Fulvio Ricci

Harmonic Analysis on Nilpotent Groups

Università di Padova

A.A. 2017-18

Contents

CHAPTER 1

Vector fields and nilpotent Lie algebras

1. Vector fields, flows, exponentials

Let Ω be an open subset of $\mathbb{R}^n$. A *vector field* X on Ω is a first-order differential operator on Ω without 0-th order term,

$$X = \sum_{j=1}^n a_j(x)\partial_{x_j} \ . \tag{1.1.1}$$

For $x \in \Omega$, we set

$$X_x = \big(a_1(x), \dots, a_n(x)\big) \ ,$$

the vector determining the directional derivative that the vector field computes at x.

We shall assume that the vector field is real and smooth, i.e. that the coefficients a_j are real-valued C^∞-functions on Ω.

Given a point $x_0 \in \Omega$, the Cauchy problem

$$\begin{cases} \gamma'(t) = X_{\gamma(t)} \\ \gamma(0) = x_0 \ , \end{cases} \tag{1.1.2}$$

admits a local solution $\gamma : (-\varepsilon, \varepsilon) \to \Omega$ for some $\varepsilon > 0$. The following result is well known.

Theorem 1.1.1. *Let X be a smooth vector field on Ω. Then*

(i) *for every $x_0 \in \Omega$ the problem* (1.1.2) *has a unique solution $\gamma_{x_0}(t)$ defined on a maximal open interval I_{x_0} containing* 0*;*
(ii) *given $K \subset \Omega$ compact, there is $\varepsilon_K > 0$ such that γ_x is defined for $|t| < \varepsilon_K$ for every $x \in K$;*
(iii) *the map $\Phi_X(x,t) = \gamma_x(t)$ is smooth on its domain $D = \big\{(x,t) : t \in I_x\big\}$;*
(iv) *more generally, if*

$$X_y = \sum_{j=1}^n a_j(x,y)\partial_{x_j}$$

is a family of vector fields with coefficients depending smoothly on x and on a parameter $y \in \mathbb{R}^m$, and $\gamma_{y,x_0}(t)$ is the solution of the Cauchy problem (1.1.2) *relative to X_y, then the map $(x,y,t) \longmapsto \gamma_{y,x}(t)$ is smooth in all variables.*

Notice that (ii) is a consequence of (i) and (iii); it is stated for better clarity.

For fixed $t \in \mathbb{R}$, let $\Omega_t \subseteq \Omega$ consist of the elements x such that $t \in I_x$, and let $\varphi_{X,t} : \Omega_t \to \Omega$ be given by

$$\varphi_{X,t}(x) = \gamma_x(t) \ .$$

Then Ω_t is open and $\varphi_{X,t}$ is smooth. The following properties hold:

(i) $\varphi_{X,0} = \mathrm{Id}$;
(ii) $\varphi_{X,t} \circ \varphi_{X,t'} = \varphi_{X,t+t'}$, when defined;
(iii) in particular $\varphi_{X,t}$ is invertible on Ω_t and $\varphi_{X,t}^{-1} = \varphi_{X,-t}$;

The last identity is equivalent to saying that, if $x' = \gamma_x(t)$, then $\gamma_{x'}(t) = \gamma_x(t+t')$, which follows from uniqueness of solutions of Cauchy problems..

The maps $\varphi_{X,t}$ form the *flow* of the vector field X on Ω.

When analyzing the flow of a vector field locally, that is on a compact subset K of Ω, it is sometimes convenient to replace the vector field X by $X' = \eta X$, with $\eta \in \mathcal{D}(\Omega)$ and identically equal to 1 on a neighborhood of K. Then the flow of X' coincides with that of X on K, but has the advantage that the map $(x,t) \mapsto \varphi_{X',t}(x)$ is defined for every x and every t. When this happens, we say that a vector field is *complete*.

The uniqueness of the solution of Cauchy problems gives the following characterization of the flow.

Proposition 1.1.2. *Let Ω be an open set and $\delta > 0$. For each $t \in (-\delta, \delta)$, let $\varphi_t : \Omega_t \to \Omega$ be a smooth map, with $\Omega_t \subseteq \Omega$, and assume that*

(i) *$\Omega_0 = \Omega$ and $\varphi_0 = \mathrm{Id}$;*
(ii) *for each compact subset K of Ω there is $\delta(K) > 0$ such that*
- *$K \subset \Omega_t$ for $|t| < \delta(K)$,*
- *the composition $\varphi_t \circ \varphi_{t'}$ is defined on K for $|t| + |t'| < \delta(K)$,*
- *for $|t| + |t'| < \delta(K)$, $\varphi_t \circ \varphi_{t'} = \varphi_{t+t'}$ on K;*

(iii) *the map $\Phi(x,t) = \varphi_t(x)$ is smooth on its domain.*

For every $f \in C^\infty(\Omega)$, let

$$Xf = \frac{d}{dt}_{|_{t=0}} f \circ \varphi_t \ . \tag{1.1.3}$$

Then X is a smooth vector field on Ω, and φ_t coincides with the flow of X restricted to Ω_t.

Proof. Let $K \subset \Omega$ be compact and $x \in K$. For $|t| < \delta(K)$, let $\gamma_x(t) = \varphi_t(x)$ and define X_x as the tangent vector $\gamma_x'(0)$. The definition does not depend on the choice of K containing x. Formula (1.1.3) follows from the chain rule.

To prove that the curves γ_x are integral curves of X, fix $x_0 \in K$, $|t| < \delta(K)$ and let $x = \gamma_{x_0}(t) = \varphi_t(x_0)$. Then

$$\begin{aligned}
\gamma_{x_0}'(t) &= \lim_{h\to 0} \frac{1}{h}\big(\varphi_{t+h}(x_0) - \varphi_t(x_0)\big) \\
&= \lim_{h\to 0} \frac{1}{h}\big(\varphi_h(x) - x\big) \\
&= \lim_{h\to 0} \frac{1}{h}\big(\gamma_x(h) - x\big) \\
&= X_x = X_{\gamma_{x_0}(t)} \ .
\end{aligned}$$

Since $\gamma_{x_0}(0) = x_0$, γ_{x_0} is the (possibly non-maximal) solution of (1.1.2). □

We will use the following simple property, a direct consequence of the chain rule.

Lemma 1.1.3. *Let X be a smooth vector field on Ω with flow $\{\varphi_{X,t}\}$. Then the flow of the constant multiple sX of X is $\varphi_{sX,t}(x) = \varphi_{X,st}(x)$.*

Definition. *We call* exponential of tX *the operator* $\exp(tX)$ *acting on a function f compactly supported in Ω as*

$$\exp(tX)f(x) = f\big(\varphi_{X,t}(x)\big) \ . \tag{1.1.4}$$

We also denote by $\mathcal{D}(K)$ the space of smooth function on $\mathbb{R}^n$ supported on K.

Lemma 1.1.4. *Given $K \subset \Omega$ compact,* $\exp(tX)f$ *is defined for $|t| < \delta_K$ and $f \in \mathcal{D}(K)$. The exponential of X satisfies the following properties:*

(i) $\exp(0X)f = f$;
(ii) $\exp(-tX) = \exp(tX)^{-1}$;

(iii) $\exp\big((t+s)X\big) = \exp(tX)\exp(sX)$;
(iv) $\exp\big(t(sX)\big) = \exp\big((ts)X\big)$;
(v) *if* f *is* C^1, $\frac{d}{dt}\exp(tX)f = X\exp(tX)f = \exp(tX)Xf$;
(vi) *if* $f \in \mathcal{D}(K)$ *and* $k \in \mathbb{N}$,

$$\exp(tX)f(x) = \sum_{j=0}^{k} \frac{t^j}{j!} X^j f(x) + O(t^{k+1}) \ , \tag{1.1.5}$$

where the remainder term, as an element of $\mathcal{D}(\Omega)$, *depends continuously on* f *and* X.

Proof. By the remark following Theorem 1.1.1, we can assume that X is complete. Statements (i) and (iii) follow directly from the corresponding identities for the flow of X. Then (ii) is a direct consequence. (iv) follows from Lemma 1.1.3.

In order to prove (v), observe that for every function f and x in the domain of $\varphi_{X,t}$,

$$Xf(x) = \frac{d}{ds}_{|s=0} f\big(\varphi_{X,s}(x)\big) \ .$$

Hence

$$\begin{aligned} \frac{d}{dt}\exp(tX)f(x) &= \frac{d}{ds}_{|s=0} \exp\big((s+t)X\big)f(x) \\ &= \frac{d}{ds}_{|s=0} \exp(sX)\exp(tX)f(x) \\ &= X\big(\exp(tX)f\big)(x) \ , \end{aligned}$$

but also

$$\begin{aligned} \frac{d}{dt}\exp(tX)(x) &= \frac{d}{ds}_{|s=0} \exp(tX)\exp(sX)f(x) \\ &= \exp(tX)Xf(x) \ . \end{aligned}$$

Finally (vi) follows easily from (v). □

We consider now two smooth vector fields

$$X = \sum_{j=1}^{n} a_j(x)\partial_{x_j} \ , \qquad Y = \sum_{j=1}^{n} b_j(x)\partial_{x_j}$$

on Ω and we want to understand the interactions between their exponentials.

Some key remarks come from the analysis of the composition

$$\exp(sX)\exp(tY)\exp(-sX) \ ,$$

i.e. the conjugation of $\exp(tY)$ by $\exp(sX)$.

Observe that

$$\exp(sX)\exp(tY)\exp(-sX)f(x) = f\big(\varphi_{X,-s}\circ\varphi_{Y,t}\circ\varphi_{X,s}(x)\big) \ ,$$

and that for fixed s, the maps $\varphi_t(x) = \varphi_{X,-s}\varphi_{Y,t}\varphi_{X,s}(x)$ satisfy the hypotheses of Proposition 1.1.2 with Ω replaced by some subdomain $\Omega_s \subseteq \Omega$. Therefore there is a smooth vector field Y_s on Ω_s such that $\varphi_t = \varphi_{Y_s,t}$, i.e.

$$\exp(sX)\exp(tY)\exp(-sX) = \exp(tY_s) \ . \tag{1.1.6}$$

We call Y_s the *adjoint of* Y *by* $\exp(sX)$ and write

$$Y_s = \mathrm{Ad}\big(\exp(sX)\big)Y \ .$$

The expression of Y_s is obtained by differentiating in t at $t=0$:

$$Y_s = \exp(sX)Y\exp(-sX) \ .$$

Hence, for f smooth,

$$\begin{aligned}\big[\mathrm{Ad}\,\big(\exp(sX)\big)Y\big]f(x) &= \exp(sX)Y\exp(-sX)f(x)\\ &= \big(Y(f\circ\varphi_{X,-s})\big)\big(\varphi_{X,s}(x)\big)\ .\end{aligned} \tag{1.1.7}$$

Since the map $(s,t,x)\longmapsto\varphi_{X,-s}\varphi_{Y,t}\varphi_{X,s}(x)$ is smooth, it follows that the map

$$(s,x)\longmapsto \mathrm{Ad}\,\big(\exp(sX)\big)Yf(x)$$

is also smooth. We can therefore define a new vector field,

$$\mathrm{ad}\,(X)Y=\frac{d}{ds}_{|s=0}\mathrm{Ad}\,\big(\exp(sX)\big)Y\ .$$

Because the domains of $\exp(tY)$ and $\exp(sX)$ invade Ω as $s,t\to 0$, $\mathrm{ad}\,(X)Y$ is defined on all of Ω. By (1.1.7) we obtain the following description.

Proposition 1.1.5. *We have*

$$\mathrm{ad}\,(X)Y=XY-YX\stackrel{\mathrm{def}}{=}[X,Y]\ .$$

Observe that, even though XY and YX are separately second order differential operators, their difference is only first order, more precisely a vector field:

$$[X,Y]=\sum_{j=1}^n\left(\sum_{k=1}^n\big(a_k\partial_{x_k}b_j-b_k\partial_{x_k}a_j\big)\right)\partial_{x_j}\ .$$

Proof. Let

$$h(u,v,x)=\exp(uX)Y\exp(vX)f(x)\ .$$

Since the right-hand side is smooth in u and v,

$$\mathrm{ad}\,(X)Yf(x)=\frac{d}{ds}_{|s=0}h(s,-s,0)=(\partial_u-\partial_v)h(0,0,x)\ .$$

We have

$$\begin{aligned}\partial_u h(0,0,x)&=\frac{d}{du}_{|u=0}\exp(uX)Yf(x)\\ &=X(Yf)(x)\ .\end{aligned}$$

Moreover,

$$\begin{aligned}\partial_v h(0,0,x)&=\frac{d}{dv}_{|v=0}Y\exp(vX)f(x)\\ &=Y\frac{d}{dv}_{|v=0}\exp(vX)f(x)\\ &=Y(Xf)(x)\ ,\end{aligned}$$

since for each v the coefficients of Y are evaluated at x. Hence,

$$\mathrm{ad}\,(X)Yf(x)=(XY-YX)f(x)\ . \qquad \square$$

Proposition 1.1.6. *The following are equivalent:*

(i) $\exp(sX)\exp(tY)=\exp(tY)\exp(sX)$ *for* $|s|,|t|<\delta$ *for some* $\delta>0$ *(and hence every* δ *for which the* $\exp$ *are defined);*
(ii) $Y\exp(sX)=\exp(sX)Y$ *for every* s *(as above);*
(iii) $XY=YX$*;*
(iv) $\exp(sX)\exp(tY)=\exp(sX+tY)$ *for every* s *and* t *(as above).*

PROOF. If (i) holds,
$$\exp(tY_s) = \exp(tY)$$
for every t. Differentiating at $t = 0$, we obtain that $Y_s = Y$, which gives (ii).

From (ii), we obtain (iii) differentiating in s at $s = 0$.

Going backwards, if (iii) holds, differentiating in s at any s,
$$\begin{aligned}\frac{d}{ds}\mathrm{Ad}\big(\exp(sX)\big)Y &= \frac{d}{ds}\exp(sX)Y\exp(-sX)\\ &= \exp(sX)XY\exp(-sX) - \exp(sX)YX\exp(-sX)\\ &= \exp(sX)[X,Y]\exp(-sX)\\ &= 0\ .\end{aligned}$$

Therefore $\mathrm{Ad}\big(\exp(sX)\big)Y$ does not depend on s, hence it is constantly equal to Y, its values at $s = 0$. This gives (ii).

Now, if (ii) holds, then $\mathrm{Ad}\big(\exp(sX)\big)Y = Y$, and (1.1.6) gives (i).

Suppose now that (iii) holds. Then $[X, \alpha Y] = [X, X + \alpha Y] = [\alpha Y, X + \alpha Y] = 0$ for every $\alpha \in \mathbb{R}$. By (i), the product $T(s) = \exp(-sX)\exp\big(s(X + \alpha Y)\big)\exp(-s\alpha Y)$ does not depend of the order of the three factors. Therefore,
$$T'(s) = (X + \alpha Y - X - \alpha Y)T(s) = 0\ .$$

Hence $T(s) = \mathrm{Id}$, i.e.
$$\exp\big(s(X + \alpha Y)\big) = \exp(sX)\exp(s\alpha Y)\ ,$$
which gives (iv).

Conversely, if (iv) holds, $T(s)$ is constant, so that
$$\begin{aligned}T'(s) &= -\exp(-sX)X\exp\big(s(X + \alpha Y)\big)\exp(-s\alpha Y)\\ &\qquad + \exp(-sX)\exp\big(s(X + \alpha Y)\big)(X + \alpha Y)\exp(-s\alpha Y)\\ &\qquad - \alpha\exp(-sX)\exp\big(s(X + \alpha Y)\big)Y\exp(-s\alpha Y)\\ &= -\exp(-sX)X\exp\big(s(X + \alpha Y)\big)\exp(-s\alpha Y)\\ &\qquad + \exp(-sX)\exp\big(s(X + \alpha Y)\big)X\exp(-s\alpha Y)\\ &= 0\ .\end{aligned}$$

Therefore, $X\exp\big(s(X + \alpha Y)\big) = \exp\big(s(X + \alpha Y)\big)X$, which implies that $[X, X + \alpha Y] = \alpha[X, Y] = 0$ for every α. This gives (iii). □

All the results of this Section remain valid if Ω is replaced by a smooth (i.e. C^∞) manifold M and a vector field X on M is defined as an operator which acts on smooth functions on M and which can be expressed as in (1.1.1) in any set of local coordinates.

2. Lie algebras and the Baker-Campbell-Hausdorff formula

The space of smooth real vector fields on a manifold has a natural linear structure; in addition, the operation

(1.2.1)
$$[X, Y] = XY - YX\ ,$$

called the *commutator* of X and Y, is defined on this space.

We concentrate our attention on some algebraic notions connected with this operation.

DEFINITION. *A* Lie algebra *over* $\mathbb{R}$ *(resp.* $\mathbb{C}$*)*[1] *is a real (resp. complex) vector space* $\mathfrak{g}$*, endowed with a bilinear map (called* Lie bracket*)*
$$[\ ,\ ] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

[1]We shall mainly be interested in real Lie algebras.

such that

(i) $[x,y] = -[y,x]$ *for every* $x, y \in \mathfrak{g}$*;*

(ii) *the* Jacobi identity

$$\big[[x,y],z\big] + \big[[y,z],x\big] + \big[[z,x],y\big] = 0 \tag{1.2.2}$$

holds for every $x, y, z \in \mathfrak{g}$.

Since the commutator of vector fields satisfies the Jacobi identity, smooth vector fields on a given manifold M form a real Lie algebra (similarly, complex vector fields on M form a complex Lie algebra). In general, this is an infinite-dimensional Lie algebra. Here are a few examples of finite-dimensional Lie algebras.

Examples 1.

(a) On $M = \mathbb{R}^2$ consider the vector fields

$$X = \partial_x\ , \qquad Y = x\partial_y\ , \qquad Z = \partial_y\ .$$

Since $[X,Y] = Z$ and $[X,Z] = [Y,Z] = 0$, the 3-dimensional vector space that they generate is closed under commutator, and therefore a Lie algebra.

(b) On $M = \mathbb{R}$ let

$$X = d/dx\ , \qquad Y = x\, d/dx\ .$$

Then $[X,Y] = X$, so that X and Y span a 2-dimensional Lie algebra.

(c) On $\mathbb{R}^3$, let

$$X_1 = x_3\partial_{x_2} - x_2\partial_{x_3}\ , \qquad X_2 = x_1\partial_{x_3} - x_3\partial_{x_1}\ , \qquad X_3 = x_2\partial_{x_1} - x_1\partial_{x_2}\ .$$

Their linear span is a Lie algebra.

(d) The notion of Lie algebra is abstract, and a Lie algebra need not consist necessarily of vector fields. For instance, every associative algebra becomes a Lie algebra under the commutator $[x,y] = xy - yx$. In particular, this is true for the algebra of $n \times n$ real (or complex) matrices.

Definition. *A Lie algebra* $\mathfrak{g}$ *is called* abelian *if* $[x,y] = 0$ *for every* $x, y \in \mathfrak{g}$*. It is called* nilpotent *if there is a* k *such that every iterated Lie bracket*

$$\big[\cdots\big[[x_0,x_1],x_2\big]\cdots, x_k\big]\ . \tag{1.2.3}$$

is zero. If k *is the smallest integer for which this happens, one says that* $\mathfrak{g}$ *is* nilpotent of step k.

A subalgebra *of a Lie algebra* $\mathfrak{g}$ *is a linear subspace which is closed under the Lie bracket. If* $S = \{x_i\}_{i\in I}$ *is a subset of* $\mathfrak{g}$*, the* subalgebra generated by S *is the smallest subalgebra of* $\mathfrak{g}$ *containing* S.

Clearly, the Lie subalgebra generated by S contains all iterated Lie brackets (1.2.3) with the $x_j \in S$. It also contains Lie brackets of Lie brackets, like

$$\Big[\big[[x_1,x_2],x_3\big],[x_4,x_5]\Big]\ ,$$

or more complicated expressions of this kind. It is possible, however, to reduce such expressions to linear combinations of iterated commutators.

Lemma 1.2.1. *The subalgebra generated by* S *is the linear span of the iterated Lie brackets* (1.2.3) *of elements of* S.

Proof. By induction on the number of nested brackets, it is sufficient to verify that, by the Jacobi identity,

$$\big[[a,b],[c,d]\big] = \Big[\big[[a,b],c\big],d\Big] - \Big[\big[[a,b],d\big],c\Big]\ . \qquad \square$$

We go back to vector fields and their exponentials.

Let X and Y be two smooth vector fields on a manifold M. We have seen in Section 1 that the identity $\exp(sX)\exp(tY) = \exp(sX+tY)$ does not hold in general, unless $[X,Y] = 0$. It turns out that one can obtain good approximations of the product $\exp(sX)\exp(tY)$ for small values of s and t by single exponentials. We shall not give proofs, but we try at least to motivate some formal identities. For a closer insight, see the appendix.

Consider the formal power series

$$\exp(tX) = \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k \ ,$$

derived from the Taylor expansion (1.1.5), and observe that the expansions of the two expressions (with non-commuting variables)

$$\exp(sX)\exp(tY) \ , \qquad \exp\Big(sX + tY + \frac{st}{2}[X,Y]\Big)$$

coincide up to terms of degree 2.

A more concrete statement (whose proof we skip) is that

$$\partial_s^j \partial_t^k \big(\exp(sX)\exp(tY)f(x)\big)_{|s=t=0} = \partial_s^j \partial_t^k \Big(\exp\Big(sX + tY + \frac{st}{2}[X,Y]\Big)f(x)\Big)_{|s=t=0}$$

for $j + k \leq 2$.

A longer computation shows that $\exp(sX)\exp(tY)$ and

$$\exp\bigg(sX + tY + \frac{st}{2}[X,Y] + \frac{st^2}{12}\big[[X,Y],Y\big] - \frac{s^2t}{12}\big[[X,Y],X\big]\bigg)$$

have the same formal expansion up to terms of degree 3.

This formal procedure can be pushed to any degree. In this way one constructs a formal power series

$$S(s,t;X,Y) = sX + tY + \sum_{j,k\geq 1} s^j t^k Z_{j,k}(X,Y) \ , \tag{1.2.4}$$

where each $Z_{j,k}(X,Y)$ is a fixed linear combination of iterated commutators of X and Y, containing X j times and Y k times. This power series satisfies the formal identity

$$\exp(sX)\exp(tY) = \exp\big(S(s,t;X,Y)\big) \ . \tag{1.2.5}$$

This is the (formal) *Baker-Campbell-Hausdorff* formula.

THEOREM 1.2.2. *Let X and Y be smooth vector fields on a manifold M. For any integer N, let $S_N(s,t;X,Y)$ be the sum* (1.2.4) *truncated to $j + k \leq N$. Then, given any compact subset K of M, there is $\varepsilon = \varepsilon(K,N) > 0$ such that* $\exp\big(S_N(s,t;X,Y)\big)f(x)$ *is well defined for $x \in K$ and $|s|,|t| < \varepsilon$. Moreover*

$$\partial_s^j \partial_t^k \big(\exp(sX)\exp(tY)f(x)\big)_{|s=t=0} = \partial_s^j \partial_t^k \Big(\exp\big(S_N(s,t;X,Y)\big)f(x)\Big)_{|s=t=0}$$

for $j + k \leq N$, so that

$$\exp(sX)\exp(tY)f(x) = \exp\big(S_N(s,t;X,Y)\big)f(x) + O\big(|s||t|(|s|+|t|)^{N-1}\big) \ .$$

This statement is useful for many purposes, but it is very weak in terms of exact identities. For instance, by itself, it does not even imply that $\exp(sX)\exp(tY) = \exp(sX + tY)$ when X and Y commute.

We state, without proof, an important sufficient condition. It involves the Lie algebra generated by the two vector fields X and Y. This is of course the linear span of them and of their iterated commutators.

THEOREM 1.2.3. *Assume that X and Y generate a finite dimensional Lie algebra. For any relatively compact open subset U of M there is $\varepsilon = \varepsilon(U) > 0$ such that, for $x \in U$ and $|s|,|t| < \varepsilon$, the series of vector fields*

$$sX + tY + \sum_{j,k\geq 1} s^j t^k Z_{j,k}(X,Y) = S(s,t;X,Y)$$

is convergent on U and

$$\exp(sX)\exp(tY)f(x) = \exp\big(S(s,t;X,Y)\big)f(x) \ .$$

We mention for future reference that the associative law $(e^X e^Y)e^Z = e^X(e^Y e^Z)$ is respected by the formal power series (1.2.4). This means that, taking for simplicity $s = t = 1$ and setting $S(X,Y) = S(1,1;X,Y)$,

$$S\big(S(X,Y),Z\big) = S\big(X,S(Y,Z)\big) \ . \tag{1.2.6}$$

The situation described above becomes even simpler if the Lie algebra generated by X and Y is not only finite dimensional, but also nilpotent. In this case, in fact, the series (1.2.4) contains only finitely many terms.

Finally, other formal identities can be given a precise content, along the same lines of Theorems 1.2.2 and 1.2.3. Examples of these are

$$\exp(sX)\exp(tY)\exp(-sX) = \exp\left(tY + st[X,Y] - \frac{s^2t}{2}\big[[X,Y],X\big] + \cdots\right) \ , \tag{1.2.7}$$

$$\exp(sX)\exp(tY)\exp(-sX)\exp(-tY) = \exp\left(st[X,Y] - \frac{s^2t}{2}\big[[X,Y],X\big] - \frac{st^2}{2}\big[[X,Y],Y\big] + \cdots\right) \ . \tag{1.2.8}$$

3. Lie groups and Lie algebras

Finite-dimensional Lie algebras of vector fields on a manifold M naturally appear when the flows are required to satisfy special properties, such as being isometries (if M has a Riemannian structure), or being holomorphic (if M has a complex structure).

Another important case occurs when M has a group structure and the flows consist of group translations. In fact there is a strict relation between the theory of finite-dimensional Lie algebra and the theory of Lie groups.

Definition. *A* Lie group *is a smooth manifold G, which is also a group and the map from $G \times G$ to G that assign to the pair (x,y) the element xy^{-1} is smooth.*

Proposition 1.3.1. *The following transformations on a Lie group G are smooth:*

(i) *the* inversion $x \mapsto x^{-1}$*;*
(ii) *the* left translations $\ell_a(x) = ax$ *and the* right translations $r_a(x) = xa^{-1}$*, with $a \in G$.*

Consequently, the following operators transform smooth functions into smooth functions:

(i) *the* "check" *operation $f \mapsto \check{f}$, where $\check{f}(x) = f(x^{-1})$;*
(ii) *the* left and right translation operators $L_a f(x) = f(a^{-1}x)$*, $R_a f(x) = f(xa)$.*

Observe that the definitions of ℓ_a, r_a, L_a, R_a are given in such a way to respect in each case the multiplicative rule $\ell_{ab} = \ell_a \ell_b$, $r_{ab} = r_a r_b$, etc.

At this point we must recall the notion of *tangent vector* at a given point x on a manifold M. The idea is that once a coordinate system near x has been chosen, each tangent vector corresponds to a directional derivative at x.

In order to give an intrinsic definition (i.e. independent of the coordinates), one must define a tangent vector at x as a real linear functional v defined on smooth real functions in a neighborhood of x, such that

(i) $v(f) = v(g)$ if f and g coincide on a neighborhood of x;
(ii) $v(fg) = v(f)g(x) + f(x)v(g)$ for every f and g.

Clearly tangent vectors at x form a vector space, denoted by T_xM. We assume that $n = \dim M$.

Proposition 1.3.2. *Let $\varphi : \Omega \to M$, $\Omega \subseteq \mathbb{R}^n$, define a coordinate system near x, with $x = \varphi(0)$. For every $v \in T_xM$ there is a unique real n-tuple $(a_1, \dots, a_n)$ such that*

$$v(f) = \sum_{j=1}^n a_j \partial_j (f \circ \varphi)(0) \ .$$

In particular, the dimension of T_xM is n.

(We omit the proof.)

If X is a smooth vector field on M and f is a smooth function, the value of Xf at a given point x defines a tangent vector at x, that we call X_x.

We return now to Lie groups.

DEFINITION. *A vector field X on G is called* left-invariant *(resp.* right-invariant*) if*

$$L_a(Xf) = X(L_a f)\ ,\ \ (resp.\ R_a(Xf) = X(R_a f))$$

for every $a \in G$.

PROPOSITION 1.3.3. *Given any tangent vector v at the identity element e of G, there are a unique left-invariant vector field X such that $X_e = v$, and a unique right-invariant vector field X' such that $X'_e = v$. They are given by*

$$Xf(x) = v(L_{x^{-1}}f)\ , \qquad X'f(x) = v(R_x f)\ . \tag{1.3.1}$$

PROOF. Since $Xf(x) = L_{x^{-1}}(Xf)(e)$, in order that X be left-invariant, it must be $Xf(x) = X(L_{x^{-1}}f)(e)$, so that the first identity in (1.3.1) is forced. We must prove now that (1.3.1) actually defines a left-invariant vector field. Given $a \in G$, we have

$$L_a(Xf)(x) = Xf(a^{-1}x) = v(L_{x^{-1}a}f) = v(L_{x^{-1}}L_a f) = X(L_a f)(x)\ .$$

The proof for X' is similar. □

The commutator of two left-invariant (resp.right-invariant) vector fields is also left-invariant (resp.right-invariant).

DEFINITION. *The Lie algebra $\mathfrak{g}$ of left-invariant vector fields on G is called the* Lie algebra of G*, also denoted by* Lie(G)*.*

By Proposition 1.3.3, $\dim(\mathrm{Lie}\,(G)) = \dim(G)$.

Because of the identification between left-invariant vector fields and tangent vectors at e, it is customary to take T_eG as the underlying vector space for Lie(G). The computation of the Lie bracket $[v,w]$ of two tangent vectors at e then involves their extension to left-invariant vector fields and the computation of the commutator of these. We must remark that if one extends v and w to right-invariant vector fields instead, the result is different (even though the resulting Lie algebra on T_eG is isomorphic to $\mathfrak{g}$).

Another consequence of Proposition 1.3.3 is that the left translations by x establishes a natural 1-1 correspondence between tangent vectors at e and tangent vectors at x, given by

$$v \in T_eG \leftrightarrow v' \text{ where } v'(f) = v(L_{x^{-1}}f) \in T_xG\ . \tag{1.3.2}$$

We discuss now the flow on G generated by a left-invariant vector field X.

DEFINITION. *A* one-parameter group *in G is a smooth map $\gamma : \mathbb{R} \to G$ such that $\gamma(s+t) = \gamma(s)\gamma(t)$ for every $s,t \in \mathbb{R}$.*

It can be proved that the weaker assumption that γ be continuous implies that γ is smooth.

THEOREM 1.3.4. *Let $\{\varphi_t\}$ be the flow on G generated by a left-invariant vector field X. Then φ_t is defined on all of G for every $t \in \mathbb{R}$. Moreover $\gamma(t) = \varphi_t(e)$ is a one-parameter group and*

$$\varphi_t(x) = x\gamma(t) \tag{1.3.3}$$

for every $x \in G$ and $t \in \mathbb{R}$.

Conversely, given any one-parameter group $\gamma(t)$ in G, there is a left-invariant vector field X whose flow is given by (1.3.3).

PROOF. By Theorem 1.1.1, $\gamma(t)$ is defined at least on some interval $(-\delta,\delta)$. For $x\in G$, consider

$$\begin{aligned}\frac{d}{dt}f\big(x\gamma(t)\big)&=\frac{d}{dt}(L_{x^{-1}}f)\big(\gamma(t)\big)\\&=X(L_{x^{-1}}f)\big(\gamma(t)\big)\\&=L_{x^{-1}}(Xf)\big(\gamma(t)\big)\\&=Xf\big(x\gamma(t)\big)\ .\end{aligned}$$

Therefore

$$\frac{d}{dt}\Big(\exp(-tX)f\big(x\gamma(t)\big)\Big)=0\ ;$$

this implies that the expression in parenthesis is constantly equal to $f(x)$, so that

$$\exp(tX)f(x)=f\big(x\gamma(t)\big)\ .$$

This gives (1.3.3) for $|t|<\delta$. In particular, for $|t|<\delta$, φ_t is defined on all of G. The identity $\varphi_{s+t}=\varphi_s\circ\varphi_t$ allows a unique extension of the definition of φ_t to $t\in\mathbb{R}$ such that this property is preserved. It follows from Proposition 1.1.2 that this extension is the flow generated by X. Also,

$$\gamma(s+t)=\varphi_{s+t}(e)=\varphi_s\big(\varphi_t(e)\big)=\varphi_s\big(\gamma(t)\big)=\gamma(t)\gamma(s)\ ,$$

i.e. γ is a one-parameter group.

To prove the converse, let γ be a one-parameter group. Then the maps $\varphi_t(x)=x\gamma(t)$ satisfy the assumptions of Proposition 1.1.2. Therefore they give the flow generated by a vector field X. We must show that X is left-invariant. But

$$\begin{aligned}X(L_af)(x)&=\frac{d}{dt}_{|t=0}L_af\big(x\gamma(t)\big)=\frac{d}{dt}_{|t=0}f\big(a^{-1}x\gamma(t)\big)\\&=Xf(a^{-1}x)=L_a(Xf)(x)\ ,\end{aligned}$$

for every $a\in G$. □

COROLLARY 1.3.5. *There is a one-to-one correspondence between left-invariant vector fields on G and one-parameter groups in G. It assigns to every one-parameter group $\gamma(t)$ the vector field X generating the flow $\varphi_t(x)=x\gamma(t)$.*

4. The group exponential

Let G be a Lie group, v a tangent vector at e, X the left-invariant vector field such that $X_e=v$, and $\gamma_v(t)$ the corresponding one-parameter group, according to Corollary 1.3.5. At this stage it is convenient to use T_eG as the underlying vector space of $\mathfrak{g}=\mathrm{Lie}\,(G)$.

DEFINITION. *The* exponential map $\exp_G:\mathfrak{g}\to G$ *is given by*

$$\exp_G(v)=\gamma_v(1)\ .$$

The connection with the exponential of a left-invariant vector field is given by the identities

$$\exp(tX)f(x)=f\big(x\exp_G(tv)\big)=R_{\exp_G(tv)}f(x)\ ,$$
$$Xf(x)=\frac{d}{dt}_{|t=0}f\big(x\exp_G(tv)\big)\ .$$

Observe that

$$\exp(sX)\exp(tY)f(x)=R_{\exp_G(sX_e)}R_{\exp_G(tY_e)}f(x)=R_{\exp_G(sX_e)\exp_G(tY_e)}f(x)\ ,$$

so that the two exponential notations are consistent. As a matter of fact, since $\exp_G$ is defined on the "abstract" Lie algebra, it is as well correct to write $\exp_G(X)$ instead of $\exp_G(X_e)$, once the Lie algebra is thought of as consisting of left-invariant vector fields. It is also customary to drop the subscript in $\exp_G$

(unless more than one group is present). This makes the notation $\exp(X)$ ambiguous, but it works fine in most cases.

PROPOSITION 1.4.1. *The group exponential is a smooth map, and it is a diffeomorphism from a neighborhood of* 0 *in* $\mathfrak{g}$ *onto a neighborhood of* e *in* G.

Before giving the proof, we recall the notion of *differential at* $t_0 \in \mathbb{R}^n$ of a smooth map φ from a neighborhood Ω of t_0 to a manifold M. It is the linear map $d\varphi(t_0) : \mathbb{R}^n \to T_{\varphi(t_0)}M$ that assigns to $u \in \mathbb{R}^n$ the tangent vector v defined by

$$v(f) = \partial_u (f \circ \varphi)(t_0)$$

for f smooth on a neighborhood of $\varphi(t_0)$.

Of course, $\mathbb{R}^n$ can be replaced by any finite-dimensional vector space ($\mathfrak{g}$ in the present instance).

PROOF. In the notation of Theorem 1.1.1 (iv), the map

$$(x, v, t) \longmapsto \gamma_{v,x}(t) = x \exp_G(tv)$$

is smooth, because the left-invariant vector fields X_v depend linearly on v. Restricting this map to $x = e$ and $t = 1$, we obtain the first part of the statement.

By the inverse mapping theorem, the second part of the statement follows if we show that $d\exp_G(0) : \mathfrak{g} \to T_e G \sim \mathfrak{g}$ is invertible. For $u \in \mathfrak{g}$ and f smooth on G, we have

$$\big(d\exp_G(0)u\big)(f) = \frac{d}{ds}_{|_{s=0}} f\big(\exp_G(su)\big) = u(f) \ .$$

Hence $d\exp_G(0) = \mathrm{Id}$. $\square$

It follows that the exponential mapping defines a system of local coordinates on G near e. These are called *canonical coordinates of the first kind*. There are other ways to use the exponential mapping in order to define coordinate systems near e.

The general *canonical coordinates of the second kind* are defined as follows. Decompose $\mathfrak{g}$ as the direct sums of linear subspaces,

$$\mathfrak{g} = \mathfrak{h}_1 \oplus \mathfrak{h}_2 \oplus \cdots \oplus \mathfrak{h}_k \ ,$$

and define $\varphi : \mathfrak{g} \to G$ as

$$\varphi(v_1 + v_2 + \cdots + v_k) = \exp_G(v_1)\exp_G(v_2)\cdots\exp_G(v_k) \ ,$$

if $v_j \in \mathfrak{h}_j$ for every j.

The previous proof shows that $d\varphi(0)u = u$ for $u \in \mathfrak{h}_j$, so that again $d\varphi(0) = \mathrm{Id}$.

A basic fact in Lie theory is that every (abstract) finite dimensional real Lie algebra is isomorphic to the Lie algebra of a Lie group. We shall not prove this fact in general, but only for nilpotent Lie algebras.

Let then $\mathfrak{g}$ be a finite-dimensional nilpotent Lie algebra over $\mathbb{R}$. For $x, y \in \mathfrak{g}$, let

$$\begin{aligned} S(x,y) &= x + y + \frac{1}{2}[x,y] + \frac{1}{12}\big[[x,y],y\big] - \frac{1}{12}\big[[x,y],x\big] + ... \\ &= \sum_{j+k\geq 1} Z_{j,k}(x,y) \end{aligned}$$

be the formal expression in (1.2.4) for $s = t = 1$. Since $\mathfrak{g}$ is nilpotent, the sum is finite, so that S is a polynomial mapping from $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$. This means that if we introduce linear coordinates on $\mathfrak{g}$, and write $x = (x_1, \dots, x_n)$ etc., then

$$S(x,y) = \big(S_1(x,y), \dots, S_n(x,y)\big) \ ,$$

where each S_j is a polynomial in x and y.

THEOREM 1.4.2. *Introduce on $\mathfrak{g}$ the composition law*

$$x \cdot y = S(x,y) \ . \tag{1.4.1}$$

This defines a Lie group structure having $\mathfrak{g}$ as its underlying manifold. If we denote by G this Lie group, then Lie(G) *is isomorphic to $\mathfrak{g}$.*

PROOF. The composition law is associative by (1.2.6). Clearly, $S(0,x) = S(x,0) = 0$, so that 0 is the identity element. Since $[x,-x] = 0$, $S(x,-x) = 0$, so that $x^{-1} = -x$. Similarly, we obtain that

$$(sv) \cdot (tv) = S(sv,tv) = (s+t)v \ ,$$

showing that $\gamma(tv) = tv$ is a one-parameter group. By Theorem 1.3.4, these are all the one-parameter groups in G. Moreover $\exp_G(tv) = tv$.

