manuscripta math. 63, 469 - 478 (1989)

manuscripta mathematica

Optimal Nonlinear Approximation

RONALD A. DEVORE[1], RALPH HOWARD[2], AND CHARLES MICCHELLI

We introduce a definition of nonlinear n-widths and then determine the n-widths of the unit ball of the Sobolev space W_p^r in L_q. We prove that in the sense of these widths the manifold of splines of fixed degree with n free knots is optimal for approximating functions in these Sobolev spaces.

1. Introduction. There are many known classes of functions which can be approximated by nonlinear families such as rational functions or splines with free knots better than they can be approximated by the elements of linear spaces such as polynomials. Perhaps the simplest example of this is that functions f with $f' \in L_p(\Omega)$, $\Omega := [0,1]$, $1 \le p \le \infty$ can be approximated in the ***uniform norm*** by piecewise constants with n free knots ($1 \le p \le \infty$, see [5]) or by rational functions of degree n ($1 < p \le \infty$, see [9] or [3]) with an error of approximation $O(n^{-1})$. At the same time, polynomials of degree $< n$ or for that matter any n dimensional space can not yield an error of approximation better than $O(n^{-3/2+1/p})$ for $1 \le p \le 2$ (see [6]). The purpose of the present paper is to discuss in what sense these and other estimates for nonlinear approximation are optimal.

We begin by discussing which nonlinear families will be considered in our approximation. Let X be a Banach space and let M be a mapping from $\mathbb{R}^n$ into X which associates with each $a \in \mathbb{R}^n$ the element $M_n(a) \in X$. We shall approximate the elements $f \in X$ by the elements of $\mathcal{M}_n := \{M_n(a) : a \in \mathbb{R}^n\}$. If $f \in X$, the error of approximation of f is

$$E(f,\mathcal{M}_n)_X := \inf_{a\in\mathbb{R}^n} \|f - M_n(a)\|_X. \tag{1.1}$$

More generally for a set K of elements of X, we have

$$E(K,\mathcal{M}_n)_X := \sup_{f\in K} E(f,\mathcal{M}_n)_X. \tag{1.2}$$

We are interested in some sense in the best manifolds $\mathcal{M}_n$ of dimension n for approximating the elements of K. That is, we would like to choose $\mathcal{M}_n$ so that (1.2) is as small as possible. If we were to operate in strict analogy with the case of linear approximation, we would define the nonlinear n-width of K as the infimum of (1.2) over all manifolds $\mathcal{M}_n$ of dimension n. However, this is too general to be of any use. In fact, this width is zero for all K and all separable X. Indeed, for $n = 1$, there is a space filling manifold (even with M_n continuous). Namely, let $\{x_k\}_{k=-\infty}^{\infty}$ be dense in X and define

[1]This author was supported by NSF Grant DMS 8620108
[2]This author was supported by NSF Grant DMS 8803585

$M_n(a) := (a-k)x_{k+1} + (k+1-a)x_k$ for $k \le a \le k+1$. Then, M_n is continuous and for the corresponding manifold $\mathcal{M}_n$, we have (1.2) is zero for all K.

One possiblility to circumvent the triviality described in the previous paragraph is to assume smoothness for the manifold $\mathcal{M}_n$. However, it is easy to see that this would exclude the classical manifolds such as rational functions and splines with free knots. It turns out that a more reasonable approach is to impose conditions on how the approximation by the elements of $\mathcal{M}_n$ takes place.

We recall that a quasi-norm $||.||$ satisfies the usual properties of a norm except that the triangle inequality is replaced by; $||f+g|| \le C(||f||+||g||)$ with C an absolute constant. We say that a mapping $\bar{a}$ from K into $\mathbb{R}^n$ is a continuous selection for K if it is continuous in the topology of some quasi-norm. We recall that if K is a compact set then all quasi-norms are equivalent on K and so we can simply say that $\bar{a}$ is continuous on K. Given such an $\bar{a}$, for each $f \in K$, $M_n(\bar{a}(f))$ is an approximation to f from $\mathcal{M}_n$. We define

$$E(K,\bar{a},\mathcal{M}_n)_X := \sup_{f \in K} ||f - M_n(\bar{a}(f))|| \tag{1.3}$$

