TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 178, April 1973

ALTERNATING CHEBYSHEV APPROXIMATION

BY

CHARLES B. DUNHAM

ABSTRACT. An approximating family is called alternating if a best Chebyshev approximation is characterized by its error curve having a certain number of alternations. The convergence properties of such families are studied. A sufficient condition for the limit of best approximation on subsets to converge uniformly to the best approximation is given: it is shown that this is often (but not always) a necessary condition. A sufficient condition for the Chebyshev operator to be continuous is given: it is shown that this is often (but not always) a necessary condition.

Let $[\alpha, \beta]$ be a nondegenerate closed interval and $C[\alpha, \beta]$ be the space of continuous functions on $[\alpha, \beta]$. For $g \in C[\alpha, \beta]$ define

$$\|g\| = \sup\{|g(x)| : \alpha \leq x \leq \beta\}.$$

Let F be an approximating function with parameter and P be its parameter space such that $F(A, \cdot) \in C[\alpha, \beta]$ for all $A \in P$. The Chebyshev problem is: given $f \in C[\alpha, \beta]$ find $A^* \in P$ such that $e(A) = \|f - F(A, \cdot)\|$ attains its infimum $\rho(f) = \inf\{e(A): A \in P\}$. Such a parameter A^* is called best and $F(A^*, \cdot)$ is called a best approximation to f on $[\alpha, \beta]$.

Alternation.

Definition. $g \in C[\alpha, \beta]$ is said to *alternate* n times if there exists an $n + 1$ point set $\{x_0, \cdots, x_n\}$, $\alpha \leq x_0 < \cdots < x_n \leq \beta$ such that, for $i = 0, \cdots, n$,

$$|g(x_i)| = \|g\|, \qquad g(x_i) = (-1)^i g(x_0).$$

The set $\{x_0, \cdots, x_n\}$ is called an *alternant* of g. g is said to alternate exactly n times if g alternates n times and does not alternate $n + 1$ times.

Definition. F is said to have property NS of degree n at A if a necessary and sufficient condition for $F(A, \cdot)$ to be best to $f \in C[\alpha, \beta]$ is that $f - F(A, \cdot)$ alternate n times.

The pair (F, P) is called an *alternating family* if F has property NS of positive degree at all $A \in P$. The problem of characterizing alternating families was solved by Rice in [9, pp. 325–327] for a constant number of alternations and in [12, pp. 15–21] for a variable number of alternations.

Received by the editors October 28, 1971.

AMS (MOS) subject classifications (1970). Primary 41A50.

Key words and phrases. Chebyshev approximation, alternation, subsets, Chebyshev operator.

Definition. F has property Z of degree m at A, if the fact that $F(A, \cdot) - F(B, \cdot)$ has m zeros implies $F(A, \cdot) \equiv F(B, \cdot)$.

Definition. F has property $\mathfrak{A}$ of degree n at A, if for any integer $m < n$, any sequence $\{x_1, \cdots, x_m\}$ with $\alpha = x_0 < x_1 < \cdots < x_{m+1} = \beta$, any sign σ, and any real ϵ with $0 < \epsilon < \min\{x_{j+1} - x_j : j = 0, \cdots, m\}/2$, there exists a $B \in P$, such that $\|F(A, \cdot) - F(B, \cdot)\| < \epsilon$,

$$\begin{aligned} \operatorname{sgn}(F(A, x) - F(B, x)) &= \sigma, \qquad \alpha \leq x \leq x_1 - \epsilon, \\ &= \sigma(-1)^j, \qquad x_j + \epsilon \leq x \leq x_{j+1} - \epsilon, \\ &= \sigma(-1)^m, \qquad x_m + \epsilon \leq x \leq \beta. \end{aligned}$$

In case $m = 0$, the above sign condition reduces to

$$\operatorname{sgn}(F(A, \cdot) - F(B, \cdot)) = \sigma.$$

Definition. F has *degree* n at A if F has property Z of degree n at A and property $\mathfrak{A}$ of degree n at A. We denote the degree of F at A by $d(A)$.

Theorem (Rice). *Let (F, P) be an alternating family. A necessary and sufficient condition that F have property* NS *of degree n at A is that F have degree n at A.*

We assume henceforth that (F, P) is an alternating family.

Uniqueness.

Theorem 1. *Let $F(A, \cdot)$ be a best approximation to f, then $F(A, \cdot)$ is a unique best approximation to f.*

This is a special case of more general theorems proven in [4, end of §2] and [5, Theorem 3].

Convergence properties of alternating families. We note that alternating families have no strong topological structure, since any dense subset of an alternating family is an alternating family. However, if convergence of a sequence does occur, we have results about the manner of convergence and the nature of the limit.

Theorem 2. *If $\{F(A_k, \cdot)\}$ is a sequence of approximants such that the degree of F at A_k is no more than n and $\{F(A_k, \cdot)\}$ converges pointwise to $F(A, \cdot)$ on a set of n points, the degree of F at A is no more than n.*

Proof. Suppose first that pointwise convergence occurs on $n + 1$ points, $x_0 < \cdots < x_n$. If the degree of F at A is greater than n, then by property $\mathfrak{A}$ at A, there exists $B \in P$ such that

$$\operatorname{sgn}(F(A, x_i) - F(B, x_i)) = (-1)^i, \qquad i = 0, \cdots, n.$$

Let

$$\mu = \min\{|F(A, x_i) - F(B, x_i)| : i = 0, \cdots, n\}.$$

As $\{F(A_k, \cdot)\}$ converges uniformly to $F(A, \cdot)$ on this finite point set, there exists K such that

$$|F(A, x_i) - F(A_k, x_i)| < \mu/2, \qquad k > K, \ i = 0, \cdots, n,$$

hence

$$\operatorname{sgn}(F(A_k, x_i) - F(B, x_i)) = (-1)^i, \qquad k > K, \ i = 0, \cdots, n.$$

Hence $F(A_k, \cdot) - F(B, \cdot)$ has n zeros, contradicting F having degree no more than n at A_k. Hence the theorem is proven for the case of pointwise convergence on $n + 1$ points.

