Introduction to Analysis on Lie Groups:

Lectures by W. Hebisch.

Notes by J. Inglis.

December 12, 2007

Contents

1 Examples of Lie Groups

We first of all give some examples of simple Lie groups. Each example given below is the simplest possible example of a given class of Lie group. They also illustrate properties (or lack of them) important for analysis on Lie groups.

1. $\mathrm{SL}(2,\mathbb{R})$, the group of 2×2 matrices of determinant 1. This is a semisimple Lie group.

2. The Heisenberg group, $\mathbb{H}$. This group has several representations. For example, it can be represented as the group of matrices

$$\begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}, \quad x, y, z \in \mathbb{R}.$$

Perhaps, for calculations, a more useful representation is given by $\mathbb{R}^3$ equipped with the the group action

$$(x_1, y_1, z_1) \circ (x_2, y_2, z_2) = (x_1 + x_2, y_1 + y_2, z_1 + z_2 + x_1 y_2 - x_2 y_1).$$

The Heisenberg group is the simplest example of a nilpotent Lie group.

3. The "$ax + b$" group, of invertible affine transformations of $\mathbb{R}^3$. This group can be represented by matrices of the form

$$\begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}, \quad a \in \mathbb{R}\backslash\{0\}, \quad b \in \mathbb{R}.$$

The "$ax + b$" group is the simplest example of a non-commutative solvable group, the simplest example of a non-unimodular group, and the simplest example of a Lie group with exponential growth i.e. the volume of a ball in this group grows exponentially with the radius. Note that this group sits inside $\mathrm{SL}(2,\mathbb{R})$, and it can be shown that $\mathrm{SL}(2,\mathbb{R})$ shares some of the properties as the "$ax + b$" group, for example exponential growth.

4. $\mathrm{SO}(3)$, the group of 3×3 orthogonal matrices of determinant 1. This is a compact group.

5. The group of matrices

$$\begin{pmatrix} \cos\alpha & \sin\alpha & x \\ -\sin\alpha & \cos\alpha & y \\ 0 & 0 & 1 \end{pmatrix}, \quad x, y \in \mathbb{R}, -\pi < \alpha \leq \pi.$$

This group of matrices represents the motion group of the plane, and is the simplest example of a non-nilpotent group of polynomial growth.

Lie groups are differentiable manifolds equipped with smooth group action. Loosely, it is therefore natural to investigate “objects” that are invariant under left or right translation by the group law. For example, the set of all left (or right) invariant vector fields on a Lie group is defined to be the Lie algebra of a Lie group, and a left (or right) invariant measure is said to be a left (or right) Haar measure (see appendix for details).

Consider again the Heisenberg group, $\mathbb{H}$, defined above. We can compute the left invariant vector fields on $\mathbb{H}$. We have

$$\begin{aligned} \partial_s f((s,0,0)\circ(x,y,z)) &= \partial_s f(x+s, y, z+sy), \\ &= (\partial_x + y\partial_z) f. \end{aligned}$$

Thus $X = \partial_x + y\partial_z$ is a left invariant vector field on $\mathbb{H}$. Similarly,

$$\begin{aligned} \partial_s f((0,s,0)\circ(x,y,z)) &= \partial_s f(x, y+s, z-xs), \\ &= (\partial_y - x\partial_z) f, \end{aligned}$$

and

$$\begin{aligned} \partial_s f((0,0,s)\circ(x,y,z)) &= \partial_s f(x, y, z+s), \\ &= \partial_z f, \end{aligned}$$

so that $Y = \partial_y - x\partial_z$ and $Z = \partial_z$ are also left invariant vector fields. We can also compute the commutator $[X,Y]$, noting that partial derivatives commute,

$$\begin{aligned} [X,Y] &= (\partial_x + y\partial_z)(\partial_y - x\partial_z) - (\partial_y - x\partial_z)(\partial_x + y\partial_z) \\ &= \partial_x(-x\partial_z) + y\partial_z\partial_y - \partial_y(y\partial_z) + x\partial_z\partial_x \\ &= -2\partial_z \\ &= -2Z. \end{aligned}$$

Remark 1 *In Euclidean space, it is clear that if we have just two directions to travel in, we will always remain in a 2-dimensional plane. This is not true in the case of the Heisenberg group. Indeed, if we travel a distance t, first in the direction of the vector field X, then in direction Y, then in direction $-X$ and finally in direction $-Y$, then by the formula*

$$\exp(tX)\exp(tY)\exp(-tX)\exp(-tY) \approx t^2[X,Y],$$

we see that in effect we have moved a distance t^2 in the direction $[X,Y]$. But by above, we have therefore moved in the direction Z. Thus, moving in just two directions, we can get to any point in the group.

2 Carnot-Carathéodory Distance

Let $X_1, \ldots, X_m$ be smooth right invariant vector fields on a finite dimensional Lie group $\mathbb{G}$. We say $X_1, \ldots, X_m$ satisfy Hörmander's condition if

$X_1, \ldots, X_m, [X_1, X_2], \ldots, [X_1, X_m]+$ *all repeated commutators up to order* $k < \infty$

span the tangent space at the identity e. Note that $\mathbb{G}$ is finite dimensional, so this tangent space is also finite dimensional.

Suppose $X_1, \ldots, X_m$ satisfy Hörmander's condition. We say an absolutely continuous curve $\gamma \colon [0, 1] \to \mathbb{G}$ is admissible if $\exists\ a_1, \ldots, a_m$, such that

$$\gamma'(t) = \sum_{i=1}^{m} a_i(t) X_i(\gamma(t))$$

i.e $\gamma'(t) \in \mathrm{sp}\{X_1(\gamma(t)), \ldots, X_m(\gamma(t))\}$.

For such an admissible curve γ, define the length of γ in following way

$$|\gamma|^2 = \inf_{a_1,\ldots,a_m} \int_0^1 \sum_{i=1}^{m} a_i^2(t)\mathrm{d}t.$$

Finally define the function $d(x) = \inf\{|\gamma| : \gamma \text{ is admissible}, \gamma(0) = e, \gamma(1) = x\}$. This is the *Carnot-Carathéodory* distance of x from the identity, or the *optimal control* distance, which generates the sub-Riemannian geometry on $\mathbb{G}$.

Remark 2 *The function $d(x)$ is well defined by Chow's theorem, which asserts that there exists at least one admissible curve in $\mathbb{G}$ from the identity to x. For a proof of Chow's theorem see [2].*

Remark 3 *We can see that $d(x \cdot y) \leq d(x) + d(y)$ using right invariance of the vector fields. Indeed, let γ_1 be an admissible curve from e to y, and γ_2 be an admissible curve from e to x. Then by right invariance $\gamma_2 \cdot y$ is an admissible curve from y to $x \cdot y$, with length equal to $|\gamma_2|$. Therefore γ, formed by γ_1 followed by $\gamma_2 \cdot y$ (re-parametrising), is an admissible curve from e to $x \cdot y$. Since $|\gamma| = |\gamma_1| + |\gamma_2|$, it is clear from the definition that $d(x \cdot y) \leq |\gamma_1| + |\gamma_2|$. Taking infimums over γ_1 and γ_2 gives us the result.*

Now, define $\nabla f = (X_1 f, \ldots, X_m f)$ to be the *sub-gradient* of a suitable function f (see section 3 for more details). We have

Proposition 1 $|\nabla d| \leq 1$ *a.e.*

PROOF. Note that $\forall\ i$, we have $t \to \exp(tX_i)$ is an admissible curve, and hence by definition $d(\exp(tX_i)) \leq t$. Therefore, using Remark 3,

$$|d(x\exp(tX_i)) - d(x)| \leq t.$$

Thus d is Lipschitz in the direction X_i, and we can adjust Rademacher's classical theorem to the case of Lie groups (see appendix, and [4]), to get that d is almost everywhere differentiable in direction X_i, and $|X_i d| \leq 1$ *a.e.* (with respect to the measure induced on $\mathbb{G}$ by the Lebesgue measure, so that a set in $\mathbb{G}$ is null if and only if its image in all coordinate charts is null in $\mathbb{R}^n$).

We know that $|X_i d| \leq 1$ *a.e.*, but to check that $|\nabla d| \leq 1$ *a.e.* we must handle all directions. Consider $\nabla d.(c_1, \ldots, c_m) = \sum c_i X_i d$, where $\sum c_i^2 = 1$, and the c_i's are rational.

We use the result that if $|v.c| \leq 1$ for all c such that c_i are rational and $\sum c_i^2 = 1$ then $|v| \leq 1$ *a.e.* Since $|\nabla d.c| = (\sum c_i^2 (X_i d)^2)^{\frac{1}{2}} \leq 1$, the result follows.

□

Consider now $\mathbf{B}(r) = \{x \in \mathbb{G} :\ d(x) < r\}$, the unit ball in $\mathbb{G}$ with respect to the distance d. We can recursively form subspaces

$$\begin{aligned} V_1 &= \mathrm{sp}\{X_1, \ldots, X_m\} \\ V_j &= V_{j-1} + [V_1, V_{j-1}], \quad \text{for } j \geq 2, \end{aligned}$$

where $[V_1, V_{j-1}] = \mathrm{sp}\{[Y, Z] \colon Y \in V_1, Z \in V_{j-1}\}$. By Hörmander's condition, there exists smallest n such that V_n is the tangent space at the identity, i.e. V_n is the Lie algebra of $\mathbb{G}$.

Set $d_1 = \mathrm{dim}V_1, d_i = \mathrm{dim}V_i - \mathrm{dim}V_{i-1}$ for $i = 2, \ldots, n$. Then it can be shown that (see [3])

$$\mathrm{Vol}\ \mathbf{B}(r) \approx r^{d_1 + \cdots + d_n}.$$

For the Heisenberg group, $\mathbb{H}$, it can be shown by direct computation that $\partial\mathbf{B}(r)$ is not a smooth submanifold of $\mathbb{H}$. Precisely it has a singularity in the vertical direction Z.

Also note that

$$\delta_s \colon (x, y, z) \mapsto (sx, sy, s^2 z)$$

is a group homomorphism (called a dilation) on $\mathbb{H}$. Therefore $\mathbf{B}(r) = \delta_r(\mathbf{B}(1))$, and $\mathrm{Vol}\ \mathbf{B}(r) = r^4 \mathrm{Vol}\ \mathbf{B}(1)$.