For $v \in \mathfrak{g}$, let X_v be the left-invariant vector field

$$\begin{aligned} X_v f(x) &= \frac{d}{dt}_{|t=0} f\big(x \cdot (tv)\big) \\ &= \frac{d}{dt}_{|t=0} f\Big(x + tv + \frac{t}{2}[x,v] + \frac{t^2}{12}\big[[x,v],v\big] - \frac{t}{12}\big[[x,v],x\big] + ...\Big) \ . \end{aligned}$$

If $\partial_u f$ denotes the directional derivative of f in the direction u, we have

$$\begin{aligned} X_v f(x) &= \partial_v f(x) + \frac{1}{2}\partial_{[x,v]} f(x) - \frac{1}{12}\partial_{\big[[x,v],x\big]}(x) + \cdots \\ &= \partial_v f(x) + \sum_{j=1}^k \partial_{Z_{j,1}(x,v)} f(x) \ . \end{aligned} \tag{1.4.2}$$

It must be understood that the first term involves a derivative in the same direction v independently of x, whereas each term in the sum contains a derivative in a direction depending on x. For better clarity, let us introduce linear coordinates on $\mathfrak{g}$, so that $\partial_u f(x) = \langle u, \nabla f(x)\rangle$ and

$$X_v f(x) = \langle v, \nabla f(x)\rangle + \frac{1}{2}\langle [x,v], \nabla f(x)\rangle + \sum_{j=2}^k \langle Z_{j,1}(x,v), \nabla f(x)\rangle \ . \tag{1.4.3}$$

Hence the first term is a constant coefficient operator, the second term has linear coefficients in x, and, similarly, the j-th term in the sum has coefficients which are homogeneous polynomial of degree j.

In order to prove that the Lie algebra of G is isomorphic to $\mathfrak{g}$, we show that the map $v \longmapsto X_v$ is an isomorphism. Clearly, this map is linear. It is also bijective, because X_v is the unique left-invariant vector field such that $(X_v)_0 = \partial_v$. It remains to prove that $[X_v, X_w] = X_{[v,w]}$ for every $v, w \in \mathfrak{g}$. Equivalently, we must prove that $[X_v, X_w] f(0) = \partial_{[v,w]} f(0)$.

Since $X_u g(0) = \partial_u g(0)$ for every g, and keeping in mind that all terms involving second-order derivatives must cancel with each other, we have

$$\begin{aligned} [X_v, X_w] f(0) &= \partial_v X_w f(0) - \partial_w X_v f(0) \\ &= \frac{1}{2}\langle \big(\partial_v [x,w]\big)_{|x=0}, \nabla f(0)\rangle + \sum_{j=2}^k \langle \big(\partial_v Z_{j,1}(x,w)\big)_{|x=0}, \nabla f(0)\rangle \\ &\quad - \frac{1}{2}\langle \big(\partial_w [x,v]\big)_{|x=0}, \nabla f(0)\rangle - \sum_{j=2}^k \langle \big(\partial_w Z_{j,1}(x,v)\big)_{|x=0}, \nabla f(0)\rangle \ . \end{aligned}$$

Since $[x,w]$ is linear in x, $\partial_v [x,w] = [v,w]$. On the other hand, $Z_{j,1}(x,w)$ is an n-tuple of homogeneous polynomials of degree j, so that $\partial_v Z_{j,1}(x,w)$ is an n-tuple of polynomials homogeneous of degree $j-1$, and they will vanish at the origin. Ultimately,

$$[X_v, X_w] f(0) = \frac{1}{2}\langle [v,w], \nabla f(0)\rangle - \frac{1}{2}\langle [w,v], \nabla f(0)\rangle = \partial_{[v,w]} f(0) \ . \qquad \square$$

It can be shown that the group G so constructed is the "universal covering" of every connected Lie group having a Lie algebra isomorphic to $\mathfrak{g}$, in the following sense.

THEOREM 1.4.3. *Let G' be any connected Lie group whose Lie algebra is isomorphic to the nilpotent Lie algebra $\mathfrak{g}$. Then G' is isomorphic to a quotient of the group G of Theorem 1.4.2, modulo a central discrete subgroup of G. If G' is also simply connected, then it is isomorphic to G.*

A Lie group is also called *nilpotent* if its Lie algebra is nilpotent.

EXAMPLES 2.

(a) If $\mathfrak{g}$ is abelian, then $S(v,w) = v + w$, and G coincides with $\mathfrak{g}$ itself with the abelian group structure underlying its linear structure. Hence $G \sim \mathbb{R}^n$, and every other n-dimensional abelian, connected Lie group is isomorphic to $\mathbb{R}^k \times \mathbb{T}^{n-k}$ for some k, $0 \le k \le n$.

(b) The smallest non-commutative nilpotent Lie algebra is the three-dimensional *Heisenberg algebra* $\mathfrak{h}_1$, spanned by a basis e_1, e_2, e_3 satisfying the following relations:

$$[e_1, e_2] = e_3 \ , \qquad [e_1, e_3] = [e_2, e_3] = 0 \ .$$

If $v = xe_1 + ye_2 + te_3$ and $w = x'e_1 + y'e_2 + t'e_3$, then

$$S(v,w) = v + w + \frac{1}{2}[v,w] = (x + x')e_1 + (y + y')e_2 + \Big(t + t' + \frac{1}{2}(xy' - x'y)\Big)e_3 \ .$$

The corresponding connected and simply connected Lie group H_1, called the three-dimensional *Heisenberg group* can then be regarded as $\mathbb{R}^3$ with product

$$(x,y,t) \cdot (x',y',t') = \Big(x + x', y + y', t + t' + \frac{1}{2}(xy' - x'y)\Big) \ . \tag{1.4.4}$$

The basis of left-invariant vector fields X, Y, T on H_1 such that $X_0 = \partial_x$, $Y_0 = \partial_y$, $T_0 = \partial_t$ is

$$X = \partial_x - \frac{y}{2}\partial_t \ , \qquad Y = \partial_y + \frac{x}{2}\partial_t \ , \qquad T = \partial_t \ . \tag{1.4.5}$$

5. Generating systems of vector fields

Let M be an n-dimensional smooth manifold, and let $X_1, \dots X_k$ be smooth vector fields on M (with no relation between k and n). Let also $\mathcal{L}(X_1, \dots, X_k)$ be the Lie algebra of vector fields on M that they generate. We also set

$$\mathcal{L}_x(X_1, \dots, X_k) = \big\{ Y_x : Y \in \mathcal{L}(X_1, \dots, X_k) \big\} \subseteq T_x M \ .$$

DEFINITION. *We say that the set $\{X_1, \dots, X_k\}$ is a* generating system of vector fields *on M (or that it satisfies* the Hörmander condition*) if, for every $x \in M$, $\mathcal{L}_x(X_1, \dots, X_k) = T_x M$.*

EXAMPLES 3.

(a) The vector fields X, Y in (1.4.5) form a generating system on $\mathbb{R}^3$.

(b) The vector fields $X_1 = \partial_x$ and $X_2 = f(x)\partial_y$ form a generating system on $\mathbb{R}^2$ if and only if for every $x \in \mathbb{R}$ there is a k such that $f^{(k)}(x) \neq 0$.

The interest in generating systems comes from two different (but not completely unrelated) properties, one being of geometric nature, the other more strictly analytic. Since both properties are local, we shall restrict ourselves to $M = \Omega$, an open connected subset of $\mathbb{R}^n$.

5.1. Sub-unitary curves and control distance.

If $I = (i_0, i_1, i_2, \dots, i_j) \in \{1, \dots, k\}^{j+1}$, we denote by $X^{[I]}$ the j-th order commutator

$$X^{[I]} = \big[\cdots\big[[X_{i_0}, X_{i_1}], X_{i_2}\big], \cdots, X_{i_j}\big] \ .$$

We also set $|I| = j$ and let N_j the number of distinct commutators $X^{[I]}$ with $|I| \le j$. If the X_j form a generating system, for each $x \in \Omega$ there is m such that $(X^{[I]})_x$ with $|I| \le m$ span $\mathbb{R}^n$.

LEMMA 1.5.1. *Let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω. Given $x \in \Omega$, let m be such that the commutators $X^{[I]}$ with $|I| \le m$ generate $\mathbb{R}^n$. Then the map*

$$\Phi(t_1, \dots, t_{N_m}) = \varphi_{X^{[I_1]}, t_1} \circ \cdots \circ \varphi_{X^{[I_{N_m}]}, t_{N_m}}(x)$$

is, for $\delta > 0$ small enough, a C^∞ surjection from the ball $|t| < \delta$ in $\mathbb{R}^{N_m}$ onto a neighborhood of x in Ω.

PROOF. Since Φ is smooth on a neighborhood of $t = 0$, it is sufficient to prove that $D\Phi(0)$ has rank n. But this is immediate, since $\partial_{t_j}\Phi(0) = (X^{[I_j]})_x$. □

LEMMA 1.5.2. *Let X, Y be two vector fields in Ω and let $x \in \Omega$. Then the curve*

$$\gamma(t) = \varphi_{Y,\sqrt{t}} \circ \varphi_{X,\sqrt{t}} \circ \varphi_{Y,-\sqrt{t}} \circ \varphi_{X,-\sqrt{t}}(x)$$

is C^1 on some interval $[0, \delta)$ and $\gamma'(0) = [X, Y]_x$.

PROOF. Clearly, γ is continuous and C^∞ for $t > 0$. If f is a smooth function on Ω,

$$f\big(\gamma(t)\big) = \exp(-\sqrt{t}X)\exp(-\sqrt{t}Y)\exp(\sqrt{t}X)\exp(\sqrt{t}Y)f(x) \ ,$$

hence

$$\begin{aligned}
\frac{d}{dt} f\big(\gamma(t)\big) = \frac{1}{2\sqrt{t}}\Big(&-\exp(-\sqrt{t}X)X\exp(-\sqrt{t}Y)\exp(\sqrt{t}X)\exp(\sqrt{t}Y) \\
&-\exp(-\sqrt{t}X)\exp(-\sqrt{t}Y)Y\exp(\sqrt{t}X)\exp(\sqrt{t}Y) \\
&+\exp(-\sqrt{t}X)\exp(-\sqrt{t}Y)X\exp(\sqrt{t}X)\exp(\sqrt{t}Y) \\
&+\exp(-\sqrt{t}X)\exp(-\sqrt{t}Y)\exp(\sqrt{t}X)Y\exp(\sqrt{t}Y)\Big)f(x) \\
= \exp(-\sqrt{t}X)&\frac{1}{2\sqrt{t}}\big(-X\exp(-\sqrt{t}Y) + \exp(-\sqrt{t}Y)X\big)\exp(\sqrt{t}X)\exp(\sqrt{t}Y)f(x) \\
+ \exp(-\sqrt{t}X)&\exp(-\sqrt{t}Y)\frac{1}{2\sqrt{t}}\big(-Y\exp(\sqrt{t}X) + \exp(\sqrt{t}X)Y\big)\exp(\sqrt{t}Y)f(x) \ ,
\end{aligned}$$

for $t > 0$. Moreover

$$\begin{aligned}
\lim_{t\to 0^+}\frac{1}{2\sqrt{t}}&\big(-X\exp(-\sqrt{t}Y) + \exp(-\sqrt{t}Y)X\big) \\
&= \lim_{t\to 0^+}\frac{1}{2\sqrt{t}}\big(-Y\exp(\sqrt{t}X) + \exp(\sqrt{t}X)Y\big) \\
&= \frac{1}{2}[X, Y] \ ,
\end{aligned}$$

so that

$$\lim_{t\to 0^+}\frac{d}{dt}f\big(\gamma(t)\big) = [X, Y]f(x) \ .$$

Therefore

$$\frac{d}{dt}_{|t=0} f\big(\gamma(t)\big) = [X, Y]f(x) \ ,$$

and the derivative is continuous at $t = 0$.

For $f(x) = x_j$, we then have

$$\frac{d}{dt}_{|t=0}\big(\gamma(t)\cdot e_j\big) = [X, Y]_x \cdot e_j \ ,$$

since $\nabla f = e_j$. $\square$

Denote the curve γ above as $\gamma^x_{X,Y}(t)$. Then $\gamma^x_{X,-Y}(-t)$ is C^1 on $(-\delta, 0]$ and its tangent vector at $t = 0$ is again $[X,Y]$. Therefore we can extend $\gamma^x_{X,Y}$ to a C^1 curve on $(-\delta, \delta)$ as

$$\gamma^x_{X,Y}(t) = \gamma^x_{(\operatorname{sgn} t)X,Y}(|t|) \ .$$

More generally, for fixed $\beta \in (0,1)$, define

$$\gamma^{x,\beta}_{X,Y}(t) = \varphi_{Y,t^\beta} \circ \varphi_{X,t^{1-\beta}} \circ \varphi_{Y,-t^\beta} \circ \varphi_{X,-t^{1-\beta}}(x)$$

for $t \geq 0$ and $\gamma^{x,\beta}_{X,Y}(t) = \gamma^{x,\beta}_{X,-Y}(-t)$ for $t < 0$. The same proof as above shows that $\gamma^{x,\beta}_{X,Y}$ is C^1 on some interval $(-\delta, \delta)$ and its tangent vector at $t = 0$ is $[X,Y]$.

We omit the proof of the following fact.

Lemma 1.5.3. *The function $\psi^\beta_{X,Y}(x,t) = \gamma^{x,\beta}_{X,Y}(t)$ is C^1 in both variables.*

We now extend the above construction to higher order commutators of vector fields. For a third-order commutator $\big[[X,Y],Z\big]$, start with

$$\gamma^{x,1/3}_{[X,Y],Z}(t) = \varphi_{Z,t^{1/3}} \circ \varphi_{[X,Y],t^{2/3}} \circ \varphi_{Z,-t^{1/3}} \circ \varphi_{[X,Y],-t^{2/3}}(x) \qquad (t \geq 0) \ ,$$

(and similarly for $t < 0$).

Adopting the notation $\psi^\beta_{X,Y;t}(x) = \psi^\beta_{X,Y}(x,t)$ replace the factors $\varphi_{[X,Y],t^{2/3}}$ by $\psi^{1/2}_{X,Y;t^{2/3}}$ to obtain a new curve,

$$\begin{aligned}
\gamma^x_{X,Y,Z}(t) &= \varphi_{Z,t^{1/3}} \circ \psi^{1/2}_{X,Y;t^{2/3}} \circ \varphi_{Z,-t^{1/3}} \circ \psi^{1/2}_{X,Y;-t^{2/3}}(x) \\
&= \varphi_{Z,t^{1/3}} \circ \varphi_{Y,t^{1/3}} \circ \varphi_{X,t^{1/3}} \circ \varphi_{Y,-t^{1/3}} \circ \varphi_{X,-t^{1/3}} \\
&\qquad \circ \varphi_{Z,-t^{1/3}} \circ \varphi_{Y,t^{1/3}} \circ \varphi_{X,-t^{1/3}} \circ \varphi_{Y,-t^{1/3}} \circ \varphi_{X,t^{1/3}}(x) \ .
\end{aligned}$$

An application of the Baker-Campbell-Hausdorff formula shows that this curve is C^1 on some interval $[0, \delta)$ and satisfies $\gamma'(0) = \big[[X,Y],Z\big]$. It follows from Lemma 1.5.2 that

$$\psi_{X,Y,Z}(x,t) = \gamma^x_{X,Y,Z}(t)$$

is C^1 in both variables.

Let $X_1, \dots, X_k$ be a given set of vector fields. For any $I = (i_1, \dots, i_p)$, we define inductively $\gamma^x_I(t)$ and $\psi_I(x,t)$ as follows[2].

If $|I| = 1$, we set $\gamma^x_i(t) = \psi_i(x,t) = \varphi_{X_i}(x,t)$. If $p = |I| > 1$, calling $I' = (i_1, \dots, i_{p-1})$, we set

$$\gamma^x_I(t) = \psi_I(x,t) = \varphi_{X_{i_p},t^{1/p}} \circ \psi_{I';t^{(p-1)/p}} \circ \varphi_{X_{i_p},-t^{1/p}} \circ \psi_{I';-t^{(p-1)/p}}(x)$$

for $t \geq 0$, and

$$\gamma^x_I(t) = \psi_I(x,t) = \varphi_{-X_{i_p},|t|^{1/p}} \circ \psi_{I';|t|^{(p-1)/p}} \circ \varphi_{-X_{i_p},-|t|^{1/p}} \circ \psi_{I';-|t|^{(p-1)/p}}(x)$$

for $t < 0$.

One can also see inductively that the explicit expression of $\psi_I(x,t)$ contains the composition of a number[3] $q = q(p)$ of factors $\varphi_{X_i,\pm|t|^{1/p}}$.

It remains true that ψ_I is C^1 in both variables. We can then conclude as follows.

Proposition 1.5.4. *Let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω, and let $x \in \Omega$. Let also $I_1, \dots, I_n$, with $|I_j| \leq m$ for every j, be such that $(X^{I_1})_x, \dots, (X^{I_n})_x$ form a basis of $\mathbb{R}^n$. Then the map*

$$\Psi_x(t_1, \dots, t_n) = \psi_{I_1,t_1} \circ \dots \circ \psi_{I_n,t_n}(x) \tag{1.5.1}$$

is, for $\delta > 0$ small enough, a C^1-diffeomorphism of the ball $|t| < \delta$ onto a neighborhood U_x of x in Ω.

---

[2]We keep the notation $\psi_{I;t}(x) = \psi_I(x,t)$.

[3]Precisely, $q(p) = 3 \cdot 2^{p-1} - 2$.

COROLLARY 1.5.5. *Let $p_j = |I_j|$. Every element $y = \Psi_x(t_1, \dots, t_n) \in U_x$ can be written as*

$$y = \varphi_{X_{i_s}, \varepsilon_s |t_{j_s}|^{1/p_{j_s}}} \circ \cdots \circ \varphi_{X_{i_1}, \varepsilon_1 |t_{j_1}|^{1/p_{j_1}}}(x) \ , \tag{1.5.2}$$

where $s = q(p_1) + \cdots + q(p_n)$, and for $1 \le \ell \le s$, $1 \le i_\ell \le k$, $1 \le j_\ell \le n$ and $\varepsilon_\ell = \pm 1$.

We have the tools now to present the control distance on Ω.

DEFINITION. *Let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω. A* sub-unitary curve *in Ω is a piecewise smooth curve γ such that $\gamma'(t)$, whenever defined, can be expressed as a linear combination*

$$\gamma'(t) = \sum_{j=1}^{k} c_j(t) (X_j)_{\gamma(t)} \ ,$$

with $\sum c_j^2(t) \le 1$.

THEOREM 1.5.6. *Assume that Ω is connected, and let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω. Then every pair of points in Ω can be joined by a sub-unitary curve.*

PROOF. Consider first two points x, y with $y = \Psi_x(t_1, \dots, t_n) \in U_x$. From (1.5.2) we obtain a sub-unitary curve $\gamma : [0, T] \to U$ joining x and y as follows.

Take $T = |t_{j_1}|^{1/p_1} + \cdots + |t_{j_s}|^{1/p_s}$, and split $[0, T]$ into adjacent subintervals $[\tau_{\ell-1}, \tau_\ell]$ with $\tau_\ell - \tau_{\ell-1} = |t_\ell|^{1/p_\ell}$. Calling

$$y_\ell = \varphi_{X_{i_\ell}, \varepsilon_\ell |t_{j_\ell}|^{1/p_{j_\ell}}} \circ \cdots \circ \varphi_{X_{i_1}, \varepsilon_1 |t_{j_1}|^{1/p_{j_1}}}(x) \ ,$$

we set $\gamma(0) = x$, and, inductively,

$$\gamma(\tau) = \varphi_{\varepsilon_\ell X_{i_\ell}, \tau - \tau_{\ell-1}}(y_{\ell-1}) \ ,$$

for $\tau_{\ell-1} \le \tau \le \tau_\ell$. In other words, on the interval $[\tau_{\ell-1}, \tau_\ell]$, γ follows the integral curve of $\pm X_{i_\ell}$ joining $y_{\ell-1}$ to y_ℓ. In particular, γ is sub-unitary.

So the subset of points in Ω that can be joined to a fixed point x_0 is open and non-empty. But also its complement is open. Since Ω is connected, every point can be joined to x_0 by a sub-unitary curve. ☐

DEFINITION. *Let $\Gamma_{x,y}$ be the family of all sub-unitary curves $\gamma : [0, T] \to \Omega$ joining x and y. The* control distance *$d(x, y)$ on Ω, induced by the system $\{X_1, \dots, X_k\}$, is the infimum of the lengths T of the domain of all the $\gamma \in \Gamma_{x,y}$.*

DEFINITION. *The control distance is a true distance, and it induces the euclidean topology on Ω. More precisely, given $x \in \Omega$, let m be such that the vectors $(X^I)_x$ with $|I| \le m$ span $\mathbb{R}^n$. Then there are constants c_1, c_2 such that, for y in a sufficiently small neighborhood of x,*

$$c_1 |x - y| \le d(x, y) \le c_2 |x - y|^{\frac{1}{m}} \ .$$

PROOF. Let B be a Euclidean closed ball centered at x and contained in Ω. By compactness, there is $M > 0$ such that $|(X_j)_{x'}| \le M$ for every j and every $x' \in B$. It follows easily that, for every sub-unitary curve γ, $|\gamma'(t)| \le \sqrt{k} M = M'$ whenever $\gamma(t) \in B$.

Let γ be a sub-unitary curve joining x and y, defined on the interval $[0, T]$. If γ is entirely contained in B, then the Euclidean length of γ is not larger than $M'T$, so that

$$T \ge \frac{|x - y|}{M'} \ .$$

Otherwise, let T_0 be the smallest t such that $\gamma(t)$ on the boundary of B. Then $\gamma_0 = \gamma_{|_{[0,T_0]}}$ is entirely contained in B, so that, by the previous argument,

$$T > T_0 \ge \frac{r}{M'} \ ,$$

if r is the radius of B. It follows that

$$d(x, y) \ge \frac{\min\{|x - y|, r\}}{M'} \ .$$

This proves the first inequality. To prove the second inequality, suppose that $y \in U_x = \Psi_x\big(B(0,\delta)\big)$, with $\delta \leq 1$, and consider the curve γ introduced in the proof of Theorem 1.5.6. Then $d(x,y) \leq T = |t_{j_1}|^{1/p_1} + \cdots + |t_{j_s}|^{1/p_s}$. Since $|t_{j_\ell}| \leq |t| = \big|(t_1,\dots,t_n)\big| < 1$ and $p_\ell \leq m$ for every ℓ, we have that

$$d(x,y) \leq s|t|^{1/m} \ .$$

On the other hand, since Ψ_x is bi-Lipschitz in x, $|x-y| \sim |t|$. □

A small refinement of the argument given above shows that, given any compact $K \subset \Omega$, an inequality of the form $c_1|x-y| \leq d(x,y) \leq c_2|x-y|^{\frac{1}{m}}$ holds for all $x, y \in K$, with m, c_1, c_2 depending on K.

EXAMPLES 4.

(a) Consider the control distance in $\mathbb{R}^2$ induced by the vector fields $X_1 = \partial_x$, $X_2 = x\partial_y$. If a curve γ goes through the point $P = (x,y)$ with $x \neq 0$ and has a tangent vector $v_P = (v_1, v_2)$ at this point, then the components of v_P w.r. to $(X_1)_P$ and $(X_2)_P$ are v_1 and $\frac{v_2}{x}$ respectively.

Suppose, for instance, we want to compute the distance between two points on the y-axis, e.g. $O = (0,0)$ and $A = (0,a)$. Any sub-unitary curve, defined on $[0,T]$ and joining O to A must have a horizontal tangent vector at both these points. Let $\gamma(t) = \big(\gamma_1(t), \gamma_2(t)\big)$ be such a curve. Then

$$\gamma_1'(t)^2 + \frac{\gamma_2'(t)^2}{\gamma_1(t)^2} \leq 1$$

for all but a finite number of values of t. In particular, $|\gamma_1'(t)| \leq 1$, so that $|\gamma_1(t)| \leq |t|$ for every t. It follows that $|\gamma_2'(t)| \leq |\gamma_1(t)| \leq |t|$, hence $|\gamma_2(t)| \leq \frac{t^2}{2}$ for every t. So, in order to have $\gamma_2(T) = a$, we must have $T \geq \sqrt{2a}$.

On the other hand, let γ be formed by three segments, the first joining O to $(b,0)$ with tangent vector $(1,0)$ in time b, the second joining $(b,0)$ to (b,a) with tangent vector $(0,b)$ in time $\frac{a}{b}$, and the third joining (b,a) with tangent vector $(-1,0)$ in time b. This is a sub-unitary curve, and the total time is $T = 2b + \frac{a}{b}$, which is minimal for $b = \sqrt{a/2}$, with $T = \sqrt{8a}$. Therefore

$$\sqrt{2a} \leq d(O,A) \leq \sqrt{8a} \ .$$

5.2. Hypoelliptic operators.

The results that we state here are due to L. Hörmander[4].

The Sobolev norm of order $s \in \mathbb{R}$ of a smooth function f on $\mathbb{R}^n$ with compact support is given by

$$\|f\|_{(s)} = \left(\int_{\mathbb{R}^n} |\hat f(\xi)|^2 (1+|\xi|^2)^s \, d\xi\right)^{1/2} \ . \tag{1.5.3}$$

The Sobolev space $H^s(\mathbb{R}^n)$ is the completion of $\mathcal{D}(\mathbb{R}^n)$ w.r. to the Sobolev norm of order s. Clearly,

$$H^{s'}(\mathbb{R}^n) \subseteq H^s(\mathbb{R}^n) \subseteq L^2(\mathbb{R}^n) \ , \tag{1.5.4}$$

if $0 < s < s'$. If k is an integer, the Plancherel formula shows that H^k consists of the L^2-functions whose distributional derivatives are L^2 up to order k, and that

$$\|f\|_{(k)} \sim \sum_{|\alpha| \leq k} \|\partial^\alpha f\|_2 \ . \tag{1.5.5}$$

If X is a vector field on $\Omega \subseteq \mathbb{R}^n$, it is pretty obvious that for every compact $K \subset \Omega$ there is a constant $C(K) > 0$ such that, for every smooth f supported on K,

$$\|Xf\|_2 \leq C(K)\|f\|_{(1)} \ .$$

The next statement says that a generating system of vector fields is sufficient to control some Sobolev norm.

[4] The proof can be found in L. Hörmander, *Hypoelliptic second-order differential equations*, Acta Math. vol.119 (1967), p.147-171. See the Appendix for a proof.

THEOREM 1.5.7. *Let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω. Let m be such that, at every $x \in \Omega$, the iterated commutators $(X^I)_x$ with $|I| \leq m$ span $\mathbb{R}^n$, and let $s < 1/m$. Then, for every compact $K \subset \Omega$, there is a constant $C(K) > 0$ such that, for every smooth f supported on K,*

$$\|f\|_{(s)} \leq C(K)\Big(\|f\|_2 + \sum_{j=1}^{k} \|X_j f\|_2\Big) .$$

A closely related issue is hypoellipticity of differential operators obtained by composition of vector fields.

DEFINITION. *Let L be a linear differential operator on Ω with smooth coefficients. Then L is said to be* hypoelliptic *if for any distribution $u \in \mathcal{D}'(\Omega)$ and any $\Omega' \subseteq \Omega$, the condition $Lu \in C^\infty(\Omega')$ implies that $u \in C^\infty(\Omega')$.*

For a distribution $u \in \mathcal{D}'(\Omega)$, the *singular support* of u, denoted by $\operatorname{sing\,supp} u$, is defined as the complement of the largest open set Ω' where u is C^∞. Then the hypoellipticity condition is equivalent to requiring that

$$\operatorname{sing\,supp} u \subseteq \operatorname{sing\,supp}(Lu) . \tag{1.5.6}$$

The terminology comes from the fact that elliptic linear operators with smooth coefficients satisfy (1.5.6). However there are non-elliptic hypoelliptic operators, such as the heat operator $\partial_t - \Delta_x$ on $\mathbb{R}^{n+1}$ (with $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$).

THEOREM 1.5.8. *Let $\{X_1, \dots, X_k\}$ be a generating system of vector fields on Ω. Then the operators*

$$L_1 = \sum_{j=1}^{k} X_j^2 , \qquad L_2 = X_1 - \sum_{j=2}^{k} X_j^2$$

are hypoelliptic.

An operator of the form L_1 is given different names: *sub-Laplacian*, *Hörmander sum of squares*, or others. Among the operators of the form L_2, particularly important are the heat operators $\partial_t - L_1$, defined on $\Omega \times \mathbb{R}$, L_1 as above.

In Chapter 5 we will prove hypoellipticity of L_1, following Hörmander's original proof.

CHAPTER 2

Structure of nilpotent Lie algebras and homogeneous groups

1. Central series and filtrations

We call a *Lie bracket of order k* in a Lie algebra an expression of the form

$$\big[\cdots\big[[x_0, x_1], x_2\big]\cdots, x_k\big] \ .$$

Definition.

(i) *A subalgebra $\mathfrak{h}$ of a Lie algebra $\mathfrak{g}$ is called an* ideal *if $[x,y] \in \mathfrak{h}$ for every $x \in \mathfrak{g}$ and $y \in \mathfrak{h}$.*

(ii) *If $\mathfrak{h}$ is an ideal of $\mathfrak{g}$, the* quotient Lie algebra *is the quotient space $\mathfrak{g}/\mathfrak{h}$ with bracket*

$$[x+\mathfrak{h}, y+\mathfrak{h}] = [x,y] + \mathfrak{h} \ . \tag{2.1.1}$$

(iii) *The* center *of a Lie algebra $\mathfrak{g}$ consists of those elements x such that $[x,y]=0$ for every $y \in \mathfrak{g}$.*

(iv) *A nilpotent Lie algebra $\mathfrak{g}$ is called* step k *(or* of length k*) if k is the smallest integer for which all Lie brackets of order k are zero.*

The requirement that $\mathfrak{h}$ be an ideal in the definition of a quotient Lie algebra is necessary, in order that (2.1.1) be a good definition.

We define the *descending central series* of a Lie algebra $\mathfrak{g}$.

We use the following notation: if V, W are linear subspaces of a Lie algebra $\mathfrak{g}$, we denote by $[V,W]$ the linear subspace of $\mathfrak{g}$ spanned by the Lie brackets $[v,w]$ with $v \in V$, $w \in W$. We also tend to reserve gothic letters to subspaces that are also Lie subalgebras.

We define inductively

$$\mathfrak{g}^1 = \mathfrak{g} \ , \qquad \mathfrak{g}^{j+1} = [\mathfrak{g}, \mathfrak{g}^j] \ ;$$

i.e. $\mathfrak{g}^j$ is the subspace spanned by the Lie brackets of order $j-1$.

Clearly, $\mathfrak{g}^{j+1} \subseteq \mathfrak{g}^j$. Hence each $\mathfrak{g}^j$ is an ideal of $\mathfrak{g}$. Moreover,

$$[\mathfrak{g}^i, \mathfrak{g}^j] \subseteq \mathfrak{g}^{i+j} \ .$$

The following statement is a direct consequence of the definitions.

Proposition 2.1.1. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if there is k such that $\mathfrak{g}^k = \{0\}$. If this is the case, the step of $\mathfrak{g}$ is the largest m for which $\mathfrak{g}^m \neq \{0\}$.*

An example where the condition in Proposition 2.1.1 is not verified is the following. Let $\mathfrak{g}$ be $\mathbb{R}^3$ with the wedge product, i.e., $[x,y] = x \wedge y$. It is easily verified that $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$, hence $\mathfrak{g}^j = \mathfrak{g}$ for every j. A Lie algebra satisfying the condition $[\mathfrak{g},\mathfrak{g}] = \mathfrak{g}$ is called *semisimple*.

The descending central series of a nilpotent Lie algebra is useful to identify natural generating systems of $\mathfrak{g}$.

Proposition 2.1.2. *Let $\mathfrak{g}$ be a nilpotent Lie algebra and let W be a linear subspace of $\mathfrak{g}$ complementary to $[\mathfrak{g},\mathfrak{g}]$, i.e., $\mathfrak{g} = W \oplus [\mathfrak{g},\mathfrak{g}]$, and let $\{e_1, \dots, e_k\}$ be a basis of W. Then $\{e_1, \dots, e_k\}$ generates $\mathfrak{g}$.*

Proof. By the direct sum decomposition in the statement, any element $x \in \mathfrak{g}$ decomposes uniquely as $x = w + r_2$ with $w \in W$ and $r_2 \in [\mathfrak{g},\mathfrak{g}] = \mathfrak{g}^2$. Then r_2 is a finite sum of terms of the form $[y'_j, y''_j]$, with $y'_j, y''_j \in \mathfrak{g}$. Decompose, as above, $y'_j = w'_j + z'_j$ and $y''_j = w''_j + z''_j$ with $w'_j, w''_j \in W$ and $z'_j, z''_j \in \mathfrak{g}^2$. Then

$$[y'_j, y''_j] = [w'_j, w''_j] + \text{ terms in } \mathfrak{g}^3 \ ,$$

hence

$$x = w + \sum_j [w'_j, w''_j] + r_3 \ , \text{ with } r_3 \in \mathfrak{g}^3 \ .$$

This procedure can be iterated obtaining, at each step, higher order Lie brackets involving elements of W plus a remainder term $r_j \in \mathfrak{g}^j$. When we arrive at $j = m+1$ the remainder term is zero. The last step is the expansion of any element of W in terms of the basis $\{e_i\}$ and use of bilinearity of the Lie bracket. □

The descending central series is a notable example of a *descending filtration* (or simply a filtration) of a Lie algebra $\mathfrak{g}$. By definition is a finite family $\{\mathfrak{v}_j\}_{1\le j\le m+1}$ of subspaces such that

(i) $\mathfrak{v}_{j+1} \subseteq \mathfrak{v}_j$ for $1 \le j \le m$;
(ii) $\mathfrak{v}_1 = \mathfrak{g}$, $\mathfrak{v}_{m+1} = \{0\}$ (one then allows $\mathfrak{v}_k = \{0\}$ also for $k > m+1$);
(iii) for every i, j, $[\mathfrak{v}_i, \mathfrak{v}_j] \subseteq \mathfrak{v}_{i+j}$.

The definition implies that each $\mathfrak{v}_j$ is in fact an ideal of $\mathfrak{g}$. By induction,

$$\mathfrak{g}^j \subseteq \mathfrak{v}_j \ , \tag{2.1.2}$$

so that filtrations only exist in nilpotent Lie algebras and the length of a filtration cannot be smaller than the step of the Lie algebra.

One also defines the *ascending central series* of a Lie algebra $\mathfrak{g}$ as follows. We start with the *center of* $\mathfrak{g}$,

$$\mathfrak{z}(\mathfrak{g}) = \{x \in \mathfrak{g} : [x,y] = 0 \ \forall\, y \in \mathfrak{g}\} \ ,$$

and define inductively

$$\mathfrak{g}_1 = \mathfrak{z}(\mathfrak{g}) \ , \qquad \mathfrak{g}_{j+1} = \{x \in \mathfrak{g} : [x,y] \in \mathfrak{g}_j \ \forall\, y \in \mathfrak{g}\} \ .$$

Using the Jacobi identity, one proves by induction that each $\mathfrak{g}_j$ is an ideal of $\mathfrak{g}$ and that $\mathfrak{g}_j \subseteq \mathfrak{g}_{j+1}$ for every j. Introducing the quotient algebras $\mathfrak{h}_j = \mathfrak{g}/\mathfrak{g}_j$ with canonical projection π_j, it can easily be verified that $\pi_{j+1}(\mathfrak{g}_j) = \mathfrak{z}(\mathfrak{h}_{j+1})$ for every j.

Proposition 2.1.3. *A Lie algebra $\mathfrak{g}$ is nilpotent if and only if there is k such that $\mathfrak{g}_k = \mathfrak{g}$. If this is the case, the step of $\mathfrak{g}$ is the smallest m for which $\mathfrak{g}_m = \mathfrak{g}$.*

Proof. Assume that $\mathfrak{g}$ is nilpotent, and that all iterated k-fold brackets are zero. Then the center $\mathfrak{z}(\mathfrak{g})$ contains all the Lie brackets of order k. It follows that $\mathfrak{g}_1 = \mathfrak{g}/\mathfrak{z}$ is nilpotent of step $\le k-1$. The converse is also true, as one can see immediately.

By induction, the fact that $\mathfrak{g}$ is nilpotent of step $\le k$ is equivalent to $\mathfrak{g}_k = \{0\}$. The conclusion follows easily. □

2. Dilations and homogeneous groups

Let G_1, G_2 be Lie groups and $\psi : G_1 \to G_2$ be a smooth group homomorphism. If v is in the Lie algebra $\mathfrak{g}_1$ of G_1, then $\psi(\exp_{G_1} tv)$ is a one-parameter group in G_2. By Corollary 1.3.5 of Chapter 1, there is a unique $v' = \psi_*(v) \in \mathfrak{g}_2$ such that $\psi(\exp_{G_1} tv) = \exp_{G_2}(tv')$.

Lemma 2.2.1. *The map $\psi_* : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism, i.e. it is linear and*

$$\psi_*\big([u,v]\big) = \big[\psi_*(u), \psi_*(v)\big] \ .$$

If G_1 is connected, ψ_* uniquely determines ψ.

PROOF. The fact that $\psi_*(sv) = s\psi_*(v)$ for $s \in \mathbb{R}$ is obvious. Let X be the left-invariant vector field[1] on G_1 such that $X_{e_1} = v$, and let Y be the left-invariant vector field on G_2 such that $Y_{e_2} = v' = \psi_*(v)$.

Since

$$\begin{aligned} X(f\circ\psi)(x) &= \frac{d}{dt}_{|t=0} f\circ\psi\big(x\exp_{G_1}(tv)\big) \\ &= \frac{d}{dt}_{|t=0} f\Big(\psi(x)\psi\big(\exp_{G_1}(tv)\big)\Big) \\ &= \frac{d}{dt}_{|t=0} f\big(\psi(x)\exp_{G_1}(tv')\big) \\ &= Yf\big(\psi(x)\big)\ , \end{aligned}$$

for every smooth function f on G_2, the following identity holds:

$$X(f\circ\psi) = (Yf)\circ\psi\ . \tag{2.2.1}$$

Moreover Y is the only left-invariant vector field on G_2 for which this identity holds, because (2.2.1) implies that

$$Yf(e_2) = \frac{d}{dt}_{|t=0} f\circ\psi\big(\exp_{G_1}(tv)\big) = \frac{d}{dt}_{|t=0} f\big(\exp_{G_1}(tv')\big) = v'(f)\ .$$

It follows directly from (2.2.1) that if X' is another left-invariant vector field on G_1 and Y' is the corresponding vector field on G_2, then $X + X'$ corresponds to $Y + Y'$ and $[X, X']$ to $[Y, Y']$.

