Gromov–Hausdorff limits of compact Heisenberg manifolds with sub-Riemannian metrics

Kenshiro Tashiro

Abstract

In this dissertation, we study a moduli space of left invariant metrics on compact Heisenberg manifolds. It is an analogy of the classical moduli space of flat metrics on tori. A key idea is a new volume form on compact Heisenberg manifolds, called a minimal Popp's volume form, which is continuous under the canonical topology of the moduli space.

In this setting, we show a version of Mahler's compactness theorem for compact Heisenberg manifolds (Theorem 6.1). To be precise, a subspace of the moduli space is precompact in the canonical topology if total measure with respect to the minimal Popp's volume is uniformly bounded above and the systole is uniformly bounded below. Moreover we show that the canonical topology of the moduli space coincides with the Gromov–Hausdorff topology (Theorem 7.1). This concludes that non-collapsed Gromov–Hausdorff limits of compact Heisenberg manifolds are isometric to again compact Heisenberg manifolds.

Acknowledgement

First of all, the author would like to express his great thank to his supervisor Prof. Koji Fujiwara for drawing his attention to this theme, and for many helpful comments and suggestions. Throughout 6 years he taught not only technical contents but also how to go on as a researcher. These will be the author's asset.

The author is grateful to Prof. Enrico Le Donne for having much discussion with him on sub-Riemannian geometry and nilpotent group theory, and for hospitality during his stay at Jyväskylä.

The author thanks Prof. Takumi Yokota for giving him a chance to discuss the interaction of Riemannian and sub-Riemannian geometry. The author appreciate Prof. Ryokichi Tanaka, who taught him the classical Mahler's theorem and its background from group theory. The author thanks Prof. Shin-ichi Ohta, who taught him the regularity problem on Gromov–Hausdorff limit spaces. This work is supported by JSPS KAKENHI Grant Number JP20J13261.

Contents

1 Introduction

A sub-Riemannian manifold is a triple (M, D, g), where M is a connected smooth manifold, D is a subspace in the tangent bundle TM, and g a metric on D. Recently sub-Riemannian manifolds are actively studied from a viewpoint of optimal transport theory and optimal control theory. In this dissertation, we study Gromov–Hausdorff limits of a sequence of sub-Riemannian manifolds.

1.1 Background

1.1.1 The Gromov–Hausdorff topology

We recall backgrounds of Riemannian manifolds with a lower Ricci curvature bound.

Myers Theorem is the most fundamental fact which relates the Ricci curvature and the topology of manifolds. It asserts that if a complete n-dimensional Riemannian manifold has a positive lower bound $K > 0$ of the Ricci curvature, then it is compact and its diameter is bounded above by a constant $C = C(n, K)$ [38]. If we assume $K = 0$, then the Cheeger–Gromoll's splitting theorem holds. It asserts that if a complete Riemannian manifold with non-negative Ricci curvature has a line, then it isometrically splits to the direct product of the line and a lower dimensional Riemannian manifold. As a corollary, the fundamental group of such a manifold is almost abelian and the first Betti number is not greater than the dimension n. One of the important lemmas in its proof is now called the Laplacian comparison theorem. Laplace–Beltrami operator has a deep relationship with the Ricci curvature lower bound. For example, Bakry–Émery's curvature dimension condition gives an equivalence between the inequality (1) and the Ricci curvature lower bound.

As we can see from the Myers theorem and the Cheeger–Gromoll's splitting theorem, a lower bound of the Ricci curvature implies that a manifold is 'small'. The following Gromov's precompactness theorem generalizes this idea to a broader setting. Let $M(n, K, d)$ be a family of n-dimensional Riemannian manifolds with the Ricci curvature lower bound K and the diameter upper bound $d > 0$. He showed in [29] that $M(n, K, d)$ is precompact under the Gromov-Hausdorff topology. Here the Gromov–Hausdorff topology is defined as follows.

Definition 1.1 (Gromov–Hausdorff topology)**.** *We say that a sequence of compact metric spaces $\{(X_i, d_i)\}$ converges to a compact metric space (X_∞, d_∞) in the Gromov–Hausdorff topology if there are a sequence of (not necessarily continuous) mappings $\{f_i : X_i \to X_\infty\}$ and a sequence of positive numbers $\{\epsilon_i\}$ such that*

- $\epsilon_i \to 0$,
- $|d_i(x, y) - d_\infty(f_i(x), f_i(y))| < \epsilon_i$,
- *For all* $y \in X_\infty$*, there is* $x \in X_i$ *such that* $d_\infty(y, f_i(x_i)) < \epsilon_i$.

Remark 1.1. *We can define the Gromov–Hausdorff convergence for a sequence of non-compact proper metric space. Here a metric space is proper if every closed ball is compact. We say that a sequence of pointed proper metric spaces* $\{(X_i, d_i, p_i)\}$ *converges to* $(X_\infty, d_\infty, d_\infty)$ *in the pointed Gromov–Hausdorff topology if the metric balls* $\{B_{X_i}(x_i, R)\}$ *converges to* $B_{X_\infty}(x_\infty, R)$ *for any* $R > 0$ *in the Gromov–Hausdorff topology.*

The Gromov–Hausdorff topology means the closeness of two metric spaces ignoring their topologies and differential structure.

The following Bishop–Gromov's inequality is important in the proof of Gromov's precompactness theorem (Lemma 5.3 in [31]). Assume that a n-dimensional compact Riemannian manifold M has the Ricci curvature lower bound $-(n-1)K$. Then the ball $B_M(x, r)$ of radius $r > 0$ centered at $x \in M$ satisfies

$$\frac{vol\,(B_M(x, R))}{vol(B_M(x, R'))} \geq \frac{vol\,(B_{M_K}(x_0, R))}{vol(B_{M_K}(x_0, R'))} \quad \text{for} \quad R \leq R',$$

where M_K is the simply connected n-dimensional Riemannian manifold with a constant sectional curvature K and $x_0 \in M$. By using the Biship–Gromov's inequality, we can prove that the number of an ϵ-net in the manifold M is less than the constant $C = C(n, d, K, \epsilon)$. From the precompactness of ϵ-nets, we can show the precompactness of $M(n, -(n-1)K, d)$.

We will study the properties of spaces in $\overline{M(n, K, d)}$. Since the Gromov–Hausdorff topology represents the closeness of metric spaces, Gromov–Hausdorff limit spaces may reflect the property of the Ricci curvature lower bound. However, a limit space is generally not a manifold. Thus it is difficult to define the curvature with the differential structure on a manifold. To address this issue, Cheeger and Colding introduce the notion of rectifiability on a limit space in [21, 22, 23]. They show in [24] that the splitting theorem for a limit space holds. As an application of the splitting theorem, they studied the infinitesimal structure of a limit space and established the rectifiability. The important idea in the proof of the splitting theorem is to show the quantitive version of Laplacian comparison theorem in a limit space.

1.1.2 Curvature dimension conditions

Recently, the synthetic notion of the Ricci curvature lower bound is actively studied from a viewpoint of optimal transport theory. In this research one studies a metric measure space (X, d, m), where (X, d) is a complete separable metric space and m is a Borel measure on X. It is a generalization of normalized compact orientable Riemannian manifold $(M, g, \frac{1}{vol_g(M)} vol_g)$. In this setting, Lott–Villani and Sturm introduced the curvature dimension condition $CD(K, N)$ respectively in [34] and [43, 44]. In [39], Ohta introduced the measure contraction property $MCP(K, N)$ which is weaker than the condition $CD(K, N)$. Here the constant $K \in \mathbb{R}$ represents the Ricci curvature lower bound and the constant $N \in [1, \infty]$ represents the upper bound of a dimension. In fact, a Riemannian manifold satisfies $CD(K, N)$ if and only if the Ricci curvature is bounded below by K and the topological dimension is N. The same statement follows if we replace the condition $CD(K,N)$ with $MCP(K,N)$. Both conditions yield the Myers theorem and the Bishop–Gromov inequality on a metric measure space. Moreover, a family of metric measure spaces satisfying $CD(K,N)$ or $MCP(K,N)$ is compact under the measured Gromov–Hausdorff topology. Hence we can study Gromov–Hausdorff limits in the same setting. Here the measured Gromov–Hausdorff topology is induced from the Gromov–Hausdorff convergence and the weak convergence of measures. For more detailed information, see Definition 0.2 in [27].

Another curvature dimension condition was studied by Bakry, Émery and many other researchers. Let us explain its framework called Γ-calculus. Let (X,d,m) be a metric measure space and L a self-adjoint operator on the space of square integrable functions $L^2(X)$. We define the *carré du champ operator* $\Gamma : L^2(X) \times L^2(X) \to L^2(X)$ by

$$\Gamma(f_1, f_2) = \frac{1}{2}\left(L(f_1 f_2) - f_1 L f_2 - f_2 L f_1\right)$$

and define the *iterated carré du champ operator* $\Gamma_2 : L^2(X) \times L^2(X) \to L^2(X)$ by

$$\Gamma_2(f_1, f_2) = \frac{1}{2}\left(L\Gamma(f_1, f_2) - \Gamma(f_1, Lf_2) - \Gamma(Lf_1, f_2)\right).$$

We say that a metric measure space (X,d,m) with the operator L satisfies the Bakry–Émery's curvature dimension condition $BE(K,N)$ if

$$\Gamma_2(f,f) \geq K\Gamma(f,f) + \frac{1}{N}(Lf)^2. \tag{1}$$

In fact, a N-dimensional normalized Riemannian manifold with the extension of the Laplace–Beltrami operator satisfies $BE(K,N)$ if and only if it has the Ricci curvature lower bound by K (Proposition 6.2 in [9]). For more detailed information on the Γ-calculus, see [10].

On a metric measure space, we choose the self-adjoint operator L as follows. Let $Ch : L^2(X) \to \mathbb{R}$ be the function defined by

$$Ch(f) = \frac{1}{2}\inf\left\{ \varliminf_{j\to\infty} \int_X |\nabla f_j|^2 dm \;\middle|\; f_j \in Lip_b(X),\ f_j \to f \ \text{ in } L^2(X) \right\}.$$

This function Ch is called the *Cheeger energy functional*. We say that a metric measure space is *infinitesimally Hilbertian* if the Cheeger energy functional is a quadratic form. A metric measure space is said to satisfy $RCD(K,N)$ if it is infinitesimally Hilbertian and satisfy $CD(K,N)$. We can define the gradient flow of Ch on $L^2(X)$ in an appropriate way. Let P_t be the heat semigroup of the gradient flow and L the generator of P_t. For this operator L, $BE(K,N)$ is equivalent to $RCD(K,N)$ ([7] for $N=\infty$ and [8, 26] for $N<\infty$).

1.1.3 The sub-Riemannian curvature

We will explain backgrounds on the Ricci curvature on a sub-Riemannian manifold. Intrinsic Ricci curvature on a sub-Riemannian manifold is defined by using the techniques of optimal control theory. First of all, We recall the setting of optimal control theory. Let M be a n-dimensional manifold, $U \subset \mathbb{R}^k$ an open region, $f : M \times U \to Vec(M)$ a smooth mapping, and $\varphi : M \times U \to \mathbb{R}$ a positive continuous function. For a square integrable mapping $u : [0,T] \to M$, define the mapping $x_u : [0,T] \to M$ by the solution of the equation

$$\dot{x}_u = f_u(x), \qquad x_u(0) = x_0.$$

Define the functional J by

$$J(u) = \int_0^T \varphi(x_u(t), u(t))dt.$$

This J is called the *cost functional*. The *optimal control problem* asks what u minimizes the cost functional J under the assumption

$$x_u(0) = x_0 \quad \text{and} \quad x_u(T) = x_1.$$

We call such a u the *optimal control* and the associated x_u the *optimal trajectory*.

Example 1.1. Assume that a Riemannian manifold (M, g) has a global orthonormal frame $\{f_1, \ldots, f_n\}$. Set $U = \mathbb{R}^n$, $f(x,u) = \sum_{i=1}^n u_i f_i(x)$ for $u = (u_1, \ldots, u_n)$ and $\varphi(x,u) = g_x(f(x,u), f(x,u))$. If a mapping u is an optimal control, then the associated optimal trajectory is a length minimizing geodesic.

We can generalize the above example to the sub-Riemannian setting by letting $U = \mathbb{R}^{\dim D}$. For more detailed information on optimal control theory, please see [5].

After the pioneering results by Agrachev–Gamkrelidze [3] and Agrachev–Zelenko [4], the sub-Riemannian Ricci curvature is defined by Zelenko–Li in [49]. They define a Jacobi curve on T^*M as a generalization of Jacobi fields, and introduce the curvature along geodesics by using the coefficient function of the Jacobi equation. We do not pursue its definition in this dissertation. For a detailed information, please see [11].

The model space for the sub-Riemannian Ricci curvature is the *linear quadratic optimal control problem* (later call LQ problem). Roughly speaking, LQ problem is a optimal control problem such that $M = \mathbb{R}^n$, $U = \mathbb{R}^k$, f is linear and φ is a quadratic form. In [11], Barilari–Rizzi compared the cut time of trajectories in a sub-Riemannian manifold and that in a model LQ problem. As a corollary they gave a version of Myers Theorem.

The sub-Riemannian Ricci curvature is intrinsically defined, however it is quite difficult to compute. It is also difficult to give a comparison theorem which reflects geometric properties of sub-Riemannian manifolds. One of the reason is that LQ problems do not give a metric structure on $M = \mathbb{R}^n$.

1.1.4 Sub-Riemannian manifolds and the curvature dimension conditions

From a viewpoint of optimal transport theory, the following sub-Riemannian manifolds are known to satisfy the measure contraction property $MCP(K, N)$. In [32], Juillet showed that the Heisenberg Lie group H_n with a sub-Riemannian metric satisfies $MCP(0, 2n+3)$. In subsequent researches, many Carnot groups are shown to satisfy $MCP(0, N)$ with $N \in [1, \infty)$ in [40, 41, 12]. An example other than Carnot groups is given by Agrachev–Lee. In [6], they showed that if a 3-dimensional contact sub-Riemannian manifold has a positive sub-Riemannian Ricci curvature, then it satisfies $MCP(0, 5)$. The proof is done in the following way. For $x \in M$, let $\exp_x : T^*M \to M$ be the sub-Riemannian exponential map (Definition 1.6), and $A \subset T_x^*M$ a measurable set. We calculate the Jacobian of $\exp_x$, and integrate it on tA for $t \in [0, 1]$. Then we can check the $MCP(0, N)$ condition for the negligible cut loci case. This proof follows only if the complement of the image of the sub-Riemannian exponential map $(Im(\exp_x))^c$ has zero measure. However, it is not known whether this property holds for every sub-Riemannian manifold. In fact, the set $(Im(\exp_x))^c$ is contained in the *abnormal set* (Definition 1.7). It is an important open problem in sub-Riemannian geometry whether the abnormal set has zero measure. This is why we cannot check the $MCP(K, N)$ condition for a general sub-Riemannian manifold.

On the other hand, Juillet also showed that the Heisenberg group does not satisfy $CD(K, N)$ for all $K \in \mathbb{R}$ and $N \in [1, \infty]$ in [32]. Namely he showed that the Heisenberg group does not satisfy the geodesic Brunn–Minkowski inequality. Other examples are not published, however it is expected that all sub-Riemannian manifolds do not satisfy $MCP(K, N)$.

Sub-Riemannian version of the Bakry–Émery's curvature dimension condition $BE(K, N)$ is established by Baudoin–Garofalo. In [13], they introduced the generalized curvature dimension condition for a class of sub-Riemannian manifolds. This class contains many important sub-Riemannian manifolds such as Sasakian manifolds. They studied the Γ-calculus for the locally sub-elliptic diffusion operator L. For the carré du champ operator Γ associated to L, they assume the existence of the vertical carré du champ operator Γ^Z. The generalized curvature dimension inequality is given by the iterated carré du champ operators Γ_2 and Γ_2^Z. This condition is easier to calculate than the sub-Riemannian Ricci curvature. Under this condition, sub-Riemannian manifolds satisfy the Myers theorem (Theorem 10.1). Moreover Baudoin–Bonnefont–Garofalo–Munive showed in [14] that such manifolds also satisfy the doubing property. Here a metric measure space (X, d, m) satisfies the *doubling property* if there exists a constant $D > 0$ such that for any $x \in X$ and any $r > 0$ the ball $B_X(x, r)$ satisfies

$$m(B_X(x, 2r)) \leq Dm(B_X(x, r)).$$

This implies that a family of compact sub-Riemannian manifolds with the generalized curvature dimension inequality is precompact under the Gromov–Hausdorff topology.

