Approximation of critical regularity functions on stratified homogeneous groups

Eduard Curcă

Approximation of critical regularity functions on stratified homogeneous groups

Eduard Curcă

Abstract

Let G be a stratified homogeneous group with homogeneous dimension Q and whose Lie algebra is generated by the left-invariant vector fields $X_1,...,X_{d_1}$. Let $1 < p, q < \infty$, $\alpha = Q/p$ and $\delta > 0$. We prove that for any function $f \in \dot{F}_q^{\alpha,p}(G)$ there exists a function $F \in L^\infty(G) \cap \dot{F}_q^{\alpha,p}(G)$ such that

$$\begin{aligned} \sum_{i=1}^{\mathbf{k}} \|X_i(f-F)\|_{\dot{F}_q^{\alpha-1,p}(G)} &\leq \delta \|f\|_{\dot{F}_q^{\alpha,p}(G)}, \\ \|F\|_{L^\infty(G)} + \|F\|_{\dot{F}_q^{\alpha,p}(G)} &\leq C_\delta \|f\|_{\dot{F}_q^{\alpha,p}(G)} \end{aligned}$$

where $\mathbf{k}$ is the largest integer smaller than $min(p, d_1)$ and C_δ is a positive constant only depending on δ. Here, $\dot{F}_q^{\alpha,p}(G)$ is a Triebel-Lizorkin type space adapted to G.

This generalizes earlier results of Bourgain, Brezis [4] and of Bousquet, Russ, Wang, Yung [6] in the Euclidean case and answers an open problem in [6].

Contents

1 Introduction

Let $B \subset \mathbf{R}^d$ ($d \geq 2$) be a Euclidean ball. It is well-known that, if $f \in L^p_{loc}(B, \mathbf{R})$ with $1 < p < \infty$, then the equation $\operatorname{div} Y = f$ has a solution $X \in W^{1,p}_{loc}(B, \mathbf{R}^d)$. When $p = d$, this Y "almost" belongs to $L^\infty_{loc}(B, \mathbf{R}^d)$. A striking result obtained by Bourgain and Brezis (in [3]) asserts that is possible to find $Y \in W^{1,d}_{loc}(B, \mathbf{R}^d) \cap L^\infty_{loc}(B, \mathbf{R}^d)$. Their argument relies on a new type of approximation results.

This seminal work has been followed by a number of approximation results of similar type [4], [5], [13], [6]. Our work is primarily motivated by two types of developments of the results in [13], [6] concerning functions in critical Sobolev spaces that barely fail the embedding in L^∞.

The first of these results (Lemma 1.7 in [13]) deals with the extension of the approximation result given in [4] (Theorem 11) in the Euclidean case, to the more general case of stratified homogeneous groups. Somewhat informally this reads (see Section 2 for definitions):

Theorem 1.1 *Suppose G is a stratified homogeneous group whose homogeneous dimension is Q and let $X_1, ..., X_{n_1}$ be a minimal family of vector fields generating the Lie algebra of G. Then, for any Schwartz function f on G and any $\delta > 0$ there exists a function F such that:*

$$\begin{aligned} \sum_{i=1}^{n_1 - 1} \|X_i(f - F)\|_{L^Q(G)} &\leq \delta \|\nabla_b f\|_{L^Q(G)}, \\ \|F\|_{L^\infty(G)} + \|\nabla_b F\|_{L^Q(G)} &\leq C_\delta \|\nabla_b f\|_{\dot{F}^{\alpha,p}_q(G)}, \end{aligned}$$

where C_δ is a constant only depending on δ.

Here, $\nabla_b f = (X_1 f, ..., X_{n_1} f)$. Theorem 11 in [4] corresponds to the Euclidean case.

On the other hand, it was proved in [6] (Theorem 1.1) that Theorem 11 in [4] remains true, in the Euclidean case if we replace the critical Sobolev space $\dot{W}^{1,d}(\mathbf{R}^d)$ by more general critical spaces such as $\dot{F}^{d/p,p}_q(\mathbf{R}^d)$. More precisely, we have the following:

Theorem 1.2 *Consider the parameters $1 < p, q < \infty$, $\alpha = d/p$ and let $\mathbf{k}$ be the largest positive integer with $\mathbf{k} < \min(p, d)$. Then, for every $\delta > 0$ there exists a constant $C_\delta > 0$ only depending on δ, such that for every function $f \in \dot{F}^{\alpha,p}_q(\mathbf{R}^d)$ there exists $F \in L^\infty(\mathbf{R}^d) \cap \dot{F}^{\alpha,p}_q(\mathbf{R}^d)$ satisfying the following estimates:*

$$\begin{aligned} \sum_{i=1}^{\mathbf{k}} \|\partial_i(f - F)\|_{\dot{F}^{\alpha-1,p}_q(\mathbf{R}^d)} &\leq \delta \|f\|_{\dot{F}^{\alpha,p}_q(\mathbf{R}^d)}, \\ \|F\|_{L^\infty(\mathbf{R}^d)} + \|F\|_{\dot{F}^{\alpha,p}_q(\mathbf{R}^d)} &\leq C_\delta \|f\|_{\dot{F}^{\alpha,p}_q(\mathbf{R}^d)}. \end{aligned}$$

Note that here we have a somewhat unnatural technical condition on $\mathbf{k}$, which does not seem to be optimal. Namely, we impose $\mathbf{k} < \min(p, d)$ instead of only imposing $\mathbf{k} < d$. (See [6] for a discussion on this assumption.)

The purpose of this paper is to find a common roof to Theorem 1.1 and Theorem 1.2 and to give an affirmative answer to Open question 1.4 in [6]. Our generalisation is an adaptation of Theorem 1.2 above to the stratified homogeneous groups context of Theorem 1.1. In this case the role of the Euclidean dimension is played by the homogeneous dimension Q of the group and the critical regularity becomes, in this case, $\alpha = Q/p$. The role of the derivatives is played by the vector fields that generate the full Lie algebra of G.

The statement of our main result is:

Theorem 1.3 *Consider the parameters $1 < p, q < \infty$, $\alpha = Q/p$ and let $\mathbf{k}$ be the largest positive integer with $\mathbf{k} < \min(p, d_1)$. Then, for every $\delta > 0$ there exists a constant $C_\delta > 0$ only depending on δ, such that, for every function $f \in \dot{F}_q^{\alpha,p}(G)$, there exists $F \in L^\infty(G) \cap \dot{F}_q^{\alpha,p}(G)$ satisfying the following estimates:*

$$\begin{aligned} \sum_{i=1}^{\mathbf{k}} \|X_i(f-F)\|_{\dot{F}_q^{\alpha-1,p}(G)} &\leq \delta \|f\|_{\dot{F}_q^{\alpha,p}(G)}, \\ \|F\|_{L^\infty(G)} + \|F\|_{\dot{F}_q^{\alpha,p}(G)} &\leq C_\delta \|f\|_{\dot{F}_q^{\alpha,p}(G)}. \end{aligned}$$

We will give in Section 2 precise definition of the function spaces we consider on G. For the time being, let us mention that we cover the case of the more familiar anisotropic homogeneous Sobolev spaces $\dot{N}L^{m,p}$, defined informally as containing the functions f on G for which $\nabla_b^m f \in L^p$.

Despite the fact that we also have the unnatural restriction $\mathbf{k} < \min(p, d_1)$, as in the Euclidean case, this suffices for some applications to divergence-like systems. Basically, all the applications to such systems presented in [4] can be easily adapted to the stratified homogeneous group setting and higher order Sobolev spaces. We give one example, formulated for simplicity for spaces of integer regularity.

Theorem 1.4 *Let $m < Q$ be a positive integer. Suppose $f \in \dot{N}L^{m-1,Q/m}(G)$ and there exist functions $v_1, ..., v_{d_1} \in \dot{N}L^{m,Q/m}(G)$ such that*

$$X_1 v_1 + ... + X_{d_1} v_{d_1} = f.$$

Then, there exist $u_1, ..., u_{d_1} \in L^\infty(G) \cap \dot{N}L^{m,Q/m}(G)$ such that

$$X_1 u_1 + ... + X_{d_1} u_{d_1} = f.$$

The paper is divided into two parts. The first one (Section 2) deals with the construction of the Triebel-Lizorkin spaces on stratified homogeneous groups. We mention that the Euclidean analogues of these spaces coincide with the classical ones and that in the general stratified homogeneous group setting, they also satisfy similar interpolation and duality properties as their classical analogues.

Spaces of a similar kind were already defined and studied for example in [1], [10] and other works (see also [9] for a construction of inhomogenous spaces in the more general context of Lie groups of polynomial volume growth). Our construction is very similar to the one given in [10] (it turns out that our spaces essentially coincide with the ones introduced in [10], as a consequence of our Proposition 2.10). While the construction in [10] is based on spectral decomposition of sublaplacians, our construction is based only on the relatively elementary technique developed in [13] for obtaining a Littlewood-Paley decomposition for functions defined on the group. (We also notice that our purpose is *not* to explore the properties of these spaces, but rather to prove a minimal number of their properties, required in the proof of Theorem 1.3.)

While in [13] Littlewood-Paley decomposition is obtained by a Calderón reproducing formula with two convolutions, we will also need similar reproducing formulas with three convolutions (we will prove that all the definitions of the spaces with two or more convolutions coincide). This allows us to prove the full analogue of the Littlewood-Paley inequality as well as other inequalities needed in the proof of Theorem 1.3.

The second part (Sections 3 and 4) is devoted to the proof of Theorem 1.3. We follow closely the proof in [6]. Several relatively minor modifications were made in order to simplify the exposition. Some more substantial adaptations were required in order to bypass the lack of commutativity of the vector fields. In some cases the arguments are easily adapted to the group setting, and in these situations we only sketch the arguments or refer to the proofs in [6]. In the Appendix we recall the Calderón-Zygmund theory on stratified homogeneous groups in order to give a direct proof of an inequality (Proposition A1) whose Euclidean analogue was proved in the Appendix of [6] by similar but more complicated means.

2 Function spaces on stratified homogeneous groups

2.1 Basic facts on stratified homogeneous groups

Here, we follow mainly Folland and Stein [8] and Stein [12]. We also present some auxiliary results, possibly known to experts, that we will need in order to develop the Littlewood-Paley theory of function spaces on stratified homogeneous groups. We will consider homogeneous groups as defined in ([12], p. 618). For such a group G, we write the following decomposition of its Lie algebra $\mathfrak{g}$:

$$\mathfrak{g} = V_1 \oplus V_2 \oplus ... \oplus V_\ell,$$

where $V_1, ..., V_\ell$ are vector spaces of left-invariant vector fields such that

(i) $[V_i, V_j] \subseteq V_{i+j}$ (making the convention that V_ℓ is not trivial and any V_j with $j > \ell$ is trivial),

(ii) V_1 generates the whole algebra $\mathfrak{g}$ (this is the so called Hörmander condition).

Dimension. We write $d_j := \dim V_j$ and $d := d_1 + ... + d_\ell$; the number $Q = d_1 + 2d_2 + ... + \ell d_\ell$ is called the homogeneous dimension of G. As sets, we identify G with $\mathbf{R}^d$. In view of this identification, we consider the following dilation rule: if $x = (x_1, ..., x_d) \in G$ and $\lambda > 0$, then $\lambda x = (\lambda^{a_1} x_1, ..., \lambda^{a_d} x_d)$, where

$$a := (a_1, ..., a_d) = (1, ..., 1, 2, ..., 2, ..., \ell, ..., \ell) \tag{2.1}$$

is the vector of the homogeneities, each $j \in \{1, ..., \ell\}$ appearing d_j times. The dilations are known to be automorphisms of G and, with respect to them, the following "norm" on G is homogeneous:

$$\|x\|_G := \left(\sum_{j=1}^{\ell} \sum_{d_1+...+d_{j-1} < i \leq d_1+...+d_j} |x_i|^{\frac{2\ell!}{j}} \right)^{\frac{1}{2\ell!}}. \tag{2.2}$$

We have also the quasi-triangle inequality

$$\|x \cdot y\|_G \lesssim \|x\|_G + \|y\|_G \,, \text{ for } x, y \in G.$$

Subgradient. We write $X_1, X_2, ..., X_d$ for the left-invariant vector fields forming the standard basis of $\mathfrak{g}$, with $X_1, X_2, ..., X_{d_1}$ forming a basis of V_1. We will call full gradient and subgradient respectively the following operators

$$\nabla := (X_1, X_2, ..., X_d) \,, \quad \nabla_b := (X_1, X_2, ..., X_{d_1}) \,.$$

Note that, whenever f is a Schwartz function on $\mathbf{R}^d$ with $\nabla_b f \equiv 0$ then, thanks to the Hörmander condition, we get $\nabla f \equiv 0$. Hence, in a sense, the subgradient encodes all the differential information about f. We will always be concerned with the subgradient of functions rather than with the full gradient. We will consider for example the Sobolev-type space $\dot{N}L^{1,Q}$, which informally is a space of functions on G whose subgradient is in L^Q. Note that this space is not the same as $\dot{W}^{1,Q}$ on G seen as a manifold.

Similar considerations hold for right-invariant vector fields. We will write X_j^R for the right-invariant analogue of X_j.

An important aspect is that, with the identification $G = \mathbf{R}^d$, we have that $x \cdot y$ is a polynomial in x, y and $(x \cdot y)_k = x_k + y_k$ for any $x, y \in G$ as long as $1 \leq k \leq d_1$. Also we have $x^{-1} = -x$ for all $x \in G$.

Balls and the maximal function. We consider balls on G defined by the quasimetric ρ on G, given by

$$\rho(x, y) = \left\| y^{-1} \cdot x \right\|_G$$

for $x, y \in G$. The open ball centered in x and of radius $\delta > 0$ is the set

$$B(x, \delta) = \{ y \in G | \rho(y, x) < \delta \},$$

whose Lebesgue measure is $|B(x, \delta)| \sim \delta^Q$. For all balls $B = B(x, \delta)$ and $\lambda > 0$ we will write $\lambda B = B(x, \lambda\delta)$.

We also consider the Hardy-Littlewood maximal function M on G, defined by

$$Mf(x) = \sup_{B \ni x} \frac{1}{|B|} \int_B |f(y)| \, dy,$$

for all functions $f \in L^1_{loc}(G)$, where the supremum is taken over all balls $B \subset G$ containing x.

We recall the following classical facts (for proofs see Chapter 2 in [12]):

Proposition 2.1 *(i) If φ is a nonnegative decreasing function on $[0, \infty)$, such that $C = \int_G \varphi(\|y\|_G) dy < \infty$ and ϕ is a measurable function on G such that $|\phi(y)| \leq \varphi(\|y\|_G)$ on G, then*

$$|f * \phi| \lesssim CMf \quad \text{on} \quad G,$$

for any Schwartz f. Here the convolution on G is defined by

$$f * \phi(x) = \int_{\mathbf{R}^d} f(y) \phi(y^{-1} \cdot x) dy = \int_{\mathbf{R}^d} f(x \cdot y^{-1}) \phi(y) dy.$$

(ii) M is of weak type $(1, 1)$ and of strong type (p, p) for all $1 < p \leq \infty$.

(iii) (the Fefferman-Stein inequality) Consider a sequence of Schwartz functions $(f_j)_{j \in \mathbf{Z}}$. Then, for $1 < p, q < \infty$, we have

$$\left\| \|Mf_j\|_{l_j^q} \right\|_{L^p} \lesssim_{p,q} \left\| \|f_j\|_{l_j^q} \right\|_{L^p}.$$

Vector fields and polynomials. We remind the following elementary formula ([12], p. 621):

$$X_j f(x) := \left. \frac{\partial f(x \cdot y)}{\partial y_j} \right|_{y=0} = \partial_j f(x) + \sum_{k > j} q_{j,k}(x) \partial_k f(x) \tag{2.3}$$

where $y = (0, ...0, y_j, 0, ...0)$ and $q_{j,k}$ are homogeneous polynomials of degree $a_k - a_j$.

Another elementary fact is that the integral of the functions of the form $X_j f$, where f is a Schwartz function is, as in the Euclidean case, equal to 0. Here is a proof of this fact. For any $y = (0, ...0, y_j, 0, ...0) \in G$, with $y_j \neq 0$, using the fact that the Lebesgue measure on $\mathbf{R}^d$ is a bi-invariant Haar measure on G ([8], Proposition (1.2), p. 3), we have

$$\int_{\mathbf{R}^d} \frac{f(x \cdot y) - f(x)}{y_j} dx = \frac{1}{y_j} \left(\int_{\mathbf{R}^d} f(x \cdot y) dx - \int_{\mathbf{R}^d} f(x) dx \right) = 0.$$

Using now the formula (2.3), the classical mean value theorem in the (Euclidean) $\mathbf{R}^d$ and the dominated convergence theorem, we can pass to the limit when $y_j \to 0$ in the above formula to obtain

$$\int_{\mathbf{R}^d} X_j f(x) dx = 0.$$

A similar formula holds for right-invariant vector fields. As an immediate consequence of this and the Leibniz rule we get the formula (see [8], p. 21)

$$\int_{\mathbf{R}^d} (X_j f)\, g dx = - \int_{\mathbf{R}^d} f\, (X_j g)\, dx \tag{2.4}$$

whenever f and g are Schwartz functions or one of them is Schwartz and the other one is polynomial.

Before going to the next step let us fix some notation. For a real valued function f sufficiently smooth on G and a positive integer m, we write $\nabla_b^m f$ for the vector valued function whose components are

$$\nabla_b^\gamma f := \left(X_1^{\gamma_1^1} X_2^{\gamma_2^1} ... X_{d_1}^{\gamma_{d_1}^1} \right) \left(X_1^{\gamma_1^2} X_2^{\gamma_2^2} ... X_{d_1}^{\gamma_{d_1}^2} \right) ... \left(X_1^{\gamma_1^m} X_2^{\gamma_2^m} ... X_{d_1}^{\gamma_{d_1}^m} \right) f \tag{2.5}$$

listed in the lexicographic order given by $\gamma = \left(\gamma_1^1, ..., \gamma_{d_1}^m\right) \in \mathbf{N}^{d_1} \times ... \times \mathbf{N}^{d_1}$ (m times) with $|\gamma| = \sum_{i,j} \gamma_j^i = m$. Note that by embedding $\mathbf{N}^{d_1} \times ... \times \mathbf{N}^{d_1}$ in $\left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$, we can define $\nabla_b^\gamma f$ by the above formula whenever $|\gamma| < \infty$.

We will also use many times the notation $\nabla_b^m \cdot \varphi$ where $\varphi = \left(\varphi_\gamma\right)_{|\gamma|=m}$ is a finite family of Schwartz functions. This has the following meaning

$$\nabla_b^m \cdot \varphi := \sum_{|\gamma|=m} \nabla_b^\gamma \varphi_\gamma. \tag{2.6}$$

Let us see that high powers of the subgradient are able to anihilate low degree polynomials. More specifically,

Proposition 2.2 *Suppose $p \in \mathbf{R}[x_1, ..., x_d]$ is a polynomial and consider $m \in \mathbf{N}^*$. Then $\nabla_b^m p$ is a vector valued polynomial with $\deg \nabla_b^m p \leq \ell \deg p - m$. In particular, if m is such that $m > \ell \deg p$, then we have that $\nabla_b^m p \equiv 0$.*

The similar assertion for the right-invariant subgradient also holds.

