APPROXIMATION WITH BRUSHLET SYSTEMS

LASSE BORUP AND MORTEN NIELSEN

Abstract. We consider an orthonormal basis for $L_2(\mathbf{R})$ consisting of functions that are well localized in the spatial domain and have compact support in the frequency domain. The construction is based on smooth local cosine bases and is inspired by F. Meyer and R. Coifman's brushlets, which are of local exponential type in the frequency domain. We show that some of these brushlet-type bases constitute unconditional bases for $L_p(\mathbf{R})$, $1 < p < \infty$, and $B_q^s(L_p(\mathbf{R}))$, $1 < p, q < \infty$, $s > 0$, and that the norm in these spaces can be expressed by the expansion coefficients. In $L_p(\mathbf{R})$ we construct greedy brushlet-type bases and derive Jackson and Bernstein inequalities. One of the appealing properties of local trigonometric bases is that given one basis it is easy to construct a whole library of bases. In this setting, we discuss brushlet packets constructions. Finally, we investigate a natural bivariate extension leading to ridgelet-type bases for $L_2(\mathbf{R}^2)$.

1. Introduction

Local sine and cosine bases for $L_2(\mathbf{R})$ were introduced by R. Coifman and Y. Meyer [3] and has proven to be a useful tool in signal processing. A typical atom from such a basis has the form

$$b_I(x) \cos\left[\pi(n + \tfrac{1}{2})\frac{x - x_I}{|I|}\right] \tag{1.1}$$

with I an interval from any fixed segmentation of the real axis, x_I is the left endpoint of I, and b_I is a smooth bell function with compact support around I. The basis functions thus have perfect localization in time and are well localized in frequency depending on the smoothness properties of b_I. The flexibility in the choice of segmentation is quite overwhelming and it is natural to try to exploit this to adapt the segmentation to the given signal. One tractable way to do this is to combine adjacent intervals from a given segmentation or split an interval from the segmentation. Each of these operations give rise to

Date: June 26, 2001.
Key words and phrases. local trigonometric bases, brushlet bases, unconditional bases, nonlinear approximation.

new orthonormal basis and it can be shown that the families of bases obtained this way generally fit into the setup for the best basis algorithm of Coifman and Wickerhauser [4] thus allowing for fast adaptive expansion of functions according to certain cost functionals.

Due to certain technical reasons the construction by Coifman and Meyer does not work for general one-periodic systems such as the trigonometric system, i.e. the cosine term in (1.1) cannot be replaced by e^{ikx}. This deficiency was later overcome by V. Wickerhauser [22]. Moreover, it was pointed out by Wickerhauser that such local trigonometric bases have fast implementations based on the FFT which makes the bases even more appealing from a practical point of view.

It is clearly possible to map any such local orthonormal basis by the Fourier transform to a new type of orthonormal basis well localized in time and with compact support in the frequency domain. This was first noticed by E. Laeng [11]. This idea was further developed by R. Coifman and F. G. Meyer [12]. They considered bases constructed using the local trigonometric bases of Wickerhauser and called such objects brushlets. Tensor products of such objects combined with the adaptive expansion from the best basis algorithm has turned out to be quite a successful tool for image compression [12]. This is in part due to the fact that tensor product of brushlets have only one peak in frequency compared to say tensor products of wavelet packets with four peaks in the frequency domain. The brushlets thus have a better angular resolution than separable wavelet packets.

In the present paper we consider brushlet bases from the point of view of their approximation properties. So far the brushlet bases have only been considered from a very practical point of view where the $L_2(\mathbf{R})$ theory is quite sufficient, but it is clear that under certain restrictions the brushlets will be well behaved in other classical function spaces. For technical reasons all results in the present paper will be proved for bases of the type considered by E. Laeng which we will call brushlet type bases but it should be noted that they will hold true for the brushlets too.

We obtain sufficient conditions on the brushlets to be unconditional bases for $L_p(\mathbf{R})$, $1 < p < \infty$, and for the Besov spaces $B^s_q(L_p(\mathbf{R}))$, $1 < p, q < \infty$, $s > 0$. Among the bases that satisfy the conditions are a class of wavelet like bases with the twist that the “mother wavelet” has two humps, however the gain is that the construction works for any expanding dilation factor even irrational ones. This positive result should be compared to the fact that it is not possible to obtain unconditional bases of local trigonometric functions for $L_p(\mathbf{R})$, $p \neq 2$, due to a classical result by Orlicz [23].

Jackson and Bernstein inequalities are derived for certain brushlet systems and we prove that under certain conditions the brushlets form so-called greedy bases for $L_p(\mathbf{R})$, which means that near best N-term nonlinear approximation of $L_p(\mathbf{R})$ functions can be obtained by thresholding the expansion coefficients.

Finally we consider an extension of the brushlet bases to $L_2(\mathbf{R}^2)$, with the basis being separable in polar coordinates. Such bases resemble the Ridgelet construction by D. Donoho, see [8], but with a more adaptable frequency localization.

We should also point out that we realized that the mentioned results should be true by reading the interesting paper [21] by L. Villemoes where he proves that for "nice" segmentations of the line the number of different bell functions needed in (1.1) is finite.

2. Local Cosine Bases

Local trigonometric bases were introduced by Coifman and Meyer [3] and studied in more detail in [1, 9]. Local exponential bases were introduced by Wickerhauser in [22]. Here we will just give a short summary of the results about such bases in the special setup that we will need in the subsequent sections. The reader can consult [1] for more information on the bases.

Given a discrete countable set $E \subset \mathbf{R}$, let $\mathcal{P}$ be a countable collection of pairwise disjoint intervals $I = [\alpha_I^l, \alpha_I^r)$ which covers $\mathbf{R}\backslash E$. Assume furthermore that each interval in $\mathcal{P}$ has one adjacent interval in $\mathcal{P}$ on both sides. A brushlet is essentially going to have support on one of the intervals in $\mathcal{P}$ in the frequency domain. Thus the purpose of the set E is to allow brushlets with arbitrary fine frequency localization. We call E the *set of accumulation points* for $\mathcal{P}$.

To each interval $I \in \mathcal{P}$ assign a cutoff radius $\varepsilon_I^l > 0$ at the left endpoint and a cutoff radius $\varepsilon_I^r > 0$ at the right endpoint. Given two adjacent intervals $I, I' \in \mathcal{P}$ with $\alpha_I^r = \alpha_{I'}^l$ (I' is to the right of I), we require that $\varepsilon_I^r = \varepsilon_{I'}^l$ called the *compatibility condition.* The frequency localization of a brushlet will be on an interval of the form $[\alpha_I^l - \varepsilon_I^l, \alpha_I^r + \varepsilon_I^r]$. In order to make the brushlets orthogonal (in $L_2(\mathbf{R})$) we require that only two such intervals overlap i.e.

$$\varepsilon_I^l + \varepsilon_I^r \leq |I|. \tag{2.1}$$

Take a non-negative ramp function $\rho \in C^d(\mathbf{R})$, $d \geq 2$ such that

$$\rho(\xi) = \begin{cases} 0 & \text{for } \xi \leq -1 \\ 1 & \text{for } \xi \geq 1 \end{cases}$$

and with the property that $\rho(\xi)^2+\rho(-\xi)^2=1$ for all $\xi\in\mathbf{R}$. Assign to each interval $I=[\alpha_I^l,\alpha_I^r)\in\mathcal{P}$ a *bell function*

$$b_I(\xi)=\rho\Big(\frac{\xi-\alpha_I^l}{\varepsilon_I^l}\Big)\rho\Big(\frac{\alpha_I^r-\xi}{\varepsilon_I^r}\Big). \tag{2.2}$$

Notice that $\operatorname{supp}(b_I)\subseteq[\alpha_I^l-\varepsilon_I^l,\alpha_I^r+\varepsilon_I^r]$. Thus b_I and $b_{I'}$ only overlap if I and I' are adjacent intervals. Furthermore, if $I,I'\in\mathcal{P}$ are adjacent intervals with $\alpha_I^r=\alpha_{I'}^l$, then $b_I(\xi)^2+b_{I'}(\xi)^2=1$ for $\alpha_I^l+\varepsilon_I^l\le\xi\le\alpha_{I'}^r-\varepsilon_{I'}^r$.

Now the set of local cosine functions

$$\hat{w}_{n,I}(\xi)=\sqrt{\frac{2}{|I|}}b_I(\xi)\cos\Big(\pi\big(n+\tfrac{1}{2}\big)\frac{\xi-\alpha_I^l}{|I|}\Big),\quad I\in\mathcal{P},\quad n\in\mathbf{N}_0, \tag{2.3}$$

form an orthonormal basis for $L_2(\mathbf{R})$ (see [1]). We call the collection $\{w_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ a *brushlet system*. The brushlets also have an explicit representation. Define a set of *modified bell functions* $\{g_I\}_{I\in\mathcal{P}}$ by

$$b_I(\xi)=\hat{g}_I\Big(\frac{\xi-\alpha_I^l}{|I|}\Big). \tag{2.4}$$

Then a direct calculation shows that

$$w_{n,I}(x)=\sqrt{2|I|}e^{i\alpha_I^l x}\big\{g_I\big(|I|x+\pi\big(n+\tfrac{1}{2}\big)\big)+g_I\big(|I|x-\pi\big(n+\tfrac{1}{2}\big)\big)\big\}. \tag{2.5}$$

Remark 2.1. Instead of the cosine term $\cos\big(\pi\big(n+\frac{1}{2}\big)\cdot\big)$ in (2.3) we could have used terms like $\sin\big(\pi\big(n+\frac{1}{2}\big)\cdot\big)$, $\cos\big(\pi n\cdot\big)$ or $\sin\big(\pi n\cdot\big)$. The only requirement is that two brushlets corresponding to adjacent intervals have opposite polarity (see [1]). Thus in particular it is possible to use different local trigonometric functions on parts of $\mathbf{R}$ separated by an element in E.

Remark 2.2. Another possibility is to use $e^{in\cdot}$. If we let $e_{n,I}(\xi)=|I|^{-1/2}e^{in\frac{\xi-\alpha_I^l}{|I|}}$, then the set of functions $\{v_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ given by

$$\begin{aligned}\hat{v}_{n,I}(\xi)&=b_I(\xi)\big\{b_I(\xi)e_{n,I}(\xi)+b_I(2\alpha_I^l-\xi)e_{n,I}(2\alpha_I^l-\xi)\\&\quad-b_I(2\alpha_I^r-\xi)e_{n,I}(2\alpha_I^r-\xi)\big\},\end{aligned}$$

would be the brushlets constructed in [12]. All the results contained in this paper also holds for these functions, but for simplicity we choose to work with a construction based on local cosines.

From (2.5) we see that if the modified bell function g_I is well localized at zero, a brushlet essentially consists of two peaks localized at $\pm\frac{\pi(n+1/2)}{|I|}$. The question is when g_I is well localized. The following example gives a partial answer.