We will concentrate on nilpotent Lie groups, but first of it is useful to look at a more general group.

Example The"$ax + b$" group has group law

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 + x_2, e^{x_2} y_1 + y_2).$$

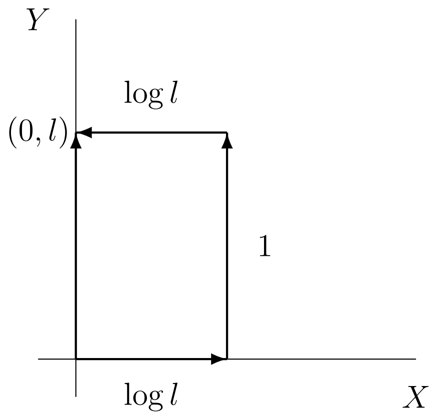
Now we compute the left invariant vector fields, as before

$$\partial_s f((s,0)\cdot(x,y)) = \partial_s f(s+x,y) = \partial_x f$$

$$\partial_s f((0,s)\cdot(x,y)) = \partial_s f(x, se^x + y) = e^x \partial_y f.$$

Hence $X = \partial_x$ and $Y = e^x\partial_y$ are the two left invariant vector fields, (and it is clear that this group is not nilpotent).

Suppose we move a distance l in the Y direction to a point denoted by $(0,l)$, as represented in following diagram.



This is clearly the same as moving a distance $\log l$ in the X direction, then 1 in the Y direction, then $\log l$ in the $-X$ direction. So we can see that $d((0,l)) \leq 2\log l + 1$. Hence, the shortest route from 0 to $(0,l)$ is not necessarily found by going a distance l in the Y direction, as in the Euclidean case. For example, if $l = e^2$ it is clear that it is shorter to take the alternative route, since we have

$$d((0,e^2)) \leq 2\log e^2 + 1 = 5 < e^2.$$

3 L^2-Spaces and Weighted L^2-Spaces

Let $\mathbb{G}$ be a Lie group and let $X_1, \ldots, X_m$ be smooth right invariant vector fields on $\mathbb{G}$. Let dx be the left Haar measure on $\mathbb{G}$ (which is unique up to multiplying by a constant). We take left invariant Haar measure here, since our vector fields are right invariant (see below). Note that in the case of nilpotent Lie groups, any left invariant Haar measure is also right invariant, and vice versa (i.e. the Lie group is unimodular).

Any right invariant vector field X generates a one parameter subgroup

$$\{\exp(tX) : t \in \mathbb{R}\}.$$

Moreover, for any $f \in C^\infty(\mathbb{G})$, we have

$$Xf(x) = \lim_{s\to 0} \frac{f(\exp(sX)x) - f(x)}{s}.$$

Since, dx is left invariant, we have for all i

$$\int_{\mathbb{G}} f(\exp(sX_i)x)g(x)dx = \int_{\mathbb{G}} f(x)g(\exp(-sX_i)x)dx$$

and hence

$$\begin{aligned}
\int_{\mathbb{G}} X_i f(x)g(x)dx &= \int_{\mathbb{G}} \lim_{s\to 0} \frac{f(\exp(sX_i)x)g(x) - f(x)g(x)}{s} dx \\
&= \int_{\mathbb{G}} \lim_{s\to 0} \frac{f(x)g(\exp(-sX_i)x) - f(x)g(x)}{s} dx \\
&= -\int_{\mathbb{G}} f(x) \lim_{s\to 0} \frac{g(\exp(-sX_i)x) - g(x)}{-s} dx \\
&= -\int_{\mathbb{G}} f(x) X_i g(x) dx.
\end{aligned}$$

So $\langle X_i f, g\rangle = -\langle f, X_i g\rangle$.

Recall that $\nabla f = (X_1 f, \ldots, X_m f)$. ∇ can be treated as a closed operator from $L^2(\mathbb{G})$ to $L^2(\mathbb{G}, \mathbb{C}^m)$, the set of vector valued L^2 functions. Indeed ∇ defined on compactly supported smooth functions is closable. Alternatively we get the same operator if we define the domain of ∇ to be all L^2 functions such that $X_i f$ is also in L^2 (as a weak derivative).

Then we define $\Delta = \nabla^* \nabla$, which is clearly self-adjoint. We say that Δ is the *sub-Laplcian* on $\mathbb{G}$.

Remark 4 *$\Delta \approx -\sum X_i^2$, but we do not define it in this way, since it is not clear that this is self-adjoint after closure. Later on we will see that we can indeed write $\Delta = -\sum X_i^2$, but for now we can develop the theory by defining Δ as above.*

Remark 5 *This definition makes sense and gives a self-adjoint operator in the more general setting of manifolds.*

Consider the equation

$$\partial_t f = -\Delta f. \quad (*)$$

In L^2 we can use spectral theory to see that given a self-adjoint operator L and the corresponding resolution of the identity $E(\lambda)$, we have

$$\begin{aligned} e^{-tL} &= \int_0^\infty e^{-t\lambda} dE(\lambda) \\ \Rightarrow \partial_t e^{-tL} &= \int_0^\infty -\lambda e^{-t\lambda} dE(\lambda) = -L \int_0^\infty e^{-t\lambda} dE(\lambda) \\ &= -Le^{-tL}. \end{aligned}$$

Therefore, since Δ is self-adjoint by definition, we have that $f(t,x) = e^{-t\Delta} f_0(x)$ solves equation $(*)$, subject to condition $f(0,x) = f_0(x)$. Note the semigroup relations

$$e^{-t_1\Delta} e^{-t_2\Delta} = e^{-(t_1+t_2)\Delta} = e^{-t_2\Delta} e^{-t_1\Delta},$$

and

$$\partial_t e^{-t\Delta} f = \lim_{s\to 0} \frac{e^{-(t+s)\Delta} f - e^{-t\Delta} f}{s} = -\Delta e^{-t\Delta} f.$$

Now consider the family of inner products on $L^2(\mathbb{G})$ defined by

$$\langle f, g\rangle_s = \langle f, e^{sd} g\rangle$$

where d is the Carnot-Carathéodory distance defined in the previous section. We will study these for $s < 0$.

We show that $e^{sd}g$ remains in the domain of ∇, so that it makes sense to consider $X_i(e^{sd}g)$. It suffices to show this for compactly supported smooth g, since passing to the closure will give the result on whole domain of ∇. We first approximate d by convolutions (see section 4), by noting that there exist smooth ϕ_n such that $d_n = d * \phi_n$ tends to d locally in L^2 and ∇d_n tends to ∇d locally in L^2.

Passing to a subsequence we may assume that ∇d_n tends to ∇d almost everywhere, and so, using the fact that g has compact support in L^2,

$$e^{sd_n} g \to e^{sd} g \text{ uniformly},$$

and

$$\nabla e^{sd_n} g \to \nabla e^{sd} g \text{ almost everywhere}.$$

Now $\nabla e^{sd_n} g$ is bounded (since $s < 0$), so $\nabla e^{sd_n} g \to \nabla e^{sd} g$ in L^2. Since ∇ is closed, therefore $e^{sd}g$ belongs to the domain of ∇. Note that as a by-product we also get the Leibniz formula.

Now we can compute

$$\begin{aligned}\langle X_i f, X_i(e^{sd} g)\rangle &= \langle X_i f, s(X_i d)e^{sd} g\rangle + \langle X_i f, e^{sd} X_i g\rangle \\ &= \langle X_i f, s(X_i d) g\rangle_s + \langle X_i f, X_i g\rangle_s,\end{aligned}$$

and so

$$\begin{aligned}\langle \Delta f, f\rangle_s &= \langle \Delta f, e^{sd} f\rangle \\ &= \langle \nabla f, \nabla(e^{sd} f)\rangle \\ &= \sum_{i=1}^{m} \langle X_i f, X_i(e^{sd} f)\rangle \\ &= \sum_{i=1}^{m} \langle X_i f, X_i f\rangle_s + s \sum_{i=1}^{m} \langle X_i f, (X_i d) f\rangle_s \\ &= \|\nabla f\|_s^2 + s\langle \nabla f, (\nabla d) f\rangle_s.\end{aligned}$$

So

$$|\mathrm{Im}\ \langle \Delta f, f\rangle_s| = |s|\, |\mathrm{Im}\ \langle \nabla f, (\nabla d) f\rangle_s| \leq |s|\, \|\nabla f\|_s \|(\nabla d) f\|_s$$

and since $s < 0$

$$\mathrm{Re}\ \langle \Delta f, f\rangle_s \geq \|\nabla f\|_s^2 + s\|\nabla f\|_s \|(\nabla d) f\|_s$$

Now

$$|s| \|\nabla f\|_s \|(\nabla d) f\|_s \leq \|\nabla f\|_s^2/2 + s^2/2\|f\|_s^2$$

using the inequality $ab \leq a^2/2 + b^2/2$ and the fact that $|\nabla d| \leq 1$. So

$$\begin{aligned}\mathrm{Re}\ \langle \Delta f, f\rangle_s &\geq \|\nabla f\|_s^2 + s\|\nabla f\|_s \|(\nabla d) f\|_s \\ &\geq \|\nabla f\|_s^2 - \|\nabla f\|_s^2/2 - s^2/2\|f\|_s^2 \\ &= \|\nabla f\|_s^2/2 - s^2/2\|f\|_s^2,\end{aligned}$$

which gives

$$\mathrm{Re}\ \langle \Delta f, f\rangle_s + s^2/2\|f\|_s^2 \geq \|\nabla f\|_s^2/2.$$

Finally

$$\begin{aligned}|\mathrm{Im}\ \langle \Delta f, f\rangle_s| &\leq \|\nabla f\|_s^2/2 + s^2/2\|f\|_s^2 \\ &\leq \mathrm{Re}\ \langle \Delta f, f\rangle_s + s^2/2\|f\|_s^2 + s^2/2\|f\|_s^2 \\ &= \mathrm{Re}\ \langle \Delta f, f\rangle_s + s^2\|f\|_s^2. \qquad (1)\end{aligned}$$

Consider now the semigroup $e^{-(\Delta+s^2)t}$.
Note $\langle (\Delta + s^2) f, f\rangle_s \in S_{\frac{\pi}{4}}$, where $S_{\frac{\pi}{4}} = \{z \in \mathbb{C}\colon \arg(z) \in [-\frac{\pi}{4}, \frac{\pi}{4}]\}$. Indeed,

$$\begin{aligned}\mathrm{Re}\ \langle (\Delta + s^2) f, f\rangle_s &= \mathrm{Re}\ \langle \Delta f, f\rangle_s + s^2\|f\|_s^2 \\ &\geq |\mathrm{Im}\ \langle \Delta f, f\rangle_s| = \left|\mathrm{Im}\ \langle (\Delta + s^2) f, f\rangle_s\right|\end{aligned}$$

by above. This shows that the semigroup $e^{-(\Delta+s^2)t}$ is analytic in the sector $S_{\frac{\pi}{4}}$.

