TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 305, Number 1, January 1988

INTERPOLATION OF BESOV SPACES

RONALD A. DEVORE AND VASIL A. POPOV

ABSTRACT. We investigate Besov spaces and their connection with dyadic spline approximation in $L_p(\Omega)$, $0 < p \leq \infty$. Our main results are: the determination of the interpolation spaces between a pair of Besov spaces; an atomic decomposition for functions in a Besov space; the characterization of the class of functions which have certain prescribed degree of approximation by dyadic splines.

1. Introduction. The Besov space $B_q^\alpha(L_p)$ is a set of functions f from L_p which have smoothness α. The parameter q gives a finer gradation of smoothness (see (2.4) for a precise definition). These spaces occur naturally in many fields of analysis. Many of their applications require a knowledge of their interpolation properties, i.e. a description of the spaces which arise when the real method of interpolation is applied to a pair of these spaces.

There are two definitions of Besov spaces which are currently in use. One uses Fourier transforms in its definition and the second uses the modulus of smoothness of the function f. These two definitions are equivalent only with certain restrictions on the parameters; for example they are different when $p < 1$ and α is small. The first and simplest interpolation theorems for Besov spaces, were for interpolation between a pair $B_q^\alpha(L_p)$ and $B_r^\beta(L_p)$ with $p \geq 1$ fixed. In this case, the real method of interpolation for the parameters (θ, s) applied to these spaces gives the Besov space $B_s^\gamma(L_p)$ with $\gamma = \theta\alpha + (1-\theta)\beta$. Hence, when p is held fixed the Besov spaces are invariant under interpolation.

More interesting and somewhat more difficult to describe are the interpolation spaces when p is not fixed. Such a program has been carried out in the book of Peetre [**P**] using the Fourier transform definition of the Besov spaces. The main tool in describing these interpolation spaces is to correspond to each f in the Besov space a sequence of trigonometric polynomials obtained from the Fourier series of f. In this way, the Besov space $B_q^\alpha(L_p)$ is identified with a weighted sequence space $l_q^\alpha(L_p)$. Interpolation properties of the Besov spaces are then derived from the interpolation between sequence spaces (when these are known). The success of this approach when $p < 1$ rests on the fact that the corresponding Besov spaces are defined using H_p norms so that each f in the Besov space is a distribution and therefore has a Fourier series.

Received by the editors January 1, 1987.

1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 46E39; Secondary 41A15, 41A63.

Key words and phrases. Besov spaces, real interpolation spaces, dyadic splines, degree of approximation.

The first author was supported by NSF Grant DMS 8320562.

The Besov spaces defined by the modulus of smoothness occur more naturally in many areas of analysis including approximation theory. Especially important is the case when $p < 1$ since these spaces are needed in the description of approximation classes for the classical methods of nonlinear approximation such as rational approximation and approximation by splines with free knots (see [**Pt** and **D-P**]).

The purpose of the present paper is to describe the interpolation of the Besov spaces defined by the modulus of smoothness. This is established by developing the connections between Besov spaces and approximation by dyadic splines. We shall show that a function is in $B_q^\alpha(L_p)$ if and only if it has a certain rate of approximation by dyadic splines (§4). In this way, we can identify $B_q^\alpha(L_p)$ with certain sequence spaces in a manner similar to that described above for the Fourier transform definition. While the basic ideas behind such an identification is rather simple, some of the details become technical when dealing with the case $p < 1$. One of the main difficulties encountered is that in contrast to the Fourier transform case, the mapping which we use to associate to each $f \in L_p$ a dyadic spline is nonlinear when $p < 1$.

In the process of proving our interpolation theorem, we shall develop several interesting results about dyadic spline approximation and about the representation of a function $f \in B_q^\alpha(L_p)$ as a series of dyadic splines (see the atomic decomposition in Corollary 5.3).

2. Besov spaces. Let Ω be the unit cube in $\mathbf{R}^d$. If $f \in L_p(\Omega)$, $0 < p \le \infty$, we let $\omega_r(f,t)_p$, $t > 0$, denote the modulus of smoothness of order r of $f \in L_p(\Omega)$:

$$\omega_r(f,t)_p := \sup_{|h|\le t} \|\Delta_h^r(f,\cdot)\|_p(\Omega(rh)) \tag{2.1}$$

where $|h|$ is the Euclidean length of the vector h; Δ_h^r is the rth order difference with step $h \in \mathbf{R}^d$; and the norm in (2.1) is the L_p "norm" on the set $\Omega(rh) := \{x\colon x, x + rh \in \Omega\}$. Of course, when $p < 1$, this is not really a norm, it is only a quasi-norm, i.e. in place of the triangle inequality, we have

$$\|f+g\|_p \le 2^{1/p}[\|f\|_p + \|g\|_p]. \tag{2.2}$$

Also useful is the fact that for any $\mu \le \min(1,p)$ and any sequence (f_i)

$$\left\|\sum f_i\right\|_p \le \left(\sum \|f_i\|_p^\mu\right)^{1/\mu}. \tag{2.3}$$

If $\alpha, p, q > 0$, we say f is in the Besov space $B_q^\alpha(L_p)$ whenever

$$|f|_{B_q^\alpha(L_p)} := \left(\int_0^\infty (t^{-\alpha}\omega_r(f,t)_p)^q \frac{dt}{t}\right)^{1/q} \tag{2.4}$$

is finite. Here, r is any integer larger than α. When $q = \infty$, the usual change from integral to sup is made in (2.4).

We also define the following "norm" for $B_q^\alpha(L_p)$:

$$\|f\|_{B_q^\alpha(L_p)} := \|f\|_p + |f|_{B_q^\alpha(L_p)}. \tag{2.5}$$

Different values of $r > \alpha$ result in norms (2.5) which are equivalent. This is proved by establishing inequalities between the moduli of smoothness ω_r and $\omega_{r'}$

when $r' \le r$. A simple inequality is $\omega_r \le c\omega_{r'}$ which follows readily from (2.2). In the other direction, we have the Marchaud type inequality:

$$\omega_{r'}(f,t)_p \le ct^{r'}\left[\|f\|_p + \left(\int_t^\infty (s^{-r'}\omega_r(f,s))^\mu \frac{ds}{s}\right)^{1/\mu}\right] \tag{2.6}$$

which holds for every $\mu \le \min(1,p)$. This inequality can be proved by using standard identities for differences. In §4, we give a different proof of (2.6) using dyadic spline approximation. Using these two inequalities for moduli together with the Hardy inequality [**B-B**, p. 199], one shows that any two norms (2.5) are equivalent provided that $r > \alpha$.

There are many other norms which are equivalent to (2.5). We shall have occasion to use several of these which we describe in later sections. A simple observation is

$$\|f\|_{B_q^\alpha(L_p)} \simeq \|f\|_p + \left(\sum_{k=1}^\infty [2^{k\alpha}\omega_r(f,2^{-k})_p]^q\right)^{1/q}. \tag{2.7}$$

In fact, since ω_r is bounded, the integral in (2.4) is equivalent to the integral of the same integrand taken over $[0,1]$. Now, the monotonicity properties of ω_r allow us to discretize the integral and obtain that (2.7) is equivalent to (2.4).

3. Local polynomial approximation. We want to show that $\omega_r(f,2^{-k})_p$ in (2.7) can be replaced by the error of dyadic spline approximation. This requires inequalities between the modulus of smoothness and the degree of spline approximation. These will be given in §4. To estimate the degree of spline approximation by the modulus of smoothness, we first need estimates for local polynomial approximation. We define the local error of approximation by polynomials by

$$E_r(f,I)_p := \inf_{\deg(Q)<r} \|f-Q\|_p(I). \tag{3.1}$$

with $\deg(Q)$ the *coordinate degree* of Q. We use the notation $\| \ \|_p(I)$ to denote the L_p norm over I; when I is omitted the norm is understood to be taken over Ω.

