JOURNAL OF APPROXIMATION THEORY 37, 125–136 (1983)

Simultaneous Approximation, Interpolation and Birkhoff Systems

R. A. LORENTZ

Gesellschaft für Mathematik und Datenverarbeitung mbH, Schloß Birlinghoven, Postfach 1240, 5205 St. Augustin 1, West Germany

Communicated by E. W. Cheney

Received August 20, 1981

By simultaneous approximation we mean the approximation of a function $f \in C^{(k_p)}[a, b]$ by elements of the finite-dimensional subspace $H \subset C^{(k_p)}[a, b]$ with respect to the semi-norm $\|\cdot\|_F$ given by

$$\|f\|_F = \max_{k_i \in F} \|f^{(k_i)}\|,$$

where $F = \{k_1, ..., k_p\}$ and the k_i are integers satisfying $0 \leqslant k_1 < k_2 < \cdots < k_p$. $\|\cdot\|$ is the uniform norm on $[a, b]$. This is thus the simultaneous approximation of a function and its derivatives.

The set $\Omega(f)$ of best simultaneous approximations to f, which is never empty, usually consists of more than one element. We want to determine its precise dimension. We would also like to know when the problem of approximating with $\|\cdot\|_F$ can be replaced by approximations with a simpler semi-norm. For example, if $F = \{0, 1\}$, when do we find best approximations to f with respect to $\|\cdot\|_F$ just by finding the best Chebycheff approximation to f' by elements h', $h \in H$ and then integrating. Both questions are answered here.

In [8], the case that H is the set of algebraic polynomials was dealt with. Keener [5] has given some first results for subspaces H satisfying $\dim H^{(i)} = n - i$, $i = 0, 1, ..., m$ (or $m - 1$) for the norm given by $F = \{0, 1, ..., m\}$. The spaces $H^{(i)}$, which is the space of ith derivatives of the elements of H, are all assumed to be Haar subspaces for $i \leqslant m$. He concludes that $\Omega^{(m)}(f)$, the set of mth derivatives of best simultaneous approximations, consists only of one element (or, under the weaker assumption that $\dim H^{(m)} = n - m + 1$, $\Omega^{(m)}(f)$ is at most two-dimensional).

The dimension of $\Omega(f)$ is, however, usually much smaller than this conclusion seems to indicate. Indeed, the best simultaneous approximation may be unique. By using the ideas of [5, 8], we give the precise dimension of

0021-9045/83/020125–12$03.00/0

$\Omega(f)$ and, at the same time, show that some of the assumptions of Keener's theorems may be omitted. We also show that if H is a Birkhoff system, a concept introduced by Lorentz [6], which satisfies the above dimensionality conditions, then $\Omega(f)$ enjoys the same properties as in the case that H is the set of algebraic polynomials.

The proofs depend heavily on existence theorems for Birkhoff interpolation from H; that is, the existence of an $h \in H$ whose values and derivatives interpolate given data at given knots.

The sets from which we will be approximating have a fairly concrete structure for, as Ikebe [3] pointed out, dim $H^{(i)} = n - i$ implies that $\Pi_{i-1} \subset H$ (Π_n is the set of algebraic polynomials of degree not exceeding n). Moreover, if dim $H^{(i)} = n - i$ and $H^{(i)}$ is a Haar space, then $H^{(j)}$ is a Haar space for all $0 \leqslant j \leqslant i$.

LEMMA 1. *Let H be an n-dimensional subspace of $C^{(m)}[a, b]$ such that* dim $H^{(m)} = n - m$ *and such that $H^{(m)}$ is a Haar subspace. Then* dim $H^{(i)} = n - i$ *and $H^{(i)}$ is a Haar subspace for $i = 0, 1, \dots, m$.*

Thus if H satisfies the above assumptions, it has a representation

$$h = P_{m-1} + \sum_{i=1}^{n-m} a_i h_i,$$

where $P_{m-1} \in \Pi_{m-1}$ and span $\{h_i^{(m)}\}$ is a Haar subspace.

Let

$$E_F(f; H) = \inf_{h \in H} \|f - h\|_F.$$

Then

$$\Omega(f) = \Omega_F(f, H) = \{h \in H \mid \|f - h\|_F = E_H(f)\}.$$

One easily sees that $\Omega(f)$ is convex and not empty. By dim $\Omega(f)$ we mean the dimension of the linear hull of the set $\{g - h \mid g \in \Omega(f)\}$ for some arbitrary but fixed $h \in \Omega(f)$.

