J. Eur. Math. Soc. 16, 649–693

DOI 10.4171/JEMS/443

Yi Wang · Po-Lam Yung

A subelliptic Bourgain–Brezis inequality

Received December 22, 2011 and in revised form August 20, 2012

Abstract. We prove an approximation lemma on (stratified) homogeneous groups that allows one to approximate a function in the non-isotropic Sobolev space $\dot{NL}^{1,Q}$ by L^∞ functions, generalizing a result of Bourgain–Brezis [BB2]. We then use this to obtain a Gagliardo-Nirenberg inequality for $\overline{\partial}_b$ on the Heisenberg group $\mathbb{H}^n$.

Keywords. div-curl, compensation phenomena, critical Sobolev embedding, homogeneous groups

1. Introduction

In this paper, we study some subelliptic compensation phenomena on homogeneous groups, that have to do with divergence, curl and the space L^1 of Lebesgue integrable functions or differential forms. In the elliptic cases they were discovered by Bourgain–Brezis, Lanzani–Stein and van Schaftingen around 2004. Also lying beneath our results is the failure of the critical Sobolev embedding of the non-isotropic Sobolev space $\dot{NL}^{1,Q}$ into L^∞. In particular, we prove an approximation lemma that describes how functions in $\dot{NL}^{1,Q}$ can be approximated by functions in L^∞.

To begin with, let us describe the elliptic results on $\mathbb{R}^n$ ($n \geq 2$) upon which our results are based. We denote by d the Hodge–de Rham exterior derivative, and by d^* its (formal) adjoint. The theory discovered by Bourgain–Brezis, Lanzani–Stein and van Schaftingen has three major pillars, each best illustrated by a separate theorem. The first involves the solution of d^*:

Theorem 1.1 (Bourgain–Brezis [BB2]). *Suppose $q \neq n-1$. Then for any q-form f with coefficients on $L^n(\mathbb{R}^n)$ that is in the image[1] of d^*, there exists a $(q+1)$-form Y with coefficients in $L^\infty(\mathbb{R}^n)$ such that*

$$d^*Y = f$$

in the sense of distributions, and $\|Y\|_{L^\infty(\mathbb{R}^n)} \leq C\|f\|_{L^n(\mathbb{R}^n)}$.

Y. Wang: Department of Mathematics, Stanford University, Stanford, CA 94305, USA;
e-mail: wangyi@math.stanford.edu

P.-L. Yung: Department of Mathematics, Rutgers, the State University of New Jersey,
Piscataway, NJ 08854, USA; e-mail: pyung@math.rutgers.edu

Mathematics Subject Classification (2010): Primary 35H20; Secondary 42B35

[1] By this we mean f is the d^* of some form with coefficients in $\dot{W}^{1,n}(\mathbb{R}^n)$, where $\dot{W}^{1,n}(\mathbb{R}^n)$ is the (homogeneous) Sobolev space of functions that have one derivative in $L^n(\mathbb{R}^n)$.

In particular, we have

Corollary 1.2 (Bourgain–Brezis [BB1]). *For any function $f \in L^n(\mathbb{R}^n)$, there exists a vector field Y with coefficients in $L^\infty(\mathbb{R}^n)$ such that*

$$\operatorname{div} Y = f$$

in the sense of distributions, and $\|Y\|_{L^\infty(\mathbb{R}^n)} \leq C\|f\|_{L^n(\mathbb{R}^n)}$.

The second pillar is a Gagliardo–Nirenberg inequality for differential forms:

Theorem 1.3 (Lanzani–Stein [LS]). *Suppose u is a q-form on $\mathbb{R}^n$ that is smooth with compact support. We have*

$$\|u\|_{L^{n/(n-1)}(\mathbb{R}^n)} \leq C(\|du\|_{L^1(\mathbb{R}^n)} + \|d^*u\|_{L^1(\mathbb{R}^n)})$$

unless d^*u is a function or du is a top form. If d^*u is a function, one needs to assume $d^*u = 0$; if du is a top form, one needs to assume $du = 0$. Then the above inequality remains true.

Since d of a 1-form is its curl and d^* of a 1-form is its divergence, this is sometimes called the div-curl inequality.

The third theorem is the following compensation phenomenon:

Theorem 1.4 (van Schaftingen [vS1]). *If u is a C_c^∞ 1-form on $\mathbb{R}^n$ with $d^*u = 0$, then for any 1-form ϕ with coefficients in $C_c^\infty(\mathbb{R}^n)$, we have*

$$\int_{\mathbb{R}^n} u \cdot \phi \, dx \leq C\|u\|_{L^1(\mathbb{R}^n)}\|\phi\|_{\dot{W}^{1,n}(\mathbb{R}^n)}.$$

If $\dot{W}^{1,n}(\mathbb{R}^n)$ were embedded into $L^\infty(\mathbb{R}^n)$, Theorem 1.1 would be trivial by Hodge decomposition, and so would be Theorem 1.4 by Hölder's inequality. It is remarkable that these theorems remain to hold even though the desired Sobolev embedding fails.

It turns out that all three theorems above are equivalent by duality. Van Schaftingen [vS1] gave a beautiful elementary proof of Theorem 1.4, thereby proving all of them.

We mention here that these results seem to be quite different from the more classical theory of compensated compactness; no connection between them is known so far.

We also refer the reader to the work of Brezis–van Schaftingen [BvS], Chanillo–van Schaftingen [CvS], Maz'ya [Ma], Mironescu [Mi], Mitrea–Mitrea [MM], van Schaftingen [vS2], [vS3] and Amrouche–Nguyen [AN] for some interesting results related to these three theorems. In particular, Chanillo–van Schaftingen proved in [CvS] a generalization of Theorem 1.4 to general homogeneous groups.

On the other hand, in [BB2], Bourgain–Brezis proved the following remarkable theorem, strengthening all three theorems above:

Theorem 1.5 (Bourgain–Brezis [BB2]). *In Theorem 1.1 and Corollary 1.2, the space $L^\infty(\mathbb{R}^n)$ can be replaced by the smaller Banach space $L^\infty(\mathbb{R}^n) \cap \dot{W}^{1,n}(\mathbb{R}^n)$. Also, in Theorems 1.3 and 1.4, the spaces $L^1(\mathbb{R}^n)$ can be replaced by the bigger Banach space $L^1(\mathbb{R}^n) + (\dot{W}^{1,n}(\mathbb{R}^n))^*$. (Here X^* denotes the dual of a Banach space X.)*

They proved this by giving a direct constructive proof of the analog of Theorem 1.1, where the space $L^\infty(\mathbb{R}^n)$ is replaced by $L^\infty(\mathbb{R}^n) \cap \dot{W}^{1,n}(\mathbb{R}^n)$; they then deduced the rest by duality. In the former argument they used the following approximation lemma, which is another remedy for the failure of the critical Sobolev embedding, and which is of independent interest:

Lemma 1.6 (Bourgain–Brezis [BB2]). *Given any $\delta > 0$ and any function $f \in \dot{W}^{1,n}(\mathbb{R}^n)$, there exist a function $F \in L^\infty(\mathbb{R}^n) \cap \dot{W}^{1,n}(\mathbb{R}^n)$ and a constant $C_\delta > 0$, with C_δ independent of f, such that*

$$\sum_{i=2}^{n} \|\partial_i f - \partial_i F\|_{L^n(\mathbb{R}^n)} \leq \delta \|\nabla f\|_{L^n(\mathbb{R}^n)}$$

and

$$\|F\|_{L^\infty(\mathbb{R}^n)} + \|\nabla F\|_{L^n(\mathbb{R}^n)} \leq C_\delta \|\nabla f\|_{L^n(\mathbb{R}^n)}.$$

Here one should think of F as an $L^\infty(\mathbb{R}^n) \cap \dot{W}^{1,n}(\mathbb{R}^n)$ function whose derivatives approximate those of the given f in all but one direction.

In this paper, we prove an analog of the above approximation lemma on any homogeneous group G. To describe our result we need some notations. First, let $\mathfrak{g}$ be a Lie algebra (over $\mathbb{R}$) that is *graded*, in the sense that $\mathfrak{g}$ admits a decomposition

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_m$$

into direct sums of subspaces such that

$$[V_{j_1}, V_{j_2}] \subseteq V_{j_1+j_2}$$

for all j_1, j_2, where V_j is understood to be zero if $j > m$. We assume that $V_m \neq \{0\}$. It is immediate that $\mathfrak{g}$ is nilpotent of step m. We introduce a natural family of dilations on $\mathfrak{g}$, by letting

$$\lambda \cdot v = \lambda v_1 + \lambda^2 v_2 + \cdots + \lambda^m v_m$$

if $v = v_1 + \cdots + v_m$, $v_i \in V_i$ and $\lambda > 0$. This defines a one-parameter family of algebra automorphisms of $\mathfrak{g}$. Furthermore, we assume that $\mathfrak{g}$ is *stratified*, in the sense that V_1 generates $\mathfrak{g}$ as a Lie algebra. Let G be the connected and simply connected Lie group whose Lie algebra is $\mathfrak{g}$. Such a Lie group G with stratified $\mathfrak{g}$ is then called a *homogeneous group*. It carries a one-parameter family of automorphic dilations, given by $\lambda \cdot \exp(v) := \exp(\lambda \cdot v)$ where $\exp\colon \mathfrak{g} \to G$ is the exponential map. In the following we fix such a group G.

Now define the *homogeneous dimension* Q of G by

$$Q := \sum_{j=1}^{m} j \cdot n_j$$

where $n_j := \dim V_j$. We also pick a basis $X_1, \ldots, X_{n_1}$ of V_1. Any linear combination of these will then be a left-invariant vector field of degree 1 on G. If f is a function on G, we define its *subelliptic gradient* as the n_1-tuple

$$\nabla_b f := (X_1 f, \ldots, X_{n_1} f).$$

The *homogeneous non-isotropic Sobolev space* $\dot{NL}^{1,Q}(G)$ is then the space of functions on G whose subelliptic gradient is in $L^Q(G)$. Here in defining the $L^Q(G)$ space, we use the Lebesgue measure on $\mathfrak{g}$, which we identify with G via the exponential map. In the following, we will denote the function spaces on G by $\dot{NL}^{1,Q}$, L^Q, L^∞ etc. for simplicity unless otherwise specified.

It is well-known that $\dot{NL}^{1,Q}$ fails to embed into L^∞. Nonetheless, we prove the following approximation lemma for functions in $\dot{NL}^{1,Q}$:

Lemma 1.7. *Given any $\delta > 0$ and any function f on G with $\|\nabla_b f\|_{L^Q} < \infty$, there exist a function $F \in L^\infty$ with $\nabla_b F \in L^Q$, and a constant $C_\delta > 0$ independent of f, such that*

$$\sum_{k=2}^{n_1} \|X_k f - X_k F\|_{L^Q} \leq \delta \|\nabla_b f\|_{L^Q},$$

$$\|F\|_{L^\infty} + \|\nabla_b F\|_{L^Q} \leq C_\delta \|\nabla_b f\|_{L^Q}.$$

Specializing this result to the Heisenberg group $\mathbb{H}^n$, we deduce, for instance, the following result about the solution of $\overline{\partial}_b$:

Theorem 1.8. *Suppose $Q = 2n+2$ and $q \neq n-1$. Then for any $(0, q)$-form f on $\mathbb{H}^n$ that has coefficients in L^Q and that is the $\overline{\partial}_b^*$ of some other form with coefficients in $\dot{NL}^{1,Q}$, there exists a $(0, q+1)$-form Y on $\mathbb{H}^n$ with coefficients in $L^\infty \cap \dot{NL}^{1,Q}$ such that*

$$\overline{\partial}_b^* Y = f$$

in the sense of distributions, with $\|Y\|_{L^\infty} + \|\nabla_b Y\|_{L^Q} \leq C\|f\|_{L^Q}$.

We then have a Gagliardo–Nirenberg inequality for $\overline{\partial}_b$ on $\mathbb{H}^n$:

Theorem 1.9. *Suppose $Q = 2n + 2$. If u is a $(0, q)$ form on $\mathbb{H}^n$ with $2 \leq q \leq n-2$, then*

$$\|u\|_{L^{Q/(Q-1)}} \leq C(\|\overline{\partial}_b u\|_{L^1+(\dot{NL}^{1,Q})^*} + \|\overline{\partial}_b^* u\|_{L^1+(\dot{NL}^{1,Q})^*}). \tag{1.1}$$

Also, if $n \geq 2$ and u is a function on $\mathbb{H}^n$ that is orthogonal to the kernel of $\overline{\partial}_b$, then

$$\|u\|_{L^{Q/(Q-1)}} \leq C\|\overline{\partial}_b u\|_{L^1+(\dot{NL}^{1,Q})^*}. \tag{1.2}$$

There is also a version of this result for $(0, 1)$ forms and $(0, n-1)$ forms, analogous to the last part of Theorem 1.3.

A weaker version of this theorem, with $L^1 + (\dot{NL}^{1,Q})^*$ above replaced by L^1, can also be deduced easily from the work of Chanillo–van Schaftingen [CvS] (cf. also [Y]).

Several difficulties need to be overcome when we prove Lemma 1.7 on a general homogeneous group. The first is that we no longer have a Fejér kernel as in the Euclidean spaces, which served as the building block of a good reproducing kernel K_j in the original proof of Bourgain–Brezis. As a result, we need to find an appropriate variant of that. What we do is take the heat kernels S_j, and use S_{j+N}, where N is large, as our approximate reproducing kernel. In other words, we use $S_{j+N}\Delta_j f$, where N is large, to approximate $\Delta_j f$, where $\Delta_j f$ is a Littlewood–Paley piece of the function f. Since the heat kernel does not localize perfectly in "frequency", we need, in the preparational stage, some extra efforts to deal with additional errors that come up in that connection.

Our second difficulty, which is also the biggest challenge, is that our homogeneous group is in general not abelian. Hence we must carefully distinguish between left- and right-invariant derivatives when we differentiate a convolution (which is defined in (2.3)): $X_k(f * K)$ is equal to $f * (X_k K)$, and not to $(X_k f) * K$, if X_k is left-invariant and K is any kernel (cf. Proposition 4.1 in Section 4). To get around that, several ingredients are involved. One of them is to explore the relationship between left- and right-invariant vector fields, which we recall in Section 4. Another is to introduce two different auxiliary controlling functions ω_j and $\tilde{\omega}_j$. These are functions that dominate $|\Delta_j f|$ pointwise (at least morally), and both $X_k \omega_j$ and $X_k \tilde{\omega}_j$, for $k = 2, \dots, n_1$, will be better controlled than $X_1 \omega_j$ and $X_1 \tilde{\omega}_j$. The key point here, on the other hand, is that $\tilde{\omega}_j$ is frequency localized, and it dominates ω_j. Also, ω_j satisfies

$$\Big\| \sup_j 2^j \omega_j \Big\|_{L^Q} \lesssim 2^{\sigma(Q-1)/Q} \|\nabla_b f\|_{L^Q}.$$

On the contrary, $\tilde{\omega}_j$ will not satisfy the analog of this inequality, and ω_j will not be frequency localized. In defining such ω_j and $\tilde{\omega}_j$, instead of taking an "L^∞ convolution" as in the definition of ω_j used by Bourgain–Brezis, we will take a discrete convolution in l^Q and an honest convolution for ω_j and $\tilde{\omega}_j$ respectively. (The precise definition of ω_j and $\tilde{\omega}_j$ can be found in Section 7.) We then use ω_j to control the part of f where the high frequencies are dominating, and use $\tilde{\omega}_j$ to control the other part where the low frequencies are dominating.

Finally, we will need two slightly different versions of Littlewood–Paley theories on a homogeneous group. One is chosen such that $f = \sum_j \Delta_j f$, and the other is such that the reverse Littlewood–Paley inequality holds (as in Proposition 5.2).

We will now proceed as follows. In Section 2–5 we describe some preliminaries about homogeneous groups. This includes some mean-value type inequalities on G, some tools that allow us to mediate between left- and right-invariant derivatives, as well as a refinement of Littlewood–Paley theory on G. In Section 6 we give some algebraic preliminaries needed in the proof of Lemma 1.7, in Section 7 we give an outline of that proof, and Sections 8–11 contains the details. Finally in Section 12 we prove Theorems 1.8 and 1.9.

2. Preliminaries

Let G be a homogeneous group, $n_j := \dim V_j$, and $X_1, \dots, X_{n_1}$ be a basis of V_1 as above. We introduce a coordinate system on G. First write

$$n := n_1 + \cdots + n_m,$$

and extend $X_1, \dots, X_{n_1}$ to a basis $X_1, \dots, X_n$ of $\mathfrak{g}$ such that $X_{n_{j-1}+1}, \dots, X_{n_j}$ is a basis of V_j for all $1 \le j \le m$ (with n_0 understood to be 0). Then for $x = [x_1, \dots, x_n] \in \mathbb{R}^n$, we identify x with $\sum_{i=1}^n x_i X_i \in \mathfrak{g}$. We will also identify $\mathfrak{g}$ with G via the exponential map.

