Pure and Applied Functional Analysis
Volume 8, Number 2, 2023, 457–476

Yokohama Publishers
ISSN 2189-3764 ONLINE JOURNAL

ANISOTROPIC BESOV REGULARITY OF PARABOLIC PDES

STEPHAN DAHLKE* AND CORNELIA SCHNEIDER†

Dedicated to Prof. Dr. Ron DeVore on the occasion of his 80th birthday.

Abstract. This paper is concerned with the regularity of solutions to parabolic evolution equations. Special attention is paid to the smoothness in the specific anisotropic scale $B^{r\mathbf{a}}_{\tau,\tau}$, $\frac{1}{\tau}=\frac{r}{d}+\frac{1}{p}$ of Besov spaces where $\mathbf{a}$ measures the anisotropy. The regularity in these spaces determines the approximation order that can be achieved by fully space-time adaptive approximation schemes. In particular, we show that for the heat equation our results significantly improve [3].

1. Introduction

This paper is concerned with the study of anisotropic Besov regularity estimates for parabolic partial diffenential equations (PDEs). Regularity estimates in Besov spaces are always important since they determine the convergence order of adaptive and other nonlinear constructive approximation schemes for the corresponding, unknown solution. In contrast to this, it is the classical Sobolev smoothness that determines the convergence order of more conventional, uniform schemes. We refer e.g. to DeVore [14] and [7]. For elliptic PDEs, a lot of results in this direction have been achieved in recent years, see [7, 8, 11] and many others. In all these cases, the Besov smoothness was generically higher than the Sobolev regularity which justifies the use of adaptive algorithms. However, most of these results are concerned with *isotropic* Besov estimates which fit perfectly to the stationary character of elliptic partial differential equations. Quite recently, also Besov regularity results for parabolic PDEs have been investigated, see e.g. [10, 23]. In these works, the authors derived time-dependent Besov regularity in space. In particular, they determine the convergence order of space-adaptive numerical schemes such as classical time-marching schemes for parabolic equations. For good reasons, in recent years the development of numerical schemes working on the whole space-time cylinder has become more and more important [24]. In many cases, these schemes are simply more efficient. However, if we take the whole space-time cylinder into account, then

2020 *Mathematics Subject Classification. Primary:* 35B65, 46E35. *Secondary:* 35K05, 65M12.
Key words and phrases. Parabolic PDEs, Lipschitz domains, anisotropic Sobolev spaces, anisotropic Kondratiev spaces, anisotropic Besov spaces, extension operator, heat equation, adaptivity.
*The work of S. Dahlke has been supported by Deutsche Forschungsgemeinschaft (DFG), Grant No. DA 360/22-1.
†The work of C. Schneider has been supported by Deutsche Forschungsgemeinschaft (DFG), Grant No. SCHN 1509/1-2.

anisotropic structures occur, since we have (in the simplest case) one derivative in time but two derivatives in space. Therefore, anisotropic Besov spaces might be reasonable choices for the regularity spaces. First results for the heat equation have been obtained by Aimar and Gomez [3]. The results in this very interesting paper rely on a certain interpolation technique in scales with $p > 1$, which naturally limits the applicability of their approach. Therefore, in this paper, we follow a different line: In the meantime, it has turned out that a very efficient way to establish Besov regularity for the solution to a PDE is first to study the regularity in weighted Sobolev spaces, the so-called Kondratiev spaces [9]. The reason is that very sharp embeddings of Kondratiev spaces into Besov spaces habe been derived. The whole program has, e.g., very efficiently been carried out in [8, 23]. Usually, Kondratiev spaces can be used for a very precise description of the singularities of the solutions. For our purposes, clearly anisotropic Kondratiev spaces are needed. Therefore, in our setting, the following tasks have to be solved:

- Define suitable anisotropic Kondratiev spaces and establish embeddings into anisotropic Besov spaces.
- Establish anisotropic Kondratiev regularity for the problem under consideration.

In our case, we define the anisotropic Kondratiev spaces simply by means of anisotropic weigths, whereas the anisotropic Besov spaces are defined by tensor products of differently scaled wavelets, where the scaling is compatible with the anisotropy. In this setting, the desired embedding is possible and constitutes our first main result Theorem 4.1. Moreover, we show that for the heat equation the regularity problem in anisotropic Kondratiev spaces is solvable. Combining these facts yields our second main result formulated in Theorem 5.3.
This paper is organized as follows: In Section 2 we recall the notation used throughout the paper. Section 3 is dedicated to anisotropic function spaces and their relations. In particular, we deal with anisotropic Sobolev and Besov spaces. Moreover, we introduce anisotropic Kondratiev spaces and study in Section 4 their relations with anisotropic Besov spaces via embeddings. Finally, in Section 5 we use our obtained results in order to investigate the regularity of solutions of the heat equation in anisotropic Besov spaces and compare the outcome with the results from [3].

2. Preliminaries

We collect some notation used throughout the paper. As usual, we denote by $\mathbb{N}$ the set of all natural numbers, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and $\mathbb{R}^d$, $d \in \mathbb{N}$, the d-dimensional real Euclidean space with $|x|$, for $x \in \mathbb{R}^d$, denoting the Euclidean norm of x. By $\mathbb{Z}^d$ we denote the lattice of all points in $\mathbb{R}^d$ with integer components. For $a \in \mathbb{R}$, let $\lfloor a \rfloor$ denote its integer part and $a_+ := \max(a, 0)$.
Moreover, c stands for a generic positive constant which is independent of the main parameters, but its value may change from line to line. The expression $A \lesssim B$ means that $A \leq c\,B$. If $A \lesssim B$ and $B \lesssim A$, then we write $A \sim B$.

Given two quasi-Banach spaces X and Y, we write $X \hookrightarrow Y$ if $X \subset Y$ and the natural embedding is bounded. By supp f we denote the support of the function f. Moreover, $\mathcal{S}(\mathbb{R}^d)$ denotes the Schwartz space of rapidly decreasing functions. The set of distributions on Ω will be denoted by $\mathcal{D}'(\Omega)$, whereas $\mathcal{S}'(\mathbb{R}^d)$ denotes the set of tempered distributions on $\mathbb{R}^d$. The terms *distribution* and *generalized function* will be used synonymously. Furthermore, let $\hat{f}$ stand for the Fourier transform on $\mathcal{S}'(\mathbb{R}^d)$ with inverse $f^\vee$.

For the application of a distribution $u \in \mathcal{D}'(\Omega)$ to a test function $\varphi \in \mathcal{D}(\Omega)$ we write (u, φ). The same notation will be used if $u \in \mathcal{S}'(\mathbb{R}^d)$ and $\varphi \in \mathcal{S}(\mathbb{R}^d)$ (and also for the inner product in $L_2(\Omega)$). For $u \in \mathcal{D}'(\Omega)$ and a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d$, we write $D^\alpha u$ for the α-th *generalized* or *distributional derivative* of u with respect to $x = (x_1, \ldots, x_d) \in \Omega$, i.e., $D^\alpha u$ is a distribution on Ω, uniquely determined by the formula

$$(D^\alpha u, \varphi) := (-1)^{|\alpha|}(u, D^{(\alpha)}\varphi), \qquad \varphi \in \mathcal{D}(\Omega),$$

where $D^{(\alpha)}$ stands for the corresponding classical derivative. In particular, if $u \in L^1_{\mathrm{loc}}(\Omega)$ and there exists a function $v \in L^1_{\mathrm{loc}}(\Omega)$ such that

$$\int_\Omega v(x)\varphi(x)\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega u(x) D^{(\alpha)}\varphi(x)\mathrm{d}x \qquad \text{for all} \qquad \varphi \in \mathcal{D}(\Omega),$$

we say that v is the *α-th weak derivative* of u and write $D^\alpha u = v$. We also use the notation $\frac{\partial^k}{\partial x_j^k} u := D^\beta u$ as well as $D_j^k u := \partial_{x_j^k} u := D^\beta u$, for some multi-index $\beta = (0, \ldots, k, \ldots, 0)$ with $\beta_j = k$, $k \in \mathbb{N}$.

3. Anisotropic function spaces

Compared to classical (isotropic) function spaces, the smoothness properties of an element in an anisotropic function space depend on a chosen direction in $\mathbb{R}^d$. In order to capture this phenomenon, let us fix throughout the paper an *anisotropy* $\mathbf{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d_+$ normalized by

$$\left(\frac{1}{a_1} + \ldots + \frac{1}{a_d}\right) = d. \tag{3.1}$$

Moreover, we denote by

$$|x|_{\mathbf{a}} := \sum_{j=1}^d |x_j|^{a_j}, \qquad x = (x_1, \ldots, x_d) \in \mathbb{R}^d, \tag{3.2}$$

the *anisotropic pseudo-distance* corresponding to $\mathbf{a}$.