Suppose now that G_1 is connected, and ψ, η are smooth homomorphisms from G_1 to G_2 such that $\psi_* = \eta_*$. Then

$$\psi\big(\exp_{G_1}(tv)\big) = \exp_{G_2}\big(t\psi_*(v)\big) = \eta\big(\exp_{G_1}(tv)\big)\ .$$

By Proposition 1.4.1 of Chapter 1, the image of the exponential map of G_1 contains a full neighborhood U of the unit element e_1. Therefore $\psi = \eta$ on U. It follows that $\psi = \eta$ on the subgroup H of G_1 generated by U. If $x \in H$, then $xU \subset H$, so that H is open. But

$$G_1 \setminus H = \bigcup_{x\notin H} xH\ ,$$

which is also open. Since G_1 is connected, $H = G_1$. □

One says that the Lie algebra homomorphism ψ_* is *induced by* the group homomorphism ψ. Observe that ψ_* is nothing but the *differential of* ψ at the identity of G_1.

It is not true in general that any homomorphism from $\mathfrak{g}_1$ to $\mathfrak{g}_2$ is induced by a group homomorphism from G_1 to G_2. This statement becomes true if we assume that G_1 is connected and simply connected. We give the proof for nilpotent groups.

PROPOSITION 2.2.2. *Let G_1, G_2 be connected nilpotent Lie groups, with Lie algebras $\mathfrak{g}_1, \mathfrak{g}_2$, and let G_1 be simply connected. If $\lambda : \mathfrak{g}_1 \to \mathfrak{g}_2$ is a Lie algebra homomorphism, then there is a unique smooth group homomorphism ψ from G_1 to G_2 such that $\psi_* = \lambda$.*

PROOF. We assume first that also G_2 is simply connected. We identify G_j with $\mathfrak{g}_j$ endowed with the product (1.4.1) in Chapter 1, and define $\psi = \lambda$. It is clear that

$$Z_{j,k}\big(\lambda(x), \lambda(y)\big) = \lambda\big(Z_{j,k}(x,y)\big)\ ,$$

because $Z_{j,k}$ is a linear combination of iterated Lie brackets. Therefore

$$S\big(\lambda(x), \lambda(y)\big) = \lambda\big(S(x,y)\big)\ ,$$

which says that ψ is a group homomorphism. Smoothness is obvious and uniqueness follows from Lemma 2.2.1.

If Z is a central discrete subgroup of G_2 and π is the canonical projection of G_2 onto G_2/Z, then $\pi\circ\psi$ is a smooth homomorphism whose differential is λ. By Theorem 1.4.3 in Chapter 1, this covers the general case. □

[1] Here we are implicitely identifying $\mathfrak{g}$ with the tangent space T_eG.

Let V be a real vector space. A family $\{\delta_t\}_{t>0}$ of linear maps of V to itself is called a *set of dilations* on V if there are real numbers $\lambda_j > 0$ and subspaces W_{λ_j} of V such that V is the direct sum of the W_{λ_j} and

$$(\delta_t)_{|W_{\lambda_j}} = t^{\lambda_j}\mathrm{Id}$$

for every j.

DEFINITION. *Let $\mathfrak{g}$ be a Lie algebra and $\{\delta_t\}_{t>0}$ be a set of dilations on its underlying vector space. If each δ_t is an automorphism of $\mathfrak{g}$, then the pair $\big(\mathfrak{g}, \{\delta_t\}\big)$ is called a* homogeneous Lie algebra.

A homogeneous Lie group *is a connected Lie group G endowed with a family $\{D_t\}_{t>0}$ of automorphisms such that its Lie algebra $\mathfrak{g}$ is homogeneous under the $\delta_t = (D_t)_*$.*

It must be observed that the same Lie algebra can have different homogeneous structures.

EXAMPLES 5.

(a) If $\mathfrak{g}$ is abelian, any set of dilations on it makes it a homogeneous Lie algebra.

(b) Let $\mathfrak{h}_n$ be the $(2n+1)$-dimensional *Heisenberg algebra*, with basis $e_1, \dots, e_n$, $f_1, \dots, f_n$, h and Lie brackets

$$[e_i, e_j] = [f_i, f_j] = 0 \ , \quad [e_i, f_j] = \delta_{i,j} h \ , \quad [e_i, h] = [f_i, h] = 0 \ , \tag{2.2.2}$$

for every i, j. Setting

$$\delta_t(e_j) = t^{\alpha_j} e_j \ , \qquad \delta_t(f_j) = t^{\beta_j} f_j \ , \qquad \delta_t(h) = t^{\gamma} h \ ,$$

with $\alpha_j, \beta_j, \gamma > 0$, the $\{\delta_t\}$ make $\mathfrak{h}_n$ a homogeneous algebra if and only if $\alpha_j + \beta_j = \gamma$ for every j.

(c) Let $\mathfrak{g}$ be the Lie algebra consisting of the real upper triangular matrices

$$A = \begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ 0 & 0 & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \ ,$$

with Lie bracket $[A, B] = AB - BA$. Then

$$\delta_t(A) = (t^{j-i} a_{i,j})$$

is a set of dilations on $\mathfrak{g}$.

LEMMA 2.2.3. *A homogeneous Lie algebra is nilpotent. A homogeneous Lie group is simply connected.*

PROOF. Let $x \in W_{\lambda_j}$ and $y \in W_{\lambda_k}$. Then

$$\delta_t[x, y] = \big[\delta_t(x), \delta_t(y)\big] = t^{\lambda_j + \lambda_k}[x, y] \ ,$$

i.e. $[x, y] \in W_{\lambda_j + \lambda_k}$. This implies that if $m \min \lambda_j > \max \lambda_j$, any Lie bracket of order m in $\mathfrak{g}$ is zero.

Assume now that G is homogeneous. By Theorem 1.4.3 in Chapter 1, we can assume that $G = \tilde{G}/Z$, where $\tilde{G}$ has the same Lie algebra of G, it is connected and simply connected, and Z is a discrete normal subgroup of $\tilde{G}$.

The dilations D_t on G induce dilations δ_t of $\mathfrak{g}$, and these in turn induce dilations $\tilde{D}_t$ on $\tilde{G}$, by Proposition 2.2.2. If π is the canonical projection of $\tilde{G}$ onto G, we claim that

$$D_t \circ \pi = \pi \circ \tilde{D}_t \ , \tag{2.2.3}$$

for every t. This follows from Lemma 2.1, once we observe that π_* is the identity map of $\mathfrak{g}$, so that $(D_t \circ \pi)_* = (\pi \circ \tilde{D}_t)_*$.

But (2.2.3) implies that $D_t(Z) \subseteq Z$ for every t. Since Z is discrete, it must be trivial. $\square$

Let G be a homogeneous Lie group, and let X be a left-invariant vector field such that $X_e = v \in W_{\lambda_j}$. The left-invariant vector field Y such that $Y_e = \delta_t v = t^{\lambda_j} v$ is clearly $t^{\lambda_j} X$.

Therefore, by (2.2.1),

$$X(f \circ D_t)(x) = t^{\lambda_j} (Xf)(D_t x) \ , \tag{2.2.4}$$

which generalizes the identity

$$\frac{d}{dx} f(tx) = t f'(tx)$$

for functions of one variable.

3. Graded and stratified algebras

Definition. *A* gradation *on* $\mathfrak{g}$ *is a decomposition of* $\mathfrak{g}$ *as the direct sum of linear subspaces* $\{W_j\}_{1 \le j \le m}$ *such that* $[W_j, W_k] \subseteq W_{j+k}$ *(or* $[W_j, W_k] = \{0\}$ *if* $j + k > m$*). A Lie algebra endowed with a gradation is called a* graded Lie algebra*, and the associated connected and simply connected Lie group a* graded Lie group.

From a gradation $\{W_j\}$, one constructs a filtration $\{\mathfrak{v}_j\}$ setting

$$\mathfrak{v}_j = W_j \oplus \cdots \oplus W_m \ . \tag{2.3.1}$$

However, not every filtration can be obtained from a gradation as in (2.3.1). An example is given by the filtration $\mathfrak{v}_3 \subset \mathfrak{v}_2 \subset \mathfrak{v}_1$ on the Heisenberg algebra $\mathfrak{h}_2$, with

$$\mathfrak{v}_1 = \mathfrak{h}_2 \ , \quad \mathfrak{v}_2 = \operatorname{span}\{e_2, h\} \ , \quad \mathfrak{v}_3 = \mathbb{R}h \ .$$

The proof is left to the reader.

If $\{W_j\}$ is a gradation of $\mathfrak{g}$, the dilations

$$\delta_t(x) = t^j x \ , \qquad \text{if } x \in W_j$$

are automorphism, so that $\mathfrak{g}$ canonically inherits a homogeneous structure.

Conversely, if the dilations δ_t on a homogeneous Lie algebra have eigenvalues t^j (i.e. with integer exponents), the eigenspaces W_j relative to the eigenvalues t^j form a gradation of $\mathfrak{g}$.

A nilpotent Lie algebra may admit more than one gradation, but there are nilpotent Lie algebras that cannot be graded. In fact, there are nilpotent Lie algebras that even do not admit any set of (automorphic) dilations. As a curiosity, the lowest dimensional nilpotent Lie algebra without dilations is spanned by elements $e_1, \dots, e_7$ with the following non-trivial Lie brackets among the basis elements[2]:

$$[e_1, e_j] = e_{j+1} \ \ (2 \le j \le 6), \quad [e_2, e_3] = e_6 \ , \quad [e_2, e_4] = [e_5, e_2] = [e_3, e_4] = e_7 \ .$$

Definition. *A* stratified Lie algebra *is a graded Lie algebra* $\mathfrak{g}$ *such that* W_1 *generates* $\mathfrak{g}$.

If $\mathfrak{g}$ is stratified, then

$$W_j = \underbrace{\big[\cdots[W_1, W_1], W_1 \cdots\big]}_{j \text{ times}} \ . \tag{2.3.2}$$

Examples of stratified Lie algebras are:

- the Heisenberg algebra $\mathfrak{h}_n$, with $W_1 = \operatorname{span}\{e_j, f_j\}_{1 \le j \le n}$;
- the algebra in Example 5(b) of Section 2, with W_1 consisting of the matrices $(a_{i,j})$ such that $a_{i,j} = 0$ unless $j = i + 1$.

[2] See R. Goodman, *Nilpotent Lie groups: Structure and Applications to Analysis.*

4. Free Lie algebras

Given two integers $m \geq 1$ and $n \geq 2$, there is a canonical way to construct an m-step nilpotent Lie algebra with k generators. This is called the *free nilpotent Lie algebra* $\mathfrak{f}_{n,m}$ and satisfies a universal property that we will see below.

Given a set $E = \{e_1, \dots, e_n\}$ with n elements, let $\mathcal{A}_E$ be the *free associative algebra* generated by E over $\mathbb{R}$. This may be viewed as the linear span of all the "monomials"

$$e^I = e_{i_1} e_{i_2} \cdots e_{i_k} \ ,$$

where n is an arbitrary integer, $I = (i_1, \dots, i_k)$ is a k-tuple of indices, $1 \leq i_j \leq n$. The product is the natural multiplication with non-commuting variables.

Observe that $\mathcal{A}_E$ admits an infinite gradation: if $\mathcal{A}_k$ is the linear span of the monomials of degree k, then $\mathcal{A}_k \mathcal{A}_j \subseteq \mathcal{A}_{k+j}$.

As any associative algebra (see Example 1(d) in Chapter 1), $\mathcal{A}_E$ is a Lie algebra with bracket $[x, y] = xy - yx$. Clearly, $[\mathcal{A}_k, \mathcal{A}_j] \subseteq \mathcal{A}_{k+j}$.

Let $\mathfrak{g}_E$ be the Lie subalgebra of $\mathcal{A}_E$ generated by E, i.e. the linear span of the iterated Lie brackets of elements of E. Since $E \subset \mathcal{A}_1$, any k-fold iterated Lie bracket belongs to $\mathcal{A}_k$. Therefore, if we set $W_k = \mathfrak{g}_E \cap \mathcal{A}_k$, we have that

$$\mathfrak{g}_E = \bigoplus_{k=1}^{\infty} W_k \ ,$$

and $[W_j, W_k] \subseteq W_{k+j}$.

Let $\mathfrak{h}_m$ be the sum of the W_k with $k > m$. Then $\mathfrak{h}_m$ is an ideal in $\mathfrak{g}_E$, and the quotient Lie algebra $\mathfrak{g}_E/\mathfrak{h}_m$ is clearly finite dimensional and nilpotent of step m.

DEFINITION. *The Lie algebra* $\mathfrak{f}_{n,m} = \mathfrak{g}_E/\mathfrak{h}_m$ *is called the* free nilpotent Lie algebra *of step m with n generators.*

It follows from the construction that the projections $\tilde{W}_k$ of the W_k with $k \leq m$ in $\mathfrak{g}_E/\mathfrak{h}_m$ give a stratification of $\mathfrak{f}_{n,m}$.

Computing the dimension of $\mathfrak{f}_{n,m}$ requires computing the dimension of each W_k. For $k = 2$ this equals $n(n-1)/2$, because a basis is given by the elements $e_{i,j} = e_i e_j - e_j e_i = [e_i, e_j]$ for $i < j$. For $k \geq 3$ the computation is harder, because one has to take into account the linear dependence relations among the iterated Lie brackets produced by the Jacobi identity.

A better "model" for $\mathfrak{f}_{n,2}$ is the direct sum $V \oplus \Lambda^2 V$, where $V = W_1$ is an n-dimensional vector space and $\Lambda^2 V = W_2$ is its exterior product, with Lie bracket

$$[v, v'] = v \wedge v'$$

from $W_1 \times W_1$ to W_2.

Observe that $\mathfrak{f}_{2,2}$ is isomorphic to the Heisenberg algebra $\mathfrak{h}_1$.

We state without proof the following result[3].

LEMMA 2.4.1. *Let φ be any function from the set E to a Lie algebra $\mathfrak{g}'$. Then φ extends uniquely to a Lie algebra homomorphism from $\mathfrak{g}_E$ to $\mathfrak{g}'$.*

From this we derive the following consequences.

THEOREM 2.4.2. *Let φ be any function from the set E to an m-step nilpotent Lie algebra $\mathfrak{g}'$. Then φ extends uniquely to a Lie algebra homomorphism from $\mathfrak{f}_{n,m}$ to $\mathfrak{g}'$.*

Every m-step nilpotent Lie algebra $\mathfrak{g}'$ with n generators is isomorphic to a quotient of $\mathfrak{f}_{n,m}$.

[3]It can be found in N. Jacobson, *Lie algebras.*

Proof. It follows from Lemma 2.4.1 that φ extends uniquely to a homomorphism $\tilde\varphi : \mathfrak{g}_E \to \mathfrak{g}'$. Since $\mathfrak{g}'$ is m-step nilpotent, $\tilde\varphi(W_k) = \{0\}$ for $k > m$. Therefore $\tilde\varphi$ factors modulo $\mathfrak{h}_m$.

If $x_1, \dots, x_n$ generate $\mathfrak{g}'$, take $\varphi : E \to \mathfrak{g}'$ as $\varphi(e_j) = x_j$. The induced homomorphism from $\mathfrak{f}_{n,m}$ to $\mathfrak{g}'$ is clearly surjective. $\square$

5. Homogeneous norms

Definition. *Let G be a homogeneous group with dilations $\{D_t\}$. A* homogeneous norm *on G, relative to the given dilations, is a continuous function $|\ | : G \to [0, +\infty)$ such that*

(i) $|x| = 0$ *if and only if x is the identity element;*
(ii) $|x^{-1}| = |x|$ *for every x;*
(iii) $|D_t(x)| = t|x|$ *for every $x \in G$ and $t > 0$.*

The existence of homogeneous norms on any homogeneous group G can be shown by the following argument.

We can assume that G has its Lie algebra $\mathfrak{g}$ as underlying manifold, with the product $xy = S(x, y)$. In this setting, $\delta_t = (D_t)_* = D_t$. Let W_j be the eigenspaces of each δ_t, with eigenvalues t^{λ_j}. If $\|\ \|_j$ is any vector space norm on W_j, and x_j denotes the W_j-component of $x \in \mathfrak{g}$, we define

$$|x| = \sum_{j=1}^{m} \|x_j\|_j^{1/\lambda_j} \ . \tag{2.5.1}$$

The properties of a homogenous norm are easily checked.

Proposition 2.5.1. *Every homogeneous group G admits a homogeneous norm that is smooth away from the unit element.*

Two homogeneous norms $|\ |$ and $|\ |'$ on G are mutually equivalent, in the sense that there are constants $A, B > 0$ such that

$$A|x| \le |x|' \le B|x| \ ,$$

for every x in the group.

If $|\ |$ is a homogeneous norm on G, there is a constant $C > 0$ such that

$$|xy| \le C\big(|x| + |y|\big) \ , \tag{2.5.2}$$

for every $x, y \in G$.

Proof. As before, we can restrict ourselves to the case of a homogenous Lie algebra $\mathfrak{g}$ with product $xy = S(x, y)$.For every j choose a Euclidean norm $\|\ \|_j$ on W_j and let

$$\|x\| = \left(\sum_{j=1}^{m} \|x_j\|_j^2 \right)^{1/2} \ , \tag{2.5.3}$$

and consider the function

$$F(t, x) = \|D_t(x)\|^2 = \sum_{j=1}^{m} t^{2\lambda_j} \|x_j\|_j^2 \ .$$

For fixed $x \neq 0$, $F(t, x)$ is continuous and strictly increasing in t, it tends zero for $t \to 0$ and to infinity for $t \to \infty$. Therefore there is a unique $t > 0$ such that $\|D_t x\| = 1$. Define

$$|x| = t^{-1} \ . \tag{2.5.4}$$

But $F(t, x)$ is smooth from $(0, +\infty) \times (\mathfrak{g} \setminus \{0\})$ to $(0, +\infty)$ and $|x|^{-1}$ is the implicit function for $F(t, x) = 1$. Since $\partial_t F(x, t)$ is always different from zero, the implicit function is smooth.

Setting $|0| = 0$, the properties of a homogeneous norm are satisfied, and this proves the first statement.

For the second statement, it is sufficient to prove that any homogeneous norm is equivalent to (2.5.4). Observe that the unit sphere Σ in the Euclidean norm (2.5.3) coincides with the set where $|x| = 1$ for the homogeneous norm (2.5.4).

If $|\ |'$ is any other homogeneous norm, its restriction to Σ is never zero. By compactness, there are constants $A, B > 0$ such that

$$A \le |x|' \le B\ ,$$

for any $x \in \Sigma$. If $y \neq 0$, let $t^{-1} = |y|$, so that $D_t(y) \in \Sigma$. Then

$$A \le |D_t(y)|' \le B\ ,$$

and the conclusion follows.

For the last statement, consider the closed unit ball $B = \{|x| \le 1\}$. It follows from the previous proof that this is a compact set for any homogeneous norm. Because the product map $xy = S(x,y)$ is continuous, the set $B^2 = \{xy : x, y \in B\}$ is also compact. Therefore there is a constant $C > 0$ such that $|xy| \le C$ for every $x, y \in B$.

Let $x, y \in \mathfrak{g}$. If both of them are 0, there is nothing to prove. If not, let $t^{-1} = |x| + |y| > 0$. Then $D_t(x)$ and $D_t(y)$ are in B, so that

$$t|xy| = |D_t(xy)| = |D_t(x)D_t(y)| \le C\ ,$$

and this concludes the proof. □

Let $|\ |$ be a homogeneous norm on G. Then

$$d(x,y) = |x^{-1}y| \tag{2.5.5}$$

satisfies the following properties:

(1) $d(x,y) = d(y,x)$ for every $x, y \in G$;
(2) $d(x,y) = 0$ if and only if $x = y$;
(3) $d(x,z) \le C\big(d(x,y) + d(y,z)\big)$ for every $x, y, z \in G$;
(4) $d(zx, zy) = d(x,y)$ for every $x, y, z \in G$;
(5) $d\big(D_t(x), D_t(y)\big) = t\, d(x,y)$ for every $x, y \in G$ and $t > 0$.

Such a function d is the *left-invariant homogeneous quasi-distance* induced by the given norm. Clearly, every left-invariant homogeneous quasi-distance on G is induced by a homogeneous norm through (2.5.5).

The function

$$d'(x,y) = |xy^{-1}|$$

is obviously a right-invariant homogeneous quasi-distance.

It is natural to ask if a homogeneous group G admits a left-(or right-)invariant homogeneous distance, i.e. satisfying the quasi-triangular inequality (3) with $C = 1$. This is equivalent to asking if G admits a homogeneous norm such that $|xy| \le |x| + |y|$.

As we are going to see, this is true for stratified groups and can be proved using *control distances.*

As a preliminary remark, observe that if $\gamma(t)$ is a smooth curve, its tangent vector $\gamma'(t_0)$ is the element of $T_{\gamma(t_0)}G$ such that

$$\gamma'(t_0)(f) = \frac{d}{dt}_{|t=t_0} f\big(\gamma(t)\big)\ .$$

It will be convenient for us to identify the tangent vector with an element of the Lie algebra. We therefore use (1.3.2)in Chapter 1 to "transport" $\gamma'(t_0)$ to T_eG as

$$\tilde\gamma'(t_0)(f) = \frac{d}{dt}_{|t=t_0} f\big(\gamma(t_0)^{-1}\gamma(t)\big) \in T_eG\ . \tag{2.5.6}$$

This will allow us to apply any norm defined on the Lie algebra to tangent vectors at any point.

We assume now that the homogeneous structure on G derives from a stratification $\{W_j\}$ of $\mathfrak{g}$.

THEOREM 2.5.2. *Let $\{X_1, \ldots, X_k\}$ be a basis of W_1, and let $d(x,y)$ be the induced control distance on G. Then d is a left-invariant homogeneous distance.*

PROOF. We must only prove properties (4) and (5). As to (4), it is sufficient to prove that, if $\gamma(s)$ is a sub-unitary curve, then also $\gamma_1(s) = z\gamma(s)$ is sub-unitary for every $z \in G$. But, applying (2.5.6), we see that

$$\tilde{\gamma}_1'(s) = \tilde{\gamma}'(s)$$

for any s, and the conclusion is immediate.

In order to prove (5), given a sub-unitary curve $\gamma(s)$, we prove that $\gamma_2(s) = D_t\big(\gamma(s/t)\big)$ is also sub-unitary for every $t > 0$. By (2.5.6) we have

$$\begin{aligned}
\tilde{\gamma}_2'(s_0)(f) &= \frac{d}{ds}_{|_{s=s_0}} f\big(\gamma_2(s_0)^{-1}\gamma_2(s)\big) \\
&= \frac{d}{ds}_{|_{s=s_0}} f \circ D_t\big(\gamma(s_0/t)^{-1}\gamma(s/t)\big) \\
&= \frac{1}{t}\frac{d}{du}_{|_{u=s_0/t}} f \circ D_t\big(\gamma(s_0/t)^{-1}\gamma(u)\big) \\
&= \frac{1}{t}\tilde{\gamma}'(s_0/t)(f \circ D_t) \\
&= \frac{1}{t}\delta_t\big(\tilde{\gamma}'(s_0/t)\big)(f) \ .
\end{aligned}$$

Hence, $\tilde{\gamma}_2'(s_0) \in W_1$ and its norm is not larger than 1. This provides a 1-1 correspondence between the sub-unitary curves joining x and y and those joining $D_t x$ and $D_t y$. The lengths of corresponding curves coincide up to a factor t. □

6. Integration on nilpotent groups

As in Theorem 1.4.2 of Chapter 1, we parametrize the elements of a connected and simply connected nilpotent Lie group G by the elements of its Lie algebra $\mathfrak{g}$ with the product $xy = S(x,y)$.

The first result we are going to prove is the fact that the Lebesgue measure on $\mathfrak{g}$ is invariant under both left and right translations of G. It is a well-known fact that every locally compact group (in particular any Lie group) admits a positive, regular Borel measure that is invariant under left translations, called the *left Haar measure*, which is unique up to scalar multiples; similarly such a group admits a unique, up to scalar multiples, *right Haar measure*. For general groups, right and left Haar measures do not always coincide.

The existence of a Haar measure on G which is both left- and right-invariant is expressed by saying that G is *unimodular*.

Let $\{\mathfrak{v}_j\}_{1\le j\le m+1}$ be any filtration on $\mathfrak{g}$. We construct a basis of $\mathfrak{g}$ as follows.

For each $j \le m$, we fix a subspace W_j such that $\mathfrak{v}_j = \mathfrak{v}_{j+1} \oplus W_j$, so that

$$\mathfrak{g} = W_1 \oplus W_2 \oplus \cdots \oplus W_m \ .$$

We write $x \in \mathfrak{g}$ as $x = x^1 + \cdots + x^m$ with $x^j \in W_j$. Recalling that

$$xy = S(x,y) = x + y + \sum_{\ell,k\ge 1} Z_{\ell,k}(x,y) \ , \tag{2.6.1}$$

and that each $Z_{\ell,k}(x,y)$ is a linear combination of iterated Lie brackets containing ℓ times x and k times y, we see that

$$Z_{\ell,k}(x,y) \in \mathfrak{g}^{\ell+k} \subseteq \mathfrak{v}_{\ell+k} \ ,$$

by (**??**). Therefore, if $(xy)^j$ and $Z^j_{\ell,k}$ denote the W_j-components of the indicated objects, we see that $Z^j_{\ell,k} = 0$ if $j < \ell + k$. In particular,

$$(xy)^1 = x^1 + y^1 \ .$$

The W_2-component is

$$(xy)^2 = x^2 + y^2 + \frac{1}{2}[x,y]^2 = x^2 + y^2 + \frac{1}{2}[x^1,y^1] \ ,$$

because of the bilinearity of the Lie bracket and the fact that $[x^i,y^j] \in \mathfrak{v}_{i+j}$.

For the general component, we must observe that

$$Z_{\ell,k}(x,y) = Z_{\ell,k}(x^1 + \cdots + x^m, y^1 + \cdots + y^m) \ ,$$

being a combination of iterated Lie brackets, decomposes by multilinearity into a sum of terms, each of them being an iterated Lie bracket containing ℓ times components of x and k times components of y. In order that an individual term have a non-trivial component in W_j, it is necessary that the sum of the indices of the various components be at most equal to j. This implies that, for $\ell, k \geq 1$, only the i-th components with $i < j$ can appear. Therefore we can write

$$(xy)^j = x^j + y^j + P_j(x^1,\ldots,x^{j-1},y^1,\ldots,y^{j-1}) \ , \tag{2.6.2}$$

where P_j is a polynomial function[4] from $W_1 \times \cdots \times W_{j-1} \times W_1 \times \cdots \times W_{j-1}$ to W_j.

THEOREM 2.6.1. *The Lebesgue measure dx on $\mathfrak{g}$ satisfies the following property: if f is an integrable function on G and $a \in G$, then*

$$\int_G f(xa)\,dx = \int_G f(ax)\,dx = \int_G f(x)\,dx \ .$$

PROOF. By (2.6.2),

$$\int_G f(xa)\,dx = \int_{\mathfrak{g}} f\big(x^1+a^1, x^2+a^2+P_2(x^1,a^1), x^3+a^3+P_3(x^1,x^2,a^1,a^2),\ldots$$
$$\ldots, x^m + a^m + P_m(x^1,\ldots,x^{m-1},a^1,\ldots,a^{m-1})\big)\,dx^1\cdots dx^m \ .$$

The consecutive changes of variables

$$\begin{aligned} y^m &= x^m + a^m + P_m(x^1,\ldots,x^{m-1},a^1,\ldots,a^{m-1}) \\ y^{m-1} &= x^{m-1} + a^{m-1} + P_m(x^1,\ldots,x^{m-2},a^1,\ldots,a^{m-2}) \\ &\text{etc.} \end{aligned}$$

show that

$$\int_G f(xa)\,dx = \int_G f(x)\,dx \ .$$

The other identity is proved in the same way. □

REMARK. It is convenient to have a more abstract interpretation of this result. The notion of Haar measure on a Lie group is intrinsic and does not depend on the choice of coordinates. From this point of view, Theorem 2.6.1 is not an intrinsic statement, but relative to a particular choice of coordinates on a connected and simply connected nilpotent Lie group G.

To be more precise, let $\mathfrak{g}$ be the Lie algebra of G. The fact that

$$\exp_G(x)\exp_G(y) = \exp_G\big(S(x,y)\big) \ ,$$

means that the product (2.6.1) is the product on G expressed in the canonical coordinates of the first kind. Therefore, the "intrinsic" formulation of Theorem 2.6.1 is that the Haar measure on a connected and simply connected nilpotent Lie group is equal to the Lebesgue measure on its Lie algebra, once expressed in canonical coordinates of the first kind.

It is not difficult to prove that, if we use instead the canonical coordinates of the second kind on G relative to the decomposition of $\mathfrak{g}$ in the subspaces W_j as above, the product on $\mathfrak{g}$ takes a form similar to (2.6.2), only with different polynomials. Therefore the proof of Theorem 2.6.1 applies, and the Haar measure on G is again expressed by the Lebesgue measure on $\mathfrak{g}$.

---

[4]This means a map that is expressed by polynomials with respect to linear coordinates in the domain and in the codomain.

An interesting example connected with this remark is the following. Consider the canonical coordinates of the second kind on the Heisenberg group H_1 induced by the decomposition of $\mathfrak{h}_1$ into $W_1 = \mathbb{R}e_1$, $W_2 = \operatorname{span}\{e_2, e_3\}$ (in the notation of Example (4.b) in Chapter 1). The product on $\mathbb{R}^3$ then becomes

$$(x,y,z)(x',y',z') = (x+x', y+y', z+z'+xy') \ . \tag{2.6.3}$$

If we represent (x,y,z) by the matrix

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

then (2.6.3) is the ordinary matrix product. This given an alternative (and frequent) description of H_1. It is straightforward to verify that the Haar measure is represented by the Lebesgue measure also in these coordinates.

Assume now that G is homogeneous, with dilations $\{D_t\}$, and let $t^{\lambda_1}, \dots, t^{\lambda_m}$ be the eigenvalues of $\delta_t = (D_t)_*$, with eigenspaces $W_1, \dots, W_m$. For every Lebesgue-measurable set $E \subset \mathfrak{g}$,

$$m\big(\delta_t(E)\big) = t^Q m(E) \ , \tag{2.6.4}$$

with

$$Q = \sum_{j=1}^m \lambda_j \mathrm{dim} W_j \ .$$

The number Q is called the *homogeneous dimension* of G with respect to the given dilations. It follows from (2.6.4) that for any integrable function f on G,

$$\int_G f\big(D_t(x)\big)\,dx = t^{-Q} \int_G f(x)\,dx \ . \tag{2.6.5}$$

Proposition 2.6.2. *Let $|\ |$ be a homogeneous norm on G. The integral*

$$\int_{|x|<1} |x|^{-\alpha}\,dx$$

is convergent if and only if $\alpha < Q$, and the integral

$$\int_{|x|>1} |x|^{-\alpha}\,dx$$

is convergent if and only if $\alpha > Q$.

Proof. Let $c = m\big(\{x : 1 < |x| < 2\}\big)$. Then $c > 0$ and

$$\int_{2^j<|x|<2^{j+1}} |x|^{-\alpha}\,dx \sim 2^{-j\alpha} m\big(\{x : 2^j < |x| < 2^{j+1}\}\big) = c2^{j(Q-\alpha)} \ ,$$

where $\sim$ denotes that the two sides bound each other up to a multiplicative constant.

The conclusion follows by summing over $j < 0$ in the first case and over $j \geq 0$ in the second case. $\square$

CHAPTER 3

Homogeneous hypoelliptic operators on homogeneous groups

1. General properties of hypoelliptic operators

Let K be a compact subset of $\mathbb{R}^n$. We denote by $\mathcal{D}_k(K)$ the space of C^k functions on $\mathbb{R}^n$ supported on K, with the C^knorm, and by

$$\mathcal{D}(K) = \bigcap_{k\in\mathbb{N}} \mathcal{D}_k(K)$$

the space of C^∞ functions on $\mathbb{R}^n$ supported on K. It is well-known that $\mathcal{D}(K)$ is a Fréchet space under the family of C^k norms, for $k \in \mathbb{N}$.

Lemma 3.1.1. *For every C^∞ function f supported on a ball of radius r and every $k \in \mathbb{N}$,*

$$\|f\|_k \leq 2r\|f\|_{k+1} \ .$$

Proof. We can assume that the support of f is contained in the cube $Q = [-r, r]^n$. For any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n)$ of length $|\alpha| \leq k$, let $\alpha' = (\alpha_1 + 1, \ldots, \alpha_n)$. If $x \in Q$,

$$\begin{aligned}\left|\partial^\alpha f(x)\right| &= \left|\int_{-r}^{x_1} \partial^{\alpha'} f(t, x_2, \ldots, x_n)\, dt\right| \\ &\leq 2r\|\partial^{\alpha'} f\|_\infty \ ,\end{aligned}$$

so that $\|f\|_k \leq 2r\|f\|_{k+1}$. □

Let now L be a linear differential operator with smooth coefficients on an open subset Ω of $\mathbb{R}^n$. Denote by ${}^tL$, the *transpose* of L, the operator such that

$$\int_\Omega Lf(x)g(x)\, dx = \int_\Omega f(x)\, {}^tLg(x)\, dx \ , \tag{3.1.1}$$

for every pair of test functions f, g. If

$$Lf(x) = \sum_{|\alpha|\leq m} a_\alpha(x)\partial^\alpha f(x) \ ,$$

then

$${}^tLf(x) = \sum_{|\alpha|\leq m} \partial^\alpha(a_\alpha f)(x) = \sum_{|\alpha|\leq m} b_\alpha(x)\partial^\alpha f(x) \ ,$$

where the b_α are linear combinations of derivatives of the a_α, in particular smooth functions.

For K a compact subset of Ω and $\nu \in \mathbb{N}$, define

$$V^\nu(K) = \left\{f \in \mathcal{D}_\nu(K) : Lf \in \mathcal{D}(K)\right\} \ .$$

Observe that if ν is smaller than the order m of L, Lf is defined as a distribution:

$$\langle Lf, \varphi\rangle = \int_K f(x)\, {}^tL\varphi(x)\, dx \ ,$$

for $\varphi \in \mathcal{D}(\Omega)$.

Consider the family of norms

$$\|f\|_{V^\nu,k} = \|f\|_\nu + \|Lf\|_k \tag{3.1.2}$$

on $V^\nu(K)$.

LEMMA 3.1.2. *$V^\nu(K)$ is a Fréchet space. If L is hypoelliptic, $V^\nu(K) = \mathcal{D}(K)$.*

PROOF. Let $\{f_j\}$ be a Cauchy sequence with respect to each one of the norms (3.1.2). In particular, the f_j converge, in the C^ν-norm, to a function $f \in \mathcal{D}_\nu(K)$, and, by the completeness of $\mathcal{D}(K)$, the Lf_j converge to $g \in \mathcal{D}(K)$ in every C^k-norm.

In order to prove the first statement, we must show that $g = Lf$ as a distribution on Ω. If $\varphi \in \mathcal{D}(\Omega)$, then

$$\begin{aligned}\langle Lf, \varphi\rangle &= \langle f, {}^tL\varphi\rangle \\ &= \int_K f(x)\, {}^tL\varphi(x)\, dx \\ &= \lim_{j\to\infty} \int_K f_j(x)\, {}^tL\varphi(x)\, dx \\ &= \lim_{j\to\infty} \int_K Lf_j(x)\varphi(x)\, dx \\ &= \langle g, \varphi\rangle \ .\end{aligned}$$

If L is hypoelliptic and $f \in V^\nu(K)$, then f itself is C^∞ on Ω. But it is supported on K, so that $f \in \mathcal{D}(K)$. The inclusion $\mathcal{D}(K) \subseteq V^\nu(K)$ is obvious. □

THEOREM 3.1.3. *Let L be hypoelliptic in Ω. Given $x \in \Omega$ and $k \in \mathbb{N}$, there are a compact neighborhood U_k of x in Ω and $k' \in \mathbb{N}$ such that, for every $f \in \mathcal{D}(U_k)$,*

$$\|f\|_k \le C_k \|Lf\|_{k'} \ .$$

PROOF. We can assume that $k \ge 1$. If K is a compact neighborhood of x, the identity map

$$i : \mathcal{D}(K) \longrightarrow V^{k-1}(K)$$

is obviously continuous and, by Lemma 3.1.2, is onto. By the open mapping theorem[1] i is a homeomorphism.

It follows that the inclusion

$$j : V^{k-1}(K) \longrightarrow \mathcal{D}_k(K)$$

is continuous. Hence there are $k' \in \mathbb{N}$ and $C > 0$ such that

$$\|f\|_k \le C_k\big(\|f\|_{k-1} + \|Lf\|_{k'}\big)$$

for every $f \in \mathcal{D}(K)$. Let $U_k \subset K$ be a ball centered at x with radius $r < 1/4C_k$. By Lemma 3.1.1, if $f \in \mathcal{D}(U_k)$,

$$C_k\|f\|_{k-1} < \frac{1}{2}\|f\|_k \ ,$$

so that

$$\|f\|_k \le 2C_k\|Lf\|_{k'} \ . \qquad \square$$

COROLLARY 3.1.4. *If L, x and U_k are as in Theorem 3.1.3, then L is injective on $\mathcal{D}(U_k)$.*

As a direct consequence of Theorem 3.1.3, we prove a result concerning local solvability of ${}^tL$.

DEFINITION. *Let L be a linear differential operator with smooth coefficients on Ω. We say that L is* locally solvable *at $x \in \Omega$ if, for every k, x has an open neighborhood V_k such that, for every distribution $\psi \in \mathcal{D}'_k(\Omega)$, there is a distribution $u \in \mathcal{D}'(V)$ such that $Lu = \psi$ on V_k.*

[1] For the validity of the open mapping theorem for continuous linear maps between Fréchet spaces, see Trèves, *Topological Vector Spaces, Distributions and Kernels*.

THEOREM 3.1.5. *Let L be hypoelliptic in Ω. Then ${}^tL$ is locally solvable at every point of Ω.*

PROOF. Given $x \in \Omega$, let U_0 be a compact neighborhood of x. Given $\psi \in \mathcal{D}'(\Omega)$, there are $k \in \mathbb{N}$ and $C > 0$ such that

$$\big|\langle \psi, f\rangle\big| \le C\|f\|_k$$

for every $f \in \mathcal{D}(U_0)$. Let $U = U_k$ be as in Theorem 1.3. We can assume that $U \subset U_0$.

Let

$$X = \big\{Lg : g \in \mathcal{D}(U)\big\} \subseteq \mathcal{D}(U) \ .$$

Define the linear functional $\lambda : X \to \mathbb{C}$ as

$$\lambda(Lg) = \langle \psi, g\rangle \ .$$

This is a good definition by Corollary 3.1.4. By Theorem 3.1.3,

$$\big|\lambda(Lg)\big| \le C\|g\|_k \le C'\|Lg\|_{k'} \ .$$

By the Hahn-Banach theorem, λ extends to a continuous linear functional $\tilde\lambda$ on $\mathcal{D}_{k'}(U)$. If $V \mathrel{\overset{\circ}{=\!\!\!\rightarrow}} U$ and $g \in \mathcal{D}(V)$, there is a distribution u on V, of order k', such that $\tilde\lambda(f) = \langle u, f\rangle$ for $f \in \mathcal{D}(V)$; in particular

$$\langle u, Lg\rangle = \lambda(Lg) = \langle \psi, g\rangle \ ,$$

for every $g \in \mathcal{D}(V)$. This means that ${}^tLu = \psi$ on V. □

2. Convolution on nilpotent groups

Let G be a connected and simply connected nilpotent Lie group.