Example 2.3. Given a $\Lambda \geq 1$, let $\mathcal{P}_\Lambda \subseteq \mathcal{P}$ be the collection of intervals I satisfying

$$\Lambda^{-1} \leq \frac{|I|}{|I'|} \leq \Lambda \tag{2.6}$$

for all adjacent intervals $I' \in \mathcal{P}$. Let

$$\varepsilon_I^l = (1+\Lambda)^{-1}|I|,$$

and notice that

$$\varepsilon_I^l + \varepsilon_{I'}^l = (1+\Lambda)^{-1}(|I| + |I'|) \leq |I|,$$

for $I' \in \mathcal{P}_\Lambda$ adjacent to I, so condition (2.1) is satisfied. In this case the modified bell function g_I defined in (2.4) is given by

$$\hat{g}_I(\xi) = \rho\big((1+\Lambda)\xi\big)\rho\bigg(\frac{\varepsilon_I^l(1+\Lambda)}{\varepsilon_I^r}(1-\xi)\bigg). \tag{2.7}$$

Notice that $|\text{supp}(\hat{g}_I)| \leq 2$. Assume $\|\rho^{(n)}\|_{L^\infty} \leq C$ for $n = 0, 1, \ldots, d$. Then

$$\|\hat{g}_I^{(d)}\|_{L^\infty} \leq C^2(1+\Lambda)^d \sum_{n=0}^{d} \binom{d}{n} \Lambda^n, \tag{2.8}$$

and thus $|g_I(x)| \leq C\min\{1, |x|^{-d}\}$, where C depends only on ρ and Λ.

Remark 2.4. A related situation as in Example 2.3 was investigated by L. Villemoes in [21]. He considered what he calls a good partition - a dyadic partition satisfying (2.6) with $\Lambda = 2$. With this restriction he reduced the set of different modified bell functions to three.

Given a bell function b_I we define an operator P_I by

$$\begin{aligned} \widehat{P_I f}(\xi) &= b_I(\xi)\big\{b_I(\xi)\hat{f}(\xi) + b_I(2\alpha_I^l - \xi)\hat{f}(2\alpha_I^l - \xi) \\ &\quad - b_I(2\alpha_I^r - \xi)\hat{f}(2\alpha_I^r - \xi)\big\}. \end{aligned} \tag{2.9}$$

It can be verified that P_I is an orthonormal projection, mapping $L_2(\mathbf{R})$ onto $\overline{\text{span}}\{w_{n,I} : n \geq 0\}$. Notice that for two adjacent intervals $I, I' \in \mathcal{P}$ with $\alpha_I^r = \alpha_{I'}^l$, we have $\widehat{P_I f}(\xi) + \widehat{P_{I'} f}(\xi) = \hat{f}(\xi)$ a.e. for $\alpha_I^l + \varepsilon_I^l \leq \xi \leq \alpha_{I'}^r - \varepsilon_{I'}^r$.

We need to know when P_I is a bounded operator in $L_p(\mathbf{R})$, $1 < p < \infty$. This is naturally connected to the same question for the bell functions as Fourier multipliers. Consider the operator S_I, $I \in \mathcal{P}$, given by

$$\widehat{S_I f}(\xi) = b_I(\xi)\hat{f}(\xi).$$

Then the Hörmander-Mihlin multiplier theorem implies that $\|S_I f\|_{L_p} \leq CA\|f\|_{L_p}$, if $|(\xi-\alpha)\frac{d}{d\xi}b_I(\xi)| \leq A$ for some $\alpha \in \mathbf{R}$ and $A > 0$, where C only depends on p (see [2]). Since

$$P_I f = S_I\{S_I f + e^{i2\alpha_I^l \cdot} S_I f(-\cdot) + e^{i2\alpha_I^r \cdot} S_I f(-\cdot)\},$$

we have the following result.

Lemma 2.5. *Suppose there exist an $\alpha \in \mathbf{R}$ and an absolute constant $A > 0$ such that $|(\xi-\alpha)\frac{d}{d\xi}b_I(\xi)| \leq A$. Then $\|P_I f\|_{L_p} \leq C_p A^2 \|f\|_{L_p}$, $1 < p < \infty$, where C_p depends only on p.*

Remark 2.6. Notice that $|\xi - \alpha_I^l| \leq 2|I|$ for $\xi \in \operatorname{supp}(b_I)$. Thus if we have the situation as in Example 2.3, then $|(\xi-\alpha_I^l)\frac{d}{d\xi}b_I(\xi)| \leq C$ using (2.4) and (2.8).

3. Brushlet bases in L^p

The way we have defined the brushlet bases, they share many properties with wavelets. In this section we consider the brushlets as basis functions in $L_p(\mathbf{R})$. It turns out that we need a slightly stronger condition on the bell functions than in Lemma 2.5. More precisely we need to impose that the bell functions satisfies $|\frac{d}{d\xi}b_I(\xi)| \leq C(\operatorname{dist}(\xi, E))^{-1}$, $\xi \in \mathbf{R}$, where E is the set of accumulation points. One can ask if this restriction is possible in our setting. The following example shows that the condition is satisfied if any pair of adjacent intervals in the partition $\mathcal{P}$ have comparable length.

Example 3.1. Fix two adjacent numbers $\eta_1 < \eta_2$ from the set of accumulation points E, and let $\{\alpha_k\}_{k\in\mathbf{Z}}$ be the sequence of numbers such that the collection $\{I \in \mathcal{P} : I \subset (\eta_1, \eta_2)\}$ is given by $\{I_k\}_{k\in\mathbf{Z}}$, $I_k = [\alpha_k, \alpha_{k+1})$.

Suppose the length of the intervals decays exponentially as they get closer to η_2, more precisely suppose

$$1 < \lambda \leq \frac{|I_k|}{|I_{k+1}|} \leq \Lambda < \infty$$

for all $k \geq 0$. With this assumption we have

$$|\eta_2 - \alpha_k| = \sum_{j=k}^{\infty} |I_j| \leq |I_k| \sum_{j=0}^{\infty} \lambda^{-j} = \frac{\lambda}{\lambda - 1}|I_k|, \quad k \in \mathbf{N}_0.$$

Take $\xi \in \operatorname{supp}(b_{I_k})$. Then $|\eta_2 - \xi| < |\eta_2 - \alpha_k| + |I_k| \leq \frac{2\lambda-1}{\lambda-1}|I_k|$. Thus, if

$$\varepsilon_{I_k}^l = (1+\Lambda)^{-1}|I_k|, \quad k \in \mathbf{N}_0,$$

we have from Example 2.3 that

$$|\tfrac{d}{d\xi}b_{I_k}(\xi)| \leq C|I_k|^{-1} < C\frac{2\lambda-1}{\lambda-1}|\eta_2-\xi|^{-1}, \quad k \in \mathbf{N}_0,\, \xi \in \mathbf{R},$$

with C depending only on ρ and Λ.

The same estimates holds true if the intervals decays exponentially as they get closer to η_1. If $\eta_0 := \max\{|\eta| : \eta \in E\} < \infty$, similar assumptions gives $|\frac{d}{d\xi}b_I(\xi)| \leq C|\xi-\eta_0|^{-1}$ for $I \subset (\eta_0, \infty)$.

We can now state the first result on brushlet bases in $L_p(\mathbf{R})$.

Proposition 3.2. *Let $\{b_I\}_{I\in\mathcal{P}}$ be a collection of bell functions of the type (2.2). Suppose the set of accumulation points E is finite and there exist an absolute constant C such that $|\frac{d}{d\xi}b_I(\xi)| \leq C(dist(\xi,E))^{-1}$, for all $I \in \mathcal{P}$, $\xi \in \mathbf{R}$. Then for $1 < p < \infty$,*

$$\|f\|_{L_p} \asymp \Big\| \Big(\sum_{I\in\mathcal{P}} |P_I f|^2\Big)^{1/2} \Big\|_{L_p}. \tag{3.1}$$

Proof. Fix an $\eta \in E$ and let $\mathcal{P}_\eta$ be the subset of $\mathcal{P}$ such that $|\frac{d}{d\xi}b_I(\xi)| \leq C|\xi-\eta|^{-1}$ for all $I \in \mathcal{P}_\eta$, $\xi \in \mathbf{R}$. We will first show one of the inequalities in (3.1) with $\mathcal{P}_\eta$ instead of $\mathcal{P}$ and then use that E is finite.

Define the function $m : \mathbf{R} \to \mathcal{L}(\ell_2,\ell_2)$, $\ell_2 = \ell_2(\mathcal{P}_\eta)$, by $\xi \to \{b_I(\xi)a_I\}_I$, $\{a_I\}_{I\in\mathcal{P}_\eta} \in \ell_2(\mathcal{P}_\eta)$. Clearly $\|m\|_{\mathcal{L}(\ell_2,\ell_2)} \leq \sup_I \|b_I\|_{L^\infty} \leq 1$. Since

$$\sum_{I\in\mathcal{P}_\eta} |\tfrac{d}{d\xi}b_I(\xi)a_I|^2$$

contains at most two terms for a fixed $\xi \in \mathbf{R}$, the derivative $\frac{d}{d\xi}m$ is given by $\xi \to \{\frac{d}{d\xi}b_I(\xi)a_I\}_{I\in\mathcal{P}_\eta}$, and we have the bound $\|\frac{d}{d\xi}m\|_{\mathcal{L}(\ell_2,\ell_2)} \leq \sup_{I\in\mathcal{P}_\eta} |\frac{d}{d\xi}b_I(\xi)| \leq C|\xi-\eta|^{-1}$. Define the operator T_m by

$$\widehat{T_m f}(\xi) = \{b_I(\xi)\hat{f}_I(\xi)\}_{I\in\mathcal{P}_\eta} \quad \text{for all } f \in L_2(\mathbf{R},\ell_2).$$

Then the vector valued Hörmander-Mihlin multiplier theorem imply [2, 16]

$$\left\{ \int_{-\infty}^{\infty} \Big(\sum_{I\in\mathcal{P}_\eta} |(T_m f)_I(x)|^2\Big)^{p/2} dx \right\}^{1/p} \leq C_p \|f\|_{L_p(\mathbf{R},\ell_2)}. \tag{3.2}$$

Now define the function $\tilde{m} : \mathbf{R} \to \mathcal{L}(\mathbf{C},\ell_2)$ by $\xi \to \{b_I(\xi)a\}_{I\in\mathcal{P}_\eta}$, $a \in \mathbf{C}$. Notice that

$$\|\tilde{m}\|_{\mathcal{L}(\mathbf{C},\ell_2)} = \Big[\sum_{I\in\mathcal{P}_\eta} |b_I(\xi)|^2\Big]^{1/2} \leq 2 \quad \text{and} \quad \|\tfrac{d}{d\xi}\tilde{m}\|_{\mathcal{L}(\mathbf{C},\ell_2)} \leq 2C|\xi-\eta|^{-1},$$

since only two bell functions overlap at a given ξ. Thus using the Hörmander-Mihlin theorem once more we have $\|T_{\tilde{m}}f\|_{L_p(\mathbf{R},\ell_2)} \leq C'_p\|f\|_{L_p(\mathbf{R})}$, where $T_{\tilde{m}}$ is the operator given by

$$\widehat{T_{\tilde{m}}f}(\xi) = \{b_I(\xi)\hat{f}(\xi)\}_{I\in\mathcal{P}_\eta} \quad \text{for all } f \in L_2(\mathbf{R}).$$

Given $f \in L_2(\mathbf{R})\cap L_p(\mathbf{R})$ we define three functions $f^i \in L_2(\mathbf{R}, \ell_2(\mathcal{P}_\eta))$, $i = 1, 2, 3$, by $f^1_I = (T_{\tilde{m}}f)_I$, $f^2_I(x) = e^{i2\alpha^l_I x}f^1_I(-x)$, and $f^3_I(x) = e^{i2\alpha^r_I x}$ $f^1_I(-x)$, $I \in \mathcal{P}_\eta$. Notice that

$$\widehat{f^2_I}(\xi) = b_I(2\alpha^l_I - \xi)\hat{f}(2\alpha^l_I - \xi), \qquad \text{and} \qquad \widehat{f^3_I}(\xi) = b_I(2\alpha^r_I - \xi)\hat{f}(2\alpha^r_I - \xi).$$