Proof. It suffices to prove the statement when p is a monomial. Suppose $p(x) = x^\alpha = x_1^{\alpha_1} ... x_d^{\alpha_d}$ for some $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbf{N}^d$ and consider the function $q = \nabla_b^m p$. We can see from the formula (2.3) that q is a vector valued polynomial on $\mathbf{R}^d$. Writing λx for the group dilation of

$x \in G$ with the parameter $\lambda > 0$, we immediately see from the definition of the subgradient that $\nabla_b^m (p(\lambda x)) = \lambda^m \nabla_b^m p(\lambda x)$. Also, we have

$$p(\lambda x) = (\lambda^{a_1} x)_1^{\alpha_1} ... (\lambda^{a_d} x)_d^{\alpha_d} = \lambda^{\langle a, \alpha \rangle} p(x).$$

From this we conclude that, for all $x \in G$,

$$q(\lambda x) = (\nabla_b^m p)(\lambda x) = \lambda^{-m} \nabla_b^m (p(\lambda x)) = \lambda^{\langle a, \alpha \rangle - m} \nabla_b^m p(x) = \lambda^{\langle a, \alpha \rangle - m} q(x).$$

If cx^β is a monomial ($c \neq 0$) of maximum degree in q, as before we get $(\lambda x)^\beta = \lambda^{\langle a, \beta \rangle} x^\beta$ for all $\lambda > 0$. Choosing from these monomials one for which $\langle a, \beta \rangle$ is maximum, we get by the above formula that $\langle a, \beta \rangle = \langle a, \alpha \rangle - m$ and hence $\deg q = |\beta| \leq \langle a, \beta \rangle \leq \ell |\alpha| - m$. □

Let us next recall a fundamental formula that makes a connection between the derivatives on $\mathbf{R}^d$ and the vector fields from $\mathfrak{g}$. More specifically, for any $1 \leq i \leq d$ we have ([8], p. 25)

$$\partial_i = \sum_{k=1}^{d} P_{k,i} X_k, \tag{2.7}$$

where $P_{k,i}$ are homogeneous polynomials of degree $a_k - a_i$.

We will also need the following.

Proposition 2.3 *We have that*

$$\partial_i = \sum_{k=1}^{d_1} X_k D_{k,i}^*. \tag{2.8}$$

where the operators $D_{k,i}^*$ are the adjoints of some operators of the form $\sum_\gamma p_\gamma \nabla_{\mathbf{R}^d}^\gamma$ for appropriate polynomials p_γ and multi-indexes γ in a finite subset of $\mathbf{N}^d$.

Proof. Since the vector fields $X_1, X_2, ..., X_{d_1}$ are generating the full Lie algebra of the group, we can write each X_j in terms of $X_1, X_2, ..., X_{d_1}$ using commutators, which are linear combinations of expressions of the form $\nabla_b^\gamma = \nabla_b^{\gamma'} X_k$ for some $1 \leq k \leq d_1$ and some indexes γ, $\gamma' \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$. Keeping the last vector field from such an expression and using (2.3) to express $\nabla_b^{\gamma'}$ in terms of derivatives on $\mathbf{R}^d$ and polynomials, we can rewrite (2.7) as

$$\partial_i = \sum_{k=1}^{d_1} D_{k,i} X_k, \tag{2.9}$$

where each operator $D_{k,i}$ is of the form $\sum_\gamma p_\gamma \nabla_{\mathbf{R}^d}^\gamma$ for some polynomials p_γ and γ in a finite subset of $\mathbf{N}^d$.

Now, if f and g are arbitrary Schwartz functions we can write (see (2.4)):

$$\int_{\mathbf{R}^d} f \partial_i g dx = -\int_{\mathbf{R}^d} (\partial_i f) g dx = -\sum_{k=1}^{d_1} \int_{\mathbf{R}^d} (D_{k,i} X_k f) g dx = \sum_{k=1}^{d_1} \int_{\mathbf{R}^d} f X_k \left(D_{k,i}^* g\right) dx$$

and hence, by identification,

$$\partial_i = \sum_{k=1}^{d_1} X_k D_{k,i}^*,$$

which proves the Proposition 2.3. □

Proposition 2.4 *Let $m \in \mathbf{N}$ and f be a Schwartz function.*

(i) If $f = \nabla_b^m \cdot \varphi$ for a family of Schwartz functions φ, then for any polynomial p with $\deg p < m/\ell$ we have $\int_G pf dx = 0$.

(ii) There exists an $m' \in \mathbf{N}$ only depending on m and G such that if we have $\int_G pf dx = 0$ for any polynomial p with $\deg p \leq m'$, then there exists a family of Schwartz functions φ such that $f = \nabla_b^m \cdot \varphi$.

The same is also true in the case of the right-invariant subgradient.

Remarks. (1) Since the assertion of (ii) in the above proposition remains true for any integer larger than m', when applying this part of the proposition, we will assume for technical reasons that $m' > m\ell$.

(2) In particular, Proposition 2.4 gives the following (informally speaking): if φ_1 is a Schwartz family, then there exists another Schwartz family φ_2 such that:

$$\left(\nabla_b^R\right)^{m'} \cdot \varphi_1 = \nabla_b^m \cdot \varphi_2.$$

This property will be used several times.

Proof. Part (i) follows from Proposition 2.2 and by a repeated application of the formula (2.4). Part (ii) will be proved by induction on m. The case $m = 0$ is trivial (we take by convention $m' = 0$). Fix $m \geq 1$ and suppose we have the statement of (ii) for $m-1$. Consider the number $m' := (m-1)' + M + 2$, where M is the maximum degree reached by a polynomial p_γ entering in the expression of the operators $D_{k,i}$ that occur in (2.9). If $\int_G pf dx = 0$ for any polynomial p of degree at most m', then we can use the well-known fact that in the Euclidean case there exists a collection of Schwartz families $(\phi_i)_{1 \leq i \leq d}$ such that

$$f = \sum_{i=1}^{d} \partial_i \left(\nabla_{\mathbf{R}^d}^{m'-1} \cdot \phi_i \right).$$

Using now formula (2.8) we can write:

$$f = \sum_{i=1}^{d} \partial_i \left(\nabla_{\mathbf{R}^d}^{m'-1} \cdot \phi_i \right) = \sum_{i=1}^{d} \sum_{k=1}^{d_1} X_k D_{k,i}^* \left(\nabla_{\mathbf{R}^d}^{m'-1} \cdot \phi_i \right) = \sum_{k=1}^{d_1} X_k \left(\sum_{i=1}^{d} D_{k,i}^* \nabla_{\mathbf{R}^d}^{m'-1} \cdot \phi_i \right) = \sum_{k=1}^{d_1} X_k \tilde{\phi}_k,$$

where $\tilde{\phi}_k$ are the Schwartz functions $\tilde{\phi}_k = \sum_{i=1}^{d} D_{k,i}^* \nabla_{\mathbf{R}^d}^{m'-1} \cdot \phi_i$. It is easy to see that $\int_G p \tilde{\phi}_k dx = 0$ for all polynomials p of degree at most $(m-1)'$. By the induction hypothesis, we get that for each k there exists a family of Schwartz functions φ_k such that $\tilde{\phi}_k = \nabla_b^{m-1} \cdot \varphi_k$. From this and the above formula, we get the conclusion. □

Convolutions. We recall that, for two Schwartz functions f, g their convolution is defined by the formula:

$$f * g(x) = \int_{\mathbf{R}^d} f(y) g(y^{-1} \cdot x) dy = \int_{\mathbf{R}^d} f(x \cdot y^{-1}) g(y) dy.$$

It can be verified directly that the convolution is associative.

Concerning the interaction of vector fields with the convolution, it is known that ([8], p. 22):

Proposition 2.5 *For all Schwartz functions f, g we have:*

$$X_j (f * g) = f * (X_j g), \quad X_j^R (f * g) = \left(X_j^R f\right) * g, \qquad (X_j f) * g = f * \left(X_j^R g\right). \tag{2.10}$$

We have also the following elementary fact.

Proposition 2.6 *If Φ_1, Φ_2 are two Schwartz functions, then $\Phi_1 * \Phi_2$ is also Schwartz.*

Proof. We can easily observe that, since each component of $x \cdot y$ is a polynomial in x and y, we can find a large number $n_G \in \mathbf{N}^*$ such that $1 + |x \cdot y| \lesssim (1 + |x|)^{n_G} (1 + |y|)^{n_G}$ for all $x, y \in \mathbf{R}^d$. This implies that, for example, we have

$$\begin{aligned}
\sup_x (1 + |x|)^N |\Phi_1 * \Phi_2(x)| &\leq \sup_x \int_{\mathbf{R}^d} \left(1 + \left|x \cdot y^{-1} \cdot y\right|\right)^N \left|\Phi_1(x \cdot y^{-1})\right| |\Phi_2(y)| \, dy \\
&\lesssim \sup_x \int_{\mathbf{R}^d} \left(1 + \left|x \cdot y^{-1}\right|\right)^{Nn_G} \left|\Phi_1(x \cdot y^{-1})\right| (1 + |y|)^{Nn_G} |\Phi_2(y)| \, dy \\
&\lesssim \int_{\mathbf{R}^d} (1 + |y|)^{Nn_G} |\Phi_2(y)| \, dy < \infty.
\end{aligned}$$

More generally, the estimate of $\sup_x (1 + |x|)^N \left|\partial^\beta (\Phi_1 * \Phi_2)(x)\right|$ is reduced to the above calculation using the connection between the derivatives and the vector fields on G via (2.7) and (2.3). □

2.2 The Littlewood-Paley decomposition

We introduce the following notation. Whenever Λ is a Schwartz function on G and j is an integer, we write Λ_j for the function defined by $\Lambda_j(x) = 2^{jQ}\Lambda(2^j x)$. Also, if f is another Schwartz function, we write $\Lambda_j f = f * \Lambda_j$.

Proposition 2.7 *Given $m \in \mathbf{N}$, there exist Schwartz families Λ^1, Λ^2, Λ^3 on $\mathbf{R}^d$ such that $\int_{\mathbf{R}^d} P(x)\Lambda^1(x)dx = \int_{\mathbf{R}^d} P(x)\Lambda^2(x)dx = \int_{\mathbf{R}^d} P(x)\Lambda^3(x)dx = 0$ for all the polynomials P of degree $\leq m'$ (with m' as in Proposition 2.4) and such that for all Schwartz functions f we have*

$$f = \sum_{j \in \mathbf{Z}} f * \Lambda^1_j * \Lambda^2_j * \Lambda^3_j = \sum_{j \in \mathbf{Z}} \Lambda^3_j \Lambda^2_j \Lambda^1_j f, \tag{2.11}$$

the convergence being in any $L^p(\mathbf{R}^d)$ for $1 < p < \infty$. In particular, according to Proposition 2.4 (ii), there exist families of Schwartz families φ_i, ϕ_i ($i = 1, 2, 3$) such that $\Lambda^i = \nabla_b^m \cdot \varphi_i = \left(\nabla_b^R\right)^m \cdot \phi_i$ for each $i = 1, 2, 3$.

Remark. Some explanations are in order. The proposition literally states that there exist three finite Schwartz families $\Lambda^i = (\Lambda^{i,a})_{a \in A}$ (A is a finite set), $i = 1, 2, 3$, such that all the moments of order up to m' of each $\Lambda^{i,a}$ are zero and

$$f = \sum_{j \in \mathbf{Z}} \sum_{a \in A} f * \Lambda^{1,a}_j * \Lambda^{2,a}_j * \Lambda^{3,a}_j = \sum_{j \in \mathbf{Z}} \sum_{a \in A} \Lambda^{3,a}_j \Lambda^{2,a}_j \Lambda^{1,a}_j f.$$

The last assertion means that there exists $6|A|$ Schwartz families $\varphi_{i,a}$, $\phi_{i,a}$ such that

$$\Lambda^{i,a} = \nabla_b^m \cdot \varphi_{i,a} = \left(\nabla_b^R\right)^m \cdot \phi_{i,a}$$

for all $a \in A$ and $i = 1, 2, 3$ (see (2.6)). Since the use of the familly A leads to heavy notation, we prefer the form of the above proposition which turns out to be more convenient in the calculations

that follow. This can be compared with the summation convention in geometry. We also note that the absolute value of expressions like $\Lambda_j f$, where $\Lambda = (\Lambda_a)_{a\in A}$ is a Schwartz family, will have the following meaning:

$$\left|\Lambda_j^1 f\right| = \sum_{a\in A} \left|\Lambda_j^{1,a} f\right|.$$

Similarly, we set

$$\left|\Lambda_j^2 \Lambda_j^1 f\right| = \sum_{a\in A} \left|\Lambda_j^{2,a} \Lambda_j^{1,a} f\right|,$$

and so on.

These conventions, together with (2.6), will enable us to estimate expressions involving Schwartz families as if they were *functions*. We will also abuse the notation in other situations, where the distinction between functions and finite families of functions will be clearly irrelevant (see also the conventions in [13]).

Proof. This proof follows the lines of Proposition 5.5 in [13]. We consider a radial Schwartz function Ψ with $\hat{\Psi} \equiv 1$ on $B_{\mathbf{R}^d}(0,1)$ and $supp\ \hat{\Psi} \subseteq B_{\mathbf{R}^d}(0,2)$ (here $B_{\mathbf{R}^d}(0,1)$ and $B_{\mathbf{R}^d}(0,2)$ are Euclidean balls). We need now the easy argument of Proposition 5.1 from [13] which we reproduce below for the convenience of the reader.

Lemma 2.8 *Let Φ be a Schwartz function on $\mathbf{R}^d$ such that $\int_{\mathbf{R}^d} \Phi dx = 1$ and fix some $1 < p < \infty$. Then, for any Schwartz function f, we have*

$$f = \sum_{j\in\mathbf{Z}} f * (\Phi_j - \Phi_{j-1}),$$

the convergence being in L^p.

Proof. We have, for any $N \in \mathbf{N}^*$,

$$\sum_{|j|\le N} f * (\Phi_j - \Phi_{j-1}) = f * \Phi_N - f * \Phi_{-N-1}.$$

Hence it remains to see that $f * \Phi_N \to f$ and $f * \Phi_{-N} \to 0$ in L^p when $N \to \infty$. In order to prove the first claim we write, using Minkowski's inequality,

$$\begin{aligned} \|f * \Phi_N - f\|_{L^p} &= \left\| \int_{\mathbf{R}^d} \left(f(x \cdot (2^{-N}y)^{-1}) - f(x) \right) \Phi(y) dy \right\|_{L^p_x} \\ &\le \int_{\mathbf{R}^d} \left\| f(x \cdot (2^{-N}y)^{-1}) - f(x) \right\|_{L^p_x} |\Phi(y)|\, dy \to 0 \end{aligned}$$

(this can be seen by using the dominated convergence theorem).

In order to prove the second claim, again by Minkowski's inequality we have

$$\|f * \Phi_{-N}\|_{L^p} \le \|f\|_{L^1} \|\Phi_{-N}\|_{L^p} = 2^{-NQ(1-1/p)} \|f\|_{L^1} \to 0,$$

proving the lemma. □

Proof of Proposition 2.7 continued.

The above Lemma applied to $\Phi = \Psi * \Psi * \Psi$ (see Proposition 2.6) yields

$$\begin{aligned} f &= \sum_{j\in\mathbf{Z}} f * \Big((\Psi*\Psi*\Psi)_j - (\Psi_{-1}*\Psi_{-1}*\Psi_{-1})_j \Big) = \sum_{j\in\mathbf{Z}} f * (\Psi*\Psi*\Psi - \Psi_{-1}*\Psi_{-1}*\Psi_{-1})_j \\ &= \sum_{j\in\mathbf{Z}} f * (\Psi*\Psi*(\Psi-\Psi_{-1}) + \Psi*(\Psi-\Psi_{-1})*\Psi_{-1} + (\Psi-\Psi_{-1})*\Psi_{-1}*\Psi_{-1})_j\,, \qquad (2.12) \end{aligned}$$

the convergence being in $L^p(\mathbf{R}^d)$ with $1 < p < \infty$. Since we have $\hat{\Psi} - \hat{\Psi}_{-1} \equiv 0$ in a neighborhood of 0, the function $\Psi - \Psi_{-1}$ is orthogonal to all polynomials. By applying Proposition 2.4 (ii) we can find a Schwartz family φ such that $\Psi - \Psi_{-1} = \left(\nabla_b^R\right)^{2n_2} \cdot \varphi$, with $n_2 := (2n_1)'$ where $n_1 := (m')'$. Using (2.10) we can write schematically, abusing the notation,

$$\begin{aligned} \Psi*\Psi*(\Psi-\Psi_{-1}) &= \Psi*\Psi*\left(\nabla_b^R\right)^{2n_2}\cdot\varphi = \Psi*\nabla_b^{n_2}\Psi*\left(\nabla_b^R\right)^{n_2}\varphi \\ &= \Psi*\left(\nabla_b^R\right)^{2n_1}\tilde{\Psi}*\left(\nabla_b^R\right)^{n_2}\varphi = \nabla_b^{n_1}\Psi*\left(\nabla_b^R\right)^{n_1}\tilde{\Psi}*\left(\nabla_b^R\right)^{n_2}\varphi, \end{aligned}$$

where $\tilde{\Psi}$ is a Schwartz family such that $\left(\nabla_b^R\right)^{2n_1}\tilde{\Psi} = \nabla_b^{n_2}\Psi$; this can be seen to exist thanks to Proposition 2.4 (see Remark (2)). The other terms in (2.12), namely $\Psi*(\Psi-\Psi_{-1})*\Psi_{-1}$ and $(\Psi-\Psi_{-1})*\Psi_{-1}*\Psi_{-1}$ can be handled in a similar way. We find that each one of them is a finite sum in which each term is of the form $Y_1^{(m')'}\cdot\phi_1 * Y_2^{(m')'}\cdot\phi_2 * Y_3^{(m')'}\cdot\phi_3$ where ϕ_i are Schwartz families and Y_i is ∇_b or ∇_b^R. This implies (2.11) via Proposition 2.4 (i), once we note that $(m')' > m'\ell$ (see the Remark (1) after Proposition 2.4). □

Remarks. (1) We will use sometimes the function $\Delta = \Psi*\Psi*\Psi - \Psi_{-1}*\Psi_{-1}*\Psi_{-1}$ for which, as we can see in the above proof, we have the estimate $|\Delta_j f| \le \left|\Lambda_j^3\Lambda_j^2\Lambda_j^1 f\right|$ for all integers j and all Schwartz functions f. From (2.12), we have

$$f = \sum_{j\in\mathbf{Z}} \Delta_j f \quad \text{in } L^p,\ 1 < p < \infty.$$

Schematically we write $\Delta = \Lambda^3\Lambda^2\Lambda^1$. We will also consider its weaker analogue,

$$\Delta^1 := \Lambda^2\Lambda^1. \qquad (2.13)$$

(2) It is easy to see that we can obtain decompositions of the form

$$f = \sum_{j\in\mathbf{Z}} \Lambda_j^k ... \Lambda_j^3\Lambda_j^2\Lambda_j^1 f,$$

with arbitrary $k \ge 1$ and Λ^1,..., Λ^k as in Proposition 2.7. It turns out that, for the estimates we need in this work, convolutions involving $k \ge 3$ terms are in some cases *very convenient*. Note that a decomposition formula as above with $k \ge 2$ convolutions implies a decomposition with $k-1$ convolutions. In this regard we note that even if in most cases a decomposition formula with two convolutions suffices (to define Triebel-Lizorkin spaces and to prove several of their properties), the proof of Theorem 1.3 relies on decomposition formulas with three convolutions (this will be used, for example to prove the Bernstein type inequalities (3.3)).

