Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Nonlinear n-term approximation for the solution of the dirichlet problem

supervised by
Dr. Markus Hansen,
Prof. Dr. Christoph Schwab.

Jakob Zech
September 20, 2013

Abstract

In this semester paper we establish the advantage of nonlinear approximation over linear approximation methods. In particular we consider approximations of solutions of the Dirichlet problem on bounded Lipschitz domains. The paper is structured into three parts: In the first section we introduce multidimensional wavelet bases, Besov spaces and interpolation spaces. We continue with proving a regularity result for the mentioned Dirichlet problem. Finally the approximation spaces of n-term wavelet approximation are characterized via Besov spaces, and we conclude that they are larger than for a linear approximation method.

1 Introduction

In this section we introduce the basic notation and state several lemmas and theorems which we need later on. Throughout this paper we will use the notation $f(x) \sim g(x)$ to indicate that there exist $c_1, c_2 > 0$ such that

$$c_1 f \leq g \leq c_2 f. \tag{1}$$

1.1 Wavelets

We start by constructing a multidimensional wavelet basis. In this paper $\varphi \in L^2(\mathbb{R})$ always denotes a scaling function, which generates a multiresolution analysis, and $\psi \in L^2(\mathbb{R})$ is its corresponding wavelet. For any such pair, $\boldsymbol{x} \in \mathbb{R}^d,\ d > 1,\ \alpha \in \left\{\beta \in \{0,1\}^d : \beta \neq 0^d\right\}$ we define

$$\varphi_d(\boldsymbol{x}) := \bigotimes_{i=1}^{d} \varphi(x_i), \tag{2}$$

$$\psi_{d,\alpha}(\boldsymbol{x}) := \bigotimes_{\{i:\alpha_i=0\}} \varphi(x_i) \bigotimes_{\{i:\alpha_i\neq 0\}} \psi(x_i). \tag{3}$$

Now consider the set of functions:

$$\mathcal{B}_1^1 := \left\{\varphi_d(\boldsymbol{x}-k) : k \in \mathbb{Z}^d\right\}, \tag{4}$$

$$\mathcal{B}_1^2 := \left\{2^{\frac{ld}{2}}\psi_{d,\alpha}(2^l\boldsymbol{x}-k) : \alpha \neq 0^d,\ l \in \mathbb{N}_0, k \in \mathbb{Z}^d\right\}, \tag{5}$$

$$\mathcal{B}_1 := \mathcal{B}_1^1 \cup \mathcal{B}_1^2. \tag{6}$$

For $\eta \in \mathcal{B}$ we define $I_\eta := 2^{-l}(k + [0,1]^d)$ if η has the scaling parameters l, k, which we call **level** (l) and **shift parameter** (k). Notice that $|I_\eta| \sim |\operatorname{supp}(\eta)|$.

Lemma 1. *The set $\mathcal{B}_1$ forms a basis of $L^2(\mathbb{R}^d)$. Moreover if φ is an orthonormal scaling function, i.e. integer shifts of the function are orthonormal, then we get an orthonormal basis of $\mathbb{R}^d$ as well.*

Proof. It is well known that the span of stepfunctions of the type $\mathbb{1}_{\bigotimes_{i=1}^d A_i}$, with measurable $A_i \subset \mathbb{R},\ |A_i| < \infty$, is dense in $L^2(\mathbb{R}^d)$, where $|A_i|$ denotes the Lebesgue measure of the set A_i.

These functions can be written as the product $\prod_{i=1}^{d} \mathbb{1}_{A_i}(x_i)$. Since $\mathcal{C} := \{\varphi(x-k) : k \in \mathbb{Z}\} \cup \left\{2^{\frac{l}{2}}\psi(2^l x - k) : l \in \mathbb{N}_0, k \in \mathbb{Z}\right\}$ is a basis of $L^2(\mathbb{R})$ we can now choose $f_i \in \operatorname{span}(\mathcal{C})$ such that $\|f_i - \mathbb{1}_{A_i}\|_{L^2(\mathbb{R})} < \varepsilon_i$. Then for $d = 2$

$$\begin{aligned}\|f_1 f_2 - \mathbb{1}_{A_1}\mathbb{1}_{A_2}\|_{L^2(\mathbb{R}^2)} &\leq \|f_1 f_2 - f_1 \mathbb{1}_{A_2}\|_{L^2(\mathbb{R}^2)} + \|f_1 \mathbb{1}_{A_2} - \mathbb{1}_{A_1}\mathbb{1}_{A_2}\|_{L^2(\mathbb{R}^2)} \\ &\leq \|\mathbb{1}_{A_2} - f_2\|_{L^2(\mathbb{R})} \|f_1\|_{L^2(\mathbb{R})} + \|\mathbb{1}_{A_1} - f_1\|_{L^2} \|\mathbb{1}_{A_2}\|_{L^2(\mathbb{R})}. \end{aligned} \tag{7}$$

If we choose $\varepsilon_1 := \varepsilon/2$ and $\varepsilon_2 := \varepsilon/(2\|f_1\|_{L^2})$, we see that the span of the functions $\eta_1 \otimes \eta_2$ is dense in $L^2(\mathbb{R}^2)$, if we identify $\eta_1 \otimes \eta_2$ with the function $\eta_1(x_1)\eta_2(x_2)$. Hence $L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) = L^2(\mathbb{R}^2)$, in the sense that there exists an isomorphism between these spaces. Still for $d = 2$ in (4) and (5), and with

$$\mathcal{V}_0 := \operatorname{span}\{\mathcal{B}_1^1\} \tag{8}$$

$$\mathcal{W}_l := \operatorname{span}\left\{2^l \Psi_{2,\alpha}(2^l \boldsymbol{x} - k) : \alpha \neq 0^d,\ k \in \mathbb{Z}^d\right\} \tag{9}$$

$$\mathcal{V}_l := \mathcal{V}_0 \oplus \bigoplus_{j=0}^{l-1} \mathcal{W}_l,\ \forall l > 0, \tag{10}$$

we have $\overline{\bigcup_{l=0}^{\infty} \mathcal{V}_l} = L^2(\mathbb{R})$. It is a property of the functions φ, ψ, that the above sums are in fact direct. Since the $\mathcal{V}_l$ are nested we also get $\overline{\bigcup_{l=0}^{\infty} \mathcal{V}_l \otimes \mathcal{V}_l} = L^2(\mathbb{R}) \otimes L^2(\mathbb{R}) = L^2(\mathbb{R}^2)$. It holds that $\mathcal{V}_l = \mathcal{V}_{l-1} \oplus \mathcal{W}_{l-1}$, and we have for $l = 1$

$$\begin{aligned}\mathcal{V}_l \otimes \mathcal{V}_l &= \mathcal{V}_1 \otimes \mathcal{V}_1 = (\mathcal{V}_0 \oplus \mathcal{W}_0) \otimes (\mathcal{V}_0 \oplus \mathcal{W}_0) \\ &= (\mathcal{V}_0 \otimes \mathcal{V}_0) \oplus (\mathcal{V}_0 \otimes \mathcal{W}_0) \oplus (\mathcal{W}_0 \otimes \mathcal{V}_0) \oplus (\mathcal{W}_0 \otimes \mathcal{W}_0) \\ &= (\mathcal{V}_0 \otimes \mathcal{V}_0) \oplus \bigoplus_{j=0}^{l-1} \left((\mathcal{W}_j \otimes \mathcal{W}_j) \oplus (\mathcal{V}_j \otimes \mathcal{W}_j) \oplus (\mathcal{W}_j \otimes \mathcal{V}_j)\right). \end{aligned} \tag{11}$$

An induction step in l now easily proves (11) for all $l > 1$. Therefore a basis of $\mathcal{V}_l \otimes \mathcal{V}_l$ is given by all functions in $\mathcal{B}_1$ with level less than l. Unifying over all levels l then gives the set $\mathcal{B}_1$. According to the above considerations the span of this set must be dense in $L^2(\mathbb{R}^2)$, and therefore $\mathcal{B}_1$ is a basis of $L^2(\mathbb{R}^2)$. Finally, an induction step in d proves the claim for all $d > 1$. It is obvious from the definition that the basis is orthonormal if φ is an orthonormal scaling function. □

