Journal of Functional Analysis 286 (2024) 110344

Contents lists available at ScienceDirect

Journal of Functional Analysis

journal homepage: www.elsevier.com/locate/jfa

Regular Article

Sobolev embeddings for kinetic Fokker-Planck equations

Andrea Pascucci *, Antonello Pesce

Dipartimento di Matematica, Università di Bologna, Bologna, Italy

A R T I C L E   I N F O

Article history:
Received 2 December 2022
Accepted 9 January 2024
Available online 29 January 2024
Communicated by Luis Silvestre

Dedicated to Ermanno Lanconelli on the occasion of his 80th birthday

Keywords:
Sobolev embeddings
Fokker-Planck equations
Weak Hörmander condition
Langevin kinetic model

A B S T R A C T

We introduce intrinsic Sobolev-Slobodeckij spaces for a class of ultra-parabolic Kolmogorov type operators satisfying the weak Hörmander condition. We prove continuous embeddings into Lorentz and intrinsic Hölder spaces. We also prove approximation and interpolation inequalities by means of an intrinsic Taylor expansion, extending analogous results for Hölder spaces. The embedding at first order is proved by adapting a method by Luc Tartar which only exploits scaling properties of the intrinsic quasi-norm, while for higher orders we use uniform kernel estimates.

1. Introduction

In this paper we develop a functional framework for the study of kinetic Fokker-Planck equations. Specifically, we introduce intrinsic Sobolev spaces suitably related to a system of Hörmander's vector fields: our main results are embedding, interpolation and approximation theorems that are the basic tools in many problems concerning partial differential equations.

* Corresponding author.
E-mail addresses: andrea.pascucci@unibo.it (A. Pascucci), antonello.pesce2@unibo.it (A. Pesce).

https://doi.org/10.1016/j.jfa.2024.110344

Let (t, x) denote a point in $\mathbb{R} \times \mathbb{R}^N$ and, for fixed $d \leqslant N$, consider the vector fields

$$\partial_{x_1}, \ldots, \partial_{x_d} \quad \text{and} \quad Y := \langle Bx, \nabla_x \rangle + \partial_t, \tag{1.1}$$

where B is a constant $N \times N$ matrix and $\nabla_x = (\partial_{x_1}, \ldots, \partial_{x_N})$. We assume the Hörmander's condition is satisfied:

$$\text{rank Lie}(\partial_{x_1}, \ldots, \partial_{x_d}, Y) = N + 1. \tag{1.2}$$

The classical example we have in mind is the Langevin kinetic model, given by the system of stochastic differential equations

$$\begin{cases} dV_t = dW_t, \\ dP_t = V_t dt, \end{cases} \tag{1.3}$$

where W is a d-dimensional Brownian motion. Here the processes V and P represent the velocity and position of a system of d particles. The forward Kolmogorov (or Fokker-Planck) operator of (1.3), written in terms of the variables $x = (v, p) \in \mathbb{R}^d \times \mathbb{R}^d$, is in the form of a sum of squares of the vector fields $\partial_{v_1}, \ldots, \partial_{v_d}$ plus a drift (or transport term) Y_0, precisely

$$\frac{1}{2} \sum_{i=1}^{d} \partial_{v_i}^2 - Y_0, \qquad Y_0 := \langle v, \nabla_p \rangle + \partial_t. \tag{1.4}$$

In this example, $N = 2d$ and

$$B = \begin{pmatrix} 0_d & 0_d \\ I_d & 0_d \end{pmatrix}$$

where I_d and 0_d denote the $d \times d$ identity and null matrices respectively. Operator (1.4) satisfies the Hörmander's condition, is hypoelliptic and has a Gaussian fundamental solution that is the transition density of the Markov process (V, P) in (1.3).

The literature on generalized Sobolev spaces for Hörmander's vector fields is vast (see, for instance, [25], [26], [10], [31], [17]). When dealing with the regularity properties of PDEs modeled on the vector fields (1.1)-(1.2), as for example the equation in divergence form

$$\sum_{i,j=1}^{d} \partial_{x_i} \left(a_{ij} \partial_{x_j} u \right) - Yu = 0, \tag{1.5}$$

it is standard to assign a formal weight to each of the vector fields, that is one for the directions $\partial_{x_1}, \ldots, \partial_{x_d}$ of diffusion and is two for the drift Y; also, consistently with the structure of the equation, Y should be interpreted as a *second order derivative* in intrinsic sense. As earlier noted in [25] among others, this fact raises a question about the role of Y in the definition of *first order* intrinsic Sobolev space $W^{1,p}$: indeed, in the degenerate case $d < N$, the regularity properties of (1.5) strongly rely on Hörmander's condition and involve the second order derivative Y in a crucial way. Many remarkable results have been proven for weak solutions of (1.5), defined as functions u such that Yu belongs to L^2, in addition to the minimal assumptions $u, \partial_{x_1} u, \dots, \partial_{x_d} u \in L^2$ needed to write the equation (1.5) in the sense of distributions: we refer for instance to [3], [4], [19], [5], [6], [1] and [9]. In [22] a first $L^2 - L^\infty$ estimate has been proven by using Moser's approach; moreover, in [12] a Harnack inequality for kinetic Fokker-Planck equations with rough coefficients has been proven extending the De Giorgi-Nash-Moser theory.

In Section 2 we introduce intrinsic Sobolev-Slobodeckij spaces for (1.1), denoted by $W_B^{k,p}$, where at first order (i.e. $k = 1$) the vector field Y appears as a fractional derivative of order 1/2: this approach is coherent with the scaling properties of the Hörmander vector fields and therefore seems suitable for the study of (1.5). In particular, we can give a natural definition of weak solution u of (1.5) in the Sobolev space $W_B^{1,2}$ without requiring $Yu \in L^2$ as it is usually done in the literature: as far as we know, this is the first result in this direction.

We mention that the use of fractional derivatives makes it difficult to prove embedding results by means of representation formulas in terms of a parametrix, at least for $k = 1$, as in [22] or [5]. Indeed, for the proof of our main embedding result, Theorem 1.1, we use a remarkable method developed by Tartar [28], that is only based on scaling arguments and a characterization of Lorentz spaces given in Lemma A.5.

In the following statement $\mathbf{d}$ denotes the homogeneous dimension of $\mathbb{R}^{N+1}$ induced by the vector fields (1.1), whose precise definition is given in (2.6): to fix ideas, $\mathbf{d} = 4d + 2$ for the Fokker-Planck operator (1.4).

Theorem 1.1 *($W_B^{1,p}$ embeddings).*

i) For $1 \leqslant p < \mathbf{d}$ we have

$$W_B^{1,p} \subseteq L^{q,p}, \qquad p \leqslant q \leqslant p^*, \qquad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{\mathbf{d}}, \tag{1.6}$$

where $L^{p,q}$ denotes the Lorentz space. In particular, $W_B^{1,p} \subseteq L^q$ for $p \leqslant q \leqslant p^*$;

ii) for $\mathbf{d} < p < \infty$ we have

$$W_B^{1,p} \subseteq C_B^{0,1-\frac{\mathbf{d}}{p}}; \tag{1.7}$$

iii) for $p = \mathbf{d}$ we have

$$W_B^{1,\mathbf{d}} \subseteq L^{q,\mathbf{d}} \subseteq L^q, \qquad q \geqslant \mathbf{d}. \tag{1.8}$$

Moreover, if $u \in W_B^{1,\mathbf{d}}$ then for every $\lambda, \delta > 0$ we have

$$\int_{|u|>\delta} e^{\lambda |u(z)|^{\frac{\mathbf{d}}{\mathbf{d}-1}}} dz < \infty. \tag{1.9}$$

The Morrey embedding (1.7) is given in terms of the optimal generalized Hölder spaces $C_B^{k,\alpha}$ only recently introduced in [18] together with an intrinsic Taylor formula. Embeddings for higher order spaces $W_B^{k,p}$ are provided in Theorem 7.1. Remarkably, estimate (1.9) extends Trudinger's result [30]. Embedding results for Kolmogorov equations were also proved in [7] and more recently in [11].

We acknowledge that Tartar himself applied his approach to the Langevin operator (1.4): according to [28], Appendices II and III, he proved that for a function $f = f(t, x, v)$ on $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d$, with $f, \nabla_v f, Y_0 f := (\partial_t + v \cdot \nabla_x) f \in L^p$, one can first prove the "crude" embedding estimate

$$\|f\|_q \lesssim \|f\|_p + \|\nabla_v f\|_p + \|Y_0 f\|_p,$$

for some $q > p$ and then get the embeddings:

- in $L^{p^{**},p}$ if $1 \leqslant p < 3d + 1$, with $\frac{1}{p^{**}} = \frac{1}{p} - \frac{1}{3d+1}$;
- in L^∞ if $p > 3d + 1$;
- in L^r if $p = 3d + 1$, for any $p \leqslant r < \infty$.

As a main motivation, our study is a first step in the development of a theory of generalized Besov spaces for possible applications to *stochastic* partial differential equations: we mention that recent results for stochastic kinetic equations were established in [21] and [32]. Secondly, even for deterministic kinetic equations, our results improve the known regularity estimates available in the literature by providing the natural functional framework for weak solutions of kinetic Fokker-Planck equations.

The paper is structured as follows. In Section 2 we state the precise assumptions, introduce the intrinsic Sobolev and Hölder spaces and collect some preliminary result concerning the geometric structure induced on $\mathbb{R}^{N+1}$ by the vector fields (1.1). In Section 3 we prove a first interpolation result, Proposition 3.5, that provides a simplified and equivalent definition of intrinsic Sobolev quasi-norm. In Section 4 we show an intrinsic Taylor expansion, Theorem 4.1, for functions in $W_B^{k,p}$ which extends the analogous results for intrinsic Hölder spaces proved in [18]. Crucial approximation and interpolation results, Theorems 5.2 and 5.3, are proven in Section 5. Section 6 contains the proof of our main result, Theorem 1.1, on the embeddings of $W_B^{1,p}$. Eventually, in Section 7 we prove Theorem 7.1 on the higher order embeddings. For reader's convenience, in the Appendix we recall some basic result about interpolation and Lorentz spaces.

In the context of our proofs we will often use the notation $A \lesssim B$, meaning that $A \leqslant cB$ for some positive constant c which may depend on the quantities specified in the corresponding statement.

2. Preliminaries

2.1. Assumptions

We recall that Hörmander's condition is equivalent to the well-known Kalman rank condition for controllability of linear systems (cf., for instance, Section 9.5 in [20]); also, it was shown in [13] that, up to a change of basis, condition (1.2) is equivalent to the following

Assumption 2.1 *(Hörmander's condition).* The matrix B takes the block-form

$$B = \begin{pmatrix} * & * & \cdots & * & * \\ B_1 & * & \cdots & * & * \\ 0 & B_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_r & * \end{pmatrix} \tag{2.1}$$

where B_j is a $(d_{j-1} \times d_j)$-matrix of rank d_j with

$$d \equiv d_0 \geqslant d_1 \geqslant \cdots \geqslant d_r \geqslant 1, \qquad \sum_{j=0}^{r} d_j = N.$$

In general, the $*$-blocks in (2.1) are arbitrary. Our second standing assumption is the following

Assumption 2.2 *(Homogeneity).* All the $*$-blocks in (2.1) are null.

As proven in [13], Assumption 2.2 is equivalent to the fact that the kinetic Fokker-Planck operator

$$\mathscr{K} := \frac{1}{2} \sum_{i=1}^{d} \partial_{x_i}^2 - Y \tag{2.2}$$

is homogeneous of degree two with respect to the family of dilations defined as follows: first of all, consistently with the block decomposition (2.1) of B, we write $x \in \mathbb{R}^N$ as the direct sum $x = x^{[0]} + \cdots + x^{[r]}$ where $x^{[i]} \in \mathbb{R}^N$ is defined as

$$x_k^{[i]} = \begin{cases} x_k & \text{if } \bar{d}_{i-1} < k \leqslant \bar{d}_i, \\ 0 & \text{otherwise,} \end{cases} \qquad \bar{d}_i := \sum_{j=0}^{i} d_j, \quad \bar{d}_{-1} := 0, \quad i = 0, \dots, r.$$

Then, we have $\mathscr{K}(u(D_\lambda)) = \lambda^2 (\mathscr{K} u)(D_\lambda)$ where

$$D_\lambda(t,x) := (\lambda^2 t, \hat{D}_\lambda x), \qquad \hat{D}_\lambda x := \sum_{i=0}^{r} \lambda^{2i+1} x^{[i]}. \tag{2.3}$$

For instance, the Langevin operator (1.4) is homogeneous with respect to the dilation group $D_\lambda(t,v,p) = (\lambda^2 t, \lambda v, \lambda^3 p)$ in $\mathbb{R}^{2d+1}$.

2.2. Intrinsic Hölder and Sobolev spaces

In this section we recall the definition of intrinsic Hölder space as given in [18] and introduce a notion of intrinsic Sobolev space, naturally associated to the system of vector fields (1.1).

Let $h \mapsto e^{hX}z$ denote the integral curve of a Lipschitz vector field X starting from $z \in \mathbb{R}^{N+1}$, defined as the unique solution of

$$\begin{cases} \frac{d}{dh} e^{hX}z = X\left(e^{hX}z\right), & h \in \mathbb{R}, \\ e^{hX}z|_{h=0} = z. \end{cases}$$

For the vector fields in (1.1), we have

$$e^{h\partial_{x_i}}(t,x) = (t, x + h\mathbf{e}_i), \qquad e^{hY}(t,x) = (t+h, e^{hB}x),$$

where $\mathbf{e}_i$ is the i-th element of the canonical basis of $\mathbb{R}^N$.

Definition 2.3. Let m_X be a formal weight associated to the vector field X. For $\alpha \in]0, m_X]$, we say that $u \in C^\alpha_X$ if the quasi-norm

$$\|u\|_{C^\alpha_X} := \sup_{\substack{z\in\mathbb{R}^{N+1}\\ h\in\mathbb{R}\setminus\{0\}}} \frac{\left|u\left(e^{hX}z\right) - u(z)\right|}{|h|^{\frac{\alpha}{m_X}}}$$

is finite.

Hereafter, we set *the formal weight of the vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ equal to one and the formal weight of Y equal to two*, which is coherent with the homogeneity of the Fokker-Planck operator $\mathscr{K}$ with respect to the dilations D_λ in (2.3). From [18] we recall the following

Definition 2.4 *(Intrinsic Hölder spaces).* For $\alpha \in]0,1]$ we define the Hölder quasi-norms

$$\|u\|_{C^{0,\alpha}_B} := \sup_{\mathbb{R}^{N+1}} |u| + \sum_{i=1}^{d} \|u\|_{C^\alpha_{\partial_{x_i}}} + \|u\|_{C^\alpha_Y},$$

$$\|u\|_{C^{1,\alpha}_B} := \sup_{\mathbb{R}^{N+1}} |u| + \|\nabla_d u\|_{C^{0,\alpha}_B} + \|u\|_{C^{\alpha+1}_Y},$$

where $\nabla_d := (\partial_{x_1}, \dots, \partial_{x_d})$ and inductively, for $n \geqslant 2$,

$$\|u\|_{C^{n,\alpha}_B} := \sup_{\mathbb{R}^{N+1}} |u| + \|\nabla_d u\|_{C^{n-1,\alpha}_B} + \|Yu\|_{C^{n-2,\alpha}_B}.$$

Next we introduce the intrinsic Sobolev spaces. First, as in [17], for any $u \in L^p$, with $p \geqslant 1$, we define the fractional Sobolev-Slobodeckij quasi-norm of order $s \in]0,1[$ along a Lipschitz vector field X as

$$[u]_{X,s,p} := \left(\int_{\mathbb{R}^{N+1}} dz \int_{|h|\leqslant 1} \frac{|u(e^{hX}z) - u(z)|^p}{|h|^{ps+1}} dh \right)^{\frac{1}{p}}.$$

Definition 2.5. For $p \geqslant 1$ we set

$$\begin{aligned} |u|_{1,p,B} &:= \|\nabla_d u\|_p + [u]_{Y,\frac{1}{2},p}, \\ |u|_{2,p,B} &:= |\nabla_d u|_{1,p,B} + \|Yu\|_p, \end{aligned}$$

and inductively, for $n \geqslant 3$,

$$|u|_{n,p,B} := |\nabla_d u|_{n-1,p,B} + |Yu|_{n-2,p,B}.$$

Definition 2.6 *(Intrinsic Sobolev spaces).* For $p \geqslant 1$ we define the Sobolev quasi-norms

$$\begin{aligned} \|u\|_{W_B^{1,p}} &:= \|u\|_p + |u|_{1,p,B}, \\ \|u\|_{W_B^{2,p}} &:= \|u\|_p + \|\nabla_d u\|_{W_B^{1,p}} + \|Yu\|_p, \end{aligned}$$

and inductively, for $n \geqslant 3$,

$$\|u\|_{W_B^{n,p}} := \|u\|_p + \|\nabla_d u\|_{W_B^{n-1,p}} + \|Yu\|_{W_B^{n-2,p}}.$$

The following alternative definition of Sobolev quasi-norm is sometimes useful.

