Geometric and Spectral Properties of Hypoelliptic Operators

Master Thesis

Stine M. Berge

May 25, 2017

Acknowledgements

First of all I want to thank my three supervisors; Erlend Grong, Alexander Vasil'ev and Irina Markina for helping me: Alexander and Irina for taking good care me and for basically raising me during the last three years, with all the work and patience it implies: My main supervisor Erlend for putting up with my awkwardness and for being extremely encouraging and kind.

Additionally, I want to thank my husband for his love and support and for telling me when I'm wrong. Also a HUGE thanks to my friends and family for without my life would be a lot more dull (yes Mauricio, this includes you in whatever category you like).

Contents

1 Introduction

There is a canonical second order differential operator on any Riemannian manifold which reflects its geometry, as well as standing as a prototype example of an elliptic operator; *the Laplacian*. The most familiar example is $\mathbf{R}^n$ endowed with the standard Riemannian metric, in which case the Laplacian is $\Delta = \sum_{i=1}^n \partial_i^2$. For every elliptic operator L on a manifold there exists a unique Riemannian metric $\mathbf{g}$ such that L is the sum of the Laplacian of $\mathbf{g}$ and first and zero order terms. From this we can conclude that the Laplacian determines the Riemannian metric uniquely. Additionally, the Laplacian is invariant under isometries, which is one of the many connections between the Laplace operator and Riemannian geometry. We will consider the Laplace operator acting on the subspace of smooth compactly supported functions in the L^2-space with respect to the Riemannian volume measure. In this setting the Laplacian is a negative symmetric operator with respect to the L^2 inner-product. The Laplacian is also the only elliptic operator inducing the Riemannian metric with these properties. When given a compact Riemannian manifold it is well known, see e.g. [Jos11], that the eigenvalues of the Laplacian $\{\lambda_i\}_{i\in\mathcal{I}}$, can be ordered in the following way:

$$0 = \lambda_0 > \lambda_1 \geq \lambda_2 \geq \cdots,$$

where $\lambda_i \to -\infty$. In general, finding the eigenvalues explicitly is difficult. However, we can we can use the geometry of the Riemannian manifold to get more information about the spectrum. For example Weyl's law, found in e.g. [Jos11], tells us that $\operatorname{Vol}(M) = (2\pi)^n \lim_{r\to\infty} \frac{N(r)}{r^{n/2}}$, where $N(r)$ is the number of eigenvalues greater that $-r$ and $\operatorname{Vol}(M)$ denotes the volume of M. A corollary of Weyl's law is that isospectrical manifolds, i.e. manifolds where the spectrum of the Laplacian operator coincides, have the same volume. We will however focus on the Lichnerowicz estimate of the first eigenvalue, which states that if there exists a positive constant K such that the Ricci curvature satisfies $\operatorname{Ric}(v,v) \geq K\|v\|_{\mathbf{g}}^2$ then

$$\lambda_1 \leq \frac{-n}{n-1}K.$$

The reason for this is that we want to study generalizations of this result for the sub-Laplacian.

A differential operator L is called hypoelliptic if Lu is smooth implies that u is smooth. This class of differential operators include all elliptic operators. In much the same way Riemannian geometry and elliptic operators are connected, we have a connection between hypoelliptic operators and sub-Riemannian geometry. Informally, sub-Riemannian geometry can be viewed as Riemannian geometry where we have imposed restrictions on the velocities. This is formalized by defining a distribution $\mathcal{H}$ on the manifold and require all permissible curves to be tangent to $\mathcal{H}$. We call $\mathcal{H}$ for the horizontal bundle, and curves tangent to $\mathcal{H}$ are called horizontal. To get the notion of length of horizontal curves we define a fiber metric $\mathbf{h}$ on $\mathcal{H}$, which is referred to as a sub-Riemannian metric. We will additionally impose the bracket generating condition on $\mathcal{H}$, i.e. the tangent bundle is equal to the span of iterated brackets of vector fields in $\mathcal{H}$. This condition has many important consequences for the geometry, for example ensuring that it becomes a metric space and that locally minimizing curves always exist. For us, however, the most important consequence is that the analogue of the Laplacian, called the sub-Laplacian, is hypoelliptic, which follows from the conditions given in [Hör67]. When considered as an operator on L^2 with respect to some fixed volume measure, we again get that the sub-Laplacian is the unique symmetric operator inducing the sub-Riemannian metric $\mathbf{h}$. Under the assumption that the manifold is compact, the eigenvalues of the sub-Laplacian can be ordered in the same way as for the Laplace operator. However, in the case of the sub-Laplacian less is know about the interplay between the spectrum and the sub-Riemannian geometry. Therefore, a large part of this thesis will be devoted to exploring the extensions of the Lichnerowicz estimate to sub-Riemannian geometry.

This thesis is divided into three main sections and two appendixes. The appendixes will be used to introduce the prerequisites for the main parts. In Appendix A we begin with a quick introduction to spectral theory and self-adjoint operators, before moving on to second order differential operators. As we will see both the Laplacian and the sub-Laplacian are examples of such kinds of operators. Appendix B will begin by defining Lie groups and the Lie derivative, which will be needed when computing examples. Since we will do a lot of computations in L^2, we give a short introduction to integration on manifolds. Lastly, we give some properties of second order differential operators on manifolds.

In Section 2 we start by reviewing some concepts from Riemannian geometry and developing theory which we will need later. Afterwards, we define the Laplacian and give some basic properties, for instance that the Laplacian is preserved under isometries and that it is essentially self-adjoint on complete manifolds. To give some background and to motivate our approach to the sub-Riemannian Lichnerowicz estimate we will give the complete proof of the Lichnerowicz estimate in the Riemannian setting.

Section 3 begins with an introduction to sub-Riemannian geometry, giving some of the relevant definitions and theorems for our results. We will discuss the bracket generating condition and its consequences, which includes that every two point can be connected by a horizontal curve, and the local existence of length minimizing curves. In this setting we will define the sub-Laplacian and give some related properties, most of which can be found [Hör67], [FP83] and [GT16a]. For example the fact that the sub-Laplacian is hypoelliptic, that on complete manifolds it is essentially self-adjoint and the sub-Riemannian analogue of the Bochner formula.

Building upon this theory we will in Section 4 state and prove a sub-Riemannian analogue of the Lichnerowicz estimate, improving the result in [BK16] and [BKW16]. The new Lichnerowicz estimate is valid in cases where the orthogonal compliment of $\mathcal{H}$ is not integrable and the Yang-Mills condition does not hold, which were some of the requirements in [BK16] and [BKW16]. This makes us able to, not only estimate the spectral gap of more examples, but also to consider families of sub-Riemannian metrics on a distribution $\mathcal{H}$. This is done in Section 4.4 where we consider $SU(2)$ with a two dimensional bracket generating distribution $\mathcal{H}$ with a sub-Riemannian metric $\mathbf{h}$. Then we can define a family of sub-Riemannian metrics on $SU(2)$ by conformal change $e^{2f}\mathbf{h}$, where we can estimate the first eigenvalue for all f where the constant K, which depend on f, is positive. In this setting K becomes the analogue of the bound on the Ricci curvature from below.

In Section 4.5 we give an example which is not totally geodesic. The example is on $SU(2) \times SU(2)$, where we consider a family of distributions $(\mathcal{H}_a, \mathbf{h}_a)$ with some fixed orthogonal compliment $(\mathcal{V}, \mathbf{g}_{\mathcal{V}})$. In this example we can in general not use the Bott connection, since it fails to be compatible with the taming metric. Our result however still hold with another choice of connection. When we can not use the Bott connection the formula for estimating the spectral gap becomes more complicated, see Theorem 4.6. The example illustrates that the approach given in this thesis gives us greater freedom in the choice of orthogonal compliment, which makes the theory more flexible.

2 Riemannian Manifolds and the Laplacian

For completeness and to fix notation, let us begin by defining Riemannian manifolds. To us a manifold will always be a second countable connected topological Hausdorff space which is locally euclidean and endowed with a maximal smooth structure. Let M be a manifold, let $\pi : B \to M$ be a smooth vector bundle, and let $\Gamma(B)$ denote the set of all smooth sections on $\pi : B \to M$. A **fiber metric g** on $\pi : B \to M$ is a smooth section of the symmetric bundle $\mathrm{Sym}^2(B^*)$ satisfying $\mathbf{g}_p(v,v) > 0$ for all nonzero elements $v \in B_p$. A more intuitive way to look at fiber metrics is that they give an inner-product on each fiber of the bundle. If the fiber metric $\mathbf{g}$ is defined on the tangent bundle, then $\mathbf{g}$ is called a **Riemannian metric** and we call the pair $(M, \mathbf{g})$ a **Riemannian manifold**. Every Riemannian metric $\mathbf{g}$ induces a canonical map $\flat^{\mathbf{g}} : TM \to T^*M$ defined fiberwise by $(\flat^{\mathbf{g}}(v))(w) = \mathbf{g}(v,w)$, for $v, w \in T_pM$. Since $\mathbf{g}$ is non-degenerate, $\flat^{\mathbf{g}}$ is invertible, with inverse denoted by $\sharp^{\mathbf{g}}$. A Riemannian metric induces a fiber metric on every tensor bundle $T^k_l M$ as follows: On T^*M the fiber metric is defined by $\langle \alpha, \beta \rangle_{\mathbf{g}^*} = \langle \sharp^{\mathbf{g}}\alpha, \sharp^{\mathbf{g}}\beta \rangle_{\mathbf{g}}$, and on $T^k_l M$ by

$$\langle \alpha^1 \otimes \cdots \otimes \alpha^l \otimes v_1 \otimes \cdots \otimes v_k, \tilde{\alpha}^1 \otimes \cdots \otimes \tilde{\alpha}^l \otimes \tilde{v}_1 \otimes \cdots \otimes \tilde{v}_k \rangle_{\mathbf{g}^k_l} = \Pi^l_{j=1} \langle \alpha^j, \tilde{\alpha}^j \rangle_{\mathbf{g}^*} \Pi^k_{i=1} \langle v_i, \tilde{v}_i \rangle_{\mathbf{g}}.$$

2.1 Affine Connections and Curvature

To define essential objects like geodesic, curvature and the (sub-)Laplacian, we need to define connections. Some of the theory developed in this section will not be used immediately, but will be important later. The proofs of the statements not proved in this section can be found in standard books on Riemannian geometry, see e.g. [Lee97].

Definition 2.1. Let $\pi : B \to M$ be a vector bundle. An **affine connection** on B is a map $\nabla : \Gamma(TM) \times \Gamma(B) \to \Gamma(B)$ satisfying for all $X_1, X_2 \in \Gamma(TM)$, $Y_1, Y_2 \in \Gamma(B)$, $a, b \in \mathbf{R}$ and $f, g \in C^\infty(M)$

1. $\nabla_{fX_1 + gX_2} Y_1 = f\nabla_{X_1} Y_1 + g\nabla_{X_2} Y_1$,
2. $\nabla_{X_1}(aY_1 + bY_2) = a\nabla_{X_1} Y_1 + b\nabla_{X_1} Y_2$ and
3. $\nabla_{X_1}(fY_1) = f\nabla_{X_1} Y_1 + (X_1 f) Y_1$.

It can be shown that $\nabla_X Y(p)$ only depends on X at the point p, and on Y along a smooth curve tangent to X_p. Let $X_1, \ldots, X_n$ and $E_1, \ldots, E_r$ be local frames for TM and $\pi : B \to M$, respectively. Define the **Christoffel symbols** $\Gamma^\alpha_{i\beta}$ of the connection with respect to the local frames by $\nabla_{X_i} E_\beta = \sum_{\alpha=1}^r \Gamma^\alpha_{i\beta} E_\alpha$. The Christoffel symbols determine the affine connection uniquely in the domain of the local frames by the formula

$$\nabla_X Y = \sum_{\alpha=1}^r \left(X(Y^\alpha) + \sum_{i=1}^n \sum_{\beta=1}^r X^i Y^\beta \Gamma^\alpha_{i\beta} \right) E_\alpha,$$

where $X = \sum_{i=1}^n X^i X_i$ and $Y = \sum_{\alpha=1}^r Y^\alpha E_\alpha$.

Given an affine connection ∇ on $\pi : B \to M$ and a smooth curve $\gamma : (a,b) \to M$, we define the **parallel translation** of a vector $v \in T_{\gamma(t_0)} M$, where $t_0 \in (a,b)$, along γ to be the unique solution $v(t) \in T_{\gamma(t)} M$ of the differential equation $v(t_0) = v$ and $\nabla_{\dot{\gamma}(t)} v(t) = 0$.

Proposition 2.2. *Let ∇ be an affine connection on the bundle $\pi : B \to M$ of rank r. Then at every fixed point p there exists a local frame $E_1, \ldots, E_r$ about p such that $(\nabla_X E_\alpha)_p = 0$, for any vector field X.*

Proof. It is enough to show that the Christoffel symbol vanishes at p, since

$$\nabla_X E_\alpha = \sum_{i=1}^{n} \sum_{\beta=1}^{r} X^i \Gamma_{i\alpha}^{\beta} E_\beta.$$

Let $\{e_\alpha\}$ be a basis for the fiber B_p, and let $(x_1, \ldots, x_n)$ be a coordinate frame centered at the point p which maps to some star shaped domain in $\mathbf{R}^n$. If $y = (y_1, \ldots, y_n)$ is any point in the coordinate frame written in local coordinates, then γ_y denotes the curve $\gamma_y(t) = ty$. For every point y in the domain of the chart we define the extension of e_α, denoted $E_\alpha(ty)$, by parallel transport e_α along the radial curve γ_y. Doing this along all radial curves create a local frame on $\pi : B \to M$ which we denote by $\{E_\alpha\}$. This fact, in turn, means that

$$\nabla_{\dot{\gamma}_y(1)} E_\alpha = \sum_{i=1}^{n} \nabla_{y_i \partial_i} E_\alpha = \sum_{i=1}^{n} y_i \nabla_{\partial_i} E_\alpha = 0,$$

by using the definition of parallel translation together with $\dot{\gamma}_y(1) = \sum_{i=1}^{n} y_i \partial_i$. Hence the Christoffel symbols satisfy

$$y_1 \Gamma_{\beta 1}^{\alpha} + \cdots + y_n \Gamma_{\beta n}^{\alpha} = 0.$$

Since this is true for any y we have that

$$x_1 \Gamma_{\beta 1}^{\alpha} + \cdots + x_n \Gamma_{\beta n}^{\alpha} = 0.$$

Taking the derivative with respect to x_j, we get that

$$\Gamma_{\beta j}^{\alpha} = -x_1 \partial_j \Gamma_{\beta 1}^{\alpha} - \cdots - \widehat{x_j \partial_j \Gamma_{\beta j}^{\alpha}} - \cdots - x_n \partial_j \Gamma_{\beta n}^{\alpha},$$

where $\widehat{\ }$ indicates that the term is omitted from the sum. Since the frame was centered at p we have that $\Gamma_{\beta j}^{\alpha}(p) = 0$, which completes the proof. □

We say that a local frame $E_1, \ldots, E_r$ on the vector bundle $\pi : B \to M$ is **parallel at the point** p if $(\nabla_X E_i)_p = 0$ for all $X \in \Gamma(TM)$. Often when we are given a fiber metric we would like the given local frame to be orthonormal with respect to the metric. Given a fiber metric $\mathbf{g}$ on $\pi : B \to M$, we say that ∇ is **compatible** with $\mathbf{g}$ if

$$X(\mathbf{g}(Y, Z)) = \mathbf{g}(\nabla_X Y, Z) + \mathbf{g}(Y, \nabla_X Z).$$

Corollary 2.3. *Let $\mathbf{g}$ be a fiber metric on the bundle $\pi : B \to M$ compatible with the affine connection ∇ and let $p \in M$. Then there exists an orthonormal local frame which is parallel at the point p.*

Proof. Let $v, w \in B_p$, let γ be a smooth curve through p and V and W the parallel translates of v and w along γ, respectively. Then

$$\frac{d}{dt}(\mathbf{g}(V, W)) = \mathbf{g}(\nabla_{\dot{\gamma}} V, W) + \mathbf{g}(V, \nabla_{\dot{\gamma}} W) = 0,$$

hence $\mathbf{g}(V, W)$ is constant along γ. This means that if $\{e_\alpha\}$ is an orthonormal basis in the proof of Proposition 2.2, then $\{E_\alpha\}$ is orthonormal. □

A special class of affine connections are the **linear connections**, which are connections on the tangent bundle.

Lemma 2.4. *Given a linear connection ∇ on M, there is a unique affine connection defined on each tensor bundle $T_l^k M$, also denoted by ∇, which coincides with the original on TM and satisfying;*

1. *on $C^\infty(M)$, $\nabla_X f = X(f)$,*
2. *$\nabla_X (F \otimes G) = (\nabla_X F) \otimes G + F \otimes (\nabla_X G)$ and*
3. *$\nabla_X \mathrm{tr}(F) = \mathrm{tr} \nabla_X F$ where the trace is taken over any pair consisting of a higher and a lower index.*

Additionally, we get that ∇ has the following properties

- *$\nabla_X (\omega(Y)) = (\nabla_X \omega)(Y) + \omega(\nabla_X Y)$, for $\omega \in \Gamma(T^*M)$ and $X, Y \in \Gamma(TM)$.*
- *if $F \in \Gamma(T_l^k M)$, then the affine connection can be computed as follows;*

$$\begin{aligned}(\nabla_X F)\left(\omega^1, \dots, \omega^l, Y_1, \dots, Y_k\right) &= X\left(F\left(\omega^1, \dots, \omega^l, Y_1, \dots, Y_k\right)\right) \\ &\quad - \sum_{i=1}^{l} F\left(\omega^1, \dots, \nabla_X \omega^i, \dots, \omega^l, Y_1, \dots, Y_k\right) - \sum_{j=1}^{k} F\left(\omega^1, \dots, \omega^l, Y_1, \dots, \nabla_X Y_j, \dots, Y_k\right).\end{aligned}$$

Using affine and linear connections ∇ we can define several tensor fields. The most important for us are the torsion, Hessian and curvature.

The **torsion** with respect to a linear connection ∇ is the $(2,1)$-tensor field T^∇ defined by

$$T^\nabla(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y].$$

As is easily seen from the definition, the torsion is skew-symmetric. For us the following property of the torsion will be useful for computations.

Lemma 2.5. *For any $f \in C^\infty(M)$ and for any linear connection ∇ we have that*

$$(\nabla_X df)(Y) - (\nabla_Y df)(X) = -df\left(T^\nabla(X, Y)\right).$$

Proof. If we use the definition of how ∇ acts on forms we get

$$\begin{aligned}(\nabla_X df)(Y) - (\nabla_Y df)(X) &= X(df(Y)) - df(\nabla_X Y) - Y(df(X)) + df(\nabla_Y X) \\ &= -df(\nabla_X Y - \nabla_Y X) - Y(X(f)) + X(Y(f)) \\ &= -df(\nabla_X Y - \nabla_Y X) - (-[X, Y](f)) \\ &= -df\left(T^\nabla(X, Y)\right).\end{aligned}$$

□

A linear connection ∇ is said to be **symmetric** if its torsion tensor is identically zero, i.e. $T^\nabla \equiv 0$. If the linear connection is symmetric, then

$$(\nabla_X df)(Y) = (\nabla_Y df)(X),$$

by the previous lemma.

Given a linear connection ∇ we define the **Hessian** of a tensor F to be

$$\nabla^2_{X,Y}F = \nabla_X \nabla_Y F - \nabla_{\nabla_X Y} F,$$

where $X, Y \in \Gamma(TM)$. The Hessian is tensorial in X and Y, i.e. linear over $C^\infty(M)$. A short calculation shows that $\nabla^2_{X,Y} f = (\nabla_X df)(Y)$ and using Lemma 2.5, we have that

$$T^\nabla(X,Y)(f) = \nabla^2_{Y,X} f - \nabla^2_{X,Y} f. \tag{2.1}$$

It follows that ∇^2 is symmetric when acting on functions if and only if the connection is symmetric. Equivalently, the Hessian has an anti-symmetric part if and only if the linear connection has torsion. Later in this thesis we will use the Hessian of a connection to define the (sub-)Laplacian.

The **curvature endomorphism** is defined on a tensor F by

$$R^\nabla(X,Y)F = \nabla_X \nabla_Y F - \nabla_Y \nabla_X F - \nabla_{[X,Y]} F = [\nabla_X, \nabla_Y] F - \nabla_{[X,Y]} F.$$

We can also define the curvature endomorphism over any bundle by using the same formula. Intuitively the curvature tells us how far the Lie bracket and the affine connection are from commuting. If they do commute, i.e. the curvature endomorphism is identically zero, then we say that the affine connection is **flat**. The next proposition is proved in e.g. [Tay11].

Proposition 2.6. *If ∇ is a flat affine connection on $\pi : B \to M$, then there exists a parallel local frame $\{E_\alpha\}$, i.e. a local frame such that $\nabla_X E_\alpha = 0$ for all X.*

In the next section we will define the Levi-Civita connection. The term flat comes from the theorem which tells us that if the Levi-Civita connection is flat, then the Riemannian manifold is locally isometric to Euclidean space, see e.g. [Lee97].

Lemma 2.7. *Assume that we are given an affine connection ∇ on some bundle $\pi : B \to M$ compatible with a fiber metric $\mathbf{g}$. Then*

$$\langle R^\nabla(X,Y)E, F\rangle_{\mathbf{g}} = -\langle E, R^\nabla(X,Y)F\rangle_{\mathbf{g}},$$

for all $E, F \in \Gamma(B)$ and $X, Y \in \Gamma(TM)$. In other words, $R^\nabla(X,Y)$ is skew-symmetric with respect to $\mathbf{g}$ on each fiber.

Proof. By direct computation we obtain

$$\begin{aligned}
&\langle \nabla_X \nabla_Y E, F\rangle_{\mathbf{g}} - \langle \nabla_Y \nabla_X E, F\rangle_{\mathbf{g}} - \langle \nabla_{[X,Y]} E, F\rangle_{\mathbf{g}} \\
&= X(\langle \nabla_Y E, F\rangle_{\mathbf{g}}) - Y(\langle \nabla_X F, E\rangle_{\mathbf{g}}) + \langle E, \nabla_Y \nabla_X F\rangle_{\mathbf{g}} - Y(\langle \nabla_X E, F\rangle_{\mathbf{g}}) \\
&\quad + X(\langle \nabla_Y F, E\rangle_{\mathbf{g}}) - \langle E, \nabla_X \nabla_Y F\rangle_{\mathbf{g}} - [X,Y](\langle E, F\rangle_{\mathbf{g}}) + \langle E, \nabla_{[X,Y]} F\rangle_{\mathbf{g}} \\
&= X(Y(\langle E, F\rangle_{\mathbf{g}})) - X(\langle E, \nabla_Y F\rangle_{\mathbf{g}}) - Y(X(\langle E, F\rangle_{\mathbf{g}})) + Y(\langle \nabla_X E, F\rangle_{\mathbf{g}}) \\
&\quad - Y(\langle \nabla_X E, F\rangle_{\mathbf{g}}) + X(\langle \nabla_Y F, E\rangle_{\mathbf{g}}) - [X,Y](\langle E, F\rangle_{\mathbf{g}}) - \langle E, R^\nabla(X,Y)F\rangle_{\mathbf{g}} \\
&= -\langle E, R^\nabla(X,Y)F\rangle_{\mathbf{g}}.
\end{aligned}$$

$\square$

Given the local frames $X_1, \ldots, X_n$ and $E_1, \ldots, E_r$ on the tangent bundle and the bundle $\pi : B \to M$, respectively, the curvature can be written as

$$R^{\nabla}(X, Y) E = \sum_{i,j=1}^{n} \sum_{k,l=1}^{r} X^i Y^j E^k R_{ijk}{}^l E_l,$$

where $X = \sum_{i=1}^n X^i X_i, Y = \sum_{j=1}^n Y^j X_j$ and $E = \sum_{l=1}^r E^l E_l$.

When we want to compute the curvature of any affine connections we will use the following lemma. To make the result easier to state we will first introduce the notation for cyclic sum. We use the symbol $\circlearrowleft$ as an abbreviation for the cyclic sum, e.g.

$$\circlearrowleft f(X_1, X_2, X_3) = f(X_1, X_2, X_3) + f(X_3, X_1, X_2) + f(X_2, X_3, X_1).$$

Lemma 2.8 (First Bianchi Identity for Affine Connections with Torsion)**.** *Let ∇ be a linear connection on M. Then*

$$\circlearrowleft R^{\nabla}(X_1, X_2) X_3 = \circlearrowleft \left(\nabla_{X_1} T^{\nabla}\right)(X_2, X_3) - \circlearrowleft T^{\nabla}\left(X_1, T^{\nabla}(X_2, X_3)\right),$$

for any $X_1, X_2, X_3 \in \Gamma(TM)$.

Proof. Writing out the right hand side we get

$$\begin{aligned}
&\circlearrowleft \left(\nabla_{X_1} T^{\nabla}\right)(X_2, X_3) - \circlearrowleft T^{\nabla}\left(X_1, T^{\nabla}(X_2, X_3)\right) \\
&= \circlearrowleft \nabla_{X_1}(T(X_2, X_3)) + \circlearrowleft T^{\nabla}(\nabla_{X_1} X_2, X_3) - \circlearrowleft T^{\nabla}(\nabla_{X_1} X_3, X_2) \\
&\quad - \circlearrowleft \nabla_{X_1}(T(X_2, X_3)) + \circlearrowleft \nabla_{T(X_2, X_3)} X_1 + \circlearrowleft [X_1, T(X_2, X_3)] \\
&= - \circlearrowleft T^{\nabla}(\nabla_{X_1} X_2, X_3) + \circlearrowleft T^{\nabla}(\nabla_{X_1} X_3, X_2) + \circlearrowleft \nabla_{T(X_2, X_3)} X_1 + \circlearrowleft [X_1, T(X_2, X_3)].
\end{aligned}$$

Using that

$$\circlearrowleft f(X_1, X_2, X_3) = \circlearrowleft f(X_3, X_1, X_2) = \circlearrowleft f(X_2, X_3, X_1),$$

along with the definition of torsion gives

$$\begin{aligned}
&- \circlearrowleft T^{\nabla}(\nabla_{X_1} X_2, X_3) + \circlearrowleft T^{\nabla}(\nabla_{X_1} X_3, X_2) + \circlearrowleft \nabla_{T(X_2, X_3)} X_1 + \circlearrowleft [X_1, T(X_2, X_3)] \\
&= - \circlearrowleft \nabla_{\nabla_{X_1} X_2} X_3 + \circlearrowleft \nabla_{X_3} \nabla_{X_1} X_2 + \circlearrowleft [\nabla_{X_1} X_2, X_3] + \circlearrowleft \nabla_{\nabla_{X_1} X_3} X_2 - \circlearrowleft \nabla_{X_2} \nabla_{X_1} X_3 \\
&\quad + \circlearrowleft [X_2, \nabla_{X_1} X_3] + \circlearrowleft \nabla_{\nabla_{X_2} X_3} X_1 - \circlearrowleft \nabla_{\nabla_{X_3} X_2} X_1 - \circlearrowleft \nabla_{[X_2, X_3]} X_1 + \circlearrowleft [X_1, T(X_2, X_3)] \\
&= \circlearrowleft \nabla_{X_3} \nabla_{X_1} X_2 + \circlearrowleft [\nabla_{X_1} X_2, X_3] - \circlearrowleft \nabla_{X_2} \nabla_{X_1} X_3 + \circlearrowleft [X_2, \nabla_{X_1} X_3] \\
&\quad - \circlearrowleft \nabla_{[X_2, X_3]} X_1 + \circlearrowleft [X_1, \nabla_{X_2} X_3] - \circlearrowleft [X_1, \nabla_{X_3} X_2] \\
&= \circlearrowleft R^{\nabla}(X_1, X_2) X_3.
\end{aligned}$$

$\square$

When the linear connection is symmetric the formula above is traditionally called the first Bianchi identity.

2.1.1 The Levi-Civita Connection

In Riemannian geometry we can define a unique linear connection, called the Levi-Civita connection, which is not necessarily the case in sub-Riemannian geometry. The uniqueness of this connection is proved in standard books on Riemannian geometry, e.g. [Lee97].

Definition 2.9. The **Riemannian connection**, also called the **Levi-Civita connection**, is the unique linear connection which is both symmetric and compatible with the metric.

In the rest of the section on Riemannian geometry we are going to assume that the given linear connection is the Levi-Civita connection. Let $X_1, \ldots, X_n$ be a local frame for TM, and denote $\mathbf{g}(X_i, X_j)$ by g_{ij}. Then the Christoffel symbols can be computed locally by

$$\Gamma_{ij}^k = \frac{1}{2}(X_i(g_{jk}) + X_j(g_{ki}) - X_k(g_{ij}) + \mathbf{g}([X_i, X_j], X_k) - \mathbf{g}([X_j, X_k], X_i) - \mathbf{g}([X_i, X_k], X_j)). \quad (2.2)$$

We will use this formula when doing examples in Section 4.3, 4.4 and 4.5.

Let

$$\mathcal{T}(\gamma) = \left\{V : (a, b) \to TM : V(t) \in T_{\gamma(t)}M\right\},$$

where $V \in \mathcal{T}(\gamma)$ are smooth the curve $\gamma : (a, b) \to M$. We say that an element $V \in \mathcal{T}(\gamma)$ is **extendable** if there exists a smooth extension of V defined on an open set containing the curve γ. Given a linear connection ∇ and a curve $\gamma : I \to M$ there exists a unique operator $D_t : \mathcal{T}(\gamma) \to \mathcal{T}(\gamma)$ defined by the properties;

1. $D_t(aV + bW) = aD_t(V) + bD_t(W)$ for $a, b \in \mathbf{R}$ and $V, W \in \mathcal{T}(\gamma)$,
2. $D_t(fV) = \dot{f}V + fD_tV$ for $f \in C^\infty(I)$,
3. $D_tV(t) = \nabla_{\dot{\gamma}(t)}\tilde{V}$, where $\tilde{V}$ is an extension (if possible) of V.