2.3 Definition of function spaces on stratified homogeneous groups

Let $s \in \mathbf{R}$, $p, q \in (1,\infty)$ and fix $m > |s|$ and some Schwartz families Λ^1, Λ^2 whose moments up to order m' are zero (see Proposition 2.7 and the Remarks after) and such that we have the following decomposition formula with two convolutions:

$$f = \sum_{j\in\mathbf{Z}} \Lambda_j^2\Lambda_j^1 f,$$

for any Schwartz function f.

We define the spaces $\dot{F}_q^{s,p}$ and $\dot{B}_q^{s,p}$ as being the spaces of tempered distributions f on $\mathbf{R}^d$ whose (semi)norms, respectively defined as:

$$\begin{aligned}\|f\|_{\dot{F}_q^{s,p}} &= \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sjq} \left|\Lambda_j^1 f\right|^q \right)^{1/q} \right\|_{L^p}, \\ \|f\|_{\dot{B}_q^{s,p}} &= \left(\sum_{j\in\mathbf{Z}} 2^{sjq} \left\|\Lambda_j^1 f\right\|_{L^p}^q \right)^{1/q},\end{aligned}$$

are finite.

We notice that at first sight these definitions seem to depend on the families Λ^1,Λ^2. We will show however (Proposition 2.10), that the definition of $\dot{F}_q^{s,p}$ (and of $\dot{B}_q^{s,p}$) does not depend on Λ^1, Λ^2. We will also show (Proposition 2.14) that, as expected, the space $\dot{F}_2^{n,p}$ with n a nonnegative integer, is the same as the more "classical" Sobolev space $\dot{N}L^{n,p}$.

Independence of the definition. We will need the following simple lemma:

Lemma 2.9 *Consider a sequence $(f_k)_{k\in\mathbf{Z}}$ of Schwartz functions such that all but a finite number of them are zero. Consider also an $s\in\mathbf{R}$, an integer $m>|s|$ and two finite Schwartz families Λ and Θ for which all the moments up to the order m' are zero. Then, for $1<p,q<\infty$, we have:*

$$\left\| \left(\sum_k 2^{skq} \left| \Lambda_k \sum_j \Theta_j f_j \right|^q \right)^{1/q} \right\|_{L^p} \lesssim \left\| \left(\sum_k 2^{skq} |f_k|^q \right)^{1/q} \right\|_{L^p}. \tag{2.14}$$

Proof. From the assumptions on Θ and Λ, and Proposition 2.4, we know there are some Schwartz families ϕ and φ such that $\Theta = \nabla_b^m \cdot \phi$ and $\Lambda = \left(\nabla_b^R\right)^m \cdot \varphi$. With compact notation,

$$\Theta_j * \Lambda_k = \left(\Theta * \Lambda_{k-j}\right)_j = \left(\nabla_b^m \cdot \phi * \Lambda_{k-j}\right)_j = 2^{m(k-j)} \left(\phi * \left(\left(\nabla_b^R\right)^m \Lambda \right)_{k-j} \right)_j,$$

hence,

$$\Theta_j * \Lambda_k = 2^{m(k-j)} \phi_j * \left(\left(\nabla_b^R\right)^m \Lambda \right)_k. \tag{2.15}$$

In a similar way, we get

$$\Theta_j * \Lambda_k = 2^{m(j-k)} \left((\nabla_b)^m \Theta \right)_j * \varphi_k. \tag{2.16}$$

Note that, if g, ϕ and ψ are Schwartz and j, k are two integers, then

$$\left| g * \phi_j * \psi_k \right| \lesssim M(g * \phi_j) \lesssim MMg,$$

where the implicit multiplicative constants only depend on ϕ and ψ. Using this observation and (2.15), (2.16), we can write

$$|\Lambda_k \Theta_j f_j| \lesssim 2^{-m|k-j|} MMf_j.$$

Choosing $\beta\in(0,1)$ such that $\beta m > |s|$, and using Hölder's inequality, we can write:

$$\begin{aligned}\sum_k 2^{skq} \left| \sum_j \Lambda_k \Theta_j f_j \right|^q &\lesssim \sum_k 2^{skq} \left(\sum_j 2^{-m|k-j|} MMf_j \right)^q \\ &= \sum_k 2^{skq} \left(\sum_j 2^{-(1-\beta)m|k-j|} 2^{-\beta m|k-j|} MMf_j \right)^q \\ &\lesssim \sum_k 2^{skq} \sum_j 2^{-q\beta m|k-j|} |MMf_j|^q = \sum_j \left(\sum_k 2^{skq} 2^{-q\beta m|k-j|} \right) (MMf_j)^q,\end{aligned}$$

where we had used, in the third line, the fact that

$$\left(\sum_j 2^{-q'(1-\beta)m|k-j|}\right)^{q/q'} < \infty.$$

We have now, for all $j \in \mathbf{Z}$,

$$\sum_k 2^{skq} 2^{-q\beta m|k-j|} = \sum_{k\geq j} ... + \sum_{k<j} ... = \sum_{k\geq 0} 2^{sqj} 2^{(s-\beta m)qk} + \sum_{k<0} 2^{sqj} 2^{(s+\beta m)qk} \sim 2^{sjq}$$

and, as a consequence of the above inequality,

$$\left(\sum_k 2^{skq} \left|\sum_j \Lambda_k \Theta_j f_j\right|^q\right)^{1/q} \lesssim \left(\sum_j 2^{sjq} \left(MMf_j\right)^q\right)^{1/q}.$$

Applying twice the Fefferman-Stein inequality we get (2.14). □

Now we can see that the above lemma implies the independence of the definition of the spaces of Triebel-Lizorkin type with respect to the choice of Λ^1, Λ^2. (The following statement is similar to Theorem 7 in [10].)

Proposition 2.10 *Given the parameters $s \in \mathbf{R}$, $p, q \in (1, \infty)$, the space $\dot{F}_q^{s,p}$ does not depend on the auxiliary functions Λ^1, Λ^2.*

Proof. Indeed, let $s \in \mathbf{R}$, $p, q \in (1, \infty)$, and $m_1, m_2 > |s|$. Consider, as in the definition of the Triebel-Lizorkin spaces, two couples of functions Λ^1, Λ^2 and Θ^1, Θ^2 corresponding to m_1, m_2 respectively. We can construct, using the first and the second couples of functions, the spaces $\left(\dot{F}_q^{s,p}\right)_\Lambda$ and $\left(\dot{F}_q^{s,p}\right)_\Theta$ respectively. Using Proposition 2.7 and Lemma 2.9 for $\Lambda = \Lambda^1$, $\Theta = \Theta^2$ and $f_j = \Theta_j^1 f$ for a Schwartz function f, we get, after a limiting argument that:

$$\|f\|_{\left(\dot{F}_q^{s,p}\right)_\Lambda} = \left\|\left(\sum_k 2^{skq} \left|\Lambda_k^1 \sum_j \Theta_j^2 \left(\Theta_j^1 f\right)\right|^q\right)^{1/q}\right\|_{L^p} \lesssim \left\|\left(\sum_k 2^{skq} \left|\Theta_k^1 f\right|^q\right)^{1/q}\right\|_{L^p} = \|f\|_{\left(\dot{F}_q^{s,p}\right)_\Theta}.$$

Note that in a similar way we can obtain the converse inequality. Hence, by density, we have that $\left(\dot{F}_q^{s,p}\right)_\Lambda = \left(\dot{F}_q^{s,p}\right)_\Theta$ with equivalent norms. □

Remarks. (1) The same type of independence can be proved, in a very similar way, for the Besov spaces $\dot{B}_q^{s,p}$. In this case the analogue of Lemma 2.9 is

Lemma 2.11 *Consider a sequence $(f_k)_{k\in\mathbf{Z}}$ of Schwartz functions such that all but a finite number of them are zero. Consider also an $s \in \mathbf{R}$, an integer $m > |s|$ and two finite Schwartz families Λ and Θ for which all the moments up to the order m' are zero. Then, for $1 < p, q \leq \infty$, we have:*

$$\left(\sum_k 2^{skq} \left\|\Lambda_k \sum_j \Theta_j f_j\right\|_{L^p}^q\right)^{1/q} \lesssim \left(\sum_k 2^{skq} \|f_k\|_{L^p}^q\right)^{1/q}.$$

Note that here we allow the values $p = \infty$, $q = \infty$. This is due to the fact that the Fefferman-Stein inequality is no longer needed.

(2) Lemma 2.9 can also be used to prove real and complex interpolation results for the Triebel-Lizorkin spaces with the same method of retract as for the classical spaces. In this case, the extension and retract operators $E : \dot{F}_q^{s,p} \to L^p(\dot{l}_s^q)$ and $R : L^p(\dot{l}_s^q) \to \dot{F}_q^{s,p}$ are defined by $Ef := (\Lambda_k^1 f)_{k\in\mathbf{Z}}$ and $R\,(f_k)_{k\in\mathbf{Z}} := \sum_{j\in\mathbf{Z}} \Lambda_j^2 f_j$. Lemma 2.9 is used to prove that R is well-defined and bounded, while these properties are obvious for E. Similarly for Besov spaces, relying on Lemma 2.11.

Inspecting the above proof of Proposition 2.10, we can see immediately that, by a very similar reasoning, we get the following:

Corollary 2.12 *Consider some parameters $1 < p, q < \infty$, $s \in \mathbf{R}$. Also consider an integer $m > |s|$ and a Schwartz family $\tilde{\Lambda}$ such that all its moments of order up to m' are zero. Then, for any Schwartz function f, we have:*

$$\left\| \left(\sum_k 2^{skq} \left| \tilde{\Lambda}_k f \right|^q \right)^{1/q} \right\|_{L^p} \lesssim \|f\|_{\dot{F}_q^{s,p}} .$$

The lifting property. Let us now see how Corollary 2.12 implies the lifting property for the spaces $\dot{F}_q^{s,p}$ (the following statement is similar to Corollary 21 in [10]).

Proposition 2.13 *For any Schwartz function f, we have*

$$\|\nabla_b f\|_{\dot{F}_q^{s,p}} \sim \|f\|_{\dot{F}_q^{s+1,p}} .$$

Proof. Consider some Schwartz functions Λ_j^1, Λ_j^2 for which all the moments of order up to m' are zero ($s \in \mathbf{R}$ and the integer $m > |s|$ being fixed) and such that

$$f = \sum_{j\in\mathbf{Z}} \Lambda_j^2 \Lambda_j^1 f,$$

for any Schwartz function f. Combining the definition of the Triebel-Lizorkin spaces, Proposition 2.5 and Corollary 2.12, we have

$$\begin{aligned}
\|\nabla_b f\|_{\dot{F}_q^{s,p}} &\sim \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sqj} \left| (\nabla_b f) * \Lambda_j^1 \right|^q \right)^{1/q} \right\|_{L^p} = \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sqj} \left| f * \left(\nabla_b^R \Lambda_j^1\right) \right|^q \right)^{1/q} \right\|_{L^p} \\
&= \left\| \left(\sum_{j\in\mathbf{Z}} 2^{(s+1)qj} \left| f * \left(\nabla_b^R \Lambda^1\right)_j \right|^q \right)^{1/q} \right\|_{L^p} = \left\| \left(\sum_{j\in\mathbf{Z}} 2^{(s+1)qj} \left| \tilde{\Lambda}_j^1 f \right|^q \right)^{1/q} \right\|_{L^p} \\
&\lesssim \|f\|_{\dot{F}_q^{s+1,p}} ,
\end{aligned}$$

where $\tilde{\Lambda}^1 = \nabla_b^R \Lambda^1$.

For the opposite inequality, using Proposition 2.4 and the independence of the definition (Proposition 2.10), we can assume that $\Lambda^1 = \nabla_b^R \phi$ where $\phi = \left(\nabla_b^R\right)^{m'} \psi$ for some Schwartz function ψ, and then we have:

$$\begin{aligned}
\|f\|_{\dot{F}_q^{s+1,p}} &\sim \left\| \left(\sum_{j\in\mathbf{Z}} 2^{(s+1)qj} \left| f * \Lambda_j^1 \right|^q \right)^{1/q} \right\|_{L^p} = \left\| \left(\sum_{j\in\mathbf{Z}} 2^{(s+1)qj} \left| f * \left(\nabla_b^R \phi\right)_j \right|^q \right)^{1/q} \right\|_{L^p} \\
&= \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sqj} \left| f * \nabla_b^R \phi_j \right|^q \right)^{1/q} \right\|_{L^p} = \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sqj} \left| \nabla_b f * \phi_j \right|^q \right)^{1/q} \right\|_{L^p} \\
&= \left\| \left(\sum_{j\in\mathbf{Z}} 2^{sqj} \left| \phi_j * (\nabla_b f) \right|^q \right)^{1/q} \right\|_{L^p} \lesssim \|\nabla_b f\|_{\dot{F}_q^{s,p}} .
\end{aligned}$$

Hence, for all Schwartz functions f we have $\|\nabla_b f\|_{\dot{F}_q^{s,p}} \sim \|f\|_{\dot{F}_q^{s+1,p}}$. □

The identification $\dot{F}_2^{n,p} = \dot{N}L^{n,p}$. The following statement is a generalisation of Proposition 5.7 in [13].

Proposition 2.14 *Fix an $m \in \mathbf{N}^*$ and consider Schwartz families Λ^1, Λ^2 corresponding to m as in Proposition 2.7. Then, for any Schwartz function f we have*

$$\left\| \left(\sum_{j\in\mathbf{Z}} 2^{2nj} \left| \Lambda_j^1 f \right|^2 \right)^{1/2} \right\|_{L^p} \sim \|\nabla_b^n f\|_{L^p} ,$$

for all $n \in \mathbf{N}$ with $n \leq m-1$ and $1 < p < \infty$. In other words, we have $\dot{F}_2^{n,p} = \dot{N}L^{n,p}$ with equivalent norms.

Proof. We follow the lines of Proposition 5.7 in [13], which proves a similar statement in the case $n = 1$. The estimate "$\lesssim$" easily follows by writing $\Lambda^1 = \left(\nabla_b^R\right)^{n+1} \cdot \varphi$ for a Schwartz family φ and then applying Proposition 5.4 in [13], whose statement is reproduced below in a simplified form (see also [12], Chapter 13, section 5.3):

Lemma 2.15 *If D is a Schwartz function such that $\int_G D dx = 0$, then for a fixed $1 < p < \infty$ and any Schwartz function f we have:*

$$\left\| \left(\sum_{j\in\mathbf{Z}} |D_j f|^2 \right)^{1/2} \right\|_{L^p} \lesssim \|f\|_{L^p} .$$

Using this we immediately obtain:

$$\left\| \left(\sum_{j\in\mathbf{Z}} \left| 2^{nj} \Lambda_j^1 f \right|^2 \right)^{1/2} \right\|_{L^p} = \left\| \left(\sum_{j\in\mathbf{Z}} \left| \nabla_b^n f * \left(\nabla_b^R \cdot \varphi\right)_j \right|^2 \right)^{1/2} \right\|_{L^p} \lesssim \|\nabla_b^n f\|_{L^p} .$$

For the reverse estimate we need to observe that, according to the proof of Proposition 5.5 in [13], whenever we have a decomposition of the form $f = \sum_j f * \Lambda_j * \Theta_j$ with Λ and Θ Schwartz and having zero integral, we get for any Schwartz function f that

$$\|f\|_{L^p} \lesssim \left\| \left(\sum_{j \in \mathbf{Z}} |\Lambda_j f|^2 \right)^{1/2} \right\|_{L^p}. \tag{2.17}$$

Before going further, we sketch, for the convenience of the reader the standard duality argument to prove (2.17). For all Schwartz functions g write, using the above Lemma 2.15,

$$\begin{aligned} \langle f, g \rangle &= \sum_j \langle \Theta_j \Lambda_j f, g \rangle = \sum_j \langle \Lambda_j f, \Theta_j^* g \rangle \leq \int_G \left(\sum_j |\Lambda_j f|^2 \right)^{1/2} \left(\sum_j \left|\Theta_j^* g\right|^2 \right)^{1/2} dx \\ &\leq \left\| \left(\sum_{j \in \mathbf{Z}} |\Lambda_j f|^2 \right)^{1/2} \right\|_{L^p} \left\| \left(\sum_{j \in \mathbf{Z}} \left|\Theta_j^* g\right|^2 \right)^{1/2} \right\|_{L^{p'}} \lesssim \left\| \left(\sum_{j \in \mathbf{Z}} |\Lambda_j f|^2 \right)^{1/2} \right\|_{L^p} \|g\|_{L^{p'}}. \end{aligned} \tag{2.18}$$

We obtain (2.17) by taking, in (2.18), the supremum over g such that $\|g\|_{L^{p'}} \leq 1$.

Using (2.17) with $\Lambda = \Lambda^1$ and $\Theta = \Lambda^2$, replacing f by $\nabla_b^n f$ and using (2.10) together with Corollary 2.12, we obtain:

$$\begin{aligned} \|\nabla_b^n f\|_{L^p} &\lesssim \left\| \left(\sum_{j \in \mathbf{Z}} \left|\nabla_b^n f * \Lambda_j^1\right|^2 \right)^{1/2} \right\|_{L^p} = \left\| \left(\sum_{j \in \mathbf{Z}} \left| f * \left(\left(\nabla_b^R\right)^n \Lambda_j^1 \right) \right|^2 \right)^{1/2} \right\|_{L^p} \\ &= \left\| \left(\sum_{j \in \mathbf{Z}} \left| 2^{jn} f * \left(\left(\nabla_b^R\right)^n \Lambda^1 \right)_j \right|^2 \right)^{1/2} \right\|_{L^p} = \left\| \left(\sum_{j \in \mathbf{Z}} \left| 2^{jn} \tilde{\Lambda}_j^1 f \right|^2 \right)^{1/2} \right\|_{L^p} \\ &\lesssim \|f\|_{\dot{F}_2^{n,p}} \sim \left\| \left(\sum_{j \in \mathbf{Z}} \left| 2^{jn} \Lambda_j^1 f \right|^2 \right)^{1/2} \right\|_{L^p}, \end{aligned}$$

where $\tilde{\Lambda}^1 = \left(\nabla_b^R\right)^n \Lambda^1$. This proves the proposition. □

3 Estimates of the auxiliary functions

3.1 Remark concerning the approximations

Following [6], our purpose is to prove the approximation property stated in Theorem 1.3. In the remaining part of the paper we will use decompositions formulas with three convolutions, as in Proposition 2.7.