Definition 2.7. For $n \in \mathbb{N}$ and $p \geqslant 1$ we set

$$|||u|||_{W_B^{n,p}} := \|u\|_p + |u|_{n,p,B}.$$

Clearly we have $\|\cdot\|_{W_B^{n,p}} \geqslant |||\cdot|||_{W_B^{n,p}}$. In Section 3, Proposition 3.5, we prove that $\|\cdot\|_{W_B^{n,p}}$ and $|||\cdot|||_{W_B^{n,p}}$ are equivalent and therefore define the same functional spaces. This means that the intermediate orders quasi-norms are not needed to characterize $W_B^{n,p}$.

Remark 2.8. Let

$$\lfloor u \rfloor_{Y,s,p} := \left(\int_{\mathbb{R}^{N+1}} dz \int_{\mathbb{R}} \frac{|u(e^{hY}z) - u(z)|^p}{|h|^{ps+1}} dh \right)^{\frac{1}{p}}$$

and notice that

$$[u]_{Y,s,p} \leqslant \lfloor u \rfloor_{Y,s,p} \leqslant [u]_{Y,s,p} + c_{p,s}\|u\|_p, \qquad c_{p,s} := \left(\int\limits_{|h|>1} \frac{2}{|h|^{1+ps}} dh \right)^{\frac{1}{p}}.$$

Then, if we replace $[u]_{Y,s,p}$ by $\lfloor u \rfloor_{Y,s,p}$ in Definition 2.6, we get equivalent norms.

2.3. Dilation and translation groups

Besides the homogeneity with respect to D_λ in (2.3), operator $\mathscr{K}$ in (2.2) has also the remarkable property of being invariant with respect to the left translations in the group law

$$(t,x) \circ (s,\xi) = (t+s, e^{sB}x + \xi), \qquad (t,x),(s,\xi) \in \mathbb{R}^{N+1}.$$

Indeed, a simple computation shows that, for any $z, \zeta \in \mathbb{R}^{N+1}$,

$$\zeta^{-1} \circ e^{\delta Y} z = e^{\delta Y}(\zeta^{-1} \circ z), \qquad \zeta^{-1} \circ e^{\delta \partial_{x_i}} z = e^{\delta \partial_{x_i}}(\zeta^{-1} \circ z), \qquad i = 1, \dots, d, \tag{2.4}$$

where $(t,x)^{-1} = (-t, -e^{-tB}x)$. Analogously, we have (see, for instance, [23])

$$D_\lambda e^{\delta Y}(z) = e^{\delta \lambda^2 Y}(D_\lambda z), \qquad D_\lambda e^{\delta \partial_{x_i}}(z) = e^{\delta \lambda^{2j+1} \partial_{x_i}} D_\lambda z, \qquad i = \bar{d}_{j-1}+1, \dots, \bar{d}_j. \tag{2.5}$$

A D_λ-homogeneous norm on $\mathbb{R}^{N+1}$ is defined as

$$\|(t,x)\|_B = |t|^{\frac{1}{2}} + |x|_B, \qquad |x|_B = \sum_{i=0}^{r} |x^{[i]}|^{\frac{1}{2i+1}},$$

and

$$\mathbf{d} := 2 + \sum_{k=0}^{r} (2k+1) d_k \tag{2.6}$$

is usually called the homogeneous dimension of $\mathbb{R}^{N+1}$ with respect to D_λ.

Lemma 2.9 *([16], Proposition* 5.1*). There exists $m = m(B) \geqslant 1$ such that*

$$\|\zeta^{-1} \circ z\|_B \leqslant m(\|\zeta\|_B + \|z\|_B) \quad m^{-1}\|z\|_B \leqslant \|z^{-1}\|_B \leqslant m\|z\|_B \quad z, \zeta \in \mathbb{R}^{N+1}. \tag{2.7}$$

Remark 2.10. Since $e^{\delta Y} z = z \circ (\delta, 0)$, by (2.7) we have

$$\frac{1-mc}{m}\|z\|_B \leqslant \|e^{\delta Y} z\|_B \leqslant m(1+c)\|z\|_B, \tag{2.8}$$

for any $|\delta|^{\frac{1}{2}} \leqslant c\|z\|_B$ with $c \in \left]0, \frac{1}{m}\right[$.

Remark 2.11. The matrix B is nilpotent of degree $r+1$. In particular, for any $n \leqslant r$ we have

$$B^n = \begin{pmatrix} 0_{\bar{d}_{n-1}\times d_0} & 0_{\bar{d}_{n-1}\times d_1} & \cdots & 0_{\bar{d}_{n-1}\times d_{r-n}} & 0_{\bar{d}_{n-1}\times(\bar{d}_r-\bar{d}_{r-n})} \\ \prod_{j=1}^{n} B_j & 0_{d_n\times d_1} & \cdots & 0_{d_n\times d_{r-n}} & 0_{d_n\times(\bar{d}_r-\bar{d}_{r-n})} \\ 0_{d_{n+1}\times d_0} & \prod_{j=2}^{n+1} B_j & \cdots & 0_{d_{n+1}\times d_{r-n}} & 0_{d_{n+1}\times(\bar{d}_r-\bar{d}_{r-n})} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{d_r\times d_0} & 0_{d_r\times d_1} & \cdots & \prod_{j=r-n+1}^{r} B_j & 0_{d_r\times(\bar{d}_r-\bar{d}_{r-n})} \end{pmatrix},$$

where

$$\prod_{j=1}^{n} B_j = B_n B_{n-1}\cdots B_1,$$

and $B^n = 0$ for $n > r$. Thus

$$e^{\delta B} = I_N + \sum_{j=1}^{r} \frac{B^j}{j!}\delta^j$$

is a lower triangular matrix with diagonal $(1,\dots,1)$ and therefore it has determinant equal to 1.

Lemma 2.12. *For any $n \in \mathbb{N}$ and $u \in L^p$, with $p \geqslant 1$, we have*

$$\|u(D_\lambda)\|_p = \lambda^{-\frac{\mathbf{d}}{p}}\|u\|_p, \qquad |u(D_\lambda)|_{n,p,B} = \lambda^{n-\frac{\mathbf{d}}{p}}|u|_{n,p,B} \tag{2.9}$$

Proof. The first equality follows by a simple change of variable. Next, for $i = 1,\dots,d$ we have, by (2.5)

$$\|\partial_{x_i} u(D_\lambda)\|_p^p = \int_{\mathbb{R}^{N+1}} |\partial_{x_i} u(D_\lambda z)|^p dz = \int_{\mathbb{R}^{N+1}} |\lambda(\partial_{x_i} u)(D_\lambda z)|^p dz =$$

(by the change of variable $z' = D_\lambda z$)

$$= \lambda^p \int_{\mathbb{R}^{N+1}} |(\partial_{x_i} u)(z')|^p \lambda^{-\mathbf{d}} dz' = \lambda^{p-\mathbf{d}}\|\partial_{x_i} u\|_p^p. \tag{2.10}$$

Similarly

$$[u(D_\lambda)]^p_{Y,\frac{1}{2},p} = \int\limits_{\mathbb{R}^{N+1}} \int\limits_{\mathbb{R}} \left|u\left(D_\lambda(e^{hY}z)\right) - u(D_\lambda z)\right|^p \frac{dh}{|h|^{1+\frac{p}{2}}} dz$$
$$= \int\limits_{\mathbb{R}^{N+1}} \int\limits_{\mathbb{R}} \lambda^{2+p} |u(e^{h\lambda^2 Y} D_\lambda z) - u(D_\lambda z)|^p \frac{dh}{|\lambda^2 h|^{1+\frac{p}{2}}} dz =$$

(by the change of variables $(h', z') = (\lambda^2 h, D_\lambda z)$)

$$= \lambda^p \int\limits_{\mathbb{R}^{N+1}} \int\limits_{\mathbb{R}} |u(e^{h'Y}z') - u(z')|^p \frac{dh'}{|h'|^{1+\frac{p}{2}}} \lambda^{-\mathbf{d}} dz = \lambda^{p-\mathbf{d}} [u]^p_{Y,\frac{1}{2},p}. \tag{2.11}$$

(2.10) and (2.11) give the second equality for $n = 1$. The case $n = 2$ is analogous and the general case $n > 2$ follows by induction. □

3. Alternative Sobolev norms and a first interpolation result

3.1. Intrinsic weak derivatives in $W^{n,p}_B$

By definition, the quasi-norm $|\cdot|_{n,p,B}$ only controls weak derivatives of order n and $n-1$, which are made up of compositions of $\partial_{x_1}, \dots, \partial_{x_d}$ and Y for any possible permutation. We show that actually a function $u \in W^{n,p}_B$ supports all the weak derivatives of intrinsic order l, $l \leqslant n$, and for $k \in \mathbb{N}_0$ and $\beta \in \mathbb{N}_0^N$ we have

$$Y^k \partial^\beta u \in W^{n-l,p}_B, \quad 2k + \langle\beta\rangle_B = l, \tag{3.1}$$

where

$$\partial^\beta = \partial^{\beta_1}_{x_1} \cdots \partial^{\beta_N}_{x_N}, \qquad \langle\beta\rangle_B := \sum_{i=0}^{r} (2i+1) \sum_{k=1+\bar{d}_{i-1}}^{\bar{d}_i} \beta_k. \tag{3.2}$$

Indeed these derivatives can be recovered by taking appropriate iterated commutators of the vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and Y: exploiting these commutators, we can also rearrange the terms appearing in $|\cdot|_{n,p,B}$ and provide a more explicit characterization which only make use of the intrinsic derivatives in the form (3.1).

First we recall some preliminary notions from [18], Section 4. By the structure of the matrix B, for any $n = 0, \dots, r$ and $v \in \mathbb{R}^N$ we have

$$B^n v \in \bigoplus_{i=n}^{r} V_i, \qquad V_i := \{x^{[i]} \mid x \in \mathbb{R}^N\}, \tag{3.3}$$

and $B^n = 0$ for $n > r$. In particular, if $v \in V_0$ then we have

$$B^n v \in V_n, \qquad n = 0, \dots, r.$$

Moreover there exist subspaces

$$V_{0,r} \subseteq V_{0,r-1} \subseteq \cdots \subseteq V_{0,1} \subseteq V_{0,0} := V_0$$

such that the linear map

$$\psi_n : V_{0,n} \longrightarrow V_n, \qquad \psi_n(v) := B^n v \tag{3.4}$$

is bijective. For $v \in V_0$, we introduce the following iterated commutators

$$X_v^{(0)} := \sum_{k=1}^{d} v_k \partial_{x_k},$$

and recursively

$$X_v^{(n)} := [X_v^{(n-1)}, Y] = X_v^{(n-1)} Y - Y X_v^{(n-1)}, \qquad n \in \mathbb{N}. \tag{3.5}$$

Lemma 3.1. *Let $u \in W_B^{n,p}$. Then, for any $i \in \mathbb{N}_0$, $2i+1 \leqslant n$, and $v \in V_0$ we have*

$$X_v^{(i)} u \in W_B^{n-2i-1,p}. \tag{3.6}$$

Proof. We use an induction argument on n. If $n \leqslant 2$ there is nothing to prove because $X_v^{(0)}$ is a linear combination of the vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and the thesis follows by definition.

Assume (3.6) is true for some fixed $n > 2$ and let us prove it for $n+1$. We proceed by induction on i. For $i = 0$, again there is nothing to prove. We assume (3.6) for some $i > 0$ such that $2(i+1)+1 \leqslant n+1$ and prove it for $i+1$. If $u \in W_B^{n+1,p}$ then

$$X_v^{(i+1)} u = X_v^{(i)} Y u - Y X_v^{(i)} u.$$

Here $Yu \in W_B^{n-1,p}$ by definition and therefore $X_v^{(i)} Y u \in W_B^{n-1-(2i+1),p}$ by the inductive hypothesis on n. On the other hand $X_v^{(i)} u \in W_B^{n+1-(2i+1),p}$ by the inductive hypothesis on i and therefore $Y X_v^{(i)} u \in W_B^{n+1-(2i+1)-2,p}$ by definition. Then (3.6) holds for $n+1$, for any i, $2i+1 \leqslant n+1$, and this concludes the proof. $\square$

Proposition 3.2. *Let $u \in W_B^{n,p}$. Then for any $i \in \mathbb{N}_0$, $2i+1 \leqslant n$, we have*

$$\partial_{x_j} u \in W_B^{n-2i-1,p}, \qquad j = 1 + \bar{d}_{i-1}, \dots, \bar{d}_i.$$

Proof. By induction it is not difficult to prove that

$$X_v^{(i)} \varphi = \langle B^i v, \nabla_x \varphi \rangle, \qquad \varphi \in C^\infty.$$

Next, since ψ_i in (3.4) is bijective, for every $j = 1 + \bar{d}_{i-1}, \dots, \bar{d}_i$ there exists $w_j \in V_{0,i}$ such that $B^i w_j = \mathbf{e}_j \in V_i$. Then $f := X^{(i)}_{w_j} u \in W^{n-2i-1,p}_B$ is such that

$$\int_{\mathbb{R}^{N+1}} f(z)\varphi(z)dz = -\int_{\mathbb{R}^{N+1}} u(z)X^{(i)}_{w_j}\varphi(z)dz = -\int_{\mathbb{R}^{N+1}} u(z)\partial_{x_j}\varphi(z)dz, \quad \varphi \in C_0^\infty,$$

which means that f is the weak derivative $\partial_{x_j} u$. $\square$

By Proposition 3.2 and the definition of intrinsic Sobolev spaces we eventually infer the following:

Corollary 3.3. *Let $u \in W^{n,p}_B$. Then, for any $k \in \mathbb{N}_0$, $\beta \in \mathbb{N}_0^\beta$ such that $2k+\langle\beta\rangle_B = l \leqslant n$, we have*

$$Y^k \partial^\beta u \in W^{n-l,p}_B.$$

Corollary 3.4. *The following quasi-norms are equivalent:*

i) $|u|_{n,p,B}$;

ii)

$$\sum_{2k+\langle\beta\rangle_B=n} \|Y^k\partial^\beta u\|_p + \sum_{2k+\langle\beta\rangle_B=n-1} [Y^k\partial^\beta u]_{Y,\frac{1}{2},p}; \tag{3.7}$$

iii)

$$\begin{cases} \sum\limits_{2k+\langle\beta\rangle_B=n-1} |Y^k\partial^\beta u|_{1,p,B}, & n = 2l+1,\ l \in \mathbb{N}, \\ \sum\limits_{2k+\langle\beta\rangle_B=n-1} |Y^k\partial^\beta u|_{1,p,B} + \|Y^l u\|_p, & n = 2l,\ l \in \mathbb{N}. \end{cases} \tag{3.8}$$

Proof. By induction it is not difficult to check that $|u|_{n,p,B}$ controls all the L^p-norms of the nth-order derivatives that are compositions of Y and $\partial_{x_1}, \dots, \partial_{x_d}$ for any possible permutation, as well as the fractional quasi-norms of the $(n-1)$th-order derivatives. Then it suffices to note that, by (3.5) we have

$$X^{(0)}_v Y^n = \sum_{i=0}^n \binom{n}{i} Y^i X^{(n-i)}_v, \qquad v \in V_0; \tag{3.9}$$

then, proceeding as in the proof of Proposition 3.2 to rearrange the derivatives, we get (3.7).