The operator D_t is called the **covariant derivative along** γ. A **geodesic** with respect to the affine connection ∇ is a curve satisfying $D_t\dot{\gamma} = 0$, i.e. the covariant derivative of $\dot{\gamma}$ along γ is zero. Geodesics are unique in the sense that all geodesics (defined on a connected interval) starting at a point $\gamma(t_0) = p$ and with initial velocity $\dot{\gamma}(t_0)$ coincides on common domain. This gives us the concept of a maximal geodesic starting at the point p with initial velocity $v \in T_pM$, denoted by γ_v.

The **Riemannian exponential map** $\mathrm{Exp} : \mathcal{E} \to M$, where

$$\mathcal{E} = \{v \in TM : [0, 1] \subset \mathcal{D}(\gamma_v)\}$$

is an open set containing the zero section, is defined by $\mathrm{Exp}(v) = \gamma_v(1)$. It can be shown that the exponential map is smooth and $\gamma_v(t) = \mathrm{Exp}(tv)$. The **restricted Riemannian exponential map** Exp_p is defined by restricting Exp to $\mathcal{E}_p = T_pM \bigcap \mathcal{E}$. About every point p in M, there are neighborhoods $\mathcal{U} \subset \mathcal{E}_p$ and $p \in \mathcal{V} \subset M$ such that $\mathrm{Exp}_p : \mathcal{U} \to \mathcal{V}$ is a diffeomorphism. Using this fact we can define what is known as **normal coordinates about a point** p, which are coordinates centered at p with the additional property that if $q \in \mathcal{V}$ then the geodesic from p to q, written in coordinates, is a straight line.

Define a continuous curve $\gamma : [a, b] \to M$ to be **absolutely continuous** if its derivative exists almost everywhere and in any coordinate frame satisfies the fundamental theorem of calculus.

Let $\gamma : [a,b] \to M$ be an absolutely continuous curve, then we can define the length of γ to be $l(\gamma) = \int_a^b \|\dot{\gamma}(t)\|_{\mathbf{g}} dt$. Denote by $C_{ac}(A,B)$ the set of all absolutely continuous curves starting in A and ending in B. Then the distance between two points A and B is defined by

$$d(A,B) = \inf_{\gamma \in C_{ac}(A,B)} l(\gamma).$$

It can be showed that the topology induced by the distance function is the same as the original topology of the manifold, see for instant [Lee97]. We use the notation

$$\mathrm{diam}\,(M) = \sup_{x,y \in M} d(x,y),$$

to denote the **diameter** of the manifold.

An **isometry** between two Riemannian manifolds $(M, \mathbf{g})$ and $(N, \tilde{\mathbf{g}})$ is a diffeomorphism $\varphi : (M, \mathbf{g}) \to (N, \tilde{\mathbf{g}})$ where $\varphi^* \tilde{\mathbf{g}} = \mathbf{g}$. Since isometries preserves the Riemannian metric, we have that it also preserves the distance. It is true, but not trivial to show, that distance preserving maps are isometries. Any isometry $\varphi : (M, \mathbf{g}) \to (N, \tilde{\mathbf{g}})$ preserves the Levi-Civita connection, i.e. $\nabla^{\mathbf{g}}_X Y = \nabla^{\tilde{\mathbf{g}}}_{\varphi_* X} \varphi_* Y$, where $\nabla^{\mathbf{g}}$ and $\nabla^{\tilde{\mathbf{g}}}$ are the Levi-Civita connections on M and N, respectively. This is showed by proving that the pullback of the Levi-Civita connection is also compatible with the metric and symmetric. Then equality will follow by uniqueness. As a corollary we obtain that isometries preserves the curvature endomorphism and Hessian.

A **minimizing curve** is a curve whose length is equal to the distance between its endpoints. There is a very strong connection between curves of minimal length and geodesics, namely that every unit speed minimizing curve is a geodesic, and the partially converse statement, every geodesic is locally a minimizing curve. If A, B are two arbitrary points on the manifold, then there might exist everything between none and infinitely many minimizing curves between the points. We say that a Riemannian manifold is **geodesically complete** if every maximal geodesic is defined for all $t \in \mathbf{R}$.

Theorem 2.10 (Hoph-Rinow)**.** *A Riemannian manifold is geodesically complete if and only if it is complete as a metric space with respect to the distance defined above.*

The proof of the above theorem can be found in [Lee97]. Using the Hoph-Rinow Theorem, one gets that on complete manifolds there always exists a minimizing curve between any two points.

The **Ricci curvature** $\mathrm{Ric}^{\nabla} : \Gamma(TM) \otimes \Gamma(TM) \to C^{\infty}(M)$ is defined to be the trace of the first and the last component of the curvature endomorphism with respect to the Riemannian metric $\mathbf{g}$. Given a local orthonormal frame $X_1, \ldots, X_n$ the Ricci curvature can be written as

$$\mathrm{Ric}^{\nabla}(X,Y) = \sum_{i=1}^{n} \langle R^{\nabla}(X_i, X) Y, X_i \rangle_{\mathbf{g}}.$$

Writing it relative to the given local orthonormal frame we get the components $\mathrm{Ric}_{ij} = \sum_{k=1}^{n} R_{kij}{}^{k}$. By using the first Bianchi identity together with Lemma 2.7 we get that

$$\langle R^{\nabla}(X,Y) Z, X \rangle_{\mathbf{g}} = \langle R^{\nabla}(X,Z) Y, X \rangle_{\mathbf{g}},$$

hence it follows that the Ricci curvature is symmetric. The proof of the next theorem can be found in e.g. [Lee97].

Theorem 2.11 (Bonnet-Myers's Theorem)**.** *Let $(M, \mathbf{g})$ be a complete connected Riemannian manifold, and assume that* $\mathrm{Ric}\,(X, X) \geq \frac{n-1}{R^2}\|X\|_{\mathbf{g}}^2$, *where X is a vector field. Then M is a compact manifold with finite fundamental group and of diameter at most πR.*

The question may also arise what happens if under the conditions of the Bonnet-Myers's theorem the diameter is maximal, i.e. πR. The next theorem answers that question and is proven in [Che75].

Theorem 2.12. *Assume that $(M, \mathbf{g})$ is a complete Riemannian manifold where the Ricci curvature satisfies* $\mathrm{Ric}\,(X, X) \geq \frac{n-1}{R^2}\|X\|_{\mathbf{g}}^2$. *Furthermore, assume that the diameter of the manifold is maximal, i.e.* $\mathrm{diam}\,(M) = \pi R$. *Then $(M, \mathbf{g})$ is isometric to S^n with the round metric of radius R.*

2.2 Laplacian on Riemannian Manifolds

As the title indicates, we will in this section define the Laplacian and explore some basic properties. Mostly we will be concerned with the properties related to the spectrum of the Laplacian.

Definition 2.13. Let $(M, \mathbf{g})$ be a Riemannian manifold. The **gradient** $\mathrm{grad}_{\mathbf{g}} : C^\infty(M) \to \Gamma(TM)$ is defined by the property $\mathbf{g}\left(\mathrm{grad}_{\mathbf{g}}\varphi, X\right) = d\varphi(X)$. In other words, the gradient is defined to be the map sending $\varphi \mapsto \sharp^{\mathbf{g}}(d\varphi)$.

Locally the gradient can be written as $\mathrm{grad}_{\mathbf{g}}\varphi = \sum_{i=1}^n X_i(\varphi) X_i$, where $X_1, \ldots, X_n$ is some orthonormal frame. Some properties of the gradient are that it is linear over $\mathbf{R}$, and it satisfies the product rule

$$\mathrm{grad}_{\mathbf{g}}(\varphi \cdot \psi) = \varphi\, \mathrm{grad}_{\mathbf{g}}\psi + \psi\, \mathrm{grad}_{\mathbf{g}}\varphi.$$

Definition 2.14. The **Laplacian** on a Riemannian manifold $(M, \mathbf{g})$ is the operator $\Delta_{\mathbf{g}} : C^\infty(M) \to C^\infty(M)$ defined by $\Delta_{\mathbf{g}} = \mathrm{div}_{\mathbf{g}} \circ \mathrm{grad}_{\mathbf{g}}$.

For the definition of divergence, see Appendix B. Let $X_1, \ldots, X_n$ be a local orthonormal frame, then the Laplace operator can be written locally on the form

$$\Delta_{\mathbf{g}} f = \sum_{i=1}^n X_i^2(f) + \sum_{i,j=1}^n X_i(f) \langle X_j, [X_j, X_i]\rangle_{\mathbf{g}}.$$

Hence it is an elliptic second order operator. If we are given a chart $(x_1, \ldots, x_n)$, then we can write the Laplacian in the form

$$\Delta_{\mathbf{g}}\varphi = \frac{1}{\sqrt{|\det(g_{ij})|}} \sum_{i,j=1}^n \partial_i \left(g^{ij} \sqrt{|\det(g_{ij})|}\partial_j \varphi \right),$$

which is sometimes taken to be the definition.

Proposition 2.15. *On a Riemannian manifold $(M, \mathbf{g})$ we have that*

$$\mathrm{div}_{\mathbf{g}} \circ \mathrm{grad}_{\mathbf{g}} f = \mathrm{tr}_{\mathbf{g}} \nabla^2_{\times,\times} f.$$

Proof. Let $X_1, \ldots, X_n$ be an orthonormal local frame. The divergence of a vector field $X = \sum_{i=1}^{n} X^i X_i$ is given by

$$\operatorname{div}_{\mathbf{g}} X = \sum_{i=1}^{n} X_i \left(X^i\right) + \sum_{i,j=1}^{n} X^i \langle X_j, [X_j, X_i]\rangle_{\mathbf{g}},$$

by the formula presented in Proposition B.1 e). Hence

$$\begin{aligned}
\operatorname{div}_{\mathbf{g}} \left(\sharp^{\mathbf{g}} df\right) &= \sum_{i=1}^{n} X_i \left(X_i \left(f\right)\right) + \sum_{i,j=1}^{n} X_j \left(f\right) \langle X_i, [X_i, X_j]\rangle_{\mathbf{g}} \\
&= \sum_{i=1}^{n} \nabla_{X_i} \nabla_{X_i} f - \sum_{i,j=1}^{n} X_j \left(f\right) \langle X_i, \nabla_{X_j} X_i - \nabla_{X_i} X_j\rangle_{\mathbf{g}} \\
&= \sum_{i=1}^{n} \nabla_{X_i} \nabla_{X_i} f - \sum_{i,j=1}^{n} X_j \left(f\right) \left(\Gamma_{ji}^{i} + \langle \nabla_{X_i} X_i, X_j\rangle_{\mathbf{g}}\right) \\
&= \sum_{i=1}^{n} \nabla_{X_i} \nabla_{X_i} f - \sum_{i=1}^{n} \nabla_{X_i} X_i \left(f\right) \\
&= \operatorname{tr}_{\mathbf{g}} \nabla^2_{\times,\times} f.
\end{aligned}$$

□

Notice that in the euclidean case, the Laplacian is $\Delta_{\mathbf{g}_{\mathbf{R}^n}} = \sum_{i=1}^{n} \partial_i^2$ when using the standard euclidean coordinates. The above proposition tells us that, as in $\mathbf{R}^n$ with the standard metric, the divergence of the gradient is the same as the trace of the Hessian. We have that the Laplacian is linear over the real numbers and by combining the product laws for the divergence and the gradient, we obtain

$$\Delta_{\mathbf{g}} \left(\varphi\psi\right) = \psi \Delta_{\mathbf{g}} \varphi + \varphi \Delta_{\mathbf{g}} \psi + 2\langle \operatorname{grad}_{\mathbf{g}} \varphi, \operatorname{grad}_{\mathbf{g}} \psi\rangle.$$

Viewing the Laplacian as the trace of the Hessian, makes it easier to show that the Laplacian is preserved under isometries.

Corollary 2.16. *Let $\varphi : (M, \mathbf{g}_M) \to (N, \mathbf{g}_N)$ be an isometry. Then*

$$\Delta_{\mathbf{g}_M} \left(f \circ \varphi\right) = \left(\Delta_{\mathbf{g}_N} f\right) \circ \varphi,$$

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

Proof. It is not difficult to see, using the symmetry and compatibility of the Levi-Civita connection, that $\varphi^* \left(\nabla_X Y\right) = \nabla_{\varphi_* X} \left(\varphi_* Y\right)$. Let $f \in C^{\infty}(N)$ and $X \in \Gamma(M)$, then

$$\begin{aligned}
\left(\nabla^2_{X,X} f \circ \varphi\right) &= X \left(X \left(f \circ \varphi\right)\right) - \nabla_X X \left(f \circ \varphi\right) \\
&= X \left(X \left(f \circ \varphi\right) \circ \varphi^{-1} \circ \varphi\right) \circ \varphi^{-1} \circ \varphi - \left(\nabla_X X \left(f \circ \varphi\right)\right) \circ \varphi^{-1} \circ \varphi \\
&= \tilde{X} \left(\tilde{X} \left(f\right)\right) \circ \varphi - \left(\varphi^* \left(\nabla_X X\right) \left(f\right)\right) \circ \varphi \\
&= \tilde{X} \left(\tilde{X} \left(f\right)\right) \circ \varphi - \nabla_{\tilde{X}} \tilde{X} \left(f\right) \circ \varphi,
\end{aligned}$$

where $\tilde{X}_m \left(f\right) = X_{\varphi^{-1}(m)} \left(f \circ \varphi\right)$. The result follows from taking the trace on both sides, together with the fact that φ preserves orthonormal frames. □

By Proposition B.7 in the appendix, the Laplace operator is the unique operator with symbol $\mathbf{g}^*$ which is symmetric with respect to the volume density. This means that

$$\int_M f\Delta_{\mathbf{g}}(g)\, d\,\mathrm{vol}_{\mathbf{g}} = -\int_M \mathbf{g}^*(df, dg)\, d\,\mathrm{vol}_{\mathbf{g}},$$

when f, g are compactly supported. Since $\mathbf{g}$ is positive definite, this makes $\Delta_{\mathbf{g}}$ a negative operator. If M is compact, we can define the Laplace on $C^\infty(M)$ viewed as a subspace of $L^2(\mathbf{g}^0)$ without any problems. Since $\overline{C^\infty(M)} = L^2(M)$, the Laplace operator is a densely defined symmetric negative operator. In the case where M is a non-compact manifold, the Laplacian operator can be defined on the space of smooth compactly supported function $C_0^\infty(M)$, and is again densely defined. This means that by Theorem A.9 there exists a self-adjoint extension of the Laplace operator in $L^2(M)$ when M. It is in general several different self-adjoint extensions of the Laplacian different from the Friedrichs extension given in Appendix A. However, in the case when M is complete the following theorem assures us that the extension is unique.

Theorem 2.17. *[Str83] Let M be a complete Riemannian manifold, then the Laplacian is essentially self-adjoint.*

Proof. Since there exists at least one self-adjoint extension, the Laplacian is closable. Denote this temporarily by $\overline{\Delta_{\mathbf{g}}}$, hence we only need to show that $\overline{\Delta_{\mathbf{g}}}$ is self-adjoint. By Theorem A.10 it suffices to show that the adjoint of $\overline{\Delta_{\mathbf{g}}}$ does not have positive eigenvalues. Assume that $\overline{\Delta}_{\mathbf{g}}^* u = \lambda u$ for some positive constant λ. Then u is smooth, since by using integration by parts the adjoint of $\overline{\Delta_{\mathbf{g}}}$ is an elliptic operator. Fix a point $p \in M$. Assume that $\varphi_{r,s}$ is a function which takes the value 1 in the ball with radius s about p, the value 0 outside the ball of radius r about p, and satisfies the estimate $\|\operatorname{grad}_{\mathbf{g}}(\varphi_{r,s})\|_{\mathbf{g}} \le C\frac{1}{r-s}$. This functions can be constructed by first defining

$$\varphi(x) = \begin{cases} 1 & \text{if } x \le 1 \\ 2-x & \text{if } 1 < x < 2 \\ 0 & \text{if } x \ge 2 \end{cases},$$

and then define $\varphi_{r,s}(x) = \varphi\left(\frac{d(p,x)+s-2r}{s-r}\right)$. Note that if the manifold is complete then $\varphi_{r,s}$ has compact support, which may not be the case for non-complete manifolds. We have that

$$\begin{aligned} 0 \ge \lambda\langle \varphi_{r,s}^2 u, u\rangle_{L^2(\mathbf{g}^0)} &= \langle \varphi_{r,s}^2 u, \Delta_{\mathbf{g}} u\rangle_{L^2(\mathbf{g})} = -\langle \operatorname{grad}_{\mathbf{g}}\left(\varphi_{r,s}^2 u\right), \operatorname{grad}_{\mathbf{g}} u\rangle_{L^2(\mathbf{g})} \\ &= -2\langle u\varphi_{r,s} \operatorname{grad}_{\mathbf{g}}\varphi_{r,s}, \operatorname{grad}_{\mathbf{g}} u\rangle_{L^2(\mathbf{g})} - \varphi_{r,s}^2\langle \operatorname{grad}_{\mathbf{g}} u, \operatorname{grad}_{\mathbf{g}} u\rangle_{L^2(\mathbf{g})}. \end{aligned}$$

Using Cauchy-Schwartz,

$$\begin{aligned} \|\varphi_{r,s} \operatorname{grad}_{\mathbf{g}} u\|_{L^2(\mathbf{g})}^2 &\le 2|\langle u\varphi_{r,s} \operatorname{grad}_{\mathbf{g}}\varphi_{r,s}, \operatorname{grad}_{\mathbf{g}} u\rangle_{L^2(\mathbf{g})} \le 2\|u \operatorname{grad}_{\mathbf{g}}\varphi_{r,s}\|_{L^2(\mathbf{g})}\|\varphi_{r,s} \operatorname{grad}_{\mathbf{g}} u\|_{L^2(\mathbf{g})} \\ &\le \|\operatorname{grad}_{\mathbf{g}}\varphi_{r,s}\|_{L^\infty(\mathbf{g})}\|u\|_{L^2(\mathbf{g}^0)}\|\varphi_{r,s} \operatorname{grad}_{\mathbf{g}} u\|_{L^2(\mathbf{g})}, \end{aligned}$$

where L^∞ is the essential supremum of the function. Hence we obtain the inequality

$$\|\varphi_{r,s} \operatorname{grad}_{\mathbf{g}} u\|_{L^2(\mathbf{g})} \le \|\operatorname{grad}_{\mathbf{g}}\varphi_{r,s}\|_{L^\infty(\mathbf{g}^0)}\|u\|_{L^2(\mathbf{g}^0)}.$$

Letting $r \to \infty$, we get that $\|\varphi_{r,s} \operatorname{grad}_{\mathbf{g}} u\|_{L^2(\mathbf{g})}^2 = 0$. Thus u is constant, and since $\Delta_{\mathbf{g}}$ of a constant function is zero we are done. □

There are several ways to extend the Laplacian to forms and tensors. For example note that if δ is the adjoint of d in L^2, then we can write the Laplacian on functions by $-\delta \circ df$. Then the extension on forms, called the **Hodge Laplacian**, is defined by

$$-\delta \circ d\alpha - d \circ \delta\alpha.$$

We will however work with another extension referred to as the **Rough Laplacian**, or the **Bochner Laplacian**. This extension is defined for any tensor field F by the formula

$$\Delta_{\mathbf{g}} F = \operatorname{tr}_{\mathbf{g}} \nabla^2_{\times,\times} F.$$

On forms these two extensions does not in general coincide, but are related by the Bochner-Weitzenböck identity, see e.g. [Jos11]. Additionally, both extensions of the Laplacian are essentially self-adjoint. For a proof of this see [Str83].

2.3 Bochner Formula and Spectral Gap

On a compact Riemannian manifold the constant functions are always eigenvectors of the Laplacian with eigenvalue 0. In this case the eigenvalues can be ordered in a decreasing sequence $0 = \lambda_0 > \lambda_1 \geq \lambda_2 \geq \cdots$ where $\lambda_n \to -\infty$, which is proven in e.g. [Jos11]. Note that if the manifold is compact the eigenspace at 0 consists of the constant functions, since

$$\langle \Delta_{\mathbf{g}} f, f \rangle_{L^2(\mathbf{g}^0)} = -\langle df, df \rangle_{L^2(\mathbf{g}^*)}.$$

The value of λ_1 is called the **spectral gap** of the Laplacian. More generally, on a manifold the value of the first nonzero eigenvalue is called the spectral gap.

In this section we aim to prove the Lichnerowicz estimate: Which states that if (M, d) is a complete Riemannian manifold such that there exists a positive number K such that $\operatorname{Ric}^{\nabla}(v, v) \geq K\|v\|^2_{\mathbf{g}}$ then

$$-\lambda_1 \geq \frac{n}{n-1} K,$$

where n is the dimension of the manifold. By the Bonnet-Myers Theorem (Theorem 2.11) the boundedness Ricci curvature from bellow by a positive constant forces the manifold to be compact.

The first step in both the Riemannian and sub-Riemannian setting is the Bochner formula, which gives us a concrete formula for the commutator of the differential and the Laplace operator.

Proposition 2.18 (Bochner's Formula)**.** *For any Riemannian manifold $(M, \mathbf{g})$ with Laplacian $\Delta_{\mathbf{g}}$, we have that*

$$\Delta_{\mathbf{g}} df - d\Delta_{\mathbf{g}} f = \operatorname{Ric}^{\nabla}\left(\operatorname{grad}_{\mathbf{g}} f, \cdot\right).$$

Proof. Let $X_1, \ldots, X_n$ be an orthonormal local frame parallel at the point x. Doing the computations at this point we get

$$\begin{aligned}
\left(\Delta_{\mathbf{g}} df - d\Delta_{\mathbf{g}} f\right)(X_j) &= \sum_{i=1}^{n} \left(\nabla_{X_i}\nabla_{X_i} df\right)(X_j) - d\left(\nabla_{X_i}\nabla_{X_i} f - \nabla_{\nabla_{X_i}X_i} f\right)(X_j) \\
&= \sum_{i=1}^{n} \nabla_{X_i}\left(\left(\nabla_{X_i} df\right)(X_j)\right) - X_j\left(\nabla_{X_i}\left(df\left(X_i\right)\right) - df\left(\nabla_{X_i}X_i\right)\right) \\
&= \sum_{i=1}^{n} \left(\nabla_{X_i}\nabla_{X_j} df\right)(X_i) - \left(\nabla_{X_j}\nabla_{X_i} df\right)(X_i) \\
&= \sum_{i=1}^{n} \left(\sharp^{\mathbf{g}}\nabla_{X_i}\nabla_{X_j} df - \sharp^{\mathbf{g}}\nabla_{X_j}\nabla_{X_i} df\right)\left(\flat^{\mathbf{g}} X_i\right) \\
&= \operatorname{Ric}^{\nabla}\left(X_j, \operatorname{grad}_{\mathbf{g}} f\right).
\end{aligned}$$

Since the point was arbitrary and the Ricci curvature is symmetric we are done. □

Next we want to evaluate both sides in df and integrate. The only problem is the part containing $\langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*}$, which is the only term which can not be written as some constant times $\int_M \|df\|_{\mathbf{g}^*} d\,\mathrm{vol}_{\mathbf{g}}$. Hence we will need the following lemma.

Lemma 2.19. *We have that*

$$\frac{1}{2}\Delta_{\mathbf{g}}\left(\|df\|^2_{\mathbf{g}^0_1}\right) - \|\nabla^2 f\|^2_{\mathbf{g}^0_2} = \langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*}.$$

Proof. Using the basis given in Proposition 2.2 about the point $x \in M$ and doing the computation at p we obtain

$$\begin{aligned}
\Delta_{\mathbf{g}}\left(\langle df, df\rangle_{\mathbf{g}^*}\right) &= \sum_{i=1}^{n} X_i\left(X_i\langle \sharp^{\mathbf{g}} df, \sharp^{\mathbf{g}} df\rangle_{\mathbf{g}}\right) = \sum_{i=1}^{n} X_i\left(2\langle \nabla_{X_i}\sharp^{\mathbf{g}} df, \sharp^{\mathbf{g}} df\rangle_{\mathbf{g}}\right) \\
&= \sum_{i=1}^{n} X_i\left(2\langle \sharp^{\mathbf{g}}\nabla_{X_i} df, \sharp^{\mathbf{g}} df\rangle_{\mathbf{g}}\right) \\
&= 2\sum_{i=1}^{n} \langle \nabla_{X_i}\sharp^{\mathbf{g}}\nabla_{X_i} df, \sharp^{\mathbf{g}} df\rangle_{\mathbf{g}} + \langle \sharp^{\mathbf{g}}\nabla_{X_i} df, \nabla_{X_i}\sharp^{\mathbf{g}} df\rangle_{\mathbf{g}} \\
&= 2\sum_{i=1}^{n} \langle \nabla_{X_i}\nabla_{X_i} df, df\rangle_{\mathbf{g}^*} + \langle \nabla_{X_i} df, \nabla_{X_i} df\rangle_{\mathbf{g}^*}.
\end{aligned}$$

Recognizing the term on the right to be $\langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*}$ we get

$$\begin{aligned}
2\langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*} + 2\sum_{i,j=1}^{n}\left(\nabla_{X_i} df\left(X_j\right)\right)^2 &= 2\langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*} + 2\sum_{i,j=1}^{n}\left(X_i\left(df\left(X_j\right)\right) - df\left(\nabla_{X_i}X_j\right)\right)^2 \\
&= 2\langle \Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*} + 2\langle \nabla^2 f, \nabla^2 f\rangle_{\mathbf{g}^0_2}.
\end{aligned}$$

□

Next thing to do is to give a bound for the term $\|\nabla^2 f\|^2_{\mathbf{g}^0_2}$.

Lemma 2.20. *We have the following bound on the Hessian*

$$\|\nabla^2 f\|^2_{\mathbf{g}^0_2} \geq \frac{1}{n}\left(\Delta_{\mathbf{g}} f\right)^2.$$

Proof. Denote by $\alpha_{ij} = \nabla^2_{X_i,X_j} f$, where X_i is an orthonormal frame given by Proposition 2.2 about the point $x \in M$. Using this notation $\nabla^2 f = \sum_{i,j=1}^n \alpha_{ij}\omega^i \otimes \omega^j$, where ω^i is the dual frame for X_i. Then

$$\|\nabla^2 f\|^2_{\mathbf{g}^0_2} = \sum_{i,j=1}^n \alpha_{ij}^2 = \|\left(\alpha_{ij}\right)\|_2^2,$$

where $\|\cdot\|_2$ denotes the Hilbert-Schmidt norm of a matrix. By definition of $\Delta_{\mathbf{g}}$ we have that, at the point x,

$$\left(\Delta_{\mathbf{g}} f\right)^2 = \left(\sum_{i=1}^n \nabla^2_{X_i,X_i} f\right)^2 = \left(\sum_{i=1}^n \alpha_{ii}\right)^2 = \operatorname{tr}\left(\alpha_{ij}\right)^2.$$

Hence it is enough to show that the inequality holds at the point x, i.e. show that it holds for matrices endowed with the Hilbert-Schmidt norm. Define the projection $P : M_n(\mathbf{R}) \to M_n(\mathbf{R})$ by $P(A) = \frac{\operatorname{tr}(A)}{n} I$. It is straightforward to show that all non-zero projections have operator norm equal to 1. Hence we have $\frac{\operatorname{tr}(A)}{\sqrt{n}} = \|PA\|_2 \leq \|P\|\|A\|_2 = \|A\|_2$. □

Finally we are in a position to prove the Lichnerowicz estimate first proven in [Lic58].

Theorem 2.21 (Lichnerowicz estimate)**.** *Let $(M, \mathbf{g})$ be a complete Riemannian manifold of dimension n such that $\operatorname{Ric}^\nabla(v,v) \geq K\|v\|^2_{\mathbf{g}}$, for a positive constant K. Then*

$$-\lambda_1 \geq \frac{n}{n-1}K.$$

Proof. Using Lemma 2.19 together with the Bochner formula we have that

$$\frac{1}{2}\Delta_{\mathbf{g}}\left(\langle df, df\rangle_{\mathbf{g}^*}\right) - \langle d\Delta_{\mathbf{g}} f, df\rangle_{\mathbf{g}^*} = \|\nabla^2 f\|^2_{\mathbf{g}^0_2} + \operatorname{Ric}^\nabla\left(\operatorname{grad}_{\mathbf{g}}(f), \operatorname{grad}_{\mathbf{g}}(f)\right). \tag{2.3}$$

Hence we get the inequality

$$\frac{1}{2}\Delta_{\mathbf{g}}\|df\|^2_{\mathbf{g}^*} - \langle d\Delta_{\mathbf{g}} f, df\rangle_{\mathbf{g}^*} \geq \frac{1}{n}\left(\Delta_{\mathbf{g}} f\right)^2 + K\|df\|^2_{\mathbf{g}^*},$$

by using Lemma 2.20. If we let f be an eigenvector of λ_1 and integrate with respect to the volume density we have that

$$\begin{aligned}
\frac{1}{2}\int_M \Delta_{\mathbf{g}}\|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} - \int_M \langle d\Delta_{\mathbf{g}} f, df\rangle_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} &\geq \int_M \frac{1}{n}\left(\Delta_{\mathbf{g}} f\right)^2 d\operatorname{vol}_{\mathbf{g}} + K\int_M \|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} \\
-\lambda_1 \int_M \langle df, df\rangle_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} &\geq \frac{\lambda_1}{n}\int_M \left(\Delta_{\mathbf{g}} f\right) f d\operatorname{vol}_{\mathbf{g}} + K\int_M \|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} \\
-\lambda_1 \int_M \|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} &\geq \frac{-\lambda_1}{n}\int_M \|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} + K\int_M \|df\|^2_{\mathbf{g}^*} d\operatorname{vol}_{\mathbf{g}} \\
-\lambda_1 &\geq \frac{-\lambda_1}{n} + K \\
-\lambda_1 &\geq \frac{n}{n-1}K.
\end{aligned}$$

□

The sphere S^n endowed with the standard round metrics, i.e. the induced metric on S^n from $\mathbf{R}^{n+1}$, is one of the examples of a manifold which satisfies all the conditions in the theorem above. In this case it can be showed that the bound is sharp, which is proven in [Oba72].