We can use (1) to estimate the s-norm of the semigroup by the following calculation.

$$\begin{aligned}
\partial_t \|e^{-t(\Delta+s^2)}f\|_s^2 &= \partial_t \langle e^{-t(\Delta+s^2)}f, e^{-t(\Delta+s^2)}f\rangle_s \\
&= -2\text{Re } \langle(\Delta+s^2)e^{-t(\Delta+s^2)}f, e^{-t(\Delta+s^2)}f\rangle_s \\
&= -2(\text{Re } \langle \Delta g, g\rangle_s + s^2\|g\|_s^2), \quad \text{where } g = e^{-t(\Delta+s^2)}f \\
&\leq -2|\text{Im } \langle \Delta g, g\rangle_s| \leq 0.
\end{aligned}$$

Therefore $\|e^{-t(\Delta+s^2)}f\|_s^2 \leq \|f\|_s^2$, and so

$$\|e^{-t\Delta}\|_s \leq e^{ts^2}.$$

Finally, by duality of $L^2(\mathbb{G}, e^{sd}\mathrm{d}x)$ and $L^2(\mathbb{G}, e^{-sd}\mathrm{d}x)$, we have that this result also holds for $s > 0$.

4 Convolutions

Let μ_1 and μ_2 be finite measures on a Lie group $\mathbb{G}$. Define

$$\int_{\mathbb{G}} h(x)d(\mu_1 * \mu_2)(x) = \int_{\mathbb{G}\times\mathbb{G}} h(y\cdot z)d\mu_1(y)d\mu_2(z).$$

It is clear that $\mu_1 * (\mu_2 * \mu_3) = (\mu_1 * \mu_2) * \mu_3$, and

$$\|\mu_1 * \mu_2\| \leq \|\mu_1\|\|\mu_2\|$$

with equality for positive measures (where the norm is total variation norm on measures).

We use left Haar measure to identify functions with measures and in that way define convolutions of functions

$$\int_{\mathbb{G}} h(x)(f_1 * f_2)(x)dx = \int_{\mathbb{G}\times\mathbb{G}} h(y\cdot z)f_1(y)f_2(z)dydz.$$

Putting $z = y^{-1}x$, and using left invariance we have

$$\int_{\mathbb{G}} h(x)(f_1 * f_2)(x)dx = \int_{\mathbb{G}\times\mathbb{G}} h(x)f_1(y)f_2(y^{-1}x)dydx.$$

So

$$f_1 * f_2(x) = \int_{\mathbb{G}} f_1(y)f_2(y^{-1}x)dy.$$

Therefore we have $(\delta_x * f)(y) = f(x^{-1}y)$, where δ_x is the Dirac delta function. Moreover, neglecting the role of the modular function for notational sake, or assuming the Lie group is unimodular, $(f * \delta_x)(y) = f(yx^{-1})$. Hence

$$Xf(y) = \lim_{s \to 0} \frac{f(\exp(sX)y) - f(y)}{s} = \lim_{s \to 0} \frac{(\delta_{\exp(sX)} * f - f)(y)}{s}.$$

Remark 6 *We can use convolutions to smooth functions on a Lie group, as in the Euclidean case. Indeed, it can be shown that if f is in L^p and $\varphi \in C^\infty(\mathbb{G})$, then $f * \varphi$ is smooth. We sketch this argument.*

Define the function $(R_g f)(x) = f(xg)$. By the continuity of the modular function, if we are working in a compact set K, there exists C such that

$$\|R_g f\|_{L^p} \leq C\|f\|_{L^p}.$$

So if $\varphi \in L^1$ and $\operatorname{supp} \varphi \subset K$, then using the two identities $f * \delta_g = m(g) R_{g^{-1}} f$, where m is the modular function, and $\varphi = \int \varphi(g) \delta_g dg$ for $\varphi \in L^1$, we have

$$\begin{aligned} \|f * \varphi\| &= \|\int \varphi(g) f * \delta_g\| \\ &\leq \|\varphi\|_{L^1} C \|f\|_{L^p}. \end{aligned}$$

We can similarly show that all right derivatives of $f * \varphi$ are bounded, and using the same methods described in section 7 below, this can be used to show smoothness of $f * \varphi$.

Remark 7 *By definition, convolution is only defined for integrable functions. However, we have the the inequality*

$$\|f_1 * f_2\|_{L^\infty} \leq \|f_1\|_{L^1} \|f_2\|_{L^\infty}$$

and hence for $p \geq 1$

$$\|f_1 * f_2\|_{L^p} \leq \|f_1\|_{L^1} \|f_2\|_{L^p},$$

and so we can define convolutions for functions in all L^p spaces.

5 The Heat Kernel

Theorem 1 *Δ is essentially self-adjoint on $C_0^\infty(\mathbb{G})$ i.e. the closure of Δ defined on $C_0^\infty(\mathbb{G})$ by $\Delta = -\sum X_i^2$ in $L^2(\mathbb{G})$ is self-adjoint.*

Proof. We recall the result that if the range of $(I+\Delta)$ is dense in $L^2(\mathbb{G})$, then Δ is essentially self-adjoint (see appendix).

So suppose $(\Delta+I)(C_0^\infty(\mathbb{G}))$ is not dense in $L^2(\mathbb{G})$. Then there exists $f \neq 0 \in L^2(\mathbb{G})$ such that

$$\langle (\Delta+I)\varphi, f\rangle = 0$$

for all $\varphi \in C_0^\infty(\mathbb{G})$.

We may assume that f is smooth (otherwise replace f by $f * \eta$ for $\eta \in C_0^\infty(\mathbb{G})$ such that $f*\eta$ approximates f, and use the fact that Δ is invariant under right translation).

Look at $\psi_R^2 f$ where

$$\psi_R(x) = \begin{cases} 1 - \frac{d(x)}{R} & \text{when } d(x) \leq R \\ 0 & \text{otherwise.} \end{cases}$$

Then $|\nabla \psi_R(x)| \leq \frac{1}{R}$, and $\psi_R^2 f \in D(\nabla)$. By assumption we have

$$0 = \langle (\Delta+I)\varphi, f\rangle = \langle \varphi, f\rangle + \langle \nabla\varphi, \nabla f\rangle$$

for all $\varphi \in C_0^\infty(\mathbb{G})$. Let $\varphi_n \in C_0^\infty(\mathbb{G})$ be such that $\varphi_n \to \psi_R^2 f$ in $L^2(\mathbb{G})$, and $\nabla\varphi \to \nabla(\psi_R^2 f)$ in $L^2(\mathbb{G})$. Note that we can do this using convolutions, since $\psi_R^2 f$ is compactly supported. Hence by continuity we have

$$\begin{aligned} 0 &= \langle \psi_R^2 f, f\rangle + \langle \nabla(\psi_R^2 f), \nabla f\rangle \\ &= \langle \psi_R^2 f, f\rangle + 2\langle (\nabla\psi_R)\psi_R f, \nabla f\rangle + \langle \psi_R^2 \nabla f, \nabla f\rangle \\ &= \|\psi_R f\|^2 + 2\langle (\nabla\psi_R) f, \psi_R \nabla f\rangle + \|\psi_R \nabla f\|^2 \end{aligned} \tag{2}$$

Now since $|\nabla\psi_R(x)| \leq \frac{1}{R}$,

$$|\langle (\nabla\psi_R) f, \psi_R\nabla f\rangle| \leq \frac{1}{R}\|f\|\|\psi_R\nabla f\| \leq \frac{1}{2R^2}\|f\|^2 + \frac{1}{2}\|\psi_R\nabla f\|^2,$$

again using the inequality $ab \leq a^2/2 + b^2/2$.

From (2) we see that $\langle (\nabla\psi_R) f, \psi_R\nabla f\rangle$ is real, and so therefore we have

$$\langle (\nabla\psi_R) f, \psi_R\nabla f\rangle \geq -\frac{1}{2R^2}\|f\|^2 - \frac{1}{2}\|\psi_R\nabla f\|^2.$$

Hence

$$\begin{aligned} 0 &\geq \|\psi_R f\|^2 + 2(-\frac{1}{2R^2}\|f\|^2 - \frac{1}{2}\|\psi_R\nabla f\|^2) + \|\psi_R\nabla f\|^2 \\ &= \|\psi_R f\|^2 - \frac{1}{R^2}\|f\|^2. \end{aligned}$$

Taking the limit as $R \to \infty$ in this we see that we have $0 \geq \|f\|^2$, and thus $f = 0$, which is a contradiction.

$\square$

Corollary 1 *By the above theorem, we can in fact write* $\Delta = -\sum X_i^2$.

In section 3 we defined $\exp(-t\Delta)$ not only on $L^2(\mathbb{G})$ but also on the weighted space $L^2(e^{sd}dx)$. Let us note that if s is big enough then the weighted space $L^2(e^{-sd}dx)$ contains L^∞. So the semigroup is well defined on L^∞. Similarly, if λ is large enough

$$(\Delta + \lambda I)^{-1} = \int_0^\infty \exp(-t\Delta)e^{-t\lambda}dt$$

is continuous from L^∞ to $L^2(e^{-sd}dx)$. We are going to prove that in fact $(\Delta+\lambda I)^{-1}$ and $\exp(-t\Delta)$ are contractions on L^∞.

