JOURNAL OF APPROXIMATION THEORY 41, 329–344 (1984)

An Algorithm for the Computation of Strict Approximations in Subspaces of Spline Functions

HANS STRAUSS

Institut für Angewandte Mathematik, Universität Erlangen-Nürnberg, 8520 Erlangen, West Germany

Communicated by G. Meinardus

Received March 15, 1983

An algorithm is developed which computes strict approximations in subspaces of spline functions of degree $m-1$ with k fixed knots. The strict approximation is a unique best Chebyshev approximation for a problem defined on a finite set which can be considered as the "best" of the best approximations. Moreover, a sequence of strict approximations defined on certain subsets of an interval I converges to a best approximation on I if $k \leqslant m$ and at least to a nearly best approximation on I if $k > m$.

INTRODUCTION

We consider the problem of approximating a given function f in $C(I)$ by subspaces of spline functions of degree $m-1$ with k fixed knots, with respect to the supremum norm. Schumaker [6] has observed that the idea of the classical Remez algorithm can also be used for spline subspaces. Recently, Nürnberger and Sommer [3] developed a Remez type algorithm for computing best spline approximations.

In this paper we determine strict approximations. Rice [4] defines the strict approximation as a particular unique best Chebyshev approximation for problems defined on a finite set. It can be considered as the "best" of the best Chebyshev approximations.

In order to compute best approximations on an interval I it is often useful to replace the interval by finite sets. Then the continuous problem is replaced by discrete problems. It seems to be natural to determine strict approximations on the finite subsets. The chief purpose of this paper is to develop an algorithm which computes the strict approximation in subspaces of spline functions. For this purpose a characterization of strict approximations which is established in [9] will be very important.

Then we define a Remez type algorithm. A sequence of strict approximations defined on certain finite subsets converges to a best approximation

0021-9045/84 $3.00

on I if $k \leqslant m$ and at least to a nearly best approximation which for most practical cases is a best approximation of f if $k > m$. In Section 1 we present some results concerning the characterization of strict approximations. In Section 2 we develop the algorithm which determines strict approximations. We shall show that it is possible to divide the approximation problem into subproblems where exchange rules can be applied as in the classical Remez algorithm for Haar subspaces. In Section 3 we state some results concerning the convergence of sequences of strict approximations. Finally, we shall give some numerical examples in Section 4.

1. Preliminaries

Let T be a compact subset of $\mathbb{R}$ and $C(T)$ be the normed linear space of all continuous real-valued functions defined on T. Let the space $C(T)$ be normed by

$$\|f\| = \max_{x \in T} |f(x)|.$$

Suppose that G is an n-dimensional subspace of $C(T)$. Then we denote the set of *best approximations* to a function f in $C(T)$ out of G by

$$P_G(f) := \{g_0 \in G : \|f - g_0\| = \inf\{\|f - g\| : g \in G\}.$$

We shall also consider approximation problems on a subset U of T. A function g_0 in G is called a *best approximation* to f on U if

$$\max_{x \in U} |f(x) - g_0(x)| = \inf_{g \in G} \{\max_{x \in U} |f(x) - g(x)|\}.$$

We use the following notations: We denote by $E(f)$ the set of *extremal points* of the function f on T

$$E(f) = \{x \in T : |f(x)| = \|f\|\}.$$

A function f is said to *alternate* on the points $t_1 < \cdots < t_h$ in T if $f(t_i) \cdot f(t_{i+1}) < 0$, $i = 1,..., h$, and we call points $t_1 < \cdots < t_h$ in T *alternating extremal points* of f if $\mu(-1)^i f(t_i) = \|f\|$, $i = 1,..., h$, $\mu \in \{-1, 1\}$.

A subset $R = \{u_i\}_{i=1}^{n+1}$ of T is called a *reference* if

$$A\begin{pmatrix} u_1,..., u_{n+1} \\ g_1,..., g_n \end{pmatrix}$$

has rank n, where $G = \operatorname{span}\{g_1,..., g_n\}$ and

$$A\begin{pmatrix} t_1,..., t_h \\ g_1,..., g_n \end{pmatrix} = \begin{pmatrix} g_1(t_1) & \cdots & g_n(t_1) \\ \vdots & & \vdots \\ g_1(t_h) & \cdots & g_n(t_h) \end{pmatrix},$$

$a \leqslant t_1 < \cdots < t_h \leqslant b$. Let g_0 be a best approximation to f on the reference R then $\gamma = \|(f - g_0)|_R\|$ is called the *reference deviation.*

A subspace G satisfies the *Haar condition* if $g \in G$, $g(x) = 0$ at n distinct points of T implies $g \equiv 0$. In this case the best approximation is always unique.

We shall consider approximation problems for subspaces of polynomial spline functions. These subspaces do not satisfy the Haar condition. Therefore the best approximations are not uniquely determined and it is natural to consider conditions which single out one of the best approximations.

Rice [5] defines a unique "strict approximation" for functions defined on a finite set. Strauss [9] has established a characterization theorem for the strict approximation to a function f out of a subspace of polynomial spline functions. This theorem will be very useful for developing an algorithm which determines the strict approximation. We shall establish such an algorithm.

First we need the following notations: Let Δ denote the partition $a = x_0 < x_1 < \cdots < x_k < x_{k+1} = b$ on the interval $[a, b]$. The subspace $S_m(\Delta)$ of *polynomial spline functions* of degree $m - 1$ $(m \geqslant 2)$ with simple fixed knots at Δ is defined by

$$S_m(\Delta) = \{s \in C^{m-2}[a, b] : s|_{[x_i, x_{i+1}]} \in \pi_{m-1}, i = 0,..., k\}$$

where π_{m-1} denotes all polynomials of degree $\leqslant m - 1$. We shall denote by I_{pq} a subinterval with boundary points x_p and x_q, where x_p and x_q are points of Δ.