3.1. **Anisotropic Sobolev spaces.** Let $\mathcal{O} \subset \mathbb{R}^d$ be a domain, $1 < p < \infty$, and $\boldsymbol{\ell} = (l_1, \ldots, l_d) \in \mathbb{N}_0^d$. Then

(3.3)
$$W_p^{\boldsymbol{\ell}}(\mathcal{O}) = \left\{ f \in L_p(\mathcal{O}) : \ \|f|W_p^{\boldsymbol{\ell}}(\mathcal{O})\| := \|f|L_p(\mathcal{O})\| + \sum_{i=1}^d \left\| \frac{\partial^{l_i} f}{\partial x_i^{l_i}} \Big| L_p(\mathcal{O}) \right\| < \infty \right\}$$

is an anisotropic Sobolev space. If $l_1 = \ldots = l_d = l$, then $W_p^{\boldsymbol{\ell}}(\mathcal{O}) = W_p^l(\mathcal{O})$ is the usual (isotropic Sobolev space). We see that in contrast to the usual Sobolev spaces, the smoothness properties of an element of an anisotropic Sobolev space depend in general on the chosen direction in $\mathbb{R}^d$. For $\boldsymbol{\alpha} = \alpha\mathbf{a}$ with $\alpha \in \mathbb{R}$ and $\mathbf{a}$

as in (3.1), corresponding anisotropic Bessel potential spaces $H^{\boldsymbol{\alpha}}(\mathbb{R}^d)$ $(=H^{\alpha\mathbf{a}}(\mathbb{R}^d))$ can be defined via

$$H^{\alpha\mathbf{a}}(\mathbb{R}^d) := \left\{ f \in L_2(\mathbb{R}^d):\ \|f|H^{\alpha\mathbf{a}}(\mathbb{R}^d)\| := \left\|(1+|\xi|^2_{\mathbf{a}})^{\alpha/2}\hat{f}(\xi)|L_2(\mathbb{R}^d)\right\| < \infty \right\}.$$

Remark 3.1. For the regularity studies in [3] the authors were mainly interested in solutions of the homogeneous heat equation $\partial_t u - \Delta u = 0$. Therefore, special attention was paid to the anisotropic Sobolev spaces $W^{2,1}_p(\Omega)$ normed by

$$\begin{aligned}\left\|u|W^{2,1}_p(\Omega)\right\| := \|u|L_p(\Omega)\| &+ \sum_{i=1}^d \left\| \frac{\partial}{\partial x_i} u \Big| L_p(\Omega)\right\| \\ &+ \sum_{i,j=1}^d \left\| \frac{\partial^2}{\partial x_i \partial x_j} u \Big| L_p(\Omega)\right\| + \left\| \frac{\partial}{\partial t} u \Big| L_p(\Omega)\right\|,\end{aligned}$$

defined on the space-time cylinder $\Omega = D \times [0,T]$, where $D \subset \mathbb{R}^d$ is some Lipschitz domain. In particular, these spaces coincide with our anisotropic Sobolev spaces $W^{\boldsymbol{\ell}}_p(\mathcal{O})$ if we replace $\mathcal{O} \subset \mathbb{R}^d$ by $\Omega \subset \mathbb{R}^{d+1}$ in (3.3) and put $\boldsymbol{\ell} = (2,\dots,2,1) \in \mathbb{N}_0^{d+1}$.

3.2. **Anisotropic Besov spaces, wavelet decompositions.** We first recall the definition of anisotropic Besov spaces on $\Omega \subset \mathbb{R}^d$. Whenever f is a function in Ω, we denote by $\Delta^k_h f$ the difference of order $k \geq 1$ and step $h \in \mathbb{R}^d$, defined iteratively via

$$(\Delta^\Omega_h f)(x) = \begin{Bmatrix} f(x+h) - f(x), & \text{if } x, x+h \in \Omega \\ 0, & \text{otherwise} \end{Bmatrix}$$

$$\text{and} \quad (\Delta^{k,\Omega}_h f)(x) = \Delta^\Omega_h(\Delta^{k-1,\Omega}_h f)(x), \quad x \in \Omega.$$

If $\mathbf{k} = (k_1,\dots,k_d)$ is a multi-index with $k_i \geq 0$, we define the *iterated difference of order* $\mathbf{k}$ by

$$\Delta^{\mathbf{k},\Omega}_h f(x) = \left(\Delta^{k_1,\Omega}_{h_1 e_1} \circ \cdots \circ \Delta^{k_d,\Omega}_{h_d e_d} f\right)(x),$$

where $e_1,\dots,e_d$ denotes the canonical basis of $\mathbb{R}^d$. Moreover, let $\boldsymbol{\alpha} = \alpha\mathbf{a} = (\alpha_1,\dots,\alpha_d)$ with $\alpha > 0$, $\mathbf{a}$ as in (3.1) and let $0 < p, q < \infty$. We say that $f \in L_p(\Omega)$ belongs to the anisotropic Besov space $\mathbf{B}^{\boldsymbol{\alpha}}_{p,q}(\Omega)$ if the semi-norm

$$|f|_{\mathbf{B}^{\boldsymbol{\alpha}}_{p,q}(\Omega)} = \sum_{i=1}^d \left(\int_0^\infty t^{\alpha_i} \|\Delta^{k_i,\Omega}_{te_i} f|L_p(\Omega)\|^q \frac{dt}{t}\right)^{1/q}$$

is finite (here k_i are integers such that $k_i > \alpha_i$, $i = 1,\dots,d$). Moreover, the norms

$$\|f|\mathbf{B}^{\boldsymbol{\alpha}}_{p,q}(\Omega)\| := \|f|L_p(\Omega)\| + |f|_{\mathbf{B}^{\boldsymbol{\alpha}}_{p,q}(\Omega)},$$

are known to be equivalent for any choice $k_i > \alpha_i$. Finally, the isotropic Besov spaces $B^s_{p,q}(\Omega)$ are nothing but $\mathbf{B}^{\boldsymbol{\alpha}}_{p,q}(\Omega)$ if $\boldsymbol{\alpha} = (s,\dots,s)$. For our studies below it will be convenient to use another approach and define anisotropic Besov spaces $B^{\alpha\mathbf{a}}_{p,q}(\Omega)$ via wavelet decompositions, valid for the whole range $0 < p, q < \infty$. In particular, our wavelet approach is based on compactly supported wavelets and a dilation adapted to the anisotropy of the spaces. Such a characterization of anisotropic Besov spaces was developed in [15] with the forerunners [16, 18]. Note that we adapt the results presented there according to our needs. The wavelet system we are looking for will be dilated by a matrix M, where

$$M := \operatorname{diag}\left(\lambda^{1/a_1}, \ldots, \lambda^{1/a_d}\right) \quad \text{for some} \quad \lambda > 1, \tag{3.4}$$

which is 'compatible' with the anisotropy $\mathbf{a}$ in the sense that one recovers the correct homogeneity over Besov semi-norms, i.e.,

$$|\det M|^{1/p} |f(M\cdot)|_{\mathbf{B}^{\alpha\mathbf{a}}_{p,q}} = \lambda^{\alpha} |f|_{\mathbf{B}^{\alpha\mathbf{a}}_{p,q}}.$$

In particular, also with this approach we recover the isotropic Besov spaces $B^s_{p,q}(\Omega)$ based on dyadic dilations by setting $a_1 = \ldots = a_d = 1$, $\alpha = s$, and $\lambda = 2$.

We briefly recall our wavelet approach based on multi-resolution analysis: For our definition of the anisotropic Besov spaces, we will use compactly supported wavelets constituting Riesz-bases in $L_2(\mathbb{R})$ that are obtained by dilating, translating and scaling a fixed function, the so–called *mother wavelet* ψ. This mother wavelet is usually constructed by means of a *multiresolution analysis (MRA)* that is, a sequence $\{V_j\}_{j\in\mathbb{Z}}$ of shift-invariant, closed subspaces of $L_2(\mathbb{R})$ whose union is dense in L_2 while their intersection is trivial. Moreover, all the spaces are related via dilation, and the space V_0 is spanned by the translates of a fixed function ϕ, called the *generator* or *father wavelet.* We put $\psi^0 := \phi$ and $\psi^1 := \psi$ and denote by U the nontrivial vertices of the square $[0,1]^d$. Then by taking tensor products, i.e.,

$$\psi^u(x_1, \ldots, x_d) := \prod_{j=1}^{d} \psi^{u_j}(x_j), \quad u = (u_1, \ldots, u_d) \in U,$$

a compactly supported basis for $L_2(\mathbb{R}^d)$ can be constructed. In contrast to the isotropic case our wavelets are constructed such that they are well adapted to the anisotropy $\mathbf{a}$, which is achieved by using the diagonal dilation Matrix M from (3.4) compatible with $\mathbf{a}$. For this reason we will call them M-wavelets in the sequel.

The existence of compactly supported scaling functions (and wavelets) for an arbitrary dilation matrix M is a delicate matter. Concrete examples when M has a relatively simple form can be found in [4, 17, 19]. However, since we consider tensor products of wavelets the situation simplifies considerably in our context. In this case M is diagonal and we only dilate differently in different directions. Additionally, we may assume that M is integer valued and put $m = |\det M| = \lambda^d$. Note that from the discussion in [16, Sect. 3.3] it follows that this is not a severe restriction in our construction since for all anisotropies $\mathbf{a} \in \mathbb{Q}^d_+$ there exists a number $\lambda > 1$ such that $\lambda^{1/a_1}, \ldots, \lambda^{1/a_d} \in \mathbb{N}$.

We now explain what we call an *admissible biorthogonal M-wavelet bases* in the sequel. For the precise construction we refer to [15, 16]. Let ϕ be a compactly supported scaling function, the *father wavelet*, of tensor product type on $\mathbb{R}^d$ having sufficiently high smoothness and let $\Psi' = \{\psi_i : \ i = 1, \ldots, m-1\}$ be the set containing the corresponding multivariate mother wavelets such that, for a given $L \in \mathbb{N}$ with $L > d/2$ and some $N > 0$ the following requirements hold: For all

$\psi \in \Psi'$,

$$\text{(3.5)} \qquad \operatorname{supp} \phi,\ \operatorname{supp} \psi \ \subset\ [-N, N]^d,$$

$$\text{(3.6)} \qquad \phi,\ \psi\ \in\ H^{L\mathbf{a}}(\mathbb{R}^d),$$

$$\text{(3.7)} \qquad \psi\ \perp\ \Pi_{L-1} := \operatorname{span}\left\{x^{\boldsymbol{\ell}} = x_1^{l_1} \cdots x_d^{l_d} :\ |\boldsymbol{\ell}| = l_1 + \ldots + l_d \le L-1\right\}.$$

