JOURNAL OF APPROXIMATION THEORY 26, 219–231 (1979)

A Generalization of the Bramble-Hilbert Lemma and Applications to Multivariate Interpolation

DAVID S. WATKINS

Department of Mathematics, UMC 41, Utah State University, Logan, Utah 84322

Communicated by Arthur Sard

Received January 16, 1978

The Bramble–Hilbert lemma is a useful tool for proving error bounds for multivariate interpolation and approximation. It has two versions, a basic one and a more general "sharpened" version. In this paper we prove a generalization of the sharpened form of the lemma. The proof given here is much simpler than the original. We consider several multivariate interpolation schemes and compare the error bounds given by the (basic) lemma, its sharpened form, and the generalization proved here.

1. INTRODUCTION

The Bramble–Hilbert lemma [1, Theorem 2] and its sharpened form [2, Theorem 2] have been widely applied in proofs of error bounds. For example, the lemma was used in [3] to obtain finite element error bounds, and the sharpened form was applied to multivariate Hermite interpolation in [2]. The difference between the basic and the sharpened form of the lemma is this: In the basic form, the error bounds contain all derivatives of u of some order m, where u is the function which is being approximated. The sharpened form allows sometimes to discard certain of the m-th derivatives. In this paper we prove a generalization which allows us (in some cases) to throw away even more m-th derivatives at the expense of adding one or more higher order derivatives. Our compensation for adding higher order derivatives (and thus increasing the smoothness requirement on u) is that the terms associated with the higher derivatives become insignificant as the mesh parameter h tends to zero.

Our applications are to interpolation schemes in two and three dimensions. The interpolation schemes are of the type used in the finite element method [8, 10].

0021-9045/79/070219–13$02.00/0

2. Notation and Definitions

We will use the usual multiindex notation. A *multiindex* is an n-tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ whose entries are nonnegative integers. The *order* of α is $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$. The symbol D^α will denote the formal differential operator

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}.$$

Let Ω be a bounded domain in $\mathbb{R}^n$, and let $1 \leqslant p < \infty$. The norm on $L_p(\Omega)$ will be denoted $\|\cdot\|_{0,p}$. For any nonnegative integer m the *Sobelev space* $W^{m,p}(\Omega)$ is the space of functions $u \in L_p(\Omega)$ whose distributional derivatives $D^\alpha u$ are also in $L_p(\Omega)$, for all α of order less than or equal to m. (We have $W^{0,p}(\Omega) = L_p(\Omega)$.) The norm for $W^{m,p}(\Omega)$ is

$$\| u \|_{m,p} = \left[\sum_{|\alpha| \leqslant m} \| D^\alpha u \|_{0,p}^p \right]^{1/p}.$$

With this norm $W^{m,p}(\Omega)$ is a complete space. We also define a seminorm

$$| u |_{m,p} = \left[\sum_{|\alpha| = m} \| D^\alpha u \|_{0,p}^p \right]^{1/p}.$$

The seminorm differs from the norm in that the sum in the seminorm includes only those α whose order is exactly m.

For any nonnegative integer j, P_j will denote the space of n-variate polynomials of total degree not exceeding j.

Note that $P_{m-1} = \{u \in W^{m,p}(\Omega) \mid | u |_{m,p} = 0\}$.

If $k < m$ the space $W^{m,p}(\Omega)$ is a subset of $W^{k,p}(\Omega)$, and the embedding map $I: W^{m,p}(\Omega) \to W^{k,p}(\Omega)$ is clearly bounded. If Ω is bounded and has a continuous boundary [4] the embedding is even compact. This means that every sequence which is bounded in the norm of $W^{m,p}(\Omega)$ has a subsequence which is convergent in $W^{k,p}(\Omega)$. From this point on we will assume that Ω is a domain for which this *compact embedding theorem* holds. See [4, p. 108, Theorem 6.3] for a proof.

Later we will have to make the stronger assumption that Ω satisfies the *strong cone condition*. That is, there exist open sets $S_1, \ldots, S_i$ covering Ω and cones $C_1, \ldots, C_i$ with vertices at the origin such that for each $x \in S_j \cap \Omega$, the cone $x + C_j$ is contained in Ω. In our applications Ω will be a square or a cube. These domains certainly satisfy the strong cone condition.

Throughout this paper the letter C will be used to denote a generic constant whose value will generally not be the same from one place to the next.