to be the error of approximation for the set K by the nonlinear method of approximation $M_n(\bar{a}(.))$. To find the "best" nonlinear method for K, we consider all manifolds $\mathcal{M}_n$ and all continuous selections $\bar{a}$ and define

$$d_n(K)_X := \inf_{\bar{a},\, \mathcal{M}_n} E(K,\bar{a},\mathcal{M}_n)_X \tag{1.4}$$

to be the ***continuous*** nonlinear n-width of K. Then $d_n(K)_X$ is a nondecreasing function of n and if $K \subset \tilde{K}$, $d_n(K)_X \le d_n(\tilde{K})_X$.

The purpose of the present paper is to determine (asymptotically) $d_n(K)_X$ for certain X, K. In complete analogy with the linear case, we establish lower bonds for $d_n(K)_X$ for general K and X in §3 in terms of the Bernstein width of K and then we apply this in §4. Upper bounds for $X = L_q$ and K a set determined by a smoothness condition in L_p are given in §5. The upper and lower bounds serve to determine the n-widths of these sets.

Before proceeding to the main results of this paper, we discuss in §2 some ramifications of our defintion. In particular, we examine the condition that the approximation is made through a continuous selection $\bar{a}$.

2. Remarks on the definition of d_n. We want to point out that for certain "good" manifolds $\mathcal{M}_n$, the requirement that the approximation takes place through a continuous selection $\bar{a}$ is not an essential restriction. We say that an element $M_n(a)$ is a ***near best approximation*** to f with constant λ, if

$$||f - M_n(a)||_X \le \lambda E(f,\mathcal{M}_n)_X. \tag{2.1}$$

It is an interesting question to decide for which $\mathcal{M}_n$, X, and K, there exists a continuous selection $\bar{a}$ such that $M_n(\bar{a}(f))$ is a near best approximation with fixed constant λ for all $f \in K$. When this is the case, we have

$$(2.2)\qquad E(K,\bar{a},\mathcal{M}_n)_X \le \lambda E(K,\mathcal{M}_n)_X.$$

For such manifolds, $\bar{a}$ could be dropped in the definition (1.4) with the resulting quantity differing from $d_n(K)_X$ by at most the multiplicative constant λ. In other words, in this case, the selection $\bar{a}$ plays no essential role. On the other hand, $\bar{a}$ has a taming effect on the more bizarre manifolds. We give now some examples where $\bar{a}$ plays no essential role.

We shall often make use of the following remarks about a metric space Y. We denote by $B(f,\eta)$ the ball centered at f of radius η. If a collection of balls $B_\nu := B(f_\nu,\eta)$ cover Y (or some subset K of Y), then from the paracompactness of Y [7, p.160], there is a locally finite collection $\{U\}$ of open sets which are a refinement of $\{B_\nu\}$ and which cover Y (or K). Locally finite means that for each $f \in Y$ (or K), there is a ball $B(f,\eta)$, $\eta > 0$, which intersects at most a finite number of the U. For the covering $\{U\}$, there is a partition of unity $\{\alpha_U\}$ subordinate to $\{U\}$ (see [7, p.171]). That is, the functions α_U are nonnegative, continuous and supported on U and $\sum_U \alpha_U \equiv 1$ on Y (or K). If for each U, a_U is a point in $\mathbb{R}^n$, then the function

$$(2.3)\qquad \bar{a}(f) := \sum_U \alpha_U(f) a_U$$

is a continuous function on Y taking values in $\mathbb{R}^n$. Indeed, given any f in Y, we choose a ball $B := B(f,\eta)$, $\eta > 0$, small enough that it intersects at most a finite number of the sets U. Then on B, the sum (2.3) involves only a finite number of terms and is therefore continuous.

We first prove that linear manifolds (i.e. M is a linear function) admit a continuous near best selection $\bar{a}$.

THEOREM 2.1. *If $\mathcal{M}_n$ is an n-dimensional linear manifold and X is a Banach space, then for any $\epsilon > 0$ there exists a continuous selection $\bar{a}$ such that*

$$(2.4)\qquad \|f - M_n(\bar{a}(f))\|_X \le E(f,\mathcal{M}_n)_X + \epsilon,$$

for all f in X.

