SIAM J. MATH. ANAL.
Vol. 35, No. 5, pp. 1203–1212

THE BRAMBLE–HILBERT LEMMA FOR CONVEX DOMAINS*

S. DEKEL† AND D. LEVIATAN‡

Abstract. The Bramble–Hilbert lemma is a fundamental result on multivariate polynomial approximation. It is frequently applied in the analysis of finite elements methods (FEM) used for numerical solutions of PDEs. However, this classical estimate depends on the geometry of the domain and may "blow up" for simple examples such as a sequence of triangles of equivalent diameter that become thinner and thinner. Thus, in FEM applications one usually requires that the mesh has "quasi-uniform" geometry. This assumption is perhaps too restrictive when one tries to obtain estimates of nonlinear approximation methods that use piecewise polynomials.

Our main result that improves upon this point is the following. Let $\Omega \subset \mathbb{R}^n$ be a bounded convex domain and let $g \in W_p^m(\Omega)$, $m \in \mathbb{N}$, $1 \le p \le \infty$, where $W_p^m(\Omega)$ is the Sobolev space. Then there exists a polynomial P of total degree $m-1$ for which

$$|g - P|_{k,p} \le C(n,m)\big(\operatorname{diam}\Omega\big)^{m-k}|g|_{m,p}, \quad k = 0, 1, \dots, m,$$

where $|\cdot|_{k,p} := \sum_{|\alpha|=k}\|D^\alpha \cdot\|_{L_p(\Omega)}$ is the Sobolev seminorm of order k. As a consequence we get that for $f \in L_p(\Omega)$,

$$E_{m-1}(f,\Omega)_p \approx K_m\Big(f, \big(\operatorname{diam}\Omega\big)^m\Big)_p,$$

where $E_{m-1}(f,\Omega)_p := \inf_{P\in\Pi_{m-1}}\|f-P\|_{L_p(\Omega)}$ is the error of polynomial approximation of degree $m-1$ and $K_m(\ ,\ )_p$ is the K-functional associated with the pair $(L_p(\Omega), W_p^m(\Omega))$, and where the constants of equivalence depend only on m and n.

For the case of convex domains (elements) this extends a recent result for $p = 2$, and for $m = 1$ and $2 < p \le \infty$. This also improves previous results where the constant in the estimate further depends on the geometry of the domain, or where there is a constraint $p > n(\ge 2)$.

Key words. Bramble–Hilbert lemma, multivariate nonlinear approximation, finite element methods

AMS subject classifications. 41A10, 41A25, 41A63, 65M60

DOI. 10.1137/S0036141002417589

1. Introduction.

We begin by recalling classical smoothness measures over multivariate domains. Here and throughout the paper we assume that the domain $\Omega \subset \mathbb{R}^n$ is compact with a nonempty interior. A first notion of smoothness uses the *Sobolev spaces* $W_p^m(\Omega)$. These are spaces of functions $g \in L_p(\Omega)$ which have all their distributional derivatives of order up to m, $D^\alpha g := \frac{\partial^k g}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$, $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha \in \mathbb{Z}_+^n$, $|\alpha| := \sum_{i=1}^n \alpha_i = k$, $0 \le k \le m$, in $L_p(\Omega)$. The seminorm of $W_p^m(\Omega)$ is given by $|g|_{m,p} := \sum_{|\alpha|=m}\|D^\alpha g\|_{L_p(\Omega)} < \infty$ and may be regarded as a measure of the smoothness of order m of a function, provided the function is in the appropriate Sobolev space. The *K-functional* of order m of $f \in L_p(\Omega)$ (see, e.g., [De], [BeSh]) is defined by

$$K_m(f,t)_p := K\big(f, t, L_p(\Omega), W_p^m(\Omega)\big) := \inf_{g\in W_p^m(\Omega)}\{\|f-g\|_p + t|g|_{m,p}\}. \tag{1.1}$$

---

*Received by the editors November 12, 2002; accepted for publication (in revised form) August 8, 2003; published electronically January 6, 2004.

http://www.siam.org/journals/sima/35-5/41758.html

†RealTimeImage, 6 Hamasger St., Or-Yehuda 60408, Israel (shai.dekel@turboimage.com).

‡School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel (leviatan@math.tau.ac.il).

Since we assume Ω to be compact we may denote

$$K_m(f,\Omega)_p := K_m(f,d^m)_p, \tag{1.2}$$

where $d := \operatorname{diam}\Omega$.