3. The Bramble-Hilbert Lemma and Its Generalizations

Theorem 1 (Bramble–Hilbert Lemma). *Let $A\colon W^{m,p}(\Omega) \to Y$ be a bounded linear operator with domain $W^{m,p}(\Omega)$ and range in a normed linear space $(Y, \|\cdot\|)$. (Thus there exists a constant $\|A\|$ such that $\|Au\| \leqslant \|A\| \cdot \|u\|_{m,p}$ for all $u \in W^{m,p}(\Omega)$.) Suppose also that $A(P_{m-1}) = 0$. Then there is a constant C (which depends on Ω, m, and p, but not on A) such that*

$$\|Au\| \leqslant C \cdot \|A\| \cdot |u|_{m,p} \qquad \forall u \in W^{m,p}(\Omega).$$

The original statement of the Bramble–Hilbert lemma refers to a functional rather than an operator. The switch from functional to operator makes application of the theorem easier and does not cause any complications in the proof. Theorem 1 is a special case of Theorem 3, which we prove below.

Before we can state the sharpened form of the Bramble–Hilbert lemma, we must introduce some new notation. Let K_1 be the set of all multiindices of order exactly m, and let K_0 be the subset consisting of all multiindices α such that $\alpha_j = m$ for some j, and $\alpha_i = 0$ if $i \neq j$. For any intermediate set K (i.e., $K_0 \subseteq K \subseteq K_1$) we define

$$P_K = \{u \in W^{m,p}(\Omega) \mid D^\alpha u = 0 \ \forall \alpha \in K\}.$$

It is easy to see that $P_{K_1} = P_{m-1}$ and P_{K_0} is the set of all polynomials whose degree *in each variable* is at most $m - 1$. (For example, if $n = 2$ and $m = 4$, P_{K_0} is the space of bicubic polynomials.) For any K with $K_0 \subseteq K \subseteq K_1$ we have $P_{K_1} \subseteq P_K \subseteq P_{K_0}$. Thus P_K has finite dimension, for $\dim P_K \leqslant \dim P_{K_0} = m^n$. With each K we associate a norm and a seminorm:

$$\|u\|_{K,p} = \left[\|u\|_{0,p}^p + \sum_{\alpha \in K} \|D^\alpha u\|_{0,p}^p\right]^{1/p} \tag{1}$$

$$|u|_{K,p} = \left[\sum_{\alpha \in K} \|D^\alpha u\|_{0,p}^p\right]^{1/p}. \tag{2}$$

Note that $P_K = \{u \in W^{m,p}(\Omega) \mid |u|_{K,p} = 0\}$.

We will now assume that Ω satisfies strong cone condition. Under this assumption the following result of Aronszajn and Smith holds: *The norm $\|\cdot\|_{K_0,p}$ is equivalent to the Sobolev norm $\|\cdot\|_{m,p}$.* The main ideas of the proof can be found in [6] or [7]. It follows easily that $\|\cdot\|_{K,p}$ is equivalent to $\|\cdot\|_{m,p}$ for any K for which $K_0 \subseteq K \subseteq K_1$.

Theorem 2 (Sharpened Bramble–Hilbert Lemma). *Let K be a set of multiindices with $K_0 \subseteq K \subseteq K_1$. Let $A\colon W^{m,p}(\Omega) \to Y$ be a bounded linear operator* *with domain* $W^{m,p}(\Omega)$ *and range in a normed linear space* $(Y, \|\cdot\|)$. (*Thus there exists a constant* $\|A\|$ *such that* $\|Au\| \leqslant \|A\| \cdot \|u\|_{K,p}$ *for all* $u \in W^{m,p}(\Omega)$.) *Suppose also that* $A(P_K) = 0$. *Then there is a constant* C (*which depends on* Ω, K, *and* p, *but not on* A) *such that*

$$\|Au\| \leqslant C \cdot \|A\| \cdot |u|_{K,p} \qquad \forall u \in W^{m,p}(\Omega).$$

Theorem 2 reduces to Theorem 1 in the case $K = K_1$. Theorem 2 is also a special case of Theorem 3.