We shall need an estimate for the local error of polynomial approximation in terms of the smoothness of f. One such estimate is Whitney's theorem:

$$E_r(f,I)_p \le c\omega_r(f,l_I)_p \tag{3.2}$$

with l_I the side length of I. Here and in what follows, c is a constant which depends only on r, d (and p, if p appears) unless otherwise stated. The value of c may vary at each appearance.

Whitney's theorem is best known for univariate functions and $p \ge 1$. It has also been proved by Yu. Brudnyi [**Br**] for multivariate functions and $p \ge 1$. A proof of (3.2) for all p and all dimensions d can be found in the paper of Storozhenko and Oswald [**S-O**]. We would also like to mention that the ideas used in the univariate proof for $p \ge 1$ carry over to the general case. For example, in the forthcoming book of Popov and Petrushev [**P-P**], the reader will find a proof of this type for $p < 1$ for univariate functions.

The modulus of smoothness is not suitable when we want to add up estimates over several intervals. We therefore introduce the following modified modulus:

$$w_r(f,t)_p := w_r(f,t,I)_p := \left(t^{-d}\int_{\Omega_t}\int_{I(rs)} |\Delta_s^r(f,x)|^p\,dx\,ds\right)^{1/p}, \tag{3.3}$$

where $\Omega_t := [-t,t]^d$. Using identities for differences, it can be shown that w_r and ω_r are equivalent, i.e., $c_1 w_r(f,t)_p \le \omega_r(f,t)_p \le c_2 w_r(f,t)_p$, with constants $c_1, c_2 > 0$ which depend only on r, p and d (see **[P-P]** for a proof of this in the univariate case; the same proof applies to the multivariate case as well). From this, we have the following result.

LEMMA 3.1. *If $f \in L_p(I)$, with $0 < p \le \infty$ and if I is a cube with side length l_I, then*

$$E_r(f,I)_p \le c w_r(f,l_I,I)_p. \tag{3.4}$$

This result in a slightly different form can also be found in **[S-O]**.

There always exist polynomials Q of best $L_p(I)$ approximation of coordinate degree $< r$: $\|f-Q\|_p(I) = E_r(f,I)_p$. In the present paper we shall also find it very useful to use the concept of "near best" approximation. We say Q is a near best $L_p(I)$ approximation to f from polynomials of coordinate degree $< r$ with constant A if

$$\|f-Q\|_p(I) \le A E_r(f,I)_p. \tag{3.5}$$

It follows that if P is any polynomial of coordinate degree $< r$, then

$$\|f-Q\|_p(I) \le A\|f-P\|_p(I). \tag{3.6}$$

One method for constructing near best approximants of f is as follows. We let $\rho \le p$ and we let Q_ρ be any polynomial of *near best* $L_\rho(I)$ approximation to f of coordinate degree $< r$, i.e. $\|f-Q_\rho\|_\rho(I) \le AE_r(f,I)_\rho$.

LEMMA 3.2. *If $\rho \le p$, and Q_ρ is as above, we have*

$$\|f-Q_\rho\|_p(I) \le cAE_r(f,I)_p \tag{3.7}$$

with the constant c depending only on r, d and ρ.

PROOF. Let Q be a *best* $L_p(I)$ approximation to f of coordinate degree $< r$. Then, from elementary properties of polynomials (see **[D-Sh**, §3], we have with $\theta := 1/p - 1/\rho$,

$$\begin{aligned}\|f-Q_\rho\|_p &\le c(E_r(f,I)_p + \|Q-Q_\rho\|_p)\\ &\le c(E_r(f,I)_p + |I|^\theta\|Q-Q_\rho\|_\rho)\\ &\le c(E_r(f,I)_p + |I|^\theta[\|f-Q\|_\rho(I) + \|f-Q_\rho\|_\rho(I)])\\ &\le c(E_r(f,I)_p + |I|^\theta(A+1)\|f-Q\|_\rho(I))\\ &\le c(E_r(f,I)_p + (A+1)\|f-Q\|_p(I)) \le cAE_r(f,I)_p.\end{aligned}$$

Here, the first inequality uses the quasi-norm property (2.2); the second inequality is a comparison of polynomial norms; the third again uses (2.2); the fourth uses (3.6); and the fifth inequality is Hölder's inequality.

We introduce the following notation. If I is any cube, we let P_I denote any *near best* $L_\rho(I)$ approximation to f from polynomials of coordinate degree $< r$ with constant A. The following lemma shows that P_I is also a near best approximation on larger cubes.

LEMMA 3.3. *For any $p \geq \rho$ and any cube $J \supset I$ with $|J| \leq a|I|$, we have*

$$\|f - P_I\|_p(J) \leq cE_r(f, J)_p, \tag{3.8}$$

with c depending at most on r, d, a and A.

PROOF. If P is the best L_p approximation to f on J from polynomials of coordinate degree $< r$, then from (2.2) and Lemma 3.2,

$$\begin{aligned}\|P_I - P\|_p(I) &\leq c[\|f - P_I\|_p(I) + \|f - P\|_p(I)] \\ &\leq c[E_r(f, I)_p + E_r(f, J)_p] \\ &\leq cE_r(f, J)_p.\end{aligned}$$

This estimate can be enlarged to J (see [**D-Sh**, §3]):

$$\|P_I - P\|_p(J) \leq cE_r(f, J)_p.$$

Hence,

$$\|f - P_I\|_p(J) \leq c[\|f - P\|_p(J) + \|P - P_I\|_p(J)] \leq cE_r(f, J). \quad \square$$

4. Dyadic spline approximation. We want in this section to describe the connection between Besov spaces and dyadic spline approximation. Our main goal is to show that ω_r in (2.6) can be replaced by an error in dyadic spline approximation with a resulting equivalent seminorm. This means that the Besov spaces $B_q^\alpha(L_p)$ are the approximation spaces for the approximation by dyadic splines in L_p. Such characterizations are known when $p \geq 1$ (see [**C**]; also [**D-S**]) and when $p < 1$ and $d = 1$ (see [**O**]).

We let $\mathbf{D}_k$ denote the collection of dyadic cubes of $\mathbf{R}^d$ of side length 2^{-k} and we let $\mathbf{D}_k(\Omega)$ denote the set of those cubes $I \in \mathbf{D}_k$ with $I \subseteq \Omega$. We introduce two spline spaces for this partition. The first of these is $\Pi_k := \Pi_k(r)$, the space of all piecewise polynomials of coordinate degree $< r$ on the partition $\mathbf{D}_k$. That is, $S \in \Pi_k$ means that in the interior of each cube $I \in \mathbf{D}_k$, S is a polynomial of coordinate degree $< r$. We denote by $\Pi_k(\Omega)$ the restrictions of splines S in Π_k to Ω.

A best (or near best) approximation S_k to f in $L_p(\Omega)$ from $\Pi_k(\Omega)$ is gotten by simply taking $S := P_I$, $x \in I$, where P_I, $I \in \mathbf{D}_k(\Omega)$, is a best (or near best) approximation to f in $L_p(I)$ by polynomials of coordinate degree $< r$ on each cube I from $\mathbf{D}_k(\Omega)$. However, we shall also need to construct good approximations from $\Pi_k(\Omega)$ which have smoothness. For this, we shall use the tensor product B-splines and the quasi-interpolants of de Boor-Fix.