PROOF: Let $\eta > 0$ and consider balls $B(f_\nu,\eta)$, $\nu \in \Lambda$, which are a covering for X. Here Λ is some index set. Let $\{U\}$ be the covering and $\{\alpha_U\}$ be the partition of unity described above. For each U, we choose an f_U from U and let a_U be such that $M_n(a_U)$ is a best approximation to f_U. Then the function $\bar{a}$ of (2.3) is a continuous selection. Now let f be in X and let a_0 be such that $M_n(a_0)$ is a best approximation to f. If $\alpha_U(f) \ne 0$, then f and f_U are both in U and hence $\|f - f_U\|_X < 2\eta$. Therefore, $\|f_U - M_n(a_U)\|_X \le \|f_U - M_n(a_0)\|_X \le E + 2\eta$, where $E := E(f,\mathcal{M}_n)_X$. It follows that

$$(2.5)\qquad \|f - M_n(a_U)\|_X \le E + 4\eta.$$

Since $M(\sum \alpha_U a_U) = \sum \alpha_U M_n(a_U)$, we have

$$(2.6)\quad \|f - M_n(\bar{a}(f))\|_X = \|\sum_U \alpha_U(f)[f - M_n(a_U)]\|_X \le (E + 4\eta)\sum_U \alpha_U(f) = E + 4\eta.$$

Since η is arbitrary, we have proved the theorem. ■

COROLLARY 2.2. *Under the hypotheses of Theorem 2.1, if K is a compact set contained in X for which $E(K,\mathcal{M}_n)_X \neq 0$, then for each $\lambda > 1$, there is a selection $\bar{a}$ such that*

$$(2.7)\qquad E(K,\bar{a},\mathcal{M}_n)_X \le \lambda E(K,\mathcal{M}_n)_X.$$

PROOF: We can take $\epsilon := (\lambda - 1)E(K,\mathcal{M}_n)_X$ and apply Theorem 2.1.

In the case that X is separable, we have the following strengthening of (2.7).

THEOREM 2.3. *If X is a separable Banach space and $\mathcal{M}_n$ is a linear manifold, then for each $\lambda > 1$, there is a continuous selection $\bar{a}$ defined on X such that*

$$(2.8)\qquad \|f - M_n(\bar{a}(f))\|_X \le \lambda E(f,\mathcal{M}_n)_X, \quad f \in X.$$

PROOF: We use the fact that for each λ there is a strictly convex norm $\|.\|_0$ on X which satisfies

$$(2.9)\qquad \|f\|_X \le \|f\|_0 \le \lambda\|f\|_X, \quad f \in X.$$

This follows from the Clarkson renormalization lemma (see e.g. [4, pg. 107]. Now, let $f \in X$, and let M_f be its best approximation from $\mathcal{M}_n$ in $\|.\|_0$. Then the mapping $f \to M_f$ is well known to be continuous on X (see e.g. [2, pg. 23]). Now the manifold $\mathcal{M}_n$ is a translate of a linear space X_n; $\mathcal{M}_n = g_0 + X_n$. We take a basis $P_1,\dots,P_n$ of X_n and we parametrize $\mathcal{M}_n$ by the coefficients of the P_k, that is, $M(a) := g_0 + \sum_k a_k P_k$ for $a := (a_k)$. We can write $M_f = M(\bar{a}(f))$ where $M_f = g_0 + \sum_k a_k(f)P_k$. Then $\bar{a}(f) := (a_k(f))$ is continuous on X. Indeed, the Euclidean norm of (a_k) is equivalent to $\|\sum_k a_k P_k\|_X$. Finally, we have

$$\|f - M(\bar{a}(f))\|_X \le \|f - M(\bar{a}(f))\|_0 = \inf_{M\in\mathcal{M}_n} \|f - M\|_0 \le \lambda \inf_{M\in\mathcal{M}_n} \|f - M\|. ■$$

We can also show that continuous selections $\bar{a}$ exist for more general manifolds provided that they are sufficiently smooth. Let $\mathcal{M}_n$ satisfy

$$(2.10)\qquad C_1\|a - b\| \le \|M_n(a) - M_n(b)\|_X \le C_2\|a - b\|,$$

for some norm $\|.\|$ on $\mathbb{R}^n$ and some constants $C_1, C_2 > 0$ (independent of a, b). Then if f is in X, the function $F(a) := \|f - M_n(a)\|_X$ is continuous on $\mathbb{R}^n$ and $F(a) \to \infty$ as $\|a\| \to \infty$. Hence, F attains its minimum and therefore there exists a best approximation to f from $\mathcal{M}_n$.