From (3.2) we have

$$\int_{-\infty}^{\infty}\Big(\sum_{I\in\mathcal{P}_\eta}|(T_m f^i)_I(x)|^2\Big)^{p/2}dx \leq C^p_p\|T_{\tilde{m}}f\|^p_{L_p(\mathbf{R},\ell_2)} \leq (C_pC'_p)^p\|f\|^p_{L_p},$$

$i = 1, 2, 3$. However,

$$P_If = T_m(f^1 + f^2 + f^3), \quad \text{for } I \in \mathcal{P}_\eta,$$

so

$$\Big\|\Big(\sum_{I\in\mathcal{P}_\eta}|P_If|^2\Big)^{1/2}\Big\|_{L_p} \leq 3C_pC'_p\|f\|_{L_p}.$$

Finally, since E is a finite set and $\ell_2 \hookrightarrow \ell_1$ we have

$$\begin{aligned}
&\Big\|\Big(\sum_{I\in\mathcal{P}}|P_If|^2\Big)^{1/2}\Big\|_{L_p} \\
&\qquad \leq \Big\{\int_{-\infty}^{\infty}\Big(\sum_{\eta\in E}\Big(\Big\{\sum_{I\in\mathcal{P}_\eta}|P_If(x)|^2\Big\}^{1/2}\Big)^2\Big)^{p/2}dx\Big\}^{1/p} \\
&\qquad \leq \Big\{\int_{-\infty}^{\infty}\Big(\sum_{\eta\in E}\Big\{\sum_{I\in\mathcal{P}_\eta}|P_If(x)|^2\Big\}^{1/2}\Big)^p dx\Big\}^{1/p} \\
&\qquad \leq (\#E)3C_pC'_p\|f\|_{L_p},
\end{aligned}$$

using the Minkowski inequality in the last step.

We now turn to the converse inequality. For $f \in L_2(\mathbf{R})$ we have $\|f\|^2_{L_2} = \sum_{I\in\mathcal{P}}\|P_If\|^2_{L_2}$ so by polarization, for $g \in L_2(\mathbf{R}) \cap L_{p'}(\mathbf{R})$

with $1/p + 1/p' = 1$,

$$\begin{aligned}
|\langle f, g\rangle| &= \left| \sum_{I\in\mathcal{P}} \int_{-\infty}^{\infty} (P_I f)(x)\overline{(P_I g)(x)}\, dx \right| \\
&\le \int_{-\infty}^{\infty} \Big(\sum_{I\in\mathcal{P}} |(P_I f)(x)|^2 \Big)^{1/2} \Big(\sum_{I\in\mathcal{P}} |(P_I g)(x)|^2 \Big)^{1/2} dx \\
&\le \Big\| \Big(\sum_{I\in\mathcal{P}} |P_I f|^2 \Big)^{1/2} \Big\|_{L_p} \Big\| \Big(\sum_{I\in\mathcal{P}} |P_I g|^2 \Big)^{1/2} \Big\|_{L^q}.
\end{aligned}$$

Taking the supremum of this inequality with the restriction $\|g\|_{L^q} \le 1$ gives

$$c_p \|f\|_{L_p} \le \Big\| \Big(\sum_{I\in\mathcal{P}} |P_I f|^2 \Big)^{1/2} \Big\|_{L_p}$$

and the result follows. □

Remark 3.3. The result in Proposition 3.2 is essentially an orthogonal version of the Littlewood-Paley decomposition with smooth multipliers.

Using Proposition 3.2 we can show even more. By introducing a square function based on the brushlet coefficients the $L_p(\mathbf{R})$ norm of a function can be calculated from the size of these coefficients.

Proposition 3.4. *Let $\{w_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ be a brushlet system with associated partition $\mathcal{P}$ and bell functions satisfying the conditions in Proposition 3.2. Moreover, suppose there is an absolute constant $C > 0$ such that the set of modified bell functions $\{g_I\}_{I\in\mathcal{P}}$ given by (2.4) satisfies*

$$|g_I(x)| \le C(1 + x^2)^{-1}, \quad x \in \mathbf{R}. \tag{3.3}$$

Then $\{w_{n,I}\}_{I\in\mathcal{P},n>0}$ form an unconditional basis for $L_p(\mathbf{R})$, $1 < p < \infty$, and we have the characterization

$$\|f\|_{L_p} \asymp \Big\| \Big(\sum_{I\in\mathcal{P},n\in\mathbf{N}_0} |\langle f, w_{n,I}\rangle|^2 |I| \chi_{E_{(n,I)}} \Big)^{1/2} \Big\|_{L_p}, \tag{3.4}$$

where $E_{(n,I)} := \big\{x \in \mathbf{R} : |I|x - \pi\big(n + \frac{1}{2}\big) \in (-1, 1)\big\}$.

Remark 3.5. Notice that sufficient conditions such that (3.3) is true is given in Example 2.3.

Proof. We first notice that $E_{(n,I)}$, $n \in \mathbf{N}_0$, are disjoint intervals for a fixed $I \in \mathcal{P}$, so given $x \in \mathbf{R}$ there is at most one $m \in \mathbf{N}_0$ such that

$x \in E_{(m,I)}$. Using (2.5), we have

$$\begin{aligned}|I|^{1/2}\chi_{E_{(m,I)}}(x)|\langle f, w_{m,I}\rangle| &= |I|^{1/2}\chi_{E_{(m,I)}}(x)|\langle P_I f, w_{m,I}\rangle| \\ &\leq 2^{1/2}\chi_{E_{(m,I)}}(x)\int_{-\infty}^{\infty}|P_I f(y)||I|\big|g_I\big(|I|y-\pi(m+\tfrac{1}{2})\big)\big|\,dy \\ &\quad +2^{1/2}\chi_{E_{m,I}}(x)\int_{-\infty}^{\infty}|P_I f(y)||I|\big|g_I\big(|I|y+\pi(m+\tfrac{1}{2})\big)\big|\,dy \\ &\leq C\big[\mathcal{M}(P_I f)(x)+\mathcal{M}(\widetilde{P_I f})(-x)\big],\end{aligned}$$

where $\tilde{g}(x) = g(-x)$, $\mathcal{M}$ is the Hardy-Littlewood maximal operator and we used [17, p. 57] in the last step. By the Fefferman-Stein maximal inequality we have

$$\Big\|\Big(\sum_{I\in\mathcal{P}}|\mathcal{M}(P_I f)|^2\Big)^{1/2}\Big\|_{L_p} \leq C_p\Big\|\Big(\sum_{I\in\mathcal{P}}|P_I f|^2\Big)^{1/2}\Big\|_{L_p} \leq C_p\|f\|_{L_p},$$

and since $\mathcal{M}(\tilde{g})(-x) = \mathcal{M}(g)(x)$ we conclude,

$$\Big\|\Big(\sum_{I\in\mathcal{P},n\in\mathbf{N}_0}|\langle f, w_{n,I}\rangle|^2|I|\chi_{E_{(n,I)}}\Big)^{1/2}\Big\|_{L_p} \leq c_p\|f\|_{L_p}.$$

To get the converse inequality, we consider the linear operator W : $L_2(\mathbf{R}) \to \ell_2(\mathbf{N}_0 \times \mathcal{P})$ defined by

$$Wf = \big\{\langle f, w_{n,I}\rangle|I|^{1/2}\chi_{E_{(n,I)}}\big\}_{I\in\mathcal{P},n\in\mathbf{N}_0}.$$

By a direct calculation using Parsevals relation, we see that for $f, g \in L_2(\mathbf{R})$,

$$\int_{-\infty}^{\infty}\langle Wf, Wg\rangle_{\ell_2}(x)\,dx = 2\langle f, g\rangle.$$

Thus, for $g \in L_2(\mathbf{R}) \cap L_{p'}(\mathbf{R})$,

$$\begin{aligned}2|\langle f,g\rangle| &= \Big|\int_{-\infty}^{\infty}\langle Wf, Wg\rangle_{\ell_2}(x)\,dx\Big| \\ &\leq \big\|\sqrt{\langle Wf, Wf\rangle_{\ell_2}}\big\|_{L_p}\big\|\sqrt{\langle Wg, Wg\rangle_{\ell_2}}\big\|_{L^{p'}} \\ &\leq C_q\big\|\sqrt{\langle Wf, Wf\rangle_{\ell_2}}\big\|_{L_p}\|g\|_{L^{p'}}. \qquad (3.5)\end{aligned}$$

Taking the supremum of (3.5) for $\{g : \|g\|_{L_{p'}} \leq 1\}$ yields the desired inequality.

Now we can complete the proof and show that the brushlet system is an unconditional basis for $L_p(\mathbf{R})$. All that remains is to verify that the system has dense span in $L_p(\mathbf{R})$. Suppose $g \in L_{p'}(\mathbf{R})$ is such that $\langle g, w_{n,I}\rangle = 0$ for all I, n. It follows from the characterization of the

$L_{p'}(\mathbf{R})$-norm by (3.4) that $g = 0$ and using the Hahn-Banach theorem we conclude that the span of the brushlets is indeed dense in $L_p(\mathbf{R})$. $\square$

From Proposition 3.2 and 3.4 we notice a clear similarity with wavelet expansions. The main difference is that the brushlets allow a more flexible decomposition of the Fourier domain. Notice also that the above results easily extends to $L_p(\mathbf{R}^d)$, $d > 1$, by constructing separable brushlet bases. We leave the easy proof, based on Fubini's theorem, to the reader.

Corollary 3.6. *Let* $\{w^i_{n,I}\}_{I\in\mathcal{P}_\rangle, n\in\mathbf{N}_0}$, $i = 1, \ldots, d$, *be a set of brushlet systems each satisfying the conditions in Proposition 3.4. Fix* $d \geq 1$ *and define* $w_{n,Q}(x) := w^1_{n_1,I_1}(x_1)\otimes\cdots\otimes w^d_{n_d,I_d}(x_d)$, $Q = I_1\times\cdots\times I_d$, $n = (n_1, \ldots, n_d)$. *Then* $\{w_{n,Q}\}_{n,Q}$ *form an unconditional basis for* $L_p(\mathbf{R}^d)$, $1 < p < \infty$.

3.1. **Nonlinear approximation with brushlets.** In this section we will consider the question of approximating a function in $L_p(\mathbf{R})$ by a finite brushlet expansion. For this purpose we need some basic terminology from approximation theory.