We may regard G as its Lie algebra $\mathfrak{g}$ with the product $xy = S(x, y)$. Hence the spaces $\mathcal{D}(G)$ and $\mathcal{S}(G)$ of test functions on G are well defined, as well as their dual spaces $\mathcal{D}'(G)$ and $\mathcal{S}'(G)$ of distributions and tempered distributions respectively.

We fix a Lebesgue measure dx on $\mathfrak{g}$, which also serves as a Haar measure on G. The spaces $L^p(G)$ are referred to this measure.

The *convolution* of two functions f, g on G is given by

$$f * g(x) = \int_G f(xy^{-1})g(y)\, dy = \int_G f(y)g(y^{-1}x)\, dy \ . \tag{3.2.1}$$

In general, if G is non-commutative, $f * g \ne g * f$.

We state without proof some basic facts[2].

THEOREM 3.2.1. *The integral* (3.2.1) *is absolutely convergent for almost every x if $f \in L^p(G)$, $g \in L^q(G)$, and $1/p + 1/q \ge 1$. If these conditions are satisfied, and*

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1 \ , \tag{3.2.2}$$

then $f * g \in L^r(G)$ (and continuous if $r = \infty$) and the Young inequality *holds:*

$$\|f * g\|_r \le \|f\|_p\|g\|_q \ . \tag{3.2.3}$$

If $p, q \in (1, \infty)$ are such that $1/p + 1/q > 1$, $f \in L^p(G)$, and g satisfies the weak-L^q condition

$$\sup_{s>0} s^q m\Big(\big\{x : |g(x)| > s\big\}\Big) = \|g\|_{\mathrm{w}-q}^q < \infty \ , \tag{3.2.4}$$

then $f * g \in L^r(G)$ with r as in (3.2.2), *and the* generalized Young inequalities *hold:*

$$\|f * g\|_r \le \|f\|_p\|g\|_{\mathrm{w}-q} \ , \qquad \|g * f\|_r \le \|f\|_p\|g\|_{\mathrm{w}-q} \ . \tag{3.2.5}$$

Concerning smoothness properties of convolutions, we give the following statement.

[2]The proofs are almost the same as in $\mathbb{R}^n$. One can find them in any good textbook on distribution theory.

Proposition 3.2.2. *If $f,g\in\mathcal{D}(G)$, then $f*g\in\mathcal{D}(G)$, and the map $(f,g)\longmapsto f*g$ is continuous from $\mathcal{D}(G)\times\mathcal{D}(G)$ to $\mathcal{D}(G)$. The same statement remains true if we replace $\mathcal{D}(G)$ by $\mathcal{S}(G)$.*

If $\Phi\in\mathcal{D}'(G)$ and $f\in\mathcal{D}(G)$, one defines[3]

$$\begin{aligned}\Phi*f(x)&=\langle\Phi,L_x\check f\rangle\ ,\\ f*\Phi(x)&=\langle\Phi,R_{x^{-1}}\check f\rangle\ .\end{aligned}\tag{3.2.6}$$

The same definitions make sense for $\Phi\in\mathcal{S}'(G)$ and $f\in\mathcal{S}(G)$.

Theorem 3.2.3. *If $\Phi\in\mathcal{D}'(G)$ and $f\in\mathcal{D}(G)$, then $\Phi*f$ and $f*\Phi$ are C^∞ functions on G. If $\Phi\in\mathcal{S}'(G)$ and $f\in\mathcal{S}(G)$, they are C^∞ functions with polynomial growth, together with all their derivatives.*

The rule

$$\operatorname{supp}(f*g)\subseteq(\operatorname{supp}f)(\operatorname{supp}g)$$

is respected also by convolution between a function and a distribution.

Relevant identities are:

(i) $\delta_a*f=L_af$, $f*\delta_a=R_{a^{-1}}f$ for every $a\in G$; in fact,

$$\delta_a*f(x)=\langle\delta_a,L_x\check f\rangle=L_x\check f(a)=f(a^{-1}x)=L_af(x)\ ,$$

and similarly for the other identity.

(ii) $f*(\partial_u\delta_0)=X_uf$, where $u\in\mathfrak{g}(=T_0G)$ and X_u is the corresponding left-invariant vector field. In fact,

$$\begin{aligned}f*(\partial_u\delta_0)(x)&=\langle\partial_u\delta_0,R_{x^{-1}}\check f\rangle=-\partial_u(R_{x^{-1}}\check f)(0)\\&=-\partial_u(L_{x^{-1}}f)\check{}(0)=\partial_u(L_{x^{-1}}f)(0)\\&=X_uf(x)\ .\end{aligned}$$

The convolution of two distributions is not always defined. However, if $\Phi,\Psi\in\mathcal{D}'(G)$ and one of them, say Φ, has compact support, then one can set, for $f\in\mathcal{D}(G)$,

$$\begin{aligned}\langle\Phi*\Psi,f\rangle&=\langle\Psi,\check\Phi*f\rangle\ ,\\ \langle\Psi*\Phi,f\rangle&=\langle\Psi,f*\check\Phi\rangle\ ,\end{aligned}\tag{3.2.7}$$

where $\check\Phi$ is the distribution such that

$$\langle\check\Phi,f\rangle=\langle\Phi,\check f\rangle\ .$$

Then $\Phi*\Psi$ and $\Psi*\Phi$ are in $\mathcal{D}'(G)$. If $\Psi\in\mathcal{S}'(G)$, then they are also in $\mathcal{S}'(G)$. If any two of the distributions Φ,Ψ,Λ have compact support, then the associative property holds:

$$(\Phi*\Psi)*\Lambda=\Phi*(\Psi*\Lambda)\ .$$

For $\Phi\in\mathcal{D}'(G)$, define $L_a\Phi$, $R_a\Phi$ by

$$\langle L_a\Phi,f\rangle=\langle\Phi,L_{a^{-1}}f\rangle\ ,\qquad\langle R_a\Phi,f\rangle=\langle\Phi,R_{a^{-1}}f\rangle\ .$$

This definition is motivated by the fact that, by the invariance properties of the Lebesgue measure, for any pair of test functions,

$$\langle L_af,g\rangle=\int_Gf(a^{-1}x)g(x)\,dx=\int_Gf(x)g(ax)\,dx=\langle f,L_{a^{-1}}g\rangle\ ,$$

and similarly for right translations.

Proposition 3.2.4. *$L_a\Phi=\delta_a*\Phi$, $R_a\Phi=\Phi*\delta_{a^{-1}}$.*

If X is a left-invariant vector field and $u=X_0$, then $X\Phi=\Phi*(\partial_u\delta_0)$ and the following identities hold:

$$\langle X\Phi,f\rangle=-\langle\Phi,Xf\rangle\ ,\tag{3.2.8}$$

$$X(\Phi*\Psi)=\Phi*(X\Psi)\ ,\tag{3.2.9}$$

whenever the convolution is defined.

[3] As in Proposition 1.3.1 in Chapter 1, we set $\check f(x)=f(x^{-1})$.

PROOF. The first statement follows immediately from (i) above.

By invariance of the Lebesgue measure, if $f, g \in \mathcal{D}(G)$, the integral

$$\int_G f(xa)g(xa)\, dx$$

does not depend on a. Taking $a = \exp_G(tu)$ and differentiating at $t = 0$, we obtain that

$$\int_G \big(Xf(x)g(x) + f(x)Xg(x)\big)\, dx = 0 \ ,$$

so that

$$\langle Xf, g\rangle = -\langle f, Xg\rangle \ ,$$

i.e.

$$ {}^tX = -X \ . \tag{3.2.10}$$

By definition,

$$\langle X\Phi, f\rangle = \langle \Phi, {}^tXf\rangle = -\langle \Phi, Xf\rangle \ .$$

Hence, by (ii) and (3.2.7),

$$\begin{aligned} \langle X\Phi, f\rangle &= -\langle \Phi, f * (\partial_u\delta_0)\rangle \\ &= -\langle \Phi * (\partial_u\delta_0)\check{}, f\rangle \\ &= \langle \Phi * (\partial_u\delta_0), f\rangle \ . \end{aligned}$$

This also proves (3.2.8), and (3.2.9) follows easily. □

Consider a convolution operator of the form

$$Tf = f * K \ ,$$

where the kernel K is a tempered distribution and $f \in \mathcal{S}(G)$. This satisfies the identity

$$T(L_af) = L_a(Tf)$$

for every $a \in G$, i.e. it commutes with left translations, or, otherwise said, it is a *left-invariant operator*. In fact,

$$T(L_af) = (\delta_a * f) * K = \delta_a * (f * K) = L_a(Tf) \ .$$

The following theorem, that we shall not prove, says that the converse is also true, under very mild hypotheses.

THEOREM 3.2.5. *Let $T : \mathcal{D}(G) \to \mathcal{D}'(G)$ (resp. $T : \mathcal{S}(G) \to \mathcal{S}'(G)$) be a left-invariant operator. Then there is $K \in \mathcal{D}'(G)$ (resp. $K \in \mathcal{S}'(G)$) such that $Tf = f * K$.*

3. Homogeneous distributions on homogeneous groups

Let G be a homogeneous group with dilations $\{D_t\}_{t>0}$, homogeneous dimension Q, and Haar measure dx.

If Φ is a distribution on G, we define $\Phi \circ D_t$ as the distribution such that for every test function f

$$\langle \Phi \circ D_t, f\rangle = \langle \Phi, t^{-Q} f \circ D_{t^{-1}}\rangle \ . \tag{3.3.1}$$

This definition is justified by the fact that, if Φ is a locally integrable function φ, i.e.

$$\langle \Phi, f\rangle = \int_G \varphi(x)f(x)\, dx \ ,$$

then

$$\int_G \varphi(D_tx)f(x)\, dx = t^{-Q}\int_G \varphi(x)f(D_{t^{-1}}x)\, dx \ .$$

DEFINITION. *A distribution* Φ *on* G *is* homogeneous of degree $\alpha \in \mathbb{C}$ *if*

$$\Phi \circ D_t = t^\alpha \Phi$$

for every $t > 0$.

EXAMPLES 6.

3.3.a If $|\ |$ is a homogeneous norm and $\Re\alpha > -Q$, then $\Phi(x) = |x|^\alpha$ is a locally integrable distribution, homogeneous of degree α.

3.3.b Let $t^{\lambda_1}, \ldots, t^{\lambda_m}$ be the eigenvalues of the dilations $\delta_t = (D_t)_*$ on $\mathfrak{g}$, and denote by W_{λ_j} be the corresponding eigenspaces. We fix a basis $\{e_1, \ldots, e_n\}$ of $\mathfrak{g}$ such that each e_i belongs to one of the W_{λ_j}.

Let $S_i(x,y)$ be the components of $xy = S(x,y)$ with respect to this basis. If $e_i \in W_{\lambda_j}$, then S_i, as a function on $\mathfrak{g} \times \mathfrak{g}$, is homogeneous of degree λ_j. Since S_i is a polynomial, this means that each monomial appearing in it must have degree λ_j.

3.3.c The Dirac delta δ_0 at the unit element is homogeneous of degree $-Q$.

3.3.d Let $u \in W_{\lambda_j}$ (i.e. $\delta_t(u) = t^{\lambda_j} u$ for every $t > 0$), and let X_u be the corresponding left-invariant vector field. If a distribution Φ is homogeneous of degree α, then $X_u\Phi$ is homogeneous of degree $\alpha - \lambda_j$. In fact, using (2.2.4) in Chapter 2,

$$\begin{aligned}\langle (X_u\Phi) \circ D_t, f\rangle &= t^{-Q}\langle X_u\Phi, f \circ D_t\rangle = -t^{-Q}\langle \Phi, X_u(f \circ D_{t^{-1}})\rangle \\ &= -t^{-Q-\lambda_j}\langle \Phi, (X_u f) \circ D_{t^{-1}}\rangle = -t^{-\lambda_j}\langle \Phi \circ D_t, X_u f\rangle \\ &= -t^{\alpha-\lambda_j}\langle \Phi, X_u f\rangle \\ &= t^{\alpha-\lambda_j}\langle X_u\Phi, f\rangle \ .\end{aligned}$$

THEOREM 3.3.1. *Let* $Tf = f * K$ *be a left-invariant operator, with* $f \in \mathcal{S}(G)$. *Then*

$$T(f \circ D_t) = t^{-\alpha}(Tf) \circ D_t \tag{3.3.2}$$

if and only if K *is homogeneous of degree* $-Q + \alpha$.

PROOF. Assume that K is homogeneous of degree $-Q + \alpha$. Then

$$T(f \circ D_t)(x) = \langle K, R_{x^{-1}}(f \circ D_t)\check{}\rangle \ .$$

But $R_{x^{-1}}(f \circ D_t)\check{}(y) = R_{(D_t x)^{-1}}\check{f}(D_t y)$, so that

$$\begin{aligned}T(f \circ D_t)(x) &= \langle K, (R_{(D_t x)^{-1}}\check{f}) \circ D_t\rangle \\ &= t^{-Q}\langle K \circ D_{t^{-1}}, R_{(D_t x)^{-1}}\check{f}\rangle \\ &= t^{-\alpha}\langle K, R_{(D_t x)^{-1}}\check{f}\rangle \\ &= t^{-\alpha}(Tf)(D_t x) \ .\end{aligned}$$

Conversely, if T satisfies (3.3.2), then

$$\langle K, f \circ D_t\rangle = T(\check{f} \circ D_t)(0) = t^{-\alpha}(T\check{f})(0) = t^{-\alpha}\langle K, f\rangle \ ,$$

which means that K is homogeneous of degree $-Q + \alpha$. $\square$

4. Left-invariant differential operators

Let G be a connected and simply connected nilpotent Lie group. For simplicity, we assume that G is its own Lie algebra $\mathfrak{g}$ with the product $xy = S(x,y)$.

If $X_1, \ldots, X_k$ are left-invariant vector fields on G, then $L = X_1X_2 \cdots X_k$ is a left-invariant differential operator on G, and such is any linear combination of compositions of this kind. We shall prove the converse statement.

We fix a basis $(e_1, \dots, e_n)$ of $\mathfrak{g}$, and denote by X_j be the left-invariant vector field such that $(X_j)_0 = \partial_{e_j}$. If $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, we set

$$X^\alpha = X_1^{\alpha_1} X_2^{\alpha_2} \cdots X_n^{\alpha_n} .$$

THEOREM 3.4.1. (Poincaré-Birkhoff-Witt)[4] *Let L be a left-invariant differential operator on G. Then L can be written in one and only one way as*

$$L = \sum_{|\alpha| \le m} c_\alpha X^\alpha . \tag{3.4.1}$$

PROOF. Let $(x_1, \dots, x_n)$ be the coordinates on $\mathfrak{g}$ induced by the fixed basis. Then

$$Lf(0) = \sum_{|\alpha| \le m} a_\alpha \partial^\alpha f(0) .$$

The proof goes by induction on m. If $m = 0$, then $Lf(0) = af(0)$ and

$$Lf(x) = L_{x^{-1}}(Lf)(0) = L(L_{x^{-1}} f)(0) = aL_{x^{-1}} f(0) = af(x) .$$

If $m = 1$,

$$Lf(0) = \sum_{j=1}^n a_j \partial_{x_j} f(0) + a_0 f(0) .$$

If $L' = L - \sum_{j=1}^n a_j X_j$, then $L'f(0) = a_0 f(0)$. Since L' is also left-invariant, it follows that

$$L = \sum_{j=1}^n a_j X_j + a_0 .$$

Assume that the statement is true for $m - 1$. We set

$$L' = L - \sum_{|\alpha|=m} a_\alpha X^\alpha .$$

Observe that for each j,

$$X_j = \partial_{x_j} + \sum_{k=1}^n b_{j,k}(x) \partial_{x_k} ,$$

where every $b_{j,k}$ vanishes at the origin. Therefore

$$X^\alpha f(x) = \partial^\alpha f(x) + \cdots$$

where the other terms either vanish at 0 or are lower-order terms. Hence $X^\alpha f(0) = \partial^\alpha f(0) +$ lower-order derivatives of f at 0.

It follows that $L'f(0)$ is a combination of derivaties of f at 0 of order not exceeding $m - 1$, and we can use the inductive assumption.

The proof shows that the coefficients c_α in (3.4.1) coincide with a_α if $|\alpha| = m$. By induction, the representation (3.4.1) is unique. □

The uniqueness part of Theorem (3.4.1) depends heavily on the fact that the vector fields X_j have been *ordered* and that this ordering is respected when composing the monomials X^α. If this restriction is removed, then the same operator L can have more than one representation as in (3.4.1). If, for instance, $X_3 = [X_1, X_2]$ and $L = X_2 X_1$, its correct expression according to (3.4.1) is

$$L = X_1 X_2 - X_3 .$$

[4]Theorem (3.4.1) is valid on any Lie group, and the proof is essentially the same. The only reason for restricting to nilpotent groups is that we can take advantage of a global coordinate system.

5. Fundamental solutions

Let G be a connected and simply connected Lie group.

DEFINITION. *Let L be a left-invariant differential operator on G. A distribution K on G is called a* (global) fundamental solution *of L if $LK = \delta_e$. A distribution K_0 on a neighborhood V of 0 is called a* local fundamental solution *of L if $LK_0 = \delta_0$ on V.*

LEMMA 3.5.1. *If L has a global fundamental solution K, then for every $\psi \in \mathcal{D}'(G)$ with compact support the convolution*

$$u = \psi * K$$

satisfies $Lu = \psi$ on all of G.

L has a local fundamental solution if and only if it is locally solvable at any point of G.

PROOF. As to the first statement, it is sufficient to observe that, by (3.2.9),

$$L(\psi * K) = \psi * (LK) = \psi \ .$$

If L is locally solvable at 0, by definition, there are a neighborhood V of 0 and $K_0 \in \mathcal{D}'(V)$ such that $LK_0 = \delta_0$ on V.

Assume now that L has a local fundamental solution $K_0 \in \mathcal{D}(V)$.

Take $\eta \in \mathcal{D}(V)$ such that $\eta(x) = 1$ on a neighborhood V' of 0, and set $K' = \eta K_0$. If $\varphi \in \mathcal{D}(V')$,

$$\langle LK', \varphi\rangle = \langle K_0, \eta\, {}^tL\varphi\rangle = \langle K_0, {}^tL\varphi\rangle \ ,$$

because η is identically 1 on the support of $L\varphi$. Therefore

$$\langle LK', \varphi\rangle = \langle LK_0, \varphi\rangle = \varphi(0) \ ,$$

so that $LK' = \delta_0$ on V'.

Observe that K' can be regarded as a distribution on G with compact support, because, for $\varphi \in \mathcal{D}(G)$ we can set

$$\langle K', \varphi\rangle \overset{\text{def}}{\rightarrow}= \langle K_0, \eta\varphi\rangle \ ,$$

which makes sense since $\eta\varphi \in \mathcal{D}(V)$. Then LK' is defined on all of G, and

$$LK' = \delta_0 + \Phi \ , \tag{3.5.1}$$

where Φ is a distribution whose support does not intersect V'.

Let now W be a neighborhood of 0 such that $W^{-1}W \subset V'$, and let $\tilde\eta \in \mathcal{D}(W)$ be equal to 1 on a neighborhood W' of 0.

If now $\psi \in \mathcal{D}'(G)$, the convolution $u = (\tilde\eta\psi) * K'$ is well defined by (3.2.7). By (3.2.9) and (3.5.1),

$$Lu = (\tilde\eta\psi) * LK' = \tilde\eta\psi + (\tilde\eta\psi) * \Phi \ .$$

The last term is supported on $W(G\setminus V')$, and this set does not intersect W by the assumption $W^{-1}W \subset V'$. Therefore $Lu = \tilde\eta\psi$ on W, hence $Lu = \psi$ on W'. □

PROPOSITION 3.5.2. *Let L be a left-invariant hypoelliptic operator on G. Then ${}^tL$ is also left-invariant and it has a local fundamental solution.*

PROOF. The first statement follows easily from the definition of ${}^tL$, and the second from Theorem 3.1.5. □

6. Global fundamental solutions of homogeneous hypoelliptic operators

Assume now that G is a homogeneous group. We say that a left-invariant differential operator L is *homogeneous of order* μ if it satisfies (3.3.2) with $\alpha = \mu$, i.e. if

$$L(f \circ D_t) = t^\mu (Lf) \circ D_t \ . \tag{3.6.1}$$

It follows from (2.2.4) that if X is a left-invariant vector field with X_e in some eigenspace W_λ of the dilations, then X is homogeneous of order λ.

Let $\{X_1, \dots, X_n\}$ be a basis of $\mathfrak{g}$ consisting of homogeneous vector fields, with orders $\lambda_1, \dots, \lambda_n$, not necessarily all different.

The operator

$$X^\alpha = X_1^{\alpha_1} X_2^{\alpha_2} \cdots X_n^{\alpha_n}$$

is homogeneous of order

$$d(\alpha) = \sum_{j=1}^n \alpha_j \lambda_j \ . \tag{3.6.2}$$

The following statement is a direct consequence of the Poincaré-Birkhoff-Witt Theorem.

Proposition 3.6.1. *A left-invariant differential operator L is homogeneous of order μ if and only if*

$$L = \sum_{d(\alpha)=\mu} c_\alpha X^\alpha \ .$$

Example 1.

3.6.a On the Heisenberg group H_1, let X, Y, T be the vector fields in (1.4.5). The operators X^2+Y^2, $XY+T$ and T are all homogeneous of order 2 with respect to the dilations $D_s(x, y, t) = (sx, sy, s^2t)$. The last two ones are homogeneous of order 3 with respect to the dilations $D_s(x, y, t) = (sx, s^2y, s^3t)$. The operator $X^2 + Y^2 + T^2$ cannot be homogeneous under any set of (automorphic) dilations.

Theorem 3.6.2. *Let L be a left-invariant differential operator on G, homogeneous of order $\mu < Q$, where Q is the homogeneous dimension of G. Assume that L and ${}^tL$ are hypoelliptic. Then L has a global fundamental solution, which is smooth away from the origin and homogeneous of degree $-Q + \mu$.*

Proof. By Proposition 3.5.2, L has a local fundamental solution H, defined on some neighborhood V of 0. Since L is hypoelliptic and $LH = 0$ on $V \setminus \{0\}$, then H is C^∞ away from 0. As in the proof of Lemma 3.5.1, we take $\eta \in \mathcal{D}(V)$ which is identically 1 in a neighborhood of 0, and set $K_1 = \eta H$. Then K_1 is defined on all of G, is smooth away from 0 and, as in (3.5.1),

$$LK_1 = \delta_0 + \Phi \ ,$$

where the support of Φ is compact and does not contain 0. Because of the smoothness of K_1, $\Phi \in \mathcal{D}(G)$.

If $|\ |$ is a homogeneous norm on G, there are $a, b > 0$ such that

$$\operatorname{supp} \Phi \subset \{x : a < |x| < b\} \ .$$

We set

$$K_t = t^{\mu - Q} K_1 \circ D_{t^{-1}} \ , \qquad \Phi_t = t^{-Q} \Phi \circ D_{t^{-1}} \ .$$

By (3.6.1),

$$LK_t = t^{-Q}(LK_1) \circ D_{t^{-1}} = \delta_0 + \Phi_t \ ,$$

because δ_0 is homogeneous of degree $-Q$. Now

$$\operatorname{supp} \Phi_t = D_t(\operatorname{supp} \Phi) \subset \{x : ta < |x| < tb\} \ ,$$

so that K_t is also a local fundamental solution of L. Observe that

$$\lim_{t \to \infty} \Phi_t = 0 \ ,$$

in the sense of distributions, so that, if we prove that $\lim_{t\to\infty} K_t = K$ exists in the sense of distributions, then K will be a global fundamental solution. Furthermore, we will have

$$s^{Q-\mu} K \circ D_s = \lim_{t\to\infty} s^{Q-\mu} K_t \circ D_s = \lim_{t\to\infty} K_{t/s} = K \ ,$$

showing that K is homogeneous of degree $-Q+\mu$.

For $t > 1$, we want to write

$$K_t = K_1 + \int_1^t \frac{dK_s}{ds}\, ds \ , \tag{3.6.3}$$

but we must first discuss that the integrand is a well defined distribution.

We begin with the derivative K_1' of K_s at $s = 1$. If $\varphi \in \mathcal{D}(G)$,

$$\begin{aligned}\lim_{s\to 1} \frac{1}{s-1} \langle K_s - K_1, \varphi\rangle &= \lim_{s\to 1} \frac{1}{s-1} \langle K_1, s^\mu \varphi \circ D_s - \varphi\rangle \\ &= \left\langle K_1, \frac{d}{ds}_{|_{s=1}} (s^\mu \varphi \circ D_s)\right\rangle \ .\end{aligned}$$

Let $(x_1, \dots, x_n)$ be coordinates on G such that $D_s x = (s^{\lambda_1} x_1, \dots, s^{\lambda_n} x_n)$. Then

$$\frac{d}{ds}_{|_{s=1}} \varphi(D_s x) = \sum_{j=1}^n \lambda_j x_j \partial_{x_j} \varphi(x) = E\varphi(x) \ ,$$

a modified Euler operator. Hence

$$\frac{d}{ds}_{|_{s=1}} \langle K_s, \varphi\rangle = \langle K_1, \mu\varphi + E\varphi\rangle \ ,$$

i.e.

$$K_1' = \mu K_1 + {}^t E K_1 = (\mu - Q) K_1 - E K_1 \ .$$

We next show that $K_1' \in \mathcal{D}(G)$. We have

$$\begin{aligned} LK_1' &= \frac{d}{ds}_{|_{s=1}} (LK_s) \\ &= \frac{d}{ds}_{|_{s=1}} (\delta_0 + \Phi_s) \\ &= -Q\Phi - E\Phi \ ,\end{aligned}$$

which is a smooth function.

Since L is hypoelliptic, this implies that K_1' is smooth on G. Clearly, K_1' has compact support.

At this point, at any $s > 1$,

$$\begin{aligned}\frac{dK_s}{ds} &= s^{-1} \frac{d}{du}_{|_{u=1}} K_{su} \\ &= s^{-1} \frac{d}{du}_{|_{u=1}} s^{\mu-Q} K_u \circ D_{s^{-1}} \\ &= s^{\mu-Q-1} K_1' \circ D_{s^{-1}} \ .\end{aligned}$$

If $\varphi \in \mathcal{D}(G)$,

$$\int_1^t \langle s^{\mu-Q-1} K_1' \circ D_{s^{-1}}, \varphi\rangle\, ds = \int_1^t s^{\mu-Q-1} \int_G K_1'(D_{s^{-1}} x) \varphi(x)\, dx\, ds \ .$$

Since $\mu < Q$,

$$\int_1^\infty s^{\mu-Q-1} \int_G |K_1'(D_{s^{-1}} x)| |\varphi(x)|\, dx\, ds \le C \|\varphi\|_1 \ ,$$

showing that the integral in (3.6.3) has a limit in the sense of distributions for $t \to \infty$.

Finally, the smoothness of K away from 0 follows from the hypoellipticity of L, or otherwise, from the fact that, for $x \neq 0$,

$$K(x) = K_1(x) + \int_1^\infty s^{\mu-Q-1} K_1'(D_{s^{-1}}x)\, ds\ . \qquad \square$$

Obviously, in the hypotheses of Theorem 3.6.2, also ${}^tL$ has a homogeneous global fundamental solution, smooth away from the origin.

Corollary 3.6.3. *The fundamental solution K of L constructed in the proof of Theorem 6.2 is its only homogeneous fundamental solution.*

Proof. Let H be a fundamental solution of L, homogeneous of degree α. For $f \in \mathcal{D}(G)$,

$$\begin{aligned} f \circ D_t &= \big(L(f * H)\big) \circ D_t \\ &= t^{-\mu} L\big((f * H) \circ D_t\big) \\ &= t^{Q-\mu} L\big((f \circ D_t) * t^\alpha H\big) \\ &= t^{Q-\mu+\alpha} f \circ D_t\ , \end{aligned}$$

which implies that $\alpha = -Q + \mu$.

Hence $K - H$ is also homogeneous of degree $-Q + \mu < 0$ and satisfies $L(K - H) = 0$. Since L is hypoelliptic, $K - H$ must be smooth also at the origin. Necessarily $K - H = 0$. $\square$

7. Sub-Laplacians on stratified groups

The considerations developed in the last sections apply in particular to the case where G is a stratified group and

$$L = \sum_{j=1}^k X_j^2\ , \tag{3.7.1}$$

assuming that the X_j form a basis of the generating subspace W_1 of $\mathfrak{g}$. The operator (3.7.1) is called a *sub-Laplacian* on G.

Under the natural dilations on G as a stratified group, L is homogenous of order 2. By virtue of Hörmander's Theorem, L is hypoelliptic. By (3.2.10) ${}^tL = L$, so that the assumptions of Theorem 3.6.2 are satisfied.

If $Q > 2$, then L admits a unique fundamental solution K on G which is homogeneous of degree $-Q+2$.

In general, one does not have explicit formulas for K. A sporadic example of an explicit fundamental solution on a non-abelian group can be given for the Heisenberg group. Before giving the construction, we recall the abelian case, i.e. $G = \mathbb{R}^n$.

The general sub-Laplacian is then an elliptic constant coefficient operator,

$$L = \sum_{j=1}^k \partial_{v_j}^2 = \sum_{i,j} a_{i,j} \partial_{x_i} \partial_{x_j}\ ,$$

where the vectors v_j span $\mathbb{R}^n$ and hence the matrix $(a_{i,j})$ is poditive definite. By a linear change of variables, one reduces to the case where $L = \Delta$ is the ordinary Laplacian. Since $Q = n$, we shall impose that $n \geq 3$.

Lemma 3.7.1. *The homogeneous fundamental solution of Δ in dimension $n \geq 3$ is*

$$K(x) = \frac{c}{|x|^{n-2}}\ , \tag{3.7.2}$$

for some[5] $c < 0$.

[5] The value of the constant can be computed explicitly, but we will not need it.

PROOF. The operator Δ is invariant under rotations, i.e.

$$(\Delta f) \circ \rho = \Delta(f \circ \rho)$$

for any orthogonal transformation ρ of R^n. This implies that if K is a fundamental solution of Δ, then also $K \circ \rho$ is a fundamental solution. By Corollary 3.6.3, the unique homogeneous fundamental solution of Δ must be a radial function, i.e.

$$K(x) = \frac{c}{|x|^{n-2}} \ .$$

Clearly $c \neq 0$. Assume that $c > 0$ and let χ be the characteristic function of the unit ball. If $u = \chi * K$, it is clear that u assumes its maximum at $x = 0$. Since $\Delta u = \chi$, u is smooth for $|x| \neq 1$, and therefore $\partial^2_{x_j} u(0) \leq 0$ for every j. But this is in contradiction with the fact that $\Delta u(0) = \chi(0) = 1$. □

If $n = 2$, all the fundamental solutions of Δ have the form

$$K(x) = c \log |x| + h(x) \ ,$$

where h is any harmonic function on $\mathbb{R}^2$ and c is a positive constant. This example shows that the hypothesis that the order of the operator be smaller than Q in the statement of Theorem 3.6.2 cannot be removed.

Consider then the Heisenberg Lie algebra $\mathfrak{h}_n$ of Example 2.2.b and the corresponding connected and simply connected Lie group H_n, i.e. $\mathbb{R}^{2n+1}$ with product

$$(x, y, t)(x', y', t') = \Big(x + x', y + y', t + t' + \frac{1}{2}(x \cdot y' - x' \cdot y)\Big) \ ,$$

where $x, y, x', y' \in \mathbb{R}^n$ and $\cdot$ denotes the Euclidean inner product.

A basis of left-invariant vector fields is given by

$$X_j = \partial_{x_j} - \frac{y_j}{2}\partial_t \ , \qquad Y_j = \partial_{y_j} + \frac{x_j}{2}\partial_t \ , \qquad T = \partial_t \ ,$$

with $1 \leq j \leq n$. The subspaces $W_1 = \operatorname{span}\{X_j, Y_j : 1 \leq j \leq n\}$ and $W_2 = \mathbb{R}T$ give a stratification of $\mathfrak{h}_n$ and

$$L = \sum_{j=1}^n (X_j^2 + Y_j^2)$$

is a sub-Laplacian. Since the homogeneous dimension of H_n under the dilations associated to the given stratification is $Q = 2n + 2$, there is a fundamental solution of L homogeneous of degree $-2n$.

THEOREM 3.7.2. *Writing $z = x + iy \in \mathbb{C}^n$, the homogeneous fundamental solution of L is*

$$K(z, t) = c \frac{1}{\big(|z|^4 + 16t^2\big)^{n/2}} \ ,$$

for some $c < 0$.

PROOF. The explicit expression of L is

$$L = \sum_{j=1}^n \big(\partial^2_{x_j} + \partial^2_{y_j}\big) + \sum_{j=1}^n \big(x_j \partial_{y_j} - y_j \partial_{x_j}\big)\partial_t + \frac{|x|^2 + |y|^2}{4}\partial_t^2 \ , \tag{3.7.3}$$

where the terms $x_j \partial_{y_j} - y_j \partial_{x_j}$ are the angular derivatives in the plane (x_j, y_j).

We shall exploit another invariance property of L, which is better expressed by introducing complex coordinates $z_j = x_j + iy_j$. The group law becomes

$$(z, t)(z', t') = \Big(z + z', t + t' - \frac{1}{2}\Im\mathrm{m}(z|z')\Big)$$

on $\mathbb{C}^n \times \mathbb{R}$, where $(\ |\ )$ is the Hermitian product

$$(z|z') = z_1 \bar{z}'_1 + \cdots + z_n \bar{z}'_n \ .$$

If ρ is a unitary trasformation of $\mathbb{C}^n$, i.e. a linear transformation preserving the Hermitian product, then the map

$$\psi_\rho(z,t) = \big(\rho(z), t\big)$$

is a group automorphism.

In particular, this is true for

$$\psi_\theta(z,t) = (e^{i\theta_1}z_1, \dots, e^{i\theta_n}z_n, t)$$

where $\theta = (\theta_1, \dots, \theta_n) \in [0, 2\pi)^n$. Then

$$(Lf) \circ \psi_\theta = L(f \circ \psi_\theta) \ ,$$

just because each term in (3.7.3) satisfies the same identity.

Therefore, if K is the homogeneous fundamental solution, $K \circ \psi_\theta = K$ for every θ. This implies that K is radial in each variable z_j, and the angular derivatives of K are zero. Hence, for $(z,t) \neq (0,0)$,

$$LK = \Delta_z K + \frac{|z|^2}{4}\partial_t^2 K \ , \tag{3.7.4}$$

where Δ_z is the Laplacian in the x, y variables.

But then it is clear that

$$(LK) \circ \psi_\rho = L(K \circ \psi_\rho)$$

for every unitary transformation ρ.

Therefore, $K \circ \psi_\rho = K$ for every ρ. This implies that K is radial in z, and we can write

$$K(z,t) = H\Big(\frac{|z|^2}{4}, t\Big) \ . \tag{3.7.5}$$

Then $H(s,t)$ is a smooth function on the half-plane of $\mathbb{R}^2$ where $s > 0$, continuous for $s \geq 0$ and $(s,t) \neq (0,0)$, and homogeneous of degree $-n$ with respect to the isotropic dilations $\delta_u(s,t) = (us, ut)$.

We want to express the condition $LK(z,t) = 0$ for $(z,t) \neq (0,0)$ in terms of H. By (3.7.4), we just have to observe that

$$\begin{aligned}
\Delta_z H\Big(\frac{|z|^2}{4}, t\Big) &= 4\sum_{j=1}^n \partial_{z_j}\partial_{\bar z_j} H\Big(\frac{|z|^2}{4}, t\Big) \\
&= \sum_{j=1}^n \partial_{z_j}\Big(z_j \frac{\partial H}{\partial s}\Big(\frac{|z|^2}{4}, t\Big)\Big) \\
&= \frac{|z|^2}{4}\frac{\partial^2 H}{\partial s^2}\Big(\frac{|z|^2}{4}, t\Big) + n\frac{\partial H}{\partial s}\Big(\frac{|z_j|^2}{4}, t\Big) \ .
\end{aligned}$$

Hence

$$s\partial_s^2 H + n\partial_s H + s\partial_t^2 H = 0 \ . \tag{3.7.6}$$

The first two terms correspond to the radial part of the Laplacian in $n+1$ dimensions. In fact, consider on $\mathbb{R}^{n+1}$ a radial function $f(y) = g\big(|y|\big)$, and set $s = |y|$. Then

$$\Delta f(y) = \partial_s^2 g\big(|y|\big) + \frac{n}{s}\partial_s g\big(|y|\big) \ .$$

This means that, if we define $\Psi(y,t) = H\big(|y|, t\big)$ on $\mathbb{R}^{n+2}$, then $\Delta\Psi = 0$ for $y \neq 0$. Observe that Ψ is continuous on $\mathbb{R}^{n+2} \setminus \{0\}$. By standard argument involving the mean value theorem, this implies that Ψ is smooth on its whole domain and $\Delta\Psi = 0$.

Since Ψ is homogeneous of degree $-n$, its Laplacian as a distribution is a homogeneous distribution of degree $-n-2$ supported at the origin. The only possibility is that $\Delta\Psi = \alpha\delta_0$. The constant α cannot be zero, since otherwise Ψ would be smooth at the origin. By Lemma 3.7.1, there is a constant c such that

$$\Psi(y,t) = \frac{c}{\big(|y|^2 + t^2\big)^{n/2}} \ .$$

Hence $H(s,t) = c(s^2 + t^2)^{-n/2}$, and K is as stated.

As in the proof of Lemma 3.7.1, the fact that $c < 0$ can be proved by contradiction. Suppose $c > 0$, and let χ be the characteristic function of the set B where $|z|^4 + 16t^2 \leq 1$. Then $u = \chi * K$ is continuous on H_n and smooth for $|z|^4 + 16t^2 \neq 1$, since $Lu = \chi$. □

8. The strong maximum principle

In this section we will present the strong maximum principle for an operator

$$L = \sum_{j=1}^{k} X_j^2 \ ,$$

where the X_j are smooth vector fields on a connected open set $\Omega \subseteq \mathbb{R}^n$ forming a generating system[6].

To begin with, we observe that the following weak form of the maximum principle holds.

Lemma 3.8.1. *Let u be a function on Ω attaining a local maximum at a point x_0. Then $Lu(x_0) \leq 0$.*

Proof. If $v_j = X_j(x_0)$, then

$$X_j = \partial_{v_j} + \sum_{k=1}^{n} a_{j,k}(x)\partial_{x_k} \ ,$$

with $a_{j,k}(x_0) = 0$. Therefore

$$Lu(x_0) = \sum_{j=1}^{k} \partial_{v_j}^2 u(x_0) + \text{ first-order terms } = \sum_{j=1}^{k} \partial_{v_j}^2 u(x_0) \ , \tag{3.8.1}$$

and this is non-positive because such are all pure second derivatives at a maximum point. □

The strong form we are interested in is the following.

Theorem 3.8.2. *Assume that $Lu \geq 0$ on Ω, and that u takes its maximum at a point $x_0 \in \Omega$. Then u is constant.*

The proof is based on a lemma, that we do not prove here, and whose statement requires a couple of definitions.