Thus we write x for the point $\exp(\sum_{i=1}^{n} x_i X_i) \in G$. This defines a coordinate system on G. The group identity of G is $0 = [0, \dots, 0]$, and the dilation on G is given explicitly by

$$\lambda \cdot x = [\lambda x_1, \dots, \lambda x_{n_1}, \lambda^2 x_{n_1+1}, \dots, \lambda^2 x_{n_2}, \dots, \lambda^m x_{n_{m-1}+1}, \dots, \lambda^m x_n]$$

for $\lambda > 0$ and $x = [x_1, \dots, x_n]$.

For $x, y \in G$, we write $x \cdot y$ for their group product in G. By the Campbell–Hausdorff formula, this group law is given by a polynomial map when viewed as a map from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}^n$. More precisely, the map $(x, y) \mapsto x \cdot y$ can be computed by

$$\begin{aligned} x \cdot y &= \exp\Big(\sum_{i=1}^{n} x_i X_i\Big) \cdot \exp\Big(\sum_{i=1}^{n} y_i X_i\Big) \\ &= \exp\Big(\sum_{i=1}^{n} x_i X_i + \sum_{i=1}^{n} y_i X_i + \frac{1}{2}\Big[\sum_{i=1}^{n} x_i X_i, \sum_{i=1}^{n} y_i X_i\Big] + \dots\Big). \end{aligned} \tag{2.1}$$

It follows that for $1 \le k \le n_1$, the k-th coordinate of $x \cdot y$ is $x_k + y_k$.

The dilations on G are automorphisms of the group: in particular,

$$\lambda \cdot (x \cdot y) = (\lambda \cdot x) \cdot (\lambda \cdot y) \tag{2.2}$$

for all $\lambda > 0$ and all $x, y \in G$.

A function $f(x)$ on G is said to be *homogeneous* of degree l if $f(\lambda \cdot x) = \lambda^l f(x)$ for all $x \in G$ and $\lambda > 0$. From (2.1) we see that for all $n_j < k \le n_{j+1}$, the k-th coordinate of $x \cdot y$ is equal to $x_k + y_k + P_k(x, y)$ where P_k is a homogeneous polynomial of degree j on $G \times G$. On G one can define the homogeneous norm

$$\|x\| = \Big(\sum_{j=1}^{m} \sum_{n_{j-1} < k \le n_j} |x_k|^{2m!/j}\Big)^{1/2m!}.$$

It is a homogeneous function of degree 1 on G, and satisfies a quasi-triangle inequality

$$\|x \cdot y\| \le C(\|x\| + \|y\|)$$

for all $x, y \in G$, where C is a constant depending only on G. We also have $\|x\| = \|x^{-1}\|$ for all $x \in G$, since if $x = [x_1, \dots, x_n]$ then $x^{-1} = [-x_1, \dots, -x_n]$.

Any element X of $\mathfrak{g}$ can be identified with a left-invariant vector field on G. It will be said to be homogeneous of degree l if $X(f(\lambda \cdot x)) = \lambda^l (Xf)(\lambda \cdot x)$ for all C^1 functions f. Then $X_1, \dots, X_{n_1}$ is a basis of left-invariant vector fields of degree 1 on G. We remind the reader that we write $\nabla_b f = (X_1 f, \dots, X_{n_1} f)$, and call this the subelliptic gradient of f.

By the form of the group law on G, one can see that if $n_{j-1} < k \le n_j$, then X_k can be written as

$$X_k = \frac{\partial}{\partial x_k} + \sum_{p=j+1}^{m} \sum_{n_{p-1} < k' \le n_p} P_{k,k'}(x) \frac{\partial}{\partial x_{k'}}$$

where $P_{k,k'}(x)$ is a homogeneous polynomial of degree $p - j$ if $n_{p-1} < k' \le n_p$.

If X is a left-invariant vector field on G, we write X^R for the right-invariant vector field on G that agrees with X at the identity (namely 0). We also write $\nabla_b^R f$ for the n_1-tuple $(X_1^R f, \dots, X_{n_1}^R f)$.

The Lebesgue measure dx on $\mathbb{R}^n$ is a Haar measure on G if we identify $x \in \mathbb{R}^n$ with a point in G as we have always done. It satisfies $d(\lambda \cdot x) = \lambda^Q dx$ for all positive λ, where $Q = \sum_{j=1}^m j \cdot n_j$ is the homogeneous dimension we introduced previously.

With the Haar measure we define the L^p spaces on G. If f and g are two L^1 functions on G, then their convolution is given by

$$f * g(x) = \int_G f(x \cdot y^{-1}) g(y)\, dy, \tag{2.3}$$

or equivalently

$$f * g(x) = \int_G f(y) g(y^{-1} \cdot x)\, dy.$$

The non-isotropic Sobolev space $N\dot{L}^{1,Q}$ is the space of functions f such that $\|\nabla_b f\|_{L^Q} < \infty$.

3. Some basic inequalities

To proceed further, we collect some basic inequalities that will be useful on a number of occasions.

In this and the next section, f and g will denote two general C^1 functions on G. The first proposition is a mean-value inequality.

Proposition 3.1. *There exist absolute constants $C > 0$ and $a > 0$ such that*

$$|f(x \cdot y^{-1}) - f(x)| \le C\|y\| \sup_{\|z\| \le a\|y\|} |(\nabla_b f)(x \cdot z^{-1})|$$

for all $x, y \in G$.

For a proof of this proposition, see Folland–Stein [FS, p. 33, (1.41)].

There is also a mean-value inequality for right translations, whose proof is similar and we omit it:

Proposition 3.2. *There exist absolute constants $C > 0$ and $a > 0$ such that*

$$|f(y^{-1} \cdot x) - f(x)| \le C\|y\| \sup_{\|z\| \le a\|y\|} |(\nabla_b^R f)(z^{-1} \cdot x)| \quad \text{for all } x, y \in G.$$

Next we have some integral estimates:

Proposition 3.3. *If*

$$|\nabla_b f(x)| \le (1 + \|x\|)^{-M} \quad \text{and} \quad |g(x)| \le (1 + \|x\|)^{-M} \tag{3.1}$$

for all $M > 0$, then for any non-negative integer k,

$$\int_G |f(x \cdot y^{-1}) - f(x)|\, |g_k(y)|\, dy \le C 2^{-k} (1 + \|x\|)^{-M} \tag{3.2}$$

for all M, where $g_k(y) := 2^{kQ} g(2^k \cdot y)$.

The key here, as well as in the next two propositions, is that we get a small factor 2^{-k} on the right hand side of our estimates.

Proof. We split the integral into two parts:

$$\int_G |f(x\cdot y^{-1}) - f(x)|\,|g_k(y)|\,dy = \int_{\|y\|\le c\|x\|} dy + \int_{\|y\|>c\|x\|} dy = I + II,$$

where c is a constant chosen such that if $\|y\| \le c\|x\|$ and $\|z\| \le a\|y\|$ then $\|x\cdot z^{-1}\| \ge \frac{1}{2}\|x\|$. Here a is the constant appearing in the statement of Proposition 3.1, and such a c exists by (3.3) below. We then apply Proposition 3.1 twice. First in I, the integrand can be bounded by

$$|f(x\cdot y^{-1}) - f(x)|\,|g_k(y)| \le C(1+\|x\|)^{-M}\|y\|\,|g_k(y)|$$

for all M, and

$$\int_G \|y\|\,|g_k(y)|\,dy = C2^{-k}.$$

Also, in II, the integrand can be bounded by

$$|f(x\cdot y^{-1}) - f(x)|\,|g_k(y)| \le C\|y\|\,|g_k(y)|,$$

and

$$\begin{aligned}\int_{\|y\|\ge c\|x\|} \|y\|\,|g_k(y)|\,dy &\le \int_{\|y\|\ge c\|x\|} \|y\|(1+2^k\|y\|)^{-Q-M-1}2^{kQ}\,dy \\ &\le C2^{-k}(1+\|x\|)^{-M}\end{aligned}$$

for all M. Combining the estimates concludes the proof. □

We remark here that if we want (3.2) to hold for a specific M, then we only need condition (3.1) to hold with M replaced by $Q+M+1$.

In particular, we have

Proposition 3.4. *If f, g are as in Proposition 3.3, and in addition $\int_G g(y)\,dy = 0$, then for any non-negative integer k, we have*

$$|f * g_k(x)| \le C2^{-k}(1+\|x\|)^{-M} \quad \textit{for all } M.$$

Proof. One can write

$$f * g_k(x) = \int_G (f(x\cdot y^{-1}) - f(x))g_k(y)\,dy$$

since $\int_G g(y)\,dy = 0$ implies $\int_G g_k(y)\,dy = 0$. Then taking absolute values and using Proposition 3.3, one yields the desired claim. □

Similarly, suppose $f_k(x) := 2^{kQ} f(2^k\cdot x)$. Using the representation

$$f_k * g(x) = \int_G f_k(y)g(y^{-1}\cdot x)\,dy,$$

and invoking Proposition 3.2 instead of Proposition 3.1, we can estimate $f_k * g$ as well.

Proposition 3.5. *Suppose*

$$|\nabla_b^R g(x)| \le (1 + \|x\|)^{-M} \quad \text{and} \quad |f(x)| \le (1 + \|x\|)^{-M} \quad \text{for all } M > 0.$$

Suppose further that $\int_G f(y)\,dy = 0$. *Then for any non-negative integer k, we have*

$$|f_k * g(x)| \le C2^{-k}(1 + \|x\|)^{-M} \quad \text{for all } M.$$

Finally, let σ be a non-negative integer, and adopt the shorthand $x_\sigma := 2^{-\sigma} \cdot x^\sigma$, where $x^\sigma := [2^\sigma x_1, x_2, \dots, x_n]$ if $x = [x_1, \dots, x_n]$. We will need the following mean-value type inequality for $\|x_\sigma\|$.

Proposition 3.6. *For any* $x, \theta \in G$,

$$|\, \|(x \cdot \theta)_\sigma\| - \|x_\sigma\| \,| \le C\|\theta\|.$$

Here the constant C is independent of σ. *Similarly* $|\, \|(\theta \cdot x)_\sigma\| - \|x_\sigma\| \,| \le C\|\theta\|$.

In particular, taking $\sigma = 0$, the norm function satisfies

$$|\, \|x \cdot \theta\| - \|x\| \,| \le C\|\theta\| \tag{3.3}$$

for all $x, \theta \in G$.

Proof of Proposition 3.6. We prove the desired inequalities using scale invariance. The key fact is that the function $x \mapsto \|x_\sigma\|$ is homogeneous of degree 1 and smooth away from 0; in fact, $(\lambda \cdot x)_\sigma = \lambda \cdot x_\sigma$, and the homogeneity of the above map follows:

$$\|(\lambda \cdot x)_\sigma\| = \|\lambda \cdot x_\sigma\| = \lambda \|x_\sigma\|.$$

By scaling x and θ simultaneously, without loss of generality, we may assume that $\|x\| = 2$. Now to prove the first inequality, we consider two cases: $\|\theta\| \le 1$ and $\|\theta\| \ge 1$. If $\|\theta\| \le 1$, then the (Euclidean) straight line joining x and $x \cdot \theta$ stays in a compact set not containing 0. Then we apply the Euclidean mean-value inequality to the function $x \mapsto \|x_\sigma\|$, which is smooth in this compact set and satisfies $|\nabla \|x_\sigma\|\,| \lesssim 1$ there uniformly in σ. It follows that if $\|\theta\| \le 1$, we have

$$|\, \|(x \cdot \theta)_\sigma\| - \|x_\sigma\| \,| \lesssim |\theta| \le C\|\theta\|.$$

Here $|\theta|$ is the Euclidean norm of θ. On the other hand, if $\|\theta\| \ge 1$, then

$$|\, \|(x \cdot \theta)_\sigma\| - \|x_\sigma\| \,| \le \|(x \cdot \theta)_\sigma\| + \|x_\sigma\| \le \|x \cdot \theta\| + \|x\| \lesssim C(\|x\| + \|\theta\|) \lesssim \|\theta\|,$$

where the second to last inequality follows from the quasi-triangle inequality. Thus we have the desired inequality either case. One can prove the second inequality similarly. □

4. Left- and right-invariant derivatives

Next we describe how one mediates between left- and right-invariant derivatives when working with convolutions on G. First we have the following basic identities.

Proposition 4.1.

$$X_k(f*g) = f*(X_k g), \quad (X_k f)*g = f*(X_k^R g), \quad \textit{and} \quad X_k^R(f*g) = (X_k^R f)*g,$$

assuming f, g and their derivatives decay sufficiently rapidly at infinity.

A proof can be found in Folland–Stein [FS, p. 22]. Since our groups are not abelian in general, one has to be careful with these identities; one does not have, for instance, the identity between $X_k(f*g)$ and $(X_k f)*g$.

We also have the following flexibility of representing coordinate and left-invariant derivatives in terms of right-invariant ones.

Proposition 4.2. (a) *For $1 \le i \le n$, the coordinate derivative $\partial/\partial x_i$ can be written as*

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^{n_1} X_k^R D_{i,k}$$

where $D_{i,k}$ are homogeneous differential operators of degree $j-1$ if $n_{j-1} < i \le n_j$.

(b) *In fact any $\partial/\partial x_i$, $1 \le i \le n$, can be written as a linear combination of $(\nabla_b^R)^\alpha$ with coefficients that are polynomials in x, where α ranges over a finite subset of the indices $\{1,\dots,n_1\}^{\mathbb{N}}$.*

This proposition holds because our homogeneous groups are stratified. Since this proposition is rather well-known, we omit its proof. We point out that a similar statement is proved in Stein [S, p. 608, Lemma in Section 3.2.2].

Next, we have the following lemma that allows one to write the left-invariant derivative of a bump function as sums of right-invariant derivatives of some other bumps.

Proposition 4.3. *Suppose ϕ is a Schwartz function on G (by which we mean a Schwartz function on the underlying $\mathbb{R}^n$).*

(a) *If $\int_G \phi(x)dx = 0$, then there exist Schwartz functions $\varphi^{(1)},\dots,\varphi^{(n_1)}$ such that*

$$\phi = \sum_{k=1}^{n_1} X_k^R \varphi^{(k)}.$$

(b) *If furthermore $\int_G x_k\phi(x)dx = 0$ for all $1 \le k \le n_1$, then one can take the $\varphi^{(k)}$'s such that $\int_G \varphi^{(k)}(x)dx = 0$ for all $1 \le k \le n_1$. This will be the case, for instance, if ϕ is the left-invariant derivative of another Schwartz function whose integral is zero.*

Proof. To prove part (a), first we claim that any Schwartz function ϕ on $\mathbb{R}^n$ that has integral zero can be written as

$$\phi = \sum_{i=1}^{n} \frac{\partial \phi^{(i)}}{\partial x_i}$$

for some Schwartz functions $\phi^{(1)},\dots,\phi^{(n)}$.

To see that we have such a representation, we use the Euclidean Fourier transform on $\mathbb{R}^n$. First, we observe that since the integral of ϕ is zero, which implies $\widehat{\phi}(0) = 0$, we have, for all $\xi \in \mathbb{R}^n$,

$$\widehat{\phi}(\xi) = \int_0^1 \frac{d}{ds}\widehat{\phi}(s\xi)\,ds = \sum_{i=1}^n \xi_i \int_0^1 \frac{\partial \widehat{\phi}}{\partial \xi_i}(s\xi)\,ds. \tag{4.1}$$

Taking inverse Fourier transform, one can write ϕ as a sum of coordinate derivatives of some functions. The problem is that $\int_0^1 \frac{\partial \widehat{\phi}}{\partial \xi_i}(s\xi)\,ds$, while smooth in ξ, does not decay as $\xi \to \infty$. So the above expression is only good for small ξ. But for large ξ, we have

$$\widehat{\phi}(\xi) = \sum_{i=1}^n \xi_i \frac{\xi_i}{|\xi|^2}\widehat{\phi}(\xi). \tag{4.2}$$

(Here $|\xi|$ is the Euclidean norm of ξ.) Hence if we take a smooth cut-off $\eta \in C_c^\infty(\mathbb{R}^n)$ with $\eta \equiv 1$ near the origin, then combining (4.1) and (4.2), we have

$$\begin{aligned}\widehat{\phi}(\xi) &= \eta(\xi)\widehat{\phi}(\xi) + (1-\eta(\xi))\widehat{\phi}(\xi)\\ &= \sum_{i=1}^n \xi_i \left(\eta(\xi)\int_0^1 \frac{\partial \widehat{\phi}}{\partial \xi_i}(s\xi)\,ds + (1-\eta(\xi))\frac{\xi_i}{|\xi|^2}\widehat{\phi}(\xi)\right).\end{aligned}$$

Taking inverse Fourier transform, we get the desired decomposition in our claim above.