Let $\Psi := \{\psi_k\}_{k\in\mathbf{N}}$ be a normalized basis of a Banach space X. Given a function $f \in X$ with associated expansion coefficients $\{c_k\}_{k\in\mathbf{N}}$, we consider approximating f using N elements from Ψ. The task is to minimize the difference between f and the approximation in X. Denote the lower bound by $\sigma_N(f, \Psi, X)$, i.e.

$$\sigma_N(f, \Psi, X) := \inf_{\Gamma\subset\mathbf{N}, \#\Gamma=N} \Big\| f - \sum_{j\in\Gamma} c_j\psi_j \Big\|_X. \tag{3.6}$$

Depending on the behavior of this bound for different values of N, the functions in X are divided into different approximation spaces. We define the approximation spaces $\mathcal{A}^s_q(\Psi, X)$, $0 < q < \infty$, $s \geq 0$, as all $f \in X$ with

$$\|f\|_{\mathcal{A}^s_q(\Psi,X)} := \Big(\sum_{k=1}^{\infty} \big(k^s \sigma_{k-1}(f, \Psi, X)\big)^q \frac{1}{k} \Big)^{1/q} < \infty,$$

where $\sigma_0(f, \Psi, X) := \|f\|_X$. Assume there exist a subspace $Y \hookrightarrow X$ and constants $C, C' > 0$ such that the following Jackson and Bernstein inequalities holds for some $r > 0$:

$$\text{(Jackson)} \qquad \sigma_N(f, \Psi, X) \leq C\|f\|_Y N^{-r}, \quad \text{for all } f \in Y$$

and
(Bernstein)

$$\Big\|\sum_{j\in\Gamma} c_j\psi_j\Big\|_Y \le C'\Big\|\sum_{j\in\Gamma} c_j\psi_j\Big\|_X N^r, \quad \text{for all } \Gamma\subset\mathbf{N} \text{ with } \#\Gamma = N.$$

Then it is well known, that the approximation spaces $\mathcal{A}^s_q(\Psi, X)$, $0 < s < r$, $0 < q < \infty$, are given by the interpolation spaces (see [7, Chapter 7]),

$$\mathcal{A}^s_q(\Psi, X) = (X, Y)_{s/r,q}. \tag{3.7}$$

In section 4 we will derive Jackson and Bernstein inequalities for brushlet approximation in case $X = L_p(\mathbf{R})$ and $Y = B^s_q(L_p(\mathbf{R}))$ for some values of s, p and q.

For certain bases the limit in (3.6) can be reached (up to an absolute constant independent of f and N) simply by picking the N largest coefficients in the expansion. If a basis satisfy this property it is called a *greedy basis*. More precisely, define for each $f \in X$ with expansion coefficients $\{c_j\}_{j\in\mathbf{N}}$, and $N \in \mathbf{N}$, the function

$$G_N(f) := \sum_{j\in A} c_j\psi_j,$$

where $A \subset \mathbf{N}$ is a set of cardinality N such that $|c_j| \ge |c_k|$ for all $j \in A$ and $k \notin A$ (if A is not unique take any such set).

Definition 3.7. A basis Ψ is called greedy if there exists a constant C independent of f and N such that $\|f - G_N(f)\| \le C\sigma_N(f, \Psi, X)$ for all $f \in X$.

The reader can consult [10] for more details on greedy bases. From a practical point of view greedy bases are very desirable since thresholding is a much simpler operation than trying to minimize (3.6) directly. Since it is difficult to verify if a basis satisfy Definition 3.7, we can use another characterization of greedy bases given by S.V. Konyagin and V.N. Temlyakov in [10]. We need to define a democratic basis.

Definition 3.8. $\{\psi_k\}_{k\in\mathbf{N}}$ is called *democratic* if there exists a constant $C > 0$ such that

$$\Big\|\sum_{k\in P}\psi_k\Big\|_X \le C\Big\|\sum_{k\in Q}\psi_k\Big\|_X,$$

for any two finite set of indices P and Q with the same cardinality, $\#P = \#Q$.

Konyagin and Temlyakov proved that a basis is greedy if and only if it is democratic and unconditional [10, Theorem 1]. From Proposition 3.4 we know that certain brushlet bases are unconditional bases in $L_p(\mathbf{R})$, $1 < p < \infty$. We want to show that they are democratic too.

This is equivalent to proving the Temlyakov-type inequalities given in Lemma 3.10 and Lemma 3.13 below. First we need to define a special class of partitions of $\mathbf{R}$.

Definition 3.9. Assume the set of accumulation points E is finite, $E = \{\eta_j\}_{j=1}^J$, and let $\mathcal{P}$ be a partition of $\mathbf{R}\backslash E$. Fix a $\lambda > 1$. Assume there exist an associated set $E' = \{\eta'_j\}_{j=0}^J$, with $\eta'_0 = -\infty$, $\eta'_J = \infty$, and $\eta'_{j-1} < \eta_j < \eta'_j$, $j = 1, 2, \dots, J$, such that $\lambda \leq \frac{|I|}{|I'|}$, for all adjacent $I, I' \in \mathcal{P}$ with either $\eta'_{j-1} < \alpha^r_I = \alpha^l_{I'} < \eta_j$ or $\eta_j < \alpha^r_{I'} = \alpha^l_I < \eta'_j$, $j = 1, 2, \cdots, J$. Then we call $\mathcal{P}$ an *exponential partition* of $\mathbf{R}$ of order λ.

We will show that brushlet bases associated with exponential partitions are greedy bases.

Lemma 3.10. *Let $\{w_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ be a brushlet system satisfying the conditions in Proposition 3.4 where $\mathcal{P}$ is an exponential partition of order $\lambda > 1$. Consider $f = \sum_{(n,I)\in Q} c_{n,I} w_{n,I}$, $Q \subset \mathbf{N}_0 \times \mathcal{P}$, $\#Q = N$. Let $1 < p < \infty$. Assume $\|c_{n,I} w_{n,I}\|_{L_p} \leq 1$, $(n, I) \in Q$. Then*

$$\|f\|_{L_p} \leq C N^{1/p},$$

with C depending only on p and λ.

In order to prove this lemma we need the following observation.

Lemma 3.11. *Let $0 < a_1 < a_2 < \cdots < a_s$ be a set of numbers satisfying $1 < \lambda \leq \frac{a_{j+1}}{a_j}$ and let $E_j \subset \mathbf{R}$, $j = 1, 2, \dots, s$ be measurable sets. Then*

$$\int_{-\infty}^{\infty} \Big[\sum_{j=1}^{s} a_j^{1/q} \chi_{E_j}(x)\Big]^q dx \leq C \sum_{j=1}^{s} a_j |E_j|,$$

where C depends only on q and λ.

The proof is basically given in [19], but will be shown here for completeness.

Proof. Write $F(x) := \sum_{j=1}^s a_j^{1/q} \chi_{E_j}(x)$, and define sets

$$E_\ell^- := E_\ell \backslash \cup_{k=\ell+1}^s E_\ell.$$

Then for $x \in E_\ell^-$ we have $F(x) \leq \sum_{j=1}^\ell a_j^{1/q}$. Define $\beta_j := a_j^{1/q}$, and notice $\tilde{\lambda} := \lambda^{1/q} \leq \frac{\beta_{j+1}}{\beta_j}$. Since $\beta_\ell \geq \tilde{\lambda}^{\ell-j}\beta_j$ we have

$$\sum_{j=1}^{\ell} \beta_j \leq \beta_\ell \sum_{j=0}^{\ell-1} \tilde{\lambda}^{-j} \leq \frac{\tilde{\lambda}}{\tilde{\lambda}-1}\beta_\ell.$$

Thus $F(x) \le C a_\ell^{1/q}$ for $x \in E_\ell^-$. Notice that $\cup_{j=1}^s E_j = \cup_{\ell=1}^s E_\ell^-$. Hence

$$\int_{-\infty}^{\infty} \Big[\sum_{j=1}^s a_j^{1/q}\chi_{E_j}(x)\Big]^q dx \le C\sum_{\ell=1}^s a_\ell |E_\ell^-| \le C\sum_{j=1}^s a_j|E_j|.$$

□

We can now prove Lemma 3.10.

Proof. From Proposition 3.4 we have

$$\|f\|_{L_p} \asymp \Big\| \Big[\sum_{(n,I)\in Q} |c_{n,I}|^2 |I| \chi_{E_{(n,I)}}\Big]^{1/2}\Big\|_{L_p},$$

where $E_{(n,I)} = \big((\pi(n+1/2)-1)|I|^{-1}, (\pi(n+1/2)+1)|I|^{-1}\big)$. Notice that

$$\|c_{n,I}w_{n,I}\|_{L_p} \asymp |c_{n,I}||I|^{1/2}|E_{(n,I)}|^{1/p} = 2^{1/p}|c_{n,I}||I|^{1/2-1/p},$$

i.e. $|c_{n,I}| \le C_p'|I|^{1/p-1/2}$. Thus

$$\|f\|_{L_p} \le C_p''\Big\|\Big[\sum_{(n,I)\in Q}|I|^{2/p}\chi_{E_{(n,I)}}\Big]^{1/2}\Big\|_{L_p} \le C_p''\Big\|\sum_{(n,I)\in Q}|I|^{1/p}\chi_{E_{(n,I)}}\Big\|_{L_p}.$$

Given $\eta \in E$ let $\eta' \in E'$ be the associated point from Definition 3.9 with $\eta < \eta'$. Let $Q_\eta \subset Q$ be the collection of indicies (n,I) with $I \subset (\eta,\eta')$. We will show that

$$\Big\|\sum_{(n,I)\in Q_\eta}|I|^{1/p}\chi_{E_{(n,I)}}\Big\|_{L_p} \le CN^{1/p},$$

and since the set E is finite, this suffices to prove the lemma. Recall that $E_{(n,I)} \cap E_{(n',I)} = \emptyset$ for $n \ne n'$. Introduce the sets

$$E_I := \bigcup_n E_{(n,I)},$$

where the union is taken over all n such that $(n,I) \in Q_\eta$. Since $|E_{(n,I)}| = 2|I|^{-1}$, we have $N \ge 2\sum_I |I||E_I|$. Hence, using Lemma 3.11,

$$\begin{aligned}\Big\|\sum_{(n,I)\in Q_\eta}|I|^{1/p}\chi_{E_{(n,I)}}\Big\|_{L_p} &\le C_p'''\Big\|\sum_I |I|^{1/p}\chi_{E_I}\Big\|_{L_p}\\ &\le C\Big[\sum_I |I||E_I|\Big]^{1/p} \le CN^{1/p},\end{aligned}$$

where C depends only on p and λ. □

Remark 3.12. Notice that for a finite collection P, $\#P = J$,

$$\Big\| \sum_{(n,I)\in P} |I|^{1/p} \chi_{E_{(n,I)}} \Big\|_{L_p} \leq \sum_{(n,I)\in P} (|I||E_{(n,I)}|)^{1/p} \leq J2^{1/p},$$

so we could relax the conditions in Lemma 3.10 a bit by assuming $\mathcal{P}$ is an exponential partition of order $\lambda > 1$ for all but a finite number of intervals.

From Lemma 3.10 we derive an inverse estimate by duality (a proof can be found in [19]).

Lemma 3.13. *Let $\{w_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ be a brushlet system satisfying the conditions in Proposition 3.4, where $\mathcal{P}$ is an exponential partition of order $\lambda > 1$. Consider $f = \sum_{(n,I)\in Q} c_I w_{n,I}$, $\#Q = N$. Let $1 < p < \infty$. Assume $\|c_{n,I} w_{n,I}\|_{L_p} \geq 1$, $(n,I) \in Q$. Then*

$$\|f\|_{L_p} \geq CN^{1/p},$$

with C depending only on p and λ.

Now combining Proposition 3.4, Lemma 3.10 and Lemma 3.13 we have the following result.

Corollary 3.14. *If $\{w^p_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ is a brushlet system normalized in $L_p(\mathbf{R})$, with associated exponential partition and satisfying the conditions in Proposition 3.4. Then the brushlet system is a greedy basis in $L_p(\mathbf{R})$, $1 < p < \infty$.*

The following Corollary of Lemma 3.10 will be used to show a Jackson inequality for N term brushlet approximation (see Corollary 4.5).

