ASYMPTOTIC PROPERTIES OF BEST $L_2[0,1]$ APPROXIMATION BY SPLINES WITH VARIABLE KNOTS*

By

D. L. BARROW† AND P. W. SMITH**

Texas A&M University

Abstract. Let $S_N{}^k$ be the set of kth-order splines on [0,1] having at most $N-1$ interior knots, counting multiplicities. We prove the following sharp asymptotic behavior of the error for the best $L_2[0, 1]$ approximation of a sufficiently smooth function f by the set $S_N{}^k$:

$$\lim_{N\to\infty} N^k \operatorname{dist}(f, S_N{}^k) = (|B_{2k}|/(2k)!)^{1/2}\left(\int_0^1 |f^{(k)}(\tau)|^\sigma d\tau\right)^{1/\sigma},$$

where $\sigma = (k + 1/2)^{-1}$ and B_{2k} is the $2k$th Bernoulli number. Similar results have previously been obtained for piecewise polynomial (i.e., with no continuity constraints) approximation, but with different constant before the integral term. The approach we use is first to study the asymptotic behavior of $\operatorname{dist}(f, S_N{}^k(t))$, where $S_N{}^k(t)$ is the linear space of kth-order splines having simple knots determined from the fixed function t by the rule $t_i = t(i/N)$, $i = 0, \cdots, N$.

1. Introduction. In this paper we derive an expression for the exact asymptotic value of the error when approximating a smooth function by splines, as the number of knots tends to infinity. The approximation we have in mind is nonlinear and may be described for a given function f as follows: for a fixed positive integer k and $N = 1, 2, 3, \cdots$, find the function s^* which minimizes $\|f - s\|$ (the $L_2[0, 1]$ norm) among all functions s in the class of splines $S_N{}^k$. $S_N{}^k$ is the L_2-closure of the set of all splines of order k having at most $N-1$ simple knots in the interior of [0, 1]; equivalently, $S_N{}^k$ is the manifold of kth-order splines having at most $N-1$ interior knots, counting multiplicities. Thus, $\|f - s^*\| = \operatorname{dist}(f, S_N{}^k)$. A main result of this paper, Theorem 2, states that for f in $C^k[0, 1]$,

$$\lim_{N\to\infty} N^k \operatorname{dist}(f, S_N{}^k) = C_k \|f^{(k)}\|_\sigma, \tag{1.1}$$

where $C_k = (|B_{2k}|/(2k)!)^{1/2}$ (B_{2k} is the $2k$th Bernoulli number), $\sigma = (k + 1/2)^{-1}$, and the "norm" is that of $L_\sigma[0, 1]$.

Previous authors (Sacks and Ylvisaker [12, 13], Burchard [8], and McClure [11]) have given similar sharp error estimates, but with $S_N{}^k$ replaced by $P_N{}^k$, the space of *piecewise polynomials* of order k (degree $< k$) having at most $N-1$ break points in [0, 1]. Actually, these authors and others (e.g., Wahba [15], Dodson [9], and de Boor [4]) have given

* Received November 8, 1977; revised version received March 3, 1978.

† The work of this author was supported by the National Science Foundation under Grant MCS-77-02464.

** The research of this author was partially supported by the U. S. Army Research Office under Grant DAHC 04-75-G-0816.

results more general, and in other contexts, than best L_2 approximation from $P_N{}^k$, but this setting seems best for illustrating the essential difference between earlier work on variable knot approximations and the present paper.

Splines, after all, are piecewise polynomials, so what is the primary difference between approximating from $S_N{}^k$ and from $P_N{}^k$? To explain, we first fix a knot sequence $\boldsymbol{t} = (t_0, t_1, \cdots, t_n): 0 = t_0 < t_1 < \cdots < t_N = 1$. Then $S_N{}^k(\boldsymbol{t})$ (resp. $P_N{}^k(\boldsymbol{t})$) will denote the linear space of kth-order splines having simple knots at $t_1, \cdots, t_{N-1}$ (resp., functions which are polynomials of degree $< k$ on $(t_0, t_1), (t_1, t_2), \cdots, (t_{N-1}, t_N)$). The problem of finding the best L_2 approximant from $S_N{}^k$ (resp., $P_N{}^k$) can be solved by first finding the orthogonal projection $s(\boldsymbol{t}) \in S_N{}^k(\boldsymbol{t})$ (resp., $p(\boldsymbol{t}) \in P_N{}^k(\boldsymbol{t})$) of f, and then choosing $\boldsymbol{t}$ to minimize $\|f - s(\boldsymbol{t})\|$ (resp., $\|f - p(\boldsymbol{t})\|$). The difficulty in the spline case arises from the fact that $s(\boldsymbol{t})$ depends globally on f, while $p(\boldsymbol{t})$ depends only locally on f; that is, $p(\boldsymbol{t})\,|_{(t_i, t_{i+1})}$ depends only on $f|_{(t_i, t_{i+1})}$. Thus, the exact error $\|f - s(\boldsymbol{t})\|$ is more difficult to analyze than $\|f - p(\boldsymbol{t})\|$.

It is possible to obtain (asymptotic) bounds for N^k dist$(f, S_N{}^k)$ by observing that

$$S_N{}^k \subset P_N{}^k \subset S^k_{Nk-k+1}. \tag{1.2}$$

For example, McClure ([11], p. 24), using his results that N^k dist$(f, P_N{}^k) \to L_k \|f^{(k)}\|_\sigma$, obtains, for $N \to \infty$, $L_k \|f^{(k)}\|_\sigma \leq \underline{\lim}\, N^k$ dist$(f, S_N{}^k) \leq \underline{\lim}\, N^k$ dist$(f, S_N{}^k) \leq U_k \|f^{(k)}\|_\sigma$. To give an idea of the improvement available using (1.1) we compare in Table 1 some of the constants L_k, C_k, and U_k.