1.1.5 Problems

We have explained some classes of sub-Riemannian manifolds which are precompact under the Gromov–Hausdorff topology. However, it is not known what spaces appear as Gromov–Hausdorff limit spaces. In the author's research we ask

(1) when the limit space has a differential structure,

(2) how we can explain singular points on the limit space,

(3) if the limit space has singular points, how we justify the sub-Riemannian structure.

In this dissertation we study limits of a special class of sub-Riemannian manifolds. It is the first step toward our objectives.

1.2 Preliminaries from sub-Riemannian geometry

In this section we prepare notation on sub-Riemannian geometry.

1.2.1 Sub-Riemannian structure

Let M be a connected orientable smooth manifold, $(E, \tilde{g})$ a metric vector bundle on M, and $f : E \to TM$ be a fiberwise linear smooth map. For $x \in M$, denote by D_x the image of $f|_{E_x}$. We call the collection of subspaces $D = \{D_x\}_{x \in M}$ the *distribution*. On each subspace D_x we define the inner product g_x by

$$g_x(u, v) = \inf \{\tilde{g}(U, V) | u = f(U), v = f(V)\}.$$

Definition 1.2 (Sub-Riemannian structure)**.** *A sub-Riemannian manifold is a triple (M, D, g). The pair (E, f) is called the sub-Riemannian structure on M.*

We say that a vector field on M is *horizontal* if it is a section of the distribution D.

Example 1.2. Let G be a connected Lie group, $\mathfrak{g}$ the associated Lie algebra, $V \subset \mathfrak{g}$ a subspace and $\langle \cdot, \cdot \rangle$ an inner product on V. For $x \in G$, denote by $L_x : G \to G$ the left translation by x. Define a sub-Riemannian structure on G by

$$D_x = (L_x)_* V, \qquad g_x(u, v) = \langle L_{x*}^{-1} u, L_{x*}^{-1} v \rangle.$$

Such a sub-Riemannian structure (D, g) is called left invariant.

The associated distance function is given in the same way to Riemannian distance. We say that an absolutely continuous path $c : [0, 1] \to M$ is *admissible* if $\dot{c}(t) \in D_{c(t)}$ a.e. $t \in [0, 1]$. We define the length of an admissible path by

$$\ell(c) = \int_0^1 \sqrt{g(\dot{c}(t), \dot{c}(t))} dt.$$

For $x, y \in M$, define the distance function by

$$d(x, y) = \inf \{\ell(c) \mid c(0) = x, c(1) = y, c \text{ is admissible}\}.$$

In general not every pair of points in M is joined by an admissible path. This implies that the value of the function d may be the infinity. The following *bracket generating* condition ensure that any two points are joined by an admissible path.

Definition 1.3 (Bracket generating distribution)**.** *For every $i \in \mathbb{N}$, let D^i be the submodule in $Vec(M)$ inductively defined by*

$$D^1 = D, \qquad D^{i+1} = D^i + [D, D^i],$$

and set $D^i_x = \left\{X(x) \mid X \in D^i\right\}$.

We say that a distribution D is bracket generating if for all $x \in M$ there is $r = r(x) \in \mathbb{N}$ such that $D^r_x = T_x M$.

Theorem 1.1 (Chow–Rashevskii's theorem, Theorem 3.31 in [1])**.** *Let (M, D, g) be a sub-Riemannian manifold with a bracket generating distribution. Then the following two assertions hold.*

(1) (M, d) is a metric space,

(2) the topology induced by (M, d) is equivalent to the manifold topology.

In particular, $d : M \times M \to \mathbb{R}$ is continuous.

Assume that the metric space (M, d) is proper, that is every closed ball in (M, d) is compact. With the help of the Ascoli–Alzera theorem, we can show the existence of a length minimizing path joining any two points (Theorem 3.43 in [1]).

1.2.2 Length minimizing paths

Let us consider length minimizing paths on a sub-Riemannian manifold (M, D, g).

For simplicity we assume that the dimension of D_x is equal to m for all $x \in M$ and that we have a family of globally defined m smooth vector fields $\{f_1, \dots, f_m\}$ such that $\{f_1(x), \dots, f_m(x)\}$ is an orthonormal basis of (D_x, g_x). We call such a family a *generating family.*

Let σ be the canonical symplectic form on T^*M. For a given function $h : T^*M \to \mathbb{R}$, there is a unique vector field $\vec{h}$ on T^*M defined by

$$\sigma(\cdot, \vec{h}(\lambda)) = dh.$$

We call the vector $\vec{h}$ the *Hamiltonian vector field of h.*

Let $h_i : T^*M \to \mathbb{R}$, $i = 1, \dots, m$ be the function on the cotangent bundle defined by

$$h_i(\lambda) = \langle \lambda | f_i(x) \rangle,$$

where $\langle \cdot | \cdot \rangle$ is the canonical pairing of covectors and vectors. Then length minimizing paths on a sub-Riemannian manifold are explained with the Hamiltonian vector fields of h_i's as follows.

Theorem 1.2 (The Pontryagin maximal principle, Theorem 4.20 in [1]). *Let $\gamma : [0, T] \to M$ be a length minimizing path parametrized by constant speed, and write its differential as*

$$\dot{\gamma}(t) = \sum_{i=1}^{m} u_i(t) f_i(\gamma(t)).$$

Then there is a Lipschitz curve $\lambda : [0, T] \to M$ such that

$$\begin{cases} \lambda(t) \in T^*_{\lambda(t)} M, \\ \dot{\lambda}(t) = \sum_{i=1}^{m} u_i \vec{h}_i(\lambda(t)) \qquad a.e.\ t \in [0, T], \end{cases} \tag{2}$$

and one of the following conditions satisfied:

(N) $h_i(\lambda(t)) = u_i(t), \quad i = 1, \ldots, m, \quad t \in [0, 1],$

(A) $h_i(\lambda(t)) = 0, \quad i = 1, \ldots, m.$

Definition 1.4 (Normal extremal). *The Lipschitz curve λ with the condition (N) is called a normal extremal, and its projection γ is called a normal trajectory.*

Definition 1.5 (Abnormal extremal). *The Lipschitz curve λ with the condition (N) is called an abnormal extremal, and its projection γ is called an abnormal trajectory.*

Let $D_x^{\perp} \subset T_x^* M$ be the subspace defined by

$$D_x^{\perp} = \{ \lambda \in T_x^* M \mid \langle \lambda | u \rangle = 0 \ \text{ for all } u \in D_x \}.$$

With this notation, we can say that an extremal $\lambda(t) = (x(t), p(t))$ is abnormal if $\lambda(t) \in D_{x(t)}^{\perp}$.

Example 1.3. Suppose that (M, D, g) is a Riemannian manifold. Then the condition (A) implies that $\lambda(t) \in D_{x(t)}^{\perp} = \{0\}$. Combined with the second equality in (2), such a Lipschitz curve λ is a constant curve. This argument shows that every abnormal trajectories on a Riemannian manifold is a constant curve.

Remark 1.2. *An abnormal extremal and a normal extremal may project to the same trajectory. Hence a trajectory may be normal and abnormal simultaneously.*

1.2.3 Normal extremals

Normal extremals are characterized by a solution of the differential equation, called the Hamiltonian system. Let $H : T^*M \to \mathbb{R}$ be the function defined by

$$H(\lambda) = \frac{1}{2}\sum_{i=1}^{m} h_i(\lambda)^2.$$

This function is called the *sub-Riemannian Hamiltonian.*

Remark 1.3. *If a manifold (M, D, g) is Riemannian, then we have the canonical metric g^* on T^*M induced by the musical isomorphism. The Riemannian Hamitonian is defined by using this metric.*

However, sub-Riemannian manifolds have no canonical isomorphism between the tangent bundle and cotangent bundle. Hence we use the canonical pairing of covectors and vectors h_i's instead of metrics.

Theorem 1.3 (Theorem 4.25 in [1])**.** *A Lipschitz curve $\lambda : [0, T] \to T^*M$ is a normal extremal if and only if it is a solution of the Hamiltonian system*

$$\dot{\lambda}(t) = \vec{H}(\lambda(t)), \qquad t \in [0, T],$$

where $\vec{H}$ is the Hamiltonian vector field of H.

Moreover, the corresponding normal trajectory γ is smooth and has a constant speed satisfying

$$\frac{1}{2}\|\dot{\gamma}(t)\|^2 = H(\lambda(t)).$$

A straightforward computation shows the following local expression for $\lambda(t) = (x(t), p(t))$ by

$$\begin{cases} \dot{x}(t) = \frac{\partial H}{\partial p}, \\ \dot{p}(t) = -\frac{\partial H}{\partial x}. \end{cases}$$

Theorem 1.3 asserts that a normal extremal λ is the orbit of the Hamiltonian flow with the initial data $(x, p) \in T^*M$. Hence we can define the exponential map of a sub-Riemannian manifold.

Definition 1.6 (Sub-Riemannian exponential map)**.** *The exponential map exp_x : $T_x^*M \to M$ is defined by*

$$\exp_x(p) = \gamma_p(1),$$

where γ_p is the normal trajectory associated to the initial data $(x, p) \in T^*M$.

Remark 1.4. *In Riemannian geometry, the domain of the exponential map* $\exp_x$ *is the tangent T_xM. We cannot define the sub-Riemannian exponential map with this domain. In fact, we always have two normal trajectories γ_1, γ_2 : $[0, T] \to M$ such that $\gamma_1(0) = \gamma_2(0)$, $\dot{\gamma}_1(0) = \dot{\gamma}_2(0)$ and $\gamma_1(t) \neq \gamma_2(t)$ for all $t > 0$. Hence the initial data of the direction do not determine the normal trajectories.*

Theorem 1.2 gives a necessary condition for an admissible path being length minimizing. In fact every normal trajectory is a geodesic. Here we say that an admissible path is a *geodesic* if it is constant speed and for every $t \in [0, T]$, there is a neighborhood I of t in $[0, T]$ such that $\ell(\gamma|_I)$ is equal to the distance between its endpoints.

Theorem 1.4 (Theorem 4.64 in [1])**.** *Let $\gamma : [0, T] \to M$ be a sub-Riemannian normal trajectory. Then for every $\tau \in [0, T)$ there is ϵ_0 such that for $0 < \epsilon < \epsilon_0$*

(1) $\gamma|_{[\tau,\tau+\epsilon]}$ is a length minimizing path,

(2) $\gamma|_{[\tau,\tau+\epsilon]}$ is the unique length minimizing path joining $\gamma(\tau)$ and $\gamma(\tau+\epsilon)$ up to reparametrization.

1.2.4 Abnormal extremals

The computation of abnormal extremals is quite hard. For example, we cannot characterize an abnormal extremal as a solution of ordinary differential equation, thus it is difficult to determine the regularity of abnormal trajectories. In [36], Montgomery showed that there exists a length minimizing abnormal trajectory which is not normal. Moreover, it is an open problem that the following *abnormal set* has measure zero.

Definition 1.7 (Abnormal set)**.** *For $x \in M$, let $Abn(x)$ be the set of endpoints of abnormal trajectories issuing from x. This is called the abnormal set of x.*

Hence we cannot ignore abnormal trajectories in general setting.

However, one has interesting classes of sub-Riemannian manifolds which has no abnormal trajectories.

Definition 1.8 (Fat distribution)**.** *We say that a distribution D is fat if for any $x \in M$ and any horizontal vector field X with $X(x) \neq 0$,*

$$T_xM = D_x + \{[X, Y]_x \mid Y : \text{ horizontal vector field}\}.$$

For example, a contact manifold (M, θ) with the distribution $D = Ker\theta$ is a fat distribution. For more detailed information on a fat distribution, see Section 5.6 of [37].

1.2.5 The Popp's volume

On a Riemannian manifold, one has a canonical Riemannian volume form defined by

$$v = \nu_1 \wedge \cdots \wedge \nu_n,$$

where $\nu_1, \cdots \nu_n$ is a dual coframe of an orthonormal basis. In sub-Riemannian geometry, we also have a canonical volume form, called *Popp's volume* introduced in [37]. It is defined under the following assumption.

Definition 1.9 (Equiregular distribution)**.** *A sub-Riemannian manifold* (M, D, g) *is equiregular if for any* $i \in N$ *the dimension of the subspace* D_x^i *is independent of the choice of* $x \in M$*.*

If $D_x^r = T_x M$, we say that a sub-Riemannian manifold is r-step. For simplicity, we define the Popp's volume in the 2-step case.

Definition 1.10 (Nilpotentization)**.** *The nilpotentization of* D *at the point* $x \in M$ *is the graded vector space*

$$gr_x(D) = D_x \oplus D_x^2 / D_x.$$

On the vector space $gr_x(D)$ we can define a new Lie bracket $[\cdot, \cdot]'$ by

$$[X \ mod \ D, Y \ mod \ D]' = [X, Y] \quad \mathrm{mod} \ \ D.$$

The new Lie bracket rule induces a different Lie algebra structure from the original one.

From the inner product on D_x, we obtain the inner product on the nilpotentization $gr_x(D)$ of D. Let $\pi : D_x \otimes D_x \to D_x^2 / D_x$ be a linear map given by

$$\pi(u \otimes v) = [U, V]_x \ \ mod \ \ D_x,$$

where U, V are horizontal extensions of u, v. Define the norm $\| \cdot \|_2$ on D_x^2 / D_x by

$$\|z\|_2 = \min \left\{ \|U(x)\| \|V(x)\| \mid [U, V]_x = z \ mod \ D, \ U, V : \text{ horizontal vector fields} \right\}.$$

This norm satisfies the parallelogram law, thus we obtain the inner product $\langle \cdot, \cdot \rangle_2$ on D_x^2 / D_x. By adding the original inner product on the distribution D, we obtain the new inner product $\langle \cdot, \cdot \rangle_x'$ on the nilpotentization $gr_x(D)$.

Let $\omega_x \in \wedge^n gr_x(D)^*$ be the volume form obtained by wedging the elements of orthonormal dual basis in $(gr_x(D), \langle \cdot, \cdot \rangle_x')$. It is defined up to sign.

By the following lemma, the volume $\omega_x \in \wedge^n gr_x(D)^*$ is transported to the volume on $\wedge^n T_x^* M$.

Lemma 1.1 (Lemma 10.4 in [37])**.** *Let* E *be a vector space of dimension* n *with a filtration by linear subspaces* $F_1 \subset F_2 \subset \cdots F_l = E$*. Let* $Gr(F) = F_1 \oplus F_2/F_1 \oplus \cdots F_l / F_{l-1}$ *be the associated graded vector space. Then there is a canonical isomorphism* $\theta : \wedge^n E^* \simeq \wedge^n gr(F)^*$*.*

Let $\theta : \wedge^n T_x^* M \to \wedge^n gr_x(D)^*$ be the isomorphism obtained by Lemma 1.1.

Definition 1.11 (Popp's volume)**.** *The Popp's volume form* P *of* (M, D, g) *is defined by*

$$P_x = \theta^* \omega_x, \qquad x \in M.$$

Trivially the Popp's volume of a Riemannian manifold is the canonical Riemannian volume form.

The Popp's volume has a useful expression via the structure constant. We say that a local frame $X_1, \dots, X_n$ is adapted if $X_1, \dots, X_m$ are orthonormal. Define the smooth functions on M by

$$[X_i, X_j] = \sum_{l=1}^{n} c_{ij}^l X_l.$$

We call them *structure constants*. We define the $n - m$ dimensional square matrix B by

$$B_{hl} = \sum_{i,j=1}^{m} c_{ij}^h c_{ij}^l.$$

Theorem 1.5 (Theorem 20.6 in [1]). *Let $X_1, \dots, X_n$ be a local adapted frame, and $\nu^1, \dots, \nu^n$ the dual coframe. Then the Popp's volume P satisfies*

$$P = (\det B)^{-\frac{1}{2}} \, \nu^1 \wedge \cdots \wedge \nu^n.$$

Example 1.4 (The Heisenberg Lie group). Let H_n be the Lie group diffeomorphic to the $2n+1$-dimensional Euclidean space with the law of group operation

$$(x_1, \dots, x_{2n}, z)(y_1, \dots, y_{2n}, w) = \left(x_1 + y_1, \dots, x_{2n} + y_{2n}, z + w + \frac{1}{2} \sum_{i=1}^{n} (x_i y_{i+n} - x_{i+n} y_i) \right).$$

The group H_n is called the n-Heisenberg Lie group.

