

THE CHARACTERIZATION OF BEST NONLINEAR TCHEBYCHEFF APPROXIMATIONS

BY
JOHN R. RICE⁽¹⁾

1.1. Let two real functions f and F be given such that $f=f(x)$ is a continuous function for $x \in [0, 1]$ and $F=F(a^1, a^2, \dots, a^n, x)$ is a continuous function of n parameters and $x \in [0, 1]$. For simplicity of notation the point (a^1, a^2, \dots, a^n) in Euclidean n -space, E_n , is denoted by a . The domain of the parameters of F is denoted by P , a subset of E_n .

The main problem in the theory of approximation of continuous functions may be stated as follows: Determine $a^* \in P$ so that the deviation of the function $F(a, x)$ from $f(x)$ shall be minimized. Naturally one must define the deviation of $F(a, x)$ from $f(x)$ and different definitions lead to different theories. In this paper the deviation of $F(a, x)$ from $f(x)$ is taken to be $\max_{x \in [0, 1]} |F(a, x) - f(x)|$. All maxima and minima are taken over $x \in [0, 1]$ unless otherwise stated. $F(a^*, x)$ is said to be a best approximation to $f(x)$ if $\max |F(a^*, x) - f(x)| \leq \max |F(a, x) - f(x)|$ for all $a \in P$. The results of this paper remain valid if $[0, 1]$ is replaced by any other closed and bounded interval.

By 1920 an elegant theory of Tchebycheff approximation had been developed for $F(a, x)$ depending linearly on the parameters [2; 3; 7]. Some of the main points of this linear theory are given in the following resume.

Let T be a set of n functions $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$ continuous on $[0, 1]$ and let $P(a, x) = \sum_{i=1}^n a^i \phi_i(x)$ be called a T -polynomial. $P(a, x)$ is said to be a nontrivial T -polynomial if $\sum_{i=1}^n |a^i| > 0$. T is said to be a *Tchebycheff Set* if every nontrivial polynomial has at most $n-1$ zeros. $\max |F(a, x) - f(x)|$ is said to *alternate n times* if there are $n+1$ points $0 \leq x_1 < x_2 < \dots < x_{n+1} \leq 1$ such that $F(a, x_j) - f(x_j) = -[F(a, x_{j+1}) - f(x_{j+1})] = \pm \max |F(a, x) - f(x)|$. Alternance is a basic concept in the theory of Tchebycheff approximations. The following theorem answers the principal questions of a general nature in the theory of linear Tchebycheff approximations.

THEOREM. *Let T be a Tchebycheff set and let $f(x)$ be an arbitrary function continuous on $[0, 1]$. Then*

- A. *$f(x)$ possesses a best approximation,*
- B. *a necessary and sufficient condition that $P(a_0, x)$ be a best approximation to $f(x)$ is that $\max |P(a_0, x) - f(x)|$ alternates at least n times,*
- C. *the best approximation to $f(x)$ is unique.*

Received by the editors October 22, 1959.

⁽¹⁾ Submitted in partial fulfillment for the degree of Doctor of Philosophy at the California Institute of Technology.

It would be desirable to have an explicit method for determining best approximations, but none exists. Theorem B, which characterizes best approximations, provides the principal means of attacking the actual approximation problem [5; 6].

It is natural to try to extend the theory to cases where $F(a, x)$ does not depend linearly on the parameters. Strangely enough little has been done in this direction except that the existence of best approximations has been shown for large classes of nonlinear F and the special case

$$F(a, x) = \left(\sum_{k=1}^n a^k x^k \right) / \left(\sum_{k=n+1}^{n+m+2} a^k x^{k-n-1} \right)$$

has been treated by linear techniques.

The first steps towards a nonlinear theory were made by T. S. Motzkin [4] and L. Tornheim [8]. Tornheim introduced the concept of n -parameter families of functions and Motzkin introduced the equivalent concept of unisolvence. A family of functions is said to form an n -parameter family if given a set of n points, $(x_1, y_1), \dots, (x_n, y_n)$ with $0 \leq x_1 < \dots < x_n \leq 1$, then there is a unique member of the family passing through these points. Motzkin's definition of unisolvence involves more assumptions but these may be shown to be redundant. Motzkin and Tornheim noted that if the family of functions defined by $F(a, x)$ were unisolvent (n -parameter) then results analogous to A, B, C hold. These and related concepts will play an important role in this paper.

This paper will give some general results on the characterization of best nonlinear Tchebycheff approximations. It will go further than to give conditions on F sufficient for the generalization of the classical linear theory to be valid. Under various basic assumptions on F , it will give criteria which are both necessary and sufficient for the generalization of Theorem B to hold.

A similar problem in the linear theory has been considered by A. Haar [3]. He has shown for $F(a, x) = \sum_{k=1}^n a^k \phi_k(x)$ that a necessary and sufficient condition on F for best approximations to be unique is that $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$ form a Tchebycheff set. This is to say that a necessary and sufficient condition for Theorem C to be valid is that $\phi_1(x), \phi_2(x), \dots, \phi_n(x)$ form a Tchebycheff set.

Let Y be a set of basic assumptions on F . A natural generalization of Theorem B is: *Let F satisfy Y and have Property X. Then the alternation of $\max |F(a^*, x) - f(x)|$ at least n times is a necessary and sufficient condition for $F(a^*, x)$ to be a best approximation to $f(x)$.* The problem considered in the second part of this paper is as follows: given a set Y determine a Property X which is both necessary and sufficient for this theorem to hold for all continuous $f(x)$.

This problem is solved for three sets of basic assumptions on F . The first set essentially assumes only that F is continuous. The resulting Property X

is quite complicated. The second set assumes that the family of curves defined by F is closed under pointwise convergence. The resulting Property X is very closely related to unisolvence. The third set assumes a stronger closure property for the family of curves defined by F . The resulting Property X is unisolvence.

In the third part of this paper the degree of unisolvence is defined as a function of the parameters. F is said to be unisolvent of degree k at $a^* \in P$ if F is unisolvent for all sets of k points on the graph of $F(a^*, x)$, $x \in [0, 1]$. The extension of Theorem B is then taken to be: *Let F satisfy Y and have Property X. Then a necessary and sufficient condition that $F(a^*, x)$ be a best approximation to $f(x)$ is that $\max |F(a^*, x) - f(x)|$ alternates at least k times.* Here k is not assumed to be a constant. The second set of conditions from the second part are taken for Y. It is then shown that Property X is "unisolvence of degree k at a^* " and is both necessary and sufficient for this theorem to be valid for all $f(x)$. This theory is of great practical interest and includes many common nonlinear approximating functions.

2.1. Euclidean n -dimensional space is denoted by E_n ; points in E_n are denoted by a, b , etc. and the coordinates of a are (a^1, a^2, \dots, a^n) . Curly brackets, $\{ \}$, denote a set and $\{x | \dots\}$ is read as "the set of x such that \dots ." Sequences are denoted by square brackets.

The real function $F = F(a, x)$ is defined for $x \in [0, 1]$ and $a \in P$ where P is a subset of E_n . F is continuous in the sense that given $a_0 \in P$, $x_0 \in [0, 1]$ and $\epsilon > 0$ there is a $\delta > 0$ such that $a \in P$, $x \in [0, 1]$, $|a_0 - a| + |x_0 - x| < \delta$ implies that $|F(a_0, x_0) - F(a, x)| < \epsilon$. It is assumed that if $a \neq a^*$ then $F(a, x) \neq F(a^*, x)$ for some $x \in [0, 1]$. The range R of F is defined to be the set $\{(x, F(a, x)) | x \in [0, 1], a \in P\}$ in the x, y plane. $f(x)$, $g(x)$, etc. will denote continuous functions on $[0, 1]$.

The *Tchebycheff Approximation Problem* for a continuous function, $f(x)$, may be stated as: *Determine $a^* \in P$ such that $\max |F(a^*, x) - f(x)| \leq \max |F(a, x) - f(x)|$ for all $a \in P$.* A solution, $F(a^*, x)$, to the approximation problem is called a *best approximation* to $f(x)$.

The concept of "alternance" plays a major role in the classical theory of Tchebycheff approximation developed for polynomials and Tchebycheff sets. The aim of this part of this paper is to study "alternance" in more general situations and to characterize it there. The following definitions are required.

DEFINITION 1. $\max |F(a^*, x) - f(x)|$ is said to *alternate k times* if there are $k + 1$ points $0 \leq x_1 < x_2 < \dots < x_{k+1} \leq 1$, such that $F(a^*, x_i) - f(x_i) = -(F(a^*, x_{i+1}) - f(x_{i+1})) = \pm \max |F(a^*, x) - f(x)|$. If $\max |F(a^*, x) - f(x)|$ alternates k times but not $k + 1$ times then $\max |F(a^*, x) - f(x)|$ is said to *alternate exactly k times*.