Note that as k increases, so does C_k/L_k. This illustrates the effect of the higher-order continuity constraint on functions in $S_N{}^k$, as opposed to no continuity constraints for $P_N{}^k$. One should also keep in mind, when comparing the above constants, that $k + (k + 1)(N - 1)$ parameters are needed to represent an arbitrary function in $P_N{}^k$, but only $k + 2(N - 1)$ for $S_N{}^k$.

It is of interest to note that L_k and C_k agree when $k = 1$ or 2. The reason for this when $k = 1$ is clear, since then $P_N{}^1 = S_N{}^1$. The case $k = 2$ is something of an exception, and deserves special comment. In case $f'' > 0$, the best approximants s^* and p^* from $S_N{}^2$ and $P_N{}^2$, respectively, are the same; that is, p^* is continuous, as may be verified by differentiating the error $\|f - p(t)\|^2$ with respect to the knots t. Thus, best approximation (of a convex function) by piecewise linear functions with variable knots is equivalent to best approximation by (continuous) second-order splines. Sacks and Ylvisaker ([12, 13]) have previously noted a similar equivalence in the case $k = 2$ ($k = 1$ in their notation).

Sec. 2 is concerned with the (linear) spline spaces $S_N{}^k(t)$ whose knots are determined by a fixed *distribution function t* according to the rule $t_i = t(i/N)$, $i = 0, 1, \cdots, N$. Sequences chosen this way have been called *regular sequences* by Sacks and Ylvisaker [12]. One

TABLE 1.

k	L_k	C_k	$U_k (= k^k L_k)$
1	2.89 (−1)	2.89 (−1)	2.89 (−1)
2	3.73 (−2)	3.73 (−2)	1.49 (−1)
4	1.98 (−4)	9.09 (−4)	5.08 (−2)
6	4.17 (−7)	2.30 (−5)	1.95 (−2)
8	4.67 (−10)	5.82 (−7)	7.84 (−3)
10	3.25 (−13)	1.47 (−8)	3.25 (−3)
12	1.54 (−16)	3.74 (−10)	1.38 (−3)

benefit of choosing knots this way is that one can (at least formally) interchange the operations of "finding optimal knots" and "letting $N \to \infty$". That is, one can first let $N \to \infty$, obtaining the asymptotic error as a functional in t, and then find the distribution t which minimizes the functional. Theorem 1 gives such a functional.

Theorem 2 (Eq. (1.1)) is proved in Sec. 3. In this section we also give a result, Theorem 3, on the limiting distribution, as $N \to \infty$, of the knots of best approximations in $S_N{}^k$ to a function f. Perhaps surprisingly, this limiting distribution turns out to be identical to the one which McClure [11] found to be the limit of the knots of best approximations from $P_N{}^k$.

Sec. 4 contains some concluding remarks about these results and discusses some generalizations and open problems. We state a conjecture (which has been supported by numerical experiments) that a certain interpolation projection onto $S_N{}^k(t)$ is asymptotically equivalent to the orthogonal projection.

2. L_2-projection onto the spline subspaces $S_N{}^k(t)$. In this section we derive an expression for the limiting value, as the number of knots tends to infinity, of the error of best $L_2[0, 1]$ approximation from certain spline spaces. These subspaces, denoted by $S_N{}^k(t)$, are described as follows. Let $t \in C^1[0, 1]$ satisfy $t(0) = 0$, $t(1) = 1$, and $t' > 0$ on $[0, 1]$. Let N and k be positive integers. Define the knot sequence $t = (t_0, t_1, \cdots, t_N)$ by $t_i = t(i/N)$, $i = 0, \cdots, N$. Then $S_N{}^k(t)$ is defined to be the vector space of splines of order k having simple knots at $t_1, t_2, \cdots, t_{N-1}$; that is, s belongs to $S_N{}^k(t)$ if and only if s belongs to the smoothness class $C^{k-2}[0, 1]$ and its restriction to each (t_i, t_{i+1}) is a polynomial of degree at most $k - 1$.

It will be convenient to introduce the following notation: for each fixed N, and $i = 0, 1, \cdots, N - 1$, let

$$h_i = t_{i+1} - t_i, \quad h = \max_i h_i, \quad \text{and} \quad \Delta = h/\min_i h_i .$$

Letting $0 < m = \min t' \le \max t' = M$, we see that

$$h \le M/N \quad \text{and} \quad \Delta \le M/m. \tag{2.1}$$

We will also make use of the *Bernoulli polynomials*

$$B_k(x) = x^k + \binom{k}{1}B_1\, x^{k-1} + \cdots + B_k, \quad 0 \le x \le 1,$$

where k is a non-negative integer and the B_s are the Bernoulli numbers. The reader is referred to [14] for a discussion of these functions. Their properties that we wish to note are, for $k \ge 1$,

$$\int_0^1 B_k(x)dx = 0, \tag{2.2}$$

$$B_k'(x) = k\, B_{k-1}(x), \tag{2.3}$$

and

$$B_k^{(j)}(0) = B_k^{(j)}(1) \quad \text{for} \quad j = 0, 1, \cdots, k-2 \quad \text{and} \quad k \ge 2. \tag{2.4}$$

Let $P_N(t): L_2[0, 1] \to S_N{}^k(t)$ be the orthogonal projection operator and let

$$C_k = (|B_{2k}|/(2k)!)^{1/2}.$$

In the following, all norms are those of $L_2[0, 1]$ unless otherwise noted. The main result of this section is the following theorem.