Let $\mathfrak{h}_n$ be the associated Lie algebra. We have a canonical basis $\{X_1, \dots, X_{2n}, Z\}$ of $\mathfrak{h}_n$ determined by

$$X_i(e) = \frac{\partial}{\partial x_i}, \quad (i = 1, \dots, 2n) \quad \text{and} \quad Z(e) = \frac{\partial}{\partial z}.$$

A straightforward computation shows that $[X_i, X_{i+n}] = Z$ for all $i = 1, \dots, n$ and the other brackets are zero.

Let $\exp : \mathfrak{h}_n \to H_n$ be the exponential map. It is well known that the exponential map is a diffeomorphism. The exponential map sends $X_1, \dots, Z \in \mathfrak{h}_n$ to $e_1, \dots, e_{2n+1} \in H_n$, where $e_1, \dots, e_{2n+1}$ is the canonical basis in the above coordinates. The Campbell–Baker–Hausdorff formula for the Heissenberg Lie group asserts

$$\exp(X) \cdot \exp(Y) = \exp(X + Y + \frac{1}{2}[X, Y]), \qquad X, Y \in \mathfrak{h}_n.$$

In particular we obtain

$$[\exp(X), \exp(Y)]_c = \exp([X, Y]),$$

where the bracket in the left hand side is the commutator. Hence we can identify the Heisenberg Lie group H_n to its Lie algebra $\mathfrak{h}_n$ via the exponential map.

Let V be a vector subspace in the Lie algebra $\mathfrak{h}_n$. The left translation of the subspace V induces a left invariant distribution D_V on H_n. From the Lie bracket rule given above, the distribution D_V is bracket generating if and only if $\pi : V \to V_0 = Span\{X_1, \dots, X_{2n}\}$ is surjective.

Moreover if D_V is bracket generating, then it is a fat distribution. Hence all length minimizing paths are normal trajectories. Explicit formula of normal trajectories on the Heisenberg Lie group is given in Section 4.

From the left invariance of the distribution, it is also equiregular. Thus we can define the Popp's volume form. Explicit formula of the Popp's volume on the Heisenberg Lie group is given in Section 5.

1.3 Main results

1.3.1 Moduli spaces and precompact theorems

Let N be a simply connected nilpotent Lie group, and Γ a lattice in N. We call the quotient space $\Gamma\backslash N$ a *compact nilmanifold*. We say that a metric d on a compact nilmanifold $\Gamma\backslash H_n$ is left invariant if the pullback metric $\tilde{d}$ on N is left invariant, that is $\tilde{d}(gx, gy) = \tilde{d}(x, y)$ for all $g, x, y \in N$.

Definition 1.12 (Moduli space)**.** *We denote by $\mathcal{M}(\Gamma\backslash N)$ the set of left invariant sub-Riemannian metrics on $\Gamma\backslash N$. We call it the moduli space of left invariant sub-Riemannian metrics on $\Gamma\backslash N$.*

As we will see in Definition 3.2, we can endow the canonical topology in $\mathcal{M}(\Gamma\backslash N)$, called the quotient topology.

In this dissertation, we study left invariant sub-Riemannian metrics on a compact Heisenberg manifold, which is a quotient space of the Heisenberg Lie group $\Gamma\backslash H_n$.

In his doctorial thesis [18] (or [17] in arXiv), Boldt showed that a family of compact Heisenberg manifolds with left invariant Riemannian metrics is precompact in the quotient topology of the moduli space under the assumptions on metric tensors (Theorem 6.2), where the moduli space is the one given by Gordon and Wilson in Remark 2.6 of [28]. For more detailed information, see Section 6.

We can regard his result as a variation of the Mahler's compactness theorem (Theorem 2.4). The Mahler's compactness theorem asserts that a family of flat tori is precompact in the quotient topology of the moduli space if the systoles are bounded below by $s > 0$ and the total measure is bounded above by $v > 0$. In fact, recall that a compact Heisenberg manifold has a fiber bundle structure $S^1 \to \Gamma\backslash H_n \to T^{2n}$. Two of the Boldt's assumptions on metric tensors coincide with those of Mahler for the base flat tori.

We improve the Boldt's result to the sub-Riemannian setting. We fix a compact Heisenberg manifold $\Gamma\backslash H_n$ in the following.

Theorem 1.6 (The former part of Theorem 1.1 in [45])**.** *For positive numbers $v, s > 0$, let $\mathcal{A}(v, s) \subset \mathcal{M}(\Gamma\backslash H_n)$ be the set of compact Heisenberg manifolds endowed with left invariant sub-Riemannian metrics of corank* 0 *or* 1 *such that*

- *the minimal Popp's volumes (Definition 5.2) are bounded above by v, and*
- *the systoles are bounded below by s.*

Then $\mathcal{A}(v, s)$ is precompact in the quotient topology of the moduli space $\mathcal{M}(\Gamma\backslash H_n)$.

Here we denote by the corank of a sub-Riemannian metric the codimension of the distributions in the tangent spaces. The Riemannian metrics are the corank 0 sub-Riemannian metrics.

By using Theorem 1.6, we compute a Gromov–Hausdorff limit of a sequence of compact Heisenberg manifolds as follows.

Theorem 1.7 (The latter part of Theorem 1.1 in [45])**.** *The quotient topology of the moduli space coincides with the Gromov–Hausdorff topology.*

In particular, $\mathcal{A}(v, s)$ is precompact under the Gromov–Hausdorff topology and its Gromov–Hausdorff closure is in the moduli space.

By Theorem 1.7, a Gromov–Hausdorff limit of a sequence in $\mathcal{A}(v, s) \subset \mathcal{M}(\Gamma\backslash H_n)$ is diffeomorphic to $\Gamma\backslash H_n$. If we remove one of the two assumption, then the topology of a limit space may differ. For example, if the minimal Popp's volume diverges to the infinity, then the diameter also diverges by Proposition 5.2. Hence the limit space is not compact.

The following example shows that compact Heisenberg manifolds may collapse to a flat torus.

Example 1.5. Let $\Gamma = \langle \exp(X_1), \ldots, \exp(Z)\rangle$, and $\{d_k\}$ a sequence of left invariant corank 1 sub-Riemannian metrics on $\Gamma\backslash H_n$ whose orthonormal basis is $\{kX_1, X_2, \ldots, X_{2n}\}$. Then the length of the closed curve $c(t) = \Gamma \exp(tX_1)$ ($t \in [0, 1]$) is equal to $\frac{1}{k}$. Hence its length goes to zero as k goes to the infinity.

Moreover, the distance from Γe to $\Gamma \exp\left(\frac{1}{2}Z\right)$ also goes to zero. Indeed let $c_k : [0, 4] \to H_n$ be the path inductively defined by

$$c_k(t) = \begin{cases} \exp(t\sqrt{\frac{k}{2}}X_1) & (t \in [0,1]), \\ c_k(1)\exp\left((t-1)\frac{1}{\sqrt{2k}}X_{n+1}\right) & (t \in [1,2]), \\ c_k(2)\exp\left(-(t-2)\sqrt{\frac{k}{2}}X_1\right) & (t \in [2,3]), \\ c_k(3)\exp\left(-(t-3)\frac{1}{\sqrt{2k}}X_{n+1}\right) & (t \in [3,4]). \end{cases}$$

By the Campbell–Baker–Hausdorff formula, the endpoint of c_k is $\exp(\frac{1}{2}Z)$. By definition the length of c_k is $\frac{2\sqrt{2}}{\sqrt{k}}$. Since the distance from e to $\exp(\frac{1}{2}Z)$ is less than the length of c_k, the distance from Γe to $\Gamma \exp\left(\frac{1}{2}Z\right)$ goes to zero as $k \to \infty$. It implies that the diameter of the circle fiber $\{\Gamma \exp(tZ) \mid t \in [0, 1]\}$ goes to zero. These arguments show that the Gromov–Hausdorff limit of the sequence $\{(\Gamma\backslash H_n, d_k)\}$ is the $2n - 1$-dimensional flat torus.

In general, a Gromov–Hausdorff limit of compact Riemannian manifolds is not a manifold. For example, for any flat orbifold O, there is a sequence of flat manifolds which converges to O. This example is considered in Section 8.3.10.(b) in [31] and Section 3 in [15].

In this dissertation, collapsed limits and singular points in limit spaces are not studied. These are the next goal of our research.

We explain the difference and the difficulty of the sub-Riemannian setting compared to the Riemannian setting that is dealt by Boldt [17]. He showed a precompactness in the quotient topology under the following four assumptions (Theorem 6.2).

- the systole of a base flat torus are bounded from below by a positive number $s > 0$,
- the Riemannian volume of a base flat torus is bounded from above by $v > 0$,
- the length of a circle fiber is in a compact interval $I \subset \mathbb{R}_{>0}$,
- the length of a shortest closed curve homotopic to a circle fiber is bounded from above by $d > 0$.

The second and the third condition imply that the Riemannian volume of a compact Heisenberg manifold is bounded from above. To prove a compactness theorem for the sub-Riemannian setting, if we bound the Riemannian volume of a compact Heisenberg manifold from above, it does not contain an interesting example of a convergent sequence. Namely it does not contain a sequence of compact Riemannian Heisenberg manifolds which converges to a sub-Riemannian one, since the Riemannian volume diverges to infinity.

To address this issue, we introduce a new volume form, called the minimal Popp's volume form (see Definition 5.2). If we replace the assumption on the Riemannian volume form to the one on the minimal Popp's volume, then it contains the above example. With a careful computation of Riemannian and sub-Riemannian geodesics on compact Heisenberg manifolds, we show an upper bound on the minimal Popp's volume and a positive lower bound on the systole of a compact Heisenberg manifold imply almost all the assumptions by Boldt (Lemma 6.1, Lemma 6.2 and Remark 5.1). The rest of the proof is by adapting the argument by [17] to the sub-Riemannian setting.

On Theorem 1.7, we prove that if a sequence of compact Heisenberg manifold converges in the Gromov–Hausdorff topology, then the minimal Popp's volume and the systole are bounded from above and below respectively. This concludes that a Gromov–Hausdorff limit coincides with the limit in the moduli space.

1.3.2 Outline of the paper

The outline of this paper is as follows.

In Section 2, we recall the moduli space and Mahler's compactness theorem for flat tori.

In Section 3, we introduce the moduli space of compact Heisenberg manifolds with left invariant sub-Riemannian metrics of various corank. Our basic idea is the construction of the moduli space of flat tori in Section 3. Main tool is the characterization of isometry class of compact nilmanifolds established by Gordon and Wilson in [28], and the affineness of isometries of nilpotent Lie groups given by Kivioja and Le Donne in [33].

In Section 4, we recall the Riemannian geodesics and sub-Riemannian normal geodesics on the Heisenberg Lie group. We use Eberlein's calculation for Riemannian case in [25]. We use Agrachev–Barilari–Boscain's one for sub-Riemannian case in [2].

In Section 5, we introduce the minimal Popp's volume form on the Heisenberg Lie group. In sub-Riemannian geometry, establishing a canonical volume form with respect to its sub-Riemannian structure is important. One already has a natural volume form, called Popp's volume, which is defined for every equiregular sub-Riemannian manifold. However we also know that if a sequence of Riemannian manifolds converges to a sub-Riemmanian manifold in the Gromov–Hausdorff topology, then its Riemannian volume diverges, although the limit space has a finite Popp's volume. That is why we need to introduce an appropriate volume form which is continuous in the moduli space.

In Section 6, we prove a compactness theorem for compact Heisenberg manifolds with sub-Riemannian metrics in the topology of the moduli space (Theorem 1.6). Technical improvements from Boldt's argument are Lemma 6.1 and 6.2, where we show that the assumption in Theorem 1.6 implies those of Boldt, see Theorem 6.2.

In Section 7, we show that the quotient topology of the moduli space coincides with the Gromov–Hausdorff topology (Theorem 1.7).

2 Moduli space and Mahler's compactness theorem for flat tori

In this section, we recall the moduli space of flat tori and Mahler's compactness theorem.

Let $\{e_1, \ldots, e_n\}$ be a canonical basis of $\mathbb{R}^n$, and $\langle \cdot, \cdot \rangle_0$ the inner product on $\mathbb{R}^n$ with respect to that basis. We identify $A \in GL_n(\mathbb{R})$ to the uniform lattice $A\mathbb{Z}^n$. It is easy to show that $A, B \in GL_n(\mathbb{R})$ induce the same lattice if and only if there is $Z \in GL_n(\mathbb{Z})$ such that $AZ = B$. Let $(A\mathbb{Z}^n \backslash \mathbb{R}^n, g_A)$ be the flat torus such that the natural projection $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_0) \to (A\mathbb{Z}^n \backslash \mathbb{R}^n, g_A)$ is a Riemannian covering map. Then $(A\mathbb{Z}^n \backslash \mathbb{R}^n, g_A)$ is isometric to $(B\mathbb{Z}^n \backslash \mathbb{R}^n, g_B)$ if and only if there is $R \in O(n)$ such that $RA = B$. Now the moduli space of flat tori is given as follows.

Theorem 2.1 (Theorem 1 in [48]). *The isometry classes of n-dimensional flat tori is parametrized by*

$$\mathcal{M}(\mathbb{T}^n) = O(n) \backslash GL_n(\mathbb{R}) / GL_n(\mathbb{Z}). \tag{3}$$

It is well known that there is a one-to-one correspondance between double coset spaces $O(n)\backslash GL_n(\mathbb{R})/GL_n(\mathbb{Z})$ and $GL_n(\mathbb{Z})\backslash GL_n(\mathbb{R})/O(n)$ via the mapping $O(n)xGL_n(\mathbb{Z}) \mapsto GL_n(\mathbb{Z})x^{-1}O(n)$. This space has the following geometric description.

We identify $A \in GL_n(\mathbb{R})$ to the inner product $\langle \cdot, \cdot \rangle_A$ on $\mathbb{R}^n$ such that the orthonormal basis is $\{Ae_1, \ldots, Ae_n\}$. For two matrices $A, B \in GL_n(\mathbb{R})$, the inner product $\langle \cdot, \cdot \rangle_A$ coincides with $\langle \cdot, \cdot \rangle_B$ if and only if there is $R \in O(n)$ such that $AR = B$. In other words, $GL_n(\mathbb{R})/O(n)$ is the space of inner products on $\mathbb{R}^n$. Let g_A be the flat metric on $\mathbb{T}^n = \mathbb{Z}^n\backslash\mathbb{R}^n$ such that the natural projection $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_A) \to (\mathbb{T}^n, g_A)$ is a Riemannian covering. Then $(\mathbb{T}^n, g_A)$ is isometric to $(\mathbb{T}^n, g_B)$ if and only if there is $Z \in GL_n(\mathbb{Z})$ such that $Z^*\langle \cdot, \cdot \rangle_A = \langle \cdot, \cdot \rangle_B$. In the terminology of matrices, it is equivalent to $ZA = B$. Thus the following definition is also valid to describe isometry classes of flat tori.

Theorem 2.2 (Another formulation of Theorem 2.1)**.** *The isometry classes of n-dimensional flat tori is parametrized by the double coset space*

$$\hat{\mathcal{M}}(\mathbb{T}^n) = GL_n(\mathbb{Z})\backslash GL_n(\mathbb{R})/GL_n(\mathbb{Z}). \tag{4}$$

We call it the *dual moduli space of n-dimensional flat tori.*

We endow $\mathcal{M}(\mathbb{T}^n)$ and $\hat{\mathcal{M}}(\mathbb{T}^n)$ the quotient topology of $GL_n(\mathbb{R}) \subset \mathbb{R}^{n^2}$. Notice that the canonical bijection gives a homeomorphism between them.

Let us pass to explain the Mahler's compactness theorem. The original statement is the following.

