'_p agree with the corresponding components of A^* .

Let us assume (7) is not satisfied for some x_j . We will show

this leads to a contradiction. For definiteness assume (7) does not hold at x_{N+1} . Then $c = |F(A^*, x_{N+1}) - F(x_{N+1})| > 0$. We restrict ourselves to positive ε less than $c/2$. By Lemma 1 for $\varepsilon > \delta > 0$ and δ sufficiently small there is an A such that

$$(8) \quad \begin{aligned} (a) \quad & F(A, x_j) = F(A^*, x_j) + (-1)^{j+1} \left(\frac{\delta}{2}\right) \quad j = 1, \dots, N. \\ (b) \quad & \|F(A) - F(A^*)\| \leq \frac{c}{2} = \frac{1}{2} |F(A^*, x_{N+1}) - F(x_{N+1})|. \end{aligned}$$

We only consider p so large that

$$(9) \quad |F(A_p, x_j) - F(x_j)| \leq \frac{\delta}{4} \quad j = 1, \dots, N + 1.$$

Then by construction,

$$(10) \quad \text{sign}(F(A, x_j) - F(A_p, x_j)) = -\text{sign}(F(A, x_{j-1}) - F(A_p, x_{j-1})) \quad j = 2, \dots, N + 1.$$

For by (6) (8a) and (9), (10) surely holds for $j = 2, \dots, N$ and from (6) (8b) and (9), it also holds for $j = N + 1$.

Thus from (10) it follows that $F(A, x) - F(A_p, x)$ has at least N -zeros in $[0, 1]$. But since N is the maximal degree it follows that $F(A, x) \equiv F(A_p, x)$. But by construction $F(A, x)$ and $F(A_p, x)$ disagree at x_1, \dots, x_{N+1} . This is the desired contradiction and the result follows.

THEOREM 3. *Let $F(A^*, x)$ be the best approximation to g from V , where $F(A^*, x)$ is normal. Then there is an $\alpha > 0$ such that for each $A \in P$*

$$\|g - F(A)\| \geq \|g - F(A^*)\| + \alpha \|F(A) - F(A^*)\|.$$

Proof. The result is trivial if $g \in V$. Hence we assume $g \notin V$. Now if the conclusion is false one can find a sequence $\{A_n\} \in P$ and a sequence of positive numbers $\{\alpha_n\}$ converging to zero so that $F(A_n) \not\equiv F(A^*)$ and such that

$$(11) \quad \|g - F(A_n)\| = \|g - F(A^*)\| + \alpha_n \|F(A_n) - F(A^*)\|.$$

We claim the sequence $\{\|F(A_n)\|\}$ is bounded. This can be seen by considering the following expression derived from (11):

$$(12) \quad \begin{aligned} \|F(A^*) - F(A_n)\| - \|g - F(A^*)\| &\leq \|g - F(A^*)\| \\ &+ \alpha_n \|F(A^*) - F(A_n)\|. \end{aligned}$$

If one divides both sides of (12) by $\|F(A^*) - F(A_n)\|$ and assumes $\{\|F(A_n)\|\}$ is not bounded, then the assumption that $\alpha_n \rightarrow 0$ is con-

tradicted. Hence by (11) and the boundedness of $\{\|F(A_n)\|\}$

$$(13) \quad \lim_{n \rightarrow \infty} \|g - F(A_n)\| = \|g - F(A^*)\| .$$

Therefore by Theorem 2 there is a sequence $\{B_k\} \in P$ converging to A^* where the sequence is equivalent to a subsequence of $\{A_n\}$ and the last $M - N$ components of each B_k agree with the corresponding components of A^* . Note that for the sequence $\{B_k\}$ (11) remains valid. Let

$$Y = \{x \in [0, 1]: |g(x) - F(A^*, x)| = \|g - F(A^*)\|\}$$

$$\sigma(x) = \text{sign}(g(x) - F(A^*, x)) .$$

Since the $\{B_k\}$ satisfy (11), it follows that for each $x \in Y$,

$$(14) \quad \begin{aligned} \alpha_k \|F(B_k) - F(A^*)\| &= \|g - F(B_k)\| - \|g - F(A^*)\| \\ &\geq \sigma(x)[g(x) - F(B_k, x)] \\ &\quad - \sigma(x)[g(x) - F(A^*, x)] \\ &= \sigma(x)[F(A^*, x) - F(B_k, x)] . \end{aligned}$$

We claim there is a $\gamma > 0$ such that for all k

$$(15) \quad \max_{x \in Y} \sigma(x)[-F(B_k, x) + F(A^*, x)] \geq \gamma \|B_k - A^*\| .$$