THEOREM 1. For $f \in C^k[0, 1]$,

$$\lim_{N\to\infty} N^k \|f - P_N(t)f\| = C_k\left(\int_0^1 |f^{(k)}(t(x))|^2(t'(x))^{2k+1}dx\right)^{1/2} \equiv (J(t))^{1/2}. \tag{2.5}$$

Proof: We will show that the error function $f - P_N(t)f$ is "close" to the function M_N defined by

$$M_N(\tau) = f^{(k)}(t_i)\frac{h_i^k}{k!}B_k\left(\frac{\tau - t_i}{h_i}\right), \quad t_i \le \tau < t_{i+1}, \quad i = 0, \cdots, N-1.$$

A calculation, using repeated integrations by parts and the properties (2.2)–(2.4) of Bernoulli polynomials, shows that $N^{2k}\|M_N\|^2$ is a Riemann sum for the integral term in (2.5), so that as $N \to \infty$, $N^{2k}\|M_N\|^2 \to J(t)$. Introducing $q_N = f - M_N$, we have

$$f - P_N(t)f = M_N + (q_N - P_N(t)q_N) - P_N(t)M_N.$$

Theorem 1 will be proved when we show that the second and third terms in the above equation are appropriately small:

LEMMA 1.

$$\lim_{N\to\infty} N^k\|q_N - P_N(t)q_N\| = 0.$$

LEMMA 2.

$$\lim_{N\to\infty} N^k\|P_N(t)M_N\| = 0.$$

Proof of Lemma 1: For $i = 1, 2, \cdots, N-1$, and $j = 0, 1, \cdots, k-2$, let J_i^j denote the magnitude of the jump in the jth derivative of M_N (and hence also of q_N) at t_i. Letting B_k^j denote $|B_k^{(j)}(0)|/k!$ and recalling (2.4), we have

$$\begin{aligned} J_i^j &= |M_N^{(j)}(t_i^+) - M_N^{(j)}(t_i^-)| = B_k^j|f^{(k)}(t_i)h_i^{k-j} - f^{(k)}(t_{i-1})h_{i-1}^{k-j}| \\ &= B_k^j|h_i^{k-j}(f^{(k)}(t_i) - f^{(k)}(t_{i-1})) + f^{(k)}(t_{i-1})(h_i^{k-j} - h_{i-1}^{k-j})| \\ &\le B_k^j(h^{k-j}\omega(f^{(k)}, h) + C\,h^{k-j-1}|h_i - h_{i-1}|), \end{aligned}$$

where $\omega(g,\cdot)$ denotes the modulus of continuity of a function g and C is some constant depending on f, k and j but not on N or i. Since t' is continuous, we have

$$|h_i - h_{i-1}| = |t'(\xi_i) - t'(\xi_{i-1})|/N \le \omega(t', 2/N)/N.$$

It follows that

$$J_i^j = o(h^{k-j}). \tag{2.6}$$

The notation $o(h^p)$ is to be interpreted as follows: for any $\epsilon > 0$, there is a number $\delta > 0$ (equivalently, an integer N_0), which may depend on f, k, and j but not on N or i, such that if $h \le \delta$ ($N \ge N_0$), then $|o(h^p)| < \epsilon h^p$.

We prove Lemma 1 by establishing the following stronger result: for $j = 0, 1, \cdots, k-1$,

$$\operatorname{dist}_\infty(q_N^{(j)}, S_N^{k-j}(t)) = o(h^{k-j}); \tag{2.7}$$

dist_∞ denotes distance in the norm of $L_\infty[0, 1]$, and the function $q_N^{(j)}$ is defined by differentiation on the subintervals (t_i, t_{i+1}), its value at the knots t_i being immaterial.

First consider $j = 0$. In view of (2.6) we can find a piecewise linear function l_0 such that $\|l_0\|_\infty = o(h^k)$ and $q_N + l_0$ is continuous. Furthermore, in view of the quasi-uniformity conditions (2.1), we have $\|l_0'\|_\infty = o(h^{k-1})$. Hence, if we mimic a proof of de Boor's ([4], Lemma 1), we have

$$\begin{aligned}\mathrm{dist}(q_N, S_N^k(t)) &\le \|l_0\|_\infty + \mathrm{dist}_\infty(q_N + l_0, S_N^k(t))\\ &\le \|l_0\|_\infty + \frac{hk}{2}\,\mathrm{dist}_\infty(q_N' + l_0', S_N^{k-1}(t))\\ &\le \|l_0\|_\infty + \frac{hk}{2}[\|l_0'\|_\infty + \mathrm{dist}_\infty(q_N', S_N^{k-1}(t))].\end{aligned}$$

Hence, (2.7) holds for $j = 0$ if it holds for $j = 1$. To prove it for $j = 1$, we find a piecewise linear function l_1, $\|l_1\|_\infty = o(h^{k-1})$, $\|l_1'\|_\infty = o(h^{k-2})$, such that $q_N' + l_1$ is continuous, and argue as above. Continuing in this manner, we reduce to the case when $j = k - 1$. To prove (2.7) when $j = k - 1$, we use Taylor expansions to obtain, for $t_i < \tau < t_{i+1}$,

$$f^{(k-1)}(\tau) = f^{(k-1)}(t_i) + f^{(k)}(\xi)(\tau - t_i) = f^{(k-1)}(t_i) + f^{(k)}(t_i)(\tau - t_i) + o(h),$$

and

$$M_N^{(k-1)}(\tau) = M_N^{(k-1)}(t_i) + M_N^{(k)}(t_i)(\tau - t_i) = M_N^{(k-1)}(t_i) + f^{(k)}(t_i)(\tau - t_i),$$

so that

$$q_N^{(k-1)}(\tau) = f^{(k-1)}(t_i) - M_N^{(k-1)}(t_i) + o(h).$$

Hence, (2.7) holds in this case, and Lemma 1 is proved.