For $f \in L_p(\Omega)$, $1 \le p \le \infty$, $h \in \mathbb{R}^n$, and $m \in \mathbb{N}$, we recall the mth order difference operator $\Delta_h^m(f,\cdot) : \Omega \to \mathbb{R}$

$$\Delta_h^m(f,x) := \Delta_h^m(f,\Omega,x) := \begin{cases} \sum_{k=0}^m (-1)^{m-k}\binom{m}{k} f(x+kh), & [x,x+mh] \subset \Omega, \\ 0 & \text{otherwise,} \end{cases}$$

where $[x,y]$ denotes the line segment connecting any two points $x,y \in \mathbb{R}^n$. The *modulus of smoothness* (see, e.g., [De], [BeSh]) is defined by

$$\omega_m(f,t)_p := \sup_{|h|\le t} \|\Delta_h^m(f,\Omega,\cdot)\|_{L_p(\Omega)}, \qquad t>0, \tag{1.3}$$

where for $h \in \mathbb{R}^n$, $|h|$ denotes the norm of h. We also denote

$$\omega_m(f,\Omega)_p := \sup_{h\in\mathbb{R}^n} \|\Delta_h^m(f,\Omega,\cdot)\|_{L_p(\Omega)}. \tag{1.4}$$

It is known that the above two notions of smoothness, (1.1) and (1.3), are sometimes equivalent (see section 5.4 in [BeSh] for the case $\Omega = \mathbb{R}^n$ and [JS] for the case of multivariate Lipschitz domains). That is, there exist $C_1, C_2 > 0$, such that for any $t > 0$

$$C_1 K_m(f,t^m)_p \le \omega_m(f,t)_p \le C_2 K_m(f,t^m)_p. \tag{1.5}$$

However, while it is easy to show that C_2 in (1.5) depends only on m (see [BeSh, (5.4.33)]), the constant C_1 may further depend on the geometry of Ω.

Let $\Pi_{m-1} := \Pi_{m-1}(\mathbb{R}^n)$ denote the multivariate polynomials of total degree $m-1$ (order m) in n variables. Given a "nontrivial" multivariate domain, our goal is to estimate the degree of approximation of a function $f \in L_p(\Omega)$, $1 \le p \le \infty$,

$$E_{m-1}(f,\Omega)_p := \inf_{P\in\Pi_{m-1}} \|f-P\|_{L_p(\Omega)},$$

using one of the above notions of smoothness. One of the classical results in this direction is the *Bramble–Hilbert lemma* [BrHi]. To introduce it we require the following definitions.

A domain Ω is *star-shaped with respect to a ball* $B \subseteq \Omega$ if for each point $x \in \Omega$, the closed convex-hull of $\{x\} \cup B$ is contained in Ω. Let $\rho_{\max} = \max\{\rho : \Omega$ is star-shaped with respect to a ball $B \subseteq \Omega$ of radius $\rho\}$. The *chunkiness parameter* of Ω is defined by

$$\gamma := \frac{d}{\rho_{\max}} \qquad (d = \operatorname{diam}\Omega). \tag{1.6}$$

This leads to the following formulation of the Bramble–Hilbert lemma (a weaker formulation estimates, instead, sublinear functionals; see Corollary 1.5).

BRAMBLE–HILBERT LEMMA. *Let Ω be star-shaped with respect to some ball B and let $g \in W_p^m(\Omega)$, $1 \le p \le \infty$, $m \in \mathbb{N}$. Then there exists a polynomial $P \in \Pi_{m-1}$ for which*

$$|g-P|_{k,p} \le C(n,m,\gamma) d^{m-k} |g|_{m,p}, \qquad k = 0,1,\dots,m. \tag{1.7}$$

See Chapter 4 in [BrSc] for a proof of this result and [H] for a slightly stronger version of (1.7). Obviously the main drawback of (1.7) is that the constant depends on the chunkiness parameter (1.6) which "blows up," for example, in the case of a sequence of triangles of equivalent diameter that become thinner and thinner. This problem is usually resolved in the finite elements methods (FEM) literature by assuming that the mesh is *quasi-uniform*, i.e., that the collection of domains (elements) used to discretize the given problem has a uniformly bounded chunkiness parameter.