Moreover, by definition we have

$$\sum_{2k+\langle\beta\rangle_B=n-1} |Y^k\partial^\beta u|_{1,p,B} = \sum_{2k+\langle\beta\rangle_B=n-1} \left(\sum_{j=1}^d \|\partial_{x_j} Y^k\partial^\beta u\|_p + [Y^k\partial^\beta u]_{Y,\frac{1}{2},p} \right),$$

and the fractional part of the quasi-norm coincides with (3.7). As for the first term in the sum we need to distinguish two cases: if $n = 2l + 1$ for some $l \in \mathbb{N}$, then we get an equivalence with (3.7) by rearranging the derivatives as in the proof of Proposition 3.2; indeed, compared to $W_B^{n-1,p}$ we have the additional set of Euclidean derivatives ∂_{x_j}, $j = 1 + \bar{d}_{l-1}, \dots, \bar{d}_l$ which can be recovered from $\sum_{j=1}^{d} \partial_{x_j} Y^l u$ by (3.9), and similarly for the mixed derivatives. If $n = 2l$ we have the derivatives $Y^l u$ that cannot be written as sums of iterated commutators and thus we get (3.8). □

3.2. Interpolation inequality and equivalence of the norms $\|\cdot\|_{W_B^{n,p}}$ and $|||\cdot|||_{W_B^{n,p}}$

Proposition 3.5. *Let $1 \leqslant n < m$ and $p \geqslant 1$. There exists $c = c(m, p, B)$ such that*

$$|u|_{n,p,B} \leqslant c\left(\varepsilon |u|_{m,p,B} + \varepsilon^{-\frac{n}{m-n}} \|u\|_p\right), \qquad u \in W_B^{m,p}, \quad \varepsilon > 0. \tag{3.10}$$

In particular the norms $\|\cdot\|_{W_B^{n,p}}$ and $|||\cdot|||_{W_B^{n,p}}$ are equivalent.

Proof. The proof is based on a two-step induction.

Step 1: case $n = 1$ and $m = 2$. The estimate

$$\|\partial_{x_i} u\|_p \lesssim \|u\|_p + \|\partial_{x_i x_i} u\|_p, \qquad i = 1, \dots, d,$$

is standard (cf. for instance [2], Chapter 5). On the other hand, by Fubini's Theorem we have

$$[u]^p_{Y,\frac{1}{2},p} = \int_{-1}^{1} \frac{dh}{|h|^{1+\frac{p}{2}}} \int_{\mathbb{R}^{N+1}} J_p(z,h)dz, \qquad J_p(z,h) := |u(e^{hY}z) - u(z)|^p.$$

By the mean value theorem along the vector field Y, for every $z \in \mathbb{R}^{N+1}$ and $h \in [-1,1]$, $h \neq 0$, there exists $|\bar{h}| \leqslant |h|$ such that $|J_p(z,h)| \leqslant |Yu(e^{\bar{h}Y}z)|^p |h|^p$: then, by a change of variable and recalling Remark 2.11,

$$[u]^p_{Y,\frac{1}{2},p} \leqslant 2\|Yu\|_p^p \int_0^1 |h|^{\frac{p}{2}-1} dh \leqslant \frac{4}{p} \|Yu\|_p^p.$$

Thus we obtain

$$|u|_{1,p,B} \lesssim |u|_{2,p,B} + \|u\|_p. \tag{3.11}$$

The thesis follows by a scaling argument: indeed, applying (3.11) to $u(D_{\varepsilon^{-1}} \cdot)$, by (2.9) we get

$$\varepsilon^{1-\frac{\mathbf{d}}{p}} |u|_{1,p,B} \lesssim \varepsilon^{2-\frac{\mathbf{d}}{p}} |u|_{2,p,B} + \varepsilon^{-\frac{\mathbf{d}}{p}} \|u\|_p.$$

Step 2: induction on n, m with $m = n + 1$. We first prove the preliminary interpolation inequality:

$$\|Yu\|_p \lesssim \varepsilon[Yu]_{Y,\frac{1}{2},p} + \varepsilon^{-1}[u]_{Y,\frac{1}{2},p}, \qquad u \in W_B^{3,p}, \quad \varepsilon > 0. \tag{3.12}$$

We have

$$u(e^Y z) - u(z) - Yu(z) = \int_0^1 \left(Yu(e^{\delta Y} z) - Yu(z)\right) d\delta,$$

and therefore

$$\|Yu\|_p^p = \int_{\mathbb{R}^{N+1}} \left| u(e^Y z) - u(z) - \int_0^1 \left(Yu(e^{\delta Y} z) - Yu(z)\right) d\delta \right|^p dz \leqslant$$

(by the triangular and Hölder inequalities)

$$\lesssim \int_{\mathbb{R}^{N+1}} |u(e^Y z) - u(z)|^p dz + \int_{\mathbb{R}^{N+1}} \int_0^1 |Yu(e^{\delta Y} z) - Yu(z)|^p d\delta dz =: I_1 + I_2,$$

where

$$I_2 = \int_{\mathbb{R}^{N+1}} \int_0^1 \frac{|Yu(e^{\delta Y} z) - Yu(z)|^p}{\delta^{1+\frac{p}{2}}} \delta^{1+\frac{p}{2}} d\delta dz \leqslant \frac{1}{2}[Yu]_{Y,\frac{1}{2},p}$$

and

$$I_1 \leqslant \int_{\mathbb{R}^{N+1}} \int_0^1 |u(e^Y z) - u(e^{\delta Y} z)|^p d\delta dz + \int_{\mathbb{R}^{N+1}} \int_0^1 |u(e^{\delta Y} z) - u(z)|^p d\delta dz =$$

(by the change of variables $z' = e^Y z$ and $\delta' = \delta - 1$)

$$= \int_{\mathbb{R}^{N+1}} \int_{-1}^0 |u(e^{\delta' Y} z') - u(z')|^p d\delta' dz' + \int_{\mathbb{R}^{N+1}} \int_0^1 |u(e^{\delta Y} z) - u(z)|^p d\delta dz \leqslant [u]_{Y,\frac{1}{2},p},$$

reasoning as for I_2 in the last step. Then (3.12) follows by a scaling argument.

Next we prove that if, for some $\bar{n} \in \mathbb{N}$, (3.10) holds with $n = \bar{n}$, $m = \bar{n} + 1$ then it also holds with $n = \bar{n} + 1$ and $m = \bar{n} + 2$. By Step 1, (3.8) and (3.12), if $\bar{n} + 1$ is *even* we have

$$
\begin{aligned}
|u|_{\bar n+1,p,B} &\lesssim \sum_{2k+\langle\beta\rangle_B=\bar n} |Y^k\partial^\beta u|_{1,p,B} + \|Y^{\frac{\bar n+1}{2}}u\|_p \\
&\leqslant c_1 \sum_{2k+\langle\beta\rangle_B=\bar n} \Big(\varepsilon\Big(|Y^k\partial^\beta u|_{2,p,B} + [Y^{\frac{\bar n+1}{2}}u]_{Y,\frac12,p}\Big) \\
&\qquad + \varepsilon^{-1}\Big(\|Y^k\partial^\beta u\|_p + [Y^{\frac{\bar n-1}{2}}u]_{Y,\frac12,p}\Big)\Big) \\
&\leqslant c_1\left(\varepsilon|u|_{\bar n+2,p,B} + \varepsilon^{-1}|u|_{\bar n,p,B}\right) \leqslant
\end{aligned}
$$

(by the inductive hypothesis)

$$
\leqslant c_1\varepsilon|u|_{\bar n+2,p,B} + c_1c_2\varepsilon^{-1}\varepsilon_1|u|_{\bar n+1,p,B} + c_1c_2\varepsilon^{-1}\varepsilon_1^{-\bar n}\|u\|_p.
$$

If $\bar n+1$ is *odd*, by (3.8) we derive the same estimate only exploiting Step 1. To conclude it suffices to take $\varepsilon_1 = \frac{\varepsilon}{2c_1c_2}$.

Step 3: backward induction on n. Let $m\in\mathbb{N}$, $m>2$, be fixed. We prove that if (3.10) is true for m and $n=\bar n$ for some $\bar n\in\{2,\dots,m-1\}$, then it is also true for m and $n=\bar n-1$.

By Step 2 we have

$$
|u|_{\bar n-1,p,B} \lesssim \varepsilon_1|u|_{\bar n,p,B} + \varepsilon_1^{-(\bar n-1)}\|u\|_p \lesssim
$$

(by the inductive hypothesis)

$$
\lesssim \varepsilon_1\left(\varepsilon_2|u|_{m,p,B} + \varepsilon_2^{-\frac{\bar n}{m-\bar n}}\|u\|_p\right) + \varepsilon_1^{-(\bar n-1)}\|u\|_p.
$$

Letting now $\varepsilon=\varepsilon_1\varepsilon_2$ and $\varepsilon_1=\varepsilon^{\frac{1}{m-(\bar n-1)}}$, we get

$$
|u|_{\bar n-1,p,B} \lesssim \varepsilon|u|_{m,p,B} + \varepsilon^{-\frac{\bar n-1}{m-(\bar n-1)}}\|u\|_p,
$$

which concludes the proof. □

4. Taylor expansion in $W_B^{n,p}$

According to [18], the *n-th order B-Taylor polynomial of u around* $\zeta=(s,\xi)$ is formally defined as

$$
T_nu(\zeta,z) := \sum_{0\leqslant 2k+\langle\beta\rangle_B\leqslant n} \frac{(t-s)^k(x-e^{(t-s)B}\xi)^\beta}{k!\,\beta!} Y^k\partial_\xi^\beta u(s,\xi), \qquad z=(t,x)\in\mathbb{R}^{N+1}, \tag{4.1}
$$

with $\langle\beta\rangle_B$ as in (3.2). The main result of this section is the following.

Theorem 4.1. *Let $n \in \mathbb{N}_0$ and $p \geqslant 1$. There exists $c = c(n, p, B)$ such that, for any $u \in W_B^{n+1,p} \cap C^\infty$ we have*

$$\|u - T_n u(\cdot \circ \zeta, \cdot)\|_p \leqslant c\|\zeta\|_B^{n+1} \|u\|_{W_B^{n+1,p}}, \qquad \zeta \in \mathbb{R}^{N+1}. \tag{4.2}$$

The proof is based on an induction procedure developed in [18] to derive the $C_B^{n,\alpha}$ estimate of the remainder. For completeness, here we give a fairly comprehensive presentation of the main lines, and refer to [18] for the details of the construction. To simplify the exposition we first split the proof in different steps, corresponding to particular cases of (4.2).

Lemma 4.2. *There exists $c = c(n, p)$ such that, for any $u \in W_B^{n+1,p} \cap C^\infty$ and $\delta \in \mathbb{R}$, we have*

$$\|u(e^{\delta Y} \cdot) - \sum_{k=0}^{[n/2]} \frac{\delta^k}{k!} Y^k u\|_p \leqslant c|\delta|^{\frac{n+1}{2}} \|u\|_{W_B^{n+1,p}}, \tag{4.3}$$

where $[n/2]$ denotes the integer part of $n/2$.

Proof. We first check that, for $u \in W_B^{1,p}$, we have

$$\|u(e^{\delta Y} \cdot) - u\|_p \lesssim |\delta|^{\frac{1}{2}} [u]_{Y,\frac{1}{2},p}, \qquad |\delta| \leqslant 1. \tag{4.4}$$

Without loss of generality, we assume $\delta \in (0, 1]$. Adding and subtracting $u(e^{hY} \cdot)$ and integrating on $h \in [0, \delta]$, we have

$$\delta \|u(e^{\delta Y} \cdot) - u\|_p^p \lesssim \int_0^\delta \|u(e^{\delta Y} \cdot) - u(e^{hY} \cdot)\|_p^p dh + \int_0^\delta \|u(e^{hY} \cdot) - u\|_p^p dh =: I_1 + I_2.$$

Then we have

$$I_2 = \int_{\mathbb{R}^{N+1}} \int_0^\delta \frac{|u(e^{hY} z) - u(z)|^p}{h^{1+\frac{p}{2}}} h^{1+\frac{p}{2}} dh\, dz \leqslant \frac{1}{2} |\delta|^{1+\frac{p}{2}} [u]_{Y,\frac{1}{2},p}^p.$$

By the change of variable $\bar{z} = e^{\delta Y} z$, $\bar{h} = h - \delta$ and Remark 2.11, the term I_1 is analogous, and thus we get (4.4). Similarly, we also see that if $u \in W_B^{2,p}$ then

$$\|u(e^{\delta Y} \cdot) - u\|_p \lesssim |\delta| \|Y u\|_p \lesssim |\delta| \|u\|_{W_B^{2,p}}. \tag{4.5}$$

More generally, for $u \in W_B^{n+1,p} \cap C^\infty$, by the mean-value theorem along the vector field Y, for some $\bar{\delta}$ such that $|\bar{\delta}| \leqslant |\delta|$ we have

$$u(e^{\delta Y}z) - \sum_{i=0}^{[n/2]} \frac{\delta^i}{i!} Y^i u(z) = \frac{\delta^{[n/2]}}{[n/2]!} \left(Y^{[n/2]} u(e^{\bar{\delta} Y} z) - Y^{[n/2]} u(z) \right). \tag{4.6}$$

Now, if $n = 2h$ for some $h \in \mathbb{N}$ then $Y^h u \in W_B^{1,p}$: thus (4.3) follows by combining (4.6) with (4.4) applied to $Y^h u$. Similarly, if $n = 2h + 1$ for some $h \in \mathbb{N}$ then (4.3) follows by combining (4.6) with (4.5) applied to $Y^h u \in W_B^{2,p}$. □

Lemma 4.3. *There exists $c = c(p, B)$ such that*

$$\|u(\cdot \circ (0, \xi)) - u\|_p \leqslant c|\xi|_B \|u\|_{W_B^{1,p}}, \qquad u \in W_B^{1,p}, \quad \xi \in \mathbb{R}^N. \tag{4.7}$$

Proof. We have the standard inequality

$$\|u(e^{\delta \partial_{x_i}} \cdot) - u\|_p \leqslant |\delta| \|\partial_{x_i} u\|_p, \qquad i = 1, \dots, d. \tag{4.8}$$

Notice that, for $u \in W_B^{1,p}$, we have L^p-bounds only on the first d spatial derivatives. Thus, in order to prove (4.7) we must exploit estimate (4.4) and connect any arbitrary point $z = (t, x) \in \mathbb{R}^{N+1}$ to $z \circ (0, \xi) = (t, x + \xi)$ through a chain of integral curves associated only to the vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and Y.

To do so, we define a sequence of points $(z_k = (t, x(k)))_{k=0,\cdots,r}$ adjusting, at any step k, the set of variables of the layer V_k in (3.3). Following [18], Lemma 4.22 we set

$$v_0 = \frac{\xi^{[0]}}{|\xi^{[0]}|}, \qquad |\delta_0| = |\xi^{[0]}|,$$

and

$$z_{-1} = z, \qquad z_0 = \gamma_{v_0, \delta_0}^{(0)}(z_{-1}) := e^{\delta_0 \langle v_0, \nabla \rangle} z = (t, x + \xi^{[0]}).$$

For $k = 1, \dots, r$ let

$$z_k = \gamma_{v_k, \delta_k}^{(k)}(z_{k-1}) := e^{-\delta_k^2 Y} \left(\gamma_{v_k, -\delta_k}^{(k-1)} \left(e^{\delta_k^2 Y} \left(\gamma_{v_k, \delta_k}^{(k-1)}(z_{k-1}) \right) \right) \right), \qquad \delta_k = |w_k|^{\frac{1}{2k+1}},$$

where $v_k = w_k / |w_k|$ and w_k is the unique vector in $V_{0,k} \subset V_0$ such that $B^k w_k = \xi^{[k]} + x^{[k]} - x^{[k]}(k-1)$. Importantly, it can be proven by induction that, for any $v \in V_0$ we have

$$\gamma_{v,\delta}^{(k)}(t, x) = (t, x + S_k(\delta) v), \qquad S_k(\delta) v = (-1)^k \sum_{\substack{h \in \mathbb{N}^k \\ |h| \leqslant r}} \frac{(-B)^{|h|}}{h!} \delta^{2|h|+1} v \in \bigoplus_{j=k}^{r} V_j.$$

In other words the flow $\gamma_{v,\delta}^{(k)}$ only affects the set of variables $[k : r]$. Moreover, $\gamma_{v_r, \delta_r}^{(r)}(z_{r-1}) = (t, x + \xi)$ by construction. Notice also that at any step, δ_k does not depend on x and also, the specific choice of w_k implies $\delta_k \leqslant c_B |\xi|_B$ (cf. [18], Lemma 4.22).