Now if f is real and smooth, and $f(x_0)$ is maximal, then clearly by the above corollary

$$(\Delta f)(x_0) \geq 0.$$

So

$$\|(\Delta + \lambda I)f\|_{L^\infty} \geq \lambda \|f\|_{L^\infty},$$

for all smooth functions f that are less than any given ε outside a big enough ball, and $\lambda > 0$. Indeed, if f is such a function then there exists an x_0 such that $f(x_0)$ is maximal and then

$$(\Delta f + \lambda f)(x_0) = \Delta f(x_0) + \lambda f(x_0) \geq \lambda f(x_0) = \lambda \|f\|_{L^\infty},$$

and so $\|(\Delta + \lambda I)f\|_{L^\infty} \geq \lambda \|f\|_{L^\infty}$. Then

$$\|(\Delta + \lambda I)^{-1}(\Delta + \lambda I)f\|_{L^\infty} = \|f\|_{L^\infty} \leq \frac{1}{\lambda}\|(\Delta + \lambda I)f\|_{L^\infty}$$

for f in the range of $\Delta + \lambda I$. Since Δ commutes with right translations, the range of $\Delta + \lambda I$ contains all $g = f * \varphi$ where f is in the range of $\Delta + \lambda I$ and $\varphi \in C_0^\infty(\mathbb{G})$. By density of the range of $\Delta + \lambda I$ in $L^2(\mathbb{G})$ this implies that the inequality above remains valid for $f = g * \varphi$ with $g \in L^2(\mathbb{G})$ and $\varphi \in C_0^\infty(\mathbb{G})$. We then approximate $f \in L^2(\mathbb{G}) \cap L^\infty(\mathbb{G})$ by convolutions. Finally, for all $f \in L^\infty$ we can find a sequence of $f_n \in L^2(\mathbb{G}) \cap L^\infty(\mathbb{G})$ such that $\|f_n\|_{L^\infty(\mathbb{G})} \to \|f\|_{L^\infty(\mathbb{G})}$ and f_n converges to f in $L^2(e^{-sd}dx)$, so

$$\|(\Delta + \lambda I)^{-1}\|_{L^\infty} \leq \frac{1}{\lambda} \Rightarrow \|(t\Delta + I)^{-1}\|_{L^\infty} \leq 1$$

for all $t > 0$.

We now define

$$\exp(-t\Delta) = \lim_{n\to\infty}(I + \frac{t}{n}\Delta)^{-n}.$$

Hence by above $\|\exp(-t\Delta)\|_{L^\infty} \leq 1$, i.e. $\exp(-t\Delta)$ is a contraction on L^∞. Note that $\Delta\mathbf{1} = 0$, so $\exp(-t\Delta)\mathbf{1} = \mathbf{1}$. This together with $\|\exp(-t\Delta)\|_{L^\infty} \leq 1$ means that $\exp(-t\Delta)$ maps non-negative functions to non-negative functions. Indeed, let f be a continuous function decaying at infinity, and suppose that $f \geq 0$. Then since $e^{-t\Delta}$ is a contraction on L^∞, we have

$$\begin{array}{rrcl} & \|e^{-t\Delta}(\|f\|_\infty - f)\|_\infty & \leq & \|\|f\|_\infty - f\|_\infty \\ \Rightarrow & e^{-t\Delta}\|f\|_\infty - e^{-t\Delta}f & \leq & \|\|f\|_\infty - f\|_\infty \\ \Rightarrow & \|f\|_\infty - e^{-t\Delta}f & \leq & \|\|f\|_\infty - f\|_\infty \\ \Rightarrow & e^{-t\Delta}f & \geq & \|f\|_\infty - \|\|f\|_\infty - f\|_\infty, \end{array}$$

where we used the fact that $e^{-t\Delta}\mathbf{1} = \mathbf{1}$ to get that $e^{-t\Delta}\|f\|_\infty = \|f\|_\infty$. Now since $f \geq 0$ we have $\|\|f\| - f\|_\infty \leq \|f\|_\infty$, and therefore $e^{-t\Delta}f \geq 0$. So $H_t = e^{-t\Delta}$ is a Markov semigroup on continuous functions decaying at infinity, and is called the *heat diffusion* semigroup.

It follows from the right invariance of Δ that we can write

$$H_t f = \rho_t * f$$

where ρ_t is a non-negative measure, called the *heat kernel* associated to Δ on $L^2(\mathbb{G}, dx)$.

We want to study this heat kernel, and our aim for the remainder of these notes will be to deduce bounds for $\rho_t(x)$, and then to look at its smoothness properties. For a detailed treatment of the heat kernel in a general setting, see [1].

6 Two-sided Estimate for the Heat Kernel

In this section we closely follow [5] which should be referred to for the details. Another account of these methods with more emphasis in Lie groups can be found in [7]. We deduce that the so called scale-invariant parabolic Harnack principle implies that the heat kernel satisfies the following two sided Gaussian estimate

$$\frac{1}{|\mathbf{B}(e, t^{\frac{1}{2}})|} c_1 \exp\left(\frac{-d^2(x)}{c_2 t}\right) \leq \rho_t(x) \leq \frac{1}{|\mathbf{B}(e, t^{\frac{1}{2}})|} c_3 \exp\left(\frac{-d^2(x)}{c_4 t}\right),$$

where $\mathbf{B}(x, r)$ is ball of radius r centered at x $(= \mathbf{B}(r) \cdot x)$. In fact this two-sided bound is equivalent to the parabolic Harnack principle and holds if and only if the following two properties are satisfied:

- there exists a constant C such that volume growth has the following doubling property
$$\forall x \in \mathbb{G}, \forall r > 0, |\mathbf{B}(x, 2r)| \leq C|\mathbf{B}(x, r)|$$

- there exists a constant P such that the Poincaré inequality

$$\forall x \in \mathbb{G}, \forall r > 0, \int_{\mathbf{B}(x,r)} \left|f - Av_{\mathbf{B}(x,r)} f\right|^2 dx \leq Pr^2 \int_{\mathbf{B}(x,2r)} |\nabla f|^2 dx,$$

is satisfied, where $A_{\mathbf{B}(x,r)} f$ is the average of f over the ball $\mathbf{B}(x, r)$.

We will also show that a very general class of Lie groups satisfy these two conditions.

6.1 Poincaré Inequality

The Poincaré inequality in a ball states that

$$\int_{\mathbf{B}(x,r)} \left|f - Av_{\mathbf{B}(x,r)} f\right|^2 dx \leq Pr^2 \int_{\mathbf{B}(x,2r)} |\nabla f|^2 dx,$$

for all $f \in C^\infty(\mathbf{B}(x, r))$ and some $P > 0$, where $Av_{\mathbf{B}(x,r)} f$ is the average of f over the ball $\mathbf{B}(x, r)$. One crucial aspect of the inequalities is that they are assumed to hold for all $f \in C^\infty(\mathbf{B}(x, r))$, instead of merely $f \in C_0^\infty(\mathbf{B}(x, r))$. So the natural question to ask is for what kind of Lie groups does this hold?

We have the following theorem.

Theorem 2 *Suppose $\mathbb{G}$ is unimodular. Then*

1. *For all $0 < r \leq 1$ and all balls $\boldsymbol{B}$ of radius r, the Poincaré inequality*

$$\forall f \in C^\infty(\boldsymbol{B}), \int_{\boldsymbol{B}} |f - Av_{\boldsymbol{B}} f|^2 dx \leq Pr^2 \int_{\boldsymbol{B}} |\nabla f|^2 dx$$

is satisfied.

2. *If there exists C such that the volume growth doubling condition is satisfied then the Poincaré inequality*

$$\forall f \in C^\infty(\boldsymbol{B}), \int_{\boldsymbol{B}} |f - Av_{\boldsymbol{B}} f|^2 dx \leq Pr^2 \int_{\boldsymbol{B}} |\nabla f|^2 dx$$

is satisfied for all balls $\boldsymbol{B}$ of radius $r > 0$.

This theorem follows from the following lemma.

Lemma 1 *If $d(x) < r$, then $\|(\delta_x - \delta_e) * f\|_{L^p(\boldsymbol{B}(e,r))} \leq r\|\nabla f\|_{L^p(\boldsymbol{B}(e,2r))}$.*

PROOF.
Let $\gamma(s)$ be an admissible curve from e to x of length smaller then r.
Fix s_0, and write $\gamma(s) = \gamma(s_0)(\gamma(s_0)^{-1}\gamma(s))$. Then

$$\|\partial_s(\delta_{\gamma(s)} * f)\|_{L^p(\mathbf{B}(e,r))} = \|\partial_s \delta_{\gamma(s_0)^{-1}\gamma(s)} * f\|_{L^p(\mathbf{B}(\gamma(s_0),r))}$$

and

$$\partial_s(\delta_{\gamma(s_0)^{-1}\gamma(s)} * f)(y) = \partial_s f(\gamma(s)^{-1}\gamma(s_0)y).$$

Evaluated at $s = s_0$ this is equal to $(\gamma'(s_0)f)(y) = -\sum a_i(s_0)(X_i f)(y)$. So

$$\|(\delta_x - \delta_e) * f\|_{L^p(\mathbf{B}(e,r))} \leq \|\sum a_i(s_0)(X_i f)(\cdot)\|_{L^p(\mathbf{B}(e,2r))} \leq r\|\nabla f\|_{L^p(\mathbf{B}(e,2r))}.$$

$\square$

We now prove a weaker form of the result stated in Theorem 2.

PROOF. We aim to prove that

$$\|f - Av_{\mathbf{B}(e,r)} f\|_{L^2(\mathbf{B}(e,r))} \leq Pr\|\nabla f\|_{L^2(\mathbf{B}(e,4r))}.$$

This inequality follows from a more general inequality, valid for all $1 \leq p \leq \infty$:

$$\|f - Av_{\mathbf{B}(e,r)} f\|_{L^p(\mathbf{B}(e,r))} \leq Pr\|\nabla f\|_{L^p(\mathbf{B}(e,4r))}.$$

We will prove the inequality above using interpolation between $p = 1$ and $p = \infty$.