Perhaps another limitation of (1.7) is that it is too restrictive to be applied in estimates in nonlinear approximation by piecewise polynomials. For instance, let $f \in L_p([0,1]^2)$ and define $S_N^m(\mathbb{R}^2)$ to be the collection

$$\sum_{k=1}^{N} \mathbf{1}_{\Delta_k} P_k,$$

where Δ_k are triangles with disjoint interiors and $P_k \in \Pi_{m-1}(\mathbb{R}^2)$, and we wish to estimate (see [KP], [DLS])

$$\sigma_{N,m}(f)_p := \inf_{\varphi \in S_N^m} \|f - \varphi\|_{L_p([0,1]^2)}.$$

Thus, there have been quite a few attempts at removing the dependence of the constants on the geometry of Ω, and of estimating them. Perhaps the most significant result has recently been obtained by Verfürth [V], in the case of convex domains and $p = 2$. Using the notation $H^m := W_2^m$, Verfürth has proved the following proposition.

PROPOSITION (see [V]). *Let Ω be a convex domain and let $g \in H^m(\Omega)$, $m \in \mathbb{N}$. Then there exists a polynomial $P \in \Pi_{m-1}$ for which*

$$|g - P|_{H^k} \le C(n,m) d^{m-k} |g|_{H^m}, \qquad k = 0, 1, \dots, m-1. \tag{1.8}$$

Also if $m = 1$, and if $g \in W_p^1$, $2 < p \le \infty$, then

$$\|g - P\|_{L_p(\Omega)} \le C(n,p) d |g|_{W_p^1}. \tag{1.9}$$

Verfürth gives concrete estimates of the above constants and has some further results for star-shaped domains as well.

Earlier, Dechevski, and Quak [DQ] improved the Bramble–Hilbert lemma in some cases. Their result applies to the larger class of domains that are star-shaped with respect to a point. A domain Ω is *star-shaped with respect to a point* $x_0 \in \Omega$ if for any point $x \in \Omega$ the line segment $[x_0, x]$ is contained in Ω. The following is a modified version of their result.

PROPOSITION (see [DQ]). *Let Ω be a Lipschitz domain, which is star-shaped with respect to a point $x_0 \in \Omega$. Then for $m \in \mathbb{N}$ and $2 \le n < p \le \infty$, there exists a polynomial $P \in \Pi_{m-1}$ for which*

$$|g - P|_{k,p} \le C(n,m,p) d^{m-k} |g|_{m,p}, \qquad k = 0, 1, \dots, m. \tag{1.10}$$

Although the constant in (1.10) does not depend on geometrical parameters such as (1.6), the above proposition assumes the constraint $n < p$ that does not cover one of the most common cases in applications of the FEM, namely, $n = p = 2$.

Our approach differs from previous work in one crucial detail. For convex domains we can construct an approximating polynomial that is more adaptive to the shape of the domain. Thus, instead of constructing a polynomial using either some center point $x_0 \in \Omega$ or some maximal but relatively small ball $B \subset \Omega$, our construction uses John's "maximal" ellipsoid (see Proposition 3.2) combined with a simple affine transformation argument. Our main result is

THEOREM 1.1. *Let $\Omega \subset \mathbb{R}^n$ be convex, and let $g \in W_p^m(\Omega)$, $m \in \mathbb{N}$, $1 \le p \le \infty$. Then there exists a polynomial $P \in \Pi_{m-1}$ for which*

$$|g - P|_{k,p} \le C(n,m) d^{m-k} |g|_{m,p}, \quad k = 0, 1, \ldots, m. \tag{1.11}$$

We emphasize that our proof of Theorem 1.1 is constructive and we are going to specify the polynomial P which yields (1.11). In fact we show that one may take $P(x) := Q^m(g(A\cdot)(A^{-1}x)$, where Q^m is the averaged Taylor polynomial over the ball $B(0,1) \subset \mathbb{R}^n$, and A is an affine transformation related to Ω (see definitions and details in sections 2 and 3).

A direct consequence of Theorem 1.1 is the following.