Let N be the univariate B-spline of degree $r - 1$ which has knots at the points $0, 1, \ldots, r$, i.e., $N(x) := r[0, 1, \ldots, r](x - \cdot)_+^{r-1}$ with the usual divided difference notation. For higher dimensions, we define N by

$$N(x) := N(x_1) \cdots N(x_d). \tag{4.1}$$

These are the tensor product B-splines. They are piecewise polynomials of coordinate degree $< r$ which have continuous derivatives $D^\nu N$, $0 \leq \nu \leq \mathbf{r-2}$, and derivatives $D^\nu N$ in L_∞ for $0 \le \nu < \mathbf{r} - \mathbf{1}$. We use the notation $\mathbf{k} := (k, k, \ldots, k)$. The splines N are nonnegative and are supported on the cube $[0, r]^d$.

To get splines in the space Π_k, we let

$$N_k(x) := N(2^k x), \qquad k = 0, 1, \ldots, \tag{4.2}$$

and

$$N_{j,k}(x) := N_k(x - x_j), \qquad j \in Z^d, \tag{4.3}$$

where the $x_j := 2^{-k} j$ are the vertices of the cubes in $\mathbf{D}_k$. The B-splines $N_{j,k}$ are a partition of unity, i.e.

$$\sum_{j \in Z^d} N_{j,k} \equiv 1, \quad \text{on } \mathbf{R}^d. \tag{4.4}$$

Each spline S in the span of the $N_{j,k}$ can be written in a B-spline series:

$$S = \sum_{j \in Z^d} \alpha_j(S) N_{j,k} \tag{4.5}$$

with the $\alpha_j := \alpha_{j,k}$ the dual functionals of the $N_{j,k}$. The functionals α_j can be expressed in terms of the univariate functionals:

$$\alpha_j(S) = \alpha_{j_1}((\cdots \alpha_{j_d}(S))) \tag{4.6}$$

where the univariate functional α_{j_ν} is applied to a multivariate function g by considering g as a function of x_ν with the other variables held fixed.

There are many representations for the functionals α_j. We mention in particular, the de Boor-Fix formula [**B-F**]. This representation gives that for *any point* ξ_j in the supp(N_j), we can write

$$\alpha_j(S) = \sum_{0 \le \nu \le \mathbf{r} - \mathbf{1}} a_\nu D^\nu(S)(\xi_j), \qquad j \in \Lambda, \tag{4.7}$$

for certain coefficients a_ν depending on ξ_j and r.

For approximation on Ω, we need only the B-splines $N_{j,k}$ which do not vanish identically on Ω. We let $\Lambda := \Lambda(k)$ denote the set of j for which this is the case and we let $\Sigma_k := \Sigma_k(\Omega)$ denote the linear span of the B-splines $N_{j,k}$, $j \in \Lambda$. Then any $S \in \Sigma_k$ can be written

$$S = \sum_{j \in \Lambda} \alpha_j(S) N_{j,k}. \tag{4.8}$$

For the representation of α_j, $j \in \Lambda$, we shall choose the points ξ_j as the center of a cube $J_j := J_{j,k} \in \mathbf{D}_k$ such that

$$\xi_j \in J_j \subset \operatorname{supp}(N_j) \cap \Omega, \qquad j \in \Lambda. \tag{4.9}$$

With this choice, we can define $\alpha_j(f)$ for any f which is suitably differentiable at ξ_j. In particular, in this way, we have that α_j is defined for any S in Π_k.

From (4.7), it is easy to estimate the coefficients of a spline $S \in \Pi_k$.

LEMMA 4.1. *We have for any $0 < p \leq \infty$ and any $S \in \Pi_k$,*

$$|\alpha_j(S)| \leq c2^{kd/p}\|S\|_p(J_j). \tag{4.10}$$

PROOF. This is well known for one variable and $p \geq 1$. A similar proof applies in the general case. For example, from Markov's inequality applied to S on J_j and estimates for the coefficients a_ν (see [**D**]), it follows that

$$|\alpha_j(S)| \leq c\|S\|_\infty(J_j). \tag{4.11}$$

Since $|J_j| = 2^{-kd}$, (4.10) follows from (4.11) and the well-known inequality between L_p and L_∞ norms for polynomials (see e.g. [**D-Sh**, §3]). □

Closely related to (4.10) is the following.

LEMMA 4.2. *If $S = \sum_{j\in\Lambda} \alpha_j N_{j,k}$ is in Σ_k, then for any $0 < p \leq \infty$, we have*

$$c_1\|S\|_p \leq \left(\sum_{j\in\Lambda} |\alpha_j(S)|^p 2^{-kd}\right)^{1/p} \leq c_2\|S\|_p \tag{4.12}$$

with c_1, c_2 depending at most on d and r.

PROOF. Again this is well known (see [**B**]) when $p \geq 1$ and the general case is proved in the same manner. For example, since $\Sigma_k \subset \Pi_k$, the right side of (4.12) follows from (4.10) and the fact that a point x falls in at most r^d of the cubes J_j. For the left inequality, we use the fact that at most r^d terms in the representation of S are nonzero at a given point x. Hence

$$|S(x)|^p \leq c\sum_{j\in\Lambda} |\alpha_j|^p N_{j,k}(x)^p.$$

Integrating with respect to x and using the fact that the integral of $N_{j,k}^p$ is less than $c2^{-kd}$ (because $N_{j,k} \leq 1$) gives the desired result. □

Now, let f be any function which is $r-1$ times continuously differentiable at each of the points ξ_j. Then $\alpha_j(f)$ is defined for all j and we define

$$Q_k(f) := \sum_{j\in\Lambda} \alpha_j(f) N_{j,k}. \tag{4.13}$$

The Q_k are called quasi-interpolant operators. In particular Q_k is defined for all $S \in \Pi_k$, and it follows that Q_k is a projector from Π_k onto Σ_k: $Q_k(S) = S$ whenever $S \in \Sigma_k$.

We want to examine the approximation properties of the Q_k. For this, we introduce the following notation. If $I \in \mathbf{D}_k$, we let $\tilde{I}$ be the smallest cube which contains each of the J_j for which $\operatorname{supp} N_{j,k} \cap I \neq \varnothing$. Then, $\tilde{I} \subset \Omega$ and $|\tilde{I}| \leq c|I|$ with c depending only on d and r.

LEMMA 4.3. *If $S \in \Pi_k$ and $0 < p \leq \infty$, then for each $I \in \mathbf{D}_k(\Omega)$, we have*

$$\|Q_k(S)\|_p(I) \leq c\|S\|_p(\tilde{I}), \tag{4.14}$$

and

$$\|S - Q_k(S)\|_p(I) \leq cE_r(S, \tilde{I})_p. \tag{4.15}$$

PROOF. We let Λ_I be the set of j such that $N_{j,k}$ does not vanish identically on I. We use the representation (4.13) and the estimate (4.10) for the functionals α_j, to find

$$\|Q_k(S)\|_p(I) \le \max_{j\in\Lambda_I} |\alpha_j(S)| \left\| \sum_{j\in\Lambda_I} N_{j,k} \right\|_p (I) \tag{4.16}$$
$$\le c|I|^{1/p} \max_{j\in\Lambda_I} 2^{kd/p} \|S\|_p(J_j) \le c\|S\|_p(\tilde{I}),$$

because of (4.4). This is (4.14).