Next suppose $F(A_k, \cdot)$ converges pointwise to $F(A, \cdot)$ on $\{z_1, \cdots z_n\}$ and does not converge pointwise to $F(A, \cdot)$ at a point z. There exists a subsequence (assume without loss of generality it is $\{F(A_k, \cdot)\}$) which is bounded away from $F(A, \cdot)$ by $\epsilon > 0$ at z, say $F(A_k, z) - F(A, z) > \epsilon$. Let $\{x_0, \cdots, x_n\}$ be the set $\{z_1, \cdots, z_n\} \cup z$ ordered so that $x_0 < \cdots < x_n$ and let $z = x_j$. If the degree of F at A is more than n then by property $\mathfrak{A}$ there exists $B \in P$ such that $\|F(A, \cdot) - F(B, \cdot)\| < \epsilon/2$ and

$$\begin{aligned} \operatorname{sgn}(F(B, x_i) - F(A, x_i)) &= (-1)^{i+j+1}, \qquad i \neq j, \\ &= 1, \qquad i = j. \end{aligned}$$

Choose k such that

$$|F(A_k, x_i) - F(A, x_i)| < \mu/2, \qquad i = 0, \cdots, n, \ i \neq j.$$

We have

$$\operatorname{sgn}(F(B, x_i) - F(A_k, x_i)) = (-1)^{i+j+1}$$

and $F(B, \cdot) - F(A_k, \cdot)$ has n zeros, contradicting F having property Z of degree n at A_k. The theorem is proven. It is helpful in following the proof to draw a diagram.

Corollary. *The set $\{F(A, \cdot): A \in P, d(A) \leq n\}$ is a closed subset of the approximating family.*

It follows that for any positive integer n, $\{F(A, \cdot): A \in P, d(A) > n\}$ is an open subset of the approximating family (Rice proved this for varisolvent F [10, p. 301]), hence

Corollary. *For any $A \in P$, there exists $\delta(A) > 0$ such that if $B \in P$ and $\|F(A, \cdot) - F(B, \cdot)\| < \delta(A)$, $d(B) \geq d(A)$.*

We now give a generalization of a theorem of Tornheim ([11, pp. 72–73], [15, pp. 460–462]). A special case of this was given but not proven in [3].

Theorem 3. *Let* $\alpha \leq x_1 < \cdots < x_n \leq \beta$ *and let* $\{x_i^k\} \to x_i$, $i = 1, \cdots, n$. *Let* $F(A_k, x_i^k) \to F(A, x_i)$, $i = 1, \cdots, n$. *Let* F *have degree* n *at* A_k, $k = 1, \cdots,$ *and at* A. *Then* $\{F(A_k, \cdot)\}$ *converges uniformly to* $F(A, \cdot)$.

Proof. If the theorem is false there exists $\eta > 0$ and a subsequence such that

$$\|F(A, \cdot) - F(A_{k(j)}, \cdot)\| > \eta.$$

Assume without loss of generality that $k(j) = k$, then there exists a sequence $\{y_k\}$ such that

$$|F(A, y_k) - F(A_k, y_k)| > \eta, \qquad k = 1, \cdots.$$

Assume without loss of generality that

$$F(A_k, y_k) - F(A, y_k) > \eta, \qquad k = 1, \cdots.$$

The sequence $\{y_k\}$ has an accumulation point y. Assume without loss of generality that $\{y_k\} \to y$. The proof depends on where y is in relation to the n points.

Consider the case where y lies between them:

$$x_1 < \cdots < x_m \leq y < x_{m+1} < \cdots < x_n.$$

In case $y = x_m$ we assume that $x_m^k < y^k$ for all k. Define

$$z_k = (x_k + x_{k+1})/2, \qquad k \neq m,\ k = 1, \cdots, n-1.$$

We now construct $F(C, \cdot)$ such that $F(A, \cdot) - F(C, \cdot)$ has exactly $n-2$ zeros, all of them simple and one in each ϵ-neighbourhood of z_k, $k = 1, \cdots, n-1$, $k \neq m$. This gives $F(A, \cdot) - F(C, \cdot)$ the desired behaviour of signs on $\{x_1, \cdots, x_n\}$ and we then forget about $\{z_k\}$. To get $F(C, \cdot)$ we apply property $\mathfrak{A}$ of degree n at A with $\{z_1, \cdots, z_{m-1}, z_{m+1}, \cdots, z_n\}$ as the point set

$$\epsilon = (1/10)\min\{x_{m+1} - y, \min\{x_i - x_j, i > j\}, \eta\},$$

and $\sigma = (-1)^m$ to obtain the existence of $C \in P$ such that $\|F(A, \cdot) - F(C, \cdot)\| < \epsilon$ and

$$\begin{aligned} \operatorname{sgn}[F(A, w) - F(C, w)] &= (-1)^{m+i-1}, \qquad |w - x_i| \leq \epsilon,\ i \leq m, \\ &= (-1)^{m+i}, \qquad |w - x_i| \leq \epsilon,\ i > m. \end{aligned}$$

Let $W = \{w: a \leq w \leq b, \operatorname{dist}(w, \{x_1, \cdots, x_n\}) \leq \epsilon\}$ and let

$$\mu = \inf\{|F(A, w) - F(C, w)| : w \in W\}.$$

As W is a compact set containing no zeros of $F(A, \cdot) - F(C, \cdot)$, $\mu > 0$. Select k such that

$$|F(A_k, x_i^k) - F(A, x_i^k)| < \mu/2, \qquad i = 1, \cdots, n,$$
$$|y_k - y| < \epsilon,$$
$$|x_i^k - x_i| < \epsilon.$$