DEFINITION 2. F has *Property N* if for every continuous $f(x)$, the alternance n times of $\max |F(a^*, x) - f(x)|$ is a necessary condition for $F(a^*, x)$ to be a best approximation to $f(x)$.

DEFINITION 3. F has Property S if for every continuous $f(x)$, the alternance n times of $\max |F(a^*, x) - f(x)|$ is a sufficient condition for $F(a^*, x)$ to be a best approximation to $f(x)$.

DEFINITION 4. F has Property NS if F has both Property N and Property S.

From the last definition it is seen that the problem of characterizing "alternance" is the problem of characterizing Property NS.

2.2. The two properties needed to characterize Property NS are given in Definitions 5 and 6.

DEFINITION 5. F has Property Z of degree k if $a \neq a^*$ implies that $F(a, x) - F(a^*, x)$ has at most $k-1$ zeros in $[0, 1]$.

The phrase "of degree k " is omitted if $k=n$ is understood.

DEFINITION 6. F has Property A if given $a^* \in P$, $k < n$,

$$\{x_j \mid 0 = x_0 < x_1 < \cdots < x_{k+1} = 1\}$$

and ϵ with $0 < \epsilon < 2^{-1} \min_j (x_{j+1} - x_j)$, $j=0, 1, \dots, k$ then (i) there exist $a_1, a_2 \in P$ such that for $x \in [0, 1]$ $F(a^*, x) - \epsilon < F(a_1, x) < F(a^*, x) < F(a_2, x) < F(a^*, x) + \epsilon$ (ii) there exist $a_3, a_4 \in P$ such that $|F(a_3, x) - F(a^*, x)| < \epsilon$, $|F(a_4, x) - F(a^*, x)| < \epsilon$ for $x \in [0, 1]$ and $F(a_3, x) - F(a^*, x)$, $F(a_4, x) - F(a^*, x)$ change sign from $x_j - \epsilon$ to $x_j + \epsilon$, and have no zeros outside $[x_j - \epsilon, x_j + \epsilon]$, $j=1, 2, \dots, k$. Further $F(a_3, 0) > F(a^*, 0) > F(a_4, 0)$.

The first theorem characterizing Property NS can now be stated.

THEOREM 1. F has Property NS if and only if F has Property A and Property Z.

The proof of Theorem 1 follows from a series of four lemmas.

LEMMA 1. If F has Property A then F has Property N.

Proof. Assume that F has Property A but does not have Property N, i.e., $F(a^*, x)$ is a best approximation to $f(x)$ and $\max |F(a^*, x) - f(x)|$ does not alternate n times. If $\max |F(a^*, x) - f(x)|$ does not alternate at all, let x_0 be a point where $|F(a^*, x) - f(x)|$ assumes its maximum. Say $F(a^*, x_0) - f(x_0) > 0$, then by (i) of Definition 6 there is an $a_2 \in P$ so that $0 < F(a^*, x) - F(a_2, x) < \epsilon$. For ϵ sufficiently small it is seen that $F(a_2, x)$ is a better approximation to $f(x)$ than $F(a^*, x)$. Hence $\max |F(a^*, x) - f(x)|$ must alternate at least once if F has Property A.

The interval $[0, 1]$ may be divided into l subintervals by $0 = x_0 < x_1 < \cdots < x_l = 1$, $l \leq n$ so that (i) $F(a^*, x_j) - f(x_j) = 0$, $j=1, 2, \dots, l-1$ and (ii) $\max |F(a^*, x) - f(x)|$ alternates exactly once on any two adjacent subintervals but does not alternate in any one subinterval. Let $\alpha_j = \max (F(a^*, x) - f(x)) - \min (F(a^*, x) - f(x))$ with the maximum and minimum taken over $[x_j, x_{j+1}]$, $j=0, 1, \dots, l-1$. Let $\eta_j = \max |F(a^*, x) - f(x)| - 2^{-1}\alpha_j$, $j=0, 1, \dots, l-1$. Let $\epsilon' = 2^{-1} \min |x_{j+1} - x_j|$. About each x_j , $j=1, 2, \dots, l-1$

there is an interval $[x_j - \epsilon_j, x_j + \epsilon_j]$ such that in this interval $|F(a^*, x) - f(x)| \leq 4^{-1} \max |F(a^*, x) - f(x)|$. Let

$$\epsilon = \min(\epsilon_1, \dots, \epsilon_{l-1}, \epsilon', \eta_0, \dots, \eta_{l-1}, 4^{-1} \max |F(a^*, x) - f(x)|)$$

and determine $a \in P$ by (ii) of Definition 6 so that $F(a, x) - F(a^*, x)$ changes sign from $x_j - \epsilon$ to $x_j + \epsilon$, $j = 1, 2, \dots, l-1$ with $\max |F(a^*, x) - F(a, x)| < \epsilon$. Further choose the sign of $F(a^*, 0) - F(a, 0)$ so that $|F(a^*, x) - f(x)| > |F(a, x) - f(x)|$ in $(0, x_1 - \epsilon)$; the same inequality will then hold outside of $[x_j - \epsilon, x_j + \epsilon]$, $j = 1, 2, \dots, l-1$. $F(a, x)$ is a better approximation to $f(x)$ than $F(a^*, x)$ which contradicts the assumption and proves the lemma.

LEMMA 2. *If F has Property Z then F has Property S.*

Proof. Assume that $\max |F(a^*, x) - f(x)|$ alternates n times and that there is an $a \in P$ such that $\max |F(a, x) - f(x)| < \max |F(a^*, x) - f(x)|$. Let $\{x_j | j = 1, 2, \dots, n+1, x_j < x_{j+1}\}$ be a set of points in $[0, 1]$ such that $F(a^*, x_j) - f(x_j) = -(F(a^*, x_{j+1}) - f(x_{j+1})) = \pm \max |F(a^*, x) - f(x)|$. Then $\text{sgn}(F(a^*, x_j) - F(a, x_j)) = -\text{sgn}(F(a^*, x_{j+1}) - F(a, x_{j+1}))$ which implies that $F(a^*, x) - F(a, x)$ has at least n zeros contradicting Property Z.

LEMMA 3. *If F has Property NS then F has Property Z.*

Proof. Assume that F has Property NS and that there are $a, a^* \in P$ such that $F(a, x) - F(a^*, x)$ has n or more zeros. Let $M(x) = 2^{-1}[F(a, x) + F(a^*, x)]$; $\eta = 2^{-1} \max |F(a, x) - F(a^*, x)|$ and let $0 \leq x_1 < x_2 < \dots < x_{n+1} \leq 1$ be n points where $F(a, x) - F(a^*, x) = 0$ along with one point where $F(a, x) - F(a^*, x) \neq 0$, say $F(a^*, x_2) > F(a, x_2)$ for concreteness. Set

$$\delta = 4^{-1} \min(x_1, x_2 - x_1, \dots, x_{n+1} - x_n, 1 - x_{n+1})$$

with x_1 or $1 - x_{n+1}$ omitted if they are zero. A function $f(x)$ continuous on $[0, 1]$ will be defined as follows: $f(x) = M(x)$ in the intervals

$$[x_j + \delta, x_{j+1} - \delta], j = 1, 2, \dots, n;$$

$f(x) = M(x)$ in $[0, x_1 - \delta]$ if $x_1 > 0$; $f(x) = M(x)$ in $[x_{n+1} + \delta, 1]$ if $x_{n+1} < 1$; in the remaining intervals $f(x)$ satisfies $f(x_j) = F(a^*, x_j) + (-1)^{j+1}(3\eta/2)$ and $|F(a, x) - f(x)| < 3\eta/2$, $|F(a^*, x) - f(x)| < 3\eta/2$ for $x \neq x_j$.

Now $\max |F(a^*, x) - f(x)|$ alternates exactly n times and

$$\max |F(a, x) - f(x)|$$

alternates exactly $n-2$ times. By Property S $F(a^*, x)$ is a best approximation to $f(x)$ and since $\max |F(a, x) - f(x)| = \max |F(a^*, x) - f(x)|$ so is $F(a, x)$. This contradicts Property N and proves the lemma.