Our generalization of Theorem 2 will allow sets K which contain multiindices of order other than m. Specifically, we will allow any finite set K which contains K_0. Such a K can be expressed in the form $K = K_0 \cup L$, where L is a (possibly empty) finite set of multiindices. For any such set let $W^{K,p}(\Omega)$ denote the set of all $u \in W^{m,p}(\Omega)$ such that the distributional derivatives $D^\alpha u$ are in $L_p(\Omega)$ for all $\alpha \in L$. In one of the applications which follow, the members of L will all have order less than m. In this case we have $W^{K,p}(\Omega) = W^{m,p}(\Omega)$. But in most of our applications the members of L will have order greater than m, in which case $W^{K,p}(\Omega)$ is a proper subset of $W^{m,p}(\Omega)$. If all members of L have order exactly m, we have the situation of Theorem 2. The expressions (1) and (2) define a norm $\|\cdot\|_{K,p}$ and seminorm $|\cdot|_{K,p}$ on $W^{K,p}(\Omega)$. With this norm $W^{K,p}(\Omega)$ is a complete space. As before let $P_K = \{u \in W^{K,p}(\Omega) \mid |u|_{K,p} = 0\}$. Then P_K is a finite dimensional space of polynomials satisfying $P_K \subseteq P_{K_0}$. Although L can be any set of multiindices, in any useful application every $\alpha \in L$ will satisfy $\alpha_i \leqslant m$, $i = 1,..., n$. The reason for this is that if β is any multiindex satisfying $\beta_i > m$ for some i, then $D^\beta u = 0$ for all $u \in P_{K_0}$, and therefore we could not change P_K by adjoining β to L.

THEOREM 3. *Let* K *be any set of multiindices containing* K_0. *Let* A: $W^{K,p}(\Omega) \to Y$ *be a bounded linear operator from* $W^{K,p}(\Omega)$ *into a normed linear space* $(Y, \|\cdot\|)$. (*Thus there exists a constant* $\|A\|$ *such that* $\|Au\| \leqslant \|A\| \cdot \|u\|_{K,p}$ *for all* $u \in W^{K,p}(\Omega)$.) *Suppose also that* $A(P_K) = 0$. *Then there is a constant* C (*which depends on* Ω, K, *and* p, *but not on* A) *such that*

$$\|Au\| \leqslant C \cdot \|A\| \cdot |u|_{K,p} \qquad \forall u \in W^{K,p}(\Omega).$$

Since P_K is finite dimensional, there exists a closed subspace X_K of $W^{K,p}(\Omega)$ such that

$$W^{K,p}(\Omega) = P_K \oplus X_K.$$

Theorem 3 is an easy consequence of the following key lemma.

LEMMA. *There is a constant* C *such that* $\|v\|_{K,p} \leqslant C\, |v|_{K,p}$ *for all* $v \in X_K$.

Thus $|\cdot|_{K,p}$ is a norm on X_K equivalent to $\|\cdot\|_{K,p}$. The constant which appears here is the same constant as in the statement of Theorem 3. It is clearly independent of A. We will deduce Theorem 3 first, then prove the lemma.

Proof of Theorem 3. Let $u \in W^{K,p}(\Omega)$. Then there exist unique $q \in P_K$ and $v \in X_K$ such that $u = q + v$. Since $A(P_K) = 0$ we have $Au = Av$. Thus

$$\| Au \| = \| Av \| \leqslant \| A \| \cdot \| v \|_{K,p}.$$

Now if we apply the lemma to v we get

$$\| Au \| \leqslant C \cdot \| A \| \cdot | v |_{K,p}.$$

If we can show that $| u |_{K,p} = | v |_{K,p}$, we will be done. But this follows immediately from the fact that $D^\alpha q = 0$ if $\alpha \in K$. ∎

Proof of Lemma. We use a variant of a well known argument. Suppose that no such constant exists. Then there is a sequence (w_j) of functions in X_K such that $\| w_j \|_{K,p} > j \, | w_j |_{K,p}$ for all j. We may assume that $\| w_j \|_{K,p} = 1$ for all j. Thus $| w_j |_{K,p} \to 0$. Since (w_j) is a bounded sequence in $W^{K,p}(\Omega) \subseteq W^{m,p}(\Omega)$, and $\| w_j \|_{m,p} \leqslant C \| w_j \|_{K,p}$, it follows by the compact embedding theorem that (w_j) has a subsequence (v_j) which is a Cauchy sequence in $L_p(\Omega)$. This subsequence is also a Cauchy sequence in $W^{K,p}(\Omega)$, for $| v_i - v_j |_{K,p} \to 0$ as $i, j \to \infty$. Thus as $W^{K,p}(\Omega)$ is complete, there exists $v \in W^{k,p}(\Omega)$ such that $\| v_j - v \|_{K,p} \to 0$. Now $\| v \|_{K,p} = \lim_{j\to\infty} \| v_j \|_{K,p} = 1$, so in particular $v \neq 0$. But note that $v \in X_K$ because X_K is closed. On the other hand, $| v |_{K,p} = \lim_{j\to\infty} | v_j |_{K,p} = 0$, so $v \in P_K$. Thus $v \in P_K \cap X_K = (0)$, a contradiction. ∎

This proof is not constructive, as it does not give us an upper bound for C. Dupont and Scott [11] have recently obtained a constructive proof of the Bramble–Hilbert lemma.