It suffices to prove this approximation property for functions of a special form:

$$f_J := \sum_{|j| \leq J} \Lambda_j^3 \Lambda_j^2 \Lambda_j^1 f = \sum_{|j| \leq J} \Delta_j f,$$

where Λ_j^1, Λ_j^2, Λ_j^3 and $m > \alpha$ are fixed. (This particular form of the functions f_J will ensure, as we will see, that some expressions involving infinite sums and products are well-defined.) Indeed,

suppose that f is a fixed Schwartz function and for each positive integer J we can find an F_J satisfying the estimates:

$$\begin{aligned}\sum_{i=1}^{\mathbf{k}} \|X_i(f_J - F_J)\|_{\dot{F}_q^{\alpha-1,p}} &\leq \delta \|f_J\|_{\dot{F}_q^{\alpha,p}}, \\ \|F_J\|_{L^\infty} + \|F_J\|_{\dot{F}_q^{\alpha,p}} &\leq C_\delta \|f_J\|_{\dot{F}_q^{\alpha,p}}.\end{aligned}$$

Note that Lemma 2.9 immediately implies that $\|f - f_J\|_{\dot{F}_q^{\alpha,p}} \to 0$ when $J \to \infty$. By the sequential Banach-Alaoglu theorem, we can choose a subsequence $(J_k)_{k\geq 1}$ such that F_{J_k} converges weakly star in L^∞ to a function $F \in L^\infty$. Together with the last estimate and the above observation, this easily implies that $F \in \dot{F}_q^{\alpha,p}$ as follows. For any positive integer N and any compact set $K \subset G$ we have

$$\left\| \left(\sum_{|j|\leq N} 2^{\alpha q j} \left|\Lambda_j^1 F_{J_k}\right|^q \right)^{1/q} \right\|_{L^p(K)} \leq C_\delta \|f_{J_k}\|_{\dot{F}_q^{\alpha,p}} \lesssim_\delta \|f\|_{\dot{F}_q^{\alpha,p}},$$

where by $\lesssim_\delta$ we indicate that the implicit multiplicative constany may depend on δ.

Since, $\|F_{J_k}\|_{L^\infty} \lesssim_\delta \|f\|_{\dot{F}_q^{\alpha,p}}$ we get $\left\|\Lambda_j^1 F_{J_k}\right\|_{L^\infty} \lesssim_\delta \|f\|_{\dot{F}_q^{\alpha,p}}$ for all j. We also can see that $\Lambda_j^1 F_{J_k}(x) \to \Lambda_j^1 F(x)$ for every $x \in G$. Hence, the above inequality and the dominated convergence theorem imply that

$$\left\| \left(\sum_{|j|\leq N} 2^{\alpha q j} \left|\Lambda_j^1 F\right|^q \right)^{1/q} \right\|_{L^p(K)} \lesssim_\delta \|f\|_{\dot{F}_q^{\alpha,p}},$$

and from this we get the claim. Also we obtain that

$$\|F\|_{L^\infty} + \|F\|_{\dot{F}_q^{\alpha,p}} \lesssim_\delta \|f\|_{\dot{F}_q^{\alpha,p}}$$

and, in a similar way,

$$\sum_{i=1}^{\mathbf{k}} \|X_i(f - F)\|_{\dot{F}_q^{\alpha-1,p}} \leq \delta \|f\|_{\dot{F}_q^{\alpha,p}}.$$

From now, we consider J is a fixed positive integer.

3.2 Definitions and properties of some auxiliary functions

For a real number σ and $x \in G$ we will write $x_\sigma := (2^{-\sigma}x_1, ..., 2^{-\sigma}x_{\mathbf{k}}, x_{\mathbf{k}+1}, ..., x_d)$. Consider the functions $S, E : G \to \mathbf{R}$ defined by:

$$S(x) := \min(1, \|x\|_G^{-Q-1}) \text{ and } E(x) := \exp\left(-(1 + \|x_\sigma\|_G^{2\ell!})^{1/2\ell!}\right),$$

We will also consider the functions

$$S_j(x) := 2^{jQ} S(2^j x), \quad E_j(x) := 2^{jQ} E(2^j x)$$

and set $S_j f = f * S_j$. With this notation we introduce the new functions (where Δ^1 was defined in (2.13)):

$$\omega_j(x) := \left(\int_{\mathbf{R}^d} \left[\left(S_j \left|\Delta_j^1 f\right|\right)(2^{-j}r) E(r^{-1} \cdot (2^j x)) \right]^p dr \right)^{1/p}, \text{ if } |j| \leq J \text{ and 0 otherwise.}$$

Consider a smooth function $\zeta : [0,\infty) \to [0,1]$ such that $\zeta \equiv 1$ on $[0,1/2]$ and $\zeta \equiv 0$ on $[1,\infty)$. Following [13], we define the functions ζ_j as follows:

$$\zeta_j := \begin{cases} \zeta\left(\frac{2^{\alpha j}\omega_j}{\sum_{k<j,k\equiv j(\mathrm{mod}\, R)} 2^{\alpha k}\omega_k}\right), & \text{if } \sum_{k<j,k\equiv j(\mathrm{mod}\, R)} 2^{\alpha k}\omega_k \neq 0, \\ 0\,, & \text{otherwise,} \end{cases}$$

where R is a large positive integer that will be chosen later.

Using the ζ_j's, we decompose a finite sum $f_J = \sum_{|j|\leq J} \Delta_j f$ as follows:

$$f_J = \sum_{|j|\leq J} \Delta_j f = \sum_{|j|\leq J} (1-\zeta_j)\Delta_j f + \sum_{|j|\leq J} \zeta_j \Delta_j f = \sum_j h_j + \sum_j g_j = h + g$$

where

$$h := \sum_j h_j, \text{ with } h_j := (1-\zeta_j)\Delta_j f \ \text{ if } |j| \leq J \text{ and } 0 \text{ otherwise,}$$
$$g := \sum_j g_j, \text{ with } g_j := \zeta_j \Delta_j f \ \text{ if } |j| \leq J \text{ and } 0 \text{ otherwise.}$$

Then we put

$$\tilde{h} := \sum_j h_j \prod_{j'>j} (1-U_{j'}), \quad \text{with} \quad U_j := (1-\zeta_j)\omega_j,$$

$$\tilde{g} := \sum_{c=0}^{R-1} \sum_{j\equiv c(\mathrm{mod}\, R)} g_j \prod_{\substack{j'>j \\ j'\equiv c(\mathrm{mod}\, R)}} (1-G_{j'}), \ \text{ with } G_j := \sum_{\substack{t>0 \\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t}\omega_{j-t}.$$

The heart of the proof of Theorem 1.3 consists in establishing the fact that $F_J := \tilde{h} + \tilde{g}$ is a "good approximation" of $f_J = h + g$.

Pointwise and integral estimates on ω_j. Here we collect several useful estimates on ω_j in which we will see an instance of the role played by the critical condition on the exponents: $\alpha p = Q$.

In what follows we will need the following elementary approximation property proved in [13] (Proposition 3.6):

Proposition 3.1 *For any $\sigma \in \mathbf{R}$ and $x, \theta \in G$ we have:*

$$\left| \|(x\cdot\theta)_\sigma\|_G - \|x_\sigma\|_G \right| \leq C\,\|\theta\|_G \quad \text{and} \quad \left| \|(\theta\cdot x)_\sigma\|_G - \|x_\sigma\|_G \right| \leq C\,\|\theta\|_G\,.$$

In particular

$$\left| \|x\cdot\theta\|_G - \|x\|_G \right| \leq C\,\|\theta\|_G \ \text{ and } \ \left| \|\theta\cdot x\|_G - \|x\|_G \right| \leq C\,\|\theta\|_G\,.$$

Proposition 3.2 *Let $\sigma > 0$. With the above notation we have:*

(i) $\omega_j \lesssim E_j S_j \left|\Delta_j^1 f\right| \lesssim 2^{Q\sigma} MM\left(\Delta_j^1 f\right)$ for all $j \in \mathbf{Z}$;
(ii) $|\Delta_j f| \lesssim \omega_j$ for all $j \in \mathbf{Z}$;
(iii) $\|\omega_j\|_{L^\infty} \lesssim 2^{\mathbf{k}\sigma}\,\|f\|_{\dot{F}_q^{\alpha,p}}$ for all $j \in \mathbf{Z}$;
(iv) $\|U_j\|_{L^\infty} \lesssim 2^{\mathbf{k}\sigma}\,\|f\|_{\dot{F}_q^{\alpha,p}}$ for all $j \in \mathbf{Z}$;

(v) $\left\| \left\| 2^{\alpha j} \omega_j \right\|_{l_j^q} \right\|_{L^p} \lesssim 2^{Q\sigma} \left\| f \right\|_{\dot{F}_q^{\alpha,p}}$.

Proof. It is not hard to see that there exist measurable pairwise disjoint sets $M_1, M_2, ...$ covering G, such that we have $B_i \subseteq M_i \subseteq 3C \cdot B_i$ for some balls B_i of radius $1/3$ in G, where $3C \cdot B_i$ is the ball of the same center as B_i and of radius $3C$. (Here $C > 1$ is a constant such that $\rho(x,y) \leq C(\rho(x,z) + \rho(z,y))$ for all $x, y, z \in G$.) Indeed, let $(x_n)_{n\geq 1}$ be a C-net in G. That is, the balls $(B(x_n, C))_{n\geq 1}$ cover G, and $\rho(x_i, x_j) \geq C$ for all $i \neq j$. We note that, if $i \neq j$, then the balls $B(x_i, 1/3)$ and $B(x_j, 1/3)$ are disjoint. Now we put $B_i := B(x_i, 1/3)$ and $M_1 := B(x_1, C) \setminus (\cup_{j\neq 1} B_j)$, and $M_k := (B(x_k, C) \setminus (M_1 \cup ... \cup M_{k-1})) \setminus (\cup_{j\neq k} B_j)$ for all $k \geq 2$.

We observe that Proposition 3.1 implies that, for each $x, \theta \in G$ with $\left\|\theta\right\|_G \lesssim 1$ we have $E(x\cdot\theta) \sim E(\theta\cdot x) \sim E(x)$ and $S(x\cdot\theta) \sim S(\theta\cdot x) \sim S(x)$. It follows, that

$$\begin{aligned} S_j \left|\Delta_j^1 f\right|(x\cdot\theta) &= 2^{jQ} \int_{\mathbf{R}^d} \left|\Delta_j^1 f\right|(y) S\left(\left(2^j y^{-1}\right)\cdot\left(2^j x\right)\cdot\left(2^j\theta\right)\right) dy \\ &\sim 2^{jQ} \int_{\mathbf{R}^d} \left|\Delta_j^1 f\right|(y) S\left(\left(2^j y^{-1}\right)\cdot\left(2^j x\right)\right) dy \\ &= S_j \left|\Delta_j^1 f\right|(x), \end{aligned}$$

for all $x \in G$, provided $\left\|\theta\right\|_G \lesssim 2^{-j}$.

If r_i is the center of B_i, then for all r in $2B_i$, and hence for all r in M_i, we can write $r = r_i \cdot \theta$ for some θ depending on r with $\left\|\theta\right\|_G \leq 2$. Now, considering the above estimates and the decomposition $G = \bigcup_i M_i$ we can write, since $|M_i| \sim 1$,

$$\begin{aligned} \omega_j(x) &= \left(\sum_{i=1}^{\infty} \int_{M_i} \left(S_j \left|\Delta_j^1 f\right|(2^{-j} r) E(r^{-1}\cdot(2^j x))\right)^p dr\right)^{1/p} \\ &\sim \left(\sum_{i=1}^{\infty} \left(S_j \left|\Delta_j^1 f\right|(2^{-j} r_i) E(r_i^{-1}\cdot(2^j x))\right)^p\right)^{1/p} \\ &\leq \sum_{i=1}^{\infty} S_j \left|\Delta_j^1 f\right|(2^{-j} r_i) E(r_i^{-1}\cdot(2^j x)) \\ &\sim \sum_{i=1}^{\infty} \int_{M_i} S_j \left|\Delta_j^1 f\right|(2^{-j} r) E(r^{-1}\cdot(2^j x)) dr \\ &= \int_G S_j \left|\Delta_j^1 f\right|(2^{-j} r) E(r^{-1}\cdot(2^j x)) dr = E_j S_j \left|\Delta_j^1 f\right|(x). \end{aligned} \tag{3.1}$$

Next we note that $E(x) \leq \tilde{E}(x) := \exp(-\left\|2^{-\sigma} x\right\|_G)$ and therefore (using Proposition 2.1)

$$E_j S_j \left|\Delta_j^1 f\right| \leq \tilde{E}_j S_j \left|\Delta_j^1 f\right| \lesssim \tilde{E}_j M \left|\Delta_j^1 f\right| \lesssim \left\|\tilde{E}_j\right\|_{L^1} MM \left|\Delta_j^1 f\right| \lesssim 2^{Q\sigma} MM \left|\Delta_j^1 f\right|. \tag{3.2}$$

We obtain (i), from (3.1) and (3.2).

Now we prove (ii). By the change of variables $s^{-1} = r^{-1}\cdot(2^j x)$ we can write, as before,

$$\begin{aligned} \omega_j(x) &= \left(\sum_{i=1}^{\infty} \int_{M_i} \left(S_j \left|\Delta_j^1 f\right|\left(x\cdot(2^{-j} s)\right) E(s^{-1})\right)^p ds\right)^{1/p} \\ &\geq \left(\int_{M_1} \left(S_j \left|\Delta_j^1 f\right|\left(x\cdot(2^{-j} s)\right) E(s^{-1})\right)^p ds\right)^{1/p} \sim S_j \left|\Delta_j^1 f\right|(x). \end{aligned}$$

To conclude we observe that, for all $j \in \mathbf{Z}$,

$$\begin{aligned} |\Delta_j f| &\leq \left|\Lambda_j^3 \Lambda_j^2 \Lambda_j^1 f\right| = \left|\left(\Lambda_j^2 \Lambda_j^1 f\right) * \Lambda_j^3\right| \leq \left|\Lambda_j^2 \Lambda_j^1 f\right| * \left|\Lambda_j^3\right| \\ &= \left|\Delta_j^1 f\right| * \left|\Lambda_j^3\right| \lesssim \left|\Delta_j^1 f\right| * S_j = S_j \left|\Delta_j^1 f\right|, \end{aligned}$$

where we used the fact that, since Λ^3 is Schwartz, we have $|\Lambda^3| \lesssim S$ and hence $\left|\Lambda_j^3\right| \lesssim S_j$.

In order to prove (iii) we observe that, since $\alpha p = Q$,

$$\begin{aligned} \left\|\Delta_j^1 f\right\|_{L^\infty} &\lesssim \left\|\Lambda_j^1 f * \Lambda_j^2\right\|_{L^\infty} \lesssim \left\|\Lambda_j^1 f\right\|_{L^p} \left\|\Lambda_j^2\right\|_{L^{p'}} \\ &\lesssim 2^{\alpha j} \left\|\Lambda_j^1 f\right\|_{L^p} \lesssim \|f\|_{\dot{F}_q^{\alpha,p}}, \end{aligned} \tag{3.3}$$

which together with (i), the fact that $\|E_j\|_{L^1} \lesssim 2^{\mathbf{k}\sigma}$ and the Young inequality gives the estimate.

Item (iii) and the definition of U_j immediately imply (iv).

In order to prove (v), we observe that

$$\begin{aligned} \left\|\left\|2^{\alpha j} \Delta_j^1 f\right\|_{l_j^q}\right\|_{L^p} &= \left\|\left\|2^{\alpha j} \Lambda_j^2 \Lambda_j^1 f\right\|_{l_j^q}\right\|_{L^p} \lesssim \left\|\left\|2^{\alpha j} M \Lambda_j^1 f\right\|_{l_j^q}\right\|_{L^p} \\ &\lesssim \left\|\left\|2^{\alpha j} \Lambda_j^1 f\right\|_{l_j^q}\right\|_{L^p} = \|f\|_{\dot{F}_q^{\alpha,p}}, \end{aligned} \tag{3.4}$$

which, again, together with (i) and the Fefferman-Stein inequality, gives the estimate. □

Remark. Items (i), (ii) and (v) do not use the fact that $\alpha = Q/p$. In contrast, (iii) and (iv) require $\alpha = Q/p$.

Proposition 3.3 *We have* $\left\|\sup_{j\in\mathbf{Z}} 2^{\alpha j} \omega_j\right\|_{L^p} \lesssim \sigma 2^{\frac{\mathbf{k}\sigma}{p}} \|f\|_{\dot{F}_q^{\alpha,p}}$.

Proof. We follow the proof in [6] of Proposition 4.7. We have

$$\begin{aligned} \left|\sup_{j\in\mathbf{Z}} 2^{\alpha j} \omega_j(x)\right|^p &= \sup_{j\in\mathbf{Z}} \int_{\mathbf{R}^d} \left(2^{\alpha j} S_j \left|\Delta_j^1 f\right| (2^{-j} r) E(r^{-1} \cdot (2^j x))\right)^p dr \\ &\leq \int_{\mathbf{R}^d} E^p(r^{-1}) \left(\sup_{j\in\mathbf{Z}} 2^{\alpha j} S_j \left|\Delta_j^1 f\right| \left(x \cdot \left(2^{-j} r\right)\right)\right)^p dr \\ &\leq \int_{\mathbf{R}^d} E^p(r^{-1}) \left\|2^{\alpha j} S_j \left|\Delta_j^1 f\right| \left(x \cdot \left(2^{-j} r\right)\right)\right\|_{l_j^q}^p dr. \end{aligned}$$

We note that, according to (3.4), $\left\|\left\|2^{\alpha j} \Delta_j^1 f\right\|_{l_j^q}\right\|_{L^p} \lesssim \|f\|_{\dot{F}_q^{\alpha,p}}$, and hence, using Proposition A1 (see the Appendix) we get

$$\begin{aligned} \left\|\sup_{j\in\mathbf{Z}} 2^{\alpha j} \omega_j\right\|_{L^p}^p &\leq \int_{\mathbf{R}^d} E^p(r^{-1}) \left\|\left\|2^{\alpha j} S_j \left|\Delta_j^1 f\right| \left(x \cdot \left(2^{-j} r\right)\right)\right\|_{l_j^q}\right\|_{L_x^p}^p dr \\ &\lesssim \left(\int_{\mathbf{R}^d} E^p(r^{-1}) \ln^p(2 + \|r\|_G) dr\right) \|f\|_{\dot{F}_q^{\alpha,p}}^p. \end{aligned}$$

By a change of variables, we can write

$$\int_{\mathbf{R}^d} E^p(r^{-1}) \ln^p(2 + \|r\|_G) dr = 2^{\mathbf{k}\sigma} \int_{\mathbf{R}^d} \exp\left(-p(1 + \|y\|_G^{2\ell!})^{1/2\ell!}\right) \ln^p(2 + \|y_{-\sigma}\|_G) dy \lesssim \sigma^p 2^{\mathbf{k}\sigma},$$

and we get the claim. □

To make the notation more compact we introduce the functions $\mathbf{I}_m(x) = 1_{A_m}(x)$, where

$$A_m := \left\{ y \in \mathbf{R}^d \left| 2^{\alpha m}\omega_m(y) > \frac{1}{2} \sum_{k<m,\ k\equiv m(\mathrm{mod}\, R)} 2^{\alpha k}\omega_k(y) \right. \right\},\ \ m \in \mathbf{Z}.$$

With this we have:

Proposition 3.4 $\left\| \|2^{\alpha m}\omega_m \mathbf{I}_m\|_{l^q_m} \right\|_{L^p} \lesssim R\sigma 2^{\frac{\mathbf{k}\sigma}{p}} \|f\|_{\dot{F}^{\alpha,p}_q}$.

