AALBORG UNIVERSITY

Nonlinear approximation with general wave packets

by

Lasse Borup and Morten Nielsen

February 2004

R-2004-05

DEPARTMENT OF MATHEMATICAL SCIENCES
AALBORG UNIVERSITY
Fredrik Bajers Vej 7G ▪ DK-9220 Aalborg Øst ▪ Denmark
Phone: +45 96 35 80 80 ▪ Telefax: +45 98 15 81 29
URL: www.math.auc.dk/research/reports/reports.htm

ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

NONLINEAR APPROXIMATION WITH GENERAL WAVE PACKETS

LASSE BORUP AND MORTEN NIELSEN

ABSTRACT. We study nonlinear approximation in the Triebel-Lizorkin spaces with dictionaries formed by dilating and translating one single function g. A general Jackson inequality is derived for best m-term approximation with such dictionaries. In some special cases where g has a special structure, a complete characterization of the approximation spaces is derived.

1. INTRODUCTION

The purpose of this paper is to study the approximation properties of systems with a structure similar to wavelet systems such as wavelet frames and more general systems.

Wavelets and wavelet frames have a common structure, they are generated by dilating and translating a finite number of functions,

$$\mathcal{D} = \{2^{jd/2}\psi^i(2^j \cdot -k)\}_{j,k}, \qquad i = 1, 2, \dots, L,$$

and $\mathcal{D}$ forms a stable system in $L_2(\mathbf{R}^d)$, i.e., $\mathcal{D}$ is a frame for L_2 or better.

The nonlinear approximation properties of wavelet systems have been studied extensively [8, 7], and recently the same properties for wavelet frames have been investigated [4, 12]. It turns out that the results for such systems have the same flavor, the approximation spaces associated with best m-term approximation in L_p are essentially Besov spaces. One might expect this to be a consequence of a combination of the following facts:

(1) The systems have the same structure, i.e., each system is generated by dilating and translating a finite number of functions.
(2) Each generator of the system has a number of vanishing moments.
(3) Each system is stable in L_2, i.e, it forms a frame.

The main point of this paper is to show that for the purpose of nonlinear approximation, items (2) and (3) above are not really important. It is not the fact that the functions form a frame for the classical function spaces, or the

---

Date: February 9, 2004.

2000 *Mathematics Subject Classification.* Primary 41A17; Secondary 41A30, 41A25, 41A46.

Key words and phrases. Wavelet, tight wavelet frame, Besov space, Triebel-Lizorkin space, nonlinear approximation.

This work is in part supported by the Danish Technical Science Foundation, Grant no. 9701481.

vanishing moments of the generator, but rather the general structure of the function system that is important. Consequently, many of the approximation results hold true for a much larger class of generators than just wavelet and tight wavelet frames. Support for this fact can be found in the work of Frazier-Jawerth [10] and Petrushev-Kyriazis [17, 15]. Frazier and Jawerth study expansions of functions with wavelet like systems generated by a smooth function with a prescribed number of vanishing moments, while Petrushev and Kyriazis consider approximation with wavelet systems generated by a smooth function with non-zero integral such as the Gaussian. However, many interesting functions do have a few vanishing moments. For example, one type of function often encountered in applications is one that is very smooth, but with a comparatively small number of vanishing moments. Often generators of wavelet frames will have these characteristics, see *e.g.* [5, 6]. In this paper we extend the results in [15] to cover more general functions, including generators of wavelet frames. The idea is the same as in [15], namely to build nice wave packets from the initial function, but the technique is different; we utilize more directly the powerful machinery build by Frazier and Jawerth.

The structure of the paper is as follows. In Section 2 we give the definition of an approximation space associated with best m-term approximation in a semi-(quasi-)normed abelian group with a general dictionary. We are particularly interested in measuring the approximation error in a homogeneous Triebel-Lizorkin space. We also discuss the Jackson and Bernstein inequalities necessary in order to characterize the approximation space.

In Section 3 we review some of the results on the ϕ-transform given by Frazier and Jawerth. More precisely, we state sufficient conditions for a function g, such that dilations and translations of g give wavelet type expansions of homogeneous Besov- and Triebel-Lizorkin spaces. In Section 4 we show that we can get rid of some of the conditions used by Frazier and Jawerth. In particular, we show that the number of vanishing moments of the generator is not important.

Finally, in the last two sections we study non-linear approximation given by (oversampled) wavelet-type dictionaries $\mathcal{D}_g$ generated by the quite general functions g studied in Section 4. Section 5 is devoted to proving a Jackson inequality for best m-term approximation with $\mathcal{D}_g$ in a homogeneous Triebel-Lizorkin space, while in the final section of the paper, Section 6, we give a Bernstein inequality in L_p for m-term approximation with $\mathcal{D}_g$ when g is associated to an MRA. For such a dictionary we thus have a complete characterization of the approximation spaces associated with best m-term approximation in L_p.

2. NON-LINEAR APPROXIMATION ON ABELIAN GROUPS

In this section we introduce the notation and results from approximation theory needed to state the main results more precisely. We study approximation of smooth functions with m-term approximants formed by dilating and translating a single function $g \in L_2(\mathbf{R}^d)$, where we measure the approximation error in a general Triebel-Lizorkin semi-(quasi-)norm. That is, given a function $g \in L_2(\mathbf{R}^d)$, we consider the following dictionaries of dilations and translations of g:

$$\mathcal{D}_g^R := \{g(2^j \cdot -k/R) | j \in \mathbf{Z}, k \in \mathbf{Z}^d\},$$

and

$$\overline{\mathcal{D}}_g := \{g(a \cdot -b) | a \in \mathbf{R}_+, b \in \mathbf{R}^d\},$$

where $R \geq 1$ is an "oversampling ratio". For a given dictionary, $\mathcal{D} = \mathcal{D}_g^R$, $R \geq 1$, or $\mathcal{D} = \overline{\mathcal{D}}_g$, we consider the associated nonlinear manifold of all possible m-term expansions by elements from $\mathcal{D}$ given by

$$\Sigma_m(\mathcal{D}) = \left\{ S : S = \sum_{j=1}^m a_j g_j, \quad \text{with } a_j \in \mathbf{C}, g_j \in \mathcal{D} \right\}.$$

The error of the best m-term approximation from $\Sigma_m(\mathcal{D})$ of a function f from a semi-(quasi-)normed abelian group $X \subset \mathcal{S}'(\mathbf{R}^d)$, with $\mathcal{D} \subset X$, is given by

$$\sigma_m(f, \mathcal{D})_X := \inf_{S \in \Sigma_m(\mathcal{D})} |f - S|_X.$$

We let $\mathcal{A}_q^\gamma(X, \mathcal{D})$, $\gamma > 0$, $0 < q \leq \infty$, denote the approximation space of all functions f such that

$$|f|_{\mathcal{A}_q^\gamma(X,\mathcal{D})} := \left(\sum_{m=1}^\infty \left(m^\gamma \sigma_m(f, \mathcal{D})_X\right)^q \frac{1}{m} \right)^{1/q} < \infty,$$

with the following standard modification when $q = \infty$:

$$|f|_{\mathcal{A}_\infty^\gamma(X,\mathcal{D})} := \sup_{m \in \mathbf{N}} m^\gamma \sigma_m(f, \mathcal{D}) < \infty.$$

Now the fundamental question is whether it is possible to characterize $\mathcal{A}_q^\gamma(X, \mathcal{D})$ in terms of well known spaces. The answer clearly depends on the properties of g and the semi-(quasi-)norm on X.