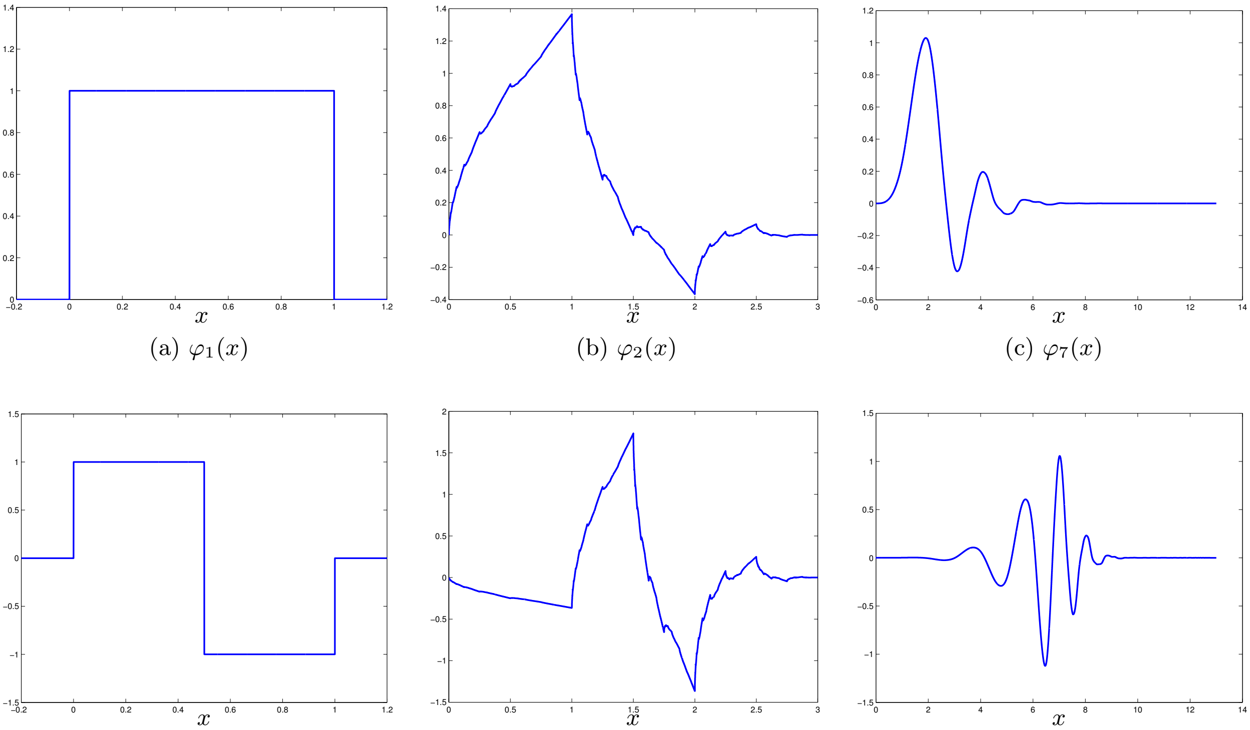
We will also work with another basis, which in contrast to the above one, cannot be defined on bounded domains (cf. [Woj97, Prop. 5.2]):

Lemma 2. *Suppose the scaling function φ generates a multiresolution analysis. Then the set of functions*

$$\mathcal{B}_2 := \left\{2^{\frac{ld}{2}} \psi_{d,\alpha}(2^l \boldsymbol{x} - k) : \alpha \neq 0^d, l \in \mathbb{Z}, k \in \mathbb{Z}^d\right\} \tag{12}$$

forms a Riesz basis of $L^2(\mathbb{R}^d)$.

We will especially work with the **Daubechies wavelets** (see Figure 1), to which we refer as φ_m, ψ_m. They are constructed in [Dau92, Chapter 6]. The scaling function φ_m generates a multiresolution analysis and ψ_m is its wavelet. They have the following properties: They are orthogonal, have compact support and m vanishing moments. This means ψ_m is orthogonal to polynomials of degree less than m. Moreover as m goes to infinity they become arbitrarily smooth.



(a) $\varphi_1(x)$ (b) $\varphi_2(x)$ (c) $\varphi_7(x)$

(d) $\psi_1(x)$ (e) $\psi_2(x)$ (f) $\psi_7(x)$

Figure 1: A few examples of the Daubechies scaling functions φ_m and wavelets ψ_m. As m grows, the support and smoothness of these functions grow as well.

1.2 Besov spaces

Definition 3. *For any Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$ the* **Besov space** $B^s_{p,q}(\Omega)$, $0 < p \leq \infty$, $0 < q < \infty$, $0 < s < \infty$ *can be defined as follows:*

$$\|f\|_{B^s_{p,q}(\Omega)} := \|f\|_{L^p(\Omega)} + \underbrace{\left(\int_0^\infty (t^{-s}\omega_{[s]+1}(f,t,\Omega)_p)^q \frac{dt}{t}\right)^{\frac{1}{q}}}_{=:|f|_{B^s_{p,q}}}, \tag{13}$$

$$B^s_{p,q}(\Omega) := \left\{ f \in L^p(\Omega) : \|f\|_{B^s_{p,q}(\Omega)} < \infty \right\}, \tag{14}$$

where

$$\omega_n(f,t,\Omega)_p := \sup_{\substack{h\in\mathbb{R}^d\\|h|\le t}} \|\Delta_h^n(f)\|_{L^p(\Omega_{h,n})}, \tag{15}$$

$$\Delta_h^n(f)(x) := \sum_{k=0}^{n}\left(\binom{n}{k}(-1)^{n-k}f(x+kh)\right),$$

$$= \Delta_h^{n-1}(\Delta_h f), \quad \text{if } \Delta_h(f)(x) := f(x+h)-f(x), \tag{16}$$

$$\Omega_{h,n} := \{y\in\Omega : y+kh\in\Omega\ \forall 0\le k\le n\}. \tag{17}$$

Moreover if $p=q$ we use the shorter notation $B_p^s(\Omega) := B_{p,p}^s(\Omega)$, and if the context is clear we furthermore omit the Ω. Using a partition of unity we can define $B_{p,q}^s(\partial\Omega)$ as the set of all measurable functions g, which locally can be written as $g(x,\phi(x)) := f(x)$ where $f\in B_{p,q}^s(\mathbb{R}^{d-1})$, and $\partial\Omega$ is the graph of the function ϕ.

Here is a characterization of these spaces in terms of their wavelet coefficients [DD97, Prop. 2.1, Prop 2.2] and [DeV98, Remark 7.4].

Lemma 4. *Let $\varphi,\psi\in C^r(\mathbb{R})$ be a scaling function and its associated wavelet, and $r>s\ge 0$. Then for $p\in(0,\infty)$*

$$\|f\|_{B_p^s(\mathbb{R}^d)} \sim \left\|\sum_{\eta\in\mathcal{B}_1^1}\langle f,\eta\rangle\eta\right\|_{L^p(\mathbb{R}^d)} + \left(\sum_{\eta\in\mathcal{B}_1^2}|I_\eta|^{-\frac{ps}{d}}|\langle f,\eta\rangle|^p\right)^{\frac{1}{p}}. \tag{18}$$

And therefore for $\tau := (s/d+1/2)^{-1}$

$$\|f\|_{B_\tau^s(\mathbb{R}^d)} \sim \left\|\sum_{\eta\in\mathcal{B}_1^1}\langle f,\eta\rangle\eta\right\|_{L^\tau(\mathbb{R}^d)} + \left(\sum_{\eta\in\mathcal{B}_1^2}|\langle f,\eta\rangle|^\tau\right)^{\frac{1}{\tau}}, \tag{19}$$

as well as

$$|f|_{B_\tau^s(\mathbb{R}^d)} \sim \left(\sum_{\eta\in\mathcal{B}_2}|\langle f,\eta\rangle|^\tau\right)^{\frac{1}{\tau}}. \tag{20}$$

Remark 5. The Besov spaces fullfill the following well known embeddings: For $p\ge 1$ and $s>0$ $B_{p,\min(p,2)}^s(\mathbb{R}^d)\hookrightarrow W^{s,p}(\mathbb{R}^d)\hookrightarrow B_{p,\max(p,2)}^s(\mathbb{R}^d)$, where $W^{s,p}(\mathbb{R}^d)$ is the (fractional) Sobolev space with smoothness s in the L^p norm. The notation "$\hookrightarrow$" indicates a continuous embedding. In particular that means $B_2^s = H^s$. Also $B_{p',q'}^{s'}(\mathbb{R}^d)\hookrightarrow L^p(\mathbb{R}^d)$ if the point $(1/p',s')$ lies above the line with slope d throught the point $(1/p,0)$ in the $1/p\times s$ plane that is for $s'>d(1/p'-1/p)$ (see Figure 2).

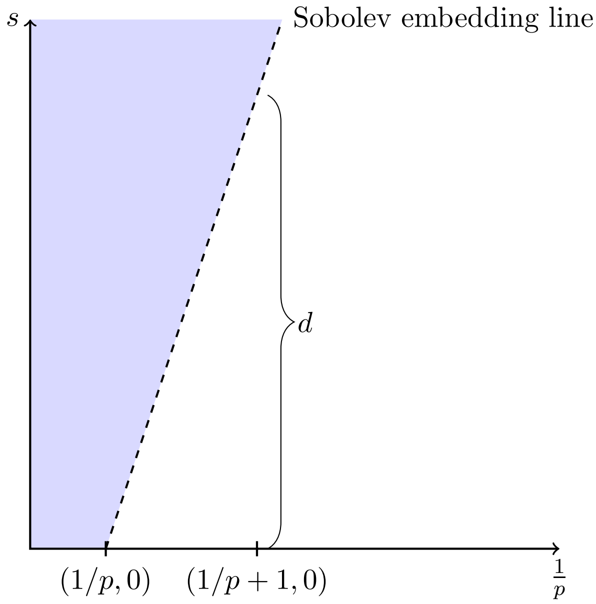


Figure 2: Besov spaces $B^{s'}_{p',q'}(\mathbb{R}^d)$ correspond to the points $(1/p', s')$ in this figure. Points in the light blue area above the dashed line with slope d, are compactly embedded in the space $L^p(\mathbb{R}^d)$. Points in the white area below the dashed line are never embedded in $L^p(\mathbb{R}^d)$ and points on the dashed line may or may not be embedded in $L^p(\mathbb{R}^d)$ depending on their parameter q' that cannot be seen in this picture.