THEOREM 2.4. *If X is a Banach space and $\mathcal{M}_n$ is a manifold satisfying (2.8), then for each $\epsilon > 0$ there is a continuous selection $\bar{a}$ on X such that*

$$(2.11) \qquad \|f - M_n(\bar{a}(f))\|_X \leq CE(f,\mathcal{M}_n)_X + \epsilon$$

for a constant $C \leq 1 + 2C_2C_1^{-1}$.

PROOF: We let $\eta, E, \{U\}, \{\alpha_U\}, a_0$ and $\bar{a}$ be as in the proof of Theorem 2.1. From (2.5), $\|M_n(a_U) - M_n(a_0)\|_X \leq 2E + 4\eta$ whenever $\alpha_U(f) \neq 0$. Hence, from (2.10), $\|a_U - a_0\| \leq C_2(2E + 4\eta)$ and therefore using the definition of the partition of unity α_U, we have $\|\bar{a}(f) - a_0\| \leq C_1^{-1}(2E + 4\eta)$. Then using the upper inequality in (2.10), $\|M_n(\bar{a}(f)) - M_n(a_0)\|_X \leq C_1^{-1}C_2(2E + 4\eta)$. Finally,

$$\begin{aligned}\|f - M_n(\bar{a}(f))\|_X &\leq \|f - M_n(a_0)\|_X + \|M_n(a_0) - M_n(\bar{a}(f))\|_X \\ &\leq (1 + 2C_1^{-1}C_2)E + 4C_1^{-1}C_2\eta. \blacksquare\end{aligned}$$

To apply Theorem 2.4 to nonlinear n-widths, one would want the constants C_1, C_2 of (2.10) to be independent of n. We cannot show that the manifolds of rational functions or free knot splines satisfy conditions like this but we do show later in §5 that for the compact sets of interest to us, we can find a continuous selection satisfying (2.2) for the manifold of free knot splines.

3. A lower bound for d_n. For a quasi-normed linear space Y, we shall denote by $U(Y) := \{y : \|y\| \leq 1\}$ the unit ball of Y and by $\partial(U(Y))$ its boundary. A similar definition applies when $\|.\|$ is a semi-norm or even a quasi-semi-norm. The Bernstein width of a subset K of the quasi-normed linear space X is

$$(3.1) \qquad b_n(K)_X := \sup_{X_{n+1}} \sup\{\rho : \rho U(X_{n+1}) \subset K\},$$

with the first *sup* taken over all $n+1$ dimensional linear subspaces of X. It is well known that the Bernstein width of K provides a lower bound for the linear n-width of K (see [8, p.13]. The following shows (with the same proof) that this remains valid for our definition of nonlinear n-width.

THEOREM 3.1. *Let X be a normed linear space and let $K \subset X$. Then*

$$(3.2) \qquad d_n(K)_X \geq b_n(K)_X.$$

If X is a quasi-normed linear space then

$$(3.3) \qquad d_n(K)_X \geq c_0 b_n(K)_X$$

for an absolute constant c_0.

PROOF: Let $\rho < b_n(K)_X$ and let X_{n+1} be an $n+1$ dimensional subsapce of X such that $\rho U(X_{n+1}) \subset K$. If $\mathcal{M}_n$ is any n dimensional manifold and $\bar{a}$ is any continuous selection for K, we let $\tilde{a}(f) := \bar{a}(f) - \bar{a}(-f)$. If $\|.\|$ is the quasi-norm involved in the continuity of $\bar{a}$, then on X_{n+1}, $\|.\|$ is equivalent to $\|.\|_X$ (the quotient of these two quasi-norms is a continuous nonvanishing function on $\partial(U(X))$). Thus, $\tilde{a}(f)$ is a continuous mapping of $\partial(\rho U(X_{n+1}))$ into $\mathbb{R}^n$ and $\tilde{a}$ is odd, i.e. $\tilde{a}(-f) = -\tilde{a}(f)$. Hence, by Borsuk's antipodality theorem [1], there is an f_0 in $\partial(\rho U(X_{n+1}))$ for which $\tilde{a}(f) = 0$, i.e. $\bar{a}(-f_0) = \bar{a}(f_0)$. We write $2f_0 = (f_0 - M_n(\bar{a}(f_0)) - (-f_0 - M_n(\bar{a}(-f_0))$ and find