It is well known that the main tool in the characterization of $\mathcal{A}_q^\alpha(X, \mathcal{D})$ comes from the link between approximation theory and interpolation theory (see *e.g.* [9, 2]). Let Y be an abelian group with semi-(quasi-)norm $|\cdot|_Y$ continuously embedded in X. Given $\alpha > 0$, the Jackson inequality

$$\sigma_m(f, \mathcal{D})_X \leq C m^{-\alpha} |f|_Y, \qquad \forall f \in Y, \forall m \in \mathbf{N} \tag{2.1}$$

and the Bernstein inequality

$$|S|_Y \le C' m^{\alpha} |S|_X, \qquad \forall S \in \Sigma_m(\mathcal{D}) \tag{2.2}$$

(with some constants C and C' independent of f, S and m) imply, respectively, the continuous embedding

$$(X,Y)_{\beta/\alpha,q} \hookrightarrow \mathcal{A}_q^{\beta}(X,\mathcal{D})$$

and the converse embedding

$$(X,Y)_{\beta/\alpha,q} \hookleftarrow \mathcal{A}_q^{\beta}(X,\mathcal{D})$$

for all $0 < \beta < \alpha$ and $q \in (0,\infty]$. Here $(X,Y)_{\theta,q}$ denotes the interpolation space between X and Y obtained using the real method. We refer the reader to [1] for the definition of the real method of interpolation.

Whenever g is an orthonormal wavelet and $X = L_p$, $1 < p < \infty$, it is known that $\mathcal{A}_q^{\beta}(L_p, \mathcal{D}_g^1)$ is essentially a Besov space [8]. Kyriazis and Petrushev [15] considered the problem of (partially) characterizing $\mathcal{A}_q^{\gamma}(L_p, \overline{\mathcal{D}}_g)$ in [15] for more general functions g. For $r \in \mathbf{N}$, they obtained the Jackson embedding (for the definition of $\dot{B}_{\tau}^{s}(L_\tau)$, see below)

$$\dot{B}_{\tau}^{s}(L_\tau) \hookrightarrow \mathcal{A}_{\infty}^{\gamma}(L_p, \overline{\mathcal{D}}_g),$$

with $0 < s < r$, $0 < p < \infty$, and $1/\tau := s/d + 1/p$, provided that the function g satisfies

- $g \in C^{r+1}(\mathbf{R}^d)$
- $|g^{(\alpha)}(x)| \le C(1+|x|)^{-d-r-\varepsilon}$, for some $\varepsilon > 0$ and $|\alpha| \le r+1$
- $\int_{\mathbf{R}^d} g(x)\,dx = 1$.

For the case $d = 1$ and $H(x) = (1+x^2)^{-N}$ there is a complete characterization

$$\mathcal{A}_q^{\gamma}(L_p, \overline{\mathcal{D}}_H) = \big(L_p, \dot{B}_{\tau}^{s}(L_\tau)\big)_{\gamma/s,q}, \qquad 1/\tau = s + 1/p,$$

where the inverse Bernstein estimate needed to get the characterization can be derived directly from the work of Pekarskii [16].

We will consider approximation in a more general space X, and not only $X = L_p$. Let us briefly recall the definition of the homogeneous Triebel-Lizorkin and Besov spaces (see also, *e.g.* [18]).

Let $\{\phi_j\}_{j\in\mathbf{Z}}$ be a collection of functions in $\mathcal{S}(\mathbf{R}^d)$ with $\operatorname{supp}(\phi_j) \subset \{x \mid 2^j \le |x| \le 2^{j+1}\}$. Then

- For $0 < p \le \infty$, $s \in \mathbf{R}$, and $0 < q \le \infty$, we define the Besov seminorm for $f \in \mathcal{S}'(\mathbf{R}^d)$,

$$|f|_{\dot{B}_q^s(L_p(\mathbf{R}^d))} := \Big(\sum_{j\in\mathbf{Z}} 2^{jsq} \|\mathcal{F}^{-1}\phi_j \mathcal{F} f\|_{L_p(\mathbf{R}^d)}^q \Big)^{1/q},$$

with the appropriate modification when $q = \infty$, and the homogeneous Besov is defined as

$$\dot{B}^s_q(L_p(\mathbf{R}^d)) := \{f : f \in \mathcal{S}'(R^d), |f|_{\dot{B}^s_q(L_p(\mathbf{R}^d))} < \infty\}.$$

- For $0 < p < \infty$, $s \in \mathbf{R}$, and $0 < q \leq \infty$, we define the Triebel-Lizorkin semi-norm for $f \in \mathcal{S}'(\mathbf{R}^d)$,

$$|f|_{\dot{F}^s_q(L_p(\mathbf{R}^d))} := \left\| \left(\sum_{j\in\mathbf{Z}} 2^{jsq} |\mathcal{F}^{-1}\phi_j \mathcal{F} f(\cdot)|^q \right)^{1/q} \right\|_{L_p(\mathbf{R}^d)},$$

with the appropriate modification when $q = \infty$, and the homogeneous Triebel-Lizorkin space is defined as

$$\dot{F}^s_q(L_p(\mathbf{R}^d)) := \{f : f \in \mathcal{S}'(R^d), |f|_{\dot{F}^s_q(L_p(\mathbf{R}^d))} < \infty\}.$$

One can check that the kernel of the semi-(quasi-)norm on $\dot{F}^\beta_t(L_p)$ and $\dot{B}^\beta_t(L_p)$, respectively, is exactly the space of polynomials $\mathcal{P}$ on $\mathbf{R}^d$. It is also well-known that for $1 < p < \infty$, $\dot{F}^0_2(L_p) \approx L_p(\mathbf{R}^d)$ (the spaces are identical modulo $\mathcal{P}$), and for $0 < p \leq 1$, $\dot{F}^0_2(L_p) \approx H_p(\mathbf{R}^d)$. We will therefor measure the error of the best m-term approximation in $X = \dot{F}^\beta_t(L_p)$ since it covers the "classical" case $X = L_p$ as well. The main candidates to characterize $\mathcal{A}^\gamma_q(\dot{F}^\beta_t(L_p), \mathcal{D}^R_g)$ and $\mathcal{A}^\gamma_q(\dot{F}^\beta_t(L_p), \overline{\mathcal{D}}_g)$ are (essentially) homogeneous Besov spaces. However, it is only for functions g with a special structure we can address the problem of getting a characterization of $\mathcal{A}^\gamma_q(\dot{F}^\beta_t(L_p), \mathcal{D}^R_g)$. The main obstacle is the lack of a Bernstein inequality for a general $\mathcal{D}_g$. The conditions on g to ensure the existence of such an inequality is an open and likely very hard problem.