COROLLARY 1.2. *For all convex domains $\Omega \subset \mathbb{R}^n$ and functions $f \in L_p(\Omega)$, $1 \le p \le \infty$,*

$$E_{m-1}(f, \Omega)_p \approx K_m(f, \Omega)_p,$$

where $K_m(f, \Omega)_p$ is defined in (1.2)*, and the constants of equivalency depend only on m and n.*

We wish to point out a recent result of Karaivanov and Petrushev [KP] who showed that if $\Delta \subset \mathbb{R}^2$ is a triangle and $f \in L_p(\Delta)$, $0 < p \le \infty$, then for any $m \in \mathbb{N}$

$$E_{m-1}(f, \Delta)_p \le C(m,p) \omega_m(f, \Delta)_p, \tag{1.12}$$

where $\omega_m(f, \Delta)_p$ is defined in (1.4). This implies that for all triangles $\Delta \subset \mathbb{R}^2$ and functions $f \in L_p(\Delta)$, $1 \le p \le \infty$, we have the equivalence

$$E_{m-1}(f, \Delta)_p \approx \omega_m(f, \Delta)_p \approx K_m(f, \Delta)_p,$$

where the constants of equivalence depend only on p and m. Indeed, it is this result that motivated us to try to find shape-independent estimates.

We also get the following formulation of the Bramble–Hilbert lemma.

COROLLARY 1.3. *Let $\Omega \subset \mathbb{R}^n$ be convex, and let l be a sublinear functional given on $W_p^m(\Omega)$, $m \in \mathbb{N}$, $1 \le p \le \infty$, with the following properties.*

(i) *There exists a constant $\widetilde{C}$ such that for all $g \in W_p^m(\Omega)$, $|l(g)| \le \widetilde{C} \sum_{k=0}^m d^k |g|_{k,p}$;*
(ii) *$l(P) = 0$ for all $P \in \Pi_{m-1}$.*

Then for all $g \in W_p^m(\Omega)$,

$$|l(g)| \le C(n, m, \widetilde{C}) d^m |g|_{m,p}.$$

Section 2 reviews the averaged Taylor polynomial approach to the classical Bramble–Hilbert lemma (see Chapter 4 in [BrSc]). In section 3 we introduce John's theorem and explain how this tool can be applied in the case of convex domains via an affine transformation argument. Finally, in section 4 we assemble all the above tools to give a constructive proof of Theorem 1.1. We also define the notion of "almost convex" domains and note that our results extend to this case too.

2. The averaged Taylor polynomial.

We recall some basic definitions of multivariate polynomials, differentials, and Taylor series. Throughout this section we use the notation of Chapter 4 in [BrSc]. For a multi-index $\alpha \in \mathbb{Z}_+^n$ let $\alpha! = \prod_{i=1}^n \alpha_i!$, and denote by $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$ the *multivariate monomial of total degree* $|\alpha|$. Denote the set of all multivariate polynomials of total degree $m-1$ by

$$\Pi_{m-1}(\mathbb{R}^n) := \Big\{ \sum_{|\alpha|\le m-1} c_\alpha x^\alpha \Big\}.$$

The classical *Taylor polynomial of order* m (*degree* $m-1$) of a function $g \in C^m(\Omega)$ at $x \in \Omega$, about the point $y \in \Omega$, is given by

$$T_y^m g(x) := \sum_{|\alpha|<m} \frac{D^\alpha g(y)}{\alpha!}(x-y)^\alpha. \tag{2.1}$$

The *Taylor remainder of order* m of a function $g \in C^m(\Omega)$ at $x \in \Omega$, about the point $y \in \Omega$, is given by

$$TR_y^m g(x) := m \sum_{|\alpha|=m} \frac{(x-y)^\alpha}{\alpha!} \int_0^1 s^{m-1} D^\alpha g\big(x + s(y-x)\big) ds. \tag{2.2}$$

It is meaningful provided the segment $[y,x]$ is contained in Ω. Then we have

$$g(x) = T_y^m g(x) + TR_y^m g(x).$$

Next we introduce the averaged Taylor polynomial. It can be shown that for a ball $B(x_0,\rho) := \{z \in \mathbb{R}^n : |z - x_0| \le \rho\}$ there exists a *cut-off function* ϕ_B with the following properties:

(i) $\int_{\mathbb{R}^n} \phi_B(x)dx = 1$,
(ii) $\mathrm{supp}(\phi_B) = B$,
(iii) $\phi_B \in C^\infty(\mathbb{R}^n)$,
(iv) $\|\phi_B\|_\infty \le \rho^{-n}$.