LEMMA 4. *If F has Property NS then F has Property A.*

Proof. Let $a^* \in P$, $\{x_j | 0 = x_0 < x_1 < \cdots < x_{l+1} = 1\}$, $l \leq n-1$ and ϵ be given with $0 < \epsilon < 2^{-1} \min_j (x_{j+1} - x_j)$, $j = 0, 1, \cdots, l$. A construction showing the existence of a_3 of Definition 6 will be described, the existence of a_1, a_2 and a_4 follow by similar constructions. A function $f(x)$ will be defined as follows: $f(x)$ is continuous on $[0, 1]$; in $[x_j + \epsilon, x_{j+1} - \epsilon]$, $j = 1, 2, \cdots, l-1$, $f(x) = F(a^*, x) + (-1)^{j+1}(\epsilon/3)$, and $f(x) - F(a^*, x)$ is linear in each of the remaining intervals. Since $\max |F(a^*, x) - f(x)|$ alternates at most l times, $F(a^*, x)$ is not a best approximation to $f(x)$ by Property N. Hence there is an $a_3 \in P$ which gives a better approximation to $f(x)$ than $F(a^*, x)$. This implies that

$$\max |F(a_3, x) - F(a^*, x)| < 2\epsilon/3.$$

Further $F(a_3, x) - F(a^*, x)$ has no zeros outside of the intervals $[x_j - \epsilon, x_j + \epsilon]$, $j = 1, 2, \cdots, l$ and alternates in sign from $x_j - \epsilon$ to $x_j + \epsilon$, $j = 1, 2, \cdots, l$.

It is clear that Theorem 1 follows from these four lemmas.

A difficult problem associated with this theorem is the problem of possible improvement. It is difficult to know whether there is some other property, apparently simpler or more restrictive, which, along with Property Z, is equivalent to Property NS. The following example shows that it is unlikely that Theorem 1 can be improved upon in this sense. This example shows that there exists a one parameter function F with Property NS which is unlikely to satisfy any set of conditions simpler than those given for Property A.

Let the number p be of the form $c/2^m$ where c is odd. Then the index $m(p)$ of p is defined to be m . The multiplication of a set P of real numbers by a constant a is defined by $aP = \{ap | p \in P\}$ and $p_0 + aP$ means $\{p_0 + ap | p \in P\}$.

Let $P_1 = \{\pm 2^{-j} | j = 1, 2, \cdots\} \cup \{0\}$. P_k is defined recursively from P_{k-1} as follows: let $p \in P_{k-1}$ and form the set $p + P_1/2^{2+m(p)}$, P_k is defined to be the union of all such sets for $p \in P_{k-1}$. Let $P = \bigcup_{k=1}^{\infty} P_k$. It may be shown that P is a nowhere dense countable set.

P is taken as the parameter space of F with $F(a, x) = a$ for $a \in P$. Every point of P is the limit from both the left and right of points in P , giving F Property A. Hence F so defined has Property A, Property Z and Property NS.

2.3. With further assumptions on the function F a much simpler and more elegant characterization of Property NS may be given. As a preliminary a theorem in the field of Real Variables is proved, a theorem of some interest in itself.

THEOREM 2. Let $\{f_\alpha(x)\}$ be a uniformly bounded infinite set of functions continuous on $[0, 1]$ with Property Z. Then $\{f_\alpha(x)\}$ contains a pointwise convergent subsequence.

The proof is broken into four parts. In the following when a sequence of

functions is chosen whose function values converge at a particular point, it will always be chosen so that the convergence is monotonic at that point.

LEMMA 5. *Let $\{f_\alpha(x)\}$ be a uniformly bounded infinite set of functions continuous in $[0, 1]$ which are mutually nonintersecting. Then $\{f_\alpha(x)\}$ contains a pointwise convergent sequence.*

Proof. Consider $\{f_\alpha(0)\}$ which is a bounded set of real numbers and hence contains a convergent monotonic sequence $[f_k(0) | k=1, 2, \dots]$. Since $[f_k(0)]$ is monotonic so is $[f_k(x)]$ monotonic for every $x \in [0, 1]$. Since $[f_k(x)]$ is bounded it follows that $[f_k(x)]$ is convergent for every $x \in [0, 1]$.

LEMMA 6. *Let $\{f_\alpha(x)\}$ be a uniformly bounded infinite set of functions continuous on $[0, 1]$ with Property Z of degree 2. Then $\{f_\alpha(x)\}$ contains a pointwise convergent sequence.*

Proof. For simplicity the sequences $f_{1l}(x), f_{2l}(x), \dots$ and $g_{1l}(x), g_{2l}(x), \dots$ are denoted by F_l and G_l and their limits at x are denoted by $F_l(x)$ and $G_l(x)$ if these limits exist and if F_l, G_l are monotonic at x . $F_l \subset G_l$ will be used to indicate that F_l is a subsequence of G_l with the order preserved.

Choose a sequence F_0 from $\{f_\alpha(x)\}$ such that $F_0(0)$ and $F_0(1)$ exist. If F_0 is increasing at both 0 and 1 or decreasing at both points then the functions of $\{f_{k0}(x)\}$ must be mutually nonintersecting and Lemma 5 states that F_0 is a point-wise convergent sequence.

If F_0 converges in opposite directions at 0 and 1, then choose $F_1 \subset F_0$ such that $F_1(1/2)$ exists. The direction of convergence of F_1 at $x=1/2$ must agree with the direction of convergence of F_1 at either 0 or 1. Hence a subsequence $G_1 \subset F_1$ may be chosen which is pointwise convergent in an interval of length $1/2$ as in Lemma 5.

Choose $F_2 \subset G_1$ so that F_2 converges at the midpoint of the remaining interval. Again the direction of convergence of F_2 at the midpoint must agree with the direction of convergence of F_2 at one of the endpoints. As in Lemma 5 there is a subsequence $G_2 \subset F_2$ which is pointwise convergent except in an interval of length $1/4$.

In this way an infinite number of subsequences $G_1 \supset G_2 \supset \dots \supset G_l \supset \dots$ may be found where G_l is pointwise convergent except possibly in an interval of length 2^{-l} . The diagonal sequence of this sequence of sequences may be formed and it converges at all points of $[0, 1]$ with possibly one exception. In case the diagonal sequence fails to converge at one point then a further subsequence is chosen that converges there also.

LEMMA 7. *Let $\{f_\alpha(x)\}$ be a uniformly bounded infinite set of functions continuous on $[0, 1]$ with Property Z and such that for every $x \in [0, 1]$, every monotonic infinite sequence from $\{f_\alpha(x)\}$ is monotonic decreasing. Then $\{f_\alpha(x)\}$ contains a pointwise convergent sequence.*

Proof. The proof is by induction on the degree of Property Z. The lemma has been established for $k=1$ and $k=2$. It will be shown that if the lemma is true for Property Z of degree k then it is true for Property Z of degree $k+2$.

Select a sequence F_0 from $\{f_\alpha(x)\}$ which converges at $x=0, 1/2, 1$. F_0 is monotonically decreasing at these points. The assertion is now made that F_0 either has a subsequence which is convergent in $[0, 1/2]$ or one which is convergent in $[1/2, 1]$.

Assume that there is no convergent subsequence in $[1/2, 1]$. Then there is an $x_1 \in [1/2, 1]$ such that $F_0(x_1)$ does not exist. Choose subsequences $F_1 \subset F_0$, $G_1 \subset F_0$ such that $F_1(x_1) < G_1(x_1)$. For some k_1 , $f_{k_1}(x_1) < G_1(x_1)$ and therefore there is a subsequence $H_1 \subset G_1$ such that every function of H_1 intersects $f_{k_1}(x)$ in $(1/2, x_1)$ and $(x_1, 1)$.

H_1 does not converge in $[1/2, 1]$ and hence there is an x_2 , say in $[x_1, 1]$, such that $H_1(x_2)$ does not exist. Choose $F_2 \subset H_1$, $G_2 \subset H_1$ such that $F_2(x_2) < G_2(x_2)$. Again for some k_2 , $f_{k_2}(x_2) < G_2(x_2)$ and hence there is a subsequence $H_2 \subset G_2$ such that every function of H_2 intersects $f_{k_2}(x)$ in (x_1, x_2) and $(x_2, 1)$.

In this way an infinite sequence $F_0^* = [f_{k_l}(x) | l=1, 2, \dots, f_{k_l}(x) \in F_l]$ is obtained such that every member of this sequence intersects every other member at least twice in $[1/2, 1]$. Therefore in $[0, 1/2]$ F^* has Property Z of degree k . The induction hypothesis is applied in $[0, 1/2]$ to obtain a pointwise convergent subsequence $G_0^* \subset F_0^*$.

It has thus been shown that there is a sequence G_0^* from $\{f_\alpha(x)\}$ which converges in an interval of length $1/2$. The entire construction is now repeated to obtain $G_1^* \subset G_0^*$ which converges in half of the remaining interval. Continuing in this way a sequence of sequences $G_0^* \supset G_1^* \supset \dots$ is obtained which converge except in intervals of length $1/2, 1/4, \dots$. The diagonal process is used as in Lemma 6 to obtain a subsequence which is pointwise convergent in $[0, 1]$.

It is clear that this lemma is also true when decreasing is replaced by increasing.

Proof of Theorem 2. The proof is by induction. The induction hypothesis is that the theorem is true for Property Z of degree k . For $k=1$ the theorem follows from Lemmas 5 and 6.