The main focus in the first part of the present paper is to obtain Jackson estimates for $\overline{\mathcal{D}}_g$ for fairly general g that have some smoothness and decay. The main tool is to create functions that can be used in the Frazier-Jawerth theory of ϕ-transforms. In section 3 we introduce the basic results of the ϕ-transform, and we show how to create "wave-packets" from $\overline{\mathcal{D}}_g$ that form decompositions of the Triebel-Lizorkin spaces $\dot{F}^\beta_t(L_p)$. In section 5 we use the decompositions from Section 3 to prove the following Jackson estimate: For $\gamma > \beta$ and $1/\tau := (\gamma - \beta)/d + 1/p$, there exists a finite constant C such that

$$\sigma_m(f, \overline{\mathcal{D}}_g)_{\dot{F}^\beta_t(L_p)} \leq C m^{-(\gamma-\beta)/d} |f|_{\dot{B}^\gamma_\tau(L_\tau)}, \qquad \forall m \in \mathbf{N}, f \in \dot{B}^\gamma_\tau(L_\tau),$$

which is valid for a large class of functions g with some smoothness and decay.

In section 6 we study a more restricted class of functions g on $\mathbf{R}$. For functions g based on a refinable function much more can be said about $\mathcal{A}^\gamma_q(\dot{F}^\beta_t(L_p), \mathcal{D}^R_g)$. In fact, a complete characterization of $\mathcal{A}^\gamma_q(L_p, \mathcal{D}^{2^K}_g)$ is given for fairly general functions g based on a refinable function for large values of K, i.e., for a sufficiently high oversampling ratio 2^K. The characterization is based on a Bernstein inequality for $\mathcal{D}_g^{2^K}$ which we also prove in Section 6. For an even more restricted class of functions g, based on a refinable function, we derive a characterization of $\mathcal{A}_q^\gamma(L_p, \mathcal{D}_g^2)$. This result is also derived in Section 6.

3. The ϕ-transform of Frazier and Jawerth

The main tool we use to obtain a general Jackson inequality in Section 4 is the ϕ-transform of Frazier and Jawerth. In this section we will recall some well known results about the ϕ-transform. Let D denote the set of dyadic cubes $Q = Q_{\nu k} = 2^{-\nu}([0,1]^d + k)$, $\nu \in \mathbf{Z}$, $k \in \mathbf{Z}^d$. We will use two index notations in this paper. For $Q = Q_{\nu k}$ we let $\phi_Q(x) := 2^{-\nu d/2}\phi(2^\nu x - k)$, while $\{\phi^Q\}_{Q\in D}$ just denotes a sequence indexed by the dyadic cubes, but not necessarily given by dilates and translates of a single function.

Let us briefly recall the definition of the discrete Besov and Triebel-Lizorkin spaces.

For $\alpha \in \mathbf{R}$ and $0 < p, q \leq \infty$ we let $\dot{b}_q^\alpha(L_p)$ be the collection of all complex-valued sequences $s = \{s_Q\}_{Q\in D}$ such that

$$\|s\|_{\dot{b}_q^\alpha(L_p)} := \Big(\sum_{\nu\in\mathbf{Z}} 2^{\nu q(\alpha+\frac{d}{2}-\frac{d}{p})}\Big(\sum_{Q\in D:\ |Q|=2^{-\nu d}} |s_Q|^p\Big)^{\frac{q}{p}}\Big)^{\frac{1}{q}} < \infty,$$

with the usual changes when $p = \infty$ or $q = \infty$. Likewise, for $\alpha \in \mathbf{R}$, $0 < p < \infty$ and $0 < q \leq \infty$ we define the space $\dot{f}_q^\alpha(L_p)$ by

$$\|s\|_{\dot{f}_q^\alpha(L_p)} := \Big\|\Big(\sum_{Q\in D} (|Q|^{-\frac{\alpha}{d}-\frac{1}{2}}|s_Q|\chi_Q)^q\Big)^{\frac{1}{q}}\Big\|_{L_p} < \infty,$$

again with the usual change when $q = \infty$.

The following theorem from [11, Theorem 6.16] shows the connection between the sequence spaces $\dot{b}_q^\alpha(L_p)$ and $\dot{f}_q^\alpha(L_p)$ and the classical Besov and Triebel-Lizorkin spaces. For notational convenience we only consider the Triebel-Lizorkin spaces and corresponding sequence space $\dot{f}_q^\alpha(L_p)$, but a similar result holds for the correspondence between Besov spaces and the sequence spaces $\dot{b}_q^\alpha(L_p)$.

Theorem 3.1 ([11]). *Given $\alpha \in \mathbf{R}$, $0 < q \leq \infty$, $0 < p < \infty$. Suppose $\phi \in S(\mathbf{R}^d)$ satisfies*

$$\operatorname{supp}\hat{\phi} \subseteq \{\xi \in \mathbf{R}^d : \tfrac{1}{2} \leq |\xi| \leq 2\}, \tag{3.1}$$

$$|\hat{\phi}(\xi)| \geq c > 0 \qquad \text{if } \tfrac{3}{5} \leq |\xi| \leq \tfrac{5}{3}. \tag{3.2}$$

Then there exists $\tilde{\phi} \in S(\mathbf{R}^d)$ *satisfying the same conditions as* ϕ *such that each* $f \in \dot{F}_q^\alpha(L_p)$ *can be decomposed as*

$$f = \sum_{\nu \in \mathbf{Z}} \sum_{Q \in D_\nu} \langle f, \phi_Q \rangle \tilde{\phi}_Q,$$

with $\|\{\langle f, \phi_Q \rangle\}\|_{\dot{f}_q^\alpha(L_p)} \leq c|f|_{\dot{F}_q^\alpha(L_p)}$ *for some constant* c *independent of* f*. Moreover, for any sequence* $s = \{s_Q\}_{Q \in D}$*, we have*

$$|\sum_{Q \in D} s_Q \tilde{\phi}_Q|_{\dot{F}_q^\alpha(L_p)} \leq c\|s\|_{\dot{f}_q^\alpha(L_p)}.$$

Remark 3.2. The result in Theorem 3.1 holds true for Besov spaces too, even for $p = \infty$.