Corollary 3.15. *Let $\{w^p_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ be a brushlet system satisfying the conditions in Lemma 3.10. Let $s > 0$ and $1/\tau = s + 1/p$. Then for $f \in L_p(\mathbf{R})$ with associated brushlet coefficients $\{c^p_{n,I}\}$ we have*

$$\sigma_N(f, \{w^p_{n,I}\}, L_p) \leq C\|\{c^p_{n,I}\}\|_{\ell^\tau} N^{-s},$$

with C depending only on p, λ and s.

A proof entirely based on the inequality in Lemma 3.10 can be found in [5]. We leave the calculations to the reader.

3.2. **Brushlet Packets in $L_p(\mathbf{R})$.** One is often interested in adaptive methods to expand functions to get more efficient representations. A popular approach is to use the best basis algorithm of Coifman and Wickerhauser, see [4]. The brushlets are ideally suited for such an algorithm. The most elementary way to implement the algorithm in this setting is to decompose the frequency axis using intervals of equal size giving us a predetermined "finest" frequency resolution. Then one collapses adjacent intervals according to an information cost functional as explained in [4]. This approach is very similar to the idea of wavelet packets and we call such functions brushlet packets. The brushlets associated with the finest frequency localization are called basis brushlet packets. In this section we will look more closely on the $L_p(\mathbf{R})$ convergence of brushlet packet expansions. Since one possible expansion is in the basis brushlet packets there is no hope of unconditional convergence due to the uniform support of the basic brushlet packets in frequency. Still the situation is better than for general wavelet packets where it can be shown that the functions are unstable in $L_p(\mathbf{R})$ [14]. We have the following result.

Proposition 3.16. *Let $\mathcal{P} = \{I_n\}_{n\in\mathbf{Z}}$ with $I_n = \gamma[-1/2+n, 1/2+n)$, for some $\gamma > 0$. For $\varepsilon_I = \beta \le \gamma/2$, the associated brushlet system $\{w_{n,I}\}$ form a Schauder basis for $L_p(\mathbf{R})$ in the sense that*

$$\sum_{|n|\le N} P_{I_n} f + \sum_{n=0}^{K} \langle f, w_{I_{N+1},n}\rangle w_{I_{N+1},n} + \sum_{n=0}^{K} \langle f, w_{I_{-N-1},n}\rangle w_{I_{-N-1},n} \longrightarrow f,$$

as $N, K \to \infty$.

Proof. First we consider $T_N := \sum_{|n|\le N} P_n$ on $L_p(\mathbf{R})$. It is clear that T_N is just the projection onto the space generated by the bell function which is 1 on $[-N+\beta, N-\beta]$ and equal to the bell function associated with P_{-N} on $[-N-\beta, -N+\beta]$ and equal to the bell function associated with P_N on $[N-\beta, N+\beta]$. Thus from Lemma 2.5 it follows that $T_N : L_p(\mathbf{R}) \to L_p(\mathbf{R})$ with bound independent of N. By (2.9) each P_N is spanned by $\{w_{[N-1/2,N+1/2),n}\}_n$. Thus

$$\Big\| \sum_{n=0}^{K} \langle f, w_{I_{N+1},n}\rangle w_{I_{N+1},n} \Big\|_{L_p} \le C_p \|P_{N+1} f\|_{L_p},$$

with C_p independent of N. To complete the proof we notice that the functions with compactly supported Fourier transform are dense in $L_p(\mathbf{R})$ so the brushlet system is also dense in $L_p(\mathbf{R})$. This completes the proof. $\square$

Remark 3.17. Using the Proposition it is easy to show that the appropriate partial sums associated with other brushlet packets also converge in $L_p(\mathbf{R})$ since such partial sums are just special summation methods of the partial sums associated with the basic brushlet packets. The situation is similar to the wavelet packet case, see [15].

4. Brushlet bases in $B_q^s(L_p(\mathbf{R}))$

In this section we consider brushlets in the Besov spaces. We will show that some brushlet systems constitute unconditional bases for the inhomogeneous Besov spaces $B_q^s(L_p(\mathbf{R}))$, $1 < p, q < \infty$, $s > 0$. The main assumption we need in order to characterize a Besov space by brushlet coefficients is that the length of the intervals in the partition $\mathcal{P}$ essentially growth exponentially as their location gets further away from zero. Thus we need to assume that the set of accumulation points is bounded,

$$\eta_0 := \max\{|\eta| : \eta \in E\} < \infty.$$

Given a real positive number α, we define

$$\mathcal{Q}_\alpha = \{I \in \mathcal{P} : I \cap (-\alpha, \alpha) = \emptyset\}.$$

Proposition 4.1. *Let $\{w_{n,I}\}_{I\in\mathcal{P},n>0}$ be a brushlet system with associated partition $\mathcal{P}$. Suppose there exists an absolute constant C such that the bell functions $\{b_I\}_{I\in\mathcal{P}}$ satisfy $|(\xi - \alpha_I^l)\frac{d}{d\xi} b_I(\xi)| \le C$, $\xi \in \mathbf{R}$, and suppose there exists an $\alpha > \eta_0$ such that*

$$1 < \lambda \le \frac{|I'|}{|I|} \le \Lambda < \infty$$

for all adjacent intervals $I, I' \in \mathcal{Q}_\alpha$ with $\alpha_I^r = \alpha_{I'}^l \ge 0$ or $\alpha_I^l = \alpha_{I'}^r < 0$. Then for $1 < p, q < \infty$, $s > 0$,

$$\|f\|_{B_q^s(L_p(\mathbf{R}))} \asymp \Big(\sum_{I\in\mathcal{P}} \big(|I|^s \|P_I f\|_{L_p}\big)^q\Big)^{1/q}. \tag{4.1}$$

Given a partition of unity $\{\hat\phi_k\}_{k\in\mathbf{N}_0}$ satisfying $\mathrm{supp}(\hat\phi_0) \subseteq [-2, 2]$, $\hat\phi_k(\xi) = \hat\phi_k(-\xi)$, $\mathrm{supp}(\hat\phi_k) \subseteq [-2^{k+1}, -2^{k-1}] \cup [2^{k-1}, 2^{k+1}]$ for $k > 0$ and $|\xi\hat\phi_k'(\xi)| \le C$, we will prove that the RHS of (4.1) is equivalent to

$$\Big(\sum_{k\in\mathbf{N}_0} \big(2^{ks}\|\phi_k * f\|_{L_p}\big)^q\Big)^{1/q},$$

In order to prove this we need the following technical lemma.

Lemma 4.2. *Let $\{\alpha_k\}_{k\in\mathbf{N}_0}$ be a strictly increasing sequence of real non negative numbers satisfying $\lim_{k\to\infty} \alpha_k = \infty$. Define a partition $\{I_k\}_{k\in\mathbf{N}_0}$ of $[\alpha_0, \infty)$ by $I_k = [\alpha_k, \alpha_{k+1})$ and let $\{b_k\}_{k\in\mathbf{N}_0}$ be associated window functions. Assume $\lambda \le |I_{k+1}|/|I_k| \le \Lambda$ for all $k \in \mathbf{N}_0$ and define*

$$A_k = \{k' \in \mathbf{Z} : supp(\hat\phi_{k'}) \cap supp(b_k) \neq \emptyset\}$$

and

$$B_k = \{k' \in \mathbf{Z} : supp(b_{k'}) \cap supp(\hat\phi_k) \neq \emptyset\}$$

for $k > 0$. Then $\#A_k \leq d_A < \infty$ and $\#B_k \leq d_B < \infty$ independent of k. Furthermore, there exists constants $0 < c_A, c_B, C_A, C_B < \infty$ such that $c_A 2^{k'} \leq |I_k| \leq C_A 2^{k'}$ for all $k' \in A_k$ and $c_B|I_{k'}| \leq 2^k \leq C_B|I_{k'}|$ for all $k' \in B_k$, independent of k.

Proof. We claim that

$$(\lambda-1)(\alpha_k-\alpha_0)+|I_0| \leq |I_k| \leq (\Lambda-1)(\alpha_k-\alpha_0)+|I_0|, \tag{4.2}$$

which will be proved by induction. Clearly (4.2) holds for $k = 0$. Assume that it holds for $0 \leq n \leq k-1$. Then since $\alpha_k = \alpha_{k-1}+|I_{k-1}|$, we have that

$$(\lambda-1)(\alpha_k-\alpha_0-|I_{k-1}|)+|I_0| \leq |I_{k-1}| \leq (\Lambda-1)(\alpha_k-\alpha_0-|I_{k-1}|)+|I_0|$$

or equivalently that

$$((\lambda-1)(\alpha_k-\alpha_0)+|I_0|)/\lambda \leq |I_{k-1}| \leq ((\Lambda-1)(\alpha_k-\alpha_0)+|I_0|)/\Lambda.$$

This inequality together with the restriction that $\lambda \leq |I_k|/|I_{k-1}| \leq \Lambda$ gives the inequality for k.

We can now prove that $\#A_k \leq d_A$. Using (4.2) for a fixed $k > 0$, we have

$$I_k = [\alpha_k, \alpha_k+|I_k|) \subseteq [\alpha_k, \alpha_k+(\Lambda-1)(\alpha_k-\alpha_0)+|I_0|) \subseteq [\alpha_k, \alpha_k(\Lambda+1)),$$

Let k_i', $i = 0, 1$ be the smallest respectively largest positive integer such that $2^{k_0'+1} \geq \alpha_k$ and $2^{k_1'-1} \leq \alpha_k(\Lambda+1)$ (i.e. $\mathrm{supp}(\hat{\phi}_{k'}) \cap I_k \neq \emptyset$ imply $k_0' \leq k' \leq k_1'$). Then

$$(k_1'-k_0'-2) \leq \log_2\left(\frac{\alpha_k(\Lambda+1)}{\alpha_k}\right) = \log_2(1+\Lambda)$$

for $k > 0$, and since $b_k \subset I_{k-1} \cup I_k \cup I_{k+1}$, we have

$$\#A_k \leq 3\log_2(1+\Lambda)+6.$$

Notice also that (4.2) imply $c_1\alpha_k \leq |I_k| \leq c_2\alpha_k$, where $c_1 = \min\{\lambda-1, \frac{\lambda|I_0|}{|I_0|+\alpha_0}\}$ and $c_2 = \max\{\Lambda-1, \frac{\Lambda|I_0|}{|I_0|+\alpha_0}\}$, and since $2^{k_0'-1} \leq \alpha_k \leq 2^{k_1'+1}$, we have

$$c_1 2^{-d_A-1}2^{k'} \leq |I_k| \leq c_2 2^{d_A+1}2^{k'}$$

for all $k' \in A_k$.