Definition. *Let F be a closed subset of Ω and let $x_0 \in F$. We say that a vector $v \in \mathbb{R}^n$ is* normal to F at x_0 *if there is $r > 0$ such that the closed ball of radius r and center $x_0 + rv$ intersects F only at x_0.*

A vector w is said to be tangent to F at x_0 *it it is orthogonal to all normal vectors to F at x_0.*

Lemma 3.8.3. *Let F be closed in Ω, and let X be a smooth vector field[7]. If $X(x)$ is tangent to F for every $x \in F$, then every integral curve of X intersecting F is entirely contained in F.*

Proof of Theorem 3.8.2. Let F be the set where u attains its maximum. We claim that each of the X_j is tangent to F at any of its points. If this is not true, there are $x_0 \in F$ and v be normal to F at x_0, such that $X_j(x_0) \cdot v \neq 0$. We set $v_j = X_j(x_0)$.

Let $y_0 = x_0 + rv$ and $B = \{x : |x - y_0| \leq r\}$ be such that $F \cap B = \{x_0\}$. The function

$$\varphi(x) = e^{-\alpha|x-y_0|^2} - e^{-\alpha r^2}$$

[6] See, J.-M. Bony, *Principe du maximum, inégalité de Harnack at unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier Gren. 19 (1969), 277-304.

[7] Here it is sufficient that X be Lipschitz (see the reference above).

is strictly positive on $\overset{\circ}{\to} B$ and negative outside. By (3.8.1),

$$
\begin{aligned}
L\varphi(x_0) &= \sum_{j=1}^{k} \partial_{v_j}^2 \varphi(x_0) \\
&= e^{-\alpha r^2} \sum_{j=1}^{k} \Big(4\alpha^2 \big((x_0 - y_0) \cdot v_j\big)^2 - 2\alpha |v_j|^2 \Big) \\
&= e^{-\alpha r^2} \left(4\alpha^2 r^2 \sum_{j=1}^{k} (v \cdot v_j)^2 - 2\alpha \sum_{j=1}^{k} |v_j|^2 \right) .
\end{aligned}
$$

Since the coefficient of α^2 is strictly positive, we can take α large enough so that $L\varphi > 0$ in a closed neighborhood U of x_0.

For $\lambda > 0$, consider the function $\tilde{u} = u + \lambda\varphi$ on U. Then $L\tilde{u} > 0$. By a compactness argument, we can find λ small enough so that $\tilde{u} < u(x_0)$ on the part of the boundary of U intersecting B. Clearly, $\tilde{u} < u \le u(x_0)$ on the remaining part of the boundary of U. Hence $\tilde{u}$ attains a maximum in the interior of U, but this is in contradiction with Lemma 3.8.1.

By Lemma 3.8.3, the integral curves of the vector fields X_j intersecting F are all contained in F. It follows from Proposition 1.5.4 that, if $x \in F$, then F contains a full neighborhood of x. Hence F is both closed and open. □

This fact has a series of consequences for sub-Laplacians on stratified groups.

Corollary 3.8.4. *Let L be a sub-Laplacian on a stratified group G, and let K be its homogeneous fundamental solution. Then*

(i) *If $f \in \mathcal{D}(G)$, then $(Lf) * K = f$;*
(ii) *$K = \check{K}$;*
(iii) *$K(x) < 0$ for every $x \neq 0$.*

Proof. The function $g = (Lf) * K$ tends to zero at infinity. In fact, assume that the support of f, and hence of Lf, is contained on the set where $|x| < r$, w.r. to some homogeneous norm. Then

$$
|g(x)| \le \int_{|y|<r} \big|K(y^{-1}x)\big|\, dy \le C \int_{|y|<r} |y^{-1}x|^{-Q+2}\, dy \ .
$$

Let $c > 0$ be the constant in the triangular inequality for the norm. Then

$$
|x| \le c\big(|y| + |y^{-1}x|\big) \ ,
$$

so that, if $|x| > 2cr$ and $|y| < r$,

$$
|y^{-1}x| \ge \frac{1}{c}|x| - |y| > \frac{1}{2c}|x| \ .
$$

Hence, if $|x| > 2cr$, $|g(x)| < C_r |x|^{-Q+2}$.

Consider the function $h = f - g$. It tends to zero at infinity and

$$
Lh = Lf - (Lf) * (LK) = 0 \ .
$$

By Theorem 3.8.2 it must be zero. This proves (i).

In order to prove (ii), take $f, g \in \mathcal{D}(G)$. Then, using (i),

$$
\begin{aligned}
\langle f, g \rangle &= \langle (Lf) * K, g \rangle \\
&= \langle Lf, g * \check{K} \rangle \\
&= \langle f, L(g * \check{K}) \rangle \ .
\end{aligned}
$$

This shows that $L(g * \check{K}) = g$ for every $g \in \mathcal{D}(G)$, i.e. $L\check{K} = \delta_0$. Hence $\check{K}$ is a homogeneous fundamental solution of L. By Corollary 3.6.3, $\check{K} = K$.

To prove (iii), we first show that $K(x) \leq 0$ for all $x \neq 0$. Suppose $K(x_0) > 0$, and let U be a symmetric neighborhood of the origin such that $K(x) > 0$ for every $x \in Ux_0$. Let $\varphi \geq 0$ be a non-trivial smooth function supported in U. Then

$$\varphi * K(x_0) = \int_U \varphi(y) K(y^{-1}x_0)\, dy > 0 \ .$$

Since $\varphi * K(x)$ tends to zero at infinity, it would attain a maximum at some point. But $L(\varphi * K) = \varphi \geq 0$ on all of G, so it would be constant, in fact identically equal to zero, and this is a contradiction.

If we had now $K(x_0) = 0$ for some $x_0 \neq 0$, it would be a maximum of K on $\Omega = G \setminus \{0\}$. Hence K would be identically zero, which is again a contradiction. $\square$

CHAPTER 4

Sub-elliptic estimates for homogeneous hypoelliptic operators on homogeneous groups

1. L^p-norm inequalities for convolutions with homogeneous distributions

Let G be a nilpotent Lie group, K a distribution on G, and T the operator $Tf = f * K$. Initially, T is defined only on $\mathcal{D}(G)$ and takes values into the space of C^∞ functions on G.

Assume we know that, for every $f \in \mathcal{D}(G)$, $Tf \in L^q(G)$ for some $q \in [1,\infty]$, and that there is $p \in [1,\infty)$ such that

$$\|Tf\|_q \le C\|f\|_p \tag{4.1.1}$$

for some $C > 0$. Then T extends by continuity to a continuous operator from $L^p(G)$ to $L^q(G)$.

The next statement says that only the case $p \le q$ is interesting.

Proposition 4.1.1. *Assume that* (4.1.1) *holds for every $f \in \mathcal{D}(G)$, and that $p > q$. Then $T = 0$, i.e. $K = 0$.*

The proof is based on the following lemma.

Lemma 4.1.2. *Let $f \in L^p(G)$, with $1 \le p < \infty$. Then*

$$\lim_{x\to\infty} \|f - L_x f\|_p = 2^{1/p}\|f\|_p\ .$$

Proof. Assume first that f is continuous with compact support, and let E be its support. For $x \in G$, $L_x f$ is supported on xE. Therefore, if $x \notin EE^{-1}$, f and $L_x f$ have disjoint supports, and

$$\begin{aligned}\|f - L_x f\|_p &= \left(\int_E |f(y)|^p\,dy + \int_{xE} |f(x^{-1}y)|^p\,dy\right)^{1/p}\\ &= 2^{1/p}\|f\|_p\ .\end{aligned}$$

Given now a generic $f \in L^p(G)$ and $\varepsilon > 0$, there is f_ε, continuous with compact support E_ε, such that $\|f - f_\varepsilon\|_p < \varepsilon$. For $x \notin E_\varepsilon E_\varepsilon^{-1}$,

$$\begin{aligned}\big|\|f - L_x f\|_p - 2^{1/p}\|f\|_p\big| &\le \big|\|f - L_x f\|_p - 2^{1/p}\|f_\varepsilon\|_p\big| + 2^{1/p}\big|\|f_\varepsilon\|_p - \|f\|_p\big|\\ &\le \big|\|f - L_x f\|_p - \|f_\varepsilon - L_x f_\varepsilon\|_p\big| + 2^{1/p}\|f_\varepsilon - f\|_p\\ &\le \|f - L_x f - f_\varepsilon + L_x f_\varepsilon\|_p + 2^{1/p}\varepsilon\\ &\le \|f - f_\varepsilon\|_p + \|L_x f - L_x f_\varepsilon\|_p + 2^{1/p}\varepsilon\\ &\le (2 + 2^{1/p})\varepsilon\ . \qquad \square\end{aligned}$$

Proof. Proof of Proposition 4.1.1 Since $T \circ L_x = L_x \circ T$, it follows that

$$\|L_x Tf - Tf\|_q = \|T(L_x f - f)\|_q \le \|T\|\|L_x f - f\|_p$$

For every $f \in \mathcal{D}(G)$ and $x \in G$. In our hypotheses, both p and q are finite, so that, letting x go to infinity on both sides, we obtain that

$$2^{1/q}\|Tf\|_q \leq \|T\| 2^{1/p}\|f\|_p \ .$$

But then $\|T\| \leq 2^{1/p-1/q}\|T\|$. But $2^{1/p-1/q} < 1$, hence $T = 0$. □

Assume now that G is homogeneous, and K is homogeneous of degree $-Q+\alpha$. We know from Theorem 4.1 that

$$T(f \circ D_t) = t^{-\alpha}(Tf) \circ D_t \ , \tag{4.1.2}$$

for $f \in \mathcal{D}(G)$.

Proposition 4.1.3. *Assume that T satisfies* (4.1.1) *and* (4.1.2) *for every $f \in \mathcal{D}(G)$. Then*

$$\frac{1}{p} - \frac{1}{q} = \frac{\Re e \alpha}{Q} \ .$$

Proof. This depends on the fact that

$$\|f \circ D_t\|_p = t^{-\frac{Q}{p}}\|f\|_p \ .$$

If $T = 0$, there is nothing to prove. Assume therefore that $T \neq 0$. By (4.1.2),

$$t^{-\frac{Q}{q}}\|Tf\|_q \leq t^{-\Re e\alpha - \frac{Q}{p}}\|T\|\|f\|_p \ ,$$

hence

$$\|T\| \leq t^{-\Re e\alpha - \frac{Q}{p} + \frac{Q}{q}}\|T\| \ ,$$

for every t. Since $\|T\| > 0$, this forces the exponent of t to be zero. □

We will be concerned only with real values of α.

Combining together Proposition 4.1.1 and Proposition 4.1.3, we see that it makes sense to restrict one's attention to the values of α with $0 \leq \alpha \leq Q$.

The case $\alpha = 0$ is the most delicate and we leave it aside for the moment. We shall discuss instead the case $0 < \alpha < Q$, assuming that K coincides with a continuous function away from the origin[1].

Theorem 4.1.4. *Let K be a distribution homogeneous of degree $-Q+\alpha$, with $0 < \alpha < Q$ and continuous away from the origin. Then K is a locally integrable function, satisfying the inequality*

$$|K(x)| \leq \frac{C}{|x|^{Q-\alpha}} \tag{4.1.3}$$

for some constant $C > 0$.

If $1 < p < q < \infty$, and $(1/p) - (1/q) = \alpha/n$, then T is bounded from $L^p(G)$ to $L^q(G)$.

Proof. Let K_0 be the function detrmined by K on $G \setminus \{0\}$. Then K_0 is obviously homogeneous of degree $-Q+\alpha$, and it follows easily that it satisfies (4.1.3). Hence it is locally integrable on G, and defines a distribution on G.

Consider therefore $K' = K - K_0$. This is a distribution supported at the origin, also homogeneous of degree $-Q+\alpha$. Due to the compact support of K', $\big|\langle K', f\rangle\big|$ is controlled by some C^k norm of f on a fixed small neighborhood of the origin. But, because of its homogeneity,

$$\langle K', f\rangle = t^{\alpha}\langle K', f \circ D_t\rangle \ .$$

Letting t tend to 0, the C^k norms of $f \circ D_t$ remain bounded, so that $\langle K', f\rangle = 0$. This shows that $K = K_0$.

For the second part, in view of Theorem 3.2.1, it is sufficient to prove that K satisfies (4.2.4) there, with $q = \frac{Q}{Q-\alpha}$. But, if $s > 0$ and $|K(x)| > s$, then $|x| < s^{-\frac{1}{Q-\alpha}}$, and the volume of the ball with this radius has volume equal to a constant times $s^{-\frac{Q}{Q-\alpha}}$. □

[1] If $\alpha = Q$, (4.1.1) can only hold for $p = 1$ and $q = \infty$. If K is continuous and homogeneous of degree 0, then it is bounded and (4.1.1) is easily verified.

Theorem 4.1.4 has interesting applications to homogeneous hypoelliptic operators. Hypoellipticity means that regularity of Lf on some open set implies regularity of f on the same set. The definition involves C^∞-regularity, but one can measure different degrees of "regularity" considering different scales of function spaces.

For instance, Sobolev spaces take into account the order of differentiability in L^2, Lipschitz-Hölder spaces describe the continuity properties, and L^p spaces the order of integrability.

We give results concerning Lebesgue spaces. This type of estimates are called *sub-elliptic estimates.*

Corollary 4.1.5. *Suppose that L is a homogeneous operator of homogeneous order $\mu < Q$, with both L and ${}^tL$ hypoelliptic. Let $\{X_1, \dots, X_n\}$ be a basis of $\mathfrak{g}$ consisting of homogeneous vector fields. In the notation of Section 6 of Chapter 2, let*

$$X^\alpha = X_1^{\alpha_1} \cdots X_n^{\alpha_n}$$

be homogeneous of order $d(\alpha)$, with $0 \le d(\alpha) < \mu$.

If $p < \frac{Q}{\mu - d(\alpha)}$, and

$$\frac{1}{p} - \frac{1}{q} = \frac{\mu - d(\alpha)}{Q} \ ,$$

given $f \in \mathcal{D}'(G)$ with compact support,

$$Lf \in L^p(G) \Longrightarrow X^\alpha f \in L^q(G) \ .$$

If $L = \sum_j X_j^2$ is a homogeneous sub-Laplacian on a stratified group, under the same conditions on p and q, corresponding to $\mu = 2$ and $d(\alpha) = 0, 1$, we alos have the estimate

$$\|X^\alpha f\|_r \le C \|Lf\|_p \ ,$$

for some $C > 0$ independent of f.

Proof. Let K be the homogeneous fundamental solution of L. The distribution $X^\alpha K$ is then homogeneous of degree $-Q + \mu - d(\alpha)$ and sommoth away from the origin. By Theorem 1.4, $(Lf) * (X^\alpha K) \in L^q(G)$.

Take $g = f - (Lf) * K$. Since $Lg = Lf - Lf = 0$, g is a C^∞ function. Hence

$$X^\alpha f = X^\alpha g + (Lf) * (X^\alpha K)$$

is locally in L^q. But it has compact support, so it is in $L^q(G)$.

If L is a sub-Laplacian, we can use Corollary 3.8.4 to say that $f = (Lf) * K$. Hence

$$X^\alpha f = (Lf) * (X^\alpha K) \ ,$$

and the conclusion follows again from Theorem 4.1.4. □

The second result is a refinement of the previous one, of a local nature.

Corollary 4.1.6. *In the hypotheses of Corollary 4.1.5, assume that $u \in \mathcal{D}'(G)$ and Lu coincides with an L^p_{loc} function on some open set Ω, with $p < \frac{Q}{\mu - d(\alpha)}$. Then $X^\alpha u \in L^q_{\mathrm{loc}}(\Omega)$, with*

$$\frac{1}{p} - \frac{1}{q} = \frac{\mu - d(\alpha)}{Q} \ .$$

The proof requires a preliminary lemma.

Lemma 4.1.7. *Let Φ be a distribution on G which coincides with a smooth function away from the origin, and let Ψ be a distribution with compact support. Then $\Psi * \Phi$ is smooth away from the support of Ψ.*

Proof. The assumption on Φ is that there is a smooth function $\Phi(x)$ on $G \setminus \{0\}$ such that

$$\langle \Phi, f \rangle = \int \Phi(x) f(x) \, dx$$

whenever $f \in \mathcal{D}(G)$ and $0 \notin \operatorname{supp} f$.

Let now E be a compact set disjoint from $F = \operatorname{supp} \Psi$. Since $0 \notin F^{-1}E$, we can take $\eta \in \mathcal{D}(G)$ equal to 1 on a neighborhood U of $F^{-1}E$ and equal to 0 on a neighborhood of 0.

If $g \in \mathcal{D}(E)$, then $\check{\Psi} * g \in \mathcal{D}(G)$ and its support is contained in $F^{-1}E$. Hence

$$\begin{aligned} \langle \Psi * \Phi, g\rangle &= \langle \Phi, \check{\Psi} * g\rangle \\ &= \int \Phi(x)\check{\Psi} * g(x)\, dx \\ &= \int \Phi(x)\eta(x)\check{\Psi} * g(x)\, dx \\ &= \langle \Psi * (\eta\Phi), g\rangle \ . \end{aligned}$$

Therefore $\Psi * \Phi$ and $\Psi * (\eta\Phi)$ coincide on $\overset{\circ}{\to} E$. Since $\eta\Phi \in \mathcal{D}(G)$, also $\Psi * (\eta\Phi) \in \mathcal{D}(G)$. Hence $\Psi * \Phi$ is C^∞ on $\overset{\circ}{\to} E$. □

Proof of Corollary 4.1.6. Fix $x_0 \in \Omega$, and let $\varphi \in \mathcal{D}(G)$ be supported in Ω and equal to 1 on a neighborhood V of x_0. If we prove that $X^\alpha(\varphi u) \in L^q(W)$ for some sub-neighborhood W of V, then $X^\alpha u \in L^q(W)$. Because $X^\alpha(\varphi u) = L(\varphi u) * (X^\alpha K)$ plus a smooth function, it is sufficient to prove that $L(\varphi u) * (X^\alpha K) \in L^q(V)$.

Applying Leibniz's rule, one sees that $L(\varphi u) = \varphi L u$ plus terms containing derivatives of φ. Therefore, all these other terms are zero on V, so that

$$L(\varphi u) = \varphi L u + \Psi \ ,$$

where Ψ is a distribution with compact support contained in $G \setminus V$.

Then

$$\big(L(\varphi u)\big) * (X^\alpha K) = (\varphi L u) * (X^\alpha K) + \Psi * (X^\alpha K) \ , \tag{4.1.4}$$

where K is the homogeneous fundamental solution of L. Since $\varphi L u \in L^p(G)$, it follows from Theorem 1.4 that $(\varphi L u) * (X^\alpha K) \in L^q(G)$.

By Lemma 4.1.7, $\Psi * (X^\alpha K)$ is smooth on V. Hence, if W is a neighborhood of x_0 with $\bar{W} \subset V$, we conclude that $L(\varphi u) * (X^\alpha K) \in L^q(W)$. □

2. Principal value distributions

In Corollaries 4.1.5 and 4.1.6 we estimated L^r norms of "derivatives" $X^\alpha u$ in terms of L^p norms of Lu, under the assumption that $d(\alpha)$ be smaller than μ, the homogeneous order of L. This was mainly a technical restriction, beacuse we wanted the derivative $X^\alpha K$ of the homogeneous fundamental solution of L to be locally integrable.

It is possible, however, to extend the scope of Corollaries 4.1.5 and 4.1.6 so to include also the limiting case $d(\alpha) = \mu$. In this case the distribution $X^\alpha K$ is homogeneous of degree $-Q$ and must be handled with special attention. It is one of the situations where the notions of "function" and of "distribution" tend to diverge and confusion between the two may lead to wrong or contradictory statements.

Consider a *function* $K(x)$, smooth[2] on $G \setminus \{0\}$ and homogeneous of degree $-Q$. The question we pose concerns the existence of a *distribution* on G, that we temporarily denote by $\tilde{K}$, such that

$$\langle \tilde{K}, f\rangle = \int K(x) f(x)\, dx \ , \tag{4.2.1}$$

whenever $0 \notin \operatorname{supp} f$ and $f \in \mathcal{D}(G)$. An equivalent way of stating the question is the following. The function K defines a distribution on $G \setminus \{0\}$ (i.e. a linear functional on $\mathcal{D}(G \setminus \{0\})$) by ordinary integration. Can this functional be extended to $\mathcal{D}(G)$ as a homogeneous distribution?

One first remark is that, if such an extension exists, it is not unique. The reason is that the Dirac delta δ_0 at the origin is homogeneous of degree $-Q$, so that if $\tilde{K}$ is a distribution satisfying (4.2.1) when 0 is not in the support of f, the same is true for $\tilde{K} + c\delta_0$ for any constant c.

[2] Smoothness is not a necessary requirement. Local integrability away from the origin would be sufficient at this stage.

The other remark is that the answer is negative in general. A simple example is given by the function $K(x) = 1/|x|^Q$, defined in terms of some fixed homogeneous norm on G. Assume that $\tilde{K}$ exists. Homogeneity of degree $-Q$ means that

$$\langle \tilde{K}, f \circ D_t \rangle = \langle \tilde{K}, f \rangle \ ,$$

for every $t > 0$ and every f. Take $f \in \mathcal{D}(G)$, with $0 \le f(x) \le 1$, identically equal to 1 on a neighborhood U of the origin, and supported on $D_2 U$. Then $f - f \circ D_2 \ge 0$, is not identically zero, and it is supported on $D_2 U \setminus D_{1/2} U$. Hence

$$\langle \tilde{K}, f - f \circ D_2 \rangle = \int \frac{f(x) - f(D_2 x)}{|x|^Q} \, dx > 0 \ ,$$

in contrast with the fact that

$$\langle \tilde{K}, f - f \circ D_2 \rangle = \langle \tilde{K}, f \rangle - \langle \tilde{K}, f \circ D_2 \rangle = 0 \ .$$

In order to specify the conditions under which the answer to our question is positive, we must first present a "polar decomposition" of the Lebesgue measure".

Lemma 4.2.1. *Let $|\ |$ be a homogeneous norm on G, and let S be the unit sphere. There is a positive Borel measure σ on S such that*

$$\int_G f(x) \, dx = \int_0^\infty \int_S f\big(D_t(x)\big) \, d\sigma(x) \, t^{Q-1} \, dt \ ,$$

for every integrable function f.

Proof. If E is a Borel subset of S, let

$$E^\sharp = \{ D_t(x) : x \in E \ , \ t \le 1 \} \ ,$$

and define

$$\sigma(E) = Q m(E^\sharp) \ .$$

For $0 < a < b$, let

$$E_{a,b} = \{ D_t(x) : x \in E \ , \ a < t \le b \} = D_b(E^\sharp) \setminus D_a(E^\sharp) \ .$$

Then

$$m(E_{a,b}) = \frac{b^Q - a^Q}{Q} \sigma(E) = \int_{E \times [a,b]} t^{Q-1} dt \, d\sigma \ .$$

Standard measure-theoretic arguments give the conclusion. □

If $K(x)$ is a homogeneous function of degree $-Q$, locally integrable away from the origin, we can then talk of its mean value on S. We show that its value does not depend on S (i.e. on the fixed homogeneous norm).

Lemma 4.2.2. *Let $K(x)$ is a homogeneous function of degree $-Q$, locally integrable away from the origin. The integral*

$$\int_S K(x) \, d\sigma(x) = \mu(K) \tag{4.2.2}$$

does not depend on the homogeneous norm.

Proof. It follows from Lemma 4.2.1 that, if $0 < a < b$,

$$\int_{a < |x| < b} K(x) \, dx = \log(b/a) \mu(K) \ .$$

Let $|\ |'$ be another homogeneous norm. By Proposition 2.5.1, there is $r > 0$ such that the ball B'_r centered at the origin and with radius r in the norm $|\ |'$ is contained in the unit ball B_1 in the norm $|\ |$. Then

$$\begin{aligned}\int_{B_2\setminus B_1} K(x)\,dx &= \int_{B_2\setminus B'_r} K(x)\,dx - \int_{B_1\setminus B'_r} K(x)\,dx \\ &= \int_{B_2\setminus B'_{2r}} K(x)\,dx + \int_{B'_{2r}\setminus B'_r} K(x)\,dx - \int_{B_1\setminus B'_r} K(x)\,dx\ .\end{aligned}$$

But, setting $x = D_2 x'$, we see that

$$\int_{B_2\setminus B'_{2r}} K(x)\,dx = \int_{B_1\setminus B'_r} K(x)\,dx\ .$$

Hence

$$\int_{r<|x|'<2r} K(x)\,dx = \int_{1<|x|<2} K(x)\,dx\ ,$$

and this gives the conclusion. □

Proposition 4.2.3. *Let K be a smooth homogeneous function of degree $-Q$ on $G\setminus\{0\}$, and let $|\ |$ be a homogeneous norm on G. Then K extends to a homogeneous distribution on G if and only if $\mu(K) = 0$.*

The proof requires a preliminary lemma.

Lemma 4.2.4. *Let $|\ |$ be a homogeneous norm on $G(=\mathfrak{g})$, and $\|\ \|$ a vector space norm on $\mathfrak{g}$. There is a constant $C > 0$ such that*

$$\|x\| \le C|x|^\gamma \qquad \text{if } |x| < 1\ ,$$

where $\gamma = \min_j \lambda_j$.

Proof. Let W_{λ_j} be the eigenspaces of the dilations on $\mathfrak{g}$, with eigenvalues t^{λ_j}. We can assume, without loss of generality, that

$$|x| = \sum_j \|x_j\|^{1/\lambda_j}\ ,$$

and that

$$\|x\| = \sum_j \|x_j\|\ ,$$

if x_j denotes the component on x in W_{λ_j}. If $|x| < 1$, then $\|x_j\| < 1$ for every j. Hence

$$\|x\| < \sum_j \|x_j\|^{\gamma/\lambda_j} \le C\Big(\sum_j \|x_j\|^{1/\lambda_j}\Big)^\gamma = |x|^\gamma\ ,$$

as required. □

Proof of Proposition 4.2.3. If $f \in \mathcal{D}(G)$, we claim that the limit

$$\lim_{\varepsilon\to 0} \int_{|x|>\varepsilon} K(x) f(x)\,dx$$

exists and is finite. We check the Cauchy condition for $0 < \varepsilon < \varepsilon'$:

$$\begin{aligned}\Big|\int_{|x|>\varepsilon} K(x)f(x)\,dx - \int_{|x|>\varepsilon'} K(x)f(x)\,dx\Big| &= \Big|\int_{\varepsilon<|x|<\varepsilon'} K(x)f(x)\,dx\Big| \\ &= \Big|\int_{\varepsilon<|x|<\varepsilon'} K(x)\big(f(x)-f(0)\big)\,dx\Big| \\ &\le \int_{\varepsilon<|x|<\varepsilon'} |K(x)|\big|f(x)-f(0)\big|\,dx\ .\end{aligned}$$

The mean value theorem implies that

$$\big|f(x)-f(0)\big| \le \|\nabla f\|_\infty \|x\| \ ,$$

which implies, for $|x|<1$, that

$$\big|f(x)-f(0)\big| \le C\|\nabla f\|_\infty |x|^\gamma \ .$$

Hence, if $\varepsilon'<1$,

$$\begin{aligned}\bigg|\int_{|x|>\varepsilon} K(x)f(x)\,dx - \int_{|x|>\varepsilon'} K(x)f(x)\,dx\bigg| \\ \le C\|\nabla f\|_\infty \int_{\varepsilon<|x|<\varepsilon'} |x|^{-Q+\gamma}\,dx \\ \le C\|\nabla f\|_\infty {\varepsilon'}^\gamma \ .\end{aligned} \tag{4.2.3}$$

This implies the Cauchy condition.

We can then define, for $f\in\mathcal{D}(G)$,

$$\langle \tilde K, f\rangle = \lim_{\varepsilon\to 0}\int_{|x|>\varepsilon} K(x)f(x)\,dx \ . \tag{4.2.4}$$

This is clearly a linear functional. We must prove that it is continuous. Restricting to $f\in\mathcal{D}(E)$, with E compact in G, we have, by (4.2.3),

$$\begin{aligned}\bigg|\int_{|x|>\varepsilon} K(x)f(x)\,dx\bigg| &\le \bigg|\int_{\varepsilon<|x|<1} K(x)f(x)\,dx\bigg| + \bigg|\int_{|x|>1} K(x)f(x)\,dx\bigg| \\ &\le C\|\nabla f\|_\infty + C\int_{|x|>1\,,\,x\in E} |f(x)|\,dx \\ &\le C(E)\|f\|_{C^1} \ .\end{aligned}$$

To prove the "only if" part, assume that a function K, with $\mu(K)\neq 0$, extends to a homogeneous distribution $\tilde K$. We can decompose K as

$$K(x) = K_0(x) + \frac{\mu(K)}{\sigma(S)}|x|^{-Q} \ ,$$

where $\mu(K_0)=0$. Since also K_0 extends to a homogeneous distribution, the same would be true for $|x|^{-Q}$, in contradiction with our previous remark. □

Remark. Even though the condition $\mu(K)=0$ is independent of the homogeous norm, the distribution defined in (4.2.4) depends on the particular norm. As an example, consider the function

$$K(re^{i\theta}) = \frac{\cos 4\theta}{r^2}$$

on $\mathbb{R}^2(=\mathbb{C})$, the dilation D_t being scalar multiplication by t. Take $|x+iy| = (x^2+y^2)^{1/2}$ and $|x+iy|' = |x|+|y|$. Denoting by B_ε and B'_ε the balls of radius ε in the respective norms, we have

$$\begin{aligned}\int_{|x|'>\varepsilon} K(x)f(x)\,dx - \int_{|x|>\varepsilon} K(x)f(x)\,dx &= \int_{B_\varepsilon\setminus B'_\varepsilon} K(x)f(x)\,dx \\ &= \int_{B_1\setminus B'_1} K(x)f(\varepsilon x)\,dx \ ,\end{aligned}$$

which tends to

$$f(0)\int_{B_1\setminus B'_1} K(x)\,dx \ .$$

Simple geometric considerations show that this integral is strictly negative. Therefore the two distributions defined by (4.2.4) in terms of the two norms differ by a non-zero multiple of δ_0.

The distribution (4.2.4) is called a *principal value distribution*, denoted by

$$\mathrm{p.v.}K(x) \ .$$

Even though the notation is ambiguous unless a homogeneous norm is specified, this ambiguity is irrelevant for the purpose of discussing boundedness on L^p of the corresponding convolution operator. Observe in fact that, by Proposition 4.1.3, we have to consider $q = p$. The Dirac delta produces the identity operator, and its presence does not make any difference. The same harmless ambiguity is present in the following statement.

THEOREM 4.2.5. *Let K be a distribution on G, homogeneous of degree $-Q$ and smooth away from the origin, and let K_0 be the corresponding homogeneous function on $G \setminus \{0\}$. Then $\mu(K_0) = 0$ and*

$$K = p.v.K_0 + c\delta_0 \ .$$

PROOF. By Proposition 4.2.3, $\mu(K_0) = 0$. Then $H = K - \mathrm{p.v.}K_0$ is also homogeneous of degree $-Q$ and supported at the origin. We adapt the proof of Theorem 4.1.4.

Take $f \in \mathcal{D}(G)$ with $f(0) = 0$. Due to the compact support of H, $\big|\langle H, f\rangle\big|$ is controlled by some C^k norm of f on a fixed small neighborhood of the origin. But, because of its homogeneity,

$$\langle H, f\rangle = \langle H, f \circ D_t\rangle \ .$$

If t tends to 0, the C^k norms of $f \circ D_t$ tend to 0, so that $\langle K, f\rangle = 0$.

Fix now $\varphi \in \mathcal{D}(G)$ with $\varphi(0) = 1$. If f is a general function in $\mathcal{D}(G)$,

$$\langle H, f\rangle = \langle H, f - f(0)\varphi\rangle + f(0)\langle H, \varphi\rangle = f(0)\langle H, \varphi\rangle \ .$$

Hence $H = \langle H, \varphi\rangle\delta_0$. □

3. The almost orthogonality principle

For convolution operators by distributions homogeneous of degree $-Q$ it makes sense to discuss boundedness from $L^p(G)$ to itself. In this section and in the next one we will develop the tools that will allow us to prove boundedness on L^2 for principal value distributions that are smooth away from the origin and homogeneous of degree $-Q$.

In this section we present the *almost orthogonality principle* for general linear operator on Hilbert spaces.

Let $\{T_k\}_{k\in\mathbb{Z}}$ be a sequence of continuous linear operators from a Hilbert space H to itself, with

$$\|T_k\| \le C \ , \tag{4.3.1}$$

for every k. Let

$$V_k = (\ker T_k)^\perp \ , \qquad W_k = \overline{T_k H} \ ,$$

and assume that

$$V_j \perp V_k \ , \qquad W_j \perp W_k \ , \tag{4.3.2}$$

for $j \neq k$.

Observe that the two conditions in (4.3.2) are respectively equivalent to

$$T_j^* T_k = 0 \ , \qquad T_j T_k^* = 0 \ . \tag{4.3.3}$$

Let $S_N = \sum_{|k|\le N} T_k$. For $v \in H$, calling P_k the orthogonal projection onto V_k, we have

$$\begin{aligned} \|S_N v\|^2 &= \Big\| \sum_{|k|\le N} T_k v \Big\|^2 = \sum_{|k|\le N} \|T_k v\|^2 \\ &= \sum_{|k|\le N} \|T_k P_k v\|^2 \le C \sum_{|k|\le N} \|P_k v\|^2 \\ &\le C\|v\|^2 \ , \end{aligned}$$

because the $P_k v$ are mutually orthogonal. Hence $\|S_N\| \leq C$, independently of the number of operators. The same proof shows that $\{S_N v\}$ is a Cauchy sequence in H.

We then have proved the following “orthogonality principle”.

PROPOSITION 4.3.1. *Let $\{T_k\}$ be a sequence of linear operators on H satisfying* (4.3.1) *and* (4.3.2). *Then the series*

$$\sum_{k\in\mathbb{Z}} T_k$$

converges in the strong topology to an operator S satisfying $\|S\| \leq C$.

Motivated by these considerations, we can formulate the almost orthogonality principle.

THEOREM 4.3.2. *Let $\{T_k\}_{k\in\mathbb{Z}}$ be a sequence of bounded linear operators on H, and let $a = \{a(n)\}_{n\in\mathbb{Z}}$ be a summable sequence of positive numbers.*

Assume that

$$\|T_j^* T_k\|^{1/2} \leq a(j-k) \ , \qquad \|T_j T_k^*\|^{1/2} \leq a(j-k) \ , \tag{4.3.4}$$

for every j, k. Then the series

$$\sum_{k\in\mathbb{Z}} T_k$$

converges in the strong topology to an operator S satisfying $\|S\| \leq A = \|a\|_1$.

PROOF. We recall that, for any bounded linear operator T on H, $\|T^*T\| = \|T\|^2$. The inequality

$$\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$$

is obvious, and on the other hand, denoting the inner product in H by $(\ |\ )$,

$$\begin{aligned}
\|T^*T\| &= \sup_{\|v\|\leq 1, \|w\|\leq 1} \big|(T^*Tv|w)\big| \\
&= \sup_{\|v\|\leq 1, \|w\|\leq 1} \big|(Tv|Tw)\big| \\
&\geq \sup_{\|v\|\leq 1} (Tv|Tv) \\
&= \|T\|^2 \ .
\end{aligned}$$

In particular,

$$\|T_k\| = \|T_k^* T_k\|^{1/2} \leq a(0) \leq A \ .$$

Let $U = S_N^* S_N$. Since $U = U^*$, it follows that $\|U^2\| = \|U\|^2$, and inductively, $\|U^{2^m}\| = \|U\|^{2^m}$, i.e.

$$\|S_N^* S_N\| = \big\|(S_N^* S_N)^{2^m}\big\|^{1/2^m} \ .$$

Setting $n = 2^m$, we have

$$(S_N^* S_N)^n = \sum_{-N\leq i_1, i_2, \dots, i_{2n}\leq N} T_{i_1}^* T_{i_2} \cdots T_{i_{2n-1}}^* T_{i_{2n}} \ .$$

Each summand can be majorized in two different ways:

$$\begin{aligned}
\|T_{i_1}^* T_{i_2} \cdots T_{i_{2n}}\| &\leq \|T_{i_1}^* T_{i_2}\| \cdots \|T_{i_{2n-1}}^* T_{i_{2n}}\| \\
&\leq a(i_1 - i_2)^2 a(i_3 - i_4)^2 \cdots a(i_{2n-1} - i_{2n})^2 \\
\|T_{i_1}^* T_{i_2} \cdots T_{i_{2n}}\| &\leq \|T_{i_1}^*\|\|T_{i_2} T_{i_3}^*\| \cdots \|T_{i_{2n}}\| \\
&\leq A^2 a(i_2 - i_3)^2 \cdots a(i_{2n-2} - i_{2n-1})^2 \ .
\end{aligned}$$

Moltiplying side by side and extracting square roots, we obtain

$$\|T_{i_1}^* T_{i_2} \cdots T_{i_{2n}}\| \leq A a(i_1 - i_2) a(i_2 - i_3) \cdots a(i_{2n-1} - i_{2n}) \ .$$

Therefore, summing first over i_1, then over i_2, etc., we have

$$\begin{aligned}\|(S_N^*S_N)^n\| &= \sum_{-N\le i_1,i_2,\dots,i_{2n}\le N} \|T_{i_1}^*T_{i_2}\cdots T_{i_{2n}}\| \\ &\le A \sum_{-N\le i_1,i_2,\dots,i_{2n}\le N} a(i_1-i_2)a(i_2-i_3)\cdots a(i_{2n-1}-i_{2n}) \\ &\le A^2 \sum_{-N\le i_2,\dots,i_{2n}\le N} a(i_2-i_3)\cdots a(i_{2n-1}-i_{2n}) \\ &\dots\dots \\ &\le A^{2n-1} \sum_{-N\le i_{2n-1},i_{2n}\le N} a(i_{2n-1}-i_{2n}) \\ &\le (2N+1)A^{2n} \ .\end{aligned}$$

Hence

$$\|S_N^*S_N\| \le (2N+1)^{1/2^m}A^2 \ .$$

Taking the limit for $m\to\infty$, we see that $\|S_N^*S_N\|\le A^2$, i.e. $\|S_N\|\le A$ for every N.

We prove now that, for every $v\in H$, $\{S_Nv\}$ is a Cauchy sequence. Assume first that $v=T_j^*u$ for some $u\in H$. If $w\in H$, and $M<N$,

$$\begin{aligned}\big|\big((S_N-S_M)v|w\big)\big| &\le \sum_{M+1\le|k|\le N} \big|(T_kT_j^*u|w)\big| \\ &\le \|u\|\|w\| \sum_{M+1\le|k|\le N} a(k-j)^2 \\ &\le A\|u\|\|w\| \sum_{M+1\le|k|\le N} a(k-j) \ .\end{aligned}$$

Hence,

$$\big\|(S_N-S_M)v\big\| \le A\|u\| \sum_{|k|\ge M+1-|j|} a(k) \ ,$$

which can be made smaller than any $\varepsilon>0$ by taking M large enough.

By linearity, $\{S_Nv\}$ is a Cauchy sequence for every $v\in H_0=\sum_{j\in\mathbb{Z}}T_j^*H$. If we take now $v\in\overline{H_0}$, given $\varepsilon>0$ there is $v'\in H_0$ with $\|v-v'\|<\varepsilon$. Then

$$\big\|(S_N-S_M)v\big\| \le \big\|(S_N-S_M)v'\big\| + 2A\varepsilon \ ,$$

and this implies that $\{S_Nv\}$ is again a Cauchy sequence.