Given $g \in C^m(\Omega)$, the *averaged Taylor polynomial of order* m (*degree* $m-1$) (averaged over a ball $B \subseteq \Omega$) is defined by

$$Q^m g(x) := \int_B T_y^m g(x)\phi_B(y)dy, \qquad x \in \Omega. \tag{2.3}$$

We also define the *averaged Taylor remainder*, namely,

$$R^m g(x) := g(x) - Q^m g(x). \tag{2.4}$$

The following lemma is a special case of the classical Bramble–Hilbert lemma which estimates the (simultaneous) degree of approximation of the averaged Taylor polynomial in a "normalized" setting. For the proof see Theorem 4.3.8 in [BrSc]; observe that the chunkiness parameter (1.6) in this case depends only on n.

LEMMA 2.1. *Let $B(0,1) \subseteq \Omega \subseteq B(0,n)$ be star-shaped with respect to $B(0,1)$. Then for any $g \in C^m(\Omega)$, $m \in \mathbb{N}$, and $1 \le p \le \infty$, we have*

$$|g - Q^m g|_{k,p} \le C(n,m)|g|_{m,p}, \qquad k = 0,1,\dots,m,$$

where Q^m is averaged over $B(0,1)$.

3. John's theorem.

Definition 3.1. *An ellipsoid E is the image of the closed unit ball in $\mathbb{R}^n$ under a nonsingular affine mapping $A(x) = Mx + b$, $M \in M_{n\times n}(\mathbb{R})$, $b \in \mathbb{R}^n$. The center of E is $b = A(0)$.*

The next result [J] (see also [Ba]) is the crucial ingredient that is missing in previous work. Let $c + n(E - c) := \{c + n(x - c) : x \in E\}$.

Proposition 3.2 (John's theorem). *Let $\Omega \subset \mathbb{R}^n$ be convex. Then there exists an ellipsoid $E \subseteq \Omega$ such that if x_0 is the center of E, then*

$$E \subseteq \Omega \subseteq x_0 + n(E - x_0).$$

By Definition 3.1, John's theorem implies that for each convex domain Ω we can find a nonsingular affine mapping A such that

$$B(0, 1) \subseteq A^{-1}(\Omega) \subseteq B(0, n).$$

It is interesting to note that John's ellipsoid is the ellipsoid $E \subseteq \Omega$ with maximal volume. In some sense this means that E "covers" Ω sufficiently well.

To use John's maximal ellipsoid (or equivalently, John's optimal affine transformation), we apply the following commutativity of Taylor polynomials and differentiation.

Lemma 3.3. *Let $A(x) = Mx + b$, $M \in M_{n\times n}(\mathbb{R})$, $b \in \mathbb{R}^n$, be a nonsingular affine mapping, and let $g \in C^m(\Omega)$. Then for any $x \in \Omega$, $y \in A^{-1}(\Omega)$, and $\alpha \in \mathbb{Z}^n_+$, $1 \le |\alpha| \le m - 1$, we have*

$$D^\alpha_x \Big[T^m_y \big(g(A\cdot)\big)(A^{-1}x) \Big] = T^{m-|\alpha|}_y \big((D^\alpha g)(A\cdot)\big)(A^{-1}x). \tag{3.1}$$

Proof. Observe that it is sufficient to prove that for any $1 \le k \le m - 1$ and $1 \le s \le n$,

$$D^{e_s}_x \Bigg[\sum_{|\beta|=k} \frac{D^\beta_y \tilde{g}(y)}{\beta!} (A^{-1}x - y)^\beta \Bigg] = \sum_{|\gamma|=k-1} \frac{D^\gamma_y \widetilde{g_{x_s}}(y)}{\gamma!} (A^{-1}x - y)^\gamma, \tag{3.2}$$

where $\tilde{g} := g(A\cdot)$, $\widetilde{g_{x_s}} := g_{x_s}(A\cdot)$, $g_{x_s} := \frac{\partial g}{\partial x_s}$, and $\{e_s\}_{s=1,\dots,n}$ is the standard basis of $\mathbb{R}^n$. The case of a general multivariate derivative D^α_x follows by repeated applications of (3.2), and the Taylor series formulation (3.1) is obtained by adding all the degrees $1 \le k \le m - 1$. To prove the above let $M =: (a_{i,j})_{1\le i,j\le n}$ and $M^{-1} =: (b_{i,j})_{1\le i,j\le n}$. In the calculations below, if $\beta_i = 0$, then differentiating $(A^{-1}x - y)^\beta$ with respect to x_s does not produce the term $\beta_i b_{i,s}(A^{-1}x - y)^{\beta - e_i}$; rather we have 0, and it does not appear in the summation. Hence in this case we regard $\beta_i b_{i,s}(A^{-1}x - y)^{\beta - e_i} := 0$ and $(\beta - e_i)! = \infty$, and again the term is not there. This takes care of itself automatically when we switch the summation below from β to $\gamma = \beta - e_i$.

