A proof of Hörmander's theorem for sublaplacians on Carnot groups

Marco Bramanti
Dipartimento di Matematica
Politecnico di Milano
Via Bonardi 9. 20133 Milano, Italy
e-mail: marco.bramanti@polimi.it

Luca Brandolini
Dipartimento di Ingegneria gestionale, dell'informazione e della produzione
Università degli Studi di Bergamo
Viale Marconi 5. 24044 Dalmine BG, Italy
e-mail: luca.brandolini@unibg.it

December 12, 2014

Abstract

We give a self-contained analytical proof of Hörmander's hypoellipticity theorem in the case of left invariant sub-laplacians of Carnot groups. The proof does not rely on pseudodifferential calculus and provides intermediate regularity estimates expressed in terms of Sobolev spaces induced by right invariant vector fields.

1 Introduction

A linear differential operator $\mathcal{L}$ with smooth coefficients is said to be hypoellitic in an open set $\Omega \subseteq \mathbb{R}^N$ if, for every open set $\Omega' \subseteq \Omega$, whenever a distribution $u \in \mathcal{D}'(\Omega')$ is such that $\mathcal{L}u \in C^\infty(\Omega')$ then $u \in C^\infty(\Omega')$. The celebrated hypoellipticity theorem of Hörmander [5] provides an almost complete characterization of second order hypoelliptic operators with real coefficients.

After noting that every hypoelliptic second order differential operator has necessarily semi-definite principal part, Hörmander proves that in any open set where the rank of the coefficient matrix is constant, the operator (or its opposite) can be rewritten in the form

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2 + X_0 + c \qquad (1.1)$$

where $X_0, X_1, \ldots, X_q$ are real smooth vector fields (that is, first order differential operators) and c is a smooth function. Motivated by this preliminary analysis, Hörmander studies operators already written in the form (1.1) and proves that $\mathcal{L}$ is hypoelliptic in $\Omega \subseteq \mathbb{R}^N$ if the vector fields of the Lie algebra generated by $X_0, X_1, \ldots, X_q$ span the whole $\mathbb{R}^N$ at every point $x \in \Omega$ (an assumption which has been labeled "Hörmander's condition"). Conversely, if in an open set $U \subset \Omega$ the rank of the Lie algebra is constant and strictly less than N, then the operator $\mathcal{L}$ is not hypoelliptic in U, hence Hörmander's condition is "almost necessary" for hypoellipticity. A somewhat easier proof of Hörmander's theorem was given by Kohn [6], deeply exploiting techniques of pseudodifferential operators. The paper by Kohn emphasizes an intermediate result which has great independent interest, namely the so-called subelliptic estimates, allowing to control a fractional Sobolev norm $H^{s+\varepsilon}$ of u (for some small $\varepsilon > 0$) in terms of a norm H^s of $\mathcal{L}u$ and some other less regular norm of u.

Although throughout the years several authors have written other proofs of Hörmander's theorem, it is remarkable that, apart from the probabilistic approach due to Malliavin [7], all the analytic proofs of this result substantially go along the line of Kohn, based on subelliptic estimates and pseudodifferential operators.

After the seminal hypoellipticity result by Hörmander, the theory of "Hörmander's operators" (1.1) has undergone a tremendous development. In particular, much more precise estimates for the regularity properties of operators of this kind have been obtained: Folland [4] and Rothschild-Stein [8] proved, in increasing generality, a priori estimates in L^p ($1 < p < \infty$) for the highest order derivatives *with respect to the vector fields* involved in the equations, that is estimates on $X_i X_j u$ ($i, j = 1, 2, ..., q$), $X_0 u$, in terms of $\mathcal{L}u$ and u, and analogous higher order estimates. More in detail, Folland [4] studied "Hörmander's operators on homogeneous groups", that is operators of the kind

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2 + X_0$$

where the vector fields are assumed to be left invariant with respect to a Lie group in $\mathbb{R}^N$ ("translations") and homogeneous with respect to a family of group automorphisms ("dilations"), so that $\mathcal{L}$ be 2-homogeneous and left invariant (precise definitions will be given later). For these operators Folland showed the existence of a global homogeneous fundamental solution, which is a key tool in proving sharp a priori estimates. Rothschild-Stein [8] proved analogous a priori estimates for general Hörmander's operators, also exploiting the fundamental solution built by Folland on groups. It is worthwhile noting that Folland's construction of a homogeneous fundamental solution for Hörmander's operators on homogeneous groups exploits, among other results, the hypoellipticity of these operators, a fact which, instead, is not directly used any longer in the theory developed by Rothschild-Stein. Hence we can say that Hörmander's theorem plays a special role in the context of homogeneous groups. Also, from different points of view one can say that Hörmander's operators on groups serve as models

2

for general Hörmander's operators. The corresponding theory, and particularly that of *sublaplacians on stratified homogeneous groups* (nowadays called *Carnot groups*), that is left invariant 2-homogeneous Hörmander's operators of the kind

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2$$

has been widely studied in the last decades. The books [1] and [12] contain lots of material on this subject and also references to the original papers.

In this paper we give a new proof of Hörmander's hypoellipticity theorem for sublaplacians in Carnot groups (as far as we know, the first proof in such simplified context). The double simplification consisting in assuming an underlying structure of Carnot group and avoiding the presence of the "drift term" X_0 allows to devise a fairly elementary proof. Actually, we completely avoid the use of the theory of pseudodifferential operators. We prove a kind of subelliptic estimates which are not shaped on Euclidean Sobolev spaces of fractional order, but on Sobolev spaces induced by the vector fields themselves, hence more in the spirit of the subsequent developments of the theory of Hörmander's operators. Our proof will consist in two steps. First, in section 3, we will prove a priori estimates stating that if $u \in W_{X,loc}^{1,2}$ is a local weak solution to $\mathcal{L}u = f$ in a domain (see section 2 for precise definition and notation), then the L^2 norm of any fixed number of X_i^R-derivatives of u (X_i^R being right invariant vector fields) can be controlled in terms of the L^2 norm of u and those of a (*much larger*) number of derivatives of f. Although this kind of estimate is not sharp, it has the same independent interest of Kohn's subellipticity estimates, because not only it implies that if f is smooth also u is smooth, but also expresses a quantitative a priori control on any finite number of derivatives, which is useful in several situations. The second part of the proof (section 4) consists in weakening the assumption on u, from $u \in W_{X,loc}^{1,2}$ to general distribution, which implies the hypoellipticity of $\mathcal{L}$. The first part of the proof (section 3) is nearly self-contained: we will make use of some standard facts from the theory of Carnot groups, collected in section 2. In the second part of the proof (section 4) we also have to make use of some well-known facts from distribution theory. Our proof will make use both of the Sobolev spaces induced by the vector fields X_i (as well as their right-invariant counterparts X_i^R) and of some seminorms defined in terms of finite differences that measure the Hölder continuity in L^2 sense along the vector fields of the first layer, and will be introduced and discussed in section 3. We will sketch the strategy of each of the two steps of the proof at the begenning of the corresponding section.

3

2 Some known facts about Carnot groups

In this section we recall a number of standard definitions and results that will be useful in the following. For the proofs of these facts the reader is referred to [4], [1, Chap.1], [10, Chap.XIII, §5], [2].

A *homogeneous group* (in $\mathbb{R}^N$) is a Lie group $(\mathbb{R}^N, \circ)$ (where the group operation $\circ$ will be thought as a "translation") endowed with a one parameter family $\{D_\lambda\}_{\lambda > 0}$ of group automorphisms ("dilations") which act this way:

$$D_\lambda(x_1, x_2, ..., x_N) = (\lambda^{\alpha_1} x_1, \lambda^{\alpha_2} x_2, ..., \lambda^{\alpha_N} x_N) \tag{2.1}$$

for suitable integers $1 = \alpha_1 \leqslant \alpha_2 \leqslant ... \leqslant \alpha_N$. We will write $\mathbb{G} = (\mathbb{R}^N, \circ, D_\lambda)$ to denote this structure. The number

$$Q = \sum_{i=1}^{N} \alpha_i$$

will be called *homogeneous dimension* of $\mathbb{G}$.

Under the change of coordinates $x = D_\lambda(y)$ the volume element transforms according to

$$dx = \lambda^Q dy \tag{2.2}$$

which justifies the name of homogeneous dimension for Q. Note that we always have $Q \geqslant N$, and $Q = N$ only if the dilations are the Euclidean ones.

A *homogeneous norm* on $\mathbb{G}$ (also called a *gauge* on $\mathbb{G}$) is a continuous function

$$\|\cdot\| : \mathbb{R}^N \to [0, +\infty),$$

such that, for some constant $c > 0$, for every $x, y \in \mathbb{R}^N$,

$$
\begin{array}{ll}
(i) & \|x\| = 0 \iff x = 0 \\
(ii) & \|D_\lambda(x)\| = \lambda \|x\| \ \forall \lambda > 0 \\
(iii) & \|x^{-1}\| \leqslant c \|x\| \\
(iv) & \|x \circ y\| \leqslant c(\|x\| + \|y\|).
\end{array}
$$

We will always use the symbol $\|\cdot\|$ to denote a homogeneous norm, and the symbol $|\cdot|$ to denote the Euclidean norm.

Concrete ways to define a homogeneous norm on $\mathbb{G}$ are for instance the following:

$$\|x\| = \max_{k=1,2,...,N} |x_k|^{\frac{1}{\alpha_k}}$$

or

$$\|x\| = \left(\sum_{k=1}^{N} |x_k|^{\frac{Q}{\alpha_k}} \right)^{1/Q}.$$

It can be proved that any two homogeneous norms $\|\cdot\|_1, \|\cdot\|_2$ on $\mathbb{G}$ are equivalent: there exist two positive constants k_1, k_2 such that

$$k_1 \|x\|_1 \leqslant \|x\|_2 \leqslant k_2 \|x\|_1 \ \forall x \in \mathbb{R}^N.$$

Also, the following *local* comparison with the Euclidean norm holds: for any $R > 0$ there exist two positive constants c_1, c_2 (depending on R) such that

$$c_1 |x| \leqslant \|x\| \leqslant c_2 |x|^{1/Q} \quad \text{if } |x| \leqslant R \text{ (or } \|x\| \leqslant R) \tag{2.3}$$

Although in general $x - y \neq y^{-1} \circ x$, they are locally equivalent near the origin. More precisely there exist positive constants c_1 and c_2 and a neighborhood of the origin V such that for $x, y \in V$ we have

$$c_1 |x - y| \leqslant \left| y^{-1} \circ x \right| \leqslant c_2 |x - y| . \tag{2.4}$$

To see this let $P_y(x) = y^{-1} \circ x$. It follows from the structure of the group operation on a homogeneous group that $\frac{\partial P_y}{\partial x}(y)$ is close to the identity in a neighborhood of the origin. Since

$$y^{-1} \circ x = P_y(y) + \frac{\partial P_y}{\partial x}(y)(y - x) + o(y - x)$$

and $P_y(y) = 0$, (2.4) readly follows.

We say that a smooth function f in $\mathbb{R}^N \setminus \{0\}$ is D_λ-*homogeneous of degree* $\beta \in \mathbb{R}$ (or simply "β-homogeneous") if

$$f(D_\lambda(x)) = \lambda^\beta f(x) \quad \forall \lambda > 0, x \in \mathbb{R}^N \setminus \{0\} .$$

Given any differential operator P with smooth coefficients on $\mathbb{R}^N$, we say that P is *left invariant* if for every $x, y \in \mathbb{R}^N$

$$P(L_y f)(x) = L_y(Pf(x))$$

for every smooth function f, where

$$L_y f(x) = f(y \circ x) .$$

Analogously one defines the notion of *right invariant* differential operator. Also, P is β-*homogeneous* (for some $\beta \in \mathbb{R}$) if

$$P(f(D_\lambda(x))) = \lambda^\beta (Pf)(D_\lambda(x))$$

for every smooth function f, $\lambda > 0$ and $x \in \mathbb{R}^N$.

A *vector field* is a first order differential operator

$$X = \sum_{i=1}^N c_i(x) \partial_{x_i} .$$

Let $\mathfrak{g}$ be the *Lie algebra of left invariant vector fields* over $\mathbb{G}$, where the Lie bracket of two vector fields is defined as usual by

$$[X, Y] = XY - YX .$$

Let us denote by $X_1, X_2, \ldots, X_N$ the *canonical base* of $\mathfrak{g}$, that is for $i = 1, 2, \ldots, N$, X_i is the only left invariant vector field that agrees with ∂_{x_i} at the origin. Also, $X_1^R, X_2^R, \ldots, X_N^R$ will denote the right invariant vectors fields that agree with $\partial_{x_1}, \partial_{x_2}, \ldots, \partial_{x_N}$ (and hence with $X_1, X_2, \ldots, X_N$) at the origin.

We assume that for some integer $q < N$ the vector fields $X_1, X_2, \ldots, X_q$ are 1-homogeneous and the Lie algebra generated by them is $\mathfrak{g}$. If s is the maximum length of commutators

$$\left[X_{i_1}, \left[X_{i_2}, \ldots, \left[X_{i_{s-1}}, X_{i_s} \right] \right] \right], \quad i_j \in \{1, 2, \ldots, q\}$$

required to span $\mathfrak{g}$, then we will say that $\mathfrak{g}$ is a *stratified Lie algebra of step s*, $\mathbb{G}$ is a *Carnot group* (or a stratified homogeneous group) and $X_1, X_2, \ldots, X_q$ satisfy Hörmander's condition at step s in $\mathbb{R}^N$.

Under these assumptions, let us denote by

$$\mathcal{L} = \sum_{j=1}^{q} X_j^2$$

the *canonical sublaplacian* on $\mathbb{G}$. It will be sometimes useful the compact notation

$$|\nabla_X f| = \sum_{j=1}^{q} |X_j f|$$

$$|\nabla_{X^R} f| = \sum_{j=1}^{q} \left| X_j^R f \right|$$

We will make use of the *Sobolev spaces* $W_X^{k,p}(\mathbb{G})$, $W_{X^R}^{k,p}(\mathbb{G})$ *induced by the systems of vector fields*

$$X = \{X_1, X_2, \ldots, X_q\}, X^R = \left\{ X_1^R, X_2^R, \ldots, X_q^R \right\},$$

respectively. More precisely, given an open subset Ω of $\mathbb{R}^N$, we say that $f \in W_X^{1,p}(\Omega)$ $(1 \leqslant p \leqslant \infty)$ if $f \in L^p(\Omega)$ and there exist, in weak sense, $X_j f \in L^p(\Omega)$ for $j = 1, 2, \ldots, q$. Inductively, we say that $f \in W_X^{k,p}(\Omega)$ for $k = 2, 3, \ldots$ if $f \in W_X^{k-1,p}(\Omega)$ and any weak derivative of order $k-1$ of f, $X_{j_1} X_{j_2} \ldots X_{j_{k-1}} f$, belongs to $W_X^{1,p}(\Omega)$. We set

$$\|f\|_{W_X^{k,p}(\Omega)} = \|f\|_{L^p(\Omega)} + \sum_{h=1}^{k} \sum_{j_i = 1, 2, \ldots, q} \|X_{j_1} X_{j_2} \ldots X_{j_h} f\|_{L^p(\Omega)}.$$

The space $W_{X^R}^{k,p}(\Omega)$ has a similar definition.