The result can be generalized to a larger class of atoms by a mapping of the nice functions ϕ and $\tilde{\phi}$. We only need to ensure that the corresponding matrix is almost diagonal.

Definition 3.3. *The infinite matrix* $A = \{a_{QQ'}\}_{Q,Q' \in D}$ *is said to be almost diagonal on* $\dot{f}_q^\alpha(L_p)$ *or* $\dot{b}_q^\alpha(L_p)$ *if there exist constants* $\varepsilon, c > 0$ *such that,*

$$|a_{QQ'}| \leq c 2^{(\nu'-\nu)\alpha} \left(1 + \frac{|2^{-\nu}k - 2^{-\nu'}k'|}{2^{-\min(\nu,\nu')}}\right)^{-J-\varepsilon} \min\left(2^{\frac{(\nu'-\nu)(d+\varepsilon)}{2}}, 2^{(\nu-\nu')(\frac{\varepsilon-d}{2}+J)}\right),$$

With $Q = Q_{\nu k}$ *and* $Q' = Q'_{\nu' k'}$*, where* $J = d/\min(1,p,q)$ *for* $\dot{f}_q^\alpha(L_p)$ *and* $J = d/\min(1,p)$ *for* $\dot{b}_q^\alpha(L_p)$*.*

We denote by $\mathbf{ad}_{p,q}^\alpha$ the family of almost diagonal matrices on $\dot{f}_q^\alpha(L_p)$ or $\dot{b}_q^\alpha(L_p)$. The following result from [11] shows why these matrices are of particular interest.

Theorem 3.4 ([11])**.** *Suppose* $\alpha \in \mathbf{R}$, $0 < p, q \leq \infty$*, and* $A \in \mathbf{ad}_{p,q}^\alpha$*. Then* A *is bounded on* $\dot{b}_q^\alpha(L_p)$ *and, if* $p < \infty$*, on* $\dot{f}_q^\alpha(L_p)$*.*

Definition 3.5. *Let* $\phi \in S(\mathbf{R}^d)$ *be a function satisfying* (3.1) *and* (3.2)*. We say that a collection of functions* $\{v^Q\}_{Q \in D}$ *is an* $\mathbf{ad}_{p,q}^\alpha$*-family if the bi-infinite matrix* $\{\langle v^Q, \phi_{Q'} \rangle\}_{Q,Q' \in D}$ *belongs to* $\mathbf{ad}_{p,q}^\alpha$*.*

It can be verified that the definition is independent of the choice of ϕ. Using the notion of almost diagonal operators we can give a result like in Theorem 3.1 for a larger class of atoms. We recall the following result given by Frazier and Jawerth in [10]. Again, we only state the result for the Triebel-Lizorkin spaces, but the result holds true for Besov spaces too, see Remark 3.7. For $x \in \mathbf{R}$ we let $\lfloor x \rfloor$ denote the integer satisfying $x - 1 < \lfloor x \rfloor \leq x$.

Theorem 3.6 ([10])**.** *Let* $\alpha \geq 0$, $0 < q \leq \infty$, $0 < p < \infty$, $J = d/\min(1,p,q)$*, and* $N = \max(\lfloor J - d - \alpha \rfloor, -1)$*. Suppose* δ *satisfies* $\alpha - \lfloor \alpha \rfloor < \delta \leq 1$*, and* *suppose $M > J$. Let u be a function satisfying the four conditions:*

$$|\hat{u}(\xi)| \geq c > 0 \qquad \text{if } 2^{-1} \leq |\xi| \leq 2, \tag{3.3}$$

$$\int x^\gamma u(x)\,dx = 0 \qquad \text{if } |\gamma| \leq N, \tag{3.4}$$

$$|\partial^\gamma u(x)| \leq c_\gamma (1+|x|)^{-M} \qquad \text{if } |\gamma| \leq \lfloor \alpha+1 \rfloor, \tag{3.5}$$

$$|\partial^\gamma u(x) - \partial^\gamma u(y)| \leq \sup_{|z|\leq|y-x|} \frac{|x-y|^\delta}{(1+|x-z|)^M} \qquad \text{if } |\gamma| = \lfloor \alpha+1 \rfloor. \tag{3.6}$$

Given $K \in \mathbf{Z}$, let $\psi(x) = 2^{Kd}u(2^K x)$. Then there is a $K_0 \leq 0$ with the property that if $K \leq K_0$, there exists a family of functions $\{\tilde{\psi}^Q\}_{Q\in D}$, such that for all $f \in \dot{F}^\alpha_q(L_p)$, we have

$$f = \sum_{Q\in D} \langle f, \tilde{\psi}^Q \rangle \psi_Q,$$

with $\|\{\langle f, \tilde{\psi}^Q\rangle\}\|_{\dot{f}^\alpha_q(L_p)} \leq c|f|_{\dot{F}^\alpha_q(L_p)}$, and for any sequence $s = \{s_Q\}_{Q\in D}$, we have

$$\Big| \sum_{Q\in D} s_Q \psi_Q \Big|_{\dot{F}^\alpha_q(L_p)} \leq c\|s\|_{\dot{f}^\alpha_q(L_p)}.$$

If, in addition, u satisfies

$$\int x^\gamma u(x)\,dx = 0 \qquad \text{if } |\gamma| \leq \lfloor \alpha-1 \rfloor, \tag{3.7}$$

and

$$|u(x) - u(y)| \leq \sup_{|z|\leq|y-x|} \frac{|x-y|^\rho}{(1+|x-z|)^{\max(M,M+d+\alpha-J)}}, \tag{3.8}$$

then $\{\psi_Q\}_{Q\in D}$ is an $\mathbf{ad}^\alpha_{p,q}$-family.

Remark 3.7. If $J = d/\min(1,p)$ the result in Theorem 3.6 holds true for Besov spaces too, even for $p = \infty$.

We notice that (3.5) and (3.6) basically means that $u \in C^{\alpha+1}(\mathbf{R}^d)$ and that the partial derivatives decay like $|x|^{-M}$ as $|x| \to \infty$.

4. General wave packet bases

Theorem 3.6 gives sufficent conditions for a function ψ to generate an atomic decomposition of the Triebel-Lizorkin spaces. However, many interesting functions do not have a lot of vanishing moments or they may fail to satisfy condition (3.3). For example, generators of tight wavelet frames can have few vanishing mements compared to their smoothness. We can apply Theorem 3.6 directly to such functions, but we will only get atomic decompositions valid for a more restricted range af smoothness parameters than one would expect from the smoothness of the generator. Here we extend the results in [15] to cover more general functions, including generators of wavelet frames using the powerful Theorem 3.6.