In particular, (3.7) tells us that the mother wavelets ψ are orthogonal to the polynomials Π_{L-1} of order less than L, which is possible by the assumptions (3.5) and (3.6), cf. [16, Prop. 3.3]. Moreover, by $\mathcal{D}^+$ we denote the set of all cuboids in $\mathbb{R}^d$ with measure at most 1 of the form

$$\mathcal{D}^+ := \left\{I \subset \mathbb{R}^d :\ I = M^{-j}([0,1]^d + k),\ j \in \mathbb{N}_0,\ k \in \mathbb{Z}^d\right\}$$

and we set $\mathcal{D}_j := \{I \in \mathcal{D}^+ :\ |I| = \lambda^{-jd}\}$. For the shifts and dilations of the father wavelet and the corresponding wavelets we use the abbreviations

$$\text{(3.8)} \qquad \phi_k(x) := \phi(x-k) \quad \text{and} \quad \psi_I(x) := |\det M|^{j/2}\psi(M^j x - k),$$

where $j \in \mathbb{N}_0,\ k \in \mathbb{Z}^d,\ \psi \in \Psi'$, and $I = M^{-j}([0,1]^d + k) \in \mathcal{D}^+$. It follows that

$$\left\{\phi_k,\ \psi_I :\ k \in \mathbb{Z}^d,\ I \in \mathcal{D}^+,\ \psi \in \Psi'\right\}$$

is a Riesz basis in $L_2(\mathbb{R}^d)$. Furthermore, we assume that there exists a dual basis also constructed by means of an MRA $\{\tilde{V}_j\}_{j\in\mathbb{Z}}$, i.e., functions $\tilde{\phi}$ and $\tilde{\psi} \in \tilde{\Psi}' = \{\tilde{\psi}_i : i = 1, \ldots, m-1\}$ satisfying

$$\text{(3.9)} \qquad \langle \tilde{\phi}_k, \psi_I\rangle = \langle \tilde{\psi}_I, \phi_k\rangle = 0,$$

$$\text{(3.10)} \qquad \langle \tilde{\phi}_k, \phi_l\rangle = \delta_{k,l} \qquad \text{(Kronecker symbol)},$$

$$\text{(3.11)} \qquad \langle \tilde{\psi}_I, \psi_{I'}\rangle = \delta_{I,I'}.$$

The dual Riesz basis should fulfil the same requirements as the primal Riesz basis, i.e.,

$$\text{(3.12)} \qquad \operatorname{supp} \tilde{\phi},\ \operatorname{supp} \tilde{\psi} \subset [-N, N]^d,$$

$$\text{(3.13)} \qquad \tilde{\phi},\ \tilde{\psi} \in H^{L\mathbf{a}}(\mathbb{R}^d),$$

$$\text{(3.14)} \qquad \tilde{\psi} \perp \Pi_{L-1}.$$

Denote by $Q(I)$ some cuboid (of minimal size) such that $\operatorname{supp} \psi_I \subset Q(I)$ for every $\psi \in \Psi'$. Then we may assume that $Q(I) = M^{-j}k + M^{-j}Q$ for some cuboid Q. Put $\Lambda' = \mathcal{D}^+ \times \Psi'$. Then every function $f \in L_2(\mathbb{R}^d)$ can be written as

$$f = \sum_{k\in\mathbb{Z}^d} \langle f, \tilde{\phi}_k\rangle \phi_k + \sum_{(I,\psi)\in\Lambda'} \langle f, \tilde{\psi}_I\rangle \psi_I.$$

It will be convenient to include ϕ into the set Ψ'. We use the notation $\phi_I := 0$ for $|I| < 1$, $\phi_I = \phi(\cdot - k)$ for $I = k + [0,1]^d$, and can simply write

$$\text{(3.15)} \qquad f = \sum_{(I,\psi)\in\Lambda} \langle f, \tilde{\psi}_I\rangle \psi_I, \qquad \Lambda = \mathcal{D}^+ \times \Psi, \quad \Psi = \Psi' \cup \{\phi\}.$$

The two systems $\{\phi_k, \psi_I\}_{k,I}$ and $\{\tilde{\phi}_k, \tilde{\psi}_I\}_{k,I}$ constructed as above are said to be a *pair of admissible biorthogonal M-wavelet bases* and they may be used to obtain decompositions of many classical function spaces. In particular, according to [15, Thm. 1.2] and [16, Thm. 1.2] anisotropic Besov spaces on $\mathbb{R}^d$ can be characterized by decay properties of the wavelet coefficients, if the parameters fulfill certain conditions. This characterization motivates the following definition.

Definition 3.2 (Anisotropic Besov spaces, wavelet decompositions)**.**
Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d_+$ and $0 < p, q < \infty$. Moreover, let $\boldsymbol{\alpha} = \alpha\mathbf{a}$, where $\alpha > \max\{0, d(1/p - 1)\}$ and the anisotropy $\mathbf{a}$ is normalized as in (3.1). Let M be a dilation matrix compatible with $\mathbf{a}$. We assume that $\{\phi_k, \psi_I\}_{k,I}$ and $\{\tilde{\phi}_k, \tilde{\psi}_I\}_{k,I}$ is a pair of biorthogonal admissible M-wavelet bases with $\phi, \tilde{\phi}, \psi, \tilde{\psi} \in H^{L\mathbf{a}}(\mathbb{R}^n)$ for some integer $L > \max\{d/2, \alpha_1, \ldots, \alpha_n\}$. Then the Besov space $B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d)$ $(= B^{\boldsymbol{\alpha}}_{p,q}(\mathbb{R}^d))$ is defined as the set of all functions $f \in L_p(\mathbb{R}^d)$ satisfying

$$f = \sum_{k\in\mathbb{Z}^d} \langle f, \tilde{\phi}_k\rangle \phi_k + \sum_{(I,\psi)\in\Lambda'} \langle f, \tilde{\psi}_I\rangle \psi_I \tag{3.16}$$

(convergence in $\mathcal{S}'(\mathbb{R}^d)$) with

$$\|f|B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d)\| \sim \left(\sum_{k\in\mathbb{Z}^d} |\langle f, \tilde{\phi}_k\rangle|^p\right)^{1/p} +$$

$$\left(\sum_{j=0}^{\infty} |\det M|^{j\left(\frac{\alpha}{d}+\left(\frac{1}{2}-\frac{1}{p}\right)\right)q} \left(\sum_{(I,\psi)\in\mathcal{D}_j\times\Psi'} |\langle f, \tilde{\psi}_I\rangle|^p\right)^{q/p}\right)^{1/q} < \infty. \tag{3.17}$$

Remark 3.3. (i) In particular, for the adaptivity scale $B^{\alpha\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d)$ with $\alpha = d\left(\frac{1}{\tau} - \frac{1}{p}\right)$, we see that the quasi-norm (3.17) becomes

$$\|f|B^{\alpha\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d)\| \sim \left(\sum_{k\in\mathbb{Z}^d} |\langle f, \phi_k\rangle|^\tau\right)^{1/\tau} + \left(\sum_{j=0}^{\infty} |\det M|^{j\left(\frac{1}{2}-\frac{1}{p}\right)\tau} \sum_{(I,\psi)\in\mathcal{D}_j\times\Psi'} |\langle f, \psi_I\rangle|^\tau\right)^{1/\tau}.$$

(ii) From [16, Thm. 1.2] we deduce that $\mathbf{B}^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d) = B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d)$ for the range of parameters

$$\alpha > 0, \quad 1 \le p, q < \infty,$$

whereas [15, Thm. 1.2] additionally covers the case

$$\mathbf{B}^{\alpha\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d) = B^{\alpha\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d), \qquad \alpha > \max\left\{0, d\left(\frac{1}{\tau} - 1, 0\right)\right\}, \quad \frac{1}{\tau} = \frac{\alpha}{d} + \frac{1}{p}, \quad 0 < \tau < \infty.$$

Thus, we see that the range of spaces we consider in Definition 3.2 is larger. The restriction $\alpha > \max\{0, d(1/p - 1)\}$ is necessary since it guarantees that our anisotropic Besov spaces considered in Definition 3.2 satisfy $B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d) \hookrightarrow L_{\max\{1+\varepsilon,p\}}(\mathbb{R}^d)$, see also [15, Cor. 5.4] in this context.

(iii) Interpretation: From the above construction of the anisotropic spaces we see that α describes mean smoothness and $\mathbf{a}$ measures the anisotropy.

(iv) If $p = q = 2$, then $B^{\alpha\mathbf{a}}_{2,2}(\mathbb{R}^d)$ coincides with the anisotropic Bessel potential space, i.e., we have

$$(\mathbf{B}^{\alpha\mathbf{a}}_{2,2}(\mathbb{R}^d) =)\ B^{\alpha\mathbf{a}}_{2,2}(\mathbb{R}^d) = H^{\alpha\mathbf{a}}(\mathbb{R}^d).$$

Furthermore, if $\alpha\mathbf{a} = \boldsymbol{\ell} = (l_1, \ldots, l_d)$ is an integer-valued multi-index, one recovers the anisotropic Sobolev spaces

$$(\mathbf{B}^{\boldsymbol{\ell}}_{2,2}(\mathbb{R}^d) =)\ B^{\boldsymbol{\ell}}_{2,2}(\mathbb{R}^d) = H^{\boldsymbol{\ell}}(\mathbb{R}^d) = W^{\boldsymbol{\ell}}_2(\mathbb{R}^d).$$

(v) As already mentioned before, we adapted the results from [15, 16, Thms. 1.2] slightly. In particular, we consider wavelets which form a biorthogonal basis for $L_2(\mathbb{R}^d)$ instead of $L_p(\mathbb{R}^d)$ which leads to the weight factor $|\det M|^{j\left(\frac{\alpha}{d}+(\frac{1}{2}-\frac{1}{p})\right)}$ regarding the decay of the wavelet coefficients $\langle f, \tilde{\psi}_I\rangle$ in (3.17) instead of $|\det M|^{\frac{j\alpha}{d}}$ in [16, Thm. 1.2].
Moreover, the additional condition that $\phi, \tilde{\phi} \in H^{L\mathbf{a}}(\mathbb{R}^d) \cap B^{\alpha_0\mathbf{a}}_{p,q}(\mathbb{R}^d)$ for some $\alpha_0 > 0$ can be circumvented by choosing L large enough. This can be seen as follows: Since $\phi, \tilde{\phi}$ have compact support we deduce from the the definition of the spaces that for $p < 2$, $0 < q \leq \infty$, and $L > \alpha_0$,

$$\phi, \tilde{\phi} \in H^{L\mathbf{a}}(\mathbb{R}^d) = B^{L\mathbf{a}}_{2,2}(\mathbb{R}^d) \hookrightarrow B^{\alpha_0\mathbf{a}}_{p,q}(\mathbb{R}^d).$$