1.3 Interpolation spaces

We shortly describe the concept of real interpolation. Let X, Y be a pair of Banach spaces such that Y is continuously embedded in X. Then we define the K-functional as

$$K(f,t,X,Y) := \inf_{g \in Y} \|f-g\|_X + t\,\|g\|_Y\,, \tag{21}$$

and generally use the shorter notation $K(f,t) := K(f,t,X,Y)$. We are now able to construct intermediate spaces using the K-functional. Let $\theta \in (0,1)$ and $q \in [1,\infty]$, then

$$(X,Y)_{\theta,q} := \left\{ f \in X+Y : \left(\int_{(0,\infty)} (t^{-\theta} K(f,t))^q \frac{dt}{t} \right)^{\frac{1}{q}} < \infty \right\}, \tag{22}$$

$$\|f\|_{(X,Y)_{\theta,q}} := \left(\int_{(0,\infty)} (t^{-\theta} K(f,t))^q \frac{dt}{t} \right)^{\frac{1}{q}}. \tag{23}$$

It is sometimes handy to work with the discretized norm.

Lemma 6. *Let $q \in [1,\infty]$, $\theta \in (0,1)$ and $\rho > 1$. In the above setting we have*

$$\left\| \left(\rho^{j\theta} K(f, \rho^{-j}) \right)_{j \geq 0} \right\|_{l^q} \sim \|f\|_{(X,Y)_{\theta,q}}\,. \tag{24}$$

Proof. Let $a > 0$, $t > a$ and $g \in Y$ then

$$\begin{aligned} K(f,t) &\leq \|f\|_X \\ &\leq \|f-g\|_X + \|g\|_X \\ &\leq c_1(\|f-g\|_X + a\,\|g\|_Y), \end{aligned} \tag{25}$$

since Y is continuously embedded in X. This holds for any $g \in Y$ and hence $K(f,t) \leq c_1 K(f,a)$ $\forall t > a$. For $a/2 < t < a$

$$\begin{aligned} K(f,t) &\geq \inf_{g\in Y} \|f-g\|_X + \frac{a}{2}\,\|g\|_Y \\ &\geq \frac{1}{2}K(f,a), \end{aligned} \tag{26}$$

and therefore

$$\int_{\frac{a}{2}}^{a} (K(f,t)t^{-\theta})^q \frac{1}{t}dt \geq c_2 K(f,a)^q. \tag{27}$$

We obtain from (25) and (27)

$$\begin{aligned} \int_0^\infty (K(f,t)t^{-\theta})^q \frac{1}{t}dt &= \int_0^a (K(f,t)t^{-\theta})^q \frac{1}{t}dt + \int_a^\infty (K(f,t)t^{-\theta})^q \frac{1}{t}dt \\ &\leq \int_0^a (K(f,t)t^{-\theta})^q \frac{1}{t}dt + c_1 K(f,a)^q \\ &\leq c_3 \int_0^a (K(f,t)t^{-\theta})^q \frac{1}{t}dt. \end{aligned} \tag{28}$$

Thus it is equivalent to take the integral from zero to any positive constant a, in particular we now choose $a = 1$. To discretize this integral we remark that $K(f,t)$ is monotonuously decreasing for $t \to 0$. For $t \in [\rho^{-(j+1)}, \rho^{-j}]$ and with the notation $g(t) := t^{-\theta}K(f,t)$,

$$\rho^{j\theta}K(f,\rho^{-(j+1)}) \leq g(t) \leq \rho^{(j+1)\theta}K(f,\rho^{-j}). \tag{29}$$

Like in (26) we have $K(f,\rho^{-(j+1)}) \geq (1/\rho)K(f,\rho^{-j})$:

$$\begin{aligned} & \rho^{j\theta-1}K(f,\rho^{-j}) \leq g(t) \leq \rho^{(j+1)\theta}K(f,\rho^{-j}) \\ \Leftrightarrow\quad & \rho^{(j\theta-1)}(g(\rho^{-j})\rho^{-j\theta}) \leq g(t) \leq \rho^{(j+1)\theta}(g(\rho^{-j})\rho^{-j\theta}) \\ \Leftrightarrow\quad & \rho^{-1}g(\rho^{-j}) \leq g(t) \leq g(\rho^{-j})\rho^{\theta}, \end{aligned} \tag{30}$$

and therefore

$$\begin{aligned} & c_4 \int_{\rho^{-(j+1)}}^{\rho^{-j}} g(\rho^{-j})^q \frac{1}{t}dt \leq \int_{\rho^{-(j+1)}}^{\rho^{-j}} g(t)\frac{1}{t}dt \leq c_5 \int_{\rho^{-(j+1)}}^{\rho^{-j}} g(\rho^{-j})^q \frac{1}{t}dt \\ \Rightarrow\quad & c_4 g(\rho^{-j})^q \leq \int_{\rho^{-(j+1)}}^{\rho^{-j}} g(t)^q \frac{1}{t}dt \leq c_5 g(\rho^{-j})^q. \end{aligned} \tag{31}$$

Summing j from 0 to ∞ and taking the q-th root we conclude

$$c_4 \left(\sum_{j=0}^{\infty} \rho^{j\theta q} K(f, \rho^{-j})^q \right)^{\frac{1}{q}} \leq \left(\int_0^1 (t^{-\theta} K(f,t))^q \frac{1}{t} dt \right)^{\frac{1}{q}} \leq c_5 \left(\sum_{j=0}^{\infty} \rho^{j\theta q} K(f, \rho^{-j})^q \right)^{\frac{1}{q}} . \tag{32}$$

□

Remark 7. A similar proof gives [DL93, p. 56]

$$|f|_{B^s_{p,q}} \sim \left(\sum_{j=0}^{\infty} 2^{js} (\omega_{[s]+1}(f, 2^{-j}, \mathbb{R}^d)_p)^q \right)^{\frac{1}{q}} . \tag{33}$$

In certain situations the discretized norm (24) can be slightly modified. A typical application of the following lemma is the case $X = L^p$, $Y = W^{k,p}$ and $|\cdot|_Y = |\cdot|_{W^{k,p}}$.

Lemma 8. *Let X, Y be a pair of Banach spaces such that Y is densely and continuously embedded in X and $\|g\|_Y \sim \|g\|_X + |g|_Y$, where $|\cdot|_Y$ is a seminorm on Y. With $\tilde{K}(f,t) := \inf_{g \in Y} \|f - g\|_X + t|g|_Y$ and for any $\rho > 1$ we can define the following equivalent norm on $(X,Y)_{\theta,q}$*

$$\underbrace{\left(\sum_{j \geq 0} (\rho^{j\theta} \tilde{K}(f, \rho^{-j}))^q \right)^{\frac{1}{q}}}_{=:|f|_{(X,Y)_{\theta,q}}} + \|f\|_X \sim \|f\|_{(X,Y)_{\theta,q}} . \tag{34}$$

Proof. Since $|f|_Y \leq \|f\|_Y$, it is clear that $|f|_{(X,Y)_{\theta,q}} \leq \|f\|_{(X,Y)_{\theta,q}}$ (cf. (21) and (24)). Also notice that due to the continuous embedding of Y in X

$$K(f,1) = \inf_{g \in Y} \|f - g\|_X + \|g\|_Y \geq c(\|f - g\|_X + \|g\|_X) \geq c \|f\|_X , \tag{35}$$

which is why we can absorb $\|f\|_X$ in $\|f\|_{(X,Y)_{\theta,q}}$, and obtain that the LHS is bounded by a constant multiplied with the RHS in (34). For the other direction we remark that $\left\|(\rho^{-j(1+\theta)} \|f\|_X)_{j \geq 0}\right\|_{l^q} \sim \|f\|_X$ and

$$\left\| \rho^{-j(1+\theta)} \|f\|_X \right\|_{l^q} + \left\| \rho^{j\theta} \tilde{K}(f, \rho^{-j}) \right\|_{l^q} \geq \left\| \rho^{j\theta} (\tilde{K}(f,t) + \rho^{-j} \|f\|_X) \right\|_{l^q} . \tag{36}$$