Now we return to the group setting. Let ϕ be a function on G with integral zero. Using the above claim, identifying G with the underlying $\mathbb{R}^n$, we write ϕ as

$$\phi = \sum_{i=1}^n \frac{\partial \phi^{(i)}}{\partial x_i}$$

for some Schwartz functions $\phi^{(1)}, \dots, \phi^{(n)}$. Now by Proposition 4.2(a), for all $1 \le i \le n$, one can express the coordinate derivatives $\partial/\partial x_i$ in terms of the right-invariant derivatives of order 1. Hence by rearranging the above identity we obtain Schwartz functions $\varphi^{(1)}, \dots, \varphi^{(n_1)}$ such that

$$\phi = \sum_{k=1}^{n_1} X_k^R \varphi^{(k)},$$

as was claimed in (a).

Finally, if we had in addition $\int_G x_k \phi(x)\,dx = 0$ for all $1 \le k \le n_1$, then one can check, in our construction above, that $\int_G \phi^{(i)}(x)\,dx = 0$ for all $1 \le i \le n_1$. It follows that $\int_G \varphi^{(i)}(x)\,dx = 0$ for all $1 \le i \le n_1$; in fact $\varphi^{(i)}$ is just $\phi^{(i)}$ plus a sum of derivatives of Schwartz functions, which integrates to zero. The rest of the proposition then follows. □

We point out here that in the above two propositions, left-invariant derivatives could have worked as well as right-invariant ones. More precisely:

Proposition 4.4. *Propositions 4.2 and 4.3 remain true if one replaces all right-invariant derivatives by their left-invariant counterparts.*

In what follows, we will consistently denote the operator $f \mapsto f * K$ by Kf if K is a kernel. If K is a Schwartz function, then $\nabla_b(Kf) = f * (\nabla_b K)$, where (each component of) $\nabla_b K$ is a Schwartz function with integral 0. Thus Proposition 4.3 can be applied to $\nabla_b K$; then one gets some kernels $\tilde{K}^{(k)}$ that are Schwartz functions, and satisfy

$$\nabla_b K = \sum_{k=1}^{n_1} X_k^R \tilde{K}^{(k)}.$$

Schematically we write $\nabla_b K = \nabla_b^R \tilde{K}$, and conclude that

$$\nabla_b(Kf) = f * (\nabla_b^R \tilde{K}) = (\nabla_b f) * \tilde{K}.$$

Again writing $\tilde{K} f$ for $f * \tilde{K}$, we obtain the identity

$$\nabla_b(Kf) = \tilde{K}(\nabla_b f).$$

If in addition $\int_G K(y)\,dy = 0$, then also $\int_G \tilde{K}(y)\,dy = 0$, by the last part of Proposition 4.3.

5. Littlewood–Paley theory and a refinement

We now turn to Littlewood–Paley theory for G. We actually need two versions of it. First, let Ψ be a Schwartz function on G such that $\int_G \Psi(x)\,dx = 1$, and $\int_G x_k \Psi(x)\,dx = 0$ for all $1 \leq k \leq n_1$. Such a function exists; in fact one can just take a function Ψ on $\mathbb{R}^n$ whose Euclidean Fourier transform is identically 1 near the origin, and think of it as a function on G. Now let $\Delta(x) = 2^Q \Psi(2 \cdot x) - \Psi(x)$, and $\Delta_j(x) = 2^{jQ} \Delta(2^j \cdot x)$. Also write $\Delta_j f = f * \Delta_j$. Then for $f \in L^p$, $1 < p < \infty$, we have

$$f = \sum_{j=-\infty}^{\infty} \Delta_j f$$

with convergence in L^p and

$$\left\| \Big(\sum_{j=-\infty}^{\infty} |\Delta_j f|^2 \Big)^{1/2} \right\|_{L^p} \leq C_p \|f\|_{L^p}. \tag{5.1}$$

This holds because $\int_G \Delta(y)\,dy = 0$. In fact we have the following more refined Littlewood–Paley theorem:

Proposition 5.1. *If D is a Schwartz function on G, $\int_G D(x)\,dx = 0$, and*

$$|D(x)| \leq A(1 + \|x\|)^{-(Q+2)},$$

$$\left| \frac{\partial}{\partial x_k} D(x) \right| \leq A(1 + \|x\|)^{-(Q+r+1)} \quad \text{if } n_{r-1} < k \leq n_r,\ 1 \leq r \leq m,$$

then defining $D_j f = f * D_j$ *where* $D_j(x) = 2^{jQ} D(2^j \cdot x)$, *we have*

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} |D_j f|^2 \Big)^{1/2} \Big\|_{L^p} \le C_p A \|f\|_{L^p}, \qquad 1 < p < \infty,$$

where C_p *is a constant independent of the kernel* D.

Later we will need the fact that the constant on the right hand side of the Littlewood–Paley inequality depends only on A but not otherwise on the kernel D. Applying this proposition to Δ_j yields our claim (5.1).

Proof. Without loss of generality we may assume $A = 1$. The proof relies on vector-valued singular integral theory on G, which is presented, for instance, in [S, Chapter 13, Section 5.3]. By our assumptions, it is readily checked that

$$\Big(\sum_{j=-\infty}^{\infty} |D_j(x)|^2 \Big)^{1/2} \le C \|x\|^{-Q}, \tag{5.2}$$

and

$$\Big(\sum_{j=-\infty}^{\infty} \Big| \frac{\partial}{\partial x_k} D_j(x) \Big|^2 \Big)^{1/2} \le C \|x\|^{-Q-r} \quad \text{if } n_{r-1} < k \le n_r,\ 1 \le r \le m; \tag{5.3}$$

in fact by scale invariance, it suffices to check this when $\|x\| \simeq 1$. For example, to bound $\sum_{j=-\infty}^{\infty} \big| \frac{\partial}{\partial x_k} D_j(x) \big|^2$ where $n_{r-1} < k \le n_r$, it suffices to split the sum into $\sum_{j \ge 0}$ and $\sum_{j<0}$; for the second sum, one bounds each term by $C 2^{j(Q+r)}$, and for the first sum, one bounds each term by $C 2^{j(Q+r)} 2^{-j(Q+r+1)}$. Putting these together yields the desired bound (5.3); (5.2) can be obtained similarly.

Furthermore, we need to check that for any normalized bump function Φ supported in the unit ball,

$$\Big(\sum_{j=-\infty}^{\infty} \Big| \int_G D_j(x) \Phi(R \cdot x)\, dx \Big|^2 \Big)^{1/2} \le C \quad \text{for all } R > 0. \tag{5.4}$$

By scale invariance we may assume that $R \simeq 1$. Now when $j < 0$,

$$\Big| \int_G D_j(x) \Phi(R \cdot x)\, dx \Big| \le 2^{jQ} \|\Phi(R \cdot x)\|_{L^1} \le C 2^{jQ},$$

since $|D_j(x)| \le 2^{jQ}$ and $R \simeq 1$. When $j \ge 0$,

$$\begin{aligned} \Big| \int_G D_j(x) \Phi(R \cdot x)\, dx \Big| &\le \int_G |D_j(x)|\, |\Phi(R \cdot x) - \Phi(0)|\, dx \\ &\le C \int_G |D_j(x)| R \|x\|\, dx \le C 2^{-j}, \end{aligned}$$

with the first inequality following from $\int_G D(x)\,dx = 0$, and the second inequality following from Proposition 3.1. Putting these together, we get the desired estimate (5.4).

From (5.2)–(5.4), the vector-valued singular integral theory mentioned above applies, and this gives the bounds in our current proposition. Since none of the constants C in (5.2)–(5.4) depends on the kernel D, neither does the bound of our conclusion. □

We record here that the assumption $\int_G x_k \Psi(x)\,dx = 0$ for all $1 \le k \le n_1$ guarantees that

$$\int_G x_k \Delta(x)\,dx = 0 \quad \text{for all } 1 \le k \le n_1. \tag{5.5}$$

Next, for the reverse Littlewood–Paley inequality, we need the second version of Littlewood–Paley projections, given by the following proposition:

Proposition 5.2. *There are $2n_1$ functions $\Lambda^{(1)}, \dots, \Lambda^{(2n_1)}$, each with zero integral on G, such that if $\Lambda_j^{(l)}(x) = 2^{jQ}\Lambda^{(l)}(2^j \cdot x)$ and $\Lambda_j^{(l)} f = f * \Lambda_j^{(l)}$, then for all $f \in L^p$,*

$$\sum_{l=1}^{2n_1} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j^{(l)} f|^2 \Big)^{1/2} \Big\|_{L^p} \simeq_p \|f\|_{L^p} \quad \textit{for all } 1 < p < \infty.$$

Since the sum over l is usually irrelevant for the estimates, we will abuse notation, and simply write

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j f|^2 \Big)^{1/2} \Big\|_{L^p} \simeq_p \|f\|_{L^p}.$$

Note that we do not claim $f = \sum_{l,j} \Lambda_j^{(l)} f$ here.

Proof. The key is to construct $2n_1$ Schwartz functions $\Lambda^{(1)}, \dots, \Lambda^{(2n_1)}$ and another $2n_1$ Schwartz functions $\Xi^{(1)}, \dots, \Xi^{(2n_1)}$, each of integral zero, such that the delta function at the identity 0 can be represented by

$$\delta_0 = \sum_{l=1}^{2n_1} \sum_{j=-\infty}^{\infty} \Lambda_j^{(l)} * \Xi_j^{(l)} \tag{5.6}$$

where $\Lambda_j^{(l)}(x) = 2^{jQ}\Lambda^{(l)}(2^j \cdot x)$ and $\Xi_j^{(l)}(x) = 2^{jQ}\Xi^{(l)}(2^j \cdot x)$. Here (5.6) must be interpreted to mean that for every $f \in L^p$, $1 < p < \infty$, we have

$$f = \sum_{l=1}^{2n_1} \sum_{j=-\infty}^{\infty} f * \Lambda_j^{(l)} * \Xi_j^{(l)}$$

with convergence in L^p; similarly for all calculations below involving the δ_0 function. Once we have such functions, we can write

$$\begin{aligned}(f, g) &= \sum_{l=1}^{2n_1} \sum_{j=-\infty}^{\infty} (\Xi_j^{(l)} \Lambda_j^{(l)} f, g) = \sum_{l=1}^{2n_1} \sum_{j=-\infty}^{\infty} (\Lambda_j^{(l)} f, \Xi_j^{(l)*} g) \\ &\le \sum_{l=1}^{2n_1} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j^{(l)} f|^2 \Big)^{1/2} \Big\|_{L^p} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Xi_j^{(l)*} g|^2 \Big)^{1/2} \Big\|_{L^{p'}},\end{aligned}$$

where (f, g) denotes the inner product on $L^2(G)$ and $\Xi_j^{(l)*}$ is the adjoint of $\Xi_j^{(l)}$ with respect to this inner product, which is also given by convolution against a Schwartz function of integral zero. Here p' is the dual exponent to p. Hence if $1 < p < \infty$, we can estimate the $L^{p'}$ norm above by Proposition 5.1, and get

$$(f, g) \le C_p \|g\|_{L^{p'}} \sum_{l=1}^{2n_1} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j^{(l)} f|^2 \Big)^{1/2} \Big\|_{L^p}$$

which is the desired reverse inequality since g is arbitrary. The forward inequality follows already from Proposition 5.1.

To construct Schwartz functions $\Lambda^{(1)}, \dots, \Lambda^{(2n_1)}$ and $\Xi^{(1)}, \dots, \Xi^{(2n_1)}$ with integral zero and satisfying (5.6), we proceed as follows. Let Ψ be as at the beginning of this section. Then $\Psi * \Psi$ is a Schwartz function (here $*$ is still the group convolution), and $\int_G \Psi * \Psi(x)\, dx = 1$. Let $\Psi_j(x) = 2^{jQ}\Psi(2^j \cdot x)$. Then

$$\begin{aligned} \delta_0 &= \sum_{j=-\infty}^{\infty} (\Psi_j * \Psi_j - \Psi_{j-1} * \Psi_{j-1}) \\ &= \sum_{j=-\infty}^{\infty} (\Psi_j * (\Psi_j - \Psi_{j-1}) + (\Psi_j - \Psi_{j-1}) * \Psi_{j-1}). \end{aligned} \tag{5.7}$$

Note that in the smaller brackets, we have the L^1 dilation of $\Psi_0 - \Psi_{-1}$. Now $\Psi_0 - \Psi_{-1}$ has integral zero, and the moments $\int_G x_k (\Psi_0(x) - \Psi_{-1}(x))\, dx$ are zero for $1 \le k \le n_1$. Hence by Propositions 4.3 and 4.4, we can write $\Psi_0 - \Psi_{-1}$ as either

$$\Psi_0 - \Psi_{-1} = \sum_{k=1}^{n_1} X_k^R \varphi^{(k)} \quad \text{or} \quad \Psi_0 - \Psi_{-1} = \sum_{k=1}^{n_1} X_k \psi^{(k)}$$

for some Schwartz functions $\varphi^{(1)}, \dots, \varphi^{(n_1)}$ and $\psi^{(1)}, \dots, \psi^{(n_1)}$, each of integral zero. Plugging the first identity back into the first term of (5.7) and the second identity into the second term, and integrating by parts using Proposition 4.1, we get

$$\delta_0 = \sum_{k=1}^{n_1} \sum_{j=-\infty}^{\infty} ((X_k \Psi)_j * \varphi_j^{(k)} + \psi_j^{(k)} * (X_k^R \Psi)_{j-1/2}).$$

Renaming the functions, we obtain the desired decomposition of δ_0 as in (5.6). □

To proceed further, we consider the maximal function on G defined by

$$Mf(x) = \sup_{r>0} \frac{1}{r^Q} \int_{\|y\| \le r} |f(x \cdot y^{-1})|\, dy.$$

We need the following properties of M:

Proposition 5.3. (a) *M is bounded on L^p for all $1 < p \le \infty$.*
(b) *M satisfies a vector-valued inequality, namely*

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} |Mf_j|^2 \Big)^{1/2} \Big\|_{L^p} \le C_p \Big\| \Big(\sum_{j=-\infty}^{\infty} |f_j|^2 \Big)^{1/2} \Big\|_{L^p} \quad \text{for all } 1 < p < \infty.$$

(c) *Moreover, if* $|\phi(y)| \le \varphi(\|y\|)$ *for some decreasing function* φ, *and* $A = \int_G \varphi(\|y\|)\, dy$, *then* $|f * \phi(x)| \le CAMf(x)$ *where* C *is a constant depending only on* G *but not on* ϕ.

(d) *In particular,*

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j f_j|^2 \Big)^{1/2} \Big\|_{L^p} \le C_p \Big\| \Big(\sum_{j=-\infty}^{\infty} |f_j|^2 \Big)^{1/2} \Big\|_{L^p} \quad \textit{for all } 1 < p < \infty.$$

The proof of these can be found in Stein [S, Chapter 2], once we notice that the group $(G, \|\cdot\|, dx)$ satisfies the real-variable structures set out in Chapter 1 of [S].

We also need a Littlewood–Paley inequality for derivatives:

Proposition 5.4. *We have*

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} (2^j |\Delta_j f|)^2 \Big)^{1/2} \Big\|_{L^p} \le C_p \|\nabla_b f\|_{L^p} \quad \textit{if } \nabla_b f \in L^p, \textit{ for all } 1 < p < \infty.$$

Proof. Just notice that using the notations introduced at the end of Section 4,

$$2^j \Delta_j f = 2^j f * (\nabla_b^R \tilde{\Delta})_j = (\nabla_b f) * \tilde{\Delta}_j,$$

with $\int \tilde{\Delta}(x)\, dx = 0$ by (5.5) and the second part of Proposition 4.3. Since $\nabla_b f \in L^p$, $1 < p < \infty$, we have

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} (2^j \Delta_j f)^2 \Big)^{1/2} \Big\|_{L^p} = \Big\| \Big(\sum_{j=-\infty}^{\infty} (\tilde{\Delta}_j (\nabla_b f))^2 \Big)^{1/2} \Big\|_{L^p} \le C_p \|\nabla_b f\|_{L^p}$$

for $1 < p < \infty$. □

The following is a Bernstein-type inequality for our Littlewood–Paley decomposition Δ_j:

Proposition 5.5. *If* $f \in N\dot{L}^{1,Q}$, *then for all* $j \in \mathbb{Z}$,

$$\|\Delta_j f\|_{L^\infty} \le C \|\nabla_b f\|_{L^Q}$$

where C *is independent of both* f *and* j.

Proof. Suppose $\nabla_b f \in L^Q$. Using the notations in the proof of Proposition 5.4,

$$\Delta_j f = (\nabla_b f) * (2^{-j} \tilde{\Delta}_j).$$

Since $\|2^{-j} \tilde{\Delta}_j\|_{L^{Q/(Q-1)}}$ is a constant C independent of j, we see that

$$\|\Delta_j f\|_{L^\infty} \le C \|\nabla_b f\|_{L^Q}$$

as desired. □

Finally, we need the "heat kernels" which we define as follows. Let S be a non-negative Schwartz function on G which satisfies

$$\int_G S(y)\,dy = 1 \quad \text{and} \quad S(x) \simeq e^{-\|x\|} \quad \text{for all } x \in G.$$

(For instance, $S(x) = ce^{-(1+\|x\|^{2m!})^{1/2m!}}$ will do for a suitable c, since

$$(1 + \|x\|^{2m!})^{1/2m!} - \|x\| \to 0$$

as $\|x\| \to \infty$.) We write

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

and as usual let $S_j f := f * S_j$.