Then we have

$$\operatorname{sgn}[F(A_k, x_i^k) - F(C, x_i^k)] = (-1)^{m+i-1}, \quad i \leq m,$$
$$= (-1)^{m+i}, \qquad i > m.$$

Since

$$F(A_k, y_k) - F(A, y_k) > \eta > \epsilon, \qquad F(C, y_k) - F(A, y_k) < \epsilon,$$

we have

$$\operatorname{sgn}(F(A_k, y_k) - F(C, y_k)) = +1.$$

It follows that $F(A_k, \cdot) - F(C, \cdot)$ alternates in sign on the ordered set $x_1^k < \cdots < x_m^k < y_k < y_{m+1}^k < \cdots < x_n^k$ and so has at least n zeros. But this contradicts property Z of degree n at A_k, proving the theorem in the case considered. An almost identical argument is used for other cases. For example in the case $y = x_m$ and $y_k < x_m^k$ we need merely reverse all points to use our argument. It is helpful to draw a diagram in following the proof. Such a diagram appears in Rice [11, p. 73].

Degeneracy.

Definition. F is degenerate at A if every neighbourhood of $F(A, \cdot)$ contains an element of (F, P) with higher degree.

If F is of maximum degree at A then F is nondegenerate at A. In most cases of interest all elements not of maximum degree are degenerate: this is true for families $R_m^n[\alpha, \beta]$ of polynomial rational functions or families V_n of exponential sums [8, pp. 312–313]. However, there are cases where the degree is variable and all elements are nondegenerate.

Example. Let $[\alpha, \beta]$ be finite and define $F(A, x) = a_1 + a_2 x$, $P_1 = \{(a_1, 0): a_1 < 0\}$, $P_2 = \{A: F(A, \cdot) > 0\}$, $P = P_1 \cup P_2$. F has degree 1 (2) at all $A \in P_1$ (P_2), and F is nondegenerate at all $A \in P$.

Theorem 4. *Let F have bounded degree, then the set of degenerate approximations is a closed nowhere-dense subset of (F, P).*

Proof. By the second corollary to Theorem 2 the set of nondegenerate approximants is a union of open sets and hence open, making the set of degenerate approximants closed. Suppose that S is an open subset of (F, P) containing

only degenerate approximants. Let $F(A, \cdot)$ be an element of S of maximal degree. By the second corollary there is a neighbourhood of $F(A, \cdot)$ containing only approximants of degree $d(A)$. But this contradicts degeneracy of F at A and the theorem is proven.

The limit of best approximation on subsets.

Definition. We say $\{X_k\} \to [\alpha, \beta]$ if X_k is a closed subset of $[\alpha, \beta]$ and for any $x \in [\alpha, \beta]$, there is a sequence $\{x_k\} \to x$ such that $x_k \in X_k$.

We will study the behaviour of the error norm and of best approximations on X_k for $\{X_k\} \to [\alpha, \beta]$.

Definition. Let Y be a closed subset of $[\alpha, \beta]$ and $g \in C[\alpha, \beta]$; then

$$\|g\|_Y = \sup\{|g(x)| : x \in Y\}.$$

Let us define for fixed $f \in C[\alpha, \beta]$ and a closed subset Y of $[\alpha, \beta]$,

$$\sigma(Y) = \inf\{\|f - F(A, \cdot)\|_Y : A \in P\},$$

$$S(Y) = \{F(A, \cdot) : \|f - F(A, \cdot)\|_Y = \sigma(Y),\ A \in P\}.$$

We will only consider subsets Y such that there is no $A \in P$ with $d(A) \geq \operatorname{card}(Y)$. In this case we have from [7]

Theorem 5. *A necessary and sufficient condition for $F(A, \cdot)$ to be best to f on Y is that $f - F(A, \cdot)$ alternate $d(A)$ times on Y. Best approximations are unique.*

It has been conjectured that $\sigma(X_k) \to \sigma[\alpha, \beta]$ if $\{X_k\} \to [\alpha, \beta]$. This may not be true even if (F, P) is an alternating family of constant degree.

Example. Let $[\alpha, \beta] = [0, 1]$, $F(a, x) = (1 + a)/(1 + x/a)$, $P = \{a : a > 0\}$. For $0 < a < b$ we have $F(a, \cdot) < F(b, \cdot)$ and if $\{a_k\} \to a > 0$, $\|F(a_k, \cdot) - F(a, \cdot)\| \to 0$. Hence F has degree 1 at all $A \in P$. Let $f = 0$, then since $F(a, 0) = 1 + a$, $\sigma[\alpha, \beta] = 1$. Let Y be a closed subset of $[0, 1]$ not containing zero, then $F(1/k, \cdot)$ converges on Y to zero, hence $\sigma(Y) = 0$.

It is proven later that if $S(X_k) \to S[\alpha, \beta]$ then $\sigma(X_k) \to \sigma[\alpha, \beta]$.

Definition. The distance between a subset Y of $[\alpha, \beta]$ and $[\alpha, \beta]$ is

$$\operatorname{dist}(Y, [\alpha, \beta]) = \sup\{\inf\{|x - y| : y \in Y\} : x \in [\alpha, \beta]\}.$$

We have $\{X_k\} \to [\alpha, \beta]$ if and only if $\operatorname{dist}(X_k, [\alpha, \beta]) \to 0$.

