J. Math. Kyoto Univ. (JMKYAZ)
43-3 (2003), 509–522

Geodesics and isometries of Carnot groups

By

Iwao Kishimoto

Abstract

We will investigate geodesics in 2-step Carnot groups. And, by applying this, we will prove that every isometry of 2-step Carnot group N fixing the identity is an automorphism of Lie group N. Our argument will be elementary.

Finally, we will see that the same conclusion holds for general Carnot groups.

1. Introduction

Let M be a connected manifold and D be a subbundle of the tangent bundle TM. Fix a metric on D. The curve in M is said to be *horizontal* if it is tangent to D at almost all points.

The distribution D is called *bracket-generating* if, for every point p in M, local sections of D near p span together with all their commutators the tangent space T_pM of M at p. By Chow's theorem, every two points can be joined by a horizontal curve if D is bracket-generating ([4]).

Using the metric on D, we can measure the length of horizontal curves. Then, for any two points x, y in M, we define a distance between them by the infimum of the lengths of all the horizontal curves joining them, which is denoted by $d(x, y)$. The metric space (M, d) is said to be a *Carnot-Carathéodory space*.

A horizontal curve γ is said to be a *normal geodesic* if there exists a lift $\tilde{\gamma}$ of γ to the cotangent bundle T^*M such that $\tilde{\gamma}$ satisfies the Hamiltonian equation determined by D and its metric. The normal geodesics are geodesics, that is, locally length-minimizing curves ([5], [13], [14]).

But this converse is not true. In fact, some examples of geodesics that is not normal are known ([1], [11]). If the set of horizontal curves joining fixed two points is a "submanifold" of the path space which consists of all the smooth curves, then every geodesic is normal. However, such case is rare, because some assumptions for D is needed, for instance, the strongly bracket-generating hypothesis ([13]). This is why we have difficulty in solving the variational problem.

1991 *Mathematics Subject Classification(s).* Primary 53C17; Secondary 57S25
Received January 22, 2002
Revised March 24, 2003

For a metric space (X, d) and a point p in X, we define the metric space (X_p, d_p) to be a tangent cone of X at p, if the family of the pointed metric spaces (X, td, p) parametrized by the positive number t converges to a pointed metric space (X_p, d_p, o_p) as $t \to \infty$ with respect to the Gromov-Hausdorff distance.

The tangent cone of a Riemannian manifold is isomorphic to Euclidean space with the same dimension. On the other hand, the tangent cone of a Carnot-Carathéodory space is isomorphic to a nilpotent Lie group with a left-invariant Carnot-Carathéodory distance. Its group structure does not depend on the metric on D, but only on D ([9]). These nilpotent groups are called to be *Carnot groups* ([12]).

If we say that Euclidean space is the most basic model in Riemannian geometry, we may say that Carnot group is the most basic model in Carnot-Carathéodory geometry.

In [13] and [14], Strichartz investigated the smooth isometries of Carnot-Carathéodory spaces under the strongly bracket-generating hypothesis. Pansu showed, in the proof of his rigidity theorem ([12]), that every isometry of the quaternionic Heisenberg groups is a group-automorphism. Hamenstädt proved in [5] that every isometry of Carnot groups is a group-automorphism. Her argument, however, needs the assumption that every geodesic is normal. We will prove this fact without the normality assumption. By restricting our argument to 2-step Carnot groups, we can use more elementary methods.

This paper is organized as follows. In Section 2, we will investigate horizontal curves and geodesics in 2-step Carnot groups after the precise definition of Carnot group is given. In Section 3, we will determine the isometry group of 2-step Carnot groups without using the differentials of isometries. For the r-step Carnot groups with $r > 2$, in Section 4, we first establish the differentiability of isometries and by applying this fact we determine the isometry groups.

2. Geodesics and lines of Carnot groups

2.1. Carnot groups

Let N be a simply-connected nilpotent Lie group. Assume that the Lie algebra $\mathfrak{g}$ of N satisfies

$$\mathfrak{g} = \mathfrak{g}^1 \oplus \cdots \oplus \mathfrak{g}^r, \qquad [\mathfrak{g}^1, \mathfrak{g}^j] = \mathfrak{g}^{j+1} \qquad (1 \leq j \leq r-1).$$

We fix an inner product of $\mathfrak{g}^1$. A curve $c(t)$ in N is said to be *horizontal* if the tangent vector $\dot{c}(t)$ belongs to $\mathfrak{g}^1$ for almost every t.