Proof of Lemma 2: Let $\{N_{i,k}(\cdot)\}_{i=-k+1}^{N-1}$ be the normalized B-spline basis for $S_N^k(t)$ (see de Boor [2]). This basis has the following properties (with $t_{-k+1} = t_{-k+2} = \cdots = t_0 = 0$ and $t_N = t_{N+1} = \cdots = t_{N+k-1} = 1$):

$$N_{i,k}(\tau) = 0 \qquad \text{unless } t_i < \tau < t_{i+k},$$

$$N_{i,k}(\tau) \ge 0, \text{ and } \sum_i N_{i,k}(\tau) \equiv 1.$$

In the following we will delete the subscript k. Since the partitions determined by t are quasi-uniform, there is a constant C independent of N and i so that

$$\|N_i^{(j)}\|_\infty \le C\,N^j, \qquad j = 0, 1, \cdots, k - 1. \tag{2.8}$$

Now consider those $N_i(\cdot)$'s such that $i = 0, 1, \cdots, N - k$. For these functions,

$$N_i^{(j)}(t_i) = N_i^{(j)}(t_{i+k}) = 0, \qquad j = 0, 1, \cdots, k - 2.$$

We wish to show that M_N is "nearly" orthogonal to these functions.

$$\begin{aligned}\langle M_N, N_i\rangle &= \sum_{l=i}^{i+k-1} \int_{t_l}^{t_{l+1}} M_N(\tau)N_i(\tau)d\tau = \sum_l \int_{t_l}^{t_{l+1}} f^{(k)}(t_l)\frac{h_l^k}{k!}B_k\left(\frac{\tau - t_l}{h_l}\right)N_i(\tau)d\tau\\ &= \sum_l f^{(k)}(t_l)\left\{\left[\frac{h_l^{k+1}}{(k+1)!}B_{k+1}\left(\frac{\tau - t_l}{h_l}\right)N_i(\tau)\right]_{t_l}^{t_{l+1}} - \frac{h_l^{k+1}}{(k+1)!}\int_{t_l}^{t_{l+1}} B_{k+1}\left(\frac{\tau - t_l}{h_l}\right)N_i'(\tau)d\tau\right\}.\end{aligned}$$

We continue integrating by parts until the integrals are of the form

$$\int_{t_l}^{t_{l+1}} B_{2k-1}\left(\frac{\tau - t_l}{h_l}\right) N_i^{(k-1)}(\tau)\, d\tau .$$

These integrals are zero by (2.2), since $N_i^{(k-1)}$ is constant on each (t_l, t_{l+1}). The boundary terms give a finite sum of the terms

$$B_{k+j}(0) N_i^{(j-1)}(t_l)[h_{l-1}^{k+j} - h_l^{k+j}],\ j = 1, \cdots, k-1,\ l = i+1, \cdots, i+k,$$

each of which can be shown to be of order $o(h^{k+1})$ by arguing as in Lemma 1 and using (2.8). Hence, we have shown

$$\max_{0 \le i \le N-k} |\langle M_N, N_i\rangle| = o(h^{k+1}). \tag{2.9}$$

At this point we introduce the basis $\{\bar{N}_i\}_{i=-k+1}^{N-1}$ for $S_N{}^k(t)$ defined by

$$\bar{N}_i(\tau) = N_i(\tau)/(t_{i+k} - t_i)^{1/2}.$$

This basis is convenient because the following inequality is known to hold (cf. [3]): there is a constant $c^* > 0$ (depending only on the order k) such that for any coefficient vector $a = (a_{-k+1}, \cdots, a_{N-1}) \in R^{N+k-1}$ having Euclidean norm $||a||$,

$$c^* ||a|| \le \left|\left| \sum_{i=-k+1}^{N-1} a_i \bar{N}_i(\cdot) \right|\right| \le ||a||. \tag{2.10}$$

As is well known, the projection

$$P_N(t) M_N \equiv \sum_{i=-k+1}^{N-1} x_i \bar{N}_i$$

is obtained by solving the *normal equations*

$$A_N\, \boldsymbol{x} = \boldsymbol{c}, \tag{2.11}$$

where $\boldsymbol{c} \in R^{N+k-1}$, $c_i = \langle M_N, \bar{N}_i\rangle$, and A_N is the Gram matrix of the $\{\bar{N}_i\}$; that is, $A_N(i, j) = \langle \bar{N}_i, \bar{N}_j\rangle$.

Lemma 2 will follow, using (2.10), when we show two things:

(i) $||\boldsymbol{c}|| = o(h^k)$,

(ii) $||A_N{}^{-1}||$ has a bound independent of N.

To see (i), we must show that

$$||\boldsymbol{c}||^2 = \sum_{i=-k+1}^{-1} c_i^2 + \sum_{i=0}^{N-k} c_i^2 + \sum_{i=N-k+1}^{N-1} c_i^2$$

is $o(h^{2k})$. In the first and third sums,

$$c_i^2 = \langle M_N, \bar{N}_i\rangle^2 = \left(\int_{t_i}^{t_{i+k}} M_N(\tau)\cdot \bar{N}_i(\tau) d\tau\right)^2 = O(h^{2k+1}).$$

The middle sum has roughly $N \simeq 1/h$ terms, each of size $o(h^{2k+1})$ by (2.9), and hence (i) holds.