Theorem 2.3 (Theorem 2 and 3 in [35] or see Theorem 4. IV of [19])**.** *For $v, s > 0$, let $\mathcal{A}(v, s) \subset \mathcal{M}(\mathbb{T}^n)$ be a subset such that*

1. *for all $A \in \mathcal{A}$,*
$$|\det(A)| < v,$$
2. *for all $A \in \mathcal{A}$,*
$$\inf\{\|Az\|_0 \mid z \in \mathbb{Z}^n\} > s,$$
where $\|\cdot\|_0$ is the Euclidean norm.

Then $\mathcal{A}$ is precompact in the quotient topology.

The first condition implies that the full volume of the flat torus $(A\mathbb{Z}^n\backslash\mathbb{R}^n, g_A)$ is bounded above by v. The second condition reads that the systole of the torus is bounded below by s.

From the idenification map $\mathcal{M}(\mathbb{T}^n) \to \hat{\mathcal{M}}(\mathbb{T}^n)$, we also have the following statement.

Theorem 2.4 (Equivalent to Theorem 2.3)**.** *For $v, s > 0$, let $\hat{\mathcal{A}}(v, s) \subset \hat{\mathcal{M}}(\mathbb{T}^n)$ be a subset such that*

1. *for all $A \in \hat{\mathcal{A}}$,*
$$|\det(A)|^{-1} = |\det(A^{-1})| < v,$$

2. for all $A \in \hat{\mathcal{A}}$,

$$\inf \{\|z\|_A \mid z \in \mathbb{Z}^n\} > s.$$

Then $\hat{\mathcal{A}}$ is precompact in the quotient topology.

Since the Riemannian volume form is given by $\det\left({}^tA^{-1}\right) dx_1 \wedge \cdots \wedge dx_n$, the first inequality gives a bound of the full volume. The second condition also implies a lower bound of the systole.

3 Moduli space of compact nilmanifolds with left invatiant sub-Riemannian metrics

At the first part of this section, we give a parametrization of a fixed compact nilmanifold endowed with left invariant sub-Riemannian metrics of various corank. Secondly we compute a parametrization of the moduli space for compact Heisenberg manifolds in detail.

3.1 On compact nilmanifolds

Let N be a simply conneced nilpotent Lie group, $\mathfrak{n}$ the associated Lie algebra, and Γ a lattice in N. We shall construct the moduli space of $\Gamma\backslash H_n$ with any left invariant sub-Riemannian metrics. The idea is based on the dual moduli space of flat tori given in Theorem 2.2.

First of all, we recall the classification of compact Riemannian nilmanifolds given by Gordon and Wilson in [29].

Theorem 3.1 (Theorem 5.4 in [29])**.** *Let $Stab(\Gamma)$ be the group of automorphisms of N fixing Γ, and $Inn(N)$ the group of inner automorphisms of N. Let $\tilde{g}_1$ and $\tilde{g}_2$ be left invariant Riemannian metrics on N and g_1 and g_2 the induced metrics on $\Gamma\backslash N$. Then $(\Gamma\backslash N, g_1)$ is isometric to $(\Gamma\backslash N, g_2)$ if and only if $\tilde{g}_1 = \Phi^*\tilde{g}_2$ for some $\Phi \in Inn(N) \cdot Stab(\Gamma)$.*

We generalize this theorem to sub-Riemannian ones. In order to pursue their works, we introduce data which later correspond to isometry classes.

Definition 3.1 (cf. Section 3 in [47])**.** *Let $(\mathfrak{n}, V, \langle\cdot,\cdot\rangle)$ be a triple consisting of a nilpotent Lie algebra, a bracket generating subspace and an inner products on V. We call it a data triple.*

For two data triples $(\mathfrak{n}_i, V_i, \langle\cdot,\cdot\rangle_i)$, we say that $(\mathfrak{n}_1, V_1, \langle\cdot,\cdot\rangle_1)$ is isomorphic to $(\mathfrak{n}_2, V_2, \langle\cdot,\cdot\rangle)$ if there is a Lie isomorphism $\varphi : \mathfrak{n}_1 \to \mathfrak{n}_2$ such that

- $\varphi_*^{-1}(V_1) = V_2$,
- $\varphi^*(\langle\cdot,\cdot\rangle_1) = \langle\cdot,\cdot\rangle_2$.

Isometry classes of the associated simply connected nilpotent Lie groups are characterized by the isomorphism class of the data triples.

Lemma 3.1 (cf. Theorem 3 in [47]). *Let (N_i, d_i) be two simply connected nilpotent Lie groups endowed with left-invariant sub-Riemannian metrics induced from $(V_i, \langle\cdot,\cdot\rangle_i)$. Then (N_1, d_1) is isometric to (N_2, d_2) if and only if the triple $(\mathfrak{n}_1, V_1, \langle\cdot,\cdot\rangle_1)$ is isomorphic to $(\mathfrak{n}_2, V_2, \langle\cdot,\cdot\rangle_2)$.*

For its proof, we use the following fact.

Theorem 3.2 (Theorem 2 in [33]). *Let (N_i, d_i) $(i = 1, 2)$ be pairs of connected nilpotent Lie groups and left-invariant metrics which induce the manifold topology. Then every isometry from (N_1, d_1) to (N_2, d_2) is affine.*

Here we say that an isometry between Lie groups is *affine* if it is a composition of a Lie isomorphism and a left translation.

Proof of Lemma 3.1. It follows by definition that (N_1, d_1) is isometric to (N_2, d_2) if the associated data triples are isomorphic.

Let $f : (N_1, d_1) \to (N_2, d_2)$ be an isometry. By Theorem 3.2, there is a Lie group isomorphism $\Phi : N_2 \to N_1$ such that $f = L_{f(e)} \circ \Phi$. By left-invariance, the isomorphism $\Phi : N_2 \to N_1$ is also an isometry. Since a sub-Riemannian isometry preserves metric tensors, the differential of Φ induces an isomorphism between the data triples.

□

By using Lemma 3.1, we show the sub-Riemannian analogy of Theorem 3.1.

Theorem 3.3 (cf. Theorem 5.4 in [29]). *Let $(\Gamma\backslash N, E_i, g_i)$ be compact nilmanifolds endowed with left invariant sub-Riemannian metrics. Then $(\Gamma\backslash N, E_1, g_1)$ is isometric to $(\Gamma\backslash N, E_2, g_2)$ if and only if there is $\Phi \in Inn(N) \cdot Stab(\Gamma)$ such that*

- $\Phi_*^{-1}(V_1) = V_2$,
- $\Phi^*(\tilde{g}_1) = \tilde{g}_2$,

where V_i is the fiber of E_i at Γe, and $\tilde{g}_i$ is the lift of g_i to the universal cover N.

Proof. Suppose $\Phi_*^{-1}(V_1) = V_2$ and $\Phi^*(\tilde{g}_1) = \tilde{g}_2$ with $\Phi = L_x \circ R_x^{-1} \circ \varphi$ for some $x \in N$ and $\varphi \in Stab(\Gamma)$. Since $\tilde{g}_1$ and V_1 are left-invariant, $\Phi_*^{-1}(V_1) = (R_x^{-1} \circ \varphi)_*^{-1}(V_1)$ and $\Phi^*(\tilde{g}_1) = (R_x^{-1} \circ \varphi)^*(\tilde{g}_1)$. However $R_x^{-1} \circ \varphi$ is the lift of the mapping $\Gamma\backslash N \to \Gamma\backslash N$. Hence $(\Gamma\backslash N, g_1)$ is isometric to $(\Gamma\backslash N, g_2)$ via the mapping $R_x^{-1} \circ \varphi$.

Conversely let $F : (\Gamma\backslash H_n, E_2, g_2) \to (\Gamma\backslash H_n, E_1, g_1)$ be an isometry. Then F lifts to an isometry $\tilde{F} : (N, E_2, \tilde{g}_2) \to (N, E_1, \tilde{g}_1)$. By Lemma 3.1, there is $\Phi \in Aut(N)$ such that $\Phi_*^{-1}(V_1) = V_2$ and $\Phi^*(\tilde{g}_1) = \tilde{g}_2$. Hence $F \circ \Phi^{-1}$ is an isometry of $(N, \tilde{g}_1)$. Choose $x \in N$ so that $\sigma = L_x \circ F \circ \Phi^{-1}$ is an isometry of $(N, \tilde{g}_1)$ preserving the identity. By Theorem 3.1, $\sigma \in Aut(N)$. Thus we obtain

$$R_x \circ F = L_x^{-1} \circ R_x \circ \sigma \circ \Phi \in Aut(N). \tag{5}$$

However $R_x \circ F$ is the lift of a mapping $\Gamma\backslash N \to \Gamma\backslash N$, thus $R_x \circ F \in Stab(\Gamma)$. Combining with (5), $\sigma\circ\Phi \in Inn(N)\cdot Stab(\Gamma)$. Since σ is an isometry of $(N, \tilde{g}_1)$,

$$(\sigma\circ\Phi)^*\tilde{g}_1 = \Phi^*\sigma^*\tilde{g}_1 = \Phi^*\tilde{g}_1 = \tilde{g}_2,$$
$$(\sigma\circ\Phi)_*^{-1}(V_1) = \Phi_*^{-1}\sigma_*^{-1}(V_1) = \Phi_*^{-1}(V_1) = V_2.$$

□

We are ready to define the moduli space of compact nilmanifolds $\Gamma\backslash N$ with any left invariant sub-Riemannian metrics. Fix a basis $\{X_i\}_{i=1,\ldots,n}$ of $T_{\Gamma e}(\Gamma\backslash N) \simeq \mathfrak{n}$, and set $\langle\cdot,\cdot\rangle_0$ to be the canonical inner product on $\mathfrak{n}$ with respect to the basis $\{X_1, \ldots, X_n\}$.

Denote by $\mathcal{X}_k$ the space of $n\times n$-matrices such that their kernels are written by $Span\{X_{i_1}, \ldots, X_{i_k}\}$ for some $1 \le i_1 < \cdots < i_k \le n$. We identify $A \in \mathcal{X}_k$ to the inner product $\langle\cdot,\cdot\rangle_A$ on $\mathfrak{n}$ such that an orthonormal basis of $\langle\cdot,\cdot\rangle_A$ is $\{AX_1, \ldots, AX_n\}$. Thus $(ImA, \langle\cdot,\cdot\rangle_A)$ determines a sub-Riemannian structure of corank k on $\Gamma\backslash N$.

Let $\mathcal{Y}_k \subset \mathcal{X}_k$ be the subset of bracket generating distributions. It is easy to see that $\mathcal{Y}_k$ is non-empty only if $k \ge r = \dim N/[N,N]$. Two matrices A and $B \in \mathcal{Y}_k$ determine the same distribution if and only if $ImA = ImB$. Moreover A and B determine the same inner product if and only if there is $R \in O(n)$ such that $AR = B$. In fact, if $\langle\cdot,\cdot\rangle_A = \langle\cdot,\cdot\rangle_B$, then for $u, v \in (KerA)^\perp$

$$\langle u, v\rangle_0 = \langle Au, Av\rangle_A = \langle Au, Av\rangle_B = \langle (B|_{KerB^\perp})^{-1}Au, (B|_{KerB^\perp})^{-1}Av\rangle_0.$$

Thus we have the orthogonal matrix R such that

$$R|_{(KerA)^\perp} = (B|_{(KerB)^\perp})^{-1}A|_{(KerA)^\perp}, \quad R(KerA) = KerB.$$

This implies the equality $BR = A$. The converse is trivial.

The above argument says that the set of equivalence classes $\bigcup \mathcal{Y}_k/O(n)$ is the space of sub-Riemannian metrics on $\Gamma\backslash N$.

We have given a condition for two sub-Riemannian nilmanifolds being isometric in Theorem 3.3, hence we can classify the isometry classes as follows.

Proposition 3.1. *The isometry classes of nilmanifolds with left invariant sub-Riemannian metrics of various corank is parametrized by*

$$\mathcal{M}(\Gamma\backslash N) = (Inn(N)\cdot Stab(\Gamma))_* \backslash \bigcup \mathcal{Y}_k/O(n).$$

Definition 3.2 (The quotient topology)**.** *We call the topology induced by the quotient map* $\mathbb{R}^{n^2} \supset \bigcup \mathcal{Y}_k \to \mathcal{M}(\Gamma\backslash N)$ *the canonical topology.*

3.2 The moduli space of compact Heisenberg manifolds

Let H_n be the n-Heisenberg Lie group, $\mathfrak{h}_n$ the associated Lie algebra, and $\{X_1, \ldots, X_{2n}, Z\}$ be a canonical basis of $\mathfrak{h}_n$ such that $[X_i, X_{i+n}] = Z$ for each

$i = 1, \ldots, n$ and the other brackets equal zero. In this section, we give a specific computation of the moduli space of compact Heisenberg manifolds, which was completed by Gordon and Wilson for Riemannian case in [28].

Let D_n be the set of n-tuples of integers $\boldsymbol{r} = (r_1, \ldots, r_n)$ such that $r_i \mid r_{i+1}$ for all $i = 1, \ldots, n$. For each $\boldsymbol{r} \in D_n$, let $\tilde{\boldsymbol{r}}$ be the $(2n+1)$-tuple of integers given by

$$\begin{cases} \tilde{r}_i = r_i & \text{for } \ 1 \leq i \leq n, \\ \tilde{r}_i = 1 & \text{for } \ n+1 \leq i \leq 2n+1. \end{cases}$$

For $\boldsymbol{r} \in D_n$, let $\Gamma_{\boldsymbol{r}} < H_n$ be a subgroup defined by

$$\Gamma_{\boldsymbol{r}} = \langle \exp(\tilde{r}_1 X_1), \ldots, \exp(\tilde{r}_{2n} X_{2n}), \exp(\tilde{r}_{2n+1} Z) \rangle.$$

This gives a characterization of lattices in the Heisenberg Lie group.

Theorem 3.4 (Theorem 2.4 in [28])**.** *Any uniform lattice $\Gamma < H_n$ is isomorphic to $\Gamma_{\boldsymbol{r}}$ for some $\boldsymbol{r} \in D_n$.*

Moreover, $\Gamma_{\boldsymbol{r}}$ is isomorphic to $\Gamma_{\boldsymbol{s}}$ if and only if $\boldsymbol{r} = \boldsymbol{s}$.

Fix a lattice $\Gamma_{\boldsymbol{r}}$. The following canonical matrix form is useful for analysis on the Heisenberg Lie groups.

Lemma 3.2. *For $A \in \mathcal{Y}_0 \cup \mathcal{Y}_1$, there is $R \in O(2n+1)$ and $P \in Inn(H_n)_*$ such that*

$$PAR = \begin{pmatrix} \tilde{A} & O \\ O & \rho_A \end{pmatrix}, \tag{6}$$

where $\tilde{A}$ is a $2n \times 2n$ invertible matrix and $\rho_A \in \mathbb{R}$.

Moreover, let $P' \in Inn(H_n)_*$ and $R' \in O(2n+1)$ be other matrices such that (6) hold. Then they satisfy

- $R' = R \begin{pmatrix} \tilde{R} & O \\ O & \pm 1 \end{pmatrix}$ *for some $\tilde{R} \in O(2n)$,*
- $P' = P$.

Proof. By multiplicating an appropriate orthogonal group $R \in O(2n+1)$, we can assume that AR sends Z to the Z-axis. It implies that the $(2n+1)$-row of AR is ${}^t\begin{pmatrix} 0 & \cdots & 0 & \rho_A \end{pmatrix}$ for some $\rho_A \in \mathbb{R}$. Then we can write AR as

$$AR = \begin{pmatrix} \tilde{A} & O \\ \vec{a} & \rho_A \end{pmatrix} \tag{7}$$

with some $\tilde{A} \in GL_{2n}(\mathbb{R})$ and $\vec{a} \in \mathbb{R}^{2n}$.

Let $x = \exp(\sum_{i=1}^{2n} x_i X_i + zZ) \in H_n$. We have the matrix representation of the differential of the inner automorphism $P_x = (i_x)_*$ by

$$P_x = \begin{pmatrix} I_{2n} & O \\ \tilde{x} & 1 \end{pmatrix},$$

where $\tilde{x} = (-x_{n+1}, \ldots, -x_{2n}, x_1, \ldots, x_n)$ and I_{2n} is the identity matrix.