The length of a horizontal curve $c : [a, b] \to N$ is defined to be

$$L(c) = \int_a^b \|\dot{c}(t)\| \, dt,$$

where $\| \cdot \|$ is the metric on $\mathfrak{g}^1$.

Any two points in N can be joined by a horizontal curve since the subspace $\mathfrak{g}^1$ generates the Lie algebra $\mathfrak{g}$. The distance $d_N(x,\, y)$ of two points $x,\, y$ in N is defined to be the infimum of the lengths of horizontal curves joining these two points. We call the metric space $(N,\, d_N)$ the (r-step) *Carnot group*.

The mappings $A_s : \mathfrak{g} \to \mathfrak{g}$, and $a_s : N \to N$ ($s \in \mathbf{R} \setminus \{0\}$) are defined to be $A_s|_{\mathfrak{g}^j} = s^j \cdot \mathrm{Id}_{\mathfrak{g}^j}$, and $a_s(\exp X) = \exp(A_s(X))$. Since A_s is automorphisms of Lie algebra $\mathfrak{g}$, the mappings a_s is automorphisms of N. Moreover, a_s multiply lengths of horizontal curves by s. Hence $d_N(a_s(x),\, a_s(y)) = |s|\, d_N(x,\, y)$ for any two points $x,\, y$ in N.

2.2. Horizontal curves

From this subsection to the next section, the Carnot group N is assumed to be 2-step.

Let $c(t)$ be a curve in N and $\sigma(t)$ be the curve in $\mathfrak{g}$ such that $c(t) = \exp \sigma(t)$. The decomposition $\sigma = \sigma^1 + \sigma^2$ corresponds to the decomposition $\mathfrak{g} = \mathfrak{g}^1 \oplus \mathfrak{g}^2$. By the Campbell-Hausdorff formula,

$$\begin{aligned} c(t)^{-1} c(t+\varepsilon) &= \exp(-\sigma(t)) \exp(\sigma(t+\varepsilon)) \\ &= \exp\left(\sigma(t+\varepsilon) - \sigma(t) - \frac{1}{2}[\sigma(t),\, \sigma(t+\varepsilon)]\right) \\ &= \exp\Big(\sigma^1(t+\varepsilon) - \sigma^1(t) + \sigma^2(t+\varepsilon) - \sigma^2(t) \\ &\qquad - \frac{1}{2}[\sigma^1(t),\, \sigma^1(t+\varepsilon) - \sigma^1(t)]\Big). \end{aligned}$$

Differentiating at $\varepsilon = 0$, we obtain $\dot\sigma^1(t) + \dot\sigma^2(t) - (1/2)[\sigma^1(t),\, \dot\sigma^1(t)]$. Hence we have

Lemma 2.1. *The curve $c(t) = \exp \sigma(t)$ is horizontal if and only if*

$$\dot\sigma^2(t) - \frac{1}{2}[\sigma^1(t),\, \dot\sigma^1(t)] = 0$$

for almost every t.

Let c_s be a family of horizontal curves and $X(t)$ be a vector field along $c\,(= c_0)$ such that $X(t) = \partial c_s(t)/\partial s|_{s=0}$. Next we determine the equation that the vector field $X(t)$ should satisfy.

Lemma 2.2. *A vector field X along a curve c is determined by a family of horizontal curves if and only if*

$$\dot X^2 + [\dot\sigma^1,\, X^1] = 0. \tag{2.1}$$

A vector field X along a horizontal curve is called to be *horizontal* if X satisfies the above equation.