Hence we compare the elements of the sequences in the l^q norms in (24) respectively (36):

$$\begin{aligned} K(f,\rho^{-j}) &= \inf_{g\in Y} \|f-g\|_X + \rho^{-j}\|g\|_Y \\ &\le \inf_{g\in Y} \|f-g\|_X + c\rho^{-j}(|g|_Y + \|g\|_X) \\ &\le \inf_{g\in Y} \|f-g\|_X + c\rho^{-j}(|g|_Y + \|f-g\|_X + \|f\|_X) \\ &\le 2(c+1) \inf_{g\in Y} \|f-g\|_X + c\rho^{-j}(|g|_Y + \|f\|_X) \\ &\le 2(c+1)(\tilde{K}(f,2^{-j}) + \rho^{-j}\|f\|_X). \end{aligned} \tag{37}$$

This finishes the proof. □

Let us make a few observations:

- $X \cap Y \subset (X,Y)_{\theta,q} \subset X+Y$, with $\|h\|_{X+Y} := \inf_{f+g=h} \|f\|_X + \|g\|_Y$, and $\|f\|_{X\cap Y} = \max(\|f\|_X, \|f\|_Y)$.
- Denote by $B(X,Y)$ all bounded linear operators from a Banach space X to a Banach space Y. If $T \in B(X_1,Y_1) \cap B(X_2,Y_2)$ then $T \in B((X_1,X_2)_{\theta,q},(Y_1,Y_2)_{\theta,q})$. This is called the **interpolation property**.
- $(X,Y)_{s_1,q_1} \subset (X,Y)_{s_2,q_2}$ if either $s_1 > s_2$ or $s_1 = s_2 \wedge q_1 \le q_2$.

The three claims above follow very straightforward from the definition of the norm in the interpolation spaces, respectively the discretized version (24), and the fact that any $n \in \mathbb{N}$, $n \ge [s]+1$ instead of $[s]+1$ defines an equivalent norm in (13). The first two points above characterize **interpolation spaces**, and show that the spaces we constructed are in fact interpolation spaces. The following theorem is a bit more involved. It is an important result and known as the **reiteration theorem** [DL93, Theorem 7.3, p. 195]:

Theorem 9. *Let X,Y be a pair of Banach spaces. For $0<\theta<1$, $0<\theta_1,\theta_2<1$, $q\in[1,\infty]$ we have the identities*

$$((X,Y)_{\theta_1,q},Y)_{\theta,q} = (X,Y)_{\theta_1+\theta(1-\theta_1),q}, \tag{38}$$

$$(X,(X,Y)_{\theta_2,q})_{\theta,q} = (X,Y)_{\theta\theta_2,q}, \tag{39}$$

$$((X,Y)_{\theta_1,q},(X,Y)_{\theta_2,q})_{\theta,q} = (X,Y)_{\theta_1+\theta(\theta_2-\theta_1),q}. \tag{40}$$

To complete this discussion of interpolation spaces we give an interesting example of an interpolation space [DL93, p. 196]:

Theorem 10. *Let $\theta\in(0,1)$, $q\in[1,\infty]$, $r>0$ and $p\in[1,\infty]$. Then*

$$(L^p(\mathbb{R}), W^{r,p}(\mathbb{R}))_{\theta,q} = B^{\theta r}_{p,q}(\mathbb{R}). \tag{41}$$

2 Regularity

In this section we will procede as in the papers [JK95], [DD97] and the references given therein. Let us now consider the problem

$$\begin{aligned} \Delta v &= 0 \quad \text{on } \Omega, \\ v &= g \quad \text{on } \partial\Omega, \end{aligned} \tag{42}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded Lipschitz domain, and $d \geq 3$. In this setting we have the following result [JK95, Thm. 5.1, Thm 5.15]:

Theorem 11. *Let $p = 2$, $d \geq 3$, $s \in [0,1]$, let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, and $g \in B_p^s(\partial\Omega)$. Then there exists a unique weak solution $v \in B_p^{s+1/p}(\Omega)$ for the Dirichlet problem (42).*

Remark 12. Theorem 5.1 in [JK95] is stated for the largest set of pairs $p > 0, s \in (0,1)$ such that Theorem 11 holds. This set is a polygonal domain in the $1/p \times s$ plane. The proof of the existence is done by showing the statement for Sobolev functions on certain parts of the boundary, and then using interpolation techniques and the fact that Besov spaces are interpolation spaces. Let us motivate this for our case where $p = 2$. We need:

$$g \in L^2(\partial\Omega) \Rightarrow u \in B_2^{1/2}(\Omega), \tag{43}$$

$$g \in W^{1,2}(\partial\Omega) \Rightarrow u \in B_2^{3/2}(\Omega). \tag{44}$$

This follows more or less directly from results by Dahlberg [Dah77], [Dah80], Jerison and Kenig [JK81] and characterizations of Besov spaces [JK95, Thm. 4.1, Thm. 4.2] (cf. proof of Theorem 5.15 [JK95]). Now, with (41) for bounded Lipschitz domains and for any $s \in (0,1)$

$$\left(L^2(\partial\Omega), W^{1,2}(\partial\Omega)\right)_{s,2} = B_2^s(\partial\Omega), \tag{45}$$

$$\left(B_2^{1/2}(\Omega), B_2^{3/2}(\Omega)\right)_{s,2} = B_2^{s+1/2}(\Omega). \tag{46}$$

Since the solution operator, which maps the boundary function g to the solution v, is continuous and linear in (43) and (44), we obtain Theorem 11 from the interpolation property, i.e. this operator is then also continuous for the interpolated spaces, that is, from (45) to (46).

The regularity result for the solution of the Dirichlet problem (42) can be further improved for harmonic functions. In order to do so, we need the following theorem [DD97, Theorem 3.1]:

Theorem 13. *Let $v \in B_2^s(\Omega)$ be a harmonic function, for some $s > 0$ and a bounded Lipschitz domain Ω. For every integer $m > s$ we have*

$$\left\| d(x,\partial\Omega)^{m-s} D^m v(x) \right\|_{L^2(\Omega)} \leq c \left\| v \right\|_{B_2^s(\Omega)} \tag{47}$$

for a constant $c > 0$.

Now we can improve Theorem 11.

Theorem 14. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $v \in B_2^s(\Omega)$ a harmonic function and $s > 0$. Then*

$$v \in B_\tau^t(\Omega), \text{ where} \tag{48}$$

$$\tau := \left(\frac{t}{d} + \frac{1}{2}\right)^{-1} \text{ and } 0 < t < \frac{sd}{d-1}. \tag{49}$$

Proof. Throughout this proof we work with the Daubechies wavelets $\varphi_m,\ \psi_m$, where m is large enough such that $\varphi_m,\ \psi_m$ is in $B_\tau^s(\mathbb{R}^d)$. Since Ω is a bounded Lipschitz domain there is a bounded linear extension operator mapping $v \in B_2^t(\Omega)$ to $\hat{v} \in B_2^s(\mathbb{R}^d)$. We now define

$$\tilde{\mathcal{B}} := \{\eta \in \mathcal{B}_1 : \Omega \cap \operatorname{supp} \eta \neq \emptyset\}, \tag{50}$$

$$\tilde{\mathcal{B}}_1 := \tilde{\mathcal{B}} \cap \mathcal{B}_1^1, \tag{51}$$

$$\tilde{\mathcal{B}}_2 := \tilde{\mathcal{B}} \cap \mathcal{B}_1^2, \tag{52}$$

and gain the function

$$\tilde{v} := \sum_{\eta \in \tilde{\mathcal{B}}_1} \langle \hat{v}, \eta \rangle \eta + \sum_{\eta \in \tilde{\mathcal{B}}_2} \langle \hat{v}, \eta \rangle \eta. \tag{53}$$

From the definition it is clear that $\tilde{v}|_\Omega = v$. With $\tilde{\Omega} := \operatorname{supp}(\tilde{v})$ there exist constants c_1 and c_2 such that

$$\|v\|_{B_2^s(\Omega)} \leq c_1 \|\tilde{v}\|_{B_2^s(\tilde{\Omega})} \leq c_1 \|\hat{v}\|_{B_2^s(\mathbb{R}^d)} \leq c_2 \|v\|_{B_2^s(\Omega)}. \tag{54}$$

The first two inequalities follow from (18) and the orthogonality of the Daubechies wavelets, and the last inequality is the continuity of the extension operator. Therefore all three terms are equivalent, and we continue working with $\tilde{v}$. The first sum in (53) is in $B_\tau^t(\tilde{\Omega})$ because it is a finite sum, since $|\tilde{\Omega}| < \infty$ and $\tilde{\mathcal{B}}_1$ consists of integer shifts of a compactly supported function, and all η are in B_τ^t. The main part of the proof is thus to show that the second sum in the expansion (53) belongs to $B_\tau^t(\tilde{\Omega})$. According to Lemma 4 this is equivalent to

$$\left(\sum_{\eta \in \tilde{\mathcal{B}}_2} |\langle \tilde{v}, \eta \rangle|^\tau\right)^{\frac{1}{\tau}} < \infty. \tag{55}$$

In order to estimate this sum we need to distinguish between functions with support on the boundary and the rest. We sort the functions whose support has empty intersection with $\partial\Omega$ into sets according to their level l and the distance of the support to the boundary:

$$\mathcal{D}_l := \left\{\eta \in \tilde{\mathcal{B}}_2 : |I_\eta| = 2^{-ld}\right\}, \tag{56}$$

$$\mathcal{D}_{l,j} := \left\{\eta \in \mathcal{D}_l : j2^{-l} \leq d(\operatorname{supp}(\eta), \partial\Omega) < (j+1)2^{-l}\right\}, \tag{57}$$

$$\mathcal{D}_l^\circ := \mathcal{D}_l \backslash \mathcal{D}_{l,0}. \tag{58}$$

The sets $D_l^\circ = \bigcup_{j \geq 1} \mathcal{D}_{l,j}$ form a partition of the inner functions, i.e. the functions with support in $\mathring{\Omega}$. The rest of the proof consists of first estimating the sum over the inner functions, and then the sum over functions with support on the boundary.