Finally, every $v\in H$ decomposes as $v=v_1+v_2$, with $v_1\in\overline{H_0}$ and $v_2\in\overline{H_0}^{\perp}$. But, for every $j\in\mathbb{Z}$ and $w\in H$,

$$(T_jv_2|w)=(v_2|T_j^*w)=0 \ ,$$

so that $T_jv_2=0$. Hence $S_Nv=S_Nv_1$ and the proof is complete. □

4. Convolution with principal value distributions

Let K be a distribution on G, homogeneous of degree $-Q$ and smooth away from the origin. We are interested in the convolution operator $Tf=f*K$.

We know from Section 2 that[3] $\mu(K)=0$, and that, once we fix a homogeneous norm $|\ |$ on G, the distribution K equals

$$\text{p.v.}\,K(x)+c\delta_0$$

[3] With an abuse of notation, we use the same letter K to denote the distribution on G and the function on $G\setminus\{0\}$.

for some constant c. We can then concentrate our attention on the first term, i.e. we will assume that $K = \text{p.v.}\, K(x)$.

Then, if $f \in \mathcal{D}(G)$,

$$Tf(x) = \lim_{\varepsilon \to 0} \int_{|y|>\varepsilon} f(xy^{-1})K(y)\,dy \ .$$

If we set

$$D_k = \{x : 2^{-k} \leq x \leq 2^{-k+1}\} \ , \qquad K_k(x) = K(x)\chi_{D_k}(x) \ , \qquad T_k f = f * K_k \ ,$$

then

$$Tf = \sum_{k\in\mathbb{Z}} T_k f \ .$$

We shall show that the almost orthogonality principle applies to the T_k as operators on $L^2(G)$. But before, we must establish a mean value theorem for homogeneous functions.

Even though this is not strictrly necessary, we assume that the homogeneous norm satisfies the triangular inequality

$$|xy| \leq |x| + |y| \ .$$

LEMMA 4.4.1. *Let f be a C^1 function on $G \setminus \{0\}$, homogenous of degree $\alpha \in \mathbb{R}$. There is a constant $A > 0$ such that, if $|y| < |x|/2$, then*

$$\big|f(xy) - f(x)\big| \leq A|y|^\gamma |x|^{\alpha-\gamma} \ ,$$

where $\gamma = \min_j \lambda_j$.

PROOF. Take x with $|x| = 1$ and $|y| < 1/2$. The map $y \longmapsto f(xy) - f(x)$ is smooth, it vanishes for $y = 0$, and depends smoothly on x. By the mean value theorem and compactness,

$$\big|f(xy) - f(x)\big| \leq C\|y\| \ ,$$

where C is independent of x and $\|\ \|$ is a vector space norm on $\mathfrak{g}$. By Lemma 4.2.4,

$$\big|f(xy) - f(x)\big| \leq A|y|^\gamma \ .$$

For general $x \neq 0$, let $t = |x|$ and $x' = D_{t^{-1}}x$. Then

$$\begin{aligned} \big|f(xy) - f(x)\big| &= t^\alpha \big|f(x'D_{t^{-1}}y) - f(x')\big| \\ &\leq At^\alpha |D_{t^{-1}}y|^\gamma \\ &\leq A|x|^{\alpha-\gamma}|y|^\gamma \ . \qquad \square \end{aligned}$$

THEOREM 4.4.2. *Let K be a distribution on G, homogeneous of degree $-Q$ and smooth away from the origin. Then the operator $Tf = f * K$ is bounded on $L^2(G)$.*

PROOF. We prove that the T_k satisfy the hypotheses of Theorem 4.3.2. Since $K_k \in L^1(G)$ for every k, each T_k is bounded on $L^2(G)$, and

$$\begin{aligned} \|T_k\| &\leq \|K_k\|_1 \\ &= \int_{D_k} |K(x)|\,dx \\ &= \int_{D_0} |K(x)|\,dx \\ &= \|K_0\|_1 \ . \end{aligned} \tag{4.4.1}$$

We shall now concentrate our attention on the first of the two conditions (4.3.4), the second admitting an identical proof.

In order to describe T_k^*, we consider that, for $f, g \in \mathcal{D}(G)$,

$$\begin{aligned}(T_k^* f|g) &= (f|T_k g)\\ &= \langle f, \bar{g} * \bar{K}_k\rangle\\ &= \langle f * (\bar{K}_k)\check{}, \bar{g}\rangle \ .\end{aligned}$$

Hence, if

$$K_k^*(x) = \overline{K_k(x^{-1})} \ ,$$

we see that $T_k^* f = f * K_k^*$. In order to prove the first condition (4.3.4), we observe that $T_j^* T_k f = f * K_k * K_j^*$, so that

$$\|T_j^* T_k\| \leq \|K_k * K_j^*\|_1 \ , \tag{4.4.2}$$

and we are led to estimate this L^1 norm.

For $|j-k| \leq 1$ we use Young's inequality, i.e.

$$\|K_k * K_j^*\|_1 \leq \|K_k\|_1 \|K_j^*\|_1 = \|K_0\|_1^2 \ ,$$

so that we can take $a(0) = a(\pm 1) = \|K_0\|_1$.

We consider now the case $|j-k| \geq 2$. We show how to estimate $\|K_k * K_j^*\|_1$ when $k - j \geq 2$, the other estimates being quite similar. The proof is based on the following lemma. □

LEMMA 4.4.3. *Let φ, ψ be two integrable functions, and assume that*

(i) $\operatorname{supp}\varphi \subset B(0,r)$ *and* $\int_G \varphi(x)\,dx = 0$;

(ii) *there are constants $C, \delta > 0$ such that, if $|y| \leq r$,*

$$\int_G \big|\psi(y^{-1}x) - \psi(x)\big|\,dx \leq C|y|^\delta \ .$$

Then $\|\varphi * \psi\|_1 \leq C r^\delta \|\varphi\|_1$.

PROOF. Since φ has integral zero,

$$\begin{aligned}\varphi * \psi(x) &= \int_G \varphi(y)\psi(y^{-1}x)\,dy\\ &= \int_G \varphi(y)\big(\psi(y^{-1}x) - \psi(x)\big)\,dy \ .\end{aligned}$$

Hence

$$\begin{aligned}\|\varphi * \psi\|_1 &= \int_G \left|\int_G \varphi(y)\big(\psi(y^{-1}x) - \psi(x)\big)\,dy\right|\,dx\\ &\leq \int_{|y|<r} |\varphi(y)| \int_G \big|\psi(y^{-1}x) - \psi(x)\big|\,dx\,dy\\ &\leq C\int_{|y|<r} |\varphi(y)||y|^\delta\,dy\\ &\leq C r^\delta \|\varphi\|_1 \ . \qquad \square\end{aligned}$$

□

We apply Lemma 4.4.3 to $\varphi = K_k$ and $\psi = K_j^*$. Condition (i) is satisfied with $r = 2^{-k+1}$. In order to verify (ii), we first take $j = 0$ and show that, for $|y| < 1/2$,

$$\int_G \big|K_0^*(y^{-1}x) - K_0^*(x)\big|\,dx \leq B|y|^\delta \ , \tag{4.4.3}$$

where $\delta = \min\{1, \gamma\}$, γ being as in Lemma 4.4.1.

We simplify the notation observing that, if we change x into x^{-1},

$$\int_G \big|K_0^*(y^{-1}x) - K_0^*(x)\big|\,dx = \int_G \big|K_0(xy) - K_0(x)\big|\,dx \ .$$

The integral is extended to $D_0 \cup (D_0 y^{-1})$, and we split it as

$$\begin{aligned}\int_{D_0\cap(D_0y^{-1})} \big|K(xy)-K(x)\big|\,dx &+ \int_{D_0\setminus(D_0y^{-1})}\big|K(x)\big|\,dx \\ &+ \int_{(D_0y^{-1})\setminus D_0}\big|K(xy)\big|\,dx \\ &= \int_{D_0\cap(D_0y^{-1})} \big|K(xy)-K(x)\big|\,dx + 2\int_{D_0\setminus(D_0y^{-1})}\big|K(x)\big|\,dx\ .\end{aligned}$$

For the first integral we apply Lemma 4.1 to obtain

$$\begin{aligned}\int_{D_0\cap(D_0y^{-1})} \big|K(xy)-K(x)\big|\,dx &\le |y|^\gamma \int_{D_0}|x|^{Q-\gamma}\,dx \\ &\le C|y|^\gamma\ .\end{aligned}$$

For the second integral, observe that

$$D_0\setminus(D_0y^{-1}) \subseteq \{x : 1-|y|\le |x|\le 1\}\ ,$$

so that

$$\begin{aligned}\int_{D_0\setminus(D_0y^{-1})}\big|K(x)\big|\,dx &\le C\int_{1-|y|<|x|<1}|x|^{-Q}\,dx \\ &= C'\int_{1-|y|}^1 \frac{dt}{t} \\ &\le 2C'|y|\ .\end{aligned}$$

This proves (4.4.3). Going back to condition (ii), observe that, if $y\in\operatorname{supp}K_k$, then $|D_{2^j}y|\le 2^{j-k+1}\le 1/2$. Then, changing variables and using the homogeneity of K,

$$\begin{aligned}\int_G\big|K_j^*(y^{-1}x)-K_j^*(x)\big|\,dx &= \int_G\big|K_0^*\big((D_{2^j}y)^{-1}x\big)-K_0^*(x)\big|\,dx \\ &\le B2^{j\delta}|y|^\delta\ .\end{aligned}$$

Hence (ii) is verified with $C=B2^{j\delta}$. It follows from Lemma 4.4.3 that

$$\|K_k * K_j^*\|_1 \le 2^\delta B 2^{-(k-j)\delta}\|K_k\|_1 \le B'2^{-(k-j)\delta}\ .$$

Finally, taking $a(n)=B'2^{-\delta|n|}$ if $|n|\ge 2$, and the assumptions of Theorem 4.3.2 are satisfied.

We state without proof the following sharper result. Its proof requires an adaptation of the Calderón-Zygmund theory of singular integrals, that we cannot present here[4].

Theorem 4.4.4. *Let K be a distribution on G, homogeneous of degree $-Q$ and smooth away from the origin. If $1<p<\infty$, the operator $Tf=f*K$ is bounded on $L^p(G)$.*

Using Theorem 4.4.4 instead of Theorem 4.1.4 in the proofs, we can then extend Corollaries 4.1.5 and 4.1.6.

Corollary 4.4.5. *The conclusions in Corollaries 4.1.5 and 4.1.6 extend to $d(\alpha)=\mu$, $r=p\in(1,\infty)$.*

[4] It can be found in E.M. Stein *Harmonic Analysis.*

CHAPTER 5

Proof of Hörmander's theorem for sublaplacians

Ouraim is to prove Theorem 1.5.7 for sub-Laplacians (i.e., operators of the form L_1). The proof is based on a comparison between Sobolev and Lipschitz norms.

1. Lipschitz, Besov and Sobolev norms

Let $0 < \alpha < 1$. The *p-Lipschitz norm of order* α of a function $f \in L^p(\mathbb{R}^n)$ is defined as

$$\|f\|_{\Lambda^p_\alpha} = \|f\|_p + \sup_{h \neq 0} |h|^{-\alpha} \|\tau_h f - f\|_p \ , \tag{5.1.1}$$

where we have set $\tau_h f(x) = f(x-h)$. The space $\Lambda^p_\alpha(\mathbb{R}^n)$ is defined as the subspace of those functions in $L^p(\mathbb{R}^n)$ for which $\|f\|_{\Lambda^p_\alpha} < \infty$.

Notice that, for every $a > 0$, one obtains an equivalent norm on $\Lambda^p_\alpha(\mathbb{R}^n)$ by replacing $\sup_{h\neq 0}$ with $\sup_{0<|h|<a}$. This follows easily from the fact that, for $|h| \geq a$, $|h|^{-\alpha}\|\tau_h f - f\|_p \leq 2a^{-\alpha}\|f\|_p$.

Proposition 5.1.1. *The space Λ^∞_α consists of the bounded Hölder-continuous functions of order α.*

Proof. It is quite clear that bounded Hölder-continuous functions of order α belong to $\Lambda^\infty_\alpha(\mathbb{R}^n)$. Suppose therefore that $f \in \Lambda^\infty_\alpha(\mathbb{R}^n)$. If we prove that f is continuous, the rest follows easily.

Let φ be a continuous function supported on the unit ball and with $\int_{\mathbb{R}^n} \varphi = 1$. The functions $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi(\varepsilon^{-1}x)$ form an approximate identity for $\varepsilon \to 0$. Since $\int_{\mathbb{R}^n} \varphi_\varepsilon = 1$,

$$\begin{aligned} f * \varphi_\varepsilon(x) - f(x) &= \int_{\mathbb{R}^n} \big(f(x-y) - f(x)\big)\varphi_\varepsilon(y)\, dy \\ &= \int_{\mathbb{R}^n} \big(\tau_y f(x) - f(x)\big)\varphi_\varepsilon(y)\, dy \ , \end{aligned}$$

so that

$$\|f * \varphi_\varepsilon - f\|_\infty \leq \int_{\mathbb{R}^n} \|\tau_y f - f\|_\infty |\varphi_\varepsilon(y)|\, dy \ .$$

From the inequality

$$\|\tau_y f - f\|_\infty \leq \|f\|_{\Lambda^\infty_\alpha} |y|^\alpha \ ,$$

and the fact that φ_ε is supported on the ball of radius ε, we obtain that

$$\|f * \varphi_\varepsilon - f\|_\infty \leq \|f\|_{\Lambda^\infty_\alpha} \varepsilon^\alpha \int_{\mathbb{R}^n} |\varphi_\varepsilon(y)|\, dy = \|f\|_{\Lambda^\infty_\alpha} \varepsilon^\alpha \int_{\mathbb{R}^n} |\varphi(y)|\, dy \ .$$

This gives the uniform convergence of $f * \varphi_\varepsilon$ to f. Since the φ_ε are continuous, the same is true for f. □

Remark. *We motivate the restriction* $0 < \alpha < 1$. *If the quantity in* (5.1.1) *is finite for some* $\alpha > 1$, *then* $f = 0$. *The exclusion of* $\alpha = 1$ *is less simple to explain. An indication that some problems arise with* (5.1.1) *is in Proposition 5.1.2 below, whose statement is false*[1] *for* $\alpha = 1$.

A more general class of norms are the (p, q)-*Besov norms of order* α. For $1 \leq q < \infty$, one sets[2]

$$\|f\|_{\Lambda_\alpha^{p,q}} = \|f\|_p + \left(\int_{|h|<1} \left(|h|^{-\alpha} \|\tau_h f - f\|_p \right)^q \frac{dh}{|h|^n} \right)^{\frac{1}{q}} , \tag{5.1.2}$$

and the Besov space $\Lambda_\alpha^{p,q}(\mathbb{R}^n)$ is defined as the space of L^p-functions for which (5.1.2) is finite. For $q = \infty$, one sets $\Lambda_\alpha^{p,\infty}(\mathbb{R}^n) = \Lambda_\alpha^p(\mathbb{R}^n)$.

We shall be mainly interested in the case $p = 2$, with $q = 2$ or $q = \infty$.

Proposition 5.1.2. *For* $0 < \alpha < 1$, $\Lambda_\alpha^{2,2}(\mathbb{R}^n)$ *coincides with the Sobolev space* $H^\alpha(\mathbb{R}^n)$ *and the two norms are equivalent.*

Proof. Observe that, denoting by

$$\hat{f}(\xi) = \int_{\mathbb{R}}^{n} f(x) e^{-ix\cdot\xi} \, dx$$

the Fourier transform of f, then $\widehat{\tau_h f}(\xi) = e^{-ih\cdot\xi} \hat{f}(\xi)$. Hence, for $f \in L^2(\mathbb{R}^n)$ and using Plancherel's formula,

$$\begin{aligned} \|f\|_{\Lambda_\alpha^{2,2}}^2 &\sim \|f\|_2^2 + \int_{\mathbb{R}^n} \|\tau_h f - f\|_2^2 \frac{dh}{|h|^{n+2\alpha}} \\ &\sim \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \, d\xi + \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 \frac{|e^{-ih\cdot\xi} - 1|^2}{|h|^{n+2\alpha}} \, d\xi \, dh \, . \end{aligned} \tag{5.1.3}$$

Interchanging the order of integration, we are led to consider first the integral

$$I(\xi) = \int_{\mathbb{R}^n} \frac{|e^{-ih\cdot\xi} - 1|^2}{|h|^{n+2\alpha}} \, dh \, .$$

Using the inequalities $|e^{-ih\cdot\xi} - 1| \leq |h||\xi|$ for $|h|$ small, and $|e^{-ih\cdot\xi} - 1| \leq 2$ for $|h|$ large, we see that the integral is convergent for $0 < \alpha < 1$.

Obviously, $I(0) = 0$. If $\xi \neq 0$, decompose ξ in polar coordinates as $\xi = |\xi|\omega$, with ω in the unit sphere. Changing variable, $|\xi| h = h'$, we find that

$$I(\xi) = |\xi|^{2\alpha} \int_{\mathbb{R}^n} \frac{|e^{-ih'\cdot\omega} - 1|^2}{|h'|^{n+2\alpha}} \, dh' = |\xi|^{2\alpha} I(\omega) \, .$$

If A is an orthogonal transformation of $\mathbb{R}^n$, the change of variable $h' = Ah''$ shows that $I(\omega) = I(A\omega)$. Hence $I(\omega)$ is constant, and its constant value is non-zero. Putting this in (5.1.3), we find that

$$\|f\|_{\Lambda_\alpha^{2,2}}^2 \sim \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 (1 + |\xi|^{2\alpha}) \, d\xi \sim \|f\|_{(\alpha)} \, . \qquad \square$$

Lemma 5.1.3. *We have the following continuous inclusions*[3]

$$\Lambda_\alpha^{2,2}(\mathbb{R}^n) \subset \Lambda_\alpha^{2,\infty}(\mathbb{R}^n) \subset \Lambda_\beta^{2,2}(\mathbb{R}^n) \, ,$$

---

[1] For $\alpha = 1$ it is more appropriate to replace the "first-order difference" $\tau_h f - f$ with the "second-order difference" $\tau_{2h} f - 2\tau_h f + f$. With this modification, Proposition 5.1.2 becomes true also for $\alpha = 1$, and in fact it can be extended to all $\alpha < 2$. For larger values of α, one must use higher-order differences.

[2] As for Lipschitz norms, the restriction $|h| < 1$ in the domain of integration can be replaced by $|h| < a$ for any a positive or even infinite.

[3] More generally, one has the following continuous inclusions:

$$\Lambda_\alpha^{p,q_1}(\mathbb{R}^n) \subset \Lambda_\alpha^{p,q_2}(\mathbb{R}^n) \, ,$$

if $q_1 < q_2$, and

$$\Lambda_\alpha^{p,\infty}(\mathbb{R}^n) \subset \Lambda_\beta^{p,1}(\mathbb{R}^n) \, ,$$

if $0 < \beta < \alpha < 1$.

PROOF. Take $f \in \Lambda_\alpha^{2,2}(\mathbb{R}^n)$. From the inequality $|e^{it} - 1| \le C_\alpha |t|^\alpha$, we have

$$\begin{aligned}\|\tau_h f - f\|_2^2 &\sim \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |e^{-ih\cdot\xi} - 1|^2 \, d\xi \\ &\lesssim \int_{\mathbb{R}^n} |\hat{f}(\xi)|^2 |h|^{2\alpha} |\xi|^{2\alpha} \, d\xi \\ &\le |h|^{2\alpha} \|f\|_{(\alpha)}^2 \ .\end{aligned}$$

This gives the first inclusion by Proposition 5.1.2.

Take now $f \in \Lambda_\alpha^{2,\infty}(\mathbb{R}^n)$. Then

$$\|\tau_h f - f\|_2 \le \|f\|_{\Lambda_\alpha^{2,\infty}} |h|^\alpha \ .$$

Therefore,

$$\begin{aligned}\int_{|h|<1} |h|^{-2\beta} \|\tau_h f - f\|_2^2 \frac{dh}{|h|^n} &\le \|f\|_{\Lambda_\alpha^{2,\infty}}^2 \int_{|h|<1} |h|^{2(\alpha-\beta)-n} \, dh \\ &\lesssim \|f\|_{\Lambda_\alpha^{2,\infty}}^2 \ .\end{aligned}$$

Adding the L^2-norm, one gets to the conclusion. □

COROLLARY 5.1.4. *If $s < \alpha < 1$, $\Lambda_\alpha^2(\mathbb{R}^n) \subset H^s(\mathbb{R}^n)$, and the inclusion is continuous.*

2. Lipschitz norms defined in terms of flows

Lipschitz and Besov norms are defined in terms of the translations τ_h. One can interpret translations in $\mathbb{R}^n$ as generated by the flow of n very special vector fields.

If $h = (h_1, \dots, h_n) = h_1 e_1 + \dots + h_n e_n$, then $\tau_h = \tau_{h_1 e_1} \circ \dots \circ \tau_{h_n e_n}$; setting $X_j = -\partial_{x_j}$,

$$\tau_{h_j e_j} = \exp(h_j X_j) \ ,$$

and

$$\tau_h = \exp(h_1 X_1) \cdots \exp(h_n X_n) \ .$$

We shall discuss how more general vector fields can be used to introduce adapted Lipschitz norms, and to compare these among themselves and with the standard Lipschitz norms presented above. We shall restrict ourselves to $p = 2$.

Let Ω be an open set, X a (smooth) vector field on Ω, and let K, K' be two compact subsets of Ω with $K \overset{\circ}{\subset\to} K'$. We fix $a = a(X, K, K') > 0$ small enough so that

(i) the flow $\varphi_{X,t}(x)$ is defined on K for $|t| \le a$;
(ii) for $|t| \le a$, $\varphi_{X,t}$ maps K diffeomorphically into K'.

For $f \in L^2(\Omega)$ with $\operatorname{supp} f \subset K$, we define the norm

$$\|f\|_{X,\alpha} = \|f\|_2 + \sup_{|t|\le a} |t|^{-\alpha} \|\exp(tX) f - f\|_2 \ . \tag{5.2.1}$$

LEMMA 5.2.1. *Given K, the norm* (5.2.1) *does not depend, up to equivalence, on the choice of a, as long as (i) and (ii) are satisfied, for some K', by both a and a'.*

PROOF. For the purpose of this proof, let us specify the value of a in (5.2.1) and write $\|f\|_{X,\alpha,a}$. Obviously, if $a < b$, $\|f\|_{X,\alpha,a} \le \|f\|_{X,\alpha,b}$.

---

if $\beta < \alpha$. These are the relevant inclusions with p fixed. See E.M. Stein, *Singular integrals and differentiability properties of functions*, Chap. V, Sect. 5.

On the other hand, let K_b be a compact subset of Ω such that (ii) is satisfied for $|t|<b$. For $a<|t|\le b$, we use the trivial majorization $\|\exp(tX)f-f\|_2\le\|\exp(tX)f\|_2+\|f\|_2$. Noticing that $\operatorname{supp}\big(\exp(tX)f\big)=\varphi_{X,t}^{-1}(\operatorname{supp} f)\subset K_b$, we have

$$\|\exp(tX)f\|_2^2=\int_{K_b}\big|f\big(\varphi_{X,t}(x)\big)\big|^2\,dx=\int_K|f(x)|^2J_{\varphi_{X,-t}}(x)\,dx\ .$$

By compactness, $J_{\varphi_{X,-t}}$ is bounded on K uniformly for $|t|\le b$. Therefore,

$$\|\exp(tX)f\|_2\le C\|f\|_2\ . \tag{5.2.2}$$

This easily gives the inequality $\|f\|_{X,\alpha,b}\le C'\|f\|_{X,\alpha,a}$. □

We shall not mention the choice of a in (5.2.2) unless it will be necessary.

Notice that, with $\Omega=\mathbb{R}^n$ and $X_j=-\partial_{x_j}$, the 2-Lipschitz norm $\|f\|_{\Lambda^2_\alpha}$ is equivalent to $\sum_{j=1}^n\|f\|_{X_j,\alpha}$.

The first comparison we make is between the Lipschitz norms associated with a vector field X and those associated with a modified vector field $\tilde X=\eta X$, for some $\eta\in\mathcal{D}(\Omega)$.

We expect that the trajectories $\tilde\gamma_x(t)$ of the flow generated by $\tilde X$ are contained in the trajectories $\gamma_x(t)$ of the flow generated by X, only travelling at a different speed. To be more precise, we expect that there is a real-valued function $\tau(x,t)$ such that

$$\tilde\gamma_x(t)=\gamma_x\big(\tau(x,t)\big)\ , \tag{5.2.3}$$

and $\tau(x,0)=0$. Differentiating in t, we obtain the equation

$$\tilde X_{\tilde\gamma_x(t)}=\partial_t\tau(x,t)X_{\tilde\gamma_x(t)}\ .$$

If $X_{\tilde\gamma_x(t)}\ne0$, this means

$$\eta\circ\gamma_x\big(\tau(x,t)\big)=\partial_t\tau(x,t)\ , \tag{5.2.4}$$

and surely (5.2.4) implies (5.2.3). For $x\in\Omega$ fixed, the function τ satisfies (5.2.4) if and only if $\tau_x(t)=\tau(x,t)$ is a solution of the one-dimensional Cauchy problem

$$\begin{cases}\frac{d}{dt}\tau_x=\eta\circ\gamma_x(\tau_x)\\ \tau_x(0)=0\ .\end{cases} \tag{5.2.5}$$

By Theorem 1.1.1 (iii), $\gamma_x(t)$ is smooth in both x and t. It follows from (iv) in the same theorem that (5.2.5) has a unique solution which is smooth in both x and t.

It will be convenient soon to set the notation

$$\varphi_X(x,t)=\varphi_{X,t}(x)=\gamma_x(t)\ .$$

Proposition 5.2.2. *Given K and $\eta\in\mathcal{D}(\Omega)$, there is a constant $C=C(K,\eta)$ such that, for every $\alpha\in(0,1)$ and $f\in L^2(\Omega)$ supported on K,*

$$\|f\|_{\eta X,\alpha}\le C\|f\|_{X,\alpha}\ .$$

Proof. If s,t are small enough and $|s|\le|t|$,

$$\begin{aligned}\|\exp(t\eta X)f-f\|_2&\le\|\exp(t\eta X)f-\exp(sX)f\|_2+\|\exp(sX)f-f\|_2\\&\le\|\exp(t\eta X)f-\exp(sX)f\|_2+|t|^\alpha|f|_{X,\alpha}\ .\end{aligned}$$

Hence

$$\|\exp(t\eta X)f-f\|_2\le|t|^\alpha|f|_{X,\alpha}+\inf_{|s|\le|t|}\|\exp(t\eta X)f-\exp(sX)f\|_2\ .$$

We have reduced matters to proving that

$$\inf_{|s|\le|t|}\|\exp(t\eta X)f-\exp(sX)f\|_2\lesssim|t|^\alpha\|f\|_{X,\alpha}\ . \tag{5.2.6}$$

We have

$$
\begin{aligned}
&\inf_{|s|\le|t|}\|\exp(t\eta X)f-\exp(sX)f\|_2^2\\
&\qquad=\inf_{|s|\le|t|}\int_\Omega\left|f\circ\varphi_X\big(x,\tau(x,t)\big)-f\circ\varphi_X(x,s)\right|^2dx \qquad (5.2.7)\\
&\qquad\le\frac{1}{2|t|}\int_{|s|\le|t|}\int_\Omega\left|f\circ\varphi_X\big(x,\tau(x,t)\big)-f\circ\varphi_X(x,s)\right|^2dx\,ds\ .
\end{aligned}
$$

We fix a compact subset K' of Ω containing K in its interior, and we impose that $|s|,|t|\le\delta$, with δ small enough so that $K\subset\big\{\varphi_X\big(x,\tau(x,t)\big):x\in K',|t|\le\delta\big\}$ and $K\subset\big\{\varphi_X(x,s):x\in K',|s|\le\delta\big\}$

For each $t\in[-\delta,\delta]$, we want to make a change of variables $(x,s)\mapsto(y,u)$ in such a way that the integrand becomes $\left|f\circ\varphi_X(y,u)-f(y)\right|^2$. We must then set $y=\varphi_X(x,s)$. Observing that the identity $\varphi_{X,u+s}=\varphi_{X,u}\circ\varphi_{X,s}$ can be written as

$$\varphi_X(x,u+s)=\varphi_X\big(\varphi_X(x,s),u\big)=\varphi_X(y,u)\ ,$$

we must also set $u=\tau(x,t)-s$. The change of variables is then

$$(y,u)=\Phi_t(x,s)=\big(\varphi_X(x,s),\tau(x,t)-s\big)\ ,$$

which is a smooth function in all variables x,s,t.

The Jacobian matrix $J_{x,s}\Phi_t$ in the variables x,s is

$$J_{x,s}\Phi_t(x,s)=\begin{pmatrix}J_x\varphi_X(x,s) & X_{\varphi_X(x,s)}\\ \nabla_x\tau(x,t) & -1\end{pmatrix}\ .$$

At $t=0$, $\tau(x,0)=0$ for every x, therefore $\nabla_x\tau(x,0)=0$. At $s=0$, $\varphi_X(x,0)=x$, so that $J_x\varphi_X(x,0)=I$. Therefore

$$\det J_{x,s}\Phi_0(x,0)=-1\ .$$

It follows that for $|s|,|t|$ smaller than some $\delta'>0$, the determinant of $J_{x,s}\Phi_t(x,s)$ is bounded from below in absolute value, uniformly in $x\in K'$. Hence, for each point $x\in K'$ there is a neighborhood $U_x\times(-\delta_x,\delta_x)$, with $\delta_x<\delta'$, on which every Φ_t with $|t|<\delta_x$ is invertible. By compactness, it is possible to cover K' with a finite number $U_{x_1},\dots,U_{x_m}$ of such neighborhoods.

If $\{\psi_j\}_{j=1}^m$ is a smooth partition of unity on K' subordinated to the U_{x_j}, we set $f_j=f\psi_j$. We then choose a final $\bar\delta$, smaller than δ' and δ_{x_j} for every j, such that $\operatorname{supp}f_j\circ\varphi_{X,s}\subset U_{x_j}$ for $|s|<\bar\delta$. Notice that $\bar\delta$ does not depend either on f or α, but only on K.

Also notice that, since $\tau(x,0)=0$, there is a constant A such that $|\tau(x,t)|\le A|t|$ for $|t|\le\delta'$ and $x\in K'$. Therefore $|u|\le(A+1)|t|$. Putting everything together, we have

$$
\begin{aligned}
&\frac{1}{2|t|}\int_{|s|\le|t|}\int_\Omega\left|f_j\circ\varphi_X\big(x,\tau(x,t)\big)-f_j\circ\varphi_X(x,s)\right|^2dx\,ds\\
&\qquad=\frac{1}{2|t|}\int_{|u|\le(A+1)|t|}\int_\Omega\left|f_j\circ\varphi_X(y,u)-f_j(y)\right|^2\left|\det J_{x,s}\Phi_t|_{\Phi_t^{-1}(y,u)}\right|^{-1}dy\,du\\
&\qquad\lesssim\sup_{|u|\le(A+1)|t|}\|\exp(uX)f_j-f_j\|_2^2\\
&\qquad\lesssim|t|^\alpha\|f_j\|_{X,\alpha}\ .
\end{aligned}
$$

Summing over j and noticing that $\|f\|_{X,\alpha}\sim\sum_j\|f_j\|_{X,\alpha}$, we obtain the same inequality for f. This and (5.2.7) imply (5.2.6). $\square$

We show next that, on compact sets, the norm (5.2.1) is controlled by the Λ^2_α-norm. This will follow from the following general lemma, that shall be used again later on.

Lemma 5.2.3. *Let K' be a compact neighborhood of K in Ω, and let $\varphi(x,t)$ be defined on $K'\times[0,\delta]\subset\Omega\times\mathbb{R}$, with values in Ω, such that*

(i) *φ is C^∞ in x, and all its derivatives are continuous in t;*

(ii) $\varphi(x,t) - x = O(t^\mu)$ *as* $t \to 0$ *uniformly in* x, *for some* $\mu \in \mathbb{R}^+$.

Then there is a constant $C = C(K, K', \varphi)$ *such that, for* $\alpha \in (0,1)$, t *small enough and* $f \in L^2(\Omega)$ *supported on* K,

$$\int_\Omega |f \circ \varphi(x,t) - f(x)|^2\,dx \le Ct^{2\mu\alpha} \|f\|^2_{\Lambda^2_\alpha} \ .$$

PROOF. In analogy with the previous proof, we first notice that, taking t small and $h \in \mathbb{R}^n$ with $|h| \le t^\mu$,

$$\int_\Omega |f \circ \varphi(x,t) - f(x)|^2\,dx \lesssim \int_\Omega |f \circ \varphi(x,t) - f(x-h)|^2\,dx + \|\tau_h f - f\|_2^2 \ ,$$

and therefore

$$\begin{aligned}\int_\Omega & |f \circ \varphi(x,t) - f(x)|^2\,dx \\ &\lesssim t^{-\mu n} \int_{|h|<t^\mu} \int_\Omega |f \circ \varphi(x,t) - f(x-h)|^2\,dx + t^{2\mu\alpha} \|f\|^2_{\Lambda^2_\alpha} \ .\end{aligned}$$

For t fixed, we set $x - h = y$ and $\varphi(x,t) = y + u$. If $\Phi_t(x,h) = \big(x-h, \varphi(x,t) - x + h\big)$, we have

$$J_{x,h}\Phi_t(x,h) = \begin{pmatrix} I & -I \\ J_x\varphi(x,t) - I & I \end{pmatrix} \ .$$

Since $\varphi(x,0) = x$, $J_x\varphi(x,0) = I$ and therefore $\det J_{x,h}\Phi_0(x,h) = 1$. The proof can be concluded as before, once we observe that

$$|u| \le |\varphi(x,t) - x| + |h| \lesssim t^\mu \ . \qquad \Box$$

$\Box$

PROPOSITION 5.2.4. *Given a smooth vector field* X *on* Ω *and* K *compact in* Ω, *there is a constant* $C = C(X,K)$ *such that, for every* $f \in L^2(\Omega)$ *supported on* K,

$$\|f\|_{X,\alpha} \le C\|f\|_{\Lambda^2_\alpha} \ .$$

PROOF. Apply Lemma (5.2.3) with $\mu = 1$ and $\varphi = \varphi_X$. $\Box$

3. Variations on the Baker-Campbell-Hausdorff formula

At this stage we introduce an algebraic formalism. We work with "abstract" non-commuting indeterminates $x_1, dots, x_k$. We are allowed to construct formal power series in $x_1, \dots, x_k$, keeping track of the order of factors in each monomial. We set $[x,y] = xy - yx$. Formal power series can be added amd multiplied among themeselves, and multiplied by scalars[4]. We denote as $\mathbb{R}[\![x_1, \dots, x_k]\!]$, or briefly as A, the algebra of formal power series generated by $x_1, \dots, x_k$. For any $a \in A$, the exponential series

$$e^a = \sum_{n \ge 0} \frac{1}{n!} a^n$$

is well defined and gives an invertible element of A, with $(e^a)^{-1} = e^{-a}$. The Baker-Campbell-Hausdorff formula says that, for $a, b \in A$,

$$e^a e^b = e^{s(a,b)} \ , \tag{5.3.1}$$

where s is itself a formal power series, containing only commutators

$$\begin{aligned} s(a,b) &= \sum_{j,k \ge 1} z_{j,k}(a,b) \\ &= a + b + \frac{1}{2}[a,b] + \frac{1}{12}\big[[a,b],b\big] - \frac{1}{12}\big[[a,b],a\big] + \cdots \ , \end{aligned} \tag{5.3.2}$$

[4]The scalar field can be $\mathbb{R}$ or $\mathbb{C}$, but consider real scalars here.

where each $z_{j,k}(a,b)$ denotes a *fixed* linear combination of iterated commutators of a and b, each containing a j times and b k times. When $(a,b)=(x,y)$, we shall usually drop the arguments (x,y), and write $z_{j,k}$ as well as w_j, r_N etc. to denote expression involving commutators of x and y in *fixed* finite or infinite combinations.

A first instance of identity derived from (5.3.2) is the following conjugation formula:

$$e^x e^y e^{-x} = e^{y+[x,y]+\sum_{n\geq 2} b_n} \ , \tag{5.3.3}$$

where each b_n is a linear combination of commutators of x and y of order n (see (1.2.7)).

We shall perform some manipulations of this formula aimed to obtain expressions of $e^{[x,y]}$ and e^{x+y} as products of exponentials, where each exponent contains either x or y alone, or, alternatively, commutators of x and y of a sufficiently high order.

We first reduce the problem concerning the sum to a problem concerning commutators [5].

Lemma 5.3.1. *There are two sequences of elements w_j and r_N of A such that*

(i) *each w_j is a (finite) linear combination of commutators of x and y of order j;*
(ii) *$r_N=\sum_{k=N+1}^{\infty} u_k^{(N)}$, where each $u_k^{(N)}$ is a linear combination of commutators of x and y of order k;*
(iii) *for every N,*

$$e^{x+y} = e^x e^y e^{w_2}\cdots e^{w_N} e^{r_N} \ . \tag{5.3.4}$$

Proof. Calling $z_n=\sum_{j+k=n} z_{j,k}$, we have

$$e^x e^y = e^{\sum_{n\geq 1} z_n} \ ,$$

and each z_n is a linear combination of commutators of x and y of order n.

For $N=1$, multiply both sides sides by e^{-x-y} on the left. By the same formula,

$$\begin{aligned} e^{-x-y}e^x e^y &= e^{-x-y}e^{x+y+\sum_{n\geq 2} z_n} \\ &= e^{\sum_{n\geq 2} z'_n} \ , \end{aligned}$$

where each z'_n is again a linear combination of commutators of x and y of order n. Therefore

$$e^{x+y} = e^x e^y e^{-\sum_{n\geq 2} z'_n} \ .$$

This proves the case $N=1$.

Inductively, suppose that (5.3.4) holds for a given N, with r_N as in (ii). Then there are elements $u_k^{(N+1)}$ as in (ii) and such that

$$\begin{aligned} e^{-u_{N+1}^{(N)}}e^{r_N} &= e^{-u_{N+1}^{(N)}}e^{\sum_{k\geq N+1} u_k^{(N)}} \\ &= e^{\sum_{k\geq N+2} u_k^{(N+1)}} \ . \end{aligned}$$

This concludes the proof. □

For exponentials of commutators, we have the following initial result. We change the notation slightly, writing x_1, x_2 in place of x, y.

Lemma 5.3.2. *Let $c_p=\big[\cdots[x_1,x_2],\ldots,x_p\big]$, with $p\geq 2$. Then*

$$e^{c_p} = e^{\pm x_{i_1}} e^{\pm x_{i_2}} \cdots e^{\pm x_{i_q}} e^{\sum_{n\geq p+1} v_n^{(p)}} \ , \tag{5.3.5}$$

where $i_1,\ldots,i_q\in\{1,\ldots,p\}$, $q=3\cdot 2^{p-1}-2$, and each $v_n^{(p)}$ a linear combination of commutators of $x_1,\ldots,x_p$ of order n.