The extremal sets $U_i(f, h)$ of an approximation to f are

$$U_i(f, h) = \{x \in [a, b] \mid |f^{(k_i)}(x) - h^{(k_i)}(x)| = \|f - h\|_F\}.$$

LEMMA 2. *There exists an $h \in \Omega(f)$ with the property that*

$$U_i(f, h) \subset U_i(f, g)$$

and

$$h^{(k_i)}(x) = g^{(k_i)}(x) \qquad \text{for} \quad x \in U_i(f, h)$$

for any $g \in \Omega(f)$ and all $i = 1, \dots, p$.

Any best simultaneous approximation with the above property will be called a minimal best approximation. Each point in the relative interior of $\Omega(f)$ is necessarily a minimal best approximation. The sets $U_i(f, h)$ are the same for each choice of the minimal best approximation h and we will denote these common sets by $U_i(f)$.

Let $q = \min\{k_i \mid U_i(f) \neq \varnothing\}$. q then depends on f. By $F - q$ we denote the set $\{k_i - q \mid i = 1,..., p;\ k_i - q \geqslant 0\}$. Since $0 \in F - q$, $\|\cdot\|_{F-q}$ is a norm. It is easy to see that if $h \in \Omega_F(f, H)$, then $h^{(q)} \in \Omega_{F-q}(f(q), H^{(q)})$. Moreover

LEMMA 3. *Let* $f \in C^{(k_p)}$ *and*

$$G = \{k_i \mid k_i \in F,\ U_i(f) \neq \varnothing\}.$$

Let

$$q = \min_{k_i \in G} k_i.$$

Then

$$E_G(f; H) = E_F(f; H),$$

$$\Omega_G(f) \supset \Omega_F(f)$$

and

$$\Omega_F^{(q)}(f, H) \subset \Omega_{G-q}(f^{(q)}, H^{(q)}).$$

Proof. Since $G \subset F$,

$$E_G(f; H) = \inf_{h \in H} \|f - H\|_G = \inf_{h \in H} \max_{k_i \in G} \|f^{(k_i)} - h^{(k_i)}\|$$

$$\leqslant \inf_{h \in H} \max_{k_i \in F} \|f^{(k_i)} - h^{(k_i)}\| = E_F(f; H).$$

If $h \in \Omega_F(f)$, then

$$\|f^{(k_i)} - h^{(k_i)}\| < E_F(f; H)$$

for all $k_i \in F \backslash G$. Therefore, for this h,

$$E_F(f; H) = \|f - h\|_F = \max_{k_i \in F} \|f^{(k_i)} - h^{(k_i)}\|$$

$$= \max_{k_i \in G} \|f^{(k_i)} - h^{(k_i)}\| = \|f - G\|.$$

Suppose that h does not belong to $\Omega_G(f)$. Then there is a $v \in H$ such that

$$\|f - v\|_G < \|f - h\|_G = E_F(f; H).$$

But then for all λ, $0 < \lambda < 1$, and for $k_i \in G$,

$$\|f^{(k_i)} - [(1-\lambda)h + \lambda v]^{(k_i)}\| < \|f^{(k_i)} - h^{(k_i)}\| = E_F(f; H).$$

Also, for all $\lambda > 0$ sufficiently small

$$\|f^{(k_i)} - [(1-\lambda)h + \lambda v]^{(k_i)}\| \leqslant (1-\lambda)\|f^{(k_i)} - h^{(k_i)}\| + \lambda \|f^{(k_i)} - v^{(k_i)}\| < E_F(f; H).$$

Thus

$$\|f - [(1-\lambda)h + \lambda v]\|_F < \|f - h\|_F$$

for all $\lambda > 0$ sufficiently small, a contradiction. It follows that $h \in \Omega_G(f)$ and therefore also that $E_G(f; H) = E_F(f; H)$.

To prove the last inclusion, we note that if $h \in \Omega_F(f; H)$, then there is no $v \in H$ with

$$\|f^{(k_i)} - v^{(k_i)}\| < \|f^{(k_i)} - h^{(k_i)}\| = E_F(f; H)$$

for all $k_i \in G$. Thus there is no $w \in H^{(q)}$ satisfying

$$\|[f^{(q)}]^{(k_i - q)} - w^{(k_i - q)}\| < E_F(f; H)$$

for all $k_i - q \in G - q$; i.e., $h^{(q)} \in \Omega_{G-q}(f^{(q)}; H^{(q)})$.