Theorem 2.22 (Obata Sphere Theorem)**.** *Let $(M, \mathbf{g})$ be a complete manifold with $\operatorname{Ric}^{\nabla}(v,v) \geq K\|v\|_{\mathbf{g}}^2$ where the first eigenvalue is exactly $\lambda_1 = -\frac{n}{n-1}K$. Then $(M, \mathbf{g})$ is isometric to S^n with the round metric of radius $R = \sqrt{\frac{n-1}{K}}$.*

Proof. Let us for simplicity only do the proof when $R = 1$. Let f be an eigenvector with eigenvalue $-n$. Then

$$\begin{aligned} 0 &\leq \|\nabla^2 f\|_{\mathbf{g}_2^0}^2 - \frac{1}{n}\left(\nabla_{\mathbf{g}} f\right)^2 + \operatorname{Ric}\,(df, df) - K\|df\|_{\mathbf{g}^*}^2 \\ &= \frac{1}{2}\Delta_{\mathbf{g}}\left(\|df\|_{\mathbf{g}^*}^2\right) - \langle d\Delta_{\mathbf{g}} df, df\rangle_{\mathbf{g}^*} - \frac{1}{n}\left(\Delta_{\mathbf{g}} f\right)^2 - K\|df\|_{\mathbf{g}^*}^2, \end{aligned}$$

by using equation 2.3 together with the assumptions. Integrating the right hand side we get that it is zero, hence the inequality presented in equation 2.3 is an equality. Then

$$\frac{1}{2}\Delta_{\mathbf{g}}\left(\|df\|_{\mathbf{g}}^2 + f^2\right) = 0, \tag{2.4}$$

which implies that $\|df\|_{\mathbf{g}}^2 + f^2$ is constant, and without loss of generality we can assume $\|df\|_{\mathbf{g}}^2 + f^2 = 1$. Since the manifold is compact, the function f obtain a minimum and a maximum. Since $\|df\|_{\mathbf{g}}^2 + f^2 = 1$, the maximum of f is 1 and the minimum is -1. Let γ be a minimal unit speed geodesic from q to p. Then we have that

$$l\left(\gamma\right) = \int_{\gamma} ds = \int_{\gamma} \frac{\|df\|_{\mathbf{g}^*}}{\sqrt{1-f^2}} \geq \int_{-1}^{1} \frac{dt}{\sqrt{1-t^2}} dt = \pi,$$

since $\|df\|^2 \leq 1dt$. Hence we can use Theorem 2.12 and we have the result. ☐

There are several other bounds on the spectral gap of the Laplacian. For instant in [ZY84] it is proven that if $(M, \mathbf{g})$ is a compact manifold with Ricci curvature satisfying $\operatorname{Ric}\,(v,v) \geq 0$ for $v \in TM$, then the following bound on the first eigenvalue $\lambda_1 \leq -\frac{\pi^2}{\operatorname{diam}(M)^2}$ holds. There are also several comparing estimates using the volume. For other bounds, and more information on the eigenvalues of the Laplace operator see for instant [Xia13] and [Cha84].

3 Sub-Riemannian Geometry

Sub-Riemannian geometry is more general than Riemannian geometry in the way that we allow the fiber metric to only be defined on a subbundle of the tangent bundle. Intuitively, this potentially give us restrictions on the "legal" directions we can move in when requiring the velocity to be tangent to the subbundle.

3.1 Definitions in Sub-Riemannian Geometry

Definition 3.1. Let M be a manifold, together with a subbundle $\mathcal{H}$ of the tangent bundle. Then a **sub-Riemannian manifold** is a triple $(M, \mathcal{H}, \mathbf{h})$, where $\mathbf{h}$ is a fiber-metric on $\mathcal{H}$. In this setting, the subbundle $\mathcal{H}$ is called the **horizontal distribution** and $\mathbf{h}$ is called the sub-Riemannian metric.

We will use the letter d to refer to the rank of $\mathcal{H}$, and n to refer to the dimension of the manifold. The vector fields in $\Gamma(\mathcal{H})$ are called **horizontal vector fields**. Given a sub-Riemannian metric $\mathbf{h}$ we can define the analogue of the flat and sharp operator as in Riemannian geometry. The definition of flat with respect to $\mathbf{h}$ then becomes

$$\flat^{\mathbf{h}} : \mathcal{H} \to \mathcal{H}^*,\ \flat^{\mathbf{h}}(v) = \mathbf{h}(v, \cdot),$$

for $v \in \mathcal{H}$, while the sharp operator with respect to $\mathbf{h}$ is defined to be

$$\sharp^{\mathbf{h}} : T^*M \to \mathcal{H},\ \sharp^{\mathbf{h}}(\omega) = \left(\flat^{\mathbf{h}}\right)^{-1}(\omega|_{\mathcal{H}}),$$

for $\omega \in T^*M$.

Using the sharp operator we can define the **cometric** with respect to $\mathbf{h}$ by

$$\mathbf{h}^* : T^*M \times T^*M \to T^*M,\ \langle \alpha, \beta \rangle_{\mathbf{h}^*} = \alpha\left(\sharp^{\mathbf{h}}(\beta)\right).$$

The cometric $\mathbf{h}^*$ is symmetric, since given any two elements $\alpha, \beta \in T^*M$ such that $\sharp^{\mathbf{h}}(\alpha) = v$ and $\sharp^{\mathbf{h}}(\beta) = w$ we have that

$$\langle v, w \rangle_{\mathbf{h}} = \langle v, \cdot \rangle_{\mathbf{h}}(w) = \alpha|_{\mathcal{H}}\left(\sharp^{\mathbf{h}}\beta\right) = \langle \alpha, \beta \rangle_{\mathbf{h}^*}.$$

Note that $\sharp^{\mathbf{h}}(\alpha) = \mathbf{h}^*(\alpha, \cdot)$. An additional property of the cometric is that $\mathbf{h}^*$ is zero on the annihilator of $\mathcal{H}$ in T^*M, i.e. the subbundle of T^*M given by

$$\{\alpha \in T^*M : \alpha(v) = 0 \,\forall\, v \in \mathcal{H}\}.$$

It is also possible to define the cometric by the following two properties: $\mathbf{h}^*$ is zero on the annihilator of $\mathcal{H}$, and

$$\left\langle \sharp^{\mathbf{h}}\alpha, \sharp^{\mathbf{h}}\beta \right\rangle_{\mathbf{h}} = \langle \alpha, \beta \rangle_{\mathbf{h}^*},$$

for all $\alpha, \beta \in T^*M$. The main advantage of working with the cometric instead of the sub-Riemannian metric is that the cometric is defined on the entire T^*M instead of a subbundle. Later we will also encounter the cometric as the symbol of the sub-Laplacian, which gives us another reason for preferring the cometric rather than the sub-Riemannian metric.

Let $\gamma : [a, b] \to M$ be a continuous curve such that $\dot{\gamma} \in \mathcal{H}$ almost everywhere. We say that γ is a **horizontal absolutely continuous** curve if it satisfy

$$\frac{d}{dx}\int_a^x \|\dot{\gamma}(t)\|_{\mathbf{h}} dt = \|\dot{\gamma}(x)\|_{\mathbf{h}}$$

almost everywhere. The **length** of γ is then defined to be

$$l(\gamma) = \int_a^b \|\dot{\gamma}\|_{\mathbf{h}} dt.$$

Denote by $C_{ac}^{\mathcal{H}}(A, B)$ the set of all horizontal absolutely continuous curves connecting the two points A and B on the manifold. We define the **distance** between A and B, denoted $d(A, B)$, to be

$$d(A, B) = \inf_{\gamma \in C_{ac}^{\mathcal{H}}(A,B)} l(\gamma),$$

where we use the convention $\inf \varnothing = \infty$. There are many cases in which the distance between two points may be infinite. For instance, if $\mathcal{H}$ is an integrable distribution, i.e. $[X, Y] \in \Gamma(\mathcal{H})$ for all $X, Y \in \Gamma(\mathcal{H})$, which by Frobenius Theorem (see e.g. [War83]), ensures us that the motion, when restricted to be horizontal, can not leave a submanifold of dimension d. In Section 3.2 we will discuss Chows theorem which gives us a criteria for when the distance between any two points is finite.

3.2 Bracket Generating

Definition 3.2. Let M be a manifold with a distribution $\mathcal{H}$ defined on M. Define the **Lie hull** to be the span of all iterated brackets of horizontal vector fields, i.e.

$$\operatorname{Lie}(\mathcal{H}) = \operatorname{Span}\{\Gamma(\mathcal{H}), [\Gamma(\mathcal{H}), \Gamma(\mathcal{H})], [\Gamma(\mathcal{H}), [\Gamma(\mathcal{H}), \Gamma(\mathcal{H})]], \ldots\}.$$

A distribution $\mathcal{H}$ is said to be **bracket generating** if $\operatorname{Lie}(\mathcal{H})_x = T_xM$ for all $x \in M$.

Define $\mathcal{H}_1 = \Gamma(\mathcal{H})$ and define $\mathcal{H}_i$ inductively by

$$\mathcal{H}_i = \mathcal{H}_{i-1} + [\mathcal{H}_1, \mathcal{H}_{i-1}].$$

Then assuming $\mathcal{H}$ is bracket generating, there exists a minimal number $r(x)$ such that

$$\mathcal{H}_1(x) \subset \mathcal{H}_2(x) \subset \cdots \subset \mathcal{H}_r(x) = T_xM,$$

where $\mathcal{H}_i(x) = \{X(x) : X \in \mathcal{H}_i\}$. Denote by $n_i(x)$ the dimension of $\mathcal{H}_i(x)$. The tuple of natural number $(n_1(x), \ldots, n_r(x))$ is called the **growth vector** at the point x. We say that x is a **regular point** if the growth vector is constant in some neighborhood of x. If every point in M is regular, i.e. the growth vector is constant, then $(M, \mathcal{H}, \mathbf{h})$ is called **equiregular**. In this case the number r is called the **step** of the distribution $\mathcal{H}$.

In Section 4 we are going to assume that the horizontal distribution $\mathcal{H}$ is bracket generating. Assuming this have many important consequences as we will see.

Proposition 3.3. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold where the distribution $\mathcal{H}$ is bracket generating. If the horizontal gradient of a smooth function f is zero, i.e. $\sharp^{\mathbf{h}} df = 0$, then the function f is constant.*

Proof. Since $\mathcal{H}$ is bracket generating, it suffices to show that for all $X_1, \ldots, X_m \in \Gamma(\mathcal{H})$ we have that

$$df\left([X_1, [X_2, [\cdots, [X_{n-1}, X_n]]]]\right) = [X_1, [X_2, [\cdots, [X_{n-1}, X_n]]]]\,(f) = 0.$$

Since $X_i\,(f) = 0$, and

$$\begin{aligned}
[X_1, [X_2, [\cdots, [X_{n-1}, X_n]]]]\,(f) &= X_1\,[X_2, \cdots, [X_{n-1}, X_n]]\,(f) - [X_2, \cdots, [X_{n-1}, X_n]]\,X_1\,(f) \\
&= X_1\,[X_2, \ldots, [X_{n-1}, X_n]]\,(f) \\
&\;\;\vdots \\
&= X_1 X_2 \cdots X_n\,(f) - X_1 X_2 \cdots X_{n-2} X_n X_{n-1}\,(f) = 0,
\end{aligned}$$

we are done. □

A more important consequence is the Chow-Rashevskii Theorem below, which was first proved in [Ras38] and independently by [Cho39]. A proof of the theorem can also be found in [Mon02]. The proof of the theorem, at least in the version given in [Mon02], uses the fact that the endpoint map becomes an open map whenever the distribution $\mathcal{H}$ is bracket generating.

Theorem 3.4 (Chow-Rashevskii Theorem)**.** *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold, where the distribution $\mathcal{H}$ is assumed to be bracket generating. Then it is possible to connect any two point with a horizontal curve.*

This means that the distance between any pair of points is finite, and (M, d) becomes a metric space. It is also possible to show that the topology on M induced by the distance function d is the same as its original topology when $\mathcal{H}$ is bracket generating, see [Mon02]. As in any metric space we can define the Hausdorff measure and dimension. In sub-Riemannian geometry the Hausdorff dimension in general is bigger than the topological dimension of the manifold.

Theorem 3.5 (Michell's Theorem)**.** *For a sub-Riemannian manifold $(M, \mathcal{H}, \mathbf{h})$ where $\mathcal{H}$ is bracket generating we have that around every regular point x the Hausdorff dimension is given by $\sum_{i=1}^{r} i\,(n_i(x) - n_{i-1}(x))$.*

Example 3.6 (The Heisenberg Group)**.** Let us consider $\mathbf{R}^3$, with coordinates denoted by (x, y, z), and let X, Y, Z be the vector spaces defined by $X = \partial_x - \frac{1}{2}y\partial_z$, $Y = \partial_y + \frac{1}{2}x\partial_z$ and $Z = \partial_z$. Then the vector fields X, Y, Z satisfies the bracket relations $[X, Y] = -Z$, $[X, Z] = [Y, Z] = 0$. Define $\mathcal{H} = \operatorname{Span}\{X, Y\}$, and let the sub-Riemannian metric $\mathbf{h}$ be the sub-Riemannian metric making X, Y an orthonormal frame. Then $\mathcal{H}$ is bracket generating. It is also easily seen that the growth vector is $(2, 3)$ at every point, hence $\left(\mathbf{R}^3, \mathcal{H}, \mathbf{h}\right)$ is equiregular.

Example 3.7. Let G be any Lie group with Lie algebra $\mathfrak{g}$. Assume $\mathfrak{h}$ is a subspace of $\mathfrak{g}$ such that the minimal ideal containing $\mathfrak{h}$ is $\mathfrak{g}$. Define $\mathcal{H}$ by left translating $\mathfrak{h}$. Then $\mathcal{H}$ is bracket generating. Let $x_1, \ldots, x_d$ be a basis for $\mathfrak{h}$, again, define $X_1, \ldots, X_d$ by left translation. Then we can define $\mathbf{h}$ such that $X_1, \ldots, X_d$ is an orthonormal basis. We will encounter several examples of this structure when we are calculating the spectral gap. See the subsections 4.3, 4.4 and 4.5.

3.3 Geodesics

As in Section 2, we say that an absolutely continuous horizontal curve $\gamma \in C_{ac}^{\mathcal{H}}(A, B)$ is **minimizing** if it minimizes the distance.

Definition 3.8. An absolutely continuous horizontal curve $\gamma \in C^{\mathcal{H}}_{ac}(A, B)$ is called a (sub-Riemannian) **geodesic** if γ locally is a minimizing curve.

Assuming that $\mathcal{H}$ is bracket generating implies the following result. The proof of this result can be found in [Mon02].

Theorem 3.9. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold, where $\mathcal{H}$ is bracket generating. Then for any fixed point p there exists a neighborhood U of p such that any point $q \in U$ can be connected to p by a minimizing curve. If we additionally assume that (M, d) is a complete metric space, then the previous statement holds for $U = M$.*

On any manifold we can define a canonical symplectic structure on the cotangent bundle T^*M. If $(x_1, \ldots, x_n, \mathcal{U})$ is a chart on M, then it induces the chart $(x_1, \ldots, x_n, p_1, \ldots, p_n)$ on T^*M, where $p_i = dx_i$. Then the canonical symplectic form on T^*M is defined by $\omega = \sum_{i=1}^n dx_i \wedge dp_i$. The **sub-Riemannian Hamiltonian** is the function $H : T^*M \to \mathbf{R}$ defined by

$$H(\alpha) = \frac{1}{2}\|\alpha\|^2_{\mathbf{h}^*}.$$

By using polarization, the sub-Riemannian Hamiltonian uniquely defines the sub-Riemannian structure. Define the Hamiltonian vector field of the Hamiltonian H by $\omega\left(\vec{H}, \cdot\right) = dH$. Locally the integral curves of $\vec{H}$ can be written as solution of

$$\begin{cases} \dot{x}_i = \frac{\partial H}{\partial p_i}, \\ \dot{p}_i = -\frac{\partial H}{\partial x_i}. \end{cases}$$

The equations above is called the Hamiltonian equations of the Hamiltonian function H. One can show that if $(\gamma(t), \xi(t)) \in T^*M$ is an integral curve of the Hamiltonian equations then $\gamma(t)$ is a geodesic, which is showed in [Mon02]. The reverse statement is however not true. Hence we will divide geodesics into **normal geodesics** which are projections of integral curves of the sub-Riemannian Hamiltonian equations, and **abnormal geodesics** which are the geodesics which are not normal. An open problem in sub-Riemannian geometry is whether all abnormal geodesics are smooth. The normal geodesics are smooth, being the projections of integral curves of a smooth vector field. Another feature of normal geodesics is that locally they are the unique geodesic connecting the endpoints. In other words if γ is a normal geodesic, then at point $x = \gamma(t_0)$ there exists an ϵ such that γ is the unique minimizing curve connecting the points $\gamma(t_0 - \varepsilon)$ and $\gamma(t_0 + \varepsilon)$.

3.4 Hypoelliptic operators

For us the most important consequence of $\mathcal{H}$ being bracket generating is the relation to hypoelliptic operators given in Theorem 3.11.

Definition 3.10. Let L be a second order differential operator defined on a manifold M. The operator L is said to be **hypoelliptic** if for all open subsets $V \subset M$, $f \in C^\infty(V)$ and distributions u such that $Lu = f$ implies that $u \in C^\infty(V)$.

As noted in the appendix, all elliptic operators are hypoelliptic by Corollary A.19. The vector fields $X_0, \ldots, X_m$ is said to be **bracket generating** if they span a bracket generating distribution. Hörmander gave in the 60's the following classification of hypoelliptic operators in [Hör67].

Theorem 3.11. *Let L be a second order differential operator on the form $L = \sum_{i=1}^{k} X_i^2 + X_0 + c$, where $X_0, \ldots, X_k$ are vector fields defined on an open bounded set M, with the additional property that $Lie\,(X_0, \ldots, X_k)$ is bracket generating. Then L is hypoelliptic.*

Note that in the Theorem of Hörmander the vector fields can be degenerate, i.e. there might exist points $x \in M$ where $X_{ix}(f) = 0$ for all $f \in C^{\infty}(U)$. If given an operator L on the form in Theorem 3.11, then $L - \lambda$ is hypoelliptic. It follows that if λ is an eigenvalue with eigenvector u, then u has to be smooth, since $(L - \lambda)\, u = 0$ is a smooth function.

When we define sub-Laplacian in Section 3.5, the bracket generating condition on the metric will make the operator hypoelliptic. From here on out, we will therefore assume that whenever $(M, \mathcal{H}, \mathbf{h})$ is a sub-Riemannian manifold then $\mathcal{H}$ is bracket generating.

3.5 Sub-Laplacian and More Theory on Affine Connections

In Riemannian geometry we defined the Levi-Civita connection and used it to define other geometric invariants like the curvature and geodesics. For sub-Riemannian geometry, we can not ask for a symmetric and compatible affine connection without making the distribution integrable. If we do assume that the connection is symmetric and compatible with $\mathbf{h}^*$, this will imply that $\mathcal{H}$ is an integrable distribution. Since we want the affine connection to reflect the sub-Riemannian structure we are going to ask for the affine connection to be compatible with the sub-Riemannian metric $\mathbf{h}$, i.e. $\nabla \mathbf{h}^* = 0$.

Lemma 3.12. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold, and let ∇ be a linear connection on M. Then the following are equivalent*

a) $\mathbf{h}$ *is compatible with* ∇.

b) for all $X \in \Gamma(TM)$ and all $\alpha \in \Omega(M)$ we have that $\sharp^{\mathbf{h}} \nabla_X \alpha = \nabla_X \sharp^{\mathbf{h}} \alpha$.

c) for all $X \in \Gamma(TM)$ and $Y, Z \in \Gamma(\mathcal{H})$ we have that $\nabla_X Y \in \Gamma(\mathcal{H})$, and furthermore

$$X\,(\mathbf{h}\,(Y, Z)) = \mathbf{h}\,(\nabla_X Y, Z) + \mathbf{h}\,(Y, \nabla_X Z)\,.$$

Proof. To show the equivalence of a) and b) we compute

$$\nabla_X\,(\mathbf{h}^*\,(\alpha, \cdot)) = (\nabla_X \mathbf{h}^*)\,(\alpha, \cdot) + \mathbf{h}^*\,(\nabla_X \alpha, \cdot)\,.$$

Note that if we denote $\sharp^{\mathbf{h}} \alpha = Y$ and $\beta \in \Omega(M)$, then $\beta(Y) = \beta\left(\sharp^{\mathbf{h}} \alpha\right) = \mathbf{h}^*(\alpha, \beta)$. Hence we get that $\mathbf{h}^*(\alpha, \cdot) = \sharp^{\mathbf{h}} \alpha$. Using this

$$\nabla_X \sharp^{\mathbf{h}} \alpha = (\nabla_X \mathbf{h}^*)\,(\alpha, \cdot) + \sharp^{\mathbf{h}} \nabla_X \alpha,$$

and it follows that a) and b) are equivalent.

For the implication either a) or b) imply c), note that any element in $Y \in \Gamma(\mathcal{H})$ can be written as $\sharp^{\mathbf{h}} \alpha$ for some 1-form. The affine connection preserves $\mathcal{H}$ since $\nabla_X Y = \sharp^{\mathbf{h}} \nabla_X \alpha$, where the right hand side is horizontal. For the last part, denote by α_1, α_2 1-forms such that $\sharp^{\mathbf{h}} \alpha_1 = Y_1, \sharp^{\mathbf{h}} \alpha_2 = Y_2$. Then

$$\begin{aligned}
\nabla_X \langle Y_1, Y_2 \rangle_{\mathbf{h}} &= \nabla_X \langle \alpha_1, \alpha_2 \rangle_{\mathbf{h}^*} \\
&= (\nabla_X \mathbf{h}^*)\,(\alpha_1, \alpha_2) + \langle \nabla_X \alpha_1, \alpha_2 \rangle_{\mathbf{h}^*} + \langle \alpha_1, \nabla_X \alpha_2 \rangle_{\mathbf{h}^*} \\
&= \langle \nabla_X Y_1, Y_2 \rangle_{\mathbf{h}} + \langle Y_1, \nabla_X Y_2 \rangle_{\mathbf{h}}.
\end{aligned}$$

Lastly, let us show that c) implies b). Since ∇ preserves $\mathcal{H}$, both $\sharp^{\mathbf{h}}\nabla_X\alpha$ and $\nabla_X\sharp^{\mathbf{h}}\alpha$ is horizontal, hence it is enough to prove that for all $Y \in \Gamma(\mathcal{H})$ we have that $\mathbf{h}\left(\sharp^{\mathbf{h}}\nabla_X\alpha, Y\right) = \mathbf{h}\left(\nabla_X\sharp^{\mathbf{h}}\alpha, Y\right)$. We will also use that ∇ restricted to $\mathcal{H}$ is an affine connection, which means that we can define an affine connection on $\mathcal{H}^*$ by the formula $(\nabla_X\beta)(Y) = X(\beta(Y)) - (\nabla_X Y)(\beta)$, for $\beta \in \mathcal{H}^*$ and $Y \in \mathcal{H}$. In fact we can define an affine connection on any tensor bundle over $\mathcal{H}$ in the same way as in Lemma 2.4. Let $Z = \sharp^{\mathbf{h}}\alpha$, then by assumption we get that

$$\begin{aligned}\mathbf{h}\left(\sharp^{\mathbf{h}}\alpha, Y\right) &= \mathbf{h}(Z, Y) = X(\mathbf{h}(Y, Z)) - (\nabla_X Y)\left(\flat^{\mathbf{h}} Z\right) \\ &= \left(\nabla_X \flat^{\mathbf{h}} Z\right)(Y) = (\nabla_X\alpha)(Y) = \mathbf{h}\left(\sharp^{\mathbf{h}}\nabla_X\alpha, Y\right),\end{aligned}$$

where the last equality follows from that $\nabla_X\flat^{\mathbf{h}}Z = \nabla_X\alpha|_{\mathcal{H}}$. □

If $\mathbf{h}$ is a sub-Riemannian metric on $\mathcal{H}$, then one can always extend it to a Riemannian metric. Such an extension is known as a **taming metric**. In other words a taming metric is a Riemannian metric $\mathbf{g}$ which satisfies $\mathbf{g}|_{\mathcal{H}} = \mathbf{h}$, and in this setting we say that **g tames h**. Note that (unless $\mathcal{H} = TM$) there is not a unique choice of taming metric, and hence the taming metric will always depend upon choice. With a taming metric in place it makes sense to talk about an orthogonal compliment to $\mathcal{H}$. The orthogonal compliment of $\mathcal{H}$, which we will denote by $\mathcal{V}$, is called the **vertical bundle**. In this setting we can decompose $TM = \mathcal{H} \oplus \mathcal{V}$. When $\mathbf{g}$ is restricted to $\mathcal{V}$ we get a sub-Riemannian metric on the vertical bundle which we will denote by $\mathbf{g}_{\mathcal{V}}$. When given a decomposition of the tangent bundle $TM = \mathcal{H} \oplus \mathcal{V}$ into horizontal and vertical part, we will denote by $\mathrm{pr}_{\mathcal{H}}$ and $\mathrm{pr}_{\mathcal{V}}$ the projection down on the horizontal and vertical bundle, respectively. This means that we can extend $\mathbf{h}$ (and $\mathbf{g}_{\mathcal{V}}$) to the entire tangent bundle by first projecting down to $\mathcal{H}$ (and $\mathcal{V}$). With this in mind, we can define a (local) **adapted frame** to be an orthonormal frame $X_1, \ldots, X_d, Z_1, \ldots, Z_{n-d}$ on the tangent bundle such that $X_i \in \Gamma(\mathcal{H})$ and $Z_j \in \Gamma(\mathcal{V})$. In Section 4 we are going to assume that ∇ is compatible with a taming metric and preserves the horizontal distribution, i.e. for $X, Y \in \Gamma(\mathcal{H})$ we have that $\nabla_X Y \in \Gamma(\mathcal{H})$.

Lemma 3.13. *Assume ∇ is a linear connection on the sub-Riemannian manifold $(M, \mathcal{H}, \mathbf{h})$, which is compatible with a taming metric $\mathbf{g}$ and additionally preserves the horizontal distribution. Then*

i) ∇ is compatible with $\mathbf{h}$ and $\mathbf{g}_{\mathcal{V}}$ and

ii) the affine connection preserves the orthogonal compliment of $\mathcal{H}$, $\mathcal{V}$, with respect to the taming metric.

Proof. Since the proof is the same for $\mathbf{h}$ and $\mathbf{g}_{\mathcal{V}}$ we will only show it for $\mathbf{h}$. By using Lemma 3.12 we need to show that for all $Y_1, Y_2 \in \Gamma(\mathcal{H})$ and all $X \in \Gamma(TM)$ we have that

$$\nabla_X(\mathbf{h}(Y_1, Y_2)) = \langle \nabla_X Y_1, Y_2\rangle_{\mathbf{h}} + \langle Y_1, \nabla_X Y_2\rangle_{\mathbf{h}}.$$

Since

$$\nabla_X(\mathbf{h}(Y_1, Y_2)) = \nabla_X(\mathbf{g}(Y_1, Y_2)) = \langle \nabla_X Y_1, Y_2\rangle_{\mathbf{g}} + \langle Y_1, \nabla_X Y_2\rangle_{\mathbf{g}} = \langle \nabla_X Y_1, Y_2\rangle_{\mathbf{h}} + \langle Y_1, \nabla_X Y_2\rangle_{\mathbf{h}},$$

it follows that $\nabla\mathbf{h}^* = 0$.

For the second part we want to show that if $Y \in \Gamma(\mathcal{V})$ then $\nabla_X Y \in \Gamma(\mathcal{V})$, which is the same as showing that for all $Y \in \Gamma(\mathcal{V})$ and $Z \in \Gamma(\mathcal{H})$ we have that $\mathbf{h}(\nabla_X Y, Z) = 0$. By compatibility of the metric

$$0 = \nabla_X \mathbf{h}(Y, Z) = \mathbf{h}(\nabla_X Y, Z) + \mathbf{h}(Y, \nabla_X Z),$$

and since $\mathbf{h}(Y, \nabla_X Z) = 0$ by assumption, we have the result. $\square$

According to the previews lemma; if we restrict the affine connection to either the horizontal or vertical bundle, then we have again an affine connection. By Corollary 2.3, we can find an orthonormal local frame parallel at the point p for $\mathcal{H}$ and $\mathcal{V}$, respectively. Hence there exist an orthonormal adapted frame parallel at the point p.

Definition 3.14. Denote by $(M, \mathcal{H}, \mathbf{h})$ a sub-Riemannian manifold, and let ∇ be a linear connection defined on M. Then we define the **rough sub-Laplacian** L^∇ with respect to the connection ∇ to be $L^\nabla = \mathrm{tr}_{\mathbf{h}} \nabla^2_{\times,\times}$.

When acting on functions this is a second order linear operator. Hence we can calculate the symbol. Let $X_1, \dots, X_d$, $Z_1, \dots, Z_{n-d}$ be an adapted orthonormal frame which is parallel at the point x, then

$$\begin{aligned}
\sigma_{L^\nabla}(df, dg) &= \frac{1}{2}\left(L^\nabla(fg) - fL^\nabla(g) - gL^\nabla(f)\right) \\
&= \frac{1}{2}\sum_{i=1}^{d} X_i(X_i(fg)) - fX_i(X_i(g)) - gX_i(X_i(f)) \\
&= \sum_{i=1}^{d} X_i(g) X_i(f) = \mathbf{h}^*(df, dg).
\end{aligned}$$

We also note that if we write L^∇ in a basis, then it has the form given in Theorem 3.11. Since we have assumed that the given sub-Riemannian manifold is bracket generating, L^∇ becomes hypoelliptic, which among other things imply that the eigenvectors are smooth.