---

[5] A simpler equivalent statement of Lemma 3.1 would consist in the following identity:

$$e^{x+y} = e^x e^y \prod_{j=2}^{\infty} e^{w_j} \ ,$$

with the w_j as in (i). For our purposes it is preferable to focus on the properties of the remainders r_N such that$e^{r_N} = \prod_{j=N+1}^{\infty} e^{w_j}$.

PROOF. From (5.3.3) we obtain that

$$e^{x_1}e^{x_2}e^{-x_1}e^{-x_2} = e^{x_2+[x_1,x_2]+\sum_{n\geq 3} b_n}e^{-x_2} = e^{[x_1,x_2]+\sum_{n\geq 3} b_n'} ,$$

where the b_n' have the same properties as the b_n. Therefore,

$$e^{-[x_1,x_2]}e^{x_1}e^{x_2}e^{-x_1}e^{-x_2} = e^{\sum_{n\geq 3} b_n''} ,$$

which easily gives (5.3.5) with $q=4$.

Inductively, assume that (5.3.5) holds up to some p. Here we explicitly write $v_n^{(p)}(x_1,x_2)$ to emphasize that each $v_n^{(p)}$ is a fixed expression in x_1,x_2. Then $c_{p+1}=[c_p,x_{p+1}]$. We apply (5.3.5) with $p=2$ to obtain

$$e^{c_{p+1}} = e^{c_p}e^{x_{p+1}}e^{-c_p}e^{-x_{p+1}}e^{\sum_{n\geq 3} v_n^{(1)}(c_p,x_2)} .$$

Since $v_n^{(1)}(c_p,x_2)$ is a commutator of c_p and x_{p+1} of order n, it follows from Lemma 1.2.1, that $v_n^{(1)}(c_p,x_{p+1})$ is a linear combination of commutators of $x_1,\dots,x_{p+1}$ of order at least $p+n-1$. Hence

$$e^{c_{p+1}} = e^{c_p}e^{x_{p+1}}e^{-c_p}e^{-x_{p+1}}e^{\sum_{k\geq p+2} d_k(x_1,\dots,x_{p+1})} .$$

We then insert the expansion (5.3.5) of e^{c_p} and e^{-c_p}. If each expansion contains q factors equal to $e^{\pm x_j}$ for some j, we obtain altogether $2q+2$ factors of this kind, plus two factors $e^{\pm\sum_{n\geq p+1} v_n^{(p)}}$ placed among them.

We can then shift these extra terms to the far right using the conjugation formula

$$e^{-x_j}e^{a_{p+1}+\sum_{n\geq p+2} a_n}e^{x_j} = e^{a_{p+1}+\sum_{n\geq p+2} a_n'} ,$$

derived from (5.3.3) (notice that the leading term in the series remains unchanged). After doing so, we obtain that

$$\begin{aligned} e^{c_{p+1}} &= e^{\pm x_{i_1}}e^{\pm x_{i_2}}\cdots e^{\pm x_{i_{2q+2}}} \\ &\quad\times e^{v_{p+1}^{(p)}+\sum_{n\geq p+1}\alpha_n}e^{-v_{p+1}^{(p)}+\sum_{n\geq p+2}\beta_n}e^{\sum_{k\geq p+2} d_k} \\ &= e^{\pm x_{i_1}}e^{\pm x_{i_2}}\cdots e^{\pm x_{i_{2q+2}}}e^{\sum_{k\geq p+2} d_k'} , \end{aligned}$$

with $\alpha_n=\alpha_n(x_1,\dots,x_{p+1})$ etc. This gives (5.3.5). □

COROLLARY 5.3.3. *Let c_p be as above, and let $N\geq p+1$. Then there is $q=q(p,N)$ such that*

$$e^c = e^{\pm x_{i_1}}e^{\pm x_{i_2}}\cdots e^{\pm x_{i_q}}e^{\sum_{n\geq N} v_n} , \tag{5.3.6}$$

where $i_1,\dots,i_q\in\{1,\dots,p\}$, and each $v_n=v_n^{(p,N)}$ a linear combination of commutators of $x_1,\dots,x_p$ of order n.

PROOF. We use induction on N. The case $N=p+1$ is Lemma 5.3.2. Suppose (5.3.6) holds for N. We use the usual argument to extract the leading term v_N from the last exponent:

$$e^{-v_N}e^{\sum_{n\geq N} v_n} = e^{\sum_{n\geq N+1} v_n'} ,$$

so that

$$e^{c_p} = e^{\pm x_{i_1}}e^{\pm x_{i_2}}\cdots e^{\pm x_{i_q}}e^{v_N}e^{\sum_{n\geq N+1} v_n'} .$$

We write $v_N=\sum_{j=1}^r \lambda_j c_j$, where the c_j are commutators of $x_1,\dots,x_p$ of order N, and the coefficients λ_j do not depend on $x_1,\dots,x_p$. Making repeated use of (5.3.4) with $N=1$, we obtain that

$$e^{v_N} = e^{\lambda_1 c_1}\cdots e^{\lambda_r c_r}e^{\sum_{\ell\geq 2} u_\ell} ,$$

where each u_ℓ is a linear combination of commutators of $c_1,\dots,c_r$ of order ℓ, hence a linear combination of commutators of $x_1,\dots,x_p$ of order greater than or equal to $2N$. Hence,

$$e^{c_p} = e^{\pm x_{i_1}}e^{\pm x_{i_2}}\cdots e^{\pm x_{i_q}}e^{\lambda_1 c_1}\cdots e^{\lambda_r c_r}e^{\sum_{n\geq N+1} v_n''} .$$

We next expand each factor $e^{\lambda_j c_j}$ according to Lemma 5.3.2 to derive the conclusion. □

In the same way we prove the following improvement of Lemma 5.3.1.

COROLLARY 5.3.4. *For every $N, p \geq 2$ there are an integer $q = q(N,p)$ and a sequence $\{\omega_n^{(N)}\}_{n \geq N}$ of linear combinations of commutators of $x_1, \dots, x_p$ of order n such that*

$$e^{x_1+\cdots+x_p} = e^{\pm x_{i_1}} e^{\pm x_{i_2}} \cdots e^{\pm x_{i_q}} e^{\sum_{n\geq N} \omega_n^{(N)}} ,$$

with $i_1, \dots, i_q \in \{1, \dots, p\}$.

4. Operations on vector fields and Lipschitz norms

If X, Y are two commuting vector fields on Ω, it is quite easy to establish the inequality

$$\|f\|_{X+Y,\alpha} \leq C(K)\big(\|f\|_{X,\alpha} + \|f\|_{Y,\alpha}\big)$$

when f is supported on a compact subset K. It is sufficient to observe that, by (5.2.2),

$$\begin{aligned}\big\| \exp\big(t(X+Y)\big)f - f\big\|_2 &= \|\exp(tX)\exp(tY)f - f\|_2 \\ &\leq \big\|\exp(tX)\big(\exp(tY)f - f\big)\big\|_2 + \|\exp(tX)f - f\|_2 \\ &\leq C(K)\|\exp(tY)f - f\|_2 + \|\exp(tX)f - f\|_2 \\ &\leq C(K)|t|^\alpha\big(\|f\|_{X,\alpha} + \|f\|_{Y,\alpha}\big) .\end{aligned}$$

The situation is much more complicated if X and Y do not commute. In this case we are also interested in norms like $\|f\|_{[X,Y],\alpha}$ and similar with higher-order commutators.

We shall use the formulas obtained in the previous section. When applied to exponentials of vector fields, these formulas do not make sense as such, because they contain infinite series that do not converge in general.

At each stage, we must replace infinite sums by truncations of sufficiently high order and introduce remainder terms, in complete analogy with the use of Taylor expansions with non-analytic functions. The starting point is the formulation of the Baker-Campbell-Hausdorff formula given in Theorem 2.2 in Chapter I. We must also introduce a parameter t tending to zero to which compare the remainders. An indeterminate x is then replaced by tX, with X a given vector field. Notice that a commutator of order $p+1$ is then replaced by the same commutator of the involved vector fields, multiplied by t^{p+1}.

We then obtain the following formulations of Corollaries 5.3.3 and 5.3.4.

COROLLARY 5.4.1. *Let $X_1, \dots, X_p$ be smooth vector fields on Ω.*

(i) *If $C_p = \big[\cdots[X_1, X_2], \dots, X_p\big]$ and $N \geq p+1$, let q, $i_1, \dots, i_q \in \{1, \dots, p\}$ be as in Corollary 5.3.3. For any compact $K \subset \Omega$ there is $\delta > 0$ such that, for $|t| < \delta$ and $f \in \mathcal{D}(\Omega)$ supported on K,*

$$\exp(t^pW)f(x) = \exp(\pm tX_{i_1})\exp(\pm tX_{i_2})\cdots\exp(\pm tX_{i_q})f(x) + O(t^N) , \tag{5.4.1}$$

uniformly in x.

(ii) *Given $N \geq 2$, let q, $i_1, \dots, i_q \in \{1, \dots, p\}$ be as in 5.3.4. For any compact $K \subset \Omega$ there is $\delta > 0$ such that, for $|t| < \delta$ and $f \in \mathcal{D}(\Omega)$ supported on K,*

$$\exp\big(t(X_1 + \cdots + X_p)\big)f(x) = \exp(\pm tX_{i_1})\exp(\pm tX_{i_2})\cdots\exp(\pm tX_{i_q})f(x) + O(t^N) . \tag{5.4.2}$$

Before applying these formulas, we must add some considerations on the transformations of Ω induced by the corresponding compositions of flows.

Let us consider (5.4.1) on a compact set $K' \supset K$ in Ω. It can be rewritten as

$$\exp(\mp tX_{i_q})\cdots\exp(\mp tX_{i_2})\exp(\mp tX_{i_1})\exp(t^pC_p)f(x) = f(x) + O(t^N) . \tag{5.4.3}$$

Composition of the flows in each factor gives rise to a function $\varphi(x,t)$, smooth on $K' \times (-\delta, \delta)$ and with values in Ω, such that

$$\exp(\mp tX_{i_q})\cdots\exp(\mp tX_{i_2})\exp(\mp tX_{i_!})\exp(t^pW)f(x) = f\big(\varphi(x,t)\big) . \tag{5.4.4}$$

Take $f \in \mathcal{D}(\Omega)$, supported on K' and equal to a coordinate function x_j on K. Plugging it into (5.4.3), we see that the j-th component φ_j of φ satisfies

$$\varphi_j(x,t) = x_j + O(t^N) \ . \tag{5.4.5}$$

This shows that φ satisfies the hypotheses of Lemma 5.2.3.

THEOREM 5.4.2. *Let X_1, X_2 be smooth vector fields on Ω. Given $\sigma \in (0,1)$, we have the following estimates, for every $\alpha < 1$ and $f \in L^2(\Omega)$ supported on K:*

(i) *if W is a commutator of X_1, X_2 of order p,*

$$\|f\|_{W,\frac{\alpha}{p}} \le C_{p,\sigma}\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big) + C'\|f\|_{\Lambda^2_\sigma} \ ; \tag{5.4.6}$$

(ii)

$$\|f\|_{X_1+X_2,\alpha} \le C_\sigma\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big) + C'\|f\|_{\Lambda^2_\sigma} \ . \tag{5.4.7}$$

The constants $C_{p,\sigma}$ in (5.4.6) *and C_σ in* (5.4.7) *do not depend on α or on the specific vector fields X_1, X_2.*

PROOF. We prove only (5.4.6) only. Take N such that $N\sigma \ge 1$, and let φ the function in (5.4.4); for $|t|$ small,

$$\begin{aligned}
\exp(t^pW)f(x) - f(x)& \\
&= \exp(\pm tX_{i_1})\cdots\exp(\pm tX_{i_{q-1}})\exp(\pm tX_{i_q})f\big(\varphi(x,t)\big) - f(x)\\
&= \exp(\pm tX_{i_1})\cdots\exp(\pm tX_{i_{q-1}})\exp(\pm tX_{i_q})\Big(f\big(\varphi(x,t)\big) - f(x)\Big)\\
&\quad + \exp(\pm tX_{i_1})\cdots\exp(\pm tX_{i_{q-1}})\big(\exp(\pm tX_{i_q})f(x) - f(x)\big)\\
&\quad + \cdots\\
&\quad + \exp(\pm tX_{i_1})f(x) - f(x) \ .
\end{aligned}$$

Choosing $|t| < \delta$ small enough and using (5.2.2), we obtain that

$$\|\exp(t^pW)f - f\|_2 \le 2q\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big)|t|^\alpha + C_N\|f\|_{\Lambda^2_\sigma}|t|^{N\sigma} \ ,$$

$q = q(p,N)$ being as in Corollary 5.3.3. Since $|t|$ is small, $C_N|t|^{N\sigma} \le |t|^\alpha$. This shows that, for $s > 0$ and small enough,

$$\|\exp(sW)f - f\|_2 \le 2q\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big)s^{\frac{\alpha}{p}} + \|f\|_{\Lambda^2_\sigma}s^{\frac{\alpha}{p}} \ .$$

For $s < 0$, it is sufficient to notice that $-W$ is also a commutator of X_1, X_2 of order p (just interchange X_1 and X_2 in the innermost Lie bracket). □

We can proceed now to the proof the main result about Lipschitz norms, which will be preceded by a lemma.

LEMMA 5.4.3. *Let $0 < \beta < \alpha < 1$. Given $\varepsilon > 0$, there is a constant C_ε such that*

$$\|f\|_{\Lambda^2_\beta} \le C_\varepsilon\|f\|_2 + \varepsilon\|f\|_{\Lambda^2_\alpha} \ .$$

PROOF. Fix $a > 0$. If $0 < \delta < a$,

$$\begin{aligned}
\|f\|_{\Lambda^2_\beta} &\le \|f\|_2 + \sup_{\delta \le |h| < a}|h|^{-\beta}\|\tau_hf - f\|_2 + \sup_{|h|<\delta}|h|^{-\beta}\|\tau_hf - f\|_2\\
&\le (1 + 2\delta^{-\beta})\|f\|_2 + \delta^{\alpha-\beta}\sup_{|h|<\delta}|h|^{-\alpha}\|\tau_hf - f\|_2\\
&\le (1 + 2\delta^{-\beta})\|f\|_2 + \delta^{\alpha-\beta}\|f\|_{\Lambda^2_\alpha} \ .
\end{aligned}$$

It is then sufficient to take $\delta = \varepsilon^{\frac{1}{\alpha-\beta}}$. □

THEOREM 5.4.4. *Let K be a compact subset of Ω, and suppose that the smooth vector fields $X_1, \ldots, X_k$, together with their iterated commutators X^I with $|I| \leq m$, span $\mathbb{R}^n$ at each point of K. Then, for every $\alpha \in (0,1)$, there is a constant C_α such that, for every $f \in L^2(\Omega)$ with support contained in K,*

$$\|f\|_{\Lambda^2_{\frac{\alpha}{m}}} \leq C_\alpha \sum_{j=1}^{k} \|f\|_{X_j,\alpha} \ . \tag{5.4.8}$$

PROOF. Given $x \in K$, there are multi-indices $I_1, \ldots, I_n$, with $|I_k| \leq m$ for each k, such that the vectors $\{X_x^{I_k}\}$ form a basis of $\mathbb{R}^n$. By continuity, the same holds for the vectors $\{X_{x'}^{I_k}\}$ for each x' in a neighborhood U_x of x. By the inverse function theorem, there are smooth functions $\eta_{j,k}$ on U_x such that

$$\partial_{x_j} = \sum_{k=1}^{n} \eta_{j,k}(x) X^{I_k} \ ,$$

on U_x, for $j = 1, \ldots, n$.

Covering K with a finite number of such neighborhoods, and with the aid of a subordinated partition of unity, we then conclude that there are smooth functions $\eta_{j,I}$ on a neighborhood U of K such that

$$\partial_{x_j} = \sum_{|I| \leq m} \eta_{j,I}(x) X^I \ ,$$

on U, for $j = 1, \ldots, n$. By restricting Ω if necessary, we can as well assume that $U = \Omega$.

Denote by E_j the coordinate vector field ∂_{x_j}. Applying (5.4.7) and Proposition 5.2.2, we obtain that

$$\begin{aligned}\|f\|_{E_j,\frac{\alpha}{m}} &\leq C_\sigma \sum_{|I| \leq m} \|f\|_{\eta_{j,I}X^I,\frac{\alpha}{m}} + C'\|f\|_{\Lambda^2_\sigma} \\ &\leq C_{X_1,\ldots,X_k,\sigma} \sum_{|I| \leq m} \|f\|_{X^I,\frac{\alpha}{m}} + C'\|f\|_{\Lambda^2_\sigma} \ ,\end{aligned}$$

for every $\sigma > 0$. Since $|I| \leq m$, we also have, by (5.4.6)

$$\|f\|_{X^I,\frac{\alpha}{m}} \leq C\|f\|_{X^I,\frac{\alpha}{|I|}} \leq C_{m,\sigma}\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big) + C'\|f\|_{\Lambda^2_\sigma} \ .$$

Since

$$\|f\|_{\Lambda^2_{\frac{\alpha}{m}}} \sim \sum_{j=1}^{n} \|f\|_{E_j,\frac{\alpha}{m}} \ ,$$

we conclude that

$$\|f\|_{\Lambda^2_{\frac{\alpha}{m}}} \leq C_{X_1,\ldots,X_k,\sigma}\big(\|f\|_{X_1,\alpha} + \|f\|_{X_2,\alpha}\big) + C'\|f\|_{\Lambda^2_\sigma} \ .$$

We finally fix $\sigma < \frac{\alpha}{m}$ and apply Lemma 4.3, with $\beta = \sigma$, α replaced by $\frac{\alpha}{m}$ and $\varepsilon = \frac{1}{2C'}$, to majorize the last term. □

Theorem 1.5.7 is now an immediate consequence of this result, once we observe that

$$\|f\|_{X_j,\alpha} \leq C\big(\|f\|_2 + \|X_j f\|_2\big) \ ,$$

for any $\alpha < 1$. This follows from the mean value theorem, because

$$\big|\exp(tX_j)f(x) - f(x)\big| \leq \int_0^t \big|\exp(sX_j)X_j f(x)\big|\, ds \ ,$$

so that, for $|t| \leq a$,

$$\big\|\exp(tX_j)f - f\big\|_2 \leq Ca\|X_j f\|_2 \ .$$

Given $s < \frac{1}{m}$, choose α such that $s < \frac{\alpha}{m} < \frac{1}{m}$, and apply Corollary 5.1.4.

5. Hypoellipticity of sub-Laplacians: preliminaries

We adopt the following notation. If D is a differential operator with smooth coefficients in Ω, denote by D^* the *formal adjoint* of D, i.e. the differential operator such that

$$\int_\Omega D^* f(x)\overline{g(x)}\,dx = \int_\Omega f(x)\overline{Dg(x)}\,dx \ ,$$

for $f,g\in\mathcal{D}(\Omega)$.

Suppose that $\mathcal{X}=\{X_1,\dots X_k\}$ is a generating system of vector fields on Ω. Let

$$L = -\sum_{j=1}^k X_j^2 \ .$$

Introduce the norm

$$|||f|||_{\mathcal{X}} = \Big(\|f\|_2^2 + \sum_{j=1}^k \|X_j f\|_2^2\Big)^{\frac12}$$

on $\mathcal{D}(\Omega)$. For Ω' open in Ω, we consider also the dual norm

$$|||u|||'_{\mathcal{X},\Omega'} = \sup_{f\in\mathcal{D}(\Omega'):|||f|||_{\mathcal{X}}\le 1} \big|\langle u,f\rangle\big|$$

on $\mathcal{D}'(\Omega')$. Clearly,

$$|||f|||'_{\mathcal{X},\Omega'} \le \|f\|_2 \le |||f|||_{\mathcal{X}} \ , \tag{5.5.1}$$

whenever applicable.

The reason for introducing these norms is the following.

First of all, notice that, by definition of hypoellipticity, L is hypoelliptic on Ω if and only if it is hypoelliptic on every relatively compact subdomain. Let Ω' be one such subdomain.

If

$$X_j = \sum_{\ell=1}^n a_{j,\ell}(x)\partial_{x_\ell} \ ,$$

then, for functions $f,g\in\mathcal{D}(\Omega')$,

$$\begin{aligned}
\int_{\Omega'} X_j f(x)\overline{g(x)}\,dx &= \sum_{\ell=1}^n \int_{\Omega'} a_{j,\ell}(x)\partial_{x_\ell} f(x)\overline{g(x)}\,dx \\
&= -\sum_{\ell=1}^n \int_{\Omega'} f(x)\overline{\partial_{x_\ell}\big(a_{j,\ell}(x)g(x)\big)}\,dx \\
&= -\int_{\Omega'} f(x)\overline{\big(X_j + b_j(x)\big)g(x)}\,dx \ ,
\end{aligned}$$

with $b_j = \sum_{\ell=1}^n \partial_{x_\ell} a_{j,\ell}$. In other words, $X_j^* = -X_j - b_j$. Notice that the functions $a_{j,\ell}$ and b_j are real valued.

Next, denote by $\langle\ ,\ \rangle$ the inner product in $L^2(\Omega')$. For $f\in\mathcal{D}(\Omega')$,

$$\begin{aligned}
\langle Lf,f\rangle &= \sum_{j=1}^k \langle X_j f, (X_j+b_j)f\rangle \\
&= \sum_{j=1}^k \|X_j f\|_2^2 + \sum_{j=1}^k \langle X_j f, b_j f\rangle \ .
\end{aligned}$$

Hence,

$$\Re\mathrm{e}\langle Lf,f\rangle = \sum_{j=1}^k \|X_jf\|_2^2 + \sum_{j=1}^k \Re\mathrm{e}\langle X_jf,b_jf\rangle\ .$$

Now,

$$\begin{aligned}\Re\mathrm{e}\langle X_jf,b_jf\rangle &= \frac12\big(\langle X_jf,b_jf\rangle + \langle b_jf,X_jf\rangle\big)\\ &= \frac12\big(\langle X_jf,b_jf\rangle - \langle (X_j+b_j)b_jf,f\rangle\big)\\ &= -\frac12\langle (X_jb_j+b_j^2)f,f\rangle\\ &= -\frac12\int_{\Omega'}(b_j^2+X_jb_j)|f|^2\,dx\ .\end{aligned}$$

It follows that

$$\begin{aligned}|\!|\!| f|\!|\!|_{\mathcal X}^2 &= \|f\|_2^2 + \Re\mathrm{e}\langle Lf,f\rangle - \sum_{j=1}^k \Re\mathrm{e}\langle X_jf,b_jf\rangle\\ &= \frac12\int_{\Omega'}(2+b_j^2+X_jb_j)|f|^2\,dx + \Re\mathrm{e}\langle Lf,f\rangle\\ &\le C\|f\|_2^2 + |\!|\!| Lf|\!|\!|'_{\mathcal X,\Omega'}|\!|\!| f|\!|\!|_{\mathcal X}\ .\end{aligned}$$

Using the inequality $\|f\|_2^2 \le \|f\|_2|\!|\!| f|\!|\!|_{\mathcal X}$ we can simplify by a factor $|\!|\!| f|\!|\!|_{\mathcal X}$, obtaining that

$$|\!|\!| f|\!|\!|_{\mathcal X} \le C\|f\|_2 + |\!|\!| Lf|\!|\!|'_{\mathcal X,\Omega'}\ .$$

Observe that the hypotheses of Theorem 5.4.4 are satisfied on all of Ω' with the same $m=m(\Omega')$ and the same constants. By Theorem 1.5.7, it follows that, for $f\in\mathcal D(\Omega')$ and $s<\frac1m$,

$$\|f\|_{(s)} \le C_s\big(\|f\|_2 + |\!|\!| Lf|\!|\!|'_{\mathcal X,\Omega'}\big)\ . \tag{5.5.2}$$

The hardest part of the proof will be the extension of (5.5.2) to general L^2-functions on Ω' with compact support. This will be done in Section 7. In Section 6 we introduce calculus and estimates with Bessel potentials.

Once this is done, we shall prove a "bootstrapping" argument which extends this implication to Sobolev norms of arbitrary orders. This will be done in Section 8, and it will easily lead us to the conclusion of the proof.

6. Bessel potentials and pseudo-differential operators

For $\gamma\in\mathbb R$, let K_γ be the distribution defined as the inverse Fourier transform of $(1+|\xi|^2)^\gamma$, i.e. such that

$$\langle K_\gamma,\varphi\rangle = \frac{1}{(2\pi)^n}\int_{\mathbb R^n}(1+|\xi|^2)^\gamma\hat\varphi(\xi)\,d\xi\ ,$$

for $\varphi\in\mathcal S(\mathbb R^n)$. The convolution operator

$$\varphi\longmapsto K_\gamma*\varphi = \mathcal F^{-1}\big((1+|\xi|^2)^\gamma\hat\varphi(\xi)\big) \tag{5.6.1}$$

maps $\mathcal S(\mathbb R^n)$ continuously onto itself. By duality, it also maps $\mathcal S'(\mathbb R^n)$ onto itself continuously.

Notice that $K_\gamma*K_{\gamma'}*\varphi = K_{\gamma+\gamma'}\varphi$ and that, for $j\in\mathbb N$,

$$K_j*\varphi = (1-\Delta)^j\varphi\ .$$

For this reason the operator (5.6.1) is denoted by $(1-\Delta)^\gamma$ for arbitrary[6] γ. The identity

$$\big((1-\Delta)^\gamma\varphi\big)\hat{}\,(\xi) = (1+|\xi|^2)^\gamma\hat\varphi(\xi)$$

[6]All we said is true also for complex exponents, but we shall not need them here.

extends from integer to real values of γ.

The operator $(1-\Delta)^{-\frac{s}{2}}$ is called the *Bessel potential of order s.*

The following facts follow directly from the definition of Sobolev spaces and the Plancherel formula.

LEMMA 5.6.1. *The Sobolev space $H^s(\mathbb{R}^n)$ consists of those $f \in \mathcal{S}'(\mathbb{R}^n)$ such that $(1-\Delta)^{\frac{s}{2}} f \in L^2(\mathbb{R}^n)$. In other words,*

$$H^s(\mathbb{R}^n) = (1-\Delta)^{-\frac{s}{2}}\big(L^2(\mathbb{R}^n)\big) \ .$$

If $f \in L^2(\mathbb{R}^n)$ and $0 \le |\alpha| \le s$,

$$\left\|(1-\Delta)^{-\frac{s}{2}}\partial^\alpha f\right\|_2 \le \|f\|_2 \ . \tag{5.6.2}$$

In particular, $(1-\Delta)^{-\frac{s}{2}}$ is bounded on $L^2(\mathbb{R}^n)$ for $s \ge 0$.

The properties of Bessel potentials that we shall need can be better understood in the more general context of *pseudo-differential operators* ("ψdo" in short). More precisely, we will consider "localized" versions $M_\chi(1-\Delta)^{-\frac{s}{2}}$ of the Bessel potential, where M_χ is the multiplication operator by a function $\chi \in \mathcal{D}(\Omega)$. Localized Bessel potentials of order s are examples of pseudo-differential operators of order s. In order to understand the definition of a ψdo, consider first a differential operator $D = \sum_{|\alpha|\le m} a_\alpha(x)\partial^\alpha$ of order $\le m$ with coefficients $a_\alpha \in \mathcal{S}(\mathbb{R}^n)$. Passing to the Fourier transform, we have, for $f \in \mathcal{S}(\mathbb{R}^n)$,

$$\begin{aligned}\widehat{Df}(\xi) &= \sum_{|\alpha|\le m} i^{|\alpha|}\widehat{a_\alpha} * (\xi^\alpha \hat{f})(\xi)\\ &= \int_{\mathbb{R}^n}\sum_{|\alpha|\le m} i^{|\alpha|}\widehat{a_\alpha}(\xi-\eta)\eta^\alpha \hat{f}(\eta)\,d\eta \ .\end{aligned}$$

Denote, as usual, by $\mathcal{F}$ the Fourier transform, and set

$$k(\xi,\eta) = \sum_{|\alpha|\le m} i^{|\alpha|}\widehat{a_\alpha}(\xi)\eta^\alpha \ .$$

We have shown that D is the conjugation $D = \mathcal{F}^{-1}T_k\mathcal{F}$, where T_k is the integral operator with kernel $k(\xi-\eta,\eta)$. It is easy to verify that k satisfies the estimates

$$\left|\partial_\eta^\beta k(\xi,\eta)\right| \le C_{\beta,N}(1+|\xi|)^{-N}(1+|\eta|)^{m-|\beta|} \ , \tag{5.6.3}$$

for every multi-index β and every $N > 0$.

DEFINITION. *We call special ψdo of order $m \in \mathbb{R}$ an operator if $D = \mathcal{F}^{-1}T_k\mathcal{F}$, where*

$$T_k f(\xi) = \int_{\mathbb{R}^n} k(\xi-\eta,\eta) f(\eta)\,d\eta \ , \tag{5.6.4}$$

and k is a smooth function on $\mathbb{R}^n \times \mathbb{R}^n$ satisfying (5.6.3) for every β and N.

This definition is *ad hoc*, in the sense that our hypotheses are stronger than in the general theory. [7] The following remarks are rather obvious.

(1) Differential operators with coefficients in $\mathcal{S}(\mathbb{R}^n)$ are special ψdo of the same order.
(2) For $\chi \in \mathcal{S}(\mathbb{R}^n)$ and $\gamma \in \mathbb{R}$, $M_\chi(1-\Delta)^\gamma$ is a special ψdo of order 2γ.

---

[7] One usually defines a ψdo as an operator of ther form

$$Df(x) = (2\pi)^{-n}\int_{\mathbb{R}^n} a(x,\eta)\hat{f}(\eta)e^{i\eta\cdot x}\,d\eta \ ,$$

where the function a (called the *symbol* of D) satisfies the condition

$$\left|\partial_x^\alpha\partial_\eta^\beta a(x,\eta)\right| \le C_{\alpha,\beta}(1+|\eta|)^{m-|\beta|} \ ,$$

for every pair of multi-indices α,β. The kernel k appearing in our definition is

$$k(\xi,\eta) = \int_{\mathbb{R}^n} a(x,\eta)e^{-ix\cdot\xi}\,dx \ .$$

(3) As a particular case of the previous two, pointwise multiplication by a Schwartz function is a special ψdo of order 0.
(4) If k satisfies (5.6.3) for some m, then $\partial_\eta^\beta k$ satisfies (5.6.3) with m replaced by $m-|\beta|$.

Lemma 5.6.2. *Let D be a special ψdo of order m. Then D maps $\mathcal{S}(\mathbb{R}^n)$ into $\mathcal{S}'(\mathbb{R}^n)$ continuously, and its adjoint $D^*:\mathcal{S}(\mathbb{R}^n)\to\mathcal{S}'(\mathbb{R}^n)$, defined by the condition*[8]

$$\langle D^*f,g\rangle=\langle f,Dg\rangle$$

for $f,g\in\mathcal{S}(\mathbb{R}^n)$ is also a special ψdo of order m.

Proof. Let $D=\mathcal{F}^{-1}T_k\mathcal{F}$, with k as in (5.6.4), Because the Fourier transform is a homeomorphism of $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$, it is sufficient to prove that T_k is continuous from $\mathcal{S}(\mathbb{R}^n)$ to $\mathcal{S}'(\mathbb{R}^n)$. But, for $N>n$,

$$\begin{aligned}|T_kf(\xi)|&\le C_N\int_{\mathbb{R}^n}(1+|\xi-\eta|)^{-N}(1+|\eta|)^m|f(\eta)|\,d\eta\\&\le C_N\sup_{\eta\in\mathbb{R}^n}(1+|\eta|)^m|f(\eta)|\int_{\mathbb{R}^n}(1+|\tau|)^{-N}\,d\tau\\&\le C'_N\sup_{\eta\in\mathbb{R}^n}(1+|\eta|)^m|f(\eta)|\ ,\end{aligned}$$

where the right-hand side is one of the norms defining the Schwartz topology. Therefore, T_k is continuous from $\mathcal{S}(\mathbb{R}^n)$ to $L^\infty(\mathbb{R}^n)$. Since the inclusion of $L^\infty(\mathbb{R}^n)$ into $\mathcal{S}'(\mathbb{R}^n)$ is also continuous, we have proved the first statement.

Using the extended Parseval identity $\langle\Phi,f\rangle=(2\pi)^{-n}\langle\hat\Phi,\hat f\rangle$ and setting

$$k^*(\xi,\eta)=\overline{k(-\xi,\eta+\xi)}\ ,$$

we have

$$\begin{aligned}\langle D'f,g\rangle&=(2\pi)^{-n}\langle T_k\hat f,\hat g\rangle\\&=(2\pi)^{-n}\int_{\mathbb{R}^n}\int_{\mathbb{R}^n}k(\xi-\eta,\eta)\hat f(\eta)\,d\eta\,\overline{\hat g(\xi)}\,d\xi\\&=(2\pi)^{-n}\int_{\mathbb{R}^n}\hat f(\eta)\int_{\mathbb{R}^n}k(\xi-\eta,\eta)\overline{\hat g(\xi)}\,d\xi\,d\eta\\&=(2\pi)^{-n}\int_{\mathbb{R}^n}\hat f(\eta)\int_{\mathbb{R}^n}\overline{k^*(\eta-\xi,\xi)\hat g(\xi)}\,d\xi\,d\eta\\&=(2\pi)^{-n}\langle\hat f,T_{k'}\mathcal{F}^{-1}g\rangle\\&=\langle f,\mathcal{F}^{-1}T_{k^*}\mathcal{F}g\rangle\ .\end{aligned}$$

Then for every N,

$$|\partial_\eta^\beta k^*(\xi,\eta)|\le C_N(1+|\xi|)^{-N}(1+|\eta+\xi|)^{m-|\beta|}\ .$$

If $|\xi|\le\frac{|\eta|}{2}$, then $|\eta+\xi|\sim|\eta|$, and $(1+|\eta+\xi|)^m\sim(1+|\eta|)^m$ for m both positive and negative.
If $|\xi|>\frac{|\eta|}{2}$ and $m-|\beta|\ge0$,

$$\begin{aligned}(1+|\xi|)^{-N}(1+|\eta+\xi|)^{m-|\beta|}&\le(1+|\xi|)^{-N}(1+|\eta|+|\xi|)^{m-|\beta|}\\&\le3^{m-|\beta|}(1+|\xi|)^{-N+m-|\beta|}\\&\le3^{m-|\beta|}(1+|\xi|)^{-N+m-|\beta|}(1+|\eta|)^{m-|\beta|}\ .\end{aligned}$$

[8] We use here the sesquilinear pairing $\langle\Phi,f\rangle=\overline{\langle f,\Phi\rangle}=\Phi(\bar f)$ between a distribution Φ and a function f.

If $|\xi| > \frac{|\eta|}{2}$ and $m - |\beta| < 0$,

$$\begin{aligned}(1+|\xi|)^{-N}(1+|\eta+\xi|)^{m-|\beta|} &\leq (1+|\xi|)^{-N} \\ &\leq (1+|\xi|)^{-N-m+|\beta|}\Big(1+\frac{|\eta|}{2}\Big)^{m-|\beta|} \\ &\leq 2^{-m+|\beta|}(1+|\xi|)^{-N+m-|\beta|}(1+|\eta|)^{m-|\beta|} \ .\end{aligned}$$

Since N is arbitrary, k^* satisfies (5.6.3). ☐

COROLLARY 5.6.3. *If D is a special ψdo of order m, then $D(1-\Delta)^\gamma$ and $(1-\Delta)^\gamma D$ are special ψdo of order $m+2\gamma$.*

PROOF. If $D = \mathcal{F}^{-1}T_k\mathcal{F}$, then $D(1-\Delta)^\gamma = \mathcal{F}^{-1}T_{k'}\mathcal{F}$, with $k'(\xi,\eta) = k(\xi,\eta)(1+|\eta|^2)^\gamma$, and the verification of (5.6.3) with m replaced by $m+2\gamma$ is fairly simple.

For the other operator, it is sufficient to consider its adjoint, equal to $D^*(1-\Delta)^\gamma$, and we are therefore in the previous situation. ☐

The composition of two ψdo's $D_1 = \mathcal{F}^{-1}T_{k_1}\mathcal{F}$ and $D_2 = \mathcal{F}^{-1}T_{k_2}\mathcal{F}$ can be formally defined as $D_1D_2 = \mathcal{F}^{-1}T_{k_1}T_{k_2}\mathcal{F}$. It is not clear however that this formal definition corresponds to the actual composition of the individual factors when applied to Schwartz functions. In order to settle this point, we show that special ψdo's can be continuously extended to Sobolev spaces of any order.