From the last part of this lemma it follows that if $\Omega_{G-q}(f^{(q)}, H^{(q)})$ consists of one element, then $h^{(q)}$, for any $h \in \Omega_F(f, H)$, is this element, i.e., if the best simultaneous approximation to $f^{(q)}$ from $H^{(q)}$ with respect to $\|\cdot\|_{G-q}$ is unique, then it is the qth derivative of any best simultaneous approximation to f from H with respect to $\|\cdot\|_F$. In particular, for any g, $h \in H$, $g^{(q)} = h^{(q)}$.

LEMMA 4. *Let $h \in \Omega(f)$.*

(A) *If h and f have continuous derivatives of order up to $k_p + 1$, then*

$$h^{(k_i+1)}(x) = f^{(k_i+1)}(x)$$

for all $x \in U_i(f, h) \cap (a, b)$ and $i = 1, \ldots, p$.

(B) *If $k_{i+1} = k_i + 1$ for some i, then*

$$U_i(f, h) \cap U_{i+1}(f, h) \cap (a, b) = \varnothing.$$

In proving uniqueness, we will use the following technique. We show that the difference $h - g$ of a minimal and an arbitrary best approximation is the homogeneous solution of a set of equations of the form

$$(h-g)^{(k_i)}(x) = 0, \qquad x \in U_i(f).$$

This will allow us to conclude that $h - g \equiv 0$, i.e., the best approximation is unique.

The theory of Birkhoff interpolation (see, e.g., [7]) and its extensions by Hausmann [2] and Keener [5] give us sufficient conditions that the homogeneous solution of the above interpolation problem vanish identically.

Let $E = (e_{ij})$, $i = 1,..., M$, $j = 0, 1,..., N-1$, be an $M \times N$ matrix whose entries e_{ij} are only zeros and ones. Let H be a subspace of $C^{(N-1)}[a, b]$. The Birkhoff interpolation problem for E, given data b_{ij} and knots x_i, $i = 1,..., M$, is to find an $h \in H$ such that

$$h^{(j)}(x_i) = b_{ij}$$

for all those (i,j) such that $e_{ij} = 1$. E is called the incidence matrix of the problem and $X = \{x_1,..., x_M\}$ the set of knots.

If, for a given incidence matrix E, the interpolation problem has a unique solution for each set of data $\{b_{ij}\}$ and knots $a \leqslant x_1 < x_2 < \cdots < x_M \leqslant b$, then E is said to be regular. Otherwise it is singular. If, for a given set of knots X, there is a unique solution for each set of data, we say that (E, X) is regular.

A sequence of ones given by $e_{ij} = 0$, $e_{i,j+1} = e_{i,j+2} = \cdots = e_{i,j+q} = 1$ and $e_{i,j+q+1} = 0$ (or is undefined) is said to be supported if there exist (k_1, l_1), (k_2, l_2) with $k_1 < i$, $k_2 > i$ and $l_1, l_2 < j+1$ for which $e_{k_1 l_1} = e_{k_2 l_2} = 1$. The sequence is even or odd if q is even, respectively, odd. The incidence matrix E satisfies the Polya condition if, for each $k = 1, 2,..., N$, the number of ones contained in the first k columns is at least k. It satisfies the Birkhoff condition, if, for each $k = 1, 2,..., N-1$, the number of ones contained in the first k columns is at least $k+1$.

To avoid confusion, we would like to emphasize that the columns of E are numbered from 0 to $N-1$. Thus the first k columns of E are the columns numbered from 0 to $k-1$.

The well-known theorem of Aktinson and Sharma [1] states that if E has n ones, satisfies the Pólya condition and has no odd supported sequences, then E is regular for Π_{n-1}, the subspace of algebraic polynomials of degree not exceeding $n-1$.

Using these definitions, we may state the following theorem on Birkhoff interpolation which is due primarily to Keener but whose proof needs some ideas of the original theorem of Haussmann. Since the modifications are minor, we omit the proof.