We next show that the $||A_N{}^{-1}||$ are uniformly bounded. Let $g \in L_2[0, 1]$, and set

$$P_N(t)\, g = \sum_{i=-k+1}^{N-1} y_i \bar{N}_i,$$

so that

$$A_N y = \boldsymbol{g} = (\cdots, \langle g, \bar{N}_i\rangle, \cdots)^T.$$

Using (2.10),

$$||P_N(t)g|| = ||\textstyle\sum y_i \bar{N}_i|| \geq c^* \, ||y|| = c^* \, ||A_N^{-1}\boldsymbol{g}||.$$

Taking the supremum over all g in $\{g: ||g|| \leq 1\}$, we obtain

$$1 \geq c^* \sup_{||\boldsymbol{g}|| \leq 1} ||A_N^{-1} g|| \geq c^* \epsilon \, ||A_N^{-1}||, \tag{2.12}$$

provided that for some $\epsilon > 0$, there is an ϵ-ball

$$B(0, \epsilon) = \{\boldsymbol{g}: ||\boldsymbol{g}|| \leq \epsilon\} \subset \{\boldsymbol{g}: g_i = \langle g, \bar{N}_i\rangle, \, ||g|| \leq 1\}. \tag{2.13}$$

The inequality (2.12) would give the desired result, if such an $\epsilon > 0$ exists and is independent of N. To see that this is true, let $\boldsymbol{a} \in R^{N+k-1}$, $||\boldsymbol{a}|| \leq \epsilon$, and define the linear functional $\lambda_{\boldsymbol{a}}$ on $S_N{}^k(t)$ by

$$\lambda_{\boldsymbol{a}}(\textstyle\sum b_i \bar{N}_i) = \sum a_i b_i \, .$$

Then since

$$|\lambda_{\boldsymbol{a}}(\textstyle\sum b_i \bar{N}_i)| \leq ||\boldsymbol{a}|| \; ||\boldsymbol{b}|| \leq \dfrac{1}{c^*} ||\sum b_i \bar{N}_i|| \; ||\boldsymbol{a}||,$$

it follows that $||\lambda_{\boldsymbol{a}}|| \leq ||\boldsymbol{a}||/c^*$. By the Hahn-Banach theorem λ_a can be extended to all of $L_2[0, 1]$ without increasing its norm; by the Riesz representation theorem there is some $g \in L_2[0, 1]$ such that

$$\lambda_{\boldsymbol{a}}(\textstyle\sum b_i \bar{N}_i) = \sum b_i \langle g, \bar{N}_i\rangle;$$

hence $a_i = \langle g, \bar{N}_i\rangle$ and $||g|| \leq ||\boldsymbol{a}||/c^* \leq \epsilon/c^*$. Hence, $||g|| \leq 1$ if $\epsilon = c^*$ and we have shown that (2.13) holds. This completes the proof of Lemma 2, and hence also of Theorem 1.

3. Asymptotic properties of best $L_2[0, 1]$ approximants from the spline manifolds $S_N{}^k$. Denote by $S_N{}^k$ the set of all kth-order splines on $[0, 1]$ having at most $N-1$ interior knots, multiplicities counted. For $f \in L_2[0, 1]$, we call $f_N \in S_N{}^k$ a *best $L_2[0, 1]$ approximant* to f if

$$||f - f_N|| = \inf_{s \in S_N{}^k} ||f - s|| = \operatorname{dist}(f, S_N{}^k) \, .$$

The existence of best approximants to a function f follows easily from the fact that $S_N{}^k$ is a closed set. Questions concerning their unicity, however, are more difficult to answer; the reader is referred to [1] and [10] for some partial results in this direction.

This section is concerned with asymptotic properties, as $N \to \infty$, of a sequence of best approximants $\{f_N\}$. Specifically, we will show that for f in $C^k[0, 1]$, N^k dist$(f, S_N{}^k)$ converges to a prescribed limit, and that the knots of the $\{f_N\}$ tend to a certain unique distribution. We begin with the following lemma, which is a corollary to Theorem 1.

LEMMA 3. If $f \in C^k[0, 1]$, then

$$\overline{\lim_{N\to\infty}} N^k \operatorname{dist}(f, S_N{}^k) \leq C_k \left(\int_0^1 |f^{(k)}(\tau)|^\sigma d\tau\right)^{1/\sigma} \equiv C_k ||f^{(k)}||_\sigma, \tag{3.1}$$

where $\sigma = (k + 1/2)^{-1}$ and $C_k = (|B_{2k}|/(2k)!)^{1/2}$, as before.

Proof: Instead of proving (3.1) directly, we first give the following motivation. Suppose that $f^{(k)}(x) \geq \delta > 0$ for $0 \leq x \leq 1$. Redefine the functional $J(t)$ of Theorem 1 in terms of the "knot density" function u, where u is related to t by $u(\tau) = (t^{-1})'(\tau)$. Thus,

$$\bar{J}(u) = J(t) = C_k^2 \int_0^1 |f^{(k)}(\tau)|^2/u(\tau)^{2k} d\tau, \tag{3.2}$$

where u is continuous, positive and $\int_0^1 u = 1$. By the methods of the calculus of variations, one finds that $\bar{J}$ is minimized uniquely by the function

$$u^*(\tau) = |f^{(k)}(\tau)|^{\sigma} \Big/ \int_0^1 |f^{(k)}(\tau)|^{\sigma} d\tau, \tag{3.3}$$

and

$$\bar{J}(u^*)^{1/2} = C_k \, ||f^{(k)}||_{\sigma}.$$

If we now define $t^* \in C^1[0, 1]$ by $(t^{*-1})' = u^*$, (3.1) follows immediately from Theorem 1 and the fact that $\operatorname{dist}(f, S_N^k) \leq ||f - P_N(t^*)f||$.

A further interesting observation in this case is that if t^* is substituted into (2.5), the integral term has a constant integrand. This may be interpreted as evidence supporting the notion that "good" knot sequences are ones which "balance the error", at least for large N.