As usual, one can define the Sobolev spaces of functions vanishing at the boundary of a domain, $W_{0,X}^{1,p}(\Omega)$ or $W_{0,X^R}^{1,p}(\Omega)$ taking the closure of $C_0^\infty(\Omega)$ in the norm $W_X^{1,p}(\Omega)$ or $W_{X^R}^{1,p}(\Omega)$, respectively. It can be proved that for $1 \leqslant p < \infty$ one has $W_{0,X}^{1,p}(\mathbb{R}^N) = W_X^{1,p}(\mathbb{R}^N)$ and analogously for X^R-spaces.

We will also use local Sobolev spaces. For example, we will say that $f \in W^{k,p}_{X,loc}(\Omega)$ if for every $\varphi \in C^\infty_0(\Omega)$, we have $\varphi f \in W^{k,p}_X(\Omega)$.

The validity of Hörmander's condition at step s implies the following important:

Proposition 2.1 (Regularity by means of Sobolev spaces) *Under the above assumptions we have:*

1. For any $1 \leqslant p \leqslant \infty$,

$$\bigcap_{k=1}^\infty W^{k,p}_X(\Omega) \subset C^\infty(\Omega).$$

2. For any positive integer k and $p \in [1, \infty]$ and any $\Omega' \Subset \Omega$ there exists a constant $c > 0$ such that, for every $u \in W^{ks,p}_X(\Omega)$ we have

$$\|u\|_{W^{k,p}(\Omega')} \leqslant c \, \|u\|_{W^{ks,p}_X(\Omega)} \, ,$$

where $W^{k,p}(\Omega')$ denotes the standard Sobolev space.

Let us also recall the definition of *control distance* induced by the vector fields $X_1, X_2, ..., X_q$ in $\mathbb{R}^N$. For any $\delta > 0$, let $C_{x,y}(\delta)$ be the class of absolutely continuous mappings $\varphi : [0, 1] \longrightarrow \mathbb{R}^N$ which satisfy

$$\varphi'(t) = \sum_{i=1}^q a_i(t)(X_i)_{\varphi(t)} \text{ a.e.}$$

with $a_i : [0, 1] \to \mathbb{R}$ measurable functions,

$$|a_i(t)| \leqslant \delta \text{ a.e.}$$
$$\varphi(0) = x, \varphi(1) = y.$$

Then define

$$d(x, y) = \inf \{\delta > 0 : \exists \varphi \in C_{x,y}(\delta)\} \, ,$$

with the convention $\inf \emptyset = +\infty$. We also define the d-balls

$$B_X(x, r) = \{y \in \Omega : d(x, y) < r\} \, .$$

A consequence of Hörmander's condition is that actually $d(x, y)$ is finite for every couple of points (*Chow-Rashevskii connectivity theorem*). The control distance d on a Carnot group has several properties:

i) d is left invariant:

$$d(z \circ x, z \circ y) = d(x, y) \quad \forall x, y, z \in \mathbb{R}^N \tag{2.5}$$

and satisfies

$$d(x_1 \circ x_2 \circ ... \circ x_k, 0) \leqslant \sum_{j=1}^k d(x_j, 0) \, ; \tag{2.6}$$

ii) d is 1-homogeneous:

$$d\left(D_\lambda\left(x\right),D_\lambda\left(y\right)\right)=\lambda d\left(x,y\right)\quad\forall x,y\in\mathbb{R}^N,\lambda>0.$$

iii) the function
$$\|x\|=d\left(x,0\right)$$

is a homogeneous norm. More precisely, it also satisfies the stronger properties

$$\begin{aligned}\left\|x^{-1}\right\|&=\|x\|\\\|x\circ y\|&\leqslant\|x\|+\|y\|\,.\end{aligned}\tag{2.7}$$

Throughout the following we will need to use a homogeneous norm satisfying the triangle inequality (2.7). The previous discussion shows that such a homogeneous norm always exists.

Actually this particular choice of a homogeneous norm is relevant only within the proofs, while the statements of all our results will still hold, with possibly different constants, if we replace this homogeneous norm with a different one.

Let $\mathrm{Exp}:\mathfrak{g}\to\mathbb{G}$ be the exponential map, defined as usual letting

$$\mathrm{Exp}\left(X\right)=\exp\left(X\right)\left(0\right)$$

where $\exp\left(Y\right)\left(x_0\right)=F\left(1\right)$ for F the solution to the Cauchy problem $F'\left(\tau\right)=Y_{F\left(\tau\right)};F\left(0\right)=x_0$. One has

$$\exp\left(X\right)\left(x\right)=x\circ\mathrm{Exp}\left(X\right).\tag{2.8}$$

Also recall that for every smooth function f and every $t\in\mathbb{R}$

$$\frac{d}{dt}\left(f\left(x\circ\mathrm{Exp}\left(tX\right)\right)\right)=Xf\left(x\circ\mathrm{Exp}\left(tX\right)\right).\tag{2.9}$$

Let us pointing out a relation between left and right invariant operators which will be very useful in the following.

Proposition 2.2 *Let* $\mathcal{L},\mathcal{R}$ *be any two differential operators with smooth coefficients, left and right invariant, respectively. Then* $\mathcal{L}$ *and* $\mathcal{R}$ *commute:*

$$\mathcal{L}\mathcal{R}f=\mathcal{R}\mathcal{L}f$$

for any smooth function f.

Proof. We have

$$\begin{aligned}\mathcal{L}f\left(x\right)&=\mathcal{L}_0^y\left[f\left(x\circ y\right)\right]|_{y=0}\\\mathcal{R}f\left(x\right)&=\mathcal{R}_0^z\left[f\left(z\circ x\right)\right]|_{z=0}\end{aligned}$$

where $\mathcal{L}_0, \mathcal{R}_0$ are constant coefficient differential operators (namely, they are the variable operators $\mathcal{L}, \mathcal{R}$ evaluated at the origin). Then:

$$\begin{aligned}
\mathcal{L}\mathcal{R}f(x) &= \mathcal{L}_0^y \left[\mathcal{R}f(x \circ y)\right]\big|_{y=0} = \mathcal{L}_0^y \left[\mathcal{R}_0^z \left[f(z \circ x \circ y)\right]\big|_{z=0}\right]\big|_{y=0} \\
&= \mathcal{L}_0^y \mathcal{R}_0^z \left[f(z \circ x \circ y)\right]\big|_{(z,y)=(0,0)} = \mathcal{R}_0^z \mathcal{L}_0^y \left[f(z \circ x \circ y)\right]\big|_{(z,y)=(0,0)} \\
&= \mathcal{R}\mathcal{L}f(x)
\end{aligned}$$

since $\mathcal{R}_0^z$, $\mathcal{L}_0^y$ obviously commute. ∎

We will now introduce on $\mathbb{G}$ the convolution between test functions and distributions. Throughout the paper we will denote this operation with the symbol $*$. Since the group $\mathbb{G}$ is in general noncommutative also $*$ is noncommutative. However, most of its properties are straightforward adaptations of the same properties that hold for the standard convolution, see e.g. [9, Chaper 6]. Hence, we will not give the proofs of the following two propositions.

For any given couple of test functions $\varphi, \psi \in C_0^\infty(\mathbb{G})$ we define

$$\varphi * \psi(x) = \int_{\mathbb{G}} \varphi(y) \psi\left(y^{-1} \circ x\right) dy.$$

It is a simple exercise to check that $\varphi * \psi \in C_0^\infty(\mathbb{G})$ and that given $\varphi, \psi, \omega \in C_0^\infty(\mathbb{G})$ we have

$$\varphi * (\psi * \omega) = (\varphi * \psi) * \omega.$$

Using the left translation operator $L_x \varphi(y) = \varphi(x \circ y)$ and the inversion operator $\check{\varphi}(x) = \varphi\left(x^{-1}\right)$ we can also write

$$\varphi * \psi(x) = \int \varphi(y) L_{x^{-1}} \check{\psi}(y) \, dy$$

so that given $f \in \mathcal{D}'(\mathbb{G})$ (the space of distributions on $\mathbb{G}$), it is natural to define

$$f * \psi(x) = \left\langle f, L_{x^{-1}} \check{\psi} \right\rangle. \tag{2.10}$$

For $f \in \mathcal{D}'(\mathbb{G})$ let also define $\check{f}$ requiring that for every $\varphi \in C_0^\infty(\mathbb{G})$,

$$\left\langle \check{f}, \varphi \right\rangle = \langle f, \check{\varphi} \rangle.$$

One can prove the following:

Proposition 2.3 *Let $f \in \mathcal{D}'(\mathbb{G})$ and let $\psi, \varphi \in C_0^\infty(\mathbb{G})$. Then we have:*
i)

$$\langle f * \psi, \varphi \rangle = \left\langle f, \varphi * \check{\psi} \right\rangle = \left\langle \check{f}, \psi * \check{\varphi} \right\rangle$$

ii) if $\mathcal{P}$ is a left invariant differential operator then

$$\mathcal{P}(f * \psi) = f * \mathcal{P}\psi,$$

9

iii) if $\mathcal{P}$ is a right invariant differential operator then

$$\mathcal{P}(f * \psi) = \mathcal{P}f * \psi$$

Remark 2.4 *From ii) it follows that $f * \psi \in C^\infty(\mathbb{G})$.*

Assume now that $f \in \mathcal{D}'(\mathbb{G})$ has compact support. In this case f has a unique extension to a linear functional on $C^\infty(\mathbb{G})$ and therefore (2.10) can be used to define the convolution with functions $\psi \in C^\infty(\mathbb{G})$.

Also in this case if $\mathcal{P}$ is a left invariant differential operator we have $\mathcal{P}(f * \psi) = f * \mathcal{P}\psi$.

Let now $f \in \mathcal{D}'(\mathbb{G})$ with compact support and $g \in \mathcal{D}'(\mathbb{G})$. For every $\varphi \in C_0^\infty(\mathbb{G})$ we define

$$\langle f * g, \varphi \rangle = \left\langle \breve{f}, g * \breve{\varphi} \right\rangle$$

where the inversion operator for distribution is defined in the obvious way:

$$\left\langle \breve{f}, \varphi \right\rangle = \langle f, \breve{\varphi} \rangle.$$

Note that $g * \breve{\varphi} \in C^\infty(\mathbb{G})$ so that $\left\langle \breve{f}, g * \breve{\varphi} \right\rangle$ is well defined.

The following holds:

Proposition 2.5 *Let $f, g \in \mathcal{D}'(\mathbb{G})$ and assume that f has compact support. We have the following:*
i) If $\mathcal{P}$ is a left invariant differential operator then

$$\mathcal{P}(f * g) = f * \mathcal{P}g$$

ii) If $\mathcal{P}$ is a right invariant differential operator then

$$\mathcal{P}(f * g) = \mathcal{P}f * g.$$

3 Regularity estimates for the sublaplacian in $W_X^{1,2}$

In order to show that whenever $\mathcal{L}f$ is smooth also f is smooth we will prove estimates that allow to control the Sobolev norm of f in terms of the Sobolev norm of $\mathcal{L}f$.

Our starting point is the following elementary computation. Let $f \in C_0^\infty(\mathbb{G})$, then

$$\|\nabla_X f\|_2^2 = \int_\mathbb{G} \sum_{i=1}^q (X_i f(x))^2 \, dx$$

$$= - \int_\mathbb{G} \mathcal{L}f(x) f(x) \, dx$$

$$\leqslant \|\mathcal{L}f\|_2 \|f\|_2 \leqslant \frac{1}{2} \left(\|\mathcal{L}f\|_2^2 + \|f\|_2^2 \right)$$

which gives the estimate

$$\|f\|_{W_X^{1,2}} \leqslant c\left(\|\mathcal{L}f\|_2 + \|f\|_2\right) \ \forall f \in C_0^\infty\left(\mathbb{G}\right). \tag{3.1}$$

Assume we would know, instead of (3.1), the following apparently similar estimate

$$\|f\|_{W_{X^R}^{1,2}} \leqslant c\left(\|\mathcal{L}f\|_2 + \|f\|_2\right) \ \forall f \in C_0^\infty. \tag{3.2}$$

where the Sobolev norm in the LHS is computed using the right invariant vector fields. Then exploiting the fact that the left invariant operator $\mathcal{L}$ commutes with the right invariant operators X_i^R (see Proposition 2.2) we could apply (3.2) to the function $X_{i_1}^R X_{i_2}^R ... X_{i_k}^R f$, obtaining

$$\left\|X_{i_1}^R X_{i_2}^R ... X_{i_k}^R f\right\|_{W_{X^R}^{1,2}} \leqslant c\left(\left\|\mathcal{L}X_{i_1}^R X_{i_2}^R ... X_{i_k}^R f\right\|_2 + \left\|X_{i_1}^R X_{i_2}^R ... X_{i_k}^R f\right\|_2\right)$$
$$= c\left(\left\|X_{i_1}^R X_{i_2}^R ... X_{i_k}^R \mathcal{L}f\right\|_2 + \left\|X_{i_1}^R X_{i_2}^R ... X_{i_k}^R f\right\|_2\right)$$

hence

$$\|f\|_{W_{X^R}^{k+1,2}} \leqslant c\left(\|\mathcal{L}f\|_{W_{X^R}^{k,2}} + \|f\|_{W_{X^R}^{k,2}}\right)$$

and iteratively

$$\|f\|_{W_{X^R}^{k+1,2}} \leqslant c\left(\|\mathcal{L}f\|_{W_{X^R}^{k,2}} + \|f\|_2\right),$$

an estimate that allows to control the derivatives of f in terms of the derivatives of $\mathcal{L}f$. In other words, the idea of measuring the degree of smoothness of a solution to $\mathcal{L}u = f$ (with $\mathcal{L}$ *left* invariant) using *right* invariant derivatives apparently trivializes the problem, as if we were handling an elliptic operator with constant coefficients. However, we *do not* have the bound (3.2). Nevertheless, we will see that it is possible to control the regularity expressed using right invariant vector fields in terms of the left invariant ones, but this implies a loss of regularity. If s is the step of the Lie algebra of $\mathbb{G}$, using one derivative with respect to the left invariant vector fields, we can only control a right invariant regularity of order $1/s$. More precisely we will show that for small $h \in \mathbb{G}$ the following bound holds

$$\left(\int_\mathbb{G} |f\left(h \circ x\right) - f\left(x\right)|^2 \, dx\right)^{1/2} \leqslant c\left\|h\right\|^{1/s} \left\|\nabla_X f\right\|_2 \leqslant c\left\|h\right\|^{1/s}\left(\|\mathcal{L}f\|_2 + \|f\|_2\right). \tag{3.3}$$

In the next section we will show that iterating the above estimate it is possible to control a full derivative with respect to the right invariant vector fields using $s - 1$ derivatives of $\mathcal{L}f$. More generally we will show that it is possible to control k derivatives X_i^R of f using $k + s - 1$ derivatives of $\mathcal{L}f$ and from this we will deduce that for a function f that is locally in $W_{X,}^{1,2}$, whenever $\mathcal{L}f$ is smooth also f is smooth.