With this terminology, we can write the matrix $P_x AR$ as

$$P_x AR = \begin{pmatrix} \tilde{A} & O \\ \vec{a} + \tilde{x}\tilde{A} & \rho_A \end{pmatrix}.$$

Since $\tilde{A}$ is invertible, we can take a unique $\tilde{x}$ such that $\vec{a} + \tilde{x}\tilde{A} = 0$.

Next we will prove the uniqueness of those matrices. Let $R' \in O(2n+1)$ be another matrix such that

$$AR' = \begin{pmatrix} \tilde{A}' & O \\ \vec{a}' & \rho'_A \end{pmatrix}.$$

We will write

$$A = \begin{pmatrix} \vec{a}_1 \\ \vdots \\ \vec{a}_{2n+1} \end{pmatrix}, \qquad R = \begin{pmatrix} \vec{r}_1 & \cdots & \vec{r}_{2n+1} \end{pmatrix}, \qquad R' = \begin{pmatrix} \vec{r'}_1 & \cdots & \vec{r'}_{2n+1} \end{pmatrix}.$$

From the definition of R and R', we have

$$\begin{cases} \vec{a}_i \cdot \vec{r}_{2n+1} = \vec{a}_i \cdot \vec{r'}_{2n+1} = 0 & \text{for } 1 \leq i \leq 2n, \\ \vec{a}_{2n+1} \cdot \vec{r}_{2n+1} = \rho_A, \\ \vec{a}_{2n+1} \cdot \vec{r'}_{2n+1} = \rho'_A \end{cases}.$$

The first equality implies that two vectors $\vec{r}_{2n+1}$ and $\vec{r'}_{2n+1}$ are unit vectors vertical to the plane spanned by $\vec{a}_1, \ldots, \vec{a}_{2n}$. Hence we have $\vec{r}_{2n+1} = \pm\vec{r'}_{2n+1}$ and $\rho_A = \pm\rho'_A$.

It is well known that $\{\vec{r}_1, \ldots, \vec{r}_{2n+1}\}$ and $\left\{\vec{r'}_1, \ldots, \vec{r'}_{2n+1}\right\}$ are orthonormal bases of $\mathbb{R}^{2n+1}$. We have shown that $\vec{r}_{2n+1} = \pm\vec{r'}_{2n+1}$, hence there is $\tilde{R} \in O(2n)$ such that

$${}^t RR' = \begin{pmatrix} \tilde{R} & O \\ O & \pm 1 \end{pmatrix}.$$

This gives the uniqueness of orthogonal matrices. Moreover, the above argument gives the matrix representation of $P_{x'}AR'$ by

$$P_{x'}AR' = \begin{pmatrix} \tilde{A}\tilde{R} & O \\ \left(\vec{a} + \tilde{x}'\tilde{A}\right)\tilde{R} & \pm\rho_A \end{pmatrix}.$$

Hence $x = x'$ gives the equality $\left(\vec{a} + \tilde{x}'\tilde{A}\right)\tilde{R} = 0$. This concludes $P' = P$. $\square$

The above argument gives a parametrization of the isometry class of compact Heisenberg manifolds by

$$\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n) = Stab(\Gamma_{\boldsymbol{r}})_* \backslash \left(GL_{2n}(\mathbb{R}) \times \mathbb{R}\right) / \left(O(2n) \times \{\pm 1\}\right).$$

We will give a matrix representation of $Stab(\Gamma_{\boldsymbol{r}})_*$. For $\boldsymbol{r} \in D_n$, let

$$diag(\boldsymbol{r}) = diag(r_1, \dots, r_n, 1, \dots, 1)$$

be a diagonal $2n \times 2n$ matrix, and define an anti-symmetric matrix

$$J_n = \begin{pmatrix} O & -I_n \\ I_n & O \end{pmatrix}.$$

We embed the group

$$\widetilde{Sp}(2n, \mathbb{R}) = \{\beta \in GL_{2n}(\mathbb{R}) \mid \beta J_n \beta = \epsilon(\beta) J_n,\ \epsilon(\beta) = \pm 1\}$$

into $GL_{2n+1}(\mathbb{R})$ via the mapping $\iota : \beta \mapsto \begin{pmatrix} \beta & 0 \\ 0 & \epsilon(\beta) \end{pmatrix}$.

With these notations, we give representations of matrices in $Stab(\Gamma_{\boldsymbol{r}})_*$ as follows.

Theorem 3.5 (Theorem 2.7 in [28])**.** *The differential of the stabilizer of $\Gamma_{\boldsymbol{r}}$ is given by*

$$\Pi_{\boldsymbol{r}} = \iota \left(G_{\boldsymbol{r}} \cap \widetilde{Sp}(2n, R) \right),$$

where $G_{\boldsymbol{r}} = diag(\boldsymbol{r}) GL_{2n}(\mathbb{Z}) diag(\boldsymbol{r})^{-1}$.

Finally we obtain a specific computation of $\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$.

Theorem 3.6. *The isometry classes of a compact Heisenberg manifold $\Gamma_{\boldsymbol{r}} \backslash H_n$ with left invariant metrics of various corank are parametrized by*

$$\Pi_{\boldsymbol{r}} \backslash \left(GL_{2n}(\mathbb{R}) \times \mathbb{R} \right) / \left(O(2n) \times \{\pm 1\} \right). \tag{8}$$

Here the action on the second factor only change its sign. Moreover, any matrices acting upon the first factor has determinant ± 1. This argument shows the following.

Lemma 3.3. *For $A, B \in \mathcal{Y}_0 \cup \mathcal{Y}_1$ with $[A] = [B]$, we have*

- $\det(\tilde{A}) = \det(\tilde{B})$,
- $|\rho_A| = |\rho_B|$.

Remark 3.1. *The double coset space $G_{\boldsymbol{r}} \backslash GL_{2n}(\mathbb{R}) / O(2n)$ is homeomorphic to $GL_{2n}(\mathbb{Z}) \backslash GL_{2n}(\mathbb{R}) / O(2n)$, the moduli space of flat tori of dimension $2n$.*

4 Length minimizing paths on the Heisenberg Lie group

In this section, we recall length minimizing paths on the Heisenberg group endowed with a Riemannian and sub-Riemannian metric defined by matrices in $\mathcal{Y}_0$ and $\mathcal{Y}_1$ respectively.

4.1 Riemannian geodesics

Let A be a matrix in $\mathcal{Y}_0$ of the canonical form, $\langle\cdot,\cdot\rangle_A$ the inner product on $\mathfrak{h}_n$ such that its orthonormal basis is $\{AX_1,\ldots,AX_{2n},AZ\}$, and $V_0\subset\mathfrak{h}_n$ the subspace given by

$$V_0 = Span\{X_1,\ldots,X_{2n}\}.$$

Let us write $A=\begin{pmatrix}\tilde{A} & O\\ O & \rho_A\end{pmatrix}$. Notice that the matrix $\tilde{A}$ is a linear transformation on V_0. Define the skew symmetric operator $j_A(c):V_0\to V_0$ for $c\in\mathbb{R}$ by

$$\overline{\langle j_A(c)(X),Y\rangle_A} = \overline{\langle cZ,[X,Y]\rangle_A},$$

where $\overline{\langle\cdot,\cdot\rangle_A}$ is the inner product on $\mathfrak{h}_n$ such that its orthonormal basis is given by $\{AX_1,\ldots,AX_{2n},Z\}$.

Remark 4.1. *In the study of Riemannian nilpotent Lie groups, one usually uses a skew symmetric operator $J_A(c):V_0\to V_0$ defined by*

$$\langle J_A(c)(X),Y\rangle_A = \langle cZ,[X,Y]\rangle_A.$$

This operator is equivalent to our $j_A(c)$ up to scalar. In fact,

$$\begin{aligned}\overline{\langle j_A(c)(X),Y\rangle_A} &= \overline{\langle cZ,[X,Y]\rangle_A}\\ &= \rho_A^2\langle cZ,[X,Y]\rangle_A\\ &= \rho_A^2\langle J_A(c)(X),Y\rangle_A\\ &= \overline{\langle\rho_A^2 J_A(c)(X),Y\rangle_A},\end{aligned}$$

thus we have $j_A(c)=\rho_A^2 J_A(c)$.

A direct calculation gives the matrix representation of $j_A(c)$ with respect to the basis $\{AX_1,\ldots,AX_{2n}\}$ by

$$j_A(c) = c\,{}^t\tilde{A}J_n\tilde{A}, \tag{9}$$

Since ${}^t\tilde{A}J_n\tilde{A}$ is skew symmetric, there is an orthogonal matrix $R\in O(2n+1)$ such that AR is of canonical matrix form and

$$\begin{aligned} j_{AR}(c) &= c\,{}^t\left(\widetilde{AR}\right)J_n\widetilde{AR} &(10)\\ &= c\begin{pmatrix} O & -diag(\lambda_1(A),\ldots,\lambda_n(A))\\ diag(\lambda_1(A)\ldots,\lambda_n(A)) & O\end{pmatrix}, &(11)\end{aligned}$$

where $\lambda_i(A)$'s are positive real numbers such that $\pm\sqrt{-1}\lambda_i(A)$ are the eigenvalues of $j_A(1)$. Later we will assume $0<\lambda_1(A)<\cdots\lambda_n(A)$. We sometimes require a canonical matrix form to satisfy (11).

Lemma 4.1. *For $A,B\in\mathcal{Y}_0$ with $[A]=[B]$, $\lambda_i(A)=\lambda_i(B)$ for all $i=1,\ldots,n$.*

Proof. From the definition of the moduli space, there is $P \in G_{\boldsymbol{r}} \cap \widetilde{Sp}(2n, \mathbb{R})$ and $R \in O(2n)$ such that $P\tilde{A}R = B$. The matrix representation of $j_{P\tilde{A}R}$ shows that

$$\begin{aligned} j_{P\tilde{A}R}(1) &= {}^t\left(P\tilde{A}R\right) J_n P\tilde{A}R \\ &= R^{-1}\ {}^t\tilde{A}{}^tPJ_nP\tilde{A}R \\ &= \pm R^{-1}\ {}^t\tilde{A}J_n\tilde{A}R \\ &= R^{-1}j_A(1)R. \end{aligned}$$

Hence the eigenvalue of $j_B(1)$ is

$$\pm\lambda_1(A), \ldots, \pm\lambda_n(A).$$

□

Later we simply write λ_i's. For $i = 1, \ldots, n$, let $W_i = Span\,\{AX_i, AX_{i+n}\}$. For a matrix in $A \in \mathcal{Y}_0$ of canonical form, let $h_i : T^*H_n \to \mathbb{R}$ be a function defined by $h_i(p) = p(AX_i(x))$ for $p \in T^*_xH_n$, and define $h_z : T^*H_n \to \mathbb{R}$ by $h_z(p) = p(Z)$.

A path $\gamma : [0, T] \to H_n$ parametrized by $\gamma(t) = \exp(\sum_{i=1}^{2n} x_i(t)AX_i + z(t)Z)$ is a *geodesic* if there is a lift $\ell : [0, T] \to T^*H_n$ such that the following Hamiltonian equation holds.

$$\begin{cases} \dot{h}_i = \lambda_i h_z h_{i+n}, \\ \dot{h}_{i+n} = -\lambda_i h_z h_i, \\ \dot{h}_z = 0, \end{cases} \qquad \begin{cases} \dot{x}_i = h_i, \\ \dot{z} = \frac{1}{2}\sum_{i=1}^n \lambda_i\left(x_i h_{i+n} - x_{i+n}h_i\right) + \rho_A^2 p_z, \end{cases}$$

where we denote $h_i(t) = h_i \circ \ell(t)$ and $h_z(t) = h_z \circ \ell(t)$. Such a lift ℓ is called an *extremal.*

Lemma 4.2 (Special case of Proposition 3.5 in [25])**.** *Let $\gamma : [0, T] \to H_n$ be the geodesic with the initial data of the extremal*

$$\gamma(0) = e, \qquad h_i(0) = p_i, \qquad h_z(0) = p_z.$$

Set $\xi_i = p_z\lambda_i$. Then γ is given as follows.

- *If $w \neq 0$, then*

$$\begin{pmatrix} x_i(t) \\ x_{i+n}(t) \end{pmatrix} = \frac{1}{\xi_i}\begin{pmatrix} \sin(\xi_i t) & \cos(\xi_i t) - 1 \\ -\cos(\xi_i t) + 1 & \sin(\xi_i t) \end{pmatrix}\begin{pmatrix} p_i \\ p_{i+n} \end{pmatrix}$$

  for each $i = 1, \ldots, n$, and

$$z(t) = \rho_A^2 p_z t + \frac{1}{2}\sum_{i=1}^{2n}\left(\frac{\lambda_i t}{\xi_i} - \frac{\lambda_i}{\xi_i^2}\sin(\xi_i t)\right)\left(p_i^2 + p_{i+n}^2\right).$$

- *If $w = 0$, then $x_i(t) = p_i t$ and $z(t) \equiv 0$.*

In particular, we can determine a length minimizing geodesic from e to a element in the center.

Lemma 4.3. *Let m be the least integer such that $\lambda_m = \lambda_n$, and $p > 0$. Then the unit speed length minimizing geodesics from e to pZ are given as follows.*

1. *If $p \leq \frac{2\pi\rho_A^2}{\lambda_n}$, then the inital datum of the geodesic is*
$$(h_1(0), \ldots, h_{2n}(0), h_z(0)) = (0, \ldots, 0, \rho_A).$$

2. *If $p \geq \frac{2\pi\rho_A^2}{\lambda_n}$, then the initial data of unit speed length minimizing geodesics are*
$$(h_1(0), \ldots, h_{2n}(0), h_z(0)) = (0, \ldots, 0, p_m, \ldots, p_n, 0, \ldots, 0, p_{m+n}, \ldots, p_{2n}, p_z),$$
where
$$\begin{cases} p_z = \sqrt{\frac{\pi}{p\lambda_n - \pi\rho_A^2}}, \\ p_m^2 + \cdots + p_{2n}^2 + \frac{p_z^2}{\rho_A^2} = 1. \end{cases}$$

For the case $p < 0$, it suffices to change the sign of p_z.

Proof. Trivially the path $\exp(t\rho_A Z)$ is a unit speed geodesic, which arrives at the point pZ at the time $\frac{p}{\rho_A}$.

Next we characterize length minimizing geodesics which join e and pZ with the initial data (p_x, p_z) with $p_x = (p_1, \ldots, p_{2n}) \neq (0, \ldots, 0)$. According to Theorem 2.9 in [30], such geodesics are length minimizing until the time $\frac{2\pi}{\xi_n}$. This implies that $p_1 = p_{1+n} = \cdots = p_{m-1} = p_{m-1+n} = 0$ if the resulting geodesic terminates at $pZ \in [H_n, H_n]$. The information of the endpoint and the velocity of the geodesic gives us the system

$$\begin{cases} p = z\left(\frac{2\pi}{\xi_n}\right) = \frac{2\pi\rho_A^2}{\lambda_n} + \frac{\pi}{p_z^2\lambda_n}|p_x|^2, \\ |p_x|^2 + (\rho_A p_z)^2 = 1 \end{cases} . \tag{12}$$

This system has a solution only if $p \geq \frac{2\pi\rho_A^2}{\lambda_n}$. This implies that if $p < \frac{2\pi\rho_A^2}{\lambda_n}$, there is no geodesic from e to pZ with the initial data $p_x \neq 0$.

Assume $p \geq \frac{2\pi\rho_A^2}{\lambda_n}$. Then we have the solution to (12), and its geodesic has length $\frac{2\pi}{\xi_n}$. On the other hand, we have another straight geodesic $\exp(t\rho_A Z)$ with length $\frac{p}{\rho_A}$. It is easy to see that $\frac{p}{\rho_A} \leq \frac{2\pi}{\xi_n}$ if and only if $p = \frac{2\pi\rho_A^2}{\lambda_n}$. This completes the lemma. □

4.2 Normal geodesics

Let A be a matrix in $\mathcal{Y}_1$ of the canonical form, and $(ImA, \langle\cdot,\cdot\rangle_A)$ the induced sub-Riemannian structure on the Heisenberg group H_n. In this section, we recall normal geodesics on the Heisenberg group with the metric induced from a matrix

$A \in \mathcal{Y}_0$. For more detailed information on this subject, see Section 4 in [1] or Section 5.2 in [2].