For the case that $2 \leq p$ we use [6, Thm. 18.4] and obtain

$$H^{L\mathbf{a}}(\mathbb{R}^d) = B^{L\mathbf{a}}_{2,2}(\mathbb{R}^d) \hookrightarrow B^{\alpha_0\mathbf{a}}_{p,q}(\mathbb{R}^d) \quad \text{if} \quad L > \alpha_0 + d\left(\frac{1}{2} - \frac{1}{p}\right).$$

(vi) In particular, if $1 \leq p, q < \infty$ our anisotropic Besov spaces $B^{\alpha\mathbf{a}}_{p,q}$ coincide with the spaces from [15, 16]. Therefore, by [15, Cor. 5.3] we have the following interpolation result:

$$\big(L_p(\mathbb{R}^d), B^{\alpha\mathbf{a}}_{p,r}(\mathbb{R}^d)\big)_{\theta,q} = B^{\theta\mathbf{a}}_{p,q}(\mathbb{R}^d), \qquad 0 < \theta < 1, \quad 0 < r < \infty. \tag{3.18}$$

Corresponding function spaces on domains $\mathcal{O} \subset \mathbb{R}^d$ can be introduced via restriction, i.e.,

$$\begin{aligned} B^{\alpha\mathbf{a}}_{p,q}(\mathcal{O}) &= \left\{ f \in \mathcal{D}'(\mathcal{O}) :\ \exists g \in B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d),\ g\big|_{\mathcal{O}} = f \right\}, \\ \|f|B^{\alpha\mathbf{a}}_{p,q}(\mathcal{O})\| &= \inf_{g|_{\mathcal{O}}=f} \|f|B^{\alpha\mathbf{a}}_{p,q}(\mathbb{R}^d)\|. \end{aligned}$$

3.3. **Domains allowing extensions.** In what follows we want to investigate anisotropic function spaces on more general domains $\Omega \subset \mathbb{R}^d$. So far we introduced anisotropic spaces on $\mathbb{R}^d$, where a lot of the tools we need (in particular, wavelet decompositions of anisotropic Besov spaces) are available. Then corresponding spaces on domains can be defined via restriction. Now, in order to truly establish our results on domains Ω, we need an extension operator for our anisotropic spaces. Such extensions of anisotropic spaces defined on Ω to the whole $\mathbb{R}^d$ are possible if Ω satisfies what is called a strong $\mathbf{r}$-horn condition. This is the proper counterpart of the cone-conditions one is familiar with from the isotropic setting. In order to explain this condition we need some notation. Let $\mathbf{r} = (r_1, \ldots, r_d)$ be a vector with positive components. Suppose that $0 < \rho \leq \infty$, $\varepsilon > 0$, and $\mathbf{b} \in \mathbb{R}^d$ with $b_i \neq 0$ for $i = 1, \ldots, d$. The set

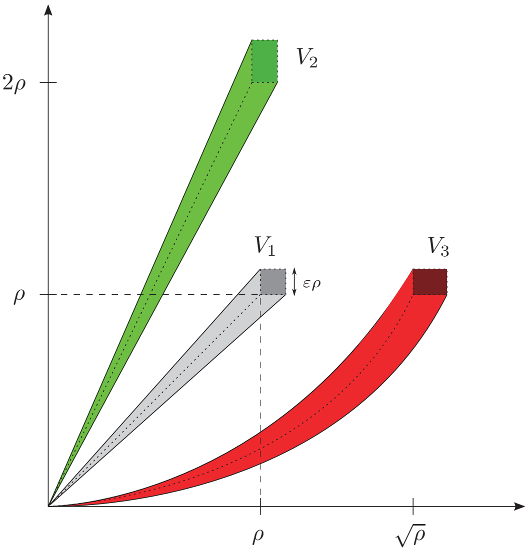
$$V(\mathbf{r}, \rho, \varepsilon, \mathbf{b}) := \bigcup_{0<\nu<\rho} \left\{ x : \frac{x_i}{b_i} > 0,\ \nu < \left(\frac{x_i}{b_i}\right)^{r_i} < (1+\varepsilon)\nu \quad \text{for} \quad i = 1, \ldots, d \right\}$$

is called an $\mathbf{r}$*-horn of radius* ρ *and opening* ε.

The diagram aside illustrates $\mathbf{r}$-horns for different parameters $\mathbf{r}$ and $\mathbf{b}$ in $\mathbb{R}^2$:

$$\begin{aligned} V_1 &= V\big((1,1), \rho, \varepsilon, (1,1)\big); \\ V_2 &= V\big((1,1), \rho, \varepsilon, (1,2)\big); \\ V_3 &= V\big((2,1), \rho, \varepsilon, (1,1)\big). \end{aligned}$$

In particular, we see that in the isotropic case (see V_1 and V_2 with $r_1 = r_2 = 1$) the horn is just a cone. If we have an anisotropy (see V_3 with $r_1 = 2$ and $r_2 = 1$) the different scaling exponent in the different directions causes the cone to become a horn. Moreover, the vector $\mathbf{b}$ specifies the exact location of the $\mathbf{r}$-horn in the coordinate system (compare V_1 with $\mathbf{b} = (1,1)$ with V_2 where $\mathbf{b} = (1,2)$).



An open set $\Omega \subset \mathbb{R}^d$ is said to satisfy a *weak* $\mathbf{r}$*-horn condition* if there is a positive integer N such that for each $j \in \{1, \ldots, N\}$, there are open sets Ω_j and an $\mathbf{r}$-horn $V_j(\mathbf{r}, \rho, \varepsilon, \mathbf{b}^{(j)})$ such that

$$\Omega = \bigcup_{j=1}^{N} \Omega_j = \bigcup_{j=1}^{N} \Big(\Omega_j + V_j(\mathbf{r}, \rho, \varepsilon, \mathbf{b}^{(j)}) \Big). \tag{3.19}$$

The relation (3.19) expresses the fact that for any point $x \in \Omega_j$, if the horn $V_j(\mathbf{r}, \rho, \varepsilon, \mathbf{b}^{(j)})$ is shifted parallel to itself in such a way that its vertex coincides with x, then the resulting shifted horn lies in Ω. If, in addition, there exists $\delta > 0$ such that

$$\Omega = \bigcup_{j=1}^{N} \Omega_j^{(\delta)}, \quad \text{where} \quad \Omega_j^{(\delta)} = \{x \in \Omega_j : \ \operatorname{dist}(x, \Omega \setminus \Omega_j) > \delta\},$$

then Ω is said to satisfy a *strong* $\mathbf{r}$*-horn* condition. Note that if $r_1 = r_2 = \ldots = r_d$, every $\mathbf{r}$-horn is a cone. It is possible in this case to show, cf. [21, p. 382], that the concept of a domain having a Lipschitz boundary coincides with the concept of a domain satisfying the $\mathbf{r}$-horn condition.

The following theorem can be found in [5, Thm. 9.6] and [21, Thm. 2, p. 382].

Theorem 3.4. *Suppose $\Omega \subset \mathbb{R}^d$ satisfies a strong* $\mathbf{r}$*-horn condition and $1 \leq p, q \leq \infty$.*

(i) Let $1 < p < \infty$. Then $W_p^{\mathbf{r}}(\Omega)$ is the set of all functions which are the restrictions to Ω of elements of $W_p^{\mathbf{r}}(\mathbb{R}^d)$. In particular, there is a bounded, linear extension map $E : W_p^{\mathbf{r}}(\Omega) \to W_p^{\mathbf{r}}(\mathbb{R}^d)$.

(ii) Furthermore, $\mathbf{B}_{p,q}^{\mathbf{r}}(\Omega)$ is the set of all functions which are the restrictions to Ω of elements of $\mathbf{B}_{p,q}^{\mathbf{r}}(\mathbb{R}^d)$. In particular, there is a bounded, linear extension map $E : \mathbf{B}_{p,q}^{\mathbf{r}}(\Omega) \to \mathbf{B}_{p,q}^{\mathbf{r}}(\mathbb{R}^d)$.

In view of Remark 3.3(ii) we immediately obtain the following result.

Corollary 3.5. *Suppose $\Omega \subset \mathbb{R}^d$ satisfies a strong* $\mathbf{r}$*-horn condition and $1 \leq p, q < \infty$. Then there is a bounded, linear extension map $E : B_{p,q}^{\mathbf{r}}(\Omega) \to B_{p,q}^{\mathbf{r}}(\mathbb{R}^d)$.*

Remark 3.6. We provide examples of domains satisfying the (weak or strong) $\mathbf{r}$-horn condition: For $d = 2$, $\Omega = \mathbb{R}^2$ as well as the rectangular parallelepiped

$$\Omega = \{(x_1, x_2) : \ |x_1| < a, \ |x_2| < b\},$$

where $a, b > 0$, satisfy the strong $\mathbf{r}$-horn condition for any $\mathbf{r}$. Moreover, the disk

$$Q = \left\{(x_1, x_2) : \ x_1^2 + x_2^2 < 1\right\}$$

satisfies the weak $\mathbf{r}$-horn condition only if $\frac{1}{2} r_1 \leq r_2 \leq 2 r_1$ and the strong $\mathbf{r}$-horn condition only if $r_1 = r_2$.
Since in the isotropic case, a Lipschitz domain satisfies the strong $\mathbf{r}$-horn condition for any $\mathbf{r} = (r, \ldots, r)$ from the product structure of the space-time cylinder $\Omega = D \times [0, T] \subset \mathbb{R}^{d+1}$, where $D \subset \mathbb{R}^d$ is a bounded Lipschitz domain, we deduce that Ω satisfies the strong $\mathbf{r}$-horn condition for any $\mathbf{r}$ of the form $\mathbf{r} = (r, \ldots, r, \frac{r}{2})$, where $r > 0$.