$$2\|f_0\|_X \leq \|f_0 - M_n(\bar{a}(f_0))\|_X + \| - f_0 - M_n(\bar{a}(-f_0))\|_X. \tag{3.4}$$

It follows that one of the two functions $f_0, -f_0$ is approximated by $M_n(\bar{a}(.))$ with an error $\geq \|f_0\|_X = \rho$. Therefore, $d_n(K)_X \geq \rho$ and (3.2) follows. If $\|.\|_X$ is only a quasi-norm, then(3.4) holds with 2 replaced by $2c_0$ on the left side. ■

4. Lower bounds for widths of smoothness classes. We shall now apply Theorem 3.1 to give lower bounds for $d_n(K)_X$ for certain sets K which are defined by a smoothness condition. The ideas here are well known and have been used previously to provide lower bounds for linear widths. We shall consider functions defined on the unit cube $\Omega := [0,1]\times\cdots\times[0,1]$ of $\mathbb{R}^d$. We let W_p^r be the Sobolev space consisting of all functions f which have weak derivatives of order $\leq r$ in L_p (all spaces and all norms here and later are over Ω unless otherwise indicated). We denote by $|.|_{W_p^r}$ and $\|.\|_{W_p^r}$ the usual semi-norm and norm for W_p^r.

We can also apply our results to Besov spaces. If r is a positive integer and $0 < p \leq \infty$, we let $\omega_r(f,t)_p := sup\{\|\Delta_h^r(f,.)\|_p(\Omega(rh)) : |h| \leq t\}$ be the L_p modulus of smoothness of f. Here we use the notation $\Omega(h)$ to denote the set of all x such that the line segment $[x, x+h] \subset \Omega$ and Δ_h^r to denote the r-th order difference operator with step h. Then, for $0 < \alpha < r$ and $0 < p, q \leq \infty$, we let $B_q^\alpha(L_p)$ denote the Besov space consisting of all functions $f \in L_p$ for which

$$|f|_{B_q^\alpha(L_p)} := \left(\int_0^\infty [s^{-\alpha}\omega_r(f,s)_p]^q ds/s\right)^{1/p} < \infty. \tag{4.1}$$

When $q = \infty$, the $L_q(dt/t)$ norm is replaced by the L_∞ norm in (4.1). We obtain the norm for $B_q^\alpha(L_p)$ by adding $\|f\|_p$ to (4.1). It is well known that different values of $r > \alpha$ give equivalent semi-norms in (4.1).

We fix the integer r and let ϕ be a $C^\infty(\mathbb{R}^d)$ function which is one on the cube $[1/4, 3/4]^d$ and vanishes outside of Ω. Furthermore, let C_0 be such that $1 \leq \|D^\nu\phi\|_\infty \leq C_0$, for $|\nu| \leq r$. We consider integers n of the form $n = m^d$ for some positive integer m and we let $Q_1, \ldots, Q_m$ be the partition of Ω into closed cubes of sidelength $1/m$. Then by applying a linear change of variables which takes Q_j to Ω, we obtain functions $\phi_1, \ldots, \phi_m$ with ϕ_j supported on Q_j and

(4.2) $$m^{|\nu|} \le \|D^\nu \phi_j\|_\infty \le C_0 m^{|\nu|}, \qquad 0 \le |\nu| \le r.$$

Let X_n be the linear span of the functions ϕ_j, $j = 1, \dots, n$.