Proof. Fix a $j \in \{0, 1, ..., R-1\}$. Since $\omega_m \equiv 0$ for all but a finite number of $m \in \mathbf{Z}$, we can choose for each $x \in G$, the largest integer $m_x \equiv j(\mathrm{mod}\, R)$ with the property that $x \in A_{m_x}$, *i.e* $2^{\alpha m_x}\omega_{m_x}(x) > \frac{1}{2}\sum_{k<m_x,\ k\equiv m_x(\mathrm{mod}\, R)} 2^{\alpha k}\omega_k(x)$. Using this, we can write

$$\begin{aligned} \sum_{m\equiv j(\mathrm{mod}\, R)} 2^{\alpha m}\omega_m(x)\mathbf{I}_m(x) &\leq 2^{\alpha m_x}\omega_{m_x}(x) + \sum_{k<m_x,\ k\equiv j(\mathrm{mod}\, R)} 2^{\alpha k}\omega_k(x) \\ &\leq 3\cdot 2^{\alpha m_x}\omega_{m_x}(x) \leq 3\sup_m 2^{\alpha m}\omega_m(x) \end{aligned}$$

and hence,

$$\left\| \|2^{\alpha m}\omega_m \mathbf{I}_m\|_{l^q_m} \right\|_{L^p} \leq \left\| \sum_m 2^{\alpha m}\omega_m \mathbf{I}_m \right\|_{L^p} \leq \sum_{j=0}^{R-1} \left\| \sum_{m\equiv j(\mathrm{mod}\, R)} 2^{\alpha m}\omega_m \mathbf{I}_m \right\|_{L^p} \leq 3R \left\| \sup_m 2^{\alpha m}\omega_m \right\|_{L^p}.$$

By using Proposition 3.3, we get the claim. □

Estimates involving derivatives. Consider a smooth function u on G and $\lambda > 0$. If $a := (1, ...1, 2, ..., 2, ...\ell, ..., \ell)$ is the vector of the homogeneities of G defined in (2.1) and γ is a multi-index, then we easily see that

$$\nabla^\gamma (u(\lambda x)) = \lambda^{\langle \gamma, a\rangle} \left(\nabla^\gamma u\right)(\lambda x)$$

(where $x \to \lambda x$ is the group dilation). If moreover u is homogeneous of degree 1, we have

$$\lambda \left(\nabla_b^\gamma u\right)(x) = \nabla_b^\gamma(\lambda u(x)) = \nabla_b^\gamma(u(\lambda x)) = \lambda^{|\gamma|}\left(\nabla_b^\gamma u\right)(\lambda x)$$

and hence

$$\left(\nabla_b^\gamma u\right)(\lambda x) = \lambda^{1-|\gamma|}\left(\nabla_b^\gamma u\right)(x), \text{ for all } x \in G \text{ and } \lambda > 0.$$

Then, for all $x \neq 0$, writing $x = \lambda\nu$ where $\lambda = \|x\|_G$, $\nu = x/\|x\|_G$, we get by the above formula that

$$\left(\nabla_b^\gamma u\right)(x) = \lambda^{1-|\gamma|}\left(\nabla_b^\gamma u\right)(\nu),$$

which implies in particular that if $\|x\|_G \geq 1$ and $|\gamma| \geq 1$ then

$$|\nabla_b^\gamma u(x)| \lesssim_\gamma \|x\|_G^{1-|\gamma|} \lesssim_\gamma 1. \tag{3.5}$$

Let us also note that if $t : \mathbf{R} \to \mathbf{R}$ and $v : G \to \mathbf{R}$ are some smooth functions, then

$$X_j\left(t\left(v\left(x\right)\right)\right) = t'\left(v\left(x\right)\right) X_j v\left(x\right) \text{ for all } 1 \leq j \leq d.$$

Iterating this, we get

$$\left|\nabla_b^{\gamma}\left(t\left(v\left(x\right)\right)\right)\right| \lesssim_{\gamma} \sum_{\substack{m_1,m_2,\ldots,m_n\geq 0\\ m_1+2m_2+\ldots+nm_n=n}} \left|t^{(m_1+\ldots+m_n)}\left(v\left(x\right)\right)\right| \sum_{\substack{\gamma_1,\ldots,\gamma_n\leq\gamma\\ |\gamma_k|=k}} \prod_{j=1}^{n} \left|\nabla_b^{\gamma_j} v\left(x\right)\right|^{m_j}, \tag{3.6}$$

for all multi-indexes $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| = n$.

These observations are the basis for proving the following proposition.

Proposition 3.5 *For every* $\gamma' \in \left(\mathbf{N}^{\mathbf{k}} \times \{0\}^{d_1-\mathbf{k}}\right)^{\mathbf{N}}$ *and* $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ *with* $|\gamma| + |\gamma'| < \infty$ *(see (2.5)), we have*

$$\left|\nabla_b^{\gamma+\gamma'}\omega_j\right| \lesssim_{\gamma,\gamma'} 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}\omega_j.$$

Proof. Replacing G with $\mathbf{R}\times G$, and considering

$$u(t,x) = \|(t,x)\|_{\mathbf{R}\times G} = (|t|^{2\ell!} + \|x\|_G^{2\ell!})^{1/2\ell!},$$

we get by the above observation (3.5) that when $t = 1$, $\left|\nabla_b^{\gamma}(1+\|x\|_G^{2\ell!})^{1/2\ell!}\right| \lesssim 1$ for all finite $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$, $\gamma \neq 0$ as above.

By (3.6) we obtain

$$\left|\nabla_b^{\gamma} \exp\left(-p(1+\|x\|_G^{2\ell!})^{1/2\ell!}\right)\right| \lesssim \exp\left(-p(1+\|x\|_G^{2\ell!})^{1/2\ell!}\right)$$

and as in [6], by differentiating the composition,

$$\left|\nabla_b^{\gamma+\gamma'} E^p(r^{-1}\cdot\left(2^j x\right))\right| \lesssim 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}E^p(r^{-1}\cdot\left(2^j x\right)).$$

Consequently we have

$$\left|\nabla_b^{\gamma+\gamma'}\omega_j^p\right| \lesssim 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}\omega_j^p$$

and by writing $\omega_j = \left(\omega_j^p\right)^{1/p}$ we get the estimate by using (3.6) again. □

Proposition 3.6 *For every* $\gamma' \in \left(\mathbf{N}^{\mathbf{k}} \times \{0\}^{d_1-\mathbf{k}}\right)^{\mathbf{N}}$ *and* $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ *with* $|\gamma|+|\gamma'| < \infty$*, we have*

$$\left|\nabla_b^{\gamma+\gamma'}\zeta_j\right| \lesssim_{\gamma,\gamma'} 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}. \tag{3.7}$$

Proof. Since the proof of (3.7) follows very closely the similar estimate in [6], we only sketch the argument.

We suppose $\zeta_j \neq 0$ and write $\zeta_j = \zeta(2^{\alpha j}\omega_j/v_j)$, where $v_j = \sum_{k<j,k\equiv j(\mathrm{mod}\,R)} 2^{\alpha k}\omega_k$. From Proposition 3.5 we get

$$\left|\nabla_b^{\gamma+\gamma'}v_j\right| \lesssim 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}v_j. \tag{3.8}$$

Since $\nabla_b^{\gamma+\gamma'}(v_j/v_j)=0$, the Leibniz rule gives us,

$$\begin{aligned}\left|v_j\nabla_b^{\gamma+\gamma'}\left(\frac{1}{v_j}\right)\right| &\lesssim \sum_{\substack{\beta\le\gamma+\gamma'\\ |\beta|<|\gamma|+|\gamma'|}}\left|\nabla_b^{\gamma+\gamma'-\beta}v_j\right|\left|\nabla_b^{\beta}\left(\frac{1}{v_j}\right)\right| \\ &\le \sum_{\substack{\beta_1\le\gamma,\beta_2\le\gamma'\\ |\beta_1|+|\beta_2|<|\gamma|+|\gamma'|}}\left|\nabla_b^{(\gamma-\beta_1)+(\gamma'-\beta_2)}v_j\right|\left|\nabla_b^{\beta_1+\beta_2}\left(\frac{1}{v_j}\right)\right|. \qquad (3.9)\end{aligned}$$

The inequality (3.9) used in conjunction with (3.8) leads by a straightforward induction on $|\gamma|+|\gamma'|$, to

$$\left|\nabla_b^{\gamma+\gamma'}\left(\frac{1}{v_j}\right)\right|\lesssim 2^{j|\gamma|}2^{(j-\sigma)|\gamma'|}\frac{1}{v_j}.$$

Using this, Proposition 3.5 and the Faà di Bruno type formula (3.6) for the functions ζ and $2^{\alpha j}\omega_j/v_j$, we can conclude as in [6]. □

4 Estimates of the approximation function

4.1 Estimates of the L^∞ norm

In this subsection we are going to verify that the functions $\tilde{h}$ and $\tilde{g}$ are well-defined and, under a smallness condition on $\|f\|_{\dot{F}_q^{\alpha,p}}$ ((4.2) below), obey the L^∞ estimates:

$$\left\|\tilde{h}\right\|_{L^\infty}\lesssim 1,\qquad \|\tilde{g}\|_{L^\infty}\lesssim R. \qquad (4.1)$$

In the remaining part of the paper we assume that f satisfies

$$\|f\|_{\dot{F}_q^{\alpha,p}}\le\eta, \qquad (4.2)$$

where η is a sufficiently small number (only depending on σ, R and δ) that will be chosen later. We also assume that $R>1/\alpha$.

In order to obtain the bounds (4.1), we will need the following observation. If $(a_k)_{k\in\mathbf{Z}}$ is a sequence of finite support, then we have the identity (Lemma 3.2 in [6]):

$$\sum_{j'>j}a_{j'}\prod_{j<j''<j'}(1-a_{j''})+\prod_{j'>j}(1-a_{j'})=1. \qquad (4.3)$$

An immediate consequence of this equality is that, whenever $a_k\in[0,1]$, we must have, for all j,

$$\sum_{j'>j}a_{j'}\prod_{j<j''<j'}(1-a_{j''})\le 1. \qquad (4.4)$$

The boundedness of $\tilde{h}$. First of all we easily see that $\tilde{h}$ is well-defined (as a consequence of the fact that only a finite number of functions h_j, ω_j and U_j are nonzero). Recalling the definition of h_j and using Proposition 3.2 (ii), we can write:

$$|h_j|=\left(1-\zeta_j\right)|\Delta_j f|\lesssim\left(1-\zeta_j\right)\omega_j=U_j.$$

If f satisfies (4.2) with small η then, by Proposition 3.2 (iv), we get $U_j \in [0,1]$ for all $j \in \mathbf{Z}$ and hence, by using (4.4) and the definition of $\tilde{h}$, we get the estimate:

$$\left|\tilde{h}\right| \leq \sum_j |h_j| \prod_{j'>j}(1-U_{j'}) \lesssim \sum_j U_j \prod_{j'>j}(1-U_{j'}) \lesssim 1.$$

The boundedness of $\tilde{g}$. Let us see first that $\tilde{g}$ is well-defined. We have that all but a finite number of the functions g_j are identically zero, hence it remains to discuss the nature of the products of the form

$$\prod_{j'>j}(1-G_{j'}). \tag{4.5}$$

Following [6], we show that these products converge uniformly. Indeed, we have $\omega_j \equiv 0$ for all $j > J$. For small η in (4.2), by Proposition 3.2 (iii), we have $|\omega_j| < 1$ and thus we can write:

$$0 \leq G_j < \sum_{\substack{t>0,\ t\geq j-J \\ t\equiv 0 (\mathrm{mod}\, R)}} 2^{-\alpha t} \leq \frac{\min\left(2^{-\alpha R}, 2^{-\alpha(j-J)}\right)}{1-2^{-\alpha R}}.$$

If j is large, then we have $G_j \lesssim_R 2^{-\alpha(j-J)}$ which proves the uniform convergence of (4.5).

Now we estimate the L^∞ norm of $\tilde{g}$. When $R > 1/\alpha$, from the above inequality we get $G_j \in [0,1]$ for all j. By the definition of ζ_j, we see that $\zeta_j(x) \neq 0$ only if

$$2^{\alpha j}\omega_j(x) \leq \sum_{\substack{k<j \\ k\equiv j (\mathrm{mod}\, R)}} 2^{\alpha k}\omega_k(x).$$

Hence,

$$|g_j(x)| \lesssim \zeta_j(x)\omega_j(x) \lesssim \sum_{\substack{k<j \\ k\equiv j (\mathrm{mod}\, R)}} 2^{\alpha(k-j)}\omega_k(x) = G_j,$$

and by using (4.4) and the definition of $\tilde{g}$ we obtain,

$$|\tilde{g}| \leq \sum_{c=0}^{R-1} \sum_{j\equiv c (\mathrm{mod}\, R)} |g_j| \prod_{\substack{j'>j \\ j'\equiv c (\mathrm{mod}\, R)}} (1-G_{j'}) \lesssim \sum_{c=0}^{R-1} \sum_{j\equiv c (\mathrm{mod}\, R)} G_j \prod_{\substack{j'>j \\ j'\equiv c (\mathrm{mod}\, R)}} (1-G_{j'}) \leq R.$$

4.2 Estimating $h - \tilde{h}$

Our goal in this subsection is to prove the following estimates:

Proposition 4.1 *Suppose α, p, q and $\mathbf{k}$ are as in Theorem 1.3. Then we have*

$$\textit{(i)}\ \sum_{i=1}^{\mathbf{k}} \left\| X_i(h-\tilde{h}) \right\|_{\dot{F}_q^{\alpha-1,p}} \lesssim R\sigma^2 2^{-\sigma\min(1,\alpha)+\frac{\mathbf{k}\sigma}{p}} \|f\|_{\dot{F}_q^{\alpha,p}} + R\sigma^2 2^{\sigma\max(1-\alpha,0)+\left(1+[\alpha]+\frac{1}{p}\right)\mathbf{k}\sigma} \|f\|_{\dot{F}_q^{\alpha,p}}^2 ;$$

$$\textit{(ii)}\ \sum_{i=1}^{d_1} \left\| X_i(h-\tilde{h}) \right\|_{\dot{F}_q^{\alpha-1,p}} \lesssim R\sigma^2 2^{\frac{\mathbf{k}\sigma}{p}} \|f\|_{\dot{F}_q^{\alpha,p}} + R\sigma^2 2^{\sigma\max(1-\alpha,0)+\left(1+[\alpha]+\frac{1}{p}\right)\mathbf{k}\sigma} \|f\|_{\dot{F}_q^{\alpha,p}}^2 .$$

(Here, $[\alpha]$ stands for the integer part of α.)

Before starting the proof, we note that, writing:

$$V_j := \sum_{j'<j} h_{j'} \prod_{j'<j''<j} (1 - U_{j''}),$$

and by using the definition of $\tilde{h}$ together with the identity (4.3) (as in [13], p. 19), one obtains

$$h - \tilde{h} = \sum_j V_j U_j. \tag{4.6}$$

In order to obtain Proposition 4.1, we first collect some estimates satisfied by U_j and V_j.

Lemma 4.2 *For every $\gamma' \in \left(\mathbf{N}^{\mathbf{k}} \times \{0\}^{d_1-\mathbf{k}}\right)^{\mathbf{N}}$ and $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| + |\gamma'| < \infty$, we have*

(i) $\left|\nabla_b^{\gamma+\gamma'} U_m\right| \lesssim 2^{m|\gamma|} 2^{(m-\sigma)|\gamma'|} \omega_m \mathbf{I}_m$;
(ii) $\|\nabla_b^{\gamma} U_m\|_{L^\infty} \lesssim 2^{m|\gamma|} 2^{\mathbf{k}\sigma} \|f\|_{\dot{F}_q^{\alpha,p}}$.

Proof. As in [6], this follows from Propositions 3.2, 3.5 and 3.6. □

Lemma 4.3 *For all $m \in \mathbf{Z}$, $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| < \infty$ we have*

$$\|\nabla_b^{\gamma} h_m\|_{L^\infty} \lesssim 2^{m|\gamma|} \|f\|_{\dot{F}_q^{\alpha,p}}.$$

Proof. This is a direct consequence of the definition of h_m, (3.7) and of the Bernstein type inequality (3.3), since we have

$$\|\Delta_j f\|_{L^\infty} = \left\|\Lambda_j^3 \Delta_j^1 f\right\|_{L^\infty} \lesssim \left\|\Delta_j^1 f\right\|_{L^\infty} \lesssim \|f\|_{\dot{F}_q^{\alpha,p}}, \tag{4.7}$$

for all j. □

Lemma 4.4 *Under the smallness assumption (4.2), we have*

(i) $|V_m| \lesssim 1$,
(ii) for all $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ *with* $|\gamma| < \infty$, $\|\nabla_b^{\gamma} V_m\|_{L^\infty} \lesssim 2^{m|\gamma|} 2^{\sigma|\gamma|\mathbf{k}} \|f\|_{\dot{F}_q^{\alpha,p}}$.

Proof. We just follow the proof in [6]. Item (i) follows directly from the construction and by using (4.4). The arguments are very similar to the ones used to prove (4.1). This item is also already proved in [13] (the inequality (6.6)).

We prove now item (ii). By induction we can write (see [6] or [13], Section 6)

$$\nabla_b^{\gamma} V_m = \sum_{m'<m} \left(\nabla_b^{\gamma} h_{m'} - \sum_{0<\beta\le\gamma} c_{\gamma',\gamma} \nabla_b^{\beta} U_{m'} \nabla_b^{\gamma-\beta} V_{m'} \right) \prod_{m'<m''<m} (1 - U_{m''}). \tag{4.8}$$

This can be seen as follows. Suppose $(A_m)_{m\in\mathbf{Z}}$ and $(B_m)_{m\in\mathbf{Z}}$ are two sequences of smooth functions on G, such that for all integers m we have

$$A_m = \sum_{m'<m} B_{m'} \prod_{m'<m''<m} (1 - U_{m''}) \tag{4.9}$$

(also we assume "good" convergence properties for all the derivatives).