To prove (4.15), we let $I \in \mathbf{D}_k$ and let P be a polynomial of best $L_p(\tilde{I})$ approximation to S of coordinate degree $< r$. Since $Q_k(P) = P$, we have by (2.2) and (4.14),

$$\|S - Q_k(S)\|_p(I) \le c[\|S - P\|_p(I) + \|Q_k(S-P)\|_p(I)] \tag{4.17}$$
$$\le c\|S - P\|_p(\tilde{I}) = cE_r(S, \tilde{I})_p. \qquad \square$$

COROLLARY 4.4. *If $0 < p \le \infty$, then $\|Q_k(S)\|_p \le c\|S\|_p$ for all $S \in \Sigma_k$.*

PROOF. This follows immediately from (4.14) when $p = \infty$. When $0 < p < \infty$, we raise both sides of (4.14) to the power p and then we sum over $I \in \mathbf{D}_k(\Omega)$. Since each point $x \in \Omega$ appears in at most c of the cubes $\tilde{I}$, with c depending only on r and d, the corollary follows. $\square$

We want to describe a class of methods for approximating each f in $L_p(\Omega)$ by smooth dyadic splines from Σ_k. For each $I \in \mathbf{D}_k$ and $f \in L_p(\Omega)$, we let $P_I := P_I(f)$ be a *near best* $L_p(I)$ approximation to f from polynomials of coordinate degree $< r$ with an absolute constant A. We then define $S_k := S_k(f) \in \Pi_k$, $k = 0, 1, \ldots$, to be the piecewise polynomial

$$S_k := P_I(x), \qquad x \in \text{interior}(I), \text{ for all } I \in \mathbf{D}_k. \tag{4.18}$$

From (3.8), we have

$$\|f - P_I\|_p(\tilde{I}) \le cE_r(f, \tilde{I})_p, \qquad I \in \mathbf{D}_k, \tag{4.19}$$

with c depending only on r, d and A.

Going further, for each $f \in L_p(\Omega)$, we define

$$T_k := T_k(f) := Q_k(S_k(f)), \qquad k = 0, 1, \ldots. \tag{4.20}$$

Then T_k is in Σ_k and we have

$$\|T_k(f)\|_p \le c\|f\|_p \tag{4.21}$$

with c depending only on r, d and A. Indeed, since P_I is a near best approximation to f, we have $\|P_I\|_p(I) \le c\|f\|_p(I)$, $I \in \mathbf{D}_k(\Omega)$. Hence, $\|S_k(f)\|_p \le c\|f\|_p$ and therefore (4.21) follows from Corollary 4.4.

THEOREM 4.5. *For any of the operators T_k in* (4.20) *and for each $f \in L_p(\Omega)$, we have*

$$\|f - T_k(f)\|_p \le c\omega_r(f, 2^{-k})_p, \qquad k = 0, 1, \ldots, \tag{4.22}$$

with c depending only on r, d, p and A.

PROOF. From (4.15), we have for each $I \in \mathbf{D}_k(\Omega)$,

$$\begin{aligned} \|f - T_k\|_p(I) &\le c[\|f - S_k\|_p(I) + \|S_k - Q_k(S_k)\|_p(I)] \\ &\le c[\|f - P_I\|_p(I) + E_r(S_k, \tilde{I})_p] \\ &\le c[E_r(f, \tilde{I})_p + E_r(S_k, \tilde{I})_p]. \end{aligned} \tag{4.23}$$

Now, for any cube $J \subseteq \tilde{I}$ with $J \in \mathbf{D}_k$, we have from (4.19)

$$\begin{aligned} \|S_k - P_I\|_p(J) = \|P_J - P_I\|_p(J) &\le c[\|f - P_J\|_p(J) + \|f - P_I\|_p(J)] \\ &\le c[E_r(f, J)_p + E_r(f, \tilde{I})_p] \le cE_r(f, \tilde{I})_p. \end{aligned} \tag{4.24}$$

Since the number of cubes $J \in \mathbf{D}_k$ contained in $\tilde{I}$ depends only on d and r, (4.24) gives $E_r(S_k, \tilde{I})_p \le cE_r(f, \tilde{I})_p$. If we use this in (4.23), we obtain

$$\|f - T_k\|_p(I) \le cE_r(f, \tilde{I})_p. \tag{4.25}$$

Now, each point $x \in \Omega$ appears in only a constant depending only on r and d number of cubes $\tilde{I}$. Hence, if we raise both sides of (4.25) to the power p and sum over all I in $\mathbf{D}_k(\Omega)$ and use (3.4), we obtain

$$\begin{aligned} \|f - T_k\|_p^p(\Omega) &\le c \sum_{I \in \mathbf{D}_k(\Omega)} w_r(f, l_{\tilde{I}}, \tilde{I})_p^p \\ &\le ct^{-d} \int_{\Omega_t} \int_{\Omega(rs)} |\Delta_s^r(f, x)|^p \, dx \, ds \end{aligned} \tag{4.26}$$

with $t := \max l_{\tilde{I}} \le c2^{-k}$. Here, we have used the fact that $w_r(f, t')_p \le cw_r(f, t)_p$ provided $t' \le t \le ct'$. Finally, (4.22) follows from (4.26) because each of the interior integrals on the right side of (4.26) does not exceed $\omega_r(f, t)_p^p$ which from the usual properties of moduli is $\le c\omega_r(f, 2^{-k})_p^p$. □

Theorem 4.5 shows that the error of dyadic approximation can be majorized by the modulus of smoothness. Namely, if we let

$$s_k(f)_p := \inf_{S \in \Sigma_k} \|f - S\|_p, \tag{4.27}$$

then we have

COROLLARY 4.6. *For each $f \in L_p(\Omega)$, and for each $r = 1, 2, \ldots,$ we have*

$$s_k(f)_p \le c\omega_r(f, 2^{-k})_p, \qquad k = 0, 1, \ldots . \tag{4.28}$$

It is also important to note that $T_k(f)$ is a near best approximation from Σ_k.

COROLLARY 4.7. *If $f \in L_p(\Omega)$, then*

$$\|f - T_k(f)\|_p \le cs_k(f)_p$$

with c depending only on r, d, p and A.

PROOF. Let S be a best $L_p(\Omega)$ approximation to f from Σ_k. Then since $Q_k(S) = S$, we have $f - T_k(f) = f - S + Q_k(S - S_k(f))$. If we use the fact that Q_k is bounded (Corollary 4.4), we obtain

$$\begin{aligned} [\|f - T_k(f)\|_p] &\le c[\|f - S\|_p + \|S - S_k(f)\|_p] \\ &\le c[\|f - S\|_p + \|f - S_k(f)\|_p] \le cs_k(f)_p. \end{aligned}$$

Here, the last inequality uses the fact that $S_k(f)$ is a near best approximation from Π_k with constant A and the error of approximating f from Π_k is smaller than the error in approximating f from Σ_k (because $\Sigma_k \subset \Pi_k$). □

We also need inverse estimates to (4.28). We let $s_{-1}(f)_p := \|f\|_p$.

THEOREM 4.8. *For each $k > 0$, and each $r = 1, 2, \dots$, we have for $\lambda := \min(r, r-1+1/p)$ and for each $f \in L_p$,*

$$\omega_r(f, 2^{-k})_p \leq c 2^{-k\lambda} \left(\sum_{j=-1}^{k} [2^{j\lambda} s_j(f)_p]^\mu \right)^{1/\mu}, \tag{4.29}$$

provided $\mu \leq \min(1, p)$.