6. Algebraic preliminaries

In this section we describe some algebraic structures we use in the proof of Lemma 1.7. First we have the following algebraic identity:

Proposition 6.1. *For any sequence $\{a_j\}$, one has*

$$1 = \sum_{j=1}^{N} a_j \prod_{1 \le j' < j} (1 - a_{j'}) + \prod_{j=1}^{N} (1 - a_j).$$

Proof. This is just saying that

$$\begin{aligned} 1 &= a_1 + (1 - a_1) = a_1 + a_2(1 - a_1) + (1 - a_1)(1 - a_2) \\ &= a_1 + a_2(1 - a_1) + a_3(1 - a_1)(1 - a_2) + (1 - a_1)(1 - a_2)(1 - a_3) \\ &= \cdots . \end{aligned}$$

□

Note that by renaming the indices, one can also write

$$1 = \sum_{j=1}^{N} a_j \prod_{j < j' \le N} (1 - a_{j'}) + \prod_{j=1}^{N} (1 - a_j).$$

Hence we have:

Proposition 6.2. *If $\{a_j\}$ is a sequence of numbers satisfying $0 \le a_j \le 1$ for all j, then*

$$0 \le \sum_{j=-\infty}^{\infty} a_j \prod_{j' > j} (1 - a_{j'}) \le 1.$$

Next, suppose we are given a function h on G such that

$$h = \sum_j h_j \tag{6.1}$$

for some functions h_j, where all h_j satisfy $\|h_j\|_{L^\infty} \le C$. We will describe a paradigm in which we approximate h by an L^∞ function that we will call $\tilde{h}$. In fact, motivated by the algebraic proposition we have above, we let

$$\tilde{h} = \sum_j h_j \prod_{j'>j} (1 - U_{j'}) \tag{6.2}$$

where U_j are some suitable non-negative functions such that

$$C^{-1}|h_j| \le U_j \le 1 \quad \text{pointwise for all } j. \tag{6.3}$$

Then at least $\|\tilde{h}\|_{L^\infty} \le C$ because

$$|\tilde{h}(x)| \le \sum_j |h_j(x)| \prod_{j'>j} (1 - U_{j'}(x)) \le C \sum_j U_j(x) \prod_{j'>j} (1 - U_{j'}(x)) \le C,$$

where the last inequality follows from Proposition 6.2. One would now ask whether this could be any sensible approximation of h; in particular, let us try to see whether $\|X_k(h - \tilde{h})\|_{L^Q}$ is small for $k = 1, \dots, n_1$. To understand this, write $h = \sum_j h_j$. Then

$$h - \tilde{h} = \sum_j h_j \Big(1 - \prod_{j'>j} (1 - U_{j'})\Big).$$

Using Proposition 6.1 to expand the latter bracket and rearranging the resulting sum, we get

$$h - \tilde{h} = \sum_j U_j V_j, \tag{6.4}$$

where V_j is defined by

$$V_j = \sum_{j'<j} h_{j'} \prod_{j'<j''<j} (1 - U_{j''}). \tag{6.5}$$

It follows that

$$X_k(h - \tilde{h}) = \sum_j (X_k U_j) V_j + \sum_j U_j (X_k V_j).$$

By Proposition 6.2 and (6.3), we have

$$|V_j| \le C \quad \text{pointwise for all } j. \tag{6.6}$$

This can be shown using the same argument we have used to bound $\|\tilde{h}\|_{L^\infty}$. Furthermore,

$$|X_k V_j| \le C \sum_{j'<j} (|X_k h_{j'}| + |X_k U_{j'}|). \tag{6.7}$$

In fact this follows from

$$X_k V_j = \sum_{j'<j} ((X_k h_{j'}) - (X_k U_{j'}) V_{j'}) \prod_{j'<j''<j} (1 - U_{j''}). \tag{6.8}$$

(6.8) holds because when one computes $X_k V_j$, the derivative hits either $h_{j'}$ or $U_{j'}$, for some $j' < j$; furthermore, the coefficient of $X_k U_{j'}$ in $X_k V_j$ is

$$-V_{j'} \prod_{j'<j''<j} (1 - U_{j''}).$$

From (6.6) and (6.7), it follows that

$$|X_k(h - \tilde{h})| \leq C\Big(\sum_j |X_k U_j| + \sum_j U_j \sum_{j'<j} (|X_k h_{j'}| + |X_k U_{j'}|)\Big); \tag{6.9}$$

we will hope to estimate this in the L^Q norm on G, if we choose U_j suitably.

In the following sections, equations (6.1)–(6.9) will form a basic paradigm of our construction.

7. Proof of Lemma 1.7: Outline

In this section we give an outline of the proof of Lemma 1.7; the details are deferred to the next four sections. The outline we give below is merely formal, neglecting some delicate issues of convergence. We will indicate how this can be fixed at the end of this section.

The proof will be in three steps. First, given $\delta > 0$ and $f \in N\dot{L}^{1,Q}$, we find a large positive integer N such that

$$\|\nabla_b f_0\|_{L^Q} \leq \frac{\delta}{3} \|\nabla_b f\|_{L^Q}, \tag{7.1}$$

where

$$f_0 := f - \sum_{j=-\infty}^{\infty} S_{j+N} \Delta_j f.$$

This is possible basically because from $f = \sum_{j=-\infty}^{\infty} \Delta_j f$, we deduce that $f_0 = \sum_{j=-\infty}^{\infty} (1 - S_{j+N}) \Delta_j f$, which is small when N is sufficiently large. We will see that N can be chosen to depend only on δ but not on f.

Next, we define auxiliary control functions ω_j and $\tilde{\omega}_j$ as follows. Let σ be a large positive integer to be chosen later, and suppose from now on we have the "smallness" condition on f:

$$\|\nabla_b f\|_{L^Q} \leq c_G 2^{-NQ} 2^{-\sigma(Q-1)}, \tag{7.2}$$

with the constant c_G depending only on the group G. For $x = [x_1, \ldots, x_n]$, we recall $x^\sigma := [2^\sigma x_1, x_2, \ldots, x_n]$ and $x_\sigma := 2^{-\sigma} x^\sigma$. Let E be the Schwartz function defined by

$$E(x) := e^{-(1+\|x_\sigma\|^{2m!})^{1/2m!}}.$$

We write Λ for the lattice $\{2^{-N} \cdot s : s \in \mathbb{Z}^n\}$ of scale 2^{-N} in G, and define ω_j by

$$\omega_j(x) := \Big(\sum_{r\in\Lambda} [S_{j+N} |\Delta_j f|(2^{-j} \cdot r) E(r^{-1} \cdot (2^j \cdot x))]^Q\Big)^{1/Q} \tag{7.3}$$

for all $x \in G$. Here N is the positive integer we chose previously. Note that ω_j is like a discrete convolution, except that we are using the l^Q norm in r rather than the sum in r. We also define $\tilde{\omega}_j$ by

$$\tilde{\omega}_j := 2^{NQ} E_j S_{j+N} |\Delta_j f| \tag{7.4}$$

where $E_j f := f * E_j$ and $E_j(y) := 2^{jQ} E(2^j y)$. The functions ω_j and $\tilde{\omega}_j$ will be used to control the Littlewood–Paley pieces $\Delta_j f$ of f; in fact they will respectively control h_j and g_j we introduce below. They will also have better derivatives in the $X_2, \dots, X_{n_1}$ directions than in the X_1 direction.

Now we decompose $f - f_0$ into the sum of two functions g and h as follows. Let $R \gg \sigma$ be another positive integer to be chosen later, and let ζ be a smooth function on $[0, \infty)$ such that $\zeta \equiv 1$ on $[0, 1/2]$, and $\zeta \equiv 0$ on $[1, \infty)$. Let

$$\zeta_j(x) := \zeta\left(\frac{2^j \omega_j(x)}{\sum_{k<j,\, k \equiv j \pmod R} 2^k \omega_k(x)} \right).$$

We remark here that $\zeta_j(x)$ is not the L^1 dilation of ζ, i.e. $\zeta_j(x) \neq 2^{jQ} \zeta(2^j \cdot x)$; it is a smooth approximation of the characteristic function of the set $\{2^j \omega_j < \sum_{k<j,\, k \equiv j \pmod R} 2^k \omega_k\}$. We then define

$$h := \sum_{j=-\infty}^{\infty} h_j, \quad g := \sum_{j=-\infty}^{\infty} g_j,$$

where

$$h_j(x) := (1 - \zeta_j(x)) S_{j+N}(\Delta_j f)(x), \quad g_j(x) := \zeta_j(x) S_{j+N}(\Delta_j f)(x).$$

It follows that

$$f = f_0 + g + h.$$

By ζ_j's definition, we can think of h as the part where "the high frequencies dominate the low frequencies", and g as the part where the reverse happens. We remark that g and h depend on R, σ and N. We are free to choose R and σ; but N is fixed once we fix δ.

To proceed further, we will approximate h by some L^∞ function $\tilde{h}$ using the paradigm of approximation we discussed in the previous section. Namely, we define

$$\tilde{h} := \sum_{j=-\infty}^{\infty} h_j \prod_{j'>j} (1 - U_{j'}) \quad \text{where} \quad U_j := (1 - \zeta_j)\omega_j.$$

We will prove that

$$\|\tilde{h}\|_{L^\infty} \leq C, \tag{7.5}$$

$$\|X_k(h - \tilde{h})\|_{L^Q} \leq C_N 2^{-\sigma/Q} R \|\nabla_b f\|_{L^Q} + C_N 2^{\sigma Q} R \|\nabla_b f\|_{L^Q}^2, \quad k = 2, \dots, n_1, \tag{7.6}$$

$$\|X_1(h - \tilde{h})\|_{L^Q} \leq C_N 2^{\sigma Q} R \|\nabla_b f\|_{L^Q}. \tag{7.7}$$

Finally, using the same paradigm, we approximate g by some $\tilde{g} \in L^\infty$, where

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

We will prove that

$$\|\tilde{g}\|_{L^\infty} \le CR, \tag{7.8}$$

$$\|\nabla_b (g - \tilde{g})\|_{L^Q} \le C_N 2^{2\sigma(Q+1)} R^2 2^{-R} \|\nabla_b f\|_{L^Q}. \tag{7.9}$$

Note that one can estimate the full ∇_b of the error here (rather than only the "good" derivatives X_k for $k = 2, \dots, n_1$). We will see in later sections that the "smallness" assumption (7.2) on the given f is used right here, in the proofs of (7.5)–(7.9).

Altogether, if we define F to be $\tilde{g} + \tilde{h}$, then by (7.5) and (7.8),

$$\|F\|_{L^\infty} \le CR,$$

and by (7.1), (7.6) and (7.9), for $k = 2, \dots, n_1$,

$$\begin{aligned} \|X_k(f - F)\|_{L^Q} &\le \|\nabla_b f_0\|_{L^Q} + \|X_k(h - \tilde{h})\|_{L^Q} + \|\nabla_b (g - \tilde{g})\|_{L^Q} \\ &\le \frac{\delta}{3} \|\nabla_b f\|_{L^Q} + C_N 2^{-\sigma/Q} R \|\nabla_b f\|_{L^Q} + C_N 2^{\sigma Q} R \|\nabla_b f\|_{L^Q}^2 \\ &\quad + C_N 2^{2\sigma(Q+1)} R^2 2^{-R} \|\nabla_b f\|_{L^Q}. \end{aligned}$$

If one now chooses $R = B\sigma$ where B is a constant $> 2(Q+1)$ (say $B = 2(Q+2)$ will do), and chooses σ to be sufficiently large with respect to N, then the above is bounded by

$$\frac{\delta}{2} \|\nabla_b f\|_{L^Q} + A_\delta \|\nabla_b f\|_{L^Q}^2 \tag{7.10}$$

where A_δ is a constant depending only on G and δ (remember N is fixed once we choose δ). The above analysis is valid whenever $\delta > 0$ and (7.2) holds, namely $\|\nabla_b f\|_{L^Q} \le c_G 2^{-NQ} 2^{-\sigma(Q-1)}$, where σ is the one that we just picked depending on δ. Note that N, σ and R are chosen independent of our given function f; they depend only on the given δ. If now a general f in $N\dot{L}^{1,Q}$ is given, and δ is sufficiently small, one will rescale f so that

$$\|\nabla_b f\|_{L^Q} = \min\{c_G 2^{-NQ} 2^{-\sigma(Q-1)}, \delta (2A_\delta)^{-1}\}.$$

The right hand side depends only on δ. Now using bound (7.10), for $k = 2, \dots, n_1$,

$$\|X_k(f - F)\|_{L^Q} \le \delta \|\nabla_b f\|_{L^Q}$$

as desired. Also,

$$\|F\|_{L^\infty} \le CR \frac{\|\nabla_b f\|_{L^Q}}{\min\{c_G 2^{-NQ} 2^{-\sigma(Q-1)}, \delta(2A_\delta)^{-1}\}} = C_\delta \|\nabla_b f\|_{L^Q}.$$

Similarly, one can derive the estimate $\|\nabla_b(f-F)\|_{L^Q} \le C_\delta \|\nabla_b f\|_{L^Q}$ by using (7.1), (7.7) and (7.9). This implies $\|\nabla_b F\|_{L^Q} \le C_\delta \|\nabla_b f\|_{L^Q}$, and completes the formal proof of Lemma 1.7.

To make rigorous the above formal argument, one needs to first approximate f by $T_K f := \sum_{|j|\le K} \Delta_j f$ in $\dot{NL}^{1,Q}$ where K is some large integer. In fact, given $\delta > 0$, if K is sufficiently large, one can have

$$\|\nabla_b(f - T_K f)\|_{L^Q} \le \delta \|\nabla_b f\|_{L^Q}.$$

Then one should define f_0 such that

$$f_0 = T_K f - \sum_{|j|\le K} S_{j+N}\Delta_j f,$$

where N is an integer large enough so that (7.1) holds. One defines ω_j, $\tilde{\omega}_j$, h_j, g_j, U_j and G_j as before if $|j| \le K$, and redefines them to be zero if $|j| > K$. The definitions of ζ_j, $\tilde{h}$, $\tilde{g}$ and F remain unchanged. One can then deduce rigorously the desired results for $\tilde{h}$ and $\tilde{g}$ as in equations (7.5)–(7.9), by following the argument presented in the next few sections, and interpreting all $\Delta_j f$ as zero whenever $|j| > K$. In particular, in the display equation immediately before (11.3), the sum $\sum_{s=-\infty}^{\infty} \nabla_b(G_s H_s)$ becomes a finite sum (only over $|s| \le K$), and would then be in L^Q (since each term is by the arguments in Section 11), so that the use of Proposition 5.2 to bound its L^Q norm is justified. The details of such checking are left to the interested reader. One then concludes the proof as above, using, instead of $f = f_0 + g + h$, that

$$f = (f - T_K f) + f_0 + g + h.$$

In what follows, C will denote constants independent of δ, N, σ and R. All dependence of constants on N, σ and R will be made clear in the notations.