Now we define the following subspaces: Let I be an interval satisfying $(x_0, x_{k+1}) \subset I \subset [x_0, x_{k+1}]$ and let T be a compact subset of I. Then

$$\begin{aligned} S_m(I) = \{s \in S_m(\Delta) :& \\ &\text{if } I = (x_0, x_{k+1}] \text{ then } s^{(i)}(x_0) = 0, i = 0,..., m-2, \\ &\text{if } I = [x_0, x_{k+1}) \text{ then } s^{(i)}(x_{k+1}) = 0, i = 0,..., m-2, \\ &\text{if } I = (x_0, x_{k+1}) \text{ then } s^{(i)}(x_0) = s^{(i)}(x_{k+1}) = 0, i = 0,..., m-2\}, \\ S_m(I, T) = \{s|_T : s \in S_m(I)\}.& \end{aligned}$$

Notice that the functions of $S_m(I)$ satisfy boundary conditions if the interval I is not closed.

A local basis of $S_m(I)$ will be very useful. Let the partition $\Delta = \{x_0, \dots, x_{k+1}\}$ be given. A partition $\tilde{\Delta} = \{x_{-m+1}, \dots, x_{m+k}\}$ *satisfying* $x_{-m+1} < \cdots < x_0 < \cdots < x_{k+1} < \cdots < x_{m+k}$ is called an *extended partition associated with* Δ. Suppose that M_i, $i = -m+1, \dots, k$ is the mth order B-spline associated with the knots $x_i, \dots, x_{i+m}$ (see [8, p. 118]). We also denote $M_i|_J$ by M_i, where J is a subset of $[x_{-m+1}, x_{m+k}]$.

Then $S_m(I) = \operatorname{span}\{M_{-m+1}, \dots, M_k\}$ if $I = [x_0, x_{k+1}]$, $S_m(I) = \operatorname{span}\{M_0, \dots, M_k\}$ if $I = (x_0, x_{k+1}]$ and $S_m(I) = \operatorname{span}\{M_0, \dots, M_{k-m-1}\}$ if $I = (x_0, x_{k+1})$.

Let the partition $\tilde{\Delta} = \{x_{-m+1}, \dots, x_n\}$, $n \geqslant 1$, be given and let $\tilde{I} = (x_{-m+1}, x_n)$. Suppose that $U = \{u_i\}_{i=1}^n$ is a subset of $\tilde{I}$. Then the matrix

$$A\begin{pmatrix} u_1 \cdots u_n \\ M_{-m+1} \cdots M_{n-m} \end{pmatrix}$$

has rank n if and only if

$$x_{-m+i} < u_i < x_i, \qquad i = 1, \dots, n. \tag{1.1}$$

Hence the subspace $S_m(\tilde{I}, T)$, $T \subset \tilde{I}$, has dimension n, iff there exists a subset $U \subset T$ satisfying (1.1).

Now we shall consider characterization theorems for best approximations to f (see [9]).

THEOREM 1.1. *Let the partition* $\tilde{\Delta} = \{x_i\}_{i=-m+1}^n$, $n \geqslant 1$, *be given and let* $\tilde{I} = (x_{-m+1}, x_n)$. *Suppose that* T *is a compact subset of* $\tilde{I}$ *such that* $\dim S_m(\tilde{I}, T) = n$.

(a) *Then* s_0 *in* $S_m(\tilde{I})$ *is a best approximation to a function* f *in* $C(T)$ *on* T *if and only if there exists a subinterval* J_R *and a subset* $R = \{u_i\}_{i=p}^{q+1} \subset T \cap J_R$ *such that* $(f - s_0)|_T$ *has alternating extremal points on* R *where* R *and* J_R *satisfy*

$$J_R = \begin{cases} (x_{-m+1}, x_n) & p = 1,\ q = n \\ [x_{p-1}, x_n) & p > 1,\ q = n \\ (x_{-m+1}, x_{-m+q+1}] & p = 1,\ q < n \\ [x_{p-1}, x_{-m+q+1}] & p > 1,\ q < n,\ q - p \geqslant m - 1 \end{cases} \quad \text{if}$$

$$u_i \in (x_{-m+i}, x_{i-1}), \qquad i = p+1, \dots, q. \tag{1.2}$$

(b) *The best approximations are uniquely determined on* J_R.

Remark. The subspace $S_m(I)$ is spanned by B-splines. But it is possible to derive characterization theorems for all kinds of boundary conditions from Theorem 1.1 (for details, see [9]). For example, if we set $n = m + k$

and $T = [x_0, x_{k+1}]$ we obtain the problem which was considered by Rice [5] and Schumaker [6] that is an approximation problem defined on $[x_0, x_{k+1}]$.

The best approximations in Theorem 1.1 are not unique in general. Therefore Rice considered the so-called *strict approximations* which are particular unique best Chebyshev approximations for problems defined on finite sets. For a definition of these approximations (see [5, p. 239]). In this paper we shall only state a characterization of strict approximations for the problem in consideration.

The following definition will be very important: Let f be an element of $C(T)$, where T is a finite subset of $\mathbb{R}$ and let G be an n-dimensional subspace of $C(T)$. Suppose that g_0 is a best approximation from G to f on T. A subset S of the extremal points of $f - g_0$ is said to be a *critical point set* if g_0 is a best approximation to f on S but is not a best approximation to f on any proper subset of S. A critical point set contains at most $n+1$ points.

Now we shall consider the approximation problem concerning spline functions. A characterization of crtitical point sets is given in [9].

THEOREM 1.2. *Let the assumptions of Theorem* 1.1 *be given and let s_0 be a best approximation from $S_m(\tilde{I}, T)$ to a function f in $C(T)$. Then the following conditions are equivalent:*

(a) *The subset $R \subset T$ is a critical point set.*

(b) *The subset $R \subset T$ is associated with a subinterval J_R satisfying the properties* (1.2) *and $f - s_0$ has alternating extremal points on R.*

Moreover, the subspace $S_m(\tilde{I}, R)$ satisfies the Haar condition if $R \subset T$ is a critical point set and there exists a reference R_1 relative to $S_m(I)$ satisfying $R \subset R_1$.