We are ready to prove (4.7): by the Minkowski inequality we have

$$\|u(\cdot \circ (0,\xi)) - u\|_p^p \lesssim \sum_{k=0}^{r} \int_{\mathbb{R}^{N+1}} |u(z_k) - u(z_{k-1})|^p dz = \sum_{k=0}^{r} \|u(\gamma^{(k)}_{v_k,\delta_k}) - u\|_p^p,$$

where we exploited the changes of variables $z' = z_{k-1} \equiv z_{k-1}(z)$ in the last step: here we use the fact that, by Remark 2.11, the Jacobian of the change of variables has determinant equal to one. Then the proof is completed once we have proved that, for any $k, i \in \{0, \dots, r\}$ we have

$$\|u(\gamma^{(k)}_{v_i,\delta_i}) - u\|_p \lesssim \delta_i \|u\|_{W^{1,p}_B} \lesssim |\xi|_B \|u\|_{W^{1,p}_B}. \tag{4.9}$$

We proceed by induction on k. The case $k = 0$ follows from (4.8). Assume now (4.9) holds for some $k \in \{0, \dots, r-1\}$: as before we have

$$\begin{aligned}\|u(\gamma^{(k+1)}_{v_i,\delta_i}) - u\|_p \lesssim & \|u(\gamma^{(k)}_{v_i,\delta_i}) - u\|_p + \|u(e^{\delta_i^2 Y} \cdot) - u\|_p \\ & + \|u(\gamma^{(k)}_{v_i,-\delta_i}) - u\|_p + \|u(e^{-\delta_i^2 Y} \cdot) - u\|_p,\end{aligned}$$

and the thesis follows from the inductive step and (4.4). $\square$

We are ready to prove Theorem 4.1.

Proof of Theorem 4.1. We prove that, for any $n \in \mathbb{N}_0$, $\zeta \in \mathbb{R}^{N+1}$ and $u \in W^{n+1,p}_B \cap C^\infty$, we have

$$\|u(\cdot \circ \zeta) - T_n u(\cdot, \cdot \circ \zeta)\|_p \leqslant c\|\zeta\|_B^{n+1} \|u\|_{W^{n+1,p}_B}. \tag{4.10}$$

Estimate (4.2) follows from (4.10) since

$$\|u - T_n u(\cdot \circ \zeta, \cdot)\|_p = \|u(\cdot \circ \zeta^{-1}) - T_n u(\cdot, \cdot \circ \zeta^{-1})\|_p.$$

Now, for $z = (t, x), \zeta = (s, \xi) \in \mathbb{R}^{N+1}$, we write

$$u(z \circ \zeta) - T_n u(z, z \circ \zeta) = \underbrace{u(z \circ \zeta) - T_n u(e^{sY} z, z \circ \zeta)}_{=:F_1(z,\zeta)} + \underbrace{T_n u(e^{sY} z, z \circ \zeta) - T_n u(z, z \circ \zeta)}_{=:F_2(z,\zeta)}.$$

By definition

$$e^{sY} z = (t + s, e^{sB} x), \qquad z \circ \zeta = (t + s, \xi + e^{sB} x).$$

Hence $F_1(z, \zeta)$ contains increments that only differ in the spatial variables, while $F_2(z, \zeta)$ contains increments that only differ along Y.

To estimate $F_2(z,\zeta)$ we first notice that the increments in the Taylor polynomials appearing in $F_2(z,\zeta)$ are given by

$$(e^{sY}z)^{-1}\circ(z\circ\zeta)=(0,\xi),\qquad z^{-1}\circ(z\circ\zeta)=(s,\xi),$$

thus we have

$$\begin{aligned}F_2(z,\zeta)&=\sum_{\langle\beta\rangle_B\leqslant n}\frac{\xi^\beta}{\beta!}(\partial^\beta u)(e^{sY}z)-\sum_{2k+\langle\beta\rangle_B\leqslant n}\frac{s^k\xi^\beta}{k!\,\beta!}(Y^k\partial^\beta u)(z)\\&=\sum_{\langle\beta\rangle_B\leqslant n}\frac{\xi^\beta}{\beta!}\left((\partial^\beta u)(e^{sY}z)-\sum_{2k\leqslant n-\langle\beta\rangle_B}\frac{s^k}{k!}(Y^k\partial^\beta u)(z)\right).\end{aligned}$$

Taking the L^p norm in dz and using (4.3) for $\partial^\beta u\in W_B^{n-\langle\beta\rangle_B}$ by Corollary 3.3, we get

$$\|F_2(\cdot,\zeta)\|_p\lesssim\sum_{\langle\beta\rangle_B\leqslant n}|\xi|_B^{\langle\beta\rangle_B}|s|^{\frac{n-\langle\beta\rangle_B+1}{2}}\|\partial^\beta u\|_{W_B^{n+1-\langle\beta\rangle_B,p}}\lesssim\|\zeta\|_B^{n+1}\|u\|_{W_B^{n+1,p}}.$$

It remains to prove

$$\|F_1(\cdot,\zeta)\|_p\lesssim\|\zeta\|_B^{n+1}\|u\|_{W_B^{n+1,p}}.\tag{4.11}$$

First notice that, by a change of variable we have

$$\|F_1(\cdot,\zeta)\|_p=\|u(\cdot\circ(0,\xi))-T_nu(\cdot,\cdot\circ(0,\xi))\|_p.$$

The case $n=0$ corresponds to Lemma 4.3. Next we assume that (4.11) holds for $n=\bar n-1$ and prove it for $n=\bar n$. We have

$$\begin{aligned}&u(z\circ(0,\xi))-T_{\bar n}u(z,z\circ(0,\xi))\\&\quad=u(t,x+\xi)-T_{\bar n}u((t,x),(t,x+\xi))\\&\quad=\underbrace{u(t,x+\xi)-T_{\bar n}u((t,\bar x),(t,x+\xi))}_{=:F_{11}(z,\zeta)}+\underbrace{T_{\bar n}u((t,\bar x),(t,x+\xi))-T_{\bar n}u((t,x),(t,x+\xi))}_{=:F_{12}(z,\zeta)},\end{aligned}$$

where $\bar x$ the point in $\mathbb{R}^N$ defined by

$$\bar x^{[i]}=\begin{cases}x^{[i]}&\text{if }2i+1\leqslant\bar n,\\x^{[i]}+\xi^{[i]}&\text{if }2i+1>\bar n.\end{cases}$$

Notice that $(x+\xi-\bar x)^\beta=\xi^\beta$ for $\langle\beta\rangle_B\leqslant\bar n$ and $|x-\xi-\bar x|_B\leqslant|\xi|_B$. For $x\in\mathbb{R}^N$, we introduce the notation

$$x^{[i:j]} = \sum_{k=i}^{j} x^{[k]} \qquad 0 \leqslant i < j \leqslant r.$$

Then

$$T_{\bar{n}} u((t,\bar{x}),(t,x+\xi)) = \sum_{\langle\beta\rangle_B \leqslant \bar{n}} \frac{1}{\beta!} \partial^\beta u(t, x + \xi^{[[\frac{\bar{n}+1}{2}]:r]}) \xi^\beta. \tag{4.12}$$

Now, since $u \in W_B^{\bar{n}+1,p}$ has weak derivatives of order $[\frac{\bar{n}+1}{2i+1}]$ in any direction of the increments $[i]$, $i \leqslant [\frac{\bar{n}}{2}]$, it is not difficult to check, similarly to (4.3) that

$$\|F_{11}(\cdot,\zeta)\|_p \lesssim |\xi|_B^{\bar{n}+1} \|u\|_{W_B^{n+1,p}}.$$

On the other hand, by (4.12), we have

$$F_{12}(z,\zeta) = \sum_{\langle\beta\rangle_B \leqslant \bar{n}} \frac{\xi^\beta}{\beta!} \left(\partial^\beta u(t,\bar{x}) - \partial^\beta u(t,x)\right).$$

Then, taking the L^p norm in dz, we have

$$\|F_{12}(\cdot,\zeta)\|_p \lesssim \sum_{\langle\beta\rangle_B \leqslant \bar{n}} \frac{|\xi|_B^{\langle\beta\rangle_B}}{\beta!} \|\partial^\beta u(\cdot \circ \xi^{[[\frac{\bar{n}+1}{2}]:r]}) - \partial^\beta u\|_p$$

Now we use the inductive hypothesis on $\partial^\beta u \in W_B^{\bar{n}-\langle\beta\rangle_B+1,p}$ for $|\beta|_B \geqslant 1$, and finally get

$$\|F_{12}(\cdot,\zeta)\|_p \lesssim \sum_{\langle\beta\rangle_B \leqslant \bar{n}} \frac{|\xi|_B^{\langle\beta\rangle_B}}{\beta!} |\xi|_B^{\bar{n}-\langle\beta\rangle_B+1} \|\partial^\beta u\|_{W_B^{\bar{n}-\langle\beta\rangle_B+1,p}} \lesssim |\xi|_B^{\bar{n}+1} \|u\|_{W_B^{n+1,p}}. \quad \square$$

Remark 4.4. By Theorem 4.1, for any $i \in \{\bar{d}_{j-1}+1, \dots, \bar{d}_j\}$ with $2j+1 > n$ we have in particular

$$\|u(e^{\delta\partial_{x_i}} \cdot) - u\|_p \lesssim |\delta|^{\frac{n+1}{2j+1}} \|u\|_{W_B^{n+1,p}}, \qquad u \in W_B^{n+1,p}.$$

Then, using Fubini's Theorem it is straightforward to check that, for any $\varepsilon > 0$

$$[u]_{\partial_{x_i}, \frac{n+1}{2j+1}-\varepsilon, p} \lesssim \varepsilon^{-1} \|u\|_{W_B^{n+1,p}}, \qquad u \in W_B^{n+1,p}.$$

Together with Corollary 3.3 this gives the expected regularity in any spatial direction, which is not prescribed a priori by the definition of the spaces. Also, by Corollary 3.3 we can further infer that, for any $k \in \mathbb{N}_0$, $\beta \in \mathbb{N}_0^N$ with $2j+1 > n-2k-\langle\beta\rangle_B \geqslant 0$ we have

$$[Y^k \partial^\beta u]_{\partial_{x_i}, \frac{n+1-2k-\langle\beta\rangle_B}{2j+1}-\varepsilon, p} \lesssim \varepsilon^{-1} \|u\|_{W_B^{n+1,p}}, \qquad u \in W_B^{n+1,p}.$$

5. Approximation and interpolation

5.1. Approximation in $W_B^{n,p}$

Let φ be a test function supported on $\|z\|_B \leqslant 1$ with unitary integral. Following [23], we define the n-th order approximation for $u \in W_B^{n,p}$ as

$$u_{n,\varepsilon}(z) := \int_{\mathbb{R}^{N+1}} T_{n-1}u(\zeta, z)\varphi\left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)\right) \frac{d\zeta}{\varepsilon^{\mathbf{d}}}, \qquad \varepsilon > 0, \tag{5.1}$$

where $T_n u(\zeta, z)$ is the B-Taylor polynomial in (4.1) and $\mathbf{d}$ the homogeneous dimension of $\mathbb{R}^{N+1}$. Notice that

$$\int_{\mathbb{R}^{N+1}} \varphi\left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)\right) \frac{d\zeta}{\varepsilon^{\mathbf{d}}} = \int_{\|\zeta\|_B \leqslant 1} \varphi(\zeta) d\zeta = 1.$$

We also recall the useful Lemma 3.2 from [23] which still holds for functions in $W_B^{n,p}$: indeed, its proof relies only on basic algebraic rules of derivation, namely the Leibniz formula and the chain rule for compositions with smooth functions.

Lemma 5.1. *For any $u \in W_B^{n,p}$ and $z, \zeta \in \mathbb{R}^{N+1}$ we have*

$$\begin{aligned} \partial_{x_i} T_n u(\zeta, z) &= T_{n-1}(\partial_i u)(\zeta, z), \qquad && n \geqslant 1,\ i = 1, \dots, d, \\ Y_z T_n u(\zeta, z) &= T_{n-2}(Yu)(\zeta, z), \qquad && n \geqslant 2. \end{aligned}$$

Theorem 5.2 *(Approximation). Let $n, m \in \mathbb{N}$ with $n < m$. There exist constants $c_1 = c(n, p, B)$ and $c_2 = c(n, m, p, B)$ such that for any $u \in W_B^{n,p}$ and $0 < \varepsilon \leqslant 1$ we have*

$$\|u - u_{n,\varepsilon}\|_p \leqslant c_1 \varepsilon^n \|u\|_{W_B^{n,p}}, \tag{5.2}$$

$$\|u_{n,\varepsilon}\|_{W_B^{m,p}} \leqslant c_2 \varepsilon^{n-m} \|u\|_{W_B^{n,p}}. \tag{5.3}$$

Proof. We denote by $\mathfrak{D}^l$ any weak derivative of intrinsic order l, that is $\mathfrak{D}^l = Y^k \partial_x^\beta$ with $2k + \langle \beta \rangle_B = l$, and let

$$I_\varepsilon^{(n,l)} u(z) := \int_{\mathbb{R}^{N+1}} (T_{n-1}u(\zeta, z) - u(z))(\mathfrak{D}^l \varphi)\left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)\right) \frac{d\zeta}{\varepsilon^{\mathbf{d}}}.$$

We prove the following preliminary estimates for $u \in W_B^{n,p} \cap C_0^\infty$:

$$\|I_\varepsilon^{(n,l)} u\|_p \lesssim \varepsilon^{n-l} \|u\|_{W_B^{n,p}}, \tag{5.4}$$

$$[u_{0,\varepsilon}]_{Y, \frac{1}{2}, p} \lesssim \varepsilon^{-1} \|u\|_p, \tag{5.5}$$

$$[I_\varepsilon^{(n,l)} u]_{Y, \frac{1}{2}, p} \lesssim \varepsilon^{n-l-1} \|u\|_{W_B^{n,p}}. \tag{5.6}$$

First observe that by (2.4), (2.5) and the change of variable $\eta = D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)$, we have

$$\begin{aligned} I_\varepsilon^{(n,l)} u(z) := & \, \varepsilon^{-l} \int_{\mathbb{R}^{N+1}} (T_{n-1}u(\zeta, z) - u(z)) \, (\mathfrak{D}^l \varphi) \left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z) \right) \frac{d\zeta}{\varepsilon^{\mathbf{d}}} \\ = & \, \varepsilon^{-l} \int_{\|\eta\|_B \leqslant 1} \left(T_{n-1}u(z \circ (D_\varepsilon \eta)^{-1}, z) - u(z) \right) \mathfrak{D}^l \varphi(\eta) d\eta. \end{aligned}$$

Then, by Minkowski integral inequality and Theorem 4.1, we get

$$\|I_\varepsilon^{(n,l)} u\|_p^p \leqslant \varepsilon^{-lp} \int_{\|\zeta\|_B \leqslant 1} \|T_{n-1}u(\cdot \circ (D_\varepsilon \zeta)^{-1}, \cdot) - u\|_p^p |\mathfrak{D}^l \varphi(\zeta)| d\zeta \leqslant$$

(since $\|D_\varepsilon \zeta\|_B = \varepsilon \|\zeta\|_B$)

$$\leqslant \varepsilon^{-lp} \int_{\|\zeta\|_B \leqslant 1} \varepsilon^{np} \|\zeta\|_B^{np} \|u\|_{W_B^{n,p}}^p |\mathfrak{D}^l \varphi(\zeta)| d\zeta \lesssim \varepsilon^{(n-l)p} \|u\|_{W_B^{n,p}}^p$$

which proves (5.4). By a similar argument, we have

$$\begin{aligned} [u_{0,\varepsilon}]_{Y,\frac{1}{2},p}^p = & \int_{\mathbb{R}^{N+1}} \int_{|h| \leqslant 1} \left| \int_{\mathbb{R}^{N+1}} u(\zeta) \left(\varphi \left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ e^{hY} z) \right) - \varphi \left(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z) \right) \right) \frac{d\zeta}{\varepsilon^{\mathbf{d}}} \right|^p \\ & \times \frac{dh}{|h|^{1+\frac{p}{2}}} dz \\ = & \int_{|h| \leqslant 1} \int_{\mathbb{R}^{N+1}} \left| \int_{\mathbb{R}^{N+1}} u(z \circ (D_\varepsilon \zeta)^{-1}) \left(\varphi(e^{\frac{h}{\varepsilon^2} Y} \zeta) - \varphi(\zeta) \right) d\zeta \right|^p dz \frac{dh}{|h|^{1+\frac{p}{2}}} \leqslant \end{aligned}$$