The goal is to give a lower bound of the first nonzero eigenvalue, similar to the Lichnerowicz estimate in the Riemannian case. In the proof of the Riemannian Lichnerowicz estimate it was crucial that the Laplace operator was symmetric with respect to the volume density. We will take the volume density given by a taming metric. In this case the sub-Laplacian might not be symmetric with respect to the chosen volume density. However, in [GT16a] they give the following condition for L to be symmetric with respect to $\mathrm{vol}_{\mathbf{g}}$.

Lemma 3.15. *Let $\beta(v) = \mathrm{tr}T^\nabla(v, \times)(\times)$, then the adjoint of $\nabla_{pr_{\mathcal{H}}}$ is given by $\iota_{\sharp^{\mathbf{h}}(\beta)} - \sum_{i=1}^{d} \iota_{X_i} \nabla_{X_i}$, where $X_1, \dots, X_d$ is a local orthonormal frame for $\mathcal{H}$. Then*

$$-L^\nabla + \nabla_{\sharp^{\mathbf{h}}\beta} = \left(\nabla_{pr_{\mathcal{H}}}\right)^* \circ \nabla_{pr_{\mathcal{H}}}.$$

In the proof of this proposition they use that the adjoint of $\nabla_{\mathrm{pr}_{\mathcal{H}}\cdot}$, sending $\alpha \mapsto \nabla_{\mathrm{pr}_{\mathcal{H}}\cdot}\alpha$ is given by

$$\left(\nabla_{\mathrm{pr}_{\mathcal{H}}}\right)^* = \iota_{\sharp^{\mathbf{h}}\beta} - \sum_{i=1}^{d} \iota_{X_i} \nabla_{X_i}, \tag{3.1}$$

where $X_1, \dots, X_d, Z_1, \dots, Z_{n-d}$ is an adapted orthonormal frame.

Corollary 3.16. *The rough sub-Laplacian is symmetric with respect to the volume density of a taming metric if and only* $\operatorname{tr}T^{\nabla}(v, \times)\times = 0$ *for all* $v \in \mathcal{H}$.

By Proposition B.7, if L^{∇} is symmetric with respect to the volume density then

$$L^{\nabla}(f) = \operatorname{div}_{\mathbf{g}}\left(\sharp^{\mathbf{h}} df\right).$$

When L^{∇} is symmetric with respect to $\operatorname{vol}_{\mathbf{g}}$, then the following have the following theorem which can be found in [Str86] together with [Str83].

Theorem 3.17. *Let* $(M, \mathcal{H}, \mathbf{h})$ *be a complete sub-Riemannian manifold with a taming metric* $\mathbf{g}$. *Assume we are given a connection* ∇ *such that* L^{∇} *is symmetric with respect to* $\operatorname{vol}_{\mathbf{g}}$. *Then* L^{∇} *is essentially self-adjoint.*

The proof is essentially the same as for Theorem 2.17 only that we get that $\sharp^{\mathbf{h}} df = 0$, which by Proposition 3.3 implies that $df = 0$.

When the horizontal distribution $\mathcal{H}$ is equiregular there is a canonical choice of measure on $(M, \mathcal{H}, \mathbf{h})$ called Popp's measure, see e.g. [ABGR09]. However, we want to be able to study operators not nessesarely symmetric with respect to this measure.

4 Sub-Riemannian Lichnerowicz Estimate

Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold, and let d be the rank of the bracket generating distribution $\mathcal{H}$. Unless otherwise stated we will assume that we are given a linear connection ∇ and a taming metric $\mathbf{g}$ such that

1. ∇ is compatible with $\mathbf{g}$,
2. ∇ preserves $\mathcal{H}$ and
3. L^∇ is symmetric with respect to $\operatorname{vol}_{\mathbf{g}}$, i.e. $\operatorname{tr} T^\nabla(v, \times)\times = 0$ for all $v \in \mathcal{H}$.

Assuming M is compact, we again have that the eigenvalues of L^∇ can be ordered in a sequence $0 = \lambda_0 > \lambda_1 \geq \lambda_2 \geq \cdots$ where $\lambda_i \to -\infty$, see [FP83]. In this setting the **spectral gap** is again defined to be λ_1.

4.1 Sub-Riemannian Bochner Formula

As in the Riemannian case, the proof of the Lichnerowicz estimate will depend on the intertwining of $L^\nabla df - dL^\nabla f$. In this case we will end up with additional terms depending on the torsion which in the Riemannian case were 0. In the proof of the Sub-Riemannian Bochner formula we will need the following lemma.

Lemma 4.1. *For any $v, w \in T_xM$ we have that*

$$\nabla^2_{v,v} df(w) = \nabla^2_{v,w} df(v) - \nabla_v df\left(T^\nabla(v,w)\right) - df\left(\left(\nabla_v T^\nabla\right)(v,w)\right).$$

Proof. Let $X_1, \ldots, X_n$ be an orthonormal frame parallel at the point x. Then we can write $v = \sum_{i=1}^n a^i X_i(x)$ and $w = \sum_{i=1}^n b^i X_i(x)$. Set $X_v = \sum_{i=1}^n a^i X_i$ and $X_w = \sum_{i=1}^n b^i X_i$, which becomes parallel at the point x by linearity of the affine connection. At the point x we get that

$$\begin{aligned}\nabla^2_{X_v,X_v} df(X_w) &= \nabla_{X_v}\nabla_{X_v} df(X_w) = \nabla_{X_v}\left(\left(\nabla_{X_v} df\right)(X_w)\right) - \left(\nabla_{X_v} df\right)\left(\nabla_{X_v} X_w\right) \\ &= \nabla_{X_v}\left(\left(\nabla_{X_w} df\right)(X_v) - df\left(T^\nabla(X_v, X_w)\right)\right),\end{aligned}$$

since by Lemma 2.5 we have that

$$\left(\nabla_{X_v} df\right)(X_w) - \left(\nabla_{X_w} df\right)(X_v) = -df\left([X_v, X_w]\right).$$

This is again equal to

$$\begin{aligned}&\nabla_{X_v}\left(\left(\nabla_{X_w} df\right)(X_v)\right) - \left(\nabla_{X_v} df\right)\left(T^\nabla(X_v, X_w)\right) - df\left(\nabla_{X_v}\left(T^\nabla(X_v, X_w)\right)\right) \\ &= \left(\nabla_{X_v}\nabla_{X_w} df\right)(X_v) + \left(\nabla_{X_w} df\right)\left(\nabla_{X_v} X_w\right) - \left(\nabla_{X_v} df\right)\left(T^\nabla(X_v, X_w)\right) - df\left(\nabla_{X_v}\left(T^\nabla(X_v, X_w)\right)\right) \\ &= \left(\nabla_{X_v}\nabla_{X_w} df\right)(X_v) - \nabla_{X_v}\left(df\left(T^\nabla(X_v, X_w)\right)\right) - df\left(\nabla_{X_v} T^\nabla(X_v, X_w)\right) - df\left(T^\nabla\left(\nabla_{X_v} X_v, X_w\right)\right) \\ &\quad - df\left(T^\nabla\left(X_v, \nabla_{X_v} X_w\right)\right) \\ &= \left(\nabla_{X_v}\nabla_{X_w} df\right)(X_v) - \left(\nabla_{X_v} df\right)\left(T^\nabla(X_v, X_w)\right) - df\left(\nabla_{X_v} T^\nabla(X_v, X_w)\right), \\ &= \left(\nabla_v \nabla_w df\right)(v) - \left(\nabla_v df\right)\left(T^\nabla(v, w)\right) - df\left(\nabla_v T^\nabla(v, w)\right),\end{aligned}$$

where the next to last equality follows from

$$df\left(T^\nabla\left(\nabla_{X_v} X_w, X_v\right)\right) = T^\nabla\left(\nabla_{X_v} X_w, X_v\right)(f) = 0$$

at x. □

In this section denote by Ric the operator

$$\text{Ric}\,(\alpha) = \text{tr}_{\mathbf{h}}\left(R^{\nabla}(\times,\cdot)\,\alpha\right)(\times)\,, \tag{4.1}$$

which will be referred to as the **Ricci curvature** with respect to a sub-Riemannian structure. Define the operators $\mathcal{A} : \Omega(M) \to \Omega(M)$, where

$$\mathcal{A}(\alpha) = \text{Ric}\,(\alpha) - \alpha\left(\text{tr}_{\mathbf{h}}\left(\nabla_{\times}T^{\nabla}\right)(\times,\cdot)\right) - \alpha\left(\text{tr}_{\mathbf{h}}T^{\nabla}\left(\times, T^{\nabla}(\times,\cdot)\right)\right), \tag{4.2}$$

and $D^T : \Omega(M) \to \Omega(M)$ by

$$D^T\alpha = \text{tr}_{\mathbf{h}}\left(\nabla_{\times}\alpha\right)\left(T^{\nabla}(\times,\cdot)\right). \tag{4.3}$$

Note that the operator $\mathcal{A}$ is a tensor field, while D^T is not a tensor, and depend on the local behavior of α. The next proposition is proved by in [GT16a].

Proposition 4.2 (Sub-Riemannian Bochner formula)**.** *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold with the taming metric $\mathbf{g}$, and let ∇ be a linear connection satisfying property 1)-3) in the beginning of Section 4. Then*

$$L^{\nabla}df - dL^{\nabla}f = -2D^T df + \mathcal{A}(df)\,,$$

for any $f \in C^{\infty}(M)$.

Proof. Let $X_1, \ldots, X_d, Z_1, \ldots, Z_{n-d}$ be an adapted orthonormal frame parallel at the point x. Doing the computation at the point x, we get that

$$\begin{aligned}
\left(L^{\nabla}df - dL^{\nabla}f\right)(X_j) &= \sum_{i=1}^{d}\left(\nabla^2_{X_iX_i}df\right)(X_j) - d\left(\nabla_{X_i}\nabla_{X_i}f\right)(X_j) + d\left(\nabla_{\nabla_{X_i}X_i}f\right)(X_j) \\
&= \sum_{i=1}^{d}\left(\nabla^2_{X_iX_j}df\right)(X_i) - \left(\nabla_{X_i}df\right)\left(T^{\nabla}(X_i,X_j)\right) \\
&\quad - df\left(\left(\nabla_{X_i}T^{\nabla}\right)(X_i,X_j)\right) - \nabla_{X_j}\left(\nabla_{X_i}\nabla_{X_i}f\right) + d\left(\nabla_{\nabla_{X_i}X_i}f\right)(X_j)\,,
\end{aligned}$$

by using Lemma 4.1. By the definition of D^T we can write the equation above as

$$\begin{aligned}
&- D^T df\,(X_j) + \sum_{i=1}^{d}\left(\nabla_{X_i}\nabla_{X_j}df\right)(X_i) - df\left(\left(\nabla_{X_i}T^{\nabla}\right)(X_i,X_j)\right) - \nabla_{X_j}\left(\nabla_{X_i}\left(df\,(X_i)\right)\right) \\
&\quad + d\left(\nabla_{X_i}X_i\,(f)\right)(X_j) \\
&= -D^T df\,(X_j) + \sum_{i=1}^{d}\left(\nabla_{X_i}\nabla_{X_j}df\right)(X_i) - df\left(\left(\nabla_{X_i}T^{\nabla}\right)(X_i,X_j)\right) - \left(\nabla_{X_j}\nabla_{X_i}df\right)(X_i) \\
&\quad - \nabla_{X_j}\left(df\left(\nabla_{X_i}X_i\right)\right) + d\left(\nabla_{X_i}X_i\,(f)\right)(X_j)\,.
\end{aligned}$$

Using the definition of Ricci curvature we get

$$\begin{aligned}
-D^T df\left(X_j\right) &+ \left(\operatorname{Ric} df\right)\left(X_j\right) + \sum_{i=1}^{d} \nabla_{[X_i,X_j]} df\left(X_i\right) - \nabla_{X_j}\left(df\left(\nabla_{X_i} X_i\right)\right) \\
&\quad - df\left(\left(\nabla_{X_i} T^{\nabla}\right)\left(X_i, X_j\right)\right) + X_j\left(\nabla_{X_i} X_i\left(f\right)\right) \\
&= -D^T df\left(X_j\right) + \left(\operatorname{Ric} df\right)\left(X_j\right) + \sum_{i=1}^{d} \nabla_{[X_i,X_j]} df\left(X_i\right) - df\left(\left(\nabla_{X_i} T^{\nabla}\right)\left(X_i, X_j\right)\right) \\
&= -D^T df\left(X_j\right) + \mathcal{A}\left(df\right)\left(X_j\right) + \sum_{i=1}^{d} df\left(T^{\nabla}\left(X_i, T^{\nabla}\left(X_i, X_j\right)\right)\right) + \left(\nabla_{[X_i,X_j]} df\right)\left(X_i\right).
\end{aligned}$$

Hence if we can show that

$$\sum_{i=1}^{d} df\left(T^{\nabla}\left(X_i, T^{\nabla}\left(X_i, X_j\right)\right)\right) = -D^T\left(df\right)\left(X_j\right) - \sum_{i=1}^{d}\left(\nabla_{[X_i,X_j]} df\right)\left(X_i\right),$$

we are done. By direct computation we have that

$$\begin{aligned}
\sum_{i=1}^{d}\left(T^{\nabla}\left(X_i, T^{\nabla}\left(X_i, X_j\right)\right)\right)\left(df\right) &= \sum_{i=1}^{d} \nabla_{X_i} T^{\nabla}\left(X_i, X_j\right)\left(df\right) - \left[X_i, T^{\nabla}\left(X_i, X_j\right)\right]\left(df\right) \\
&= \sum_{i=1}^{d} X_i\left(T^{\nabla}\left(X_i, X_j\right)\left(df\right)\right) - T^{\nabla}\left(X_i, X_j\right)\left(\nabla_{X_i} df\right) - \left[X_i, T^{\nabla}\left(X_i, X_j\right)\right]\left(df\right).
\end{aligned}$$

Since $T^{\nabla}\left(X_i, X_j\right)\left(\nabla_{X_i} df\right) = D^T\left(df\right)\left(X_j\right)$, we get

$$\begin{aligned}
&- D^T\left(df\right)\left(X_j\right) + \sum_{i=1}^{d} X_i\left(T^{\nabla}\left(X_i, X_j\right)\left(df\right)\right) - \left[X_i, T^{\nabla}\left(X_i, X_j\right)\right]\left(df\right) \\
&= -D^T\left(df\right)\left(X_j\right) + \sum_{i=1}^{d} X_i\left(T^{\nabla}\left(X_i, X_j\right)\left(f\right)\right) - X_i\left(T^{\nabla}\left(X_i, X_j\right)\left(f\right)\right) + T^{\nabla}\left(X_i, X_j\right)\left(X_i\left(f\right)\right) \\
&= -D^T\left(df\right)\left(X_j\right) + \sum_{i=1}^{d} T^{\nabla}\left(X_i, X_j\right)\left(X_i\left(f\right)\right).
\end{aligned}$$

Since the frame $X_1, \ldots, X_n$ is parallel at the point x, we have that $T\left(X_i, X_j\right) = -\left[X_i, X_j\right]$, hence

$$\begin{aligned}
-D^T\left(df\right)\left(X_j\right) - \sum_{i=1}^{d} \nabla_{[X_i,X_j]}\left(df\left(X_i\right)\right) &= -D^T\left(df\right)\left(X_j\right) - \sum_{i=1}^{d}\left(\nabla_{[X_i,X_j]} df\right)\left(X_i\right) + df\left(\nabla_{[X_i,X_j]} X_i\right) \\
&= -D^T\left(df\right)\left(X_j\right) - \sum_{i=1}^{d}\left(\nabla_{[X_i,X_j]} df\right)\left(X_i\right).
\end{aligned}$$

$\square$

Continuing in the same fashion as the proof of the Lichnerowicz estimate in the Riemannian case, we have the following result.

Corollary 4.3. *With the same assumptions as in the sub-Riemannian Bochner formula we get that*

$$\left\langle L^{\nabla} f, L^{\nabla} f\right\rangle_{L^2(\mathbf{g}^0)} = \int_M \left(\|\nabla df\|^2_{\mathbf{h}^0_2} - 2\langle D^T df, df\rangle_{\mathbf{h}^*} + \langle \mathcal{A}(df), df\rangle_{\mathbf{h}^*}\right) d\,\mathrm{vol}_{\mathbf{g}}$$

and

$$\left\langle L^{\nabla} f, \Delta_{\mathbf{g}} f\right\rangle_{L^2(\mathbf{g}^0)} = \int_M \left(\|\nabla_{pr_{\mathcal{H}}} df\|^2_{\mathbf{g}^0_2} - 2\langle D^T df, df\rangle_{\mathbf{g}^*} + \langle \mathcal{A}(df), df\rangle_{\mathbf{g}^*}\right) d\,\mathrm{vol}_{\mathbf{g}}.$$

Proof. We have that

$$\begin{aligned}
\left\langle L^{\nabla} f, L^{\nabla} f\right\rangle_{L^2(\mathbf{g}^0)} &= -\int_M \langle dL^{\nabla} f, df\rangle_{\mathbf{h}^*} d\,\mathrm{vol}_{\mathbf{g}} \\
&= -\int_M \left\langle L^{\nabla} df + 2D^T df - \mathcal{A}(df), df\right\rangle_{\mathbf{h}^*} d\,\mathrm{vol}_{\mathbf{g}} \\
&= \int_M \left\langle -L^{\nabla} df - 2D^T df + \mathcal{A}(df), df\right\rangle_{\mathbf{h}^*} d\,\mathrm{vol}_{\mathbf{g}},
\end{aligned}$$

where the first equality follows from the proof of Proposition B.6. If we can show that

$$\left\langle L^{\nabla} df, df\right\rangle_{\mathbf{h}^*} = \frac{1}{2}\left(L^{\nabla}\left(\|df\|^2_{\mathbf{h}^*}\right)\right) - \langle \nabla df, \nabla df\rangle_{\mathbf{h}^0_2}$$

we are done, since $\int_M L^{\nabla}(g)\, d\,\mathrm{vol}_{\mathbf{g}} = \int_M L^{\nabla}(1)\, g d\,\mathrm{vol}_{\mathbf{g}} = 0$ for any function g. Let $X_1, \ldots, X_d$, $Z_1, \ldots, Z_{n-d}$ be an adapted orthonormal frame parallel at the point x. Then

$$\begin{aligned}
\sum_{i=1}^{d} \langle \nabla_{X_i} \nabla_{X_i} df, df\rangle_{\mathbf{h}^*} &= \sum_{i=1}^{d} \langle \nabla_{X_i} \nabla_{X_i} \sharp^{\mathbf{h}} df, \sharp^{\mathbf{h}} df\rangle_{\mathbf{h}} \\
&= \sum_{i=1}^{d} \nabla_{X_i} \left\langle \nabla_{X_i} \sharp^{\mathbf{h}} df, \sharp^{\mathbf{h}} df\right\rangle_{\mathbf{h}} - \langle \nabla_{X_i} \sharp^{\mathbf{h}} df, \nabla_{X_i} \sharp^{\mathbf{h}} df\rangle_{\mathbf{h}} \\
&= \sum_{i=1}^{d} \frac{1}{2} \nabla_{X_i} \nabla_{X_i} \left\langle \sharp^{\mathbf{h}} df, \sharp^{\mathbf{h}} df\right\rangle_{\mathbf{h}} - \langle \nabla_{X_i} \sharp^{\mathbf{h}} df, \nabla_{X_i} \sharp^{\mathbf{h}} df\rangle_{\mathbf{h}} \\
&= \frac{1}{2} L^{\nabla} \|df\|^2_{\mathbf{h}^*} - \|\nabla df\|^2_{\mathbf{h}^0_2}.
\end{aligned}$$

For the second claim, doing almost the same computation we get that

$$\begin{aligned}
\left\langle L^{\nabla} f, \Delta_{\mathbf{g}} f\right\rangle_{L^2(\mathbf{g}^0)} &= -\int_M \langle dL^{\nabla} f, df\rangle_{\mathbf{g}^*} d\,\mathrm{vol}_{\mathbf{g}} \\
&= -\int_M \left\langle L^{\nabla} df + 2D^T df - \mathcal{A}(df), df\right\rangle_{\mathbf{g}^*} d\,\mathrm{vol}_{\mathbf{g}} \\
&= \int_M \left(\|\nabla_{\mathrm{pr}_{\mathcal{H}}} df\|^2_{\mathbf{g}^0_2} - 2\langle D^T df, df\rangle_{\mathbf{g}^*} + \langle \mathcal{A}(df), df\rangle_{\mathbf{g}^*}\right) d\,\mathrm{vol}_{\mathbf{g}}.
\end{aligned}$$

□

In the remaining part of this section the results presented will be original for this thesis, unless otherwise is stated. Define the operator $\mathcal{B} : \Gamma(T^*M) \to \Gamma(T^*M)$ to be the dual of the operator

$$\mathcal{B}^*(\alpha) = \flat^{\mathbf{g}} \mathrm{tr}_{\mathbf{h}} \left(\nabla_{\times} T^{\nabla}\right)\left(\times, \sharp^{\mathbf{h}} \alpha\right)$$

in the $L^2(\mathbf{g}^*)$. The sub-Riemannian manifold $(M, \mathcal{H}, \mathbf{h})$ with an affine connection ∇ is said to be of **Yang-Mills** type if $\mathcal{B}^* \equiv 0$.

Lemma 4.4. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian geometry with a taming metric $\mathbf{g}$. Furthermore, let ∇ be an affine connection which is compatible with the metric $\mathbf{g}$, preserves $\mathcal{H}$ and the rough sub-Laplacian is symmetric with respect to $L^2(\mathbf{g}^0)$. Then we have the following equations*

a) $\langle D^T df, df\rangle_{L^2(\mathbf{h}^*)} = -\langle df, \mathcal{B}(df)\rangle_{L^2(\mathbf{h}^*)} + \frac{1}{2}\|df(T^\nabla(\cdot,\cdot))\|^2_{L^2(\mathbf{h}^0_2)}$,

b) $\|\nabla^2 f\|^2_{\mathbf{h}^0_2} \geq \frac{1}{d}\|L^\nabla f\|^2_{L^2(\mathbf{g}^0)} + \frac{1}{4}\|df(T^\nabla(\cdot,\cdot))\|^2_{\mathbf{h}^0_2}$,

c) $\frac{d-1}{d}\|L^\nabla f\|^2_{L^2(\mathbf{g}^0)} \geq \langle(\mathcal{A}+2\mathcal{B})(df), df\rangle_{L^2(\mathbf{h}^*)} - \frac{3}{4}\|df(T^\nabla(\cdot,\cdot))\|^2_{L^2(\mathbf{h}^0_2)}$.

Proof. For the first claim we compute

$$\begin{aligned}\langle D^T(df), df\rangle_{L^2(\mathbf{h}^*)} &= \langle \mathrm{tr}_\mathbf{h}(\nabla_\times df)(T^\nabla(\times,\cdot)), df\rangle_{L^2(\mathbf{h}^*)} \\ &= \int_M \mathrm{tr}_\mathbf{h}\nabla_\times df\left(T^\nabla\left(\times, \sharp^\mathbf{h} df\right)\right) d\,\mathrm{vol}_\mathbf{g} \\ &= \int_M \left\langle \nabla_{\mathrm{pr}_\mathcal{H}\cdot} df, \flat^\mathbf{g}\left(T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^\mathbf{h} df\right)\right)\right\rangle_{\mathbf{g}^0_2} d\,\mathrm{vol}_\mathbf{g}.\end{aligned}$$

Using the adjoint of $\nabla_{\mathrm{pr}_\mathcal{H}\cdot}$ given in the equation 3.1 we have that

$$\begin{aligned}&\int_M \left\langle \nabla_{\mathrm{pr}_\mathcal{H}} df, \flat^\mathbf{g}\left(T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^\mathbf{h} df\right)\right)\right\rangle_{\mathbf{g}^0_2} d\,\mathrm{vol}_\mathbf{g} \\ &\quad= \int_M \left\langle df, -\sum_{i=1}^d \iota_{X_i}\left(\nabla_{X_i}\flat^\mathbf{g}\left(T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^\mathbf{h} df\right)\right)\right)\right\rangle_{\mathbf{g}^*} d\,\mathrm{vol}_\mathbf{g},\end{aligned}$$

where $X_1, \ldots, X_d, Z_1, \ldots, Z_{n-d}$ is an adapted orthonormal frame parallel at the point x. Then

$$\begin{aligned}-\sum_{i=1}^d \iota_\times\left(\nabla_\times\flat^\mathbf{g}\left(T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^\mathbf{h} df\right)\right)\right) &= -\sum_{i=1}^d \left(\nabla_{X_i}T^\nabla\right)\left(X_i, \sharp^\mathbf{h} df\right) + T^\nabla\left(\nabla_{X_i}X_i, \sharp^\mathbf{h} df\right) \\ &\quad + T^\nabla\left(X_i, \nabla_{X_i}\sharp^\mathbf{h} df\right) \\ &= -\sum_{i=1}^d \left(\nabla_{X_i}T^\nabla\right)\left(X_i, \sharp^\mathbf{h} df\right) + T^\nabla\left(X_i, \nabla_{X_i}\sharp^\mathbf{h} df\right).\end{aligned}$$

Observing that

$$-\sum_{i,j=1}^d (\nabla_{X_i} df)(X_j)\, df\left(T^\nabla(X_i, X_j)\right) = \frac{1}{2}\sum_{i,j=1}^d \left(df\left(T^\nabla(X_i, X_j)\right)\right)^2 = \frac{1}{2}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{\mathbf{h}^0_2},$$

we get

$$\begin{aligned}&\sum_{i=1}^d \left\langle df, -\flat^\mathbf{g}\left(\left(\nabla_{X_i}T^\nabla\right)\left(X_i, \sharp^\mathbf{h} df\right) + T^\nabla\left(X_i, \nabla_{X_i}\sharp^\mathbf{h} df\right)\right)\right\rangle_{\mathbf{g}^*} \\ &\quad= -\langle df, \mathcal{B}^*(df)\rangle_{\mathbf{g}^*} + \frac{1}{2}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{\mathbf{h}^0_2}.\end{aligned}$$

Hence

$$\left\langle D^{T}df,df\right\rangle_{L^{2}(\mathbf{h}^{*})}=-\langle df,\mathcal{B}\left(df\right)\rangle_{L^{2}(\mathbf{h}^{*})}+\frac{1}{2}\|df\left(T^{\nabla}\left(\cdot,\cdot\right)\right)\|^{2}_{L^{2}\left(\mathbf{h}^{0}_{2}\right)}.$$

For the second claim, define the symmetric operator

$$S\left(\nabla^{2}\right)(X,Y)=\frac{1}{2}\nabla^{2}_{X,Y}+\frac{1}{2}\nabla^{2}_{Y,X}$$

and the anti-symmetric operator

$$A\left(\nabla^{2}\right)(X,Y)=\frac{1}{2}\nabla^{2}_{X,Y}-\frac{1}{2}\nabla^{2}_{Y,X}.$$

Then

$$\nabla^{2}_{X,Y}=S\left(\nabla^{2}\right)(X,Y)+A\left(\nabla^{2}\right)(X,Y).$$

Also

$$\begin{aligned}\left\langle S\left(\nabla^{2}\right)f,A\left(\nabla^{2}\right)f\right\rangle_{\mathbf{h}^{0}_{2}}&=\frac{1}{4}\sum_{i,j=1}^{d}\left(\nabla^{2}_{X_i,X_j}f+\nabla^{2}_{X_j,X_i}f\right)\left(\nabla^{2}_{X_i,X_j}f-\nabla^{2}_{X_j,X_i}f\right)\\&=\frac{1}{4}\sum_{i,j=1}^{d}\left(\nabla^{2}_{X_i,X_j}f\right)^{2}-\left(\nabla^{2}_{X_j,X_i}f\right)^{2}=0,\end{aligned}$$

hence $\|\nabla^{2}f\|^{2}_{\mathbf{h}^{0}_{2}}=\|S\left(\nabla^{2}\right)f\|^{2}_{\mathbf{h}^{0}_{2}}+\|A\left(\nabla^{2}\right)f\|^{2}_{\mathbf{h}^{0}_{2}}$. Using equation 2.1 implies that

$$\begin{aligned}\|A\left(\nabla^{2}\right)f\|^{2}_{\mathbf{h}^{0}_{2}}&=\sum_{i,j=1}^{d}\left(\frac{1}{2}\nabla^{2}_{X_i,X_j}f-\frac{1}{2}\nabla^{2}_{X_j,X_i}f\right)^{2}\\&=\frac{1}{4}\sum_{i,j=1}^{d}\left(df\left(T^{\nabla}\left(X_i,X_j\right)\right)\right)^{2}=\frac{1}{4}\|df\left(T^{\nabla}\left(\cdot,\cdot\right)\right)\|^{2}_{\mathbf{h}^{0}_{2}}.\end{aligned}$$

By using the same argument as in the proof of Lemma 2.20 we have that

$$\|S\left(\nabla^{2}\right)f\|^{2}_{\mathbf{h}^{0}_{2}}\geq\frac{1}{d}\left(L^{\nabla}f\right)^{2}.$$

Hence we have proved the second claim.