Remark 3.7. The extension operator from Theorem 3.4 allows us to transfer many results (such as embeddings, interpolation, etc.), which are known for anisotropic spaces on $\mathbb{R}^d$, to domains satisfying the horn condition. In particular, it allows us to relate the regularity spaces

$$\mathbb{B}_p^s(\Omega) := \left(L_p(\Omega), W_p^{2,1}(\Omega)\right)_{\frac{s}{2}, p}, \qquad 0 < s < 1, \quad 1 < p < \infty,$$

appearing in [3] (which for general $s > 0$ can be defined via the action of the derivatives ∂_t and $\partial_{x_j x_i}$) to our spaces: According to [6, Thm. 18.9] for any $\Omega \subset \mathbb{R}^d$ satisfying an $\boldsymbol{\ell}$-horn condition we have the embedding

$$B_{p,\min(p,2)}^{\boldsymbol{\ell}}(\Omega) \hookrightarrow W_p^{\boldsymbol{\ell}}(\Omega) \hookrightarrow B_{p,\max(p,2)}^{\boldsymbol{\ell}}(\Omega). \tag{3.20}$$

Since the space-time cylinder $\Omega = D \times [0, T]$ satisfies the $\boldsymbol{\ell}$-horn condition for arbitrary $\boldsymbol{\ell}$ we deduce from (3.20) and (3.18) that

$$\mathbb{B}_p^s(\Omega) = B_{p,p}^{s,\ldots,s,\frac{s}{2}}(\Omega) = B_{p,p}^{\tilde{s}\mathbf{a}}(\Omega), \tag{3.21}$$

using the anisotropy

$$\mathbf{a} = (a_1, \ldots, a_{d+1}) = \left(\frac{d+2}{d}, \ldots, \frac{d+2}{d}, \frac{1}{2}\frac{d+2}{d}\right) = \frac{d+2}{d}\left(1, \ldots, 1, \frac{1}{2}\right) \tag{3.22}$$

together with the mean smoothness $\tilde{s} = \frac{sd}{d+2}$. Moreover, choosing $s = p = q = 2$ and $\tilde{s} = \frac{2d}{d+2}$ yields the special case

$$W^{2\dots,2,1}(\Omega) = B^{\frac{2d}{d+2}\mathbf{a}}_{2,2}(\Omega). \tag{3.23}$$

3.4. **Anisotropic Kondratiev spaces.** In this section, we introduce anisotropic Kondratiev spaces, which are special weighted anisotropic Sobolev spaces. The corresponding isotropic spaces play a central role in the regularity theory for elliptic PDEs on domains with piecewise smooth boundary, particularly polygons (2D) and polyhedra (3D). For a systematic treatment and further references we refer to [9]. In particular, in the isotropic case the weight is often chosen to be a power of the distance to the *singular set* of the boundary of a domain $\mathcal{O} \subset \mathbb{R}^d$, i.e., the set of all points $x \in \partial\mathcal{O}$ for which for any $\varepsilon > 0$ the set $\partial\mathcal{O} \cap B_\varepsilon(x)$ is not smooth (here $B_\varepsilon(x)$ denotes the open ball in $\mathbb{R}^d$ around a point x with radius $\varepsilon > 0$).
We adapt this idea and define now corresponding anisotropic Kondratiev spaces using weights which constitute powers of the anisotropic distance based on (3.2) to a singular set $S \subset \partial\mathcal{O}$.

Precisely, let $\mathcal{O} \subset \mathbb{R}^d$ be a domain and let S be a nontrivial closed subset of its boundary $\partial\mathcal{O}$. Furthermore, let $1 \le p < \infty$, $\mathbf{m}= (m_1, \dots, m_d) = m\mathbf{a} \in \mathbb{N}_0^d$ where the anisotropy $\mathbf{a} = (a_1, \dots, a_d)$ is normalized as in (3.1), and $\gamma \in \mathbb{R}$. Then the anisotropic Kondratiev space $\mathcal{K}^{\mathbf{m}}_{p,\gamma}(\mathcal{O})$ (=$\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\mathcal{O})$) is the collection of all $u \in \mathcal{D}'(\Omega)$ such that

$$\|u|\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\mathcal{O})\| := \left(\sum_{i=1}^d \sum_{\alpha_i \le m_i} \int_{\mathcal{O}} \left|(\rho_{\mathbf{a}}(x))^{m-\gamma} D_i^{\alpha_i} u(x)\right|^p \mathrm{d}x\right)^{1/p} < \infty,$$

where $\rho_{\mathbf{a}}(x) = \min(1, \mathrm{dist}_{\mathbf{a}}(x, S))$ and $\mathrm{dist}_{\mathbf{a}}$ denotes the anisotropic distance to $S \subset \partial\Omega$, i.e.,

$$\mathrm{dist}_{\mathbf{a}}(x, S) = \inf_{y \in S} |x - y|_{\mathbf{a}} \qquad \text{with} \qquad |x - y|_{\mathbf{a}} = \sum_{i=1}^d |x_i - y_i|^{a_i}.$$

Remark 3.8. Later on we want to compare our results with the ones obtained in [3] on the space-time cylinder $\Omega = D \times [0, T]$. In this context we remark that the weight appearing in the gradient estimates in [3, Thm. 4] is comparable to our weight $\rho_{\mathbf{a}}(x)$: It is (also) based on powers of the so-called parabolic distance $\delta(x, t)$, which is a special anisotropic distance to the parabolic boundary

$$S := \partial_{\mathrm{par}}\Omega := (D \times \{0\}) \cup (\partial D \times [0, T]).$$

To be precise, for $(x, t) \in \Omega = D \times [0, T]$ it is defined as

$$\delta(x, t) := \inf\{\rho((x, t), (y, s)) : \ (y, s) \in \partial_{\mathrm{par}}\Omega\}, \quad \rho((x, t), (y, s)) \sim |x-y| + \sqrt{|t - s|}.$$

Thus, for the special anisotropy (3.22) we see that

$$|(x, t) - (y, x)|_{\mathbf{a}} = \sum_{i=1}^d |x_i - y_i|^{a_i} + |t - s|^{a_{d+1}} \sim \left(\sum_{i=1}^d |x_i - y_i| + \sqrt{|t - s|}\right)^{\frac{d+2}{d}},$$

which yields

$$\rho_{\mathbf{a}}(x,t) \sim \big(\delta(x,t)\big)^{\frac{d+2}{d}}. \tag{3.24}$$

4. Embeddings between anisotropic Kondratiev and Besov spaces

Theorem 4.1 (Embeddings between Kondratiev and Besov spaces). *Let $\mathbf{m} = m\mathbf{a} \in \mathbb{N}^d$, where the anisotropy $\mathbf{a}$ is normalized as in (3.1) and $\mathbf{s} = s\mathbf{a}$, $\mathbf{r} = r\mathbf{a} \in \mathbb{R}^d_+$. Moreover, assume that the bounded domain $\Omega \subset \mathbb{R}^d$ satisfies the strong $\mathbf{s}$-horn condition. Then we have a continuous embedding*

$$\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega) \cap B^{s\mathbf{a}}_{p,p}(\Omega) \hookrightarrow B^{r\mathbf{a}}_{\tau,\tau}(\Omega), \qquad \frac{1}{\tau} = \frac{r}{d} + \frac{1}{p}, \qquad 1 < p < \infty, \tag{4.1}$$

for all $0 \le r < \min(m, \frac{sd}{d-1})$ and $\gamma > \frac{\delta}{d} r$, where δ denotes the dimension of the singularity set $S \subset \partial\Omega$.

Proof. Since for $r = 0$ the result is clear, we assume in the sequel that $r > 0$ and $0 < \tau < p$.

Step 1. The proof is based on the wavelet characterization of Besov spaces from Definition 3.2. Since our domain Ω satisfies the $\mathbf{s}$-horn condition, according to Corollary 3.5 we can extend every $u \in B^{s\mathbf{a}}_{p,p}(\Omega)$ to some function $\tilde{u} = Eu \in B^{s\mathbf{a}}_{p,p}(\mathbb{R}^d)$. From this we deduce that in order to establish embedding (4.1) it is ultimately enough to show

$$\left(\sum_{(I,\psi)\in\Lambda} |I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^{\tau} \right)^{1/\tau} \lesssim \max\{\|u|\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega)\|, \|u|B^{s\mathbf{a}}_{p,p}(\Omega)\|\}. \tag{4.2}$$

Let us give some further explanations here. We may extend the solution u to a function $\tilde{u} = E(u)$ on the whole Euclidean plane with the additional property that $\langle E(u), \tilde{\psi}_I\rangle = 0$ whenever $Q(I) \cap \Omega = \emptyset$. Then on Ω we have $u = \sum_{(I,\psi)\in\Lambda} \langle \tilde{u}, \tilde{\psi}_I\rangle \psi_I$. Therefore, if we can show that the expression on the right-hand side is contained in the Besov space $B^{r\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d)$, the same is true for its restriction to Ω. To prove this, we use Definition 3.2. Moreover, we see that the first term there which reads as

$$\sum_{k\in\mathbb{Z}^d} \langle \tilde{u}, \tilde{\phi}(\cdot - k)\rangle \phi(\cdot - k)$$

(and also emerges in (3.17)) can be incorporated in the second term (see the formulation (3.15)) and therefore does not appear on the left hand side of (4.2). This is caused by the fact that ϕ shares the same smoothness and support properties as the wavelets ψ_I for $|I| = 1$ (note that below the vanishing moments of ψ_I only become relevant for $|I| < 1$). Therefore, the coefficients $\langle \tilde{u}, \tilde{\phi}(\cdot - k)\rangle$ are incorporated in our considerations since they can be treated exactly like any of the coefficients $\langle \tilde{u}, \tilde{\psi}_I\rangle$ in Step 2.