As in the previous section, let $\overline{\langle \cdot, \cdot \rangle_A}$ be the inner product on $\mathfrak{h}_n$ such that its orthonormal basis is $\{AX_1, \dots, AX_{2n}, Z\}$. For $c \in \mathbb{R}$, we define a skew symmetric operator $j_A(c) : V_0 \to V_0$ by

$$\overline{\langle j_A(c)(X), Y \rangle_A} = \overline{\langle cZ, [X, Y] \rangle_A}.$$

We assume that $j_A(c)$ has a canonical matrix representation

$$j_A(c) = c \begin{pmatrix} O & -diag(\lambda_1, \dots, \lambda_n) \\ diag(\lambda_1, \dots, \lambda_n) & O \end{pmatrix}.$$

Let $h_i : T^*H_n \to \mathbb{R}$ be a function defined by $h_i(p) = p(AX_i(x))$ for $p \in T^*_x H_n$, and define $h_z : T^*H_n \to \mathbb{R}$ by $h_z(p) = p(Z(x))$.

By Theorem 1.3 A path $\gamma : [0, T] \to H_n$, $\gamma(t) = \exp(\sum_{i=1}^{2n} x_i(t) AX_i + z(t) Z)$ is a normal geodesic if there is a lift $\ell : [0, T] \to T^*H_n$ such that the following sub-Riemannian Hamiltonian equation follows.

$$\begin{cases} \dot{v}_i = \lambda_i \dot{w} \dot{v}_{i+n}, \\ \dot{v}_{i+n} = -\lambda_i \dot{w} \dot{v}_i, \\ \dot{w} = 0, \end{cases} \qquad \begin{cases} \dot{x}_i = v_i, \\ \dot{z} = \frac{1}{2} \sum_{i=1}^{n} \lambda_i \left(x_i v_{i+n} - x_{i+n} v_i \right), \end{cases} \tag{13}$$

where $v_i(t) = h_i \circ \ell(t)$ and $w(t) = h_z \circ \ell(t)$.

Remark 4.2. *Every bracket generating distribution on TH_n is fat. Thus every length minimizing path on H_n is a normal geodesic.*

The sub-Riemannian Hamiltonian equation gives a specific computation of normal geodesics.

Lemma 4.4 (Section 5.2 of [2] or Lemma 14 in [41])**.** *Let $\gamma : [0, T] \to H_n$ be the normal geodesic with the initial data*

$$\gamma(0) = e, \quad h_i(\ell(0)) = v_i, \quad h_z(\ell(0)) = w,$$

where $\ell : [0, T] \to T^*H_n$ is the associated normal extremal. Set $\xi_i = q(0)\lambda_i$. Then γ is parametrized as follows.

- *If $w \neq 0$, then*

$$\begin{pmatrix} x_i(t) \\ x_{i+n}(t) \end{pmatrix} = \frac{1}{\xi_i} \begin{pmatrix} \sin(\xi_i t) & -\cos(\xi_i t) + 1 \\ \cos(\xi_i t) - 1 & \sin(\xi_i t) \end{pmatrix} \begin{pmatrix} v_i \\ v_{i+n} \end{pmatrix} \tag{14}$$

  for each $i = 1, \dots, n$, and

$$z(t) = \frac{1}{2} \sum_{i=1}^{2n} \left(\frac{\lambda_i t}{\xi_i} - \frac{\lambda_i}{\xi_i^2} \sin(\xi_i t) \right) \left(v_i^2 + v_{i+n}^2 \right). \tag{15}$$

- *If $w = 0$, then $x_i(t) = v_i t$ and $z(t) \equiv 0$.*

5 The minimal Popp's volume form on the Heisenberg Lie group

In this section, we introduce a volume form which is valid for both Riemannian and sub-Riemannian metrics on the Heisenberg Lie group H_n, and give its fundamental facts.

5.1 Riemannian volume form

First of all, we recall the Riemannian volume form and the Popp's volume form on the Heisenberg Lie group.

Let A be a matrix in $\mathcal{Y}_0$. By Lemma 3.2, we can assume that A has the canonical form $A = \begin{pmatrix} \tilde{A} & O \\ O & \rho_A \end{pmatrix}$. Then its Riemannian volume form is given by

$$v_R(A) = \left|\det({}^tA^{-1})\right| dX_1 \wedge \cdots \wedge dZ = \left|\rho_A^{-1} \det(\tilde{A})^{-1}\right| dX_1 \wedge \cdots \wedge dZ. \quad (16)$$

By Lemma 3.3, the Riemannian volume form is invariant under the choice of representatives in $\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$.

5.2 Popp's volume form

Let A be a matrix in $\mathcal{Y}_1$. By multiplicating an appropriate orthogonal group from the right, we can assume that the kernel of A is $\mathbb{R}Z$. Then A has a matrix representation by

$$A = \begin{pmatrix} \tilde{A} & O \\ \vec{a} & 0 \end{pmatrix}, \quad (17)$$

where $\tilde{A}$ is an invertible matrix and $\vec{a}$ is a vector in $\mathbb{R}^{2n}$.

We define the Popp's volume as in the procedure in Section 1.2.5. The idea of the Popp's volume is to construct a canonical inner product on $\mathfrak{h}_n = ImA \oplus [\mathfrak{h}_n, \mathfrak{h}_n]$. Such an inner product is defined as follows. Let $\langle \cdot, \cdot \rangle_{AZ}$ be an inner product on $[\mathfrak{h}_n, \mathfrak{h}_n] \simeq \mathbb{R}$ given by

$$\langle Z_1, Z_2 \rangle_{AZ} = \min \left\{ \langle X_1, X_2 \rangle_A \cdot \langle Y_1, Y_2 \rangle_A \mid [X_i, Y_i] = Z_i, \quad X_i, Y_i \in ImA \right\}.$$

Combining with the inner product $\langle \cdot, \cdot \rangle_A$ on ImA, we obtain a new inner product $\underline{\langle \cdot, \cdot \rangle_A}$ on $\mathfrak{h}_n = ImA \oplus [\mathfrak{h}_n, \mathfrak{h}_n]$ by

$$\underline{\langle (U_1, V_1), (U_2, V_2) \rangle_A} = \langle U_1, U_2 \rangle_A + \langle V_1, V_2 \rangle_{AZ}. \quad (18)$$

By Definition 1.11 The Popp's volume form $v_{sR}(A)$ on $(H_n, ImA, \langle \cdot, \cdot \rangle_A)$ is a left invariant volume form $\nu^1 \wedge \cdots \wedge \nu^{2n+1}$, where $\nu^1, \ldots, \nu^{2n+1}$ is an orthonormal basis for $(\mathfrak{h}_n, \underline{\langle \cdot, \cdot \rangle_A})$.

In the following paragraph we give an explicit formula for the Popp's volume.

Let c_{ij} be real numbers defined by $[AX_i, AX_j] = c_{ij} Z$, and define $\delta(A)$ to be

$$\delta(A) = \left(\sum_{1 \le i,j \le 2n} c_{ij}^2 \right)^{\frac{1}{2}}. \tag{19}$$

As we show in the following lemma, the function δ is defined on $\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$.

Lemma 5.1. *For two matrices $A, B \in \mathcal{Y}_1$ with $[A] = [B]$, we have $\delta(A) = \delta(B)$.*

Proof. If $[A] = [B]$, then there is $P \in G_{\boldsymbol{r}} \cap \widetilde{Sp}(2n, \mathbb{R})$ and $R \in O(2n)$ such that $P\tilde{A}R = \tilde{B}$.

A direct calculation shows that $\delta(A)$ is the Hilbert–Schmidt norm of the matrix ${}^t\tilde{A}J_n\tilde{A}$. Since $P \in \widetilde{Sp}(2n, \mathbb{R})$ and $R \in O(2n)$, we have

$$\begin{aligned} \delta(B) &= \|{}^tR^t\tilde{A}^tPJ_nP\tilde{A}R\|_{HS} \\ &= \| \pm^t \tilde{A}J_n\tilde{A}\|_{HS} \\ &= \delta(A). \end{aligned}$$

□

Remark 5.1. *Suppose A is chosen to be the canonical form. Then we have*

$$\delta(A) = \sqrt{2 \sum_{i=1}^{n} \lambda_i(A)^2}.$$

In particular the numbers $\lambda_n(A)$ and $\delta(A)$ satisfy

$$\lambda_n(A) \le \delta(A) \le 2n\lambda_n(A).$$

By Theorem 1.5, we have a specific computation of the Popp's volume form by

$$v_{sR}(A) = \delta(A)^{-1} \left| \det(\tilde{A})^{-1} \right| dX_1 \wedge \cdots \wedge dX_{2n} \wedge dZ, \tag{20}$$

In other words, the unit vector of $(\mathfrak{h}_n, \underline{\langle \cdot, \cdot \rangle_A})$ along the Z-axis is $\pm\delta(A)Z$.

Remark 5.2. *By Lemma 3.3 and Lemma 5.1, Popp's volume form does not depend on the choice of the representative of $\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$.*

5.3 Introduction of the minimal Popp's volume form

The Riemannian volume and the Popp's volume are left-invariant by their definition. Hence they are scalar multiples of a fixed Haar volume form v_H. Let $c_1, c_2 \in \mathbb{R}$ be real constants. We say that the left invariant volume form $v_1 = c_1 v_H$ is less than $v_2 = c_2 v_H$ if $|c_1| < |c_2|$.

For A in $\mathcal{Y}_0 \cup \mathcal{Y}_1$ and a bracket generating subspace $W \subset \mathfrak{h}_n$, we will define the *induced Popp's volume form* of (A, W) as follows.

Definition 5.1 (Induced Popp's volume form)**.** *We define $v(A, W)$ to be the Popp's volume form of the sub-Riemannian structure $(W, \langle\cdot,\cdot\rangle_A|_W)$.*

Under these preparation, we will introduce the *minimal Popp's volume form of A* on the Heisenberg group.

Definition 5.2 (Minimal Popp's volume)**.** *For $A \in \mathcal{Y}_0 \cup \mathcal{Y}_1$, the minimal Popp's volume form of A is defined by*

$$v(A) = \min \{v(A, W) \mid W \subset ImA : \ \textit{bracket generating subspace}\}.$$

Remark 5.3. *If A is in $\mathcal{Y}_1$, the minimal Popp's volume form of A coincides the Popp's volume form.*

For a fixed $A \in \mathcal{Y}_0$, we can find a condition for the induced Popp's volume of (A, W) attaining its minimum.

Lemma 5.2. *For a fixed matrix $A \in \mathcal{Y}_0$, $v(A, W)$ attains minimum if and only if $W \perp \mathbb{R}Z$ in $\langle\cdot,\cdot\rangle_A$.*

In particular, if A is of canonical form, then $W = V_0 = Span\{X_1, \ldots, X_{2n}\}$.

Proof. By multiplicating an appropriate orthogonal group, we can assume that A sends Z into the Z-axis. Let $P : \mathfrak{h}_n \to (\mathbb{R}Z)^\perp$ be the projection with respect to the inner product $\langle\cdot,\cdot\rangle_A$. Since W is bracket generating, W does not contain the Z-axis, Hence $P|_W$ is a linear isomorphism from W to $(\mathbb{R}Z)^\perp$. Thus we have unique t_i's in $\mathbb{R}$ such that $(P|_W)^{-1}(AX_i) = AX_i + t_i Z$. A direct calculation shows that the subset $\left\{\frac{AX_1+t_1Z}{\sqrt{1+t_1^2}}, \ldots, \frac{AX_{2n}+t_{2n}Z}{\sqrt{1+t_{2n}^2}}\right\}$ forms an orthonormal basis of $(W, \langle\cdot,\cdot\rangle_A)$.

Since $\mathbb{R}Z$ is the center of $\mathfrak{h}_n$, we have

$$\left[\frac{AX_i + t_i Z}{\sqrt{1+t_i^2}}, \frac{AX_j + t_j Z}{\sqrt{1+t_j^2}}\right] = \frac{c_{ij}}{\sqrt{\left(1+t_i^2\right)\left(1+t_j^2\right)}} Z. \tag{21}$$

By the formula (20) and (19), we obtain the induced Popp's volume form of (A, W) by

$$\begin{aligned} v(A, W) &= \left(\sum_{1\le i,j\le 2n} \frac{c_{ij}^2}{(1+t_i^2)(1+t_j^2)}\right)^{-\frac{1}{2}} \prod_{k=1}^{2n} \sqrt{1+t_k^2}\, dX_1 \wedge \cdots \wedge dX_{2n} \wedge dZ \\ &= \left(\sum_{1\le i,j\le 2n} \frac{c_{ij}^2}{(1+t_i^2)(1+t_j^2)\prod_{k=1}^{2n}(1+t_k^2)}\right)^{-\frac{1}{2}} dX_1 \wedge \cdots \wedge dX_{2n} \wedge dZ. \end{aligned}$$

Hence the induced Popp's volume $v(A, W)$ attains the minimum if and only if $t_1 = \cdots = t_{2n} = 0$, that is, $W = (\mathbb{R}Z)^\perp$. □

Remark 5.4. *If we take $A \in \mathcal{Y}_0$ with $\rho_A \ge \delta(A)$, then the minimal Popp's volume of A coincides the Riemannian volume. If not, then $v(A) = v(A, (\mathbb{R}Z)^\perp)$.*

5.4 Fundamental facts on the minimal Popp's volume

Next we will state fundamental facts on the minimal Popp's volume. One is the continuity under the quotient topology of the moduli space.

Proposition 5.1. *Let $\{[A_k]\} \subset \mathcal{M}(\Gamma\backslash H_n)$ be a sequence of metrics converging to $[A_0] \in \mathcal{M}(\Gamma\backslash H_n)$. Then there is measurable maps $\phi_i : \Gamma\backslash H_n \to \Gamma\backslash H_n$ such that the push forward of the minimal Popp's volume forms $\phi_{i*}(v(A_k))$ converges to $v(A_0)$.*

Proof. By Lemma 3.2, we can assume each A_k has the form

$$A_k = \begin{pmatrix} \tilde{A}_k & O \\ O & \rho_{A_k} \end{pmatrix}, \tag{22}$$

where $\tilde{A}_k$ converges to $\tilde{A}_0$ and $\rho_{A_k} \to \rho_{A_0}$ as $k \to \infty$. Since $\tilde{A}_k$ converges to $\tilde{A}_0$, we also have the continuity of $\delta : \mathcal{M}(\Gamma_{\boldsymbol{r}}\backslash H_n) \to \mathbb{R}$.

From the definition of the minimal Popp's volume, it is trivial to see that $v(A_k)$ converges to $v(A_0)$. These canonical form of matrices are given by applying isomorphisms in $Inn(H_n) \cdot Stab(\Gamma_{\boldsymbol{r}})$, hence we can pick measurable maps ϕ_i from those isomorphisms. $\square$

The next proposition is on a boundedness of the minimal Popp's volume. It is well known that if a sequence of compact Riemannian Heisenberg manifolds converges to a sub-Riemannian Heisenberg manifold, then the sequence of Riemannian volumes diverges, although the distance function converges. In other words, a diameter bound does not give a bound of the Riemannian volume form. However, we can show that a diameter bound gives a uniform bound of minimal Popp's volume form.

Proposition 5.2. *For any $D > 0$, there is $V(D) > 0$ such that if the diameter of a compact Heisenberg manifold is bounded by D, then the minimal Popp's volume form is smaller than $V(D) dX_1 \wedge \cdots \wedge dZ$.*

Proof. Let $A \in \mathcal{Y}_0 \cup \mathcal{Y}_1$ be a matrix of canonical form, $\langle \cdot, \cdot \rangle_A$ the associated inner product on ImA, and $\langle \cdot, \cdot \rangle_0$ the inner product on $\mathfrak{h}_n$ with respect to the basis $\{X_1, \ldots, X_{2n}, Z\}$. We will denote the distance function on $\Gamma_{\boldsymbol{r}}\backslash H_n$ and on H_n by d_A and $\tilde{d}_A$ respectively. The diameter bound implies the following inequality.

$$\sup_{x \in H_n} \inf_{\gamma \in \Gamma_{\boldsymbol{r}}} \tilde{d}_A(e, \gamma x) \leq D. \tag{23}$$

First, we show the following claim.