Then, if X is a left-invariant vector field from the Lie algebra of G, we can write

$$\begin{aligned}
XA_m &= \sum_{m'<m} (XB_{m'}) \prod_{m'<m''<m} (1-U_{m''}) \\
&\quad - \sum_{\substack{m' \\ m'<m}} B_{m'} \sum_{\substack{\nu \\ m'<\nu<m}} (XU_\nu) \prod_{m'<m''<\nu} (1-U_{m''}) \prod_{\nu<m''<m} (1-U_{m''}) \\
&= \sum_{m'<m} (XB_{m'}) \prod_{m'<m''<m} (1-U_{m''}) \\
&\quad - \sum_{\substack{\nu \\ \nu<m}} (XU_\nu) \sum_{\substack{m' \\ m'<\nu}} B_{m'} \prod_{m'<m''<m} (1-U_{m''}) \prod_{\nu<m''<m} (1-U_{m''}) \\
&= \sum_{m'<m} (XB_{m'}) \prod_{m'<m''<m} (1-U_{m''}) - \sum_{\substack{\nu \\ \nu<m}} (XU_\nu)\, A_\nu \prod_{\nu<m''<m} (1-U_{m''}) \\
&= \sum_{m'<m} (XB_{m'}) \prod_{m'<m''<m} (1-U_{m''}) - \sum_{m'<m} (XU_{m'})\, A_{m'} \prod_{m'<m''<m} (1-U_{m''}),
\end{aligned}$$

and hence, we get

$$XA_m = \sum_{m'<m} \left((XB_{m'}) - (XU_{m'})\, A_{m'} \right) \prod_{m'<m''<m} (1-U_{m''}).$$

We observe that this equality is of the same form as (4.9); in the sense that, if we now define

$$A^1_m := XA_m \quad \text{and} \quad B^1_m := (XB_m) - (XU_m)\, A_m,$$

then

$$A^1_m = \sum_{m'<m} B^1_{m'} \prod_{m'<m''<m} (1-U_{m''}).$$

Appliyng this iteratively, using the definition of V_m, we get (4.8).

Now, by using Lemmas 4.3 and 4.2,

$$\begin{aligned}
\|\nabla^\gamma_b V_m\|_{L^\infty} &\lesssim \sum_{m'<m} \left(\|\nabla^\gamma_b h_{m'}\|_{L^\infty} + \sum_{0<\gamma'\le\gamma} \left\|\nabla^\beta_b U_{m'}\right\|_{L^\infty} \left\|\nabla^{\gamma-\beta}_b V_{m'}\right\|_{L^\infty} \right) \|f\|_{\dot{F}^{\alpha,p}_q} \\
&\lesssim \sum_{m'<m} \left(2^{m'|\gamma|} + \sum_{0<\beta\le\gamma} 2^{m'|\beta|} 2^{\mathbf{k}\sigma} \left\|\nabla^{\gamma-\beta}_b V_{m'}\right\|_{L^\infty} \right) \|f\|_{\dot{F}^{\alpha,p}_q}
\end{aligned}$$

and by induction on $|\gamma|$ we get the inequality in item (ii). (Recall that we work under the smallness assumption (4.2).) □

We are now in position to complete the proof of Proposition 4.1.

Proof of Proposition 4.1. We prove (i) in detail, following closely [6]. As in [6], for all $1 \leq k \leq \mathbf{k}$, we write

$$\begin{aligned}
\left\| X_k(h-\tilde{h}) \right\|_{\dot{F}_q^{\alpha-1,p}} &= \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k(h-\tilde{h}) \right\|_{l_m^q} \right\|_{L^p} \\
&= \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k \left(\sum_{j\in\mathbf{Z}} V_j U_j \right) \right\|_{l_m^q} \right\|_{L^p} \\
&= \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k \left(\sum_{r\in\mathbf{Z}} V_{r+m} U_{r+m} \right) \right\|_{l_m^q} \right\|_{L^p} \\
&\leq \sum_{r\in\mathbf{Z}} \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k (U_{r+m} V_{r+m}) \right\|_{l_m^q} \right\|_{L^p}.
\end{aligned}$$

We split this last sum in three terms $\sum_{r>\sigma}$, $\sum_{r<0}$, $\sum_{0\leq r\leq\sigma}$.

(I) Estimate of $\sum_{r>\sigma}$. Following [6] and using (2.10), we have:

$$\begin{aligned}
\left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k (U_{r+m} V_{r+m}) \right\|_{l_m^q} \right\|_{L^p} &= \left\| \left\| 2^{(\alpha-1)m} (U_{r+m} V_{r+m}) * X_k^R \Lambda_m^1 \right\|_{l_m^q} \right\|_{L^p} \\
&= \left\| \left\| 2^{\alpha m} (U_{r+m} V_{r+m}) * \left(X_k^R \Lambda^1 \right)_m \right\|_{l_m^q} \right\|_{L^p} \\
&\lesssim \left\| \| 2^{\alpha m} M(U_{r+m} V_{r+m}) \|_{l_m^q} \right\|_{L^p} \\
&\lesssim \left\| \| 2^{\alpha m} (U_{r+m} V_{r+m}) \|_{l_m^q} \right\|_{L^p} \lesssim \left\| \| 2^{\alpha m} U_{r+m} \|_{l_m^q} \right\|_{L^p} \\
&= 2^{-\alpha r} \left\| \| 2^{\alpha m} U_m \|_{l_m^q} \right\|_{L^p}. \qquad (4.10)
\end{aligned}$$

Recalling that $U_j = (1-\zeta_j)\omega_j$ and using Proposition 3.4 we get

$$\left\| \| 2^{\alpha m} U_m \|_{l_m^q} \right\|_{L^p} \lesssim \left\| \| 2^{\alpha m} \omega_m \mathbf{I}_m \|_{l_m^q} \right\|_{L^p} \lesssim R\sigma 2^{\mathbf{k}\sigma/p} \|f\|_{\dot{F}_q^{\alpha,p}},$$

and summing up,

$$\sum_{r>\sigma} \ldots \lesssim \sum_{r>\sigma} \left(2^{-\alpha r} R\sigma 2^{\mathbf{k}\sigma/p} \|f\|_{\dot{F}_q^{\alpha,p}} \right) \lesssim R\sigma 2^{-\alpha\sigma+\mathbf{k}\sigma/p} \|f\|_{\dot{F}_q^{\alpha,p}}.$$

(II) Estimate of $\sum_{r<0}$. If $a := [\alpha]$ then, as we have already seen, we can write $\Lambda^1 = \left(\nabla_b^R\right)^a \varphi$ for a Schwartz family φ, and then $\Lambda_m^1 = 2^{-ma} \left(\nabla_b^R\right)^a \varphi_m$. Hence, if X_k is a vector field in a "good" direction, i.e. $1 \leq k \leq \mathbf{k}$, we have

$$\begin{aligned}
\left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 X_k (U_{m+r} V_{m+r}) \right\|_{l_m^q} \right\|_{L^p} &= \left\| \left\| 2^{(\alpha-1)m} 2^{-ma} X_k (U_{m+r} V_{m+r}) * \left(\nabla_b^R\right)^a \varphi_m \right\|_{l_m^q} \right\|_{L^p} \\
&= \left\| \left\| 2^{(\alpha-1)m} 2^{-ma} \left[\nabla_b^a X_k (U_{m+r} V_{m+r}) \right] * \varphi_m \right\|_{l_m^q} \right\|_{L^p} \\
&\lesssim \left\| \left\| 2^{(\alpha-1)m} 2^{-ma} M \nabla_b^a X_k (U_{m+r} V_{m+r}) \right\|_{l_m^q} \right\|_{L^p} \\
&\lesssim \left\| \left\| 2^{(\alpha-1)m} 2^{-ma} \nabla_b^a X_k (U_{m+r} V_{m+r}) \right\|_{l_m^q} \right\|_{L^p} \\
&\lesssim 2^{-(\alpha-1-a)r} \left\| \left\| 2^{(\alpha-1-a)m} \nabla_b^a X_k (U_m V_m) \right\|_{l_m^q} \right\|_{L^p},
\end{aligned}$$

where we have used the Fefferman-Stein inequality in the third line.

As in [6], using the Leibniz rule and Lemmas 4.2 and 4.4, we obtain

$$\begin{aligned}
\left|\nabla_b^a X_k(U_m V_m)\right| &\lesssim \left|V_m\left(\nabla_b^a X_k U_m\right)\right| + \sum_{l=0}^{a}\left|\nabla_b^l U_m\right|\left|\nabla_b^{a+1-l} V_m\right| \\
&\lesssim 2^{ma}2^{m-\sigma}\omega_m\mathbf{I}_m + \sum_{l=0}^{a}\left(2^{ml}\omega_m\mathbf{I}_m\right)\left(2^{m(a+1-l)}2^{\mathbf{k}(a+1-l)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}\right) \\
&\lesssim 2^{m(a+1)}\left(2^{-\sigma}+2^{\mathbf{k}(a+1)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}\right)\omega_m\mathbf{I}_m.
\end{aligned}$$

Now we get, via Proposition 3.4,

$$\begin{aligned}
\left\|\left\|2^{(\alpha-1)m}\Lambda_m^1 X_k(U_{m+r}V_{m+r})\right\|_{l_m^q}\right\|_{L^p} &\lesssim 2^{-(\alpha-1-a)r}\left(2^{-\sigma}+2^{\mathbf{k}(a+1)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}\right)\left\|\left\|2^{\alpha m}\omega_m\mathbf{I}_m\right\|_{l_m^q}\right\|_{L^p} \\
&\lesssim R\sigma 2^{\frac{\mathbf{k}\sigma}{p}}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}2^{-(\alpha-1-a)r}\left(2^{-\sigma}+2^{\mathbf{k}(a+1)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}\right)
\end{aligned}$$

and, summing up,

$$\sum_{r<0}...\lesssim R\sigma\left(2^{-\sigma+\frac{\mathbf{k}\sigma}{p}}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}+2^{\mathbf{k}\left(a+1+\frac{1}{p}\right)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}^2\right).$$

(III) Estimate of $\sum_{0\leq r\leq\sigma}$. This is similar to the preceding estimate. Here, instead of taking a to be the integer part of α, we consider $a=0$. As above we conclude that

$$\left\|\left\|2^{(\alpha-1)m}\Lambda_m^1 X_k(U_{m+r}V_{m+r})\right\|_{l_m^q}\right\|_{L^p}\lesssim R\sigma 2^{\frac{\mathbf{k}\sigma}{p}}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}2^{-(\alpha-1)r}\left(2^{-\sigma}+2^{\mathbf{k}\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}\right),$$

and by summing up,

$$\sum_{0\leq r\leq\sigma}...\lesssim C_\alpha(\sigma)\,R\sigma\left(2^{-\sigma+\frac{\mathbf{k}\sigma}{p}}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}+2^{\mathbf{k}\left(1+\frac{1}{p}\right)\sigma}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}^2\right)$$

where $C_\alpha(\sigma)\sim 1$ if $\alpha>1$, $C_\alpha(\sigma)\sim\sigma$ if $\alpha=1$ and $C_\alpha(\sigma)\sim 2^{(1-\alpha)\sigma}$ if $\alpha<1$.

With this we have proved (i). The proof of (ii) follows the same lines as the one of (i). The main difference is that since we are no longer restricted to the case of derivatives in "good" directions, we have to use, instead of Lemma 4.2 (i) applied with $|\gamma'|=1$ (as in (II) and implicitly in (III) above), the weaker statement for the case $|\gamma'|=0$. This will produce almost the same estimates, the difference being that the coefficient $2^{-\sigma+\frac{\mathbf{k}\sigma}{p}}$ of $\left\|f\right\|_{\dot{F}_q^{\alpha,p}}$ in the corresponding parts (I), (II) becomes $2^{\frac{\mathbf{k}\sigma}{p}}$. □

4.3 Estimating $g-\tilde{g}$

Our goal in this section is to prove the following counterpart of Proposition 4.1.

Proposition 4.5 *Consider $1<p,q<\infty$ and $\alpha=Q/p$. Also consider $a_\alpha\in(0,\alpha]$ such that $a_\alpha<\alpha/(1-\alpha)$ if $\alpha<1$ and $a_\alpha=1$ if $\alpha\geq 1$. We have*

$$\left\|\nabla_b(g-\tilde{g})\right\|_{\dot{F}_q^{\alpha-1,p}}\lesssim 2^{Q\sigma}R2^{-\min(1,\alpha a_\alpha)R}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}+2^{([\alpha]+1)Q\sigma}R^2 2^{-\min(1,\alpha a_\alpha)R}\left\|f\right\|_{\dot{F}_q^{\alpha,p}}^2.$$

We recall the definition of G_j:

$$G_j := \sum_{\substack{t>0 \\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t}\omega_{j-t}.$$

The starting point is the identity (similar to (4.6))

$$g - \tilde{g} = \sum_j G_j H_j,$$

where

$$H_j := \sum_{\substack{j'<j \\ j'\equiv j(\mathrm{mod}\, R)}} g_{j'} \prod_{\substack{j'<j''<j \\ j''\equiv j(\mathrm{mod}\, R)}} (1 - G_{j''})$$

and $g_j = \zeta_j \Delta_j f$.

Lemma 4.6 *For all $m \in \mathbf{Z}$, $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| < \infty$,*

$$|\nabla_b^\gamma G_m| \lesssim_\gamma 2^{Q\sigma} \sum_{\substack{t>0 \\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} 2^{|\gamma|(m-t)} MM\left|\Delta_{m-t}^1 f\right|.$$

Proof. By the definition of G_m and Proposition 3.5,

$$|\nabla_b^\gamma G_m| \le \sum_{\substack{t>0 \\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \left|\nabla_b^\gamma \omega_{m-t}\right| \lesssim \sum_{\substack{t>0 \\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} 2^{|\gamma|(m-t)} \omega_{m-t}.$$

Note now that, according to Proposition 3.2,

$$\omega_{m-t} \lesssim 2^{Q\sigma} MM\left|\Delta_{m-t}^1 f\right|,$$

whence the estimate. □

Lemma 4.7 *For all $m \in \mathbf{Z}$, $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| < \infty$,*

$$|\nabla_b^\gamma g_m| \lesssim_\gamma 2^{|\gamma| m} M\left|\Delta_m^1 f\right|.$$

Proof. By Proposition 3.6 and the Leibniz rule, recalling the definition of g_m, we have

$$\begin{aligned} |\nabla_b^\gamma g_m| &\lesssim \sum_{0\le\gamma'\le\gamma} 2^{|\gamma-\gamma'|m} \left|\nabla_b^{\gamma'}\left(\Lambda_m^3\left(\Delta_m^1 f\right)\right)\right| \lesssim \sum_{0\le\gamma'\le\gamma} 2^{|\gamma-\gamma'|m} \left|\left(\Delta_m^1 f\right) * \nabla_b^{\gamma'}\Lambda_m^3\right| \\ &\lesssim \sum_{0\le\gamma'\le\gamma} 2^{|\gamma-\gamma'|m} 2^{|\gamma'|m} M\left|\Delta_m^1 f\right| \lesssim_\gamma 2^{|\gamma|m} M\left|\Delta_m^1 f\right| \end{aligned}$$

(since $|\gamma - \gamma'| = |\gamma| - |\gamma'|$ when $0 \le \gamma' \le \gamma$). □

Lemma 4.8 *For all $m \in \mathbf{Z}$, $\gamma \in \left(\mathbf{N}^{d_1}\right)^{\mathbf{N}}$ with $|\gamma| < \infty$, and under the smallness condition (4.2) on f, we have*

(i) $|H_m| \lesssim 1$,
(ii) $|\nabla_b^\gamma H_m| \lesssim 2^{|\gamma|Q\sigma} \sum_{t>0,\ t\equiv 0(\mathrm{mod}\, R)} 2^{|\gamma|(m-t)} MM\left|\Delta_{m-t}^1 f\right|$.

Proof. Item (i) follows directly from the construction. Also, it is proved in [13] (Section 11).

Item (ii) is obtained following the strategy in [6] (Lemma 6.5). The proof is similar to the one of Lemma 4.4. It is done by induction on $|\gamma|$ and using Lemmas 4.6, 4.7. □

Proof of Proposition 4.5. As in the estimate of $h-\tilde{h}$, we can write

$$\|\nabla_b (g-\tilde{g})\|_{\dot{F}_q^{\alpha-1,p}} \leq \sum_{r\in\mathbf{Z}} \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 \nabla_b (G_{r+m} H_{r+m}) \right\|_{l_m^q} \right\|_{L^p}.$$

Recalling that

$$G_{r+m} := \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \omega_{r+m-t},$$

we get

$$\begin{aligned}
\|\nabla_b (g-\tilde{g})\|_{\dot{F}_q^{\alpha-1,p}} &\leq \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{r\in\mathbf{Z}} \left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 \nabla_b (\omega_{r+m-t} H_{r+m}) \right\|_{l_m^q} \right\|_{L^p} \\
&= \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{r>a_\alpha t} ... + \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{r\leq 0} ... + \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{0<r\leq a_\alpha t} ....
\end{aligned}$$

(I) Estimate of $\sum_{r>a_\alpha t}$. Using the fact that $\|H_m\|_{L^\infty} \lesssim 1$ and Proposition 3.2 we have (as in (4.10)) :

$$\left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 \nabla_b (\omega_{r+m-t} H_{r+m}) \right\|_{l_m^q} \right\|_{L^p} \lesssim 2^{-\alpha(r-t)} \left\| \|2^{\alpha m} \omega_m\|_{l_m^q} \right\|_{L^p} \lesssim 2^{-\alpha(r-t)} 2^{Q\sigma} \|f\|_{\dot{F}_q^{\alpha,p}}.$$

Summing up we get:

$$\begin{aligned}
\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{r>a_\alpha t} ... &\lesssim \left(\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{r>a_\alpha t} 2^{-\alpha(r-t)} \right) 2^{Q\sigma} \|f\|_{\dot{F}_q^{\alpha,p}} \\
&= \left(\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} \sum_{r>a_\alpha t} 2^{-\alpha r} \right) 2^{Q\sigma} \|f\|_{\dot{F}_q^{\alpha,p}} \\
&\lesssim \sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha a_\alpha t} 2^{Q\sigma} \|f\|_{\dot{F}_q^{\alpha,p}} \lesssim 2^{-\alpha a_\alpha R} 2^{Q\sigma} \|f\|_{\dot{F}_q^{\alpha,p}}.
\end{aligned}$$

(II) Estimate of $\sum_{r\leq 0}$. Let $a \geq 0$ be an integer. As in the estimate (II) for $h-\tilde{h}$ we obtain

$$\left\| \left\| 2^{(\alpha-1)m} \Lambda_m^1 \nabla_b (\omega_{r+m-t} H_{r+m}) \right\|_{l_m^q} \right\|_{L^p} \lesssim 2^{-(\alpha-1-a)r} \left\| \left\| 2^{(\alpha-1-a)m} \nabla_b^{a+1} (\omega_{m-t} H_m) \right\|_{l_m^q} \right\|_{L^p}.$$

In order to estimate the right hand side we recall that the following estimates hold (see Proposition 3.2, Proposition 3.5 and Lemma 4.8):

$$\begin{aligned}
\omega_{m-t} &\lesssim 2^{Q\sigma} MM(\Delta_{m-t}^1 f), \quad \left|\nabla_b^l \omega_{m-t}\right| \lesssim 2^{(m-t)l} \omega_{m-t}; \\
|H_m| &\lesssim 1, \quad \left|\nabla_b^l H_m\right| \lesssim 2^{lQ\sigma} \sum_{t>0} 2^{(m-t)l} MM\left|\Delta_{m-t}^1 f\right|
\end{aligned}$$

for all $l \in \mathbf{N}$. By using now the Leibniz rule we get:

$$\begin{aligned}\left|\nabla_b^{a+1}(\omega_{m-t}H_m)\right| &\lesssim 2^{(m-t)(a+1)}\omega_{m-t} \\ &\quad +2^{(a+1)Q\sigma}\sum_{t'>0}\sum_{l=0}^{a}2^{(t'-t)l}2^{(a+1)(m-t')}MM\left(\Delta_{m-t}^1 f\right)MM\left(\Delta_{m-t'}^1 f\right).\end{aligned} \tag{4.11}$$