Proof. For a family of horizontal curves $c_s(t) = \exp \sigma_s(t)$ with $c_0 = c$,

$$c_0(t)^{-1} c_s(t) = \exp\left(\sigma_s^1(t) - \sigma_0^1(t) + \sigma_s^2(t) - \sigma_0^2(t) - \frac{1}{2}[\sigma_0^1(t),\, \sigma_s^1(t) - \sigma_0^1(t)]\right).$$

Therefore, if we denote the associated vector field by $X = X^1 + X^2$, then

$$X^1 = \left.\frac{\partial}{\partial s}\sigma_s^1\right|_{s=0}, \qquad X^2 = \left.\frac{\partial}{\partial s}\sigma_s^2\right|_{s=0} - \frac{1}{2}\left[\sigma^1,\, \left.\frac{\partial}{\partial s}\sigma_s^1\right|_{s=0}\right]. \tag{2.2}$$

By Lemma 2.1,

$$\begin{aligned}
\dot{X}^2 + [\dot{\sigma}^1,\, X^1] &= \left.\frac{\partial}{\partial s}\dot{\sigma}_s^2\right|_{s=0} + \frac{1}{2}[\dot{\sigma}^1,\, X^1] - \frac{1}{2}[\sigma^1,\, \dot{X}^1] \\
&= \frac{1}{2}\left.\frac{\partial}{\partial s}[\sigma_s^1,\, \dot{\sigma}_s^1]\right|_{s=0} - \frac{1}{2}[X^1,\, \dot{\sigma}^1] - \frac{1}{2}[\sigma^1,\, \dot{X}^1] = 0.
\end{aligned}$$

Conversely, for any vector field $X = X^1 + X^2$ satisfying (2.1), define a family of horizontal curves $c_s(t) = \exp \sigma_s(t)$ to be such that

$$\begin{aligned}
&\sigma_s^1(t) = \sigma^1(t) + sX^1(t), \\
&\sigma_s^2(t) = s\left(X^2(0) + \frac{1}{2}[\sigma^1(0),\, X^1(0)]\right) + \frac{1}{2}\int_0^t [\sigma_s^1(u),\, \dot{\sigma}_s^1(u)]\, du.
\end{aligned}$$

By Lemma 2.1, the curves $c_s(t)$ are horizontal and $X(t) = \partial c_s(t)/\partial s|_{s=0}$, since the equation (2.1) yields (2.2) □

Definition 2.1. A horizontal curve $c : [0,1] \to N$ is *regular* if, for any $\alpha \in \mathfrak{g}$, there exists a horizontal vector field X along c such that $X(0) = 0$ and $X(1) = \alpha$.

For a curve $\eta : [0,1] \to \mathfrak{g}^1$ such that $\eta(0) = \eta(1) = 0$, we define

$$\mathcal{H}_c(\eta) = -\int_0^1 [\dot{\sigma}^1,\, \eta]\, dt.$$

Then, by Lemma 2.2, we have

Lemma 2.3. *A horizontal curve c is regular if and only if the linear mapping $\mathcal{H}_c$ is surjective onto $\mathfrak{g}^2$.*

Let $\mathfrak{h}_c$ be the subspace of $\mathfrak{g}^1$ spanned by $\{\sigma^1(t) - \sigma^1(0) \mid 0 < t \leq 1\}$. Since $\mathfrak{h}_c$ is also spanned by $\{\dot{\sigma}^1(t) \mid 0 \leq t \leq 1\}$, we have

Lemma 2.4. *A horizontal curve c is regular if and only if*

$$[\mathfrak{h}_c,\, \mathfrak{g}^1] = \mathfrak{g}^2.$$

In particular, a horizontal curve c is regular if $\mathfrak{h}_c = \mathfrak{g}^1$.

Remark 1. A 2-step graded Lie algebra $\mathfrak{g}$ is *strongly bracket-generating*, if $[\xi, \mathfrak{g}^1] = \mathfrak{g}^2$ for any $\xi \in \mathfrak{g}^1 \setminus \{0\}$. In this case, all the horizontal curve are regular except constant curves. Conversely, if all the horizontal curve are regular except constant curves, then $\mathfrak{g}$ is strongly bracket-generating.

For example, Heisenberg group is strongly bracket-generating.

The above two lemmas are the special case of those in [6].