$$\begin{aligned} D^{e_s}_x \Bigg[\sum_{|\beta|=k} \frac{D^\beta_y \tilde{g}(y)}{\beta!} (A^{-1}x - y)^\beta \Bigg] &= \sum_{|\beta|=k} \frac{D^\beta_y \tilde{g}(y)}{\beta!} D^{e_s}_x \big((A^{-1}x - y)^\beta\big) \\ &= \sum_{|\beta|=k} \frac{D^\beta_y \tilde{g}(y)}{\beta!} \sum_{i=1}^n \beta_i b_{i,s} (A^{-1}x - y)^{\beta - e_i} \\ &= \sum_{|\beta|=k} \sum_{i=1}^n \frac{D^\beta_y \tilde{g}(y)}{(\beta - e_i)!} b_{i,s} (A^{-1}x - y)^{\beta - e_i} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{|\gamma|=k-1} \frac{(A^{-1}x-y)^{\gamma}}{\gamma!} \sum_{i=1}^{n} b_{i,s} D_y^{\gamma+e_i} \tilde{g}(y)\\
&= \sum_{\gamma|=k-1} \frac{(A^{-1}x-y)^{\gamma}}{\gamma!} \sum_{i=1}^{n} b_{i,s} D_y^{\gamma} \Bigl(\sum_{j=1}^{n} a_{j,i} g_{x_j}(Ay) \Bigr)\\
&= \sum_{|\gamma|=k-1} \frac{(A^{-1}x-y)^{\gamma}}{\gamma!} \sum_{j=1}^{n} D_y^{\gamma}\bigl(g_{x_j}(Ay)\bigr) \sum_{i=1}^{n} a_{j,i} b_{i,s}\\
&= \sum_{|\gamma|=k-1} \frac{(A^{-1}x-y)^{\gamma}}{\gamma!} \sum_{j=1}^{n} D_y^{\gamma}\bigl(g_{x_j}(Ay)\bigr)\delta_{j,s}\\
&= \sum_{|\gamma|=k-1} \frac{D_y^{\gamma}(\widetilde{g_{x_s}}(y))}{\gamma!} (A^{-1}x-y)^{\gamma}\ . \qquad \square
\end{aligned}
$$

By (2.3), we have the following corollary.

COROLLARY 3.4. *Let $\Omega \subset \mathbb{R}^n$, and let A be a nonsingular affine mapping such that $B(0,1) \subseteq A^{-1}(\Omega)$. Then for $g \in C^m(\Omega)$ and $\alpha \in \mathbb{Z}_+^n$, $|\alpha| = k$, $1 \le k \le m-1$,*

$$
D^{\alpha}\Bigl[Q^m\bigl(g(A\cdot)\bigr)(A^{-1}x)\Bigr] = Q^{m-k}\bigl((D^{\alpha}g)(A\cdot)\bigr)(A^{-1}x), \tag{3.3}
$$

where Q^m is averaged on $B(0,1)$.

Observing that affine transformations map convex domains onto convex domains, the following argument, when combined with John's theorem, is the main tool of our approach.

LEMMA 3.5. *Let $\Omega \subset \mathbb{R}^n$, and let A be a nonsingular affine mapping such that $B(0,1) \subseteq A^{-1}(\Omega) \subseteq B(0,n)$ and $A^{-1}(\Omega)$ is star-shaped with respect to $B(0,1)$. Then for $g \in C^m(\Omega)$, $1 \le p < \infty$, and $P(x) = Q^m(g(A\cdot))(A^{-1}x)$ (where Q^m is averaged on $B(0,1)$), we have*

$$
|g-P|_{W_p^k(\Omega)} \le C(n,m) d^{m-k} |g|_{W_p^m(\Omega)}, \qquad k = 0,1,\dots,m. \tag{3.4}
$$