Lemma 1. *Let $F(A, \cdot)$ be a best approximation to f on $[\alpha, \beta]$ and $d(A) = n$. Let $x_0, \cdots, x_n$ be an alternant of $f - F(A, \cdot)$. Let $\epsilon > 0$ be given; then there exists δ, $0 < \delta < \epsilon$, such that if $|x_i' - x_i| < \delta$, $i = 0, \cdots, n$, and*

$$\max\{|f(x_i') - F(B, x_i')| : i = 0, \cdots, n\} \leq \|f - F(A, \cdot)\|,$$

then

$$\text{(1)} \qquad \operatorname{sgn}[f(x_0) - F(A, x_0)](-1)^i[F(B, x_i') - F(A, x_i')] \geq -\epsilon, \qquad i = 0, \cdots, n.$$

Proof. $f - F(A, \cdot)$ is continuous on $[\alpha, \beta]$, hence uniformly continuous on $[\alpha, \beta]$. There exists $\delta > 0$ such that if $|x - y| < \delta$,

$$|f(x) - F(A, x) - f(y) + F(A, y)| < \epsilon.$$

Choose Y such that $\operatorname{dist}([\alpha, \beta], Y) < \delta$. Assume without loss of generality that $f(x_0) - F(A, x_0) > 0$; then

$$\begin{aligned} f(x_i) - F(A, x_i) &= (-1)^i \|f - F(A, \cdot)\|, \\ \text{(2)} \qquad (-1)^i(f(x_i') - F(A, x_i')) &\geq \|f - F(A, \cdot)\| - \epsilon, \\ (-1)^i(f(x_i') - F(B, x_i')) &\leq \|f - F(A, \cdot)\|. \end{aligned}$$

The above inequality can be changed to

$$(-1)^i(F(B, x_i') - f(x_i')) \geq -\|f - F(A, \cdot)\|$$

and we add this to (2) to get

$$(-1)^i(F(B, x_i') - F(A, x_i')) \geq -\epsilon, \qquad i = 0, \cdots, n.$$

Lemma 2. *Let F be nondegenerate at A. Then for given $\epsilon > 0$ there exists $\eta(\epsilon)$ such that $\|F(A, \cdot) - F(B, \cdot)\| < \eta(\epsilon)$ if (1) holds and $\eta(\epsilon) \to 0$ as $\epsilon \to 0$.*

Proof. We assume that $f(x_0) - F(A, x_0) > 0$. If the lemma fails there exists

(i) a sequence $\{A_k\}$ of parameters,

(ii) a sequence $\{(x_0^k, \cdots, x_n^k)\}$ of $n + 1$ tuples,

(iii) a sequence $\{y_k\}$ of points, and

(iv) a constant $\mu > 0$,

such that $\{x_i^k\} \to x_i$, $i = 0, \cdots, n$,

$$\text{(3)} \qquad (-1)^i[F(A_k, x_i^k) - F(A, x_i^k)] \geq -1/k, \qquad i = 0, \cdots, n,$$

and $|F(A_k, y_k) - F(A, y_k)| > \mu$. As $[\alpha, \beta]$ is compact, $\{y_k\}$ has an accumulation point y_0 and we can assume without loss of generality that $\{y_k\}$ converges to y_0. We can further select a subsequence such that $F(A_k, y_k) - F(A, y_k)$ has a constant sign. We will assume that this is positive, hence

$$F(A_k, y_k) - F(A, y_k) > \mu.$$

Let $\lambda = \min\{x_{i+1} - x_i : i = 0, \cdots, n - 1\}/3$. Let us define

$$\bar{x}_k = (x_k + x_{k+1})/2, \qquad k = 0, \cdots, n - 1.$$

We first study the case where y_0 is in the open interval (x_0, x_n), $x_j \leq y_0 < x_{j+1}$. There exists $\delta > 0$ such that $y_0 \in (x_j + \delta, x_{j+1} - \delta)$ and we can assume without loss of generality that y_k is in that interval for all k.

We consider the case where j is even. We construct $F(B, \cdot)$ such that $\|F(B, \cdot) - F(A, \cdot)\| < \mu/2$, F is of degree n at B, $\operatorname{sgn}(F(B, y_0) - F(A, y_0)) = 1$, and the difference $F(B, \cdot) - F(A, \cdot)$ has exactly $n - 2$ zeros, all of which are simple and located one in each interval $(\bar{x}_i - \lambda/10, \bar{x}_i + \lambda/10)$, $i = 0, \cdots, n-1$, $i \neq j-1, j$. As x_{j-1}, x_j, x_{j+1}, and y_0 are in $(\bar{x}_{j-2} + \lambda/10, \bar{x}_{j+1} - \lambda/10)$ we have

$$\operatorname{sgn}[F(B, x_i) - F(A, x_i)] = 1, \qquad i = j-1, j, j+1,$$

hence

$$\operatorname{sgn}[F(B, x_i) - F(A, x_i)] = (-1)^{j+i-1}, \qquad i = 0, \cdots, n, \ i \neq j. \tag{4}$$

As $F(A, \cdot) - F(B, \cdot)$ is nonzero outside $\bigcup_{i=0}^{n-1} (\bar{x}_i - \lambda/10, \bar{x}_i + \lambda/10)$, there is a closed interval I_i containing x_i as an interior point such that $F(A, \cdot) - F(B, \cdot)$ does not vanish on I_i, $i = 0, \cdots, n$. Without loss of generality we can assume that

$$\{x_i^k\} \subset I_i, \qquad i = 0, \cdots, n. \tag{5}$$

Let $\epsilon = \inf\{|F(A, x) - F(B, x)| : x \in \bigcup_{i=0}^{n} I_i\}$. Choose k such that $1/k < \epsilon/2$ and $x_j < y_k < x_{j+2}$. By (3) we have

$$(-1)^i [F(A, x_i^k) - F(A_k, x_i^k)] \leq 1/k < \epsilon/2, \qquad i = 0, \cdots, n. \tag{6}$$