Before stating precisely our regularity result let us clarify the meaning of $\mathcal{L}f$ for functions that are locally in $W_X^{1,2}$.

Definition 3.1 *Let $f \in W_{X,loc}^{1,2}(\mathbb{G})$. We say that $\mathcal{L}f \in L_{loc}^2(\mathbb{G})$ if there exists $F \in L_{loc}^2(\mathbb{G})$ such that for every $\phi \in C_0^\infty(\mathbb{G})$ we have*

$$-\sum_{j=1}^q \int X_j f X_j \phi \, dx = \int F\phi \, dx \quad \forall \phi \in C_0^\infty(\mathbb{G}), \tag{3.4}$$

and in this case we will say that f is a local weak solution to the equation $\mathcal{L}f = F$

Remark 3.2 *If $f \in W_X^{1,2}(\mathbb{G})$ and $\mathcal{L}f \in L^2(\mathbb{G})$ then, by density, (3.4) holds for every $\phi \in W_{X,0}^{1,2}(\mathbb{G})$.*

Notation 3.3 *For a couple of cutoff functions $\zeta, \zeta_1 \in C_0^\infty(\Omega)$ we will write*

$$\zeta \prec \zeta_1$$

to say that

$$\zeta_1 \equiv 1 \text{ on } \operatorname{sprt} \zeta.$$

Our regularity result is the following.

Theorem 3.4 (Regularity estimates) *Let $f \in W_{X,loc}^{1,2}(\mathbb{G})$ be a local weak solution to $\mathcal{L}f = F \in L_{loc}^2(\mathbb{G})$, and, for some open set $\Omega \subset \mathbb{G}$, let $\zeta, \zeta_1 \in C_0^\infty(\Omega), \zeta \prec \zeta_1$. Then:*
(i) For any $k = 1, 2, 3, \ldots$, there exists $c > 0$ such that whenever $\mathcal{L}f \in W_{X^R,loc}^{k+s^2-1,2}(\Omega)$ then $f \in W_{X^R,loc}^{k,2}(\Omega)$ and

$$\|\zeta f\|_{W_{X^R}^{k,2}(\mathbb{R}^N)} \leqslant c \left\{ \|\zeta_1 \mathcal{L}f\|_{W_{X^R}^{k+s-1,2}(\mathbb{R}^N)} + \|\zeta_1 f\|_{L^2(\mathbb{R}^N)} \right\}. \tag{3.5}$$

(ii) In particular, if $\mathcal{L}f \in C^\infty(\Omega)$ then $f \in C^\infty(\Omega)$.

The proof of Theorem 3.4 will be accomplished in §3.2. In §3.1 we develop some properties of finite difference operators that will be necessary for the proof of the theorem.

3.1 Finite difference operators and Sobolev norm

It is well known that Hörmander condition for the vector fields $X_1, X_2, \ldots, X_q$ implies that it is possible to connect two points using only integral lines of the vector fields. From an algebraic point of view this means that every element of the group $\mathbb{G}$ can be written as a product of exponentials of the first layer vector fields.

Theorem 3.5 *Let $\mathbb{G}$ be a Carnot group. There exist absolute constants $M \in \mathbb{N}$ and $c > 0$ such that for every $x \in \mathbb{G}$ there exist real numbers $t_1, t_2, \ldots, t_M$ and indices $k_1, k_2, \ldots, k_M \in \{1, 2, \ldots, q\}$ such that*

$$x = \operatorname{Exp}(t_1 X_{k_1}) \circ \operatorname{Exp}(t_2 X_{k_2}) \circ \cdots \circ \operatorname{Exp}(t_M X_{k_M})$$

and $|t_j| \leqslant c\,\|x\|$.

For a proof of this theorem see e.g. [1, 19.2.1].

A simple but important consequence of the above theorem is the fact that it is possible to control the increment of a function using $|\nabla_X f|$. We start with a definition.

Definition 3.6 (Finite difference operators) *For every $h \in \mathbb{G}$ let us define the operators:*

$$\Delta_h f(x) = f(x \circ h) - f(x)$$
$$\widetilde{\Delta}_h f(x) = f(h \circ x) - f(x).$$

Note that the operator Δ_h, which acts on functions computing the increment of f corresponding to an increment of its variable *on the right* $(x \mapsto x \circ h)$ is actually a *left invariant* operator $(L_y \Delta_h f = \Delta_h L_y f)$; analogously, $\widetilde{\Delta}_h$ computes the increment of f corresponding to an increment of its variable *on the left* and is a *right invariant* operator. This "duality" is a central point in the techniques which will be used throughout the paper.

Proposition 3.7 *There exists a constant $c = c(\mathbb{G})$ such that for any $h \in \mathbb{G}$ and any $f \in W_X^{1,2}(\mathbb{G})$ we have*

$$\|\Delta_h f\|_2 \leqslant c \|h\| \||\nabla_X f|\|_2$$

and for any $f \in W_{X^R}^{1,2}(\mathbb{G})$ we have

$$\left\|\widetilde{\Delta}_h f\right\|_2 \leqslant c \|h\| \||\nabla_{X^R} f|\|_2.$$

Proof. It is enough to prove the first assertion, since the second one is analogous. Also, it's enough to prove the assertion for $f \in C_0^\infty(\mathbb{G})$, the case $f \in W_X^{1,2}(\mathbb{G})$ then follows from the fact that $C_0^\infty(\mathbb{G})$ function are dense in $W_X^{1,2}(\mathbb{G})$.

Let $h \in \mathbb{G}$, by Theorem 3.5 we can write

$$h = \operatorname{Exp}(t_1 X_{k_1}) \circ \cdots \circ \operatorname{Exp}(t_{M-1} X_{k_{M-1}}) \circ \operatorname{Exp}(t_M X_{k_M})$$

with

$$|t_i| \leqslant c \|h\|.$$

Let $x \in \mathbb{G}$ and set

$$h_j = \operatorname{Exp}(t_j X_{k_j})$$
$$x_0 = x$$
$$x_j = x_{j-1} \circ h_j = x \circ h_1 \circ h_2 \circ \ldots \circ h_j$$
$$x_M = x \circ h.$$

For any $f \in C_0^\infty(\mathbb{G})$ using (2.9) we obtain

$$f(x \circ h) - f(x) = \sum_{j=1}^{M} f(x_j) - f(x_{j-1}) = \sum_{j=1}^{M} f(x_{j-1} \circ h_j) - f(x_{j-1})$$

$$= \sum_{j=1}^{M} \int_0^{t_j} \frac{d}{ds} \left[f\left(x_{j-1} \circ \mathrm{Exp}\left(sX_{k_j}\right)\right)\right] ds$$

$$= \sum_{j=1}^{M} \int_0^{t_j} X_{k_j} f\left(x_{j-1} \circ \mathrm{Exp}\left(sX_{k_j}\right)\right) ds.$$

Hence

$$\|\Delta_h f\|_2 \leqslant \int_0^{|t_1|} \left\|X_{k_1} f\left(x \circ \mathrm{Exp}\left(sX_{k_1}\right)\right)\right\|_2 ds$$

$$+ \sum_{j=2}^{M} \int_0^{|t_j|} \left\|X_{k_j} f\left(x \circ h_1 \circ h_2 \circ ... \circ h_{j-1} \circ \mathrm{Exp}\left(sX_{k_j}\right)\right)\right\|_2 ds.$$

Since

$$\left\|X_{k_j} f\left(x \circ h_1 \circ h_2 \circ ... \circ h_{j-1} \circ \mathrm{Exp}\left(sX_{k_j}\right)\right)\right\|_2^2 = \int_{\mathbb{G}} \left|X_{k_j} f(x)\right|^2 dx$$

we have (by Theorem 3.5)

$$\|\Delta_h f\|_2 \leqslant \max\{|t_j|\} \sum_{j=1}^{N} \left\|X_{k_j} f\right\|_2 \leqslant C \|h\| \|\nabla_X f\|_2.$$

∎

In order to prove a bound of the kind (3.3) we need to translate the information about a control on the quotient $\frac{\|\Delta_h f\|_2}{\|h\|}$ in terms of a control on $\frac{\|\widetilde{\Delta}_h f\|_2}{\|h\|}$. As already remarked, this implies a loss of regularity. The following lemma is an adaptation of the proof of Lemma 4.1 in [5].

Lemma 3.8 *Let $f \in L^2(\mathbb{G})$, $U \Subset \mathbb{G}$, sprt $f \subset U$. There exists $c > 0$, depending on U, such that*

$$\sup_{0 < \|h\| \leqslant 1} \frac{\left\|\widetilde{\Delta}_h f\right\|_2}{\|h\|^{1/s}} \leqslant c \sup_{0 < \|h\| \leqslant 1} \frac{\|\Delta_h f\|_2}{\|h\|}$$

whenever the right hand side is finite. (Here s is the step of the Lie algebra of $\mathbb{G}$).

Proof. Let $h \in \mathbb{G}$, $\|h\| \leqslant 1$ and let $\mathcal{B} = \left\{k \in \mathbb{G} : \|k\| \leqslant \|h\|^{1/s}\right\}$. For every $k \in \mathbb{G}$ we can write

$$\int_{\mathbb{G}} |f(h \circ x) - f(x)|^2 dx$$

$$\leqslant 2 \int_{\mathbb{G}} |f(h \circ x) - f(x \circ k)|^2 dx + 2 \int_{\mathbb{G}} |f(x \circ k) - f(x)|^2 dx$$

14

and since $|\mathcal{B}| = c\,\|h\|^{Q/s}$, integrating over $\mathcal{B}$ we obtain

$$\int_{\mathbb{G}} |f(h\circ x) - f(x)|^2\, dx$$

$$\leqslant c\,\|h\|^{-Q/s} \int_{\mathcal{B}} \int_{\mathbb{G}} |f(h\circ x) - f(x\circ k)|^2\, dx\, dk$$

$$+\, c\,\|h\|^{-Q/s} \int_{\mathcal{B}} \int_{\mathbb{G}} |f(x\circ k) - f(x)|^2\, dx\, dk$$

$$= c\,\|h\|^{-Q/s} \int_{\mathcal{B}} \int_{\mathbb{G}} |f(h\circ x) - f(x\circ k)|^2\, dx\, dk + c\,\|h\|^{-Q/s} \int_{\mathcal{B}} \|\Delta_k f\|_2^2\, dk$$

$$\equiv A + B.$$

We have

$$B \leqslant c\,\|h\|^{-Q/s} \sup_{0<\|k'\|\leqslant 1} \frac{\|\Delta_{k'} f\|_2^2}{\|k'\|^2} \int_{\mathcal{B}} \|k\|^2\, dk$$

$$\leqslant c\,\|h\|^{-Q/s} \sup_{0<\|k'\|\leqslant 1} \frac{\|\Delta_{k'} f\|_2^2}{\|k'\|^2} \|h\|^{2/s} \int_{\mathcal{B}} dk$$

$$= c\,\|h\|^{2/s} \sup_{0<\|k'\|\leqslant 1} \frac{\|\Delta_{k'} f\|_2^2}{\|k'\|^2}.$$

Let us consider A. We make the change of variable

$$(x,k) = \Phi_h(y,u) = \left(h^{-1}\circ y\circ u,\, u^{-1}\circ y^{-1}\circ h\circ y\right)$$

so that

$$A = c\,\|h\|^{-Q/s} \int_{\Phi_h^{-1}(\mathbb{G}\times\mathcal{B})} |f(y\circ u) - f(y)|^2\, |\det J_{\Phi_h}(y,u)|\, dy\, du.$$

Since $u = y^{-1}\circ h\circ y\circ k^{-1}$,

$$\|u\| \leqslant \|y^{-1}\circ h\circ y\| + \|k^{-1}\|$$

$$\leqslant c\left(\left|y^{-1}\circ h\circ y\right|^{1/s} + \|h\|^{1/s}\right).$$

Let $F(y,h) = y^{-1}\circ h\circ y$. Since F is smooth and $F(y,0) = 0$, it follows that $\left|y^{-1}\circ h\circ y\right| \leqslant c\,|h|$ for y ranging in a compact set. This happens because x ranges in the compact set U, $\|k\| \leqslant \|h\|^{1/s} \leqslant 1$, $\Phi_h(y,u)$ is a diffeomorphism, hence also (y,u) range in a bounded set. Hence we can write

$$\|u\| \leqslant c\left(|h|^{1/s} + \|h\|^{1/s}\right) \leqslant c\,\|h\|^{1/s}.$$

15

Moreover since the Jacobian $J_{\Phi_h}(y,u)$ is also bounded by a constant depending only on U,

$$A \leqslant c \left\| h \right\|^{-Q/s} \int_{\left\{ \left\| u \right\| \leqslant c \left\| h \right\|^{1/s} \right\}} \int_{\mathbb{G}} \left| f\left(y \circ u \right) - f\left(y \right) \right|^2 dy du$$

$$\leqslant c \left\| h \right\|^{-Q/s} \sup_{0 < \left\| k' \right\| \leqslant 1} \frac{\left\| \Delta_{k'} f \right\|_2^2}{\left\| k' \right\|^2} \int_{\left\{ \left\| u \right\| \leqslant c \left\| h \right\|^{1/s} \right\}} \left\| u \right\|^2 du$$

$$= c \sup_{0 < \left\| k' \right\| \leqslant 1} \frac{\left\| \Delta_{k'} f \right\|_2^2}{\left\| k' \right\|^2} \left\| h \right\|^{2/s}.$$

∎

Combining Proposition 3.7 with the above lemma gives

$$\left\| \widetilde{\Delta}_h f \right\|_2 \leqslant c \left\| h \right\|^{1/s} \left\| \left| \nabla_X f \right| \right\|_2 \leqslant c \left\| h \right\|^{1/s} \left(\left\| \mathcal{L} f \right\|_2 + \left\| f \right\|_2 \right) \tag{3.6}$$

for every $f \in W_X^{1,2}(\mathbb{G})$ supported in a compact U, and $\mathcal{L} f \in L^2(\mathbb{G})$, with c also depending on U. Now we need to recover a full derivative with respect to right invariant vector fields. To do that we introduce higher order difference operators and adapted seminorms.