- Inner functions: The support of the functions $\eta \in \tilde{\mathcal{B}}_2$ is the linear transformation of at most $2^d - 1$ bounded Lipschitz domains, namely the supports of the unshifted basis functions at level $l = 0$ in (3). With the Bramble-Hilbert Lemma and a scaling argument we then get the standard result that for $w \in W^{m,2}$ there exists a polynomial P of degree less than m s.t.

$$\|w - P\|_{L^2(\text{supp}(\eta))} \le c|\,\text{supp}(\eta)|^{\frac{m}{d}} |w|_{W^{m,2}(\text{supp}(\eta))}. \tag{59}$$

Recall that φ_m, ψ_m and hence all $\eta \in \tilde{\mathcal{B}}_2$ are orthogonal to polynomials of degree less than m. With the notation $\delta(x) := d(x, \partial\Omega)$, $\delta_\eta := d(\text{supp}(\eta), \partial\Omega)$ and for suitable polynomials P_η it holds that

$$\begin{aligned}
|\langle \tilde{v}, \eta\rangle| &= |\langle \tilde{v} - P_\eta, \eta\rangle| \\
&\le \|\tilde{v} - P_\eta\|_{L^2(\text{supp}(\eta))} \underbrace{\|\eta\|_{L^2(\text{supp}(\eta))}}_{=1} \\
&\overset{(59)}{\le} c|I_\eta|^{\frac{m}{d}} |\tilde{v}|_{W^{m,p}(\text{supp}(\eta))} \\
&\le c|I_\eta|^{\frac{m}{d}} \delta_\eta^{s-m} \underbrace{\left(\int_{\text{supp}(\eta)} (|\delta(x)^{m-s} D^m v(x)|)^2 dx\right)^{\frac{1}{2}}}_{=:\mu_\eta}.
\end{aligned} \tag{60}$$

With this, $|I_\eta| = 2^{-ld}$ and Hölder's inequality

$$\begin{aligned}
\sum_{\eta \in \mathcal{D}_l^\circ} |\langle \tilde{v}, \eta\rangle|^\tau &\le c \sum_{\eta \in \mathcal{D}_l^\circ} \mu_\eta^\tau 2^{-m\tau l} \delta_\eta^{(s-m)\tau} \\
&\le c \left(\sum_{\eta \in \mathcal{D}_l^\circ} (\mu_\eta)^{\tau\frac{2}{\tau}}\right)^{\frac{\tau}{2}} \left(\sum_{\eta \in \mathcal{D}_l^\circ} (2^{-m\tau l} \delta_\eta^{(s-m)\tau})^{\frac{2}{2-\tau}}\right)^{\frac{2-\tau}{2}}.
\end{aligned} \tag{61}$$

The functions in the same level are shifts of fixed lengths of finitely many compactly supported basis functions. Consequently a point x lies in the support of at most $\tilde{c}$ functions $\eta \in \mathcal{D}_l^\circ$ for some $\tilde{c} \in \mathbb{N}$. The sum over the μ_η^2 in (61) can be easily bounded:

$$\begin{aligned}
\sum_{\eta \in \mathcal{D}_l^\circ} \mu_\eta^2 &= \sum_{\eta \in \mathcal{D}_l^\circ} \int_{\text{supp}(\eta)} |\delta(x)^{m-s} D^m v(x)|^2 dx \\
&\le \tilde{c} \int_{\tilde{\Omega}} |\delta(x)^{m-s} D^m v(x)|^2 dx \\
&\le c\, \|\tilde{v}\|^2_{B_2^s(\tilde{\Omega})},
\end{aligned} \tag{62}$$

where the last inequality follows from Theorem 13. For the second sum in (61) we use that Ω is a bounded Lipschitz domain and therefore $|\mathcal{D}_{l,j}| \le c2^{l(d-1)}$. Also $\mathcal{D}_{l,j} = \emptyset$ if $\text{diam}(\Omega) < j2^{-l}$, i.e. $j > c2^l$. With this, (61), (62) and using that for $\eta \in \mathcal{D}_{l,j}$ $\delta_\eta \ge j2^{-l}$ by the definition of

$\mathcal{D}_{l,j}$, we get the estimate

$$\begin{aligned}
\sum_{\eta\in\mathcal{D}_l^\circ} |\langle \tilde{v},\eta\rangle|^\tau &\le c\left(\sum_{j=1}^{c2^l}\sum_{\eta\in\mathcal{D}_{l,j}} 2^{-\frac{2m\tau l}{2-\tau}}\delta_\eta^{\frac{(s-m)2\tau}{2-\tau}}\right)^{\frac{2-\tau}{\tau}} \\
&\le c\left(\sum_{j=1}^{c2^l} 2^{l(d-1)}2^{-\frac{2m\tau l}{2-\tau}}(j2^{-l})^{\frac{(s-m)2\tau}{2-\tau}}\right)^{\frac{2-\tau}{2}} \\
&= c\left(\sum_{j=1}^{c2^l} 2^{l(d-1-\frac{2m\tau}{2-\tau}-\frac{2\tau(s-m)}{2-\tau})} j^{\frac{(s-m)2\tau}{2-\tau}}\right)^{\frac{2-\tau}{2}} \\
&= c\left(2^{l(d-1-\frac{2s\tau}{2-\tau})}\sum_{j=1}^{c2^l} j^{\frac{2\tau(s-m)}{2-\tau}}\right)^{\frac{2-\tau}{\tau}}. \qquad (63)
\end{aligned}$$

In order for the sum in (63) to be uniformly bounded in l we need

$$\frac{2\tau(s-m)}{2-\tau} < -1. \qquad (64)$$

Notice that $0<\tau<2$ by its definition. Therefore this condition can be satisfied by choosing m large enough, and we obtain

$$\sum_{\eta\in\mathcal{D}_l^\circ} |\langle \tilde{v},\eta\rangle|^\tau \le c2^{l(\frac{(d-1)(2-\tau)}{2}-s\tau)}. \qquad (65)$$

Now we sum over all levels

$$\sum_{l=0}^{\infty}\sum_{\eta\in\mathcal{D}_l^\circ} |\langle \tilde{v},\eta\rangle|^\tau \le c\sum_{l=0}^{\infty} 2^{l(\frac{(d-1)(2-\tau)}{2}-s\tau)}. \qquad (66)$$

The above sum is now over all inner functions. It is finite iff

$$\frac{(d-1)(2-\tau)}{2} - s\tau < 0. \qquad (67)$$

Plugging in the definition of $\tau = (t/d+1/2)^{-1} = (2d)/(2t+d)$ we get

$$\begin{aligned}
& & (d-1)(2-\frac{2d}{2t+d}) - 2s\frac{2d}{2t+d} &< 0 \\
&\Leftrightarrow & (d-1)(2(2t+d)-2d) - 2s2d &< 0 \\
&\Leftrightarrow & (d-1)t - sd &< 0 \\
&\Leftrightarrow & t &< \frac{sd}{d-1}, \qquad (68)
\end{aligned}$$

which was our initial assumption (49).