Proposition 4.1. *Suppose $g \in L_2(\mathbf{R}^d)$ and $\hat{g}$ satisfies a Lipschitz condition of some order $\beta \in (0,1]$. Suppose, furthermore, there exist constants $0 \le a < b < \infty$ such that*

$$\int_a^b \hat{g}(t\xi)\,dt \neq 0 \qquad \text{if } \xi \in A, \tag{4.1}$$

where $A = \{\xi \in \mathbf{R}^d \colon 1/2 \le |\xi| \le 2\}$. Then for any $N \in \mathbf{N}$, there exists a finite set of coefficients $\{a_i, b_i, c_i\} \subset \mathbf{R}_+ \times \mathbf{R}^d \times \mathbf{R}$ such that the function $u = \sum_i c_i g(a_i \cdot -b_i)$ has N vanishing moments and satisfies (3.3).

Proof. Let $h(\xi) := \int_a^b \hat{g}(t\xi)\,dt$ for $\xi \in A$. Clearly, h is a continuous function on A, and thus bounded away from zero on A. Define for any $n \in \mathbf{N}$ the Riemann sums

$$S_n(\xi) = \frac{b-a}{n} \sum_{k=0}^{n-1} \hat{g}(t_k \xi), \qquad t_k = a + \frac{k}{n}(b-a).$$

Then $S_n(\xi) \to h(\xi)$ for each $\xi \in A$ as $n \to \infty$. Moreover, since $\hat{g}$ satisfies a Lipschitz condition of order $\beta \in (0,1]$, we have for any $\xi, \xi' \in A$

$$\begin{aligned} |S_n(\xi') - S_n(\xi)| &\le \frac{b-a}{n} \sum_{k=0}^{n-1} |\hat{g}(t_k\xi') - \hat{g}(t_k\xi)| \\ &\le C\frac{b-a}{n} \sum_{k=0}^{n-1} |t_k\xi' - t_k\xi|^\beta \\ &= C|\xi' - \xi|^\beta \frac{b-a}{n} \sum_{k=0}^{n-1} t_k^\beta \le C|\xi' - \xi|^\beta (b-a)^{1+\beta}. \end{aligned}$$

That is to say, the family of functions $\{S_n\}_{n\in\mathbf{N}}$ is equicontinuous. Now, by the Arzela-Ascoli theorem there exists a subsequence which converges uniformly on A. The limit of the subsequence is the function h, so in particular, there exists an $n_0 \in \mathbf{N}$ such that $S_{n_0}(\xi) \neq 0$ for all $\xi \in A$, since h is bounded away from zero. Hence we have a function

$$\tilde{u} = \mathcal{F}^{-1} S_{n_0} = \frac{b-a}{n_0} \sum_{k=0}^{n_0-1} t_k^{-d} g(x/t_k)$$

given by a finite linear combination of dilates of g, which satisfies (3.3).

Now, in order to obtain $N \in \mathbf{N}$ vanishing moments, we simply apply a high-pass filter which has a zero of order N at $\xi = 0$ and no zeros on A. For example, let

$$u = \Big(\sum_{j=1}^d (\Delta_{e_j} + \Delta_{-e_j})\Big)^N \tilde{u},$$

where Δ_e is the difference operator in the direction $e \in \mathbf{R}^d$, $\Delta_e f(x) = f(x+e) - f(x)$, and e_j is a unit vector in the j-th direction. Then it is easy to see that u satisfies (3.4), and since $\mathcal{F}(\Delta_{e_j} + \Delta_{-e_j})f(\xi) = 2(\cos(\xi \cdot e_j) - 1)\hat{f}$ and $|\sum_{j=1}^d \cos(\xi \cdot e_j) - 1| \geq 1 - \cos(1/2)$ on A, we have that u satisfies (3.3) too. Now, since u is given by a finite linear combination (at most $(2N+1)^d$ terms) of (integer) translates of $\tilde{u}$, the result follows. □

Notice that if $\hat{g}$ is continuous, the condition

$$\int_0^\infty \hat{g}(te)dt \neq 0$$

for all unit vectors $e \in \mathbf{R}^d$, implies (4.1). For a real-valued univariate function this can always be achieved by translating the initial function $g \not\equiv 0$. Thus, we have the following corollary of Proposition 4.1.

Corollary 4.2. *Suppose $g \in L_2(\mathbf{R})$ is a nontrivial real-valued function with $\hat{g}$ satisfying a Lipschitz condition of some order $\beta \in (0,1]$. Then for any $N \in \mathbf{N}$, there exists a finite set of coefficients $\{a_i, b_i, c_i\} \subset \mathbf{R}_+ \times \mathbf{R} \times \mathbf{R}$ such that the function $u = \sum_i c_i g(a_i \cdot - b_i)$ has N vanishing moments and satisfies* (3.3).

Suppose there exists a cone $C \subset \mathbf{R}^d$ with vertex at the origin, such that $\hat{g}(\xi) = 0$ for all $\xi \in C$. Now, since this implies that $e^{ib\cdot\xi}\hat{g}(a\xi) = 0$ for all $\xi \in C$, $b \in \mathbf{R}^d$ and $a \in \mathbf{R}$, functions with support intersecting C cannot be approximated by linear combinations of translated and dilated versions of g. Thus, (4.1) is close to being a necessary condition. One way to get rid of this condition is to allow linear combinations of modulations and/or rotations of g.

In [15], Kyriazis and Petrushev considered atoms with $\int_{\mathbf{R}^d} g(x)\,dx \neq 0$ and decay properties like (3.5). In particular, they obtained expansions of the Besov and Triebel-Lizorkin spaces like in Theorem 3.6 using an (oversampled) dictionary $\mathcal{D}_g^R$. This can also be deduced from the following result.

Proposition 4.3. *Let $g \in L_1(\mathbf{R}^d) \cap L_2(\mathbf{R}^d)$. Suppose there exists a constant $\delta > 0$ such that $|\hat{g}(\xi)| > 0$ for all $0 < |\xi| < \delta$. Then there exists an integer $K_0 \in \mathbf{Z}$ such that for any $K \geq K_0$ and any $N \in \mathbf{N}$ there exists a finite set of coefficients $\{k_i, c_i\} \subset \mathbf{Z}^d \times \mathbf{R}$ such that the function $u = \sum_i c_i g(2^K(\cdot - k_i))$ has N vanishing moments and satisfies* (3.3).