The subsets R characterizing best approximations are critical point sets. We shall use the following notations.

Suppose R is a critical point set. Then there exists a subinterval J_R which is associated with R. We call this subinterval J_R *associated with the critical point set R*. Let s_1 be a best approximation to f on a reference $U \subset T$ then there exists a unique critical point set R (see [9]). We call the subinterval J_R to be *associated with the reference.*

Now we are able to state the characterization theorem for strict approximations (see [9]).

THEOREM 1.3. *Let the partition $\Delta = \{x_i\}_{i=0}^{k+1}$ be given, let $I = [x_0, x_{k+1}]$ and $S_m(I, T)$, where T is a finite subset such that* $\dim(S_m(I, T)) = m + k$. *Suppose that f is an element of $C(T)$ and s_0 of $S_m(I, T)$. Then the following properties are equivalent:*

(a) *The function s_0 is the strict approximation to f out of $S_m(I, T)$.*

(b) *There exists a partition of the interval* $x_0 = x_{v_0} < x_{v_1} < \cdots < x_{v_h} < x_{v_{h+1}} = x_{k+1}$ *such that the subintervals* $I_i = I_{v_{i-1}, v_i}$ *satisfy the following conditions:*

(i) $I = \bigcup_{i=1}^{h+1} I_i$, $I_i \cap I_{i+1} = \varnothing$ *for all* $i = 1,..., h$.

(ii) 0 *is the unique best approximation from* $S_m(I_i)$ *to* $(f - s_0)$ *on* T_i, *where* $T_i = T \cap I_i$ *for all* $i = 1,..., h+1$ *and there exists a critical point set* R_i *associated with* I_i *relative to* $S_m(I_i)$.

(iii) *Let* $\gamma_i := \max_{x \in T_i} |(f - s_0)(x)|$. *Then for all* $i = 1,..., h$ *the following conditions will hold: If* $x_{v_i} \in I_i$ *then* $\gamma_i \geqslant \gamma_{i+1}$ *and if* $x_{v_i} \notin I_i$ *then* $\gamma_i \leqslant \gamma_{i+1}$.

For a discussion of this theorem see [9].

Remark. The approximation problem is divided into subproblems and the best approximations of these problems are always unique. Moreover, the subspaces $S_m(I_i, R_i)$ satisfy the Haar condition. This property will be very important for our later investigations.

The strict approximation can be constructed by the following inductive definition (see [9]):

DEFINITION 1.4. Let $\Delta = \{x_i\}_{i=0}^{k+1}$ and $I = [x_0, x_{k+1}]$ be given. Suppose that T is a finite subset of I such that $\dim S_m(I, T) = m + k$. Set $G_0 = S_m(I, T) = \text{span}\{M_{-m+1},..., M_k\}$, $\tilde{I}_0 = \varnothing$ and let Z_0 be the set of integers $\{-m+1,..., k\}$. Then we define for $j \geqslant 1$ the following sequence of functions s_j: Let G_j be the set of best approximations to the function

$$f - (s_1 + \cdots + s_{j-1})(\text{i.e.}, f \text{ if } j = 1) \text{ on } T \cap \{I \backslash \tilde{I}_{j-1}\}$$

out of span $\{\{M_i\}_{i \in Z_{j-1}}\}$ and let s_j be a function in G_j.

Suppose that I_j is a subinterval of $I \backslash \tilde{I}_{j-1}$ which is associated with a critical point set of $f - (s_1 + \cdots + s_j)$. Then we define $\tilde{I}_j = \tilde{I}_{j-1} \cup I_j$ and

$$Z_j = \{i \in Z_{j-1} : \{x : M_i(x) \neq 0\} \cap \tilde{I}_j = \varnothing\}.$$

This construction is continued until $Z_t = \varnothing$ for some t. We denote by $s(f)$ the function $s_1 + \cdots + s_t$.

COROLLARY 1.5. *It can be shown that* $s(f)$ *is the strict approximation and that* $\{I_i\}$ *is a partition of* I *satisfying the properties of Theorem* 1.4.

It is obvious that the subintervals I_i of Corollary 1.5 are not in natural order, in general.

2. The Algorithm

In this section we want to develop an algorithm which determines the strict approximation of our problem. The characterization theorem for strict approximations will be very important for this purpose.

The subspaces of polynomial spline functions with fixed knots do not satisfy the Haar condition. But we can show that it is possible to divide this problem into subproblems which satisfy the Haar condition on certain subsets. First we shall consider these subproblems.

Problem I. Let the partition $\Delta = \{x_i\}_{i=0}^{k+1}$ be given, let I be an interval such that $(x_0, x_{k+1}) \subset I \subset [x_0, x_{k+1}]$. Suppose that T is a finite subset of I satisfying $\dim(S_m(I, T)) = \dim(S_m(I)) = n$. Then we determine the best approximations to a function f in $C(T)$ out of $G = S_m(I, T)$.

First we state the following theorem which is similar to the case where the Haar condition is satisfied.

Theorem 2.1. *Let Problem* I *be given. Suppose that the subset $R_0 = \{u_i\}_{i=1}^{n+1}$ is a reference in T. Let s_0 denote a best approximation from $S_m(I, T)$ to f on R_0 with reference deviation γ_0 and let J_{R_0} be the unique subinterval associated with R_0.*

If there exists a point $t \in T \cap J_{R_0}$ such that $|(f - s_0)(t)| > \gamma_0$ then it is possible to exchange a point $\bar{t}$ of $R_0 \cap J_{R_0}$ to form a new reference $R_1 = \{R_0 \backslash \{\bar{t}\}\} \cup \{t\}$ satisfying $\gamma_1 > \gamma_0$, where γ_1 is the reference deviation on R_1. If J_{R_1} is the subinterval associated with R_1 then $J_{R_1} \subset J_{R_0}$.

Proof. First we shall consider the approximation problem defined on $\bar{R}_0 = R_0 \cap J_{R_0} = \{u_i\}_{i=p}^{q+1}$.