For the final claim, using Corollary 4.3 together with a) and b) implies

$$\begin{aligned}\|L^{\nabla}f\|^{2}_{L^{2}(\mathbf{g}^{0})}&=\int_{M}\left(\|\nabla df\|^{2}_{\mathbf{h}^{0}_{2}}-2\langle D^{T}df,df\rangle_{\mathbf{h}^{*}}+\langle\mathcal{A}\left(df\right),df\rangle_{\mathbf{h}^{*}}\right)d\operatorname{vol}_{\mathbf{g}}\\&=\int_{M}\left(\|\nabla df\|^{2}_{\mathbf{h}^{0}_{2}}+2\left\langle\mathcal{B}df,df\right\rangle_{\mathbf{h}^{*}}-\|df\left(T^{\nabla}\left(\cdot,\cdot\right)\right)\|^{2}_{\mathbf{h}^{0}_{2}}+\langle\mathcal{A}\left(df\right),df\rangle_{\mathbf{h}^{*}}\right)d\operatorname{vol}_{\mathbf{g}}\\&\geq\frac{1}{d}\|L^{\nabla}f\|^{2}_{L^{2}(\mathbf{g}^{0})}-\frac{3}{4}\|df\left(T^{\nabla}\left(\cdot,\cdot\right)\right)\|^{2}_{L^{2}\left(\mathbf{h}^{0}_{2}\right)}+2\left\langle\mathcal{B}df,df\right\rangle_{L^{2}(\mathbf{h}^{*})}+\langle\mathcal{A}\left(df\right),df\rangle_{L^{2}(\mathbf{h}^{*})}.\end{aligned}$$

Moving $\frac{1}{d}\|L^{\nabla}f\|^{2}_{L^{2}(\mathbf{g}^{0})}$ to the left hand side we obtain equation c). □

4.2 Lichnerowicz Estimate

We will begin by looking at the setting given in [BK16] and improved in [BKW16], and what is already known. In Section 4.2.2 we will develop another setting in which we give another Lichnerowicz estimate.

4.2.1 Lichnerowicz Estimate by F. Baudoin, B. Kim and J. Wang

Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold. Assume that there exists a taming metric $\mathbf{g}$ such that the orthogonal compliment of $\mathcal{H}$, denoted $\mathcal{V}$, becomes integrable. Additionally, assume that $\mathbf{g}$ satisfy the condition

$$(\mathcal{L}_X \mathbf{g})(V, V) = (\mathcal{L}_V \mathbf{g})(X, X) = 0,$$

for $X \in \mathcal{H}$ and $V \in \mathcal{V} = (\mathcal{H})^{\perp}$. Define the connection ∇ to be

$$\nabla_X Y = \mathrm{pr}_{\mathcal{H}} \nabla^{\mathbf{g}}_{\mathrm{pr}_{\mathcal{H}} X} \mathrm{pr}_{\mathcal{H}} Y + \mathrm{pr}_{\mathcal{H}} [\mathrm{pr}_{\mathcal{V}} X, \mathrm{pr}_{\mathcal{H}} Y] + \mathrm{pr}_{\mathcal{V}} [\mathrm{pr}_{\mathcal{H}} X, \mathrm{pr}_{\mathcal{V}} Y] + \mathrm{pr}_{\mathcal{V}} \nabla^{\mathbf{g}}_{\mathrm{pr}_{\mathcal{V}} X} \mathrm{pr}_{\mathcal{V}} Y, \qquad (4.4)$$

where $\nabla^{\mathbf{g}}$ denotes the Levi-Civita connection with respect to $\mathbf{g}$. This connection is called the **Bott connection**. Then Lemma 4.8 and Lemma 4.10 shows that the Bott connection satisfy the following conditions:

- ∇ preserves $\mathcal{H}$,
- is compatible with $\mathbf{g}$,
- the rough sub-Laplacian is symmetric with respect to $L^2\left(\mathbf{g}^0\right)$,
- $T^{\nabla}(X, Y) \in \Gamma(\mathcal{V})$ for all $X, Y \in \Gamma(TM)$,
- the Ricci curvature is symmetric and $\langle \mathrm{Ric}\,(\alpha), \beta \rangle_{\mathbf{g}}$ depends only on the horizontal part of α and β.

In [BK16] and [BKW16] they showed the following bound on the first eigenvalue.

Theorem 4.5. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold with a taming metric $\mathbf{g}$ given as above. Assume that the Bott connection ∇ satisfy the Yang-Mills condition, and that we have the following bounds*

- $\langle \mathrm{Ric}\, \alpha, \alpha \rangle_{\mathbf{h}^*} \geq \rho_1 \|\alpha\|^2_{\mathbf{h}^*}$,
- $\|\alpha\left(T^{\nabla}(\cdot, \cdot)\right)\|^2_{\mathbf{h}^0_2} \geq \rho_2 \|\alpha\|^2_{\mathbf{g}_{\mathcal{V}}}$ *and*
- $\|\mathcal{R}\left(\cdot, \sharp^{\mathbf{h}} \alpha\right)\|^2_{\mathbf{g}_{\mathcal{V}}} \leq \kappa_1 \|\alpha\|^2_{\mathbf{h}^*}$,

where $\kappa_1 \geq 0$ and $\rho_i > 0$. Then the first eigenvalue λ_1 of the sub-Laplacian L^{∇} satisfies

$$\lambda_1 \leq \frac{-\rho_1}{1 - \frac{1}{d} + \frac{3\kappa_1}{\rho_2}}.$$

4.2.2 New Lichnerowicz Estimate

It is known that the bound presented in the previous theorem is sharp, see [BK16]. The method used to show the bound does however rely on some choice of parameters. When trying to generalize the method it gets difficult to find the optimal parameters, hence we will use a different method. With the method presented we are able to lower the conditions required, and we regain the bound of Theorem 4.5 as a special case.

Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold. Assume that there exists an orthogonal compliment $\mathcal{V}$, i.e. a choice of sub-bundle such that $TM = \mathcal{H} \oplus \mathcal{V}$, such that $\mathbf{h}$ satisfy

$$\left(\mathcal{L}_V \mathrm{pr}_{\mathcal{H}}^* \mathbf{h}\right)(X, X) = 0, \tag{4.5}$$

for $V \in \Gamma(\mathcal{V})$ and $X \in \Gamma(\mathcal{H})$. In this case we will call $\mathcal{V}$ for **metric preserving**, and we say that $(M, \mathcal{H}, \mathbf{h})$ has a metric preserving compliment if such a $\mathcal{V}$ exists.

Define the **curvature** $\mathcal{R}$ and **cocurvature** $\bar{\mathcal{R}}$ of $(M, \mathcal{H}, \mathbf{h})$ with respect to $\mathcal{V}$ by

$$\mathcal{R}(X, Y) = \mathrm{pr}_{\mathcal{V}} \left[\mathrm{pr}_{\mathcal{H}} X, \mathrm{pr}_{\mathcal{H}} Y\right]$$

and

$$\bar{\mathcal{R}}(X, Y) = \mathrm{pr}_{\mathcal{H}} \left[\mathrm{pr}_{\mathcal{V}} X, \mathrm{pr}_{\mathcal{V}} Y\right],$$

respectively. Note that $\mathcal{V}$ is integrable if and only if $\bar{\mathcal{R}} \equiv 0$. We will additionally assume that

$$\mathrm{tr} \bar{\mathcal{R}}(X, \mathcal{R}(X, \times)) \times = 0, \tag{4.6}$$

which is a condition which has appeared in [GT16a], [GT16b] and [GT16c]. Note that if $\mathcal{V}$ is integrable, then the condition is always satisfied.

Let $\mathbf{g}$ be a taming metric for $\mathbf{h}$ making $\mathcal{H}$ and $\mathcal{V}$ orthogonal, and let $\mathbf{g}_{\mathcal{V}}$ be $\mathbf{g}$ restricted to $\mathcal{V}$. Assume that

$$\mathrm{tr}_{\mathbf{g}_{\mathcal{V}}} \left(\mathcal{L}_X \mathbf{g}\right)(\times, \times) = 0, \tag{4.7}$$

whenever $X \in \Gamma(\mathcal{H})$. Let ∇' be any linear connection compatible with $\mathbf{g}_{\mathcal{V}}$, and denote by $\nabla^{\mathbf{g}}$ the Levi-Civita connection with respect to $\mathbf{g}$. Define the affine connection ∇ by

$$\nabla_X Y = \nabla^{\mathbf{g}}_{\mathrm{pr}_{\mathcal{H}} X} \mathrm{pr}_{\mathcal{H}} Y + \mathrm{pr}_{\mathcal{H}} \left[\mathrm{pr}_{\mathcal{V}} X, \mathrm{pr}_{\mathcal{H}} Y\right] + \nabla'_X \mathrm{pr}_{\mathcal{V}} Y. \tag{4.8}$$

Let L^{∇} be the sub-Laplacian with respect to $\mathbf{h}$. Then $L^{\nabla} f$ depend only on $\mathrm{pr}_{\mathcal{H}} \nabla^{\mathbf{g}}_{\cdot} \cdot$ and the choice of compliment $\mathcal{V}$, since if $X_1, \dots, X_d, Z_1, \dots Z_{n-d}$ is an adapted orthonormal frame then

$$L^{\nabla} f = \sum_{i=1}^{d} X_i \left(X_i (f)\right) - \left(\mathrm{pr}_{\mathcal{H}} \nabla^{\mathbf{g}}_{X_i} X_i\right)(f).$$

Define $\mathrm{Ric}_{\mathcal{V}}$ to be

$$\begin{aligned} \mathrm{Ric}_{\mathcal{V}}(\alpha) = \alpha\left(\mathrm{tr}_{\mathbf{h}} \left(\nabla_{\times} T^{\nabla}\right)(\times, \sharp^{\mathbf{g}_{\mathcal{V}}} \cdot)\right) + \alpha\left(\mathrm{tr}_{\mathbf{h}} T^{\nabla}\left(\times, T^{\nabla}(\times, \sharp^{\mathbf{g}_{\mathcal{V}}} \cdot)\right)\right) \\ + \left\langle T^{\nabla}\left(\mathrm{pr}_{\mathcal{H}}, \sharp^{\mathbf{g}_{\mathcal{V}}} \alpha\right), T^{\nabla}\left(\mathrm{pr}_{\mathcal{H}}, \sharp^{\mathbf{g}_{\mathcal{V}}} \cdot\right)\right\rangle_{\mathbf{h} \otimes \mathbf{g}}. \end{aligned}$$

Under these assumptions we aim to show the following result.

Theorem 4.6. *Assume that $(M, \mathcal{H}, \mathbf{h})$ is a compact sub-Riemannian manifold with a taming metric $\mathbf{g}$ such that the orthogonal compliment $\mathcal{V}$ are metric preserving, and the assumptions 4.6 and 4.7 are satisfied. Let ∇ be any connection on the form of equation 4.8 satisfying*

- $\langle \mathrm{Ric}\,(\alpha), \alpha \rangle_{\mathbf{h}^*} \geq \rho_1 \|\alpha\|^2_{\mathbf{h}^*}$,
- $\|\alpha\left(T^{\nabla}(\cdot, \cdot)\right)\|^2_{\mathbf{h}^0_2} \geq \rho_2 \|\alpha\|^2_{\mathbf{g}^*_{\mathcal{V}}}$,

- $\|T^{\nabla}(\cdot, \sharp^{\mathbf{h}}\alpha)\|^2_{\mathbf{h}^*\otimes\mathbf{g}_{\mathcal{V}}} \leq \kappa_1 \|\alpha\|^2_{\mathbf{h}^*}$,
- $|\langle \mathcal{B}(\alpha), \alpha\rangle_{\mathbf{h}^*} + \langle T^{\nabla}(pr_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_{\mathcal{V}}}\alpha), T^{\nabla}(pr_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}}\alpha)\rangle_{\mathbf{g}_2^0}| \leq 2\kappa_2 \|\alpha\|_{\mathbf{g}^*_{\mathcal{V}}}\|\alpha\|_{\mathbf{h}^*}$,
- $|\langle \mathrm{Ric}_{\mathcal{V}}(\alpha), \alpha\rangle_{\mathbf{g}^*}| \leq \kappa_3 \|\alpha\|^2_{\mathbf{g}^*_{\mathcal{V}}}$,
- $|\langle T^{\nabla}(pr_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_{\mathcal{V}}}\alpha), T^{\nabla}(pr_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}}\alpha)\rangle_{\mathbf{g}_2^0} + \alpha\left(\mathrm{tr}_{\mathbf{h}} T^{\nabla}\left(\times, T^{\nabla}\left(\times, \sharp^{\mathbf{h}}\alpha\right)\right)\right)| \leq \kappa_4 \|\alpha\|_{\mathbf{h}^*}\|\alpha\|_{\mathbf{g}^*_{\mathcal{V}}}$,

where $\kappa_i \geq 0$ and $\rho_2 > 0$. Define the constants

$$K = \rho_1\rho_2 - 4\kappa_2^2 - 3\kappa_1\kappa_3 - 4\kappa_2\kappa_4 - 2(4\kappa_2 + \kappa_4)\sqrt{\kappa_1\kappa_3}$$

and $k = (4\kappa_2 + \kappa_4)\sqrt{\kappa_1}$. Then if $K > 0$ we have that

$$-\lambda_1 \geq \left(\sqrt{\frac{K}{\rho_2\left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)} + \left(\sqrt{\kappa_3} + \frac{k}{\rho_2\left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)}\right)^2} - \frac{k}{\rho_2\left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)}\right)^2 - \kappa_3,$$

where λ_1 is the first eigenvalue of L^{∇}.

Requiring that $\rho_2 > 0$ implies that $\mathcal{H}$ is of step 2, since, by the next lemma, the torsion satisfies $T^{\nabla}(\mathrm{pr}_{\mathcal{H}}\cdot, \mathrm{pr}_{\mathcal{H}}\cdot) = -\mathcal{R}(\cdot,\cdot)$.

The proof of the above theorem will be divided into several intermediate steps. We will begin by proving some general properties of connection used in this setting.

Lemma 4.7. *Let $(M, \mathcal{H}, \mathbf{h})$ be a sub-Riemannian manifold and let $\mathbf{g}$ be a taming metric such that the vertical bundle $\mathcal{V}$ is metric preserving. Then any affine connection ∇ on the form 4.8 satisfies property 1-3 in the beginning of Section 4. Furthermore, we have that the torsion satisfies $T^{\nabla}(X, Y) \in \Gamma(\mathcal{V})$ whenever either $X \in \Gamma(\mathcal{H})$ and*

$$T^{\nabla}(pr_{\mathcal{H}}\cdot, pr_{\mathcal{H}}\cdot) = -\mathcal{R}(\cdot,\cdot).$$

Proof. The fact that ∇ preserves $\mathcal{H}$ follows from the definition of ∇. When X and Y are in $\Gamma(\mathcal{H})$ the compatibility follows from the compatibility of $\nabla^{\mathbf{g}}$ with $\mathbf{h}$. If $Y \in \Gamma(\mathcal{V})$ and $X \in \Gamma(TM)$, then

$$(\nabla_X \mathbf{g})(Y, Y) = (\nabla'_X \mathbf{g}_{\mathcal{V}})(Y, Y) = 0.$$

For the final case when $X \in \Gamma(\mathcal{V})$ and $Y \in \Gamma(\mathcal{H})$ we have that

$$(\nabla_X \mathbf{g})(Y, Y) = (\mathcal{L}_X \mathbf{g})(Y, Y) = 0,$$

by assumption 4.5.

Let $X \in \Gamma(\mathcal{H})$. If $Y \in \Gamma(\mathcal{H})$ we have that

$$T^{\nabla}(X, Y) = \mathrm{pr}_{\mathcal{H}}\nabla^{\mathbf{g}}_X Y - \mathrm{pr}_{\mathcal{H}}\nabla^{\mathbf{g}}_Y X - [X, Y] = -\mathcal{R}(X, Y) \in \Gamma(\mathcal{V}).$$

On the other hand if $Y \in \Gamma(\mathcal{V})$ then we have that

$$T^{\nabla}(X, Y) = \nabla'_X Y - \mathrm{pr}_{\mathcal{H}}[Y, X] - [X, Y] = \nabla'_X Y - \mathrm{pr}_{\mathcal{V}}[X, Y] \in \Gamma(\mathcal{V}).$$

Hence $T^{\nabla}(X, Y) \in \Gamma(\mathcal{V})$.

To see that L^{∇} is symmetric we need to show that $\mathrm{tr} T^{\nabla}(X, \times) \times = 0$, whenever $X \in \mathcal{H}$ by Corollary 3.16. Let $X_1, \ldots, X_d, Z_1, \ldots, Z_{n-d}$ be an adapted orthonormal frame. Then

$$\begin{aligned}
\sum_{i=1}^{d} \langle T^{\nabla}(X, X_i), X_i \rangle_{\mathbf{g}} + \sum_{j=1}^{n-d} \langle T^{\nabla}(X, Z_j), Z_j \rangle_{\mathbf{g}} &= \sum_{j=1}^{n-d} \langle T^{\nabla}(X, Z_j), Z_j \rangle_{\mathbf{g}} \\
&= \sum_{j=1}^{n-d} \langle Z_j, \nabla_X Z_j - \nabla_{Z_j} X - [X, Z_j] \rangle_{\mathbf{g}} \\
&= -\langle Z_j, [X, Z_j] \rangle_{\mathbf{g}} = \frac{-1}{2} \mathrm{tr}_{\mathbf{g}_{\mathcal{V}}} (\mathcal{L}_X \mathbf{g})(\times, \times) = 0,
\end{aligned}$$

where the last equality follows from assumption 4.7. □

Lemma 4.8. *Assume that $(M, \mathcal{H}, \mathbf{h})$ is a sub-Riemannian manifold with a taming metric $\mathbf{g}$ which is metric preserving, and satisfy condition 4.6. Let ∇ be on the form 4.8 and let* Ric *be defined as in equation 4.1 with respect to ∇. Then*

$$\langle \mathrm{Ric}\,(\alpha), \beta \rangle_{\mathbf{g}} = \langle \mathrm{Ric}\,(pr_{\mathcal{H}^*} \alpha), \beta \rangle_{\mathbf{h}^*},$$

where α, β are 1-forms. Additionally, we have that Ric *is symmetric, i.e.*

$$\langle \mathrm{Ric}\,(\alpha), \beta \rangle_{\mathbf{g}} = \langle \mathrm{Ric}\,(\beta), \alpha \rangle_{\mathbf{g}}.$$

Proof. Note that

$$\mathrm{tr} T^{\nabla}\left(\mathrm{pr}_{\mathcal{V}} X, T^{\nabla}(\mathrm{pr}_{\mathcal{H}} X, \mathrm{pr}_{\mathcal{H}} \times)\right) \times = \mathrm{tr} \bar{\mathcal{R}}(X, \mathcal{R}(X, \times)) \times = 0,$$

by assumption 4.6. Since ∇ is compatible with $\mathbf{g}$, we have by using Lemma 2.7, that

$$\langle R^{\nabla}(X, Y) Z, Z \rangle_{\mathbf{g}} = 0,$$

for all $X, Y, Z \in \Gamma(\mathcal{H})$. Let $X, Y \in \Gamma(\mathcal{H})$ and $Z \in \Gamma(\mathcal{V})$. Using the first Bianchi identity (Lemma 2.8) for affine connections with torsion when $X, Y \in \Gamma(\mathcal{H})$ and $Z \in \Gamma(\mathcal{V})$, we have that

$$\begin{aligned}
\mathbf{g}\left(R^{\nabla}(X, Z) Y, X\right) + \mathbf{g}\left(R^{\nabla}(Y, X) Z, X\right) + \mathbf{g}\left(R^{\nabla}(Z, Y) X, X\right) &= \mathbf{g}\left(R^{\nabla}(X, Z) Y, X\right) \\
= \mathbf{g}\left(-T^{\nabla}\left(X, T^{\nabla}(Z, Y)\right) - T^{\nabla}\left(Y, T^{\nabla}(X, Z)\right) - T^{\nabla}\left(Z, T^{\nabla}(Y, X)\right), X\right) \\
+ \mathbf{g}\left(\left(\nabla_X T^{\nabla}\right)(Z, Y) + \left(\nabla_Y T^{\nabla}\right)(X, Z) + \left(\nabla_Z T^{\nabla}\right)(Y, X), X\right) \\
= -\mathbf{g}\left(T^{\nabla}\left(Z, T^{\nabla}(Y, X)\right), X\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\langle \mathrm{Ric}\,(\alpha), \beta \rangle_{\mathbf{g}^*} &= \mathrm{tr}_{\mathbf{h}} \left\langle R^{\nabla}(\times, \sharp^{\mathbf{g}} \beta) \sharp^{\mathbf{h}} \alpha, \times \right\rangle_{\mathbf{h}} \\
&= \mathrm{tr}_{\mathbf{h}} \left\langle R^{\nabla}\left(\times, \sharp^{\mathbf{h}} \beta\right) \sharp^{\mathbf{h}} \alpha, \times \right\rangle_{\mathbf{h}} - \mathrm{tr}_{\mathbf{h}} \langle R^{\nabla}(\times, \sharp^{\mathbf{g}_{\mathcal{V}}} \beta) \sharp^{\mathbf{h}} \alpha, \times \rangle_{\mathbf{h}} \\
&= \mathrm{tr}_{\mathbf{h}} \left\langle R^{\nabla}\left(\times, \sharp^{\mathbf{h}} \beta\right) \sharp^{\mathbf{h}} \alpha, \times \right\rangle_{\mathbf{h}} \\
&= \langle \mathrm{Ric}\, \mathrm{pr}_{\mathcal{H}^*} \alpha, \beta \rangle_{\mathbf{h}^*}.
\end{aligned}$$

For symmetry it is enough to show that $\langle R^{\nabla}(X,Y)Z,X\rangle_{\mathbf{h}} = \langle R^{\nabla}(X,Z)Y,X\rangle_{\mathbf{h}}$ when $X,Y,Z \in \Gamma(\mathcal{H})$, since the Ricci curvature only depends on the horizontal part. By using the first Bianchi identity we have that

$$\langle R^{\nabla}(X,Y)Z,X\rangle_{\mathbf{h}} - \langle R^{\nabla}(X,Z)Y,X\rangle_{\mathbf{h}} = 0,$$

since the torsion is vertical. □

In the next Lemma we will find a constant such that C such that $C\|df\|_{\mathbf{h}^*} \geq \|df\|_{\mathbf{g}^*_{\mathcal{V}}}$. In the proof we will use the connection $\tilde{\nabla}$ defined by $\tilde{\nabla}_X\alpha = \nabla_X\alpha - \flat^{\mathbf{g}}T^{\nabla}(X, \sharp^{\mathbf{g}_{\mathcal{V}}}\alpha)$, and the map $\tilde{D^T}$ defined by

$$\tilde{D^T}(\alpha) = \mathrm{tr}_{\mathbf{h}}\left(\tilde{\nabla}_{\times}\alpha\right)\left(T^{\nabla}(\times,\cdot)\right).$$

Lemma 4.9. *Let $(M,\mathcal{H},\mathbf{h})$ be a compact sub-Riemannian manifold which is metric preserving and satisfies assumption 4.6 and 4.7. Let ∇ be some connection on the form 4.8 and assume that the following inequalities are satisfied*

- $\|\alpha\left(T^{\nabla}(\cdot,\cdot)\right)\|^2_{\mathbf{h}^0_2} \geq \rho_2\|\alpha\|^2_{\mathbf{g}^*_{\mathcal{V}}}$,
- $\|T^{\nabla}\left(\cdot,\sharp^{\mathbf{h}}\alpha\right)\|^2_{\mathbf{h}^*\otimes\mathbf{g}_{\mathcal{V}}} \leq \kappa_1\|\alpha\|^2_{\mathbf{h}^*}$,
- $|\langle\mathcal{B}(\alpha),\alpha\rangle_{\mathbf{h}^*} + \langle T^{\nabla}(pr_{\mathcal{H}}\cdot,\sharp^{\mathbf{g}_{\mathcal{V}}}\alpha), T^{\nabla}\left(pr_{\mathcal{H}}\cdot,\sharp^{\mathbf{h}}\alpha\right)\rangle_{\mathbf{g}^0_2}| \leq 2\kappa_2\|\alpha\|_{\mathbf{g}^*_{\mathcal{V}}}\|\alpha\|_{\mathbf{h}^*}$,
- $|\langle\mathrm{Ric}_{\mathcal{V}}(\alpha),\alpha\rangle_{\mathbf{g}^*}| \leq \kappa_3\|\alpha\|^2_{\mathbf{g}^*_{\mathcal{V}}}$,

where $\kappa_i \geq 0$ and $\rho_2 > 0$. Then

$$\frac{2}{\rho_2}\left(\sqrt{\kappa_3-\lambda_1}\sqrt{\kappa_1}+2\kappa_2\right)\|df\|_{L^2(\mathbf{h}^*)} \geq \|df\|_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)}.$$

$$\left(\tilde{D^T}df\right)\left(\sharp^{\mathbf{h}}df\right) \leq \frac{2}{\rho_2}\left(\kappa_1\kappa_3 - \lambda_1\kappa_1 + 2\kappa_2\sqrt{\kappa_1}\sqrt{\kappa_3-\lambda_1}\right)\|df\|^2_{L^2(\mathbf{h}^*)}$$

and

$$\|df\left(T^{\nabla}(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}^0_2\right)} \leq \frac{4}{\rho_2}\left(\kappa_1\kappa_3 + 4\kappa_2^2 - \lambda_1\kappa_1 + 4\kappa_2\sqrt{\kappa_3-\lambda_1}\right)\|df\|^2_{L^2(\mathbf{h}^*)},$$

where $L^{\nabla}f = \lambda_1 f$,

Proof. By combining the sub-Riemannian Bochner formula (Proposition 4.2) together with the fact that

$$\frac{1}{2}L\|df\|^2_{\mathbf{g}^*_{\mathcal{V}}} = \langle Ldf, df\rangle_{\mathbf{g}^*_{\mathcal{V}}} + \|\nabla_{\mathrm{pr}_{\mathcal{H}}\cdot}df\|^2_{\mathbf{g}^*\otimes\mathbf{g}^*_{\mathcal{V}}},$$

we get that

$$\|\nabla_{\mathrm{pr}_{\mathcal{H}}\cdot}df\|^2_{L^2\left(\mathbf{g}^*\otimes\mathbf{g}^*_{\mathcal{V}}\right)} = -\langle dLf, df\rangle_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)} + 2\langle D^T df, df\rangle_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)} - \langle\mathcal{A}df, df\rangle_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)}.$$

The equality above implies that

$$\begin{aligned}\|\tilde{\nabla}_{\mathrm{pr}_{\mathcal{H}}\cdot}df\|^2_{L^2\left(\mathbf{g}^*\otimes\mathbf{g}^*_{\mathcal{V}}\right)} &= -\lambda_1\|df\|^2_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)} + \langle\mathrm{Ric}_{\mathcal{V}}df, df\rangle_{L^2(\mathbf{g}^*)}\\ &\leq (\kappa_3-\lambda_1)\|df\|^2_{L^2\left(\mathbf{g}^*_{\mathcal{V}}\right)}.\end{aligned}$$

By the definition of $\tilde{D^T}$ we have that

$$\tilde{D^T}(\alpha)(Y) = \left(D^T(\alpha)\right)(Y) - \langle T^\nabla(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_\mathcal{V}}\alpha), T^\nabla(\mathrm{pr}_{\mathcal{H}}\cdot, Y)\rangle_{\mathbf{g}_1^1}.$$

Using Cauchy-Schwarz we get

$$\begin{aligned}\left|\int_M \tilde{D^T}(df)\left(\sharp^{\mathbf{h}} df\right) d\,\mathrm{vol}_{\mathbf{g}}\right| &= \left|\left\langle \tilde{\nabla}_{\mathrm{pr}_{\mathcal{H}}\cdot} df, \flat^{\mathbf{g}} T^\nabla\left(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}} df\right)\right\rangle_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}\right| \\ &\le \sqrt{\kappa_3 - \lambda_1}\sqrt{\kappa_1}\|df\|_{L^2(\mathbf{h}^*)}\|df\|_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}.\end{aligned}$$

On the other hand we have that

$$\begin{aligned}\int_M \tilde{D^T}(df)\left(\sharp^{\mathbf{h}} df\right) d\,\mathrm{vol}_{\mathbf{g}} &= \langle -\mathcal{B}(df), df\rangle_{L^2(\mathbf{h}^*)} + \frac{1}{2}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}_2^0\right)} \\ &\quad - \langle T^\nabla\left(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}} df\right), T^\nabla\left(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_\mathcal{V}} df\right)\rangle_{L^2\left(\mathbf{g}_2^0\right)} \\ &\ge -2\kappa_2\|df\|_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}\|df\|_{L^2(\mathbf{h}^*)} + \frac{\rho_2}{2}\|df\|^2_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}.\end{aligned}$$

Setting the two inequalities together we obtain

$$\frac{2}{\rho_2}\left(\sqrt{\kappa_3 - \lambda_1}\sqrt{\kappa_1} + 2\kappa_2\right)\|df\|_{L^2(\mathbf{h}^*)} \ge \|df\|_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}.$$

For the bound on $\tilde{D}^T(df)\left(\sharp^{\mathbf{h}} df\right)$, we use that

$$\begin{aligned}\int_M \tilde{D^T}(df)\left(\sharp^{\mathbf{h}} df\right) d\,\mathrm{vol}_{\mathbf{g}} &\le \sqrt{\kappa_3 - \lambda_1}\sqrt{\kappa_1}\|df\|_{L^2\left(\mathbf{g}_\mathcal{V}^*\right)}\|df\|_{L^2(\mathbf{h}^*)} \\ &\le \frac{2}{\rho_2}\left(\kappa_1\kappa_3 - \lambda_1\kappa_1 + 2\kappa_2\sqrt{\kappa_1}\sqrt{\kappa_3 - \lambda_1}\right)\|df\|^2_{L^2(\mathbf{h}^*)}.\end{aligned}$$