If (15) is false there is a sequence of positive numbers $\{\gamma_k\}$ converging to zero and a subsequence of the $\{B_k\}$ which we do not relabel such that

$$(16) \quad \max_{x \in Y} \sigma(x) \frac{-F(B_k, x) + F(A^*, x)}{\|B_k - A^*\|} \leq \gamma_k .$$

By the mean value theorem for large k ,

$$(17) \quad \max_{x \in Y} \sigma(x) \left[\sum_{i=1}^N \frac{\partial F(A_k(x), x)}{\partial a_i} \frac{(-b_{ki} + a_i^*)}{\|B_k - A^*\|} \right] \leq \gamma_k$$

where $B_k \equiv (b_{k1}, \dots, b_{kM})$, $N = d(A^*)$, and $A_k(x) \in P$ is on the line between B_k and A^* . Set $C_k \equiv -B_k + A^*/\|B_k - A^*\|$. Since $\|C_k\| = 1$ we can assume by going to subsequences that $C_k \rightarrow C \equiv (c_1, \dots, c_M)$ where $\|C\| = 1$. Using this subsequence in (17) and taking limits, we find,

$$(18) \quad \max_{x \in Y} \sigma(x) \sum_{i=1}^N c_i \frac{\partial F}{\partial a_i}(A^*, x) \leq 0 .$$

By (18) and Theorem 1, the nonzero function $\sum_{i=1}^N c_i (\partial F(A^*, x)/\partial a_i)$ has at least N zeroes which contradicts the fact that $W(A^*)$ is a Haar subspace. Therefore (15) holds. Combining (14) and (15),

$$(19) \quad \alpha_k \|F(B_k) - F(A^*)\| \geq \gamma \|B_k - A^*\| .$$

Since $B_k \rightarrow A^*$, by the mean value theorem there is a $D > 0$ such that for sufficiently large k ,

$$(20) \quad \|F(B_k) - F(A^*)\| \leq D \|B_k - A^*\| .$$

Hence from (19) and (20), for large k

$$\alpha_k \geq \frac{\gamma}{D} > 0 .$$

This contradicts the fact that $\alpha_k \rightarrow 0$. The proof is thus complete.

THEOREM 4. *If $F(A^*, x)$ is a best approximation to $g(x)$ and $F(A^*, x)$ is normal then,*

(1) *There is a $\gamma > 0$ such that $\|f - g\| < \gamma$ implies f has a best approximation, Tf .*

(2) *Furthermore there is a $\lambda > 0$ such that for all f which have a best approximation Tf*

$$\|F(A^*) - Tf\| \leq \lambda \|g - f\| .$$

Proof. By Theorem 3 for all $A \in P$ there is a $\alpha > 0$ such that,

$$(21) \quad \alpha \|F(A) - F(A^*)\| \leq \|g - F(A)\| - \|g - F(A^*)\| .$$

For each $f(x)$ and each n consider all $F(A, x)$ such that $\|f - F(A)\| \leq 1/n + \inf_{\hat{A} \in P} \|f - F(\hat{A})\|$. Then from (21),

$$(22) \quad \begin{aligned} \alpha \|F(A) - F(A^*)\| &\leq \|g - f\| + \|f - F(A)\| - \|g - F(A^*)\| \\ &\leq \|g - f\| + \|f - F(A^*)\| + 1/n - \|g - F(A^*)\| \\ &\leq \|g - f\| + \|g - f\| + 1/n . \end{aligned}$$

Choose $\varepsilon > 0$ such that $\|A - A^*\| \leq \varepsilon$ implies $A \in P$. By Lemma 2 there is an $\delta > 0$ so that $\|F(A^*) - F(A)\| \leq \delta$ implies the existence of a A' such that $\|A^* - A'\| \leq \varepsilon$ and $F(A, x) \equiv F(A', x)$. Thus the set

$$\{F(A, x): A \in P, \|F(A) - F(A^*)\| \leq \delta\}$$

is compact. Hence from this fact and (22), there is an $\gamma > 0$ such that $\|f - g\| < \gamma$ implies Tf exists. For the analogous argument in the case of unisolvent functions see Tornheim [16], Theorem 7. The second conclusion follows directly from (22).

It should be noted that an immediate consequence of Theorem 4 and the definition of $d(\hat{A})$ is that the normal points in $C[0, 1]$ form an open set.

4. Applications. We now consider several applications of the continuity results. First, we consider the case of generalized rationals. Let P and Q be two finite dimensional subspaces of $C[0, 1]$ with bases (p_1, \dots, p_n) and (q_1, \dots, q_m) respectively. Then

$$V = \left\{ F(A, x): F(A, x) = \frac{\sum_{i=1}^n a_i p_i(x)}{\sum_{j=1}^m a_{j+n} q_j(x)}: \sum_{j=1}^m a_{j+n} q_j(x) > 0 \text{ for } x \in [0, 1] \right\} .$$

$$P = \{A = (a_1, \dots, a_{n+m}) \in E_{n+m}: F(A, x) \in V\} .$$

If we assume for each $F(A, x) \in V$ that

$$P + F(A, x)Q \equiv \{p(x) + F(A, x)q(x): p \in P, q \in Q\}$$

is a Haar subspace, then it is easy to show that V satisfies conditions (A) and (B) and $d(A) = \text{dimension}(P + F(A, x)Q)$. This is the situation in ordinary rational polynomial or trigonometric approximation [3, 8].