The subspace $S_m(I, \bar{R}_0)$ satisfies the Haar condition, i.e., the matrix

$$A = A \begin{pmatrix} u_p, ..., u_{q+1} \\ M_{-m+p}, ..., M_{-m+q} \end{pmatrix}$$

satisfies the Haar condition. Hence there exists a unique vector $\lambda \in \mathbb{R}_{q-p+2}$ (up to a scalar) satisfying

$$\lambda^T A = 0, \qquad \lambda_i \neq 0, \qquad i = 1, ..., q - p + 2.$$

In order to show the assertions of the theorem we can use the proof in [11, p. 38].

Following the lines of this proof it can be shown that there exists a point $\bar{t} \in \bar{R}_0$ such that $\bar{R}_1 = \{\bar{R}_0 \backslash \{\bar{t}\}\} \cup \{t\}$ is a reference relative to $S_m(I, T \cap J_{R_0})$ and the reference deviation γ_1 on R_1 satisfies $\gamma_1 > \gamma_0$.

The subspace $S_m(I, \bar{R}_1)$ does not satisfy the Haar condition in general. Therefore the subinterval $J_{\bar{R}_1}$ associated with the reference $\bar{R}_1$ relative to $S_m(I, T \cap J_{R_0})$ may be a proper subset of $J_{R_0} = J_{\bar{R}_0}$.

It can be easily seen that $R_1 = \bar{R}_1 \cup \{R_0 \cap \{I \backslash J_{R_0}\}\}$ is a reference relative to $S_m(I, T)$, $J_{R_1} = J_{\bar{R}_1}$ is the subinterval associated with this reference and γ_1 is the reference deviation.

This completes the proof of the theorem.

COROLLARY 2.2. *Let the assumptions of Theorem* 2.1 *be given. Then the exchange of the point is uniquely determined by the following rules:*

Let $\sigma_i = \operatorname{sign}((f - s_0)(u_i))$ *and* $\bar{\sigma} = \operatorname{sign}((f - s_0)(t))$.

(a) *Let* $t \in (u_j, u_{j+1})$, *where* $u_p \leqslant u_j < u_{j+1} \leqslant u_{q-1}$. *Then* $\bar{t} = u_j$ *if* $\sigma_j \bar{\sigma} > 0$ *elsewhere* $\bar{t} = u_{j+1}$.

(b) *Let* $t < u_p$ *then* $\bar{t} = u_p$ *if* $\sigma_p \bar{\sigma} > 0$ *elsewhere* $\bar{t} = u_{q+1}$.

(c) *Let* $t > u_{q+1}$ *then* $\bar{t} = u_{q+1}$ *if* $\sigma_{q+1} \bar{\sigma} > 0$ *elsewhere* $\bar{t} = u_p$.

Proof. Since the matrix

$$A = A \begin{pmatrix} u_p, \dots, u_{q+1} \\ M_{-m+p}, \dots, M_{m+q} \end{pmatrix}$$

satisfies the Haar condition we can apply the well-known exchange rules to this case.

THEOREM 2.3. *Let Problem* I *be given. Suppose that R is a reference in T and J_R is the subinterval associated with R.*

Then a repeated application of the exchange rules of Theorem 2.1 *yields a reference $\bar{R}$ associated with a subinterval $J_{\bar{R}}$ such that the following conditions hold:*

(a) $J_{\bar{R}} \subset J_R$.

(b) *If $\bar{s}$ is a best approximation to f on $\bar{R}$ then $\bar{s}$ is also a best approximation on $T \cap J_{\bar{R}}$ and $J_{\bar{R}}$ is associated with the critical point set $\bar{R} \cap J_{\bar{R}}$ of $f - \bar{s}$ relative to $S_m(J_{\bar{R}})$.*

Proof. We apply Theorem 2.1 to the reference $R_0 = R$ and obtain a sequence of references $R_1, \dots, R_t$ where the subintervals J_{R_i} associated with R_i satisfy $J_{R_0} \supset J_{R_1} \supset \cdots \supset J_{R_t}$. Let s_i be a best approximation on R_i with reference deviation γ_i. The process stops after a finite number of steps since T is a finite set and $\gamma_i < \gamma_{i+1}$.

Hence there exists an integer t such that $\|(f - s_t)_{T \cap J_{R_t}}\| = \gamma_t$ and we have $\bar{R} = R_t$.

Using this theorem we are able to determine a function $\bar{s}$ in $S_m(I, T)$ and a subinterval $J_{\bar{R}}$ such that $\bar{s}$ is a best approximation to f on $J_{\bar{R}}$. But $\bar{s}$ is not a best approximation on I in general.

Now we want to define functions by an inductive definition which is similar to Definition 1.4.

DEFINITION 2.4. Let Problem I be given where $I=[x_0, x_{k+1}]$. Let $\tilde{I}_0=\emptyset$ and let Z_0 be the set of integers $\{-m+1,..., k\}$. Then we define for $j \geqslant 1$ the following sequence of functions s_j in $G_{j-1}=\operatorname{span}\{\{M_i\}_{i \in Z_{j-1}}\}$. There exists a subinterval $I_j \subset I \backslash \tilde{I}_{j-1}$ such that s_j is a best approximation from G_{j-1} to $f-(s_1+\cdots+s_{j-1})$ (i.e., f if $j=1$) on the subset $T \cap I_j$ and I_j is associated with a critical point set R_j of $(f-(s_1+\cdots+s_j))|_{T \cap I_j}$ relative to G_{j-1}. Let $\gamma_j=\|(f-(s_1+\cdots+s_j))|_{I_j}\|$. Then we define $\tilde{I}_j=\tilde{I}_{j-1} \cup I_j$ and

$$Z_j=\{i \in Z_{j-1}: \{x: M_i(x) \neq 0\} \cap \tilde{I}_j=\emptyset\}.$$

The construction is continued until $Z_t=\emptyset$ for some t. Then we denote by $h(f)$ the function $s_1+\cdots+s_t$ and the set $\{(I_i, R_i, \gamma_i)\}_{i=1}^t$ corresponds to $h(f)$.