PROOF. We let U_k be a best approximation to f from Σ_k and let $u_k := U_k - U_{k-1}$, $k = 0, 1, \dots$, with $U_{-1} := 0$. If $|h| \leq r^{-1} 2^{-k}$ and $x \in \Omega(rh)$, we write

$$\Delta_h^r(f, x) = \Delta_h^r(f - U_k, x) + \sum_{j=0}^{k} \Delta_h^r(u_j, x). \tag{4.30}$$

Then, from (2.3),

$$\|\Delta_h^r(f)\|_p(\Omega(rh)) \leq c \left(s_k(f)^\mu + \sum_{j=0}^{k} \|\Delta_h^r(u_j)\|_p(\Omega(rh))^\mu \right)^{1/\mu}. \tag{4.31}$$

To estimate $\|\Delta_h^r(u_j)\|_p(\Omega(rh))$, we write u_j in its B-spline series:

$$u_j = \sum_{\nu \in \Lambda(j)} \alpha_{\nu,j}(u_j) N_{\nu,j}. \tag{4.32}$$

For any x, at most c B-splines (4.32) are nonzero at x with c depending only on r and d. Hence,

$$|\Delta_h^r(u_j, x)|^p \leq c \sum_{\nu \in \Lambda(j)} |\alpha_{\nu,j}(u_j)|^p |\Delta_h^r(N_{\nu,j}, x)|^p. \tag{4.33}$$

Now, we shall give two estimates for $\Delta_h^r(N_{\nu,j}, x)$. The first of these is for the set Γ which consists of all x such that x and $x + rh$ are in the same cube $I \in \mathbf{D}_j$ and $N_{\nu,j}$ does not vanish identically on I. Since $N_{\nu,j}$ is a polynomial on I whose rth order derivatives do not exceed $c2^{jr}$, we have

$$|\Delta_h^r(N_{\nu,j}, x)| \leq c(2^j |h|)^r, \qquad x \in \Gamma. \tag{4.34}$$

Our second estimate is for the set Γ' which consists of all x such that x and $x + rh$ are in different cubes from $\mathbf{D}_j$ and $N_{\nu,j}$ does not vanish identically on both of these cubes. Since $N_{\nu,j} \in W_\infty^{r-1}$ (Sobolev space) has $(r-1)$th derivatives whose $L_\infty(\Omega)$ norms do not exceed $c2^{j(r-1)}$, we have

$$|\Delta_h^r(N_{\nu,j}, x)| \leq c(2^j |h|)^{r-1}, \qquad x \in \Gamma'. \tag{4.35}$$

Now, the set Γ has measure $\leq c2^{-jd}$ because the support of $N_{\nu,j}$ has measure $\leq c2^{-jd}$. Also, Γ' has measure $\leq c|h| 2^{-j(d-1)}$. Indeed, for x to be in Γ', we must have $\operatorname{dist}(x, \operatorname{bound}(I)) \leq r|h|$ for the cube I which contains x. The measure of all such $x \in I$ is less than $c|h|2^{-j(d-1)}$. Since $N_{\nu,j}$ vanishes on all but c cubes with c depending only on r and d, we have $|\Gamma'| \le c|h|2^{-j(d-1)}$ as claimed.

Using these two estimates for the measure of Γ and Γ' together with (4.34) and (4.35), we obtain

$$\text{(4.36)} \quad \begin{aligned} \int_{\Omega(rh)} |\Delta_h^r(N_{\nu,j})|^p &\le c[|h|^{rp}2^{jrp}2^{-jd} + |h|^{(r-1)p}2^{j(r-1)p}|h|2^{-j(d-1)}] \\ &\le c|h|^{\lambda p}2^{j\lambda p}2^{-jd} \end{aligned}$$

because $|h|2^{-k} \le r^{-1} \le 1$.

Now, we integrate (4.33) and use (4.36) to find

$$\text{(4.37)} \quad \begin{aligned} \|\Delta_h^r(u_j)\|_p &\le c|h|^\lambda 2^{j\lambda}\left(\sum |\alpha_{\nu,j}(u_j)|^p 2^{-jd}\right)^{1/p} \\ &\le c|h|^\lambda 2^{j\lambda}\|u_j\|_p \le c|h|^\lambda 2^{j\lambda}[s_j(f)_p + s_{j-1}(f)_p], \end{aligned}$$

where the next to last inequality is (4.12) and the last inequality is the triangle inequality applied to $u_j = f - U_{j-1} - (f - U_j)$.

If we use (4.37) in (4.31), we obtain

$$\text{(4.38)} \quad \|\Delta_h^r(f)\|_p(\Omega(rh)) \le c\left(s_k(f)^\mu + |h|^{\lambda\mu}\sum_{j=-1}^{k}[2^{j\lambda}s_j(f)]^\mu\right)^{1/\mu}.$$

If we now take a sup over all $|h| \le r^{-1}2^{-k}$, (4.38) gives

$$\omega_r(f, 2^{-k})_p \le c\omega_r(f, r^{-1}2^{-k})_p \le c2^{-k\lambda}\left(2^{k\lambda\mu}s_k(f)^\mu + \sum_{j=-1}^{k}[2^{j\lambda}s_j(f)]^\mu\right)^{1/\mu}.$$

Since the term $2^{k\lambda\mu}s_k(f)^\mu$ can be incorporated into the sum, we obtain (4.29). □

It is also possible to estimate $\omega_{r'}$ for each $r' < r$:

$$\text{(4.39)} \quad \omega_{r'}(f, 2^{-k})_p \le c2^{-kr'}\left(\sum_{j=-1}^{k}(2^{jr'}s_j(f)_p)^\mu\right)^{1/\mu}.$$

Indeed, this is proved in exactly the same way as we have derived (4.29), except that, in place of (4.34) and (4.35), we use the inequality

$$\text{(4.40)} \quad |\Delta_h^{r'}(N_{\nu,j}, x)| \le c|h|^{r'}2^{jr'}$$

which follows from the fact that $N_{\nu,j}$ has all derivatives of order r' in L_∞.

With (4.39), we can easily prove the Marchaud type inequality (2.6).

COROLLARY 4.9. *There is a constant c depending only on p, r, and d such that for each $f \in L_p$, we have the inequality* (2.6).

PROOF. We have by (4.28): $s_j(f)_p \le c\omega_r(f, 2^{-j})_p$, $j = 0, 1, \ldots$. Also, $s_{-1}(f)_p := \|f\|_p$. Using this in (4.39) gives for $2^{-k-1} \le t \le 2^{-k}$,

$$\omega_{r'}(f, t)_p \le c\omega_{r'}(f, 2^{-k})_p \le c2^{-kr'}\left(\|f\|_p^\mu + \sum_{j=0}^{k}[2^{jr'}\omega_r(f, 2^{-j})_p]^\mu\right)^{1/\mu}$$

and (2.6) then follows from the monotonicity of ω_r. □

5. Other seminorms for Besov spaces. The estimates of the last section allow us to introduce several norms which are equivalent to $\|f\|_{B_q^\alpha(L_p)}$. If $\mathbf{a} := (a_k)$ is a sequence whose component functions are in the quasi-normed space X, we use the $l_q^\alpha(X)$ "norms"

$$\|\mathbf{a}\|_{l_q^\alpha(X)} := \left(\sum_{k=0}^{\infty} [2^{k\alpha}\|a_k\|_X]^q\right)^{1/q} \tag{5.1}$$

with the usual change to a supremum norm when $q = \infty$. When (a_k) is a sequence of real numbers, we replace $\|a_k\|_X$ by $|a_k|$ in (5.1) and denote the resulting norm by $\|(a_k)\|_{l_q^\alpha}$.

Useful for us will be the discrete Hardy inequalities

$$\|(b_k)\|_{l_q^\alpha} \le c\|(a_k)\|_{l_q^\alpha} \tag{5.2}$$

which hold if either

$$\begin{aligned} &\text{(i)} \quad |b_k| \le c2^{-k\lambda}\left(\sum_{j=0}^{k}[2^{j\lambda}|a_j|]^\mu\right)^{1/\mu} \quad \text{or} \\ &\text{(ii)} \quad |b_k| \le c\left(\sum_{j=k}^{\infty}|a_j|^\mu\right)^{1/\mu} \end{aligned} \tag{5.3}$$

with $\mu \le q$ and (in (i)) $\alpha < \lambda$. Here, the constant c in (5.2) depends only on r, d and $1/(\lambda - \alpha)$ in case of (i) and $1/\alpha$, in the case of (ii).