By Lemma 2.1, the curve $\sigma(t)$ is contained in the Lie subalgebra $\mathfrak{h}_c \oplus [\mathfrak{h}_c, \mathfrak{h}_c]$ if $\sigma(0) = 0$. Hence we have

Proposition 2.1. *Assume that a horizontal curve c passes through the identity of N. There exists a subgroup M of N such that the curve c is contained and is regular in Carnot group M.*

2.3. Geodesics

For any point $x \in N$ and any cotangent vector $\alpha \in T_x^* N$ at x, we define $\|\alpha\| = \sup\langle \alpha, v\rangle$ where v is a vector in $\mathfrak{g}^1$ with $\|v\| = 1$.

Let H be a Hamiltonian function on the cotangent bundle T^*N such that $H(\alpha) = (1/2)\|\alpha\|^2$.

Definition 2.2. A curve γ in Carnot group N is *normal geodesic* if there exists a lift $\tilde{\gamma}$ of γ to T^*N such that $\tilde{\gamma}$ satisfies the Hamiltonian equation determined by the function H:

$$\begin{cases} \dfrac{dq_i(\tilde{\gamma}(t))}{dt} = \left.\dfrac{\partial H}{\partial p_i}\right|_{\tilde{\gamma}(t)} \\ \dfrac{dp_i(\tilde{\gamma}(t))}{dt} = -\left.\dfrac{\partial H}{\partial q_i}\right|_{\tilde{\gamma}(t)} \end{cases},$$

where (q_i, p_i) is the standard coordinate of T^*N determined by a coordinate (q_i) of N.

Let $P : \mathfrak{g} \to \mathfrak{g}^1$ be the projection. We extend the inner product on $\mathfrak{g}^1$ to one on $\mathfrak{g}$ such that $\mathfrak{g}^1$ and $\mathfrak{g}^2$ are orthogonal.

Hereafter, we identify $\mathfrak{g}$ with the tangent space $T_x N$ and the cotangent space $T_x^* N$ at $x \in N$.

Lemma 2.5. *The Hamiltonian equation of normal geodesics is equivalent to the following equation*:

$$\dot{\gamma}(t) = P\mathrm{Ad}^*_{\gamma(t)}\xi, \qquad \xi \in \mathfrak{g}. \tag{2.3}$$

Here, Ad_g^* is the adjoint transformation of Ad_g with respect to the inner product on $\mathfrak{g}$.

The proof is to see [5] (and also [8]).

It is remarked that the solutions of the equation (2.3) do not depend on an inner product on $\mathfrak{g}^2$.

Let $\gamma(t)$ be a normal geodesic that satisfies the equation (2.3). Then, for any $\eta \in \mathfrak{g}^1$,

$$\langle \dot{\gamma}(t),\, \eta \rangle_{\mathfrak{g}^1} = \langle \mathrm{Ad}^*_{\gamma(t)}\xi,\, \eta \rangle_{\mathfrak{g}} = \langle \xi,\, \mathrm{Ad}_{\gamma(t)}\eta \rangle_{\mathfrak{g}}.$$

If $\gamma(t) = \exp \sigma(t)$, then

$$\begin{aligned} \mathrm{Ad}_{\gamma(t)}\eta &= \exp(\mathrm{ad}(\sigma(t)))\eta = \eta + \mathrm{ad}(\sigma(t))\eta \\ &= \eta + [\sigma^1(t),\, \eta]. \end{aligned}$$

Hence we have

$$\langle \dot{\gamma}(t),\, \eta \rangle_{\mathfrak{g}^1} = \langle \xi^1,\, \eta \rangle_{\mathfrak{g}^1} + \langle \xi^2,\, [\sigma^1(t),\, \eta] \rangle_{\mathfrak{g}^2}.$$

For each $w \in \mathfrak{g}^2$, we define an anti-symmetric transformation A_w of $\mathfrak{g}^1$ such that

$$\langle A_w u,\, v \rangle_{\mathfrak{g}^1} = \langle w,\, [u,\, v] \rangle_{\mathfrak{g}^2}, \qquad \text{for} \quad \text{any } u,\, v \in \mathfrak{g}^1. \tag{2.4}$$

Then we have

$$\dot{\sigma}^1(t)\, (= \dot{\gamma}(t)) = \xi^1 + A_{\xi^2}\sigma^1(t). \tag{2.5}$$

We can compute $\sigma^2(t)$ by Lemma 2.1.