Definition 3.9 *For $m = 1, 2, 3, 4, ...,$ let*

$$\Delta_h^m = \underbrace{\Delta_h \Delta_h ... \Delta_h}_{m \ times}.$$

$$\widetilde{\Delta}_h^m = \underbrace{\widetilde{\Delta}_h \widetilde{\Delta}_h ... \widetilde{\Delta}_h}_{m \ times}.$$

Then, for $\alpha > 0$ and $f \in L^2(\mathbb{G})$ we define the semi-norms

$$|f|_{m,\alpha} = \sup \left\{ \frac{\left\| \Delta_h^m f \right\|_{L^2(\mathbb{G})}}{\left\| h \right\|^\alpha} : h = \mathrm{Exp}\left(t X_i \right) \ for \ i = 1, 2, ..., q, \ t \in \mathbb{R} \ s.t. \ 0 < \left\| h \right\| \leqslant 1 \right\}$$

$$|f|_{m,\alpha}^R = \sup \left\{ \frac{\left\| \widetilde{\Delta}_h^m f \right\|_{L^2(\mathbb{G})}}{\left\| h \right\|^\alpha} : h = \mathrm{Exp}\left(t X_i \right) \ for \ i = 1, 2, ..., q, t \in \mathbb{R} \ s.t. \ 0 < \left\| h \right\| \leqslant 1 \right\}.$$

We also set for convenience

$$|f|_0 = |f|_0^R = \left\| f \right\|_{L^2(\mathbb{G})}$$

$$|f|_m = |f|_{m,m}$$

$$|f|_m^R = |f|_{m,m}^R.$$

Remark 3.10 *From Proposition 3.7 we read in particular that*

$$|f|_1 \leqslant c \left\| f \right\|_{W_X^{1,2}(\mathbb{G})}$$

$$|f|_1^R \leqslant c \left\| f \right\|_{W_{X^R}^{1,2}(\mathbb{G})}$$

for any $f \in W_X^{1,2}(\mathbb{G})$ or $W_{X^R}^{1,2}(\mathbb{G})$, respectively. It is not difficult to extend this bound to higher order seminorms:

Theorem 3.11 *Let $f \in L^2(\mathbb{G})$ and let $m = 1, 2, ...$ Then:*
1. If $f \in W_X^{m,2}(\mathbb{G})$ then

$$\sum_{k=0}^{m} |f|_k \leqslant c \, \|f\|_{W_X^{m,2}(\mathbb{G})} \tag{3.7}$$

Analogously,
2. If $f \in W_{X^R}^{m,2}(\mathbb{G})$ then

$$\sum_{k=0}^{m} |f|_k^R \leqslant c \, \|f\|_{W_{X^R}^{m,2}(\mathbb{G})} \, . \tag{3.8}$$

Proof. Let us prove (3.7), the proof of (3.8) being similar. Let $f \in W_X^{m,2}(\mathbb{G})$ and let $h = \text{Exp}(tX_j)$, for some $j = 1, 2, ..., q$, and $0 < \|h\| \leqslant 1$. First of all observe that a simple induction argument shows that

$$\Delta_h^m f(x) = \int_0^t \cdots \int_0^t X_j^m f(x \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_m X_j)) \, ds_1 \cdots ds_m.$$
$$\tag{3.9}$$

Indeed for $m = 1$ we have

$$\Delta_h f(x) = f(x \circ \text{Exp}(tX_j)) - f(x)$$
$$= \int_0^t \frac{d}{ds} f(x \circ \text{Exp}(sX_j)) \, ds$$
$$= \int_0^t X_j f(x \circ \text{Exp}(sX_j)) \, ds$$

where the last identity follows by (2.8) and the definition of exponential of a vector field.

Assume now (3.9) holds for some m. Since $\text{Exp}(sX_j)$ and $\text{Exp}(tX_j)$ commute we have

$$\Delta_h^{m+1} f(x) = \Delta_h^m f(x \circ h) - \Delta_h^m f(x)$$
$$= \int_0^t \cdots \int_0^t \left[X_j^m f(x \circ h \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_m X_j)) \right.$$
$$\left. - X_j^m f(x \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_m X_j)) \right] ds_1 \cdots ds_m$$
$$= \int_0^t \cdots \int_0^t \int_0^t \frac{d}{ds} \left[X_j^m f(x \circ \text{Exp}(sX_j) \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_m X_j)) \right] ds \, ds_1 \cdots ds_m$$
$$= \int_0^t \cdots \int_0^t \int_0^t \frac{d}{ds} \left[X_j^m f(x \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_m X_j) \circ \text{Exp}(sX_j)) \right] ds_1 \cdots ds_m ds$$
$$= \int_0^t \cdots \int_0^t X_j^{m+1} f(x \circ \text{Exp}(s_1 X_j) \circ \cdots \circ \text{Exp}(s_{m+1} X_j)) \, ds_1 \cdots ds_{m+1}.$$

Then, since $\|h\| \leqslant ct$, by Minkowski's inequality for integrals and the translation invariance of Lebesgue measure on $\mathbb{G}$, we obtain

$$\|\Delta_h^m f\|_{L^2(\mathbb{G})}$$
$$\leqslant \int_0^{|t|} \cdots \int_0^{|t|} \left\|X_j^m f\left(x \circ \operatorname{Exp}\left(s_1 X_j\right) \circ \cdots \circ \operatorname{Exp}\left(s_m X_j\right)\right)\right\|_{L^2(\mathbb{G})} ds_1 \cdots ds_m$$
$$= \int_0^{|t|} \cdots \int_0^{|t|} \left\|X_j^m f\right\|_{L^2(\mathbb{G})} ds_1 \cdots ds_m \leqslant |t|^m \left\|X_j^m f\right\|_{L^2(\mathbb{G})} \leqslant c \|h\|^m \|f\|_{W_X^{m,2}(\mathbb{G})}.$$

■

Remark 3.12 *The same proof also shows that*

$$\|\Delta_h^m f\|_{L^\infty(\mathbb{G})} \leqslant c \|h\|^m \|f\|_{W_X^{m,\infty}(\mathbb{G})}, \tag{3.10}$$

an inequality that we will sometimes apply in the following.

So far we have seen that Sobolev norms control the corresponding seminorms defined by finite difference operators. Let us prove a converse result, for the case $m = 1$:

Proposition 3.13 *Let $f \in L^2(\mathbb{G})$.*
1. If $|f|_1 < \infty$ then $f \in W_X^{1,2}(\mathbb{G})$, with

$$\|X_j f\|_2 \leqslant C |f|_1 \ \text{ for } j = 1, 2, ..., q.$$

2. If $|f|_1^R < \infty$ then $f \in W_{X^R}^{1,2}(\mathbb{G})$, with

$$\left\|X_j^R f\right\|_2 \leqslant C |f|_1^R \ \text{ for } j = 1, 2, ..., q.$$

Proof. Let $\phi \in C_0^\infty(\mathbb{G})$ and $h = \operatorname{Exp}(tX_j)$ for some t, $j = 1, 2, ..., q$. Then

$$\frac{1}{\|h\|} \left| \int \Delta_h f(x) \phi(x) \, dx \right| \leqslant |f|_1 \|\phi\|_2.$$

On the other hand since $h^{-1} = \operatorname{Exp}(-tX_j)$, we have

$$\frac{1}{\|h\|} \int \Delta_h f(x) \phi(x) \, dx = \frac{-1}{\|h\|} \int f(x) \Delta_{h^{-1}} \phi(x) \, dx$$
$$= \int f(x) \frac{\phi(x \circ \operatorname{Exp}(-tX_j)) - \phi(x)}{-t} dx$$

hence

$$\left| \int f(x) \frac{\phi(x \circ \operatorname{Exp}(-tX_j)) - \phi(x)}{-t} dx \right| \leqslant |f|_1 \|\phi\|_2$$

and for $t \to 0$

$$\left| - \int f X_j \phi \, dx \right| \leqslant |f|_1 \|\phi\|_2$$

18

which implies that there exists, in weak sense, $X_j f \in L^2(\mathbb{G})$, with $\|X_j f\|_2 \leqslant |f|_1$. The proof of 2. is similar. ∎

Let us note an important difference between the seminorms $|f|_m$ for $m = 1$ and for higher m. While for $m = 1$ the finiteness of $|f|_1$ for an L^2 function implies the finiteness of the Sobolev norm $W_X^{1,2}$, for higher m the seminorms $|f|_m$ are strictly weaker than $\|\cdot\|_{W_X^{m,2}}$. However, the equivalence in case $m = 1$ will be enough for our purposes.

3.2 Regularity estimates

We split the proof of Theorem 3.4 into several steps. Let us start rewriting (3.6) in the form

$$|f|_{1,1/s}^R \leqslant c\left(\|\mathcal{L}f\|_2 + \|f\|_2\right). \tag{3.11}$$

Recall that the constant c depends on some compact U containing the support of f.

As a first step we will show, with an iterative argument, that this estimate can be extended to higher order seminorms $|\cdot|_{m,m/s}^R$ (Theorem 3.15), which in turn allows to control $|\cdot|_{1,1}^R$ by $|\cdot|_{m,m/s}^R$, for m large enough (see Lemma 3.16). This is the estimate that is needed to implement the idea described at the beginning of Section 3.

We start with some Lemmas.

Lemma 3.14 *For every $m \in \mathbb{N}$ and $h \in \mathbb{G}$ we have*

$$\widetilde{\Delta}_h^m (fg) = \sum_{k=0}^m \binom{m}{k} \widetilde{\Delta}_h^k f \; L_h^k \widetilde{\Delta}_h^{m-k} g \tag{3.12}$$

where L_h is the left translation operator

$$L_h f(x) = f(h \circ x). \tag{3.13}$$

Proof. When $m = 1$ it is immediate to see that

$$\widetilde{\Delta}_h (fg)(x) = \widetilde{\Delta}_h f(x) L_h g(x) + f(x) \widetilde{\Delta}_h g(x).$$

The case $m > 1$ can be obtain by induction. We omit the details. ∎

Let us come to the extension of (3.11) to higher order difference operators, also in a localized version.

Theorem 3.15 *Let $\zeta_0, \zeta \in C_0^\infty(\mathbb{G})$ with $\zeta_0 \prec \zeta$. For every $m \in \mathbb{N}$ the exists $c > 0$ such that if $f \in W_{X,loc}^{1,2}(\mathbb{G})$ and $\mathcal{L}f \in L_{loc}^2(\mathbb{G})$ then*

$$|\zeta_0 f|_{m,m/s}^R \leqslant c\left(\sum_{j=0}^{m-1} |\zeta \mathcal{L} f|_j^R + \|\zeta f\|_2\right), \tag{3.14}$$

whenever the right hand side is finite.

19

Proof. Given two cut-off functions $\zeta_0, \zeta \in C_0^\infty(\mathbb{G})$ satisfying $\zeta_0 \prec \zeta$, for any fixed positive integer n we can always construct intermediate cut-off functions $\zeta_0 \prec \zeta_1 \prec \ldots \prec \zeta_n = \zeta$. Also, for a fixed positive integer m there exists $\varepsilon > 0$ such that when $\|h\| < \varepsilon$ we have $\Delta_h^{m-1} \zeta_i \prec \zeta_{i+1}$ for $i = 0, 1, \ldots, n$. The number n will vary in different steps of the proof.

We will prove the theorem by induction on m. Let $m = 1$. Applying (3.11) to $\zeta_0 f$ we obtain

$$|\zeta_0 f|_{1,1/s}^R \leqslant c \left(\|\mathcal{L}(\zeta_0 f)\|_2 + \|\zeta_0 f\|_2 \right)$$

$$\leqslant c \left(\|\zeta_0 \mathcal{L} f\|_2 + \left\| \sum_{i=1}^q X_i \zeta_0 X_i f \right\|_2 + \|f \mathcal{L} \zeta_0\|_2 + \|\zeta_0 f\|_2 \right) \qquad (3.15)$$

$$\leqslant c \left(\|\zeta_0 \mathcal{L} f\|_2 + \sum_{i=1}^q \|X_i (\zeta_1 f)\|_2 + \|\zeta_1 f\|_2 \right).$$

For the term that contains the first order derivatives we have

$$\sum_{i=1}^q \|X_i (\zeta_1 f)\|_2^2 = \sum_{i=1}^q \int X_i (\zeta_1 f) X_i (\zeta_1 f) \, dx = -\int \mathcal{L}(\zeta_1 f)(\zeta_1 f) \, dx$$

$$= -\int \zeta_1 \mathcal{L} f \cdot \zeta_1 f dx - \int (\mathcal{L}\zeta_1) f^2 \zeta_1 dx - 2 \sum_{i=1}^q \int (X_i \zeta_1)(X_i f) \zeta_1 f dx$$

$$\leqslant \|\zeta_1 \mathcal{L} f\|_2 \|\zeta_1 f\|_2 + c \|\zeta_2 f\|_2^2 - 2 \sum_{i=1}^q \int (X_i \zeta_1)(X_i (\zeta_1 f)) f dx + 2 \sum_{i=1}^q \int (X_i \zeta_1)^2 f^2 dx$$

$$\leqslant \|\zeta_1 \mathcal{L} f\|_2 \|\zeta_1 f\|_2 + c \|\zeta_2 f\|_2^2 + 2 \sum_{i=1}^q \|X_i (\zeta_1 f)\|_2 \|\zeta_2 f\|_2$$

$$\leqslant c \left(\|\zeta_1 \mathcal{L} f\|_2^2 + \|\zeta_2 f\|_2^2 \right) + \varepsilon \sum_{i=1}^q \|X_i (\zeta_1 f)\|_2^2 + \frac{c}{\varepsilon} \|\zeta_2 f\|_2^2$$

therefore

$$\sum_{i=1}^q \|X_i (\zeta_1 f)\|_2 \leqslant c \left(\|\zeta_2 \mathcal{L} f\|_2 + \|\zeta_2 f\|_2 \right) \qquad (3.16)$$

which inserted into (3.15) gives

$$|\zeta_0 f|_{1,1/s}^R \leqslant c \left(\|\zeta_2 \mathcal{L} f\|_2 + \|\zeta_2 f\|_2 \right). \qquad (3.17)$$