In the following theorem, we let $T_k := T_k(f)$ be defined as in (4.20) for a given $r = 1, 2, \ldots$ and given near best approximations P_I with constant A. We let $t_k := t_k(f) := T_k - T_{k-1}$ with $T_{-1} := 0$ and let $\lambda := \min(r - 1 + 1/p, r)$, as before.

THEOREM 5.1. *Let $0 < \alpha$ and $0 < q, p \le \infty$. If $\alpha < \lambda$, then the following norms are equivalent to $N(f) := \|f\|_{B_q^\alpha(L_p)}$ with constants of equivalency depending only on d, r and A and the constant of* (5.2):

$$\begin{aligned} &\text{(i)} \quad N_1(f) := \|(s_k(f))\|_{l_q^\alpha} + \|f\|_p, \\ &\text{(ii)} \quad N_2(f) := \|(f - T_k(f))\|_{l_q^\alpha(L_p)} + \|f\|_p, \\ &\text{(iii)} \quad N_3(f) := \|(t_k(f))\|_{l_q^\alpha(L_p)}. \end{aligned} \tag{5.4}$$

PROOF. From Theorem 4.5, $s_k(f)_p \le \|f - T_k(f)\|_p \le c\omega_r(f, 2^{-k})_p$. Hence, $N_1(f) \le N_2(f) \le cN(f)$. On the other hand, from Theorem 4.8 and the Hardy inequality (5.2) above, we have $N(f) \le cN_1(f)$. This shows that $N(f)$, $N_1(f)$ and $N_2(f)$ are all equivalent. Since $\|t_k\|_p \le c[\|f - T_k\|_p + \|f - T_{k-1}\|_p]$ we have $N_3(f) \le cN_2(f)$. In the other direction $f - T_k = \sum_{k+1}^\infty t_j$ and therefore from (2.3), we obtain for $k = -1, 0, 1, \ldots$,

$$\|f - T_k\|_p \le \left(\sum_{k+1}^{\infty}\|t_j\|_p^\mu\right)^{1/\mu}.$$

Note, when $k = -1$, this is an estimate for $\|f\|_p$. Now, from the Hardy inequality (5.2), we have $N_2(f) \le cN_3(f)$ and therefore $N_2(f)$ and $N_3(f)$ are equivalent. $\square$

The norm N_1 of Theorem 5.1 shows that a function f is in $B_q^\alpha(L_p)$ if and only if $(s_k(f))$ is in l_q^α. In the terminology of [**D-P**], we have that the approximation class A_q^α for dyadic spline approximation in L_p is the same as the Besov space $B_q^\alpha(L_p)$. Related to this is the following Bernstein type inequality for dyadic splines.

COROLLARY 5.2. *If $r = 1, 2, \ldots$ and $\alpha < \lambda$, then for each $S \in \Sigma_n$,*

$$\|S\|_{B_q^\alpha(L_p)} \le c2^{\alpha n}\|S\|_p \tag{5.5}$$

with c independent of S and n.

PROOF. Since $S \in \Sigma_n$, $s_k(S) = 0$, $k \ge n$, and for $k < n$, we have $s_k(S)_p \le \|S\|_p$. Hence, for $q < \infty$,

$$N_1(S)^q \le c \sum_{k=-1}^{n} [2^{k\alpha} s_k(S)_p]^q \le c2^{\alpha nq}\|S\|_p^q$$

and (5.5) follows from Theorem 5.1. Similarly for $q = \infty$. □

Another interesting application of Theorem 5.1 is the following atomic decomposition for functions f in $B_q^\alpha(L_p)$. According to Theorem 5.1, we can write $f = \sum t_k$ with the notation of that theorem. Since $t_k \in \Sigma_k$, we have

$$t_k = \sum_{\nu \in \Lambda(k)} \alpha_{\nu,k} N_{\nu,k} \tag{5.6}$$

with $N_{\nu,k}$ the B-splines for $\mathbf{D}_k$. Hence,

$$f = \sum_{k=0}^{\infty} \sum_{\nu \in \Lambda(k)} \alpha_{\nu,k} N_{\nu,k} \tag{5.7}$$

with convergence in the sense of L_p.

COROLLARY 5.3. *Let $0 < q, p \le \infty$ and $r = 1, 2, \ldots$. If $0 < \alpha < \lambda$, then a function $f \in L_p$ is in $B_q^\alpha(L_p)$ if and only if f can be represented as in* (5.7) *with*

$$N_4(f) := \left(\sum_{k=0}^{\infty} 2^{\alpha k q} \left(\sum_{j \in \Lambda(k)} |\alpha_{\nu,k}|^p 2^{-kd} \right)^{q/p} \right)^{1/q} < \infty \tag{5.8}$$

(*and the usual modification if either p or $q = \infty$*) *and $N_4(f)$ is equivalent to* $\|f\|_{B_q^\alpha(L_p)}$.

PROOF. From Lemma 4.2,

$$\|t_k\|_p \simeq \left(\sum_{\nu \in \Lambda_k} |\alpha_{\nu,k}|^p \right)^{1/p}.$$

Hence from Theorem 5.1, $N_4(f)$ is equivalent to $N_3(f)$ which is in turn equivalent to $N(f)$. □

A different atomic decomposition was given by M. Frazier and B. Jaewerth [**F-J**] for Besov spaces defined by the Fourier transform. In the case $d = 1$, there is also an atomic decomposition using spline functions by Ciesielski [**C**].

6. Interpolation theorems. We are now interested in proving interpolation theorems for Besov spaces. If $\alpha_0, \alpha_1 > 0$, and $0 < p_0, p_1, q_0, q_1 \le \infty$, we introduce the abbreviated notation $B_i := B_{q_i}^{\alpha_i}(L_{p_i})$ and $l_i := l_{q_i}^{\alpha_i}(L_{p_i})$, $i = 0, 1$.

We recall that if X_0, X_1 is a pair of quasi-normed spaces which are continuously embedded in a Hausdorff space X, then the K-functional

$$K(f, t, X_0, X_1) := \inf_{f = f_0 + f_1} \{\|f_0\|_{X_0} + t\|f_1\|_{X_1}\} \tag{6.1}$$

is defined for all $f \in X_0 + X_1$. This K-functional determines new function spaces. If $0 < \theta < 1$ and $0 < q \le \infty$, we define the space $X_{\theta,q} := (X_0, X_1)_{\theta,q}$ as the set of all f such that

$$\|f\|_{X_{\theta,q}} := \|f\|_{X_0 + X_1} + \left(\int_0^\infty [t^{-\theta} K(f,t)]^q \frac{dt}{t}\right)^{1/q} \tag{6.2}$$

is finite.

We wish to establish a connection between the K-functional for B_0, B_1 and the K-functional for l_0, l_1. For this, we fix a number $0 < \rho \le p_0, p_1$ and an integer r such that $\alpha_0, \alpha_1 \le r - 1$. We let $P_I(f)$ be the best $L_\rho(I)$ approximation to f from polynomials of coordinate degree $< r$. According to Lemma 3.2, if $S_k(f)$ is defined by (4.18), and $T_k(f)$ is defined by (4.20) then $N_2(f)$ and $N_3(f)$ of Theorem 5.1 are equivalent to the norm of $B_q^\alpha(L_p)$.

If $f \in L_\rho$, we let $Tf := (t_k(f))$. In this way, we associate to each $f \in L_\rho$ a sequence of dyadic splines and $f \in B_q^\alpha(L_p)$ if and only if $Tf \in l_q^\alpha(L_p)$ and from Theorem 5.1

$$\|f\|_{B_q^\alpha(L_p)} \simeq \|Tf\|_{l_q^\alpha(L_p)}, \tag{6.3}$$

for all $\alpha, q > 0$, provided $p \ge \rho$.