8. Estimating f_0

We now begin the proof of our approximation Lemma 1.7. Let N be a large positive integer. Define f_0 as in Section 7. First,

$$f_0 := \sum_{j=-\infty}^{\infty} (I - S_{j+N})\Delta_j f = \sum_{j=-\infty}^{\infty} \sum_{k\ge N} (S_{j+k+1} - S_{j+k})\Delta_j f$$

where I is the identity operator and the sum in k converges in $\dot{NL}^{1,Q}$ norm. Now let P be the kernel of the operator $S_1 - S_0$. Then P is a Schwartz function, and

$$\int_G P(y)\,dy = 0.$$

Furthermore, if we define $P_k f = f * P_k$ where $P_k(y) = 2^{kQ} P(2^k \cdot y)$, then

$$f_0 = \sum_{j=-\infty}^{\infty} \sum_{k\ge N} P_{j+k}(\Delta_j f),$$

where the convergence is in $N\dot{L}^{1,Q}$. Using the notation at the end of Section 4, one gets

$$\nabla_b f_0 = \sum_{j=-\infty}^{\infty} \sum_{k \geq N} \tilde{P}_{j+k} \tilde{\Delta}_j (\nabla_b f),$$

where the convergence is in L^Q, the kernels $\tilde{P}$ and $\tilde{\Delta}$ are Schwartz, and satisfy

$$\int_G \tilde{P}(y)\,dy = \int_G \tilde{\Delta}(y)\,dy = 0$$

since $\int_G P = \int_G \Delta = 0$. Now it follows from Proposition 5.2 that

$$\|\nabla_b f_0\|_{L^Q} \simeq \left\| \left(\sum_{r=-\infty}^{\infty} \left| \Lambda_r \sum_{j=-\infty}^{\infty} \sum_{k \geq N} \tilde{P}_{j+k} \tilde{\Delta}_j (\nabla_b f) \right|^2 \right)^{1/2} \right\|_{L^Q}.$$

To proceed further, we replace, in the right hand side of the above formula, the index j by $j + r$, and then pull out the summation over j and k. Then we obtain for $\|\nabla_b f_0\|_{L^Q}$ the following bound:

$$\sum_{j=-\infty}^{\infty} \sum_{k \geq N} \left\| \left(\sum_{r=-\infty}^{\infty} |\Lambda_r \tilde{P}_{j+r+k} \tilde{\Delta}_{j+r} (\nabla_b f)|^2 \right)^{1/2} \right\|_{L^Q},$$

which is equal to

$$\sum_{j=-\infty}^{\infty} \sum_{k \geq N} \left\| \left(\sum_{r=-\infty}^{\infty} |\Lambda_{r+j} \tilde{P}_{r+k} \tilde{\Delta}_r (\nabla_b f)|^2 \right)^{1/2} \right\|_{L^Q} \tag{8.1}$$

if we first replace the index r by $r-j$, and then replace j by $-j$. Now we split the sum into two parts, over $j < 0$ and $j \geq 0$, and show that both are bounded by $C2^{-N} \|\nabla_b f\|_{L^Q}$. The sum where $j < 0$ can be estimated using Proposition 3.5: in fact $\Lambda_{r+j} \tilde{P}_{r+k} \tilde{\Delta}_r (\nabla_b f) = (\tilde{\Delta}_r (\nabla_b f)) * (\tilde{P}_{r+k} * \tilde{\Lambda}_{r+j})$, and by Proposition 3.5,

$$|\tilde{P}_{r+k} * \tilde{\Lambda}_{r+j}|(x) = |(\tilde{P} * \tilde{\Lambda}_{k-j})_{r+j}|(x) \leq C2^{-(k-j)} 2^{Q(r+j)} (1 + 2^{r+j} \|x\|)^{-(Q+1)}$$

since $k - j > 0$, Λ and $\tilde{P}$ are Schwartz, and $\int_G \tilde{P}(y)\,dy = 0$. From this we infer, using Proposition 5.3(c), that

$$|\Lambda_{r+j} \tilde{P}_{r+k} \tilde{\Delta}_r (\nabla_b f)| \leq C2^{-(k-j)} M \tilde{\Delta}_r (\nabla_b f).$$

It follows that the sum where $j < 0$ is bounded by

$$\sum_{j<0} \sum_{k \geq N} C2^{-(k-j)} \left\| \left(\sum_{r=-\infty}^{\infty} (M \tilde{\Delta}_r (\nabla_b f))^2 \right)^{1/2} \right\|_{L^Q} \leq C2^{-N} \|\nabla_b f\|_{L^Q}$$

as desired. Next, for the sum where $j \geq 0$, one defines an auxiliary kernel D by

$$D = \tilde{\Delta} * \tilde{P}_k * \Lambda_j \quad \text{if } j \geq 0 \text{ and } k \geq N.$$

We claim the following:

Proposition 8.1. *D is a Schwartz function,*

$$|D(x)| \leq C2^{-j-k}(1+\|x\|)^{-(Q+2)},$$

$$\left|\frac{\partial}{\partial x_l} D(x)\right| \leq C2^{-j-k}(1+\|x\|)^{-(Q+r+1)} \quad \text{if } n_{r-1} < l \leq n_r,\ 1 \leq r \leq m,$$

and

$$\int_G D(x)\,dx = 0.$$

Assume the proposition for the moment. Then one can apply the refinement of Littlewood–Paley theory in Proposition 5.1 and conclude that

$$\left\|\left(\sum_{r=-\infty}^{\infty} |D_r(\nabla_b f)|^2\right)^{1/2}\right\|_{L^Q} \leq C2^{-j-k}\|\nabla_b f\|_{L^Q},$$

i.e.

$$\left\|\left(\sum_{r=-\infty}^{\infty} |\Lambda_{r+j}\tilde{P}_{r+k}\tilde{\Delta}_r(\nabla_b f)|^2\right)^{1/2}\right\|_{L^Q} \leq C2^{-j-k}\|\nabla_b f\|_{L^Q}.$$

Summing over $j \geq 0$ and $k \geq N$, we see that the sum over $j \geq 0$ in (8.1) is bounded by $C2^{-N}\|\nabla_b f\|_{L^Q}$ as well, as desired.

Proof of Proposition 8.1. It is clear that D is Schwartz and $\int_G D(x)\,dx = 0$. To prove the estimate for $|D(x)|$, first consider $\tilde{\Delta} * \tilde{P}_k$ with $k \geq N$. From Proposition 4.1, we have $\nabla_b^R(\tilde{\Delta} * \tilde{P}_k) = (\nabla_b^R \tilde{\Delta}) * \tilde{P}_k$, so by Proposition 3.4, we get

$$|(\nabla_b^R)^\alpha(\tilde{\Delta} * \tilde{P}_k)(x)| \leq C_{\alpha,M}2^{-k}(1+\|x\|)^{-M}$$

for all multi-indices α and all $M > 0$. It follows that

$$|\tilde{\Delta} * \tilde{P}_k(x)| + \left|\frac{\partial}{\partial x_l}(\tilde{\Delta} * \tilde{P}_k)(x)\right| \leq C_{M,l}2^{-k}(1+\|x\|)^{-M}$$

for any $M > 0$ and all $1 \leq l \leq n$. Here we have applied Proposition 4.2(b), which says that each $\partial/\partial x_l$ can be written as a linear combination of $(\nabla_b^R)^\alpha$ with coefficients that are polynomials in x. Applying Proposition 3.4 again, we get

$$|D(x)| \leq C_M 2^{-j-k}(1+\|x\|)^{-M}$$

for all $M > 0$, since $j \geq 0$. Similarly

$$|(\nabla_b^R)^\alpha D(x)| \leq C_{\alpha,M}2^{-j-k}(1+\|x\|)^{-M}$$

for all multi-indices α and all $M > 0$. It follows that

$$\left|\frac{\partial}{\partial x_l} D(x)\right| \leq C_M 2^{-j-k}(1+\|x\|)^{-M}$$

for all $1 \leq l \leq n$ and all $M > 0$, from which our proposition follows. □

9. Properties of ω_j and $\tilde{\omega}_j$

Suppose $\delta > 0$ is given, and let N be chosen as in the previous section. Let σ be a very large positive integer, to be chosen depending on N and thus on δ. Suppose (7.2) holds, i.e. $\|\nabla_b f\|_{L^Q} \leq c_G 2^{-NQ} 2^{-\sigma(Q-1)}$, where c_G is a sufficiently small constant depending only on G. Define ω_j and $\tilde{\omega}_j$ by (7.3) and (7.4):

$$\omega_j(x) := \Big(\sum_{r\in\Lambda}[S_{j+N}|\Delta_j f|(2^{-j}\cdot r)E(r^{-1}\cdot(2^j\cdot x))]^Q\Big)^{1/Q}$$

$$\tilde{\omega}_j(x) := 2^{NQ} E_j S_{j+N}|\Delta_j f|(x).$$

First, we need a pointwise bound for ω_j. To obtain it we observe:

Proposition 9.1. *Let S_j and E_j be defined as in Section 7. Then whenever $x, \theta \in G$ with $\|\theta\| \leq 2^{-j}$, we have*

$$S_j(x\cdot\theta) \simeq S_j(x) \quad \textit{and} \quad E_j(\theta\cdot x) \simeq E_j(x).$$

In particular,

$$S_j f(x\cdot\theta) \simeq S_j f(x)$$

if f is a non-negative function and $\|\theta\| \leq 2^{-j}$.

Proof. First we observe that

$$E(x) \simeq e^{-\|x_\sigma\|} \quad \text{for all } x \in G.$$

This is because

$$(1+\|x_\sigma\|^{2m!})^{1/2m!} - \|x_\sigma\| \to 0$$

as $\|x_\sigma\| \to \infty$. Now by Proposition 3.6, we have

$$|\,\|(\theta\cdot x)_\sigma\| - \|x_\sigma\|\,| \leq C \quad \text{if } \|\theta\| \leq 1.$$

Hence from $E(x) \simeq e^{-\|x_\sigma\|}$ and $E(\theta\cdot x) \simeq e^{-\|(\theta\cdot x)_\sigma\|}$, we get

$$E(\theta\cdot x) \simeq E(x) \quad \text{if } \|\theta\| \leq 1.$$

Scaling yields the desired claim for E_j.

Next, suppose $\|\theta\| \leq 1$. We claim that $S(x\cdot\theta) \simeq S(x)$ for all $x \in G$. This holds because $S(x\cdot\theta) \simeq e^{-\|x\cdot\theta\|}$ and $S(x) \simeq e^{-\|x\|}$ for all x, and one can apply (3.3) to compare the right hand sides. Scaling yields the claim for S_j. □

Now comes the pointwise bound for ω_j, from both above and below.

Proposition 9.2.

$$\omega_j(x) \simeq \Big(\sum_{r\in\Lambda}[S_{j+N}|\Delta_j f|(x\cdot 2^{-j}r)E(r^{-1})]^Q\Big)^{1/Q}.$$

Here the implicit constant is independent of N and σ.

Proof. Recall that by (7.3),

$$\omega_j(x) := \Big(\sum_{r\in\Lambda}[S_{j+N}|\Delta_j f|(2^{-j}\cdot r)E(r^{-1}\cdot(2^j\cdot x))]^Q\Big)^{1/Q}$$
$$= \Big(\sum_{s\in(2^j\cdot x)^{-1}\cdot\Lambda}[S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot s))E(s^{-1})]^Q\Big)^{1/Q}$$

The last identity follows from a change of variable: if $s = (2^j\cdot x)^{-1}\cdot r$, then we have $r = (2^j\cdot x)\cdot s$, so $2^{-j}\cdot r = x\cdot(2^{-j}\cdot s)$, the last identity following because dilations are group homomorphisms (cf. (2.2)). Now recall that Λ is the lattice $\{2^{-N}\cdot s : s\in\mathbb{Z}^n\}$. Hence every $s\in(2^j\cdot x)^{-1}\cdot\Lambda$ can be written uniquely as $r\cdot(2^{-N}\cdot\theta)$ for some $r\in\Lambda$ and $\theta\in G$ such that if $\theta=[\theta_1,\dots,\theta_n]$, then all $\theta_k\in[0,1)$. This defines a map from the shifted lattice $(2^j\cdot x)^{-1}\cdot\Lambda$ to the original lattice Λ, and it is easy to see that this map is a bijection. Hence if the inverse of this map is denoted by $s=s(r)$, then

$$\omega_j(x) = \Big(\sum_{r\in\Lambda}[S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot s(r)))E(s(r)^{-1})]^Q\Big)^{1/Q}.$$

But $s(r) = r\cdot(2^{-N}\cdot\theta)$ for some $\|\theta\|\le 1$. Thus by Proposition 9.1, we get

$$E(s(r)^{-1}) \simeq E(r^{-1}).$$

Also, from the same relation between $s(r)$ and r, we deduce that $2^{-j}\cdot s(r) = (2^{-j}\cdot r)\cdot(2^{-(j+N)}\cdot\theta)$ with $\|2^{-(j+N)}\cdot\theta\|\le 2^{-(j+N)}$. Thus by Proposition 9.1 again, we get

$$S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot s(r))) \simeq S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot r)).$$

Hence the proposition follows. □

By a similar token, one can prove

Proposition 9.3.

$$\tilde\omega_j(x) \simeq \sum_{r\in\Lambda} S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot r))E(r^{-1}),$$

with implicit constants independent of N and σ.

Proof. This is because by (7.4),

$$\tilde\omega_j(x) = 2^{NQ}\int_G S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot y))E(y^{-1})\,dy$$
$$= \sum_{r\in\Lambda}\int_{[0,1)^n} S_{j+N}|\Delta_j f|(x\cdot(2^{-j}\cdot r)\cdot(2^{-(j+N)}\cdot\theta))E((2^{-N}\cdot\theta)^{-1}\cdot r^{-1})\,d\theta.$$

The second equality follows from the fact that every $y\in G$ can be written uniquely as $r\cdot(2^{-N}\cdot\theta)$ for some $r\in\Lambda$ and $\theta\in[0,1)^n$, which we have already used in the proof of Proposition 9.2. Note again that if $y = r\cdot(2^{-N}\cdot\theta)$, then by the fact that dilations are group homomorphisms, we have $2^{-j} \cdot y = 2^{-j} \cdot (r \cdot (2^{-N} \cdot \theta)) = (2^{-j} \cdot r) \cdot (2^{-(j+N)} \cdot \theta)$. Also, we used $dy = 2^{-NQ} d\theta$ in the change of variables. Now one can mimic the proof of Proposition 9.2. In fact, one observes that whenever $\|\theta\| \leq 1$, one has

$$E((2^{-N} \cdot \theta)^{-1} \cdot r^{-1}) \simeq E(r^{-1})$$

and

$$S_{j+N}|\Delta_j f|(x \cdot (2^{-j} \cdot r) \cdot (2^{-(j+N)} \cdot \theta)) \simeq S_{j+N}|\Delta_j f|(x \cdot (2^{-j} \cdot r)).$$

One then concludes that

$$\tilde{\omega}_j(x) \simeq \sum_{r \in \Lambda} S_{j+N}|\Delta_j f|(x \cdot (2^{-j} \cdot r)) E(r^{-1}). \qquad \square$$

The two propositions above yield

Proposition 9.4.

$$S_{j+N}|\Delta_j f|(x) \leq C\omega_j(x) \leq C\tilde{\omega}_j(x).$$

Proof. The first inequality holds because the term corresponding to $r = 0$ on the right hand side of the equation in Proposition 9.2 is precisely $S_{j+N}|\Delta_j f|(x)$. The second inequality holds by the previous two propositions, since the l^Q norm of a sequence is always smaller than or equal to its l^1 norm. $\square$

Next we have

Proposition 9.5.

$$\|\omega_j\|_{L^\infty} \leq C\|\tilde{\omega}_j\|_{L^\infty} \leq C 2^{NQ} 2^{\sigma(Q-1)} \|\nabla_b f\|_{L^Q} \leq 1$$

if c_G in assumption (7.2) is chosen sufficiently small.

We fix this choice of c_G from now on.

Proof. The first inequality follows from the previous proposition. The second inequality follows from

$$\|\tilde{\omega}_j\|_{L^\infty} \leq 2^{NQ} \|E\|_{L^1} \|S\|_{L^1} \|\Delta_j f\|_{L^\infty}, \qquad \|E\|_{L^1} = C 2^{\sigma(Q-1)},$$

and Bernstein's inequality as in Proposition 5.5. The last inequality holds since $\|\nabla_b f\|_{L^Q} \leq c_G 2^{-NQ} 2^{-\sigma(Q-1)}$ and c_G is sufficiently small. $\square$

Proposition 9.6.

$$|X_1 \omega_j| \leq C 2^j \omega_j \quad \textit{and} \quad |X_k \omega_j| \leq C 2^{j-\sigma} \omega_j \quad \textit{for } k = 2, \ldots, n_1.$$

Proof. One just needs to recall the definition of ω_j from (7.3), namely

$$\omega_j(x) = \Big(\sum_{r\in\Lambda}[S_{j+N}|\Delta_j f|(2^{-j}\cdot r)E(r^{-1}\cdot(2^j\cdot x))]^Q\Big)^{1/Q},$$

and to differentiate it. Here it is crucial that the variable x is in the argument of E and not in $S_{j+N}|\Delta_j f|$; in other words, we could not have taken the expression in Proposition 9.2 to be the definition of ω_j, because while it is true that the continuous convolution $f * g$ can be written as $\int_G f(y^{-1})g(y\cdot x)\,dy$ or $\int_G f(x\cdot y^{-1})g(y)\,dy$ via integration by parts, the analogous statement fails for discrete convolutions. Hence if ω_j was defined by the expression in Proposition 9.2, then there would be no way of integrating by parts and letting the derivatives fall on E here.