We now continue with the proof of Lemma 3 for arbitrary $f \in C^k[0, 1]$. Define u^* by (3.3) and let $\{u_j\} \subset C[0, 1]$ be a sequence converging to u^* in such a way that $\bar{J}(u_j)$ converges to $\bar{J}(u^*)$, where each $u_j > 0$ and $\int_0^1 u_j = 1$. Letting $(t_j^{-1})' = u_j$ we deduce from Theorem 1 that for each j,

$$\overline{\lim_{N\to\infty}} \, N^k \operatorname{dist}(f, S_N^k) \leq \overline{\lim_{N\to\infty}} \, N^k \operatorname{dist}(f, S_N^k(t_j)) = J(t_j)^{1/2}.$$

Letting $j \to \infty$ gives (3.1), and Lemma 3 is proved.

As might be expected, since u^* actually minimizes $\bar{J}(u)$, the upper bound in (3.1) cannot be improved and is actually sharp, as will be shown below. Furthermore, we will show that knot sequences of best approximants from S_N^k will converge, in a sense to be made precise, to the *asymptotically optimal distribution function* t^*. It is worth noting that the distribution function t^* (or equivalently, the density u^*) is identical to that used by other authors [4, 5, 6, 9, 11, 12] for choosing knot sequences, but in contexts somewhat different from the present one.

THEOREM 2. If $f \in C^k[0, 1]$, then

$$\lim_{N\to\infty} N^k \operatorname{dist}(f, S_N^k) = C_k ||f^{(k)}||_{\sigma}. \tag{3.4}$$

Proof: In view of Lemma 3, we have only to prove

$$\underline{\lim}_{N\to\infty} N^k \operatorname{dist}(f, S_N^k) \geq C_k ||f^{(k)}||_{\sigma}. \tag{3.5}$$

Before proving (3.5) in general, we first consider some special cases.

Case (i). $f(t) = t^k/k!$. Let $\mathring{S}_N^k$ denote the set of splines on [0, 1] having at most $N - 1$ interior, simple knots and satisfying the boundary conditions $s^{(j)}(0) = s^{(j)}(1) = 0$, for $j = k$

$-2i-1$, $i = 1, 2, \cdots$. It is known that $N^k \operatorname{dist}(\bar{f}, \mathring{S}_N{}^k) = C_k$, where $\bar{f} = B_k(\cdot)/k!$, by a theorem of Schoenberg [14, Theorem 6]. A complete proof of this and more general results may be found in a paper of Jetter and Lange [10]. Since $S_N{}^k$ is a subset of the L_2 closure of $\mathring{S}_{N+2k}{}^k$, we have

$$N^k \operatorname{dist}(f, S_N{}^k) = N^k \operatorname{dist}(\bar{f}, S_N{}^k) \geq N^k \operatorname{dist}(\bar{f}, \mathring{S}_{N+2k}{}^k) = (N/N+2k)^k C_k ,$$

from which (3.5) follows.

Case (ii). $f \in C^k[0, 1]$, $|f^{(k)}| \geq \delta > 0$. Suppose (3.5) were false in this case. Then for an increasing sequence $\{N_j\}$,

$$N_j{}^k \operatorname{dist}(f, S_{N_j}{}^k) \equiv N_j{}^k \,||f - f_{N_j}|| < d_k ||f^{(k)}||_\sigma , \tag{3.6}$$

where $0 < d_k < C_k$. For ease of notation, we will delete the subscript j. For each $m = 1, 2, 3, \cdots$, and for N sufficiently large, subdivide the interval $[0, 1]$ into finitely many closed subintervals $I_l = [\alpha_l , \alpha_{l+1}]$ whose end points coincide with knots of f_N, in such a way that each I_l contains in its interior $m_l - 1$ knots of f_N, where $m \leq m_l \leq m + k - 1$. Thus, $\sum m_l \equiv \bar{N} \leq N$, with equality in case all the interior α_l are simple knots of f_N. The inequality (3.6) implies that

$$\sum_l \left(\frac{\bar{N}}{m_l}\right)^{2k} m_l{}^{2k} \int_{I_l} (f - f_N)^2 \leq N^{2k} ||f - f_N||^2 < d_k{}^2 \left(\int_0^1 |f^{(k)}|^\sigma\right)^{2k+1} . \tag{3.7}$$

We claim that (3.7) implies that for some index l,

$$m_l{}^{2k} \int_{I_l} (f - f_N)^2 < d_k{}^2 \left(\int_{I_l} |f^{(k)}|^\sigma\right)^{2k+1} . \tag{3.8}$$

Indeed, the assumption of (3.8) false for every l, together with the following lemma, contradicts (3.7).

LEMMA 4. Let $a_l > 0$, $b_l \geq 0$, $l = 1, \cdots, n$ be numbers satisfying $\sum a_l = \sum b_l = 1$. Then for any $p > 1$,

$$1 \leq \sum_{l=1}^n a_l{}^{1-p} b_l{}^p,$$

with equality holding if and only if $a_l = b_l$ for each l.

The lemma is a form of Hölder's inequality: $1 = \sum(b_l/a_l)a_l \leq (\sum(b_l/a_l)^p a_l)^{1/p}$. In the application, take $p = 2k + 1$, $a_l = m_l/\bar{N}$, and

$$b_l = \int_{I_l} |f^{(k)}|^\sigma \Big/ \int_0^1 |f^{(k)}|^\sigma .$$

Denote the interval for which (3.8) holds by I^j. We assume, by passing to a subsequence if necessary, that each m_l in (3.8) is some fixed $\bar{m}$, $m \leq \bar{m} \leq m + k - 1$. As $j \to \infty$, it must be that length $(I^j) \to 0$, as follows from the assumption $|f^{(k)}| \geq \delta > 0$. If we now introduce a change of variables into the integrals of (3.8) so that the integrations are over $[0, 1]$, and then let j tend to infinity, we obtain

$$\bar{m}^k \operatorname{dist}(t^k/k!, S_{\bar{m}}{}^k) \leq d_k . \tag{3.10}$$

Since $\bar{m}$ can be arbitrarily large, (3.10) contradicts case (i), and so case (ii) is proven.