4. Applications to Interpolation in Two and Three Dimensions

We will consider several interpolation schemes. In each case we will compare the error bounds obtained by using the Bramble–Hilbert lemma, the sharpened Bramble–Hilbert lemma, and Theorem 3.

Quadratic Interpolation. Let $R = (a, b)^2$ be a square in the x–y plane. Suppose we subdivide R into small squares of side h. We restrict our attention to squares for convenience. Assume that $h \leqslant 1$, also for convenience. Let P be the space of all polynomials of the form

$$a + bx + cy + dx^2 + exy + fy^2 + gx^2y + hxy^2.$$

Thus

$$P = P_2 \oplus \langle x^2y, xy^2 \rangle.$$

Then given any function $u \in C(\bar{\Omega})$, there exists a unique interpolant $v = Bu$ such that $v \in C(\bar{\Omega})$, the restriction of v to each small square is a polynomial in P, and v equals u at the corners and midsides of each square. (cf. [10, Sect. 7.3] or [8, Sect. 1.9]). We will call this interpolation scheme the *quadratic serendipity* interpolation scheme, in keeping with popular finite element terminology. We should write $B_h u$ instead of Bu, but we omit the h for typographic simplicity.

We wish to measure the error $u - Bu$ in Sobolev norms. This can be done by obtaining an error bound on each small square and summing the results, so let us focus on one of the squares, $S = \{(x, y) \mid c < x < c + h, d < y < d + h\}$. We will retain the symbols u and v to denote the restrictions of u and v to $\bar{S}$, and we will continue to write $v = Bu$. This causes no difficulties because v is determined locally by u. On $\bar{S}$, v is the unique polynomial in P which equals u at the four corners and four midsides of $\bar{S}$.

The error bound on S is obtained by transforming onto the unit square $\Omega = (0, 1)^2$ in the ξ–η plane and applying the Bramble–Hilbert lemma there. The affine map

$$x = c + \xi h, \qquad y = d + \eta h$$

is a one-to-one mapping of $\bar{\Omega}$ onto $\bar{S}$. This map induces a correspondence $w \leftrightarrow w'$ between functions on $\bar{S}$ and functions on $\bar{\Omega}$ as follows. Given w defined on $\bar{S}$, let w' be given by $w'(\xi, \eta) = w(x, y) = w(c + \xi h, d + \eta h)$. It is easy to verify that

$$\| D^\alpha w \|_{0,p,S} = h^{-|\alpha|} \mid J \mid^{1/p} \| D^\alpha w' \|_{0,p,\Omega} \tag{3}$$

whenever $D^\alpha w$ exists. Here $\mid J \mid = h^2$ is the Jacobian determinant of the transformation. $D^\alpha w$ is the derivative with respect to the x and y variables, whereas $D^\alpha w'$ is the derivative with respect to the ξ and η variables. We have appended an extra subscript on the norms to show that they correspond to different domains. It follows from (3) that

$$\mid w \mid_{m,p,S} = h^{-m} \mid J \mid^{1/p} \mid w' \mid_{m,p,\Omega}, \tag{4}$$

and since $h \leqslant 1$,

$$\| w \|_{r,p,S} \leqslant h^{-r} \mid J \mid^{1/p} \| w' \|_{r,p,\Omega}, \tag{5}$$

for all m and r for which the appropriate derivatives exist. It is also clear that if $v = Bu$, then v' is a polynomial in P (in the variables ξ and η), and v' is the unique polynomial in P which interpolates u' at the corners and midsides of

$\bar{\Omega}$. We define a linear operator B' by $B'u' = v'$. If we regard B' as an operator from $W^{m,p}(\Omega)$ into $W^{r,p}(\Omega)$ with $m > 2/p$, then we can use the Soblev lemma [4, p. 72, Theorem 3.8] and [5, p. 340, Theorem 1] to show that B' is bounded. Indeed, if $m > 2/p$, then $W^{m,p}(\Omega) \subseteq C(\bar{\Omega})$, so $B'u'$ is well defined for $u' \in W^{m,p}(\Omega)$. Furthermore, if we let $n_1, n_2, \ldots, n_8$ be the eight points of $\bar{\Omega}$ at which $v' = u'$ and let $\psi_1, \psi_2, \ldots, \psi_8$ be the unique functions in P such that $\psi_i(n_j) = \delta_{ij}$, then

$$v'(\xi, \eta) = \sum_{i=1}^{8} v'(n_i)\, \psi_i(\xi, \eta) = \sum_{i=1}^{8} u'(n_i)\, \psi_i(\xi, \eta).$$