Using (3.3), we estimate the double sum from the right hand side as follows:

$$\begin{aligned}\sum_{t'>0}\sum_{l=0}^{a}\ldots &\lesssim \|f\|_{\dot{F}_q^{\alpha,p}}\left(\sum_{0<t'\le t}2^{(a+1)(m-t')}MM\left(\Delta_{m-t'}^1 f\right)+\sum_{t'>t}2^{(t'-t)a}2^{(a+1)(m-t')}MM\left(\Delta_{m-t}^1 f\right)\right)\\ &\lesssim \|f\|_{\dot{F}_q^{\alpha,p}}\sum_{0<t'\le t}2^{(a+1)(m-t')}MM\left(\Delta_{m-t'}^1 f\right).\end{aligned}$$

Going back to (4.11), we obtain

$$\left|\nabla_b^{a+1}(\omega_{m-t}H_m)\right|\lesssim 2^{(m-t)(a+1)}\omega_{m-t}+2^{(a+1)Q\sigma}\|f\|_{\dot{F}_q^{\alpha,p}}\sum_{0<t'\le t}2^{(a+1)(m-t')}MM\left(\Delta_{m-t'}^1 f\right)$$

and hence, using Proposition 3.2, the term $\left\|\left\|2^{(\alpha-1-a)m}\nabla_b^{a+1}(\omega_{m-t}H_m)\right\|_{l_m^q}\right\|_{L^p}$ is bounded by

$$\begin{aligned}&2^{(\alpha-1-a)t}\left\|\left\|2^{\alpha(m-t)}\omega_{m-t}\right\|_{l_m^q}\right\|_{L^p}+2^{(a+1)Q\sigma}\|f\|_{\dot{F}_q^{\alpha,p}}B_{a,\alpha}(t)\left\|\left\|2^{\alpha m}MM\left(\Delta_m^1 f\right)\right\|_{l_m^q}\right\|_{L^p}\\ \lesssim\ &2^{(\alpha-1-a)t}\left\|\left\|2^{\alpha m}\omega_m\right\|_{l_m^q}\right\|_{L^p}+2^{(a+1)Q\sigma}\|f\|_{\dot{F}_q^{\alpha,p}}B_{a,\alpha}(t)\left\|\left\|2^{\alpha m}\Delta_m^1 f\right\|_{l_m^q}\right\|_{L^p}\\ \lesssim\ &2^{Q\sigma}2^{(\alpha-1-a)t}\|f\|_{\dot{F}_q^{\alpha,p}}+2^{(a+1)Q\sigma}\|f\|_{\dot{F}_q^{\alpha,p}}B_{a,\alpha}(t)\left\|\left\|2^{\alpha m}\Lambda_m^1 f\right\|_{l_m^q}\right\|_{L^p}\\ \lesssim\ &2^{Q\sigma}2^{(\alpha-1-a)t}\|f\|_{\dot{F}_q^{\alpha,p}}+2^{(a+1)Q\sigma}B_{a,\alpha}(t)\|f\|_{\dot{F}_q^{\alpha,p}}^2,\end{aligned}$$

where $B_{a,\alpha}(t)=\sum_{0<t'\le t}2^{(\alpha-1-a)t'}$. Finally, we obtain

$$\begin{aligned}\left\|\left\|2^{(\alpha-1)m}\Lambda_m^1\nabla_b(\omega_{r+m-t}H_{r+m})\right\|_{l_m^q}\right\|_{L^p} &\lesssim 2^{-(\alpha-1-a)r}2^{Q\sigma}2^{(\alpha-1-a)t}\|f\|_{\dot{F}_q^{\alpha,p}}\\ &\quad +2^{-(\alpha-1-a)r}2^{(a+1)Q\sigma}B_{a,\alpha}(t)\|f\|_{\dot{F}_q^{\alpha,p}}^2.\end{aligned} \tag{4.12}$$

If we choose now $a=[\alpha]$ and we observe that in this case we have $B_{a,\alpha}(t)\lesssim 1$, we can write

$$\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\,R)}}2^{-\alpha t}\sum_{r\le 0}\ldots\lesssim 2^{Q\sigma}2^{-(a+1)R}\|f\|_{\dot{F}_q^{\alpha,p}}+2^{(a+1)Q\sigma}2^{-\alpha R}\|f\|_{\dot{F}_q^{\alpha,p}}^2.$$

(III) Estimate of $\sum_{0\le r\le a_\alpha t}$. Using the estimate (4.12) above with $a=0$, we get

$$\sum_{0\le r\le a_\alpha t}\left\|\left\|2^{(\alpha-1)m}\Lambda_m^1\nabla_b(\omega_{r+m-t}H_{r+m})\right\|_{l_m^q}\right\|_{L^p}\lesssim A_\alpha(t)\left(2^{Q\sigma}2^{(\alpha-1)t}\|f\|_{\dot{F}_q^{\alpha,p}}+2^{Q\sigma}B_{0,\alpha}(t)\|f\|_{\dot{F}_q^{\alpha,p}}^2\right),$$

where

$$A_\alpha(t)\lesssim\begin{cases}2^{(1-\alpha)a_\alpha t} & \text{if } \alpha<1,\\ a_\alpha t & \text{if } \alpha=1,\\ 1 & \text{if } \alpha>1,\end{cases}\quad\text{and } B_{0,\alpha}(t)\lesssim\begin{cases}1 & \text{if } \alpha<1,\\ t & \text{if } \alpha=1,\\ 2^{(1-\alpha)t} & \text{if } \alpha>1.\end{cases}$$

Now summing up we get three possible bounds:

(1) if $\alpha < 1$, we have $a_\alpha < \frac{\alpha}{1-\alpha}$ and

$$\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{0\le r\le a_\alpha t} ... \lesssim 2^{Q\sigma} 2^{-R(1-(1-\alpha)a_\alpha)} \|f\|_{\dot{F}_q^{\alpha,p}} + 2^{Q\sigma} 2^{-R(\alpha-(1-\alpha)a_\alpha)} \|f\|^2_{\dot{F}_q^{\alpha,p}};$$

(2) if $\alpha = 1$, we have $a_\alpha = 1$ and

$$\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{0\le r\le a_\alpha t} ... \lesssim 2^{Q\sigma} R 2^{-R} \|f\|_{\dot{F}_q^{\alpha,p}} + 2^{Q\sigma} R^2 2^{-R} \|f\|^2_{\dot{F}_q^{\alpha,p}};$$

(3) if $\alpha > 1$, we have $a_\alpha = 1$ and

$$\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{0\le r\le a_\alpha t} ... \lesssim 2^{Q\sigma} 2^{-R} \|f\|_{\dot{F}_q^{\alpha,p}} + 2^{Q\sigma} 2^{-R} \|f\|^2_{\dot{F}_q^{\alpha,p}}.$$

Now from the above estimates, since $0 < a_\alpha < \alpha$, we have

$$\sum_{\substack{t>0\\ t\equiv 0(\mathrm{mod}\, R)}} 2^{-\alpha t} \sum_{0\le r\le a_\alpha t} ... \lesssim 2^{Q\sigma} R 2^{-R\min(1,\alpha a_\alpha)} \|f\|_{\dot{F}_q^{\alpha,p}} + 2^{Q\sigma} R^2 2^{-R\min(1,\alpha a_\alpha)} \|f\|^2_{\dot{F}_q^{\alpha,p}}.$$

Toghether with (I) and (II), this gives Proposition 4.5. □

4.4 Proof of Theorem 1.3

Now we can estimate the Triebel-Lizorkin norm of $f_J - F_J = (h - \tilde{h}) + (g - \tilde{g})$. By Proposition 4.1 (i) and Proposition 4.5, we have

$$\begin{aligned}\sum_{i=1}^{\mathbf{k}} \|X_i(f_J - F_J)\|_{\dot{F}_q^{\alpha-1,p}} &\lesssim \sum_{i=1}^{\mathbf{k}} \left\|X_i(h-\tilde{h})\right\|_{\dot{F}_q^{\alpha-1,p}} + \sum_{i=1}^{\mathbf{k}} \|X_i(g-\tilde{g})\|_{\dot{F}_q^{\alpha-1,p}} \\ &\lesssim \left(R\sigma^2 2^{-\sigma\min(1,\alpha)-\frac{\mathbf{k}\sigma}{p}} + 2^{Q\sigma} R 2^{-\min(1,\alpha a_\alpha)R}\right) \|f\|_{\dot{F}_q^{\alpha,p}} + D_{R,\sigma} \|f\|^2_{\dot{F}_q^{\alpha,p}},\end{aligned}$$

where $D_{R,\sigma}$ is a large constant depending on R and σ.

As in [6], for $\sigma \in \mathbf{N}$, we set

$$R = R_\sigma := \left[\frac{100Q}{\min(1, \alpha a_\alpha)}\right] \sigma.$$

If $\delta > 0$ is fixed, then it is easy to see that (using the fact that $\mathbf{k}/p < \min(1,\alpha)$) for a σ large enough, we have

$$R\sigma^2 2^{-\sigma\min(1,\alpha)+\frac{\mathbf{k}\sigma}{p}} \lesssim \delta/4 \quad \text{and} \quad 2^{Q\sigma} R 2^{-\min(1,\alpha a_\alpha)R} \lesssim \delta/4.$$

Hence, for a large D_δ we have

$$\sum_{i=1}^{\mathbf{k}} \|X_i(f_J - F_J)\|_{\dot{F}_q^{\alpha-1,p}} \le \frac{\delta}{2} \|f\|_{\dot{F}_q^{\alpha,p}} + D_\delta \|f\|^2_{\dot{F}_q^{\alpha,p}},$$

and since we assumed that $\|f\|_{\dot{F}_q^{\alpha,p}}$ is small (see (4.2)), then we may take $D_\delta \|f\|_{\dot{F}_q^{\alpha,p}} \le \delta/2$ obtaining

$$\sum_{i=1}^{\mathbf{k}} \|X_i(f_J - F_J)\|_{\dot{F}_q^{\alpha-1,p}} \le \delta \|f\|_{\dot{F}_q^{\alpha,p}}. \tag{4.13}$$

In a similar way, using Proposition 4.1 (ii) and Proposition 4.5 we get

$$\begin{aligned}\sum_{i=1}^{d_1}\|X_i(f_J-F_J)\|_{\dot F_q^{\alpha-1,p}} &\lesssim \sum_{i=1}^{d_1}\left\|X_i(h-\tilde h)\right\|_{\dot F_q^{\alpha-1,p}}+\sum_{i=1}^{d_1}\|X_i(g-\tilde g)\|_{\dot F_q^{\alpha-1,p}}\\ &\lesssim D'_{R,\sigma}\|f\|_{\dot F_q^{\alpha,p}}+D''_{R,\sigma}\|f\|^2_{\dot F_q^{\alpha,p}},\end{aligned}$$

and hence, as above,

$$\sum_{i=1}^{d_1}\|X_i(f_J-F_J)\|_{\dot F_q^{\alpha-1,p}}\lesssim_\delta \|f\|_{\dot F_q^{\alpha,p}}, \tag{4.14}$$

provided that $\|f\|_{\dot F_q^{\alpha,p}}$ is small enough. From (4.14) and the lifting property (Proposition 2.13) of the Triebel-Lizorkin norm, we get

$$\|F_J\|_{\dot F_q^{\alpha,p}}\le\|(f_J-F_J)\|_{\dot F_q^{\alpha,p}}+\|f_J\|_{\dot F_q^{\alpha,p}}\sim\sum_{i=1}^{d_1}\|X_i(f-F)\|_{\dot F_q^{\alpha-1,p}}+\|f\|_{\dot F_q^{\alpha,p}}\lesssim_\delta\|f\|_{\dot F_q^{\alpha,p}}. \tag{4.15}$$

Now (4.13) and (4.15) together with the L^∞ estimates (4.1) give Theorem 1.3 under the smallness assumption on $\|f\|_{\dot F_q^{\alpha,p}}$ (observing that the bounds proved do not depend on J and taking $J\to\infty$). We complete the proof of Theorem 1.3 via the homogeneity of the norms.

Remarks. (1) Following the same lines, it is also possible (and easier) to prove a version of Theorem 1.3 for the Besov spaces introduced in Subsection 2.3:

Theorem 4.9 *Consider the parameters $1<p<\infty$, $1<q\le\infty$, $\alpha=Q/p$ and let $\mathbf{k}$ be the largest positive integer with $\mathbf{k}<\min(p,d_1)$. Then, for every $\delta>0$ there exists a constant $C_\delta>0$ only depending on δ, such that for every function $f\in\dot B_q^{\alpha,p}(G)$ there exists $F\in L^\infty(G)\cap\dot B_q^{\alpha,p}(G)$ satisfying the following estimates:*

$$\begin{aligned}\sum_{i=1}^{\mathbf{k}}\|X_i(f-F)\|_{\dot B_q^{\alpha-1,p}(G)} &\le \delta\|f\|_{\dot B_q^{\alpha,p}(G)},\\ \|F\|_{L^\infty(G)}+\|F\|_{\dot B_q^{\alpha,p}(G)} &\le C_\delta\|f\|_{\dot B_q^{\alpha,p}(G)}.\end{aligned}$$

(2) To mention one application of Theorem 1.3, we state the following generalisation of Theorem 1.8 in [13] concerning the Hodge systems on the $(2n+1)$-dimensional Heisenberg group $\mathbf{H}^n$. Note that in this case $d=2n+1$, $d_1=2n$ and $Q=2n+2$.

Theorem 4.10 *Consider $1<p,q<\infty$, $\alpha:=(2n+2)/p$ and let r be an integer with $1\le r<\min(p,n)$. For any $(0,r)$-form φ in $\dot F_q^{\alpha,p}(\mathbf{H}^n)$, there exists a $(0,r)$-form Y in $L^\infty(\mathbf{H}^n)\cap\dot F_q^{\alpha,p}(\mathbf{H}^n)$ such that*

$$\overline\partial_b^* Y=\overline\partial_b^*\varphi$$

and

$$\|Y\|_{L^\infty(\mathbf{H}^n)}+\|Y\|_{\dot F_q^{\alpha,p}(\mathbf{H}^n)}\lesssim\left\|\overline\partial_b^*\varphi\right\|_{\dot F_q^{\alpha-1,p}(\mathbf{H}^n)}.$$

(See [13] for notation.)

This is proved by using Theorem 1.3 to approximate in an efficient way the coefficients of the form φ and then to conclude by using an iteration argument. Since the proof is very similar to the one given in [13] and its Euclidean analogue in [6] (Theorem 1.2), we omit it. Theorem 1.4 can be proved following the same lines.

5 Appendix

We collect here some facts related to the Calderón-Zygmund theory on stratified homogeneous groups for vector-valued functions. These results (Lemma A1 and Theorem A1) are well-known. However, since it is hard to find the exact statements in the literature (see for example [2] for a Euclidean version, or [7] for similar considerations on spaces of homogeneous type) hence, we have chosen to present them here.

Consider a Banach space A. In what follows we deal with functions from the space $L^p_A := L^p(G, A)$ where $1 \leq p \leq \infty$.

A first result is a Calderón-Zygmund decomposition of fuctions on G (see also Théorème 2.2 in [7], Chapitre 3), obtained via the weak $(1,1)$ estimate for the maximal operator:

Lemma A1. *Consider a function $f \in L^1_A$ and a number $\lambda > 0$. Then there exist a countable family of measurable sets $(\Omega_n)_{n\geq 1}$ which are pairwise disjoint and a decomposition $f = g + b = g + \sum_n b_n$ where $g, b, b_n \in L^1_A$ for all $n \geq 1$, and such that:*
(i) $\|g\|_{L^\infty_A} \lesssim \lambda$;
(ii) supp $b_n \subseteq \Omega_n$, $\int b_n(x)dx = 0$ and $\|b_n\|_{L^1_A} \lesssim \lambda|\Omega_n|$ for all n;
(iii) $\sum_n |\Omega_n| \lesssim \frac{1}{\lambda}\|f\|_{L^1_A}$.

Proof. We adapt the standard proof in the Euclidean case. Consider the open set $\tilde{\Omega} := \{x \in G |\ M\|f\|_A(x) > \lambda\}$. For each $x \in \tilde{\Omega}$ we consider a ball B_x centered in x and such that $B_x \subset \tilde{\Omega}$, but $2 \cdot B_x \subsetneq \tilde{\Omega}$ (here if $c > 0$ and B is a ball in G centered in x_B of radius R_B, then $c \cdot B$ is the ball in G of center x_B and of radius cR_B). Notice that, by Proposition 2.1,

$$|B_x| \leq \left|\tilde{\Omega}\right| \leq \frac{1}{\lambda}\|f\|_{L^1_A}$$

and hence, the balls B_x have uniformly bounded radii. Using the Vitali covering lemma (which has the same proof in G as in the Euclidean case), we can find a countable subfamilly of balls $(B_k)_{k\geq 1}$ of the familly $(B_x)_{x\in\tilde{\Omega}}$, which are pairwise disjoint and such that $\tilde{\Omega} = \bigcup_{x\in\tilde{\Omega}} B_x \subseteq \bigcup_{k\geq 1} C \cdot B_k$, where $C > 2$ is an absolute constant only depending on G.

We set

$$\Omega_1 := \left(\tilde{\Omega} \cap C \cdot B_1\right) \setminus \left(\bigcup_{j\neq 1} B_j\right)$$

and inductively we define

$$\Omega_k := \left(\left(\tilde{\Omega} \cap C \cdot B_k\right) \setminus \bigcup_{1\leq i\leq k-1} \Omega_i\right) \setminus \left(\bigcup_{j\neq k} B_j\right)$$

for all $k \geq 2$. We see immediately that for all $k \geq 1$ we have $B_k \subseteq \Omega_k \subseteq C \cdot B_k$ and this also give us that $|\Omega_k| \sim |B_k| \sim |C \cdot B_k|$. By definition the sets Ω_k are pairwise disjoint and $\tilde{\Omega} = \bigcup_{k\geq 1} \Omega_k$. We can define the functions:

$$g(x) := \begin{cases} f(x), \text{ if } x \notin \tilde{\Omega} \\ f_{\Omega_k}, \text{ if } x \in \Omega_k \end{cases}$$

and $b_k := (f - f_{\Omega_k})\, 1_{\Omega_k}$ for all $k \geq 1$. Here, $f_{\Omega_k} := |\Omega_k|^{-1} \int_{\Omega_k} f dx$.

To prove (i), we see that if $x \in \Omega_k$ we have

$$\|g(x)\|_A = \|f_{\Omega_k}\|_A \leq \frac{1}{|\Omega_k|}\int_{\Omega_k} \|f(y)\|_A\, dy \lesssim \frac{1}{|C \cdot B_k|}\int_{C\cdot B_k} \|f(y)\|_A\, dy \leq M\|f\|_A(x_0) \leq \lambda,$$

where x_0 is a point in $C \cdot B_k \backslash \tilde{\Omega}$. (Such a point exists since $2 \cdot B_k \subsetneq \tilde{\Omega}$ and $2 \cdot B_x \subset C \cdot B_k$.) If $x \notin \tilde{\Omega}$, by the Lebesgue differentiation theorem, which is a consequence of the weak estimate for the operator M, we can write $\|g(x)\|_A \leq M \|f\|_A (x) \leq \lambda$.