(by Minkowski inequality)

$$\leqslant \|u\|_p^p \left(\int_{|h| > \varepsilon^2} + \int_{|h| \leqslant \varepsilon^2} \right) \left(\int_{\mathbb{R}^{N+1}} \left| \varphi(e^{\frac{h}{\varepsilon^2} Y} \zeta) - \varphi(\zeta) \right| d\zeta \right)^p \frac{dh}{|h|^{1+\frac{p}{2}}} =: I_1 + I_2$$

By the triangular and Hölder inequalities we have

$$I_1 \leqslant \|u\|_p^p \int_{|h| > \varepsilon^2} \left(\int_{\|\zeta\|_B \leqslant 1} 2\varphi(\zeta) d\zeta \right)^p \frac{dh}{|h|^{1+\frac{p}{2}}} \lesssim \|u\|_p^p \|\varphi\|_p^p \int_{|h| > \varepsilon^2} \frac{dh}{|h|^{1+\frac{p}{2}}} \lesssim \varepsilon^{-p} \|u\|_p^p \|\varphi\|_p^p.$$

Next, noting that $e^{\frac{h}{\varepsilon^2} Y} \zeta = \zeta \circ (\frac{h}{\varepsilon^2}, 0)$, by (2.7) we have

$$\|e^{\frac{h}{\varepsilon^2}Y}\zeta\|_B \leqslant m\left(\|\zeta\|_B + \frac{\sqrt{|h|}}{\varepsilon}\right) \leqslant 2m$$

in the integration set of I_2, and therefore, again by Hölder inequality, we have

$$I_2 \lesssim \|u\|_p^p \int\limits_{|h|\leqslant\varepsilon^2}\int\limits_{\mathbb{R}^{N+1}} \left|\varphi(e^{\frac{h}{\varepsilon^2}Y}\zeta) - \varphi(\zeta)\right|^p d\zeta \frac{dh}{|h|^{1+\frac{p}{2}}} \leqslant \varepsilon^{-p}\|u\|_p^p[\varphi]^p_{Y,\frac{1}{2},p}$$

where the last inequality easily follows by a change of variables, and this proves (5.5).

Lastly, (5.6) requires more attention. We have

$$[I_\varepsilon^{(n,l)}u]^p_{Y,\frac{1}{2},p} \leqslant S_1 + S_2$$

where

$$S_1 = \int\limits_{\mathbb{R}^{N+1}}\int\limits_{|h|\leqslant 1}\left|\int\limits_{\mathbb{R}^{N+1}} \left(T_{n-1}u(\zeta, e^{hY}z) - u(e^{hY}z)\right) \times \right.$$
$$\left. \times \left((\mathfrak{D}^l\varphi)\left(D_{\varepsilon^{-1}}(\zeta^{-1}\circ e^{hY}z)\right) - (\mathfrak{D}^l\varphi)\left(D_{\varepsilon^{-1}}(\zeta^{-1}\circ z)\right)\right)\frac{d\zeta}{\varepsilon^{\mathbf{d}}}\right|^p \frac{dh}{|h|^{1+\frac{p}{2}}}dz,$$

$$S_2 = \int\limits_{\mathbb{R}^{N+1}}\int\limits_{|h|\leqslant 1}\left|\int\limits_{\mathbb{R}^{N+1}} J_n(\zeta,z)\mathfrak{D}^l\varphi\left(D_{\varepsilon^{-1}}(\zeta^{-1}\circ z)\right)\frac{d\zeta}{\varepsilon^{\mathbf{d}}}\right|^p \frac{dh}{|h|^{1+\frac{p}{2}}}dz$$

with

$$J_n(\zeta,z) = T_{n-1}u(\zeta, e^{hY}z) - u(e^{hY}z) - \left(T_{n-1}u(\zeta,z) - u(z)\right).$$

The term S_1 can be controlled as $[u_{0,\varepsilon}]_{Y,\frac{1}{2},p}$: for simplicity here we assume that we can control the support of the increment of $\mathfrak{D}^l\varphi$ independently of ε on the whole integration set of h to exploit a Hölder inequality (otherwise we can just split the integration set and proceed as for the terms I_1 and I_2 of $[u_{0,\varepsilon}]_{Y,\frac{1}{2},p}$), then we have

$$S_1 = \varepsilon^{-lp}\int\limits_{|h|\leqslant 1}\int\limits_{\mathbb{R}^{N+1}}\left|\int\limits_{\mathbb{R}^{N+1}} \left(T_{n-1}u(e^{hY}z\circ(D_\varepsilon\zeta)^{-1}, e^{hY}z) - u(e^{hY}z)\right)\times\right.$$
$$\left.\times\left(\mathfrak{D}^l\varphi(e^{\frac{h}{\varepsilon^2}Y}\zeta) - \mathfrak{D}^l\varphi(\zeta)\right)d\zeta\right|^p dz\frac{dh}{|h|^{1+\frac{p}{2}}}$$
$$\leqslant \varepsilon^{-lp}\int\limits_{|h|\leqslant 1}\int\limits_{\mathbb{R}^{N+1}} \|T_{n-1}u(\cdot\circ(D_\varepsilon\zeta)^{-1},\cdot) - u\|_p^p|\mathfrak{D}^l\varphi(e^{\frac{h}{\varepsilon^2}Y}\zeta) - \mathfrak{D}^l\varphi(\zeta)|^p d\zeta\frac{dh}{|h|^{1+\frac{p}{2}}} \lesssim$$

(by Theorem 4.1)

$$\lesssim \varepsilon^{(n-l)p}\|u\|^p_{W^{n,p}_B}\varepsilon^{-p}[\mathfrak{D}^l\varphi]^p_{Y,\frac{1}{2},p} \lesssim e^{(n-l-1)p}\|u\|^p_{W^{n,p}_B}.$$

On the other hand, we have $S_2 = S_{21} + S_{22}$ where

$$S_{21} := \int\limits_{\mathbb{R}^{N+1}} \int\limits_{|h|\leqslant\varepsilon^2} \Big| \int\limits_{\mathbb{R}^{N+1}} J_n(\zeta,z)\varepsilon^{-l}(\mathfrak{D}^l\varphi)\left(D_{\varepsilon^{-1}}(\zeta^{-1}\circ z)\right)\frac{d\zeta}{\varepsilon^{\mathbf{d}}}\Big|^p \frac{dh}{|h|^{1+\frac{p}{2}}}dz \leqslant$$

(reasoning as above)

$$\begin{aligned}
&\leqslant \varepsilon^{-lp}\int\limits_{\mathbb{R}^{N+1}}\int\limits_{|h|>\varepsilon^2} 2\|T_{n-1}u(\cdot\circ(D_\varepsilon\zeta)^{-1},\cdot)-u\|^p_p|\mathfrak{D}^l\varphi(\zeta)|^p\frac{dh}{|h|^{1+\frac{p}{2}}}d\zeta\\
&\lesssim \varepsilon^{(n-l)p}\|u\|^p_{W^{n,p}_B}\|\mathfrak{D}^l\varphi\|^p_p\int\limits_{|h|>\varepsilon^2}\frac{dh}{|h|^{1+\frac{p}{2}}} \lesssim \varepsilon^{(n-l-1)p}\|u\|^p_{W^{n,p}_B}
\end{aligned}$$

and

$$S_{22} := \int\limits_{\mathbb{R}^{N+1}} \int\limits_{\varepsilon^2<|h|\leqslant 1} \Big| \int\limits_{\mathbb{R}^{N+1}} J_n(\zeta,z)\varepsilon^{-l}(\mathfrak{D}^l\varphi)\left(D_{\varepsilon^{-1}}(\zeta^{-1}\circ z)\right)\frac{d\zeta}{\varepsilon^{\mathbf{d}}}\Big|^p \frac{dh}{|h|^{1+\frac{p}{2}}}dz.$$

To estimate S_{22}, assume for a moment that $n \geqslant 3$: then by Theorem 4.1 we have

$$\begin{aligned}
S_{22} \leqslant \varepsilon^{-lp}\int\limits_{|h|\leqslant\varepsilon^2}\Bigg(\int\limits_{\|\zeta\|_B\leqslant 1}\Bigg(\int\limits_{\mathbb{R}^{N+1}}|h|^p\Big[\int\limits_0^1|T_{n-3}Yu(z\circ(D_\varepsilon\zeta)^{-1},e^{\lambda hY}z)-\\
-Yu(e^{\lambda hY}z)|d\lambda\Big]^p dz\Bigg)^{\frac{1}{p}}|\mathfrak{D}\varphi(\zeta)|d\zeta\Bigg)^p\frac{dh}{|h|^{1+\frac{p}{2}}}.
\end{aligned}\tag{5.7}$$

Notice that

$$z\circ(D_\varepsilon\zeta)^{-1} = e^{\lambda hY}z\circ(-\lambda h,0)\circ(D_\varepsilon\zeta)^{-1} = e^{\lambda hY}z\circ(e^{\lambda hY}D_\varepsilon\zeta)^{-1}.$$

Then, after the change of variables $\bar{z} = e^{\lambda hY}z$ and exchanging the order of integration, the term inside the square brackets in (5.7) is bounded by

$$\int\limits_0^1\|T_{n-3}Yu(\cdot\circ(e^{\lambda hY}D_\varepsilon\zeta)^{-1},\cdot)-Yu\|^p_p d\lambda \lesssim \int\limits_0^1\|e^{\lambda hY}D_\varepsilon\zeta\|^{(n-2)p}_p\|Yu\|^p_{W^{n-2,p}_B}d\lambda.$$

By (2.5) and (2.8), recalling that $\lambda h \leqslant \varepsilon^2$ and $\|\zeta\|_B \leqslant 1$ in the current integration set, we have

$$\|e^{\lambda hY}D_\varepsilon\zeta\|_B = \|D_\varepsilon(e^{\lambda\frac{h}{\varepsilon^2}Y}\zeta)\|_B = \varepsilon\|e^{\lambda\frac{h}{\varepsilon^2}Y}\zeta\|_B \lesssim \varepsilon(1+\|\zeta\|_B) \lesssim \varepsilon.$$

Therefore, substituting in (5.7) we find

$$S_{22} \leqslant \varepsilon^{-lp} \|Yu\|^p_{W_B^{n-2,p}} \int\limits_{|h|\leqslant \varepsilon^2} \left(\int\limits_{\|\zeta\|_B \leqslant 1} \varepsilon^{(n-2)} |\mathfrak{D}\varphi(\zeta)| d\zeta \right)^p \frac{dh}{|h|^{1-\frac{p}{2}}}$$

$$\lesssim \varepsilon^{(n-2-l)p} \|Yu\|^p_{W_B^{n-2,p}} \|\mathfrak{D}^l \varphi\|^p_p \int\limits_{|h|\leqslant \varepsilon^2} \frac{dh}{|h|^{1-\frac{p}{2}}} \lesssim \varepsilon^{(n-1-l)p} \|u\|^p_{W_B^{n,p}}.$$

The cases $n = 1$ or $n = 2$ are easier: it is easy to check that $T_0 u(\zeta, e^{hY} z) = T_0 u(\zeta, z) = u(\zeta)$ and $T_1 u(\zeta, e^{hY} z) = T_1 u(\zeta, z)$, therefore it suffices to use (4.3) and proceed as above. Collecting the estimates for S_1, S_{21}, S_{22} we get (5.6).

We are ready to prove (5.2) and (5.3) for $u \in W_B^{n,p} \cap C_0^\infty$, then the general statement follows by density. Clearly

$$\|u - u_{n,\varepsilon}\|_p = \|I_\varepsilon^{(n,0)} u\|_p \lesssim \varepsilon^n \|u\|_{W_B^{n,p}}$$

by (5.4). On the other hand, by (3.7), with some slight abuse of notation, we have

$$|u_{n,\varepsilon}|_{m,p,B} \lesssim \|\mathfrak{D}^m u_{n,\varepsilon}\|_p + [\mathfrak{D}^{m-1} u_{n,\varepsilon}]_{Y,\frac{1}{2},p}. \tag{5.8}$$

Since $\mathfrak{D}_z^i T_n u(\zeta, z) = 0$ for any $i > n$ we have

$$\mathfrak{D}^m u_{n,\varepsilon}(z) = \sum_{i=0}^n \int\limits_{\mathbb{R}^{N+1}} \mathfrak{D}_z^i T_{n-1} u(\zeta, z) \mathfrak{D}_z^{m-i} \varphi(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)) \frac{d\zeta}{\varepsilon^{\mathbf{d}}},$$

meaning that $\mathfrak{D}^m$, $\mathfrak{D}^{m-i}$, $\mathfrak{D}^i$ may stand for any intrinsic derivative of order $m, m-i, i$. Then, using Lemma 5.1 and that

$$\int\limits_{\mathbb{R}^{N+1}} \mathfrak{D}^i \varphi(z) dz = 0, \qquad i > 1,$$

we can write

$$\mathfrak{D}^m u_{n,\varepsilon}(z) = \sum_{i=0}^n \int\limits_{\mathbb{R}^{N+1}} \left(T_{n-1-i} \mathfrak{D}^i u(\zeta, z) - \mathfrak{D}^i u(z) \right) \mathfrak{D}_z^{m-i} \varphi(D_{\varepsilon^{-1}}(\zeta^{-1} \circ z)) \frac{d\zeta}{\varepsilon^{\mathbf{d}}},$$

and thus

$$\|\mathfrak{D}^m u_{n,\varepsilon}\|_p \lesssim \sum_{i=0}^n \|I_\varepsilon^{(n-i,m-i)} \mathfrak{D}^i u\|_p \lesssim \sum_{i=0}^n \varepsilon^{n-m} \|\mathfrak{D}^i u\|_{W_B^{n-i,p}} \lesssim \varepsilon^{n-m} \|u\|_{W_B^{n,p}}. \tag{5.9}$$

It only remains to estimate $[\mathfrak{D}^{m-1}u_{n,\varepsilon}]_{Y,\frac{1}{2},p}$: as before, if $m > n+1$, the test function is affected by at least one derivative for any non-null term of $\mathfrak{D}^{m-1}u_{n,\varepsilon}$. Therefore we have

$$[\mathfrak{D}^{m-1}u_{n,\varepsilon}]_{Y,\frac{1}{2},p} \lesssim \begin{cases} \sum_{i=0}^{n} [I_\varepsilon^{(n-1,m-1-i)}\mathfrak{D}^i u]_{Y,\frac{1}{2},p}, & m > n+1 \\ \sum_{i=0}^{n-1} [I_\varepsilon^{(n-1,m-1-i)}\mathfrak{D}^i u]_{Y,\frac{1}{2},p} + [(\mathfrak{D}^n u)_{0,\varepsilon}]_{Y,\frac{1}{2},p}, & m = n+1. \end{cases}$$

By (5.5) and (5.6), we directly derive

$$[\mathfrak{D}^{m-1}u_{n,\varepsilon}]_{Y,\frac{1}{2},p} \lesssim \varepsilon^{n-m}\|u\|_{W_B^{n,p}},$$

and recalling (5.8), (5.9) we finally get (5.3). $\square$

5.2. Interpolation on the degree of smoothness

In this section, we establish an interpolation result. The primary definitions and key results pertaining to interpolation theory are succinctly summarized in Appendix A.