More precisely, first we observe

$$|X_k(1+\|x_\sigma\|^{2m!})^{1/2m!}| \le C2^{-\sigma} \quad \text{if } k=2,\dots,n_1,$$
$$|X_1(1+\|x_\sigma\|^{2m!})^{1/2m!}| \le C.$$

Thus

$$|X_1E(x)| \le CE(x) \quad \text{and} \quad |X_kE(x)| \le C2^{-\sigma}E(x) \quad \text{if } k=2,\dots,n_1.$$

Now since we are using left-invariant vector fields, they commute with left translations. It follows that

$$X_k(E(r^{-1}\cdot(2^j\cdot x))) = 2^j(X_kE)(r^{-1}\cdot(2^j\cdot x))$$

for all $k=1,\dots,n_1$, and using the above estimates for X_kE, one easily obtains the desired inequalities. □

Proposition 9.7.

$$|X_1\tilde{\omega}_j| \le C2^j\tilde{\omega}_j \quad \textit{and} \quad |X_k\tilde{\omega}_j| \le C2^{j-\sigma}\tilde{\omega}_j \quad \textit{for } k=2,\dots,n_1.$$

Proof. Note that $\tilde{\omega}_j$ can be written as

$$\tilde{\omega}_j = 2^{NQ}\int_G S_{j+N}|\Delta_j f|(2^{-j}\cdot y^{-1})E(y\cdot(2^j\cdot x))\,dy.$$

The proof is then almost identical to that of the previous proposition. □

Proposition 9.8.

$$\Big\|\sup_j 2^j\omega_j\Big\|_{L^Q} \le C_N 2^{\sigma(Q-1)/Q}\|\nabla_b f\|_{L^Q}.$$

Proof. This is because

$$\begin{aligned}\int_G \Big(\sup_j 2^j \omega_j\Big)^Q (x)\, dx &\le \sum_j \int_G (2^j \omega_j)^Q (x)\, dx \\ &\simeq \sum_j \sum_{r\in\Lambda} E(r^{-1})^Q \int_G [2^j S_{j+N} |\Delta_j f|(x\cdot 2^{-j} r)]^Q\, dx,\end{aligned}$$

the last line following from Proposition 9.2. Now by the translation invariance of the Lebesgue measure (which is the Haar measure on G), the integral in the last sum is independent of r. Furthermore,

$$\sum_{r\in\Lambda} E(r)^Q \simeq \sum_{r\in\Lambda} \int_{\theta\in[0,1)^N} E(r\cdot(2^{-N}\cdot\theta))^Q\, d\theta = 2^{NQ} \int_G E(y)^Q\, dy \le C 2^{NQ} 2^{\sigma(Q-1)};$$

here we have used Proposition 9.1 in the first inequality, the fact that every $y \in G$ can be written uniquely as $r\cdot(2^{-N}\cdot\theta)$ for some $r \in \Lambda$ and $\theta \in [0,1)^N$ in the middle identity, and the estimate $\|E^Q\|_{L^1} \le C2^{\sigma(Q-1)}$ in the last inequality. Altogether, this shows that

$$\begin{aligned}\int_G \Big(\sup_j 2^j \omega_j\Big)^Q (x)\, dx &\le C 2^{NQ} 2^{\sigma(Q-1)} \int_G \sum_j (2^j |\Delta_j f|(x))^Q\, dx \\ &\le C 2^{NQ} 2^{\sigma(Q-1)} \int_G \Big(\sum_j (2^j |\Delta_j f|(x))^2\Big)^{Q/2} dx \\ &\le C 2^{NQ} 2^{\sigma(Q-1)} \|\nabla_b f\|_{L^Q}^Q,\end{aligned}$$

the last inequality following from Proposition 5.4. □

10. Estimating $h - \tilde{h}$

In this section we estimate $h - \tilde{h}$. First, we recall our construction: we have $h = \sum_{j=-\infty}^{\infty} h_j$, where

$$h_j(x) = (1-\zeta_j(x)) S_{j+N}(\Delta_j f)(x).$$

We also have

$$\tilde{h} = \sum_{j=-\infty}^{\infty} h_j \prod_{j'>j} (1 - U_{j'}) \quad \text{where} \quad U_j = (1-\zeta_j)\omega_j.$$

We will estimate $\tilde{h}$ following our paradigm of approximation in Section 6. By Propositions 9.4 and 9.5, we have

$$C^{-1}|h_j| \le U_j \le 1.$$

It follows from Proposition 6.2 that $\|\tilde{h}\|_{L^\infty} \le C$, proving (7.5).

Next, following the derivation of (6.9), we have

$$|X_k(h-\tilde{h})| \leq C \sum_{j=-\infty}^{\infty} |X_k U_j| + \sum_{j=-\infty}^{\infty} |U_j| \sum_{j'<j} (\|\nabla_b h_{j'}\|_{L^\infty} + \|\nabla_b U_{j'}\|_{L^\infty})$$

for $k = 1, \dots, n_1$. But

$$|U_j(x)| \leq \omega_j(x) \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k\}}(x)$$

where χ denotes the characteristic function of a set. This is because $1-\zeta_j(x) = 0$ unless $2^j\omega_j(x) > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k(x)$. Next, for $k = 2, \dots, n_1$,

$$|X_k U_j(x)| \leq C2^{j-\sigma}\omega_j(x) \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k\}}(x)$$

because $|X_k\omega_j| \leq C2^{j-\sigma}\omega_j$ by Proposition 9.6, and

$$|X_k\zeta_j| \leq C2^{j-\sigma} \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k\}}.$$

(The last inequality follows by differentiating the definition of ζ_j, and using $|X_k\omega_j| \leq C2^{j-\sigma}\omega_j$ again.) Similarly,

$$|X_1 U_j(x)| \leq C2^j\omega_j(x) \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k\}}(x).$$

Finally, we have

$$\|\nabla_b h_j\|_{L^\infty} + \|\nabla_b U_j\|_{L^\infty} \leq C_N 2^{\sigma(Q-1)} 2^j \|\nabla_b f\|_{L^Q}.$$

(The estimate on $\nabla_b U_j$ follows from the above discussion and Proposition 9.5, while the estimate on $\nabla_b h_j$ is similar.) So altogether, for $k = 2, \dots, n_1$, we have

$$|X_k(h-\tilde{h})(x)| \leq C2^{-\sigma}\mathfrak{S}(x) + C_N 2^{\sigma(Q-1)}\mathfrak{S}(x)\|\nabla_b f\|_{L^Q},$$

where

$$\mathfrak{S}(x) := \sum_{j=-\infty}^{\infty} 2^j\omega_j(x) \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv j \,(\mathrm{mod}\, R)} 2^k\omega_k\}}(x).$$

Similarly,

$$|X_1(h-\tilde{h})(x)| \leq C_N 2^{\sigma(Q-1)}\mathfrak{S}(x).$$

To proceed further, we estimate the L^Q norm of the sum $\mathfrak{S}(x)$; this sum can be rewritten as

$$\sum_{c=0}^{R-1} \sum_{j\equiv c \,(\mathrm{mod}\, R)} 2^j\omega_j(x) \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv c \,(\mathrm{mod}\, R)} 2^k\omega_k\}}(x).$$

For each fixed c, we have

$$\left| \sum_{j\equiv c \,(\mathrm{mod}\, R)} 2^j\omega_j \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv c \,(\mathrm{mod},\ R)} 2^k\omega_k\}} \right| \leq 3\sup_j 2^j\omega_j. \tag{10.1}$$

This is true because for a fixed x, for any $N > 0$, one can pick the greatest integer $j_0 \leq N$, $j_0 \equiv c \pmod{R}$ such that $2^{j_0}\omega_{j_0} > (1/2)\sum_{k<j_0,\, k\equiv c \pmod{R}} 2^k\omega_k$. Then

$$\begin{aligned}\sum_{j\leq N,\, j\equiv c \pmod{R}} & 2^j\omega_j \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv c \pmod{R}} 2^k\omega_k\}} \\ &= \sum_{j\leq j_0,\, j\equiv c \pmod{R}} 2^j\omega_j \chi_{\{2^j\omega_j > (1/2)\sum_{k<j,\, k\equiv c \pmod{R}} 2^k\omega_k\}} \\ &\leq 2^{j_0}\omega_{j_0} + \sum_{j<j_0,\, j\equiv c \pmod{R}} 2^j\omega_j \leq 3\cdot 2^{j_0}\omega_{j_0} \leq 3\sup_j 2^j\omega_j.\end{aligned}$$

Letting $N \to \infty$ we get the inequality (10.1). Hence

$$\mathfrak{S}(x) \leq 3R\sup_j 2^j\omega_j(x),$$

and from Proposition 9.8 we conclude that

$$\|\mathfrak{S}\|_{L^Q} \leq CR2^{\sigma(Q-1)/Q}\|\nabla_b f\|_{L^Q}.$$

Altogether, for $k = 2, \ldots, n_1$, we have

$$\|X_k(h-\tilde{h})\|_{L^Q} \leq C_N 2^{-\sigma/Q}R\|\nabla_b f\|_{L^Q} + C_N 2^{\sigma Q}R\|\nabla_b f\|_{L^Q}^2, \tag{10.2}$$

and this proves (7.6).

Using the pointwise bound of $X_1(h-\tilde{h})$, and applying the same method as in (10.2), one can prove

$$\|X_1(h-\tilde{h})\|_{L^Q} \leq C_N 2^{\sigma Q}R\|\nabla_b f\|_{L^Q},$$

completing our proof of (7.7).

11. Estimating $g - \tilde{g}$

In this section we estimate $g - \tilde{g}$. Again we recall our construction: we have $g = \sum_{j=-\infty}^{\infty} g_j$, where

$$g_j(x) = \zeta_j(x)S_{j+N}(\Delta_j f)(x),$$

and

$$\tilde{g} = \sum_{c=0}^{R-1}\sum_{j\equiv c \pmod{R}} g_j \prod_{\substack{j'>j \\ j'\equiv c \pmod{R}}}(1-G_{j'}) \quad \text{where} \quad G_j = \sum_{\substack{t>0 \\ t\equiv 0 \pmod{R}}} 2^{-t}\tilde{\omega}_{j-t}.$$

Now by Proposition 9.4,

$$C^{-1}|g_j| \leq \omega_j\zeta_j.$$

But then

$$C^{-1}\omega_j\zeta_j \leq G_j \leq 1.$$

In fact, on the support of ζ_j,

$$\omega_j \le \sum_{\substack{k<j \\ k\equiv j \ (\text{mod } R)}} 2^{k-j}\omega_k = \sum_{\substack{t>0 \\ t\equiv 0 \ (\text{mod } R)}} 2^{-t}\omega_{j-t} \le C G_j,$$

and the first inequality follows. The last inequality comes from Proposition 9.5. Thus

$$C^{-1}|g_j| \le G_j \le 1,$$

and from Proposition 6.2, we have $|\tilde{g}| \le CR$. This proves (7.8).

Next

$$\begin{aligned} g - \tilde{g} &= \sum_{c=0}^{R-1} \sum_{j\equiv c \ (\text{mod } R)} g_j \Big(1 - \prod_{\substack{j'>j \\ j'\equiv c \ (\text{mod } R)}} (1 - G_{j'})\Big) \\ &= \sum_{c=0}^{R-1} \sum_{j\equiv c \ (\text{mod } R)} G_j H_j = \sum_{j=-\infty}^{\infty} G_j H_j \end{aligned}$$

where

$$H_j := \sum_{\substack{j'<j \\ j'\equiv j \ (\text{mod } R)}} g_{j'} \prod_{\substack{j'<j''<j \\ j''\equiv j \ (\text{mod } R)}} (1 - G_{j''}). \tag{11.1}$$

By Proposition 6.2, an immediate estimate of H_j is

$$|H_j| \le \sum_{j'<j} |g_{j'}| \prod_{j'<j''<j} (1 - G_{j''}) \le C \sum_{j'<j} G_{j'} \prod_{j'<j''<j} (1 - G_{j''}) \le C.$$

We now collect some estimates for $\nabla_b g_j$, $\nabla_b G_j$ and $\nabla_b H_j$. To begin with, we have

Proposition 11.1.

$$\tilde{\omega}_j \le C_N 2^{\sigma(Q+1)} MM(\Delta_j f)$$

where M is the maximal function defined before Proposition 5.3.

Proof. We have

$$E(x) \le C(1 + \|x_\sigma\|)^{-(Q+1)} \le C 2^{\sigma(Q+1)} (1 + \|x\|)^{-(Q+1)}$$

and the latter is an integrable radially decreasing function. Thus

$$\tilde{\omega}_j = 2^{NQ} E_j(S_{j+N}|\Delta_j f|) \le C 2^{NQ} 2^{\sigma(Q+1)} MM(\Delta_j f). \qquad \square$$

Proposition 11.2.

$$|\nabla_b G_j| \le C_N 2^{\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\text{mod } R)}} 2^{-t} 2^{j-t} MM(\Delta_{j-t} f).$$

Proof. One differentiates the definition of G_j and estimates the derivatives of $\tilde{\omega}_j$ using Propositions 11.1 and 9.7. □

Proposition 11.3.

$$|\nabla_b g_j| \le C_N 2^j M(\Delta_j f).$$

Proof. One differentiates $g_j(x) = \zeta_j(x) S_{j+N}(\Delta_j f)(x)$, letting the derivative hit either ζ_j or S_{j+N}, and estimates the rest by the maximal function. The worst term is when the derivative hits S_{j+N}, which gives a factor of 2^{j+N}. □

Proposition 11.4.

$$|\nabla_b H_j| \le C_N 2^{\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{j-t} MM(\Delta_{j-t} f).$$

Proof. Following the derivation of (6.8) from (6.5) in Section 6, and using the definition of H_j in (11.1), we have

$$\nabla_b H_j = \sum_{\substack{j'<j \\ j'\equiv j \ (\mathrm{mod}\ R)}} (\nabla_b g_{j'} - (\nabla_b G_{j'}) H_{j'}) \prod_{\substack{j'<j''<j \\ j'\equiv j \ (\mathrm{mod}\ R)}} (1 - G_{j''}), \tag{11.2}$$

so

$$|\nabla_b H_j| \le C \sum_{\substack{j'<j \\ j'\equiv j \ (\mathrm{mod}\ R)}} (|\nabla_b g_{j'}| + |\nabla_b G_{j'}|).$$

This, with Propositions 11.2 and 11.3, leads to

$$C_N 2^{\sigma(Q+1)} \sum_{\substack{l>0 \\ l\equiv 0 \ (\mathrm{mod}\ R)}} \sum_{\substack{t\ge 0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{j-l-t} MM(\Delta_{j-l-t} f).$$

Rearranging gives the desired bound. □

The proofs of the next two estimates are the same as those in Propositions 11.3 and 11.2, except that one differentiates once more.

Proposition 11.5.

$$|\nabla_b^2 g_j| \le C_N 2^{2j} M(\Delta_j f).$$

Proposition 11.6.

$$|\nabla_b^2 G_j| \le C_N 2^{\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{2(j-t)} MM(\Delta_{j-t} f).$$

Finally we estimate second derivatives of H_j:

Proposition 11.7.

$$|\nabla_b^2 H_j| \le C_N 2^{2\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{2(j-t)} MM(\Delta_{j-t} f).$$

Proof. Differentiating (11.2) once more, again using the way we derived (6.8) from (6.5), we get

$$\nabla_b^2 H_j = \sum_{\substack{j'<j \\ j'\equiv j \ (\mathrm{mod}\ R)}} \big(\nabla_b[\nabla_b g_{j'} - (\nabla_b G_{j'})H_{j'}] - (\nabla_b G_{j'})(\nabla_b H_{j'})\big) \prod_{\substack{j'<j''<j \\ j'\equiv j \ (\mathrm{mod}\ R)}} (1 - G_{j''}).$$

Thus $|G_j| \le 1$ and $|H_j| \le C$ imply that

$$|\nabla_b^2 H_j| \le C \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} (|\nabla_b^2 g_{j-t}| + |\nabla_b^2 G_{j-t}| + |\nabla_b G_{j-t}|\,|\nabla_b H_{j-t}|).$$

The first two terms can be estimated using Propositions 11.5 and 11.6. For the last term, Propositions 11.2 and 11.4 give

$$\begin{aligned} &|\nabla_b G_j|\,|\nabla_b H_j| \\ &\quad \le C_N 2^{2\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} \sum_{\substack{l>0 \\ l\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{2j-t-l} (MM\Delta_{j-t} f)(MM\Delta_{j-l} f). \end{aligned}$$

Now we split the sum into two parts: one where $t > l$, and the other where $l \ge t$, and use $\|\Delta_j f\|_{L^\infty} \le C\|\nabla_b f\|_{L^Q} \le 1$. In the first sum, we estimate $MM\Delta_{j-t} f$ by a constant; this is possible because $MM\Delta_{j-t} f$ is bounded by $\|\Delta_{j-t} f\|_{L^\infty}$, which is bounded by a constant by the Bernstein inequality (Proposition 5.5) and our assumption (7.2). We then sum over t to get a bound

$$C \sum_{\substack{l>0 \\ l\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-l} 2^{2(j-l)} MM\Delta_{j-l} f.$$

In the second sum, we estimate $MM\Delta_{j-l} f$ by a constant instead, and sum over l to get a bound

$$C \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{2(j-t)} MM\Delta_{j-t} f.$$

These two bounds are identical. So

$$\begin{aligned} &\sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} |\nabla_b G_{j-t}|\,|\nabla_b H_{j-t}| \\ &\qquad \le C_N 2^{2\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} \sum_{\substack{l>0 \\ l\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-l} 2^{2(j-t-l)} MM\Delta_{j-t-l} f. \end{aligned}$$

Rearranging we get the desired bound. □

Now we will estimate

$$\Big\| \sum_{s=-\infty}^{\infty} \nabla_b (G_s H_s) \Big\|_{L^Q}.$$

With the use of the Littlewood–Paley projections $\Lambda_j^{(l)}$, the above is bounded by

$$\begin{aligned}
&\Big\| \Big(\sum_{j=-\infty}^{\infty} \Big(\sum_{s=-\infty}^{\infty} |\Lambda_j \nabla_b (G_s H_s)| \Big)^2 \Big)^{1/2} \Big\|_{L^Q} \\
&\qquad = \Big\| \Big(\sum_{j=-\infty}^{\infty} \Big(\sum_{s=-\infty}^{\infty} |\Lambda_j \nabla_b (G_{j-s} H_{j-s})| \Big)^2 \Big)^{1/2} \Big\|_{L^Q} \\
&\qquad \leq \sum_{s=-\infty}^{\infty} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_j \nabla_b (G_{j-s} H_{j-s})|^2 \Big)^{1/2} \Big\|_{L^Q} \\
&\qquad = \sum_{s=-\infty}^{\infty} \Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_{j+s} \nabla_b (G_j H_j)|^2 \Big)^{1/2} \Big\|_{L^Q}.
\end{aligned} \tag{11.3}$$

We split the sum into two parts: $\sum_{s \leq R}$ and $\sum_{s > R}$. We shall pick up a convergence factor $2^{-|s|}$ or $|s| 2^{-|s|}$ for each term so that we can sum over s.