By (4), (5) we have

$$\operatorname{sgn}[F(B, x_i^k) - F(A, x_i^k)] = (-1)^{j+i-1}, \qquad i = 0, \cdots, n, \ i \neq j, \tag{7}$$

and by (5) and the definition of ϵ,

$$|F(B, x_i^k) - F(A, x_i^k)| \geq \epsilon. \tag{8}$$

For i odd, $F(B, x_i^k) - F(A, x_i^k) \geq \epsilon$ and $F(A, x_i^k) - F(A_k, x_i^k) \geq -\epsilon/2$, so

$$\operatorname{sgn}(F(B, x_i^k) - F(A_k, x_i^k)) = (-1)^{i+j-1}.$$

Similar arguments show this to be true for i even, $i \neq j$. As $\|F(A, \cdot) - F(B, \cdot)\| < \mu/2$, $F(B, y_k) - F(A_k, y_k) < -\mu/2$. We have $F(B, \cdot) - F(A_k, \cdot)$ alternating in sign on $\{x_0^k, \cdots, x_{j-1}^k, y_k, x_{j+1}^k, \cdots, x_n^k\}$, hence $F(B, \cdot) - F(A_k, \cdot)$ has n zeros. But this contradicts F having degree n at B.

Next consider the case where j is odd and $y_0 \neq x_j$. Construct $F(B, \cdot)$ such that $\|F(B, \cdot) - F(A, \cdot)\| < \mu/2$, F is of degree n at B, $\operatorname{sgn}(F(B, y_0) - F(A, y_0)) = 1$, and the difference $F(B, \cdot) - F(A, \cdot)$ has exactly $n - 2$ zeros, all of which are

simple and located one in each interval $(\bar{x}_i - \lambda/10, \bar{x}_i + \lambda/10)$, $i = 0, \cdots, n-1$, $i \neq j, j+1$. As x_j, x_{j+1}, x_{j+2} and y_0 are in $(\bar{x}_{j-1} + \lambda/10, \bar{x}_{j+2} - \lambda/10)$, we have

$$\operatorname{sgn}[F(B, x_i) - F(A, x_i)] = 1, \qquad i = j, j+1, j+2,$$

hence

$$\operatorname{sgn}[F(B, x_i) - F(A, x_i)] = (-1)^{j+i}, \qquad i = 0, \cdots, n, \ i \neq j+1. \tag{9}$$

As $F(A, \cdot) - F(B, \cdot)$ is nonzero outside $\bigcup_{i=0}^{n-1} (\bar{x}_i - \lambda/10, \bar{x}_i + \lambda/10)$, there is a closed interval I_i containing x_i as an interior point such that $F(A, \cdot) - F(B, \cdot)$ does not vanish on I_i, $i = 0, \cdots, n$. Without loss of generality we can assume that

$$\{x_i^k\} \subset I_i, \qquad i = 0, \cdots, n. \tag{10}$$

Let $\epsilon = \inf\{|F(A, x) - F(B, x)| : x \in \bigcup_{i=0}^{n} I_i\}$. Choose k such that $1/k < \epsilon/2$, $x_j < y_k < x_{j+2}$, and $y_k \notin I_j \cup I_{j+2}$. By (3) we have

$$(-1)^i [F(A, x_i^k) - F(A_k, x_i^k)] \leq 1/k < \epsilon/2, \qquad i = 0, \cdots, n. \tag{11}$$

By (9), (10) we have

$$\operatorname{sgn}[F(B, x_i^k) - F(A, x_i^k)] = (-1)^{j+i}, \qquad i = 0, \cdots, n, \ i \neq j+1, \tag{12}$$

and by (10) and the definition of ϵ,

$$|F(B, x_i^k) - F(A, x_i^k)| \geq \epsilon. \tag{13}$$

For i odd, $F(B, x_i^k) - F(A, x_i^k) \geq \epsilon$ and $F(A, x_i^k) - F(A_k, x_i^k) \geq -\epsilon/2$, so

$$\operatorname{sgn}(F(B, x_i^k) - F(A_k, x_i^k)) = (-1)^{i+j}.$$

Similar arguments show this to be true for i even, $i \neq j+1$. As $\|F(A, \cdot) - F(B, \cdot)\| < \mu/2$, $F(B, y_k) - F(A_k, y_k) < -\mu/2$. We have $F(B, \cdot) - F(A_k, \cdot)$ alternating in sign on $\{x_0^k, \cdots, x_j^k, y_k, x_{j+2}^k, \cdots, x_n^k\}$, hence $F(B, \cdot) - F(A_k, \cdot)$ has n zeros. Next consider the case where j is odd and $y_0 = x_j$. By (2)

$$(-1)[F(A_k, x_j^k) - F(A, x_j^k)] \geq -1/k,$$

hence by (3) $\{x_j^k - y_k\}$ is nonzero for all k sufficiently large. By taking a subsequence if necessary, we can assume that $\{x_j^k - y_k\}$ is of constant sign. If that sign is negative, the previous argument for j odd, $y_0 \neq x_j$ can be applied. If that sign is positive, we note that change of variable $x' = -x$, which involves approximation of $f(-x)$ by $F(A, -x)$ on $[-\beta, -\alpha]$ turns the case into one already handled.

Consider next the case where y_0 is in $[x_n, \beta]$. We can set $j = n-1$ and

y_0 is in $(x_j, \beta]$. We apply the argument developed before for the case j is even (odd), using the same B, and get a difference $F(B, \cdot) - F(A_k, \cdot)$ having n zeros.

Consider finally the case where y_0 is in $[\alpha, x_0]$. We can set $j = 0$ and y_0 is in $[\alpha, x_1]$. We apply the argument developed before for the case j is even, using the same B, and get a difference $F(B, \cdot) - F(A_k, \cdot)$ having n zeros.