Step 2. For our analysis we split the index set Λ as follows. For $j \in \mathbb{N}_0$ the refinement level j is denoted by

$$\Lambda_j := \{(I, \Psi) \in \Lambda : \ |I| = |\det M|^{-j} = \lambda^{-jd}\}.$$

Furthermore, for $k \in \mathbb{N}_0$ put

$$\Lambda_{j,k} := \{(I,\psi) \in \Lambda_j : \ k\lambda^{-j} \le \rho_{I,\mathbf{a}} < (k+1)\lambda^{-j}, \ Q(I) \subset \Omega\},$$

where

$$\rho_{I,\mathbf{a}} = \inf_{x \in Q(I)} \rho_{\mathbf{a}} = \inf_{x \in Q(I), y \in S} |x-y|_{\mathbf{a}}.$$

In particular, we have $\Lambda_j = \bigcup_{k=0}^{\infty} \Lambda_{j,k}$ and $\Lambda = \bigcup_{j=0}^{\infty} \Lambda_j$.
We consider first the situation $\rho_{I,\mathbf{a}} > 0$ corresponding to $k \ge 1$ and therefore put $\Lambda_j^0 = \bigcup_{k \ge 1} \Lambda_{k,j}$. Moreover, we require $Q(I) \subset \Omega$. Recall the anisotropic version of Whitney's estimate regarding approximation with polynomials from [22, Lem. 2.1], which states that for every I there exists a polyomial $P_I \in \Pi_{\tilde{L}-1}$, where $|\mathbf{m}| = m_1 + \ldots + m_d \le \tilde{L} - 1$, such that

$$\begin{aligned} \|\tilde{u} - P_I|L_p(Q(I))\| &\lesssim \sum_{i=1}^{d} \lambda^{-j\frac{m_i}{a_i}} \|D_i^{m_i}\tilde{u}|L_p(Q(I))\| \\ &= \lambda^{-jm} \sum_{i=1}^{d} \|D_i^{m_i}\tilde{u}|L_p(Q(I))\| \lesssim |I|^{m/d} |\tilde{u}|_{W_p^{\mathbf{m}}(Q(I))}, \end{aligned}$$

where the omitted constant is independent of I and u. Here we used the fact that $\mathbf{m} = m\mathbf{a}$, i.e., $m = \frac{m_i}{a_i}$ for all $i = 1, \ldots, d$, and put

$$|\tilde{u}|_{W_p^{\mathbf{m}}(Q(I))} := \left(\sum_{i=1}^{d} \int_{Q(I)} |D_i^{m_i}\tilde{u}(x)|^p \mathrm{d}x \right)^{1/p}.$$

Note that $\tilde{\psi}_I$ can be chosen to satisfy moment conditions up to any order, we deduce that it is orthogonal to any polynomial $P_I \in \Pi_{\tilde{L}-1}$. Thus, using Hölder's inequality with $p > 1$ we estimate

$$\begin{aligned} |\langle \tilde{u}, \tilde{\psi}_I \rangle| &= |\langle \tilde{u} - P_I, \tilde{\psi}_I \rangle| \le \|\tilde{u} - P_I|L_p(Q(I))\| \cdot \|\tilde{\psi}_I|L_{p'}(Q(I))\| \\ &\lesssim |I|^{m/d} |\tilde{u}|_{W_p^{\mathbf{m}}(Q(I))} |I|^{\frac{1}{2}-\frac{1}{p}} \\ &\le |I|^{\frac{m}{d}+\frac{1}{2}-\frac{1}{p}} \rho_{I,\mathbf{a}}^{\gamma-m} \left(\sum_{i=1}^{d} \int_{Q(I)} |\rho_{\mathbf{a}}(x)|^{m-\gamma} D_i^{m_i}\tilde{u}(x)|^p \mathrm{d}x \right)^{1/p} \\ &=: |I|^{\frac{m}{d}+\frac{1}{2}-\frac{1}{p}} \rho_{I,\mathbf{a}}^{\gamma-m} \mu_{I,\mathbf{a}}. \end{aligned} \tag{4.3}$$

Note that in the third step we use that the values of $\rho_{I,\mathbf{a}}$ and $\rho_{\mathbf{a}}$ are comparable, i.e., $\rho_{I,\mathbf{a}} \sim \sup_{x \in Q(I)} \rho_{\mathbf{a}}$, since for $k \ge 1$ we consider cuboids for which distance $Q(I)$ to the singular set S is comparable to the sidelength of $Q(I)$ ($k \ge 1$ guarantees $\mathrm{dist}_{\mathbf{a}}(Q(I), S) \gtrsim l(Q(I))$). On the refinement level j, using Hölder's inequality with

$\frac{p}{\tau} > 1$, we find

$$\begin{aligned}
&\sum_{(I,\psi)\in\Lambda_j^0} |I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^\tau \\
&\qquad \le \sum_{(I,\psi)\in\Lambda_j^0} \left(|I|^{\frac{m}{d}} \rho_{I,\mathbf{a}}^{\gamma-m} \mu_{I,\mathbf{a}}\right)^\tau \\
&\qquad \lesssim \left(\sum_{(I,\psi)\in\Lambda_j^0} \left(|I|^{\frac{m}{d}\tau} \rho_{I,\mathbf{a}}^{(\gamma-m)\tau}\right)^{\frac{p}{p-\tau}}\right)^{\frac{p-\tau}{p}} \left(\sum_{(I,\psi)\in\Lambda_j^0} \mu_{I,\mathbf{a}}^p\right)^{\tau/p}.
\end{aligned}$$

For the second factor we observe that there is a controlled overlap between the cuboids $Q(I)$, meaning each $x \in \Omega$ is contained in a finite number of cuboids independent of x, such that we get

$$\begin{aligned}
\left(\sum_{(I,\psi)\in\Lambda_j^0} \mu_{I,\mathbf{a}}^p\right)^{1/p} &= \left(\sum_{(I,\psi)\in\Lambda_j^0} \sum_{i=1}^d \int_{Q(I)} |\rho_{\mathbf{a}}^{m-\gamma}(x) D_i^{m_i} \tilde{u}(x)|^p \mathrm{d}x\right)^{1/p} \\
&\lesssim \left(\sum_{i=1}^d \int_\Omega |\rho_{\mathbf{a}}^{m-\gamma}(x) D_i^{m_i} \tilde{u}(x)|^p \mathrm{d}x\right)^{1/p} \le \|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|.
\end{aligned}$$

For the first factor, by choice of $\rho_{\mathbf{a}}$ we always have $\rho_{I,\mathbf{a}} \le 1$, hence the index k is at most λ^j for the sets $\Lambda_{j,k}$ to be non-empty. The number of elements in $\Lambda_{j,k}$ is bounded by $k^{d-1-\delta}\lambda^{j\delta}$. With this we find

$$\begin{aligned}
&\left(\sum_{(I,\psi)\in\Lambda_j^0} \left(|I|^{\frac{m}{d}\tau} \rho_{I,\mathbf{a}}^{(\gamma-m)\tau}\right)^{\frac{p}{p-\tau}}\right)^{\frac{p-\tau}{p}} \\
&\qquad \le \left(\sum_{k=1}^{\lambda^j} \sum_{(I,\psi)\in\Lambda_{j,k}} \left(\lambda^{-jm\tau} (k\lambda^{-j})^{(\gamma-m)\tau}\right)^{\frac{p}{p-\tau}}\right)^{\frac{p-\tau}{p}} \\
&\qquad \le \left(\sum_{k=1}^{\lambda^j} \sum_{(I,\psi)\in\Lambda_{j,k}} \left(\lambda^{-j\gamma\tau} k^{(\gamma-m)\tau}\right)^{\frac{p}{p-\tau}}\right)^{\frac{p-\tau}{p}} \\
&\qquad \lesssim \left(\lambda^{-j\gamma\frac{p\tau}{p-\tau}} \sum_{k=1}^{\lambda^j} k^{(\gamma-m)\frac{p\tau}{p-\tau}} k^{d-1-\delta} \lambda^{j\delta}\right)^{\frac{p-\tau}{p}} \\
&\qquad \lesssim \lambda^{-j\gamma\tau} \lambda^{j\delta\frac{p-\tau}{p}} \left(\sum_{k=1}^{\lambda^j} k^{(\gamma-m)\frac{p\tau}{p-\tau}+d-1-\delta}\right)^{\frac{p-\tau}{p}}.
\end{aligned}$$

Looking at the value of the exponent in the last sum we see that

$$(\gamma-m)\frac{p\tau}{p-\tau}+d-1-\delta>-1 \qquad \Longleftrightarrow \qquad \gamma-m+r\frac{d-\delta}{d}>0,$$

which leads to

$$\begin{aligned}&\left(\sum_{(I,\psi)\in\Lambda_j^0}\left(|I|^{\frac{m}{d}\tau}\rho_{I,\mathbf{a}}^{(\gamma-m)\tau}\right)^{\frac{p}{p-\tau}}\right)^{\frac{p-\tau}{p}}\\ &\qquad\lesssim \lambda^{-j\gamma\tau}\lambda^{j\delta\frac{p-\tau}{p}}\begin{cases}\lambda^{j\left((\gamma-m)\tau+(d-\delta)\frac{p-\tau}{p}\right)}, & \gamma-m+r\frac{d-\delta}{d}>0,\\ (j+1)^{\frac{p-\tau}{p}}, & \gamma-m+r\frac{d-\delta}{d}=0,\\ 1, & \gamma-m+r\frac{d-\delta}{d}<0.\end{cases}\end{aligned}\tag{4.4}$$

Step 3. We now put $\Lambda^0:=\bigcup_{j\geq 0}\Lambda_j^0$. Summing the first line of the last estimate over all j, we obtain

$$\begin{aligned}&\sum_{(I,\psi)\in\Lambda^0}|I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau}|\langle\tilde{u},\tilde{\psi}_I\rangle|^\tau\\ &\qquad\lesssim\sum_{j=0}^\infty\lambda^{-j(m\tau-d\frac{p-\tau}{p})}\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau\lesssim\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau<\infty,\end{aligned}$$

if the geometric series converges, which happens if

$$m\tau>d\frac{p-\tau}{p}\qquad\Longleftrightarrow\qquad m>d\frac{r}{d}\qquad\Longleftrightarrow\qquad m>r.$$