Case (iii). $f \in C^k[0, 1]$. Let $A = \{x \in [0, 1]: f^{(k)}(x) = 0\}$. If (3.5) were false, then for some increasing sequence $\{N_j\}$, (3.6) would hold, so that (we again suppress the subscript j)

$$N^{2k}\int_0^1 (f - f_N)^2 < d_k^2\left(\int_0^1 |f^{(k)}|^\sigma\right)^{2k+1} = d_k^2\left(\int_{[0,1]\setminus A} |f^{(k)}|^\sigma\right)^{2k+1}$$

$$\leq \bar{d}_k^2\left(\sum_{l=1}^{n}\int_{I_l} |f^{(k)}|^\sigma\right)^{2k+1}, \tag{3.11}$$

where $0 < d_k < \bar{d}_k < C_k$ and $\{I_l\}_{l=1}{}^n$ are closed intervals where $|f^{(k)}| \geq \delta > 0$. Let m_l be the number of knots of f_N interior to I_l, so that

$$\sum_{l=1}^{n} m_l \equiv \bar{N} \leq N - 1 < N.$$

Using Lemma 4 in an argument similar to that of case (ii), we deduce that for some index l,

$$m_l^{2k}\int_{I_l} (f - f_N)^2 < \bar{d}_k^2\left(\int_{I_l} |f^{(k)}|^\sigma\right)^{2k+1}. \tag{3.12}$$

We may assume, without loss of generality, that l is the same for each j. Furthermore, it is clear that $m_l \to \infty$ as $j \to \infty$, and so (3.12) leads to a contradiction of case (*ii*). This completes the proof of Theorem 2.

The remainder of this section will be devoted to describing how the knots of best approximations $\{f_N\}$ to a function $f \in C^k[0, 1]$ are distributed, as N tends to infinity. The treatment will be similar to that given by McClure ([9], §7) for best $L_p[0, 1]$ approximation by piecewise polynomials with variable knots.

Let $f \in C^k[0, 1]$, $f^{(k)}$ not identically zero, and let $\{f_N\}$ be a sequence of best approximations to f, $f_N \in S_N{}^k$. Denote the knots of each f_N by $\boldsymbol{t}^N = (t_0{}^N, t_1{}^N, \cdots, t_N{}^N)$, $0 = t_0{}^N < t_1{}^N \leq \cdots \leq t_{N-1}^N < t_N{}^N = 1$. Let $T^N \in C[0, 1]$ be the piecewise linear function satisfying $T^N(i/N) = t_i{}^N$, $i = 0, \cdots, N$. Denote by s_N the right-continuous function which is inverse to T^N on open intervals where T^N is strictly increasing; hence, each s_N is strictly increasing on [0, 1], and is discontinuous at points which are multiple knots of f_N. Finally, let s^* be the function

$$s^*(\tau) = \int_0^\tau |f^{(k)}(t)|^\sigma dt \Big/ \int_0^1 |f^{(k)}(t)|^\sigma dt;$$

s^* is thus the inverse (in case one exists) of the function t^* discussed earlier. We are now prepared to state

THEOREM 3. Let $f \in C^k[0, 1]$, $f^{(k)} \not\equiv 0$, and let the functions $\{s_N\}_{N=1}^\infty$ and s^* be as defined above. Then

$$\lim_{N\to\infty} s_N(\tau) = s^*(\tau), \qquad \text{uniformly on} \quad [0, 1]. \tag{3.13}$$

Proof: By Helley's selection theorem (see [16]) there is a subsequence $\{s_{N_j}\}$ and a function $\bar{s}$ such that $\lim_{j\to\infty} s_{N_j}(\tau) = \bar{s}(\tau)$ for all $\tau \in [0, 1]$. We prove the theorem by showing that $\bar{s}(\tau) = s^*(\tau)$ for all $\tau \in (0, 1)$. The uniformity of convergence follows from the continuity of s^* and the monotonicity of the $\{s_N\}$.

Let $0 < \tau < 1$. We first observe that Theorem 2 can easily be shown to hold when the interval [0, 1] is replaced by any interval [a, b], with norms and integrals changed appropriately. Secondly, $s_N(\tau)N$ is roughly the number of knots of f_N in the interval $[0, \tau]$;

i.e., within a finite number bounded by k. These two remarks establish that

$$C_k\left[\int_0^\tau |f^{(k)}|^\sigma\right]^{1/\sigma} \le \varliminf_{N\to\infty} (s_N(\tau)N)^k\left(\int_0^\tau (f-f_N)^2\right)^{1/2};$$

a similar inequality holds for the interval $[\tau, 1]$. Hence, we have

$$C_k^2\left[\int_0^1 |f^{(k)}(\tau)|^\sigma\right]^{2k+1}$$

$$= \lim_{j\to\infty} N_j^{2k}\int_0^1 (f-f_{N_j})^2 \ge \varliminf_{j\to\infty} (s_{N_j})(\tau)N_j)^{2k}\int_0^\tau (f-f_{N_j})^2/(s_{N_j}(\tau))^{2k}$$

$$+ \varliminf_{j\to\infty} ((1-s_{N_j}(\tau))N_j)^{2k}\int_\tau^1 (f-f_{N_j})^2/(1-s_{N_j}(\tau))^{2k}$$

$$\ge C_k^2\left[\int_0^\tau |f^{(k)}|^\sigma\right]^{2k+1}\Big/(\bar{s}(\tau))^{2k} + C_k^2\left[\int_\tau^1 |f^k|^\sigma\right]^{2k+1}\Big/(1-\bar{s}(\tau))^{2k}. \tag{3.14}$$

If we now divide by the left side of (3.14) and apply Lemma 4, we obtain $\bar{s}(\tau) = s^*(\tau)$ as desired, and Theorem 3 is proved.