For the last inequality we have that

$$\begin{aligned}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}_2^0\right)} &= \int_M 2D^T(df)\left(\sharp^{\mathbf{h}} df\right) + 2\langle\mathcal{B}(df), df\rangle_{\mathbf{h}^*} d\,\mathrm{vol}_{\mathbf{g}} \\ &= \int_M 2\tilde{D^T}(df)\left(\sharp^{\mathbf{h}} df\right) + 2\langle\mathcal{B}(df), df\rangle_{\mathbf{h}^*} d\,\mathrm{vol}_{\mathbf{g}} \\ &\quad + 2\langle T^\nabla(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_\mathcal{V}}\alpha), T^\nabla\left(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}}\alpha\right)\rangle_{L^2\left(\mathbf{g}_2^0\right)} \\ &\le \frac{4}{\rho_2}\left(\kappa_1\kappa_3 + 4\kappa_2^2 - \lambda_1\kappa_1 + 4\kappa_2\sqrt{\kappa_1}\sqrt{\kappa_3 - \lambda_1}\right)\|df\|^2_{L^2(\mathbf{h}^*)}.\end{aligned}$$

□

Proof of Theorem 4.6. We have that

$$\begin{aligned}\frac{d-1}{d}\|Lf\|^2_{L^2(\mathbf{g}^0)} &\ge \langle(\mathcal{A} + 2\mathcal{B})\, df, df\rangle_{L^2(\mathbf{h}^*)} - \frac{3}{4}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}_2^0\right)} \\ &= \langle \mathrm{Ric}\, df, df\rangle_{L^2(\mathbf{h}^*)} - \int_M df\left(\mathrm{tr}_{\mathbf{h}} T^\nabla\left(\times, T^\nabla\left(\times, \sharp^{\mathbf{h}} df\right)\right)\right) d\,\mathrm{vol}_{\mathbf{g}} + \langle\mathcal{B} df, df\rangle_{L^2(\mathbf{h}^*)} \\ &\quad - \frac{3}{4}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}_2^0\right)} \\ &= -\int_M \langle T^\nabla(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{g}_\mathcal{V}}\alpha), T^\nabla\left(\mathrm{pr}_{\mathcal{H}}\cdot, \sharp^{\mathbf{h}}\alpha\right)\rangle_{\mathbf{g}_2^0} + df\left(\mathrm{tr}_{\mathbf{h}} T^\nabla\left(\times, T^\nabla\left(\times, \sharp^{\mathbf{h}} df\right)\right)\right) d\,\mathrm{vol}_{\mathbf{g}} \\ &\quad + \langle \mathrm{Ric}\, df, df\rangle_{L^2(\mathbf{h}^*)} - \langle\tilde{D^T} df, df\rangle_{L^2(\mathbf{h}^*)} - \frac{1}{4}\|df\left(T^\nabla(\cdot,\cdot)\right)\|^2_{L^2\left(\mathbf{h}_2^0\right)}.\end{aligned}$$

Using the bounds above together with Lemma 4.9 we get

$$
\begin{aligned}
-\lambda_1\frac{d-1}{d} &\geq \rho_1 - \frac{1}{\rho_2}\left(3\kappa_1\kappa_3 + 4\kappa_2^2 + 4\kappa_2\kappa_4 + (8\kappa_2+2\kappa_4)\sqrt{\kappa_1}\sqrt{\kappa_3-\lambda_1} - 3\kappa_1\lambda_1\right)\\
-\lambda_1\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right) &\geq \rho_1 - \frac{1}{\rho_2}\left(3\kappa_1\kappa_3 + 4\kappa_2^2 + 4\kappa_2\kappa_4 + (8\kappa_2+2\kappa_4)\sqrt{\kappa_1}\sqrt{\kappa_3-\lambda_1}\right).
\end{aligned}
$$

If we complete the square

$$
\left(\sqrt{\kappa_3-\lambda_1} + \frac{(4\kappa_2+\kappa_4)\sqrt{\kappa_1}}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\right)^2 \geq \frac{\rho_1\rho_2 - 3\kappa_1\kappa_3 - 4\kappa_2\kappa_4 - 4\kappa_2^2}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)} + \kappa_3 + \left(\frac{(4\kappa_2+\kappa_4)\sqrt{\kappa_1}}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\right)^2.
$$

This implies that

$$
\begin{aligned}
\sqrt{\kappa_3-\lambda_1} &\geq \sqrt{\frac{\rho_1\rho_2 - 3\kappa_1\kappa_3 - 4\kappa_2\kappa_4 - 4\kappa_2^2}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)} + \kappa_3 + \left(\frac{(4\kappa_2+\kappa_4)\sqrt{\kappa_1}}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\right)^2} - \frac{(4\kappa_2+\kappa_4)\sqrt{\kappa_1}}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\\
-\lambda_1 &\geq \left(\sqrt{\frac{K}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)} + \left(\sqrt{\kappa_3} + \frac{k}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\right)^2} - \frac{k}{\rho_2\left(\frac{d-1}{d}+\frac{3\kappa_1}{\rho_2}\right)}\right)^2 - \kappa_3.
\end{aligned}
$$

□

4.2.3 Lichnerowicz Estimate and the Bott Connection

Assume that the metric $\mathbf{g}$ have the following properties

$$
(\mathcal{L}_X\mathbf{g})(Z,Z) = (\mathcal{L}_Z\mathbf{g})(X,X) = 0, \tag{4.9}
$$

for $X \in \Gamma(\mathcal{H})$ and $Z \in \Gamma(\mathcal{V})$. The Bott connection ∇ which is defined by

$$
\nabla_X Y = \mathrm{pr}_{\mathcal{H}}\nabla^{\mathbf{g}}_{\mathrm{pr}_{\mathcal{H}}X}\mathrm{pr}_{\mathcal{H}}Y + \mathrm{pr}_{\mathcal{H}}[\mathrm{pr}_{\mathcal{V}}X, \mathrm{pr}_{\mathcal{H}}Y] + \mathrm{pr}_{\mathcal{V}}[\mathrm{pr}_{\mathcal{H}}X, \mathrm{pr}_{\mathcal{V}}Y] + \mathrm{pr}_{\mathcal{V}}\nabla^{\mathbf{g}}_{\mathrm{pr}_{\mathcal{V}}X}\mathrm{pr}_{\mathcal{V}}Y, \tag{4.10}
$$

where $\nabla^{\mathbf{g}}$ is the Levi-Civita connection for the taming metric $\mathbf{g}$.

Lemma 4.10. *The Bott connection is compatible with the metric* $\mathbf{g}$ *and preserves the horizontal bundle. Additionally, we have that the torsion is given by*

$$
T^{\nabla}(X,Y) = -\mathcal{R}(X,Y) - \bar{\mathcal{R}}(X,Y).
$$

Proof. That it in fact is a linear connection which preserves $\mathcal{H}$ is straight forward to check. To see that ∇ is compatible with $\mathbf{g}$ we will divide it into several cases. For X, Y, Z horizontal or vertical ∇ inherits the compatibility from $\nabla^{\mathbf{g}}$. When $X \in \Gamma(\mathcal{H})$ and $Y, Z \in \Gamma(\mathcal{V})$ we have that

$$
\begin{aligned}
X(\mathbf{g}(Y,Z)) - \mathbf{g}(\nabla_X Y, Z) - \mathbf{g}(\nabla_X Z, Y) &= X(\mathbf{g}(Y,Z)) - \mathbf{g}([X,Y],Z) - \mathbf{g}([X,Z],Y)\\
&= (\mathcal{L}_X\mathbf{g})(Z,Y) = 0.
\end{aligned}
$$

When $X \in \Gamma(\mathcal{V})$ and $Y, Z \in \Gamma(\mathcal{H})$ the computation is nearly identical. For the case when $Y, X \in \Gamma(\mathcal{V})$ and $Z \in \Gamma(\mathcal{H})$ we have that

$$(\nabla_X \mathbf{g})(Y, Z) = X(\mathbf{g}(Y, Z)) - \mathbf{g}(\mathrm{pr}_{\mathcal{H}}[X, Z], Y) - \mathbf{g}(\mathrm{pr}_{\mathcal{V}} \nabla_X Y, Z) = 0,$$

and the same computation also shows that $Y, X \in \Gamma(\mathcal{H})$ and $Z \in \Gamma(\mathcal{V})$ is compatible.

Hence we are only left to calculate the torsion. If X is horizontal and Y is vertical we have that

$$\nabla_X Y - \nabla_Y X - [X, Y] = \mathrm{pr}_{\mathcal{V}}[X, Y] - \mathrm{pr}_{\mathcal{H}}[Y, X] - [X, Y] = 0,$$

and by the skew-symmetry of the torsion tensor we have that $T^{\nabla}(Y, X)$ is also zero. When $X, Y \in \Gamma(\mathcal{H})$ we have that

$$\begin{aligned} T^{\nabla}(X, Y) &= \mathrm{pr}_{\mathcal{H}} \nabla_X Y - \mathrm{pr}_{\mathcal{H}} \nabla_Y X - [X, Y] = \mathrm{pr}_{\mathcal{H}}[X, Y] - [X, Y] \\ &= -\mathrm{pr}_{\mathcal{V}}[X, Y] = -\mathcal{R}(X, Y), \end{aligned}$$

and similarly we get that for $X, Y \in \Gamma(\mathcal{V})$ the torsion is $-\bar{\mathcal{R}}(X, Y)$. Thus for any two vector fields X, Y we have that the torsion is $T^{\nabla}(X, Y) = -\mathcal{R}(X, Y) - \bar{\mathcal{R}}(X, Y)$. $\square$

In this case it is easily seen that $\kappa_3 = \kappa_4 = 0$, hence we get the following corollary of Theorem 4.6.

Corollary 4.11. *Assume that $(M, \mathcal{H}, \mathbf{h})$ is a compact sub-Riemannian manifold with a taming metric $\mathbf{g}$ satisfying 4.6 and 4.9, and that the Bott connection satisfies*

- $\langle \mathrm{Ric}\, \alpha, \alpha \rangle_{\mathbf{h}} \leq \rho_1 \|\alpha\|^2_{\mathbf{h}^*}$,
- $\|\alpha(T^{\nabla}(\cdot, \cdot))\|^2_{\mathbf{h}^0_2} \geq \rho_2 \|\alpha\|^2_{\mathbf{g}^*_{\mathcal{V}}}$,
- $\|T^{\nabla}(\cdot, \sharp^{\mathbf{h}} \alpha)\|^2_{\mathbf{h}^* \otimes \mathbf{g}_{\mathcal{V}}} \leq \kappa_1 \|\alpha\|^2_{\mathbf{h}^*}$,
- $|\langle \mathcal{B}(\alpha), \alpha \rangle_{\mathbf{h}^*}| \leq 2\kappa_2 \|\alpha\|_{\mathbf{g}^*_{\mathcal{V}}} \|\alpha\|_{\mathbf{h}^*}$,

where $\kappa_i \geq 0$ and $\rho_2 > 0$. Then the first eigenvalue of the sub-Laplacian satisfies

$$\sqrt{-\lambda_1} \geq \sqrt{\frac{\rho_1 \rho_2 - 4\kappa_2^2}{\rho_2 \left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)} + \left(\frac{4\kappa_2 \sqrt{\kappa_1}}{\rho_2 \left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)}\right)^2} - \frac{4\kappa_2 \sqrt{\kappa_1}}{\rho_2 \left(\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}\right)}.$$

If ∇ additionally satisfy the Yang-Mills condition, i.e. $\kappa_2 = 0$, we have that

$$\lambda_1 \leq \frac{-\rho_1}{\frac{d-1}{d} + \frac{3\kappa_1}{\rho_2}}.$$

The last inequality is the same as the one given in [BK16]. One key difference is that we do not require $\mathcal{V}$ to be integrable, hence the new result weakens the requirements for the result to hold.

In the case of the Bott connection we can weaken the compactness criteria to completeness. This is due to the Bonnet-Myers Theorem given in [GT16c], which states that if $(M, \mathcal{H}, \mathbf{h})$ is complete, equation 4.9 holds and $K = \rho_1 \rho_2 - 4\kappa_2^2 > 0$ then the manifold M is compact. It would be interesting to know if this is also the case in Theorem 4.6. Additionally, it would be fun try generalizing the result to cases which are not of step 2. To know if the bound is sharp or not would also be reassuring and if there exists an analogue of the Obata Sphere estimate for this bound.

4.3 Example with Non-Integrable Orthogonal Compliment

Let us consider the Lie group

$$SO(4) = \left\{ A \in M_{4\times 4}(\mathbf{R}) : \det(A) = 1, A^{-1} = A^T \right\}$$

with Lie algebra $\mathfrak{so}(4)$ consisting of all four by four skew-symmetric matrices. Define an inner-product on the Lie algebra by $\left\langle B^1, B^2 \right\rangle_{\mathfrak{so}(4)} = -\mathrm{tr} B^1 B^2$ and extend it to the entire tangent space left translation. Then

$$\langle [A, B], C\rangle_{\mathfrak{so}(4)} = -\mathrm{tr}(ABC) + \mathrm{tr}(BAC) = -\mathrm{tr}(BCA) + \mathrm{tr}(BAC) = -\langle B, [A, C]\rangle_{\mathfrak{so}(4)},$$

which is equivalent to the metric being bi-invariant. This implies that

$$\langle [Y, X], X\rangle_{\mathfrak{so}(4)} = -\langle X, [Y, X]\rangle_{\mathfrak{so}(4)},$$

and hence $\langle [Y, X], X\rangle_{\mathfrak{so}(4)}$. In particular, we have that if X and Y are orthogonal, left invariant vector fields of constant length, we get that $(\mathcal{L}_X \mathbf{g})(Y, Y) = 0$, which means that the Bott connection will be compatible with $\mathbf{g}$. In this case we have that

$$\lambda_1 \leq -\frac{1}{15}.$$

Define the matrix $B^{ij} = e_i e_j^T - e_j e_i^T$, and let $\mathfrak{h}$ denote the span of $B^{12}, B^{14}, B^{24}, B^{34}$ and $\mathfrak{v}$ the span of B^{13}, B^{23}. Define the horizontal and vertical bundle to be the left translation of $\mathfrak{h}$ and $\mathfrak{v}$, respectively.

By direct computation we get that the bracket is given by the formula

$$\left[B^{ij}, B^{kl}\right] = \delta_{jk} B^{il} - \delta_{ik} B^{jl} - \delta_{jl} B^{ik} + \delta_{il} B^{jk}.$$

Hence we get the following bracket relations

				B^{kl}			
	$\left[B^{ij}, B^{kl}\right]$	B^{12}	B^{14}	B^{24}	B^{34}	B^{13}	B^{23}
	B^{12}	0	$-B^{24}$	B^{14}	0	$-B^{23}$	B^{13}
	B^{14}	B^{24}	0	$-B^{12}$	$-B^{13}$	B^{34}	0
B^{ij}	B^{24}	$-B^{14}$	B^{12}	0	$-B^{23}$	0	B^{34}
	B^{34}	0	B^{13}	B^{23}	0	$-B^{14}$	$-B^{24}$
	B^{13}	B^{23}	$-B^{34}$	0	B^{14}	0	$-B^{12}$
	B^{23}	B^{13}	0	$-B^{34}$	B^{24}	B^{12}	0

.

If we normalize the given basis we get that

				j			
	$[X_i, X_j]$	X_1	X_2	X_3	X_4	X_5	X_6
	X_1	0	$-\frac{1}{\sqrt{2}}X_3$	$\frac{1}{\sqrt{2}}X_2$	0	$-\frac{1}{\sqrt{2}}X_6$	$\frac{1}{\sqrt{2}}X_5$
	X_2	$\frac{1}{\sqrt{2}}X_3$	0	$-\frac{1}{\sqrt{2}}X_1$	$-\frac{1}{\sqrt{2}}X_5$	$\frac{1}{\sqrt{2}}X_4$	0
i	X_3	$-\frac{1}{\sqrt{2}}X_2$	$\frac{1}{\sqrt{2}}X_1$	0	$-\frac{1}{\sqrt{2}}X_6$	0	$\frac{1}{\sqrt{2}}X_4$
	X_4	0	$\frac{1}{\sqrt{2}}X_5$	$\frac{1}{\sqrt{2}}X_6$	0	$-\frac{1}{\sqrt{2}}X_2$	$-\frac{1}{\sqrt{2}}X_3$
	X_5	$\frac{1}{\sqrt{2}}X_6$	$-\frac{1}{\sqrt{2}}X_4$	0	$\frac{1}{\sqrt{2}}X_2$	0	$-\frac{1}{\sqrt{2}}X_1$
	X_6	$\frac{1}{\sqrt{2}}X_5$	0	$-\frac{1}{\sqrt{2}}X_4$	$\frac{1}{\sqrt{2}}X_3$	$\frac{1}{\sqrt{2}}X_1$	0

.

If we have a bi-invariant metric the Christoffel symbols of Levi-Civita connection is given by $\Gamma_{ij}^k = \frac{1}{2}\mathbf{g}\left(\left[X_i, X_j\right], X_k\right)$, for an orthonormal basis X_i. Using this we can calculate the Bott connection as

	$\nabla_{X_i} X_j$	X_1	X_2	j X_3	X_4	X_5	X_6
	X_1	0	$-\frac{1}{\sqrt{2}^3} X_3$	$\frac{1}{\sqrt{2}^3} X_2$	0	$-\frac{1}{\sqrt{2}} X_6$	$\frac{1}{\sqrt{2}} X_5$
	X_2	$\frac{1}{\sqrt{2}^3} X_3$	0	$-\frac{1}{\sqrt{2}^3} X_1$	0	0	0
i	X_3	$-\frac{1}{\sqrt{2}^3} X_2$	$\frac{1}{\sqrt{2}^3} X_1$	0	0	0	0
	X_4	0	0	0	0	0	0
	X_5	0	$-\frac{1}{\sqrt{2}} X_4$	0	$\frac{1}{\sqrt{2}} X_2$	0	0
	X_6	0	0	$\frac{-1}{\sqrt{2}} X_4$	$\frac{1}{\sqrt{2}} X_3$	0	0

.

If we use the formula $T^{\nabla}(X, Y) = -\mathcal{R}(X, Y) - \bar{\mathcal{R}}(X, Y)$, we get that the torsion is given by

	$T^{\nabla}(X_i, X_j)$	X_1	X_2	j X_3	X_4	X_5	X_6
	X_1	0	0	0	0	0	0
	X_2	0	0	0	$\frac{1}{\sqrt{2}} X_5$	0	0
i	X_3	0	0	0	$\frac{1}{\sqrt{2}} X_6$	0	0
	X_4	0	$-\frac{1}{\sqrt{2}} X_5$	$-\frac{1}{\sqrt{2}} X_6$	0	0	0
	X_5	0	0	0	0	0	$\frac{1}{\sqrt{2}} X_1$
	X_6	0	0	0	0	$-\frac{1}{\sqrt{2}} X_1$	0

.

Hence if $\alpha = \sum_{j=1}^n \alpha_j \theta^j$, where $\theta^j = \flat^{\mathbf{g}} X_j$, we get that

$$\left\|\alpha\left(T^{\nabla}(\cdot, \cdot)\right)\right\|_{\mathbf{h}_2^0}^2 = \alpha_5^2 + \alpha_6^2,$$

implying that $\rho_2 = 1$. Additionally,

$$\left\|\mathcal{R}\left(\cdot, \sharp^{\mathbf{h}}\alpha\right)\right\|_{\mathbf{h}_1^0 \otimes \mathbf{g}_{\mathcal{V}_1^0}}^2 = \alpha_4^2 + \frac{1}{2}\alpha_2^2 + \frac{1}{2}\alpha_3^2 \leq \|\alpha\|_{\mathbf{h}^*}^2,$$

making $\kappa_1 = 1$. To find a suitable constant κ_2, we do the calculation

$$\begin{aligned}
\left\langle\left(\nabla_{\times} T^{\nabla}\right)\left(\times, \sharp^{\mathbf{h}}\alpha\right), \sharp^{\mathbf{g}}\alpha\right\rangle_{\mathbf{g}_{\mathcal{V}}} &= \sum_{i,j=1}^{4}\sum_{k=5}^{6} \alpha_j \alpha_k \langle\left(\nabla_{X_i} T^{\nabla}\right)\left(X_i, X_j\right), X_k\rangle_{\mathbf{g}} \\
&= \sum_{i,j=1}^{4}\sum_{k=5}^{6} \alpha_j \alpha_k \left\langle \nabla_{X_i}\left(T^{\nabla}\left(X_i, X_j\right)\right), X_k\right\rangle_{\mathbf{g}} \\
&= \sum_{j=1}^{4}\sum_{k=5}^{6} \alpha_j \alpha_k \left\langle \nabla_{X_1}\left(-\frac{1}{\sqrt{2}}\delta_{1,4}\delta_{j2} X_5 - \frac{1}{\sqrt{2}}\delta_{1,4}\delta_{j3} X_6\right), X_k\right\rangle_{\mathbf{g}} \\
&\quad + \alpha_j \alpha_k \left\langle \nabla_{X_1}\left(\frac{1}{\sqrt{2}}\delta_{1,2}\delta_{j4} X_5 + \frac{1}{\sqrt{2}}\delta_{1,3}\delta_{j4} X_6\right), X_k\right\rangle_{\mathbf{g}} \\
&= 0,
\end{aligned}$$

hence $\kappa_2 = 0$.

To use Corollary 4.11 the only thing left is to calculate the Ricci curvature. We have that

$$\nabla_{X_i}\nabla_{X_j}X_1 = \frac{1}{8}\delta_{i1}\delta_{j2}X_2 - \frac{1}{8}\delta_{i2}\delta_{j2}X_1 + \frac{1}{8}\delta_{i1}\delta_{j3}X_3 - \frac{1}{8}\delta_{i3}\delta_{j3}X_1,$$

$$\nabla_{X_i}\nabla_{X_j}X_2 = -\frac{1}{8}\delta_{i1}\delta_{j1}X_3 + \frac{1}{8}\delta_{i2}\delta_{j1}X_1 + \frac{1}{8}\delta_{i2}\delta_{j3}X_3 - \frac{1}{8}\delta_{i3}\delta_{j3}X_2,$$

and

$$\nabla_{X_i}\nabla_{X_j}X_3 = -\frac{1}{8}\delta_{i1}\delta_{j1}X_3 + \frac{1}{8}\delta_{j1}\delta_{i3}X_1 - \frac{1}{8}\delta_{i2}\delta_{j2}X_3 + \frac{1}{8}\delta_{i3}\delta_{j2}X_2.$$

Calculating parts of the curvature endomorphism we get that

$$\nabla_{X_i}\nabla_{X_j}X_1 - \nabla_{X_j}\nabla_{X_i}X_1 - \nabla_{[X_i,X_j]}X_1 = \frac{1}{8}\delta_{i1}\delta_{j2}X_2 + \frac{1}{8}\delta_{i1}\delta_{j3}X_3 - \frac{1}{8}\delta_{i2}\delta_{j1}X_2 - \frac{1}{8}\delta_{i3}\delta_{j1}X_3 - \nabla_{[X_i,X_j]}X_1,$$

$$\nabla_{X_i}\nabla_{X_j}X_2 - \nabla_{X_j}\nabla_{X_i}X_2 - \nabla_{[X_i,X_j]}X_2 = \frac{1}{8}\delta_{i2}\delta_{j1}X_1 + \frac{1}{8}\delta_{i2}\delta_{j3}X_3 - \frac{1}{8}\delta_{i1}\delta_{j2}X_1 - \frac{1}{8}\delta_{i3}\delta_{j2}X_3 - \nabla_{[X_i,X_j]}X_2,$$

and

$$\nabla_{X_i}\nabla_{X_j}X_3 - \nabla_{X_j}\nabla_{X_i}X_3 - \nabla_{[X_i,X_j]}X_3 = \frac{1}{8}\delta_{i3}\delta_{j1}X_1 + \frac{1}{8}\delta_{i3}\delta_{j2}X_2 - \frac{1}{8}\delta_{i1}\delta_{j3}X_1 - \frac{1}{8}\delta_{i2}\delta_{j3}X_2 - \nabla_{[X_i,X_j]}X_3.$$

The Ricci curvature hence becomes

		j			
	$\mathrm{Ric}(X_i, X_j)$	X_1	X_2	X_3	X_4
	X_1	$1/4$	0	0	0
i	X_2	0	$3/4$	0	0
	X_3	0	0	$3/4$	0
	X_4	0	0	0	1

This means that the Ricci curvature is bounded below by constant $\rho_1 = \frac{1}{4}$. Hence plugging the constants $\rho_1 = \frac{1}{4}$, $\rho_2 = 1$, $\kappa_1 = 1$ and $\kappa_2 = 0$ into Corollary 4.11 we obtain the bound $-\frac{1}{15} \geq \lambda_1$.

4.4 Example which does not satisfy the Yang-Mills Condition

Let us consider the Lie group

$$SU(2) = \left\{ A \in M_{2\times 2}(\mathbf{C}) : A = \begin{bmatrix} z_1 & z_2 \\ -\bar{z_2} & \bar{z_1} \end{bmatrix}, \det(A) = 1 \right\}.$$

Then the Lie algebra is spanned by the matrices

$$X = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}, \quad Y = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad Z = \frac{1}{\sqrt{2}}\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}.$$

We will use the same letters X, Y, Z to denote the vector fields obtained by left-translation of the matrices. Let $\mathbf{h}$ be a sub-Riemannian metric on the subspace spanned by X and Y making the vector fields orthonormal, and let $\mathbf{g}$ be a taming metric such that X, Y, Z are orthonormal. For any smooth function f such that $Z(f) = 0$ define a new sub-Riemannian metric by $e^{2f}\mathbf{h}$. Then $X_f = e^{-f}X$ and $Y_f = e^{-f}Y$ is an orthonormal frame for the new sub-Riemannian metric.

Let $\mathbf{g}_f$ be a new taming metric such that X_f, Y_f, Z are orthonormal. Then the bracket relations are given by

	$[X_i, X_j]$	j: X_f	j: Y_f	j: Z
i	X_f	0	$e^{-2f}Z - X_f(f)\,Y_f + Y_f(f)\,X_f$	$-Y_f$
i	Y_f	$X_f(f)\,Y_f - e^{-2f}Z - Y_f(f)\,X_f$	0	X_f
i	Z	Y_f	$-X_f$	0

.

The distribution $\mathcal{V} = \mathrm{Span}Z$ is metric preserving since

$$\left(\mathcal{L}_Z e^{2f}\mathbf{h}\right)(aX_f + bY_f, aX_f + bY_f) = 2e^{2f}\mathbf{h}\left([Z, bY_f], aX_f\right) + 2e^{2f}\mathbf{h}\left([Z, bX_f], aY_f\right) = 0.$$

We have that the connection becomes

	$\nabla_{X_i}X_j$	j: X_f	j: Y_f	j: Z
i	X_f	$-Y_f(f)\,Y_f$	$Y_f(f)\,X_f$	0
i	Y_f	$X_f(f)\,Y_f$	$-X_f(f)\,X_f$	0
i	Z	Y_f	$-X_f$	0

.