THEOREM 5.6.4. *A ψdo D of order m extends to a bounded operator from $H^s(\mathbb{R}^n)$ to $H^{s-m}(\mathbb{R}^n)$ for every $s \in \mathbb{R}$.*

PROOF. We first prove that a special ψdo of order 0 extends to a bounded operator on $L^2(\mathbb{R}^n)$. By the density of $\mathcal{S}(\mathbb{R}^n)$ in $L^2(\mathbb{R}^n)$ and the Plancherel formula, this is equivalent to proving that if $k(\xi,\eta)$ satisfies (5.6.3) with $m=0$, then T_k is bounded on $L^2(\mathbb{R}^n)$. If $N > n$,

$$|T_kf(\xi)| \leq C_N \int_{\mathbb{R}^n}(1+|\xi-\eta|)^{-N}|f(\eta)|\,d\eta = C_N(1+|\cdot|)^{-N} * |f|(\xi) \ .$$

Therefore,

$$\begin{aligned}\|T_kf\|_2 &\leq C_N\|f\|_2\int_{\mathbb{R}^n}(1+|\xi|)^{-N}\,d\xi \\ &\leq C_N'\|f\|_2 \ .\end{aligned}$$

For general m and s, we have to show that $D_0 = (1-\Delta)^{\frac{s-m}{2}}D(1-\Delta)^{-\frac{s}{2}}$ is bounded on $L^2(\mathbb{R}^n)$. It follows from Corollary 6.3 that D_0 is a ψdo of order 0. ☐

PROPOSITION 5.6.5. *Let D_1, D_2 be special ψdo of order m_1, m_2 respectively, then D_1D_2 is a special ψdo of order m_1+m_2.*

PROOF. If k_1, k_2 are the kernels associated to D_1, D_2 respectively, we consider the composition $T_{k_1}T_{k_2}$. If $f \in \mathcal{S}(\mathbb{R}^n)$,

$$\begin{aligned}T_{k_1}T_{k_2}f(\xi) &= \int_{\mathbb{R}^n} k_1(\xi,\tau)T_{k_2}f(\tau)\,d\tau \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} k_1(\xi-\tau,\tau)k_2(\tau-\eta,\eta)f(\eta)\,d\eta\,d\tau \\ &= \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} k_1(\xi-\tau,\tau)k_2(\tau-\eta,\eta)\,d\tau\Big)f(\eta)\,d\eta \\ &= \int_{\mathbb{R}^n}\Big(\int_{\mathbb{R}^n} k_1(\xi-\eta-\tau,\tau+\eta)k_2(\tau,\eta)\,d\tau\Big)f(\eta)\,d\eta \ .\end{aligned}$$

This shows that $T_{k_1}T_{k_2} = T_{\tilde{k}}$, with

$$\tilde{k}(\xi,\eta) = \int_{\mathbb{R}^n} k_1(\xi-\tau,\tau+\eta)k_2(\tau,\eta)\,d\tau \ . \tag{5.6.5}$$

We have to show that $\tilde{k}$ satisfies (5.6.3) for $m = m_1 + m_2$. By (5.6.3) applied to k_1, k_2,

$$|\tilde{k}(\xi,\eta)| \leq C_N(1+|\eta|)^{m_2} \int_{\mathbb{R}^n} (1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\, d\tau\ .$$

If $|\tau| < \frac{|\eta|}{2}$, then $1+|\tau+\eta| \sim 1+|\eta|$, so that

$$\begin{aligned}\int_{|\tau|<\frac{|\eta|}{2}} &(1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\, d\tau \\ &\lesssim (1+|\eta|)^{m_1} \int_{|\tau|<\frac{|\eta|}{2}} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N}\, d\tau \\ &\leq (1+|\eta|)^{m_1} \int_{\mathbb{R}^n} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N}\, d\tau\ .\end{aligned}$$

If we split the last integral in two parts, according to whether $|\tau|$ is bigger or smaller than $|\xi-\tau|$, and make the change of variable $\tau' = \xi - \tau$ in one of the two, we see that

$$\int_{\mathbb{R}^n} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N}\, d\tau = 2\int_{|\tau|<|\xi-\tau|} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N}\, d\tau\ .$$

Since $|\xi| \leq |\tau| + |\xi-\tau|$, if $|\tau| < |\xi-\tau|$ then $|\xi-\tau| > \frac{|\xi|}{2}$, and therefore

$$\begin{aligned}\int_{|\tau|<|\xi-\tau|} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N}\, d\tau &\leq \Big(1+\frac{|\xi|}{2}\Big)^{-N} \int_{\mathbb{R}^n} (1+|\tau|)^{-N}\, d\tau \\ &\leq C_N(1+|\xi|)^{-N}\ .\end{aligned}$$

We have so proved that, for any N,

$$\begin{aligned}\int_{|\tau|<\frac{|\eta|}{2}} &(1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\, d\tau \\ &\leq C_N(1+|\xi|)^{-N}(1+|\eta|)^{m_1}\ .\end{aligned} \tag{5.6.6}$$

If $|\tau| > \frac{|\eta|}{2}$, we separate the case $m_1 \geq 0$ from the case $m_1 < 0$. If $m_1 \geq 0$, we use the fact that $|\tau+\eta| \leq 3|\tau|$, so that

$$\begin{aligned}\int_{|\tau|>\frac{|\eta|}{2}} &(1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\, d\tau \\ &\lesssim \int_{|\tau|>\frac{|\eta|}{2}} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N+m_1}\, d\tau \\ &\lesssim \int_{\mathbb{R}^n} (1+|\xi-\tau|)^{-N+m_1}(1+|\tau|)^{-N+m_1}\, d\tau \\ &\lesssim (1+|\xi|)^{-N+m_1} \\ &\leq (1+|\xi|)^{-N+m_1}(1+|\eta|)^{m_1}\ .\end{aligned}$$

If $m_1 < 0$, we use the fact that $(1+|\tau+\eta|)^{m_1} \leq 1$, together with the fact that $(1+|\tau|)^{m_1} \leq (1+|\eta|/2)^{m_1}$, to obtain the inequality

$$\begin{aligned}\int_{|\tau|>\frac{|\eta|}{2}} &(1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\, d\tau \\ &\lesssim (1+|\eta|)^{m_1} \int_{|\tau|>\frac{|\eta|}{2}} (1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N-m_1}\, d\tau \\ &\lesssim (1+|\eta|)^{m_1} \int_{|\tau|>\frac{|\eta|}{2}} (1+|\xi-\tau|)^{-N-m_1}(1+|\tau|)^{-N-m_1}\, d\tau \\ &\lesssim (1+|\xi|)^{-N-m_1}(1+|\eta|)^{m_1}\ .\end{aligned}$$

The two cases can be combined together to give the inequality

$$
\begin{aligned}
(5.6.7)\qquad & \int_{|\tau|>\frac{|\eta|}{2}} (1+|\xi-\tau|)^{-N}(1+|\tau+\eta|)^{m_1}(1+|\tau|)^{-N}\,d\tau \\
&\qquad \le C_N(1+|\xi|)^{-N+|m_1|}(1+|\eta|)^{m_1}\ .
\end{aligned}
$$

Putting (5.6.6) and (5.6.7) together, we have that

$$
|\tilde{k}(\xi,\eta)| \le C_N(1+|\xi|)^{-N+|m_1|}(1+|\eta|)^{m_1+m_2}\ ,
$$

for any N.

Consider now a derivative of $\tilde{k}$ in η,

$$
\partial_\eta^\beta \tilde{k}(\xi,\eta) = \sum_{\alpha\le\beta} c_{\beta,\alpha}\int_{\mathbb{R}^n} \partial_\eta^\alpha k_1(\xi-\tau,\tau+\eta)\partial_\eta^{\beta-\alpha}k_2(\tau,\eta)\,d\tau\ .
$$

By remark (4) above, each term can be treated as before, with m_1 replaced by $m_1-|\alpha|$ and m_2 replaced by $m_2-|\beta|+|\alpha|$. ☐

Proposition 5.6.6. *Let D_1, D_2 be special ψdo of order m_1, m_2 respectively. Then $[D_1, D_2]$ is a special ψdo of order m_1+m_2-1.*

Proof. Let k_1, k_2 be the kernels associated to D_1, D_2, and let $\tilde{k}, \tilde{k}'$ be such that $T_{\tilde{k}} = T_{k_1}T_{k_2}$, $T_{\tilde{k}'} = T_{k_2}T_{k_1}$. Then $\tilde{k}$ is given by (5.6.5), and

$$
\begin{aligned}
\tilde{k}'(\xi,\eta) &= \int_{\mathbb{R}^n} k_2(\xi-\tau,\tau+\eta)k_1(\tau,\eta)\,d\tau \\
&= \int_{\mathbb{R}^n} k_2(\xi-\tau,\tau+\eta)k_1(\tau,\eta)\,d\tau \\
&= \int_{\mathbb{R}^n} k_1(\xi-\tau,\eta)k_2(\tau,\xi-\tau+\eta)\,d\tau\ .
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\tilde{k}(\xi,\eta)-\tilde{k}'(\xi,\eta) &= \int_{\mathbb{R}^n} k_1(\xi-\tau,\tau+\eta)k_2(\tau,\eta)\,d\tau \\
&\qquad - \int_{\mathbb{R}^n} k_1(\xi-\tau,\eta)k_2(\tau,\xi-\tau+\eta)\,d\tau \\
&= \int_{\mathbb{R}^n} \big(k_1(\xi-\tau,\tau+\eta)-k_1(\xi-\tau,\eta)\big)k_2(\tau,\eta)\,d\tau \\
&\qquad + \int_{\mathbb{R}^n} k_1(\xi-\tau,\eta)\big(k_2(\tau,\eta)-k_2(\tau,\xi-\tau+\eta)\big)\,d\tau \\
&= h_1(\xi,\eta)+h_2(\xi,\eta)\ .
\end{aligned}
$$

We prove that each of h_1 and h_2 satisfies (5.6.3) with $m = m_1+m_2-1$. The proof requires some modifications to that of Proposition 5.6.5.

For h_1, we first integrate for $|\tau|<\frac{|\eta|}{2}$. Since $|t\tau+\eta|\sim|\eta|$ for $0<t<1$, by the mean value theorem

$$
\begin{aligned}
\big|k_1(\xi-\tau,\tau+\eta)-k_1(\xi-\tau,\eta)\big| &\le |\tau|\sup_{0<t<1}\big|\nabla_\eta k_1(\xi-\tau,t\tau+\eta)\big| \\
&\le C_N|\tau|(1+|\xi-\tau|)^{-N}(1+|\eta|)^{m_1-1}\ .
\end{aligned}
$$

Therefore,

$$\begin{aligned}\int_{|\tau|<\frac{|\eta|}{2}} &\big|k_1(\xi-\tau,\tau+\eta)-k_1(\xi-\tau,\eta)\big|\big|k_2(\tau,\eta)\big|\,d\tau\\ &\le C_N(1+|\eta|)^{m_1+m_2-1}\int_{|\tau|<\frac{|\eta|}{2}}(1+|\xi-\tau|)^{-N}(1+|\tau|)^{-N+1}\,d\tau\\ &\le C_N(1+|\xi|)^{-N+1}(1+|\eta|)^{m_1+m_2-1}\ .\end{aligned}$$

Passing to the integral for $|\tau|>\frac{|\eta|}{2}$, we write

$$\begin{aligned}\int_{|\tau|>\frac{|\eta|}{2}} &\big|k_1(\xi-\tau,\tau+\eta)-k_1(\xi-\tau,\eta)\big|\big|k_2(\tau,\eta)\big|\,d\tau\\ &\le\int_{|\tau|>\frac{|\eta|}{2}}|k_1(\xi-\tau,\tau+\eta)||k_2(\tau,\eta)|\,d\tau\\ &\quad+\int_{|\tau|>\frac{|\eta|}{2}}|k_1(\xi-\tau,\eta)||k_2(\tau,\eta)|\,d\tau\ .,\end{aligned}$$

and estimate the two terms separately, distinguishing between the cases $m_1\ge 0$ and $m_1<0$. The proof goes as in Proposition 5.6.5[9]. The derivatives of h_1 are estimated in a similar way.

For h_2, the integral must be split according to whether $|\xi-\tau|<\frac{|\eta|}{2}$ or $|\xi-\tau|>\frac{|\eta|}{2}$, and there is no substantial difference. □

Bessel potentials can be used to approximate L^2-functions by $H^s(\mathbb{R}^n)$-functions with $s>0$.

For $\delta>0$, the operator $(1-\delta^2\Delta)^\gamma$ is naturally defined by

$$\big((1-\delta^2\Delta)^\gamma\varphi\big)\hat{}\,(\xi)=(1+\delta^2|\xi|^2)^\gamma\hat\varphi(\xi)\ .$$

Lemma 5.6.7. *Let $s>0$. If $f\in L^2(\mathbb{R}^n)$, then $(1-\delta^2\Delta)^{-\frac{s}{2}}f\in H^s(\mathbb{R}^n)$ and*

$$\lim_{\delta\to0}\big\|(1-\delta^2\Delta)^{-\frac{s}{2}}f-f\big\|_2=0\ .$$

Proof. The first statement follows from the trivial estimate $(1+\delta^2|\xi|^2)^{-\frac{s}{2}}\le C_\delta(1+|\xi|^2)^{-\frac{s}{2}}$. By Plancherel's formula,

$$\big\|(1-\delta^2\Delta)^{-\frac{s}{2}}f-f\big\|_2^2=(2\pi)^n\int_{\mathbb{R}^n}\Big|\frac{1}{(1+\delta^2|\xi|^2)^{\frac{s}{2}}}-1\Big|^2|\hat f(\xi)|^2\,d\xi\ .$$

Since $\Big|\frac{1}{(1+\delta^2|\xi|^2)^{\frac{s}{2}}}-1\Big|\le 2$, $\delta\to0$, the integral tends to 0 by dominated convergence. □

For $\delta<1$, the constant C_δ appearing in the proof above is of the order of δ^{-s}. This means that, whereas the L^2-norms of $(1-\delta^2\Delta)^{-\frac{s}{2}}f$ remain bounded as $\delta\to0$, the H^s-norms can blow up, at most like δ^{-s}. In the same way, one can see that intermediate H^r-norms (i.e. with $0<r<s$) of $(1-\delta^2\Delta)^{-\frac{s}{2}}f$ can blow up at most like δ^{-r}. In particular, if $|\alpha|\le s$,

$$\big\|\partial^\alpha(1-\delta^2\Delta)^{-\frac{s}{2}}f\big\|_2\le C\delta^{-|\alpha|}\ ,$$

for δ small.

The following statement is a generalization of this fact.

Proposition 5.6.8. *Let D be a special ψdo of order m, $0\le m\le s$. For every $f\in L^2(\mathbb{R}^n)$ and every $\delta<1$,*

$$\big\|D(1-\delta^2\Delta)^{-\frac{s}{2}}f\big\|_2\le C\delta^{-m}\|f\|_2\ ,\quad\big\|(1-\delta^2\Delta)^{-\frac{s}{2}}Df\big\|_2\le C\delta^{-m}\|f\|_2\ .\tag{5.6.8}$$

[9]Notice that at this stage of the proof of Proposition 5.6.5, the factor $(1+|\eta|)^{m_1}$ was introduced by brute force. The same can be done here with the exponent m_1-1.

PROOF. Since $\big((1-\delta^2\Delta)^{-\frac{s}{2}}D\big)^* = D^*(1-\delta^2\Delta)^{-\frac{s}{2}}$, it is sufficient to prove the first estimate. If $D = \mathcal{F}^{-1}T_k\mathcal{F}$, then $(1-\delta^2\Delta)^{-\frac{s}{2}} = \mathcal{F}^{-1}T_{k'}\mathcal{F}D\mathcal{F}^{-1}$, with

$$k'(\xi,\eta) = k(\xi,\eta)(1+\delta^2|\eta|^2)^{-\frac{s}{2}} \ .$$

Hence, for every N,

$$\begin{aligned}|k'(\xi,\eta)| &\le C_N \frac{(1+|\eta|)^m}{(1+|\xi|)^N(1+\delta|\eta|)^s}\\ &\le C_N\delta^{-m}\frac{1}{(1+|\xi|)^N(1+\delta|\eta|)^{s-m}}\\ &\le C_N\delta^{-m}\frac{1}{(1+|\xi|)^N} \ .\end{aligned}$$

Therefore,

$$|T_{k'}f(\xi)| \le C_N\delta^{-m}\int_{\mathbb{R}^n}\frac{1}{(1+|\xi-\eta|)^N}|f(\eta)|\,d\eta \ .$$

Fixing $N>n$, this implies that

$$\|T_{k'}f\|_2 \le C\delta^{-m}\|f\|_2 \ . \qquad \square$$

PROPOSITION 5.6.9. *Let D be a differential operator on $\mathbb{R}^n$ of order 1 and with coefficients in $\mathcal{D}(\mathbb{R}^n)$. For every $f\in L^2(\mathbb{R}^n)$ and every $\delta<1$,*

$$\big\|\big[(1-\delta^2\Delta)^{-1},D\big]f\big\|_2 \le C\|f\|_2 \ . \tag{5.6.9}$$

PROOF. Let $v=(1-\delta^2\Delta)^{-1}f$. Then

$$\begin{aligned}\big[(1-\delta^2\Delta)^{-1},D\big]f &= (1-\delta^2\Delta)^{-1}Df - D(1-\delta^2\Delta)^{-1}f\\ &= (1-\delta^2\Delta)^{-1}\big(D(1-\delta^2\Delta)v-(1-\delta^2\Delta)Dv\big)\\ &= \delta^2(1-\delta^2\Delta)^{-1}[\Delta,D]v \ .\end{aligned}$$

As we have observed, $[\Delta,D]$ is a differential operator of order at most 2. Applying Proposition 6.8 twice, we obtain that

$$\big\|\big[(1-\delta^2\Delta)^{-1},D\big]f\big\|_2 \le C\|v\|_2 = C\big\|(1-\delta^2\Delta)^{-1}f\big\|_2 \le C^2\|f\|_2 \ . \qquad \square$$

7. Back to L

Recall from (5.5.2) that, for $f\in\mathcal{D}(\Omega')$ and $s<\frac{1}{m}$,

$$\|f\|_{(s)} \le C_s\big(\|f\|_2 + |\!|\!|Lf|\!|\!|'_{\mathcal{X},\Omega'}\big) \ .$$

We first extend it to $f\in H^2(\mathbb{R}^n)$ with $\operatorname{supp} f\subset\Omega'$.

LEMMA 5.7.1. *Suppose that $f\in H^2(\mathbb{R}^n)$, with $\operatorname{supp} f\subset\Omega'$, and that $|\!|\!|Lf|\!|\!|'_{\mathcal{X},\Omega'}<\infty$. Then (5.5.2) holds for $s<\frac{1}{m}$.*

PROOF. Let $\varphi\in\mathcal{D}(\mathbb{R}^n)$ be supported on the unit ball and with $\int_{\mathbb{R}^n}\varphi(x)\,dx = 1$. For $\varepsilon>0$, let $\varphi_\varepsilon(x)=\varepsilon^{-n}\varphi(\varepsilon^{-1}x)$, and set $f_\varepsilon=\varphi_\varepsilon * f$.

If $\varepsilon<d(\operatorname{supp} f,\partial\Omega')$, $f_\varepsilon\in\mathcal{D}(\Omega')\subset H^2(\mathbb{R}^n)$. Moreover, for $|\alpha|\le 2$,

$$\lim_{\varepsilon\to0}\|\partial^\alpha f-\partial^\alpha f_\varepsilon\|_2 = \lim_{\varepsilon\to0}\|\partial^\alpha f-\varphi_\varepsilon*\partial^\alpha f\|_2 = 0 \ , \tag{5.7.1}$$

since the φ_ε form an approximate identity for $\varepsilon\to0$. We have so proved that $f_\varepsilon\to f$ in $H^2(\mathbb{R}^n)$, hence in $H^s(\mathbb{R}^n)$, since $s<2$.

From (5.7.1) we also obtain that

$$\lim_{\varepsilon\to0}\|Lf-Lf_\varepsilon\|_2 = 0 \ ,$$

just because L is a second-order operator with bounded coefficients on Ω'. By (5.5.1),

$$\lim_{\varepsilon\to 0} |\!|\!| Lf - Lf_\varepsilon |\!|\!|'_{\mathcal{X},\Omega'} = 0 \ .$$

We can then apply (5.5.2) to the f_ε and pass to the limit. □

LEMMA 5.7.2. *Let K be a compact subset of Ω'. If $f \in L^2(\mathbb{R}^n)$, with $\operatorname{supp} f \subseteq K$, and $|\!|\!| Lf |\!|\!|'_{\mathcal{X},\Omega'} < \infty$, then, for $s < \frac{1}{m}$, $f \in H^s(\mathbb{R}^n)$ and (5.5.2) holds with a constant $C_{s,K}$ also depending on K.*

PROOF. We can assume that the vector fields X_j have compact support in Ω (if not multiply each of them by a function in $\mathcal{D}(\Omega)$ which is identically equal to 1 on a neighborhood of $\overline{\Omega'}$; since we will apply the X_j and L only to functions compactly supported in Ω', this modification will not affect our operations). This modification allows us to extend the X_j and L to all of $\mathbb{R}^n$ (setting them identically equal to 0 outside of Ω) and to regard them as special ψdo's.

For $\delta > 0$ define

$$f_\delta = (1 - \delta^2\Delta)^{-1} f \ .$$

The discussion at the beginning of Section 7 shows that $f_\delta \in H^2(\mathbb{R}^n)$. If $\chi \in \mathcal{D}(\Omega')$ is identically equal to 1 on a neighborhood of K, then also $\chi f_\delta \in H^2(\mathbb{R}^n)$. Applying Lemma 5.7.1 to χf_δ, with K replaced by $K' = \operatorname{supp}\chi$, we have that

$$\|\chi f_\delta\|_{(s)} \le C_s\big(\|f_\delta\|_2 + |\!|\!| L\chi f_\delta |\!|\!|'_{\mathcal{X},\Omega'}\big) \ . \tag{5.7.2}$$

We prove that the norms $|\!|\!| L\chi f_\delta |\!|\!|'_{\mathcal{X},\Omega'}$ are uniformly bounded for δ small. Since $f = \chi f = \chi(1-\delta^2\Delta) f_\delta$,

$$\begin{aligned} Lf &= L\chi(1-\delta^2\Delta) f_\delta \\ &= (1-\delta^2\Delta) L\chi f_\delta - \delta^2 [L\chi, \Delta] f_\delta \ . \end{aligned}$$

Applying $(1-\delta^2\Delta)^{-1}$ to both sides, we obtain that

$$L\chi f_\delta = (1-\delta^2\Delta)^{-1} Lf + \delta^2 (1-\delta^2\Delta)^{-1}[L\chi,\Delta] f_\delta \ ,$$

so that

$$|\!|\!| L\chi f_\delta |\!|\!|'_{\mathcal{X},\Omega'} \le \big|\!\big|\!\big| (1-\delta^2\Delta)^{-1} Lf \big|\!\big|\!\big|'_{\mathcal{X},\Omega'} + \delta^2 \big|\!\big|\!\big| (1-\delta^2\Delta)^{-1}[L\chi,\Delta] f_\delta \big|\!\big|\!\big|'_{\mathcal{X},\Omega'} \ . \tag{5.7.3}$$

Now,

$$\begin{aligned} |\!|\!| (1-\delta^2\Delta)^{-1} Lf |\!|\!|'_{\mathcal{X},\Omega'} &= \sup_{\varphi\in\mathcal{D}(\Omega'):|\!|\!|\varphi|\!|\!|_X\le 1} \big|\langle (1-\delta^2\Delta)^{-1} Lf, \varphi\rangle\big| \\ &= \sup_{\varphi\in\mathcal{D}(\Omega'):|\!|\!|\varphi|\!|\!|_X\le 1} \big|\langle Lf, (1-\delta^2\Delta)^{-1}\varphi\rangle\big| \\ &= \sup_{\varphi\in\mathcal{D}(\Omega'):|\!|\!|\varphi|\!|\!|_X\le 1} \big|\langle Lf, \chi(1-\delta^2\Delta)^{-1}\varphi\rangle\big| \\ &\le |\!|\!| Lf |\!|\!|'_{\mathcal{X},\Omega'} \sup_{\varphi\in\mathcal{D}(\Omega'):|\!|\!|\varphi|\!|\!|_X\le 1} |\!|\!| \chi(1-\delta^2\Delta)^{-1}\varphi |\!|\!|_{\mathcal{X}} \ . \end{aligned}$$

By (5.6.6),

$$\|\chi(1-\delta^2\Delta)^{-1}\varphi\|_2 \le \|\varphi\|_2 \ .$$

Moreover, setting $D_j g = X_j\chi g$, we can apply Proposition 5.6.9 and obtain that, for δ small,

$$\begin{aligned} \big\| X_j\chi(1-\delta^2\Delta)^{-1}\varphi)\big\|_2 &= \|D_j(1-\delta^2\Delta)^{-1}\varphi\|_2 \\ &\le \|(1-\delta^2\Delta)^{-1} D_j\varphi\|_2 + \big\|[D_j,(1-\delta^2\Delta)^{-1}]\varphi\big\|_2 \\ &\le C_K\big(\|D_j\varphi\|_2 + \|\varphi\|_2\big) \\ &\le C_K\big(\|(X_j\chi)\varphi\|_2 + \|\chi X_j\varphi\|_2 + \|\varphi\|_2\big) \\ &\le C_K |\!|\!|\varphi|\!|\!|_{\mathcal{X}} \end{aligned}$$

(notice that the constant depends on the choice of χ, i.e. on K).

Therefore,

$$|\!|\!| \chi'(1-\delta^2\Delta)^{-1}\varphi |\!|\!|_{\mathcal{X}} \le C_K |\!|\!|\varphi|\!|\!|_{\mathcal{X}} \ ,$$

and hence

$$\|(1-\delta^2\Delta)^{-1}Lf\|'_{\mathcal{X},\Omega'} \le C_K \|Lf\|'_{\mathcal{X},\Omega'} \ . \tag{5.7.4}$$

Consider now the last term in (5.7.3),

$$\delta^2 \big\|(1-\delta^2\Delta)^{-1}[L\chi,\Delta]f_\delta\big\|'_{\mathcal{X},\Omega'} \le \delta^2 \sum_{j=1}^k \big\|(1-\delta^2\Delta)^{-1}[X_j^2\chi,\Delta]f_\delta\big\|'_{\mathcal{X},\Omega'} \ .$$

Each summand in the right-hand side can be estimated as follows:

$$\begin{aligned}
&\big\|(1-\delta^2\Delta)^{-1}[X_j^2\chi,\Delta]f_\delta\big\|'_{\mathcal{X},\Omega'} \\
&\qquad = \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_{\mathcal{X}}\le 1} \big|\big\langle f_\delta, [\Delta,\chi(X_j^*)^2](1-\delta^2\Delta)^{-1}\varphi\big\rangle\big| \\
&\qquad = \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_{\mathcal{X}}\le 1} \big|\big\langle f, (1-\delta^2\Delta)^{-1}[\Delta,\chi(X_j^*)^2](1-\delta^2\Delta)^{-1}\varphi\big\rangle\big| \\
&\qquad \le \|f\|_2 \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_{\mathcal{X}}\le 1} \big\|(1-\delta^2\Delta)^{-1}[\Delta,\chi(X_j^*)^2](1-\delta^2\Delta)^{-1}\varphi\big\|_2 \ .
\end{aligned} \tag{5.7.5}$$

We have that

$$\begin{aligned}
[\Delta,\chi(X_j^*)^2] &= \Delta\chi(X_j^*)^2 - \chi(X_j^*)^2\Delta \\
&= [\Delta,\chi](X_j^*)^2 + \chi\big(\Delta(X_j^*)^2 - (X_j^*)^2\Delta\big) \\
&= [\Delta,\chi](X_j^*)^2 + \chi\big([\Delta,X_j^*]X_j^* + X_j^*[\Delta,X_j^*]\big) \\
&= [\Delta,\chi](X_j^*)^2 + \chi\big(2[\Delta,X_j^*]X_j^* + \big[X_j^*,[\Delta,X_j^*]\big]\big) \\
&= D_1X_j^* + D_2 \ ,
\end{aligned}$$

where D_1, D_2 are second-order operators with compact support in Ω'. Therefore

$$\begin{aligned}
&\big\|(1-\delta^2\Delta)^{-1}[\Delta,\chi(X_j^*)^2](1-\delta^2\Delta)^{-1}\varphi\big\|_2 \\
&\qquad \le \big\|(1-\delta^2\Delta)^{-1}D_1X_j^*(1-\delta^2\Delta)^{-1}\varphi\big\|_2 \\
&\qquad\quad + \big\|(1-\delta^2\Delta)^{-1}D_2(1-\delta^2\Delta)^{-1}\varphi\|_2 \\
&\qquad \le \big\|(1-\delta^2\Delta)^{-1}D_1(1-\delta^2\Delta)^{-1}X_j^*\varphi\big\|_2 \\
&\qquad\quad + \big\|(1-\delta^2\Delta)^{-1}D_1[X_j^*,(1-\delta^2\Delta)^{-1}]\varphi\big\|_2 \\
&\qquad\quad + \big\|(1-\delta^2\Delta)^{-1}D_2(1-\delta^2\Delta)^{-1}\varphi\|_2 \ .
\end{aligned}$$

Since X_j, D_1 and D_2 depend only on the choice of χ, i.e. on K, we have, by Proposition 5.6.8,

$$\begin{aligned}
\big\|(1-\delta^2\Delta)^{-1}D_1(1-\delta^2\Delta)^{-1}X_j^*\varphi\big\|_2 &\le \big\|D_1(1-\delta^2\Delta)^{-1}X_j^*\varphi\big\|_2 \\
&\le C_K\delta^{-2}\big\|X_j^*\varphi\big\|_2 \\
&\le C_K\delta^{-2}\big(\|\varphi\|_2 + \|X_j\varphi\|_2\big) \\
&\le C_K\delta^{-2}\big(\|\varphi\|_2 + \|X_j\varphi\|_2\big) \\
&\le C_K\delta^{-2}\|\varphi\|_{\mathcal{X}} \ .
\end{aligned}$$

Similarly,

$$\|(1-\delta^2\Delta)^{-1}D_2(1-\delta^2\Delta)^{-1}\varphi\|_2 \le C_K\delta^{-2}\|\varphi\|_2 \le C_K\delta^{-2}\|\varphi\|_{\mathcal{X}} \ .$$

Finally, by Propositions 5.6.8 and 5.6.9 together,

$$\begin{aligned}
\big\|(1-\delta^2\Delta)^{-1}D_1[X_j^*,(1-\delta^2\Delta)^{-1}]\varphi\big\|_2 &\le C_K\delta^{-2}\big\|[X_j^*,(1-\delta^2\Delta)^{-1}]\varphi\big\|_2 \\
&\le C_K\delta^{-2}\|\varphi\|_2 \\
&\le C_K\delta^{-2}\|\varphi\|_{\mathcal{X}} \ .
\end{aligned}$$

Putting these various estimates into (5.7.5), we have

$$\left\| (1-\delta^2\Delta)^{-1}[X_j^2,\Delta]f_\delta \right\|'_{\mathcal{X},\Omega'} \le C_K \|f\|_2 \ . \tag{5.7.6}$$

Then (5.7.4) and (5.7.6) give that

$$\|Lf_\delta\|'_{\mathcal{X},\Omega'} \le C_K\big(\|Lf\|'_{\mathcal{X},\Omega'} + \|f\|_2\big) \ ,$$

and from (5.7.2) we have

$$\|\chi f_\delta\|_{(s)} \le C_{s,K}\big(\|f\|_2 + \|Lf\|'_{\mathcal{X},\Omega'}\big) \ . \tag{5.7.7}$$

It follows that the norms $\|\chi f_\delta\|_{(s)}$ are bounded for δ small. Let $\{\delta_j\}$ be a sequence tending to 0 such that χf_{δ_j} have a weak limit g in $H^s(\mathbb{R}^n)$. The norm of the weak limit is not larger than the right-hand side in (5.7.7). By compactness of the inclusion $H^s(\mathbb{R}^n) \subset L^2(\mathbb{R}^n)$, χf_{δ_j} tend to g in $L^2(\mathbb{R}^n)$. But, by Lemma 5.6.7, χf_{δ_j} tend to f in $L^2(\mathbb{R}^n)$. Hence $f = g$, so that $f \in H^s(\mathbb{R}^n)$ and satifies (5.5.2). □

8. Hypoellipticity of L

Recall that, for $r \in \mathbb{R}$,

$$H^r(\mathbb{R}^n) = (1-\Delta)^{-\frac{r}{2}}\big(L^2(\mathbb{R}^n)\big) \ .$$

The following statement has the effect of "translating" (5.5.2) from the level of $H^0(= L^2)$-functions to the level of H^r-functions. The exponent r can be both positive or negative.

Lemma 5.8.1. *Let K be a compact subset of Ω', and let $\chi \in \mathcal{D}(\Omega')$ be identically equal to 1 on a neighborhood of K. If $f \in H^r(\mathbb{R}^n)$, with $\operatorname{supp} f \subseteq K$ and $\|\chi(1-\Delta)^{\frac{r}{2}}Lf\|'_{\mathcal{X},\Omega'} < \infty$, then, for $s < \frac{1}{m}$, $f \in H^{r+s}(\mathbb{R}^n)$ and*

$$\|f\|_{(r+s)} \le C_{s,K}\big(\|f\|_{(r)} + \|\chi(1-\Delta)^{\frac{r}{2}}Lf\|'_{\mathcal{X},\Omega'}\big) \ .$$

Proof. The statement follows immediately if we show that Lemma 5.7.2 can be applied to to $g = \chi(1-\Delta)^{\frac{r}{2}}f$, with K replaced by $K' = \operatorname{supp}\chi \subset \Omega'$.

By assumption, $g \in L^2(\mathbb{R}^n)$ and $\operatorname{supp} g \subseteq K'$. We have to show that

$$\|Lg\|'_{\mathcal{X},\Omega'} < \infty \ . \tag{5.8.1}$$

Assuming, as in the proof of Lemma 5.7.2, that the X_j and L are compactly supported in Ω, in our hypotheses, (5.8.1) is equivalent to

$$\left\| \big[\chi(1-\Delta)^{\frac{r}{2}}, L\big]f \right\|'_{\mathcal{X},\Omega'} < \infty \ .$$

As in the proof of 5.7.2,

$$\begin{aligned}
&\left\| \big[\chi(1-\Delta)^{\frac{r}{2}}, L\big]f \right\|'_{\mathcal{X},\Omega'} \\
&\qquad = \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_X\le 1} \big|\big\langle f, \big[L^*, (1-\Delta)^{\frac{r}{2}}\chi\big]\varphi\big\rangle\big| \\
&\qquad = \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_X\le 1} \big|\big\langle (1-\Delta)^{\frac{r}{2}}f, (1-\Delta)^{-\frac{r}{2}}\big[L^*, (1-\Delta)^{\frac{r}{2}}\chi\big]\varphi\big\rangle\big| \\
&\qquad \le \|f\|_{(r)} \sup_{\varphi\in\mathcal{D}(\Omega'):\|\varphi\|_X\le 1} \big\|(1-\Delta)^{-\frac{r}{2}}\big[L^*, (1-\Delta)^{\frac{r}{2}}\chi\big]\varphi\big\|_2 \ .
\end{aligned}$$

Now,

$$\begin{aligned}\left[L^*,(1-\Delta)^{\frac{r}{2}}\chi\right] &= \sum_{j=1}^k \left[(1-\Delta)^{\frac{r}{2}}\chi,(X_j^*)^2\right] \\ &= \sum_{j=1}^k \left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]X_j^* + \sum_{j=1}^k X_j^*\left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right] \\ &= 2\sum_{j=1}^k \left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]X_j^* + \sum_{j=1}^k \left[X_j^*,\left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]\right] .\end{aligned}$$

By Proposition 5.6.6, $\left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]$ is a special ψdo of order r, and Corollary 5.6.3 implies that its composition with $(1-\Delta)^{-\frac{r}{2}}$ is of order 0. Therefore, by Theorem 5.6.4,

$$\left\|(1-\Delta)^{-\frac{r}{2}}\left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]X_j^*\varphi\right\|_2 \le C\|X_j^*\varphi\|_2 \le C|\!|\!|\varphi|\!|\!|_{\mathcal{X}} .$$

For the same reason,

$$\left\|(1-\Delta)^{-\frac{r}{2}}\left[X_j^*,\left[(1-\Delta)^{\frac{r}{2}}\chi,X_j^*\right]\right]\varphi\right\|_2 \le C\|\varphi\|_2 .$$

Therefore,

$$\left|\!\left|\!\left|\left[(1-\Delta)^{\frac{r}{2}}\chi,L\right]f\right|\!\right|\!\right|'_{\mathcal{X},\Omega'} \le C\|f\|_{(r)} ,$$

and we are finished. □

We can now prove the final theorem.

THEOREM 5.8.2. *L is hypoelliptic in Ω.*

PROOF. Let $u\in\mathcal{D}'(\Omega)$ be such that Lu coincides with a C^∞-function on an open subset Ω'. We can assume that Ω' is relatively compact. Multiplying, if necessary, u by a function in $\mathcal{D}(\Omega)$ that is identically equal to 1 on a neighborhood of $\overline{\Omega'}$, we can also assume that u has compact support in Ω. We then extend u to all of $\mathbb{R}^n$, by imposing that it vanishes on $\mathbb{R}^n\setminus K$. Since any distribution with compact support has finite order, there is $r\in\mathbb{R}$ such that $u\in H^r(\mathbb{R}^n)$.

For a fixed ball B in Ω', we choose two functions $\chi,\chi'\in\mathcal{D}(\Omega')$ such that χ is identically equal to 1 on a neighborhood of $\bar{B}$, and χ' is identically equal to 1 on a neighborhood of $\operatorname{supp}\chi$. In particular $\chi\chi'=\chi$.

If we prove that $|\!|\!|\chi'(1-\Delta)^{\frac{r}{2}}L(\chi u)|\!|\!|'_{\mathcal{X},\Omega'}<\infty$, it follows from Lemma 5.8.1 that $\chi u\in H^{r+s}(\mathbb{R}^n)$ for $s<\frac{1}{m}$. We can then reapply Lemma 5.8.1 repeatedly to conclude that $\chi u\in H^N(\mathbb{R}^n)$ for every N. By the Sobolev embedding theorem[10], $\chi u\in C^\infty(\mathbb{R}^n)$, i.e. $u\in C^\infty(B)$, and we are finished.

We consider separately each summand coming from application of Leibniz's formula

$$L(\chi u) = \chi Lu + (L\chi)u - 2\sum_{j=1}^k (X_j\chi)X_ju .$$

For the first term, notice that $\chi Lu\in\mathcal{D}(\mathbb{R}^n)$ by hypothesis. Therefore, applying $\chi'(1-\Delta)^{\frac{r}{2}}$ to it, we obtain a function in every Sobolev space, i.e. a C^∞-function, which also has compact support. Therefore

$$|\!|\!|\chi'(1-\Delta)^{\frac{r}{2}}(\chi Lu)|\!|\!|'_{\mathcal{X},\Omega'} \le \left\|\chi'(1-\Delta)^{\frac{r}{2}}(\chi Lu)\right\|_2 < \infty .$$

For the second term, we observe that $(L\chi)u\in H^r(\mathbb{R}^n)$, i.e. $(1-\Delta)^{\frac{r}{2}}\big((L\chi)u\big)\in L^2(\mathbb{R}^n)$, and we are done.

---

[10]Directly, we have

$$\int_{\mathbb{R}^n} |\widehat{\chi u}(\xi)|^2(1+|\xi|)^{2N}\,d\xi<\infty$$

for every N. Applying Schwartz's inequality,

$$\widehat{\partial^\alpha(\chi u)}(\xi) = (i\xi)^\alpha\widehat{\chi u}(\xi)\in L^1(\mathbb{R}^n) ,$$

for every multi-index α. Therefore $\partial^\alpha(\chi u)$ is continuous.

For each of the other terms, setting $\tilde{X}_j = (X_j\chi)X_j$, we must then prove that, for some constant $C > 0$ and every $\varphi \in \mathcal{D}(\Omega')$,

$$\left|\left\langle \chi'(1-\Delta)^{\frac{r}{2}}\tilde{X}_j u, \varphi\right\rangle\right| \leq C\|\varphi\|_{\mathcal{X}} \ .$$

We have

$$\begin{aligned}\left\langle \chi'(1-\Delta)^{\frac{r}{2}}\tilde{X}_j u, \varphi\right\rangle &= \left\langle \tilde{X}_j\big(\chi'(1-\Delta)^{\frac{r}{2}}u\big), \varphi\right\rangle + \left\langle \left[\chi'(1-\Delta)^{\frac{r}{2}}, \tilde{X}_j\right]u, \varphi\right\rangle \\ &= \left\langle (1-\Delta)^{\frac{r}{2}}u, \chi'\tilde{X}_j^*\varphi\right\rangle + \left\langle \left[\chi'(1-\Delta)^{\frac{r}{2}}, \tilde{X}_j\right]u, \varphi\right\rangle\end{aligned}$$

Then

$$\left|\left\langle (1-\Delta)^{\frac{r}{2}}u, \chi'\tilde{X}_j^*\varphi\right\rangle\right| \leq \|u\|_{H^r}\|\tilde{X}_j^* u\|_2 \leq C\|u\|_{H^r}\|\varphi\|_{\mathcal{X}} \ .$$

Since $\left[\chi'(1-\Delta)^{\frac{r}{2}}, \tilde{X}_j\right]$ is a special ψdo of order r,

$$\begin{aligned}\left|\left\langle \left[\chi'(1-\Delta)^{\frac{r}{2}}, \tilde{X}_j\right]u, \varphi\right\rangle\right| &\leq \left\|\left[\chi'(1-\Delta)^{\frac{r}{2}}, \tilde{X}_j\right]u\right\|_2\|\varphi\|_2 \\ & C\|u\|_{H^r}\|\varphi\|_2 \ .\end{aligned}$$

Putting everything together, we obtain the desired estimate. □