Non-linear approximation and wavelets

Eugenio Hernández

Department of Mathematics, Universidad Autónoma de Madrid
28049 Madrid, España

Non-linear approximation with wavelets in Sobolev and Triebel-Lizorkin spaces

Gustavo Garrigós and Eugenio Hernández

1 Motivation

In this paper we shall consider the problem of non-linear approximation with wavelet like bases of the form $\{\psi_I : I \in \mathcal{D}\}$ given in (3.1) below. The non-linearity part refers to the fact that the approximation set Σ_n consists of all functions S of the form $S = \sum_{I\in\Delta} c_I\psi_I$, where $\Delta \subset \mathcal{D}$ has at most n elements and $c_I \in \mathbb{C}$.

Traditionally, in signal or image processing, the error of approximation is measured in the $L_2(\mathbb{R}^d)$ norm. For $L_p(\mathbb{R}^d)$ norms , $1 < p < \infty$, the problem has been treated in [DJP]. In this paper, we shall consider the case where the error is measured in the norm of Sobolev spaces $W^s_p(\psi, \mathbb{R}^d)$ or, more generally, in the Triebel-Lizorkin spaces $F^s_{p,q}(\psi, \mathbb{R}^d)$.

The motivation for considering these spaces comes from signal and image processing. While $\|f - S\|_2$ is the classical way of measuring the error, it does not take into account the possible deviations of an image or signal f from its approximation S in terms of its derivatives. From a theoretical point of view the use of Sobolev spaces to measure the error should produce a compressed image S visually closer to f than the traditional method.

By considering a definition of this spaces in terms of wavelet coefficients we are able to give simple proofs of our results and prove the sharpness of some of them. This illustrates the importance of having precise characterizations for the classical spaces in terms of wavelet coefficients.

This article contains material to be presented during the author's lecture at the workshop "Wavelets and Applications", Barcelona, July 1-6, 2002.

Key Words and phrases: Approximation, Sobolev spaces, Triebel-Lizorkin spaces, Jackson inequalities, Bernstein inequalities, wavelets .

2 The problem of approximation

The problem of approximation can be described in a general setting as follows. We take the point of view presented in [DP]. Let X be an abelian group under addition with a neutral element denoted by 0. A semi-quasi-norm on X is a map $|\ \ |_X : X \to \mathbb{R}^+$ such that

(i) $|x|_X \geq 0$ for all $x \in X$;

(ii) $|0|_X = 0$;

(iii) $|-x|_X = |x|_X$ for all $x \in X$;

(iv) $|x+y|_X^\nu \leq |x|_X^\nu + |y|_X^\nu$ for all $x, y \in X$ and all ν sufficiently small.

The pair $(X, |\ \ |_X)$ is called a semi-quasi-normed abelian group. Property (iv) is equivalent to $|x+y|_X \leq C(|x|_X + |y|_X)$ for all $x, y \in X$ and some $C \geq 1$ (see [BL], page 59).

Consider a semi-quasi-normed abelian group $(X, |\ \ |_X)$, and select in X a sequence of sets $\{\Sigma_n, n = 0, 1, 2, 3, \ldots\}$ such that

i) $\Sigma_0 = \{0\}$;

ii) $0 \in \Sigma_n$ for all $n = 1, 2, 3, \ldots$

iii) $\Sigma_n \subset \Sigma_{n+1}$ for all $n = 1, 2, 3, \ldots$

iv) $\bigcup_{n=1}^\infty \Sigma_n$ is dense in X.

The family $\{\Sigma_n : n = 0, 1, 2, 3, \ldots\}$ will be called and **approximation family for** X. For an element $x \in X$ we define its approximation error from elements of $\{\Sigma_n : n = 0, 1, 2, 3, \ldots\}$ by

$$\sigma_X(x, \Sigma_n) = \inf_{S \in \Sigma_n} |x - S|_X \,. \tag{2.1}$$

Notice that $\sigma_X(x, \Sigma_0) = |x|_X$. Given $\lambda > 0$ one of the problems of approximation is to find a semi-quasi-normed abelian group $(Y, |\ \ |_Y)$ such that $\bigcup_{n=1}^\infty \Sigma_n \subset Y \subset X$ and, moreover,

$$|y|_X \leq C|y|_Y, \quad \text{for all } y \in Y, \tag{2.2}$$

and

$$\sigma_X(y, \Sigma_n) \leq C n^{-\lambda} |y|_Y \,, \qquad n = 1, 2, 3 \ldots, \quad \text{for all } y \in Y \,, \tag{2.3}$$

where C is independent of λ and y. Inequalities such as (2.3) are called **Jackson type inequalities**.

Another problem in approximation theory is to characterize the approximation space $A_\infty^\alpha(X)$, $\alpha > 0$, defined as

$$A_\infty^\alpha(X) = \{x \in X : \sup_{n \in \mathbb{N}} n^\alpha \sigma_X(x, \Sigma_n) < +\infty\} \,. \tag{2.4}$$

More generally, given $\alpha > 0$, and $0 < q < \infty$, we can ask to characterize the approximation space $A_q^\alpha(X)$ given by

$$A_q^\alpha(X) = \{x \in X : (\sum_{n=1}^{\infty} [n^\alpha \sigma_X(x, \Sigma_n)]^q \ \frac{1}{n} \ )^{1/q} < +\infty\}\,. \tag{2.5}$$

For $x \in X,\ \alpha > 0$, and $0 < q < \infty$, define

$$|x|_{A_q^\alpha(X)} = (\sum_{n=1}^{\infty} [n^\alpha \sigma_X(x, \Sigma_n)]^q \ \frac{1}{n} \ )^{1/q}\,. \tag{2.6}$$

and if $q = \infty$,

$$|x|_{A_\infty^\alpha(X)} = \sup_{n \in \mathbb{N}} \ n^\alpha \sigma_X(x, \Sigma_n)\,. \tag{2.7}$$

If we assume that there exists $c > 0$ such that

$$\Sigma_n \pm \Sigma_m \subset \Sigma_{c(n+m)} \quad \text{for all } n, m \in \mathbb{N}\,, \tag{2.8}$$

it can be proved that the spaces $A_q^\alpha(X),\ \alpha > 0,\ 0 < q \le \infty$, are semi-quasi-normed abelian groups with the semi-quasi-norm given by (2.6) if $q < \infty$ and (2.7) if $q = \infty$. It is clear that if Y satisfies (2.3), $|y|_{A_\infty^\lambda(X)} \le C|y|_Y$ for all $y \in Y$, so that $Y \subset A_\infty^\lambda(X)$. Similarly, if (2.3) is satisfied, for all $0 < \epsilon < \lambda,\ 0 < q < \infty$, and all $y \in Y$ it is easy to see that $|y|_{A_q^{\lambda-\epsilon}(X)} \le C|y|_Y$ and $Y \subset A_q^{\lambda-\epsilon}(X)$.

The general setting described above is adapted to linear as well as non-linear approximation. Suppose that $\mathcal{B} = \{e_1, e_2, \dots\}$ is a basis for X (see [HW], section 5.1). If we take $\Sigma_n = \text{span}\,\{e_1, \dots, e_n\},\ n = 1, 2, 3, \dots$, the spaces Σ_n are linear subspaces of X and we are in a situation of **linear approximation**. If we take Σ_n as the set of all functions S of the form $S = \sum_{i \in \Delta} c_i e_i$, where $\Delta \subset \mathbb{N}$ has at most n elements, then, the subsets Σ_n are not linear, since $\Sigma_n \pm \Sigma_n$ is not contained in Σ_n (we have $\Sigma_n \pm \Sigma_n \subset \Sigma_{2n}$). This is a situation of **non-linear approximation**.