Then the torsion is given by

	$T^{\nabla}(X_i, X_j)$	j: X_f	j: Y_f	j: Z
i	X_f	0	$-e^{-2f}Z$	0
i	Y_f	$e^{-2f}Z$	0	0
i	Z	0	0	0

,

We also have that

$$R^2_{211} = R^1_{122} = e^{-2f} - X_f(X_f(f)) - Y_f(Y_f(f)) - X_f(f)^2 - Y_f(f)^2 = e^{-2f}\left(1 - X(X(f)) - Y(Y(f))\right),$$

and the rest are zero. Hence we have that

$$\rho_1 = \inf_{x \in SU(2)} e^{-2f}\left(1 - \left(X^2 + Y^2\right)f\right).$$

Since the torsion of vertical vector fields are 0 we have that $\kappa_3 = \kappa_4 = \kappa_5 = 0$. Let

$$\alpha = \alpha_1 \flat^{\mathbf{g}} X_f + \alpha_2 \flat^{\mathbf{g}} Y_f + \alpha_3 \flat^{\mathbf{g}} Z.$$

Then we get that

$$\langle \mathcal{B}^*\alpha, \alpha\rangle_{\mathbf{g}^*} = \alpha_3 2fe^{-2f}\left(-Y_f(f)\,\alpha_1 + X_f(f)\,\alpha_2\right) = \alpha_3 2fe^{-3f}(X(f)\alpha_2 - Y(f)\alpha_1),$$

hence we can set

$$\kappa_2 = \sqrt{2}\sup_{x \in SU(2)}\left\{|fe^{-3f}Y(f)|, |fe^{-3f}X(f)|\right\}.$$

We have that $\kappa_1 = \sup_{x \in SU(2)} e^{-4f}$, since

$$\|T^{\nabla}\left(\cdot, \sharp^{\mathbf{h}}\alpha\right)\|_{\mathbf{h}\otimes\mathbf{g}^*_{\mathcal{V}}} = e^{-4f}\|\alpha\|^2_{\mathbf{h}^*}.$$

By the formulas we have that the Yang-Mills condition is satisfied if and only if $X_f(f) = Y_f(f) = 0$. The last constant we need is $\rho_2 = 2\inf_{x\in SU(2)} e^{-4f}$, since $\|\alpha\left(T^{\nabla}(\cdot,\cdot)\right)\|^2_{\mathfrak{h}_2^0} = \alpha_3^2 2e^{-4f}$. Let $M = \max f$ and $m = \min f$, then

$$\sqrt{-\lambda_1} \geq \sqrt{\frac{2e^{-4M}\rho_1 - 4\kappa_2^2}{e^{-4M} + 3e^{-4m}} + \frac{16\kappa_2^2 e^{-4m}}{(e^{-4M}+3e^{-4m})^2}} - \frac{4e^{-2m}\kappa_2}{e^{-4M}+3e^{-4m}}.$$

Let us find when we have a spectral gap when

$$f\left(\begin{bmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{bmatrix}\right) = a\|z_1\|^2.$$

Then if $a > 0$ we get that $\rho_2 = 2e^{-4a}$, $\kappa_1 = 1$. If we left translate X, Y, Z we get

$$X = \frac{i}{\sqrt{2}}\left(z_2\partial_{z_1} - \bar{z}_2\partial_{\bar{z}_1} + z_1\partial_{z_2} - \bar{z}_1\partial_{\bar{z}_2}\right),$$

$$Y = \frac{1}{\sqrt{2}}\left(z_2\partial_{z_1} + \bar{z}_2\partial_{\bar{z}_1} - z_1\partial_{z_2} - \bar{z}_1\partial_{\bar{z}_2}\right)$$

and

$$Z = \frac{i}{\sqrt{2}}\left(z_1\partial_{z_1} - \bar{z}_1\partial_{\bar{z}_1} + z_2\partial_{z_2} + \bar{z}_2\partial_{\bar{z}_2}\right).$$

Hence we get that $X(f) = \frac{ai}{\sqrt{2}}\left(z_2\bar{z}_1 - \bar{z}_2 z_1\right)$, $Y(f) = \frac{a}{\sqrt{2}}\left(z_2\bar{z}_1 + \bar{z}_2 z_1\right)$ and

$$X(X(f)) = Y(Y(f)) = a\left(|z_2|^2 - |z_1|^2\right).$$

The constant ρ_1 can be calculated by

$$\rho_1 = \inf_{SU(2)} e^{-2a\|z_1\|^2}\left(1 - 2a\left(|z_2|^2 - |z_1|^2\right)\right) = \inf_{SU(2)} e^{-2a\|z_1\|^2}\left(1 - 2a\left(1 - 2|z_1|^2\right)\right).$$

The minimum value is when $x = 0$ when $a < \frac{1}{2}$, hence $\rho_1 = 1 - 2a$. To find a value for κ_2, we note that $\kappa_2 = 2a^2 \sup\left\{||z_1|^2 e^{-3a|z_1|^2}\mathrm{Re}\left(z_1\bar{z}_2\right)|\right\} \leq 2a^2e^{-3a}$. Hence we can estimate the spectral gap for $1 - 2a - 8a^4e^{-2a} > 0$, which is true for at least the values of a in $a \in (0, 0.4]$. Then the spectral gap is estimated by

$$\sqrt{\frac{2(1-2a) - 16a^4e^{-2a}}{1+3e^{4a}} + \frac{4a^4e^{2a}}{(1+3e^{4a})^2}} - \frac{2a^2e^a}{1+3e^4a} < \sqrt{-\lambda_1}.$$

4.5 Example with a Non-Totally Geodesic Foliation

Make two copies of $SU(2)$ with left invariant basis X_j, Y_j, Z_j, $j = 1, 2$. Where the bracket relations between X_j, Y_j, Z_j is defined as above for j fixed, and zero if they come from different copies. Define a new set of vector fields

$$X^+ = X_1 + X_2, \qquad X^- = X_1 - X_2,$$

$$Y^+ = Y_1 + Y_2, \qquad Y^- = Y_1 - Y_2,$$

and

$$Z^+ = Z_1 + Z_2, \qquad Z^- = Z_1 - Z_2.$$

The bracket relations between these new vector fields then become

$$[X^\pm, Y^\pm] = Z^+, \qquad [Y^\pm, Z^\pm] = X^+, \qquad [Z^\pm, X^\pm] = Y^+.$$

$$[X^\pm, Y^\mp] = Z^-, \qquad [Y^\pm, Z^\mp] = X^-, \qquad [Z^\pm, X^\mp] = Y^-.$$

For any real number $a \in \mathbf{R}$, define

$$X^a = X^- + aX^+.$$

Let $\mathcal{H}^a$ be the spanned by X^a, Y^- and Z^-, and let the metric $\mathbf{h}_a$ be defined such that the vector fields are orthonormal. Define the taming metric $\mathbf{g}_a$ to be such that the vector fields X^a, Y^-, Z^-, X^+, Y^+ and Z^+ are an orthonormal frame.

Then the constant K can be set to $K = 4 - \left(11 + 8\sqrt{2}\right) a^2$, and we can calculate the spectral gap for $|a| < \frac{2}{\sqrt{11+8\sqrt{2}}}$. The spectral gap is estimated by

$$-\lambda \geq \left(\sqrt{\frac{12 - 3\left(11 + 8\sqrt{2}\right) a^2}{22} + \left(\frac{11 + 6\sqrt{2}}{11}\right)^2 a^2} - \frac{6\sqrt{2}a}{11} \right)^2 - a^2.$$

The bracket relations then become

$[A,B]$	X^a	Y^-	Z^-	X^+	Y^+	Z^+
X^a	0	$Z^+ + aZ^-$	$-Y^+ - aY^-$	0	$Z^- + aZ^+$	$-Y^- - aY^+$
Y^-	$-Z^+ - aZ^-$	0	X^+	$-Z^-$	0	$X^a - aX^+$
Z^-	$Y^+ + aY^-$	$-X^+$	0	Y^-	$aX^+ - X^a$	0
X^+	0	Z^-	$-Y^-$	0	$-Z^+$	$-Y^+$
Y^+	$-Z^- - aZ^+$	0	$X^a - aX^+$	Z^+	0	X^+
Z^+	$Y^- + aY^+$	$aX^+ - X^a$	0	Y^+	$-X^+$	0

.

We have that $\mathcal{H}^a$ is metric preserving since if $A = \alpha X^a + \beta Y^- + \gamma Z^-$ then we have that

$$\left(\mathcal{L}_{X^+} \mathrm{pr}_{\mathcal{H}}^* \mathbf{h}_a\right)(A, A) = 2\mathbf{h}\left(\mathrm{pr}_{\mathcal{H}}\left[X^+, \beta Y^- + \gamma Z^-\right], \beta Y^- + \gamma Z^-\right) = 0$$

$$\left(\mathcal{L}_{Y^+} \mathrm{pr}_{\mathcal{H}}^* \mathbf{h}_a\right)(A, A) = 2\mathbf{h}\left(\mathrm{pr}_{\mathcal{H}}\left[Y^+, \alpha X^- + \gamma Z^-\right], \alpha X^- + \gamma Z^-\right) = 0$$

and

$$\left(\mathcal{L}_{Z^+} \mathrm{pr}_{\mathcal{H}}^* \mathbf{h}_a\right)(A, A) = 2\mathbf{h}\left(\mathrm{pr}_{\mathcal{H}}\left[Z^+, \alpha X^- + \beta Y^-\right], \alpha X^- + \beta Y^-\right) = 0.$$

Note also that $\bar{\mathcal{R}} = 0$, and hence $\mathrm{tr}\bar{\mathcal{R}}\left(X, \mathcal{R}\left(X, \times\right)\right) \times = 0$. We can not use the Bott connection since

$$\left(\mathcal{L}_{Y^-} \mathrm{pr}_{\mathcal{V}}^* \mathbf{g}_{\mathcal{V}}\right)\left(\alpha X^+ + \beta Z^+, \alpha X^+ + \beta Z^+\right) = -a\alpha\beta.$$

It is also easy to see that $\mathrm{tr}_{\mathbf{h}}\left(\mathcal{L}_A \mathbf{g}_a\right)(\times, \times) = 0$ whenever A is vertical. Define the connection ∇ by

$$\nabla_X Y = \mathrm{pr}_{\mathcal{H}} \nabla^{\mathbf{g}_a}_{\mathrm{pr}_{\mathcal{H}} X} \mathrm{pr}_{\mathcal{H}} Y + \mathrm{pr}_{\mathcal{H}}\left[\mathrm{pr}_{\mathcal{V}} X, \mathrm{pr}_{\mathcal{H}} X\right] + \mathrm{pr}_{\mathcal{V}} \nabla'_{\mathrm{pr}_{\mathcal{V}} X} \mathrm{pr}_{\mathcal{V}} Y,$$

where ∇' is defined by the fact that X^+, Y^+ and Z^+ is a parallel frame. Then the connection is given by

$\nabla_A B$	X^a	Y^-	Z^-	X^+	Y^+	Z^+
X^a	0	aZ^-	$-aY^-$	0	0	0
Y^-	0	0	0	0	0	0
Z^-	0	0	0	0	0	0
X^+	0	Z^-	$-Y^-$	0	0	0
Y^+	$-Z^-$	0	X^a	0	0	0
Z^+	Y^-	$-X^a$	0	0	0	0

.

Calculating the Ricci curvature we get that $R^2_{211} = R^3_{311} = R^1_{122} = R^3_{322} = R^1_{133} = R^2_{233} = 1$, and the rest is zero. Hence $\rho_1 = 2$. The torsion of the connection is given by

$T^\nabla(A,B)$	X^a	Y^-	Z^-	X^+	Y^+	Z^+
X^a	0	$-Z^+$	Y^+	0	$-aZ^+$	aY^+
Y^-	Z^+	0	$-X^+$	0	0	aX^+
Z^-	$-Y^+$	X^+	0	0	$-aX^+$	0
X^+	0	0	0	0	Z^+	Y^+
Y^+	aZ^+	0	aX^+	$-Z^+$	0	$-X^+$
Z^+	$-aY^+$	$-aX^+$	0	$-Y^+$	X^+	0

.

Let

$$\alpha = \flat^{\mathbf{g}_a}\left(\alpha_1 X^a + \alpha_2 Y^- + \alpha_3 Z^- + \alpha_4 X^+ + \alpha_5 Y^+ + \alpha_6 Z^+\right).$$

Then $\|\alpha\left(T^\nabla(\cdot,\cdot)\right)\|^2_{\mathbf{h}^0_2} = 2\|\alpha\|^2_{\mathbf{g}_\mathcal{V}}$, and $\|T^\nabla\left(\cdot, \sharp^{\mathbf{h}}\alpha\right)\|^2_{\mathbf{h}^*}$. Hence $\rho_2 = \kappa_1 = 2$. We can set $\kappa_2 = |a|/2$, since

$$\langle T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^{\mathbf{h}_a}\alpha\right), T^\nabla\left(\mathrm{pr}_\mathcal{H}\cdot, \sharp^{\mathbf{g}_\mathcal{V}}\alpha\right)\rangle_{\mathbf{g}^1_1} = 0,$$

and $\langle B^*\sharp^{\mathbf{h}}\alpha, \alpha\rangle_{\mathbf{g}^*} = -a\left(\alpha_2\alpha_5 + \alpha_3\alpha_6\right)$.

To calculate κ_3, we see that

$$\langle \mathrm{tr}_{\mathbf{h}_a} T^\nabla\left(\times, T^\nabla\left(\times, \sharp^{\mathbf{g}_\mathcal{V}}\alpha\right)\right)\rangle = -a^2\left(\alpha_5^2 + \alpha_6^2\right),$$

$$\|T^\nabla(\cdot,\cdot)\|^2_{\mathbf{h}^0_2} = 2a^2\left(\alpha_5^2+\alpha_6^2\right)$$

and

$$\alpha\left(\left(\mathrm{tr}_{\mathbf{h}}\nabla_\times T^\nabla\right)\left(\times, \sharp^{\mathbf{g}_\mathcal{V}}\alpha\right)\right) = 0.$$

Hence we can set $\kappa_3 = a^2$.

We can set $\kappa_4 = 2|a|$, since

$$\alpha\left(\mathrm{tr}_{\mathbf{h}} T^\nabla\left(\times, T^\nabla\left(\times, \sharp^{\mathbf{h}}\alpha\right)\right)\right) = 2a\alpha_1\alpha_5 - a\alpha_2\alpha_5 - a\alpha_3\alpha_6.$$

Hence the constant K can be set to $K = 4 - \left(11 + 8\sqrt{2}\right)a^2$, and we can approximate the spectral gap if $|a| < \frac{2}{\sqrt{11+8\sqrt{2}}}$. In this case the spectral gap is given by

$$-\lambda \geq \left(\sqrt{\frac{12 - 3\left(11+8\sqrt{2}\right)a^2}{22} + \left(\frac{11+6\sqrt{2}}{11}\right)^2 a^2} - \frac{6\sqrt{2}a}{11}\right)^2 - a^2.$$

A Operator Theory for Second Order Operators

In this section we will give a short introduction to self-adjoint operators and their spectrum. The content of this section will be used when discussing the Laplacian and sub-Laplacian in Section 2, since they are, in the cases we study, self-adjoint operators. Many of the theorems in this section will not be proved, but can be found in standard books covering the theme, see e.g. [Kre89].

If $T : \mathcal{D}(T) \subset \mathbf{R}^n \to \mathbf{R}^n$ is a linear operator, then we know from linear algebra that $(T - \lambda I)^{-1}$ exists and is continuous as long as λ is not an eigenvalue of T. When we have a Banach space, there are several ways $T - \lambda I$ can fail to have a "nice" inverse. As in finite dimensions, $T - \lambda I$ can fail to be injective, and hence the inverse does not exist or the inverse may fail to be continuous. Additionally, since we want to work with partially defined operators we will use the following definition.

Definition A.1. Let X be a nontrivial complex normed space, and let $T : \mathcal{D}(T) \to X$ be a linear operator with domain $\mathcal{D}(T) \subset X$. For any complex number λ, the operator $T - \lambda I$ is denoted by T_λ, where I is the identity operator on $\mathcal{D}(T)$, and R_λ denotes $(T_\lambda)^{-1}$ whenever it exists. The **resolvent** of T is the set $\rho(T) \subset \mathbf{C}$ such that for all λ in $\rho(T)$ we have that R_λ;

1. exists,
2. is bounded and
3. is densely defined on X.

The **spectrum** of T, denoted $\sigma(T)$, is defined to be the complement of the resolvent in $\mathbf{C}$. We further divide the spectrum of an operator into three parts;

- **point spectrum** $\sigma_p(T)$ are the elements in the spectrum where property 1 fails for R_λ.
- **continuous spectrum** $\sigma_c(T)$ are the elements where 1 and 3 holds while 2 does not hold.
- **residual spectrum** $\sigma_r(T)$ are the elements where property 1 holds, while property 3 does not hold.

The elements in the point spectrum are called **eigenvalues**, and if u is a nonzero element such that $Tu = \lambda u$ (i.e. λ is in $\sigma_p(T)$), then u is called an **eigenvector** of the eigenvalue λ.

In the case where $T : \mathcal{D}(T) \subset X \to X$ is a bounded operator it is well known that $\sigma(T)$ is closed. If T additionally is defined on the entire space X, then $\sigma(T)$ is bounded by the norm of T and the spectrum is compact. Moreover, we can guarantee that the spectrum $\sigma(T)$ is nonempty, which is not necessarily the case for unbounded operators.

Definition A.2. A **compact operator** $T : X \to Y$, where X and Y are Banach spaces, is an operator where $\overline{T(B_X)}$ is compact, where B_X denotes the unit ball in X.

The set of compact operators is a closed subspace of all bounded operators from X to Y, where the closedness can be showed by using sequentially compactness together with a diagonal argument.

Theorem A.3. *Let $T : X \to X$ be a compact operator. Then the spectrum $\sigma(T)$ is at most countable and the only possible accumulation point is zero.*

A.1 Spectral Theory of Self-Adjoint Operators

In this section let $(H, \langle\cdot,\cdot\rangle)$ denote a Hilbert space.

Definition A.4. Let $T : \mathcal{D}(T) \subset H \to H$ be a densely defined operator. The **domain of the adjoint operator** is the set

$$\mathcal{D}(T^*) = \{y \in H : \text{there exists } y^* \in H \text{ such that } \langle Tx, y\rangle = \langle x, y^*\rangle \text{ for all } x \in \mathcal{D}(T)\}.$$

We define the **adjoint operator** $T^* : \mathcal{D}(T^*) \subset H \to H$ by $T^*(y) = y^*$.

An operator is said to be **symmetric** if $T = T^*$ on common domain, and **self-adjoint** if T is symmetric and $\mathcal{D}(T) = \mathcal{D}(T^*)$.

The reason for requiring the domain to be dense in H in the definition of adjoint is to ensure the uniqueness of the adjoint operator. If a self-adjoint operator is defined on H, then the operator automatically becomes bounded, see [Kre89]. Thus the only unbounded self-adjoint operators have domain which is strictly contained in H. For bounded self-adjoint operators the spectrum is real and

$$\|T\| := \sup_{\|x\|=1} \|T(x)\| = \max_{\lambda\in\sigma(T)} |\lambda|.$$

Given partially defined operators on a Hilbert space we can define a partial order by $(T, \mathcal{D}(T)) \subset (S, \mathcal{D}(S))$ if $\mathcal{D}(T) \subset \mathcal{D}(S)$ and $T = S|_{\mathcal{D}(T)}$. We call $(S, \mathcal{D}(S))$ an extension of $(T, \mathcal{D}(T))$. Assume $(T, \mathcal{D}(T))$ is a symmetric operator, then the adjoint operator $(T^*, \mathcal{D}(T^*))$ is an extension of $(T, \mathcal{D}(T^*))$. This can easily be shown by using the definition of adjoint.

Definition A.5. We call a symmetric operator $(T, \mathcal{D}(T))$ **essentially self-adjoint** if there exists a unique extension $(S, \mathcal{D}(S))$ which is self-adjoint.

Definition A.6. A linear operator $T : \mathcal{D}(T) \to H$ is called **positive** if $\langle Tx, x\rangle \geq 0$, for every $x \in \mathcal{D}(T)$.

The operators we will be working with are negative, i.e. $\langle Tx, x\rangle \leq 0$, but since whenever T is a negative operator $-T$ is a positive operator the theorems remains valid after trivial modifications.

In what follows we need the following operator, which will be surprisingly useful. Define the unitary (with respect to the usual inner-product on $H \times H$) operator $V : H \times H \to H \times H$ by $(v, w) \mapsto (-w, v)$.

Lemma A.7. *If $T : \mathcal{D}(T) \subset H \to H$ is a densely defined operator, then*

$$\Gamma(T^*) = (V(\Gamma(T)))^{\perp},$$

where $\Gamma(T)$ denotes the graph of T.

Proof. By using the definitions we have that the following are equivalent:

- $(y, z) \in \Gamma(T^*)$,
- $\langle Tx, y\rangle = \langle x, z\rangle$ for all $x \in \mathcal{D}(T)$,
- $\langle(-Tx, x), (y, z)\rangle_{H\times H} = 0$,
- and $(y, z) \in (V(\Gamma(T)))^{\perp}$.

□

Theorem A.8. *Assume $(T, \mathcal{D}(T))$ is a densely defined operator in H, then T^* is a closed operator.*

Proof. For any subspace M of a Hilbert space $M^{\perp}$ is a closed subspace of H. Hence by the Lemma A.7 we have that $\Gamma(T^*)$ is closed in $H \times H$. □

The theorem implies that all self-adjoint operators are closed operators. Define an operator to be **closable** if there exists a closed extension of the operator. Then all symmetric operator are closable, again by the previous theorem. Let $\bar{T}$ denote the smallest closed extension of T, then $\bar{T}$ is called the closure of T. It is easily checked that if $(S, \mathcal{D}(S))$ is a self-adjoint extension of T, then $\left(\bar{T}, \mathcal{D}\left(\bar{T}\right)\right) \subset (S, \mathcal{D}(S))$. Hence if the closure of T is self-adjoint, then T is essentially self-adjoint.

Theorem A.9. *If $(T, \mathcal{D}(T))$ is a positive symmetric operator then there exists a self-adjoint extension $(S, \mathcal{D}(S))$ of $(T, \mathcal{D}(T))$.*

Proof. Define a new inner-product on $\mathcal{D}(T)$ by

$$\langle v, w\rangle_T := \langle v, w\rangle + \langle Tv, w\rangle,$$

where $v, w \in \mathcal{D}(T)$ and $(H, \langle\cdot,\cdot\rangle)$ is our Hilbert space. Denote by H_T the completion of $\mathcal{D}(T)$ with respect to the inner-product $\langle\cdot,\cdot\rangle_T$. Then we have an isometry $i : (\mathcal{D}(T), \langle\cdot,\cdot\rangle_T) \to (H_T, \langle\cdot,\cdot\rangle_T)$, and also a continuous map $j : (\mathcal{D}(T), \langle\cdot,\cdot\rangle_T) \to (H, \langle\cdot,\cdot\rangle)$ by inclusion. Since $\mathcal{D}(T)$ is dense in H_T, we can extend j to the space H_T, we will denote this extension by $J : (H_T, \langle\cdot,\cdot\rangle_T) \to (H, \langle\cdot,\cdot\rangle)$.

We have that J becomes injective. Assume that $J(\varphi) = 0$ for some $\varphi \in H_T$. Since $i(\mathcal{D}(T))$ is dense in H_T, there exist a sequence $v_n \in \mathcal{D}(T)$ such that $\|i(v_n) - \varphi\|_T \to 0$. This implies that $\|v_n\| \to 0$, since

$$\|v_n\| \le \|j(i(v_n))\|_T = \|J(i(v_n) - \varphi)\|_T \le \|i(v_n) - \varphi\|_T.$$

Hence $\varphi = 0$, since

$$\|\varphi\|_T^2 = \lim_{m\to\infty}\lim_{n\to\infty} \langle i(v_m), i(v_n)\rangle_T = \lim_{m\to\infty}\lim_{n\to\infty} \left(\langle v_m, v_n\rangle + \langle Tv_m, v_n\rangle\right) = 0.$$

For all $h \in H$ associate a map $\lambda_h : H_T \to \mathbf{C}$ by $\lambda_h(v) = \langle Jv, h\rangle$. Then for each h, we have that $\|\lambda_h\varphi\| \le \|\varphi\|_T\|h\|$, and thus each $\|\lambda_h\| \le \|h\|$. For each $h \in H$ there exists a unique $Ch \in H_T$ such that $\lambda_h v = \langle v, Ch\rangle_T$, by using the Riesz representation Theorem. The map $C : H \to H_T$ becomes linear and has norm less that 1.

Define the operator $B : H \to H$ by $B = J \circ C$ which becomes bounded by 1, being the composition of bounded operators with norm less or equal to 1. The operator B is positive and self-adjoint since

$$\langle Bv, w\rangle = \langle J(Cv), w\rangle = \lambda_w Cv = \langle Cv, Cw\rangle_T = \overline{\lambda_v Cw} = \overline{\langle J(Cw), v\rangle} = \langle v, Bw\rangle.$$

If $\varphi \in \ker(B)$ then, since J is injective, $0 = \langle v, C\varphi\rangle_T = \langle Jv, \varphi\rangle$ for all $v \in H_T$. Hence $\varphi = 0$, since $\mathcal{D}(T) \subset H_T$ is dense in H. The image of C is dense in H_T, since if $\varphi \in \mathrm{Im}(C)^{\perp}$ then $0 = \langle\varphi, Ch\rangle_T = \lambda_h\varphi = \langle\varphi, h\rangle$ for all $h \in H$, resulting in that $B(H)$ is dense in $J(H_T)$. Since $J(H_T)$ contains $\mathcal{D}(T)$, we have that $B(H)$ is dense in H.

Define $Av = B^{-1}v$, which becomes a positive symmetric operator using that B is positive and self-adjoint. Let us show that A also becomes self-adjoint. Define the unitary operator $U : H \times H \to H \times H$ by $(v, w) \mapsto (w, v)$. Then

$$\begin{aligned}\Gamma(A^*) &= (V\Gamma(A))^{\perp} = (V(U\Gamma(B)))^{\perp} \\ &= (U(V(\Gamma(B))))^{\perp} = U\left((V\Gamma(B))^{\perp}\right) \\ &= U(\Gamma(B^*)) = U(\Gamma(B)) = \Gamma(A),\end{aligned}$$

where V is defined as in Lemma A.7. Hence A is self-adjoint.

Define the operator $R = T + \mathrm{id}_H$, then for $v, w \in \mathcal{D}(T)$

$$\langle w, Rv\rangle = \langle J(i(w)), Rv\rangle = \lambda_{Rv} i(w) = \langle i(w), C(Rv)\rangle_T$$

on the other hand

$$\langle w, Rv\rangle = \langle i(w), i(v)\rangle_T.$$

It follows by the density of $\mathcal{D}(T)$, that $i(v) = C(Rv)$, resulting in $B(Rv) = J(C(Rv)) = J(i(v)) = v$. Hence $R \subset A$. If we define the self-adjoint operator $S = A - \mathrm{id}_H$, then $S|_{\mathcal{D}(T)} = T$. We are only left showing that S is positive. We have that

$$\langle v, Av\rangle = \langle Bx, ABx\rangle = \langle Bx, x\rangle = \lambda_{Cx} x = \langle Cx, Cx\rangle_T = \|Cx\|_T^2 \geq \|J(Cx)\|^2 = \|Bx\|^2 = \|v\|^2,$$

it follows that S is positive. □

The extension constructed in the previous proof is called the **Friedrichs extension**. As seen from the construction, if $w \in H_T$, and v_j is a sequence in $\mathcal{D}(T)$ converging to w, then $Sw = \lim_{n\to\infty} Tw_j$ in the norm of H. The Friedrichs extension is in general not the only self-adjoint extension of $(T, \mathcal{D}(T))$, however by using the corollaries of Theorem X.1 in [RS75], we get the following.

Theorem A.10. *Let A be a closed positive symmetric operator. Then A is self-adjoint if and only if A^* do not have any negative eigenvalues.*

A.2 Distributions and Sobolev Spaces

Denote by $\mathcal{D}'(\Omega)$ the **space of distributions** on Ω, i.e., the dual space of test functions

$$C_0^\infty(\Omega) = \{f : \Omega \to \mathbf{R} : \text{all partial derivatives of } f \text{ exist, and } f \text{ has compact support}\}.$$

If $f \in L^1_{\mathrm{loc}}(\Omega)$, then we can associate the distribution $\tilde{f}(\varphi) = \int_\Omega f(x)\varphi(x)\,dx$ where $\varphi \in C_0^\infty(\Omega)$. Define the derivative of the distribution T by

$$(\partial^\alpha T)\varphi = (-1)^{|\alpha|} T(D^\alpha \varphi).$$

Then $\tilde{f}$ defined above always possesses all derivatives, by using integration by parts. We say that a distribution T belongs to $L^p(\Omega)$ if $T = \tilde{g}$ for some $g \in L^p(\Omega)$, and, by abuse of notation, we write $T \in L^p(\Omega)$.

Definition A.11. The **Sobolev space** $W^{k,p}(\Omega)$ is defined to be the set

$$W^{k,p}(\Omega) = \left\{ f \in L^p(\Omega) : \forall |\alpha| \leq k, \exists g_\alpha \in L^p(\Omega) \text{ such that } \partial^\alpha \tilde{f} = g_\alpha \right\},$$

where α is a multi-index. In the special case when $p = 2$ we are going to denote $W^{k,2}(\Omega)$ by $W^k(\Omega)$.

We can define the Sobolev (k,p) norm on $W^{k,p}(\Omega)$ by

$$\|f\|_{k,p} = \left(\sum_{|\alpha| \leq k} \|\partial^\alpha \tilde{f}\|_{L^p}^p \right)^{1/p} = \left(\sum_{|\alpha| \leq k} \|g_\alpha\|_{L^p}^p \right)^{1/p}.$$

When $p = 2$ we can also define an inner-product on $W^k(\Omega)$ by

$$\langle f, g \rangle_k = \sum_{|\alpha| \leq k} \langle \partial^\alpha \tilde{f}, \partial^\alpha \tilde{g} \rangle_{L^2}^2.$$

Theorem A.12. *The Sobolev spaces $W^{k,p}(\Omega)$ are complete. In other words they are Banach spaces, and if $p = 2$ they are Hilbert spaces.*

We are also going to use the space $W_0^{k,p}$, or $W_0^k(\Omega)$ in the case $p = 2$, which is defined to be the closure of $C_0^\infty(\Omega)$ in the Sobolev (k,p)-norm. Intuitively $W_0^{k,p}(\Omega)$ contains all the functions in $W^{k,p}(\Omega)$ which are zero on the boundary of Ω. The notation $W_{loc}^{k,p}(\Omega)$ is going to denote the set of functions which are locally in $W^{k,p}(\Omega)$. In the special case where $\Omega = \mathbf{R}^n$ we have the following useful result.

Theorem A.13. *The test functions $C_0^\infty(\mathbf{R}^n)$ are dense in $W^{k,p}(\mathbf{R}^n)$.*

Note that this is not necessarily true for any other open domain, since the test functions are zero on the boundary, while one can find functions in the Sobolev spaces which, intuitively, are nonzero on the (finite part of the) boundary. This argument can be made formal by using the trace of a function. A **mollifier** is a non-negative test function which has integral 1 and is zero outside the unit ball. Define $\psi_j(x) = j^n \psi(jx)$, where $\psi : \mathbf{R}^n \to \mathbf{R}$ is a mollifier. Then we can define a **mollification of f with radius** $\varepsilon > 0$ to be $\psi_j * f$ for some $1/j < \varepsilon$, where $*$ denotes convolution.

Lemma A.14. *If $f \in L^p(\mathbf{R}^n)$ then $f * \psi_j \to f$ in $L^p(\mathbf{R}^n)$ as $j \to \infty$.*

Using Lemma A.14 it is possible to show that smooth functions on Ω are dense in $W^{k,p}(\Omega)$.

A.3 Second Order Differential Operators

In this section we are going to study second order differential operators on the form

$$L = \sum_{i,j=1}^{n} a^{ij} \partial_j \partial_i + \sum_{i=1}^{n} b^i \partial_i + c, \tag{A.1}$$

defined on the open connected bounded set U, where $a^{ij} = a^{ji}$, $a^{ij}, b^i, c \in L^\infty(U)$, and $a^{ij} \in C^1(U)$. In the classical interpretation of the problem $Lu = g$ together with some boundary conditions, we need u to be at least twice differentiable for Lu to make sense. However, $C^2(U)$ is not a very good space for L to act on, since, among other things, proving existence of solutions becomes harder. Instead we will show how to "extend" the domain of L to $W_0^1(U)$. Define the **symbol** of a second order differential operator L to be $\sigma_L(df, dg) = \sum_{i,j=1}^n a^{ij}\partial_i(f)\partial_j(g)$, where $f, g \in \mathcal{D}(L)$. By using the product rule, we can also write the symbol of an operator on the form

$$\sigma_L(df, dg) = \frac{1}{2}\left(L(fg) - fL(g) - gL(f)\right). \tag{A.2}$$

When we are going to work with second order differential operators (see Appendix B) on manifolds definition A.2 of the symbol is more natural to work with, since it does not depend on choice of coordinates.