Proof. Since $A(x) = Mx + b$ maps $B(0,1)$ into Ω, we conclude that $\|M\|_2 \le d$. Thus with $M = (a_{i,j})_{1\le i,j\le n}$, we have that $\max_{1\le i,j\le n} |a_{i,j}| \le d$. Recalling that $\tilde{g} = g(A\cdot)$, this implies that for $y \in A^{-1}(\Omega)$, $x = Ay$, and $\alpha \in \mathbb{Z}_+^n$, $|\alpha| = i$, $i = 0,\dots,m$,

$$
|D_y^{\alpha}\tilde{g}(y)| \le d^i \sum_{|\gamma|=i} |D^{\gamma}g)(Ay)|,
$$

and hence, in particular,

$$
\sum_{|\alpha|=m} \|D_y^{\alpha}\tilde{g}\|_{L_p(A^{-1}(\Omega))} \le C(n,m) d^m \sum_{|\alpha|=m} \|(D^{\alpha}g)(A\cdot)\|_{L_p(A^{-1}(\Omega))}. \tag{3.5}
$$

We can now prove (3.4) for $k = 0$. Let $\widetilde{P} := Q^m(g(A\cdot))$; then by Lemma 2.1 and (3.5)

$$
\begin{aligned}
\|g-P\|_{L_p(\Omega)} &= |\det M|^{1/p} \|\tilde{g} - \widetilde{P}\|_{L_p(A^{-1}(\Omega))}\\
&\le C(n,m) |\det M|^{1/p} |\tilde{g}|_{W_p^m(A^{-1}(\Omega))}\\
&= C(n,m) |\det M|^{1/p} \sum_{|\alpha|=m} \|D_y^{\alpha}\tilde{g}\|_{L_p(A^{-1}(\Omega))}
\end{aligned}
$$

$$
\begin{aligned}
&\le C(n,m)|\det M|^{1/p} d^m \sum_{|\alpha|=m} \|(D^\alpha g)(A\cdot)\|_{L_p(A^{-1}(\Omega))} \\
&= C(n,m) d^m \sum_{|\alpha|=m} \|D_x^\alpha g\|_{L_p(\Omega)} \\
&= C(n,m) d^m |g|_{W_p^m(\Omega)}.
\end{aligned}
$$

For $1 \le k \le m-1$ take $\alpha \in \mathbb{Z}_+^n$, $|\alpha| = k$, $1 \le k \le m-1$, and let $h := D^\alpha g$. Then (3.3) yields

$$\|D^\alpha (g-P)\|_{L_p(\Omega)} = \left\| h(x) - Q^{m-k}\big(h(A\cdot)\big)(A^{-1}x) \right\|_{L_p(\Omega)}.$$

By the case $k=0$ proved above,

$$\left\| h(x) - Q^{m-k}\big(h(A\cdot)\big)(A^{-1}x) \right\|_{L_p(\Omega)} \le C(n,m) d^{m-k} |h|_{m-k,p},$$

which in turn implies that

$$\|D^\alpha (g-P)\|_{L_p(\Omega)} \le C(n,m) d^{m-k} |g|_{m,p}. \tag{3.6}$$

Summing up (3.6) over all $\alpha \in \mathbb{Z}_+^n$, $|\alpha| = k$, we obtain the required result. The case $k = m$ is trivial. ∎

4. Proofs of the main results.

Proof of Theorem 1.1. The proof of (1.11) for the case $p = \infty$ can be applied to star-shaped domains with respect to a point x_0, by using the classical Taylor polynomial (2.1) at the point $y = x_0$ and estimating the remainder (2.2). We leave the details to the reader and assume $1 \le p < \infty$. Let $E \subseteq \Omega$ be John's maximal ellipsoid (see Proposition 3.2) and A the corresponding affine mapping, i.e., $A(B(0,1)) = E$. John's theorem implies that

$$B(0,1) \subseteq A^{-1}(\Omega) \subseteq B(0,n).$$

First assume that $g \in C^m(\Omega)$. By Lemma 3.5 the polynomial $P(x) = Q^m(g(A\cdot))(A^{-1}x)$ is in Π_{m-1} and satisfies

$$|g-P|_{k,p} \le C(n,m) d^{m-k} |g|_{m,p}, \quad k = 0,1,\dots,m.$$

Since $C^\infty(\Omega)$ is dense in $W_p^m(\Omega)$ (see, e.g., Theorem 1.3.4 in [BrSc]), the proof of the general case follows from a standard density argument. ∎