For B_k we use the following arguments. Fix a $k > 0$ and let k_i', $i = 0, 1$ be the smallest respectively largest positive integer such that $I_{k_i'} \cap [2^{k-1}, 2^{k+1}) \neq \emptyset$, i.e. $\alpha_{k_0'}+|I_{k_0'}| \geq 2^{k-1}$ and $\alpha_{k_1'} \leq 2^{k+1}$. Notice that $k_1'-k_0'$ is maximal if $|I_{k'+1}| = \lambda|I_{k'}|$ for $k_0' \leq k' < k_1'$ in which case

$$\alpha_{k_1'} = \alpha_{k_0'} + \sum_{\ell=0}^{k_1'-k_0'-1} \lambda^\ell |I_{k_0'}| = \alpha_{k_0'} + |I_{k_0'}|(\lambda^{k_1'-k_0'}-1)(\lambda-1)^{-1}.$$

Thus

$$\frac{2^{k+1}}{2^{k-1}} \geq \frac{\alpha_{k_1'}}{\alpha_{k_0'} + |I_{k_0'}|} = \frac{\alpha_{k_0'} + |I_{k_0'}|(\lambda^{k_1'-k_0'} - 1)(\lambda-1)^{-1}}{\alpha_{k_0'} + |I_{k_0'}|}$$

which gives

$$\lambda^{k_1'-k_0'} \leq 3\frac{(\lambda-1)\alpha_{k_0'}}{|I_{k_0'}|} + 4\lambda - 3.$$

Using (4.2), and that $|I_{k_0'}| \geq |I_0|$ we get

$$\lambda^{k_1'-k_0'} \leq 3\frac{(\lambda-1)\alpha_0 - |I_0|}{|I_{k_0'}|} + 4\lambda \leq 3\frac{(\lambda-1)\alpha_0}{|I_0|} + 4\lambda - 3,$$

and thus

$$(k_1' - k_0')\log\lambda \leq \log\left(\frac{3\alpha_0}{|I_0|} + 4\right) + \log\lambda.$$

Finally, since $b_{k'} \subset I_{k'-1} \cup I_{k'} \cup I_{k'+1}$, we have the estimate

$$\#B_k \leq \frac{\log\left(\frac{3\alpha_0}{|I_0|} + 4\right)}{\log\lambda} + 3.$$

Furthermore, since $\lambda \leq \lambda^{(k_1'-k_0')} \leq \frac{|I_{k_1'}|}{|I_{k_0'}|} \leq \Lambda^{(k_1'-k_0')} \leq \Lambda^{d_B}$, and $\alpha_{k'} \in [2^{k-1}, 2^{k+1})$ implies $c_1 2^{k-1} \leq |I_{k'}| \leq c_2 2^{k+1}$, where $c_1 = \min\{\lambda - 1, \frac{\lambda|I_0|}{|I_0|+\alpha_0}\}$ and $c_2 = \max\{\Lambda - 1, \frac{\Lambda|I_0|}{|I_0|+\alpha_0}\}$, we have

$$\frac{c_1}{\Lambda^{d_B}}2^{k-1} \leq |I_{k'}| \leq \frac{c_2}{\lambda}2^{k+1}$$

for all $k' \in B_k$. □

With these results at hand we can now prove Proposition 4.1.

of Proposition 4.1. We first notice that

$$\begin{aligned}\sum_{I\in\mathcal{P}\setminus\mathcal{Q}_\alpha} \left(|I|^s \|P_I f\|_{L_p}\right)^q &\leq \sum_{I\in\mathcal{P}\setminus\mathcal{Q}_\alpha} \left(|I|^s C_p \|f\|_{L_p}\right)^q \\ &= C_p^q \|f\|_{L_p}^q \Big(\sum_{I\in\mathcal{P}\setminus\mathcal{Q}_\alpha} |I|^{sq}\Big) \leq C'\|f\|_{L_p}^q,\end{aligned}$$

by Lemma 2.5. Define

$$A_I = \{k \in \mathbf{Z}_+ \,:\, \mathrm{supp}(\hat{\phi}_k) \cap \mathrm{supp}(b_I) \neq \emptyset\}$$

and

$$B_k = \{I \in \mathcal{Q}_\alpha \,:\, \mathrm{supp}(b_I) \cap \mathrm{supp}(\hat{\phi}_k) \neq \emptyset\}.$$

Then there exists $d_A < \infty$ and $d_B < \infty$ such that $\#A_I \leq d_A$ and $\#B_k \leq d_B$ by Lemma 4.2. Notice that $\{\hat{\phi}_k\}_{k\in\mathbf{N}_0}$ is a partition of unity so $\sum_{k\in A_I}\hat{\phi}_k = 1$ on $\mathrm{supp}(b_I)$ and we can substitute $\hat{f}$ in (2.9) by $\sum_{k\in A_I}\hat{\phi}_k\hat{f}$. Thus Lemma 2.5 imply

$$\begin{aligned}\|P_I f\|_{L_p} &= \Big\|P_I\Big(\sum_{k\in A_I}\phi_k * f\Big)\Big\|_{L_p}\\ &\le C_p\Big\|\sum_{k\in A_I}\phi_k * f\Big\|_{L_p} \le C_p\sum_{k\in A_I}\|\phi_k * f\|_{L_p}.\end{aligned}$$

From Lemma 4.2 we have that $|I| \asymp 2^k$ for any $k \in A_I$, where the equivalence only depends on λ, Λ and α. Thus

$$\begin{aligned}\sum_{I\in\mathcal{Q}_\alpha}\big(|I|^s\|P_I f\|_{L_p}\big)^q &\le C\sum_{I\in\mathcal{Q}_\alpha}\Big(\sum_{k\in A_I}|I|^s\|\phi_k * f\|_{L_p}\Big)^q\\ &\le C'\sum_{I\in\mathcal{Q}_\alpha}\Big(\sum_{k\in A_I}2^{ks}\|\phi_k * f\|_{L_p}\Big)^q.\end{aligned}$$

Hölders inequality with $1 = 1/q + 1/q'$ implies

$$\begin{aligned}&\sum_{I\in\mathcal{Q}_\alpha}\Big(\sum_{k\in A_I}2^{ks}\|\phi_k * f\|_{L_p}\Big)^q\\ &\qquad\le \sum_{I\in\mathcal{Q}_\alpha}\Big(\sum_{k\in\mathbf{N}_0}(\mathbf{1}_{A_I}(k))^{q'}\Big)^{q/q'}\Big(\sum_{k\in\mathbf{N}_0}\big(\mathbf{1}_{A_I}(k)2^{ks}\|\phi_k * f\|_{L_p}\big)^q\Big)\\ &\qquad\le d_A^{(q-1)}\sum_{I\in\mathcal{Q}_\alpha}\sum_{k\in\mathbf{N}_0}\mathbf{1}_{A_I}(k)\big(2^{ks}\|\phi_k * f\|_{L_p}\big)^q\\ &\qquad= d_A^{(q-1)}\sum_{k\in\mathbf{N}_0}\Big(\sum_{I\in\mathcal{Q}_\alpha}\mathbf{1}_{A_I}(k)\Big)\big(2^{ks}\|\phi_k * f\|_{L_p}\big)^q\end{aligned}$$

Finally, since there is at most a d_B fold overlap between the A_I's, this gives

$$\begin{aligned}\sum_{I\in\mathcal{Q}_\alpha}\big(|I|^s\|P_I f\|_{L_p}\big)^q &\le C' d_A^{(q-1)}\sum_{k\in\mathbf{N}_0}\Big(\sum_{I\in\mathcal{Q}_\alpha}\mathbf{1}_{A_I}(k)\Big)\big(2^{ks}\|\phi_k * f\|_{L_p}\big)^q\\ &\le C' d_A^{(q-1)} d_B\sum_{k\in\mathbf{N}_0}\big(2^{ks}\|\phi_k * f\|_{L_p}\big)^q.\end{aligned}$$

The other inequality is proved in the same fashion. Since $\sum_{I\in B_k}\widehat{P_I f} = \hat{f}$ on $\mathrm{supp}(\hat{\phi}_k)$,

$$\|\phi_k * f\|_{L_p} = \Big\|\phi_k * \Big(\sum_{I\in B_k}P_I f\Big)\Big\|_{L_p} \le C\sum_{I\in B_k}\|P_I f\|_{L_p},$$

by the Hörmander-Mihlin theorem. We leave the details to the reader. □

We can now calculate the Besov norm of a function f from knowledge about its projections $P_I f$. In fact more can be said.

Corollary 4.3. *Given a brushlet system $\{w_{n,I}\}_{I\in\mathcal{P},n>0}$ as in Proposition 4.1. Suppose, in addition, the modified bell functions $\{g_I\}_{I\in\mathcal{P}}$ satisfies $\|g_I\|_{L^1} \le C < \infty$. Then $\{w_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ form an unconditional basis for $B^s_q(L_p(\mathbf{R}))$, $1 < p, q < \infty, s > 0$, and we have the characterization*

$$\|f\|_{B^s_q(L_p(\mathbf{R}))} \asymp \Big(\sum_{I\in\mathcal{P}}\Big(\sum_{n\in\mathbf{N}_0}\big(|I|^{s+\frac{1}{2}-\frac{1}{p}}|\langle f, w_{n,I}\rangle|\big)^p\Big)^{q/p}\Big)^{1/q}.$$

Proof. Using Proposition 4.1 it suffices to prove that

$$\|P_I f\|_{L_p} \asymp |I|^{\frac{1}{2}-\frac{1}{p}}\Big(\sum_{n\in\mathbf{N}_0}|\langle f, w_{n,I}\rangle|^p\Big)^{1/p}. \tag{4.3}$$

The fact that $\|g_I\|_{L^1} \le C$ together with the representation (2.5) imply that

$$\sup_{x\in\mathbf{R}}\sum_{n\in\mathbf{N}_0}|w_{n,I}(x)| \le C|I|^{\frac{1}{2}} \quad \text{and} \quad \sup_{n\in\mathbf{N}_0}\|w_{n,I}\|_{L_1} \le C'|I|^{-\frac{1}{2}}.$$

With these two properties (4.3) is a well known result (see e.g. [13, pp. 30-31]). □

Remark 4.4. Notice that similar arguments as in the proof of Proposition 4.1 and Proposition 3.4 yields a brushlet characterization of the Triebel-Lizorkin spaces $F^s_q(L_p(\mathbf{R}))$, $1 < p, q < \infty$, $s \ge 0$, given by

$$\|f\|_{F^s_q(L_p(\mathbf{R}))} \asymp \Big\|\Big(\sum_{n\in\mathbf{N}_0, I\in\mathcal{P}}(|\langle f, w_{n,I}\rangle||I|^{s+\frac{1}{2}})^q\chi_{E_{(n,I)}}\Big)^{1/q}\Big\|_{L_p}.$$

From the characterization of the Besov spaces as given in Corollary 4.3 it is now possible to describe the approximation spaces $\mathcal{A}^\gamma_q(\{w^p_{n,I}\}, L_p)$, $1 < p < \infty$, $0 < q < \infty$, $s > 0$, by examining the Besov norm of N-term approximations of functions in $L_p(\mathbf{R})$. As noticed in section 3 the task is to derive certain Jackson and Bernstein inequalities.

Using Corollary 3.15 we can derive the following Jackson inequality.

Corollary 4.5. *Let $\{w^p_{n,I}\}_{I\in\mathcal{P},n\in\mathbf{N}_0}$ be a brushlet system normalized in $L_p(\mathbf{R})$, $1 < p < \infty$ and with associated exponential partition of order λ. Suppose the brushlet system satisfy the conditions in Proposition 3.4* *and Corollary 4.3. Given $s > 0$ such that $\tau := (s + 1/p)^{-1}$ satisfy $1 < \tau < \infty$. Then for $f \in B^s_\tau(L_\tau(\mathbf{R}))$*

$$\sigma_N(f, \{w^p_{n,I}\}, L_p) \leq C\|f\|_{B^s_\tau(L_\tau(\mathbf{R}))} N^{-s},$$

with C depending only on p, λ and s.