THEOREM 5. *Let H be an n-dimensional subspace of $C^{(k)}[a, b]$ and m an integer with $0 \leqslant m \leqslant k$. Suppose that* $\dim H^{(m)} = n - m$ *and that $H^{(m)}$ is a Haar space. If $m < k$, assume also that the $H^{(i)}$ for $m < i \leqslant k$ are Haar subspaces. Let E be an $M \times (k+1)$ incidence matrix with N ones and $X = \{x_1,..., x_M\}$ a set of knots. Moreover suppose that*

(a) *E satisfies the Pólya condition. If* $N < k+1$, *suppose that the first* N *columns of* E *satisfy the Pólya condition.*

(b) *E has no odd supported sequences.*

(c) *Whenever* $j \geqslant m+1$ *and* $e_{ij}=1$, *then* $e_{i,j-1}=1$.

(d) *If* $m<k$ *assume that* $A_k \subset (a,b)$, *where*

$$A_k = \{x_i | x_i \in X, e_{i,k}=1\}.$$

Then if $N=n$, (E,X) *is regular with respect to* H. *If* $N \leqslant n$, *the associated interpolation problem has a (not necessarily unique) solution for any set of data and any ordered set of knots.*

Our first main theorem is

THEOREM 6. *Let* H *be an* n*-dimensional subspace of* $C^{(k_p+1)}[a,b]$ *for which* $H^{(k_p)}$ *and* $H^{(k_p+1)}$ *are Haar subspaces and* $\dim H^{(k_p)} = n-k_p$. *Let* $f \in C^{(k_p+1)}[a,b]$ *and* F *be such that* $k_1=0$. *If* $U_1(f) \neq \varnothing$, *then the best simultaneous approximation to* f *from* H *with respect to* $\|\cdot\|_F$ *is unique.*

Proof. Recall that $U_i(f)=U_i(f,h)$ for any minimal best approximation h. We will show that $g-h\equiv 0$ for any other $g \in \Omega(f)$.

From Lemma 4(A) and the definition of a minimal best approximation

$$g^{(k_i)}(x) = h^{(k_i)}(x), \qquad x \in U_i(f),$$
$$g^{(k_i+1)}(x) = f^{(k_i+1)}(x) = h^{(k_i+1)}(x), \qquad x \in U_i(f)/\{a,b\}$$

for $i=1,\dots,p$.

Let (E,X) be the incidence matrix associated with these equations. We will show that (E,X) satisfies all the assumptions of Theorem 5 with $k=k_p$.

First of all, note that in (a,b) equations come in pairs which, because of Lemma 4(B), do not overlap. Thus sequences corresponding to knots in (a,b) are even. Clearly also, $e_{i,k_p+1}=1$ only for knots in (a,b) and if $e_{i,k_p+1}=1$, then also $e_{i,k_p}=1$.

We will now show that E, which has k_{p+2} columns, satisfies the Birkhoff condition for the first k_p columns. Since $U_1(f)\neq\varnothing$ and since the constants are contained in H, card $U_1(f)\geqslant 2$.

Now let l be the smallest integer, if it exists, for which the number of ones in the first l columns is less than $l+1$. We have shown that $l \geqslant 2$. Suppose $l \leqslant k_p$. By the choice of l, the first $l-1$ columns have at least l ones. Also the first l columns have no more than l ones. Thus the first l columns of E contain exactly l ones and the lth column consists only of zeros. Let $\bar{E}$ be the incidence matrix consisting of the first l columns of E. Then $\bar{E}$ contains exactly l ones, has no odd supported sequences and satisfies the Pólya (even the Birkhoff) condition. Thus, by the Atkinson–Sharma theorem, E is regular with respect to algebraic polynomials of degree not exceeding $l-1$. It is therefore possible to find a $P \in \Pi_{l-1}$ with

$$P^{(k_i)}(x) = \{f^{(k_i)}(x) - h^{(k_i)}(x)\}$$

for $x \in U_i(f)$, $0 \leqslant k_i \leqslant l-1$. Since $l \leqslant k_p, \Pi_{l-1} \subset H$. By the usual compactness arguments, it follows that for all $\lambda > 0$ sufficiently small

$$\max_{0 \leqslant k_i \leqslant l-1} \|f^{(k_i)} - (h + \lambda P)^{(k_i)}\| < \|f - h\|_F.$$

Since $P^{(j)} \equiv 0$ for $j \geqslant l$, $h + \lambda P$ is a best simultaneous approximation to f for all $\lambda > 0$ sufficiently small. But $U_1(f, h + \lambda P) = \varnothing$ and $U_1(f) \neq \varnothing$ which contradicts the minimality of h. Consequently E satisfies the Birkhoff condition for the first k_p columns and therefore also the Pólya condition for the first $k_p + 1$ columns.