Proof. By the assumptions, there exists $K_0 \in \mathbf{Z}$ such that if $K \geq K_0$ we have $|\hat{g}(2^{-K}\xi)| \geq c > 0$ for $1/2 \leq |\xi| \leq 2$. Now, an application of the high-pass filter defined in the proof of Proposition 4.1 gives the desired function. □

Recall that for a compactly supported function g the Fourier transform $\hat{g}$ is a restriction of an analytic function. Since the zero-set of a (nontrivial) univariate analytic function is discrete, we have the following corollary of Proposition 4.3.

Corollary 4.4. *Suppose g is a (nontrivial) univariate compactly supported function. Then there exists a constant $K_0 \in \mathbf{Z}$ such that for any $K \geq K_0$ and any $N \in \mathbf{N}$ there exists a finite set of coefficients $\{k_i, c_i\} \subset \mathbf{Z} \times \mathbf{R}$ such that the function $u = \sum_i c_i g(2^K(\cdot - k_i))$ has N vanishing moments and satisfies* (3.3).

5. Jackson Inequalities for General Wave packet dictionaries

According to Theorem 3.6 the results in the previous section can be used to obtain expansions of functions in the homogeneous Triebel-Lizorkin or Besov spaces using quite general generators g. In this section we will consider Jackson inequalities for m-term approximation in $\dot{F}_p^\gamma(L_q)$ with either $\overline{\mathcal{D}}_g$ or $\mathcal{D}_g^R$ for some $R \geq 1$, where $g \in \dot{F}_p^\gamma(L_q)$ satisfies the criteria given in the previous section.

Proposition 5.1. *Suppose $0 < p < \infty$, $0 < t \leq \infty$ and $0 \leq \beta < \infty$. Let $g \in \dot{F}_t^\beta(L_p)$ be a function satisfying* (3.5), (3.6) *and* (3.8) *and the condition*

$$\int_a^b \hat{g}(t\xi)\, dt \neq 0 \qquad \text{if } 1/2 \leq |\xi| \leq 2$$

for some constants $0 \leq a < b < \infty$. Suppose $\gamma > \beta$, and define $1/\tau := (\gamma - \beta)/d + 1/p$. Then there exists a finite constant C such that

$$\sigma_m(f, \overline{\mathcal{D}}_g)_{\dot{F}_t^\beta(L_p)} \leq C m^{-(\gamma-\beta)/d} |f|_{\dot{B}_\tau^\gamma(L_\tau)}, \qquad \forall m \in \mathbf{N}, f \in \dot{B}_\tau^\gamma(L_\tau).$$

Proof. By Proposition 4.1, we can construct a system $\{\psi_Q\} \subset \Sigma_K(\overline{\mathcal{D}}_g)$ for some finite K, such that

$$f = \sum_{Q \in D} \langle f, \tilde{\psi}^Q \rangle \psi_Q,$$

for every $f \in \dot{B}_\tau^\gamma(L_\tau)$, with $\|\{\langle f, \tilde{\psi}^Q \rangle\}\|_{\dot{b}_{\tau,\tau}^\gamma} \leq c|f|_{\dot{B}_\tau^\gamma(L_\tau)}$. Fix $f \in \dot{B}_\tau^\gamma(L_\tau)$, and let η be one of the $2^d - 1$ orthonormal Meyer wavelets on $\mathbf{R}^d$. We notice that the function $\tilde{f} := \sum_{Q \in D} \langle f, \tilde{\psi}^Q \rangle \eta_Q$ belongs to $\dot{B}_\tau^\gamma(L_\tau)$ with $|\tilde{f}|_{\dot{B}_\tau^\gamma(L_\tau)} \leq C|f|_{\dot{B}_\tau^\gamma(L_\tau)}$ for some constant C independent of f. From [14], we have

$$\sigma_m(\tilde{f}, \{\eta_Q\})_{\dot{F}_t^\beta(L_p)} \leq C m^{-(\gamma-\beta)/d} |\tilde{f}|_{\dot{B}_\tau^\gamma(L_\tau)}. \tag{5.1}$$

Let $\tilde{f}_m \in \Sigma_m(\{\eta_Q\})$ be a sequence that realizes (5.1) up to the relaxed constant $2C$. We want to map $\tilde{f}_m$ to an element of $\Sigma_m(\{\psi_Q\})$. To accomplish this, we consider the operator T with kernel

$$K(x, y) := \sum_{Q \in D} \psi_Q(x) \eta_Q(y).$$

The matrix of this operator in the Meyer wavelet basis is $M = [\langle \psi_P, \eta_Q \rangle]_{P,Q \in D}$. It is easy to verify that $M \in \mathbf{ad}_{p,t}^\beta$. We notice that $T\tilde{f}_m = f_m \in \Sigma_m(\{\psi_Q\}) \subset$

$\Sigma_{Km}(\overline{\mathcal{D}}_g)$, and moreover, $f - f_m = T(\tilde{f} - \tilde{f}_m)$. However, the matrix representation of T shows that T is bounded on $\dot{F}^{\beta}_{t}(L_p)$ so

$$\begin{aligned}|f - f_m|_{\dot{F}^{\beta}_{t}(L_p)} &\leq C_1|\tilde{f} - \tilde{f}_m|_{\dot{F}^{\beta}_{\tau}(L_\tau)} \\ &\leq 2C_1Cm^{-(\gamma-\beta)/d}|\tilde{f}|_{\dot{B}^{\gamma}_{\tau}(L_\tau)} \\ &\leq C'm^{-(\gamma-\beta)/d}|f|_{\dot{B}^{\gamma}_{\tau}(L_\tau)}.\end{aligned}$$

Hence, we have the wanted Jackson inequality for the dictionary $\{\psi_Q\} \subset \Sigma_K(\overline{\mathcal{D}}_g)$, and then of course there is also a Jackson inequality for the dictionary $\overline{\mathcal{D}}_g$. □

Notice that Proposition 5.1 is a generalization of the Jackson inequality obtained in [15].

Using Corollary 4.4 we have the following stronger result for $d = 1$, based on the same arguments as given in the proof of Proposition 5.1.