Theorem 5.3 *(Interpolation). For $1 \leqslant n \leqslant m$ and $1 \leqslant p \leqslant \infty$ we have*

$$(L^p, W_B^{m,p})_{\frac{n}{m},1} \subseteq W_B^{n,p} \subseteq (L^p, W_B^{m,p})_{\frac{n}{m},\infty}. \tag{5.10}$$

Proof. The first embedding in (5.10) is a direct consequence of Proposition 3.5. Indeed, from (3.10) we deduce

$$\|u\|_{W_B^{n,p}} \lesssim \varepsilon\|u\|_{W_B^{m,p}} + \varepsilon^{-\frac{n}{m-n}}\|u\|_p, \qquad \varepsilon > 0.$$

In particular, taking the optimal $\varepsilon = \left(\|u\|_p/\|u\|_{W_B^{m,p}}\right)^{\frac{m-n}{m}}$ we get

$$\|u\|_{W_B^{n,p}} \lesssim \|u\|_{W_B^{m,p}}^{\frac{m-n}{m}}\|u\|_p^{\frac{n}{m}} \tag{5.11}$$

and, by Proposition A.1, estimate (5.11) is equivalent to the embedding $(L^p, W_B^{m,p})_{\frac{n}{m},1} \subseteq W_B^{n,p}$.

The second embedding in (5.10) is a direct consequence of Theorem 5.2. Indeed, by (5.2) and (5.3), for any $\varepsilon > 0$ we have

$$t^{-\frac{n}{m}}K(t,u;L^p,W_B^{m,p}) \lesssim t^{-\frac{n}{m}}(\|u-u_{n,\varepsilon}\|_p + t\|u_{n,\varepsilon}\|_{W_B^{m,p}}) \lesssim t^{-\frac{n}{m}}(\varepsilon^n + t\varepsilon^{n-m})\|u\|_{W_B^{n,p}}.$$

Therefore, taking $\varepsilon = t^{\frac{1}{m}}$ we get

$$t^{-\frac{n}{m}}K(t,u;L^p,W_B^{m,p}) \lesssim \|u\|_{W_B^{n,p}}. \quad \square$$

6. Proof of Theorem 1.1

The proof of Theorem 1.1 is based on some basic result from interpolation theory, which we briefly recall in Appendix A for reader's convenience. As a first step we derive a "crude" embedding which will serve as a starting point to derive the general result, through the characterization of Lorentz spaces of Lemma A.5.

Lemma 6.1. *For $p \in [1, \infty)$, let $u \in W_B^{1,p}$. There exists $r = r(p, B) > p$, such that*

$$\|u\|_q \lesssim \|u\|_p^{1-\theta} |u|_{1,p,B}^{\theta}, \qquad q \in [p, r),\ \theta = \mathbf{d}\left(\frac{1}{p} - \frac{1}{q}\right). \tag{6.1}$$

Proof. As in the proof of Theorem 5.3 we start from the decomposition $u = u_\varepsilon + (u - u_\varepsilon)$ with $u_\varepsilon = u_{\varepsilon,n}$ as in (5.1) with $n = 1$ and $\varepsilon > 0$. By (5.2)

$$\|u - u_\varepsilon\|_p \lesssim \varepsilon \|u\|_{W_B^{1,p}}.$$

On the other hand, by Young's inequality we have

$$\|u_\varepsilon\|_\infty \leqslant \varepsilon^{-\mathbf{d}} \|u\|_p \|\varphi(D_{\varepsilon^{-1}})\|_{\frac{p}{p-1}} \lesssim \varepsilon^{-\frac{\mathbf{d}}{p}} \|u\|_p, \qquad \varepsilon > 0.$$

Therefore, for K as in (A.1), we have

$$K(t, u, L^p, L^\infty) \lesssim (\varepsilon + t\varepsilon^{-\frac{\mathbf{d}}{p}}) \|u\|_{W_B^{1,p}}, \qquad \varepsilon, t > 0.$$

In particular, for $\varepsilon = t^{\frac{p}{\mathbf{d}+p}}$ we get

$$K(t, u, L^p, L^\infty) \lesssim t^{\frac{p}{\mathbf{d}+p}} \|u\|_{W_B^{1,p}}, \qquad t > 0,$$

that is

$$W_B^{1,p} \subseteq (L^p, L^\infty)_{\frac{p}{\mathbf{d}+p}, \infty} = L_w^r, \qquad r = \frac{p(\mathbf{d} + p)}{\mathbf{d}} > p,$$

by (A.6). By (A.5), this yields in particular that $W_B^{1,p} \subseteq L^q$, for any $q \in [p, r)$ and

$$\|u\|_q \lesssim \|u\|_p + |u|_{1,p,B}, \qquad u \in W_B^{1,p}. \tag{6.2}$$

Finally, by the usual scaling argument, applying (6.2) to $u(D_{\varepsilon^{-1}}(\cdot))$, and using (2.9) we get

$$\|u\|_q \lesssim \varepsilon^{\mathbf{d}(\frac{1}{p} - \frac{1}{q})} \|u\|_p + \varepsilon^{\mathbf{d}(\frac{1}{p} - \frac{1}{q}) - 1} |u|_{1,p,B}, \qquad u \in W_B^{1,p},\ \varepsilon > 0, \tag{6.3}$$

and this directly yields (6.1), optimizing on ε. $\square$

Remark 6.2. For $1 \leqslant p < \mathbf{d}$ we must have $q \leqslant p^* = \frac{\mathbf{d}p}{\mathbf{d}-p}$ in (6.3) or we would get a contradiction by letting ε tend to 0. This means that the critical exponent p^* is optimal for the embedding (1.6).

Next we use an ingenious approach, devised by Tartar [28], which consists of applying (6.1) to a suitable non-linear transformation of u. Precisely, we consider $\varphi_k(u)$ where $(\varphi_k)_{k\in\mathbb{Z}}$ is an appropriate sequence of functions involving the levels $a_k := u^*(e^k)$ of Lemma A.5: for $v \in \mathbb{R}$ and $k \in \mathbb{Z}$ we set

$$\varphi_k(v) = \begin{cases} 0 & \text{if } \ |v| \leqslant a_{k+1}, \\ |v| - a_{k+1} & \text{if } \ a_{k+1} \leqslant |v| \leqslant a_k, \\ a_k - a_{k+1} & \text{if } \ |v| \geqslant a_k. \end{cases}$$

We have the following crucial

Lemma 6.3. *For $p \in [1, \infty)$, let $u \in W_B^{1,p}$. There exists a positive constant $c = c(p, B)$ such that*

$$e^{\frac{k}{p^*}}(a_k - a_{k+1}) \leqslant c|\varphi_k(u)|_{1,p,B}, \qquad k \in \mathbb{Z}, \tag{6.4}$$

where p^* is the critical exponent in (1.6).

Proof. Notice that

$$(a_k - a_{k+1})\mathbb{1}_{(|u|\geqslant a_k)} \leqslant \varphi_k(u) \leqslant (a_k - a_{k+1})\mathbb{1}_{(|u|\geqslant a_{k+1})},$$

where $\mathbb{1}_A$ denotes the indicator function of the set A. Hence, for any $q \geqslant 1$ we have

$$\mathrm{Leb}(|u| \geqslant a_k)^{\frac{1}{q}}(a_k - a_{k+1}) \leqslant \|\varphi_k(u)\|_q \leqslant (a_k - a_{k+1})\mathrm{Leb}(|u| \geqslant a_{k+1})^{\frac{1}{q}}$$

where $\mathrm{Leb}(\cdot)$ represents the Lebesgue measure. By (A.8), which follows by construction of $(a_k)_{k\in\mathbb{Z}}$ (also recall definition (A.4) of distribution function), we get

$$e^{\frac{k-1}{q}}(a_k - a_{k+1}) \leqslant \|\varphi_k(u)\|_q \leqslant e^{\frac{k+1}{q}}(a_k - a_{k+1}) \tag{6.5}$$

From the first inequality in (6.5) and (6.1) applied to $\varphi_k(u) \in W_B^{1,p}$ with q, θ as in Lemma 6.1, we infer

$$e^{\frac{k-1}{q}}(a_k - a_{k+1}) \lesssim \|\varphi_k(u)\|_p^{1-\theta}|\varphi_k(u)|_{1,p,B}^{\theta} \lesssim$$

(by the second inequality in (6.5))

$$\lesssim e^{\frac{(k+1)(1-\theta)}{p}}(a_k - a_{k+1})^{1-\theta}|\varphi_k(u)|_{1,p,B}^{\theta}.$$

Equivalently, we have

$$e^{k\left(\frac{1}{q}-\frac{1-\theta}{p}\right)}(a_k-a_{k+1})^\theta \lesssim e^{\frac{k-1}{q}-\frac{(k+1)(1-\theta)}{p}}(a_k-a_{k+1})^\theta \lesssim |\varphi_k(u)|^\theta_{1,p,B}$$

and this concludes the proof since $\frac{1}{q}-\frac{1-\theta}{p}=\frac{\theta}{p^*}$. □

Proof of Theorem 1.1. Using that $|\varphi_k'(v)| = 1$ for $a_{k+1} < |v| < a_k$ and $\varphi_k'(v) = 0$ elsewhere, it is not difficult to prove that

$$|u|_{1,p,B} < \infty \quad \text{if and only if} \quad |\varphi_k(u)|_{1,p,B} \in \ell^p(\mathbb{Z}). \tag{6.6}$$

Thus, combining (6.4) with (6.6), we deduce that $e^{\frac{k}{p^*}}(a_k-a_{k+1}) \in \ell^p(\mathbb{Z})$ for any $u \in W^{1,p}_B$.

[*Case* $1 \leqslant p < \mathbf{d}$] A direct application of Lemma A.5 gives the improved Sobolev embedding

$$W^{1,p}_B \subseteq L^{p^*,p} \subseteq L^{p^*}.$$

In particular $W^{1,p}_B \subseteq L^q$ for any $q \in [p,p^*]$ by a standard application of the Young inequality.

[*Case* $p > \mathbf{d}$] We have $p^* < 0$ and therefore for any $\bar{k} \leqslant 0$ we have

$$\begin{aligned} a^p_{\bar{k}} &= \left(\sum_{k=-\infty}^{\bar{k}} (a_k-a_{k+1})\right)^p \\ &\leqslant \sum_{k=-\infty}^{\bar{k}} (a_k-a_{k+1})^p e^{\frac{pk}{p^*}} \left(\sum_{k=-\infty}^{\bar{k}} e^{-\frac{pk}{(p-1)p^*}}\right)^{p-1} \lesssim \sum_{k\in\mathbb{Z}} (a_k-a_{k+1})^p e^{\frac{pk}{p^*}} < \infty \end{aligned}$$

by (6.4). Being decreasing, $(a_k)_{k\in\mathbb{Z}}$ is then a bounded sequence and this yields $W^{1,p}_B \subseteq L^\infty$, that is $\|u\|_\infty \lesssim \|u\|_p + |u|_{1,p,B}$: by the usual scaling argument we have

$$\|u\|_\infty \lesssim \|u\|_p^{1-\frac{\mathbf{d}}{p}} |u|^{\frac{\mathbf{d}}{p}}_{1,p,B}. \tag{6.7}$$

Applying (6.7) to $(u(e^{h\partial_{x_i}}\cdot)-u)$, for $i=1,\dots,d$, using (4.8) and noticing that $|u(e^{h\partial_{x_i}}\cdot)-u|_{1,p,B} \leqslant 2|u|_{1,p,B}$, we get

$$\|u(e^{h\partial_{x_i}}\cdot)-u\|_\infty \lesssim |h|^{1-\frac{\mathbf{d}}{p}} \|\partial_{x_i} u\|_p^{1-\frac{\mathbf{d}}{p}} |u|^{\frac{\mathbf{d}}{p}}_{1,p,B} \lesssim |h|^{1-\frac{\mathbf{d}}{p}} |u|_{1,p,B}, \qquad i=1,\dots,d.$$

Analogously, applying (6.7) to $(u(e^{hY}\cdot)-u)$ and using (4.4) we get

$$\|u(e^{hY}\cdot)-u\|_\infty \lesssim |h|^{\frac{1}{2}(1-\frac{\mathbf{d}}{p})} [u]^{1-\frac{\mathbf{d}}{p}}_{Y,\frac{1}{2},p} |u|^{\frac{\mathbf{d}}{p}}_{1,p,B} \lesssim |h|^{\frac{1}{2}(1-\frac{\mathbf{d}}{p})} |u|_{1,p,B},$$

which proves the Morrey embedding (1.7).

[*Case* $p = \mathbf{d}$] As in the case $p < \mathbf{d}$, the embeddings (1.8) follow from Lemma A.5. To get estimate (1.9), it suffices to repeat the argument used by Tartar in [29], Chapter 30: more precisely, for $p = \mathbf{d}$ we have $1/p^* = 0$ so that $a_k - a_{k+1} \in \ell^{\mathbf{d}}(\mathbb{Z})$ by (6.4)-(6.6); applying Hölder's inequality we first prove that for every $\varepsilon > 0$ there exists a constant $c = c(\varepsilon, u) > 0$ such that

$$a_k^{\frac{\mathbf{d}}{\mathbf{d}-1}} \leqslant \varepsilon|k| + c, \qquad k \leqslant 0.$$

On the set where $a_{k+1} \leqslant |u| < a_k$, which has measure less than e^{k+1} by (A.8), we have

$$e^{\lambda|u|^{\frac{\mathbf{d}}{\mathbf{d}-1}}} \leqslant e^{\lambda a_k^{\frac{\mathbf{d}}{\mathbf{d}-1}}} \leqslant e^{\lambda(\varepsilon|k|+c)}, \qquad k \leqslant 0,$$

and by choosing $\varepsilon < \frac{1}{\lambda}$ we deduce that $e^{\lambda|u|^{\frac{\mathbf{d}}{\mathbf{d}-1}}}$ is integrable on any set where $|u| \geqslant \delta > 0$. $\square$

7. Higher orders embeddings

Embeddings for higher order Sobolev spaces are classically derived by iteration from the $n = 1$ case. In our setting, because of the qualitative difference between even and odd orders of intrinsic spaces (cf. (3.8)) resulting from the two-steps iterative definition, we need some additional work at least when $n = 2$, in order to control the Holder and Sobolev-Slobodeckij quasi-norms involving the vector field Y, in the high and low summability cases respectively. Our method here is based on the representation of a $W_B^{2,p}$ function by means of the fundamental solution of a linear Kolmogorov operator with drift matrix B (see (7.8) below), which only allows to derive the embeddings in the case $p > 1$. The main result of this section is the following

Theorem 7.1 *($W_B^{n,p}$ embeddings). Let $k \in \mathbb{N}$, $n \in \mathbb{N}_0$ and $1 < p < \infty$.*

i) If $kp < \mathbf{d}$, then

$$W_B^{n+k,p} \subseteq W_B^{n,q}, \qquad p \leqslant q \leqslant p_k^*, \qquad \frac{1}{p_k^*} = \frac{1}{p} - \frac{k}{\mathbf{d}}.$$

In particular, $W_B^{k,p} \subseteq L^q$ for $p \leqslant q \leqslant p_k^*$;

ii) if $kp > \mathbf{d} > (k-1)p$, then

$$W_B^{n+k,p} \subseteq C_B^{n,k-\frac{\mathbf{d}}{p}}.$$

7.1. Fundamental solution of Fokker-Planck equations

We recall some preliminary results about the fundamental solution of the Fokker-Planck operator $\mathscr{K}$ in (2.2).

Proposition 7.2. *Hörmander's condition (1.2) is equivalent to the fact that the matrix*

$$\mathscr{C}_t = \int_0^t e^{sB} \begin{pmatrix} I_d & 0_{d\times(N-d)} \\ 0_{(N-d)\times d} & 0_{(N-d)\times(N-d)} \end{pmatrix} e^{sB^*} ds$$

is positive definite every $t > 0$. In this case, the fundamental solution of $\mathscr{K}$ with pole at 0 is given by

$$\Gamma(t,x) = \begin{cases} \frac{1}{\sqrt{(2\pi)^N \det \mathscr{C}_t}} e^{-\frac{1}{2}\langle \mathscr{C}_t^{-1}x,x\rangle}, & t > 0 \\ 0 & t \leqslant 0. \end{cases} \tag{7.1}$$

The fundamental solution with pole at ζ is the left translation of Γ with respect to the group law, that is $\Gamma(\zeta^{-1} \circ \cdot)$.