Let E have N ones. We have shown that $N \geqslant k_p + 1$. If $N \geqslant k_p + 2$, E satisfies the Pólya condition for all of its $k_p + 2$ columns. We claim that $N \geqslant n + 1$. If $N \leqslant n$, then (E, X) would be an incidence matrix satisfying all the conditions of Theorem 5. Thus one may find a $v \in H$ satisfying

$$v^{(k_i)}(x) = \{f^{(k_i)}(x) - h^{(k_i)}(x)\}$$

for $x \in U_i(f)$, $i = 0, \ldots, p$. But then for all $\lambda > 0$ sufficiently small, $h + \lambda v$ is a better simultaneous approximation to f than h is. This is impossible, so $N \geqslant n + 1$.

But now, $h - g$ is the homogeneous solution of a regular Birkhoff interpolation problem. Thus $h = g$ and the best simultaneous approximation is unique.

THEOREM 7. *Let H be an n-dimensional subspace of $C^{(k_p+1)}[a, b]$ for which $H^{(k_p)}$ and $H^{(k_p+1)}$ are Haar spaces and* $\dim H^{(k_p)} = n - k_p$. *Let $f \in C^{(k_p+1)}[a, b]$ and*

$$q = \min\{k_i \mid U_i(f) \neq \varnothing\}.$$

Then $\dim \Omega(f) = q$ *and for any $g, h \in \Omega(f)$, $g^{(q)} = h^{(q)}$. Moreover, this unique* qth *derivative of the best simultaneous approximations to f from H with respect to $\|\cdot\|_F$ is itself the unique best simultaneous approximation to $f^{(q)}$ from $H^{(q)}$ with respect to the norm $\|\cdot\|_G$, where*

$$G = \{k_i - q \mid k_i \in F, U_i(f) \neq \varnothing\}.$$

Proof. All but the first assertion follows from Theorem 6 and Lemma 3. But $g + \lambda P \in \Omega(f)$ for all $g \in \Omega(f)$, $P \in \Pi_{q-1}$ and all λ sufficiently small since $\Pi_{q-1} \subset H$. Therefore $\dim \Omega(f) \geqslant q$. On the other hand, $g^{(q)} = h^{(q)}$ for any g, $h \in \Omega(f)$ implies that $\dim \Omega(f) \leqslant q$ which completes the proof.

An alternative formulation of this theorem is

COROLLARY 8. *Suppose that the assumptions of Theorem 7 are fulfilled. If for some i, $U_i(f, g) \neq \varnothing$ for all $g \in \Omega(f)$, then $g^{(k_i)} = h^{(k_i)}$ for any $g, h \in \Omega(f)$.*

The conclusion of Corollary 8 is of course satisfied for $i = p$. This was a fact proved by Keener under slightly more restrictive assumptions on H.

COROLLARY 9. *Suppose that the assumptions of Theorem 7 are satisfied. Then $g^{(k_p)} = h^{(k_p)}$ for any two best simultaneous approximations to f.*

If the requirement that $H^{(k_p)} = n - k_p$ is replaced by $\dim H^{(k_p - 1)} = n - k_p + 1$, one obtains a weaker conclusion about the dimension of the set of best simultaneous approximations.

THEOREM 10. *Let f, H be contained in $C^{(k_p+1)}[a, b]$. Assume that $\dim H^{(k_p-1)} = n - k_p + 1$ and that $H^{(k_p-1)}$, $H^{(k_p)}$ and $H^{(k_p+1)}$ are Haar spaces. Then $\dim \Omega(f) \leqslant k_p$. In fact, if q is the smallest k_i for which $U_i(f) \neq \varnothing$, then*

(a) $q \leqslant \dim \Omega(f) \leqslant q + 2$ *for* $0 \leqslant q \leqslant k_p - 2$,

(b) $k_p - 1 \leqslant \mathit{dim}\, \Omega(f) \leqslant k_p$ *for* $q = k_p - 1$,

(c) $\mathit{dim}\, \Omega(f) = k_p - 1$ *for* $q = k_p$ *and* $\dim H^{(k_p)} = n - k_p + 1$,

(d) $\mathit{dim}\, \Omega(f) = k_p$ *for* $q = k_p$ *and* $\dim H^{(k_p)} = n - k_p$.