Proof. Let $\{c^p_{n,I}\}$ be the brushlet coefficients such that $f = \sum c^p_{n,I} w^p_{n,I}$ for $f \in L_p(\mathbf{R})$. From Proposition 3.4 we have that

$$c^p_{n,I} \asymp |I|^{\frac{1}{2}-\frac{1}{p}} \langle f, w_{n,I}\rangle = |I|^{s+\frac{1}{2}-\frac{1}{\tau}} \langle f, w_{n,I}\rangle.$$

Thus Corollary 4.3 implies

$$\|f\|_{B^s_\tau(L_\tau(\mathbf{R}))} \asymp \|\{c^p_{n,I}\}\|_{\ell_\tau},$$

and the estimate follows from Corollary 3.15. □

Likewise we can derive a Bernstein inequality for N-term brushlet approximation.

Corollary 4.6. *Let $1 < p < \infty$, and let the assumptions of Corollary 4.5 be valid. If $f = \sum_{(n,I)\in Q} c^p_{n,I} w^p_{n,I}$ with $\#Q \leq N$, we have*

$$\|f\|_{B^s_\tau(L_\tau(\mathbf{R}))} \leq CN^s \|f\|_{L_p}.$$

The proof is similar to that of wavelet expansions (see [5, 6]).

Proof. Define

$$S(f) := \Big\{\sum_{I,n} |\langle f, w_{n,I}\rangle|^2 |I| \chi_{E_{(n,I)}}\Big\}^{\frac{1}{2}} \asymp \Big\{\sum_{I,n} |c^p_{n,I}|^2 |I|^{2/p} \chi_{E_{(n,I)}}\Big\}^{\frac{1}{2}},$$

and notice that $|c^p_{n,I}||I|^{1/p}\chi_{E(n,I)} \leq CS(f)$. Hence, using Proposition 3.4 we have

$$\begin{aligned}
\|f\|^\tau_{B^s_\tau(L_\tau)} &\leq C\int_{-\infty}^{\infty} \sum_{I,n} |c^p_{n,I}|^\tau |I| \chi_{E_{(n,I)}}\, dx \\
&= C\int_{-\infty}^{\infty} \sum_{I,n} |c^p_{n,I}|^\tau |I|^{\frac{\tau}{p}} |I|^{1-\frac{\tau}{p}} \chi_{E_{(n,I)}}\, dx \\
&\leq C'\int_{-\infty}^{\infty} S(f)^\tau \sum_{I,n} |I|^{1-\frac{\tau}{p}} \chi_{E_{(n,I)}}\, dx \\
&\leq C'\|S(f)\|^\tau_{L_p} \Big\{\int_{-\infty}^{\infty} \sum_{I,n} |I| \chi_{E_{(n,I)}}\, dx\Big\}^{1-\frac{\tau}{p}} \\
&\leq C''\|f\|^\tau_{L_p} \Big\{\sum_{I,n} |I||E_{(n,I)}|\Big\}^{s\tau} \leq 2C''\|f\|^\tau_{L_p} N^{s\tau}.
\end{aligned}$$

□

As an immediately consequence of the Jackson and Bernstein inequalities in Corollary 4.5 and 4.6 and the relation (3.7) we have the identity

$$\mathcal{A}_q^\gamma(\{w_{n,I}^p\}, L_p) = (L_p(\mathbf{R}), B_\tau^s(L_\tau(\mathbf{R})))_{\gamma/s,q},$$

where $1 < p, \tau < \infty$, $0 < q < \infty$, $1/\tau = s + 1/p$, and $0 < \gamma < s$. When $1/q = \gamma + 1/p$ the RHS equals the Besov space $B_q^\gamma(L_q(\mathbf{R}))$, see [5].

5. Polar Brushlets

The easy extension of the univariate brushlet bases to bases of function spaces in $\mathbf{R}^d$, $d > 1$, is to use functions given as tensor products of brushlets in phase space, see Corollary 3.6. This was the original construction of multivariate brushlets by Meyer and Coifman in [12]. But one could also consider other types of extensions. Instead of using a partition $\mathcal{P}$ of the hole line we could consider a partition $\mathcal{P}^+$ of $\mathbf{R}^+$. This immediately leads to an orthonormal basis $\{\hat{w}_{n,I}^+\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ for $L_2(\mathbf{R}^+)$, with $\hat{w}_{n,I}^+$ given by (2.3) and thus a basis $\{w_{n,I}^+\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ for the Hardy space $\mathcal{H}^2(\mathbf{R})$. If we combine this basis with an orthonormal basis for $L_2(\mathbf{T})$, $\mathbf{T} = [0, 2\pi)$, we are able to construct a basis for $L_2(\mathbf{R}^2)$ that are given as tensor products in polar coordinates.

Definition 5.1. Let $\{\Theta_\ell\}_{\ell=0}^\infty$ be an orthonormal basis for $L_2(\mathbf{T})$. Then we define the functions $\psi_{n,I,\ell}$, $I \in \mathcal{P}^+$, $n, \ell \in \mathbf{N}_0$, by

$$\hat{\psi}_{n,I,\ell}(\xi) = \frac{1}{\sqrt{r}}\hat{w}_{n,I}^+(r)\Theta_\ell(\theta), \quad \xi = re^{i\theta},$$

The reader will notice that the construction bears some resemblance to the Ridgelet construction by D. Donoho [8] but the construction above allows for a more dynamic decomposition of the spatial domain/phaseplane.

Proposition 5.2. *The system $\{\psi_{n,I,\ell}\}_{I\in\mathcal{P}^+,n,\ell\in\mathbf{N}_0}$ form an orthonormal basis for $L_2(\mathbf{R}^2)$.*

Proof. Recall that

$$L_2(\mathbf{R}^2) = L_2(\mathbf{R}^+, r\,dr) \otimes L_2(\mathbf{T}).$$

so we just have to verify that $\{r^{-1/2}\hat{w}_{n,I}^+(r)\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ is an orthornomal basis for $L_2(\mathbf{R}^+, r\,dr)$. The collection $\{r^{-1/2}\hat{w}_{n,I}^+(r)\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ is a well defined orthonormal system in $L_2(\mathbf{R}^+, r\,dr)$ which follows easily from the fact that

$$\int_0^\infty r^{-1/2}\hat{w}_{n,I}^+(r)\overline{r^{-1/2}\hat{w}_{n',I'}^+(r)}\,r\,dr = \int_0^\infty \hat{w}_{n,I}^+(r)\overline{\hat{w}_{n',I'}^+(r)}\,dr = \delta_{n,n'}\delta_{I,I'}.$$

We only need to verify that $\{r^{-1/2}\hat{w}_{n,I}^{+}\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ is dense in $L_2(\mathbf{R}^+, r\,dr)$. Recall that the compactly supported continuous functions $C_c(\mathbf{R}^+)$ with support away from the origin are dense in $L_2(\mathbf{R}^+, r\,dr)$. Let $f \in C_c(\mathbf{R}^+)$ and suppose that $\langle f, r^{-1/2}\hat{w}_{n,I}^{+}\rangle = 0$, for all $I \in \mathcal{P}^+, n \in \mathbf{N}_0$. But this is equivalent for saying that the Fourier coefficients of $r^{1/2}f(r) \in L_2(\mathbf{R}^+, dr)$ w.r.t. the system $\{\hat{w}_{n,I}^{+}\}_{I\in\mathcal{P}^+,n\in\mathbf{N}_0}$ are all zero. Hence, $r^{1/2}f(r) = 0$ which implies that $f = 0$ and the result follows. □

The way one would often use the system $\{\psi_{n,I,\ell}\}_{I\in\mathcal{P}^+,n,\ell\in\mathbf{N}_0}$ is to apply the Fourier transform to the function f and use Plancherel's Theorem: $\langle f, \psi_{n,I,\ell}\rangle = \langle \hat{f}, \hat{\psi}_{n,I,\ell}\rangle$ to calculate the expansion coefficients. However, if one prefers to work in the spatial domain it is actually possible to find an explicit formula for $\psi_{n,I,\ell}$ using the theory of spherical harmonics. We have the following representation.

Proposition 5.3. *Suppose that each Θ_ℓ has an expansion*

$$\Theta_\ell(\theta) = \sum_{s\in\mathbf{Z}} \beta_{\ell,s} e^{is\theta},$$

with $\{\beta_{\ell,s}\}_s \in \ell_1(\mathbf{Z})$. Then

$$\psi_{n,I,\ell}(x) = \sum_{s\in\mathbf{Z}} \beta_{\ell,s} F_{n,I,s}(R) e^{is\omega}, \quad x = Re^{i\omega},$$

where

$$F_{n,I,s}(R) = 2\pi(-i)^s \int_0^\infty \hat{w}_{n,I}^{+}(r) J_s(2\pi Rr)\, r^{1/2}\, dr,$$

and

$$J_s(t) = \frac{1}{2\pi}\int_0^{2\pi} e^{it\sin\theta} e^{-is\theta} d\theta$$

is the Bessel Function of the first kind of order s.

Proof. By definition

$$\hat{\psi}_{n,I,\ell}(\xi) = \frac{1}{\sqrt{r}}\hat{w}_{n,I}^{+}(r)\Theta_\ell(\theta) = \sum_{s\in\mathbf{Z}} \beta_{\ell,s}\frac{1}{\sqrt{r}}\hat{w}_{n,I}^{+}(r) e^{is\theta},$$

is an absolutely convergent sum of terms of the type $\hat{h}_s(r,\theta) := \frac{1}{\sqrt{r}}\hat{w}_{n,I}^{+}(r)$ $e^{is\theta}$. By the theory of spherical harmonics in $\mathbf{R}^2$ the inverse Fourier transform of $\hat{h}_s(r,\theta)$ is given by [18, p. 137]

$$h_s(x) = F_s(R)e^{is\omega}, \quad x = Re^{i\omega},$$

with

$$F_s(R) = 2\pi(-i)^s \int_0^\infty \hat{w}_{n,I}^{+}(r) J_s(2\pi Rr)\, r^{1/2}\, dr.$$

The result then follows from the linearity of the Fourier transform. □

In fact more can be said. Consider the integral

$$I(\rho,\omega) := \frac{1}{2\pi}\sqrt{\rho}\int_0^{2\pi}\Theta(\theta)e^{i\rho[\cos(\theta-\omega)-1]}d\theta,$$

for $\omega \in \mathbf{T}$ and $\rho \in \mathbf{R}^+$. Notice that

$$\psi_{n,I,\ell}(x) = \frac{1}{\sqrt{2\pi R}}\int_0^{\infty}\hat{w}^+_{n,I}(r)I(rR,\omega)e^{irR}dr, \quad x = Re^{i\omega}.$$

From the theory of oscillatory integrals [17, pp. 334–337] we have

$$I(\rho,\omega) = e^{i\pi/4}\Theta(\omega) + O(\rho^{-3/2}), \quad \rho\to\infty,$$

if Θ is supported in a neighborhood of ω. Thus, since

$$\int_0^{\infty} b_I(r)(Rr)^{-3/2}dr \leq 2R^{-3/2}\big((\alpha_I^l-\varepsilon_I^l)^{-1/2}-(\alpha_I^r+\varepsilon_I^r)^{-1/2}\big),$$

we get the approximate representation

$$\psi_{n,I,\ell}(x) = \frac{1}{\sqrt{R}}e^{i\pi/4}w^+_{n,I}(R)\Theta_\ell(\omega) + O(R^{-2}), \quad x = Re^{i\omega},$$

for large α_I^l.