The second application occurs in the problem of approximation by exponential families. Specifically consider

$$V = \left\{ F(A, x) \equiv \sum_{i=1}^n a_i e^{a_{n+i} x}: \prod_{i=1}^k a_i \neq 0, a_{k+1} = 0 \quad i = 1, \dots, n - k; a_i \neq a_j, i \neq j, n + 1 \leq i, j \leq 2n \right\} .$$

Note that we rule out the difficult case of coalescing exponents [7]. In [10] it was demonstrated that V satisfies conditions (A) and (B). For $F(A, x) \in V, d(A) = n + k$.

A rather interesting application is a slight modification of a problem posed by Dunham [5]. Let V be a family satisfying conditions (A) and (B), and let $\phi(y)$ be a real valued function whose domain is the real line and whose first derivative is continuous and strictly positive. Then we seek $F(A^*, x)$ in V which minimizes

$$\max_{x \in [0, 0]} |f(x) - \phi(F(A, x))| .$$

It is easy to show that the family

$$V' = \{g(A, x): g(A, x) = \phi(F(A, x)): x \in [0, 1]; F(A, x) \in V\}$$

satisfies conditions (A) and (B), and if $g(A, x) \equiv \phi(F(A, x))$ then the degree of g is equal to $d(A)$.

Finally, consider the following problem. Let V be a family which satisfies conditions (A) and (B). Let $0 \leq x_1 < x_2 < \dots < x_k \leq 1$, where

$k \leq \min_{A \in P} d(A)$ be fixed points. Then we seek among all $F(A, x) \in V$ which interpolate $g(x)$ at x_1, \dots, x_k , the $F(A^*, x)$ which minimizes $\|g - F(A)\|$. This problem is discussed in [1]. It is not too difficult to demonstrate that the continuity results can be extended to this setting.

BIBLIOGRAPHY

1. R. B. Barrar and H. L. Loeb, *Best non-linear uniform approximation with interpolation*, Arc. Rational Mech. Analysis **33** (1969), 231-237.
2. E. W. Cheney, *Approximation by generalized rational function*, General Motors Symposium in Approximation of Functions, Elsevier 1965.
3. E. W. Cheney and H. L. Loeb, *Generalized rational approximation*, SIAM Journal, Series B, Numerical Analysis **1** (1964), 11-25.
4. C. B. Dunham, *Necessity of alternation*, (to appear in Canad. Math. Bull.)
5. ———, *Transformed rational Chebyshev approximation*, Numerical Math. **10** (1967), 147-152.
6. ———, *Continuity of the varisolvent Chebyshev operator*, Bull. Amer. Math. Soc. **74** (1968), 606-608.
7. C. R. Hobby and J. R. Rice, *Approximation from a curve of functions*, Archives Rational Mechanics Analysis **24** (1967), 96-106.
8. H. L. Loeb, *Approximation by generalized rationals*, SIAM Numerical Analysis J. **3** (1966), 34-55.
9. H. Maehly and C. Witzgall, *Tschebyscheff—Approximationen in kleinen Intervallen*, Numerische Mathematik **2** (1960), 142-150, 293-307.
10. G. Meinardus and D. Schwedt, *Nicht-linear Approximationen*, Arch. Rational Mech. Analysis **17** (1964), 297-326.
11. T. S. Motzkin, *Approximation by curves of a unisolvent family*, Bull. Amer. Math. Soc. **55** (1949), 789-793.
12. D. J. Newman and H. S. Shapiro, *Some theorems on Chebyšev approximation*, Duke Math. J. **30** (1963), 673-682.
13. J. R. Rice, *Tchebycheff approximations of functions unisolvent of variable degree*, Trans. Amer. Math. Soc. **99** (1961), 298-302.
14. ———, *On the existence and characterization of best nonlinear Tchebycheff approximations*, Trans. Amer. Math. Soc. **110** (1969), 89-97.
15. ———, *Nonlinear approximation II*, Trans. Amer. Math. Soc. **128** (1967), 437-459.
16. L. Tornheim, *On n -parameter families and associated convex functions*, Trans. Amer. Math. Soc. **69** (1950), 457-467.

Received November 15, 1968. The first author was supported in part by NSF Grant GP-9384. The second author was supported in part by NSF Grant GP-8686.

UNIVERSITY OF OREGON