Proposition 5.2. *Suppose* $0 < p < \infty$, $0 < t \leq \infty$ *and* $0 \leq \beta < \infty$. *Let* $g \in \dot{F}^{\beta}_{t}(L_p(\mathbf{R})) \cap C^{\beta+1}(\mathbf{R})$ *be a (nontrivial) compactly supported function. Suppose* $\gamma > \beta$, *and define* $1/\tau := (\gamma - \beta) + 1/p$. *Then there exists a constant* $K_0 \in \mathbf{N}_0$ *such that for each* $K \geq K_0$ *there exists a finite constant* C *such that*

$$\sigma_m(f, \mathcal{D}^{2^K}_g)_{\dot{F}^{\beta}_{t}(L_p)} \leq Cm^{-(\gamma-\beta)}|f|_{\dot{B}^{\gamma}_{\tau}(L_\tau)}, \qquad \forall m \in \mathbf{N}, f \in \dot{B}^{\gamma}_{\tau}(L_\tau).$$

6. Refinable functions

So far we have studied Jackson estimates for $\overline{\mathcal{D}}_g$ and $\mathcal{D}^R_g$ for a general smooth function g. For functions g based on a refinable function much more can be said about the approximation properties in L_p of the dictionary $\mathcal{D}^R_g$ for sufficiently large $R \geq 1$.

Recall that ϕ is called a scaling function for a multiresolution analysis (MRA), $\{V_j\}_{j\in\mathbf{Z}}$, of $L_2(\mathbf{R}^d)$, if $\{\phi(x-k)\}_{k\in\mathbf{Z}^d}$ is a Riesz basis for V_0.

It is well known that if $\phi \in L_2(\mathbf{R})$ satisfies

(1) $\{\phi(x-k)\}_{k\in\mathbf{Z}}$ is a Riesz basis in $L_2(\mathbf{R})$
(2) $\phi(x/2) = \sum_k a_k\phi(x-k)$ with convergence in $L_2(\mathbf{R})$
(3) $|\hat{\phi}(\xi)|$ is continuous at 0 and $\hat{\phi}(0) \neq 0$,

then ϕ is a scaling function for the MRA

$$V_j := \operatorname{span}\{\phi(2^j\cdot - k)\}_{k\in\mathbf{Z}}.$$

A classical example of a scaling function is a B-spline function.

Given an MRA $\{V_j\}_{j\in\mathbf{Z}}$. The following proposition states that for some functions $g \in V_1$ it is possible to give a complete characterization of the approximation space $\mathcal{A}^{\gamma}_{q}(L_p(\mathbf{R}), \mathcal{D}^2_g)$ for the twice oversampled dictionary $\mathcal{D}^2_g$.

Proposition 6.1. *Let $\phi \in W^s(L_\infty(\mathbf{R}))$, $s > 0$, be a compactly supported scaling function for an MRA of $L_2(\mathbf{R})$ generated by the low-pass filter $m_0(\xi)$, and let g be defined by*

$$\hat{g}(2\xi) = m(\xi)\hat{\phi}(\xi),$$

with $m(\xi)$ such that $R(\xi) := m(\xi)/\overline{m_0(\xi+\pi)}$ has an extension to a Laurent polynomial with no zeroes on the unit circle.

Then, for each $K \geq 1$,

$$\mathcal{A}_q^\gamma(L_p(\mathbf{R}), \mathcal{D}_g^{2^K}) = \big(L_p(\mathbf{R}), B_\tau^\alpha(L_\tau(\mathbf{R})\big)_{\gamma,q},$$

for $1 < p < \infty$, $0 < \alpha < s$, and $1/\tau = \alpha + 1/p$.

Proof. Let $\hat{\psi}(2\xi) = e^{-i\xi}\overline{m_0(\xi+\pi)}\hat{\phi}(\xi)$ be the standard wavelet associated to ϕ. Notice that $\hat{g}(2\xi) = P(\xi)\hat{\psi}(2\xi)$ with $P(\xi) = e^{i\xi}R(\xi)$. The Fourier coefficients of $P(\xi)$ decay exponentially since $P(\xi)$ has an extension to a Laurent polynomial with no poles on the unit circle. Hence, g has an expansion in $\{\psi(\cdot - k/2)\}_k$ with coefficients with exponential decay. It follows from [3, Corollary 7.8] that for any $K \geq 1$ the Jackson inequality

$$\sigma_m(f, \mathcal{D}_g^{2^K})_{L_p} \leq \sigma_m(f, \mathcal{D}_g^2)_{L_p} \leq Cm^{-\alpha}|f|_{\dot{B}_\tau^\alpha(L_\tau)}, \qquad \forall m \in \mathbf{N}, f \in \dot{B}_\tau^\alpha(L_\tau),$$

is satisfied for $1 < p < \infty$, $\alpha < s$, and $1/\tau := \alpha + 1/p$.

By the result of Jia [13], for each $0 < \alpha < s$ the Bernstein inequality

$$|S|_{\dot{B}_\tau^\alpha(L_\tau(\mathbf{R}^d))} \leq Cm^{\alpha/d}\|S\|_{L_p(\mathbf{R}^d)}, \qquad \forall S \in \Sigma_m(\mathcal{D}_\phi^1),$$

$1/\tau := \alpha/d + 1/p$, $0 < p \leq \infty$, holds true for the system

$$\mathcal{D}_\phi^1 = \{\phi(2^j x - k)\}_{j\in\mathbf{Z}, k\in\mathbf{Z}^d}.$$

By assumption there is a finite mask $\{b_k\}_k$ such that

$$g(x) = \sum_{\ell\in\mathbf{Z}^d} b_\ell \phi(2x - \ell),$$

and since ϕ is compactly supported and refinable there is another finite mask $\{a_k\}$ such that

$$\phi(x) = \sum_{\ell\in\mathbf{Z}^d} a_\ell \phi(2x - \ell).$$

For $j \in \mathbf{Z}$ and $k \in \mathbf{Z}^d$, we have

$$
\begin{aligned}
&g(2^j x - k/2^K)\\
&\quad= \sum_{\ell_1 \in \mathbf{Z}^d} b_{\ell_1} \phi\big(2^{j+1}x - k/2^{K-1} - \ell_1\big)\\
&\quad= \sum_{\ell_1,\ell_2 \in \mathbf{Z}^d} b_{\ell_1} a_{\ell_2} \phi\big(2^{j+2}x - k/2^{K-2} - 2\ell_1 - \ell_2\big)\\
&\quad\ \ \vdots\\
&\quad= \sum_{\ell_1,\ell_2,\dots,\ell_K \in \mathbf{Z}^d} b_{\ell_1} a_{\ell_2} \cdots a_{\ell_K} \phi\big(2^{j+K}x - k - 2^{K-1}\ell_1 - 2^{K-2}\ell_2 - \cdots - \ell_K\big).
\end{aligned}
$$