Assume now that (3.14) holds for every $m' < m$ and let $h = \mathrm{Exp}(t X_j)$ for some $i = 1, \ldots, q$ such that $0 < \|h\| \leqslant \varepsilon$. Since

$$\left\| \widetilde{\Delta}_h^m (\zeta_0 f) \right\|_2 = \left\| \widetilde{\Delta}_h \widetilde{\Delta}_h^{m-1} (\zeta_0 f) \right\|_2 = \left\| \widetilde{\Delta}_h (\zeta_1 g_h) \right\|_2$$

with $g_h = \widetilde{\Delta}_h^{m-1} (\zeta_0 f) \prec \zeta_1$ we have

$$\left\| \widetilde{\Delta}_h^m (\zeta_0 f) \right\|_2 \leqslant \|h\|^{1/s} |\zeta_1 g_h|_{1,1/s}^R$$

$$\leqslant c \|h\|^{1/s} \left(\|\zeta_2 \mathcal{L} g_h\|_2 + \|\zeta_2 g_h\|_2 \right),$$

20

where we have applied our theorem for $m = 1$. Since $\mathcal{L}$ and $\widetilde{\Delta}_h$ commute we obtain

$$
\left\|\widetilde{\Delta}_h^m \left(\zeta_0 f\right)\right\|_2 \leqslant c \left\|h\right\|^{1/s} \left(\left\|\mathcal{L}\widetilde{\Delta}_h^{m-1}\left(\zeta_0 f\right)\right\|_2 + \left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0 f\right)\right\|_2\right)
$$

$$
= c \left\|h\right\|^{1/s} \left(\left\|\widetilde{\Delta}_h^{m-1}\mathcal{L}\left(\zeta_0 f\right)\right\|_2 + \left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0 f\right)\right\|_2\right)
$$

$$
= c \left\|h\right\|^{1/s} \left(\left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0\mathcal{L}f + 2\sum_j X_j\zeta_0 \cdot X_j f + f\mathcal{L}\zeta_0\right)\right\|_2 + \left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0 f\right)\right\|_2\right)
$$

$$ \tag{3.18} $$

$$
\leqslant c \left\|h\right\|^{1/s} \left(\left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0\mathcal{L}f\right)\right\|_2 + \sum_j \left\|\widetilde{\Delta}_h^{m-1}\left(X_j\zeta_0\ X_j f\right)\right\|_2\right.
$$

$$
\left. + \left\|\widetilde{\Delta}_h^{m-1}\left(f\mathcal{L}\zeta_0\right)\right\|_2 + \left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0 f\right)\right\|_2\right).
$$

From Lemma 3.14 we obtain

$$
\left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0\mathcal{L}f\right)\right\|_2 \leqslant c \sum_{k=0}^{m-1} \left\|\widetilde{\Delta}_h^k\left(\mathcal{L}f\right)\ L_h^k\widetilde{\Delta}_h^{m-k-1}\zeta_0\right\|_2
$$

$$
\leqslant c \sum_{k=0}^{m-1} \left\|L_h^k\widetilde{\Delta}_h^{m-k-1}\zeta_0\right\|_\infty \left\|\widetilde{\Delta}_h^k\left(\zeta_1\mathcal{L}f\right)\right\|_2.
$$

and by (3.10)

$$
\left\|\widetilde{\Delta}_h^{m-1}\left(\zeta_0\mathcal{L}f\right)\right\|_2 \leqslant c \sum_{k=0}^{m-1} \left\|h\right\|^{m-k-1} \left\|\widetilde{\Delta}_h^k\left(\zeta_1\mathcal{L}f\right)\right\|_2 \tag{3.19}
$$

$$
\leqslant c \left\|h\right\|^{m-1} \sum_{k=0}^{m-1} \left|\zeta_1\mathcal{L}f\right|_k^R.
$$

As for the second term in (3.18) we have

$$
\left\|\widetilde{\Delta}_h^{m-1}\left(X_j\zeta_0\ X_j f\right)\right\|_2 \leqslant c \sum_{k=0}^{m-1} \left\|\widetilde{\Delta}_h^k\left(X_j\zeta_0\right)\ L_h^k\widetilde{\Delta}_h^{m-1-k}X_j f\right\|_2 \tag{3.20}
$$

$$
\leqslant c \sum_{k=0}^{m-1} \left\|\widetilde{\Delta}_h^k\left(X_j\zeta_0\right)\right\|_\infty \left\|X_j\left(\zeta_1\widetilde{\Delta}_h^{m-1-k}f\right)\right\|_2,
$$

where we used the fact that X_j and $\widetilde{\Delta}_h^{m-1-k}$ commute and $\zeta_0 \prec \zeta_1$.

Using (3.16) and the fact that $\mathcal{L}$ and $\widetilde{\Delta}_h^{m-1-k}$ commute we have

$$
\begin{aligned}
\left\| X_j \left(\zeta_1 \widetilde{\Delta}_h^{m-1-k} f \right) \right\|_2 &\leqslant c \left(\left\| \zeta_2 \widetilde{\Delta}_h^{m-1-k} \mathcal{L} f \right\|_2 + \left\| \zeta_2 \widetilde{\Delta}_h^{m-1-k} f \right\|_2 \right) \\
&\leqslant c \left(\left\| \widetilde{\Delta}_h^{m-1-k} \left(\zeta_4 \mathcal{L} f \right) \right\|_2 + \left\| \widetilde{\Delta}_h^{m-1-k} \left(\zeta_3 f \right) \right\|_2 \right) \\
&\leqslant c \left\| h \right\|^{m-1-k} |\zeta_4 \mathcal{L} f|_{m-1-k}^R \\
&\quad + c \left\| h \right\|^{(m-1-k)/s} |\zeta_3 f|_{m-1-k,(m-1-k)/s}^R .
\end{aligned}
$$

Using the inductive assumption for any $k = 0, \ldots, m-1$, we obtain

$$
|\zeta_3 f|_{m-1-k,(m-1-k)/s}^R \leqslant c \left(\sum_{i=0}^{m-k-2} |\zeta_4 \mathcal{L} f|_i^R + \|\zeta_4 f\|_2 \right)
$$

(here we implicitly assume that for $k = m-1$ the summation is empty). Therefore

$$
\begin{aligned}
\left\| X_j \left(\zeta_1 \widetilde{\Delta}_h^{m-k-1} f \right) \right\|_2 &\leqslant c \left\| h \right\|^{m-1-k} |\zeta_4 \mathcal{L} f|_{m-1-k}^R \\
&\quad + c \left\| h \right\|^{(m-1-k)/s} \left(\sum_{i=0}^{m-k-2} |\zeta_4 \mathcal{L} f|_i^R + \|\zeta_4 f\|_2 \right).
\end{aligned}
$$

Since $\left\| \widetilde{\Delta}_h^k \zeta_j \right\|_\infty \leqslant c \left\| h \right\|^k$ using (3.20) we obtain

$$
\begin{aligned}
\left\| \widetilde{\Delta}_h^{m-1} \left(X_j \zeta \, X_j f \right) \right\|_2 &\leqslant c \left\| h \right\|^{m-1} \sum_{k=0}^{m-1} |\zeta_4 \mathcal{L} f|_k^R \\
&\quad + c \sum_{k=0}^{m-1} \left\| h \right\|^k \left\| h \right\|^{(m-1-k)/s} \left(\sum_{i=0}^{m-2} |\zeta_4 \mathcal{L} f|_i^R + \|\zeta_4 f\|_2 \right) \\
&\leqslant c \left\| h \right\|^{m-1} \sum_{k=0}^{m-1} |\zeta_4 \mathcal{L} f|_k^R \qquad\qquad (3.21) \\
&\quad + c \left\| h \right\|^{(m-1)/s} \left(\sum_{k=0}^{m-2} |\zeta_4 \mathcal{L} f|_k^R + \|\zeta_4 f\|_2 \right) \\
&\leqslant c \left\| h \right\|^{(m-1)/s} \left(\sum_{k=0}^{m-1} |\zeta_4 \mathcal{L} f|_k^R + \|\zeta_4 f\|_2 \right).
\end{aligned}
$$

As to the last two terms in (3.18) we can bound

$$
\begin{aligned}
&\left\| \widetilde{\Delta}_h^{m-1} \left(f \mathcal{L} \zeta_0 \right) \right\|_2 + \left\| \widetilde{\Delta}_h^{m-1} \left(\zeta_0 f \right) \right\|_2 \qquad\qquad (3.22) \\
&\leqslant \left\| h \right\|^{(m-1)/s} \left(|f \mathcal{L} \zeta_0|_{m-1,(m-1)/s}^R + |\zeta_0 f|_{m-1,(m-1)/s}^R \right)
\end{aligned}
$$

22

Inserting (3.19), (3.21) and (3.22) in (3.18) we get

$$\left\|\widetilde{\Delta}_h^m\left(\zeta_0 f\right)\right\|_2 \leqslant c \left\|h\right\|^{m/s} \left(\sum_{k=0}^{m-1} \left|\zeta_4 \mathcal{L} f\right|_k^R + \left\|\zeta_4 f\right\|_2 \right.$$
$$\left. + \left|f\mathcal{L}\zeta_0\right|_{m-1,(m-1)/s}^R + \left|\zeta_0 f\right|_{m-1,(m-1)/s}^R \right).$$

By the inductive assumption, we obtain

$$\left|f\mathcal{L}\zeta_0\right|_{m-1,(m-1)/s}^R + \left|\zeta_0 f\right|_{m-1,(m-1)/s}^R \leqslant c \left(\sum_{j=1}^{m-2} \left|\zeta_2 \mathcal{L} f\right|_j^R + \left\|\zeta_2 f\right\|_2 \right)$$

so that

$$\left\|\widetilde{\Delta}_h^m\left(\zeta_0 f\right)\right\|_2 \leqslant c \left\|h\right\|^{m/s} \left(\sum_{i=0}^{m-1} \left|\zeta_4 \mathcal{L} f\right|_i + \left\|\zeta_4 f\right\|_2 \right).$$

Recall now that we chose $h = \mathrm{Exp}\left(tX_i\right)$ with $0 < \left\|h\right\| \leqslant \varepsilon$. To obtain (3.14) it is enough to observe that when $\left\|h\right\| > \varepsilon$ one has the trivial estimate

$$\frac{\left\|\widetilde{\Delta}_h^m\left(\zeta_0 f\right)\right\|_2}{\left\|h\right\|} \leqslant c \frac{\left\|\zeta_0 f\right\|_2}{\varepsilon} \leqslant c \left\|\zeta_4 f\right\|_2.$$

∎

The previous theorem allows to control the regularity of a function f, measured using difference operators, by means of the regularity of $\mathcal{L}f$. Unfortunately we cannot apply directly Lemma 3.13 to bound $X_i^R f$ since it requires an estimate for the first order difference of f while (3.14) contains higher order differences. A result of M.A. Marchaud allows to bound the first order difference operator using higher order one (see e.g. [3, Chapter 2, Theorem 8.1]). In the following proposition we adapt this classical result to our setting.

Proposition 3.16 (Marchaud inequality on Carnot groups) *Let $f \in L^2\left(\mathbb{G}\right)$ and assume there exist $A > 0$, an integer $m > 1$ and $1 < \alpha < 2$ such that for every $\left\|h\right\| \leqslant 1$ we have*

$$\left\|\widetilde{\Delta}_h^m f\right\|_2 \leqslant A \left\|h\right\|^\alpha. \tag{3.23}$$

Then there exists $c > 0$, independent of f, such that for $\left\|h\right\| \leqslant 1$

$$\left\|\widetilde{\Delta}_h f\right\|_2 \leqslant c \left(A + \left\|f\right\|_2\right) \left\|h\right\|.$$

Proof. For every integer $k \geqslant 1$ let

$$Q_k\left(x\right) = \frac{1 - 2^{-k}\left(x+1\right)^k}{x - 1}$$

and observe that $Q_k\left(x\right)$ is a polynomial satisfying

$$\left(x - 1\right)^k = \left(x - 1\right)^{k+1} Q_k\left(x\right) + 2^{-k}\left(x^2 - 1\right)^k. \tag{3.24}$$

For every $h \in \mathbb{G}$ let L_h denote the left translation operator (3.13). Then cleary

$$\widetilde{\Delta} f(x) = (L_h - I) f.$$

Also note that $L_h L_h = L_{2h}$, where from now on we will write

$$nh = \underbrace{h \circ \cdots \circ h}_{n \text{ times}}.$$

Replacing in equation (3.24) the variable x with the translation operator L_h gives

$$(L_h - I)^k = 2^{-k} (L_{2h} - I)^k + Q_k (L_h) (L_h - I)^{k+1}$$

(note that all the operators involved commute). Since $\|L_h f\|_2 = \|f\|_2$ we have $\|Q_k (L_h) g\| \leqslant M_k \|g\|_2$ where M_k only depends on the coefficients of the polynomial $Q_k (x)$. It follows that

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant 2^{-k} \left\|\widetilde{\Delta}_{2h}^k f\right\|_2 + \left\|Q_k (L_h) \widetilde{\Delta}_h^{k+1} f\right\|_2$$

$$\leqslant 2^{-k} \left\|\widetilde{\Delta}_{2h}^k f\right\|_2 + M_k \left\|\widetilde{\Delta}_h^{k+1} f\right\|_2$$

Applying n times the above inequality we get

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant M_k \sum_{j=0}^{n} 2^{-jk} \left\|\widetilde{\Delta}_{2^j h}^{k+1} f\right\|_2 + 2^{-k(n+1)} \left\|\widetilde{\Delta}_{2^{n+1} h}^k f\right\|_2. \tag{3.25}$$

We will now prove that

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant \begin{cases} c_k (A + \|f\|_2) \|h\|^\alpha & \text{for } 1 < k \leqslant m, \\[2mm] c_k (A + \|f\|_2) \|h\| & \text{for } k = 1, \end{cases} \tag{3.26}$$

where m is the integer appearing in (3.23). Clearly (3.23) gives the above bound when $k = m$. We now assume that (3.26) holds for $k + 1$ and we will prove it for k (assuming $k \geqslant 1$). For $\|h\| \leqslant 1$, using (3.25) we obtain

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant M_k \sum_{j=0}^{n} 2^{-jk} (A + \|f\|_2) \left\|2^j h\right\|^\alpha + 2^{-k(n+1)} \left\|\widetilde{\Delta}_{2^{n+1} h}^k f\right\|_2$$

$$\leqslant M_k (A + \|f\|_2) \sum_{j=0}^{n} 2^{-jk} 2^{j\alpha} \|h\|^\alpha + 2^{-k(n+1)} 2^k \|f\|_2$$

where in the last inequality we used the fact that

$$\|nh\| = \underbrace{\|h \circ \cdots \circ h\|}_{n \text{ times}} \leqslant n \|h\|.$$

24

Choosing $n \in \mathbb{N}$ in such a way that $2^{-n} \leqslant \|h\| \leqslant 2^{-n+1}$ gives

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant M_k \left(A + \|f\|_2\right) \|h\|^\alpha \sum_{j=0}^n 2^{j(\alpha-k)} + \|h\|^k \|f\|_2.$$

Since $1 < \alpha < 2$, if $k \geqslant 2$ the sum is bounded independently of n so that

$$\left\|\widetilde{\Delta}_h^k f\right\|_2 \leqslant c_k \left(A + \|f\|_2\right) \|h\|^\alpha.$$

If $k = 1$ we obtain

$$\begin{aligned}
\left\|\widetilde{\Delta}_h f\right\|_2 &\leqslant M_1 \left(A + \|f\|_2\right) \|h\|^\alpha \sum_{j=0}^n 2^{j(\alpha-1)} + \|h\| \|f\|_2 \\
&\leqslant M_1 \left(A + \|f\|_2\right) \|h\|^\alpha 2^{(\alpha-1)(n+1)} + \|h\| \|f\|_2 \\
&\leqslant M_1 \left(A + \|f\|_2\right) \|h\|^\alpha c \|h\|^{1-\alpha} + \|h\| \|f\|_2 \\
&= c_1 \left(A \|h\| + \|f\|_2\right).
\end{aligned}$$

∎

Corollary 3.17 *Let $f \in L^2(\mathbb{G})$ and assume that for $\varepsilon \in (0,1)$ and some integer $m > 1$ the seminorm $|f|_{m,1+\varepsilon}^R$ is finite. Then*

$$|f|_1^R \leqslant c \left\{ |f|_{m,1+\varepsilon}^R + \|f\|_2 \right\}.$$

Proof. Since $|f|_{m,1+\varepsilon}^R$ is finite we can write

$$\left\|\widetilde{\Delta}_h^m f\right\|_2 \leqslant |f|_{m,1+\varepsilon}^R \|h\|^{1+\varepsilon}$$

and by the previous Lemma with $\alpha = 1 + \varepsilon$ and $A = |f|_{m,\varepsilon}^R$, we obtain

$$\left\|\widetilde{\Delta}_h f\right\|_2 \leqslant c \left(|f|_{m,1+\varepsilon}^R + \|f\|_2\right) \|h\|$$

and the thesis follows. ∎

Before proceding with the proof of Theorem 3.4 we still need the following.