- Functions on the boundary: It remains to estimate the sum of the wavelet coefficients belonging to basis functions that are nonzero on the boundary. Recall that these were the sets $\mathcal{D}_{l,0}$, and we already observed $|\mathcal{D}_{l,0}| \leq c2^{l(d-1)}$. Using twice Hölder for sequences we get

$$
\begin{aligned}
\sum_{l=0}^{\infty} \sum_{\eta \in \mathcal{D}_{l,0}} 1 \cdot |\langle \tilde{v}, \eta \rangle|^{\tau} &\leq \sum_{l=0}^{\infty} \left(\sum_{\eta \in \mathcal{D}_{l,0}} 1 \right)^{\frac{2-\tau}{2}} \left(\sum_{\eta \in \mathcal{D}_{l,0}} |\langle \tilde{v}, \eta \rangle|^2 \right)^{\frac{\tau}{2}} \\
&\leq c \sum_{l=0}^{\infty} 2^{l \frac{(d-1)(2-\tau)}{2}} 2^{-\tau l s} \left(2^{2ls} \sum_{\eta \in \mathcal{D}_{l,0}} |\langle \tilde{v}, \eta \rangle|^2 \right)^{\frac{\tau}{2}} \\
&\leq c \left(\sum_{l=0}^{\infty} 2^{l(\frac{(d-1)(2-\tau)}{2} - s\tau)\frac{2}{2-\tau}} \right)^{\frac{2-\tau}{2}} \left(\sum_{l=0}^{\infty} \sum_{\eta \in \mathcal{D}_{l,0}} 2^{2sl} |\langle \tilde{v}, \eta \rangle|^2 \right)^{\frac{\tau}{2}} . \qquad (69)
\end{aligned}
$$

According to Lemma 4 the second sum is bounded by $\|\tilde{v}\|_{B_2^s(\tilde{\Omega})} < \infty$. For the first sum to be finite we get the condition

$$
\begin{aligned}
& \frac{2}{2-\tau} \left(\frac{(d-1)(2-\tau)}{2} - s\tau \right) < 0 \\
\Leftrightarrow \quad & \frac{(d-1)(2-\tau)}{2} - s\tau < 0, \qquad (70)
\end{aligned}
$$

which is the same as (67), and already fulfilled by assumption. This concludes the proof that (55) holds. Therefore $\tilde{v} \in B_\tau^t(\tilde{\Omega})$ and thus $v \in B_\tau^t(\Omega)$.

□

We finish this section by summing up the above theorems to gain the following result for equation (42).

Corollary 15. *Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain and $d \geq 3$. Assume that $g \in B_2^s(\partial\Omega) = H^s(\partial\Omega)$ for $s \in (0,1)$. For any t with $0 < t < (s + (1/2))d/(d-1)$ and $\tau = (t/d + 1/2)^{-1}$ there exists a unique weak solution $v \in B_\tau^t(\Omega)$ of the Dirichlet problem (42).*

3 n-term wavelet approximation

The goal of this section is to find a characterization of approximation spaces for n-term wavelet approximation. Most results are taken from the chapters 1,4 and 7 in [DeV98] and also from [Sau12]. We start by explaining the term n-term approximation. Suppose that we have a seperable Banach space X, and a basis $(b_n)_{n\in\mathbb{N}}$ of X. With the definition

$$X_n := \operatorname{span}\{b_i : 1 \leq i \leq n\}, \qquad (71)$$

$$\mathcal{X}_n := \bigcup_{\{\Lambda \subset \mathbb{N}: |\Lambda| \leq n\}} \operatorname{span}\{b_i : i \in \Lambda\}, \qquad (72)$$

we can consider two types of approximation for an element $f \in X$: We could either approximate f in the n dimensional subspace X_n of X, or we could find an approximation of f in the n dimensional manifold $\mathcal{X}_n$, which is a subset of X. Since $\mathcal{X}_n$ is not a linear space, this type of approximation is also called **nonlinear approximation**. For example, if $X = L^2([0,1])$ and the b_n are hat functions, then the linear method is consistent with finding the approximant in the set of piecewise linear functions with n fixed nodes depending on the first n hat functions. With the nonlinear method however, we approximate our target function with a piecewise linear function where the location of the n nodes can be freely choosen within the possible nodes of our basis. This will be a dense subset of the intervall $[0,1]$. Obviously we expect a nonlinear method to perform better. The space we will be looking at is $X = L^2(\mathbb{R}^d)$, and the basis $(b_n)_{n\in\mathbb{N}}$ will be the wavelet basis $\mathcal{B}_2$ from (12). In the following we characterize functions whose error in the n-term approximation decays at a certain order s.

Definition 16. *Let X be a Banach space, and $(\mathcal{X}_n)_{n\in\mathbb{N}}$ a sequence of nested subsets of X, such that their union is dense in X. Suppose moreover that $a\mathcal{X}_n = \mathcal{X}_n\ \forall a \in \mathbb{R}$ and $\mathcal{X}_n + \mathcal{X}_n \subset \mathcal{X}_{\kappa n}$ for a fixed integer $\kappa \in \mathbb{N}$. For $0 < q < \infty$, $s > 0$ and with $d_X(f, \mathcal{X}_n) := \inf_{g\in\mathcal{X}_n} \|f - g\|_X$ we define*

$$\mathcal{A}^s(X) := \mathcal{A}^s_\infty := \left\{ f \in X : d_X(f, \mathcal{X}_n) = \mathcal{O}(n^{-s}) \text{ as } n \to \infty \right\}, \tag{73}$$

$$\mathcal{A}^s_q(X) := \left\{ f \in X : \left(\sum_{n=1}^{\infty} (n^s d_X(f, \mathcal{X}_n))^q \frac{1}{n} \right)^{\frac{1}{q}} < \infty \right\}, \tag{74}$$

and call them the **approximation spaces** *of* **approximation order** *s. We define a quasinorm on $\mathcal{A}^s_q$ by*

$$\|f\|_{\mathcal{A}^s_q} := \|f\|_X + \underbrace{\left(\sum_{n=1}^{\infty} (n^s d_X(f, \mathcal{X}_n))^q \frac{1}{n} \right)^{\frac{1}{q}}}_{=:|f|_{\mathcal{A}^s_q}}, \tag{75}$$

with the second term being a quasiseminorm on $\mathcal{A}^s_q(X)$. If there is no ground for confusion we omit the X and write $\mathcal{A}^s := \mathcal{A}^s(X)$ and anlogue for $\mathcal{A}^s_q$.

Another example for the setting of Definition 16 is the case of approximation by rational functions. It is clear that for

$$\mathcal{X}_n := \left\{ R = \frac{P}{Q} : P, Q \in \mathbb{R}[x],\ \deg(P), \deg(Q) \le n \right\}, \tag{76}$$

as well as for the n-dimensional manifolds $\mathcal{X}_n$ above, κ in Definition 16 can be chosen as 2, and the other assumptions are also fullfilled.

Remark 17. Functions in $\mathcal{A}^s$ are those for which the error of the best approximation in the spaces $\mathcal{X}_n$ decays at order s. The parameter q allows a finer distinction between these functions.

We observe for $0 < q < \infty$

$$\sum_{j=0}^{\infty}\left(\sum_{n=2^j}^{2^{(j+1)}-1}(2^{js}d_X(f,\mathcal{X}_{2^{j+1}}))^q 2^{-(j+1)}\right) \le \sum_{n=1}^{\infty}(n^s d_X(f,\mathcal{X}_n))^q\frac{1}{n}$$

$$\le \sum_{j=0}^{\infty}\left(\sum_{n=2^j}^{2^{(j+1)}-1}(2^{(j+1)s}d_X(f,\mathcal{X}_{2^j}))^q 2^{-j}\right), \qquad (77)$$

and obtain

$$|f|_{\mathcal{A}^s_q} \sim \left(\sum_{j\ge 0}(2^{js}d_X(f,\mathcal{X}_{2^j}))^q\right)^{\frac{1}{q}}. \qquad (78)$$

It is possible to describe the approximation spaces in terms of the wavelet coefficients, and gain the results of this section in a direct way. We will not do so however and use a theorem, which is stated in the general Banach space setting above. For the proof we need the following Lemma.

Lemma 18. *Let $(a_j)_{j\ge 0}$ be a sequence of nonnegative real numbers. For $0 < s < r$ and $q \in [1,\infty]$, there exists a constant c such that*

$$\left\|\left(2^{(s-r)j}\sum_{k=0}^{j}2^{kr}a_k\right)_{j\ge 0}\right\|_{l^q} \le c\left\|(2^{sj}a_j)_{j\ge 0}\right\|_{l^q} \qquad (79)$$

Proof. We start with the case $1 < q < \infty$. Let $1/q + 1/q' = 1$, then with $\alpha := (r-s)/2 > 0$ and by Hölder's inequality

$$\sum_{j\ge 0}2^{(s-r)qj}\left(\sum_{k=0}^{j}2^{rk}a_k\right)^q \le \sum_{j\ge 0}2^{(s-r)qj}\underbrace{\left(\sum_{k=0}^{j}(2^{\alpha k})^{q'}\right)^{\frac{q}{q'}}}_{\le c2^{\alpha jq}}\left(\sum_{k=0}^{j}(2^{(r-\alpha)k})^q\right)$$

$$\le c\sum_{j\ge 0}2^{-\alpha jq}\left(\sum_{k=0}^{j}(2^{(r-\alpha)k})^q\right)$$

$$= c\sum_{k\ge 0}(2^{(r-\alpha)k}a_k)^q\, \underbrace{c\sum_{j\ge k}2^{-\alpha jq}}_{\le 2^{-\alpha kq}}$$

$$\le c\sum_{k\ge 0}(2^{sk}a_k)^q. \qquad (80)$$

For $q = 1$

$$\sum_{j\geq 0} 2^{(s-r)j} \sum_{k=0}^{j} 2^{rk} a_k = \sum_{k\geq 0} 2^{rk} a_k \underbrace{\sum_{j\geq k} 2^{(s-r)j}}_{\leq c2^{(s-r)k}} \leq c \sum_{k\geq 0} 2^{sk} a_k, \tag{81}$$