Thus

$$\| B'u' \|_{r,p} = \| v' \|_{r,p} \leqslant \left(\sum_{i=1}^{8} \| \psi_i \|_{r,p} \right) \max\{ | u'(n_i) | \mid i = 1,\ldots, 8\}.$$

By the Sobolev lemma there is a constant C such that

$$\max\{ | u'(n_i) | \mid i = 1,\ldots, 8\} \leqslant C \| u' \|_{m,p}\,.$$

Therefore, if we absorb $\sum \| \psi_i \|_{r,p}$ into the constant we have

$$\| B'u' \|_{r,p} \leqslant C \| u' \|_{m,p}\,.$$

Define $A\colon W^{m,p}(\Omega) \to W^{r,p}(\Omega)$ by $Au' = u' - B'u'$. Obviously A is bounded if $m \geqslant r$ and $m > 2/p$. It is also clear that $A(P) = 0$. Thus, since $P_2 \subseteq P$, we can apply Theorem 1 (the Bramble–Hilbert lemma) with $m = 3$ to get

$$\| u' - B'u' \|_{r,p,\Omega} = \| Au' \|_{r,p,\Omega} \leqslant C \, | u' |_{3,p,\Omega}\,. \tag{6}$$

We have absorbed $\| A \|$ into the constant. We are now essentially done, for if we apply (5) with $w = u - Bu$ and (4) with $w = u$ and $m = 3$, we have

$$\| u - Bu \|_{r,p,S} \leqslant h^{-r} \, | J |^{1/p} \, \| u' - B'u' \|_{r,p,\Omega} \tag{7}$$

and

$$| u' |_{3,p,\Omega} = h^3 \, | J |^{-1/p} \, | u |_{3,p,S}\,. \tag{8}$$

We now string together (7), (6), and (8) to arrive at

$$\| u - Bu \|_{r,p,S} \leqslant Ch^{3-r} \, | u |_{3,p,S}\,.$$

Finally, taking p-th powers, summing over all small squares S, and taking p-th roots, we arrive at the following result.

THEOREM 4.1. *Let $u \in W^{3,p}(R)$, and let Bu be the quadratic serendipity interpolant of u. There exists a constant C (independent of h, u, and R) such that for $r = 0, 1$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{3-r} | u |_{3,p} .$$

We take $r \leqslant 1$ here because Bu is in $W^{1,p}(\Omega)$ but not $W^{2,p}(\Omega)$.

Before we examine how we can use Theorems 2 and 3 to refine this result, let us consider briefly another interpolation scheme, the *biquadratic* scheme [8, 10]. Let $P_{2,2}$ denote the nine-dimensional space of biquadratic polynomials, polynomials of degree at most two in each variable. Given $u \in C(\bar{R})$ there exists a unique interpolant $v = Bu$ such that $v \in C(\bar{R})$ and the restriction of v to each small square is a biquadratic polynomial which equals u at the corners, midsides, and center of each square. Since $P_2 \subseteq P_{2,2}$, we can apply the argument given above verbatim to obtain for the biquadratic scheme a theorem identical to Theorem 4.1. However, we can do better than this if we use Theorem 2, the sharpened form of the Bramble–Hilbert lemma. Since $P_{2,2} = P_{K_0}$, where $K_0 = \{(3, 0), (0, 3)\}$, we can use Theorem 2 with $m = 3$ and $K = K_0$ in place of Theorem 1. We need only replace (6) and (8) by

$$\| u' - B'u' \|_{r,p,\Omega} \leqslant C | u' |_{K_0,p,\Omega} \tag{6'}$$

and

$$| u' |_{K_0,p,\Omega} = h^3 | J |^{-1/p} [\| D^{(3,0)} u \|_{0,p}^p + \| D^{(0,3)} u \|_{0,p}^p]^{1/p} \tag{8'}$$

to obtain the result

$$\| u - Bu \|_{r,p} \leqslant Ch^{3-r} [\| D^{(3,0)} u \|_{0,p}^p + \| D^{(0,3)} u \|_{0,p}^p]^{1/p}. \tag{9}$$

Thus we have refined the error bound by removing the (2, 1) and (1, 2) derivatives from the right hand side.