Proof. We start with the case $k_1 = 0$ and $U_1 \neq \varnothing$ (i.e., $q = 0$). From Lemma 4, we know that the difference $h - g$ of a minimal and an arbitrary best simultaneous approximation to f vanishes on the incidence matrix (E, X) corresponding to the equations

$$v^{(k_i)}(x) = 0, \qquad x \in U_i(f),$$
$$v^{(k_i+1)}(x) = 0, \qquad x \in U_i(f) \backslash \{a, b\},$$

$i = 1, \ldots, p$. Moreover, since H contains the algebraic polynomials of degree up to $k_p - 2$, we may conclude that E satisfies the Birkhoff condition for the first $k_p - 1$ columns and therefore the Pólya condition for the first k_p columns. If E has no ones in the column numbered k_p (corresponding to the k_pth derivative), then all its nonzero elements are concentrated in the first k_p rows and the conclusions of Theorem 7 hold. If E does have a one in the column numbered k_p, we would like to use Theorem 5 with $k = k_p + 1$ and

$m = k_p - 1$. To do this, we must modify E. Assume that E has M rows and N ones. Let $\bar{E}$ be the $M \times (k_p + 2)$ incidence matrix obtained from E by adding a one at position $(i, k_p - 1)$ and deleting the one at position $(i, k_p + 1)$ whenever $e_{i,k_p-1} = 0$, $e_{i,k_p} = 1$ and $e_{i,k_p+1} = 1$. Moreover, a one is inserted at position $(1, k_p - 1)$ respectively at $(M, k_p - 1)$ if $e_{1,k_p-1} = 0$, $e_{1,k_p} = 1$ and $e_{1,k_p+1} = 0$ respectively if $e_{M,k_p-1} = 0$, $e_{M,k_p} = 1$ and $e_{M,k_p+1} = 0$. $\bar{E}$ has the following properties

(1) If $\bar{N}$ denotes the number of ones $\bar{E}$ has, then $\bar{N} \leqslant N + 2$.

(2) If $\bar{N} \leqslant k_p + 1$, then $\bar{E}$ satisfies the Polya condition for the first $k_p + 1$ columns. Otherwise $\bar{E}$ satisfies the Pólya condition for all of its $k_p + 2$ columns.

(3) $\bar{E}$ has only even sequences in rows corresponding to knots in the interior of $[a, b]$.

(4) Let $A_{k_p+1} = \{x_i | x_i \in X, e_{i,k_p+1} = 1\}$. Then $A_{k_p+1} \subset (a, b)$.

(5) If $e_{ij} = 1$ for $j \geqslant k_p$, then $e_{i,j-1} = 1$.

Suppose that $\bar{N} \leqslant n$. Then by Theorem 5, we may conclude that if h is some minimal best approximation, then there is a $v \in H$ for which

$$v^{(k_i)}(x) = f^{(k_i)}(x) - h^{(k_i)}(x)$$

for $x \in U_i(f)$, $i = 1,..., p$. Since these are all the extremal points of $f - h$, the usual compactness argument shows that $h + \lambda v$ is a better approximation to f than h for all $\lambda > 0$ sufficiently small. Thus $\bar{N} \geqslant n + 1$.

We now define a third incidence matrix E^* which is also a modification of E. Let E^* be the $M \times (k_p + 2)$ incidence matrix derived from E by inserting a one at position $(1, k_p - 1)$ if $e_{1,k_p-1} = 0$ and $e_{1,k_p} = 1$ and by inserting a one at position $(M, k_p - 1)$ if $e_{M,k_p-1} = 0$ and $e_{M,k_p} = 1$. E^* then has exactly as many ones as $\bar{E}$; namely, $\bar{N} \geqslant n + 1$.

Now let h be a minimal and g an arbitrary best approximation to f. We know that $h - g$ vanishes on (E, X). In the case that a one at position $(1, k_p - 1)$ has been added to E to obtain E^*, let $v_1 \in H$ be such that v_1 vanishes on (E, X) and $v_1^{(k_p-1)}(a) \neq 0$. In the case that a 1 at position $(M, k_p - 1)$ has been added to E to obtain E^*, let $v_2 \in H$ be such that v_2 vanishes on (E^*, X), except that $v_2^{(k_p-1)}(b) \neq 0$. Then, for the appropriate choice of λ_1, λ_2, $h - g - \lambda_1 v_1 - \lambda_2 v_2$ vanishes on (E^*, X). It follows that $h - g \equiv \lambda_1 v_1 + \lambda_2 v_2$, i.e., $\dim \Omega(f) = 2$. It could happen that either less than two 1's were added to E to obtain E^* or that v_1, v_2 with the required properties do not exist. In either case, the dimension of $\Omega(f)$ would be smaller. Thus $\dim \Omega(f) \leqslant 2$.