Remark. It follows from Theorem 2.3 that there exists functions s_j satisfying the properties of Definition 2.4.

DISCUSSION. In Definition 1.4 the function s_j is a best approximation to $f-(s_1+\cdots+s_{j-1})$ on the whole interval $I \backslash \tilde{I}_{j-1}$ while in Definition 2.4 the function s_j is only a best approximation on a subinterval of $I \backslash \tilde{I}_{j-1}$.

Hence the function $h(f)$ may not be a best approximation to f. But we have obtained a partition of I corresponding to $h(f)$. Now we shall use this partition in order to determine by an algorithm a partition of I which corresponds to the strict approximation. This partition must have the following properties.

THEOREM 2.5. *Let $h(f)$ and $\{(I_i, R_i, \gamma_i)\}_{i=1}^t$ be as in Definition 2.4. Suppose that the following properties are satisfied: For all subintervals I_ν, $I_\mu \subset \{I_i\}_{i=1}^t$ satisfying $I_\nu=I_{p_1,q_1}$, $I_\mu=I_{q_1,q_2}$ we have $\gamma_\nu \geqslant \gamma_\mu$ if $x_{q_1} \in I_\nu$ and $\gamma_\nu \leqslant \gamma_\mu$ if $x_{q_1} \in I_\mu$. Then $h(f)$ is the strict approximation.*

Proof. Let $I_{\nu_1},..., I_{\nu_t}$ be the natural ordering of the subintervals $\{I_i\}$ then $\{(I_{\nu_i}, R_{\nu_i}, \gamma_{\nu_i})\}_{i=1}^t$ is a partition satisfying the conditions of Theorem 1.3. Hence $h(f)$ is the strict approximation.

Now we want to determine a partition and a function satisfying the properties of this definition. The following exchange theorem will be very important.

THEOREM 2.6. *Let $h(f)$ and $\{(I_i, R_i, \gamma_i)\}_{i=1}^t$ be as in Definition 2.4.*

Suppose that I_μ and I_ν are two subintervals of $\{I_i\}_{i=1}^{t}$, where $I_\mu = [x_{p_1}, x_{q_1}]$ and $I_\nu = (x_{q_1}, x_{q_2}]$. Let $\gamma_\mu < \gamma_\nu$. Then there exists a subinterval $I_0 = [x_{p_2}, x_{q_2}] \subset \{I_\mu \cup I_\nu\}$ and a subset $R_0 \subset \{R_\mu \cup R_\nu\}$ such that R_0 is a reference relative to $S_m(I_0)$ and the reference deviation γ_0 to the function $f - h(f)$ on R_0 satisfies $\gamma_\mu < \gamma_0$.

A similar result is true if I_μ and I_ν are half-open and open subintervals, respectively.

Proof. Set $I_1 = I_\mu = [x_{p_1}, x_{q_1}]$, $I_2 = I_\nu = (x_{q_1}, x_{q_2}]$, $\gamma_1 = \gamma_\mu$, $\gamma_2 = \gamma_\nu$ *and* $h = h(f)$.

It follows from Theorem 1.1 that the critical point sets $R_1 = R_\mu = \{u_i\}_{i=p_1+1}^{m+q_1}$ and $R_2 = R_\nu = \{v_i\}_{i=m+q_1}^{m+q_2}$ satisfies the conditions

$$\begin{aligned} u_i &\in (x_{-m+i}, x_{i-1}), \qquad i = p_1 + 2, \ldots, m + q_1 - 1, \\ v_i &\in (x_{-m+i}, x_{i-1}), \qquad i = m + q_1 + 1, \ldots, m + q_2 - 1 \end{aligned} \tag{2.1}$$

and $(f - h)|_{I_i}$ has alternating extremal points on R_i for $i = 1, 2$. We define a subset $Y_1 = \{y_i\}_{i=p_1+1}^{m+q_2} \subset \{R_1 \cup R_2\}$ in the following way

$$y_i = \begin{cases} u_i, & \text{if } c > 0, \\ u_{i+1} & \text{elsewhere} \end{cases}$$

for all $i = p_1 + 1, \ldots, m + q_1 - 1$, where $c = (f - h)(u_{m+q_1})(f - h)(v_{m+q_1})$ and $y_i = v_j$, $i = m + q_1, \ldots, m + q_2$.

It follows immediately that

$$(f - h)(y_i)(f - h)(y_{i+1}) < 0, \qquad i = p_1 + 1, \ldots, m + q_2 - 1. \tag{2.2}$$

We have to distinguish the following cases:

(i) Let $x_{m+q_1-1} \leqslant v_{m+q_1}$ then $I_0 = [x_{m+q_1-1}, x_{q_2}]$ and $R_0 = \{v_i\}_{i=m+q_1}^{m+q_2}$, i.e., $p_2 = m + q_1 - 1$.

(ii) Let $v_{m+q_1} < x_{m+q_1-1}$ and $c > 0$. Then it follows from (2.1) that

$$y_i \in (x_{-m+i}, x_{i-1}), \qquad i = p_1 + 2, \ldots, m + q_2 - 1 \tag{2.3}$$

and we set $I_0 = [x_{p_1}, x_{q_2}]$ and $R_0 = Y_1$, i.e., $p_2 = p_1$.

(iii) Let $v_{m+q_1} < x_{m+q_1-1}$ and $c < 0$. Then

$$\begin{aligned} y_i &\in (x_{-m+i}, x_{i-1}), \qquad i = m + q_1, \ldots, m + q_2 - 1, \\ y_i &\in (x_{-m+i}, x_i), \qquad i = p_1 + 1, \ldots, m + q_1 - 1. \end{aligned}$$

Let p_2 be an integer in $p_1 \leqslant p_2 \leqslant q_1$ such that

$$y_{p_2+1} \geqslant x_{p_2}, \qquad y_i < x_{i-1}, \qquad i = p_2 + 2, \ldots, m + q_1 - 1.$$

Then we have

$$y_i \in (x_{-m+i}, x_{i-1}), \qquad i = p_2 + 2,..., m + q_2 - 1. \tag{2.4}$$

Set $I_0 = [x_{p_2}, x_{q_2}]$, $R_0 = \{y_i\}_{i=p_2+1}^{m+q_2}$.