Similarly, in the second case we see that

$$\begin{aligned}&\sum_{(I,\psi)\in\Lambda^0}|I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau}|\langle\tilde{u},\tilde{\psi}_I\rangle|^\tau\\ &\qquad\lesssim\sum_{j=0}^\infty\lambda^{-j(\gamma\tau-\delta\frac{p-\tau}{p})}(j+1)^{\frac{p-\tau}{p}}\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau\lesssim\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau<\infty,\end{aligned}$$

where the series converges if

$$\gamma\tau>\delta\frac{p-\tau}{p},\quad \text{i.e.,}\quad \gamma>\delta\frac{r}{d},\quad \text{i.e.,}\quad m>r\frac{d-\delta}{d}+\frac{\delta}{d}r=r,\quad \text{i.e.,}\quad m>r,$$

which is the same condition as before. Finally, in the third case we find

$$\begin{aligned}&\sum_{(I,\psi)\in\Lambda^0}|I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau}|\langle\tilde{u},\tilde{\psi}_I\rangle|^\tau\\ &\qquad\lesssim\sum_{j=0}^\infty\lambda^{-j(\gamma\tau-\delta\frac{p-\tau}{p})}\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau\lesssim\|u|\mathcal{K}_{p,\gamma}^{m\mathbf{a}}(\Omega)\|^\tau<\infty,\end{aligned}$$

whenever

$$\gamma\tau>\delta\frac{p-\tau}{p}\qquad\Longleftrightarrow\qquad\gamma>\delta\frac{r}{d}.$$

Step 4. We still need to consider the sets $\Lambda_j \setminus \Lambda_j^0$ consisting of the wavelets close to S, and those ψ_I whose support intersects $\partial\Omega$. Here, we shall make use of the assumption $\tilde{u} \in B^{s\mathbf{a}}_{p,p}(\mathbb{R}^d)$. Since the number of elements in $\Lambda_j \setminus \Lambda_j^0$ is bounded from above by $c\lambda^{j(d-1)}$ we estimate using Hölder's inequality with $\frac{p}{\tau} > 1$ and obtain

$$\sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} |I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^\tau$$

$$\lesssim \lambda^{j(d-1)\frac{p-\tau}{p}} \left(\sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} \lambda^{-jd\left(\frac{1}{p}-\frac{1}{2}\right)p} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^p \right)^{\tau/p}$$

$$= \lambda^{j(d-1)\frac{p-\tau}{p}} \lambda^{-js\tau} \left(\sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} \lambda^{j\left(s+\frac{d}{2}-\frac{d}{p}\right)p} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^p \right)^{\tau/p}.$$

Summing up over j and once more using Hölder's inequality with $\frac{p}{\tau} > 1$ gives

$$\sum_{j=0}^{\infty} \sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} |I|^{\left(\frac{1}{p}-\frac{1}{2}\right)\tau} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^\tau$$

$$\lesssim \sum_{j=0}^{\infty} \lambda^{j(d-1)\frac{p-\tau}{p}} \lambda^{-js\tau} \left(\sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} \lambda^{j\left(s+\frac{d}{2}-\frac{d}{p}\right)p} |\langle \tilde{u}, \tilde{\psi}_I\rangle|^p \right)^{\tau/p}$$

$$\lesssim \left(\sum_{j=0}^{\infty} \lambda^{j(d-1)} \lambda^{-js\tau\frac{p}{p-\tau}} \right)^{\frac{p-\tau}{p}} \cdot \left(\sum_{j=0}^{\infty} \sum_{(I,\psi)\in\Lambda_j\setminus\Lambda_j^0} \lambda^{j\left(s+\frac{d}{2}-\frac{d}{p}\right)p} |\langle \tilde{u}, \tilde{\psi}_I\rangle| \right)^{\tau/p}$$

$$\lesssim \|\tilde{u}|B^{s\mathbf{a}}_{p,p}(\mathbb{R}^d)\|^\tau \lesssim \|u|B^{s\mathbf{a}}_{p,p}(\Omega)\|^\tau,$$

provided that

$$d-1 < \frac{sp\tau}{p-\tau} \quad\Longleftrightarrow\quad \frac{s}{d-1} > \frac{1}{\tau} - \frac{1}{p} = \frac{r}{d} \quad\Longleftrightarrow\quad r < \frac{sd}{d-1}.$$

Altogether, we have proved

$$\|u|B^{r\mathbf{a}}_{\tau,\tau}(\Omega)\| \le \|\tilde{u}|B^{r\mathbf{a}}_{\tau,\tau}(\mathbb{R}^d)\| \lesssim \|u|B^{s\mathbf{a}}_{p,p}(\Omega)\| + \|u|\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega)\|,$$

with constants independent of u. □

Remark 4.2. By a close inspection of the proof of Theorem 4.1 one sees that we have actually proven for any $u \in \mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega) \cap B^{s\mathbf{a}}_{p,p}(\Omega)$ that

$$\|u\|_{B^{r\mathbf{a}}_{\tau,\tau}(\Omega)} \lesssim \max\left\{ |u|_{\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega)}, \|u|B^{s\mathbf{a}}_{p,p}(\Omega)\| \right\}, \tag{4.5}$$

where

$$|u|_{\mathcal{K}^{m\mathbf{a}}_{p,\gamma}(\Omega)} := \left(\sum_{i=1}^{d} \int_\Omega \left| (\rho_{\mathbf{a}}(x))^{m-\gamma} D_i^{m_i} u(x) \right|^p \mathrm{d}x \right)^{1/p} \tag{4.6}$$

denotes the Kondratiev semi-norm, where only the highest derivatives appear.

5. Comparison and outlook: Anisotropic regularity of the heat equation

As already said before, we wish to study the regularity of parabolic problems (in particular, the heat equation) in anisotropic Besov spaces using the embedding from Theorem 4.1 and compare our results with [3, Thm. 2].
Therefore, let the domain $\Omega = D \times [0, T]$ be a space-time cylinder, where $D \subset \mathbb{R}^d$ denotes a bounded Lipschitz domain, $S = \partial_{\text{par}}\Omega$ be the parabolic boundary which has dimension $\delta = d$, and consider the anisotropy $\mathbf{a}$ from (3.22), i.e.,

$$\mathbf{a} = (a_1, \dots, a_{d+1}) = \frac{d+2}{d}\left(1, \dots, 1, \frac{1}{2}\right).$$

Moreover, we denote by $\Theta(\Omega)$ the spaces of all temperatures

$$\Theta(\Omega) := \left\{u : \ \frac{\partial u}{\partial t} = \Delta u \text{ in } \Omega\right\}.$$

Then the result from Aimar et al. obtained in [3, Thms. 2] reads as follows:

Theorem 5.1. *Let $1 < p < \infty$, $\lambda > 0$, $\alpha > 0$, and put $\frac{1}{\tau} = \frac{1}{p} + \frac{\alpha}{d}$. Then*

$$\Theta(\Omega) \cap \mathbb{B}^{\lambda}_{p}(\Omega) \subset \bigcap_{\alpha > \varepsilon > 0} \mathbb{B}^{\alpha-\varepsilon}_{\tau}(\Omega), \quad \textit{where} \quad \alpha < \min\left(d\Big(1 - \frac{1}{p}\Big), \frac{\lambda d}{d-1}\right). \tag{5.1}$$

In particular, this result was obtained with the help of gradient estimates of temperatures. In this context we recall [3, Thm. 5], which will be useful for us in the sequel. We make use of the following notation: we write $\nabla^{2,1}u$ to denote the (d^2+1)-vector given by the d^2 second-order purely spatial derivatives of u and the first derivative of u w.r.t. time, i.e., $\nabla^{2,1}u = \left(\nabla^2 u, \frac{\partial u}{\partial t}\right)$. By $(\nabla^{2,1})^n u$, $n \in \mathbb{N}$, we denote the vector of all derivatives, where each component has the form $\partial^{(\alpha, \alpha_{d+1})}u$ with $|\alpha| + 2\alpha_{d+1} = 2n$. This way we always have in each one of these derivatives an even number of space derivatives. Moreover, $|(\nabla^{2,1})^n u|$ denotes the Euclidean length of $(\nabla^{2,1})^n u$. Then [3, Thm. 5], adapted to our situation, reads as follows.

Corollary 5.2. *Let $\Omega = D \times [0, T]$ with $D \subset \mathbb{R}^d$ be a bounded Lipschitz domain, $0 < \lambda < 2n < \lambda + d$, $n \in \mathbb{N}$, and $1 < p < \infty$. Then there exists a constant c depending on d, λ, p, and the Lipschitz character of D such that*

$$\left\|\delta^{2n-\lambda}|(\nabla^{2,1})^n u| \middle| L_p(\Omega)\right\| \le c \left\|u \middle| L_p([0,T], B^{\lambda}_{p,p}(D))\right\| \le c' \|u|\mathbb{B}^{\lambda}_{p}(\Omega)\| \tag{5.2}$$

holds for every temperature u in $\Theta(\Omega)$.

P r o o f : Corollary 5.2 in its above version is a consequence of [3, Thms. 4,5] together with [1, Lem. 5.3], where the latter is the observation that the derivative $(\nabla^{2,1})^n u$ belongs to the linear span of $\nabla^{2n}u$. □

We can now improve Theorem 5.1 as follows.