It is assumed that $\{f_\alpha(x)\}$ has Property Z of degree $k+1$. If there is a subset of $\{f_\alpha(x)\}$ satisfying the assumptions of Lemma 7 the proof is complete. If there is no such subset then there is a sequence F_0 from $\{f_\alpha(x)\}$ which has opposite directions of convergence at two points, x_1 and x_2 . Let A be the set of points in $[0, 1]$ such that all subsequences of F_0 are monotonic at these points. A is nonvoid since it contains x_1 and x_2 . If $A \neq [0, 1]$ then x_3 may be chosen in A and x_4 not in A such that $|x_3 - x_4|$ is arbitrarily small. If $A = [0, 1]$ then A may be divided into two nonvoid sets; one with all increasing subsequences and one with all decreasing subsequences. Then x_3 may be

chosen in the first set and x_4 in the second so that $|x_3 - x_4|$ is arbitrarily small. In either case there is a subsequence $F_1 \subset F_0$ and $x_3, x_4 \in [0, 1]$ such that $x_3 < x_4$, F_1 has opposite directions of convergence at x_3 and x_4 and $[x_3, x_4]$ has length less than $1/2$.

Since every member of F_1 intersects every other member of F_1 at least once in $[x_3, x_4]$, in each of the intervals $[0, x_3]$ and $[x_4, 1]$, F_1 has Property Z of degree k . By the induction hypothesis there is a subsequence $F_2 \subset F_1$ which converges pointwise in these intervals.

If F_2 satisfies the assumptions of Lemma 7 in $[x_3, x_4]$ the proof is complete. If not, a repetition of the above argument shows that there is a subsequence $F_3 \subset F_2$ which converges outside of $[x_6, x_8]$ where the length of $[x_6, x_8]$ is less than $1/4$.

This process is repeated to obtain an infinite number of sequences $F_1 \supset F_2 \supset \dots \supset F_l \supset \dots$ where F_l is pointwise convergent except in an interval of length less than 2^{-l} . The diagonal sequence may be taken as in Lemma 6 to obtain a pointwise convergent sequence in $[0, 1]$.

2.4. The assumptions on F with regard to the dependence of $F(a, x)$ on a have so far been rather weak. In cases of general interest F satisfies much more stringent conditions. These conditions are (i) a closure property of some kind for the family of functions defined by F and (ii) a regularity property for P . Two such conditions are exactly defined in the following:

DEFINITION 7. F is said to be closed if P is arcwise connected and if F is closed under pointwise limits, i.e., $\lim_{k \rightarrow \infty} F(a_k, x) = G(x)$ with $(x, G(x)) \in R$ implies that there is an $a_0 \in P$ such that $\lim_{k \rightarrow \infty} a_k = a_0$ and hence $F(a_0, x) = G(x)$.

The aim of this section is to characterize Property NS with the additional assumption of closure on F .

If P is an arcwise connected set then R is seen to be a particularly simple region. Let $L(x) = \{(x, F(a, x)) | a \in P\}$. Since F is continuous and P is connected $L(x)$ is a line segment and $R = \bigcup_{x \in [0, 1]} L(x)$. Let $L_0(x)$ be the interior of $L(x)$ in the one dimensional topology and let $R_0 = \bigcup_{x \in [0, 1]} L_0(x)$.

LEMMA 8. If P is arcwise connected and F has Property NS then given $a^* \in P$ there is an $\epsilon > 0$ such that $(x, F(a^*, x) \pm \epsilon) \in R_0$ for all x .

Proof. Since F has Property NS by Theorem 1, F has Property A. By (i) of Property A there are $a_1, a_2 \in P$ such that $F(a_1, x) < F(a^*, x) < F(a_2, x)$. Let $\epsilon_1 = \min |F(a_1, x) - F(a^*, x)|$, $\epsilon_2 = \min |F(a_2, x) - F(a^*, x)|$ and $\epsilon = 2^{-1} \min(\epsilon_1, \epsilon_2)$. Since P is arcwise connected it is clear that $(x, F(a^*, x) \pm \epsilon) \in R_0$ for all x .

The following property will be used in the characterization of Property NS.

DEFINITION 8. F is said to be locally solvent if given $0 \leq x_1 < x_2 < \dots < x_n \leq 1$, $a^* \in P$ and $\epsilon > 0$ there is a $\delta(a^*, \epsilon, x_1, x_2, \dots, x_n) > 0$ such that $|y_j - F(a^*, x_j)| < \delta$ implies the existence of a solution $a \in P$ to $F(a, x_j) = y_j$ with $\max |F(a, x) - F(a^*, x)| < \epsilon$.

DEFINITION 9. *F is said to be locally unisolvent if F is locally solvent and has Property Z.*

LEMMA 9. *If P is arcwise connected and F is locally unisolvent then $R=R_0$.*

Proof. Assume that $(x_0, y_0) \notin R_0$ with $y_0 = F(a^*, x_0)$. By the local solvency of F there is a $\delta_0(a^*, x_0, \dots) > 0$ and $a_1, a_2 \in P$ such that $F(a_1, x_0) = F(a^*, x_0) + \delta_0/2$, $F(a_2, x_0) = F(a^*, x_0) - \delta_0/2$. This is a contradiction and proves the lemma.

The purpose of Lemma 9 is to show that Property NS and local unisolvence each imply that $R=R_0$ if P is connected.

A zero x_0 of $f(x)$ is said to be a simple zero if $f(x)$ changes sign at x_0 and a double zero if $f(x)$ does not change sign at x_0 . The following lemma shows that if F is locally unisolvent then Property Z is valid counting the multiplicity of zeros. Only multiplicities of 1 or 2 are considered.

LEMMA 10. *Let F be locally unisolvent. Then $F(a, x) - F(a^*, x)$ cannot have more than $n-1$ zeros counting multiplicities.*

Proof. Let $\{x_j | j=1, 2, \dots, k\}$ be the set of zeros of $F(a, x) - F(a^*, x)$. Assume, for concreteness, that x_1 is a double zero and $F(a, x) - F(a^*, x) > 0$ near x_1 . Since F is locally solvent there is an $a_1 \in P$ such that $F(a_1, x_j) = F(a, x_j)$ $j \neq 1$, $F(a_1, x_1) < F(a, x_1)$ and $\max |F(a_1, x) - F(a, x)| < \epsilon$. For ϵ sufficiently small it is clear that (i) $\{x_j | j=2, 3, \dots, k\}$ are zeros of $F(a_1, x) - F(a^*, x)$; (ii) every double zero of $F(a, x) - F(a^*, x)$ is either a double zero of $F(a_1, x) - F(a^*, x)$ or has been replaced by two simple zeros; (iii) $F(a_1, x) - F(a^*, x)$ has two simple zeros near x_1 . Hence, counting multiplicities, the number of zeros of $F(a_1, x) - F(a^*, x)$ is the same as the number of zeros of $F(a, x) - F(a^*, x)$. This process is continued to obtain $a_m \in P$ such that $F(a_m, x) - F(a^*, x)$ has all simple zeros and the number of zeros of $F(a_m, x) - F(a^*, x)$ and $F(a, x) - F(a^*, x)$ is the same counting multiplicities. This completes the proof.

Property NS can now be characterized by

THEOREM 3. *If F is closed then local unisolvence is a necessary and sufficient condition for F to have Property NS.*

The fact that F is closed leads to the following stronger form of Lemma 4.

LEMMA 11. *Let F be closed and have Property NS. Then given $a^* \in P$, $\epsilon > 0$ and a set $\{x_j | 0 = x_0 \leq x_1 < \dots < x_k \leq x_{k+1} = 1; k < n\}$ there are $a_1, a_2 \in P$ so that*

(i) $F(a^*, x) - F(a_1, x), F(a^*, x) - F(a_2, x)$ change sign at the $x_j, j=1, 2, \dots, k$ and at no other points.

(ii) $|F(a_1, x) - F(a^*, x)| \leq \epsilon, |F(a_2, x) - F(a^*, x)| \leq \epsilon.$

(iii) For each x either $F(a_1, x) - F(a^*, x) = 0, F(a_2, x) - F(a^*, x) = 0$ or $\text{sgn}[F(a_1, x) - F(a^*, x)] = -\text{sgn}[F(a_2, x) - F(a^*, x)]$.