4. Conclusion. The idea of placing the knots optimally or near-optimally in spline approximation is certainly not new. Dodson in his thesis [9] proposed an algorithm for knot placement motivated by minimizing an upper bound for the error (as opposed to minimizing the error). In fact, his algorithm attempts to achieve the knot distribution of Theorem 3, which we have seen is (asymptotically) the best possible. Similarly, de Boor [5] proposed computational schemes for solving ordinary differential equation boundary-value problems by collocation, in which the knot sequence was chosen to minimize an error bound.

One motivation for studying asymptotic properties of variable knot spline approximations is the belief that such qualitative information might prove useful in cases involving a large or even a moderate number of knots. This was demonstrated in the numerical results of Dodson [9] using the algorithm mentioned above. Moreover, since finding the knots of a best approximant to f from S_N^k is a nonlinear computational problem, the asymptotic limit of such knot sequences might be useful for initializing an iterative procedure to compute these optimal knots.

More general versions of Theorems 1, 2, and 3 would be desirable. It is possible to prove these theorems under the slightly weaker hypotheses that $f \in C^{k-1}$ and $f^{(k)}$ is piecewise continuous. Unfortunately we have not been able to prove similar results for the important functions $f(t) = t^\alpha$, $\alpha > -1/2$. In addition, it would be nice to extend these results to the spaces $L_p[0, 1]$, $1 \le p \le \infty$; it seems, however, that some new techniques must be used since we rely heavily on the linearity of the best approximation operator from L_2 onto a subspace.

We conclude with a conjecture which is motivated by the location of the zeros of the function M_N used in the proof of Theorem 1. If we assume that the error $f - P_N(t)f$ of best L_2 approximation from $S_N^k(t)$ is approximately equal to M_N, then it seems reasonable to try to approximate $P_N(t)f$ by interpolating at the zeros of scaled Bernoulli polynomials. To be more explicit, let t as before be a distribution function on [0, 1], which, for each N determines the knots $t_i = t(i/N)$ for the spline subspace $S_N^k(t)$. Let $\{\xi_i\}_{i=1}^{N+k-1}$ be some of the zeros of the function ϕ defined by $\phi(\tau) = B_k((\tau - t_i)/(t_{i+1} - t_i))$, $t_i \le \tau < t_{i+1}$, where

$B_k(\cdot)$ is the kth Bernoulli polynomial. The points $\{\xi_i\}$ are chosen so that the operation of interpolation at these points by a function in $S_N{}^k(t)$ is well defined. Thus, $I_N f$ will denote the unique spline in $S_N{}^k(t)$ which interpolates f at the $\{\xi_i\}$. We conjecture that

$$\lim_{N\to\infty} N^k \, ||f - I_N f|| = C_k \left(\int_0^1 |f^{(k)}(t(x))|^2 (t'(x))^{2k+1} dx \right)^{1/2} . \tag{4.1}$$

This result, if true, would produce via interpolation approximants which are asymptotically as good as best L_2 approximants. This is of computational interest since interpolants are usually easier to obtain than orthogonal projections. We have some strong computational evidence indicating that (4.1) is true.

REFERENCES

[1] D. Barrow, C. Chui, P. Smith, and J. Ward, *Unicity of best mean approximation by second order splines with variable knots,* Math. Comp., to appear

[2] C. de Boor, *On calculating with B-splines,* J. Approximation Theory **6,** 50–62 (1972)

[3] C. de Boor, *The quasi-interpolant as a tool in elementary polynomial spline theory,* in *Approximation theory* (G. G. Lorentz, ed.), Academic Press, New York, 1973, 269–276

[4] C. de Boor, *Good approximation by splines with variable knots,* in *Spline functions and approximation theory,* A. Meir and A. Sharma, eds., Birkhoüser, Basel, 1972, 57–72

[5] C. de Boor, *Good approximation with variable knots, II,* in *Conference on the numerical solution of differential equations,* Dundee, 1973, Springer Lecture Notes, vol. 363, 1974, 12–20

[6] H. G. Burchard, *Splines (with optimal knots) are better,* J. Applicable Analysis **3,** 309–319 (1974)

[7] H. G. Burchard and D. F. Hale, *Piecewise polynomial approximation on optimal meshes,* J. Approximation Theory **14,** 128–147 (1975)

[8] H. G. Burchard, *On the degree of convergence of piecewise polynomial approximation on optimal meshes,* Trans. Amer. Math. Soc. **234,** 531–559 (1977)

[9] D. S. Dodson, *Optimal order approximation by polynomial spline functions,* Ph.D. thesis, Purdue Univ., Lafayette, Ind., 1972

[10] K. Jetter and G. Lange, *Die Eindeutigkeit L_2-optimaler polynomialer Monospline,* Math. Z. (to appear).

[11] D. McClure, *Nonlinear segmented function approximation and analysis of line patterns,* Quart. Appl. Math. **33,** 1–37 (1975)

[12] J. Sacks and D. Ylvisaker, *Designs for regression problems with correlated errors III,* Ann. Math. Statist. **41,** 2057–2074 (1970)

[13] J. Sacks and D. Ylvisaker, *Statistical designs and integral approximation,* in *Proceedings of the 12th Biennial Canadian Mathematical Society Seminar* (Ronald Pyke, ed.), 1971, 115–136

[14] I. J. Schoenberg, *Monosplines and quadrature formulae,* in *Theory and applications of spline functions* (T. N. E. Greville, ed.), Academic Press, New York, 1969, 157–207

[15] G. Wahba, *On the regression problem of Sacks and Ylvisaker,* Ann. Math. Statist. **42,** 1035–1053 (1971)

[16] D. V. Widder, *The Laplace transform,* Princeton University Press, Princeton, 1946