Finally, we introduce some results on the normal geodesic.

Theorem 2.1 (Strichartz [13], [14], Hamenstädt [5]). *Every normal geodesic is a geodesic, that is, a locally length-minimizing curve.*

In this paper, every geodesic is assumed to have a constant velocity. We must remark that every normal geodesic has a constant velocity.

The converse of the above theorem is not true in general ([1], [8]).

Theorem 2.2 (Hsu [6]). *Every regular geodesic is normal.*

These two theorems are true for general Carnot-Carathéodory spaces.

By this theorem and Proposition 2.1, we can easily obtain

Theorem 2.3 (Golé-Karidi [1]). *Assume that the geodesic γ passes through the identity of N. There exists a subgroup M of N such that the curve γ is contained and is normal in the Carnot group M.*

2.4. Universal Carnot groups

Given a 2-step Carnot group N, we define the *universal Carnot group* $\tilde{N}$ of N whose Lie algebra $\tilde{\mathfrak{g}} = \tilde{\mathfrak{g}}^1 \oplus \tilde{\mathfrak{g}}^2$ satisfies the following relations:

$$\tilde{\mathfrak{g}}^1 = \mathfrak{g}^1, \qquad \tilde{\mathfrak{g}}^2 = \bigwedge\nolimits^2 \mathfrak{g}^1.$$

Let $\Pi_2 : \tilde{\mathfrak{g}}^2 \to \mathfrak{g}^2$ is the surjective linear mapping such that $\Pi_2(u \wedge v) = [u,\, v]$, and $\Pi : \tilde{\mathfrak{g}} \to \mathfrak{g}$ be the surjective homomorphism of graded Lie algebras such that $\Pi(\alpha^1 + \alpha^2) = \alpha^1 + \Pi_2(\alpha^2)$.

Let $\pi : \tilde{N} \to N$ be the surjective homomorphism of Carnot groups induced by Π. It is easy to see that, for any point $\tilde{x} \in \tilde{N}$ and any horizontal curve c in N through $\pi(\tilde{x})$, there exists a unique horizontal curve $\tilde{c}$ in $\tilde{N}$ through $\tilde{x}$ such that $\pi(\tilde{c}) = c$. We call this curve $\tilde{c}$ the *π-lift* of c.

Since the mapping π is surjective and preserves the length of horizontal curves, we have

Lemma 2.6. *The π-lift $\tilde{\gamma}$ of every geodesic γ in N is a geodesic in $\tilde{N}$. Moreover, if γ is shortest, then $\tilde{\gamma}$ is also shortest.*

About geodesics in the universal Carnot groups, the following theorem has already been known ([8]).

Theorem 2.4 (Brockett). *Every geodesic in $\tilde{N}$ is normal.*

Proof. Let γ be a geodesic in $\tilde{N}$. By Theorem 2.3, there exists a subgroup M of $\tilde{N}$ such that γ is a normal geodesic in M. Thus, by the equation (2.5), we can take an element $\xi = \xi^1 + \xi^2$ in the Lie algebra $\mathfrak{h} = \mathfrak{h}^1 \oplus \mathfrak{h}^2$ of M such that

$$\dot{\sigma}^1(t) = \xi^1 + \hat{A}_{\xi^2}\sigma^1(t), \tag{2.6}$$

where $\hat{A}_{\xi^2}$ is the anti-symmetric transformation of $\mathfrak{h}^1$ defined by the equation (2.4).

We can extend $\hat{A}_{\xi^2}$ to the transformation A_{ξ^2} of $\tilde{\mathfrak{g}}^1$ by $A_{\xi^2}|_{\mathfrak{h}^1} = \hat{A}_{\xi^2}$, and $A_{\xi^2}|_{(\mathfrak{h}^1)^\perp} = 0$. The transformation A_{ξ^2} of $\tilde{\mathfrak{g}}^1$ satisfies the equation (2.4), since $\mathfrak{h}^1 \wedge (\mathfrak{h}^1)^\perp$ is orthogonal to $\mathfrak{h}^2 = \bigwedge^2 \mathfrak{h}^1$.