3 Wavelet bases and function spaces

In this paper we shall be interested in non-linear approximation for functions in Besov and Sobolev spaces, when these are described in terms of wavelet type bases. More precisely, a wavelet basis is generated by a finite family of functions $\Psi = \{\psi_1, \dots, \psi_L\} \subset L_2(\mathbb{R}^d)$ by performing to each ψ_l integer translations and dyadic dilations in the following way: for a dyadic cube of the form $I_{j,k} = 2^{-j}\,([0,1)^d + k)$, where $j \in \mathbb{Z}$ and $k \in \mathbb{Z}^d$, we define

$$\psi_{l;j,k}(x) = \psi_{l;I_{j,k}}(x) = 2^{jd/2}\,\psi_l(2^j x - k)\,. \tag{3.1}$$

To simplify the notation we shall write $\mathcal{D}$ for the collection of all dyadic cubes of $\mathbb{R}^d$ and $\Lambda = \{(l;I) : l = 1, \dots, L, I \in \mathcal{D}\}$. We say that Ψ is and **orthonormal wavelet system** when the set

$$\mathcal{S}(\Psi) = \{\psi_{l;I} : (l,I) \in \Lambda\}$$

is an orthonormal basis of $L_2(\mathbb{R}^d)$.

From a theoretical point of view, wavelet systems provide unconditional bases for many classical function spaces, such as Lebesgue spaces, Sobolev spaces, or Besov spaces. Moreover, equivalent norms can be expressed with weighted sums of the wavelet coefficients (see [Me], [HW]). This makes the systems $\mathcal{S}(\Psi)$ a suitable basis to study the approximation problem described in section 2. In order to deal with these function spaces we need to impose further regularity conditions on $\psi_l \in \Psi$. The starting point is the following proposition which characterizes the Lebesgue spaces $L_p(\mathbb{R}^d), 1 < p < \infty$ in terms of the wavelet coefficients.

Proposition 3.1. *Let $\Psi = \{\psi_1, \dots, \psi_L\}$ be an orthonormal wavelet system in $L_2(\mathbb{R}^d)$. Suppose that $\psi_l \in C^\alpha(\mathbb{R}^d)$ for some $\alpha > 0$ and all $l = 1, 2, \dots, L$, and there exists $\epsilon > 0$ such that*

$$|\psi_l(x)| \le \frac{C}{(1+|x|)^{d+\epsilon}}, \qquad x \in \mathbb{R}^d,\ l = 1, 2, \dots, L. \tag{3.2}$$

Then, for all $1 < p < \infty$, $\mathcal{S}(\Psi)$ is an unconditional basis for $L_p(\mathbb{R}^d)$. Moreover, for all $f \in L_p(\mathbb{R}^d)$,

$$\|f\|_{L_p(\mathbb{R}^d)} \sim \|(\sum_{(l,I)\in\Lambda} |\langle f, \psi_{l;I}\rangle|^2\, |I|^{-1}\chi_I(\cdot))^{1/2}\|_{L_p(\mathbb{R}^d)}. \tag{3.3}$$

This result is shown when $d = 1$ in [KL] (Chapter 6), but the same proof given there works for the case $d > 1$ (see also [Me](section 6.2), where the result is proved with stronger regularity conditions).

Suppose that Ψ is an orthonormal wavelet system as in Proposition 3.1 and let $1 < p < \infty$. Then, every $f \in L_p(\mathbb{R}^d)$ can be uniquely written as

$$f = \sum_{\lambda\in\Lambda} a_\lambda(f)\psi_\lambda \qquad \text{where} \qquad a_\lambda(f) = \langle f, \psi_\lambda\rangle.$$

We now define several subspaces of $L_p(\mathbb{R}^d)$, $1 < p < \infty$. If s is a real number such that $s \ge 0$, let

$$|f|_{W_p^s(\Psi,\mathbb{R}^d)} = \|(\sum_{(l;I)\in\Lambda} |a_{l;I}(f)|^2 |I|^{-2(\frac{s}{d}+\frac{1}{2})}\chi_I(\cdot))^{\frac{1}{2}}\|_{L_p(\mathbb{R}^d)}\,, \tag{3.4}$$

and consider the **Sobolev space** $W_p^s(\Psi,\mathbb{R}^d)$ as the set of all functions $f \in L_p(\mathbb{R}^d)$ for which $|f|_{W_p^s(\Psi,\mathbb{R}^d)} < +\infty$, with the norm given by

$$\|f\|_{W_p^s(\Psi,\mathbb{R}^d)} = \|f\|_{L_p(\mathbb{R}^d)} + |f|_{W_p^s(\Psi,\mathbb{R}^d)}\,. \tag{3.5}$$

More generally, if s and r are real numbers such that $s \geq 0$ and $0 < r \leq \infty$, we define the **Triebel-Lizorkin space** $F_{p,r}^s(\Psi,\mathbb{R}^d)$ as the set of all $f \in L_p(\mathbb{R}^d)$ such that $|f|_{F_{p,r}^s(\Psi,\mathbb{R}^d)} < +\infty$ where

$$|f|_{F_{p,r}^s(\Psi,\mathbb{R}^d)} = \|(\sum_{(l;I)\in\Lambda} |a_{l;I}(f)|^r |I|^{-r(\frac{s}{d}+\frac{1}{2})}\chi_I(\cdot))^{\frac{1}{r}}\|_{L_p(\mathbb{R}^d)}\,. \tag{3.6}$$

A quasi norm in $F_{p,r}^s(\Psi,\mathbb{R}^d)$ is given by

$$\|f\|_{F_{p,r}^s(\Psi,\mathbb{R}^d)} = \|f\|_{L_p(\mathbb{R}^d)} + |f|_{F_{p,r}^s(\Psi,\mathbb{R}^d)}\,. \tag{3.7}$$

It is clear that $F_{p,2}^s(\Psi,\mathbb{R}^d) = W_p^s(\Psi,\mathbb{R}^d)$, and by Proposition 3.1, $F_{p,2}^0(\Psi,\mathbb{R}^d) = L_p(\mathbb{R}^d)$, with equivalent norms.

For the purpose of approximation we also need to introduce a class of Besov spaces. Let β, τ and q be real numbers such that $\beta > 0$, $0 < \tau, q \leq \infty$ and $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$. Take Ψ as in Proposition 3.1 and define the **Besov space** $B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)$ as the set of all $f \in L_p(\mathbb{R}^d)\cap L_\tau(\mathbb{R}^d)$ such that $|f|_{B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)} < +\infty$ where

$$|f|_{B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)} = \left(\sum_{j\in\mathbb{Z}}\left(\sum_{l=1}^{L}\sum_{k\in\mathbb{Z}^d} |a_{l;j,k}(f)|^\tau |I_{j,k}|^{-\tau(\frac{\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}\right)^{\frac{q}{\tau}}\right)^{1/q}\,. \tag{3.8}$$

A quasi-norm in $B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)$ is given by

$$\|f\|_{B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)} = \|f\|_{L_\tau(\mathbb{R}^d)} + |f|_{B_{\tau,q}^\beta(\Psi,\mathbb{R}^d)}\,. \tag{3.9}$$

Under sufficient regularity conditions on Ψ one can show that these spaces are complete, do not depend on Ψ, and actually they coincide with the classical Sobolev, Triebel-Lizorkin and Besov spaces. There are many research papers devoted to these characterizations, but non of them with such generality that treat all cases (the case $\tau < 1$ is typically avoided), or include minimal regularity assumptions on the wavelet system Ψ (see [Bo], [GHT], and [Si]). For the purpose of this note we do not need the classical expressions of these quasi-norms, and it will suffice with the presentation given above. This illustrates in particular the power of wavelet based techniques in the theoretical study of non-linear approximation.

As a final remark we must observe that, in most applications, the success of wavelet techniques is in many cases due to the good properties related to the dyadic structure. For instance, the Multiresolution Analysis (MRA) of S. Mallat ([Ma]) provides fast numerical algorithms for the analysis and synthesis of signals. For a realistic implementation of such techniques it is essential that the functions $\psi_l \in \Psi$ are compactly supported. However, the construction of compactly supported orthonormal wavelets in $L_2(\mathbb{R}^d)$ is a difficult task for which at the moment there are no satisfactory examples (besides the tensor product of one dimensional wavelets).