Theorem 6. *Let $S[\alpha, \beta]$ be nondegenerate. Let $\{X_k\} \to [\alpha, \beta]$ and $\{S(X_k)\}$ exist. Then $S(X_k)$ converges uniformly to $S[\alpha, \beta]$.*

Proof. Let $F(A, \cdot) = S[\alpha, \beta]$ and $\{x_0, \cdots, x_n\}$ be an alternant of $f - F(A, \cdot)$. Let us choose $\epsilon > 0$; then by Lemma 1 there is δ, $0 < \delta < \epsilon$, such that if $|x_i' - x_i| < \delta$, $i = 0, \cdots, n$, and

$$\max\{|f(x_i') - F(B, x_i')| : i = 0, \cdots, n\} \leq \|f - F(A, \cdot)\|,$$

then (1) holds. Choose X_k such that $\operatorname{dist}(X_k, [\alpha, \beta]) < \delta$ and let $\{x_0', \cdots, x_n'\} \in X_k$, $|x_i' - x_i| < \delta$, $i = 0, \cdots, n$. Then (1) holds with $F(B, \cdot) = S(X_k)$. By Lemma 2, $\|F(A, \cdot) - S(X_k)\| < \eta(\epsilon)$, and $\eta(\epsilon) \to 0$ as $\epsilon \to 0$.

Lemma 3. *Let $\{X_k\} \to [\alpha, \beta]$ and $S(X_k)$ converge uniformly to $S[\alpha, \beta]$; then $\sigma(X_k) \to \sigma[\alpha, \beta]$.*

Proof. Let $S[\alpha, \beta] = F(A, \cdot)$ and $|f(x) - F(A, x)| = e(A)$. Let $\epsilon > 0$ be given. As $f - F(A, \cdot)$ is uniformly continuous, there exists $\delta > 0$ such that if $|x - y| < \delta$ then $|(f(y) - F(A, y)) - (f(x) - F(A, x))| < \epsilon/2$. Choose k such that $\operatorname{dist}(X_k, [\alpha, \beta]) < \delta$, and $\|S[\alpha, \beta] - S(X_k)\| < \epsilon/2$. Let $y \in X_k$, $|x - y| < \delta$ and $S(X_k) = F(A_k, \cdot)$; then

$$\begin{aligned} &|(f(y) - F(A_k, y)) - (f(x) - F(A, x))| \\ &\qquad < |(f(y) - F(A, y)) - (f(x) - F(A, x))| + |F(A, y) - F(A_k, y)| < \epsilon, \end{aligned}$$

hence $\sigma[\alpha, \beta] - \epsilon < \sigma(X_k) \leq \sigma[\alpha, \beta]$, the last inequality following from $X_k \subset [\alpha, \beta]$.

Definition. Let

$$M(A) = \{x : |f(x) - F(A, x)| = e(A)\}.$$

By continuity of $f - F(A, \cdot)$, $M(A)$ is closed and nonempty.

It has been proven by the author [6] that for approximation by members of a class of approximation families which includes the polynomial rational functions $R_m^n[\alpha, \beta]$ and exponential families V_n, degeneracy of $S[\alpha, \beta] = F(A, \cdot)$ and $M(A)$ being nowhere dense implies that there is a sequence of closed subsets $\{X_k\}$ such that $\{X_k\} \to [\alpha, \beta]$ and $\|S(X_k) - S[\alpha, \beta]\| \nrightarrow 0$. This is not always true.

Example. Let $[\alpha, \beta] = [0, 1]$, $F(A, x) = a_1 + a_2 x$, and $P = P_- \cup P_+ \cup (0,0)$,

where $P_- = \{A: F(A, \cdot) < 0\}$, $P_+ = \{A: F(A, \cdot) > 0\}$. F is of degree 1 at (0, 0) and of degree 2 at all $A \in P_- \cup P_+$. F is in fact unisolvent of variable degree. Let $f(x) = x - \frac{1}{2}$; then f alternates once and 0 is therefore best to f. Let a closed subset Y of $[0, 1]$ be given and let $\gamma = \inf\{x: x \in Y\}$, $\delta = \sup\{x: x \in Y\}$. We consider only the case $\gamma < \frac{1}{2} < \delta$. If $\gamma - \frac{1}{2} = \frac{1}{2} - \delta$, 0 is the best approximation on Y. If $\gamma - \frac{1}{2} \neq \frac{1}{2} - \delta$, it is easily seen that there is no best approximation on Y.

The reason why failure of uniform convergence did not occur in this example is that best approximations failed to exist on subsets for which 0 was not best. If best approximations exist on subsets, then degeneracy does imply failure of uniform convergence in the most common case of interest, namely when A is best, $f - F(A, \cdot)$ alternates exactly $d(A)$ times, and $M(A)$ is nowhere dense.

Theorem 7. *Let $f - F(A, \cdot)$ alternate exactly $d(A)$ times, F be degenerate at A, and $M(A)$ be nowhere dense. There is a sequence $\{X_k\}$ of closed subsets such that $\{X_k\} \to [\alpha, \beta]$ and if a best approximation exists on all X_k, $\|S(X_k) - S[\alpha, \beta]\| \nrightarrow 0$.*

Proof. By definition there is a sequence $\{F(A_k, \cdot)\}$ such that $\|F(A, \cdot) - F(A_k, \cdot)\| < 1/k$ and $d(A_k) > d(A)$. There exists K such that, for $k > K$, $e(A) > 2/k$. For such k let $X_k = \{x: |f(x) - F(A_k, x)|\} \leq e(A) - 1/k$. Let x be a point of $[\alpha, \beta]$ which is not in $M(A)$; then $|f(x) - F(A, x)| < e(A) - 2/k$ for all k sufficiently large, hence $|f(x) - F(A_k, x)| < e(A) - 1/k$ and $x \in X_k$ for all k sufficiently large. Let $y \in M(A)$; then $|f(y) - F(A_k, y)| \geq e(A) - 1/k$. Let z be a zero of $f - F(A, \cdot)$; then $|f(z) - F(A_k, z)| \leq 1/k$. It follows that $f - F(A_k, \cdot)$ alternates $d(A)$ times on X_k with amplitude $e(A) - 1/k$. By arguments similar to those used to prove necessity of $d(A)$ alternations, $f - F(A_k, \cdot)$ has exactly $d(A)$ alternations on X_k for all k sufficiently large. But $d(A_k) > d(A)$ so A_k is not best on X_k. Suppose a best approximation $F(B_k, \cdot)$ exists on X_k. As $|f(x) - F(B_k, x)| < |f(x) - F(A_k, x)|$ for x an extremum of $f - F(A_k, \cdot)$ on X_k, $F(B_k, \cdot) - F(A_k, \cdot)$ alternates on the extrema of $f - F(A_k, \cdot)$ on X_k. Hence $F(B_k, \cdot) - F(A_k, \cdot)$ has $d(A)$ zeros and $d(B_k) \geq d(A) + 1$. Hence $f - F(B_k, \cdot)$ alternates at least $d(A) + 1$ times on X_k. It is therefore impossible that $\|F(B_k, \cdot) - F(A, \cdot)\| \to 0$, and the theorem is proven.