THEOREM 6.1. *There are constants $c_1, c_2 > 0$ which depend only on ρ, r, d, α_0, and α_1 such that*

$$c_1 K(f, t, B_0, B_1) \le K(Tf, t, l_0, l_1) \le c_2 K(f, t, B_0, B_1), \qquad t > 0, \tag{6.4}$$

whenever $f \in B_0 + B_1$.

Proof of lower inequality. We suppose that $\mathbf{a} := (a_k) \in l_1$ is such that $Tf - \mathbf{a}$ is in l_0. We define $g_k := T_k(a_k) := Q_k(S_k(a_k))$ as in (4.20). Then by (4.21), $\|g_k\|_{p_1} \le c\|a_k\|_{p_1}$. Now, we let $g := \sum_0^\infty g_k$ with convergence in L_{p_1}. Since $\sum_0^k g_j$ is in Σ_k, we have from (2.3)

$$s_k(g)_{p_1} \le \left\|\sum_{k+1}^\infty g_j\right\|_{p_1} \le c\left(\sum_{k+1}^\infty \|a_j\|_{p_1}^\mu\right)^{1/\mu}, \qquad k = -1, 0, \ldots,$$

provided $\mu \le p_1$. Here, when $k = -1$, $s_{-1}(f)_p := \|f\|_p$, as usual. If we take also $\mu \le q_1$, we have from the Hardy inequality (5.2) and the equivalence of the norms N and N_1 in Theorem 5.1 that

$$\|g\|_{B_1} \le c\|\mathbf{a}\|_{l_1}. \tag{6.5}$$

We can prove a similar estimate for $f - g$. Namely,

$$s_k(f-g)_{p_0} \le \left\| \sum_{k+1}^{\infty} (t_j - g_j) \right\|_{p_0} \le \left(\sum_{k+1}^{\infty} \|t_j - g_j\|_{p_0}^{\mu} \right)^{1/\mu}. \tag{6.6}$$

Now $t_j = Q_j(t_j)$ because Q_j is a projector. Also since $t_j \in \Pi_j$, we have $S_j(a_j - t_j) = S_j(a_j) - t_j$. Hence,

$$\begin{aligned}\|t_j - g_j\|_{p_0} &= \|Q_j(t_j - S_j(a_j))\|_{p_0} = \|Q_j(S_j(t_j - a_j))\|_{p_0} \\ &\le c\|t_j - a_j\|_{p_0}\end{aligned}$$

because of (4.21). If we use our last inequality in (6.6) and then argue as in the proof of (6.5), we obtain

$$\|f - g\|_{B_0} \le c\|Tf - \mathbf{a}\|_{l_0}. \tag{6.7}$$

Since $\mathbf{a} \in l_1$ is arbitrary, (6.5), (6.7) and the definition of the K-functional give the lower inequality in (6.4). □

For the proof of the upper inequality in (6.4), we shall need a result about approximation in a quasi-normed space X. We suppose that Z is a *linear* subspace of X such that each element $x \in X$ has a best approximation from Z. We let

$$E(x) := \inf_{z \in Z} \|x - z\|_X. \tag{6.8}$$

We say that z is a *near best* approximation to x with constant A if

$$\|x - z\| \le AE(x). \tag{6.9}$$

LEMMA 6.2. *Let X and Z be as above. If $x \in X$ and $z \in Z$ is a near best approximation to x with constant A, then for each $y \in X$, there is a $z' \in Z$ such that z' is a near best approximation to y and $z - z'$ is a near best approximation to $x - y$ with constants c depending only on X and A.*

PROOF. Let γ be such that $\|u + v\| \le \gamma(\|u\| + \|v\|)$ for all $u, v \in X$ (all norms in this proof are for X).

Case: $E(x - y) \le E(y)$. We let $z' := z'' + z$ with z'' a best approximation to $y - x$. Then, by definition $z - z'$ is near best for $x - y$ with constant 1. On the other hand,

$$\begin{aligned}\|y - z'\| &= \|y - z - z''\| \le \gamma(\|y - x - z''\| + \|x - z\|) \le \gamma(E(x - y) + AE(x)) \\ &\le \gamma(E(x - y) + \gamma AE(y) + \gamma AE(x - y)) \le (\gamma + 2\gamma^2 A)E(y).\end{aligned}$$

Case: $E(y) \le E(x - y)$. The same as the previous case with $x - y$ and y interchanged. □

Proof of the upper inequality in (6.4). We suppose that g is any function in B_1 for which $f - g$ is in B_0. We let P_I be the polynomials which make up $S_k := S_k(f)$. Then P_I is a best $L_\rho(I)$ approximation to f from polynomials of coordinate degree $< r$. Therefore, we can apply Lemma 6.2 to obtain a near best $L_\rho(I)$ approximation Q_I to g from polynomials of coordinate degree $< r$ such that $P_I - Q_I$ is also a near best $L_\rho(I)$ approximation to $f - g$.

We let U_k, R_k be obtained from Q_I and $P_I - Q_I$, $I \in \mathbf{D}_k$, by using quasi-interpolants in the same way that T_k was defined from the P_I. Since Q_k is linear, we have $R_k = T_k - U_k$. Then, by Corollary 4.7, U_k and R_k are respectively near best L_{p_1} and L_{p_0} approximations to g and $f-g$ from Σ_k, $k = 0, 1, \ldots$.

We let $t_k := T_k - T_{k-1}$, $u_k := U_k - U_{k-1}$, $r_k := R_k - R_{k-1}$, $k = 0, 1, \ldots$, with our usual convention $R_{-1} := 0$, $U_{-1} := 0$. We then have for $k = 0, 1, \ldots$,

$$\|u_k\|_{p_1} \le c[s_k(g)_{p_1} + s_{k-1}(g)_{p_1}],$$
$$\|r_k\|_{p_0} \le c[s_k(f-g)_{p_0} + s_{k-1}(f-g)_{p_0}].$$

With $\mathbf{u} := (u_k)$, it follows from Theorem 5.1 that

$$\|Tf - \mathbf{u}\|_{l_0} + t\|\mathbf{u}\|_{l_1} \le c[\|f-g\|_{B_0} + t\|g\|_{B_1}].$$

The upper estimate in (6.4) then follows from the definition of the K-functional. □

For B_0, B_1, l_0, l_1 and Tf as above, we have for any $q > 0$ and $0 < \theta < 1$,

$$f \in (B_0, B_1)_{\theta,q} \text{ if and only if } Tf \in (l_0, l_1)_{\theta,q}. \qquad \|f\|_{(B_0,B_1)_{\theta,q}} \simeq \|Tf\|_{(l_0,l_1)_{\theta,q}}. \tag{6.10}$$

Indeed, this follows immediately from the definition of the spaces $X_{\theta,q}$.

Now (6.10) allows us to deduce information about the interpolation spaces between B_0 and B_1 from known theorems (see [**P**, p. 98]) about the interpolation between l_0 and l_1. The simplest case to describe is when $p_0 = p_1 = p$. We then have

$$\left(l_{q_0}^{\alpha_0}(L_p), l_{q_1}^{\alpha_1}(L_p)\right)_{\theta,q} = l_q^{\alpha}(L_p) \quad \text{where } \alpha = \theta\alpha_0 + (1-\theta)\alpha_1. \tag{6.11}$$

From this, (6.10), and Theorem 5.1, we obtain

COROLLARY 6.2. *If* $0 < \alpha_0, \alpha_1$ *and* $0 < p, q_0, q_1 \le \infty$, *we have for each* $0 < \theta < 1$ *and* $0 < q \le \infty$,

$$\left(B_{q_0}^{\alpha_0}(L_p), B_{q_1}^{\alpha_1}(L_p)\right)_{\theta,q} = B_q^{\alpha}(L_p), \quad \textit{with } \alpha := \theta\alpha_0 + (1-\theta)\alpha_1. \tag{6.12}$$

When $p_0 \ne p_1$, the interpolation spaces between L_{p_0} and L_{p_1} can be described in terms of the Lorentz spaces $L_{p,q}$ (see [**B-B**, p. 183] for their definition and properties). We have for $0 < q_0, q_1 \le \infty$ (see [**P**, p. 98]),

$$(l_0, l_1)_{\theta,q} = l_q^{\alpha}(L_{p,q}) \tag{6.13}$$

with $\alpha := \theta\alpha_0 + (1-\theta)\alpha_1$; $1/q := \theta/q_0 + (1-\theta)/q_1$ and $1/p := \theta/p_0 + (1-\theta)/p_1$.