The general case, $q \geqslant 1$, can be treated as before to obtain the conclusion that dim $\Omega^{(q)}(f) \leqslant 2$. If $q \leqslant k_p - 2$, then $h + \lambda P$, for small λ and $P \in \Pi_{q-1}$, is also a best approximation. Thus

$$q \leqslant \dim \Omega(f) \leqslant q + 2 \leqslant k_p.$$

If $q = k_p - 1$, we have the problem of approximating $f^{(k_p - 1)}$ by elements of $H^{(k_p - 1)}$ with respect to the norm $\| \cdot \|_G$ with $G = \{0, 1\}$ or $G = \{0\}$. It was shown by Keener that if $G = \{0, 1\}$, then dim $\Omega_G(f^{(k_p-1)}, H^{(k_p-1)}) \leqslant 1$ from which it follows that

$$\dim \Omega_F^{(k_p-1)}(f) \leqslant 1.$$

Since $\Pi_{k_p - 2} \subset H$, we have

$$k_p - 1 \leqslant \dim \Omega(f) \leqslant k_p.$$

If $q = k_p$, then $G = \{0\}$ and we have the ordinary Chebycheff approximation of a continuous function by elements of a Haar subspace. Thus dim $\Omega_G(f^{(k_p)}, H^{(k_p)}) = 0$ from which it follows

$$\dim \Omega_F^{(k_p)}(f) = 0.$$

Therefore dim $\Omega(f) = k_p - 1$ if dim $H^{(k_p)} = n - k_p + 1$ and dim $\Omega(f) = k_p$ if dim $H^{(k_p)} = n - k_p$.

Combining the separate cases, we may conclude that dim $\Omega(f) \leqslant k_p$.

Although the conclusion that dim $\Omega(f) \leqslant k_p$ seems, superficially, just as good as that of Theorem 7, one should notice that dim $\Omega^{(k_p)}(f)$ may be 2, whereas in Theorem 7, dim $\Omega^{(k_p)}(f) = 0$ always.

The more exact analysis of dimension of the set of best simultaneous approximations and the slight improvement (Theorem 5) of Keeners sufficient condition for the regularity of Birkhoff interpolation have allowed dropping the assumption about the extendibility of $H^{(k_p-1)}$, $H^{(k_p)}$ and $H^{(k_p+1)}$ which Keener made and yet obtain the better upper bound k_p (instead of $k_p + 1$) for dim $\Omega(f)$.

The concept of Birkhoff systems was introduced by Lorentz in [6] in connection with the regularity of Birkhoff interpolation problems for non-polynomial spaces. A Birkhoff system is an n-dimensional subspace B of $C^{(n-1)}[a, b]$ with the property that if $b \in B$, $b^{(i)}(x_i) = 0$, $i = 0, 1, \ldots, n-1$, for any x_i with $a \leqslant x_i \leqslant b$, then $b \equiv 0$. It can shown that any incidence matrix with n ones which satisfies the Pólya condition and has no odd supported sequence is regular for a Birkhoff system (see [6]). Examples of Birkhoff systems are algebraic polynomials and span$\{e^x, e^{2x}\}$ on $[1, 1 + a]$ for any $a < \ln 2$.

As with Haar sets, if B is a subspace of $C^{(n-1)}[a, b]$ of dimension n such that dim $B^{(m)} = n - m$ and that $B^{(m)}$ is a Birkhoff system, then $B^{(i)}$ is a

Birkhoff system of dimension $n-i$ for all $i=0, 1, \ldots, m$. This even holds in reverse order. If B is a Birkhoff system of dimension n and if $\dim B^{(m)} = n-m$, then each $B^{(i)}$, $i=0, 1, \ldots, m$, is a Birkhoff system of degree $n-i$. This was noted by Lorentz [6].