The subsets X_k of the theorem were infinite sets. If we let Y_k be a finite subset of X_k containing an alternant of $f - F(A_k, \cdot)$ on X_k, we can use exactly the same arguments.

The most promising approach to the problem of nonconvergence for approximation on subsets appears to be to abstract the properties causing nonconvergence. The author has found the property of irregularity [6] to be useful.

Continuity of the best approximation operator. Let us denote by $T(f)$ the best approximation to f if it exists, thus defining the Chebyshev operator.

Definition. T is *continuous* at $f_0 \in C[\alpha, \beta]$ if for any sequence $\{f_k\} \subset C[\alpha, \beta]$ converging uniformly to f such that $T(f_k)$ exists, $k = 0, 1, \cdots$, $\|T(f_0) - T(f_k)\| \to 0$.

Theorem 8. *Let $f \in (F, P)$ or $T(f)$ be nondegenerate, then T is continuous at f.*

Proof. The case where f is an approximant is well known. If $T(f)$ is nondegenerate, the result is an immediate consequence of the following two lemmas.

Lemma 4. *Let $F(A, \cdot)$ be the best approximation to f and F have degree n at A. Let $x_0, \cdots, x_n$ be an ordered $n+1$ point set on which $f - F(A, \cdot)$ alternates $n+1$ times. If $\|f - g\| < \delta$ and $\|f - F(B, \cdot)\| < \rho(g) + \delta$, then*

$$(-1)^i [F(B, x_i) - F(A, x_i)] \cdot \operatorname{sgn}(f(x_0) - F(A, x_0)) \geq -3\delta, \qquad i = 0, \cdots, n. \tag{14}$$

Proof. We have

$$\|g - F(B, \cdot)\| < \|g - F(A, \cdot)\| + \delta \leq \|g - f\| + \|f - F(A, \cdot)\| + \delta. \tag{15}$$

$$(-1)^i [f(x_i) - F(A, x_i)] \cdot \operatorname{sgn}(f(x_0) - F(A, x_0)) = \|f - F(A, \cdot)\|. \tag{16}$$

$$(-1)^i [g(x_i) - F(B, x_i)] \cdot \operatorname{sgn}(f(x_0) - F(A, x_0)) \leq \|g - F(B, \cdot)\|. \tag{17}$$

By subtracting (16) from (17) and then applying (15) we get

$$\begin{aligned}
&(-1)^i (F(B, x_i) - F(A, x_i)) \cdot \operatorname{sgn}(f(x_0) - F(A, x_0)) \\
&\quad \geq (-1)^i (g(x_i) - f(x_i)) \cdot \operatorname{sgn}(f(x_0) - F(A, x_0)) + \|f - F(A, \cdot)\| + \|g - F(B, \cdot)\| \\
&\quad \geq -|g(x_i) - f(x_i)| - 2\delta \geq -3\delta.
\end{aligned}$$

Lemma 5. *Let F be nondegenerate at A and $d(A) = n$, then for given $\delta > 0$ there exists $\eta(\delta)$ such that $\|F(A, \cdot) - F(B, \cdot)\| < \eta(\delta)$ if (14) holds and $\eta(\delta) \to 0$.*

Lemma 5 is a special case of Lemma 2.

It is known that in the case of approximation by generalized rational functions [2] and exponential sums [13] that T is continuous only in the cases specified by the theorem, which leads to the conjecture that T is discontinuous at all f which have a degenerate best approximation and which are not approximants. This conjecture is however false.

Example. Let $[\alpha, \beta] = [0, 1]$, $F(A, x) = a_1 + a_2 x$, and $P = P_- \cup P_+ \cup (0, 0)$, where $P_- = \{A: F(A, \cdot) < 0\}$, $P_+ = \{A: F(A, \cdot) > 0\}$. F is of degree 1 at $(0, 0]$ and of degree 2 at all $A \in P_- \cup P_+$. F is in fact unisolvent of variable degree. Let $f(x) = x - \frac{1}{2}$ then f alternates once and 0 is therefore best to f. If T is

discontinuous at f there is a sequence $\{f_k\}$ which converges uniformly to f such that $T(f_k)$ exists and $\{\|T(f_k)\|\} \nrightarrow 0$. By taking a subsequence if necessary one can assume that $T(f_k) \neq 0$, hence $T(f_k)$ is of degree 2 and $f_k - T(f_k)$ alternates twice. By a well-known result [16, p. 120, Theorem 8], $\|f_k - T(f_k)\| \to \|f - T(f)\| = \frac{1}{2}$. For k large, $f - T(f_k)$ almost alternates 2 times with amplitude near $\frac{1}{2}$. But f and $T(f_k)$ are first degree polynomials, so this is impossible.

A reason why discontinuity did not occur in this example is that best approximations do not exist to all continuous functions in a neighbourhood of f. If they do exist, then degeneracy does imply discontinuity in the most common case of interest, namely when A is best and $f - F(A, \cdot)$ alternates exactly $d(A)$ times.