LEMMA 4.1. *Let* $0 < p, q \le \infty$. *If* $S = \sum_1^n c_j \phi_j$, *then*

(4.3)
$$\text{(i)} \quad |S|_{W_p^k} \approx n^{k/d} \left(n^{-1} \sum_{j=1}^n |c_j|^p \right)^{1/p}, \qquad k = 0, \dots, r,$$
$$\text{(ii)} \quad |S|_{B_q^\alpha(L_p)} \le C n^{\alpha/d} \left(n^{-1} \sum_{j=1}^n |c_j|^p \right)^{1/p}, \qquad 0 < \alpha < r$$

with the right side replaced by $\sup_{1 \le k \le n} |c_k|$ *when* $p = \infty$ *and with* C *depending only on* r *and* ϕ.

Here and later $\approx$ means that the quantities A and B being compared satisfy $A \le const.B$ and $B \le const.A$ with the constants depending only on r.

PROOF OF LEMMA 4.1: Part (i) follows simply from (4.2) and the fact that the ϕ_j have disjoint supports. For (ii), we first observe that because of (4.2), for all x(with all constants here and later depending only on r and ϕ)

(4.4) $$|\Delta_h^r(\phi_j, x)| \le C max(1, m^r |h|^r).$$

Now for each $x \in \Omega$, at most $r+1$ terms are nonzero in the sum $\Delta_h^r(S, x) = \sum_{j=1}^n c_j \Delta_h^r(\phi_j, x)$. Therefore, from (4.4) and Hölder's inequality, we obtain

$$\|\Delta_h^r(S)\|_p^p \le C \sum_{j=1}^n |c_j|^p \|\Delta_h^r(\phi_j)\|_p^p \le C \sum_{j=1}^n |c_j|^p n^{-1} max(m^r |h|^r, 1)^p$$

because Q_j has measure $1/n$. Taking a *sup* over $|h| \le t$ gives

$$\omega_r(S, t)_p \le \left(n^{-1} \sum_{j=1}^n |c_j|^p \right)^{1/p} max(m^r t^r, 1)$$

and (4.3)(ii) follows simply from this.■

THEOREM 4.2. *Let* $X = L_q$, $1 \le q \le \infty$. *If* $K_{p,r} := \{f : |f|_{W_p^r} \le 1\}$, $1 \le p \le q$, $r = 1, \dots,$ *then*

(4.5) $$d_n(K_{p,r})_X \ge C n^{-r/d}$$

If $K_{p,\sigma,\alpha} := \{f : |f|_{B^\alpha_\sigma(L_p)} \le 1\}$, $\alpha > 0$ *and* $0 < p, \sigma \le \infty$ *and* $p \le q \le \infty$, *then*

$$d_n(K_{p,\sigma,\alpha})_X \ge Cn^{-\alpha/d}. \tag{4.6}$$

The constants C *depend only on* r.

PROOF: Consider first $K_{p,r}$ and $n = m^d$. Let ϕ and X_n be as above. If $S \in X_n$, then by (4.3) and Hölder's inequality

$$|S|_{W_p^r} \le Cn^{r/d}\left(n^{-1}\sum_{j=1}^{n}|c_j|^p\right)^{1/p} \le Cn^{r/d}\left(n^{-1}\sum_{j=1}^{n}|c_j|^q\right)^{1/q} \le C_0 n^{r/d}\|S\|_q.$$

Thus, for $\rho = C_0^{-1}n^{-r/d}$, we have $\rho U(X_n) \subset K_{p,r}$ and Theorem 3.1 says $d_{n-1}(K_{p,r})_X \ge \rho$. Since $d_n(K_{p,r})_X$ is monotone in n, this establishes this estimate for all n. A similar argument applies for $K_{p,\sigma,\alpha}$.■