In recent years, alternative constructions for wavelet systems have appeared, where the orthonormal condition is dropped in favor of pairs of **biorthogonal wavelet bases** ([CDF], [Co]). A further generalization which admits compactly supported functions are the **frame wavelet systems** ([GR]). This is a pair of systems Ψ and $\tilde{\Psi}$ such that every $f \in L_2(\mathbb{R}^d)$ can be written as

$$f = \sum_{\lambda\in\Lambda} \langle f, \tilde{\psi}_\lambda \rangle \psi_\lambda = \sum_{\lambda\in\Lambda} \langle f, \psi_\lambda \rangle \tilde{\psi}_\lambda \,, \tag{3.10}$$

with unconditional convergence in $L_2(\mathbb{R}^d)$, and so that the coefficients satisfy the following stability condition

$$\frac{1}{C}\left(\sum_{\lambda\in\Lambda} |\langle f, \tilde{\psi}_\lambda \rangle|^2\right)^{1/2} \leq \|f\|_{L_2(\mathbb{R}^d)} \leq C\left(\sum_{\lambda\in\Lambda} |\langle f, \tilde{\psi}_\lambda \rangle|^2\right)^{1/2}, \tag{3.11}$$

and similary for $\{\langle f, \psi_\lambda \rangle : \lambda \in \Lambda\}$. Frame systems were extensively studied by M. Frazier and B. Jawerth ([FJ]) showing that they also provide norm characterizations for function spaces by means of weighted sums of the coefficients $\{\langle f, \psi_\lambda \rangle : \lambda \in \Lambda\}$.

The techniques we develop in this paper are general enough to be applied to any of the systems above.

4 Approximation and real interpolation

It turns out that, in many instances, the approximation spaces $A_q^\alpha(X)$ can be identified with interpolation spaces obtained by the real method of interpolation. References to this result in particular settings can be found in [DP] (Theorem 3.1) and [Co] (Theorem 38.1; see also Theorem 28.2 and Remark 28.2).

For a pair of compatible semi-quasi-normed abelian groups $(X, |\ \ |_X)$ and $(Y, |\ \ |_Y)$, with $Y \subset X$, the **K functional** is defined as

$$K(t, x) = K(t, x; X, Y) = \inf_{y\in Y}\{|x - y|_X + t|y|_Y\}, \qquad x \in X,\ t \in \mathbb{R}. \tag{4.1}$$

To state the next result we need to consider a sort of inverse estimate to (2.3), namely

$$|S|_Y \le Cn^{\lambda}|S|_X \quad \text{for all } S \in \Sigma_n,\ n = 1, 2, 3, \ldots, , \tag{4.2}$$

which is called **Bernstein type inequality**.

Theorem 4.1. *([DP], [Co]) Let $Y \subset X$ be two semi-quasi-normed abelian groups and $\{\Sigma_n : n = 0, 1, 2, \ldots\}$ an approximation family for X as defined in section 2 with $\bigcup_{n\in\mathbb{N}} \Sigma_n \subset Y$. Suppose that there exists $\lambda > 0$ such that the Jackson inequality (2.3) and the Bernstein inequality (4.2) hold. Suppose, also, that there exists an integer k such that*

$$\Sigma_n \pm \Sigma_m \subset \Sigma_{k(n+m)}. \tag{4.3}$$

Then, for $\alpha \in (0, \lambda)$ and $0 < q \le \infty$, there exists $C > 0$ such that

$$i)\ \ \|(2^{j\alpha}\sigma_X(x, \Sigma_{2^j}))_{j\ge 0}\|_{\ell_q} \le C\|(2^{j\alpha}K(x, 2^{-\lambda j}))_{j\in\mathbb{Z}}\|_{\ell_q} \quad \text{for all } x \in X$$

and

$$ii)\ \ \|(2^{j\alpha}K(x, 2^{-\lambda j}))_{j\in\mathbb{Z}}\|_{\ell_q} \le C\left[\ |x|_X + \|(2^{j\alpha}\sigma_X(x, \Sigma_{2^j}))_{j\ge 0}\|_{\ell_q}\right], \quad \text{for all } x \in X.$$

It is easy to see, due to the monotonicity of $\sigma_X(x, \Sigma_n)$, that

$$|x|_{A_q^\alpha(X)} \approx \|(2^{j\alpha}\sigma_X(x, \Sigma_{2^j}))_{j\ge 0}\|_{\ell_q} \quad x \in X,\ \alpha > 0,\ 0 < q \le \infty\,. \tag{4.4}$$

For $0 < \theta < 1$ and $0 < q \le \infty$ define the interpolation space

$$(X, Y)_{\alpha,q} = \left\{ x \in X : |x|_{(X,Y)_{\alpha,q}} = \left(\int_0^\infty \left[t^{-\theta}\ K(x, t)\right]^q\ \frac{dt}{t}\right)^{1/q} < +\infty \right\}. \tag{4.5}$$

It can be shown (see [BL], page 41) that for any real number $b > 1$

$$\int_0^\infty \left[t^{-\theta}\ K(x, t)\right]^q\ \frac{dt}{t} \approx \sum_{j\in\mathbb{Z}} \left[\ b^{\alpha j}\ K(x, b^{-j})\right]^q\ , \qquad \text{for all } x \in X. \tag{4.6}$$

Corollary 4.2. *With the same hypothesis as in Theorem 4.1, we have*

$$A_q^\alpha(X) = (X, Y)_{\alpha/\lambda,q}, \tag{4.7}$$

and there exist C_1, $C_2 > 0$ such that

$$C_1\ |x|_{A_q^\alpha(X)} \le |x|_{(X,Y)_{\alpha/\lambda,q}}\ \le\ C_2\ \left[\ |x|_X + |x|_{A_q^\alpha(X)}\right], \tag{4.8}$$

Proof. The inclusion $A_q^\alpha(X) \subset (X, Y)_{\alpha/\lambda,q}$ and the right hand side inequality follow from (4.6) and ii) of Theorem 4.1. The reverse inclusion and the left hand side inequality follow from i) of Theorem 4.1. □

The approximation spaces $A_q^\alpha(X)$, $\alpha > 0$, $0 < q \le \infty$, form an interpolation family for the real method of interpolation (see Lemma 4.4 below). To show this we start proving that the approximation spaces $A_q^\alpha(X)$ satisfy Jackson and Bernstein's inequalities.

Lemma 4.3. *The spaces $A_q^\lambda(X)$, $\lambda > 0$, $0 < q \le \infty$, satisfy the Jackson inequality*

$$i)\ \ \sigma_X(y, \Sigma_n) \ \le\ C\ n^{-\lambda}\ |y|_{A_q^\lambda(X)}, \qquad n = 1, 2, 3, \dots, \quad \textit{for all } y \in Y\,,$$

and the Bernstein inequality

$$|S|_{A_q^\lambda(X)} \ \le\ C\ n^\lambda\ |S|_X \qquad \textit{for all } S \in \Sigma_n, \quad n = 1, 2, 3, \dots$$

Proof. i) If $q = \infty$, from (2.7) we deduce the desired result with $C = 1$. If $q < \infty$, we use the equivalence (4.4) to deduce that $2^{j\lambda}\ \sigma_X(y, \Sigma_{2^j}) \ \le\ C\ |y|_{A_q^\lambda(X)}$ for all $j \ge 0$. For any $n \in N$ choose $j \ge 0$ such that $2^j < n \le 2^{(j+1)}$. Then,

$$n^\lambda \sigma_X(y, \Sigma_n) \ \le 2^{(j+1)\lambda}\ \sigma_X(y, \Sigma_{2^j}) \ \le\ C\ |y|_{A_q^\lambda(X)}$$

which shows the desired result.