Proof. The proof is given for a_2 , the proof for a_1 follows in the same way. By Lemma 8 it is possible to take ϵ small enough that $(x, F(a^*, x) \pm \epsilon) \in R_0$ for all x . Let $1/l_0 < 4^{-1} \min (x_1, x_2 - x_1, \dots, x_k - x_{k-1}, 1 - x_k)$ where x_1 or $1 - x_k$ is omitted if they are zero. A sequence of functions

$$[f_l(x) \mid l = l_0, l_0 + 1, \dots]$$

will be defined so that: $f_l(x)$ is continuous in $[0, 1]$; in the intervals $[x_j + 1/l, x_{j+1} - 1/l]$, $j = 2, 3, \dots, k-1$, $f_l(x) = F(a^*, x) + (-1)^j(\epsilon/3)$; if $x_1 > 0$, $f_l(x) = F(a^*, x) + \epsilon/3$ in $[0, x_1 - 1/l]$; if $x_k < 1$, $f_l(x) = F(a^*, x) + (-1)^k(\epsilon/3)$ in $[x_k, 1]$; $f_l(x) - F(a^*, x)$ is linear in each of the remaining intervals. $F(a^*, x)$ is not a best approximation to any $f_l(x)$ for $\max |F(a^*, x) - f_l(x)|$ alternates at most k times. However, it will be shown that there is a $b_l \in P$ such that $F(b_l, x)$ is a best approximation to $f_l(x)$.

Consider a sequence $[F(b_{li}, x) \mid i = 1, 2, \dots]$ for which

$$\lim_{i \rightarrow \infty} \max |F(b_{li}, x) - f_l(x)| = \inf_{a \in P} \max |F(a, x) - f_l(x)|.$$

$\{F(b_{li}, x)\}$ forms a bounded infinite set of continuous functions. This set has Property Z by Lemma 3. Theorem 2 may be applied to extract a convergent subsequence of $[F(b_{li}, x)]$. By the closure of F the subsequence has a limit, $F(b_l, x)$, which is a best approximation to $f_l(x)$.

In the same way $[F(b_l, x) \mid l = l_0, l_0 + 1, \dots]$ has a limit $F(a_2, x)$. Clearly $0 \leq [F(a_2, x) - F(a^*, x)](-1)^j \leq (-1)^j \epsilon$ in (x_j, x_{j+1}) , $j = 0, 1, 2, \dots, k$. If $a_2 \neq a^*$ then it is seen that $F(a_2, x) - F(a^*, x)$ changes sign at the x_j , $j = 1, 2, \dots, k$ and nowhere else.

It remains to be shown that $a_2 \neq a^*$. Since $\max |F(b_l, x) - f_l(x)|$ alternates n times, $\max |F(b_l, x) - F(a^*, x)| > \epsilon/3$. By the closure of F $\lim_{l \rightarrow \infty} b_l = a_2$. Since $\max |F(a, x) - F(a^*, x)|$ is a continuous function of a it follows that $\lim_{l \rightarrow \infty} \max |F(b_l, x) - F(a^*, x)| = \max |F(a_2, x) - F(a^*, x)| \geq \epsilon/3$ and $a_2 \neq a^*$.

DEFINITION 10. $F(a, x)$ is said to be between $F(a_1, x)$ and $F(a_2, x)$ if $|F(a_1, x) - F(a, x)| + |F(a, x) - F(a_2, x)| = |F(a_1, x) - F(a_2, x)|$ for all $x \in [0, 1]$.

LEMMA 12. Let $F(a_1, x) - F(a_2, x)$ change sign at $n-1$ distinct points and let $F(a_1, x_0) \neq F(a_2, x_0)$. If F is closed and has Property NS then given y between $F(a_1, x_0)$ and $F(a_2, x_0)$ there is an $a \in P$ such that $F(a, x_0) = y$ and $F(a, x)$ is between $F(a_1, x)$ and $F(a_2, x)$.

Proof. Since F has Property NS it follows from Lemma 3 that F has Property Z and hence $F(a_1, x) - F(a_2, x)$ has exactly $n-1$ zeros.

For concreteness assume $F(a_1, x_0) > F(a_2, x_0)$ and let $I = [F(a_2, x_0), F(a_1, x_0)]$ and A be the subset of I for which the lemma is true. Clearly $F(a_1, x_0) \in A$.

It will now be shown that A is a closed set. Let $[y_j] \mid j = 1, 2, \dots, y_j \in A]$ be given with $\lim_{j \rightarrow \infty} y_j = y_0$. Corresponding to $[y_j]$ there is a sequence

$[F(a_j, x) | F(a_j, x_0) = y_j]$ of functions between $F(a_1, x)$ and $F(a_2, x)$. Theorem 2 is applicable to this sequence and hence $[F(a_j, x)]$ has a pointwise convergent subsequence. Since F is closed this subsequence has a limit $F(a_0, x)$ and $F(a_0, x_0) = y_0$ which implies $y_0 \in A$.

It will now be shown that $I - A$ is empty. $I - A$ is an open set. If $I - A$ is not empty it contains an open interval (y_0, y_1) with $y_0, y_1 \in A$. Since $y_0 \in A$ there is an $a_3 \in P$ so that $F(a_3, x_0) = y_0$ and $F(a_3, x)$ is between $F(a_1, x)$ and $F(a_2, x)$. By Lemma 11 there is an $a_4 \in P$ such that $F(a_4, x)$ is between $F(a_1, x)$ and $F(a_2, x)$ and $y_0 < F(a_4, x_0) < y_1$. This contradicts the existence of an open interval in $I - A$ and $A = I$.

LEMMA 13. *If F is closed and has Property NS then F is locally solvent.*

Proof. The proof is conducted by an induction argument. The statement to be verified by induction is: Let $a^* \in P$, $\epsilon > 0$ and a set

$$\{x_j | 0 \leq x_1 < x_2 < \cdots < x_n \leq 1\}$$

be given. Then there is a $\delta(a^*, \epsilon, k, x_j) > 0$ such that $|y_j - F(a^*, x_j)| \leq \delta$ $j = 1, 2, \cdots, k$ implies that there is an $a \in P$ such that $F(a, x_j) = y_j$, $j = 1, 2, \cdots, k$; $|F(a, x) - F(a^*, x)| < \epsilon$ and $F(a, x_j) = F(a^*, x_j)$, $j = k+1, \cdots, n$. The induction is on k and for $k = n$ the statement implies that F is locally solvent.

For $k = 1$ this statement will be established from Lemmas 11 and 12. Let $F(a_1, x)$ and $F(a_2, x)$ be determined by Lemma 11 so that $F(a_1, x_j) = F(a_2, x_j) = F(a^*, x_j)$, $j = 2, 3, \cdots, n$; $|F(a_1, x) - F(a^*, x)| < \epsilon$, $|F(a_2, x) - F(a^*, x)| < \epsilon$ and $F(a_1, x_1) > F(a^*, x_1) > F(a_2, x_1)$. Since $F(a_1, x)$, $F(a_2, x)$ intersect at $n - 1$ points it is not possible that $F(a_1, x_1) = F(a^*, x_1)$ or $F(a_2, x_1) = F(a^*, x_1)$. Take $\delta(a^*, \epsilon, 1, x_j) = \min(|F(a_1, x_1) - F(a^*, x_1)|, |F(a_2, x_1) - F(a^*, x_1)|)$ and the statement follows for $k = 1$ by Lemma 12.

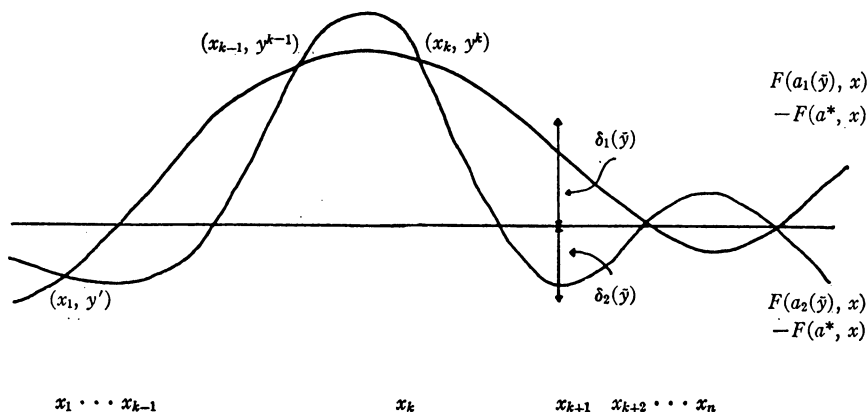
In order to proceed with the general induction step some auxiliary facts will be established. Let $\bar{y} = (y^1, y^2, \cdots, y^k)$ be a point in E_k and let $A = \{\bar{y} | |y^i - F(a^*, x_i)| \leq \delta(a^*, \epsilon, k, x_j), i = 1, 2, \cdots, k\}$. Define two subsets, $P_1(\bar{y})$ and $P_2(\bar{y})$, of P as follows: $|F(a, x) - F(a^*, x)| < \epsilon$ for $a \in P_1(\bar{y})$, $a \in P_2(\bar{y})$; $F(a, x_j) = y^j$, $j = 1, 2, \cdots, k$, $a \in P_1(\bar{y})$, $a \in P_2(\bar{y})$; $F(a, x_j) = F(a^*, x_j)$, $j = k+2, \cdots, n$, $a \in P_1(\bar{y})$, $a \in P_2(\bar{y})$; $F(a, x_{k+1}) > F(a^*, x_{k+1})$, $a \in P_1(\bar{y})$; $F(a, x_{k+1}) < F(a^*, x_{k+1})$, $a \in P_2(\bar{y})$. Let

$$\delta_1(\bar{y}) = \sup_{a \in P_1(\bar{y})} (F(a, x_{k+1}) - F(a^*, x_{k+1})),$$

$$\delta_2(\bar{y}) = \sup_{a \in P_2(\bar{y})} (F(a^*, x_{k+1}) - F(a, x_{k+1})).$$

It follows from Lemma 11 that $P_1(\bar{y})$ and $P_2(\bar{y})$ are nonempty and that $\delta_1(\bar{y}) > 0$, $\delta_2(\bar{y}) > 0$.