5. Upper bounds for $d_n(K)_X$. We shall prove an upper bound and thereby determine $d_n(K)_X$ for the sets $K = K_{p,r} := U(W_p^r)$ of the previous section. We do this only in the case of one space dimension, $\Omega = [0,1]$, (the upper bound for $K_{p,\sigma,\alpha}$ and higher dimensions requires more substantial ideas). It is well known that rational functions and splines with free knots provide errors of approximation which match the lower bounds of Theorem 4.2. However, we need to check that these can be achieved with a continuous selection. This turns out to take some care. We shall use the manifold $\mathcal{M}$ of piecewise polynomials of degree $< r$ with $2n-1$ pieces. Here, we shall specify $n-1$ of the breakpoints in advance, namely, the points k/n, $k = 1, \ldots, n-1$. The other $n-1$ breakpoints are free parameters. Hence $\mathcal{M}$ has dimension $(2n-1)r + n - 1$. To parametrize $\mathcal{M}$, we use the vector a whose first $n-1$ components $0 \le a_{-n+1} \le \cdots \le a_{-1} \le 1$ are the free breakpoints of the piecewise polynomial $M_n(a)$. It will be convenient to let $t_j := a_{-n+j}$, $j = 1, \ldots, n-1$ and $t_0 := 0$, $t_n := 1$. Notice that we allow equality in the t_j; this corresponds to a degenerate interval. We let I_j, $j = 0, \ldots, 2n-1$, be the $2n-1$ intervals determined by all the break points. The other coordintates $a_0, \ldots, a_{(2n-1)r-1}$ of a denote the coefficients of the polynomials P_j which serve to define $M_n(a)$ on I_j. Thus, $P_j := a_{rj} + a_{rj+1}x + \cdots + a_{rj+r-1}x^{r-1}$, $j = 0, \ldots, 2n-2$.

Our main result is

THEOREM 5.1. *For* X *and* $K_{p,r}$ *of Theorem 4.1, we have*

$$d_n(K_{p,r}) \approx n^{-r}. \tag{5.1}$$

PROOF: Because of Theorem 4.2, we need only prove the upper estimate $d_n(K_{p,r})_X \le Cn^{-r}$. Since $U(W_p^r) \subset U(W_1^r)$, we can restrict ourselves to the case $p = 1$. We can also assume that $q = \infty$ because the case $q < \infty$ follows from this. We shall define a continuous selection $\bar{a}$ for $K_{p,r}$ and an $\mathcal{M}$ which will produce the error of approximation $O(n^{-r})$.

We first discuss how to continuously select the free breakpoints. For $f \in K_{p,r}$ there is an integer $0 \leq k \leq n$ and points $\tau_j := \tau_j(f)$ with $0 =: \tau_0 < \tau_1 < \cdots < \tau_k \leq 1$, such that

$$(5.2) \qquad \int_0^{\tau_j} |f^{(r)}(s)|\, ds = j/n, \quad j = 1, \ldots, k \text{ and } \int_0^1 |f^{(r)}(s)|\, ds = \theta + kn^{-1}$$

and $0 < \theta \leq 1/n$. We let $\tau_j := 1$, $j = k+1, \ldots, n$ and $\tau(f) := (\tau_1(f), \ldots, \tau_n(f))$.

Let $\eta := 1/5n$. The Banach space $Y := W_1^r$ is separable. Hence there are functions $f_1, f_2, \ldots$ in $K_{p,r}$ such that the balls $B_j := B(f_j, \eta)$ (defined by $\|\cdot\|_Y$) are a cover for $K_{p,r}$. We let $\{U\}$ be the refinement and $\{\alpha_U\}$ be the partition of unity described in §2 for the sets B_j and the space Y. For each U, we let f_U be in $U \cap K_{p,r}$ and we defined $t(f) := (t_1(f), \ldots, t_n(f)) := \sum_U \alpha_U(f)\tau(f_U)$. Then t is continuous on $K_{p,r}$. We give some of its additional properties.

If g, h are in $K_{p,r}$ and $\|g - h\|_Y < 4\eta$ and $\tau_i(g) < 1$, then

$$i/n - 4/5n < \int_0^{\tau_i(g)} |h^{(r)}(s)|\, ds < i/n + 4/5n.$$

This shows that for these i,

$$(5.3) \qquad \tau_{i-1}(h) \leq \tau_i(g) \leq \tau_{i+1}(h).$$

This is also true if $\tau_i(g) = 1$. Indeed, the left side is clear and if k is as in (5.2) for g then $k < i$ and $\int_0^1 |h^{(r)}(s)|\, ds \leq \int_0^1 |g^{(r)}(s)|\, ds + 4/5n = k + \theta + 4/5n$. This implies that $\tau_{k+2}(h) = \cdots = \tau_n(h) = 1$ as desired.