Using the lemma with $p = \infty$ we have for $x \in \mathbf{B}(x, r)$

$$|f(x) - f(e)| \leq \|\delta_x * f - f\|_{L^\infty(\mathbf{B}(e,r))} \leq r\|\nabla f\|_{L^\infty(\mathbf{B}(e,2r))}$$

so

$$\begin{aligned}\|f - Av_{\mathbf{B}(e,r)} f\|_{L^\infty(\mathbf{B}(e,2r))} &\leq \|f - f(e)\|_{L^\infty(\mathbf{B}(e,2r))} + |f(e) - Av_{\mathbf{B}(e,r)} f| \\ &\leq 2\|f - f(e)\|_{L^\infty(\mathbf{B}(e,2r))} \leq 2r\|\nabla f\|_{L^\infty(\mathbf{B}(e,2r))} \\ &\leq 2r\|\nabla f\|_{L^\infty(\mathbf{B}(e,4r))}.\end{aligned}$$

For $p = 1$ we have

$$\begin{aligned}\|f - Av_{\mathbf{B}(e,r)} f\|_{L^1(\mathbf{B}(e,r))} &\leq \frac{1}{|\mathbf{B}(e,r)|} \int_{\mathbf{B}(e,r)\times\mathbf{B}(e,r)} |f(x) - f(y)| dx dy \\ &= \frac{1}{|\mathbf{B}(e,r)|} \int_{\mathbf{B}(e,r)} \int_{\mathbf{B}(y^{-1},r)} |f(gy) - f(y)| dg dy \\ &\leq \frac{1}{|\mathbf{B}(e,r)|} \int_{\mathbf{B}(e,2r)} \int_{\mathbf{B}(e,2r)} |f(gy) - f(y)| dy dg\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{|\mathbf{B}(e,r)|} \int_{\mathbf{B}(e,2r)} \|\delta_g * f - f\|_{L^1(\mathbf{B}(e,2r))} \\
&\leq \frac{1}{|\mathbf{B}(e,r)|} \int_{\mathbf{B}(e,2r)} 2r \|\nabla f\|_{L^1(\mathbf{B}(e,4r))} \\
&= 2r \frac{|\mathbf{B}(e,2r)|}{|\mathbf{B}(e,r)|} \|\nabla f\|_{L^1(\mathbf{B}(e,4r))}.
\end{aligned}
$$

The first equality is valid because the group is unimodular. The last inequality follows from the lemma with $p = 1$. $\square$

Remark 8 *The Poincaré inequality as stated in Theorem 2 implies the doubling property. On the other hand, given the doubling property, the weaker version we gave above implies inequality in Theorem 2 (see Lemma 5.3.1. in [5]).*

Note that the doubling property clearly implies that

$$
\left|\mathbf{B}(x, 2^k r)\right| \leq C^k \left|\mathbf{B}(x,r)\right|
$$

which shows $|\mathbf{B}(x,r)|$ grows at a polynomial rate. Essentially this bounds the dimension of the problem.

Relating this to our examples in the first section, we see the property holds for the Heisenberg group, SO(3), and the motion group of the plane. In fact this property is fairly straightforward to show for nilpotent Lie groups, using dilations. For our two groups of exponential growth (the"$ax + b$" group and $\mathrm{SL}(2,\mathbb{R})$), this property is not true for all r, but the following restricted doubling condition is satisfied:

$$
\exists R > 0 \textit{ and } C \textit{ such that } \forall x \in \mathbb{G}, \forall r \in (0,R), |\mathbf{B}(x,2r)| \leq C|\mathbf{B}(x,r)|.
$$

Thus on small scales the result holds for these groups too. [7] should be referred to for more details on the treatment of exponential growth Lie groups.

6.2 Local Sobolev Inequality

Very generally, we say that a Lie group $\mathbb{G}$ (of dimension n) satisfies a localized L^p Sobolev inequality with constant $C(\mathbf{B})$ and exponent $\nu > n$ if, for any geodesic ball $\mathbf{B}$ and all $f \in C_0^\infty(\mathbf{B})$

$$
\left(\int |f|^q dx\right)^{p/q} \leq C(\mathbf{B}) \frac{r(\mathbf{B})^p}{|\mathbf{B}|} \int (|\nabla f|^p + r(\mathbf{B})^{-p}|f|^p) dx
$$

where $q = p\nu/(\nu - p)$, and $r(\mathbf{B})$ is radius of ball $\mathbf{B}$.

The main result we need is the following:

Theorem 3 *Fix $1 \leq p < \infty$ and $0 < R \leq \infty$. Assume that $\mathbb{G}$ satisfies a scale invariant L^p Poincaré inequality up to radius R, and assume further that $\mathbb{G}$ satisfies the restricted doubling condition*

$$\forall x \in \mathbb{G}, \forall r \in (0, R), |\boldsymbol{B}(x, 2r)| \leq C|\boldsymbol{B}(x, r)|.$$

Then for any $K > 1$, there exist $\nu > p$ and D such that for any ball $\boldsymbol{B}$ of radius less than KR,

$$\left(\int |f|^{p\nu/(\nu-p)} \mathrm{d}x\right)^{(\nu-p)/\nu} \leq D(\boldsymbol{B}) \frac{r(\boldsymbol{B})^p}{|\boldsymbol{B}|} \int |\nabla f|^p \mathrm{d}x.$$

In other words the Poincaré inequality and the restricted doubling property imply a family of local Sobolev inequalities. Therefore by section 6.1 we have that any unimodular Lie group which satisfies the restricted doubling condition satisfies a family of local Sobolev inequalities.

Remark 9 *It can be shown that local Sobolev inequalities imply certain upper bounds on the heat kernel. However, Gaussian lower bounds cannot be obtained from such Sobolev inequalities alone. The lower bound estimate we are aiming for depends crucially on the Poincaré inequality.*

6.3 Scale-invariant Parabolic Harnack Inequality

For any $x \in \mathbb{G}$ and $s \in \mathbb{R}$, $r > 0$, let $Q = Q(x, s, r)$ be the cylinder

$$Q(x, s, r) = (s - r^2, s) \times \mathbf{B}(x, r).$$

Let Q_+, Q_- be respectively the upper and lower sub-cylinders

$$\begin{aligned} Q_+ &= (s - (1/4)r^2, s) \times \mathbf{B}(x, (1/2)r) \\ Q_- &= (s - (3/4)r^2, s - (1/2)r^2) \times \mathbf{B}(x, (1/2)r) \end{aligned}$$

We say that $\mathbb{G}$ satisfies a scale invariant parabolic Harnack principle if there exists a constant A such that for any $x \in \mathbb{G}, s \in \mathbb{R}, r > 0$ and any positive solution u of $(\partial_t + \Delta)u = 0$ in $Q = Q(x, s, r)$ we have

$$\sup_{Q_-}\{u\} \leq A \inf_{Q_+}\{u\}.$$

It turns out this scale-invariant Harnack principle carries a lot of information. In particular we have the following important theorem.

Theorem 4 *Fix $0 < R \leq \infty$ and consider the following properties:*

1. *There exists P such that for any ball $\boldsymbol{B} = \boldsymbol{B}(x,r), x \in \mathbb{G}, 0 < r < R$, and for all $f \in C^\infty(\boldsymbol{B})$,*

$$\int_{\boldsymbol{B}} |f - Av_{\boldsymbol{B}} f|^2 dx \leq Pr^2 \int_{\boldsymbol{B}} |\nabla f|^2 dx.$$

2. *There exists C such that, for any ball $\boldsymbol{B} = \boldsymbol{B}(x,r), x \in \mathbb{G}, 0 < r < R$,*

$$|\boldsymbol{B}(x, 2r)| \leq C|\boldsymbol{B}(x,r)|.$$

3. *There exists a constant A such that, for any ball $\boldsymbol{B} = \boldsymbol{B}(x,r), x \in \mathbb{G}, 0 < r < R$ and for any smooth positive solution u of $(\partial_t + \Delta)u = 0$ in the cylinder $(s - r^2, s) \times \boldsymbol{B}(x,r)$, we have*

$$\sup_{Q_-}\{u\} \leq A \inf_{Q_+}\{u\}.$$

with

$$\begin{aligned} Q_+ &= (s - (1/4)r^2, s) \times \boldsymbol{B}(x, (1/2)r) \\ Q_- &= (s - (3/4)r^2, s - (1/2)r^2) \times \boldsymbol{B}(x, (1/2)r). \end{aligned}$$

Then the conjunction of 1 and 2 is equivalent to 3.

Our final result in this section is the following.

Theorem 5 *Fix $0 < R \leq \infty$. The heat kernel $\rho_t(x)$ satisfies the two-sided Gaussian inequality*

$$\frac{1}{|\boldsymbol{B}(e, t^{\frac{1}{2}})|} c_1 \exp\left(\frac{-d^2(x)}{c_2 t}\right) \leq \rho_t(x) \leq \frac{1}{|\boldsymbol{B}(e, t^{\frac{1}{2}})|} c_3 \exp\left(\frac{-d^2(x)}{c_4 t}\right)$$

for all $x \in \mathbb{G}$ and $t \in (0, R)$ if and only if $\mathbb{G}$ satisfies a scale-invariant parabolic Harnack principle.

Corollary 2 *Combining Theorems 2, 4 and 5, we therefore have that the heat kernel ρ_t on any unimodular Lie group which satisfies the restricted doubling condition, satisfies the two-sided Gaussian inequality described above.*

7 Smoothness of the Heat Kernel

Let $X_1, \ldots, X_m$ be right invariant vector fields on a connected Lie group $\mathbb{G}$ which generate the Lie algebra i.e. which satisfy Hörmander's condition, and let $\Delta = -\sum X_i^2$, as above. Let d be the Carnot-Carathéodory distance from e. From section 3 we have that $e^{-t\Delta}$ acts on the weighted L^2 space $L^2(e^{sd})$ as an analytic semigroup in a sector. We can therefore estimate the derivative of this semigroup to get that

$$\sum_i \|X_i e^{-t\Delta} f\|_{L^2(e^{sd})} \leq c_{t,s} \|f\|_{L^2(e^{sd})}. \tag{3}$$

for some constant $c_{t,s}$, which grows exponentially with t. Indeed, using our estimate $\|e^{-t\Delta}\|_s \leq e^{ts^2}$ from section 3, we can estimate $\|\Delta e^{-t\Delta} f\|_{L^2(e^{sd})}$, and then since $\Delta = -\sum X_i^2$ and using integration by parts, we get the result. Cauchy's integral formula shows that result holds in the complex sector.