This implies that γ is a normal geodesic in $\tilde{N}$. $\square$

Now we compute the geodesic equation of $\tilde{N}$.

Let $\gamma(t) = \exp(\sigma^1(t) + \sigma^2(t))$ be a geodesic in $\tilde{N}$. By the equation (2.5), there exist $\xi^1 \in \tilde{\mathfrak{g}}^1$ and an anti-symmetric transformation A of $\tilde{\mathfrak{g}}^1$ such that

$$\dot{\sigma}^1(t) = \xi^1 + A\sigma^1(t).$$

This equation is equivalent to $\ddot{\sigma}^1(t) = A\dot{\sigma}(t)$ and $\dot{\sigma}^1(0) = \xi^1$. Hence,

$$\dot{\sigma}^1(t) = e^{tA}\xi^1. \tag{2.7}$$

It is remarked that, since the transformation A is anti-symmetric, e^{tA} is the orthogonal transformation of $\tilde{\mathfrak{g}}^1$. We now decompose the linear space $\tilde{\mathfrak{g}}^1$ into the null space of A and its orthogonal one. Under this decomposition, we can express

$$e^{tA}\xi^1 = R(t)\eta + \zeta$$

where $\xi^1 = \eta + \zeta$ is ξ^1's decomposition, and $R(t)$ is a family of orthogonal transformations whose eigenvalues are not equal to one. Integrating this, we obtain

$$\sigma^1(t) = \sigma^1(0) + S(t)\eta + t\zeta, \qquad S(t) = \int_0^t R(s)\,ds.$$

To simplify the computation, we assume that γ passes through the identity e, that is, $\sigma^1(0) = 0$, $\sigma^2(0) = 0$. Then, $\sigma^1(t) = S(t)\eta + t\zeta$. By Lemma 2.1,

$$\begin{aligned}\sigma^2(t) &= \frac{1}{2}\int_0^t \sigma^1(s) \wedge \dot{\sigma}^1(s)\,ds = \frac{1}{2}\int_0^t (S(s)\eta + s\zeta) \wedge (\dot{S}(s)\eta + \zeta)\,ds \\ &= \frac{1}{2}\int_0^t S(s)\eta \wedge \dot{S}(s)\eta\,ds + \frac{1}{2}\int_0^t S(s)\eta \wedge \zeta\,ds + \frac{1}{2}\int_0^t s\zeta \wedge \dot{S}(s)\eta\,ds.\end{aligned}$$

Since the third term is equal to $-(1/2)tS(t)\eta \wedge \zeta + (1/2)\int_0^t S(s)\eta \wedge \zeta\,ds$, we obtain

$$\begin{aligned}\sigma^1(t) &= S(t)\eta + t\zeta, \\ \sigma^2(t) &= \frac{1}{2}(-tS(t) + 2T(t))\eta \wedge \zeta + \frac{1}{2}\int_0^t S(s)\eta \wedge R(s)\eta\,ds,\end{aligned} \tag{2.8}$$

where $T(t) = \int_0^t S(s)\,ds$.

We must remark that $S(t)$ and $T(t)/t$ are bounded. We can easily see this by calculating the case $R(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$.

2.5. Lines

Hereafter, we identify $\mathfrak{g}^1$ and $N/[N, N]$. Let $p : N \to \mathfrak{g}^1$ be the natural projection, that is, $p(\exp(\xi^1 + \xi^2)) = \xi^1$.

By Lemma 2.1, for any point $x \in N$ and any curve $\underline{c}$ in $\mathfrak{g}^1$ through $p(x)$, there exists a unique horizontal curve c in N through x such that $p(c) = \underline{c}$. We call this curve c the *p-lift* of $\underline{c}$.

Since the length of the p-lift c is equal to that of $\underline{c}$, we can easily obtain

Lemma 2.7. *The p-lift of lines in $\mathfrak{g}^1$ are the geodesics.*

We define a curve γ in N to be a *line* if γ is the p-lift of a line in $\mathfrak{g}^1$. It should be remarked that there is only one line joining two points, if any.