Claim 5.1. *For any $X \in V_0$,*

$$\|X\|_A \leq 4nD \|X\|_0.$$

In particular, we obtain a lower bound $\|AX_j\|_0 \geq \frac{1}{4nD}$.

Proof. We shall show that any $X \in V_0$ with $\|X\|_0 = 1$ satisfies $\|X\|_A \leq 4nD$. Recall that the lattice $\Gamma_{\boldsymbol{r}}$ is generated by $\tilde{r}_i \exp(X_i)$'s.

Let $\gamma_i^{\frac{1}{2}} = \exp(\frac{\tilde{r}_i}{2} X_i)$. Since $\gamma_i^{\frac{1}{2}}$ is on the plane $\exp(V_0)$, the length minimizing path from e to $\gamma_i^{\frac{1}{2}}$ in H_n is given by the straight segment $\exp(tX_i)$. Moreover, the length minimizing path from $\Gamma_{\boldsymbol{r}} e$ to $\Gamma_{\boldsymbol{r}} \gamma_i^{\frac{1}{2}}$ in $\Gamma_{\boldsymbol{r}} \backslash H_n$ is the projection of $\exp(tX_i)$. In fact, elements in $\Gamma_{\boldsymbol{r}} \gamma_i^{\frac{1}{2}}$ is written as

$$\exp\left(\left(\frac{\tilde{r}_i}{2} + z\right) X_i + P_i\right),$$

where $z \in \mathbb{Z}$ and P_i are elements orthogonal to X_i in $\langle \cdot, \cdot \rangle_0$. Hence the length minimizing path from $\Gamma_{\boldsymbol{r}} e$ to $\Gamma_{\boldsymbol{r}} \gamma_i^{\frac{1}{2}}$ is realized when

$$z = \pm 1 \quad \text{and} \quad P_i = 0.$$

This shows that the length minimizing path from $\Gamma_{\boldsymbol{r}} e$ to $\Gamma_{\boldsymbol{r}} \gamma_i^{\frac{1}{2}}$ is the projection of $\exp(tX_i)$.

Combining with the diameter bound, we obtain

$$\begin{aligned} \left\| \frac{\tilde{r}_i}{2} X_i \right\|_A &= \tilde{d}_A\left(e, \gamma_i^{\frac{1}{2}}\right) \\ &= d_A\left(\Gamma_{\boldsymbol{r}} e, \Gamma_{\boldsymbol{r}} \gamma_i^{\frac{1}{2}}\right) \\ &\leq diam\left(\Gamma_{\boldsymbol{r}} \backslash H_n, ImA, \langle \cdot, \cdot \rangle_A\right) \leq D. \end{aligned}$$

Since $\tilde{r}_i \geq 1$,

$$\|X_i\|_A \leq \frac{2D}{\tilde{r}_i} \leq 2D. \tag{24}$$

Let $X = \sum_i c_i X_i$ be a element in V_0 with $\sum_{i=1}^{2n} c_i^2 = 1$. By the triangle inequality, we have

$$\|X\|_A \leq \sum_i |c_i| \|X_i\|_A \leq \sum_i 2D = 4nD.$$

□

Let us pass to give a bound of $\det(\tilde{A})$. Claim 5.1 implies that the eigenvalue of the gram matrix ${}^t\tilde{A}\tilde{A}$ is greater than $(4nD)^{-2}$. This gives the lower bound of the determinant of the matrix $\tilde{A}$ by

$$\det(\tilde{A}) \geq (4nD)^{-2n}. \tag{25}$$

Next we compute a lower bound of $\delta(A)$. By definition we have $[X_1, X_2] = Z$, hence the combination with (24) gives the inequality

$$\begin{aligned} \underline{\|Z\|_A} &= \min\{\|U\|_A \cdot \|V\|_A \mid [U, V] = Z\} \\ &\leq \|X_1\|_A \cdot \|X_2\|_A \\ &\leq (2D)^2. \end{aligned}$$

Since $\delta(A)Z$ is the unit vector, we have

$$\delta(A) \geq (4D)^{-2}. \tag{26}$$

(25) and (26) shows that the induced Popp's volume of (A, V_0) is bounded as

$$v(A, V_0) \leq 16D^2(4nD)^{2n} dX_1 \wedge \cdots \wedge dZ.$$

From the definition of minimal Popp's volume, we also obtain a bound

$$v(A) \leq 16D^2 \left(4nD\right)^{2n} dX_1 \wedge \cdots dX_{2n} \wedge dZ,$$

thus we can take $V(D) = 16D^2 \left(4nD\right)^2$.

□

6 Mahler's compactness Theorem for compact Heisenberg manifolds

The goal of this section is to give the following theorem.

Theorem 6.1. *Let $\mathcal{A}(v, s) \subset \mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$ be a family of sub-Riemannian compact Heisenberg manifolds such that the minimal Popp's volumes are bounded above by $v > 0$ and the systoles are bounded below by $s > 0$. Then $\mathcal{A}(v, s)$ is precompact with respect to the quotient topology.*

Our proof is based on that of Boldt in [17], who showed a variation of Mahler's compactness theorem for compact Riemannian Heisenberg manifolds as follows.

Theorem 6.2 (Corollary 3.14 in [17])**.** *Let $\mathcal{A}$ be a family of Riemannian compact Heisenberg manifolds such that there are positive constants $C_1, C_2, C_3 > 0$ and a compact interval $I \subset \mathbb{R} \setminus \{0\}$ such that*

- $\inf \{\|\pi(\gamma)\|_A \mid \gamma \in \Gamma_{\boldsymbol{r}}\} \geq C_1$,
- $\det(\tilde{A}) \leq C_2$,
- $\lambda_n(A) \leq C_3$,
- $\rho_A \in I$.

Then $\mathcal{A}$ is precompact in the quotient topology.

Remark 6.1. *In [17], the notations and the direction of the inequalities differ. This is due to our definition of the moduli space, where we take the inverse elements.*

6.1 Inequalities obtained from the minimal Popp's volume forms and systoles

In the following two lemmas, we will see that the assumption in Theorem 6.1 implies the four condition in Theorem 6.2 with a compact interval $I \subset \mathbb{R}$. Let $\pi : H_n \to \mathfrak{h}_n \to V_0$ be the projection map.

Lemma 6.1. *Let $A \in \mathcal{Y}_0 \cup \mathcal{Y}_1$ be a matrix of canonical form. Suppose the systole of the compact Heisenberg manifold $(\Gamma_{\boldsymbol{r}} \backslash H_n, ImA, \langle\cdot,\cdot\rangle_A)$ is bounded below by $s > 0$. Then we have*

- $\inf \{\|\pi(\gamma)\|_A \mid \gamma \in \Gamma_{\boldsymbol{r}}\} \geq s$,
- $\delta(A) \leq \frac{\sqrt{2}n}{s}$,
- $|\rho_A| \leq C_s = \max \left\{ \sqrt{\frac{n}{\sqrt{2}\pi s}}, \frac{1}{s} \right\}$.

Proof. Recall that the systole of the compact Heisenberg manifold $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_A)$ is given by

$$\inf \left\{ \tilde{d}_A(e, h^{-1}\gamma h) \mid h \in H_n,\ \gamma \in \Gamma_{\boldsymbol{r}} \right\}.$$

First assume γ is in $\Gamma_{\boldsymbol{r}} \setminus [H_n, H_n]$. Then $\left\{h^{-1}\gamma h \mid h \in H_n\right\} = \pi^{-1}(\pi(\gamma))$. Hence the distance from e to $\left\{h^{-1}\gamma h \mid h \in H_n\right\}$ is minimized when $h^{-1}\gamma h \in \exp(V_0)$, which yields

$$\inf \left\{ \tilde{d}_A(e, h^{-1}\gamma h) \mid h \in H_n \right\} = \|\pi(\gamma)\|_A.$$

This argument gives the first inequality.

Next we give a bound of $\delta(A)$. We choose $A \in \mathcal{Y}_0 \cup \mathcal{Y}_1$ so that ${}^tAJ_nA$ has the form

$$\begin{pmatrix} O & -diag(\lambda_1, \dots, \lambda_n) \\ diag(\lambda_1, \dots, \lambda_n) & O \end{pmatrix}.$$

For each $i = 1, \dots, n$, let $c_{i,T} : [0, 4T] \to H_n$ be a concatenation of four unit speed segments inductively defined by

$$c_{i,T}(t) = \begin{cases} \exp(-tAX_i) & (t \in [0,T]), \\ c_{i,T}(T) \cdot \exp(-(t-T)AX_{i+n}) & (t \in [T, 2T]), \\ c_{i,T}(2T) \cdot \exp((t-2T)AX_i) & (t \in [2T, 3T]), \\ c_{i,T}(3T) \cdot \exp((t-3T)AX_{i+n}) & (t \in [3T, 4T]). \end{cases}$$

Moreover let $c_T : [0, 4nT] \to H_n$ be the concatenation of $c_{i,T}$'s. The endpoint of c_T is $\exp\left(4T\sum_{i=1}^n \lambda_i\right)$. We choose $T > 0$ so that the endpoint of c_T is $\exp(Z)$. Then $T = \left(4\sum_{i=1}^n \lambda_i\right)^{-1}$, and the length of c_T is given by $n\left(\sum_{i=1}^n \lambda_i\right)^{-1}$. From the assumption of the systole, we obtain the bound

$$n\left(\sum_{i=1}^n \lambda_i\right)^{-1} \geq \tilde{d}_A(e, \exp(Z)) \geq s.$$

Since each λ_i are positive, we also obtain the bound of $\delta(A)$ by

$$\delta(A) = \sqrt{2\sum_{i=1}^{n} \lambda_i^2} \leq \sqrt{2}\sum_{i=1}^{n} \lambda_i \leq \frac{\sqrt{2}n}{s}.$$

We pass to a bound of $|\rho_A|$. If $A \in \mathcal{Y}_1$, then automatically $\rho_A = 0$, thus we assume $A \in \mathcal{Y}_0$.

Assume γ is in $\Gamma_{\boldsymbol{r}} \cap [H_n, H_n]$. Then every closed curve in $\Gamma_{\boldsymbol{r}} \backslash H_n$ freely homotopic to γ is homotopic to γ with fixed endpoints. This gives a bound

$$\inf \left\{ \tilde{d}_A(e, \gamma) \mid \gamma \in \Gamma_{\boldsymbol{r}} \cap [H_n, H_n] \right\} = \tilde{d}_A\left(e, \exp(Z)\right) \geq s. \tag{27}$$

If $|\rho_A|$ is smaller than 1, then we have nothing to prove. If $|\rho_A| \geq 1$. Lemma 4.3 gives a classification of a length minimizing geodesic from e to $\exp(Z)$ as follows.

(a) If $1 \leq \frac{2\pi\rho_A^2}{\lambda_n}$, then the length minimizing path is a straight segment $\exp(tZ)$.

(b) If $1 \geq \frac{2\pi\rho_A^2}{\lambda_n}$, then the length minimizing paths are not straight segments.

In the case (a), the geodesic from e to Z in H_n is $\exp(tZ)$ with the length $\frac{1}{|\rho_A|}$. Combining with the systole bound, we have

$$s \leq d(e, Z) = \frac{1}{|\rho_A|}.$$

In case (b), the condition implies $|\rho_A| \leq \sqrt{\frac{\lambda_n}{2\pi}}$. Since $\lambda_n \leq \delta(A)$, we obtain

$$|\rho_A| \leq \max\left\{ \sqrt{\frac{n}{\sqrt{2}\pi s}}, \frac{1}{s} \right\}. \tag{28}$$

$\square$

Combination with a bound of the minimal Popp's volume gives the following information.

Lemma 6.2. *Suppose that the Heisenberg manifold $(\Gamma_{\boldsymbol{r}} \backslash H_n, ImA, \langle \cdot, \cdot \rangle_A)$ satisfies the systole bound below by $s > 0$ and the minimal Popp's volume bound above by $v > 0$. Then we have the lower bound of the determinant of $\tilde{A}$ by*

$$\det(\tilde{A}) \geq \frac{\tilde{C}_s}{v} \prod_{i=1}^{n} r_i,$$

where $\tilde{C}_s = \min\left\{ C_s^{-1}, \frac{s}{\sqrt{2}n} \right\}$.

Proof. Recall that the minimal Popp's volume form is represented as

$$v(A) = \min\left\{\rho_A^{-1}, \delta(A)^{-1}\right\} det\left(\tilde{A}\right)^{-1} dX_1 \wedge \cdots \wedge dZ.$$

Lemma 6.1 gives a bound

$$\min\left\{\rho_A^{-1}, \delta(A)^{-1}\right\} \geq \tilde{C}_s,$$

where $\tilde{C}_s$ is given as in the statement.

It is well known that the full volume of the compact Heisenberg manifold $(\Gamma_{\boldsymbol{r}}\backslash H_n, ImA, \langle\cdot,\cdot\rangle_A)$ is given by

$$\int_{\Gamma_{\boldsymbol{r}}\backslash H_n} v(A) = \min\left\{|\rho_A|^{-1}, \delta(A)^{-1}\right\} \det\left(\tilde{A}^{-1}\right)\prod_{i=1}^{n} r_i.$$

Now the total measure is bounded above by v, thus we obtain a bound of $\det(\tilde{A})$ by

$$\det\left(\tilde{A}\right) \geq \frac{\tilde{C}_s}{v}\prod_{i=1}^{n} r_i. \tag{29}$$

□

6.2 Adaptation of Boldt's techniques to the sub-Riemannian setting

We prepare technical lemmas on the matrix analysis.

Lemma 6.3 (Lemma 3.10 in [17])**.** *Let $\mathcal{U}$ be a complete set of representatives for $G_{\boldsymbol{r}}/\Pi_{\boldsymbol{r}}$. For any $C > 0$, there are only finitely many $A \in \mathcal{U}$ such that $\delta(A) \leq C$.*

Proof. For any $A, B \in \mathcal{U}$ with $A \neq B$, we have $j_A(1) \neq j_B(1)$. In fact, if it were not the case then

$$^tAJ_nA = j_A(1) = j_B(1) =^t BJ_nB,$$

thus we have $B^{-1}A \in \widetilde{Sp}(2n,\mathbb{R})$. Since $\Pi_{\boldsymbol{r}} = G_{\boldsymbol{r}} \cap \widetilde{Sp}(2n,\mathbb{R})$, we have $A = B$ in $G_{\boldsymbol{r}}/\Pi_{\boldsymbol{r}}$, it is a contradiction.

By the above argument, the mapping $\iota : \mathcal{U} \to M(2n)$, $A \mapsto^t AJ_nA$ is injective. Moreover the image of ι is discrete since entries of matrices in $\mathcal{U}$ are in $\frac{1}{r_n^2}\mathbb{Z}$. On the other hand, $\delta(A)$ is the Hilbert–Schmidt norm of $\iota(A)$. Since $M(2n)$ is finite dimensional vector space, every closed ball is compact. These imply that the intersection of the image of ι and the closed ball consists of finitely many matrices. □

For a invertible matrix $A \in GL_{2n}(\mathbb{R})$, let $s_1(A), \ldots, s_{2n}(A)$ be the singular values of A, that is the eigenvalues of ^tAA. We require $s_1(A) \geq \cdots \geq s_{2n}(A)$.