It follows from (i), (ii) and (iii) that R_0 is a reference relative to $S_m(I_0)$. Let γ_0 be the reference deviation to $f-h$ on R_0.

The proof will be accomplished by showing that $\gamma_1 < \gamma_0$. This is obvious for case (i) since $\gamma_0 = \gamma_2 = \gamma_\nu > \gamma_\mu = \gamma_1$. Therefore we only must consider cases (ii) and (iii).

We conclude from (2.2) that $f-h$ alternates on R_0. It follows from (2.3) and (2.4) that R_0 is a subset which is associated with I_0 relative to $S_m(I_0)$. If s_1 is a best approximation from $S_m(I_0)$ to $f-h$ on R_0 then $(f-h)-s_1$ alternates on R_0 and γ_0 is the reference deviation.

Assume that $\gamma_0 < \gamma_1$. Then s_1 alternates on R_0. This is a contradiction because s_1 has at least $m + q_2 - p_2 - 1$ sign changes and $\dim(S_m(I_0, R_0)) = m + q_2 - p_2 - 1$.

Now we assume that $\gamma_1 = \gamma_0$. Since the best approximation $h|_{R_1}$ to f on R_1 is unique we obtain that $s_1 \equiv 0$ on $I_0 \cap I_1$. Hence $s_1|_{I_2}$ is an element of $S_m(I_2)$. Moreover, it follows from $\gamma_0 = \gamma_1 < \gamma_2$ that s_1 alternates on $R_0 \cap I_2$. Therefore s_1 has $q_2 - q_1$ sign changes on I_2 and $\dim(S_m(I_2)) = q_2 - q_1$. This contradiction completes the proof of the theorem.

Remark. The function $h(f)$ in Theorem 2.6 is a best approximation to f on $I_\mu = [x_{p_1}, x_{p_2}]$ but is not a best approximation to f on I since the error γ_ν of $f - h(f)$ on I_ν is greater than the error γ_μ on I_μ. Using the exchange rules of Theorem 2.6 we are able to determine another subinterval $I_\eta = [x_{p_2}, x_{q_2}]$ such that the reference deviation γ_η on a reference relative to $S_m(I, T \cap I_\eta)$ is greater than γ_μ. Since there is only a finite number of subintervals $I_{p,q}$ we are able to determine after a finite number of steps a subinterval which is associated with a critical point set of a best approximation. These results also enables us to develop an algorithm which computes the strict approximation.

First we shall summarize the exchange rules of Theorem 2.6 in the following result.

COROLLARY 2.7. (a) *Let the assumptions of Theorem* 2.6 *be given, Let* $R_\mu = \{u_i\}_{i=p_1+1}^{m+q_1}$ *and* $R_\nu = \{v_i\}_{i=m+q_1}^{m+q_2}$ *be critical point sets of* $f - h(f)$ *on* I_μ *and* I_ν, *respectively. Set* $c = (f-h(f))(u_{m+q_1})(f-h(f))(v_{m+q_1})$. *Then we define* $Y_1 = \{y_i\}_{i=p_1+1}^{m+q_2}$ *as*

$$y_i = \begin{cases} u_i \\ u_{i+1} \end{cases}, \quad \text{if} \quad \begin{matrix} c > 0 \\ c < 0 \end{matrix}, \quad i = p_1 + 1,..., m + q_1 - 1,$$

$$y_i = v_j, \qquad i = m + q_1,..., m + q_2.$$

Then $I_\eta = [x_{p_2}, x_{q_2}]$, $R_\eta = \{y_i\}_{i=p_2+1}^{m+q_2}$ *are associated, where* p_2 *is defined in the following way:*

(i) *Let* $x_{m+q_1-1} \leqslant v_{m+q_1}$ *then* $p_2 = m + q_1 - 1$.

(ii) *Let* $v_{m+q_1} < x_{m+q_1-1}$ *and* $c > 0$ *then* $p_2 = p_1$.

(iii) *Let* $v_{m+q_1} < x_{m+q_1-1}$ *and* $c < 0$. *Then* p_2 *is defined as the integer satisfying* $p_1 \leqslant p_2 \leqslant q_1$ *and*

$$y_{p_2+1} \geqslant x_{p_2}, \qquad y_1 < x_{i-1}, \qquad i = p_2 + 2,..., m + q_1 - 1.$$

(b) *Similar exchange rules hold if* I_μ *and* I_ν *are half-open and open subintervals, respectively, and if* $x_{q_1} \notin I_\mu$, $x_{q_1} \in I_\nu$.

We now proceed to a description of the algorithm.

ALGORITHM 2.8. *Let the partition* $\Delta = \{x_i\}_{i=0}^{k+1}$ *be given, let* $I = [x_0, x_{k+1}]$ *and let* T *be a finite subset of* I *such that* $\dim S_m(I, T) = \dim S_m(I)$:

Let $h(f)$ *be a function which is constructed according to Definition* 2.4.

Set $g_1 = h(f)$. *Let* $\{(I_i^1, R_i^1, \gamma_i^1)\}_{i=1}^{t_1}$ *be the partition corresponding to* g_1.

We define a sequence of functions g_j *for integers* $j \geqslant 1$ *as follows:*

(1) *If* g_j *satisfies the conditions of Theorem* 2.5 *then* g_j *is the strict approximation.*

(2) *If* g_j *does not satisfy these conditions then we define a function* g_{j+1} *as follows:*

Let $\{(I_i^j, R_i^j, \gamma_i^j)\}_{i=1}^{t_j}$ *be the partition corresponding to* g_j. *Then there exist subintervals* $I_\mu^j = I_{p_1,q_1}$ *and* $I_\nu^j = I_{q_1,q_2}$ *satisfying* $\gamma_\mu^j < \gamma_\nu^j$ *if* $x_{q_1} \in I_\mu^j$ *or* $\gamma_\mu^j > \gamma_\nu^j$ *if* $x_{q_1} \in I_\nu^j$. *Assume that* $x_{q_1} \in I_\mu^j$. *Then* $\mu < \nu$.