Remark 10 *In these notes we consider Δ with second order terms only. To consider a Δ with first order drift terms as well, one needs more complicated arguments than the ones we present here.*

Remark 11 *The route we take here to get smoothness of the heat kernel avoids the use of Hörmander's theorem. However, it should be noted that Hörmander's theorem gives us smoothness of the heat kernel as a direct consequence of the hypoellipticity of Δ.*

7.1 Left and Right Derivatives

We want now to compare the left and right derivatives of a smooth function f. Assume X is a right invariant vector field, and let γ be a curve with $\gamma(0) = e$ such that $\gamma'(0) = X$. Then

$$\begin{aligned} (Xf)(g) &= \partial_s f(\gamma(s)g)|_{s=0} = \partial_s (gg^{-1}\gamma(s)g)|_{s=0} \\ &= \partial_s f(gA_{g^{-1}}\gamma(s))|_{s=0} \end{aligned}$$

where, for $g, x \in \mathbb{G}$, $A_g x = g^{-1}xg$. Thus A_g is a smooth map $\mathbb{G} \to \mathbb{G}$.

Let Ad_g be the derivative of A_g with respect to x at the identity e (treating $\mathbb{G}$ as a differentiable manifold), so that Ad_g is an isomorphism of Lie algebras, $Ad_g \colon \mathfrak{g} \to \mathfrak{g}$.

For right invariant vector field $X \in \mathfrak{g} \cong T_e\mathbb{G}$, we will denote the corresponding left invariant vector field $\widetilde{X}$. Then by the calculation above we have

$$(Xf)(g) = \partial_s f(gA_{g^{-1}}\gamma(s))|_{s=0} = \widetilde{Ad_{g^{-1}}X} f(g). \tag{4}$$

Lemma 2 *We can introduce a norm on $T_e\mathbb{G}$ (as it is a finite dimensional vector space) and define the norm of Ad_g. Then one has that*

$$\|Ad_g\| \leq e^{cd(g)} \tag{5}$$

for some constant c.

Proof. Fix a scalar product on the Lie algebra ($T_e\mathbb{G}$).

Let γ be an admissible curve joining e with g. We may re-parametize γ such that $\gamma'(s) = \sum a_i(s)X_i$ and $\sum a_i(s)^2 = 1$. With respect to our fixed scalar product $|\gamma'(s)| \leq C_1$. Ad is a smooth map and norm is submultiplicative, so

$$\partial_s \|Ad_{\gamma(s)}\| \leq \|DAd_{\gamma(s)}\||\gamma'(s)|$$

$$\leq \|Ad_{\gamma(s)}\|\|DAd_e\||\gamma'(s)| \leq c\|Ad_{\gamma(s)}\|$$

where D is the full derivative, and c is some constant. Therefore

$$\begin{aligned} \partial_s \log \|Ad_{\gamma(s)}\| &\leq c, \\ \Rightarrow \log \|Ad_{\gamma(s)}\| &\leq cs, \\ \Rightarrow \|Ad_{\gamma(s)}\| &\leq \exp(cs). \end{aligned}$$

Note that norm is a Lipschitz function, so all the functions above are absolutely continuous, and such that the differential inequalities make sense a.e. Hence the final claims holds for all s.

Recall that $d(g)$ is an infimum of s such that $\gamma(s) = g$. Hence we also have

$$\|Ad_g\| \leq \exp(cd(g)),$$

as required. □

The map $g \mapsto Ad_g$ is a map of Lie groups

$$Ad\colon \mathbb{G} \to \mathbf{GL}(\mathfrak{g}),$$

where $\mathbf{GL}(\mathfrak{g})$ is the general linear group of bijective linear transformations $\mathfrak{g} \to \mathfrak{g}$ with composition as the group operation, and is called the *adjoint representation of* $\mathbb{G}$.

We can then take the derivative of Ad at the identity to get a map of Lie algebras,

$$ad : \mathfrak{g} \to \mathfrak{gl}(\mathfrak{g})$$

where $\mathfrak{gl}(\mathfrak{g})$ is Lie algebra of all linear maps on $\mathfrak{g}$, which is called the *adjoint representation of* $\mathfrak{g}$.

It can be verified that

$$ad_X Y = -[X, Y] := -(XY - YX)$$

for all $X, Y \in \mathfrak{g}$. Indeed, we have that that the exponential map is a diffeomorphism from a neighbourhood of 0 in $\mathfrak{g}$ onto a neighbourhood of the identity in $\mathbb{G}$, and

$$ad_X = \frac{d}{dt} Ad(e^{tX})\Big|_{t=0}.$$

Thus,

$$ad_X(Y) = \frac{d}{dt} Ad(e^{tX})(Y)\Big|_{t=0} = \frac{d}{dt} e^{-tX} Y e^{tX}\Big|_{t=0} = -[X, Y].$$

Moreover, one can similarly show by direct computation that

$$Ad_{\exp X} = \exp(ad_X)$$

.

One can also check that the Jacobi identity on $\mathfrak{g}$ is equivalent to the fact that ad preserves Lie brackets i.e. ad is such that

$$ad_{[X,Y]} = [ad_X, ad_Y]$$

for all $X, Y \in \mathfrak{g}$ (where on the right, the Lie bracket is the commutator of linear maps on $\mathfrak{g}$). Thus we can recover the Lie bracket from ad.

Remark 12 *Using the Campbell-Baker-Hausdorff formula and the exponential mapping we can locally reconstruct the group multiplication from the Lie algebra (see appendix). In general, for a simply connected Lie group there is exactly one group that corresponds to the associated Lie algebra.*

7.2 Sobolev Spaces

Returning to our heat semigroup $e^{-t\Delta}$ we first recall that $e^{-t\Delta} f = \rho_t * f$ for functions f decaying at infinity. Therefore we have $e^{-t\Delta}\delta_e = \rho_t * \delta_e = \rho_t$ by section 4. Moreover,

$$\begin{gathered} e^{-t\Delta}\delta_e \in L^2(e^{sd}) \\ \Rightarrow X_i e^{-2t\Delta}\delta_e \in L^2(e^{sd}), \end{gathered}$$

by (3).

We also have, assuming the group is unimodular, that

$$\sup |e^{sd}(f * g)| \leq \|f\|_{L^2(e^{sd})} \|g\|_{L^2(e^{sd})}. \tag{6}$$

Note that if the group is not unimodular, we have a similar estimate involving the modular function. This is related to Remarks 6 and 7. For general groups, by Remark 7, convolution is bounded on $L^1 \times L^\infty$ into L^∞. If the group is unimodular, then convolution is also bounded on $L^\infty \times L^1$ (into L^∞) and hence by interpolation, is bounded from $L^2 \times L^2$ into L^∞. On general groups one has to incorporate modular functions into the norm inequalities if the left factor is not in L^1.

We therefore have that

$$X_i e^{-2t\Delta}\delta_e \in L^2(e^{sd}) \Rightarrow e^{sd}|X_i e^{-2t\Delta}\delta_e| \in L^\infty,$$

and so $e^{-2tL}\delta_e$ is locally Hölder continuous with respect to the optimal control distance.

We now need to introduce some Sobolev type spaces. Let $Y_1, \dots, Y_m$ be right invariant vector fields which form the basis of the Lie algebra. We introduce the space $H_{1,s}$, which is defined to be the closure of $C^\infty(\mathbb{G})$ with respect to the norm

$$\left(\sum_i \|e^{sd} Y_i f\|^2_{L^2(dx)} + \|e^{sd} f\|^2_{L^2(dx)} \right)^{1/2}.$$

We now use interpolation theory (see [6] for details) to define $H_{\varepsilon,s}$ for all $\varepsilon > 0$, (in the same way that we arrive at classical fractional Sobolev spaces on $\mathbb{R}^n$). The optimal control distance d is comparable with the Euclidean distance at the identity, and so $e^{-2t\Delta}\delta_e$ is Hölder continuous with respect to the Euclidean distance at the identity. Thus by the classical theory of Sobolev spaces we have

$$\|e^{-2t\Delta}\delta_e\|_{H_{\varepsilon,s}} < \infty$$

for some $\varepsilon > 0$. Indeed, this uniform estimate can be obtained by covering the Lie group in small balls, shifting them to the origin, applying the classical theory at the origin, and then shifting them back.

Now for left invariant vector fields, we can similarly introduce spaces $\widetilde{H}_{\varepsilon,s}$ for $\varepsilon > 0$. Define $\widetilde{H}_{1,s}$ to be the closure of $C^\infty(\mathbb{G})$ with respect to the norm

$$\left(\sum_i \|e^{sd} \widetilde{Y}_i f\|^2_{L^2(dx)} + \|e^{sd} f\|^2_{L^2(dx)} \right)^{1/2}.$$

We then have

$$H_{1,s+c} \subset \widetilde{H}_{1,s}.$$

Indeed, using (4) and (5) for $f \in H_{1,s+c}$ we have

$$\sum_i \|e^{sd}\widetilde{X}_i f\|^2_{L^2} + \|e^{sd} f\|^2_{L^2} \leq \sum_i \|e^{sd} e^{cd} X_i f\|^2_{L^2} + \|e^{sd} e^{cd} f\|^2_{L^2},$$

and so $f \in \widetilde{H}_{1,s}$. Thus
$$e^{-2t\Delta}\delta_e \in \widetilde{H}_{\varepsilon,s-c}.$$

Differentiation from the left and right commute, and hence

$$(X_i f) * (\widetilde{X}_j h) = X_i \widetilde{X}_j (f * h),$$

for all i, j. Thus, if $f \in H_{1,s}$ and $h \in \widetilde{H}_{1,s}$, then for all i, j we have by (6)

$$\sup |e^{sd} X_i \widetilde{X}_j (f * h)| = \sup |e^{sd}(X_i f) * (\widetilde{X}_j h)| < \infty,$$

and so $e^{sd} X_i \widetilde{X}_j (f * h) \in L^\infty$. Again using (4) and (5), this in turn implies that for all i, j, $e^{(s-c)d} X_i \widetilde{X}_j (f * h) \in L^2$, from which we conclude that $f * h \in \widetilde{H}_{2,s-c}$. Finally, using interpolation theory once again, we have the result that if $f \in H_{\varepsilon_1,s}$ and $h \in \widetilde{H}_{\varepsilon_2,s}$, then
$$f * h \in \widetilde{H}_{\varepsilon_1+\varepsilon_2,s-c}.$$
Loosely, this shows that we get "more" smoothness by taking convolutions and moving all the derivatives to one side. Then by induction we can see that for all n,
$$e^{-2nt\Delta}\delta_e \in \widetilde{H}_{n\varepsilon,s-nc}.$$
Thus we can estimate all derivatives of the heat kernel, and so we get smoothness.