We recall that $\mathbf{d}$ denotes the homogeneous dimension of $\mathbb{R}^{N+1}$ defined in (2.6). Since

$$\mathscr{C}_{\lambda^2 t} = \hat{D}_\lambda \mathscr{C}_t \hat{D}_\lambda, \tag{7.2}$$

the fundamental solution Γ is homogeneous of degree $-\mathbf{d}+2$ with respect to $(D_\lambda)_{\lambda>0}$. Similarly $\partial_{x_i}\Gamma$, $\partial_{x_i x_j}\Gamma$ for $i,j = 1,\dots,d$ and $Y\Gamma$ are homogeneous of degrees $-\mathbf{d}+1$, $-\mathbf{d}$ and $-\mathbf{d}$ respectively.

Later we will exploit the global estimates of the following

Lemma 7.3. *For every $z \in \mathbb{R}^{N+1}$ we have*

$$\Gamma(z) \lesssim \|z\|_B^{-\mathbf{d}+2}, \qquad |Y\Gamma(z)| \lesssim \|z\|_B^{-\mathbf{d}}. \tag{7.3}$$

Proof. A local version of (7.3) has been proven in [8] in the general framework of non-homogeneous Fokker-Planck operators. In our setting we provide a more direct proof. Let us only consider the second estimate in (7.3) for $t > 0$: since $\mathscr{K}\Gamma = 0$ we have

$$|Y\Gamma(z)| \leqslant \frac{1}{2}\sum_{i=1}^d |\partial_{x_i x_i}\Gamma(z)| \leqslant \frac{1}{2}\sum_{i=1}^d \left(\frac{1}{2}|(\mathscr{C}_t^{-1}x)_i|^2 + |(\mathscr{C}_t^{-1})_{ii}|\right)\Gamma(z). \tag{7.4}$$

By (7.1) an (7.2) we have

$$\Gamma(z) \lesssim t^{1-\frac{\mathbf{d}}{2}} \exp\left(-\frac{1}{2}\langle \hat{D}_{\frac{1}{\sqrt{t}}} \mathscr{C}_1^{-1} \hat{D}_{\frac{1}{\sqrt{t}}} x, x\rangle\right) \lesssim t^{1-\frac{\mathbf{d}}{2}} \exp\left(-\frac{1}{2\|\mathscr{C}_1^{-1}\|}|\hat{D}_{\frac{1}{\sqrt{t}}} x|^2\right); \tag{7.5}$$

On the other hand, for $1 \leqslant i \leqslant d$, again by (7.2) we have

$$|(\mathscr{C}_t^{-1}x)_i| = |(\hat{D}_{\frac{1}{\sqrt{t}}}\mathscr{C}_1^{-1}\hat{D}_{\frac{1}{\sqrt{t}}}x)_i| \leqslant \frac{\|\mathscr{C}_1^{-1}\|}{\sqrt{t}}|\hat{D}_{\frac{1}{\sqrt{t}}}x|, \tag{7.6}$$

$$|(\mathscr{C}_t^{-1})_{ii}| = |(\mathscr{C}_t^{-1}\mathbf{e}_i)_i| \leqslant \frac{\|\mathscr{C}_1^{-1}\|}{t}. \tag{7.7}$$

Therefore, using that $\|z\|_B = t^{\frac{1}{2}}\|(1, D_{\frac{1}{\sqrt{t}}}x)\|_B = t^{\frac{1}{2}}(1 + |\hat{D}_{\frac{1}{\sqrt{t}}}x|_B)$, by (7.4), (7.5) and (7.6)-(7.7) we finally get

$$\|z\|_B^{\mathbf{d}}|Y\Gamma(z)| \lesssim t^{\frac{\mathbf{d}}{2}}(1 + |\hat{D}_{\frac{1}{\sqrt{t}}}x|_B)^{\mathbf{d}}\frac{1}{t^{\frac{\mathbf{d}}{2}}}(1 + |\hat{D}_{\frac{1}{\sqrt{t}}}x|^2)\exp\left(-\frac{1}{2\|\mathscr{C}_1^{-1}\|}|\hat{D}_{\frac{1}{\sqrt{t}}}x|^2\right) \lesssim 1.$$

The proof is completed. □

7.2. Proof of Theorem 7.1

We first examine the regularity along the vector field Y in both the high and low summability case for $W_B^{2,p}$. We recall a theorem for convolution with a homogeneous kernel proved in [27], Theorem 1, p.119.

Theorem 7.4. *For every $\alpha \in]0, \mathbf{d}[$ and $g \in L^p(\mathbb{R}^{N+1})$ with $p > 1$, the function*

$$I_\alpha(g)(z) = \int\limits_{\mathbb{R}^{N+1}} \frac{g(\zeta)}{\|\zeta^{-1} \circ z\|_B^{\mathbf{d}-\alpha}}d\zeta,$$

is a.e. defined and there exists $c = c(p, \alpha) > 0$ such that

$$\|I_\alpha(g)\|_q \leqslant c\|g\|_p, \qquad \frac{1}{p} = \frac{1}{q} + \frac{\alpha}{\mathbf{d}}.$$

Proposition 7.5. *If $p > \mathbf{d}$, then there exists $c = c(p, B)$ such that*

$$\operatorname*{ess\,sup}_{z\in\mathbb{R}^{N+1}} |u(e^{\delta Y}z) - u(z)| \leqslant c|\delta|^{1-\frac{\mathbf{d}}{2p}}\|u\|_{W_B^{2,p}}, \qquad u \in W_B^{2,p}, \quad \delta \in \mathbb{R}.$$

Proof. By the definition of fundamental solution, for $u \in C_0^\infty$ we have the representation

$$u(z) = -\int\limits_{\mathbb{R}^{N+1}} \Gamma(\zeta^{-1} \circ z)\mathscr{K}u(\zeta)d\zeta. \tag{7.8}$$

Since Γ is homogeneous of degree $-\mathbf{d} + 2$, by Theorem 7.4 and a density argument we deduce that (7.8) holds a.e. for any $u \in W_B^{2,p}$ as well. Then we have

$$|u(e^{\delta Y}z)-u(z)| \leqslant \left(\int\limits_{\|\zeta^{-1}\circ z\|_B \geqslant c\sqrt{|\delta|}} + \int\limits_{\|\zeta^{-1}\circ z\|_B < c\sqrt{|\delta|}}\right) \times |\Gamma(\zeta^{-1}\circ e^{\delta Y}z)-\Gamma(\zeta^{-1}\circ z)||\mathscr{K}u(\zeta)|d\zeta \leqslant$$

(for some $\bar\delta$, $|\bar\delta| \leqslant |\delta|$, dependent on z,ζ)

$$\leqslant \int\limits_{\|\zeta^{-1}\circ z\|_B \geqslant c\sqrt{|\delta|}} |\delta||Y\Gamma(e^{\bar\delta Y}(\zeta^{-1}\circ z))||\mathscr{K}u(\zeta)|d\zeta + \int\limits_{\|\zeta^{-1}\circ z\|_B < c\sqrt{|\delta|}} \left(\Gamma(\zeta^{-1}\circ e^{\delta Y}z)+\Gamma(\zeta^{-1}\circ z)\right)|\mathscr{K}u(\zeta)|d\zeta \leqslant S_1+S_{21}+S_{22}$$

by the uniform estimates (7.3), where

$$S_1 = \int\limits_{\|\zeta^{-1}\circ z\|_B \geqslant c\sqrt{|\delta|}} |\delta|\|e^{\bar\delta Y}(\zeta^{-1}\circ z)\|_B^{-\mathbf{d}}|\mathscr{K}u(\zeta)|d\zeta,$$

$$S_{21} = \int\limits_{\|\zeta^{-1}\circ z\|_B < c\sqrt{|\delta|}} \|e^{\delta Y}(\zeta^{-1}\circ z)\|_B^{-\mathbf{d}+2}|\mathscr{K}u(\zeta)|d\zeta,$$

$$S_{22} = \int\limits_{\|\zeta^{-1}\circ z\|_B < c\sqrt{|\delta|}} \|\zeta^{-1}\circ z\|_B^{-\mathbf{d}+2}|\mathscr{K}u(\zeta)|d\zeta.$$

Choosing c as in (2.8), we have $\|\zeta^{-1}\circ z\|_B \lesssim \|\zeta^{-1}\circ e^{\bar\delta Y}z\|_B = \|e^{\bar\delta Y}(\zeta^{-1}\circ z)\|_B$ on $\sqrt{|\delta|} \leqslant c\|\zeta^{-1}\circ z\|_B$. Then, as in [24], Lemma 2.9, we have

$$S_1 = |\delta|\sum_{k\geqslant 1} \int\limits_{c^k\sqrt{|\delta|}\leqslant\|\zeta^{-1}\circ z\|_B\leqslant c^{k+1}\sqrt{|\delta|}} \frac{|\mathscr{K}u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^{\mathbf{d}}}d\zeta$$

$$\lesssim |\delta|\sum_{k\geqslant 1}(c^k\sqrt{|\delta|})^{-\mathbf{d}} \int\limits_{\|\zeta^{-1}\circ z\|_B\leqslant c^{k+1}\sqrt{|\delta|}} |\mathscr{K}u(\zeta)|d\zeta \leqslant$$

(by Hölder's inequality)

$$\leqslant |\delta|^{1-\frac{\mathbf{d}}{2}}\sum_{k\geqslant 1} c^{-k\mathbf{d}}\|\mathscr{K}u\|_p \mathrm{Leb}\left(\|\zeta^{-1}\circ z\|_B \leqslant c^{k+1}\sqrt{|\delta|}\right)^{1-\frac{1}{p}}$$

$$\lesssim |\delta|^{1-\frac{\mathbf{d}}{2p}}\|u\|_{W_B^{2,p}} c^{\mathbf{d}(1-\frac{1}{p})}\sum_{k\geqslant 1} c^{-k\frac{\mathbf{d}}{p}},$$

where we used that $\mathrm{Leb}\left(\|\zeta^{-1}\circ z\|_B \leqslant r\right) \leqslant r^{\mathbf{d}}\mathrm{Leb}\left(\|z\|_B \leqslant 1\right)$. Similarly, we have

$$S_{22} = \sum_{k\geqslant 1} \int\limits_{c^{-k}\sqrt{|\delta|}\leqslant\|\zeta^{-1}\circ z\|_B\leqslant c^{1-k}\sqrt{|\delta|}} \frac{|\mathscr{K}u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^{\mathbf{d}-2}} d\zeta$$

$$\leqslant \sum_{k\geqslant 1}\left(\frac{c^k}{\sqrt{|\delta|}}\right)^{\mathbf{d}-2} \|\mathscr{K}u\|_p (c^{1-k}\sqrt{|\delta|})^{\mathbf{d}(1-\frac{1}{p})} \lesssim |\delta|^{1-\frac{\mathbf{d}}{2p}} \|u\|_{W_B^{2,p}} \sum_{k\geqslant 1} c^{-k(2-\frac{\mathbf{d}}{p})}.$$

For S_{21} observe now that by (2.8) on $\{\|\zeta^{-1}\circ z\|_B \leqslant c\sqrt{|\delta|}\}$ we have $\|\zeta^{-1}\circ e^{\delta Y}z\|_B \leqslant m(1+c)\sqrt{|\delta|} \equiv \bar{c}\sqrt{|\delta|}$, so that

$$S_{21} \leqslant \int\limits_{\|\zeta^{-1}\circ e^{\delta Y}z\|_B\leqslant\bar{c}\sqrt{|\delta|}} \|\zeta^{-1}\circ e^{\delta Y}z\|_B^{-\mathbf{d}+2} |\mathscr{K}u(\zeta)| d\zeta,$$

and therefore it is analogous to S_{22}. The proof is complete. □

Proposition 7.6. *If $p<\mathbf{d}$, then there exists $c=c(p,B)$ such that*

$$[u]_{Y,\frac{1}{2},p^*} \leqslant c\|u\|_{W_B^{2,p}}, \qquad u\in W_B^{2,p}.$$

Proof. Recall that

$$[u]_{Y,\frac{1}{2},p^*}^{p^*} = \int\limits_{\mathbb{R}^{N+1}} \int\limits_{|h|\leqslant 1} \frac{|u(e^{hY}z)-u(z)|^{p^*}}{|h|^{1+\frac{p^*}{2}}} dh dz.$$

Observe that, since $|h| = \|z^{-1}\circ e^{hY}z\|_B = \|(e^{hY}z)^{-1}\circ z\|_B$, possibly exchanging variables by $z' = e^{hY}z$ (whose Jacobian has determinant equal to 1), we may assume that we are integrating on a subset of

$$\{(z,h)\in\mathbb{R}^{N+1}\times[-1,1] \mid |u(z)|\geqslant|u(e^{hY}z)|\}. \tag{7.9}$$

Now, by representation (7.8) and Minkowski inequality

$$[u]_{Y,\frac{1}{2},p^*}^{p^*} = \int\limits_{\mathbb{R}^{N+1}} \int\limits_{|h|\leqslant 1} \left| \int\limits_{\mathbb{R}^{N+1}} \left(\Gamma(\zeta^{-1}\circ e^{hY}z) - \Gamma(\zeta^{-1}\circ z)\right) \mathscr{K}u(\zeta) d\zeta \right|^{p^*} \frac{dh}{|h|^{1+\frac{p^*}{2}}} dz$$

$$\leqslant \int\limits_{\mathbb{R}^{N+1}} \left(\int\limits_{\mathbb{R}^{N+1}} \left(\int\limits_{|h|\leqslant 1} |\Gamma(\zeta^{-1}\circ e^{hY}z) - \Gamma(\zeta^{-1}\circ z)|^{p^*} \frac{dh}{|h|^{1+\frac{p^*}{2}}} \right)^{\frac{1}{p^*}} |\mathscr{K}u(\zeta)| d\zeta \right)^{p^*} dz.$$

For c as in (2.8), we rewrite the inner integral of the last expression as

$$\left(\int\limits_{|h|\leqslant c\|\zeta^{-1}\circ z\|_B^2} + \int\limits_{c\|\zeta^{-1}\circ z\|_B^2\leqslant|h|\leqslant 1} \right) |\Gamma(\zeta^{-1}\circ e^{hY}z) - \Gamma(\zeta^{-1}\circ z)|^{p^*} \frac{dh}{|h|^{1+\frac{p^*}{2}}} =: S_1 + S_2.$$

Now, for some h' dependent on h, z, ζ, with $|h'| \leqslant |h|$, we have

$$S_1 = \int\limits_{|h| \leqslant c\|\zeta^{-1} \circ z\|_B^2} |Y\Gamma(e^{h'Y}(\zeta^{-1} \circ z))|^{p^*} |h|^{p^*} \frac{dh}{|h|^{1+\frac{p^*}{2}}} \lesssim$$

(by (7.3))

$$\lesssim \int\limits_{|h| \leqslant c\|\zeta^{-1} \circ z\|_B^2} \|e^{h'Y}(\zeta^{-1} \circ z)\|_B^{-p^*\mathbf{d}} \frac{dh}{|h|^{1-\frac{p^*}{2}}}$$
$$\lesssim \|\zeta^{-1} \circ z\|_B^{-p^*(\mathbf{d}-1)}$$

using that $\|\zeta^{-1} \circ z\|_B \lesssim \|e^{h'Y}(\zeta^{-1} \circ z)\|_B$ in the domain of the last integral.

On the other hand, by (7.9) and (7.3) we have

$$S_2 \leqslant 2 \int\limits_{c\|\zeta^{-1} \circ z\|_B^2 \leqslant |h|} \Gamma(\zeta^{-1} \circ z)^{p^*} \frac{dh}{|h|^{1+\frac{p^*}{2}}}$$
$$\lesssim \|\zeta^{-1} \circ z\|^{-p^*(\mathbf{d}-2)} \int\limits_{c\|\zeta^{-1} \circ z\|_B^2 \leqslant |h|} \frac{dh}{|h|^{1+\frac{p^*}{2}}} \lesssim \|\zeta^{-1} \circ z\|_B^{-p^*(\mathbf{d}-1)}.$$

Therefore we have

$$[u]_{Y,\frac{1}{2},p^*} \lesssim \|I_1(\mathscr{K}u)\|\|_{p^*},$$

with I_1 as in Theorem 7.4 and the thesis follows since $\mathscr{K}u \in L^p$ by assumption. □

We are now in position to prove Theorem 7.1.