Theorem 5.3. *Let $0 < p < \infty$, $0 < s < 2n < s + d$, $n \in \mathbb{N}$, $\alpha > 0$, and put $\frac{1}{\tau} = \frac{1}{p} + \frac{\alpha}{d}$. Then*

$$\Theta(\Omega) \cap \mathbb{B}^s_p(\Omega) \subset \mathbb{B}^\alpha_\tau(\Omega), \quad \textit{where} \quad \alpha < \min\left(2n, s\frac{d+1}{d}\right). \tag{5.3}$$

P r o o f : We can reinterpret the estimate (5.2) in terms of anisotropic Kondratiev regularity for the homogeneous heat equation as follows: The left hand side in (5.2) can be expressed via the Kondratiev semi-norm (4.6), since using (3.24) we see that for $\mathbf{m} = (2n, \ldots, 2n, n) = 2n\frac{d}{d+2}\mathbf{a} =: m\mathbf{a}$ and $s := \frac{d+2}{d}\gamma = \lambda$ we have

$$\begin{aligned}
|u|_{\mathcal{K}^{\mathbf{m}}_{p,\gamma}(\Omega)} &\sim \sum_{i=1}^d \left\|(\rho_{\mathbf{a}})^{m-\gamma} D_i^{m_i} u \middle| L_p(\Omega)\right\| \\
&\sim \sum_{i=1}^d \left\|\left(\delta^{\frac{d+2}{d}}\right)^{m-\gamma} D_i^{m_i} u \middle| L_p(\Omega)\right\| \\
&= \sum_{i=1}^d \left\|\delta^{2n-s} D_i^{m_i} u \middle| L_p(\Omega)\right\| \lesssim \left\|\delta^{2n-\lambda} |(\nabla^{2,1})^n u| \middle| L_p(\Omega)\right\|.
\end{aligned}$$

Thus, a combination of Theorem 4.1, Corollary 5.2, and the observation that $B^{\tilde{s}\mathbf{a}}_{p,p}(\Omega) = \mathbb{B}^s_p(\Omega)$ for $\tilde{s} = s\frac{d}{d+2}$ yields for a temperature $u \in \Theta(\Omega)$:

$$\|u|B^{r\mathbf{a}}_{\tau,\tau}(\Omega)\| \lesssim \max\left\{|u|_{\mathcal{K}^{\mathbf{m}}_{p,\tilde{s}}(\Omega)}, \|u|B^{\tilde{s}\mathbf{a}}_{p,p}(\Omega)\|\right\} \lesssim \|u|B^{\tilde{s}\mathbf{a}}_{p,p}(\Omega)\| \tag{5.4}$$

subject to the restriction

$$0 < r < \min\left(2n\frac{d}{d+2}, \tilde{s}\frac{d+1}{d}\right).$$

In good agreement with (3.21) we put $B^{r\mathbf{a}}_{\tau,\tau}(\Omega) := \mathbb{B}^\alpha_\tau(\Omega)$ for $\alpha = r\frac{d+2}{d}$ (i.e., the space $\mathbb{B}^\alpha_\tau(\Omega)$ with $\tau < 1$ has to be understood – in a slight abuse of notation – according to Definition 3.2) and we obtain (5.3). □

Remark 5.4. (i) Comparing (5.3) with (5.1) we conclude that the restriction $\alpha < d\big(1 - \frac{1}{p}\big)$ (resulting from the fact that the definition of the spaces $\mathbb{B}^\lambda_p(\Omega)$ and the subsequent argumentation in [3] where limited to $p > 1$) can be completely removed. However, since $\Omega \subset \mathbb{R}^{d+1}$ by replacing d by $d+1$ our approach gives a slightly worse upper bound for α $\big(s\frac{d+1}{d}$ instead of $s\frac{d}{d-1}\big)$. In particular, invoking [25, Thm. 6.2] we deduce that

$$\Theta(\Omega) \subset W^{2\ldots,2,1}(\Omega) = B^{\frac{2d}{d+2}\mathbf{a}}_{2,2}(\Omega) = \mathbb{B}^2_2(\Omega).$$

i.e., (5.3) yields for parameters $p = 2$ and $s < 2$ (with $2n = 2$ if $d = 2$) that

$$\Theta(\Omega) \subset \mathbb{B}^\alpha_\tau(\Omega), \qquad \text{where} \quad \alpha < \frac{8}{3}\ (d = 3) \quad \text{and} \quad \alpha < 2\ (d = 2).$$

On the other hand (5.1) only yields $\alpha < \frac{3}{2}$ $(d = 3)$ and $\alpha < 1$ $(d = 2)$.

(ii) Let us note that compared to [3] our approach is more flexible: It allows us to treat more general parabolic equations (also with inhomogeneous initial boundary data) as long as one has regularity results for the solution of the parabolic problem in anisotropic Kondratiev spaces. In this context we mention [20] for first results in this direction.
Moreover, we think that even better results can be achieved if one investigates regularity in anisotropic Kondratiev and Besov spaces which have different integrability w.r.t. the spacial and time variable. This interesting problem will be studied a future paper.
Finally, we remark that it is not completely clear that the anisotropic Kondratiev and Besov spaces we are dealing with in this paper are the optimal spaces for studying parabolic PDEs. Another possibility would be to have a look at the regularity of the solutions to evolution equations in Besov spaces of dominating mixed smoothness type or even Banach-valued Besov spaces.

Acknowledgements. Stephan Dahlke wants to thank Ron DeVore for numerous inspiring discussions on mathematics. The collaboration with him has influenced his mathematical life til today.

References

[1] H. Aimar, I. Gómez and B. Iaffei, *Parabolic mean values and maximal estimates for gradients of temperatures*, J. Funct. Anal. **255** (2008), 1939–1956.
[2] H. Aimar, I. Gómez and B. Iaffei, *On Besov regularity of temperatures*, *J. Fourier Anal. Appl.* **16** (2010), 1007–1020.
[3] H. Aimar and I. Gómez, *Parabolic Besov regularity for the heat equation*, Constr. Approx. **36** (2012), 145–159.
[4] A. Ayache, *Constructions of non separable dyadic compactly supported orthonormal wavelet bases for* $L^2(\mathbb{R}^2)$ *of arbitrarily high regularity*, Rev. Mat. Iberoamericana. **15** (1999), 37–58.
[5] O.V. Besov, V.P. Il'in and S.M. Nikol'skij, *Integral Representations of Functions and Imbedding Theorems. Vol. I.*, Scripta Series in Mathematics, Translated from the Russian. Halsted Press, John Wiley & Sons, New York-Toronto, 1978.
[6] O.V. Besov, V.P. Il'in and S.M. Nikol'skij, *Integral Representations of Functions and Imbedding Theorems. Vol. II.*, Scripta Series in Mathematics, Translated from the Russian. Halsted Press, John Wiley & Sons, New York-Toronto, 1979.
[7] S. Dahlke and R. A. DeVore., *Besov regularity for elliptic boundary value problems*, Comm. Partial Differential Equations **22** (1997), 1–16.
[8] S. Dahlke, M. Hansen, C. Schneider and W. Sickel, *On Besov regularity of solutions to nonlinear elliptic partial differential equations*, Nonlinear Anal. **192** (2020): 11686.
[9] S. Dahlke, M. Hansen, C. Schneider and W. Sickel, *Properties of Kondratiev spaces*, Preprint-Reihe Philipps-University Marburg, Bericht Mathematik Nr. 2018-06, 2018.
[10] S. Dahlke and C. Schneider, *Besov regularity of parabolic and hyperbolic PDEs*, Anal. Appl. **17** (2019), 235–291.
[11] S. Dahlke and W. Sickel, *Besov regularity for the Poisson equation in smooth and polyhedral cones*, Sobolev spaces in mathematics II, Int. Math. Ser. (N.Y.) **9** (2009), 123–145.
[12] I. Daubechies, *Orthonormal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **41** (1998), 909–996.
[13] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM, Philladelphia, 1992.
[14] R. A. DeVore, *Nonlinear approximation*, Acta Numer. **7** (1998), 51–150.
[15] G. Garrigós, R. Hochmuth and A. Tabacco, *Wavelet characterizations for anisotropic Besov spaces with* $0 < p < 1$, *Proc. Edinb. Math. Soc. (2)* **47** (2004), 573–595.

[16] G. Garrigós and A. Tabacco, *Wavelet decompositions of anisotropic Besov spaces*, Math. Nachr. **239-240** (2002), 80–102.

[17] W. He and M.-J. Lai, *Construction of bivariate compactly supported biorthongonal box spline wavelets with arbitrary high regularities*, Appl. Computat. Harmon. Analysis **6** (1999), 53–74.

[18] R. Hochmuth, *Wavelet characterization for anisotropic Besov spaces*, Appl. Computat. Harmon. Analysis **12** (2002), 179–208.

[19] J. Hui, S.D. Riemenschneider and Z. Shen, *Multivariate compactly supported fundamental refinable functions, duals, and biorthogonal wavelets*, Stud. Appl. Math. **102** (1999), 173–204.

[20] V.A. Kozlov and J. Rossmann, *Asymptotics of solutions of the heat equation in cones and dihedra under minimal assumptions on the boundary*, Bound. Value Probl. **142** (2012).

[21] S. M . Nikol'skij, *Approximation of Functions of Several Variables and Embedding Theorems*, Die Grundlehren der mathematischen Wissenschaften, Band 205. Translated from the Russian. Springer, New York - Heidelberg, 1975.

[22] U. Reif, *Polynomial approximation on domains bounded by diffeomorphic images of graphs*, J. Approx. Theory, **164** (2012), 954–970.

[23] C. Schneider, *Beyond Sobolev and Besov: Regularity of Solutions of PDEs and Their Traces in Function Spaces*, Lecture Notes in Mathematics, vol. 2291, Springer, 2021.

[24] C. Schwab and R. Stevenson, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comput. **78** (2009), 1293–1318.

[25] I. Wood, *Maximal L^p-regularity for the Laplacian on Lipschitz domains*, Math. Z. **255** (2007), 855–875.

Manuscript received January 5 2022
revised March 25 2022

S. Dahlke
Philipps-University Marburg, FB12 Mathematics and Computer Science, Hans-Meerwein Straße, Lahnberge, 35032 Marburg, Germany
E-mail address: `dahlke@mathematik.uni-marburg.de`

C. Schneider
Friedrich-Alexander Universität Erlangen, Applied Mathematics III, Cauerstr. 11, 91058 Erlangen, Germany
E-mail address: `schneider@math.fau.de`