and for $q = \infty$

$$2^{sj} 2^{-rj} \sum_{k=0}^{j} 2^{rk} a_k \leq 2^{sj} 2^{-rj} \left\| (2^{sj} a_j)_{j\geq 0} \right\|_{l^q} \sum_{k=0}^{j} 2^{(r-s)k} \leq c 2^{sj} \left\| (2^{sj} a_j)_{j\geq 0} \right\|_{l^q} 2^{-sj}, \tag{82}$$

which proves (79). □

Theorem 19. *Let X and Y be a pair of Banach spaces where Y is continuously and densely embedded in X, with $\|f\|_Y \sim \|f\|_X + |f|_Y$ for a seminorm $|\cdot|_Y$ on Y. Assume that $(\mathcal{X}_n)_{n\in\mathbb{N}_0}$ is a sequence of nested subsets of Y, that fullfills the assumptions in Definition 16, and there exists $r > 0$ for which the* **Jackson inequality**

$$d_X(f, \mathcal{X}_n) \leq c_1 n^{-r} |f|_Y \qquad \forall f \in Y,\ \forall n \in \mathbb{N}_0, \tag{83}$$

and the **Bernstein inequality**

$$|f|_Y \leq c_2 n^r \|f\|_X \qquad \forall f \in \mathcal{X}_n,\ \forall n \in \mathbb{N}_0, \tag{84}$$

hold for some constants $c_1, c_2 > 0$. For $q \geq 1$ and $r > s$ we then have

$$\mathcal{A}_q^s(X) = (X, Y)_{s/r,q}\,. \tag{85}$$

Proof. In this proof we work with the norm representation (78) of the interpolation space, and use the notation $\tilde{\mathcal{X}}_j := \mathcal{X}_{2^j}$. With (78) and $\rho = 2^r$ in (34) we need to show

$$\|f\|_X + \underbrace{\left(\sum_{j\geq 0} (2^{js} \tilde{K}(f, 2^{-jr}))^q \right)^{\frac{1}{q}}}_{=|f|_{(X,Y)_{s/r,q}}} \sim \|f\|_X + \underbrace{\left(\sum_{j\geq 0} (2^{js} d_X(f, \tilde{\mathcal{X}}_j))^q \right)^{\frac{1}{q}}}_{=|f|_{\mathcal{A}_q^s}}, \tag{86}$$

where $\tilde{K}(f,t) = \inf_{g\in Y} \|f-g\|_X + |g|_Y$. For the first direction we bound the error of the best approximation by the $\tilde{K}$-functional. For any $g_j \in \tilde{\mathcal{X}}_j \subset Y$ and $g \in Y$

$$\begin{aligned} d_X(f, \tilde{\mathcal{X}}_j) &\leq \|f - g_j\|_X \\ &\leq \|f - g\|_X + \|g - g_j\|_X\,. \end{aligned} \tag{87}$$

We minimize over $g_j \in \tilde{\mathcal{X}}_j$ and use the Jackson estimate (83) for the second term in (87) to obtain

$$d_X(f, \tilde{\mathcal{X}}_j) \le c(\|f - g\|_X + 2^{-jr}|g|_Y). \tag{88}$$

Minimizing over $g \in Y$ then yields

$$d_X(f, \tilde{\mathcal{X}}_j) \le c\tilde{K}(f, 2^{-jr}), \tag{89}$$

and we get $|f|_{\mathcal{A}^s_q} \le c|f|_{(X,Y)_{s/r,q}}$. For the other direction we choose elements $f_j \in \mathcal{X}_j$ with $d_X(f, f_j) \le 2d_X(f, \tilde{\mathcal{X}}_j)$ and use the triangle inequality

$$\begin{aligned}\tilde{K}(f, 2^{-jr}) &\le \|f - f_j\|_X + 2^{-jr}|f_j|_Y \\ &\le \|f - f_j\|_X + 2^{-jr}\left(|f_0|_Y + \sum_{k=1}^{j} |\underbrace{f_k - f_{k-1}}_{\in \tilde{\mathcal{X}}_{\kappa k}}|_Y\right).\end{aligned} \tag{90}$$

Now we make use of the Bernstein inequality (84) and of $d_X(f, f_k) \le 2d_X(f, \tilde{\mathcal{X}}_k)$ $\forall k = 0, \ldots, j$

$$\tilde{K}(f, 2^{-rj}) \le \|f - f_j\|_X + \tilde{c}(\kappa)2^{-rj}\left(|f_0|_Y + \sum_{k=1}^{j} 2^{rk} \underbrace{\|f_k - f_{k-1}\|_X}_{\le \|f_k - f\|_X + \|f - f_{k-1}\|_X}\right). \tag{91}$$

Again because of Bernstein $|f_0|_Y \le c_2 \|f\|_X \le c_2(\|f\|_X + \|f - f_0\|_X)$, and thus

$$\tilde{K}(f, 2^{-rj}) \le c2^{-rj}\left(\|f\|_X + \sum_{k=0}^{j} 2^{(k-j)r} d_X(f, \tilde{\mathcal{X}}_k)\right). \tag{92}$$

The next step is to multiply both sides with 2^{sj} and to take the l^q norm of the sequences with index j. Lemma 18 then finishes the proof with $a_k := d_X(f, \tilde{\mathcal{X}}_k)$

$$\begin{aligned}\left\|\left(2^{sj}\tilde{K}(f, 2^{-rj})\right)_{j\ge 0}\right\|_{l^q} &\le c\left\|\left(2^{(s-r)j}\|f\|_X + 2^{(s-r)j}\sum_{k=0}^{j} 2^{rk} d_X(f, \tilde{\mathcal{X}}_k)\right)_{j\ge 0}\right\|_{l^q} \\ &\le c\left(\left\|\left(2^{(s-r)j}\|f\|_X\right)_{j\ge 0}\right\|_{l^q} + \left\|\left(2^{(s-r)j}\sum_{k=0}^{j} 2^{rk} d_X(f, \tilde{\mathcal{X}}_k)\right)_{j\ge 0}\right\|_{l^q}\right) \\ &\le c\left(\|f\|_X + |f|_{\mathcal{A}^s_q}\right).\end{aligned} \tag{93}$$

□

To describe the approximation spaces for the n-term wavelet approximation, we now find a space Y that fits the above setting.

Lemma 20. *Let $X = L^2(\mathbb{R}^d)$, $Y = B^t_\tau(\mathbb{R}^d)$ for $t > 0$ and $1/\tau = t/d + 1/2$. Assume that the wavelet ψ with which we define the basis $\mathcal{B}_2$ in* (12) *has m vanishing moments where $m > t$ and $\psi \in B^{t+\delta}_{\tau,q'}$ for some $\delta > 0$, $q' \in (0, \infty]$. Let $\mathcal{X}_n$ be the sets*

$$\mathcal{X}_n := \bigcup_{\{\Lambda \subset \mathcal{B}_2 : |\Lambda| \leq n\}} \operatorname{span}\{\eta : \eta \in \Lambda\}. \tag{94}$$

Then there exists $c > 0$ such that the Jackson and Bernstein inequalities hold

$$d_{L^2}(f, \mathcal{X}_n) \leq c n^{-\frac{t}{d}} |f|_{B^t_\tau} \qquad \forall f \in B^t_\tau, \tag{95}$$

$$|f|_{B^t_\tau} \leq c n^{\frac{t}{d}} \|f\|_{L^2} \qquad \forall f \in \mathcal{X}_n. \tag{96}$$

Proof. First of all let us remark, that the assumptions on the wavelet ψ are sufficient so that we can use the characterization in terms of wavelet coefficients from Lemma 4 ([DeV98, p. 119]). We start with the more difficult Jackson inequality (95). According to (20) we have

$$|f|_{B^t_\tau} \sim \left(\sum_{\eta \in \mathcal{B}_2} |\langle f, \eta \rangle|^\tau \right)^{\frac{1}{\tau}}, \tag{97}$$

and therefore for any $\varepsilon > 0$

$$|\{\eta \in \mathcal{B}_2 : |\langle f, \eta \rangle| > \varepsilon\}| \leq c |f|^\tau_{B^t_\tau} \varepsilon^{-\tau}. \tag{98}$$

With $\Lambda_j := \left\{\eta \in \mathcal{B}_2 : 2^j < |\langle f, \eta \rangle| < 2^{j+1}\right\}$ we get

$$\sum_{j=k}^{\infty} |\Lambda_j| \leq c |f|^\tau_{B^t_\tau} 2^{-k\tau}, \tag{99}$$

since due to (98) $|\Lambda_j| \leq |f|^\tau_{B^t_\tau} 2^{-j\tau}$. Let

$$g_j := \sum_{\eta \in \Lambda_j} \langle \eta, f \rangle \eta, \text{ and} \tag{100}$$

$$f_k := \sum_{j \geq k} g_j, \tag{101}$$

then

$$\|f - f_k\|_{L^2} = \left\| \sum_{j=-\infty}^{k-1} g_j \right\|_{L^2} \leq \sum_{j=-\infty}^{k-1} \|g_j\|_{L^2}. \tag{102}$$