Now we determine a function g_{j+1} *and a set* $\{(I_i^{j+1}, R_i^{j+1}, \gamma_i^{j+1})\}_{i=1}^{t_{j+1}}$ *corresponding to* g_{j+1} *in the following way:*

(i) $I_i^{j+1} = I_i^j$, $R_i^{j+1} = R_i^j$, $\gamma_i^{j+1} = \gamma_i^j$, $i = 0,..., \mu - 1$.

(ii) *Let* $J_\mu^j = \bigcup_{i=1}^{\mu-1} I_i^j$. *Then* $\{I_\mu^j \cup I_\nu^j\} \subset \tilde{J}_\mu^j = I \backslash J_\mu^j$.

Using Corollary 2.7 *we determine a subinterval* I_η^j *of* $I_\mu^j \cup I_\nu^j$ *and a reference* R_η^j *corresponding to* $S_m(I_\eta^j)$. *Let* $\bar{R}_\eta^j$ *be a reference relative to* $S_m(\tilde{J}_\mu^j)$ *satisfying* $R_\eta^j = \bar{R}_\eta^j \cap I_\eta^j$. *Then we apply the construction of Definition* 2.4 *to the subspace* $S_m(\tilde{J}_\mu^j)$ *and the function* $(f - g_j)|_{\tilde{J}_\mu^j}$. *We obtain a function* $\tilde{g}_{j+1} \in S_m(\tilde{J}_\mu^j)$ *and a partition* $\{(I_i^{j+1}, R_i^{j+1}, \gamma_i^{j+1})\}_{i=\mu}^{t_{j+1}}$. *Now we define*

$$g_{j+1}(x) = \begin{cases} g_j(x), & x \in J_\mu^j, \\ \tilde{g}_{j+1}(x), & \textit{elsewhere}. \end{cases}$$

Then the partition $\{(I_i^{j+1}, R_i^{j+1}, \gamma_i^{j+1})\}_{i=1}^{t_{j+1}}$ *corresponds to* g_{j+1}. *The function* g_{j+1} *is similarly constructed if* $x_{q_1} \in I_\nu^j$.

This construction determines the strict approximation.

THEOREM 2.9. *Algorithm* 2.8 *determines after a finite number of steps the strict approximation.*

Proof. Let $\{g_j\}$ be the sequence of functions defined by Algorithm 2.8 and let $\{(I_i^j, R_i^j, \gamma_i^j)\}_{i=1}^{t_j}$ be the partitions corresponding to g_j. Hence γ_i^j is the deviation of $f-g_j$ on I_i^j for all $i=1,..., t_j$. It follows from the construction in Algorithm 2.8 that $\gamma_i^{j+1} = \gamma_i^j$, $i = 1,..., \mu-1$, and we conclude from Theorem 2.6 that $\gamma_\mu^{j+1} > \gamma_\mu^j$.

Hence the vector $\{\gamma_i^{j+1}\}_{i=1}^{t_{j+1}}$ is lexicographically greater than $\{\gamma_i^j\}_{i=1}^{t_j}$. On the other hand, there exists only a finite number of partitions. Therefore the algorithm determines after a finite number of steps the strict approximation.

Remark. It is not necessary to compute the partition on the whole interval at each iteration. If we obtain at the jth iteration two subintervals I_μ^j, I_ν^j satisfying the properties of (2) in Algorithm 2.8 then we may proceed to the $(j+1)$th iteration.

3. CHEBYSHEV APPROXIMATION FOR CONTINUOUS FUNCTIONS

Discrete approximation problems are closely related to problems defined on an interval I. In order to determine best approximations on I we define a sequence of strict approximations on finite subsets of I. Such a sequence is not convergent, in general. But the following results will hold.

First we define a *Remez type algorithm*.

ALGORITHM 3.1. *Given the partition* $\Delta: x_0 < x_1 \cdots < x_{k+1}$ *and the interval* $I = [x_0, x_{k+1}]$. *Let* f *be a function of* $C(I)$. *Suppose that* T_E *is a compact subset of* I *satisfying* $\dim S_m(I, T_E) = m+k$ *and* T_0 *is a finite subset of* T_E *such that* $\dim S_m(I, T_0) = m+k$.

At the ith *step is defined a finite subset* T_i *of* T_E *and* $s(f, T_i)$ *is a strict approximation from* $S_m(I)$ *on* T_i. *Let* $\{I_{ij}\}_{j=1}^{h_i+1}$ *be a partition of* I *corresponding to* $s(f, T_i)$. *Suppose that* y_{ij} *is a point of* $T_E \cap I_{ij}$ *such that*

$$|(f-s(f, T_i))(y_{ij})| \geqslant |(f-s(f, T_i))(x)| \quad \textit{for all} \quad x \in T_E \cap I_{ij},\ j=1,..., h_i+1.$$

Then T_{i+1} *is given by* $T_i \cup \{y_{ij}\}_{j=1}^{h_i+1}$.

We shall consider the following subsets:

Let

$$T_E = I \setminus \bigcup_{i=1}^{k} \{(x_i - \varepsilon_i, x_i) \cup (x_i, x_i + \varepsilon_i)\}$$

where $E = \{\varepsilon_1, ..., \varepsilon_k\}$, $0 \leqslant \varepsilon_i$ *for all* $i = 1, ..., k$ *and* ε_i *sufficiently small such that* $\dim S_m(I) = \dim S_m(I, T_E)$. *Then the following theorem can be similarly shown as a result in* [10].

THEOREM 3.2. *Let* T_E *be defined as follows:*

(a) $T_E = I$ *if* $k \leqslant m$ (*i.e.*, $\varepsilon_i = 0$, $i = 1, ..., k$).