In the case of the biquadratic scheme and many other interpolation schemes, Theorem 2 provides an aesthetically satisfying refinement of the error bound. However, for the quadratic serendipity scheme it does not. Certainly we will not be able to achieve (9) because the function $u(x, y) = x^2y^2$ is not interpolated exactly, yet the right hand side of (9) is zero for this function. On the other hand, we would expect to be able to improve on Theorem 4.1 because P $(= P_2 \oplus \langle x^2y, xy^2 \rangle)$ is strictly larger than P_2. Ideally, we would like to find a set of indices K of order 3 such that $P = P_K$. There is no such set. P is strictly contained in P_{K_0}, which is why we cannot attain (9). If we let $K_x = K_0 \cup \{(2, 1)\}$ and $K_y = K_0 \cup \{(1, 2)\}$, then both P_{K_x} and P_{K_y} are proper subsets of P. We can apply Theorem 2 with either

$K = K_x$ or $K = K_y$ to obtain an error bound from which the (1, 2) or (2, 1) derivative, respectively, has been deleted. Each of these results lacks symmetry but if we combine the two we get the following theorem.

THEOREM 4.2. *Let $u \in W^{3,p}(R)$, and let Bu be the quadratic serendipity interpolant of u. There exists a constant C (independent of h, u, and R) such that for $r = 0, 1$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{3-r}[\| D^{(3,0)}u \|_{0,p}^p + \| D^{(0,3)}u \|_{0,p}^p + \min\{\| D^{(1,2)}u \|_{0,p}^p , \| D^{(2,1)}u \|_{0,p}^p\}]^{1/p}.$$

Now let us see how we can use Theorem 3 to get an alternative refinement of Theorem 4.1. We have $P = P_K$, where $K = \{(3, 0), (0, 3), (2, 2)\}$, so we can apply Theorem 3 with $K = K_0 \cup L$, where $L = \{(2, 2)\}$. The space $W^{K,p}(R)$ is the set of all $u \in W^{3,p}(R)$ such that $D^{(2,2)}u \in L_p(R)$. In the error bound argument we have to replace (6) and (8) by

$$\| u' - B'u' \|_{r,p,\Omega} \leqslant C \mid u' \mid_{K,p,\Omega} \tag{6''}$$

and

$$\mid u' \mid_{K,p,\Omega} = h^3 \mid J \mid^{1/p}[\| D^{(3,0)}u \|_{0,p,S}^p + \| D^{(0,3)}u \|_{0,p,S}^p + h^p \| D^{(2,2)}u \|_{0,p,S}^p]^{1/p}. \tag{8''}$$

We used (3) to obtain (8″). We get the following result.

THEOREM 4.3. *Let $K = \{(3, 0), (0, 3), (2, 2)\}$, let $u \in W^{K,p}(R)$, and let Bu be the quadratic serendipity interpolant of u. There exists a constant C (independent of h, u, and R) such that for $r = 0, 1$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{3-r}[\| D^{(3,0)}u \|_{0,p}^p + \| D^{(0,3)}u \|_{0,p}^p + h^p \| D^{(2,2)}u \|_{0,p}^p]^{1/p}.$$

This result is almost as good as (9) because the extra term is insignificant for small h.

Quartic Interpolation. Let P be the 24-dimensional space of polynomials

$$P = P_4 \oplus \langle x^5, x^3y^2, x^2y^3, y^5, x^3y^3, x^4(3y^2 - 2y^3), x^5(3y^2 - 2y^3), (3x^2 - 2x^3)\, y^4, (3x^2 - 2x^3)\, y^5\rangle.$$

Let $u' \in C^2(\bar{\Omega})$. In [9] it is shown that there is a unique $v' \in P$ such that v' and its derivatives of order up to two (six values in all) interpolate those of u' at the four corners of $\bar{\Omega}$. We can use this local interpolation scheme and an affine transformation to define an interpolation scheme on any square: Given $u \in C^2(\bar{S})$, the affine transformation determines $u' \in C^2(\bar{\Omega})$. Let v' be the interpolant of u', and get $v = Bu$ by transforming v' back onto $\bar{S}$. Clearly v interpolates u and its derivatives of order up to two at the corners of $\bar{S}$. Now suppose $u \in C^2(\bar{R})$, where R is a large square which has been subdivided into small squares of size h. We can get an interpolant v by letting $v = Bu$ on each square. It can be shown [9] that $v \in C^1(\bar{R})$, that is, the function values *and* derivatives match up at the boundaries of the small squares. The interpolation scheme was especially designed for this, and it is this requirement which forces the nonmonomial terms such as $x^4(3y^2 - 2y^3)$ to appear in P. We will refer to this scheme as the *quartic* interpolation scheme.