7.3 A Better Estimate

We have by above that
$$|X^I e^{-t\Delta}\delta_e| \leq c_{s,t} e^{-sd}$$
where $X^I = X^{I_1} \dots X^{I_{|I|}}$, and X^{I_k} is a vector field generated by $X_1, \dots, X_m$ i.e. X^{I_k} is formed by taking commutators of $X_1, \dots, X_m$. In fact we can improve this estimate as $t \to 0$.

Consider new vector fields $X_i' = t^{1/2} X_i$, and new operator $\Delta' = -\sum (X_i')^2 = t\Delta$. Also define $d'(x) = t^{-1/2} d(x)$. As mentioned in section 2 we can recursively form subspaces

$$\begin{aligned} V_1 &= \mathrm{sp}\{X_1, \dots, X_m\} \\ V_j &= V_{j-1} + [V_1, V_{j-1}], \quad \text{for } j \geq 2. \end{aligned}$$

This is called a *filtration* of the Lie algebra generated by $X_1, \dots X_m$. Note that if $W_1 = V_1$ and W_k is the complement to V_{k-1} in V_k, then we have

$$V_k = W_1 \oplus W_2 \oplus \dots \oplus W_k$$

for all $k \geq 1$. In terms of this decomposition, we define a one parameter group of *dilations* $\{\eta_t \colon t > 0\}$ on $\oplus_k W_k$ by

$$\eta_t(\sum x_k) = \sum t^k x_k \quad (x_k \in W_k).$$

After dilations, the vectors change, but one has uniform control on the Lie brackets. Indeed, it can be shown that

$$|X^I e^{-t\Delta}\delta_e| \leq C t^{-Q/2} t^{-|I|/2} \exp\left(\frac{-cd^2}{t}\right)$$

where $Q = \sum_k k\dim W_k$.

For more general functions, note that we have

$$\Delta^{-1} = \int_0^\infty e^{-t\Delta} dt$$

and

$$(\Delta + I)^{-1} = \int_0^\infty e^{-t(\Delta+I)} dt$$

Then it is possible to show that

$$\|X^I(\Delta + I)^{-1}\delta_e\| \leq C(d^2)^{-Q/2-|I|/2+1} \sim d^{-Q-|I|+2}.$$

We can conclude that

$$\|X^I f\|_{L^2} \leq \|(\Delta + I)^{|I|/2} f\|.$$

Remark 13 *In general the estimate $\|X_i X_j f\|_{L^2} \leq \|\Delta f\|$ does not hold, for example it does not hold on the motion group of the plane. It does, however, hold on nilpotent Lie groups.*

8 Appendix

8.1 Manifolds

Let M be a topological space. A n-dimensional *chart* on M is any couple (U, φ) where U is an open subset of M and φ is a homeomorphism of U onto an open subset of $\mathbb{R}^n$. Thus a chart is an open set U in M where a local coordinate system is defined. A C-manifold of dimension n is a Hausdorff topological space with a countable base such that each of its points belongs to an n-dimensional chart. A family $\mathcal{A}$ of charts on a C-manifold is called a C^k-*atlas* (where k is a positive integer or $+\infty$) if the charts from $\mathcal{A}$ cover M, and the change of coordinates in the intersection of any two charts from $\mathcal{A}$ is given by C^k-functions. Two C^k-atlases are said to be compatible if their union is again a C^k-atlas. The union of all compatible C^k-atlases determines a C^k-structure on M. A *differentiable manifold* is a couple $(M, \mathcal{A})$ where M is a C-manifold and $\mathcal{A}$ is a C^∞-atlas on M.

A mapping $\xi\colon C^\infty(M) \to \mathbb{R}$ is a *tangent vector* to M at $p \in M$ if it is linear and for all $f, g \in C^\infty(M)$

$$\xi(fg) = \xi(f)g(p) + \xi(g)f(p).$$

The set of all such tangent vectors is denoted T_pM.

A smooth *vector field* X on M is an assignment of a tangent vector $X_p \in T_pM$ to each point $p \in M$ where for all $f \in C^\infty(M)$ the function $Xf\colon M \to \mathbb{R}$ given by $Xf(p) = X_pf$ is smooth.

Let M, N be two smooth manifolds. The derivative at $p \in M$ of the smooth map $F\colon M \to N$ is the homomorphism of tangent spaces

$$DF_p\colon T_pM \to T_{F(p)}N$$

defined by

$$DF_p(\xi)(f) = \xi(f \circ F)$$

for $f \in C^\infty(N)$.

8.2 Lie Groups and Lie Algebras

A Lie group $\mathbb{G}$ is a group which is also a differentiable manifold such that the group operations $(g, h) \mapsto gh$ and $g \mapsto g^{-1}$ are smooth maps. For each $g \in \mathbb{G}$, clearly the maps $\mathbf{L}_g$ and $\mathbf{R}_g$ from $\mathbb{G}$ to $\mathbb{G}$ defined by

$$\begin{aligned} \mathbf{L}_g(h) &= gh \\ \mathbf{R}_g(h) &= hg \end{aligned}$$

are diffeomorphisms of $\mathbb{G}$. We say the maps $\mathbf{L}_g$ and $\mathbf{R}_g$ are left and right translations on $\mathbb{G}$ respectively.

Recall that given a vector field X on $\mathbb{G}$ and diffeomorphism $k\colon \mathbb{G} \to \mathbb{G}$, we can define a new vector field on $\mathbb{G}$ using k in the following way. Firstly we define the push-forward map between tangent bundles $k_*\colon T\mathbb{G} \to T\mathbb{G}$: for $g \in \mathbb{G}, \xi \in T_g\mathbb{G}$ let

$$(k_*\xi)(f) = \xi(f \circ k)$$

for all $f \in C^\infty(\mathbb{G})$, so that $k_*\xi \in T_{k(g)}\mathbb{G}$. Now we can define the vector field k_*X by setting

$$(k_*X)(k(g)) = k_*(X(g))$$

for all $g \in \mathbb{G}$, which is well defined since k is a diffeomorphism. If X and Y are two vector fields on $\mathbb{G}$ and are such that $k_*(X(g)) = Y(k(g))$ then we write $Y = h_*X$.

A smooth vector field X on $\mathbb{G}$ is said to be *left invariant* (resp. right invariant) if for any $g \in \mathbb{G}$,

$$\mathbf{L}_{g*}X = X$$

(resp. $\mathbf{R}_{g*}X = X$), i.e if for all $g \in \mathbb{G}$, $\mathbf{L}_{g*}(X(h)) = X(gh)$, for all $h \in \mathbb{G}$ (resp. $\mathbf{R}_{g*}(X(h)) = X(hg)$ for all $h \in \mathbb{G}$).

The set of all left (or right) invariant vector fields on $\mathbb{G}$ is called the Lie algebra of $\mathbb{G}$, and is denoted $\mathfrak{g}$. The map $X \mapsto X(e)$ is a linear isomorphism between $\mathfrak{g}$ and $T_e\mathbb{G}$, the tangent space to $\mathbb{G}$ at the identity. So $\mathfrak{g}$ is a finite dimensional vector space whose dimension is the same as that of $\mathbb{G}$. A critical observation is that if X, Y are in $\mathfrak{g}$ then the commutator $[X, Y] = XY - YX$ is also in $\mathfrak{g}$ i.e $\mathfrak{g}$ is closed under the bracket operation.

Let X be a smooth vector field on $\mathbb{G}$. The *integral curve* of X passing through point $g \in \mathbb{G}$ is a curve $t \mapsto \sigma(t)$ in $\mathbb{G}$ such that $\sigma(0) = g$ and

$$\sigma_* \left(\frac{d}{dt}\right)_t = X_{\sigma(t)}$$

for all $t \in (-\varepsilon, \varepsilon)$, for some $\epsilon > 0$, where $\sigma_* \left(\frac{d}{dt}\right)_t$ is defined by

$$\sigma_* \left(\frac{d}{dt}\right)_t f = \frac{d}{dt}(f \circ \sigma(t))\bigg|_t$$

for $f \in C^\infty(\mathbb{G})$. By theory of ordinary differential equations, we see that such a curve certainly exists for some $\epsilon > 0$.

A vector field X is said to be *complete* if, at every point $g \in \mathbb{G}$, the integral curve that passes through g can be extended to an integral curve for X defined for all $t \in \mathbb{R}$. An important result is that every left (or right) invariant vector field on a Lie group is complete. Let $X \in \mathfrak{g}$. Then the unique integral curve $t \mapsto \sigma^X(t)$ of X that is defined for all $t \in \mathbb{R}$ by virtue of the result stated above, is written as

$$t \mapsto \exp(tX).$$

It can be shown that the exponential map is a local diffeomorphism from a neighbourhood of 0 in $\mathfrak{g}$ onto a neighbourhood of e in $\mathbb{G}$, and its differential map is 0 at the identity (but is not in general a global diffeomorphism nor a local homomorphism).

It is easily seen that for $X \in \mathfrak{g}, s, t \in \mathbb{R}, \exp(s+t)X = \exp(tX)\exp(sX)$. More generally if $X, Y \in \mathfrak{g}$ commute, that is $[X, Y] = 0$, then

$$\exp(X+Y) = \exp(X)\exp(Y).$$

It is of course possible to define the notion of Lie algebra independently of the notion of a Lie group. Namely an abstract Lie algebra is a vector space $\mathfrak{L}$, equipped with bilinear mapping $[\cdot,\cdot]\colon \mathfrak{L} \times \mathfrak{L} \to \mathfrak{L}$ such that
i) $[X, Y] = -[Y, X], \forall X, Y \in \mathfrak{L}$
ii) $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$ (Jacobi Identity).