Let $d_{\mathfrak{g}^1}$ be Euclidean distance of $\mathfrak{g}^1$.

Lemma 2.8. *For any two points x, y in N,*

$$d_{\mathfrak{g}^1}(p(x), p(y)) \le d_N(x, y).$$

The equality holds if and only if x and y are on the same line.

Proof. Take a shortest geodesic γ joining x and y. The curve $p(\gamma)$ joins $p(x)$ and $p(y)$ and its length is equal to γ. Thus the above inequality holds.

If the equality holds, then $p(\gamma)$ is a shortest geodesic, that is, a part of a line in $\mathfrak{g}^1$. $\square$

Now we can characterize lines by the distance-property.

Proposition 2.2. *The following conditions on a geodesic γ in N are equivalent.*

(1) *γ is a line. In other words, $\gamma(t) = \gamma(0)\, e^{tX}$ for some vector $X \in \mathfrak{g}^1$.*

(2) *There exists a positive number C such that $d_N(\gamma(0), \gamma(t)) = C|t|$ for all t.*

Proof. (i)⇒(ii): This follows from Lemma 2.7.

(ii)⇒(i): Without loss of generality, we may assume $C = 1$.

By Lemma 2.6, the π-lift of γ satisfies the condition (ii). Hence we may consider only the case $N = \tilde{N}$. Furthermore, we may also assume $\gamma(0) = e$, since the left translation of the Lie group N preserves the distance d_N.

Under these assumptions, the condition (ii) is $d_N(e, a_{1/t}\gamma(t)) = 1$, where

$$a_{1/t}\gamma(t) = \exp\left(\frac{1}{t}\sigma^1(t) + \frac{1}{t^2}\sigma^2(t)\right).$$

By Theorem 2.4, σ^1 and σ^2 satisfy the equation (2.8).

Now we compute the limit of $a_{1/t}\gamma(t)$ when $t \to \infty$.

$$\begin{aligned}
&\frac{1}{t}\sigma^1(t) = \frac{1}{t}S(t)\eta + \zeta,\\
&\frac{1}{t^2}\sigma^2(t) = \left(-\frac{1}{2t}S(t) + \frac{1}{t^2}T(t)\right)\eta \wedge \zeta + \frac{1}{2t^2}\int_0^t S(s)\eta \wedge R(s)\eta\, ds.
\end{aligned}$$

Since $\|R(t)\eta\|$, $\|S(t)\eta\|$, and $\|T(t)\eta/t\|$ are bounded,

$$\lim_{t\to\infty} \frac{1}{t}\sigma^1(t) = \zeta, \qquad \lim_{t\to\infty} \frac{1}{t^2}\sigma^2(t) = 0.$$

Therefore we have

$$\lim_{t\to\infty} a_{1/t}\gamma(t) = \exp\zeta.$$

This implies $d_N(e, \exp\zeta) = 1$, and hence $\|\zeta\| = 1$.

On the other hand, we have assumed $C = 1$. This means that the velocity of γ is equal to 1, that is,

$$\|\dot{\sigma}^1(t)\|^2 = \|\xi^1\|^2 = \|\eta\|^2 + \|\zeta\|^2 = 1.$$

Thus $\eta = 0$. This completes the proof. □

Definition 2.3. Two lines γ_1 and γ_2 are said to be *parallel* if the function $t \mapsto d_N(\gamma_1(t), \gamma_2(t))$ is bounded.

The following two lemmas are due to Hamenstädt [5].

Lemma 2.9. *If two lines are parallel, their tangent vectors are equal in $\mathfrak{g}^1$.*

Proof. Assume the lines γ_1 and γ_2 are parallel. By Lemma 2.8, the function $t \to d_{\mathfrak{g}^1}(p(\gamma_1(t)), p(\gamma_2(t)))$ is bounded. Therefore, the lines $p(\gamma_1)$ and $p(\gamma_2)$ are parallel in the Euclidean space $\mathfrak{g}^1$. □

The converse is not true in general. Let Z be the center of N.