ii) If $S \in \Sigma_{2^{j_0}}$ it is clear that $\sigma_X(S, \Sigma_{2^m}) = 0$ for all $m \ge j_0$. Thus, using (4.4) we obtain

$$|S|^q_{A_q^\lambda(X)} \ \le C \sum_{j=0}^{j_0-1} [2^{j\lambda}\ \sigma_X(S, \Sigma_{2^j})]^q \ \le C|S|^q_X \sum_{j=0}^{j_0-1} 2^{j\lambda q} \ \le C 2^{j_0\lambda q} |S|^q_X\,.$$

The desired result follows easily from this inequality. □

Lemma 4.4. *The family of spaces $\{A_q^\alpha(X) : \alpha > 0, 0 < q \le \infty\}$, is an interpolation family for the real method of interpolation, that is, if $0 < \alpha_0 < \alpha_1$, $0 < q_0, q_1 \le \infty$, and $0 < \theta < 1$ we have*

$$(A_{q_0}^{\alpha_0}(X),\ A_{q_1}^{\alpha_1}(X))_{\theta,q} = A_q^\alpha(X), \qquad \alpha = (1-\theta)\alpha_0 + \theta\alpha_1.$$

Proof. From Corollary 4.2 we conclude that $A_q^\alpha(X) = (X, A_r^\lambda)_{\alpha/\lambda,q}$ for any $\alpha \in (0, \lambda)$ and any q such that $0 < q \le \infty$. Choose $\lambda > \max\{\alpha_0, \alpha_1\}$. By the reiteration theorem for the real method of interpolation (see [BL], page 50) we obtain:

$$(A_{q_0}^{\alpha_0}(X),\ A_{q_1}^{\alpha_1}(X))_{\theta,q} = \Big((X, A_r^\lambda)_{\alpha_0/\lambda, q_0}\,, (X, A_r^\lambda)_{\alpha_1/\lambda, q_1}\Big)_{\theta,q} = (X, A_r^\lambda)_{\alpha/\lambda,q} = A_q^\alpha(X).$$

□

5 Jackson type inequalities

We want to prove inequalities of Jackson type when the error is measured in the Sobolev space $W_p^s(\Psi, \mathbb{R}^d)$ or, more generally, in the Triebel-Lizorkin space $F_{p,r}^s(\Psi, \mathbb{R}^d)$. For $s \ge 0$, $1 < p < \infty$, $0 < r \le \infty$, consider

$$\sigma_{F_{p,r}^s(\Psi,\mathbb{R}^d)}(f, \Sigma_n) = \inf_{S\in\Sigma_n}\ |f - S|_{F_{p,r}^s(\Psi,\mathbb{R}^d)}, \qquad n = 1, 2, 3, \dots \tag{5.1}$$

where Σ_n is the set of all elements of the form $\sum_{\lambda\in\Delta} c_\lambda\ \psi_\lambda$, and $\Delta \subset \Lambda$ has at most n elements. We are interested in proving inequalities of the form

$$\sigma_{F^s_{p,r}(\Psi,\mathbb{R}^d)}(f,\Sigma_n) \le Cn^{-\lambda}\ |f|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)}, \qquad n = 1,2,3,\dots \tag{5.2}$$

for all $f \in B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)$. Here we are assuming that $\{\Psi_\lambda : \lambda \in \Lambda\}$ is an orthonormal wavelet system for $L_2(\mathbb{R}^d)$ as in Proposition 3.1. The arguments in sections 5 and 6 also work, with obvious modifications, if we consider pairs of frame wavelet systems.

There is a relation among some of the parameters that appear in (5.2); this is proved in the following lemma using a homogeneity argument.

Proposition 5.1. *If Jackson's inequality (5.2) holds, we must have* $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$.

Proof. We present the proof for $\Psi = \{\psi\}$ and write

$$f = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^d} a_{j,k}(f)\psi_{j,k}, \qquad a_{j,k}(f) = \langle f, \psi_{j,k}\rangle.$$

Let $D_l f(x) = 2^{ld/2} f(2^l x)$ be the dyadic dilation operator normalized in $L_2(\mathbb{R}^d)$. Then,

$$D_l f(x) = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}^d} a_{j-l,k}(f)\psi_{j,k}. \tag{5.3}$$

Write

$$\mathcal{G}_r(x,f) = (\sum_{j\in\mathbb{Z}}\sum_{j\in\mathbb{Z}^d} |a_{j,k}(f)|^r 2^{jdr(\frac{s}{d}+\frac{1}{2})}\chi_{I_{j,k}}(x))^{\frac{1}{r}}\ .$$

which is the expression that appears in the definition of $|f|_{F^s_{p,r}(\psi,\mathbb{R}^d)}$. Then,

$$|D_l(f)|_{F^s_{p,r}(\psi,\mathbb{R}^d)} = \|(\sum_{j\in\mathbb{Z}}\sum_{j\in\mathbb{Z}^d} |a_{j,k}(f)|^r 2^{(j+l)dr(\frac{s}{d}+\frac{1}{2})}\chi_{I_{j+l,k}}(x))^{\frac{1}{r}}\|_{L_p(\mathbb{R}^d)}\ .$$

With the change of variables $x = 2^{-l}y$, if $x \in I_{j+l,k} = 2^{-(j+l)}([0,1)^d + k)$ we have $y \in I_{j,k}$, so that

$$|D_l(f)|_{F^s_{p,r}(\psi,\mathbb{R}^d)} = 2^{ld(\frac{s}{d}+\frac{1}{2})}\ 2^{-\frac{ld}{p}}\|\mathcal{G}_r(\cdot,f)\|_{L_p(\mathbb{R}^d)} = 2^{ld(\frac{s}{d}+\frac{1}{2}-\frac{1}{p})}\ |f|_{F^s_{p,r}(\psi,\mathbb{R}^d)}\ , \tag{5.4}$$

which shows the homogeneity relation for $|\ \cdot\ |_{F^s_{p,r}(\psi,\mathbb{R}^d)}$.

For the Besov spaces we have, using (5.3),

$$|D_l(f)|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} = \left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}^d} |a_{j-l,k}(f)|^\tau |I_{j,k}|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}\right)^{\frac{q}{\tau}}\right)^{1/q}$$
$$= \left(\sum_{j\in\mathbb{Z}}\left(\sum_{k\in\mathbb{Z}^d} |a_{j,k}(f)|^\tau |I_{j+l,k}|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}\right)^{\frac{q}{\tau}}\right)^{1/q}$$

$$= 2^{ld(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \; |f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \, . \tag{5.5}$$

If Jackson's inequality (5.2) holds, from (5.4) and (5.5) we deduce

$$\begin{aligned} &2^{ld(\frac{s}{d}+\frac{1}{2}-\frac{1}{p})} \; \sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)} \; (f,\Sigma_n) = \sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)} \; (D_l f,\Sigma_n) \\ &\le C \; n^{-\lambda} \; |D_l f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} = C n^{-\lambda} \; 2^{ld(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \; |f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \, , \end{aligned}$$

for all $f \in B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)$. Thus,

$$\frac{n^{\lambda}\sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)} \; (f,\Sigma_n)}{C \; |f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)}} \;\le\; 2^{ld(\frac{\beta}{d}+\frac{1}{p}-\frac{1}{\tau})} \tag{5.6}$$

for all $f \in B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)$. Since there are functions f for which the left hand side of (5.6) is non zero, letting $l \to \infty$ and $l \to -\infty$ we obtain the desired result. $\square$

We are now ready to prove the main result of this section, namely a Jackson type inequality when the error of approximation is measured in Sobolev and Triebel-Lizorkin spaces.