To estimate the first sum, we fix $s \leq R$. Then for each $j \in \mathbb{Z}$, we split G_j into a sum

$$G_j = G_j^{(1)} + G_j^{(2)},$$

where

$$G_j^{(1)} = \sum_{\substack{0 < t < |s| \\ t \equiv 0 \, (\mathrm{mod}\, R)}} 2^{-t} \tilde{\omega}_{j-t}, \quad G_j^{(2)} = \sum_{\substack{t \geq \max\{|s|, R\} \\ t \equiv 0 \, (\mathrm{mod}\, R)}} 2^{-t} \tilde{\omega}_{j-t}.$$

Note that the splitting of G_j depends on s; in particular, if $-R \leq s \leq R$, then $G_j^{(1)} = 0$ and $G_j^{(2)} = G_j$.

Now we estimate

$$\Big\| \Big(\sum_{j=-\infty}^{\infty} |\Lambda_{j+s} \nabla_b (G_j^{(1)} H_j)|^2 \Big)^{1/2} \Big\|_{L^Q}.$$

We have

$$\Lambda_{j+s} (\nabla_b (G_j^{(1)} H_j)) = (\nabla_b (G_j^{(1)} H_j)) * \Lambda_{j+s} = 2^{j+s} (G_j^{(1)} H_j) * (\nabla_b^R \Lambda)_{j+s}$$

by the compatibility of convolution with the left- and right-invariant derivatives. Hence from $|H_j| \leq C$, and

$$|G_j^{(1)}| \leq C_N 2^{\sigma(Q+1)} \sum_{\substack{0 < t < |s| \\ t \equiv 0 \, (\mathrm{mod}\, R)}} 2^{-t} MM(\Delta_{j-t} f),$$

which follows from Proposition 11.1, we have

$$|\Lambda_{j+s}(\nabla_b(G_j^{(1)}H_j))| \leq C_N 2^{\sigma(Q+1)}2^s \sum_{\substack{0<t<|s| \\ t\equiv 0 \,(\mathrm{mod}\, R)}} 2^{j-t} MMM(\Delta_{j-t}f).$$

Taking the square function in j and then the L^Q norm in space, we find that, when $s < -R$,

$$\begin{aligned}\Big\|\Big(\sum_{j=-\infty}^{\infty} |\Lambda_{j+s}\nabla_b(G_j^{(1)}H_j)|^2\Big)^{1/2}\Big\|_{L^Q} &\leq C_N 2^{\sigma(Q+1)}\frac{|s|}{R}2^s \Big\|\Big(\sum_{j=-\infty}^{\infty}(2^j|\Delta_j f|)^2\Big)^{1/2}\Big\|_{L^Q} \\ &\leq C_N 2^{\sigma(Q+1)}\frac{|s|}{R}2^s\|\nabla_b f\|_{L^Q}. \qquad (11.4)\end{aligned}$$

Here the last inequality follows from Proposition 5.4. The same norm on the left hand side above is of course zero when $-R \leq s \leq R$.

Next, we estimate

$$\Big\|\Big(\sum_{j=-\infty}^{\infty} |\Lambda_{j+s}\nabla_b(G_j^{(2)}H_j)|^2\Big)^{1/2}\Big\|_{L^Q} \leq C\Big\|\Big(\sum_{j=-\infty}^{\infty} |\nabla_b(G_j^{(2)}H_j)|^2\Big)^{1/2}\Big\|_{L^Q}.$$

Now since $|H_j| \leq C$,

$$|\nabla_b(G_j^{(2)}H_j)| \leq C(|\nabla_b G_j^{(2)}| + |G_j^{(2)}||\nabla_b H_j|).$$

We know

$$|G_j^{(2)}| \leq C_N 2^{\sigma(Q+1)} \sum_{\substack{t\geq \max\{|s|,R\} \\ t\equiv 0 \,(\mathrm{mod}\, R)}} 2^{-t} MM(\Delta_{j-t}f) \qquad (11.5)$$

by Proposition 11.1, and $|\nabla_b G_j^{(2)}|$ is bounded by

$$C \sum_{\substack{t\geq \max\{|s|,R\} \\ t\equiv 0 \,(\mathrm{mod}\, R)}} 2^{-t}2^{j-t}\tilde{\omega}_{j-t}$$

by Proposition 11.2. Therefore by Proposition 11.1 again, we have

$$|\nabla_b G_j^{(2)}| \leq C_N 2^{\sigma(Q+1)} \sum_{\substack{t\geq \max\{|s|,R\} \\ t\equiv 0 \,(\mathrm{mod}\, R)}} 2^{-t}2^{j-t} MM(\Delta_{j-t}f).$$

Hence

$$\begin{aligned}\Big\|\Big(\sum_{j=-\infty}^{\infty} |\nabla_b G_j^{(2)}|^2\Big)^{1/2}\Big\|_{L^Q} &\leq C_N 2^{\sigma(Q+1)} \sum_{\substack{t\geq \max\{|s|,R\} \\ t\equiv 0 \,(\mathrm{mod}\, R)}} 2^{-t}\Big\|\Big(\sum_{j=-\infty}^{\infty}|2^j\Delta_j f|^2\Big)^{1/2}\Big\|_{L^Q} \\ &\leq C_N 2^{\sigma(Q+1)}2^{-\max\{|s|,R\}}\|\nabla_b f\|_{L^Q}.\end{aligned}$$

Furthermore, by (11.5) and Proposition 11.4, one can estimate

$$|G_j^{(2)}|\,|\nabla_b H_j| \le C_N 2^{2\sigma(Q+1)} \sum_{\substack{t\ge\max\{|s|,R\}\\ t\equiv 0\,(\mathrm{mod}\,R)}} \sum_{\substack{m>0\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{j-t-m} MM(\Delta_{j-t}f)MM(\Delta_{j-m}f).$$

We split this sum into the sum over three regions of t and m: where $t\ge\max\{|s|,R\}$ and $m>t$; where $t\ge\max\{|s|,R\}$ and $t\ge m\ge\max\{|s|,R\}$, which is equivalent to $m\ge\max\{|s|,R\}$ and $t\ge m$; and where $0<|m|<\max\{|s|,R\}$ and $t\ge\max\{|s|,R\}$. The first two sums are basically the same; each can be bounded by

$$\sum_{\substack{m\ge\max\{|s|,R\}\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{j-m} MM(\Delta_{j-m}f)\sum_{t\ge m} 2^{-t} MM(\Delta_{j-t}f),$$

which is bounded by

$$\sum_{\substack{m\ge\max\{|s|,R\}\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{-m}2^{j-m} MM(\Delta_{j-m}f),$$

since we can bound $MM\Delta_{j-t}f$ by a constant (cf. proof of Proposition 11.7) and sum over t. The last sum is bounded by

$$C2^{-\max\{|s|,R\}} \sum_{\substack{0<m<\max\{|s|,R\}\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{j-m} MM(\Delta_{j-m}f)$$

for the same reason. Thus

$$|G_j^{(2)}|\,|\nabla_b H_j| \le C_N 2^{2\sigma(Q+1)}\Big(2\sum_{\substack{m\ge\max\{|s|,R\}\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{-m}2^{j-m} MM(\Delta_{j-m}f) + 2^{-\max\{|s|,R\}} \sum_{\substack{0<m<\max\{|s|,R\}\\ m\equiv 0\,(\mathrm{mod}\,R)}} 2^{j-m} MM(\Delta_{j-m}f)\Big).$$

Taking the l^2 norm in j and then the L^Q norm in space, we get

$$\Big\|\Big(\sum_{j=-\infty}^{\infty} |G_j^{(2)}(\nabla_b H_j)|^2\Big)^{1/2}\Big\|_{L^Q} \le C_N 2^{2\sigma(Q+1)} \max\{|s|,R\} 2^{-\max\{|s|,R\}} \|\nabla_b f\|_{L^Q},$$

It follows that

$$\Big\|\Big(\sum_{j=-\infty}^{\infty} |\Lambda_{j+s}\nabla_b(G_j^{(2)}H_j)|^2\Big)^{1/2}\Big\|_{L^Q} \le C_N 2^{2\sigma(Q+1)} \max\{|s|,R\} 2^{-\max\{|s|,R\}} \|\nabla_b f\|_{L^Q}. \tag{11.6}$$

Summing (11.4) and (11.6) over $s \leq R$, we get a bound

$$C_N 2^{2\sigma(Q+1)} R^2 2^{-R} \|\nabla_b f\|_{L^Q}$$

for the first half of the sum in (11.3).

Next we look at the second half, corresponding to the sum over all $s > R$. First,

$$|\Lambda_{j+s} \nabla_b (G_j H_j)| \leq |\Lambda_{j+s}((\nabla_b G_j) H_j)| + |\Lambda_{j+s}(G_j(\nabla_b H_j))|. \tag{11.7}$$

The first term can be written as

$$\int_G ((\nabla_b G_j)(x \cdot y^{-1}) - (\nabla_b G_j)(x)) H_j(x \cdot y^{-1}) \Lambda_{j+s}(y)\, dy + (\nabla_b G_j)(x)(\Lambda_{j+s} H_j)(x)$$
$$= I + II.$$

The second term in (11.7) can be written as

$$\int_G (G_j(x \cdot y^{-1}) - G_j(x))(\nabla_b H_j)(x \cdot y^{-1}) \Lambda_{j+s}(y)\, dy + G_j(x) \Lambda_{j+s}(\nabla_b H_j)(x)$$
$$= III + IV.$$

We estimate I, II, III, IV separately.

First, in I, we bound $|H_j| \leq C$, and write

$$(\nabla_b G_j)(x \cdot y^{-1}) - (\nabla_b G_j)(x) = 2^{NQ} \sum_{\substack{t>0 \\ t \equiv 0 \,(\mathrm{mod}\, R)}} 2^{-t} 2^{j-t}$$
$$\times \int_G S_{j+N-t} |\Delta_{j-t} f|(x \cdot z^{-1}) \big((\nabla_b E)_{j-t}(z \cdot y^{-1}) - (\nabla_b E)_{j-t}(z)\big)\, dz.$$

We put this back in I, and thus need to bound

$$\int_G |(\nabla_b E)_{j-t}(z \cdot y^{-1}) - (\nabla_b E)_{j-t}(z)|\, |\Lambda_{j+s}(y)|\, dy. \tag{11.8}$$

But

$$|\nabla_b^2 E(x)| \leq C E(x) \leq \frac{C}{(1 + 2^{-\sigma} \|x\|)^K}$$
$$\leq C 2^{\sigma K} \frac{1}{(1 + \|x\|)^K}$$

for all positive integers K. We will use this estimate with $K = 2(Q+1)$, and apply the remark after Proposition 3.3; the integral (11.8) is then bounded by

$$C 2^{2\sigma(Q+1)} 2^{-s-t} 2^{(j-t)Q} (1 + 2^{j-t} \|z\|)^{-(Q+1)}.$$

Hence

$$|I| \leq C_N 2^{2\sigma(Q+1)} 2^{-s} \sum_{\substack{t>0 \\ t \equiv 0 \,(\mathrm{mod}\, R)}} 2^{-2t} 2^{j-t} MM(\Delta_{j-t} f)(x).$$

Taking the square function in j and the L^Q norm in space, we get a bound

$$C_N 2^{2\sigma(Q+1)} 2^{-s} \|\nabla_b f\|_{L^Q}.$$

For *II*, recall the pointwise bound for $\nabla_b G_j$ from Proposition 11.2:

$$|\nabla_b G_j| \leq C_N 2^{\sigma(Q+1)} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{j-t} MM(\Delta_{j-t} f).$$

To estimate $\Lambda_{j+s} H_j$, we use of Proposition 4.3(a), and write (schematically) Λ as $\nabla_b^R \cdot \Phi$ where Φ is a ($2n$ tuple of) Schwartz function, and integrate by parts. Then

$$\begin{aligned} |\Lambda_{j+s} H_j| &= 2^{-j-s} |(\nabla_b H_j) * \Phi_{j+s}| \leq 2^{-j-s} \|\nabla_b H_j\|_{L^\infty} \\ &\leq C_N 2^{\sigma(Q+1)} 2^{-s}, \end{aligned}$$

since $\|\nabla_b H_j\|_{L^\infty} \leq C_N 2^{\sigma(Q+1)} 2^j$. Hence

$$|II| \leq C_N 2^{2\sigma(Q+1)} 2^{-s} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} 2^{j-t} MM(\Delta_{j-t} f).$$

Taking the square function in j and the L^Q norm in space, we get a contribution

$$C_N 2^{2\sigma(Q+1)} 2^{-s} \|\nabla_b f\|_{L^Q}.$$

Now to bound *III*, we follow our strategy as in I. First we bound

$$|\nabla_b H_j| \leq C_N 2^{\sigma(Q+1)} 2^j$$

by Proposition 11.4, and write

$$\begin{aligned} &G_j(x \cdot y^{-1}) - G_j(x) \\ &\quad = 2^{NQ} \sum_{\substack{t>0 \\ t\equiv 0 \ (\mathrm{mod}\ R)}} 2^{-t} \int_G S_{j+N-t} |\Delta_{j-t} f|(x \cdot z^{-1})(E_{j-t}(z \cdot y^{-1}) - E_{j-t}(z))\, dz. \end{aligned}$$

We put this back in *III*, and thus need to bound

$$\int_G |E_{j-t}(z \cdot y^{-1}) - E_{j-t}(z)|\, |\Lambda_{j+s}(y)|\, dy. \tag{11.9}$$

But

$$\begin{aligned} |\nabla_b E(x)| \leq C E(x) &\leq \frac{C}{(1 + 2^{-\sigma}\|x\|)^K} \\ &\leq C 2^{\sigma K} \frac{1}{(1 + \|x\|)^K} \end{aligned}$$

for all positive integers K. We will take $K = 2(Q+1)$, and apply the remark after Proposition 3.3; the integral (11.9) is then bounded by

$$C2^{\sigma(Q+1)}2^{-s-t}2^{(j-t)Q}(1+2^{j-t}\|z\|)^{-(Q+1)}.$$

Hence

$$|III| \leq C_N 2^{2\sigma(Q+1)}2^{-s} \sum_{\substack{t>0 \\ t\equiv 0 \,(\text{mod}\, R)}} 2^{-t}2^{j-t}MM(\Delta_{j-t}f)(x).$$

Taking the square function in j and the L^Q norm in space, we get a bound

$$C_N 2^{2\sigma(Q+1)}2^{-s}\|\nabla_b f\|_{L^Q}.$$

Finally, to estimate IV, we recall that $|G_j| \leq 1$, as was shown at the beginning of this section. Furthermore,

$$\begin{aligned} |(\Lambda_{j+s}(\nabla_b H_j))(x)| &\leq |(\nabla_b H_j) * (\nabla_b^R \Phi)_{j+s}(x)| \\ &= 2^{-j-s}|(\nabla_b^2 H_j) * \Phi_{j+s}(x)| \\ &\leq 2^{-j-s}M(\nabla_b^2 H_j)(x). \end{aligned}$$

By Proposition 11.7, this is bounded by

$$C_N 2^{2\sigma(Q+1)}2^{-s} \sum_{\substack{t>0 \\ t\equiv 0 \,(\text{mod}\, R)}} 2^{-t}2^{j-t}MMM(\Delta_{j-t}f)(x).$$

Hence

$$|IV| \leq C_N 2^{2\sigma(Q+1)}2^{-s} \sum_{\substack{t>0 \\ t\equiv 0 \,(\text{mod}\, R)}} 2^{-t}2^{j-t}MMM(\Delta_{j-t}f)(x).$$

Taking the square function in j and then the L^Q norm in space, this is bounded by

$$C_N 2^{2\sigma(Q+1)}2^{-s}\|\nabla_b f\|_{L^Q}.$$

Hence

$$\sum_{s>R}\left\|\left(\sum_{j=-\infty}^{\infty}|\Lambda_{j+s}\nabla_b(G_jH_j)|^2\right)^{1/2}\right\|_{L^Q} \leq C_N 2^{2\sigma(Q+1)}2^{-R}\|\nabla_b f\|_{L^Q}.$$

Altogether, (11.3) is bounded by

$$C_N 2^{2\sigma(Q+1)}R^2 2^{-R}\|\nabla_b f\|_{L^Q}.$$

This proves our claim (7.9), and marks the end of the proof of our approximation Lemma 1.7.