THEOREM 11. *Let $f \in C^{(k_p+1)}[a,b]$ and B be an n-dimensional Birkhoff system (which tacitly assumes that B is a subset of $C^{(n-1)}[a,b]$) such that* $\dim B^{(k_p)} = n-k_p$. *If q is the smallest k_i for which $U_i \neq \varnothing$, then $\dim \Omega(f) = q$ and $g^{(q)} = h^{(q)}$ for any g, $h \in \Omega(f)$. In particular,* $\dim \Omega(f) \leqslant k_p$.

THEOREM 12. *Let $f \in C^{(k_p+1)}[a,b]$ and B be an n-dimensional Birkhoff system such that* $\dim B^{(k_p-1)} = \dim B^{(k_p)} = n-k_p+1$ *and such that $B^{(k_p)}$ is a Haar set. If q is the smallest k_i for which $U_i \neq \varnothing$, then* $\dim \Omega(f) = q$ *for* $q \leqslant k_p - 1$ *and* $\dim \Omega(f) = k_p - 1$ *if* $q = k_p$. *In particular* $\dim \Omega(f) \leqslant k_p - 1$. *Moreover, if $g, h \in \Omega(f)$, then $g^{(k_p-1)} = h^{(k_p-1)}$.*

Proof. As in Theorem 6, we can show that the difference $h-g$ of a minimal and an arbitrary best simultaneous approximation to f vanishes on an incidence matrix (E, X) which satisfies the Pólya conditions, has no odd supported sequences and has at least $n+1$ ones. It is therefore regular with respect to B. There are no difficulties with conditions of the form $g^{(k_p)}(a) = 0$ because this does not influence the regularity of E with respect to a Birkhoff system.

This allows us to conclude that if $k_q \leqslant k_p - 1$, then $\dim \Omega(f) = k_q$ in both Theorems 11 and 12. If $k_p = q$, then the fact that $B^{(k_p)}$ is a Haar set implies that $\dim \Omega^{(k_p)}(f) = 0$. If also $\dim B^{(k_p)} = n - k_p + 1$, then necessarily $\dim \Omega(f) = k_p - 1$. The other conclusion of Theorems 11 and 12 follow from these observations. Keener [5] investigated simultaneous approximation for the space $H = \operatorname{span}\{e^x, e^{2x}\}$ on $[1, 1+\ln 2]$. For the norm $\|f\|_F = \max\{\|f\|, \|f'\|\}$, he showed that the best simultaneous approximations may not be unique although $U_1(f) \neq \varnothing$. In this case, however, the space considered is not a Birkhoff system. In fact, if the domain of definition is taken to be $[1, b]$ for any with $1 < b < \ln 2$, then H is a Birkhoff system and $U_1(f) \neq \varnothing$ implies that the best simultaneous approximation is unique.

REFERENCES

1. K. ATKINSON AND A. SHARMA, A partial characterization of poised Hermite-Birkhoff interpolation problems, SIAM *J. Numer. Anal.* **6** (1969), 230–235.
2. W. HAUSMANN, Differentiable Tchebycheff subspaces and Hermite interpolation, *Ann. Acad. Sci. Fenn. Ser. A I Math.* **4** (1978/1979), 75–83.

3. Y. IKEBE, Hermite–Birkhoff interpolation problems in Haar subspaces, *J. Approx. Theory* **8** (1973) 142–149.
4. D. J. JOHNSON, Nonuniqueness of simultaneous approximation by trigonometric polynomials, *in* "Approximation Theory II (G. G. Lorentz, C. K. Chui, and L. L. Schumaker, Eds.), pp. 411–415, Academic Press, New York, 1976.
5. L. L. KEENER, Hermite–Birkhoff interpolation and approximation of a function and its derivatives, *J. Approx. Theory* **30** (1980), 129–138.
6. G. G. LORENTZ, Independent sets of knots and singularity of interpolation matrices, *J. Approx. Theory* **30** (1980), 208–225.
7. G. G. LORENTZ, Recent progress in Birkhoff interpolation, *in* "Approximation Theory and Functional Analysis" (J. B. Prolla, Ed.), pp. 187–236, North-Holland, Amsterdam, (1979).
8. R. A. LORENTZ, Nonuniqueness of simultaneous approximation by algebraic polynomials, *J. Approx. Theory* (1975), 17–23.
9. F. J. SCHUURMAN, Uniqueness in the uniform approximation of a function and its derivatives, *SIAM J. Numer. Anal.* **6** (1969), 305–315.