Theorem 5.2. *Let us consider real numbers* $s \ge 0$, $\beta > 0$, $0 < \tau, q, r < \infty$, *and* $1 < p < \infty$, *such that* $q < p$ *and* $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$ *(d = dimension). Suppose that* $\{\Psi_\lambda : \lambda \in \Lambda\}$ *is an orthonormal wavelet system for* $L_2(\mathbb{R}^d)$ *as in Proposition 3.1.*

i) If $\tau \le q$, *for all* $f \in B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)$ *we have*

$$\sigma_{F^s_{p,r}(\Psi,\mathbb{R}^d)}(f,\Sigma_n) \;\le\; C \; n^{-(\frac{1}{q}-\frac{1}{p})} \; |f|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)} \qquad n = 1,2,3,\ldots$$

ii) If $q \le \tau$, *for all* $f \in B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)$ *we have*

$$\sigma_{F^s_{p,r}(\Psi,\mathbb{R}^d)}(f,\Sigma_n) \;\le\; C \; n^{-(\frac{1}{\tau}-\frac{1}{p})} \; |f|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)} \qquad n = 1,2,3,\ldots$$

Proof. We present the proof for $\Psi = \{\psi\}$. The proof in the general case can easily be adapted from this. We start showing that ii) follows from i). In fact, i) with $q = \tau$ gives

$$\sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)}(f,\Sigma_n) \;\le\; C \; n^{(\frac{1}{\tau}-\frac{1}{p})} \; |f|_{B^{s+\beta}_{\tau,\tau}(\psi,\mathbb{R}^d)} \qquad n = 1,2,3,\ldots$$

Since $q \le \tau$ in case ii), $\ell_q \subset \ell_\tau$ and hence $|f|_{B^{s+\beta}_{\tau,\tau}(\psi,\mathbb{R}^d)} \;\le\; |f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)}$.

By homogeneity, it is enough to show i) for $f \in B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)$ with $|f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} = 1$. Write $f = \sum_{I\in\mathcal{D}} a_I(f)\,\psi_I$. We claim that there exist at most n coefficients $a_I(f), I \in \mathcal{D}$, satisfying

$$|a_I(f)|^\tau |I|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \;\ge\; \frac{1}{n^{\tau/q}} \, . \tag{5.7}$$

To see this, suppose that there are more than n indices I satisfying (5.7). If n_j, $j = 1, 2, \dots, m$ is the number of such coefficients $a_I(f)$ belonging to the same dyadic level j we must have $n_1 + n_2 + \cdots + n_m > n$ and, since $q/\tau \geq 1$,

$$1 = |f|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \geq (\frac{n_1}{n^{\tau/q}})^{q/\tau} + \cdots + (\frac{n_m}{n^{\tau/q}})^{q/\tau} \\ \geq \frac{1}{n}(n_1 + \cdots + n_m) > 1\,,$$

which is a contradiction.

Let Δ be the set of at most n elements $I \in \mathcal{D}$ for which (5.7) is true. For each $I \in \Delta$ we have

$$|a_I(f)| \;<\; \frac{1}{n^{1/q}}\; |I|^{\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau}}\,. \tag{5.8}$$

Define $S = \sum_{I\in\Delta} a_{I(f)}\psi_I \in \Sigma_n$ and $E = f - S$. Since $\sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)}(f, \Sigma_n) \;\leq\; |f - S|_{F^s_{p,r}(\psi,\mathbb{R}^d)}$ we only need to estimate

$$|E|_{F^s_{p,r}(\psi,\mathbb{R}^d)} = \|(\sum_{I\notin\Delta} |a_I(f)|^r |I|^{-r(\frac{s}{d}+\frac{1}{2})}\; \chi_I(\cdot))^{1/r}\|_{L_p(\mathbb{R}^d)} = \|\tilde{E}\|_{L_p(\mathbb{R}^d)} \tag{5.9}$$

where

$$\tilde{E}(x) = \left(\sum_{I\notin\Delta} |a_I(f)|^r |I|^{-r(\frac{s}{d}+\frac{1}{2})}\; \chi_I(x)\right)^{1/r}.$$

For a natural number N, that will be chosen later in an appropriate way, we split the sum defining $\tilde{E}(x)$ into two parts, $\tilde{E}(x) = \tilde{E}_1(x) + \tilde{E}_2(x)$, where $\tilde{E}_1$ contains all the terms with $I \notin \Delta$ such that $\ell(I) = |I|^{1/d} \geq 2^N$ and $\tilde{E}_2$ the rest.

To estimate $\tilde{E}_1$ for $\tau \leq q$ use (5.8) and the fact that each x is exactly in one dyadic cube of any given side length to obtain

$$[\tilde{E}_1(x)]^r < \frac{1}{n^{r/q}} \sum_{\{I\notin\Delta,\,\ell(I)\geq 2^N\}} |I|^{-r(\frac{s}{d}+\frac{1}{2})}\; |I|^{r(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}\; \chi_I(x) \\ \leq \frac{1}{n^{r/q}} \sum_{j=N}^{\infty} (2^{jdr})^{\frac{\beta}{d}-\frac{1}{\tau}} = \frac{1}{n^{r/q}} \sum_{j=N}^{\infty} 2^{-jdr/p} \\ = C^r_{p,r}\frac{1}{n^{r/q}}\;\; 2^{-Ndr/p}.$$

Thus, for the large intervals we have

$$\tilde{E}_1(x) \leq C_{p,r}\frac{1}{n^{1/q}}\;\; 2^{-Nd/p}\,. \tag{5.10}$$

Let $t > 0$, and choose $N = N(t) \in \mathbb{N}$ such that

$$C_{p,r}\frac{1}{n^{1/q}}\;\; 2^{-Nd/p} \leq \frac{t}{2} < C_{p,r}\frac{1}{n^{1/q}}\;\; 2^{-(N-1)d/p}\,.$$

Let $\mu(\tilde{E},t) = |\{x \in \mathbb{R}^d : |\tilde{e}(x)| > t\}|$ be the distribution function of $\tilde{E}$. By (5.10) and our choice of N we have $\mu(\tilde{E},t) \le \mu(\tilde{E}_2, t/2)$. An estimate for $\tilde{E}_2$ is obtained in the following way:

$$\begin{aligned}\tilde{E}_2(x) &\le \left(\sum_{\{I\in\mathcal{D},\, \ell(I)<2^N\}} |a_I(f)|^r |I|^{-r(\frac{s}{d}+\frac{1}{2})} \chi_I(x) \right)^{1/r} \\ &= \left(\sum_{\{I\in\mathcal{D},\, \ell(I)<2^N\}} |a_I(f)|^r \; |I|^{r(\frac{1}{p}-\frac{1}{q}-(\frac{s}{d}+\frac{1}{2}))} \; |I|^{r(\frac{1}{q}-\frac{1}{p})} \chi_I(x) \right) \\ &\le \left[\sup_{I\in\mathcal{D}} |a_I(f)| \; |I|^{\frac{1}{p}-\frac{1}{q}-(\frac{s}{d}+\frac{1}{2})} \chi_I(x) \right] \left[\sum_{\{I\in\mathcal{D},\, \ell(I)<2^N\}} |I|^{r(\frac{1}{q}-\frac{1}{p})} \chi_I(x) \right]^{1/r}.\end{aligned}$$