Proposition 3.18 *Let $f \in L_{loc}^2(\mathbb{G})$ and assume that in the sense of distributions $\mathcal{L}f = F$ with $F \in L_{loc}^2(\mathbb{G})$, that is for every $\phi \in C_0^\infty(\mathbb{G})$,*

$$\int_{\mathbb{G}} f(x) \mathcal{L}\phi(x) \, dx = \int_{\mathbb{G}} F(x) \phi(x) \, dx.$$

If $F \in W_{X^R, loc}^{s^2, 2}(\Omega)$ then $f \in W_{X^R, loc}^{1,2}(\Omega)$ and for every $\zeta, \zeta_1 \in C_0^\infty(\Omega)$ with $\zeta \prec \zeta_1$ the following estimate holds:

$$\|\zeta f\|_{W_{X^R}^{1,2}(\mathbb{R}^N)} \leqslant c \left\{ \|\zeta_1 F\|_{W_{X^R}^{s,2}(\mathbb{R}^N)} + \|\zeta_1 f\|_{L^2(\mathbb{R}^N)} \right\}. \tag{3.27}$$

Remark 3.19 *As the reader will notice, the proof of the above proposition requires the knowledge of Theorem 3.4 for $k = 1$. In turn, the above proposition is required in the proof of the inductive step of Theorem 3.4 (but not in the proof of the case $k = 1$), so that our argument is not circular.*

Proof. Let $\phi \in C_0^\infty (\mathbb{G})$ such that

$$\phi \geqslant 0, \ \phi(x) = 0 \text{ for } \|x\| \geqslant 1 \text{ and } \int_{\mathbb{G}} \phi(x) \, dx = 1.$$

Define, for any $\varepsilon > 0$,

$$\phi_\varepsilon (x) = \varepsilon^{-Q} \phi \left(D_{\varepsilon^{-1}} x \right)$$

and

$$f_\varepsilon (x) = (\phi_\varepsilon * f)(x) = \int_{\mathbb{G}} \phi_\varepsilon (y) f \left(y^{-1} \circ x \right) dy.$$

Since by Proposition 2.3 and Remark 2.4 f_ε is smooth, we can apply to f_ε the estimate of Theorem 3.4 for $k = 1$:

$$\|\zeta f_\varepsilon\|_{W_{X^R}^{1,2} (\mathbb{G})} \leqslant c \left\{ \|\zeta_1 \mathcal{L}(f_\varepsilon)\|_{W_{X^R}^{s,2} (\mathbb{G})} + \|\zeta_1 f_\varepsilon\|_{L^2 (\mathbb{G})} \right\}. \tag{3.28}$$

By known properties of mollifiers we have

$$\|\zeta_1 f_\varepsilon\|_{L^2} \to \|\zeta_1 f\|_{L^2}.$$

Also, by Proposition 2.5, observe that $\mathcal{L}(f_\varepsilon) = \mathcal{L}(\phi_\varepsilon * f) = \phi_\varepsilon * \mathcal{L}f = (\mathcal{L}f)_\varepsilon$. Unfortunately the left convolution with ϕ_ε provides convergence in $W_X^{k,2}$ but not in $W_{X^R}^{k,2}$ and we are forced to the following rough estimate,

$$\|\zeta_1 \mathcal{L}(f_\varepsilon) - \zeta_1 \mathcal{L}f\|_{W_{X^R}^{s,2} (\mathbb{G})} \leqslant c \|\zeta_1 \mathcal{L}(f_\varepsilon) - \zeta_1 \mathcal{L}f\|_{W_X^{s^2,2} (\mathbb{G})} \tag{3.29}$$

where we used the fact that since at every point the vector fields $X_1^R, X_2^R, \ldots, X_q^R$ and their commutators up to the step s span $\mathbb{R}^N$ it is possible to represent every derivative ∂_{x_i} by means of a linear combination of right invariant differential operators of the kind $X_{j_1}^R X_{j_2}^R \cdots X_{j_k}^R$ with $k \leqslant s$. Since $\mathcal{L}f \in W_{X,loc}^{s^2,2} (\mathbb{G})$ the RHS in (3.29) vanishes as $\varepsilon \to 0^+$. In particular the RHS in (3.28) is bounded, hence the sequence ζf_ε is bounded in $W_{X^R}^{1,2} (\Omega)$. Since $W_{X^R}^{1,2} (\Omega)$ is a separable Hilbert space by Banach-Alaoglu theorem there exists a subsequence of ζf_ε converging weakly in $W_{X^R}^{1,2} (\Omega)$. This shows that $\zeta f \in W_{X^R}^{1,2} (\Omega)$. Moreover the weak convergence in $W_{X^R}^{1,2} (\Omega)$ implies

$$\|\zeta f\|_{W_{X^R}^{1,2} (\mathbb{R}^N)} \leqslant \liminf \|\zeta f_\varepsilon\|_{W_{X^R}^{1,2} (\mathbb{R}^N)} \leqslant c \left\{ \|\zeta_1 \mathcal{L}f\|_{W_{X^R}^{s,2} (\mathbb{R}^N)} + \|\zeta_1 f\|_{L^2 (\mathbb{R}^N)} \right\}$$

hence (3.27) holds. ∎

Proof of Theorem 3.4. We will prove (i) by induction on k. We start with $k = 1$. Let $f \in W^{1,2}_{X,loc}(\mathbb{G})$ such that $\mathcal{L}f = F \in W^{s^2,2}_{X^R,loc}(\Omega)$. By Corollary 3.17 and Theorem 3.15 we have

$$|\zeta f|^R_1 \leqslant c \left\{ |\zeta f|^R_{s+1,1+1/s} + \|\zeta f\|_2 \right\} \leqslant c \left\{ \sum_{j=0}^{s} |\zeta_1 \mathcal{L}f|^R_j + \|\zeta_1 f\|_2 \right\}$$

where $\zeta, \zeta_1 \in C_0^\infty(\Omega)$ such that $\zeta \prec \zeta_1$. Also, by Proposition 3.13 and Theorem 3.11

$$\|\zeta f\|_{W^{1,2}_{X^R}(\mathbb{R}^N)} \leqslant c \left\{ |\zeta f|^R_1 + \|\zeta f\|_2 \right\}$$

$$\leqslant c \left\{ \sum_{j=0}^{s} |\zeta_1 \mathcal{L}f|^R_j + \|\zeta_1 f\|_2 \right\} \leqslant c \left\{ \|\zeta_1 \mathcal{L}f\|_{W^{s,2}_{X^R}(\mathbb{R}^N)} + \|\zeta_1 f\|_2 \right\}$$

which is the case $k = 1$.

Assume now that (3.5) holds up to the integer k and let $f \in W^{1,2}_{X,loc}(\mathbb{G})$ such that $\mathcal{L}f \in W^{k+s^2,2}_{X^R,loc}(\Omega)$. By the inductive assumption we know that $f \in W^{k,2}_{X^R,loc}(\mathbb{G})$. Let X^R_I be any right invariant differential operator with $|I| \leqslant k$. Then $X^R_I f \in L^2_{loc}(\Omega)$ and since X^R_I and $\mathcal{L}$ commute we have, in the sense of distributions

$$\mathcal{L}\left(X^R_I f\right) = X^R_I \mathcal{L}f \in W^{s^2,2}_{X^R,loc}(\Omega).$$

It follows that we can apply Proposition 3.18 to $X^R_I f$ and conclude that $X^R_I f \in W^{1,2}_{X^R,loc}(\Omega)$ and satisfies the estimate

$$\left\|\zeta X^R_I f\right\|_{W^{1,2}_{X^R}(\mathbb{R}^N)} \leqslant c \left\{ \left\|\zeta_1 X^R_I \mathcal{L}f\right\|_{W^{s,2}_{X^R}(\mathbb{R}^N)} + \left\|\zeta_1 X^R_I f\right\|_{L^2(\mathbb{R}^N)} \right\}.$$

This means that $f \in W^{k+1,2}_{X^R,loc}(\Omega)$ and, introducing other cutoff functions such that $\zeta_0 \prec \zeta \prec \zeta_1 \prec \zeta_2 \prec \zeta_3$ and exploiting the inductive assumption,

$$\|\zeta_0 f\|_{W^{k+1,2}_{X^R}(\mathbb{R}^N)} = \|\zeta_0 f\|_{W^{k,2}_{X^R}(\mathbb{R}^N)} + \sum_{|I|=k} \left\|X^R_I (\zeta_0 f)\right\|_{W^{1,2}_{X^R}(\mathbb{R}^N)}$$

$$\leqslant c \left\{ \|\zeta_1 f\|_{W^{k,2}_{X^R}(\mathbb{R}^N)} + \sum_{|I|=k} \left\|\zeta_1 X^R_I f\right\|_{W^{1,2}_{X^R}(\mathbb{R}^N)} \right\}$$

$$\leqslant c \left\{ \|\zeta_2 \mathcal{L}f\|_{W^{s+k-1,2}_{X^R}(\mathbb{R}^N)} + \sum_{|I|=k} \left(\left\|\zeta_2 X^R_I \mathcal{L}f\right\|_{W^{s,2}_{X^R}(\mathbb{R}^N)} + \left\|\zeta_2 X^R_I f\right\|_{L^2(\mathbb{R}^N)} \right) \right\}$$

$$\leqslant c \left\{ \|\zeta_3 \mathcal{L}f\|_{W^{s+k,2}_{X^R}(\mathbb{R}^N)} + \|\zeta_3 f\|_{L^2(\mathbb{R}^N)} \right\}$$

so we are done. Part (ii) of the theorem follows from the fact that

$$\bigcap_{k=1}^{\infty} W^{k,2}_{X^R}(\Omega') \subset C^\infty(\Omega').$$

27

■

4 Hypoellipticity of the sublaplacian

In this section we will prove the hypoellipticity result extending Theorem 3.4 from functions in $W_X^{1,2}$ to generic distributions.

Theorem 4.1 (Hypoellipticity of the sublaplacian) *Let $\Omega \subseteq \mathbb{R}^N$ be an open set and let $f \in \mathcal{D}'(\Omega)$ such that $\mathcal{L}f \in C^\infty(\Omega)$, then $f \in C^\infty(\Omega)$. That is $\mathcal{L}$ is hypoelliptic in $\mathbb{R}^N$.*

To prove this extension we will use a regularizing operator that commutes with $\mathcal{L}$, namely the convolution with a *parametrix* (that is an approximate fundamental solution) of a right invariant *elliptic* operator. Let us consider the second order differential operator

$$L^R = \sum_{j=1}^{N} \left(X_j^R \right)^2$$

built using the whole canonical base of right invariant vector fields. From the structure of the vector fields X_i^R we can read that at the origin the principal part of L^R coincides with the classical Laplacian. Using the fundamental solution of the Laplacian we will construct a parametrix for L^R that we will name $\widetilde{\gamma}$ (Proposition 4.2). We will then study the operator $Tf = \widetilde{\gamma} * f$, and we will show that if f is a distribution, then for K large enough $T^K f \in W_{X,loc}^{1,2}$ (Proposition 4.4 and Corollary 4.5). Since T and $\mathcal{L}$ commute (Proposition 4.6), then if $\mathcal{L}f$ is smooth also $\mathcal{L}T^K f = T^K \mathcal{L}f$ is smooth and therefore by Theorem 3.4 we see that $T^K f$ is smooth. Now, if $\widetilde{\gamma}$ were the fundamental solution of L^R then just applying K times L^R to $T^K f$ we would obtain that f is smooth. The fact that $\widetilde{\gamma}$ is only an approximate fundamental solution introduces a minor difficulty that is addressed in Lemma 4.8.

Let us start with the construction of $\widetilde{\gamma}$.

Proposition 4.2 (Parametrix of a right invariant elliptic operator) *Let $V \subset \mathbb{G}$ be a neighborhood of the origin. There exist $\widetilde{\gamma} \in C^\infty(\mathbb{G} \setminus \{0\})$ and $\omega \in C^\infty(\mathbb{G} \setminus \{0\})$, both supported in V, satisfying*

$$|\widetilde{\gamma}(x)| \leqslant \frac{c}{|x|^{N-2}} \tag{4.1}$$

$$|\partial_{x_i} \widetilde{\gamma}(x)| \leqslant \frac{c}{|x|^{N-1}} \quad i = 1, 2, ..., N \tag{4.2}$$

$$|\omega(x)| \leqslant \frac{c}{|x|^{N-2}} \tag{4.3}$$

and such that in the sense of distributions

$$L^R \widetilde{\gamma} = -\delta + \omega.$$

Proof. Let $W \subset V$ be a neighborhood of the origin, such that also $W \circ W \subset V$ and let $\eta \in C_0^\infty (W)$ identically 1 in a neighborhood of the origin. Let

$$\gamma (x) = c_N \frac{1}{|x|^{N-2}}$$

be the fundamental solution of the standard Laplacian $-\Delta$ and set $\gamma_1 = \gamma \eta$. An easy computation based on the structure of the vector fields X_i^R shows that we can write

$$L^R = \Delta + \sum_{i,j} b_{ij} \frac{\partial^2}{\partial x_i \partial x_j} + \sum_i c_i \frac{\partial}{\partial x_i} \tag{4.4}$$

where $b_{ij} (x)$ and $c_i (x)$ are polynomials and $b_{ij} (0) = 0$, see [1, Corollary 1.3.19].