Since our cover is locally finite and $t_i(f)$ is a convex combination of the $\tau_i(f_U)$, we have $t_i(f) \geq \tau_i(f_U)$ for some f_U with $\|f - f_U\|_Y < \eta$ and $t_{i+1}(f) \leq \tau_{i+1}(f_{\bar{U}})$ for some $f_{\bar{U}}$ with $\|f - f_{\bar{U}}\|_Y < \eta$. We apply (5.3) to f_U and $f_{\bar{U}}$ and find

$$\tau_i(f_U) \leq t_i(f) \leq t_{i+1}(f) \leq \tau_{i+1}(f_{\bar{U}}) \leq \tau_{i+2}(f_U).$$

Hence,

$$(5.4) \qquad \int_{\tau_i(f)}^{\tau_{i+1}(f)} |f^{(r)}(s)|\, ds \leq \|f - f_U\|_Y + \int_{\tau_i(f_U)}^{\tau_{i+2}(f_U)} |f_U^{(r)}(s)|\, ds \leq 4/5n + 2/n \leq 3/n.$$

The first $n-1$ coordinates of $t(f)$ are our continuous selection for the free breakpoints. Let $I_j := [a_j, b_j]$, $j = 1, \ldots, 2n-1$ be the intervals which result when the t_j are united with the fixed breakpoints. Then the endpoints $a_j(f)$ and $b_j(f)$ vary continuously with f in Y. For the interval I_j, we let $P_j(f, .)$ be the Taylor polynomial of degree $r-1$ of f for the midpoint of I_j.

If $g \to f$ in the norm of $W_1^r(\Omega)$, then $D^\nu g \to D^\nu f$ uniformly on Ω, for all $0 \leq \nu < r$. Since the midpoints of I_j are a continuous function of f, we have that the $P_j(g, .)$ converge to the $P_j(f, .)$ as $\|f - g\|_Y \to 0$.

In summary, we have given a continuous selection $\bar{a}(f)$ defined on $K_{p,r}$. We now check the approximation by $M_n(\bar{a}(f))$. Since $|I_j| \le 1/n$ and since by (5.4)

$$\int_{I_j} |f^{(r)}| \le 3/n, \tag{5.5}$$

we have from the remainder in Taylor's formula

$$\|f - P_j\|_{L_\infty(I_j)} \le |I_j|^{r-1} \int_{I_j} |f^{(r)}| \le 3n^{-r}. \tag{5.6}$$

This shows that $d_m(K_{p,r})_X \le Cm^{-r}$ when $m = (2n-1)r + n - 1$. For other values of m, this follows from the monotonicity of d_m.■

REFERENCES

1. K. Borsuk, *Drei Sätze über die n-dimensionale euklidische Sphäre*, Fund. Math. **20** (1933), 177–191
2. E.W. Cheney, "Approximation Theory," Chelsea, New York, 1982
3. R. DeVore, *Maximal functions and their application to rational approximation*, in "Second Edmonton Conference on Approximation Theory," Z. Ditzian, A. Meir, S. Riemenschneider, A. Sharma, eds. CMS Conference Proc., Vol. 2 ,1983, pp. 143–155
4. R.B. Holmes, "A Course in Optimization and Best Approximation," Springer Verlag, New York
5. J.P. Kahane, "Teoria Constructiva de Functiones," Course Notes, University Buenos Aires, 1961
6. B.S. Kashin, *Diameters of some finite dimensional sets and classes of smooth functions*, Math. USSR Isv. **11** (1977), 317–333
7. J.L. Kelley, "General Topology," D. Van Nostrand, New York, 1955
8. A. Pinkus, "N-Widths," Springer Verlag, New York, 1986
9. V. Popov, *Uniform rational approximation of the class V_r and its applications*, Acta Math. Acad. Sci. Hung. **29** (1977), 119–129

The address of the first two authors is: Department of Mathematics, University of South Carolina, Columbia, S.C. 29208.
The address of the third author is: IBM Research Center, P.O. Box 218, Yorktown Heights, N.Y. 10598.

(Received November 24, 1988;
in revised form February 10, 1989)