It will now be shown that for any $\bar{y}_0 \in A$ there is a neighborhood, $N(\bar{y}_0)$, of \bar{y}_0 such that $\bar{y} \in N(\bar{y}_0)$ implies $\delta_1(\bar{y}) \geq \delta_1(\bar{y}_0)/2$. There is an $a_1 \in P_1(\bar{y}_0)$ such



that $F(a_1, x_{k+1}) \geq F(a^*, x_{k+1}) + \delta_1(\bar{y}_0)/2$. Hence by the induction statement applied to $F(a_1, x)$ there is a $\delta_0(a_1, \eta, k, x_j) > 0$ such that for every \bar{y} with $|y^j - y_0^j| < \delta_0$ there is an $F(a, x)$ so that (i) $F(a, x_j) = y^j, j = 1, 2, \dots, k$, (ii) $F(a, x_j) = F(a_1, x_j) = F(a^*, x_j), j = k+2, \dots, n$, (iii) $F(a, x_{k+1}) = F(a_1, x_{k+1}) \geq F(a^*, x_{k+1}) + \delta_1(\bar{y}_0)/2$. Further $|F(a, x) - F(a_1, x)| < \eta$ and for η sufficiently small $|F(a, x) - F(a^*, x)| < \epsilon$. Hence δ_0 defines a neighborhood, $N(\bar{y}_0)$ so that for every point \bar{y} of $N(\bar{y}_0)$ there is an $a \in P_1(\bar{y})$ with $F(a, x_{k+1}) \geq F(a^*, x_{k+1}) + \delta_1(\bar{y}_0)/2$.

It will now be shown that $\delta_1(\bar{y})$ is bounded away from zero for $\bar{y} \in A$. A is a compact set which has an open covering by neighborhoods, $\{N(\bar{y}) | \bar{y} \in A\}$. Hence A has a finite covering $N(\bar{y}_1), \dots, N(\bar{y}_m)$ and

$$\delta_1(\bar{y}) \geq 2^{-1} \min [\delta_1(\bar{y}_1), \delta_1(\bar{y}_2), \dots, \delta_1(\bar{y}_m)] = \delta_1 > 0.$$

A similar argument shows that $\delta_2(\bar{y}) \geq \delta_2 > 0$.

The general induction step can now be made with the preceding construction and Lemma 12. For $k > 1$ take

$$\delta(a^*, \epsilon, k+1, x_j) = \min [\delta_1/2, \delta_2/2, \delta(a^*, \epsilon, k, x_j)].$$

Given any set $\{y^j | j = 1, 2, \dots, k+1, |F(a^*, x_j) - y^j| \leq \delta(a^*, \epsilon, k+1, x_j)\}$ then there are $a_1, a_2 \in P$ such that (i) $F(a_1, x_j) = F(a_2, x_j) = y^j, j = 1, 2, \dots, k$; (ii) $F(a_1, x_{k+1}) \geq F(a^*, x_{k+1}) + \delta(a^*, \epsilon, k+1, x_j)$, $F(a_2, x_{k+1}) \leq F(a^*, x_{k+1}) - \delta(a^*, \epsilon, k+1, x_j)$; (iii) $F(a_1, x_j) = F(a_2, x_j) = F(a^*, x_j), j = k+2, \dots, n$; (iv) $|F(a_1, x) - F(a^*, x)| < \epsilon, |F(a_2, x) - F(a^*, x)| < \epsilon$. By Lemma 12 there is an $a_3 \in P$ such that $F(a_3, x_{k+1}) = y^{k+1}$ and $F(a_3, x)$ is between $F(a_1, x)$ and $F(a_2, x)$. This concludes the induction step and it follows from the induction statement for $k = n$ that F is locally solvent.

LEMMA 14. *If F is closed and locally unisolvent then F has Property N.*

Proof. Assume that $\max |F(a^*, x) - f(x)|$ alternates exactly k times, $k < n$. It will be shown that $F(a^*, x)$ cannot be a best approximation to $f(x)$.

If $\max |F(a^*, x) - f(x)|$ is not assumed at both 0 and 1 then one of the intervals $[0, \delta]$, $[1 - \delta, 1]$, say $[0, \delta]$ for concreteness, is chosen with δ determined so that for $x \in [0, \delta]$ and some $\epsilon_0 > 0$, we have $|F(a^*, x) - f(x)| < \max |F(a^*, x) - f(x)| - \epsilon_0$. Let $0 = x_0 < x_{n-k} < x_{n-k+1} < \dots < x_n = 1$ divide $[0, 1]$ into $k+1$ subintervals so that (i) $F(a^*, x_j) - f(x_j) = 0$, $j = n-k, n-k+1, \dots, n-1$; (ii) $\max |F(a^*, x) - f(x)|$ alternates exactly once in any two adjacent subintervals but does not alternate in any one subinterval. Choose $n-k-1$ distinct points, $\{x_j | j=1, 2, \dots, n-k-1, x_j < x_{j+1}\}$, in $[0, \delta]$. Let $\alpha_j = \max (F(a^*, x) - f(x)) = \min (F(a^*, x) - f(x))$ with the maximum and minimum taken over $[x_j, x_{j+1}]$, $j = n-k-1, n-k, \dots, n-1$. Let $\eta_j = \max |F(a^*, x) - f(x)| - \alpha_j/2$, $j = n-k-1, n-k, \dots, n-1$ and $\epsilon = \min (\eta_{n-k-1}, \dots, \eta_{n-1}, \epsilon_0/2)$. Let x' denote a point where $\max |F(a^*, x) - f(x)|$ is attained.

Since F is locally solvent there is an $a \in P$ such that $F(a, x) - F(a^*, x)$ changes sign at x_j , $j = 1, 2, \dots, n-1$ and $|F(a, x') - f(x')| < |F(a^*, x') - f(x')|$ with $\max |F(a, x) - F(a^*, x)| < \epsilon$. Since $F(a, x) - F(a^*, x)$ cannot have any more zeros $F(a, x)$ is uniformly closer to $f(x)$ than $F(a^*, x)$ except in $[0, \delta]$. Therefore $F(a, x)$ is a better approximation than $F(a^*, x)$ to $f(x)$.

If $\max |F(a^*, x) - f(x)|$ is assumed at both 0 and 1 let the interval $[0, \delta]$ be replaced by $[x_{n-k} - \delta, x_{n-k}]$ where $\{x_j | j=0, n-k, \dots, n\}$, $\delta, \epsilon_0, \alpha_j, \eta_j$ and ϵ are defined in a manner analogous to the preceding construction. For $n-k-1$ even choose $n-k-1$ points

$$\{x_j | x_{n-k} - \delta \leq x_1 < x_2 < \dots < x_{n-k-1} < x_{n-k}\}$$

and determine $a \in P$ so that $|F(a, 0) - f(0)| < |F(a^*, 0) - f(0)|$, $|F(a, 1) - f(1)| < |F(a^*, 1) - f(1)|$, $|F(a, x) - F(a^*, x)| < \epsilon$ and $F(a, x) - F(a^*, x)$ changes sign at the x_j , $j = 1, 2, \dots, n-1$. $F(a, x)$ is a better approximation than $F(a^*, x)$ to $f(x)$. For $n-k-2$ even choose $n-k-2$ points

$$\{x_j | x_{n-k} - \delta \leq x_2 < x_3 < \dots < x_{n-k-1} < x_{n-k}\}$$

and determine a so that $|F(a, 0) - f(0)| < |F(a^*, 0) - f(0)|$, $|F(a, 1) - f(1)| < |F(a^*, 1) - f(1)|$, $|F(a, x) - F(a^*, x)| < \epsilon$ for $x \in [0, 1]$ and $F(a, x) - F(a^*, x)$ changes sign at the x_j , $j = 2, 3, \dots, n-1$. $F(a, x) - F(a^*, x)$ cannot have another zero without having two more, counting multiplicity, and hence $F(a, x)$ is a better approximation than $F(a^*, x)$ to $f(x)$. This concludes the proof.