For every test function $\varphi \in C_0^\infty (\mathbb{R}^N)$ we have

$$\left\langle L^R \gamma_1, \varphi \right\rangle = \int \gamma_1 (x) \, L^R \varphi (x) \, dx$$

$$= \int_{\mathbb{G}} \gamma_1 (x) \, \Delta \varphi (x) \, dx + \int_{\mathbb{G}} \gamma_1 (x) \left(\sum_{i,j} b_{ij} (x) \frac{\partial^2 \varphi (x)}{\partial x_i \partial x_j} + \sum_i c_i (x) \frac{\partial \varphi (x)}{\partial x_i} \right) dx$$

$$= \int_{\mathbb{G}} \eta \gamma (x) \, \Delta \varphi (x) \, dx + \int_{\mathbb{G}} \left(\sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\gamma_1 b_{ij}) (x) - \sum_i \frac{\partial}{\partial x_i} (\gamma_1 c_i) (x) \right) \varphi (x) \, dx.$$

Observe that the integration by parts is justified by the fact that b_{ij} vanishes at the origin so that $\gamma_1 b_{ij}$ can be differentiated twice without loosing summability. A simple computation shows that

$$\eta \gamma \Delta \varphi = \gamma \Delta (\eta \varphi) - 2 \operatorname{div} (\gamma \varphi \nabla \eta) + \gamma \varphi \Delta \eta + 2 \varphi \nabla \gamma \nabla \eta$$

and therefore

$$\int_{\mathbb{G}} \eta \gamma \Delta \varphi dx = \int_{\mathbb{G}} (\gamma \Delta (\eta \varphi) - 2 \operatorname{div} (\gamma \varphi \nabla \eta) + \gamma \varphi \Delta \eta + 2 \varphi \nabla \gamma \nabla \eta) \, dx$$

$$= -\varphi (0) - 2 \int_{\mathbb{G}} \operatorname{div} (\gamma \varphi \nabla \eta) \, dx + \int_{\mathbb{G}} \varphi (\gamma \Delta \eta + 2 \nabla \gamma \nabla \eta) \, dx.$$

Since $\gamma \varphi \nabla \eta \in C_0^\infty (\mathbb{G})$ the first integral vanishes in the last expression, so that

$$L^R \gamma_1 = -\delta + \omega_1 \tag{4.5}$$

where

$$\omega_1 = (\gamma \Delta \eta + 2 \nabla \gamma \nabla \eta) + \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} (\gamma_1 b_{ij}) - \sum_i \frac{\partial}{\partial x_i} (\gamma_1 c_i) \, .$$

Observe that for small x we have

$$|\omega_1 (x)| \leqslant \frac{c}{|x|^{N-1}} \tag{4.6}$$

and that ω_1 is supported in W. Unfortunately ω_1 is still a bit too big. We need to add a term to γ_1 to improve the approximation. Let

$$\widetilde{\gamma} = \gamma_1 + \gamma_1 * \omega_1. \tag{4.7}$$

It is immediate to check that

$$L^R \left(\gamma_1 * \omega_1 \right) = L^R \gamma_1 * \omega_1 = \left(-\delta + \omega_1 \right) * \omega_1$$
$$= -\omega_1 + \omega_1 * \omega_1$$

and therefore

$$L^R \widetilde{\gamma} = L^R \gamma_1 + L^R \left(\gamma_1 * \omega_1 \right)$$
$$= -\delta + \omega$$

where

$$\omega = \omega_1 * \omega_1.$$

We are left to check that ω satisfies the properties stated in the Proposition.

Since ω_1 is supported in W it is clear that ω is supported in $W \circ W \subset V$. It remains to show that $\omega \in C^\infty \left(\mathbb{G} \backslash \{0\} \right)$ and that $\left| \omega \left(x \right) \right| \leqslant c \left| x \right|^{2-N}$. Let us fix $x_0 \neq 0$ and a neighborhood $B_r \left(x_0 \right)$ such that $0 \notin B_r \left(x_0 \right)$. Let $\psi \in C_0^\infty \left(\mathbb{G} \right)$ be a cut-off function which is identically 1 in a neighborhood of the origin and with support disjoint from $B_r \left(x_0 \right)$. We can also assume that $x \circ w^{-1}$ and $w^{-1} \circ x$ are bounded away from zero when $w \in \operatorname{sptr} \psi$ and $x \in B_r \left(x_0 \right)$.

For any $x \in B_r \left(x_0 \right)$ we write

$$\omega \left(x \right) = \int_{\mathbb{G}} \omega_1 \left(x \circ w^{-1} \right) \ \omega_1 \left(w \right) \psi \left(w \right) dw + \int_{\mathbb{G}} \omega_1 \left(x \circ w^{-1} \right) \ \omega_1 \left(w \right) \psi \left(x \circ w^{-1} \right) dw$$
$$+ \int_{\mathbb{G}} \omega_1 \left(x \circ w^{-1} \right) \ \omega_1 \left(w \right) \left[1 - \psi \left(w \right) - \psi \left(x \circ w^{-1} \right) \right] dw$$
$$= I_1 \left(x \right) + I_2 \left(x \right) + I_3 \left(x \right).$$

Let us show that the three terms are smooth.

Let $\mathcal{P}^R$ be any right invariant differential operator, since $\omega \left(x \circ w^{-1} \right)$ is not singular on the support of ψ,

$$\mathcal{P}^R I_1 \left(x \right) = \int_{\mathbb{G}}^R \mathcal{P}^R \omega_1 \left(x \circ w^{-1} \right) \ \omega_1 \left(w \right) \psi \left(w \right) dw$$

so that $I_1 \left(x \right)$ is smooth. Similarly, if $\mathcal{P}$ is any left invariant differential operator we can write

$$\mathcal{P} I_2 \left(x \right) = \int_{\mathbb{G}} \omega_1 \left(w \right) \ \mathcal{P} \omega_1 \left(w^{-1} \circ x \right) \psi \left(w \right) dw$$

and $I_2 \left(x \right)$ is smooth. Finally observe that in I_3 there are no singular terms, we can easily differentiate under the integral sign and therefore also I_3 is smooth.

As to the growth of ω, using (4.6) and (2.4) we easily obtain

$$|\omega(x)| \leqslant \int_V |x \circ w^{-1}|^{1-N} \ |w|^{1-N} \, dw$$

$$\leqslant c \int_V |x-w|^{1-N} \ |w|^{1-N} \, dw \leqslant c \, |x|^{2-N} \, .$$

Finally, let us prove (4.1) and (4.2). Recall that $\widetilde{\gamma} = \gamma_1 + \gamma_1 * \omega_1$ with $\gamma_1(x) = c_N \frac{1}{|x|^{N-2}} \eta(x)$. Since γ_1 and ω_1 are compactly supported, also $\widetilde{\gamma}$ is compactly supported and

$$|\widetilde{\gamma}(x)| \leqslant \frac{c}{|x|^{N-2}} + c \int_W \frac{1}{|x \circ w^{-1}|^{N-2}} \frac{1}{|w|^{N-2}} dw$$

$$\leqslant \frac{c}{|x|^{N-2}} + c \int_W \frac{1}{|x-w^{-1}|^{N-2}} \frac{1}{|w|^{N-2}} dw$$

$$\leqslant \frac{c}{|x|^{N-2}} \, .$$

Moreover, since,

$$|\partial_{x_i} \gamma_1(x)| \leqslant c_N \frac{1}{|x|^{N-1}}$$

and

$$\partial_{x_i} \widetilde{\gamma}(x) = \partial_{x_i} \gamma_1(x) + \partial_{x_i} (\gamma_1 * \omega_1)(x) \, .$$

we have

$$\partial_{x_i} (\gamma_1 * \omega_1)(x) = \int_{\mathbb{G}} \partial_{x_i} \left[\gamma_1 \left(x \circ w^{-1} \right) \right] \omega_1(w) \, dw$$

$$= \int_{\mathbb{G}} \sum_{k=1}^{N} (\partial_{x_k} \gamma_1) \left(x \circ w^{-1} \right) \partial_{x_i} \left[\left(x \circ w^{-1} \right)_k \right] \omega_1(w) \, dw$$

where $\partial_{x_i} \left[\left(x \circ w^{-1} \right)_k \right]$ is a polynomial. Therefore

$$\left| (\partial_{x_k} \gamma_1) \left(x \circ w^{-1} \right) \right| \leqslant \frac{c}{|x \circ w^{-1}|^{N-1}} \leqslant \frac{c}{|x-w|^{N-1}}$$

near the pole; hence

$$|\partial_{x_i} (\gamma_1 * \omega_1)(x)| \leqslant c \int_W \frac{1}{|x-w|^{N-1}} \frac{1}{|w|^{N-1}} dw \leqslant \frac{c}{|x|^{N-2}} \, .$$

∎

Let us now consider three open sets in $\mathbb{G}$, $\Omega' \Subset \Omega'' \Subset \Omega$ and let V be a neighborhood of the origin such that $V^{-1} \circ \Omega' \subset \Omega''$. Define $\widetilde{\gamma}$ as in Proposition 4.2, with $\widetilde{\gamma}$ supported in V. The convolution with $\widetilde{\gamma}$ defines a regularizing operator that acts on functions $f \in L^1_{loc}(\Omega)$ as follows. For every $x \in \Omega'$ we set

$$Tf(x) = \widetilde{\gamma} * f(x) = \int_{\mathbb{G}} \widetilde{\gamma} \left(x \circ y^{-1} \right) f(y) \, dy. \tag{4.8}$$

31

Note that
$$T : L^1(\Omega'') \longrightarrow L^1(\Omega')$$

Namely, for $x \in \Omega'$ and $x \circ y^{-1} \in \operatorname{sprt} \widetilde{\gamma}$, the point $y = \left(x \circ y^{-1} \right)^{-1} \circ x$ ranges in $V^{-1} \circ \Omega' \subset \Omega''$, hence introducing characteristic functions,

$$(\chi_{\Omega'} T f)(x) = \int_{\mathbb{G}} (\widetilde{\gamma} \chi_V) \left(x \circ y^{-1} \right) (f \chi_{\Omega''})(y) \, dy,$$

or

$$\chi_{\Omega'} T f = \widetilde{\gamma} \chi_V * f \chi_{\Omega''} \tag{4.9}$$

which by Young's inequality gives

$$\| T f \|_{L^1(\Omega')} \leqslant \| \widetilde{\gamma} \|_{L^1(V)} \| f \|_{L^1(\Omega'')} .$$

Also, T acts on distributions $f \in \mathcal{D}'(\Omega)$ as follows. For every $\varphi \in C_0^\infty(\Omega')$ we set

$$\langle T f, \varphi \rangle = \langle f, T^* \varphi \rangle \tag{4.10}$$

where

$$T^* \varphi(y) = \int_{\mathbb{G}} \widetilde{\gamma} \left(x \circ y^{-1} \right) \varphi(x) \, dx.$$

We claim that

$$T : \mathcal{D}'(\Omega) \longrightarrow \mathcal{D}'(\Omega').$$

Observe that the assumption on V implies that $T^* \varphi$ is a test function in Ω'. Namely, for $x \in \operatorname{sprt} \varphi$ and $x \circ y^{-1} \in \operatorname{sprt} \widetilde{\gamma}$ the point y ranges in $\Omega'' \Subset \Omega$. By Remark 2.4, $T^* \varphi$ is smooth and therefore the pairing (4.10) is well defined.

Remark 4.3 *The definition of the operator T depends on the choice of the neighborhood V used to define $\widetilde{\gamma}$. In the following, in order to stress this dependence, we will also write T_V.*

Proposition 4.4 (Regularizing properties of T_V) *Let $\Omega' \Subset \Omega'' \Subset \Omega$. There exists a neighborhood V of the origin such that the operator T_V defined in (4.10) has the following properties.*
  (1) Let $f \in \mathcal{D}'(\Omega)$ such that $f = D^\alpha g$, for some $g \in L^1_{loc}(\Omega)$ and multiindex α. Then $T_V f \in \mathcal{D}'(\Omega')$ and

$$T_V f = \sum_{|\beta| \leqslant |\alpha| - 1} D^\beta A_\beta \quad \text{in } \Omega'$$

for suitable $A_\beta \in L^1_{loc}(\Omega')$.
  (2) Let $f \in L^p_{loc}(\Omega)$ for some $1 \leqslant p < \frac{N}{2}$, then $T_V f \in L^{p'}(\Omega')$ for every $\frac{1}{p'} > \frac{1}{p} - \frac{2}{N}$ and

$$\| T_V f \|_{L^{p'}(\Omega')} \leqslant c \| f \|_{L^p(\Omega'')} .$$

  (3) Let $f \in L^2_{loc}(\Omega)$ then $T_V f \in W_X^{1,2}(\Omega')$.
  (4) Let $f \in C^\infty(\Omega)$ then $T_V f \in C^\infty(\Omega').$

Proof. (1) Let $f = \partial_{y_i} D^{\alpha'} g$ for some α' with $|\alpha'| = |\alpha| - 1$. Then, for $\varphi \in C_0^\infty (\Omega')$

$$\langle T_V f, \varphi \rangle = \left\langle \partial_{y_i} D^{\alpha'} g(y), \int_{\mathbb{G}} \widetilde{\gamma} \left(x \circ y^{-1} \right) \varphi(x) \, dx \right\rangle$$

$$= \left\langle D^{\alpha'} g(y), \int_{\mathbb{G}} -\partial_{y_i} \left[\widetilde{\gamma} \left(x \circ y^{-1} \right) \right] \varphi(x) \, dx \right\rangle.$$

We can write

$$-\partial_{y_i} \left[\widetilde{\gamma} \left(x \circ y^{-1} \right) \right] = -\sum_{k=1}^{N} (\partial_k \widetilde{\gamma}) \left(x \circ y^{-1} \right) \partial_{y_i} \left[x \circ y^{-1} \right] = \sum_{k=1}^{N} h_k \left(x \circ y^{-1} \right) Z_k(x, y)$$

where by (4.2) $h_k(z)$ are locally integrable functions, smooth outside the pole, and Z_k are polynomials. Hence

$$\langle T_V f, \varphi \rangle = \left\langle D^{\alpha'} g(y), \int_{\mathbb{G}} \sum_{k=1}^{N} h_k \left(x \circ y^{-1} \right) Z_k(x, y) \varphi(x) \, dx \right\rangle$$

$$= \left\langle D^{\alpha'} g(y), \int_{\mathbb{G}} \sum_{k=1}^{N} h_k(w) Z_k(w \circ y, y) \varphi(w \circ y) \, dw \right\rangle$$

since the function $y \mapsto \int_{\mathbb{G}} \sum_{k=1}^{N} h_k(w) Z_k(w \circ y, y) \varphi(w \circ y) \, dw$ belongs to $C_0^\infty (\Omega)$

$$= \left\langle g(y), \int_{\mathbb{G}} \sum_{k=1}^{N} h_k(w) (-1)^{|\alpha'|} D_y^{\alpha'} [Z_k(w \circ y, y) \varphi(w \circ y)] \, dw \right\rangle$$