Notice that $f_k \to f$ for $k \to -\infty$. To estimate $\|g_j\|_{L^2}$ we use that its wavelet coefficients are bounded by 2^{j+1} by construction, and that $\mathcal{B}_2$ is a Riesz basis in L^2, and obtain

$$\|g_j\|^2_{L^2} = \left\| \sum_{\eta \in \Lambda_j} \langle f, \eta \rangle \eta \right\|^2_{L^2} \leq c \sum_{\eta \in \Lambda_j} |\langle f, \eta \rangle|^2 \leq c 2^{2(j+1)} |\Lambda_j| \leq c 2^{j(2-\tau)} |f|^\tau_{B^t_\tau}, \tag{103}$$

where the last step follows again from (98). Summing over j from $-\infty$ to $k-1$ we conclude from (102) and (103)

$$\|f - f_k\|_{L^2} \le c|f|_{B^t_\tau}^{\frac{\tau}{2}} \sum_{j=-\infty}^{k-1} 2^{\frac{j(2-\tau)}{2}} \le c|f|_{B^t_\tau}^{\frac{\tau}{2}} 2^{k(1-\frac{\tau}{2})}. \tag{104}$$

In terms of expansion coefficients $N(k)$ of f_k with $N(k) = \mathcal{O}(|f|_{B^t_\tau}^{\tau} 2^{-k\tau})$ for $k \to -\infty$ because of (99), this means

$$\begin{aligned} d_{L^2}(f, \mathcal{X}_{N(k)}) &\le c|f|_{B^t_\tau} N(k)^{-\frac{1}{\tau}(1-\frac{\tau}{2})} \\ &\le c|f|_{B^t_\tau} N(k)^{\frac{1}{2}-\frac{1}{\tau}} = c|f|_{B^t_\tau} N(k)^{-\frac{t}{d}}. \end{aligned} \tag{105}$$

The monotone decay of the approximation error $d_{L^2}(f, \mathcal{X}_n)$, and the (at most) exponential and monotone growth of $N(k)$ imply that there is a constant $\tilde{c}$ with $d_{L^2}(f, \mathcal{X}_n) \le \tilde{c}|f|_{B^t_\tau} n^{-\frac{t}{d}}$ $\forall n \in \mathbb{N}$:

$$\begin{aligned} \sup_{n \ge N(1)} \frac{d_{L^2}(f, \mathcal{X}_n)}{n^{-\frac{t}{d}}} &\le \sup_{k \in \mathbb{N}} \frac{d_{L^2}(f, \mathcal{X}_{N(k)})}{N(k+1)^{-\frac{t}{d}}} \\ &\le c|f|_{B^t_\tau} \sup_{k \in \mathbb{N}} \left(\frac{N(k+1)}{N(k)}\right)^{\frac{t}{d}} \le \tilde{c}|f|_{B^t_\tau}, \end{aligned} \tag{106}$$

which proves (95). The Bernstein inequality (96) is particularly easy in our setting because $\mathcal{B}_2$ is a Riesz basis in $L^2(\mathbb{R}^d)$ (that is not true for L^p where $p \ne 2$). If $f \in \mathcal{X}_n$, then

$$f = \sum_{\eta \in \Lambda} \langle f, \eta \rangle \eta \tag{107}$$

for a $\Lambda \subset \mathcal{B}_2$ with $|\Lambda| \le n$. With (20) and Hölder's inequality we get for such an f

$$\begin{aligned} |f|_{B^t_\tau} &\le c \left(\sum_{\eta \in \mathcal{B}_2} |\langle f, \eta \rangle|^\tau \right)^{\frac{1}{\tau}} \\ &\le \left(\sum_{\eta \in \mathcal{B}_2} |\langle f, \eta \rangle|^{\tau \frac{2}{\tau}} \right)^{\frac{1}{\tau}\frac{\tau}{2}} \left(\sum_{\eta \in \mathcal{B}_2} 1^{\tau \frac{2}{2-\tau}} \right)^{\frac{1}{\tau}\frac{2-\tau}{2}} \\ &\le c\, \|f\|_{L^2} \, |\Lambda|^{(\frac{1}{\tau}-\frac{1}{2})} = c\, \|f\|_{L^2} \, n^{-\frac{t}{d}}. \end{aligned} \tag{108}$$

$\square$

Corollary 21. *Let $X = L^2(\mathbb{R}^d)$ and $\mathcal{X}_n$ be the spaces*

$$\mathcal{X}_n := \bigcup_{\{\Lambda \subset \mathcal{B}_2 : |\Lambda| \le n\}} \operatorname{span}\{\eta : \eta \in \Lambda\}. \tag{109}$$

Assume that the wavelet ψ with which we define the basis $\mathcal{B}_2$ in (12) *has m vanishing moments for some $m > t$ and $\psi \in B^{t+\delta}_{\tilde{\tau},q'}$ for $t > 0$, $\delta > 0$ and some $q' \in (0, \infty]$. With $1/\tilde{\tau} = t/d + 1/2$, $q \geq 1$ and for $s < t$ we then have*

$$\mathcal{A}^{s/d}_q(L^2(\mathbb{R}^d)) = \Big(L^2(\mathbb{R}^d), B^t_{\tilde{\tau}}(\mathbb{R}^d)\Big)_{s/t,q}, \tag{110}$$

and in particular for $1/\tau = s/d + 1/2$

$$\mathcal{A}^{s/d}_\tau(\mathbb{R}^d) = B^s_\tau(\mathbb{R}^d). \tag{111}$$

Proof. The assumptions in Definition 16 are easily checked. Formula (110) then follows directly from Theorem 19 and Lemma 20. Equation 111 is an interpolation result similar to Theorem 10, see [DeV98, Remark 7.6]. In fact τ is the only value for q such that the approximation space is a Besov space. $\square$

Remark 22. We obtained the results in Corollary 21 for the domain $\mathbb{R}^d$, $d \geq 1$. This can be generalized to bounded Lipschitz domains $\Omega \subset \mathbb{R}^d$, by using a bounded extension operator $E : B^s_p(\Omega) \to B^s_p(\mathbb{R}^d)$, and then considering only those wavelets that are nonzero on Ω. If we return to the setting and notation of Corollary 15, we obtain that the convergence rate of the n-term wavelet approximation for the solution of (42) is t/d. Of course we have to keep in mind that this is only true if the wavelet we are working with is smooth enough.
Now let us go back to Corollary 21. With $B^{s_0}_{p_0,q}(\mathbb{R}^d) \hookrightarrow B^{s_1}_{p_1,q}(\mathbb{R}^d)$ for $s_0 - s_1 = (d/p_0 - d/p_1)$ [Tri10, Section 2.7.1, p. 127] we obtain $H^s(\Omega) = B^s_{2,2}(\Omega) \hookrightarrow B^{s+\varepsilon}_{\tau,2}(\Omega) \hookrightarrow B^s_\tau(\Omega)$, where ε depends on d and s. Therefore $H^s(\Omega) \subsetneq \mathcal{A}^{s/d}_\infty$, and $H^s(\Omega)$ is the approximation space with order s/d for the linear method we described at the beginning of this section. That means the approximation space for the nonlinear method is larger.

References

[Dah77] Björn E. J. Dahlberg. Estimates of harmonic measure. *Arch. Rational Mech. Anal.*, 65(3):275–288, 1977.

[Dah80] Björn E. J. Dahlberg. Weighted norm inequalities for the Lusin area integral and the nontangential maximal functions for functions harmonic in a Lipschitz domain. *Studia Math.*, 67(3):297–314, 1980.

[Dau92] Ingrid Daubechies. *Ten lectures on wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[DD97] Stephan Dahlke and Ronald A. DeVore. Besov regularity for elliptic boundary value problems. *Comm. Partial Differential Equations*, 22(1-2):1–16, 1997.

[DeV98] Ronald A. DeVore. Nonlinear approximation. In *Acta numerica, 1998*, volume 7 of *Acta Numer.*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

[DL93] Ronald A. DeVore and George G. Lorentz. *Constructive approximation*, volume 303 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993.

[JK81] David S. Jerison and Carlos E. Kenig. The Neumann problem on Lipschitz domains. *Bull. Amer. Math. Soc. (N.S.)*, 4(2):203–207, 1981.

[JK95] David Jerison and Carlos E. Kenig. The inhomogeneous Dirichlet problem in Lipschitz domains. *J. Funct. Anal.*, 130(1):161–219, 1995.

[Sau12] Stefan Sauter. Wavelets. Lecture at Universitaet Zuerich in HS2012, http://www.math.uzh.ch/compmath/, 2012.

[Tri10] Hans Triebel. *Theory of function spaces*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, 2010. Reprint of 1983 edition [MR0730762], Also published in 1983 by Birkhäuser Verlag [MR0781540].

[Woj97] P. Wojtaszczyk. *A Mathematical Introduction to Wavelets*. London Mathematical Society Student Texts. Cambridge University Press, 1997.