(b) $\varepsilon_i > 0$, $i = 1, ..., k$ *if* $k > m$.

Then the sequence $s(f, T_i)$ *of Algorithm* 3.1 *converges to a function* $s(f, T_E)$ *in* $S_m(I)$.

Moreover, it is shown that the functions $s(f, T_E)$ can also be obtained as the limit of a sequence of strict approximations defined on discrete sets which "fill up" the interval.

Remarks. (1) Some improvement in the convergence rate can be obtained by modifying the method such that T_{i+1} contains all local maxima of $|f - s(f, T_i)|$ which are greater than $\|(f - s(f, T_i))|_{T_i \cap I_{ij}}\|$ on $T_E \cap I_{ij}$.

(2) There has been developed many algorithms under general assumptions which compute best approximations if the Haar condition is not satisfied. See, for example [1] and the references therein.

(3) In [2] the strict approximations are determined on finite sets. In this paper we consider a special subspace—the subspace of spline functions. Using the characterization for strict approximations we are able to develop an algorithm such that a sequence of strict approximations converges to a best approximation s_0 of f if $k \leqslant m$ and we have convergence to a nearly best approximation s_0 of f which for most practical cases is a best approximation if $k > m$.

The algorithm in [3] determines a sequence of functions which converges to a best approximation under the same assumptions. But the best approximations determined on the finite sets are no strict approximations in general.

4. EXAMPLES

The algorithm has been tested by R. Bärnreuther. Here we shall give some examples.

(1) Let the function $f(x) = \ln|1 + x|$ be defined on $I = [0, 2]$. Suppose that $S_{m,k}(I)$ has equally spaced knots. The entry in each box of the following table is the minimax error magnitude.

No. of knots	Degree of splines 3	5	7
1	0.000 813 809	0.000 036 842	0.000 001 944
2	0.000 255 263	0.000 007 444	0.000 000 270
3	0.000 101 573	0.000 002 022	0.000 000 125
4	0.000 048 111	0.000 000 993	0.000 000 055
5	0.000 025 511	0.000 000 512	0.000 000 023

All best approximations are unique. We obtain a partition of the interval I only in one iteration step of the approximation problem where $S_m(I)$ has degree 3 and 3 knots. Therefore we see that in most of the examples the algorithm works as the classical Remez algorithm for Haar subspaces.

(2) Let $f(x) = \sqrt{x}$ be defined on the interval $I = [0, 1]$. We are setting the knots near zero since f is not so well-behaved in a neighborhood of zero. The first entry in each box is the minimax error magnitude. Subsequent entries give the knot locations.

No. of knots	Degree of splines 3	5	7
1	0.012 238 0.039 5	0.009 115 0.039 5	0.007 544 0.039 5
2	0.004 940 0.006 45 0.107 6	0.003 691 0.006 45 0.107 6	0.002 120 0.006 45 0.107 6
3	0.003 119 0.002 0.02 0.15	0.002 112 0.002 0.02 0.15	0.001 748 0.002 0.02 0.15
4	0.002 386 0.001 5 0.02 0.1 0.3	0.001 785 0.001 5 0.02 0.1 0.3	
5	0.001 946 0.001 0.015 0.06 0.2 0.35		

If $m-1=3$ and $k=3$ the best approximation is not unique. We obtain a partition $[0, 0.02)$ and $[0.02, 1]$ with reference deviations 0.00283 and 0.00312, respectively. In the other cases the best approximations are unique.

(3) Let $f(x)=1/(1+x^2)$ and $I=[-5, 5]$ (Runges example). The entry in each box is the minimax error magnitude.

No. of knots	Degree of splines 3	5	7
2	0.2078	0.1432	0.0975
3	0.0329	0.5227	0.0449
4	0.1274	0.0912	0.0645
5	0.1174	0.0217	0.0225

The knots are chosen as follows: $\{-1, 1\}$ if $k=2$; $\{-1, 0, 1\}$ if $k=3$; $\{-2, -1, 1, 2\}$ if $k=4$; $\{-2, -1, 0, 1, 2\}$ if $k=5$.

REFERENCES

1. H. P. BLATT, U. KAISER, AND B. RUFFER–BEEDGEN, A multiple exchange algorithm in convex programming, *in* "Optimization: Theory and Algorithms Confolant 1981" (J.-B. Hirriart–Urruty, W. Oettli, and J. Stoer, Eds.), Lecture Notes in Pure and Applied Mathematics, Dekker, New York, in press.
2. C. S. DURIS AND M. G. TEMPLE, A finite step algorithm for determining the "strict" Chebyshev solution to $Ax=b$, *SIAM J. Numer. Anal.* **10** (1973), 690–699.
3. G. NÜRNBERGER AND M. SOMMER, A Remez type algorithm for spline functions, *Num. Math.* **41** (1983), 117–146.
4. J. R. RICE, Tchebyscheff approximlation in a compact metric space, *Bull. Amer. Math. Soc.* **68** (1962), 405–410.
5. J. R. RICE, "The Approximation of Functions," Vol. II, Addison–Wesley, Reading, Mass. 1969.
6. L. L. SCHUMAKER, Uniform approximation by Tschebycheffian spline functions, *J. Math. Mech.* **18** (1968), 369–378.
7. L. L. SCHUMAKER, Some algorithms for the computation of interpolating and approximating spline functions, *in* "Theory and Applications of Spline Functions" (T. N. E. Greville, Ed.), pp. 87–102, Academic Press, New York, 1969.
8. L. L. SCHUMAKER, "Spline Functions: Basic Theory," Wiley–Interscience, New York, 1981.
9. H. STRAUSS, Characterization of strict approximations in subspaces of spline functions, *J. Approx. Theory* **41** (1984), 297–308
10. H. STRAUSS, Chebyshev approximations in subspaces of spline functions, *Numer. Funct. Anal. and Optimiz.* **5**(4) (1982–1983), 421–448.
11. G. A. WATSON, "Approximation Theory and Numerical Methods," Wiley, New York, 1980.