Since $\ell_\tau \subset \ell_\infty$ and $\ell_q \subset \ell_\infty$ we obtain, using $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$,

$$\tilde{E}_2(x) \le \; C \; \mathcal{B}_f(x) \left[\sum_{j=-\infty}^{N+1} 2^{jd(\frac{1}{q}-\frac{1}{p})} \right] = \; C \; \mathcal{B}_f(x) 2^{Nd(\frac{1}{q}-\frac{1}{p})} \tag{5.11}$$

where

$$\mathcal{B}_f(x) = \left[\sum_{j\in\mathbb{Z}} \left(\sum_{k\in\mathbb{Z}^d} |a_{j,k}(f)|^\tau \; |I_{j,k}|^{-\tau(\frac{1}{q}+\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \chi_{I_{j.k}}(x) \right)^{q/\tau} \right]^{1/q}$$

and we have used $q < p$ to evaluate the last sum in (5.11). From the definition of N we deduce $2^{Nd} < Cn^{-p/q} t^{-p}$. From (5.11) we obtain

$$\tilde{E}_2(x) \le C \; \mathcal{B}_f(x) \; n^{-\frac{p}{q}(\frac{1}{q}-\frac{1}{p})} \; t^{-\frac{p}{q}+1}$$

where we have used again $q < p$. Returning to the distribution function we can write

$$\mu(\tilde{E},t) \le \mu(\tilde{E}_2, t/2) \le \mu(\mathcal{B}_f(x), \; Cn^{\frac{p}{q}(\frac{1}{q}-\frac{1}{p})} \; t^{\frac{p}{q}}). \tag{5.12}$$

Therefore,

$$\begin{aligned}\|\tilde{E}\|^p_{L_p(\mathbb{R}^d)} &= p \int_0^\infty \mu(\tilde{E},t)\, t^p \frac{dt}{t} \le p \int_0^\infty \mu(\mathcal{B}_f(x), \; Cn^{\frac{p}{q}(\frac{1}{q}-\frac{1}{p})} \; t^{\frac{p}{q}})\, t^p \frac{dt}{t} \\ &= C \int_0^\infty \mu(\mathcal{B}_f(x), \; u)\, n^{-p(\frac{1}{q}-\frac{1}{p})} \, u^q \frac{du}{u} = Cn^{-p(\frac{1}{q}-\frac{1}{p})} \|\mathcal{B}_f\|^q_{L_q(\mathbb{R}^d)},\end{aligned} \tag{5.13}$$

where we have done the change of variables $u = Cn^{\frac{p}{q}(\frac{1}{q}-\frac{1}{p})} \; t^{\frac{p}{q}}$. Observe that

$$\begin{aligned}\|\mathcal{B}_f(x)\|^q_{L_q(\mathbb{R}^d)} &= \sum_{j\in\mathbb{Z}} \int_{\mathbb{R}^d} \left(\sum_{k\in\mathbb{Z}^d} |a_{j,k}(f)|^\tau \; |I_{j,k}|^{-\tau(\frac{1}{q}+\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \chi_{I_{j.k}}(x) \right)^{q/\tau} dx \\ &= \sum_{j\in\mathbb{Z}} \left[\sum_{k\in\mathbb{Z}^d} \int_{I_{j,k}} |a_{j,k}(f)|^q \; |I_{j,k}|^{-q(\frac{1}{q}+\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \chi_{I_{j.k}}(x)\, dx \right]\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j\in\mathbb{Z}} \left[\sum_{k\in\mathbb{Z}^d} |a_{j,k}(f)|^q \; |I_{j,k}|^{-q(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \right] \\
&\le C \sum_{j\in\mathbb{Z}} \left[\sum_{k\in\mathbb{Z}^d} |a_{j,k}(f)|^\tau \; |I_{j,k}|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \right]^{q/\tau} \\
&= C \; |f|^q_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \qquad (5.14)
\end{aligned}
$$

where we have used in the second line that for j fixed the intervals $I_{j,k}$ are disjoint an that $\tau \le q$ in the next to the last line ($\ell_{\tau/q} \subset \ell_1$).

Hence, by (5.9), (5.13) and (5.14)

$$
\begin{aligned}
\sigma_{F^s_{p,r}(\psi,\mathbb{R}^d)}(f,\Sigma_n) \le |E|_{F^s_{p,r}(\psi,\mathbb{R}^d)} &= \|\tilde{E}\|_{L_p(\mathbb{R}^d)} \\
&= Cn^{-(\frac{1}{q}-\frac{1}{p})} \, \|\mathcal{B}_f\|^{q/p}_{L_q(\mathbb{R}^d)} \le Cn^{-(\frac{1}{q}-\frac{1}{p})} \, |f|^{q/p}_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \\
&= Cn^{-(\frac{1}{q}-\frac{1}{p})},
\end{aligned}
$$

which is the desired result. □

Remark. Taking $r = 2$ in Theorem 5.2 we obtain i) and ii) with $F^s_{p,r}(\Psi, \mathbb{R}^d)$ replaced by the Sobolev space $W^s_p(\Psi, \mathbb{R}^d)$. When $r = 2$, $s = 0$, and $1 < p < \infty$ and Ψ is an orthonormal wavelet system satisfying (3.2), we have $|f|_{F^s_{p,r}(\Psi,\mathbb{R}^d)} \approx \|f\|_{L^p(\mathbb{R}^d)}$. Thus, Theorem 5.2 gives Jackson's inequality when the error of approximation is measured in $L^p(\mathbb{R}^d)$, $1 < p < \infty$. This improves Theorem 2.1 in [DJP].

Proposition 5.3. *Let us consider real numbers $s \ge 0, \beta > 0\,,\; 0 < \tau, q, r < \infty$, and $1 < p < \infty$, such that $q < p$ and $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$. Suppose that $\{\Psi_\lambda : \lambda \in \Lambda\}$ is an orthonormal wavelet system for $L_2(\mathbb{R}^d)$ as in Proposition 3.1. Then, there exists $C > 0$ such that*

$$
|f|_{F^s_{p,r}(\psi,\mathbb{R}^d)} \;\le\; C \; |f|_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)} \qquad \textit{for all } f \in B^{\beta+s}_{\tau,q}(\psi, \mathbb{R}^d)\,. \qquad (5.15)
$$

In particular, there exist $C > 0$ such that

$$
\|f\|_{L_p(\mathbb{R}^d)} \;\le\; C \; |f|_{B^{\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \qquad \textit{for all } f \in B^{\beta}_{\tau,q}(\psi, \mathbb{R}^d)\,. \qquad (5.16)
$$

Proof. This proposition correspond to the case $n = 0$ of Theorem 5.2. The proof of (5.15) is similar to that of Theorem 5.2, except that, assuming $|f|_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)} = 1$, estimate (5.8) is replaced by $|a_I(f)| < |I|^{\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau}}$ for each $I \in \mathcal{D}$. We omit the details. Inequality (5.16) follows by taking $s = 0$ and $r = 2$ in (5.15) and using Proposition 3.1. □

6 Bernstein type inequalities

We want to prove inequalities of Bernstein type for Sobolev spaces $W^s_p(\Psi, \mathbb{R}^d)$, or more generally for Triebel-Lizorkin spaces $F^s_{p,r}(\Psi, \mathbb{R}^d)$. That is, we are interested in proving inequalities of the form

$$|S|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)} \;\leq\; C\, n^{\lambda}\; |S|_{F^{s}_{p,r}(\Psi,\mathbb{R}^d)}\,, \tag{6.1}$$

for all $S \in \Sigma_n,\ n = 1, 2, 3, \ldots$ and some $\lambda > 0$. Here we are assuming that $\{\Psi_\lambda : \lambda \in \Lambda\}$ is an orthonormal wavelet system for $L_2(\mathbb{R}^d)$ as in Proposition 3.1. We would like to find the best λ in (6.1) as well as the relation between the parameters that appear in the above inequality. A homogeneity argument, as in the case of Jackson's inequality, gives the relation among the parameters.

Proposition 6.1. *If Bernstein's inequality (6.1) holds, we must have* $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$.