With minor modifications we can apply the same argument as for the quadratic serendipity interpolation scheme to get an error bound for the quartic interpolation scheme. Since $P_4 \subseteq P$ we can use Theorem 1 with $m = 5$ to get the bound

$$\| u - Bu \|_{r,p} \leqslant Ch^{5-r} \mid u \mid_{5,p} \qquad \forall u \in W^{5,p}(\Omega) \tag{10}$$

Here we can take r as large as two because $Bu \in W^{2,p}(R)$.

We can use Theorem 2 to eliminate the (3, 2) and (2, 3) derivatives from (10). Indeed, if $K = \{(5, 0), (4, 1), (1, 4), (0, 5)\}$, then

$$P_K = P_4 \oplus \langle x^3y^2, x^2y^3, x^3y^3 \rangle \subseteq P.$$

Therefore we can apply Theorem 2 to get

$$\| u - Bu \|_{r,p} \leqslant Ch^{5-r}[\| D^{(5,0)}u \|_{0,p}^p + \| D^{(4,1)}u \|_{0,p}^p + \| D^{(1,4)}u \|_{0,p}^p + \| D^{(0,5)}u \|_{0,p}^p]^{1/p} \qquad \forall u \in W^{5,p}(R). \tag{11}$$

This is a fairly satisfactory result. The space P_K lacks only the terms x^5, y^5, and the nonmonomial terms. The latter terms are present only to allow a C^1 interpolant. From the point of view of approximation they are worthless, so we do not mind that P_K does not contain them. However, the terms x^5 and y^5 do have approximation theoretic value, and it would be nice if we could make use of it. It turns out that we can, in fact, use Theorem 3 to make use of these terms. The procedure is somewhat different this time. We take $m = 6$ (not 5) and let $K = K_0 \cup L$, where $K_0 = \{(6, 0), (0, 6)\}$ and $L = \{(4, 1), (1, 4)\}$. Then

$$P_K = P_4 \oplus \langle x^3y^2, x^2y^3, x^3y^3, x^5, y^5 \rangle \subseteq P,$$

and we can apply Theorem 3 to get the following result.

THEOREM 5. *Let $u \in W^{6,p}(R)$, and let Bu be the quartic interpolant of u. There exists C (independent of h, u, and R) such that for $r = 0, 1, 2$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{5-r}[\| D^{(4,1)}u \|_{0,p}^{p} + \| D^{(1,4)}u \|_{0,p}^{p}$$

$$+ h^{p}\{\| D^{(6,0)}u \|_{0,p}^{p} + \| D^{(0,6)}u \|_{0,p}^{p}\}]^{1/p}.$$

Quintic Interpolation: Let P be the 32-dimensional space

$$P = \langle x^i y^j \mid (i \leqslant 5) \text{ and } (j \leqslant 5) \text{ and } (i \leqslant 3 \text{ or } j \leqslant 3)\rangle.$$

Given $u \in C^2(\bar{R})$ there is a unique $v = Bu \in C^1(\bar{R})$ such that the restriction of v to each small square is a polynomial in P which interpolates u in the sense that at each corner of the square v and all of its derivatives of order up to two and also the third derivatives v_{xxy} and v_{xyy} are equal to those of u. This can be verified by the methods of [9]. We will refer to this scheme as the *quintic* interpolation scheme. This scheme is mentioned in [9] but not discussed in detail.

Since $P_5 \subseteq P$, we can apply Theorem 1 with $m = 6$ to get

$$\| u - Bu \|_{r,p} \leqslant Ch^{6-r} \mid u \mid_{6,p} \qquad \forall u \in W^{6,p}(R). \tag{12}$$

Because P contains many monomials which are not in P_5, we would expect to be able to discard many of the derivatives which appear in the seminorm in (12). In fact, we can discard four of the seven. If we let $K = \{(6, 0), (3, 3), (0, 6)\}$, then

$$P_K = P_5 \oplus \langle x^5y, x^4y^2, x^2y^4, xy^5, x^5y^2, x^2y^5\rangle \subseteq P,$$

so we can apply Theorem 2 to get

$$\| u - Bu \|_{r,p} \leqslant Ch^{6-r}[\| D^{(6,0)}u \|_{0,p}^{p} + \| D^{(3,3)}u \|_{0,p}^{p}$$

$$+ \| D^{(0,6)}u \|_{0,p}^{p}]^{1/p} \quad \forall u \in W^{6,p}(R). \tag{13}$$

With this K, P_K still lacks five of the monomials which are in P. We can remedy this by letting $K = \{(6, 0), (0, 6), (4, 4)\}$. Then $P_K = P$, and we can apply Theorem 3 to obtain the following theorem.