Lemma 6.4 (Lemma 3.12 in [17]). *For any $A, P \in GL_{2n}(\mathbb{R})$,*

$$\lambda_n(P)s_{2n}(A) \leq \lambda_n(PA). \tag{30}$$

Proof. First recall a fundamental fact on the singular values. As a consequence of Theorem Ⅲ.4.5 in [16], one has

$$s_i(L)s_{2n-i+1}(M) \leq s_1(LM)$$

for all $L, M \in GL_{2n}(\mathbb{R})$ and $i = 1, \ldots, 2n$. We choose $i = 2n$, $L =^t A$ and $M = {}^tPJ_nPA$, and obtain

$$s_{2n}({}^tA)s_1({}^tPJ_nPA) \leq s_1({}^t(PA)J_n(PA)). \tag{31}$$

Again we choose $i = 1$, $L =^t PJ_nP$ and $M = A$, and apply to $s_1({}^tPJ_nPA)$ in the left hand side of (31) to yield

$$s_{2n}({}^tA)s_{2n}(A)s_1({}^tPJ_nP) \leq s_1\left({}^t(PA)\, J_nPA\right). \tag{32}$$

Note that ${}^tAA$ is similar to $A{}^tA$, thus

$$s_{2n}({}^tA) = s_{2n}(A). \tag{33}$$

Moreover a direct calculation shows that ${}^t({}^tPJ_nP)^t PJ_nP = -\left({}^tPJ_nP\right)^2$, which yields

$$s_1\left({}^tPJ_nP\right) = \lambda_n^2(P). \tag{34}$$

Now (32), (33) and (34) show our claim. □

Finally we give the proof of Mahler's compactness theorem for compact sub-Riemannian Heisenberg manifolds.

Proof of Theorem 6.1. Recall the moduli space of sub-Riemannian compact Heisenberg manifolds

$$\Pi_{\boldsymbol{r}}\backslash GL_{2n}(\mathbb{R}) \times \mathbb{R}/O(2n) \times \{\pm 1\}\,.$$

Here the action on the second factor by $\Pi_{\boldsymbol{r}}$ and $O(2n) \times \{\pm 1\}$ only change the sign. Moreover, the third inequality of Lemma 6.1 gives bound of the second factor. Hence we only need to check the precompactness of the first factor, later denoted by $\tilde{\mathcal{A}}$.

We will denote by $\mathcal{B}$ the quotient space $GL_{2n}(\mathbb{R})/O(2n)$, and denote the canonical projections by $p_G : \mathcal{B} \to G_{\boldsymbol{r}}\backslash\mathcal{B}$, $p_\Pi : \mathcal{B} \to \Pi_{\boldsymbol{r}}\backslash\mathcal{B}$ and $p_{\Pi G} : \Pi_{\boldsymbol{r}}\backslash\mathcal{B} \to G_{\boldsymbol{r}}\backslash\mathcal{B}$.

Note that the quotient space $G_{\boldsymbol{r}}\backslash\mathcal{B}$ is homeomorphic to the moduli space of $2n$ dimensional flat tori. Hence Theorem 2.4, Lemma 6.1 and 6.2 show that $p_{\Pi G}(\tilde{\mathcal{A}})$ is precompact. As shown in Theorem 4.4.2.1 of [46], $G_{\boldsymbol{r}}$ acts on $\mathcal{B}$ properly discontinuously. Hence we can take a precompact subset $\mathcal{K} \subset \mathcal{B}$ such that $p_G(\mathcal{K}) = p_{\Pi G}\left(\tilde{\mathcal{A}}\right)$.

From the definition of $\mathcal{K}$, there exists a subset $\mathcal{V} \subset \mathcal{U}$ such that

$$\tilde{\mathcal{A}} \subset \bigcup_{P \in \mathcal{V}} p_{\Pi}(P\mathcal{K}). \tag{35}$$

Finally we will show that we can choose a subset $\mathcal{V}$ to be finite. This assertion gives precompactness of $\tilde{\mathcal{A}}$ since $\mathcal{K}$ is precompact.

Let A be a matrix in $\mathcal{K}$. Since $\mathcal{K}$ is precompact and the function of the minimum singular value $s_{2n} : \mathcal{B} \to \mathbb{R}$ is continuous, we have $M > 0$ such that $s_{2n}(A) > M$.

Let $\mathcal{W} \subset \mathcal{U}$ be a subset defined by

$$\mathcal{W} = \left\{ P \in \mathcal{U} \;\middle|\; \lambda_n(P) \geq \frac{\sqrt{2}n}{sM} \right\}.$$

Lemma 6.4 shows that

$$\delta(PA) > \lambda_n(PA) \geq \lambda_n(P)s_{2n}(A) \geq \frac{\sqrt{2}n}{s}$$

for any $A \in \mathcal{K}$. Lemma 6.1 implies that $\bigcup_{P \in \mathcal{W}} (P\mathcal{K})$ and $\mathcal{A}$ are disjoint. Hence we can take $\mathcal{V} = \mathcal{U} \setminus \mathcal{W}$, which consists of matrices P with $\lambda_n(P) < \frac{\sqrt{2}n}{sM}$. This yields the inequality

$$\delta(P) \leq 2n\lambda_n(P) < \frac{2\sqrt{2}n^2}{sM}.$$

By Lemma 6.3, Such $\mathcal{V}$ is finite. □

7 The quotient topology and the Gromov–Hausdorff topology on the moduli space

In this section we show the following.

Theorem 7.1. *Let* $Id : (\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n), \mathfrak{O}_q) \to (\mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n), \mathfrak{O}_{GH})$ *be the identity map, where* $\mathfrak{O}_q$ *is the quotient topology and* $\mathfrak{O}_{GH}$ *is the Gromov–Hausdorff topology. Then the identity map is a homeomorphism.*

Proof. Let $\{[A_k]\}$ be a sequence in the moduli space which converges to $[A_0]$ in $\mathfrak{O}_q$. We show that $[A_k]$ converges also in $\mathfrak{O}_{GH}$.

By applying automorphisms in $(Inn(H_n) \cdot Stab(\Gamma_{\boldsymbol{r}}))_*$, We can assume that each A_k has a caninical form $A_k = \begin{pmatrix} \tilde{A}_k & O \\ O & \rho_{A_k} \end{pmatrix}$, $\tilde{A}_k \to \tilde{A}_0$ and $\rho_{A_k} \to \rho_{A_0}$ in the quotient topology. By Lemma 4.2 and 4.4, the geodesic γ_k in the metric A_k with the initial data of the associated extremal

$$\ell_{\gamma_k}(0) = (p_1, \ldots, p_z) \in T_e^*(\Gamma_{\boldsymbol{r}} \backslash H_n)$$

converges to the geodesic γ_0 in the metric A_0 of the initial data

$$\ell_{\gamma_0}(0) = (p_1, \ldots, p_z) \in T_e^*(\Gamma_{\boldsymbol{r}} \backslash H_n).$$

This proves that $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{A_k})$ converges to $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{A_0})$ in the Gromov–Hausdorff topology. Hence the identity map is continuous.

Conversely we show that the inverse of the identity map is continuous. Let $\{[B_k]\} \subset \mathcal{M}(\Gamma_{\boldsymbol{r}} \backslash H_n)$ be a sequence which converges to $[B_0]$ in the Gromov–Hausdorff topology. Then the diameter of $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{B_k})$ is uniformly bounded, hence the minimal Popp's volumes are uniformly bounded by Proposition 5.2.

Moreover, we can see that the systoles are uniformly bounded below by using the techniques in [42]. In fact, denote by $d_{GH}(d_{B_k}, d_{B_0})$ the Gromov–Hausdorff distance between two compact Heisenberg manifolds with metrics d_{B_k} and d_{B_0}. By Proposition 3.2 in [42], we can take a positive number δ_2 such that the δ_2-cover of the compact Heisenberg manifold $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{B_0})$ is the universal cover. Fix $\epsilon < \frac{\delta_2}{31}$ and set $\delta_1 = \delta_2 - 11\epsilon$. Take sufficiently large k so that $d_{GH}(d_{B_k}, d_{B_0}) \leq \epsilon$. Denote by G_k the group of deck transformation of the δ_1-cover of $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{B_k})$. From its definition, G_k is a quotient group of $\Gamma_{\boldsymbol{r}}$. On the other hand, by Theorem 3.4 in [42], there is a surjective homomorphism from G_k to $\Gamma_{\boldsymbol{r}}$. Hence G_k is isomorphic to $\Gamma_{\boldsymbol{r}}$. This argument implies that the systoles of $(\Gamma_{\boldsymbol{r}} \backslash H_n, d_{B_k})$ are uniformly bounded below by δ_1.

Now Theorem 6.1 gives us a converging subsequence in the quotient topology. Moreover the limit point is unique since the Gromov–Hausdorff topology splits isometry classes. This concludes that the inverse map of the identity map is continuous. □

References

[1] Agrachev A., Barilari D., Boscain U.: A Comprehensive Introduction to Sub-Riemannian Geometry. *Cambridge University Press, Cambridge* (2019).

[2] Agrachev A., Barilari D., Boscain U.: On the Hausdorff volume in sub-Riemannian geometry, *Calc. Var. Partial Differential Equations,* **43**(3-4),355388 (2012).

[3] Agrachev A. A., Gamkrelidze R. V.: Feedback-invariant optimal control theory and differential geometry. I. Regular extremals. *J. Dynam. Control Systems*, **3**(3) 343389 (1997). 1997.

[4] Agrachev A. A., Zelenko I.: Geometry of Jacobi curves. I. *J. Dynam. Control Systems,* **8**(1) 93140 (2002).

[5] Agrachev A. A., Sachkov Y.: Control Theory from the Geometric Viewpoint, *volume 87 of Encyclopaedia of Mathematical Sciences, Springer-Verlag*, New York-Heidelberg-Berlin (2004).

[6] Agrachev A. A., Lee P. W. Y.: Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds. *Math. Ann.*, **360**(1-2), 209253 (2014).

[7] Ambrosio N., Gigli L., Saveré G., Bakry–Émery curvature dimension condition and Riemannian Ricci curvature bound, *Ann. Prob.*, **43**(1), 339-404 (2015).

[8] Ambrosio N., Gigli L., Saveré G.: Nonlinear diffusion equations and curvature conditions in metric measure spaces, *Mem. Amer. Math. Soc.*, in press.

[9] Bakry D.: L'hypercontractivité et son utilisation en théorie des semigroupes, *Ecole d'Été de Probabilities de St-Flour*, Lecture notes in Math (1994).

[10] Bakry D., Gentil I., Ledoux M.: Analysis and geometry of Markov diffusion operator, *Grundlehren der Mathematischen Wissenschaften* **348**. Springer, Cham (2014).

[11] Barilari D., Rizzi L.: Comparison theorems for conjugate points in sub-Riemannian geometry, *ESAIM Control Optim. Calc. Var.* **22** no. 2, 439472 (2016).

[12] Barilari D., Rizzi L.: Sharp measure contraction property for generalized H-type Carnot groups, *Commun. Contemp. Math.*, **20**(6), 1750081, 24pp (2018).

[13] Baudoin F., Bonnefont M., Garofalo N., Munive I.: Volume and distance comparison theorems for sub-Riemannian manifolds, *J. Funct. Anal.*, **267** no.7, 2005-2027 (2014).

[14] Baudoin F., Garofalo N., Curvature-dimension inequalities and Ricci lower bounds for subRiemannian manifolds with transverse symmetries, *J. Eur. Math. Soc.*, **19**(1), 151-219 (2017).

[15] Bettiol R., Derdzinski A., Piccione P.: Teichmüller theory and collapse of flat manifolds. *Ann. Mat. Pura Appl.*, **197**(4), 12471268 (2018).

[16] Bhatia R.: Matrix analysis, *Springer New York* (1996).

[17] Boldt S.: Relatively compact sets of Heisenberg manifolds, *preprint*, arXiv: 2008.05328.

[18] Boldt S.: The height of compact nonsingular Heisenberg-like Nilmanifolds, Dessertation, Doi: 10.18452/18924 (2018).

[19] Cassels, J.W.S.: An introduction to the geometry of numbers. *Berlin, Heidelberg, New York:Springer* (1971).

[20] Cheeger J., Gromoll D.: The splitting theorem for manifolds of nonnegative Ricci curvature, *Journal of Differential Geometry*, **6**(1), 119128 (1971).

[21] Cheeger J., Colding T. H.: Lower bounds on Ricci curvature and the almost rigidity of warped products. *Ann. of Math.* (2)**144**, 1 189237 (1996).

[22] On the structure of spaces with Ricci curvature bounded below. I. *J. Differential Geom.* **46** 3, 406480 (1997).

[23] Cheeger J., Colding T. H.: On the structure of spaces with Ricci curvature bounded below. II. *J. Differential Geom.* **54**, 1, 1335 (2000).

[24] Cheeger J., Colding T. H.: On the structure of spaces with Ricci curvature bounded below. III, *J. Differential Geom.* **54**, 1, 3774 (2000).

[25] Eberlein P.: Geometry of 2-step nilpotent Lie groups with a left invariant metric, *Ann. sci. Ec. Norm. Sup.*, **27**, 611-660 (1994).

[26] Erbar M., Kuwada K., Sturm K. T.: On the Equivalence of the Entropic Curvature Dimension Condition and Bochner's Inequality on Metric Measure Space, *Invent. Math.*, **201**, no. 3, 9931071 (2015).

[27] Fukaya K., Collapsing of Riemannian manifolds and eigenvalues of Laplace operator, *Invent. Math.* **87**,517-547 (1987).

[28] Gordon C. S., Wilson E. N.: The spectrum of the Laplacian on Riemannian Heisenberg manifolds, *Michigan Math. J.*,**33**, 253271 (1986).

[29] Gordon C. S., Wilson E. N.: Isometry groups of Riemannian solvmanifolds, *Trans. Amer. Math. Soc.* **307**, no. 1, 245269, (1988).

[30] Gornet R., Mast M. B.: Length minimizing geodesics and the length spectrum of Riemannian two-step nilmanifolds, *J. Geom. Anal.*, **13**(1), 107143, (2003).

[31] Gromov M.: Structures métriques pour les variétés riemanniannes, Textes Mathématiques, **1**, Paris, 1-120, (1981).

[32] Juillet N.: Geometric inequalities and generalized Ricci bounds in the Heisenberg group, *Int. Math. Res. Not. IMRN*, (13), 23472373 (2009).

[33] Kivioja V., Le Donne E.: Isometries of nilpotent metric groups. *J. c. polytech. Math.*, **4**, 473482 (2017).

[34] Lott L., Villani C.: Ricci curvature for metric-measure spaces via optimal transport, *Ann. of Math.* **169**, 903991, (2009).

[35] Mahler K.: On lattice points in n-dimensional star bodies. I. Existence theorems, *Proceedings of the Royal Society of London. Series A: Mathematical, Phisical and Engineering Sciences,* **187**, 151-187, (1946).

[36] Montgomery R.: Abnormal minimizers, *SIAM J. Control Optim.*, **32**(6), 1605-1620 (1994).

[37] Montgomery R.: A Tour of Subriemannian Geometries, Their Geodesics and Applications, *American Mathematical Society* (Mathematical Surveys and Monographs Vol. **91**) (2000).

[38] Myers S. B.: Riemannian manifolds with positive mean curvature, *Duke Mathematical Journal*, **8**(2), 401404 (1941).

[39] Ohta S.: On the measure contraction property of metric measure spaces, *Comment. Math. Helv.* **82**, no. 4, 805828 (2007).

[40] Rifford L.: Ricci curvatures in Carnot groups. *Math. Control Relat. Fields*, **3**(4), 467487, (2013).

[41] Rizzi L.: Measure contraction properties of Carnot groups, *Calc. Var. Partial Differential Equations,* **55**(3):Art. 60, 20pp, (2016).

[42] Sormani C., Wei G.: Hausdorff Convergence and Universal Covers, *Trans. Amer. Math. Soc.*, **353**, 35853602 (2001).

[43] Sturm T.: On the geometry of metric measure spaces. I, *Acta Math.*, **196**, 65131 (2006).

[44] Sturm T.: On the geometry of metric measure spaces. II, *Acta Math.*, **196**, 133177 (2006).

[45] Tashiro K.: Precompactness theorem for compact Heisenberg manifolds with sub-Riemannian metrics and the Gromov–Hausdorff topology, *preprint*, arXiv:2004.09407.

[46] Terras A.: Harmonic analysis on symmetric spaces and applications. II, *Springer New York* (1988).

[47] Wilson E. N.: Isometry groups on homogeneous nilmanifolds, *Geom. Dedicata* **12**, no. 3, 337346 (1982).

[48] Wolf J.A.: Local and global equivalence for flat affine manifolds with parallel geometric structures, *Geom. Dedicata,* **2**, 127132 (1973).

[49] Zelenko I., Li C.: Differential geometry of curves in Lagrange Grassmannians with given Young diagram. *Differential Geom. Appl.*, **27**(6), 723742 (2009).