Proof of Theorem 7.1. The embeddings of $W_B^{1,p}$ follow from Theorem 1.1. Regarding $W_B^{2,p}$, the statement of the theorem can be rewritten more explicitly as follows:

1) if $p > \mathbf{d}$ then

$$W_B^{2,p} \subseteq C_B^{1,1-\frac{\mathbf{d}}{p}}; \tag{7.10}$$

2) if $\frac{\mathbf{d}}{2} < p < \mathbf{d}$ then:
   – considering $n = k = 1$, we have

$$W_B^{2,p} \subseteq W_B^{1,q}, \qquad p \leqslant q \leqslant p^*, \qquad \frac{1}{p^*} = \frac{1}{p} - \frac{1}{\mathbf{d}}; \tag{7.11}$$

– considering $n = 0$ and $k = 2$, we have

$$W_B^{2,p} \subseteq C_B^{0,2-\frac{\mathbf{d}}{p}}. \tag{7.12}$$

However, by (1.7) we have that (7.11) implies (7.12) so that it suffices to prove (7.11);

3) if $p \leqslant \frac{\mathbf{d}}{2}$, which implies $n = 0$ and $k = 2$, we have

$$W_B^{2,p} \subseteq L^q, \qquad p \leqslant q \leqslant p_2^*, \qquad \frac{1}{p_2^*} = \frac{1}{p} - \frac{2}{\mathbf{d}};$$

To prove (7.10), we notice that for $p > \mathbf{d}$ we have $u, \partial_{x_i} u \in W_B^{1,p} \subseteq C_B^{0,1-\frac{\mathbf{d}}{p}}$ for any $i = 1, \dots, d$ by (1.7) and $[u]_{C_Y^{1-\frac{\mathbf{d}}{2p}}} \lesssim \|u\|_{W_B^{2,p}}$ by Proposition 7.5.

To prove (7.11), it suffices to observe that if $\frac{\mathbf{d}}{2} < p < \mathbf{d}$ then $u, \partial_{x_i} u \in W_B^{1,p} \subseteq L^{p^*}$ by (1.6) and $[u]_{Y,\frac{1}{2},p^*} \lesssim \|u\|_{W_B^{2,p}}$ by Proposition 7.6.

Finally, if $p < \frac{\mathbf{d}}{2}$ then again $W_B^{2,p} \subseteq W_B^{1,p^*}$ with $p^* < \mathbf{d}$ and by (1.6) we get

$$W_B^{2,p} \subseteq W_B^{1,p^*} \subseteq L^{\frac{p^*\mathbf{d}}{\mathbf{d}-p^*}} = L^{p_2^*}.$$

The proof of higher order embeddings is analogous: by induction, it suffices to use iteratively the previous arguments. □

Declaration of competing interest

The authors have no relevant financial or non-financial interests to disclose.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors are members of INDAM-GNAMPA.

Appendix A. Interpolation

We briefly recall some basic tool and notion from interpolation theory: for a comprehensive presentation of the subject we refer, for instance, to [14], [29] and [2].

Given two real Banach spaces Z_1, Z_2, we write $Z_1 = Z_2$ if Z_1 and Z_2 have the same elements with equivalent norms; we write $Z_1 \subseteq Z_2$ if Z_1 is continuously embedded in Z_2. The pair (Z_1, Z_2) is called an *interpolation pair* if both Z_1 and Z_2 are continuously embedded in some Hausdorff topological vector space: in this case, the intersection $Z_1 \cap Z_2$ and the sum $Z_1 + Z_2$ endowed with the norms

$$\|u\|_{Z_1\cap Z_2} := \max\{\|u\|_{Z_1}, \|u\|_{Z_2}\}, \qquad \|u\|_{Z_1+Z_2} := \inf_{\substack{u_1\in Z_1,\, u_2\in Z_2\\ u=u_1+u_2}} (\|u_1\|_{Z_1} + \|u_2\|_{Z_2}),$$

are Banach spaces. For any $t > 0$ and $u \in Z_1 + Z_2$, we set

$$K(t,u) \equiv K(t,u;Z_1,Z_2) := \inf_{\substack{u_1\in Z_1,\, u_2\in Z_2\\ u=u_1+u_2}} (\|u_1\|_{Z_1} + t\|u_2\|_{Z_2}). \tag{A.1}$$

Any Banach space E such that

$$Z_1 \cap Z_2 \subseteq E \subseteq Z_1 + Z_2,$$

is called an *intermediate space*. Among these, for $0 < \theta < 1$ and $1 \leqslant p \leqslant \infty$, we have the *real interpolation space* $(Z_1, Z_2)_{\theta,p}$ consisting of $u \in Z_1 + Z_2$ such that

$$\|u\|_{\theta,p} := \|t^{-\theta}K(t,u)\|_{L^p_*} < \infty \tag{A.2}$$

where $L^p_* = L^p_*(\mathbb{R}_{>0})$ denotes the L^p space with respect to the measure $\frac{dt}{t}$ on $\mathbb{R}_{>0}$ and $L^\infty_* := L^\infty$.

Proposition A.1 *([15], Prop. 1.20). For an intermediate space E the following conditions are equivalent:*

i) $(Z_1, Z_2)_{\theta,1} \subseteq E$;
ii) there exists a constant c such that

$$\|u\|_E \leqslant c\|u\|_{Z_1}^{1-\theta}\|u\|_{Z_2}^{\theta}, \qquad u \in Z_1 \cap Z_2.$$

In the very particular case $Z_2 \subseteq Z_1$ (for instance, if Z_1 is an L^p space and Z_2 is some intrinsic Sobolev space $W^{m,p}_B$), we have

$$Z_1 \cap Z_2 = Z_2, \qquad Z_1 + Z_2 = Z_1, \qquad K(t,u) \leqslant \min\{\|u\|_{Z_1}, t\|u\|_{Z_2}\}. \tag{A.3}$$

Then, since $t \mapsto K(t,u)$ is bounded by (A.3), for $\|u\|_{\theta,p}$ in (A.2) to be finite, what really matters is only the behavior of $K(t,u)$ near $t = 0$.

A.1. Interpolation between L^p spaces

The *distribution* of a measurable function u on $\mathbb{R}^N$ is defined as

$$\mu_u(\lambda) := \mathrm{Leb}(|u| > \lambda), \qquad \lambda \geqslant 0, \tag{A.4}$$

while

$$u^*(t) := \inf\{\lambda \geqslant 0 \mid \mu_u(\lambda) \leqslant t\}, \qquad t \geqslant 0,$$

is called the *rearranging* of u. Distribution and rearranging are decreasing and right continuous functions. Since

$$u^*(t) > \lambda \qquad \text{if and only if} \qquad \mu_u(\lambda) > t$$

we have

$$\mathrm{Leb}(u^* > \lambda) = \mathrm{Leb}\left(0 \leqslant t < \mu_u(\lambda)\right) = \mathrm{Leb}(|u| > \lambda)$$

i.e. u and u^* are equimeasurable and consequently $\|u\|_{L^p} = \|u^*\|_{L^p}$.

Definition A.2 *(Weak L^p spaces).* For $1 \leqslant p < \infty$, the weak L^p (or Marcinkiewicz) space is defined as the space of all measurable functions u such that

$$\|u\|_{L^p_w} := \sup_{\lambda > 0} \lambda \mu_u(\lambda)^{\frac{1}{p}} < \infty$$

and $L^\infty_w := L^\infty$.

Clearly $L^p \subseteq L^p_w$ and in general the inclusion is strict: for instance, $u(x) = |x|^{-N/p} \in L^p_w(\mathbb{R}^N)$ but does not belong to any L^q. On the other hand

$$L^p_w \cap L^q_w \subseteq L^r, \qquad 1 \leqslant p < r < q \leqslant \infty. \tag{A.5}$$

Definition A.3 *(Lorentz spaces).* For $1 \leqslant p < \infty$ and $1 \leqslant q \leqslant \infty$, the Lorentz space $L^{p,q}$ is defined as the set of all measurable functions u such that the following quasi-norm

$$\|u\|_{L^{p,q}} := \|t^{1/p} u^*(t)\|_{L^q_*} = \begin{cases} \left(\int\limits_0^\infty (t^{1/p} u^*(t))^q \frac{dt}{t} \right)^{\frac{1}{q}} & \text{if } 1 \leqslant q < \infty, \\ \sup\limits_{t \geqslant 0} t^{1/p} u^*(t) & \text{if } q = \infty, \end{cases}$$

is finite. We also set $L^{\infty,\infty} = L^\infty$.

By Hölder's inequality $L^{p,q_1} \subseteq L^{p,q_2}$ if $q_1 \leqslant q_2$ and more generally we have

$$L^p \equiv L^{p,p} \subseteq L^{p,q} \subseteq L^{p,\infty} \equiv L^p_w, \qquad 1 \leqslant p \leqslant q \leqslant \infty.$$

Lorentz spaces have a classical characterization as interpolation of L^p spaces (cf. [2], Corollary 7.27).

Proposition A.4. *For any $1 \leqslant p_1 < p < p_2 \leqslant \infty$ and $1 \leqslant q \leqslant \infty$, we have*

$$L^{p,q} = (L^{p_1}, L^{p_2})_{\theta,q}\,, \qquad \frac{1}{p} = (1-\theta)\frac{1}{p_1} + \theta\frac{1}{p_2},$$

and in particular

$$L^{p,q} = \left(L^1, L^\infty\right)_{1-\frac{1}{p},q}. \tag{A.6}$$

Another characterization of Lorentz spaces has been provided by L. Tartar with the aim of studying improved Sobolev's embedding theorems.

Lemma A.5 *([29], Lemma 29.4). Given a measurable function u on $\mathbb{R}^N$ such that $\mu_u(\lambda) < \infty$ for any $\lambda > 0$, we consider a decreasing sequence $(a_k)_{k\in\mathbb{Z}}$ such that*

$$u^*(e^k) \leqslant a_k \leqslant u^*(e^k-), \qquad k \in \mathbb{Z}. \tag{A.7}$$

Then, for $1 \leqslant p < \infty$ and $1 \leqslant q \leqslant \infty$, we have

$$u \in L^{p,q} \quad \text{if and only if} \quad e^{k/p}a_k \in \ell^q(\mathbb{Z}).$$

Moreover, if $a_k \to 0$ as $k \to \infty$, then

$$u \in L^{p,q} \quad \text{if and only if} \quad e^{k/p}(a_k - a_{k+1}) \in \ell^q(\mathbb{Z}).$$

Notice that from (A.7) it follows that

$$\mu_u(a_k) \leqslant e^k \leqslant \mu_u(a_k-) \leqslant \mu_u(a_{k+1}), \qquad k \in \mathbb{Z}. \tag{A.8}$$

References

[1] F. Abedin, G. Tralli, Harnack inequality for a class of Kolmogorov-Fokker-Planck equations in non-divergence form, Arch. Ration. Mech. Anal. 233 (2) (2019) 867–900.

[2] R.A. Adams, J.J.F. Fournier, Sobolev Spaces, second ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003.

[3] F. Bouchut, Hypoelliptic regularity in kinetic equations, J. Math. Pures Appl. (9) 81 (11) (2002) 1135–1159.

[4] M. Bramanti, G. Cupini, E. Lanconelli, E. Priola, Global L^p estimates for degenerate Ornstein-Uhlenbeck operators, Math. Z. 266 (4) (2010) 789–816.

[5] F. Camellini, M. Eleuteri, S. Polidoro, A compactness result for the Sobolev embedding via potential theory, in: Harnack Inequalities and Nonlinear Operators, in: Springer INdAM Ser., vol. 46, Springer, Cham, 2021, pp. 61–91.

[6] S. Cameron, L. Silvestre, S. Snelson, Global a priori estimates for the inhomogeneous Landau equation with moderately soft potentials, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 35 (3) (2018) 625–642.

[7] M. Di Francesco, A. Pascucci, S. Polidoro, The obstacle problem for a class of hypoelliptic ultraparabolic equations, Proc. R. Soc., Lond. Ser. A Math. Phys. Eng. Sci. 464 (2089) (2008) 155–176.

[8] M. Di Francesco, S. Polidoro, Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form, Adv. Differ. Equ. 11 (11) (2006) 1261–1320.

[9] H. Dong, T. Yastrzhembskiy, Global L_p estimates for kinetic Kolmogorov-Fokker-Planck equations in nondivergence form, Arch. Ration. Mech. Anal. 245 (1) (2022) 501–564.
[10] G.B. Folland, E.M. Stein, Hardy Spaces on Homogeneous Groups, Mathematical Notes, vol. 28, Princeton University Press/University of Tokyo Press, Princeton, N.J./Tokyo, 1982.
[11] N. Garofalo, G. Tralli, Hardy-Littlewood-Sobolev inequalities for a class of non-symmetric and non-doubling hypoelliptic semigroups, Math. Ann. 383 (1–2) (2022) 1–38.
[12] F. Golse, C. Imbert, C. Mouhot, A.F. Vasseur, Harnack inequality for kinetic Fokker-Planck equations with rough coefficients and application to the Landau equation, Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5) 19 (1) (2019) 253–295.
[13] E. Lanconelli, S. Polidoro, On a class of hypoelliptic evolution operators, Rend. Sem. Mat. Univ. Politec. Torino 52 (1) (1994) 29–63.
[14] G. Leoni, A First Course in Sobolev Spaces, second ed., Graduate Studies in Mathematics, vol. 181, American Mathematical Society, Providence, RI, 2017.
[15] A. Lunardi, Interpolation Theory, Lecture Notes, Scuola Normale Superiore di Pisa, vol. 16, Edizioni della Normale, Pisa, 2018.
[16] M. Manfredini, The Dirichlet problem for a class of ultraparabolic equations, Adv. Differ. Equ. 2 (5) (1997) 831–866.
[17] D. Morbidelli, Fractional Sobolev norms and structure of Carnot-Carathéodory balls for Hörmander vector fields, Stud. Math. 139 (3) (2000) 213–244.
[18] S. Pagliarani, A. Pascucci, M. Pignotti, Intrinsic Taylor formula for Kolmogorov-type homogeneous groups, J. Math. Anal. Appl. 435 (2) (2016) 1054–1087.
[19] A. Pascucci, Hölder regularity for a Kolmogorov equation, Trans. Am. Math. Soc. 355 (3) (2003) 901–924.
[20] A. Pascucci, PDE and Martingale Methods in Option Pricing, Bocconi & Springer Series, vol. 2, Springer/Bocconi University Press, Milan/Milan, 2011.
[21] A. Pascucci, A. Pesce, On stochastic Langevin and Fokker-Planck equations: the two-dimensional case, J. Differ. Equ. 310 (2022) 443–483.
[22] A. Pascucci, S. Polidoro, The Moser's iterative method for a class of ultraparabolic equations, Commun. Contemp. Math. 6 (3) (2004) 395–417.
[23] A. Pesce, Approximation and interpolation in Kolmogorov-type groups, arXiv:2205.05340, J. Math. Anal. Appl. 535 (2) (2024) 128139.
[24] S. Polidoro, M.A. Ragusa, Sobolev-Morrey spaces related to an ultraparabolic equation, Manuscr. Math. 96 (3) (1998) 371–392.
[25] L.P. Rothschild, E.M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (3–4) (1976) 247–320.
[26] K. Saka, Besov spaces and Sobolev spaces on a nilpotent Lie group, Tohoku Math. J. (2) 31 (4) (1979) 383–437.
[27] E.M. Stein, Singular Integrals and Differentiability Properties of Functions, Princeton Mathematical Series, vol. 30, Princeton University Press, Princeton, N.J., 1970.
[28] L. Tartar, Imbedding theorems of Sobolev spaces into Lorentz spaces, Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8) 1 (3) (1998) 479–500.
[29] L. Tartar, An Introduction to Sobolev Spaces and Interpolation Spaces, Lecture Notes of the Unione Matematica Italiana, vol. 3, Springer/UMI, Berlin/Bologna, 2007.
[30] N.S. Trudinger, On imbeddings into Orlicz spaces and some applications, J. Math. Mech. 17 (1967) 473–483.
[31] N.T. Varopoulos, L. Saloff-Coste, T. Coulhon, Analysis and Geometry on Groups, Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992.
[32] X. Zhang, X. Zhang, Cauchy problem of stochastic kinetic equations, arXiv:2103.02267v1, 2022.