Proof of Theorem 3. The statement of Theorem 3 may be considered as a set of four statements, namely Lemmas 2, 3, 13 and 14, all of which have been established as true.

2.5. In Theorem 3 Property NS is characterized by mainly local properties of the function F . If a stronger closure condition is imposed on F then Property NS can be characterized by "global" properties of F . This stronger condition is satisfied by the common forms of F .

DEFINITION 11. F is said to be solvent if F is locally solvent and if given a set

$\{(x_j, y_j) | 0 \leq x_1 < x_2 < \cdots < x_n \leq 1, (x_j, y_j) \in R\}$ then there is a solution, $a \in P$, to the equations $F(a, x_j) = y_j, j = 1, 2, \cdots, n$.

DEFINITION 12. F is said to be unisolvent if F is solvent and has Property Z.

DEFINITION 13. F is n -point closed if P is connected and if F is closed under pointwise convergence at any n points, i.e., $\lim_{k \rightarrow \infty} F(a_k, x_j) = y_j, j = 1, 2, \cdots, n$ with $(x_j, y_j) \in R$ implies the existence of an $a_0 \in P$ such that $F(a_0, x_j) = y_j$.

It is seen that if F is n -point closed then F is closed.

THEOREM 4. If F is n -point closed then unisolvence is a necessary and sufficient condition for F to have Property NS.

Lemmas 2, 3 and 14 still apply here as in the proof of Theorem 3 and only Lemma 13 needs to be replaced.

LEMMA 15. If F is n -point closed and has Property NS then F is solvent.

PROOF. Let $\{x_j | x_j \in [0, 1], j = 1, 2, \cdots, n\}$ be an arbitrary set of n points which shall remain fixed through this proof. Relative to this set define $U = \{y | y \in E_n, y^j = F(a, x_j), a \in P\}$. Recall the definition of $L_0(x)$ in the discussion preceding Lemma 8. Let $L_\epsilon(x)$ be the closed interval obtained by shortening $L_0(x)$ by ϵ at each end. Let $W_\epsilon = \{y | y \in E_n, y^j \in L_\epsilon(x_j)\}$. W_ϵ is seen to be an n -dimensional parallelepiped. It will be shown that $W_\epsilon \cap U = W_\epsilon$ for all ϵ sufficiently small.

It follows from the n -point closure of F that $W_\epsilon \cap U$ is a closed set in E_n .

Since n -point closure implies closure it follows from Theorem 3 that F is locally unisolvent. Let y_0 be any point of $W_\epsilon \cap U$. Since F is locally solvent it follows that there is a $\delta(y_0, x_j) > 0$ defining an n -dimensional cube about y_0 which is in U . Hence $W_\epsilon \cap U$ is an open set relative to W_ϵ .

Since $R = R_0$ is nonempty it is clear that for some $\epsilon_0 > 0$, $W_{\epsilon_0} \cap U$ is nonempty. Therefore $W_{\epsilon_0} \cap U$ is a nonempty set which is both open and closed relative to W_{ϵ_0} and $W_{\epsilon_0} \cap U = W_{\epsilon_0}$. It is clear that for $\epsilon < \epsilon_0$, $W_\epsilon \cap U$ is nonempty and hence $W_\epsilon \cap U = W_\epsilon$ for $\epsilon \leq \epsilon_0$.

Since $R = R_0$, any set $\{y^j | y^j \in L_0(x_j), j = 1, 2, \cdots, n\}$ is represented by a point in W_ϵ for some $\epsilon \leq \epsilon_0$. Therefore there is an $a \in P$ so that $F(a, x_j) = y^j, j = 1, 2, \cdots, n$, for any set $\{y^j | (x_j, y^j) \in R, j = 1, 2, \cdots, n\}$.

3.1. Often when one is using nonlinear approximating functions it develops that Property NS is lacking. Yet the characterization of best approximations is nevertheless very closely related to Property NS. Three examples are given below which illustrate this situation. The aim of this part of this paper is to generalize unisolvence and related concepts and to consider the characterization of best approximations with these generalizations.

3.2. The following examples illustrate different forms of F which do not have Property NS but for which a simple theorem characterizing best approximations can be given.

EXAMPLE 1. *Polynomial Ratios.* F is of the form $s(x)(\sum_{k=0}^n a^k x^k) \cdot (\sum_{k=n+1}^{n+m+2} a^k x^{k-n-1})^{-1}$ where $s(x) > 0$ and is continuous on $[0, 1]$. Since the numerator and denominator may be multiplied by a common factor it is assumed that $\sum_{k=1}^{n+m+2} (a^k)^2 = 1$. P is, then, taken to be the subset of the unit sphere of E_{n+m+2} for which $\sum_{k=n+1}^{n+m+2} a^k x^{k-n-1} \neq 0$ for $x \in [0, 1]$.

The classical theorem characterizing best approximations is [1, Chapter 2].

" $F(a, x) = s(x)(a^{n-\alpha} x^{n-\alpha} + \dots + a^0)/(b^{m-\beta} x^{m-\beta} + \dots + b^0)$ where $0 \leq \alpha \leq n$, $0 \leq \beta \leq m$, $b^{m-\beta} \neq 0$, $a^{n-\alpha} \neq 0$, $b^0 \neq 0$ is the best approximation to $f(x)$ if and only if $\max |F(a, x) - f(x)|$ alternates at least $m+n+2-d$ times where $d = \min(\alpha, \beta)$; if $F(a, x) = 0$ then $d = m$."

At points of P where both the numerator and denominator are of lower degree there is an effective loss of parameters. This can be due either to the cancellation of common factors or to the coefficients of the leading terms being zero. It can be shown that F is not locally solvent near such points and therefore $\max |F(a, x) - f(x)|$ cannot alternate $n+m+2$ times for all $f(x)$.

EXAMPLE 2. *An Exponential Function.* Let F be of the form $F(a, x) = a^1 e^{a^2 x} + a^3$ and let P be defined by $|a^1| < \infty$, $|a^2| < \infty$, $|a^3| < \infty$. When $a^2 = 0$ it is seen that F depends only on $a^1 + a^3$ and F is not locally solvent near $F \equiv \text{constant}$.

EXAMPLE 3. *Restricted Polynomial.* Let F be a polynomial and P be taken as a proper subset of E_n . For example, let $F(a, x) = a^1 + a^2 x + a^3 x^2$ with $|a^1| < \infty$, $|a^2| \leq 10$, $|a^3| \leq 1$. From Theorem 6 it follows that F does not have Property NS. F fails to be locally solvent on the boundaries of P .

3.3. In each of these examples F fails to be locally solvent at some exceptional points of P . However, it is seen that there is still some degree of local solvency at these points. This leads one to the concept of degree of solvency given in the following definition:

DEFINITION 14. F is solvent of degree m at $a^* \in P$ if given a set $\{x_j | 0 \leq x_1 < x_2 < \dots < x_m \leq 1\}$ and $\epsilon > 0$ then there is a $\delta(a^*, \epsilon, x_1, \dots, x_m) > 0$ such that $|y_j - F(a^*, x_j)| < \delta$ implies that there is a solution, $a \in P$, to $F(a, x_j) = y_j$, $j = 1, 2, \dots, m$ with $\max |F(a, x) - F(a^*, x)| < \epsilon$.

Thus, the degree of solvency may vary from point to point in P . The definitions of Properties NS and Z may be generalized in an analogous manner.

DEFINITION 15. F has Property NS of degree m at $a^* \in P$ if, for every $f(x)$, the alternance m times of $\max |F(a^*, x) - f(x)|$ is a necessary and sufficient condition for $F(a^*, x)$ to be a best approximation to $f(x)$. A point, $a^* \in P$, is said to be of degree m if F has Property NS of degree $m > 0$ at a^* .

DEFINITION 16. F has Property Z of degree m at $a^* \in P$ if for any $a \neq a^*$, $F(a^*, x) - F(a, x)$ has at most $m-1$ zeros.

DEFINITION 17. F is unsolvent of degree m at $a^* \in P$ if (i) F has Property Z of degree m at a^* and F is solvent of degree m at a^* ; (ii) F is not solvent of degree $m+1$ at a^* .

The analog of the classical statement characterizing best approximations is: *Let a^* be of degree m , then $F(a^*, x)$ is a best approximation to $f(x)$ if and only if $\max |F(a^*, x) - f(x)|$ alternates m times.*

3.4. The aim of the following paragraphs is to relate the degree of a point in P and the degree of unsolvence of F at that point.

LEMMA 16. *If P is connected and if the degree is defined for every point of P then given $a^* \in P$ there is an $\epsilon > 0$ such that $(x, F(a^*, x) \pm \epsilon) \in R$ for all $x \in [0, 1]$.*

LEMMA 17. *If P is connected and if F is solvent of positive degree at every point of P then $R = R_0$.*

These two lemmas may be proved by a slight modification of the proofs of Lemmas 8 and 9.