Proof. We present the proof for $\Psi = \{\psi\}$. Let $D_l f(x) = 2^{ld/2} f(2^l x)$ be the dyadic dilation operator. If Bernstein's inequality (6.1) holds, from (5.5) and (5.4) we deduce

$$2^{ld(\frac{\beta+s}{d}+\frac{1}{2}-\frac{1}{\tau})}\; |S|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \;\leq\; C\, n^{\lambda}\; 2^{ld(\frac{s}{d}+\frac{1}{2}-\frac{1}{p})} |S|_{F^{s}_{p,r}(\psi,\mathbb{R}^d)}\,,$$

for all $S \in \Sigma_n,\ n = 1, 2, 3, \ldots$ (Notice that if $S \in \Sigma_n$, $D_l S \in \Sigma_n$). Thus, for all $S \in \Sigma_n,\ n = 1, 2, 3, \ldots$

$$2^{ld(\frac{\beta}{d}+\frac{1}{p}-\frac{1}{\tau})} \;\leq\; C\, n^{\lambda}\; \frac{|S|_{F^{s}_{p,r}(\psi,\mathbb{R}^d)}}{|S|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)}}\,.$$

Letting $l \to \infty$ we deduce $\frac{\beta}{d} + \frac{1}{p} - \frac{1}{\tau} \leq 0$. Letting $l \to -\infty$ we deduce $\frac{\beta}{d} + \frac{1}{p} - \frac{1}{\tau} \geq 0$. Thus, we have the conclusion of the Proposition. □

Theorem 6.2. *Let us consider real numbers $s \geq 0,\ \beta > 0\ ,\ 0 < \tau, q, r < \infty,\ 1 < p < \infty$ such that $\frac{1}{\tau} = \frac{1}{p} + \frac{\beta}{d}$ (d = dimension). Suppose that $\{\Psi_\lambda : \lambda \in \Lambda\}$ is an orthonormal wavelet system for $L_2(\mathbb{R}^d)$ as in Proposition 3.1.*

i) If $\tau \leq q$, for all $S \in \Sigma_n,\ n = 1, 2, 3, \ldots$ we have

$$|S|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)} \leq\; C\, n^{(\frac{1}{\tau}-\frac{1}{p})}\; |S|_{F^{s}_{p,r}(\Psi,\mathbb{R}^d)} \qquad n = 1, 2, 3, \ldots$$

ii) If $q \leq \tau (< p)$, for all $S \in \Sigma_n,\ n = 1, 2, 3, \ldots$ we have

$$|S|_{B^{s+\beta}_{\tau,q}(\Psi,\mathbb{R}^d)} \leq\; C\, n^{(\frac{1}{q}-\frac{1}{p})}\; |S|_{F^{s}_{p,r}(\Psi,\mathbb{R}^d)} \qquad n = 1, 2, 3, \ldots$$

Proof. We present the proof for $\Psi = \{\psi\}$. The proof of the general case can easily be adapted from this. It is easy to see that i) follows from ii). In fact, ii) with $q = \tau$ gives

$$|S|_{B^{s+\beta}_{\tau,\tau}(\psi,\mathbb{R}^d)} \;\leq\; C\, n^{(\frac{1}{\tau}-\frac{1}{p})}\; |S|_{F^{s}_{p,r}(\psi,\mathbb{R}^d)} \qquad \text{for all } S \in \Sigma_n\ n = 1, 2, 3, \ldots$$

Since $\tau \leq q$ in case i), $\ell_\tau \subset \ell_q$ and, hence, $|S|_{B^{s+\beta}_{\tau,q}(\psi,\mathbb{R}^d)} \;\leq\; |S|_{B^{s+\beta}_{\tau,\tau}(\psi,\mathbb{R}^d)}$.

We need to show ii). Write $S = \sum_\Gamma a_I \psi_I$ with $|\Gamma| \leq n$, and let $\Gamma = \{I_1, I_2, \ldots, I_n\}$, where $|I_j| \leq |I_{j+1}|,\ j = 1, 2, \ldots m-1$. Let J be the set of all $j \in \mathbb{Z}$ such that $\ell(I_i) = 2^j$ for some $i = 1, 2, \ldots, m$. By using $\ell_1 \subset \ell_{\tau/q}$ we can write

$$|S|_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)} = \sum_{j\in J} \left(\sum_{\ell(I)=2^j} |a_I|^\tau\; |I|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})} \right)^{q/\tau}$$

$$\le C\sum_{j\in J}\sum_{\ell(I)=2^j}|a_I|^q\ |I|^{-q(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}$$
$$= C\int_{\mathbb{R}^d}\sum_{I\in\Gamma}|a_I|^q\ |I|^{-q(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}\ \chi_I(x)\,\frac{dx}{|I|}$$
$$\le\ C\int_{\mathbb{R}^d}\left[\sup_{I\in\Gamma}\ |a_I|\ |I|^{-(\frac{s}{d}+\frac{1}{2})}\ \chi_I(x)\right]^q\ \left[\sum_{I\in\Gamma}|I|^{-q(\frac{\beta}{d}-\frac{1}{\tau})-1}\ \chi_I(x)\right]dx\,.$$

Since $\ell_r\subset\ell_\infty$ we deduce

$$|S|^q_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)}\ \le\ C\ \int_{\mathbb{R}^d}(\tilde{S}_r(x))^q\ \left(\sum_{I\in\Gamma}|I|^{\frac{q}{p}-1}\ \chi_I(x)\right)dx\,. \tag{6.2}$$

where

$$\tilde{S}_r(x)=\left(\sum_{I\in\Gamma}|a_I|^r|I|^{-r(\frac{s}{d}-\frac{1}{2})}\ \chi_I(x)\right)^{1/r}$$

and we have used $\frac{1}{\tau}=\frac{1}{p}+\frac{\beta}{d}$. Let $E_i=I_i\setminus\bigcup_{j<i}I_j$, $j=1,2,\dots,m$, so that the sets E_i are disjoint and $\bigcup_{i=1}^m E_i=\bigcup_{i=1}^m I_i$. Then, from (6.2) we obtain

$$|S|^q_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)}\ \le\ C\ \sum_{i=1}^m\int_{E_i}(\tilde{S}_r(x))^q\ \left(\sum_{I\in\Gamma}|I|^{\frac{q}{p}-1}\ \chi_I(x)\right)dx\,. \tag{6.3}$$

If $x\in E_i$,

$$\sum_{I\in\Gamma}|I|^{\frac{q}{p}-1}\ \chi_I(x)\ =\sum_{l=i}^m|I_l|^{\frac{q}{p}-1}\ \chi_{I_l}(x)\ \le\ C\ |I_i|^{\frac{q}{p}-1}\,,$$

where we have used $\frac{q}{p}-1<0$ ($q<p$). From (6.3) we deduce

$$|S|^q_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)}\ \le\ C\ \sum_{i=1}^m\int_{E_i}(\tilde{S}_r(x))^q\ |I_i|^{\frac{q}{p}-1}\,dx\,.$$

Apply Hölder's inequality with exponent $\frac{p}{q}>1$ to obtain

$$|S|^q_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)}\ \le\ C\ \left[\sum_{i=1}^m\left(\int_{E_i}(\tilde{S}_r(x))^p\,dx\right)^{q/p}|E_i|^{1-\frac{q}{p}}\ |I_i|^{\frac{q}{p}-1}\right]$$
$$\le\ C\ \sum_{i=1}^m\left(\int_{E_i}(\tilde{S}_r(x))^p\,dx\right)^{q/p}.$$

We now apply Hölder's inequality to the sum with exponent $\frac{p}{q}>1$ to obtain

$$|S|^q_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)}\ \le\ C\ \left(\sum_{i=1}^m\int_{E_i}(\tilde{S}_r(x))^p\,dx\right)^{q/p}\left(\sum_{i=1}^m 1\right)^{1-\frac{q}{p}}$$
$$\le\ C\ n^{1-\frac{q}{p}}\left(\int_{\mathbb{R}^d}(\tilde{S}_r(x))^p\,dx\right)^{q/p}.$$
$$=\ C\ n^{1-\frac{q}{p}}\ |S|^q_{F^s_{p,r}(\psi,\mathbb{R}^d)}\,,$$

which is the result we wanted to prove. □

Remark. Taking $r=2$ in Theorem 6.2 we obtain i) and ii) with $F^s_{p,r}(\Psi,\mathbb{R}^d)$ replaced by the Sobolev space $W^s_p(\Psi,\mathbb{R}^d)$. When $r=2$, $s=0$, and $1<p<\infty$ and Ψ is an orthonormal wavelet system satisfying (3.2), we have $|S|_{F^s_{p,r}(\psi,\mathbb{R}^d)} \approx \|S\|_{L^p(\mathbb{R}^d)}$. Thus, Theorem 6.2 gives gives Bernstein's inequality when the error of approximation is measuree in $L^p(\mathbb{R}^d)$, $1<p<\infty$. This improves Theorem 3.1 in [DJP].