To prove (ii) and (iii), observe that by the above inequality we have

$$\left\| \|b_k\|_A \right\|_{L^1} \leq |\Omega_k| \left(\frac{1}{|\Omega_k|} \int_{\Omega_k} \|f(y)\|_A \, dy + \|f_{\Omega_k}\|_A \right) = 2|\Omega_k| \, \|f_{\Omega_k}\|_A \lesssim |\Omega_k| \lambda, \text{ for all } k,$$

and, using the weak estimate for M,

$$\sum_k |\Omega_k| \lesssim \sum_{k=1}^{\infty} |B_k| = \left| \bigcup_{k=1}^{\infty} B_k \right| \leq \left| \tilde{\Omega} \right| \lesssim \frac{1}{\lambda} \left\| \|f\|_A \right\|_{L^1}.$$

We can also see from these inequalities that

$$\sum_k \|b_k\|_{L^1_A} \lesssim \lambda \sum_k |\Omega_k| \lesssim \|f\|_{L^1_A}. \tag{5.1}$$

This proves in particular that the series defining b is absolutely convergent in L^1_A and $b, g \in L^1_A$ with $\|g\|_{L^1_A} + \|b\|_{L^1_A} \lesssim \|f\|_{L^1_A}$. □

Theorem A1. *Suppose A_1 and A_2 are two Banach spaces and $K \in L^1_{loc}(G \backslash \{0\} \to \mathfrak{L}(A_1, A_2))$ has the following properties:*

(i) there exists a constant $c > 0$ such that $\int_{\|x\|_G \geq c\|y\|_G} \|K(x) - K(y^{-1} \cdot x)\| \, dx \leq 1$ for all $y \in G$;

(ii) the operator $Tf = f * K$ is well-defined and bounded from $L^q_{A_1}$ to $L^q_{A_2}$ *for some $q \in (1, \infty)$.*

Then, $T : L^1_{A_1} \to L^{1,\infty}_{A_2}$ is well-defined and bounded. By real interpolation and duality we get that $T : L^p_{A_1} \to L^p_{A_2}$ is well-defined and bounded for any $p \in (1, \infty)$.

(Here $\mathfrak{L}(A_1, A_2)$ stands for the space of the bounded linear operators from A_1 to A_2.)

Proof. We adapt again the proof in the Euclidean case. Using Lemma A1 we can write, for a given $f \in L^1(A_1)$ and $\lambda > 0$, the decomposition at level λ: $f = g + b$. We next note that $\left|\left\{\|Tf\|_{A_2} > 2\lambda\right\}\right| \leq \left|\left\{\|Tg\|_{A_2} > \lambda\right\}\right| + \left|\left\{\|Tb\|_{A_2} > \lambda\right\}\right|$. The size of the set $\left\{\|Tg\|_{A_2} > \lambda\right\}$ can be bounded using (ii) above and the Markov inequality:

$$\begin{aligned} \left|\left\{\|Tg\|_{A_2} > \lambda\right\}\right| &\leq \lambda^{-q} \left\| \|Tg\|_{A_2} \right\|^q_{L^q} \lesssim \lambda^{-q} \left\| \|g\|_{A_1} \right\|^q_{L^q} = \lambda^{-q} \left\| \|g\|^q_{A_1} \right\|_{L^1} \\ &\leq \lambda^{-q} \lambda^{q-1} \left\| \|g\|_{A_1} \right\|_{L^1} = \lambda^{-1} \|g\|_{L^1_{A_1}} \lesssim \lambda^{-1} \|f\|_{L^1_{A_1}}. \end{aligned}$$

To estimate the size of the set $\left\{\|Tb\|_{A_2} > \lambda\right\}$ we proceed as follows. Consider the sets Ω_k from the proof of Lemma A1; for each such Ω_k we denote by y_{B_k} the center of the ball $B_k \subset \Omega_k$ and we set $\Omega^*_k := (C_1 + C) \cdot B_k \supset \Omega_k$ where $C_1 > 0$ is a large constant only depending on G and c. We write now

$$\begin{aligned} \left|\left\{\|Tb\|_{A_2} > \lambda\right\}\right| &\leq \left| \bigcup_k \Omega^*_k \right| + \left| \left\{ x \in G \backslash \bigcup_k \Omega^*_k \mid \|Tb\|_{A_2} > \lambda \right\} \right| \\ &\lesssim \lambda^{-1} \|f\|_{L^1_{A_1}} + \lambda^{-1} \int_{G \backslash \bigcup_k \Omega^*_k} \|Tb\|_{A_2} (x) dx, \end{aligned}$$

and it remains to estimate the last term. For this purpose, we note that if $x \in G \backslash \Omega^*_k$ and $y \in \Omega_k$, then $\rho(x, y_{B_k}) = \left\| y^{-1}_{B_k} \cdot x \right\|_G \geq (C_1 + C) R_{B_k} \geq C^{-1} (C_1 + C) \rho(y, y_{B_k}) \geq C^{-1} C_1 \left\| y^{-1}_{B_k} \cdot y \right\|_G$

(R_{B_k} is the radius of B_k) and thanks to the quasinorm property of $\|\cdot\|_G$, we find a constant $C_2 > 0$ depending on G only, such that $\|y^{-1}\cdot x\|_G = \left\|y^{-1}\cdot y_{B_k}\cdot y_{B_k}^{-1}\cdot x\right\|_G \geq C_2\left\|y_{B_k}^{-1}\cdot x\right\|_G - \|y^{-1}\cdot y_{B_k}\|_G \geq (C^{-1}C_1C_2-1)\left\|y_{\Omega_k}^{-1}\cdot y\right\|_G$, where we used the equality $a^{-1} = -a$ on G. If C_1 is sufficiently large, we deduce $\|y^{-1}\cdot x\|_G \geq c\,\|y^{-1}\cdot y_{B_k}\|_G = c\left\|(y^{-1}\cdot x)\left(y_{B_k}^{-1}\cdot x\right)^{-1}\right\|_G$. As a consequence,

$$\begin{aligned}
\int_{G\setminus\bigcup_k\Omega_k^*}\|Tb\|_{A_2}(x)dx &\leq \sum_n\int_{G\setminus\bigcup_k\Omega_k^*}\left\|\int_{\Omega_n}K(y^{-1}\cdot x)b_n(y)dy\right\|_{A_2}dx\\
&= \sum_n\int_{G\setminus\bigcup_k\Omega_k^*}\left\|\int_{\Omega_n}\left(K(y^{-1}\cdot x)-K(y_{B_n}^{-1}\cdot x)\right)b_n(y)dy\right\|_{A_2}dx\\
&\leq \sum_n\int_{\Omega_n}\left(\int_{G\setminus\bigcup_k\Omega_k^*}\left\|K(y^{-1}\cdot x)-K(y_{B_n}^{-1}\cdot x)\right\|_{op}dx\right)\|b_n(y)\|_{A_1}dy\\
&\leq \sum_n\int_{\Omega_n}\|b_n(y)\|_{A_1}dy \lesssim \|f\|_{L^1_{A_1}},
\end{aligned}$$

where we have used the condition (i) above and (5.1). □

Remark. We see from the proof that if $\|T\|_{L^q_{A_1}\to L^q_{A_2}} \leq 1$ then we have $\|T\|_{L^p_{A_1}\to L^p_{A_2}} \lesssim 1$. Hence if the quantity in (i) is bounded by a number $\beta > 0$ (instead of 1) and also $\|T\|_{L^q_{A_1}\to L^q_{A_2}} \leq \beta$, then we have $\|T\|_{L^p_{A_1}\to L^p_{A_2}} \lesssim \beta$.

Lemma A2. *Suppose $\varphi \in L^1(G)$ and:*
(i) $\int_{\|y\|_G\geq R}|\varphi(y)|\,dy \lesssim R^{-1}$ for any $R \geq 1$;
(ii) $\int_G|\varphi(x^{-1}\cdot y)-\varphi(y)|\,dy \lesssim \|x\|_G$ for all $x \in G$ with $\|x\|_G \leq 1$.

If for $r \in G$ we define $k_j(x) := \varphi_j(x\cdot 2^{-j}r)$, where $\varphi_j(x) = 2^{jQ}\varphi(2^jx)$ for all $j \in \mathbf{Z}$, then, for a constant $c > 0$ only depending on G, we have

$$\int_{\|y\|_G\geq c\|x\|_G}\sum_{j\in\mathbf{Z}}\left|k_j(x^{-1}\cdot y)-k_j(y)\right|dy \lesssim \ln(2+\|r\|_G).$$

Proof. We follow the proof in [6]. We decompose the sum under the integral as follows:

$$\sum_{j\in\mathbf{Z}}\left|k_j(x^{-1}\cdot y)-k_j(y)\right| = \sum_{2^j\|x\|_G\leq 1}\ldots + \sum_{1<2^j\|x\|_G<2+\|r\|_G}\ldots + \sum_{2^j\|x\|_G\geq 2+\|r\|_G}\ldots =: I+II+III.$$

We now estimate each term. Using (ii), we can estimate the first term as follows

$$\begin{aligned}
\int_{\|y\|_G\geq c\|x\|_G} I &\leq \int_G\sum_{2^j\|x\|_G\leq 1}2^{jQ}\left|\varphi(\left(2^jx^{-1}\right)\cdot\left(2^jy\right)\cdot r)-\varphi(\left(2^jy\right)\cdot r)\right|dy\\
&\leq \int_G\sum_{2^j\|x\|_G\leq 1}\left|\varphi(\left(2^jx^{-1}\right)\cdot y)-\varphi(y)\right|dy \lesssim \sum_{2^j\|x\|_G\leq 1}2^j\|x\|_G \lesssim 1.
\end{aligned}$$

For the second term we have:

$$\int_{\|y\|_G \geq c\|x\|_G} II \;\leq\; \sum_{1<2^j\|x\|_G<2+\|r\|_G} 2\int_G |k_j(y)|\,dy$$
$$= \; 2\sum_{1<2^j\|x\|_G<2+\|r\|_G} \int_G |\varphi(y)|\,dy \lesssim \ln\left(2+\|r\|_G\right).$$

To estimate the third term we use (i), which yelds

$$\int_{\|y\|_G \geq c\|x\|_G} III \;\leq\; 2\sum_{2^j\|x\|_G\geq 2+\|r\|_G} \int_{\|y\|_G\geq c_1\|x\|_G} |k_j(y)|\,dy$$
$$= \; 2\sum_{2^j\|x\|_G\geq 2+\|r\|_G} \int_{\|y\|_G\geq 2^j c_1\|x\|_G} |\varphi(y\cdot r)|\,dy$$
$$\lesssim \sum_{2^j\|x\|_G\geq 2+\|r\|_G} \frac{1}{2^j\|x\|_G} \lesssim \frac{1}{2+\|r\|_G} \lesssim 1$$

where the constant $c_1>0$ in the first line is obtained by using the quasinorm property of $\|\cdot\|_G$.

Summing up these estimates we get the claim. □

In what follows we will need to apply the above lemma to the function $\varphi = S$. It is easy to verify that the function $S(x) = \min(1, \|x\|_G^{-Q-1})$ satisfies the conditions (i) and (ii) required by Lemma A2. Indeed, by a change of variables, we can write for all $R \geq 1$,

$$\int_{\|y\|_G\geq R} |S(y)|\,dy = R^{-1}\int_{\|y\|_G\geq 1} \|y\|_G^{-Q-1}\,dy \sim R^{-1},$$

which proves that (i) is satisfied. To verify (ii), we recall that $|\|b\cdot a\|_G - \|a\|_G| \leq C\|b\|_G$ for all $a,b\in G$ (see Proposition 3.1) and note that if $\|y\|_G \leq 1 - C\|x\|_G \leq 1$, then $\|x^{-1}\cdot y\|_G \leq \|y\|_G + C\|x\|_G \leq 1$. Also, if $\|y\|_G \geq 1 + C\|x\|_G$, then $\|x^{-1}\cdot y\|_G \geq \|y\|_G - C\|x\|_G \geq 1$. Hence, if $\|x\|_G \leq 1$, we can write

$$\int_G \left|S(x^{-1}\cdot y) - S(y)\right| dy \;=\; \int_{1-C\|x\|_G\leq\|y\|_G\leq 1+C\|x\|_G} \left|S(x^{-1}\cdot y) - S(y)\right| dy$$
$$+\int_{\|y\|_G\geq 1+C\|x\|_G} \left|S(x^{-1}\cdot y) - S(y)\right| dy$$
$$\lesssim \; \|x\|_G + \int_{\|y\|_G\geq 1+C\|x\|_G} \left|\frac{1}{\|x^{-1}\cdot y\|_G^{Q+1}} - \frac{1}{\|y\|_G^{Q+1}}\right| dy$$
$$= \; \|x\|_G + \int_{\|y\|_G\geq 1+C\|x\|_G} \left|\frac{\|x^{-1}\cdot y\|_G^{Q+1} - \|y\|_G^{Q+1}}{\|x^{-1}\cdot y\|_G^{Q+1}\,\|y\|_G^{Q+1}}\right| dy$$
$$\lesssim \; \|x\|_G + \|x\|_G \int_{\|y\|_G\geq 1+C\|x\|_G} \frac{1}{\|y\|_G^{Q+2}} dy.$$

Proposition A1. *Suppose $p,q\in(1,\infty)$. Then, for every sequence $(f_j)_{j\in\mathbf{Z}}$ in $L^p(G,l^q(\mathbf{Z}))$ and for every $r\in G$ we have*

$$\left\| \left\| S_j f_j(x\cdot(2^{-j}r)) \right\|_{l_j^q} \right\|_{L_x^p} \lesssim_{p,q} \ln(2+\|r\|_G) \left\| \|f_j\|_{l_j^q} \right\|_{L^p}.$$

Proof. As we already saw, the function S satisfies the requirements of Lemma A2. Let k_j as in Lemma A2 with $\varphi = S$. We see directly that, for any Schwartz function f, we have

$$f * k_j(x) = \int_G f(y) S_j(y^{-1} \cdot x \cdot (2^{-j} r)) dy = S_j f(x \cdot (2^{-j} r)).$$

Hence $T(f_j)_{j\in\mathbf{Z}}(x) := (f_j)_{j\in\mathbf{Z}} * K(x) = (f_j * k_j)_{j\in\mathbf{Z}}(x) = (S_j f(x \cdot (2^{-j} r)))_{j\in\mathbf{Z}}$, the operator T being initially defined for a sequence of Schwartz functions $(f_j)_{j\in\mathbf{Z}}$. Considering the Banach spaces $A_1 = A_2 = l^q(\mathbf{Z})$ we can see that the statement of the Proposition A1 is equivalent to the fact that the operator $T : L^p_{A_1} \to L^p_{A_2}$ is continuous, with its norm bounded by $\ln(2 + \|r\|_G)$. This can be obtained as follows. Consider a sequence a in the unit sphere of $l^q(\mathbf{Z})$. We have that:

$$\begin{aligned} \left\langle K(x) - K(y^{-1} \cdot x), a \right\rangle &= \sum_{j\in\mathbf{Z}} \left(k_j(x) - k_j(y^{-1} \cdot x) \right) a_j \leq \left(\sum_{j\in\mathbf{Z}} \left| k_j(x) - k_j(y^{-1} \cdot x) \right|^{q'} \right)^{1/q'} \\ &\leq \sum_{j\in\mathbf{Z}} \left| k_j(x) - k_j(y^{-1} \cdot x) \right|, \end{aligned}$$

for all $x, y \in G$. Hence $\|K(x) - K(y^{-1} \cdot x)\| \leq \sum_{j\in\mathbf{Z}} |k_j(x) - k_j(y^{-1} \cdot x)|$ and thanks to Lemma A2 we get (using the same notation):

$$\int_{\|x\|_G \geq c\|y\|_G} \left\| K(x) - K(y^{-1} \cdot x) \right\| dx \leq \int_{\|x\|_G \geq c\|y\|_G} \sum_{j\in\mathbf{Z}} \left| k_j(x) - k_j(y^{-1} \cdot x) \right| dx \lesssim \ln(2 + \|r\|_G).$$

Also we can easily see that $T : L^q_{A_1} \to L^q_{A_2}$ is bounded and of norm 1. These two last observations together with Theorem A1 and the Remark after, give us the claim. □

Remark. Proposition A1 is reminiscent of an inequality due to Bourgain (see for example [11]).

Acknowledgements

The author thanks Petru Mironescu and Emmanuel Russ for useful discussions and suggestions. He was supported by the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program "Investissements d'Avenir" (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR).

Part of this work was completed while the author was visiting the Institute of Mathematics of the Romanian Academy. He thanks the Institute for its hospitality.

References

[1] Bahouri, H., Gallagher, I., *Paraproduit sur le groupe de Heisenberg et applications.* Rev. Mat. IberoAmericana 17, 69-105, 2001.

[2] Benedek, A., Calderón, A. P., Panzone, R., *Convolution operators on Banach space valued functions.* Proc. Nat. Acad. Sci., 48(3):356–365, 1962.

[3] Bourgain, J., Brezis, H., *On the equation div $Y = f$ and application to control of phases.* J. Amer. Math. Soc., 16(2):393–426, 2003.

[4] Bourgain, J., Brezis, H., *New estimates for eliptic equations and Hodge type systems.* J. Eur. Math. Soc., 9, no. 2, 277-315, 2007.

[5] Bousquet, P., Mironescu, P., Russ, E., *A limiting case for the divergence equation.* Math. Z., 274(1-2):427–460, 2013.

[6] Bousquet, P., Russ, E., Wang, Y., Yung, P-L. *Approximation in fractional Sobolev spaces and Hodge systems.* J. Funct. Anal. 276, no. 5, 1430-1478, 2019.

[7] Coifman, R., Weiss, G., *Analyse harmonique non commutative sur certains espaces homogènes*, Lecture notes, 242, Springer Verlag., 1971.

[8] Folland, G. B., Stein, E. M., *Hardy Spaces on Homogeneous Groups*, vol. 28 of Mathematical Notes, Princeton University Press, Princeton, NJ, USA, 1982.

[9] Furioli, G., Melzi, C., Veneruso, A., *Littlewood-Paley decompositions and Besov spaces on Lie groups of polynomial growth.* Math. Nachr., 279(9-10):1028–1040, 2006.

[10] Hu, G. *Homogeneous Triebel-Lizorkin Spaces on Stratified Lie Groups.* Journal of Function Spaces and Applications, ID 475103, 2013.

[11] Hytönen, T. P., *Foundations of vector-valued singular integrals revisited-with random dyadic cubes.* Bulletin of the Polish Academy of Sciences Mathematics 60(3), 2011.

[12] Stein, E.M., *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, 695 pp. Princeton Mathematical Series 43, Princeton University Press (1993).

[13] Wang, Y., Yung, P.-L., *A subelliptic Bourgain-Brezis inequality.* J. Eur. Math. Soc. (JEMS), 16(4):649-693, 2014.