12. Proof of Theorems 1.8 and 1.9

In this section we prove Theorems 1.8 and 1.9. We first recall the $\overline{\partial}_b$ complex on the Heisenberg group $\mathbb{H}^n$.

First, $\mathbb{H}^n$ is a simply connected Lie group diffeomorphic to $\mathbb{R}^{2n+1}$. We write $[x, y, t]$ for a point on $\mathbb{R}^{2n+1}$, where $x, y \in \mathbb{R}^n$ and $t \in \mathbb{R}$. The group law on the Heisenberg group is given by

$$[x, y, t] \cdot [u, v, w] = [x + u, y + v, t + w + 2(yu - xv)],$$

where yu is the dot product of y and u in $\mathbb{R}^n$. The left-invariant vector fields of order 1 on $\mathbb{H}^n$ are then linear combinations of the vector fields $X_1, \dots, X_{2n}$, where

$$X_k = \frac{\partial}{\partial x_k} + 2y_k \frac{\partial}{\partial t} \quad \text{and} \quad X_{k+n} = \frac{\partial}{\partial y_k} - 2x_k \frac{\partial}{\partial t} \quad \text{for } k = 1, \dots, n.$$

Thus in this case, n_1 is equal to $2n$, and

$$\nabla_b f = (X_1 f, \dots, X_{2n} f).$$

The one-parameter family of automorphic dilations on $\mathbb{H}^n$ is given by

$$\lambda \cdot [x, y, t] = [\lambda x, \lambda y, \lambda^2 t] \quad \text{for all } \lambda > 0.$$

The homogeneous dimension in this case is $Q = 2n + 2$.

Now let

$$Z_k = \tfrac{1}{2}(X_k - iX_{k+n}) \quad \text{and} \quad \overline{Z}_k = \tfrac{1}{2}(X_k + iX_{k+n}), \qquad k = 1, \dots, n.$$

For $0 \le q \le n$, the $(0, q)$ forms on the Heisenberg group $\mathbb{H}^n$ are expressions of the form

$$\sum_{|\alpha|=q} u_\alpha d\overline{z}^\alpha,$$

where the sum is over all strictly increasing multi-indices $\alpha = (\alpha_1, \dots, \alpha_q)$ of length q with entries in $\{1, \dots, n\}$; in other words, $\alpha_k \in \{1, \dots, n\}$ and $\alpha_1 < \dots < \alpha_q$. Moreover, $d\overline{z}^\alpha$ is a shorthand for $d\overline{z}_{\alpha_1} \wedge \dots \wedge d\overline{z}_{\alpha_q}$, and each u_α is a smooth function on $\mathbb{H}^n$. The $\overline{\partial}_b$ complex is then defined by

$$\overline{\partial}_b u = \sum_{k=1}^{2n} \sum_{|\alpha|=q} \overline{Z}_k(u_\alpha) d\overline{z}_k \wedge d\overline{z}^\alpha \quad \text{if } u = \sum_{|\alpha|=q} u_\alpha d\overline{z}^\alpha.$$

By making the above $d\overline{z}^\alpha$ an orthonormal basis for $(0, q)$ forms at every point, one gets a Hermitian inner product on $(0, q)$ forms at every point on $\mathbb{H}^n$, with which one can define an inner product on the space of $(0, q)$ forms on $\mathbb{H}^n$ that have L^2 coefficients. One can then consider the adjoint of $\overline{\partial}_b$ with respect to this inner product, namely

$$\overline{\partial}_b^* u = \sum_{|\alpha|=q} \sum_{k\in\alpha} -Z_k(u_\alpha) d\overline{z}_k \lrcorner\, d\overline{z}^\alpha;$$

here the interior product $\lrcorner$ is just the usual one on $\mathbb{R}^{2n+1}$.

Proof of Theorem 1.8. The key idea is that when one computes $\overline{\partial}_b^*$ of a $(0, q+1)$ form on $\mathbb{H}^n$, only $2(q+1)$ of the $2n$ real left-invariant derivatives of order 1 are involved. So if $q+1 < n$, then for each component of the q form, there will be some real left-invariant derivatives of degree 1 that are irrelevant in computing $\overline{\partial}_b^*$, and we can give up estimates in those directions when we apply Lemma 1.7.

We will use the bounded inverse theorem and an argument closely related to the usual proof of the open mapping theorem.

Let $N\dot{L}^{1,Q}(\Lambda^{(0,q+1)})$ be the space of $(0, q+1)$ forms on $\mathbb{H}^n$ with $N\dot{L}^{1,Q}$ coefficients, and similarly define $L^Q(\Lambda^{(0,q)})$. Consider the map

$$\overline{\partial}_b^*\colon N\dot{L}^{1,Q}(\Lambda^{(0,q+1)}) \to L^Q(\Lambda^{(0,q)}).$$

It is bounded and has closed range. Hence it induces a bounded linear bijection between the Banach spaces $N\dot{L}^{1,Q}(\Lambda^{(0,q+1)})/\ker(\overline{\partial}_b^*)$ and $\mathrm{Image}(\overline{\partial}_b^*) \subseteq L^Q(\Lambda^{(0,q)})$. By the bounded inverse theorem, this map has a bounded inverse; hence for any $(0, q)$ form $f \in \mathrm{Image}(\overline{\partial}_b^*) \subseteq L^Q(\Lambda^{(0,q)})$, there exists $\alpha^{(0)} \in N\dot{L}^{1,Q}(\Lambda^{(0,q+1)})$ such that

$$\begin{cases} \overline{\partial}_b^* \alpha^{(0)} = f, \\ \|\nabla_b \alpha^{(0)}\|_{L^Q} \le C\|f\|_{L^Q}. \end{cases}$$

Now for $q < n-1$, if I is a multi-index of length $q+1$, then one can pick $i \notin I$ and approximate $\alpha_I^{(0)}$ by Lemma 1.7 in all but the X_i direction; more precisely, for any $\delta > 0$, there exists $\beta_I^{(0)} \in N\dot{L}^{1,Q} \cap L^\infty$ such that

$$\sum_{j\neq i} \|X_j(\alpha_I^{(0)} - \beta_I^{(0)})\|_{L^Q} \le \delta \|\nabla_b \alpha_I^{(0)}\|_{L^Q} \le C\delta\|f\|_{L^Q}$$

and

$$\|\beta_I^{(0)}\|_{L^\infty} + \|\nabla_b \beta_I^{(0)}\|_{L^Q} \le A_\delta \|\nabla_b \alpha_I^{(0)}\|_{L^Q} \le C A_\delta \|f\|_{L^Q}.$$

Then if δ is picked so that $C\delta \le \frac{1}{2}$, we have

$$\beta^{(0)} := \sum_I \beta_I^{(0)} d\overline{z}^I \in N\dot{L}^{1,Q} \cap L^\infty(\Lambda^{(0,q+1)})$$

satisfying

$$\begin{cases} \|f - \overline{\partial}_b^* \beta^{(0)}\|_{L^Q} \le \dfrac{1}{2}\|f\|_{L^Q}, \\ \|\beta^{(0)}\|_{L^\infty} + \|\nabla_b \beta^{(0)}\|_{L^n} \le A\|f\|_{L^Q} \end{cases}$$

(the first inequality holds because $\|f - \overline{\partial}_b^* \beta^{(0)}\|_{L^Q} = \|\overline{\partial}_b^*(\alpha^{(0)} - \beta^{(0)})\|_{L^Q}$, and A here is a fixed constant). In other words, we have sacrificed the property $f = \overline{\partial}_b^* \alpha^{(0)}$ by replacing $\alpha^{(0)} \in N\dot{L}^{1,Q}$ with $\beta^{(0)}$, which in addition to being in $N\dot{L}^{1,Q}$ is in L^∞. Now we repeat the process, with $f - \overline{\partial}_b^* \beta^{(0)}$ in place of f, so that we obtain $\beta^{(1)} \in N\dot{L}^{1,Q} \cap L^\infty(\Lambda^{(0,q+1)})$ with

$$\begin{cases} \|f - \overline{\partial}_b^* \beta^{(0)} - \overline{\partial}_b^* \beta^{(1)}\|_{L^Q} \le \frac{1}{2}\|f - \overline{\partial}_b^* \beta^{(0)}\|_{L^Q} \le \frac{1}{2^2}\|f\|_{L^Q}, \\ \|\beta^{(1)}\|_{L^\infty} + \|\nabla_b \beta^{(1)}\|_{L^Q} \le A\|f - \overline{\partial}_b^* \beta^{(0)}\|_{L^Q} \le \frac{A}{2}\|f\|_{L^Q}. \end{cases}$$

Iterating, we get $\beta^{(k)} \in N\dot{L}^{1,Q} \cap L^\infty(\Lambda^{(0,q+1)})$ such that

$$\begin{cases} \|\beta - \overline{\partial}_b^*(\beta^{(0)} + \cdots + \beta^{(k)})\|_{L^Q} \le \frac{1}{2^{k+1}}\|f\|_{L^Q}, \\ \|\beta^{(k)}\|_{L^\infty} + \|\nabla_b \beta^{(k)}\|_{L^Q} \le \frac{A}{2^k}\|f\|_{L^Q}. \end{cases}$$

Hence $Y = \sum_{k=0}^\infty \beta^{(k)}$ satisfies $Y \in N\dot{L}^{1,Q} \cap L^\infty(\Lambda^{(0,q+1)})$ with

$$\begin{cases} \overline{\partial}_b^* Y = f, \\ \|Y\|_{L^\infty} + \|\nabla_b Y\|_{L^Q} \le 2A\|f\|_{L^Q}, \end{cases}$$

as desired. □

We mention that by the duality between $(0, q)$ forms and $(0, n-q)$ forms, we have the following corollary for solving $\overline{\partial}_b$ on $\mathbb{H}^n$:

Corollary 12.1. *Suppose $q \ne 1$. Then for any $(0, q)$ form f on $\mathbb{H}^n$ that has coefficients in L^Q and that is the $\overline{\partial}_b$ of some other form on $\mathbb{H}^n$ with coefficients in $N\dot{L}^{1,Q}$, there exists a $(0, q-1)$ form Y on $\mathbb{H}^n$ with coefficients in $L^\infty \cap N\dot{L}^{1,Q}$ such that*

$$\overline{\partial}_b Y = f$$

in the sense of distributions, with $\|Y\|_{L^\infty} + \|\nabla_b Y\|_{L^Q} \le C\|f\|_{L^Q}$.

Proof of Theorem 1.9. We use duality and the Hodge decomposition for $\overline{\partial}_b$. Suppose first u is a C_c^∞ $(0, q)$ form on $\mathbb{H}^n$ with $2 \le q \le n-2$. We test it against a $(0, q)$ form $\phi \in C_c^\infty$. Now

$$\phi = \overline{\partial}_b^* \alpha + \overline{\partial}_b \beta$$

by Hodge decomposition for $\overline{\partial}_b$ on $\mathbb{H}^n$, where

$$\|\nabla_b \alpha\|_{L^Q} + \|\nabla_b \beta\|_{L^Q} \le C\|\phi\|_{L^Q}.$$

Applying Theorem 1.8 to $\overline{\partial}_b^* \alpha$ and Corollary 12.1 to $\overline{\partial}_b \beta$, we get

$$\phi = \overline{\partial}_b^* \tilde{\alpha} + \overline{\partial}_b \tilde{\beta}$$

where $\tilde{\alpha}$ and $\tilde{\beta}$ have coefficients in $N\dot{L}^{1,Q} \cap L^{\infty}$, with bounds

$$\|\nabla_b \tilde{\alpha}\|_{L^Q} + \|\tilde{\alpha}\|_{L^\infty} \leq C\|\overline{\partial}_b^* \alpha\|_{L^Q} \leq C\|\phi\|_{L^Q},$$

$$\|\nabla_b \tilde{\beta}\|_{L^Q} + \|\tilde{\beta}\|_{L^\infty} \leq C\|\overline{\partial}_b \beta\|_{L^Q} \leq C\|\phi\|_{L^Q}.$$

Thus

$$\begin{aligned}(u, \phi) &= (u, \overline{\partial}_b^* \tilde{\alpha}) + (u, \overline{\partial}_b \tilde{\beta}) = (\overline{\partial}_b u, \tilde{\alpha}) + (\overline{\partial}_b^* u, \tilde{\beta}) \\ &\leq \|\overline{\partial}_b u\|_{L^1 + (N\dot{L}^{1,Q})^*} \|\tilde{\alpha}\|_{L^\infty \cap N\dot{L}^{1,Q}} + \|\overline{\partial}_b^* u\|_{L^1 + (N\dot{L}^{1,Q})^*} \|\tilde{\beta}\|_{L^\infty \cap N\dot{L}^{1,Q}} \\ &\leq C(\|\overline{\partial}_b u\|_{L^1 + (N\dot{L}^{1,Q})^*} + \|\overline{\partial}_b^* u\|_{L^1 + (N\dot{L}^{1,Q})^*}) \|\phi\|_{L^Q}.\end{aligned}$$

This proves the desired inequality (1.1).

The proof of (1.2) for functions u orthogonal to the kernel of $\overline{\partial}_b$ is similar, so we omit it. □

Acknowledgments. The authors would like to thank E. Stein for originally suggesting this problem, H. Brezis for his encouragement and interest in our work, and the referee who read the manuscript very carefully and gave us many helpful suggestions.

Yi Wang is supported in part by NSF grant DMS 1205350. The second author is supported in part by NSF grant DMS 1201474.

References

[AN] Amrouche, C., Nguyen, H. H.: New estimates for the div-curl-grad operators and elliptic problems with L^1-data in the whole space and in the half-space. J. Differential Equations **250**, 3150–3195 (2011) Zbl 1213.35205 MR 2771258

[BB1] Bourgain, J., Brezis, H.: On the equation div $Y = f$ and application to control of phases. J. Amer. Math. Soc. **16**, 393–426 (2003) Zbl 1075.35006 MR 1949165

[BB2] Bourgain, J., Brezis, H.: New estimates for elliptic equations and Hodge type systems. J. Eur. Math. Soc. **9**, 277–315 (2007) Zbl 1176.35061 MR 2293957

[BvS] Brezis, H., van Schaftingen, J.: Boundary estimates for elliptic systems with L^1-data. Calc. Var. Partial Differential Equations **30**, 369–388 (2007) Zbl 1149.35025 MR 2332419

[CvS] Chanillo, S., van Schaftingen, J.: Subelliptic Bourgain–Brezis estimates on groups. Math. Res. Lett. **16**, 487–501 (2009) Zbl 1184.35119 MR 2511628

[FS] Folland, G. B., Stein, E. M.: Hardy Spaces on Homogeneous Groups. Princeton Univ. Press (1982) Zbl 0508.42025 MR 0657581

[LS] Lanzani, L., Stein, E. M.: A note on div curl inequalities. Math. Res. Lett. **12**, 57–61 (2005) Zbl 1113.26015 MR 2122730

[Ma] Maz'ya, V.: Estimates for differential operators of vector analysis involving L^1-norm. J. Eur. Math. Soc. **12**, 221–240 (2010) Zbl 1187.42011 MR 2578609

[Mi] Mironescu, P.: On some inequalities of Bourgain, Brezis, Maz'ya, and Shaposhnikova related to L^1 vector fields. C. R. Math. Acad. Sci. Paris **348**, 513–515 (2010) Zbl 1204.35082 MR 2645163

[MM] Mitrea, I., Mitrea, M.: A remark on the regularity of the div-curl system. Proc. Amer. Math. Soc. **137**, 1729–1733 (2009) Zbl 1168.35325 MR 2470831

[S] Stein, E. M.: Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton Univ. Press (1993) Zbl 0821.42001 MR 1232192

[vS1] van Schaftingen, J.: Estimates for L^1-vector fields. C. R. Math. Acad. Sci. Paris **339**, 181–186 (2004) Zbl 1049.35069 MR 2078071

[vS2] van Schaftingen, J.: Estimates for L^1 vector fields under higher-order differential conditions. J. Eur. Math. Soc. **10**, 867–882 (2008) Zbl 1228.46034 MR 2443922

[vS3] van Schaftingen, J.: Limiting fractional and Lorentz space estimates of differential forms. Proc. Amer. Math. Soc. **138**, 235–240 (2010) Zbl 1184.35012 MR 2550188

[Y] Yung, P.-L.: Sobolev inequalities for $(0, q)$ forms on CR manifolds of finite type. Math. Res. Lett. **17**, 177–196 (2010) Zbl 1238.32027 MR 2592736