Proof of Corollary 1.2. The method of proof is standard but we give it for the sake of completeness. Let $f \in L_p(\Omega)$ and $g \in W_p^m(\Omega)$ be such that

$$\|f-g\|_p + d^m |g|_{m,p} \le 2K_m(f,\Omega)_p.$$

By (1.9) with $k=0$, there exists $P \in \Pi_{m-1}$ such that

$$\|g-P\|_p \le C(n,m) d^m |g|_{m,p}.$$

Therefore

$$
\begin{aligned}
E_{m-1}(f)_p &\le \|f-P\|_p \\
&\le \|f-g\|_p + \|g-P\|_p \\
&\le \|f-g\|_p + C(n,m) d^m |g|_{m,p} \\
&\le C(n,m) K_m(f,\Omega)_p.
\end{aligned}
$$

In the other direction, it is easy to see from (1.1) that for any polynomial $Q \in \Pi_{m-1}$ and any $t > 0$,

$$K_m(f,t)_p \leq \|f - Q\|_p.$$

Consequently,

$$K_m(f,\Omega)_p \leq E_{m-1}(f)_p. \qquad \square$$

Proof of Corollary 1.3. Let $g \in W_p^m(\Omega)$, and let P be the polynomial for which (1.11) holds. Then by property (ii) of the sublinear functional l we have that $|l(g)| \leq |l(g-P)|$. Property (i) and (1.11) yield

$$\begin{aligned} |l(g)| &\leq |l(g-P)| \\ &\leq \widetilde{C} \sum_{k=0}^{m} d^k |g-P|_{k,p} \\ &\leq \widetilde{C} C(n,m) \sum_{k=0}^{m} d^k d^{m-k} |g|_{m,p} \\ &\leq C(n,m,\widetilde{C}) d^m |g|_{m,p}\ . \qquad \square \end{aligned}$$

Finally, we would like to point out a certain natural extension of our results to slightly more general types of domains.

Definition 4.1. *A compact domain $\Omega \subset \mathbb{R}^n$ with nonempty interior is almost convex if there exists a nonsingular affine mapping A, such that*

(i) $B(0,1) \subseteq A^{-1}(\Omega) \subseteq B(0,n)$,

(ii) $A^{-1}(\Omega)$ *is star-shaped with respect to* $B(0,1)$.

Indeed, John's theorem shows that every convex domain is almost convex. Furthermore, by the method used in this work (specifically Lemma 3.5) it can be seen that our main results remain valid for this type of domain.

Acknowledgments. We are grateful and indebted to the referees for bringing to our attention the paper by Verfürth. We also thank Manor Mendel for fruitful discussions of this work.

REFERENCES

[Ba] K. Ball, *Ellipsoids of maximal volume in convex bodies*, Geom. Dedicata, 41 (1992), pp. 241–250.

[BeSh] C. Bennett and R. Sharpley, *Interpolation of Operators*, Pure and Applied Mathematics 129, Academic Press, Boston, 1988.

[BrHi] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolations*, SIAM J. Numer. Anal., 7 (1970), pp. 112–124.

[BrSc] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Texts in Appl. Math 15, Springer-Verlag, New York, 1994.

[De] R. A. DeVore, *Nonlinear approximation*, Acta Numer., 7 (1998), pp. 51–150.

[DLS] S. Dekel, D. Leviatan, and M. Sharir, *On bivariate smoothness spaces associated with nonlinear approximation*, Constr. Approx., to appear.

[DQ] L. T. Dechevski and E. G. Quak, *On the Bramble-Hilbert lemma*, Numer. Funct. Anal. Optim., 11 (1990), pp. 485–495.

[H] S. M. Hudson, *Polynomial approximation in Sobolev spaces*, Indiana Univ. Math. J., 39 (1990), pp. 199–228.

[J] J. Fritz, *Extremum problems with inequalities as subsidiary conditions*, in Studies and Essays Presented to R. Courant on his 60th Birthday, Interscience, New York, 1948, pp. 187–204.

[JS] H. Johnen and K. Scherer, *On the equivalence of the K-functional and the moduli of continuity and some applications*, in Constructive Theory of Functions of Several Variables, Lecture Notes in Math. 571, Springer-Verlag, Berlin, 1976, pp. 119–140.

[KP] B. Karaivanov and P. Petrushev, *Nonlinear piecewise polynomial approximation beyond Besov spaces*, Appl. Comput. Harmon. Anal., 15 (2003), pp. 177–223.

[V] R. Verfürth, *A note on polynomial approximation in Sobolev spaces*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 715–719.