Now,

$$(-1)^{|\alpha'|} D_y^{\alpha'} [Z_k(w \circ y, y) \varphi(w \circ y)] = \sum_{|\beta| \leqslant |\alpha'|} (D^\beta \varphi)(w \circ y) a_{\beta, k}(w, y)$$

for suitable polynomials $a_{\beta, k}$, hence

$$\langle T_V f, \varphi \rangle = \int_{\mathbb{G}} g(y) \left(\int_{\mathbb{G}} \sum_{k=1}^{N} h_k(w) \sum_{|\beta| \leqslant |\alpha'|} (D^\beta \varphi)(w \circ y) a_{\beta, k}(w, y) \, dw \right) dy$$

$$= \int_{\mathbb{G}} g(y) \left(\int_{\mathbb{G}} \sum_{k=1}^{N} h_k \left(x \circ y^{-1} \right) \sum_{|\beta| \leqslant |\alpha'|} (D^\beta \varphi)(x) a_{\beta, k} \left(x \circ y^{-1}, y \right) dx \right) dy$$

$$= \int_{\mathbb{G}} \sum_{|\beta| \leqslant |\alpha'|} (D^\beta \varphi)(x) \sum_{k=1}^{N} \left(\int_{\mathbb{G}} g(y) h_k \left(x \circ y^{-1} \right) a_{\beta, k} \left(x \circ y^{-1}, y \right) dy \right) dx$$

Next, observe that

$$b_\beta(x) = \sum_{k=1}^{N} \left(\int_{\mathbb{G}} g(y) h_k \left(x \circ y^{-1} \right) a_{\beta, k} \left(x \circ y^{-1}, y \right) dy \right)$$

33

belongs to $L_{loc}^1 (\Omega')$, since $g \in L_{loc}^1 (\Omega)$, $h_k \in L^1$ and is compactly supported in V, $a_{\beta,k}$ are polynomials. Hence, letting

$$A_\beta (x) = (-1)^{|\beta|} b_\beta (x)$$

we can write

$$\langle T_V f, \varphi \rangle = \int_{\mathbb{G}} \sum_{|\beta| \leqslant |\alpha'|} (-1)^{|\beta|} A_\beta (x) \left(D^\beta \varphi \right) (x) \, dx$$

$$= \left\langle \sum_{|\beta| \leqslant |\alpha| - 1} D^\beta A_\beta, \varphi \right\rangle$$

with $A_\beta \in L_{loc}^1 (\Omega')$, so that $T_V f$ has the desired structure.

(2) follows from (4.9) and Young's inequality since, by (4.1), $\widetilde{\gamma} \in L^r (\mathbb{G})$ for $1 \leqslant r < \frac{N}{N-2}$.

(3) We know that any derivative $\partial_{x_i} \widetilde{\gamma}$ $(i = 1, 2, ..., N)$ is integrable and supported in V, hence each function $X_i^R \widetilde{\gamma}$ $(i = 1, 2, ..., N)$ is a linear combination with polynomial coefficients of integrable functions, compactly supported in V, so that $X_i^R \widetilde{\gamma} \in L^1 (\mathbb{G})$. Also, $f \chi_{\Omega''} \in L^2 (\mathbb{G})$ hence by Young's inequality

$$T_V f = \widetilde{\gamma} * (f \chi_{\Omega''}) \in L^2 (\mathbb{G}),$$

that is $T_V f \in L^2 (\Omega')$, and

$$\chi_{\Omega'} X_i^R T_V f = \left(X_i^R \widetilde{\gamma} \right) * (f \chi_{\Omega''}) \in L^2 (\mathbb{G}),$$

that is $X_i^R T_V f \in L^2 (\Omega')$. This holds for $i = 1, 2, ..., N$ (not just for the first q derivatives). Now, let us recall that the left invariant vector fields X_i $(i = 1, 2, ..., N)$ can be written as linear combinations with polynomial coefficients of the right invariant vector fields X_i^R. Hence by the boundedness of Ω' we also have

$$X_i T_V f \in L^2 (\Omega') \text{ for } i = 1, 2, ..., N$$

in particular $T_V f \in W_X^{1,2} (\Omega')$.

(4). Let $f \in C^\infty (\Omega)$. From

$$T_V f (x) = \int_{\mathbb{G}} \widetilde{\gamma} \left(x \circ y^{-1} \right) f (y) \, dy = \int_{\mathbb{G}} \widetilde{\gamma} (w) f \left(w^{-1} \circ x \right) dw$$

we read that for $x \in \Omega'$ and any left invariant differential operator $\mathcal{P}$ we can write

$$\mathcal{P} T_V f (x) = \int_{\mathbb{G}} \widetilde{\gamma} (w) \mathcal{P} f \left(w^{-1} \circ x \right) dw,$$

showing that $\mathcal{P} T_V f \in C^\infty (\Omega')$. ∎

Corollary 4.5 *Let $\Omega' \Subset \Omega$. For every distribution $f \in \mathcal{D}' (\Omega)$ there exist a neighborhood of the origin V and an integer K such that $(T_V)^K f \in W_X^{1,2} (\Omega')$.*

Proof. For fixed $\Omega' \Subset \Omega$ and positive integer K to be chosen later, there exists a neighborhood V of the origin such that

$$\underbrace{V \circ V \circ ... \circ V}_{K \text{ times}} \circ \Omega' \Subset \Omega.$$

Let

$$\Omega_j = \underbrace{V \circ V \circ ... \circ V}_{j \text{ times}} \circ \Omega' \text{ for } j = 1, 2, ..., K$$

$$\Omega_0 = \Omega'.$$

so that $\Omega_K \Subset \Omega$. If $f \in \mathcal{D}'(\Omega)$, then the restriction of f to Ω_K coincides with a finite sum of derivatives of locally integrable (actually, continuous) functions:

$$f = \sum_{j=1}^{M} D^{\alpha_j} g_j$$

for suitable $g_j \in L^1(\Omega_K)$ and multiindices α_j, see e.g [11, Corollary 3, p. 263]. Applying point (1) in Proposition 4.4 k_1 times, with k_1 large enough we find that

$$T_V^{k_1}(f) \in L^1(\Omega_{K-k_1}).$$

Applying point (2) in Proposition 4.4 k_2 times (with k_2 large enough) provides $T_V^{k_1+k_2}(f) \in L^2(\Omega_{K-k_1-1-k_2})$.

Finally, using point (3) we obtain

$$T_V^{k_1+1+k_2+1}(f) \in W_X^{1,2}(\Omega_{K-k_1-1-k_2-1}),$$

which is the desired result, for $K = k_1 + k_2 + 2$. ∎

Proposition 4.6 *Let $\Omega' \Subset \Omega$ and V small enough so that $V \circ \Omega' \Subset \Omega$. Then for any distribution $f \in \mathcal{D}'(\Omega)$ and every left invariant operator $\mathcal{P}$ we have*

$$\mathcal{P}T_V f = T_V \mathcal{P} f \text{ in } \mathcal{D}'(\Omega').$$

Remark 4.7 *The previous proposition can be obviously iterated writing*

$$\mathcal{P}T_V^K f = T_V^K \mathcal{P} f \text{ in } \mathcal{D}'(\Omega')$$

for any fixed positive integer K, provided V is chosen small enough to have

$$\underbrace{V \circ V \circ ... \circ V}_{K \text{ times}} \circ \Omega' \Subset \Omega.$$

Proof. This is essentially Proposition 2.3. We omit the details. ∎

Lemma 4.8 *Let $\Omega' \Subset \Omega'' \Subset \Omega$ and $f \in \mathcal{D}'(\Omega)$. There exists V small enough so that if*

$$T_V f \in C^\infty(\Omega'')$$

then $f \in W_X^{1,2}(\Omega')$.

Proof. The proof is similar to that of Corollary 4.5. For a fixed $\Omega' \Subset \Omega$ and a positive integer K to be chosen later let V and Ω_j as in the proof of Corollary 4.5. Let $\varphi \in C_0^\infty(\Omega_K)$, using the definition of T_V and Lemma 4.2 we obtain

$$L^R T_V f = L^R(\widetilde{\gamma} * f) = L^R \widetilde{\gamma} * f = (-\delta + \omega) * f = -f + \omega * f$$

Since $f \in \mathcal{D}'(\Omega)$ we can write

$$f = \sum_{\substack{j=1 \\ |\alpha_j| \leqslant A}}^{M} D^{\alpha_j} g_j \text{ in } \mathcal{D}'(\Omega_K).$$

Note that the kernel ω satisfies the same properties of $\widetilde{\gamma}$ in terms of support and growth estimate. Then, arguing as in the proof of Proposition 4.4 we see that

$$\omega * f = \sum_{|\beta| \leqslant A-1} D^\beta A_\beta$$

so that

$$f = \sum_{|\beta| \leqslant A-1} D^\beta A_\beta - L^R T_V f \text{ in } \mathcal{D}'(\Omega_K),$$

with $L^R T_V f \in C^\infty(\Omega_K)$ since $T_V f \in C^\infty(\Omega_K)$ by assumption. We can then start again with the identity

$$L^R T_V f = -f + \omega * f$$

where now we know in advance that $f = \sum_{|\beta| \leqslant A-1} D^\beta A_\beta$ in $\mathcal{D}'(\Omega_K)$ (the smooth function $L^R T_V f$ can be absorbed in this expression) and, applying iteratively the above argument, in k_1 steps we eventually conclude $f \in L^1_{loc}(\Omega_{K-k_1})$. Hence

$$L^R T_V f = -f + \omega * f \text{ in } \mathcal{D}'(\Omega_{K-k_1})$$

that is, modulo the smooth function $L^R T_V f$, f coincides with $\omega * f$ in $\mathcal{D}'(\Omega_{K-k_1})$.

Let us reason again like in the proof of Corollary 4.5: since by Proposition 4.2, $\omega \in L^{\frac{N-1}{N-2}}(\mathbb{G})$ we see that $f \in L^{\frac{N-1}{N-2}}(\Omega_{K-k_1-1})$; then with k_2 iterations of this argument we conclude that $f \in L^2(\Omega_{K-k_1-1-k_2})$ and with one more iteration $f \in W_X^{1,2}(\Omega_{K-k_1-1-k_2-1})$. Picking finally $K = k_1 + k_2 + 2$ we get the desidered assertion. ∎

Proof of Theorem 4.1. The proof still contains interative arguments. Fix $\Omega' \Subset \Omega'' \Subset \Omega$. By Corollary 4.5, there exist a positive integer K and a neighborhood V of the origin such that $T_V^K f \in W_X^{1,2}(\Omega'')$. Let now $U \subseteq V$ a neighborhood of the origin such that

$$\underbrace{U \circ U \circ \dots \circ U}_{K \text{ times}} \circ \Omega' \Subset \Omega''.$$

36

Let

$$\Omega_j = \underbrace{U \circ U \circ ... \circ U}_{j \text{ times}} \circ \Omega' \text{ for } j = 1, 2, ..., K$$

$$\Omega_0 = \Omega';$$

so that $\Omega_K \Subset \Omega''$. Clearly, it is still true, *a fortiori*, that $T_U^K f \in W_X^{1,2}(\Omega'')$ (having replaced the operator T_V with T_U, based on a smaller neighborhood).

By Proposition 4.6 and Remark 4.7 we have

$$\mathcal{L}\left(T_U^K f\right) = T_U^K \mathcal{L} f \text{ in } \Omega_K. \tag{4.11}$$

Since $\mathcal{L}f$ is smooth in Ω, by point (5) in Proposition 4.4 also $T_U^K \mathcal{L} f$ is smooth in Ω_K. By (4.11) then $\mathcal{L}\left(T_U^K f\right)$ is smooth in Ω_K and, since $T_U^K f \in W_X^{1,2}(\Omega_K)$, we can apply Theorem 3.4 to conclude that $T_U^K f \in C^\infty(\Omega_K)$. Applying Lemma 4.8 we see that $T_U^{K-1} f \in W_X^{1,2}(\Omega_{K-1})$. Iterating this argument K times shows that $f \in W_X^{1,2}(\Omega_0) = W_X^{1,2}(\Omega')$. Since $\mathcal{L}f \in C^\infty(\Omega)$ we can apply again Theorem 3.4 to conclude $f \in C^\infty(\Omega')$. For the genericity of Ω', we have $f \in C^\infty(\Omega)$. ■

References

[1] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: Stratified Lie groups and potential theory for their sub-Laplacians. Springer Monographs in Mathematics. Springer, Berlin, 2007.

[2] M. Bramanti, L. Brandolini: Hörmander's second order differential operators, book, in preparation.

[3] R. A. DeVore, G. G. Lorentz: Constructive approximation. Grundlehren der Mathematischen Wissenschaften, 303. Springer-Verlag, Berlin, 1993.

[4] G. B. Folland: Subelliptic estimates and function spaces on nilpotent Lie groups. Ark. Mat. 13 (1975), no. 2, 161–207.

[5] L. Hörmander: Hypoelliptic second order differential equations. Acta Math. 119 (1967), 147-171.

[6] J. J. Kohn: Pseudo-differential operators and hypoellipticity. Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), pp. 61–69. Amer. Math. Soc., Providence, R.I., 1973.

[7] P. Malliavin: Stochastic calculus of variations and hypoelliptic operators. In: Proc. Inter. Symp. on Stoch. Diff. Equations, Kyoto 1976, Wiley 1978, 195–263.

[8] L. P. Rothschild, E. M. Stein: Hypoelliptic differential operators and nilpotent groups. Acta Math. 137 (1976), no. 3-4, 247–320.

[9] W. Rudin, Functional analysis. Second edition. McGraw-Hill, Inc., New York, 1991.

[10] E. M. Stein: Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals. With the assistance of Timothy S. Murphy. Princeton Mathematical Series, 43. Monographs in Harmonic Analysis, III. Princeton University Press, Princeton, NJ, 1993.

[11] F. Treves: Topological vector spaces, distributions and kernels. 1967, Academic Press

[12] N. Th.Varopoulos, L. Saloff-Coste, T. Coulhon: Analysis and geometry on groups. Cambridge Tracts in Mathematics, 100. Cambridge University Press, Cambridge, 1992.