We now show that the exponents λ we have found in Theorem 6.2 cannot be improved.

Proposition 6.3. *If Bernstein's inequality (6.1) holds, we must have*

$$\lambda \ \geq \ \max\{\frac{1}{q}-\frac{1}{p}, \frac{1}{\tau}-\frac{1}{p}\}$$

Proof. Take $S_n=\sum_{j=1}^n a_{I_{j,k}}\,\psi_{j,k}$ with $I_{j,k}$ disjoint cubes such that $|I_{j,k}|=2^{-jd}$ and $|a_{I_{j,k}}|^\tau\ |I_{j,k}|^{-\tau(\frac{s+\beta}{d}+\frac{1}{2}-\frac{1}{\tau})}=1$. Then

$$|S_n|_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)} = (\sum_{j=1}^n 1\ )^{1/q} = n^{1/q}\,. \tag{6.4}$$

Since the $I_{j,k}$ are disjoint and $\frac{1}{\tau}=\frac{1}{p}+\frac{\beta}{d}$,

$$\begin{aligned}
|S_n|_{F^s_{p,r}(\psi,\mathbb{R}^d)} &= \left[\int_{\mathbb{R}^d}\left(\sum_{j=1}^n |a_{I_{j,k}}|^r\ |I_{j,k}|^{-r(\frac{s}{d}+\frac{1}{2})}\ \chi_{I_{j,k}}(x)\right)^{p/r} dx\right]^{1/p} \\
&= \left[\sum_{j=1}^n \int_{I_{j,k}} \left(|a_{I_{j,k}}|^r\ |I_{j,k}|^{-r(\frac{s}{d}+\frac{1}{2})}\ \right)^{p/r} dx\right]^{1/p} \\
&= \left(\sum_{j=1}^n |a_{I_{j,k}}|^p\ |I_{j,k}|^{-p(\frac{s}{d}+\frac{1}{2})+1}\right)^{1/p} \\
&= \left(\sum_{j=1}^n\ |I_{j,k}|^{p(\frac{\beta+s}{d}+\frac{1}{2}-\frac{1}{\tau})}\ |I_{j,k}|^{-p(\frac{s}{d}+\frac{1}{2})+1}\right)^{1/p} \\
&= \left(\sum_{j=1}^n\ 2^{-jdp(\frac{\beta}{d}+\frac{1}{p}-\frac{1}{\tau})}\right)^{1/p} = (\sum_{j=1}^n\ 1\ )^{1/p}\ = n^{1/p}
\end{aligned}$$

If Bernstein's inequality (6.1) holds, we must have $n^{1/q}\le Cn^\lambda n^{1/p}$. Thus, $n^{\frac{1}{q}-\frac{1}{p}}\le Cn^\lambda$ for all $n=1,2,3,\ldots$, which implies $\frac{1}{q}-\frac{1}{p}\le\lambda$.

Take now $T_n=\sum_{k=1}^n a_{0,k}\psi_{0,k}$, with $I_{0,k}$ disjoint cubes belonging to the same dyadic level such that $|a_{0,k}|=1$. Then, since $|I_{0,k}|=1$,

$$|T_n|_{B^{\beta+s}_{\tau,q}(\psi,\mathbb{R}^d)} = \left(\ \sum_{k=1}^n\ |I_{0,k}|^{-\tau(\frac{\beta+s}{d}+\frac{1}{2}-\frac{1}{\tau})}\right)^{1/\tau}\ =\ n^{1/\tau}\,. \tag{6.5}$$

Also, since the $I_{0,k}$ are disjoint

$$|T_n|_{F^s_{p,r}(\psi,\mathbb{R}^d)} = \left[\int_{\mathbb{R}^d} \left(\sum_{k=1}^{n} |I_{0,k}|^{-r(\frac{s}{d}+\frac{1}{2})} \chi_{I_{0,k}}(x)\right)^{p/r} dx\right]^{1/p}$$
$$= \left(\sum_{k=1}^{n} \int_{I_{0,k}} \chi_{I_{0,k}}(x)\, dx\right)^{1/p} = n^{1/p}$$

If Bernstein's inequality (6.1) holds, we must have $n^{1/\tau} \leq Cn^{\lambda} n^{1/p}$. Thus, $n^{\frac{1}{\tau}-\frac{1}{p}} \leq Cn^{\lambda}$ for all $n = 1, 2, 3, \ldots$, which implies $\frac{1}{\tau} - \frac{1}{p} \leq \lambda$. □

References

[BL] Bergh, J., Löfstrom, J., *Interpolation spaces. An introduction*, Springer- Verlag, (1976).

[Bo] Bourdaud, G., *Ondelettes et espaces de Besov*, Rev. Mat. Iberoamericana, **11**, 477-512, (1995).

[Co] Cohen, A., *Wavelet methods in numerical analysis*, Handbook for Numerical Analysis, Vol. VII, North-Holland, Amsterdam, 417-711, (2002).

[CDF] Cohen, A., Daubechies, I, Feauveau, J. C., *Biorthogonal bases of compactly supported waveletes*, Comm. Pure Appl. Math., **45**, 485-500, (1992).

[DJP] DeVore, R.A., Jawerth, B., Popov, V.A. *Compresion of wavelet decompositions.* Amer. J. of Math, **114**, 737–785, (1992).

[DP] DeVore, B., Popov, V.A., *Interpolation spaces and non-linear approximation*, in Function Spaces and Approximation (M. Cwikel, J. Peetre, Y. Sagher, H. Wallin, Eds.), *Springer Lecture notes in Math.*, **1302**, 191–205, (1998).

[FJ] Frazier, M., Jawerth, B., *A discrete transform and decomposition of distribution spaces*, J. Func. Anal., **93**, 34-170, (1990).

[GHT] Garrigós, G., Hochmuth, R., Tabacco, A., *Wavelet characterizations for anisotropic Besov spaces: cases* $0 < p < 1$, Preprint. Available at *http://www.uam.es/gustavo.garrigos.*

[GR] Gröchenig, K., Ron, A., *Thight compactly supported wavelet frames or arbitrary high smoothness*, Proc. Amer. Math. Soc., **126** (4), 1101-1107, (1998)

[HW] Hernández, E. and Weiss, G., *A First Course on Wavelets*, CRC Press, Boca Raton FL, (1996).

[KL] Kahane, J.P., Lemarié-Rieusset, P.G., *Fourier series and wavelets*, Gordon and Breach Publishers, (1995).

[Ma] Mallat, S., *Multiresolution approximations and wavelet orthonormal bases for* $L^2(\mathbb{R}^d)$, Trans. of Amer. Math. Soc., **315**, 69-87,(1989).

[Me] Meyer, Y., *Ondelettes et opérateurs, I: Ondelettes*, Hermann, Paris, (1990) [English tanslation: *Wavelets and operators*, Cambridge University Press (1992).]

[Si] Sickel, W., *A remark on orhtonormal bases of compactly supported wavelets in Triebel-Lizorkin spaces. The case* $0 < p, q < \infty$, Arch. Math., **57**, 281-289, (1991).

[Tr] Triebel, H., *Theory of Function Spaces*, Birkhäuser, (1983).