8.3 Haar Measure

Any Lie group carries some left (resp. right) invariant measures called left (resp. right) *Haar* measures. Any two left (resp. right) Haar measures μ_1, μ_2 are related by $\mu_1 = c\mu_2$ for some $c > 0$, so that essentially there is only one left (resp. right) Haar measure.

Let μ be a left Haar measure on $\mathbb{G}$. For $g \in \mathbb{G}$ and A a Borel subset of $\mathbb{G}$, let $\mu_g(A) = \mu(Ag)$. Then μ_g is another left Haar measure:

$$\mu_g(hA) = \mu(hAg) = \mu(Ag) = \mu_g(A)$$

for all Borel subsets A of $\mathbb{G}$. Therefore by above we have that there exists $m(g) > 0$ such that

$$\mu_g = m(g)\mu.$$

Clearly $m(gh) = m(g)m(h)$ and $m(e) = 1$. The function $m\colon \mathbb{G} \to (0, \infty)$ is called the *modular function* of $\mathbb{G}$. If $m \equiv 1$ we say that $\mathbb{G}$ is *unimodular*. It can easily be shown that $\mathbb{G}$ is unimodular if and only if any left Haar measure is also a right Haar measure.

8.4 Solvable, Semisimple and Nilpotent Lie Groups

Let $\mathbb{G}$ be a Lie group, with Lie algebra $\mathfrak{g}$. Recall that an ideal of $\mathfrak{g}$ is a subspace I of $\mathfrak{g}$ such that $[X, Y] \in I$ for all $X \in \mathfrak{g}, Y \in I$.
Define

$$\mathfrak{g}^{(1)} = [\mathfrak{g}, \mathfrak{g}] = span\{[X, Y] \colon X,\ Y \in \mathfrak{g}\}.$$

Then $\mathfrak{g}^{(1)}$ is the smallest ideal of $\mathfrak{g}$ with an abelian quotient. We can then inductively define for $k \geq 1$

$$\mathfrak{g}^{(k+1)} = [\mathfrak{g}^{(k)}, \mathfrak{g}^{(k)}]$$

so that $\mathfrak{g}^{(k+1)}$ is smallest ideal of $\mathfrak{g}^{(k)}$ whose quotient is abelian. We call the series $\mathfrak{g}^{(k)}$ the derived series of the Lie algebra $\mathfrak{g}$. Note that for all k, $\mathfrak{g}^{(k)}$ is an ideal of $\mathfrak{g}^{(k-1)}$ and

$$\mathfrak{g} \supseteq \mathfrak{g}^{(1)} \supseteq \mathfrak{g}^{(2)} \supseteq \dots.$$

We say that the Lie algebra $\mathfrak{g}$ is *solvable* if $\mathfrak{g}^{(m)} = 0$ for some $m \geq 1$, and the Lie group $\mathbb{G}$ is solvable if its Lie algebra is.

If $\mathfrak{g}$ is solvable, then the derived series of $\mathfrak{g}$ provides us with an "approximation" of $\mathfrak{g}$ by a finite series of ideals with abelian quotients. This also works the other way round, since if $\mathfrak{g}$ is a Lie algebra with ideals

$$\mathfrak{g} = I_0 \supseteq I_1 \supseteq \cdots \supseteq I_{m-1} \supseteq I_m = 0$$

such that I_{k-1}/I_k is abelian for $1 \leq k \leq m$, then $\mathfrak{g}$ is solvable.

It can be shown that if $\mathfrak{g}$ is a finite dimensional Lie algebra, then there exists a unique solvable ideal of $\mathfrak{g}$ containing every solvable ideal of $\mathfrak{g}$. This largest solvable ideal is called the *radical* of $\mathfrak{g}$, and is denoted rad $\mathfrak{g}$. A non-zero Lie algebra $\mathfrak{g}$ is said to be *semisimple* if it has no non-zero solvable ideals, or equivalently if

$$\text{rad } \mathfrak{g} = 0.$$

We now inductively define another series of subspaces of the Lie algebra $\mathfrak{g}$ by

$$\mathfrak{g}^1 = [\mathfrak{g}, \mathfrak{g}] \quad \text{and} \quad \mathfrak{g}^{k+1} = [\mathfrak{g}, \mathfrak{g}^k] \text{ for } k \geq 1.$$

Then $\mathfrak{g} \supseteq \mathfrak{g}^1 \supseteq \mathfrak{g}^2 \supseteq \dots$. In this case we have that $\mathfrak{g}^k$ is an ideal of $\mathfrak{g}$ and not just of $\mathfrak{g}^{k-1}$.

The Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if for some $m \geq 1$, we have

$$\mathfrak{g}^m = 0,$$

and we say that the Lie group $\mathbb{G}$ is nilpotent if its Lie algebra is nilpotent.

An important result is that if $\mathbb{G}$ is a nilpotent Lie group, then the so-called Baker-Campbell-Hausdorff formula is global. Indeed, then there exists a Lie polynomial $\mathcal{P} \colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ with rational coefficients, (i.e. for $X, Y \in \mathfrak{g}$, $\mathcal{P}(X, Y)$ is a finite sum of terms involving X, Y and commutators of X and Y) such that for any $X, Y \in \mathfrak{g}$,

$$\exp X \exp Y = \exp \mathcal{P}(X, Y).$$

Moreover, if $\mathbb{G}$ is simply connected then the exponential map $\mathfrak{g} \to \mathbb{G}$ is a diffeomorphism. In that case, the group law of $\mathbb{G}$ is thus fully characterized by the Lie algebra structure of $\mathfrak{g}$.

We also have that any nilpotent Lie group is unimodular.

8.5 Some Classical Functional Analysis Results

In this section we cover the basic functional analysis results and definitions that are used within the notes.

8.5.1 Rademacher's Theorem

Let f be a real valued function defined on an open subset $U \subseteq \mathbb{R}^n$ so that $f\colon U \to \mathbb{R}^m$. Then f is called *Lipschitz continuous*, or is said to satisfy the *Lipschitz condition* if there exists a constant $K \geq 0$ such that for all $x, y \in U$,

$$|f(x) - f(y)| \leq K|x - y|.$$

Theorem 6 (Rademacher) *If U is an open subset of $\mathbb{R}^n$ and $f\colon U \to \mathbb{R}^m$ is Lipschitz continuous, then f is differentiable almost everywhere.*

In 1989 Pansu extended this theorem to the setting of Lie groups endowed with the Carnot-Carathéodory distance, and it is this extension of the theorem that is used in the notes. See [4] for details.

8.5.2 Essentially Self Adjoint Operators

Let $T\colon \mathcal{D}(T) \to \mathcal{H}$ be a linear operator with domain $\mathcal{D}(T)$ that is dense in a complex Hilbert space $\mathcal{H}$. If S is another such linear operator write $S \subset T$ if T is an extension of S, i.e. if

$$\mathcal{D}(S) \subset \mathcal{D}(T) \quad \text{and} \quad S = T|_{\mathcal{D}(S)}.$$

Since $\mathcal{D}(T)$ is dense in $\mathcal{H}$, we may define the *Hilbert adjoint operator*, $T^*\colon \mathcal{D}(T^*) \to \mathcal{H}$, of T as follows. The domain $\mathcal{D}(T^*)$ of T^* consists of all $y \in \mathcal{H}$ such that there is a $y^* \in \mathcal{H}$ satisfying

$$\langle Tx, y\rangle = \langle x, y^*\rangle$$

for all $x \in \mathcal{D}(T)$. For each such $y \in \mathcal{D}(T^*)$ the Hilbert adjoint operator T^* is then defined by $T^*y = y^*$.

We say that T is *symmetric* if for all $x, y \in \mathcal{D}(T)$,

$$\langle Tx, y\rangle = \langle x, Ty\rangle.$$

It is simple to show that T is symmetric if and only if $T \subset T^*$. We say that T is *self-adjoint* if $T = T^*$.

Now define

$$G(T) = \{(x, Tx)\colon x \in \mathcal{D}(T)\}.$$

If $G(T)$ is a closed subspace of $\mathcal{H} \times \mathcal{H}$, we say that T is a *closed* linear operator. If the linear operator T has an extension T_1 which is a closed linear operator, then T is said to be *closable*, and if T_1 is the minimal such closed linear extension of T then T_1 is said to be the *closure* of T in $\mathcal{H}$.

Now suppose that $T\colon \mathcal{D}(T) \to \mathcal{H}$ is a densely defined linear operator that is symmetric. If the closure of T (which is actually equal to T^{**}) is self-adjoint, then T is said to be *essentially self-adjoint*. We have the following useful result.

Theorem 7 *Let $T\colon \mathcal{D}(T) \to \mathcal{H}$ be a densely defined, symmetric linear operator on Hilbert space $\mathcal{H}$. Suppose also that*

$$\langle Tx, x\rangle \geq 0 \quad \forall x \in \mathcal{D}(T).$$

Then T is essentially self adjoint if and only if the range of $T + I$ is dense in $\mathcal{H}$ (where I is the identity map on $\mathcal{H}$).

References

[1] Davies E.B. *Heat Kernels and Spectral Theory*, Cambridge Tracts in Mathematics, 92, Cambridge University Press, Cambridge, 1989.

[2] Gromov M. *Carnot-Carathéodory spaces seen from within, In Sub-Riemannian Geometry*, edited by A. Bellaïche and J. J.Risler, Birkhäuser, 1996.

[3] Nagel A., Stein E. M., Wainger S. *Balls and metrics defined by vector fields I: Basic properties*, Acta Math. 155 (1985), 103–147.

[4] Pansu P. *Métriques de Carnot-Carathéodory et quasi-isométries des espaces symétriques de rang un*, Ann. Math., **129**, 1-60, 1989.

[5] Saloff-Coste L. *Aspects of Sobolev-Type Inequalites*, London Mathematical Society Lecture Note Series, 289. Cambridge University Press, Cambridge, 2002.

[6] Taylor, M. E. *Pseudodifferential Operators*, Princeton University Press, 1981.

[7] Varopoulos N. Th., Saloff-Coste L., Coulhon T. *Analysis and Geometry on Groups*, Cambridge Tracts in Mathematics, 100, Cambridge University Press, Cambridge, 1992.