Lemma 6. *Let F be degenerate at A and $f - F(A, \cdot)$ alternate exactly $d(A)$ times. Then there is a sequence $\{f_k\} \to f$ such that either $T(f_k)$ does not exist or if $T(f_k) = F(C_k, \cdot)$, $d(C_k) > d(A)$.*

Proof. By definition there is $A_k \in P$ such that $d(A_k) > d(A)$ and $\|F(A, \cdot) - F(A_k, \cdot)\| < 1/k$. Let

$$f_k(x) = f(x) - F(A, x) + F(A_k, x),$$

then

$$f_k(x) - F(A_k, x) = f(x) - F(A, x),$$

hence $f_k - F(A_k, \cdot)$ alternates $d(A)$ times exactly and $\|f - f_k\| < 1/k$. As $f_k - F(A_k, \cdot)$ does not alternate $d(A_k)$ times, $F(A_k, \cdot)$ is not best to f_k. Let $\|f_k - F(B_k, \cdot)\| < \|f_k - F(A_k, \cdot)\|$, then $F(A_k, \cdot) - F(B_k, \cdot)$ alternates in sign on the alternant of $f_k - F(A_k, \cdot)$, hence $F(A_k, \cdot) - F(B_k, \cdot)$ has $d(A)$ zeros. It follows that if B_k is best to f_k, $d(B_k) > d(A)$.

Theorem 8. *Let F be degenerate at A and $f - F(A, \cdot)$ alternate exactly $d(A)$ times. If a best approximation exists to all g in a neighbourhood of f, T is discontinuous at f.*

Proof. Let $\{f_k\}$ be chosen as in the lemma and suppose $\{T(f_k)\} \to T(f)$, then $\{f_k - T(f_k)\} \to f - T(f)$. But $f - T(f)$ alternates exactly $d(A)$ times and since $T(f_k)$ is of degree greater than $d(A)$, $f_k - T(f_k)$ alternates at least $d(A) + 1$ times.

A promising approach to the discontinuity problem is to take known cases of discontinuity of T and abstract from them the properties of F which permit discontinuities to occur. The author developed the property of irregularity (for use in approximation on subsets [6]), which has since been applied by E. Schmidt to the discontinuity problem [14].

Functions with degenerate best approximations. Repeated applications of Lemma 6 give

Lemma 7. *Let $F(A, \cdot)$ be best to f and be degenerate. Let F have bounded degree and a best approximation exist to all g in a neighbourhood of f. There is a sequence $\{f_k\} \to f$ such that the best approximation to f_k is nondegenerate.*

Lemma 8. *The set of continuous functions with degenerate best approximations is a closed subset of the continuous functions with best approximations.*

Proof. Suppose $\{f_k\} \to f$, $\{T(f_k)\}$ is degenerate and $T(f)$ is nondegenerate. By Theorem 6, $\{T(f_k)\} \to T(f)$. But Theorem 4 states that the degenerate approximations are closed and we have a contradiction.

Theorem 9. *Let f have bounded degree and a best approximation exist to every $f \in C[\alpha, \beta]$. The set of functions with degenerate best approximations is nowhere dense in $C[\alpha, \beta]$.*

Proof. Suppose not then by Lemma 8 there must exist an open subset G of $C[\alpha, \beta]$ such that every $f \in G$ has a degenerate best approximation. This contradicts Lemma 7.

BIBLIOGRAPHY

1. R. B. Barrar and H. L. Loeb, *On the continuity of the nonlinear Tschebyscheff operator*, Pacific J. Math. 32 (1970), 593–601. MR 41 #4071.

2. E. W. Cheney and H. L. Loeb, *On the continuity of rational approximation operators*, Arch. Rational Mech. Anal. 21 (1966), 391–401. MR 34 #6402.

3. C. B. Dunham, *Continuity of the varisolvent Chebyshev operator*, Bull. Amer. Math. Soc. 74 (1968), 606–608. MR 36 #6853.

4. ———, *Chebyshev approximation with respect to a weight function*, J. Approximation Theory 2 (1969), 223–232. MR 40 #6137.

5. ———, *Partly alternating families*, J. Approximation Theory 6 (1972), 378–386.

6. ———, *Approximation by alternating families on subsets*, Computing (to appear).

7. ———, *Chebyshev approximation with respect to a non-continuous weight function* (submitted for publication).

8. G. Meinardus and D. Schwedt, *Nicht-Lineare Approximationen*, Arch. Rational Mech. Anal. 17 (1964), 297–326. MR 33 #482.

9. J. R. Rice, *The characterization of best nonlinear Tchebycheff approximations*, Trans. Amer. Math. Soc. 96 (1960), 322–340. MR 22 #8269.

10. ———, *Tchebycheff approximations by functions unisolvent of variable degree*, Trans. Amer. Math. Soc. 99 (1961), 298–302. MR 25 #374.

11. ———, *The approximation of functions*. Vol. 1: *Linear theory*, Addison-Wesley, Reading, Mass., 1964. MR 29 #3795.

12. ———, *The approximation of functions*. Vol. 2: *Nonlinear and multivariate theory*, Addison-Wesley, Reading, Mass., 1969. MR 39 #5989.

13. E. Schmidt, *Stetigkeitsaussagen bei der Tschebyscheff-Approximation mit Exponentialsummen*, Math. Z. 113 (1970), 159–170. MR 41 #5847.

14. E. Schmidt, Unpublished communication.

15. L. Tornheim, *On n-parameter families of functions and associated convex functions*, Trans. Amer. Math. Soc. 69 (1950), 457–467. MR 12, 395.

16. J. L. Kelley, *General topology*, Van Nostrand, Princeton, N.J., 1955. MR 16, 1136.

DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF WESTERN ONTARIO, LONDON, ONTARIO, CANADA