For uniqueness and existence problems it is easier to work with (A.1) written in the form

$$\begin{aligned} Lu &= \sum_{i,j=1}^n a^{ij}u_{x_i,x_j} + \sum_{i=1}^n b^i u_{x_i} + cu \\ &= \sum_{i,j=1}^n \left(a^{ij}u_{x_i}\right)_{x_j} + \sum_{i=1}^n \tilde{b^i} u_{x_i} + cu, \end{aligned}$$

where $\tilde{b^i} = b^i - \sum_j^n a^{ij}_{x_j}$.

To make our extension to the $W_0^1(U)$, let us begin by considering the Dirichlet problem

$$\begin{cases} -Lu = f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{A.3}$$

If we take an arbitrary test function $v \in C_0^\infty(U)$ then we have, using integration by part, that any solution of (A.3) solves

$$\langle -Lu, v\rangle_{L^2(U)} = \int_U \sum_{i,j=1}^n a^{ij}u_{x_i}v_{x_j} - \sum_{i=1}^n \tilde{b^i} u_{x_i} v - cuvdx = \langle f, v\rangle_{L^2(U)}.$$

Define the bilinear form $B[\cdot,\cdot]$ on $W_0^1(U)$ by

$$B[u, v] = \int_U \sum_{i,j=1}^n a^{ij}u_{x_i}v_{x_j} - \sum_{i=1}^n \tilde{b^i} u_{x_i} v - cuvdx.$$

This formulation of the problem is called a variational formulation. We say that u is a **weak solution** of the Dirichlet problem if $B[u, v] = \langle f, v\rangle$ for all $v \in W_0^1(U)$, where the derivatives are taken as derivatives of distributions.

Setting $c = \tilde{b^i} = 0$, L becomes a symmetric operator and additionally

$$\langle Lu, v\rangle_{L^2(U)} = -B[u, v] = -\int_U \sigma_L(du, dv)\, dx.$$

If we also assume that L is elliptic (defined in the next subsection), we get that L is a symmetric negative operator.

A.3.1 Elliptic Operators

This section is about second order elliptic operators. In Section 2 we will encounter the Laplacian which is prime examples of a second order elliptic operators. The sub-Laplacian which we will use in Section 3 and 4 is however not elliptic, even thought it share some of the properties like regularity.

Definition A.15. A second order differential operator L defined on U is called (uniformly) **elliptic** if there exists a positive constant θ such that for all $f \in C^{\infty}(U)$, we have that $\sigma_L(df, df) \geq \theta \|df\|_2^2$.

In other words the operator L is elliptic if at each point the matrix $A = \left(a^{ij}\right)$ defines an inner-product, where the smallest eigenvalue of the matrix is greater or equal to θ. When we define the Laplacian of a Riemannian manifold the matrix A is going to be nothing else than the Riemannian cometric $\mathbf{g}^*$ written in coordinates, see Section 2.

When our Riemannian manifold is some open set $U \subset \mathbf{R}^n$, endowed with the standard euclidean metric we get the "standard Laplacian" Δ. In this case $\left(a^{ij}\right) = I$ and $b^i = c = 0$. This is known as the modeling example of an elliptic PDE, where the idea is that if you have some result for Δ, it should have some kind of analogue for elliptic operators. In the case of the standard Laplacian we have uniqueness and existence of solution for the Dirichlet problem, maximum principle and regularity results, which also have analogues for elliptic operators (in the case of the maximum principle several analogues).

The next theorem deals with the uniqueness and existence of the Dirichlet problem. All the theorems in this subsection are proved in [Eva10].

Theorem A.16 (Fredholm Alternative)**.** *Let L be a uniformly elliptic operator defined on $W_0^1(U)$. Then either 0 will be an eigenvalue, or for all $f \in L^2(U)$ there exists a unique weak solution of (A.3). Also if 0 is an eigenvalue, then its eigenspace is of finite dimension.*

Let us by $\sigma_p^{\mathbf{R}}$ denote the real part of the point spectrum of an operator.

Theorem A.17. *The real part of the spectrum of an elliptic operator L defined on $W_0^1(U)$ is at most countable. Moreover, if the real part of the spectrum is countable, then the eigenvalues can be ordered $\lambda_i > \lambda_{i+1}$ and $\lambda_i \to -\infty$.*

If $\lambda \notin \sigma_p^{\mathbf{R}}(L)$, then there exists a unique weak solution to the problem

$$\begin{cases} -Lu = \lambda u + f & \text{in } U \\ u = 0 & \text{on } \partial U. \end{cases} \tag{A.4}$$

In this case the inverse L^{-1} is a bounded operator and defined on the entire $L^2(U)$.

The eigenvalue of L with biggest real part is called the **principle eigenvalue** of L. An interesting fact is that the principle eigenvalue is always real and has one dimensional eigenspace, even thought the rest of the spectrum might be complex. We use the notation $V \subset\subset U$ if V is compactly embedded in U.

Theorem A.18. *Assume that $a^{ij}, b^i, c \in C^{m+1}(U)$, where m is a non-negative integer. Assuming $f \in W^m(U)$ we get that $L^{-1}(f) \subset W^{m+2}(U)$. Additionally, if $V \subset\subset U$ there exists a constant C such that*

$$\|u\|_{W^{m+2}(V)} \leq C\left(\|f\|_{W^m(U)} + \|u\|_{L^2(U)}\right),$$

for all $f \in W^m(U)$.

Using the Rellich-Kondrachov Theorem for we have the corollary.

Corollary A.19. *Assume $a^{ij}, b^i, c \in C^\infty(U)$ and $f \in C^\infty(U)$. Then $L^{-1}(f) \subset C^\infty(U)$.*

The next theorem which we are going to address is the maximum principle.

Theorem A.20. *Assume $c = 0$ and $u \in C^2(U) \cap C(\bar{U})$. If $Lu \geq 0$ ($Lu \leq 0$) throughout U, then if u attains its maximum (minimum) at an interior point, then u is identically constant.*

B Topics From Manifolds

B.1 Lie Groups and Lie Algebras

A **Lie group** is a manifold G with a group structure such that the group operation is a smooth map[1]. Define the left multiplication with an element $g \in G$ to be the map $l_g : G \to G$ sending $h \to g \cdot h$, and define the right multiplication r_g similarly.

A vector field X is called **left invariant** if $l_{g*}(X) = X$. The **Lie algebra** of a Lie group is the set of all left-invariant vector fields together with the **Lie bracket**

$$[X, Y](f) = X(Y(f)) - Y(X(f)).$$

It can be showed that the Lie bracket of two left invariant vector fields is again left invariant. We will identify the Lie algebra of a Lie group with the tangent space at the identity, by sending $X \mapsto X(e)$, where e denotes the identity element in G. The Lie algebra of the Lie group G is denoted by $\mathfrak{g}$.

The **adjoint representation** $\mathrm{Ad} : G \to \mathrm{Aut}(\mathfrak{g})$ is defined by $\sigma \mapsto l_{\sigma *} r_{\sigma^{-1} *}$, and the differential of the adjoint representation at the identity will be denoted by $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. Define the **Lie exponential map**, denoted $\exp : \mathfrak{g} \to G$ (not to be confused with the Riemannian exponential map which we are going to denote by Exp), by sending X to $\gamma_X(1)$, where γ_X is the unique integral curve of X starting at the identity. The adjoint representation and the Lie exponential map satisfy the following relation, which is proved in e.g. [War83],

$$\begin{tikzcd} G \arrow[r, "\mathrm{Ad}"] & \mathrm{Aut}(\mathfrak{g}) \\ \mathfrak{g} \arrow[u, "\exp"] \arrow[r, "\mathrm{ad}"] & \mathrm{End}(\mathfrak{g}) . \arrow[u, "\exp"'] \end{tikzcd}$$

If our Lie group additionally has the structure of a Riemannian manifold, then we say that the metric $\mathbf{g}$ is **left invariant** if for all $g \in G$ we have that $l_g^* \mathbf{g} = \mathbf{g}$. We define right invariance similarly. If our metric $\mathbf{g}$ is both left and right invariant, then $\mathbf{g}$ is called **bi-invariant**.

B.2 Lie Derivative

One can also represent the Lie bracket with the help of the Lie derivative. Let X be a vector field, and denote the flow of the vector field by X_t, where $X_t(x)$ is the unique integral curve of X starting x evaluated at t. Then we define the **Lie derivative** of a vector field Y along X to be

$$(\mathcal{L}_X Y)_x = \lim_{t \to 0} \frac{X_{-t*}\left(Y_{X_t(x)}\right) - Y_x}{t} = \left.\frac{d}{dt}\right|_{t=0} X_{-t*}\left(Y_{X_t(x)}\right),$$

and on one forms ω

$$(\mathcal{L}_X \omega)_x = \lim_{t \to 0} \frac{X_t^*\left(\omega_{X_t(x)}\right) - \omega_x}{t} = \left.\frac{d}{dt}\right|_{t=0} X_t^*\left(\omega_{X_t(x)}\right).$$

[1] It is customary to also require that the inversion map $inv(g) = g^{-1}$ is a smooth map defined on G, even though this can be shown to be redundant.

We can extend the definition to a tensor field $X_1 \otimes \cdots \otimes X_l \otimes \omega^1 \otimes \cdots \otimes \omega^k$ by

$$\begin{aligned}&\left(\mathcal{L}_X\left(X_1 \otimes \cdots \otimes X_l \otimes \omega^1 \otimes \cdots \otimes \omega^k\right)\right)_x\\ &= \frac{d}{dt}\bigg|_{t=0} X_{-t*}\left(\left(X_1 \otimes \cdots \otimes X_l\right)_{X_t(x)}\right) \otimes X_t^*\left(\left(\omega^1 \otimes \cdots \otimes \omega^k\right)_{X_t(x)}\right).\end{aligned}$$

Let $\iota_X : \Omega^m(M) \to \Omega^{m-1}(M)$ denote the map from m-forms to $m-1$-forms defined by $\iota_X \omega = \omega(X, \cdot, \cdots, \cdot)$ which is called the interior multiplication of an m-form. The next proposition is proved in for instance [War83] and partially in [Lee13].

Proposition B.1. *The Lie derivative satisfies the following properties*

a) if $f \in C^\infty(M)$ then $\mathcal{L}_X(f) = X(f)$.

b) if Y is a vector field then $\mathcal{L}_X Y = [X, Y]$.

c) the Lie derivative is a derivation on forms, and commutes with the exterior derivative.

d) on forms we have that $\mathcal{L}_X = \iota_X \circ d + d \circ \iota_X$.

e) if $\omega \in \Omega^n(M)$ and $X_1, \cdots, X_n$ are vector fields, then

$$(\mathcal{L}_X \omega)(X_1, \cdots, X_n) = \mathcal{L}_X(\omega(X_1, \cdots, X_n)) - \sum_{i=1}^n \omega(X_1, \cdots, \mathcal{L}_X X_i, \cdots, X_n),$$

additionally, we have that the differential satisfies

$$\begin{aligned}d(\omega)(X_1, \ldots, X_{n+1}) = &\sum_{i=1}^n (-1)^{i-1} X_i\left(\omega\left(X_1, \ldots, \hat{X}_i, \ldots, X_{n+1}\right)\right)\\ &+ \sum_{i<j} (-1)^{i+j} \omega\left([X_i, X_j], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{n+1}\right).\end{aligned}$$

B.3 Integration on Manifolds

Let M denote a smooth oriented manifold with or without boundary of dimension n. Furthermore, let ω be a compactly supported n-form with support in U, and let (x, U) be a chart on M. Then we can define the integral of ω to be

$$\int_M \omega = \int_{x(U)} \operatorname{sgn}(x_i)\left(x^{-1}\right)^* \omega,$$

where $\operatorname{sgn}(x_i)$ is 1 if the chart is positively oriented, and -1 if the chart is negatively oriented. When ω is an arbitrary n-form we define the integral of ω to be

$$\int_M \omega = \sum_{i \in \mathcal{J}} \int_{x_i(U_i)} \left(x_i^{-1}\right)^* (\varphi_i \omega),$$

where $\{\varphi_i\}_{i \in \mathcal{J}}$ is some partition of unity subordinated to the atlas $\{(x_i, U_i)\}_{i \in \mathcal{J}}$. It can be showed that the integral does not depend on the choice of atlas nor the choice of partition of unity. The proof of this fact together with the next theorem can be found in e.g. [Lee13].

Theorem B.2 (Stokes Theorem)**.** *Let M be an oriented manifold with or without boundary of dimension n and $\eta \in \Omega^{n-1}(M)$ be a compactly supported $(n-1)$-form, then $\int_M d\eta = \int_{\partial M} \eta$.*

When M is an orientable manifold without boundary, then we have that $\int_M d\eta = 0$.

If we fix an n-form ω then we can define the integral of a function $f \in C^\infty(M)$ by $\int_M f\omega$. If we additionally require ω to be non-degenerate and $\int_M \omega > 0$ then we can define an inner product by $\langle f, g\rangle_\omega = \int_M fg\omega$, whenever it is finite. In this case, we can define the $L^2(\omega)$ space either by completing the space of smooth functions with finite norm, or use ω to define a measure and then construct the L^2 space with respect to this measure.

The existence of a non-vanishing n-form is equivalent to our manifold is orientable. It is, however, another way to define integration of functions without requiring the manifold to be orientable. Let V denote an n-dimensional vector space. Then a **density** on V is a function $\mu : V^n \to \mathbf{R}$ satisfying that for every linear function $A : V \to V$ we have that $\mu(Av_1, \cdots, Av_n) = |\det(A)|\mu(v_1, \cdots, v_n)$. Note that $\mu \neq 0$ is never linear since if $b_1, \cdots, b_n$ is a basis for V then $\mu(b_1, \cdots, b_n) = \mu(-b_1, \cdots, b_n)$. Denote by $D(V)$ the space of all densities over V. Then it is easily showed that $D(V)$ is a vector space, and that it is in fact 1 dimensional. If we are given any element ω in $\bigwedge^n V$, then it defines an element $|\omega|$ in $D(V)$ by

$$|\omega|(v_1, \cdots, v_n) = |\omega(v_1, \cdots, v_n)|.$$

If M is an n-dimensional manifold, we can construct its density bundle by $DM = \bigcup_{p\in M} D(T_pM)$. It can be showed, see e.g. [Lee13], that DM is a bundle of rank one isomorphic to the trivial bundle. Hence there exists global non-vanishing sections of DM. Any section of DM is called a **density**. Let $\mu \in D_pN$ and let $F : M \to N$ be a smooth map between n-dimensional manifolds. Then we define the pullback of μ by

$$F^*(\mu)(v_1, \cdots, v_n) = \mu(F_*v_1, \cdots, F_*v_n),$$

where $v_1, \cdots, v_n \in T_pM$. We define on $\mathbf{R}^n$

$$\int_D f|dx^1 \wedge \cdots \wedge dx^n| = \int_D f dx^1 \cdots dx^n.$$

Given a density μ on a manifold M we define the integral of μ to be

$$\int_M \mu = \sum_{i\in\mathcal{J}} \int_{x_i(U_i)} \left(x_i^{-1}\right)^* (\varphi_i\mu),$$

where $\{\varphi_i\}_{i\in\mathcal{J}}$ is some partition of unity subordinated to the atlas $\{(x_i, U_i)\}_{i\in\mathcal{J}}$. Again, it can be showed that the integral does not depend on the choice of atlas nor the choice of partition of unity.

When we are constructing the L^2-spaces we will need that the density we have chosen is positive, i.e. $\mu(v_1, \cdots, v_n) > 0$ for $v_1, \cdots, v_n$ independent. Since we can always chose a positive non-vanishing density, we can define the L^2-product to be $\langle f, g\rangle_\mu = \int_M fg\mu$. As before we can define the $L^2(\mu)$ space either by completing $C^\infty(M)$, or use μ to define a measure and then construct the L^2 space with respect to this measure. In Section B.4 we will use the following lemma.

Lemma B.3. *Let M be a manifold without boundary. Assuming μ is a positive density, we have that*

$$\int_M \mathcal{L}_X(\mu) = 0,$$

where locally $\mathcal{L}_X\mu = |\mathcal{L}_X\omega|$ for $\mu = |\omega|$.

Proof. Let $\{(x_i, \mathcal{U}_i)\}_{i\in\mathcal{J}}$ be an atlas and $\{\varphi_i\}_{i\in\mathcal{J}}$ a partition of unity subordinated to the chosen atlas. Then we have that

$$\begin{aligned}\int_M \mathcal{L}_X(\mu) &= \sum_{i\in\mathcal{J}} \int_{x_i^{-1}(\mathcal{U}_i)} x_i^*\varphi_i\,(d\iota_X\varphi_i\mu) = \sum_{i\in\mathcal{J}} \int_{x_i^{-1}(\mathcal{U}_i)} x_i^* d\iota_X\varphi_i\omega\\ &= \sum_{i\in\mathcal{J}} \int_{x_i^{-1}(\mathcal{U}_i)} dx_i^*\iota_X\varphi_i\omega = 0,\end{aligned}$$

where the last equality is obtained by using Stokes theorem in $\mathbf{R}^n$. $\square$

Let ω be a non-vanishing n-form. Then the **divergence with respect to** ω, denoted $\operatorname{div}_\omega : \Gamma(TM) \to C^\infty(M)$, is defined by $\operatorname{div}_\omega X \cdot \omega = d(\iota_X\omega) = \mathcal{L}_X\omega$. Since the divergence does not depend on orientation, which can be seen by replacing ω with $-\omega$ in the definition, we can define the divergence on non-vanishing densities as well. Hence if $\mu = \pm|\omega|$ in an open set U then the divergence is defined by $\operatorname{div}_\mu|_U = \operatorname{div}_\omega|_U$. The divergence satisfy the following product rule

$$\operatorname{div}_\mu(\varphi X) = X(\varphi) + \varphi\operatorname{div}_\mu(X).$$

B.3.1 Riemannian Density

If we are given a Riemannian manifold $(M, \mathbf{g})$, see Section 2 for definition of Riemannian manifold, we can construct a volume density by using the metric. The Riemannian volume density $\operatorname{vol}_\mathbf{g}$ is defined by the property that for any orthonormal frame $E_1, \cdots, E_n$ we have that $\operatorname{vol}_\mathbf{g}(E_1, \cdots, E_n) = 1$. In the case when M is orientable we have that $\operatorname{vol}_\mathbf{g}$ becomes an non-vanishing n-form.

The **divergence**, denoted $\operatorname{div}_\mathbf{g}$, is the divergence with respect to $\operatorname{vol}_\mathbf{g}$. We can locally write the divergence in the form

$$\operatorname{div}_\mathbf{g} X = \frac{1}{\sqrt{|\det(g_{ij})|}}\partial_i\left(X^i\sqrt{|\det(g_{ij})|}\right).$$

Moreover, the divergence is linear with respect to $\mathbf{R}$ and satisfy the formula

$$\operatorname{div}_\mathbf{g}(\varphi X) = \varphi\operatorname{div}_\mathbf{g} X + \langle \operatorname{grad}_\mathbf{g}\varphi, X\rangle_\mathbf{g}.$$

Theorem B.4 (Divergence Theorem)**.** *Let $(M, \mathbf{g})$ be a manifold with boundary, and let N be the outward pointing normal. Then for any compactly supported smooth vector field X we have that*

$$\int_M \operatorname{div}_\mathbf{g}(X)\, d\operatorname{vol}_\mathbf{g} = \int_{\partial M} \langle X, N\rangle_\mathbf{g}\, d\operatorname{vol}_{\tilde{\mathbf{g}}},$$

where $\operatorname{vol}_{\tilde{\mathbf{g}}}$ is the induced density on the boundary.

Given a Riemannian metric we can define an L^2-product on all the tensor bundles. If we are given two (k,l) tensor fields F and G then we define the L^2-inner product to be given by

$$\langle F, G\rangle_{L^2\left(\mathbf{g}_l^k\right)} = \int_M \langle F, G\rangle_{\mathbf{g}_l^k} \omega_{\mathbf{g}},$$

and the corresponding L^2-spaces will be denoted by $L^2_{\mathbf{g}_l^k}(M)$.

B.4 Second Order Differential Operators on Manifolds

Define a second order operator on a manifold to be a linear operator $L : C^\infty(M) \to C^\infty(M)$ which can locally be written at the form (A.1), where $c = 0$, a^{ij}, b^i are smooth and $a^{ij} = a^{ji}$. In this section we will assume that M is a manifold without boundary. The **symbol** of L is defined by

$$\sigma_L(df, dg) = \frac{1}{2}\left(L(fg) - fL(g) - gL(f)\right),$$

which is a section of the symmetric bundle over TM.

Proposition B.5. *If L and $\tilde{L}$ are two second order differential operators with the same symbol, then their difference is a vector field.*

Proof. The operator $\tilde{L} - L$ is a vector field if it sends smooth functions to smooth functions, are linear over $\mathbf{R}$ and satisfy the Leibnitz rule. Since $L, \tilde{L}$ are second order linear operators, the two first properties are trivial. Hence we only need to check that their difference satisfy the Leibnitz rule.

$$\begin{aligned}
\left(\tilde{L} - L\right)(\varphi\xi) &= \tilde{L}(\varphi\xi) - L(\varphi\xi) \\
&= 2\sigma_{\tilde{L}}(d\varphi, d\xi) + \varphi\tilde{L}(\xi) + \xi\tilde{L}(\varphi) - 2\sigma_L(d\varphi, d\xi) - \varphi L(\xi) - \varphi L(\xi) \\
&= \varphi\left(\tilde{L} - L\right)(\xi) + \xi\left(\tilde{L} - L\right)(\varphi).
\end{aligned}$$

□

Let μ denote an arbitrary positive density on the manifold M, and let σ be an section of the symmetric bundle over TM. Then we will use the notation $\langle \alpha, \beta\rangle_{L^2(\sigma)} = \int_M \sigma(\alpha, \beta)\, d\mu$. Note that, unless σ is positive definite, this does not create an inner-product space, where α and β are one forms.

Proposition B.6. *If two second order differential operators L and $\tilde{L}$ have the same symbol σ and they are symmetric with respect to μ, i.e.*

$$\int_M (Lf)\, g d\mu = \int_M f\,(Lg)\, d\mu,$$

then $L = \tilde{L}$.

Proof. Using the definition of symbol we get that

$$\begin{aligned}
\int_M \sigma_L(df, dg)\, d\mu &= \frac{1}{2}\int_M \left(L(fg) - fL(g) - gL(f)\right) d\mu \\
&= \frac{1}{2}\int_M \left(-gL(f) - gL(f)\right) d\mu \\
&= -\int_M gL(f)\, d\mu,
\end{aligned}$$

Hence for all g we have that $\int_M g\left(L-\tilde{L}\right) f d\mu = 0$, which imply that $Lf = \tilde{L}f$. $\square$

One important formula we obtain from the proof is

$$\int_M gL(f)\, d\mu = -\int_M \sigma(df, dg)\, d\mu.$$

In the special case where L is the Laplace operator this formula is known as the Green's formula. If σ be a section of the symmetric bundle over TM, then we will define $\sharp^\sigma(\alpha) = \sigma(\alpha, \cdot)$, where α is a one form.

Proposition B.7. *There exists a unique second order operator L^σ which has symbol σ and is symmetric with respect to μ. This operator is given by the formula $L^\sigma = \mathrm{div}_\mu \sharp^\sigma df$.*

Proof. We have that for all vector fields X

$$\begin{aligned}
\int_M \mathrm{div}_\mu X \cdot \varphi d\mu &= \int_M \varphi \mathcal{L}_X \mu = \int_M \mathcal{L}_X(\varphi\mu) - \int_M (\mathcal{L}_X \varphi)\, d\mu \\
&= -\int_M (\mathcal{L}_X \varphi)\, d\mu = -\int_M X(\varphi)\, d\mu = -\int_M d\varphi(X)\, d\mu \\
&= -\int_M X(d\varphi)\, d\mu,
\end{aligned}$$

where $\int_M \mathcal{L}_X(\varphi\mu) = 0$ by Lemma B.3. Hence we have that

$$\begin{aligned}
\int_M (\mathrm{div}_\mu \sharp^\sigma df)\, \varphi d\mu &= -\int_M \sharp^\sigma df\,(d\varphi)\, d\mu = -\int_M \sigma(df, d\varphi)\, d\mu \\
&= -\int_M \sigma(d\varphi, df)\, d\mu = \int_M (\mathrm{div}\sharp^\sigma d\varphi)\, f d\mu,
\end{aligned}$$

and L^σ becomes symmetric.

We are only left to check that the symbol of L^σ is σ. Computing the symbol of L^σ directly reveals

$$\begin{aligned}
\frac{1}{2}\left(\mathrm{div}_\mu \sharp^\sigma d(f\varphi) - fL^\sigma(\varphi) - \varphi L^\sigma(f)\right) &= \frac{1}{2}\left(\mathrm{div}_\mu(f\sharp^\sigma d\varphi) + \mathrm{div}_\mu(\varphi\sharp^\sigma df) - fL^\sigma(\varphi) - \varphi L^\sigma(f)\right) \\
&= \frac{1}{2}(d\varphi(\sharp^\sigma df) + df(\sharp^\sigma d\varphi) + f\mathrm{div}_\mu(\sharp^\sigma d\varphi) + \varphi\mathrm{div}_\mu(\sharp^\sigma df) \\
&\quad - fL^\sigma(\varphi) - \varphi L^\sigma(f)) \\
&= \sigma(d\varphi, df),
\end{aligned}$$

by using the product rule for the divergence. $\square$

References

[ABGR09] Andrei Agrachev, Ugo Boscain, Jean-Paul Gauthier, and Francesco Rossi. The intrinsic hypoelliptic Laplacian and its heat kernel on unimodular Lie groups. *J. Funct. Anal.*, 256(8):2621–2655, 2009.

[BK16] Fabrice Baudoin and Bumsik Kim. The Lichnerowicz-Obata theorem on sub-Riemannian manifolds with transverse symmetries. *J. Geom. Anal.*, 26(1):156–170, 2016.

[BKW16] Fabrice Baudoin, Bumsik Kim, and Jing Wang. Transverse Weitzenböck formulas and curvature dimension inequalities on Riemannian foliations with totally geodesic leaves. *Comm. Anal. Geom.*, 24(5):913–937, 2016.

[Cha84] Isaac Chavel. *Eigenvalues in Riemannian geometry*, volume 115 of *Pure and Applied Mathematics*. Academic Press, Inc., Orlando, FL, 1984. Including a chapter by Burton Randol, With an appendix by Jozef Dodziuk.

[Che75] Shiu Yuen Cheng. Eigenvalue comparison theorems and its geometric applications. *Math. Z.*, 143(3):289–297, 1975.

[Cho39] Wei-Liang Chow. Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung. *Math. Ann.*, 117:98–105, 1939.

[Eva10] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[FP83] C. Fefferman and D. H. Phong. Subelliptic eigenvalue problems. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser., pages 590–606. Wadsworth, Belmont, CA, 1983.

[GT16a] E. Grong and A. Thalmaier. Stochastic completeness and gradient representations for sub-Riemannian manifolds. *ArXiv e-prints, 1605.00785*, May 2016.

[GT16b] Erlend Grong and Anton Thalmaier. Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: part I. *Math. Z.*, 282(1-2):99–130, 2016.

[GT16c] Erlend Grong and Anton Thalmaier. Curvature-dimension inequalities on sub-Riemannian manifolds obtained from Riemannian foliations: part II. *Math. Z.*, 282(1-2):131–164, 2016.

[Hör67] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[Jos11] Jürgen Jost. *Riemannian geometry and geometric analysis*. Universitext. Springer, Heidelberg, sixth edition, 2011.

[Kre89] Erwin Kreyszig. *Introductory functional analysis with applications*. Wiley Classics Library. John Wiley & Sons, Inc., New York, 1989.

[Lee97] John M. Lee. *Riemannian manifolds*, volume 176 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. An introduction to curvature.

[Lee13] John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

[Lic58] André Lichnerowicz. *Géométrie des groupes de transformations*. Travaux et Recherches Mathématiques, III. Dunod, Paris, 1958.

[Mon02] Richard Montgomery. *A tour of subriemannian geometries, their geodesics and applications*, volume 91 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

[Oba72] Morio Obata. The conjectures on conformal transformations of Riemannian manifolds. *J. Differential Geometry*, 6:247–258, 1971/72.

[Ras38] P. K. Rashevskiĭ. On the connectability of two arbitrary points of a totally nonholonomic space by an admissible curve. *Uchen. Zap. Mosk. Ped. Inst. Ser. Fiz.-Mat. Nauk*, 3(2):83–94, 1938.

[RS75] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Str83] Robert S. Strichartz. Analysis of the Laplacian on the complete Riemannian manifold. *J. Funct. Anal.*, 52(1):48–79, 1983.

[Str86] Robert S. Strichartz. Sub-Riemannian geometry. *J. Differential Geom.*, 24(2):221–263, 1986.

[Tay11] Michael E. Taylor. *Partial differential equations II. Qualitative studies of linear equations*, volume 116 of *Applied Mathematical Sciences*. Springer, New York, second edition, 2011.

[War83] Frank W. Warner. *Foundations of differentiable manifolds and Lie groups*, volume 94 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1983. Corrected reprint of the 1971 edition.

[Xia13] Changyu Xia. *Eigenvalues on Riemannian manifolds*. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2013. 29o Colóquio Brasileiro de Matemática. [29th Brazilian Mathematics Colloquium].

[ZY84] Jia Qing Zhong and Hong Cang Yang. On the estimate of the first eigenvalue of a compact Riemannian manifold. *Sci. Sinica Ser. A*, 27(12):1265–1273, 1984.