In the special case when $q = p$, we have $L_{p,q} = L_p$ and therefore, we obtain

COROLLARY 6.3. *If* $0 < \alpha_0, \alpha_1$ *and* $0 < p_0, p_1, q_0, q_1 \le \infty$, *then for each* $0 < \theta < 1$ *and for* $1/q := \theta/q_0 + (1-\theta)/q_1$; $1/p := \theta/p_0 + (1-\theta)/p_1$, *we have*

$$\left(B_{q_0}^{\alpha_0}(L_{p_0}), B_{q_1}^{\alpha_1}(L_{p_1})\right)_{\theta,q} = B_q^{\alpha}(L_p), \quad \textit{with } \alpha := \theta\alpha_0 + (1-\theta)\alpha_1, \tag{6.14}$$

provided $p = q$.

7. An embedding theorem for Besov spaces. As an application of the results of the previous sections, we shall prove Sobolev type embedding theorems for Besov spaces. These have important applications in nonlinear approximation (see [**D-P₁**]). We fix a value of p with $0 < p < \infty$. Given $\alpha > 0$, we determine σ from the equation

$$1/\sigma = \alpha/d + 1/p. \tag{7.1}$$

We shall prove that $B_p^\alpha(L_\sigma)$ is continuously embedded in L_p. For this, we shall use the following simple inequality for splines $S \in \Pi_k(r)$:

$$\|S\|_p \le c2^{k\alpha}\|S\|_\sigma. \tag{7.2}$$

Indeed, on each cube $I \in \mathbf{D}_k$, $S = P$ with P a polynomial of coordinate degree $< r$. Hence (see [**D-Sh**, §3]), $\|S\|_p(I) \le c|I|^{1/p-1/\sigma}\|S\|_\sigma(I) = 2^{k\alpha}\|S\|_\sigma(I)$. Therefore,

$$\|S\|_p^p \le c2^{k\alpha p} \sum_{I\in\mathbf{D}_k(\Omega)} \|S\|_\sigma(I)^p \le c2^{k\alpha p}\left(\sum_{I\in\mathbf{D}_k(\Omega)} \|S\|_\sigma(I)^\sigma\right)^{p/\sigma}$$

where the last inequality uses the fact that the $l_{\sigma/p}$ norm is larger than the l_1 norm because $\sigma/p < 1$.

THEOREM 7.1. *If α, σ, p are related as in* (7.1), *then $B_p^\alpha(L_\sigma)$ is continuously embedded in L_p, that is,*

$$\|f\|_p \le c\|f\|_{B_p^\alpha(L_\sigma)} \tag{7.3}$$

holds for all $f \in B_p^\alpha(L_\sigma)$.

PROOF. We choose $r > \alpha + 1$ and let $t_j \in \Sigma_j(r)$ be as in Theorem 5.1. Then $f = \sum_{j=0}^\infty t_j$ in the sense of convergence in L_σ. From (2.3), it follows that for $\mu := \min(1,p)$,

$$\|f\|_p \le \left(\sum_{j=0}^\infty \|t_j\|_p^\mu\right)^{1/\mu} \le c\left(\sum_{j=0}^\infty (2^{\alpha j}\|t_j\|_\sigma)^\mu\right)^{1/\mu} \le c\|f\|_{B_\mu^\alpha(L_\sigma)}, \tag{7.4}$$

where the second inequality follows from (7.2) and the last from Theorem 5.1.

Inequality (7.4) shows that $B_\mu^\alpha(L_\sigma)$ is continuously embedded in L_p which is the desired result when $p \le 1$. When $p > 1$, we choose $1 \le p_0 < p < p_1 < \infty$ and for $i = 0,1$, we let α_i be determined by formula (7.1) for p_i and our σ. Then by (7.4)

$$\|f\|_{p_i} \le c\|f\|_{B_1^{\alpha_i}(L_\sigma)}, \qquad i = 0,1. \tag{7.5}$$

If we now apply Corollary 6.2 with θ chosen so that $1/p = \theta/p_0 + (1-\theta)/p_1$ and $q := p$, we obtain by interpolation

$$\|f\|_p \le c\|f\|_{B_p^{\alpha'}(L_\sigma)}$$

with $\alpha' = \theta\alpha_0 + (1-\theta)\alpha_1$. Here, we have used the fact that $L_{p,p} = L_p$. Now using (7.1) for the pairs $(\alpha,p), (\alpha_0,p_0)$ and (α_1,p_1) shows that $\alpha' = \alpha$, as desired. □

REFERENCES

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

[**B-F**] C. de Boor and G. F. Fix, *Spline approximation by quasi-interpolants*, J. Approx. Theory **8** (1973), 19–45.

[**Br**] Yu. Brudnyi, *Approximation of functions of n-variables by quasi-polynomials*, Math. USSR Izv. **4** (1970), 568–586.

[**B-B**] P. L. Butzer and H. Berens, *Semi-groups of operators and approximation*, Springer-Verlag, New York, 1967.

[**C**] Z. Ciesielski, *Constructive function theory and spline systems*, Studia Math. **52** (1973), 277–302.

[**D**] R. DeVore, *Degree of approximation*, Approximation Theory, II (G. G. Lorentz, C. K. Chui, L. L. Schumaker, eds.), Academic Press, New York, 1976, pp. 117–162.

[**D-P**] R. DeVore and V. Popov, *Interpolation and non-linear approximation*, Proc. Conf. on Interpolation and Allied Topics in Analysis, Lund, 1986 (to appear).

[**D-P$_1$**] R. DeVore and V. A. Popov, *Free multivariate splines*, Constructive Approximation **3** (1987), 239–248.

[**D-S**] R. DeVore and K. Scherer, *A constructive theory for approximation by splines with an arbitrary sequence of knot sets*, Approximation Theory, Lecture Notes in Math., vol. 556, Springer-Verlag, New York, 1976.

[**D-Sh**] R. DeVore and R. Sharpley, *Maximal functions measuring smoothness*, Mem. Amer. Math. Soc., vol. 47, no. 293, 1984.

[**F-J**] M. Frazier and B. Jawerth, *Decomposition of Besov spaces*, preprint.

[**P**] J. Peetre, *New thoughts on Besov spaces*, Duke Univ. Math. Ser. I, Durham, N.C., 1976.

[**Pt**] P. Petrushev, *Direct and converse theorems for spline and rational approximation and Besov spaces*, Proc. Conf. on Interpolation and Applied Topics in Analysis, Lund, 1986 (to appear).

[**S-O**] E. A. Storozhenko and P. Oswald, *Jackson's theorem in the spaces* $L_p(\mathbf{R}^k)$, $0 < p < 1$, Siberian Math. J. **19** (1978), 630–639.

[**P-P**] V. A. Popov and P. Petrushev, *Rational approximation of real valued functions*, Encyclopedia of Math. and Applications, vol. 28, Cambridge Univ. Press, Cambridge, 1987.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SOUTH CAROLINA, COLUMBIA, SOUTH CAROLINA 29108

INSTITUTE OF MATHEMATICS, BULGARIAN ACADEMY OF SCIENCES, SOFIA, BULGARIA