Hence, $g(2^j x - k/2^K) \in \Sigma_L(\mathcal{D}^1_\phi)$ for some uniform L depending only on K and the length of the finite masks used above. Take any $S \in \Sigma_m(\mathcal{D}^{2^K}_g)$. Then $S \in \Sigma_{Lm}(\mathcal{D}^1_\phi)$, and using the Bernstein inequality for $\mathcal{D}^1_\phi$ we obtain the wanted inequality,

$$
\begin{aligned}
|S|_{\dot{B}^\alpha_\tau(L_\tau(\mathbf{R}^d))} &\le C(Mm)^{\alpha/d}\|S\|_{L_p(\mathbf{R}^d)}\\
&\le \tilde{C} m^{\alpha/d}\|S\|_{L_p(\mathbf{R}^d)}, \qquad \forall S \in \Sigma_m(\mathcal{D}^{2^K}_g).
\end{aligned}
$$

Now it follows from Section 2 that

$$
\mathcal{A}^\gamma_q(L_p(\mathbf{R}), \mathcal{D}^{2^K}_g) = \big(L_p(\mathbf{R}), B^\alpha_\tau(L_\tau(\mathbf{R})\big)_{\gamma,q},
$$

for $1 < p < \infty$, $0 < \alpha < s$, and $1/\tau = \alpha + 1/p$. □

Remark 6.2. Notice that Proposition 6.1 easily generalizes to finite sets of functions $\{g_i\}$, provided that each function, g_i, satisfies the conditions in the proposition. In particular, Proposition 6.1 applies to a wavelet (bi-)frame system $\{\psi^i\}_{i=1}^L$ based on an s-regular MRA with scaling function ϕ. Such a system is generated by a finite number of high-pass filters $m_i(\xi)$, satisfying some extension principle, see [5, 6]. Usually the filters $m_i(\xi)$ have only a single zero at $\xi = 0$. Then the hypothesis of Proposition 6.1 simply reduces to the requirement of the framelet ψ^i given by $\hat{\psi}^i(2\xi) = m_i(\xi)\hat{\phi}(\xi)$ should have at most s vanishing moments which is most often the case.

The following proposition, which concludes the paper, uses the result from Proposition 5.2 to give a complete characterization of the approximations space $\mathcal{A}^\gamma_q(L_p(\mathbf{R}), \mathcal{D}^{2^K}_g)$ for some sufficiently large $K \in \mathbf{N}$ and a more general generator g than in Proposition 6.1.

For a finite set $\Phi \subset L_2(\mathbf{R})$ we let

$$
S(\Phi) := \operatorname{span}\{\phi(\cdot - k)\colon \phi \in \Phi,\ k \in \mathbf{Z}\}.
$$

Proposition 6.3. *Let $\Phi \subset W^s(L_\infty(\mathbf{R})) \cap C^{1+\varepsilon}(\mathbf{R})$, $s, \varepsilon > 0$, be a finite set of compactly supported refinable functions, and suppose $g \in S(\Phi)$, $g \neq 0$. Then,* *there exists a K_0 such that for each $K \geq K_0$,*

$$\mathcal{A}_q^\gamma(L_p(\mathbf{R}), \mathcal{D}_g^{2^K}) = \big(L_p(\mathbf{R}), B_\tau^\alpha(L_\tau(\mathbf{R}))\big)_{\gamma,q},$$

for $1 < p < \infty$, $0 < \alpha < s$, and $1/\tau = \alpha + 1/p$.

Proof. By Proposition 5.2 there exists a constant $K_0 \in \mathbf{N}_0$ such that for any integer $K \geq K_0$

$$\sigma_m(f, \mathcal{D}_g^{2^K})_{L_p} \leq Cm^{-\alpha}|f|_{\dot{B}_\tau^\alpha(L_\tau)}, \qquad \forall m \in \mathbf{N}, f \in \dot{B}_\tau^\alpha(L_\tau)$$

for $\alpha > 0$, $0 < p < \infty$, and $1/\tau = \alpha + 1/p$. A corresponding Bernstein inequality follows by the same arguments given in the proof of Proposition 6.1. $\square$

References

[1] C. Bennett and R. Sharpley. *Interpolation of operators*. Academic Press Inc., Boston, MA, 1988.

[2] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[3] L. Borup, R. Gribonval, and M. Nielsen. Bi-framelet systems with few vanishing moments characterize besov spaces. (Preprint available from: http://www.math.auc.dk/ research/reports/R-2003-18.pdf), 2003.

[4] L. Borup, R. Gribonval, and M. Nielsen. Tight wavelet frames in Lebesgue and Sobolev spaces. *Preprint*, 2003. (Preprint available from: http://www.math.auc.dk/ research/reports/R-2003-05.ps).

[5] C. K. Chui, W. He, and J. Stöckler. Compactly supported tight and sibling frames with maximum vanishing moments. *Appl. Comput. Harmon. Anal.*, 13(3):224–262, 2002.

[6] I. Daubechies, B. Han, A. Ron, and Z. Shen. Framelets: MRA-based constructions of wavelet frames. *Appl. Comput. Harmon. Anal.*, 14(1):1–46, 2003.

[7] R. A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

[8] R. A. DeVore, B. Jawerth, and V. Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.

[9] R. A. DeVore and G. G. Lorentz. *Constructive approximation*. Springer-Verlag, Berlin, 1993.

[10] M. Frazier and B. Jawerth. A discrete transform and decompositions of distribution spaces. *J. Funct. Anal.*, 93(1):34–170, 1990.

[11] M. Frazier, B. Jawerth, and G. Weiss. *Littlewood-Paley theory and the study of function spaces*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1991.

[12] R. Gribonval and M. Nielsen. On approximation with spline generated framelets. *Constr. Approx.*, Published online July 7, 2003. (Available from: http://www.springerlink.com/ link.asp?id=nj0hqqvlc1nytg9g).

[13] R. Q. Jia. A Bernstein-type inequality associated with wavelet decomposition. *Constr. Approx.*, 9(2-3):299–318, 1993.

[14] G. Kyriazis. Non-linear approximation and interpolation spaces. *J. Approx. Theory*, 113(1):110–126, 2001.

[15] G. Kyriazis and P. Petrushev. New bases for Triebel-Lizorkin and Besov spaces. *Trans. Amer. Math. Soc.*, 354(2):749–776 (electronic), 2002.

[16] A. A. Pekarskiĭ. Chebyshev rational approximation in a disk, on a circle and on a segment. *Mat. Sb. (N.S.)*, 133(175)(1):86–102, 144, 1987.

[17] P. Petrushev. Bases consisting of rational functions of uniformly bounded degrees or more general functions. *J. Funct. Anal.*, 174(1):18–75, 2000.

[18] H. Triebel. *The structure of functions*. Birkhäuser Verlag, Basel, 2001.

DEPARTMENT OF MATHEMATICAL SCIENCES, AALBORG UNIVERSITY, FREDRIK BAJERS VEJ 7G, DK-9220 AALBORG EAST, DENMARK

E-mail address: lasse@math.auc.dk and mnielsen@math.auc.dk