THEOREM 6. *Let $K = \{(6, 0), (0, 6), (4, 4)\}$, let $u \in W^{K,p}(R)$, and let Bu be the quintic interpolant of u. There exists C (independent of h, u, and R) such that for $r = 0, 1, 2$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{6-r}[\| D^{(6,0)}u \|_{0,p}^{p} + \| D^{(0,6)}u \|_{0,p}^{p} + h^{2p} \| D^{(4,4)}u \|_{0,p}^{p}]^{1/p}.$$

Interpolation in Three Dimensions. We will consider the three dimensional analogue of the quadratic serendipity interpolation scheme. Let R be a cube in three space, and suppose we have subdivided R into small cubes of side h. Let P be the 20-dimensional space of polynomials spanned by all monomials of the form $x^i y^j z^K$ for which $i, j, k \leqslant 2$ and two of i, j, and k are less than or equal to 1. Let $u \in C(\bar{R})$. Then there exists [10] a unique interpolant $v \in C(\bar{\Omega})$ such that the restriction of v to each of the small cubes is a polynomial in P which equals u at each of the eight vertices and twelve midsides. By *midside* we mean the middle of an edge of the cube.

Error bounds for three-dimensional (or n-dimensional) interpolation are proved in the same way as for two-dimensional interpolation. Since $P_2 \subseteq P$ we can apply Theorem 1 with $m = 3$ to get an error bound which contains all (ten) derivatives of order 3. Alternatively we can apply Theorem 3 with $K = K_0 \cup L$, where $K_0 = \{(3, 0, 0), (0, 3, 0), (0, 0, 3)\}$ and $L = \{(2, 2, 0), (2, 0, 2), (0, 2, 2)\}$. Then $P_K = P$, and we have the following theorem.

THEOREM 7. *Let $K = K_0 \cup L$, where $K_0 = \{(3, 0, 0), (0, 3, 0), (0, 0, 3)\}$ and $L = \{(2, 2, 0), (2, 0, 2), (0, 2, 2)\}$. Let $u \in W^{K,p}(R)$, and let Bu be the three-dimensional quadratic serendipity interpolant of u. There is a constant C (independent of h, u, and R) such that for $r = 0, 1$,*

$$\| u - Bu \|_{r,p} \leqslant Ch^{3-r}[| u |_{K_0,p}^p + h^p | u |_{L,p}^p]^{1/p}.$$

Thus, for $u \in W^{K,p}(R)$ we have bounded the error in terms of three third derivatives of u plus a term which tends to zero as $h \to 0$.

REFERENCES

1. J. H. BRAMBLE AND S. R. HILBERT, Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation, *SIAM J. Numer. Anal.* **7** (1970), 112–124.
2. J. H. BRAMBLE AND S. R. HILBERT, Bounds for a class of linear functionals with applications to Hermite interpolation, *Numer. Math.* **16** (1971), 362–369.
3. J. H. BRAMBLE AND M. ZLAMAL, Triangular elements in the finite element method, *Math. Comp.* **24** (1970), 809–820.
4. J. NEČAS, "Les méthodes directes en théorie des équations elliptiques," Academia, Prague, 1967.
5. V. I. SMIRNOV, "A Course of Higher Mathematics," Vol. V, "Integration and Functional Analysis," Addison–Wesley, Reading, Mass., 1964.
6. K. T. SMITH, Inequalities for formally positive integro-differential forms, *Bull. Amer. Math. Soc.* **67** (1961), 368–370.
7. K. T. SMITH, Formulas to represent functions by their derivatives, *Math. Ann.* **188** (1970), 53–77.
8. G. STRANG AND G. FIX, "An Analysis of the Finite Element Method," Prentice–Hall, Englewood Cliffs, N.J., 1973.

9. D. S. WATKINS, On the construction of conforming rectangular plate elements, *Int. J. Num. Meth. Enging.* **10** (1976), 925–933.
10. O. C. ZIENKIEWICZ, "The Finite Element Method in Engineering Science," McGraw-Hill, London, 1971.
11. T. DUPONT AND R. SCOTT, "Recent Advances in Numerical Analysis," Academic Press, New York, 1978.