Recall that for certain partitions, $w^+_{n,I}(R)$ consists of two peaks localized at $\pm\frac{\pi(n+1/2)}{|I|}$. If Θ_ℓ is well localized around ω_ℓ, $\psi_{n,I,\ell}$ essentially consists of two peaks at $\frac{\pi(n+\frac{1}{2})}{|I|}e^{i\omega_\ell}$ and $\frac{\pi(n+\frac{1}{2})}{|I|}e^{i(\omega_\ell-\pi)}$ (see figure 1).

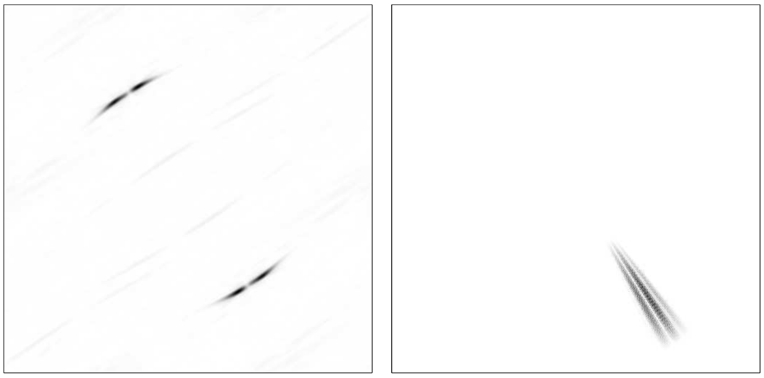

FIGURE 1. Example of a function $\psi_{n,I,\ell}$ (left) and its frequency content (right). In this example Θ_ℓ is a local cosine function.

Considering the very good localization of the basis $\{\psi_{n,I,\ell}\}$ in the phase plane it is natural to expect that such functions form an unconditional basis for the Sobolev spaces and that we have a characterization of Sobolev functions of the same type as for wavelets. That this is indeed the case will be the content of the following Proposition, which will also conclude the paper.

Proposition 5.4. *Let $\{\psi_{n,I,\ell}\}$ be a basis for which there exist a constant $\lambda > 0$ such that $\alpha_I^r/\alpha_I^l \leq \lambda$ for $I \in \mathcal{P}^+$. Then for $s > 0$ we have*

$$\|f\|_{H^s(\mathbf{R}^2)}^2 \asymp \sum_{I\in\mathcal{P}^+,n,\ell\in\mathbf{N}_0} |\langle f,\psi_{n,I,\ell}\rangle|^2(1+(\alpha_I^l)^{2s}).$$

Proof. We have

$$\|f\|_{H^s(\mathbf{R}^2)}^2 \asymp \int_0^\infty |\hat{f}(\xi)|^2(1+|\xi|^{2s})\,d\xi,$$

so it clearly suffices to verify that

$$\int_0^\infty |\hat{f}(\xi)|^2|\xi|^{2s}\,d\xi \asymp \sum_{I\in\mathcal{P}^+,n,\ell\in\mathbf{N}_0} |\langle f,\psi_{n,I,\ell}\rangle|^2(\alpha_I^l)^{2s}. \tag{5.1}$$

Put $g(r) = \int_0^{2\pi} |\hat{f}(r,\theta)|^2\,d\theta$. Then

$$\begin{aligned}
\int_0^\infty\int_0^{2\pi} |\hat{f}(r,\theta)|^2 r^{2s+1}\,d\theta\,dr &= \int_0^\infty g(r)r^{2s+1}\,dr \\
&= \sum_{I\in\mathcal{P}^+}\int_I g(r)r^{2s+1}\,dr \\
&\leq \sum_{I\in\mathcal{P}^+}(\alpha_I^r)^{2s}\int_I g(r)r\,dr
\end{aligned}$$

Given $I \in \mathcal{P}^+$, we denote by $I', I'' \in \mathcal{P}^+$ its two adjacent intervals. Notice that on the annulus $\{(r,\theta) : r \in I\}$, $\hat{f}(r,\theta)$ is given by the orthonormal expansion

$$\hat{f}_I(r,\theta) = \sum_{n,\ell\in\mathbf{N}_0}\sum_{J\in\{I,I',I''\}} \langle \hat{f},\hat{\psi}_{n,J,\ell}\rangle\hat{\psi}_{n,J,\ell}(r,\theta)$$

yielding

$$\int_I g(r)r\,dr \leq \int_0^\infty\int_0^{2\pi} |\hat{f}_I(r,\theta)|^2\,r\,dr\,d\theta = \sum_{n,\ell\in\mathbf{N}_0}\sum_{J\in\{I,I',I''\}} |\langle \hat{f},\hat{\psi}_{s,J,\ell}\rangle|^2.$$

Thus

$$\begin{aligned}
\sum_{I\in\mathcal{P}^+}(\alpha_I^r)^{2s}\int_I g(r)r\,dr &\leq \sum_{I\in\mathcal{P}^+}(\alpha_I^r)^{2s}\sum_{n,\ell\in\mathbf{N}_0}\sum_{J\in\{I,I',I''\}} |\langle \hat{f},\hat{\psi}_{s,J,\ell}\rangle|^2 \\
&\leq (\lambda^{4s}+\lambda^{2s}+1)\sum_{I\in\mathcal{P}^+,n,\ell\geq 0}(\alpha_I^l)^{2s}|\langle f,\psi_{n,I,\ell}\rangle|^2,
\end{aligned}$$

For the converse inequality we have for $I \in \mathcal{P}^+$,

$$\sum_{n,\ell \in \mathbf{N}_0} |\langle f, \psi_{n,I,\ell}\rangle|^2 \leq \sum_{J\in\{I,I',I''\}} \int_J g(r) r\, dr,$$

which implies that

$$\begin{aligned}\sum_{I\in\mathcal{P}^+, n,\ell\in\mathbf{N}_0} & (\alpha_I^l)^{2s} |\langle f, \psi_{n,I,\ell}\rangle|^2 \\ &\leq \sum_{I\in\mathcal{P}^+} \left((\alpha_I^l)^{2s} + (\alpha_{I'}^l)^{2s} + (\alpha_{I''}^l)^{2s}\right) \int_I g(r) r\, dr \\ &\leq (2+\lambda^{2s}) \sum_{I\in\mathcal{P}^+} (\alpha_I^l)^{2s} \int_I g(r) r\, dr \\ &\leq (2+\lambda^{2s}) \sum_{I\in\mathcal{P}^+} \int_I g(r) r^{1+2s}\, dr \\ &= (2+\lambda^{2s}) \int_0^\infty |\hat{f}(\xi)|^2 |\xi|^{2s}\, d\xi.\end{aligned}$$

□

We can now use a standard argument from the theory of real interpolation to extend the result to L_2 based Besov spaces $B_q^s(L_2(\mathbf{R}^2))$. We have

Corollary 5.5. *Let $\{\psi_{n,I,\ell}\}_{I\in\mathcal{P}^+,n,\ell\in\mathbf{N}_0}$ be a basis for which there exists a constant $\lambda > 0$ such that $\alpha_I^r/\alpha_I^l \leq \lambda$ for $I \in \mathcal{P}^+$. Then for $s > 0$ and $0 < q < \infty$ we have*

$$\begin{aligned}\|f\|_{B_q^s(L_2(\mathbf{R}^2))} \asymp & \sum_{I\in\mathcal{P}^+, n,\ell\geq 0} |\langle f, \psi_{n,I,\ell}\rangle|^2 \\ & + \Big(\sum_{I\in\mathcal{P}^+} (\alpha_I^l)^{sq} \Big\{ \sum_{n,\ell\in\mathbf{N}_0} |\langle f, \psi_{n,I,\ell}\rangle|^2 \Big\}^{q/2}\Big)^{1/q}.\end{aligned}$$

References

[1] Pascal Auscher, Guido Weiss, and M. Victor Wickerhauser. Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets. In *Wavelets*, pages 237–256. Academic Press, Boston, MA, 1992.

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

[3] Ronald R. Coifman and Yves Meyer. Remarques sur l'analyse de Fourier à fenêtre. *C. R. Acad. Sci. Paris Sér. I Math.*, 312(3):259–261, 1991.

[4] Ronald R. Coifman and M. Victor Wickerhauser. Entropy based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, 32:712–718, March 1992

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

[6] Ronald A. DeVore, Björn Jawerth, and Vasil Popov. Compression of wavelet decompositions. *Amer. J. Math.*, 114(4):737–785, 1992.

[7] Ronald A. DeVore and George G. Lorentz. *Constructive approximation.* Springer-Verlag, Berlin, 1993.

[8] David L. Donoho. Orthonormal ridgelets and linear singularities. *SIAM J. Math. Anal.*, 31(5):1062–1099 (electronic), 2000.

[9] Eugenio Hernández and Guido Weiss. *A first course on wavelets.* CRC Press, Boca Raton, FL, 1996. With a foreword by Yves Meyer.

[10] S. V. Konyagin and V. N. Temlyakov. A remark on greedy approximation in Banach spaces. *East J. Approx.*, 5(3):365–379, 1999.

[11] Enrico Laeng. Une base orthonormale de $L^2(\mathbf{r})$ dont les éléments sont bien localisés dans l'espace de phase et leurs supports adaptés à toute partition symétrique de l'espace des fréquences. *C. R. Acad. Sci. Paris Sér. I Math.*, 311(11):677–680, 1990.

[12] François G. Meyer and Ronald R. Coifman. Brushlets: a tool for directional image analysis and image compression. *Appl. Comput. Harmon. Anal.*, 4(2):147–187, 1997.

[13] Yves Meyer. *Wavelets and operators.* Cambridge University Press, Cambridge, 1992. Translated from the 1990 French original by D. H. Salinger.

[14] Morten Nielsen. *Size Properties of Wavelet Packets.* Ph.D. thesis, Washington University, St. Louis, May 1999

[15] Morten Nielsen. Walsh-type wavelet packet expansions. *Appl. Comput. Harmon. Anal.*, 9(3):265–285, 2000.

[16] P. Sjögren and P. Sjölin. Littlewood-Paley decompositions and Fourier multipliers with singularities on certain sets. *Ann. Inst. Fourier (Grenoble)*, 31(1):vii, 157–175, 1981.

[17] Elias M. Stein. *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals.* Princeton University Press, Princeton, NJ, 1993. With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.

[18] Elias M. Stein and Guido Weiss. *Introduction to Fourier analysis on Euclidean spaces.* Princeton University Press, Princeton, N.J., 1971. Princeton Mathematical Series, No. 32.

[19] V. N. Temlyakov. The best m-term approximation and greedy algorithms. *Adv. Comput. Math.*, 8(3):249–265, 1998.

[20] V. N. Temlyakov. Non-linear m-term approximation with regard to the multivariate Haar system. *East J. Approx.*, 4(1):87–106, 1998.

[21] Lars F. Villemoes. Adapted Bases of Time-Frequency Local Cosines. *Appl. Comput. Harmon. Anal.*, 10:139–162, 2001.

[22] Mladen Victor Wickerhauser. Smooth localized orthonormal bases. *C. R. Acad. Sci. Paris Sér. I Math.*, 316(5):423–427, 1993.

[23] P. Wojtaszczyk. *Banach spaces for analysts.* Cambridge University Press, Cambridge, 1991.

Department of Mathematical Sciences, Aalborg University, Denmark
E-mail address: lasse@math.auc.dk

Department of Mathematics, University of South Carolina, South Carolina 29208
E-mail address: nielsen@math.sc.edu