LEMMA 18. *Let F be closed and let F be unsolvent of positive degree at every point of P . If F is unsolvent of degree m at a^* then the degree at a^* is m .*

This lemma may be proved by a slight modification of the proofs of Lemmas 2 and 14.

The converse of Lemma 18 is more difficult. A series of lemmas will be established which culminate in Lemma 22 which is the converse of Lemma 18. The following lemma is a generalization of Lemma 3.

LEMMA 19. *Let the degree be defined for every point of P . If a^* is of degree m then F has Property Z of degree m at a^* .*

Proof. Assume there is an $a \in P$ such that $F(a, x) - F(a^*, x)$ has m or more zeros. Then as in the proof of Lemma 3 there is a continuous function, $f_1(x)$, such that $\max |F(a^*, x) - f_1(x)|$ alternates exactly m times, $\max |F(a, x) - f_1(x)|$ alternates exactly $m - 2$ times and $\max |F(a^*, x) - f_1(x)| = \max |F(a, x) - f_1(x)|$. Let $f_2(x)$ be the continuous function analogously defined so that $\max |F(a^*, x) - f_2(x)|$ alternates exactly $m - 2$ times, $\max |F(a, x) - f_2(x)|$ alternates exactly m times and $\max |F(a^*, x) - f_2(x)| = \max |F(a, x) - f_2(x)|$.

Since a^* is of degree m , $F(a^*, x)$ is a best approximation to $f_1(x)$ and hence $F(a, x)$ is also a best approximation to $f_1(x)$. Therefore the degree of a is less than or equal to $m - 2$. Hence $F(a, x)$ is a best approximation to $f_2(x)$ and so is $F(a^*, x)$. But since $\max |F(a^*, x) - f_2(x)|$ alternates exactly $m - 2$ times a^* cannot be of degree m .

The following lemma is the generalization of Lemma 11 and has the same proof as Lemma 11.

LEMMA 20. *Let F be closed and let the degree be defined for every point of P . Then given $a^* \in P$, $\epsilon > 0$ and a set $\{x_j | 0 = x_0 \leq x_1 < x_2 < \dots < x_{n+1} = 1\}$ there are $a_1, a_2 \in P$ such that (i) $F(a^*, x) - F(a_1, x)$, $F(a^*, x) - F(a_2, x)$ change sign at the x_j and at no other points, (ii) $|F(a_1, x) - F(a^*, x)| \leq \epsilon$, $|F(a_2, x) - F(a^*, x)|$*

$\leq \epsilon$, (iii) either $F(a_1, x) - F(a^*, x) = 0$, $F(a_2, x) - F(a^*, x) = 0$ or $\text{sgn} [F(a_1, x) - F(a^*, x)] = -\text{sgn} [F(a_2, x) - F(a^*, x)]$.

The next lemma shows an interesting property of some special subsets of P and provides a means of generalizing Lemmas 12 and 13.

LEMMA 21. *Let F be closed and let the degree be defined for every point of P . Define $P_k = \{a \mid a \in P, \text{degree of } a \geq k\}$. Then P_k is an open set.*

Proof. Let $a^* \in P_k$ be given. By Lemma 20 there is an $a_1 \in P$ such that $F(a^*, x) - F(a_1, x)$ has $k-1$ zeros in the interior of $[0, 1]$. Let

$$\{x_j \mid j = 1, 2, \dots, k-1, x_j < x_{j+1}\}$$

be the set of these zeros and set $x_0 = 0$, $x_k = 1$. Let

$$\delta = \min_j \max_{x \in [x_j, x_{j+1}]} |F(a^*, x) - F(a_1, x)|, \quad j = 0, 1, 2, \dots, k-1.$$

It is clear that for any $a \in P$ if $\max |F(a, x) - F(a^*, x)| < \delta$ then $F(a, x) - F(a_1, x)$ has at least $k-1$ zeros.

Since F is continuous there is a neighborhood of a^* , $N(a^*)$, such that if $a \in N(a^*)$ then $\max |F(a, x) - F(a^*, x)| < \delta$ and $F(a, x) - F(a_1, x)$ has at least $k-1$ zeros. By Lemma 19 the degree of a is larger than or equal to k .

LEMMA 22. *Let F be closed and let the degree be defined for every point of P . If a^* is of degree m then F is unisolvent of degree m at a^* .*

Proof. Let δ be the quantity defined in the proof of Lemma 21 and set $\Delta = \{(x, y) \mid x \in [0, 1], |F(a^*, x) - y| < \delta\}$. From the proof of Lemma 21 it is seen that every $F(a, x)$ entirely in Δ is of degree m or more. Lemma 22 is now proved by restricting consideration to functions entirely in Δ and by using the proofs of Lemmas 12 and 13.

The statement of the generalization of Lemma 12 is: "Let F be closed and let the degree be defined for every point in P . Let $a^* \in P$ be given with degree m and let $F(a^*, x) - F(a, x)$ change sign at $m-1$ points with $F(a, x_0) \neq F(a^*, x_0)$ and let $\max |F(a^*, x) - F(a, x)| \leq \delta$. Then given y between $F(a, x_0)$ and $F(a^*, x_0)$ there is an $a_0 \in P$ such that $F(a_0, x_0) = y$ and $F(a_0, x)$ is between $F(a, x)$ and $F(a^*, x)$." The only part of the statement that is not a straightforward generalization of Lemma 12 is the condition that $\max |F(a^*, x) - F(a, x)| \leq \delta$. This condition implies that a is of degree m or more and the proof of Lemma 12 may be used to prove this generalization of Lemma 12.

The generalization of Lemma 13 may be proved with a slight modification of the proof of Lemma 13. In the induction argument of the proof it is merely necessary to take $\epsilon < \delta$ and the same proof is valid for the generalization.

In view of Lemma 19, this concludes the proof of Lemma 22.

The preceding lemmas have established

THEOREM 5. *If F is closed then the following statements are equivalent:*

- (i) *F is unisolvent at every point of P and is unisolvent of degree m at a^* .*
- (ii) *The degree is defined for every point of P and the degree of a^* is m .*

3.5. Theorem 7 directly relates the characterization of best approximations to the degree of unisolvence of F . Since unisolvence is directly connected with the explicit definition of F this gives a natural means of studying the characterization of best approximations.

It is interesting to examine the variation of the degree for the three examples previously considered.

EXAMPLE 1. Let $P_k = \{a | a \in P, a^n = \dots a^{n-k} = a^{n+m} = \dots a^{n+m-k} = 0\}$ and $P' = \{a | a \in P, a^n = a^{n-1} = \dots = a^0 = 0\}$. If $a \in P_k$, $a \notin P_{k+1}$ then the degree of a is $n+m-k$ and if $a \in P'$ the degree of a is $n+1$. If a is in none of these sets then the degree of a is $n+m+1$.

EXAMPLE 2. Every point of P is of degree 3 except those for which $a^2 = 0$; these points are of degree 2. Note that F is not closed in this example for as $a^2 \rightarrow -\infty$ the limiting function is discontinuous. However, it can be shown that the theory developed here is still valid.

EXAMPLE 3. The interior points of P are of degree 3. The interior points of the two dimensional faces are of degree 2 and the points on the one dimensional edges are of degree 1.

In this example if P were defined by $|a^1| < \infty$, $|a^2| < 10$, $|a^3| < 1$ then F would have Property NS with $n=3$. However, best approximations would not exist for every continuous function as they do with the present definition of P .

REFERENCES

1. N. I. Achieser, *Theory of approximations*, New York, Ungar, 1956.
2. S. Bernstein, *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Paris, Gauthier-Villars, 1926.
3. A. Haar, *Die Minkowskische Geometrie und die Annäherung an stetige Functionen*, Math. Ann. vol. 78 (1918) pp. 244-311.
4. T. S. Motzkin, *Approximation by curves of a unisolvent family*, Bull. Amer. Math. Soc. vol. 55 (1949) pp. 789-793.
5. I. Sh. Pinsker and E. N. Novodvorskii, *On a process of equalization of maxima* (Russian), Uspehi Mat. Nauk. vol. 6 (1951) pp. 174-181.
6. Ch. de la Vallée Poussin, *Sur la méthode de l'approximation minimum*, Ann. Soc. Sci. Bruxelles ser. 1 vol. 35 (1911) pp. 1-16.
7. ———, *Leçons sur l'approximation des fonctions d'une variable réelle*, Paris, Gauthier-Villars, 1919.
8. L. Tornheim, *On n -parameter families of functions and associated convex functions*, Trans. Amer. Math. Soc. vol. 69 (1950) pp. 457-467.

CALIFORNIA INSTITUTE OF TECHNOLOGY,
PASADENA, CALIFORNIA