Lemma 2.10. *A point x is in Z if and only if, for each $v \in \mathfrak{g}^1$, the line through x tangent to v is parallel to the line through the identity e with the same direction.*

Proof. We denote $x = \exp \xi$, $\gamma_1(t) = e^{tv}$ and $\gamma_2(t) = x\, e^{tv}$. Then,

$$\begin{aligned} d_N(\gamma_1(t),\, \gamma_2(t)) &= d_N(e,\, e^{-tv} e^{\xi} e^{tv}) \\ &= d_N(e,\, \exp(\xi + t[\xi,\, v])). \end{aligned}$$

Hence two lines γ_1 and γ_2 are parallel if and only if $[\xi, v] = 0$. □

3. Isometries of 2-step Carnot groups

3.1. Invariant set

In this subsection, we aim to show the following lemma:

Lemma 3.1. *Every isometry of 2-step Carnot group N permutes the fibers of the projection p.*

By Proposition 2.2, we obtain

Lemma 3.2. *Every isometry of 2-step Carnot group N permutes the lines.*

By Lemma 2.10, we have

Lemma 3.3. *Every isometry of 2-step Carnot group N fixing the identity preserves the center.*

This lemma is due to Hamenstädt [5].

Let $\mathrm{I}(N)$ be the isometry group of the Carnot group N and $\mathrm{I}_0(N)$ be the isotropy subgroup of $\mathrm{I}(N)$ at the identity e.

We define a subset A of N to be *invariant* if $\varphi(A) = A$ for any $\varphi \in \mathrm{I}_0(N)$. Lemma 3.2 implies that the subset $\exp \mathfrak{g}^1$ is invariant and Lemma 3.3 states that the center Z is invariant.

Sublemma 1. If A is an invariant subset, then $\varphi(gA) = \varphi(g)A$ for any $g \in N$ and $\varphi \in \mathrm{I}_0(N)$.

Proof. We define the isometry ψ_g such that $\psi_g(x) = \varphi(g)^{-1}\varphi(gx)$. Then, $\psi_g(A) = A$ since $\psi_g \in \mathrm{I}_0(N)$. Thus $\varphi(gA) = \varphi(g)A$. □

To show Lemma 3.1, we only have to prove that the commutator subgroup $[N, N] = \exp \mathfrak{g}^2 = p^{-1}(0)$ is preserved by the isometries fixing e.

Sublemma 2. If A, B and A_k $(k = 1, 2, \ldots)$ are invariant subsets, then the subsets $\bigcup_k A_k$, $\bigcap_k A_k$, $A \setminus B$, AB and the closure $\bar{A}$ of A are invariant.

Proof. Since isometries are homeomorphisms, the subsets $\bigcup_k A_k$, $\bigcap_k A_k$, $A \setminus B$ and $\bar{A}$ are invariant.

By Sublemma 1, we have

$$\varphi(AB) = \bigcup_{g \in A} \varphi(gB) = \bigcup_{h \in \varphi(A)} hB = \bigcup_{h \in A} hB = AB.$$

$\square$

Let $\mathfrak{z}$ be the center of the Lie algebra $\mathfrak{g}$ and $\mathfrak{z}^1$ be the subspace of $\mathfrak{g}^1$ such that $\mathfrak{z}^1 = \mathfrak{z} \cap \mathfrak{g}^1$. Let $\mathfrak{h}^1$ be the orthogonal complement of $\mathfrak{z}^1$ in $\mathfrak{g}^1$.

Sublemma 3. The set $\exp \mathfrak{h}^1$ is invariant.

Proof. For $x \in N$, $xZ = p^{-1}(p(x) + \mathfrak{z}^1)$. Hence, by Lemma 2.8,

$$d_N(e,\, xZ) = d_{\mathfrak{g}^1}(0,\, p(x) + \mathfrak{z}^1).$$

There uniquely exsists $x' \in xZ$ such that $d_N(e,\, xZ) = d_N(e,\, x')$. In fact, if we denote the $\mathfrak{h}^1$-component of the vector $p(x)$ by v_x,

$$d_N(e,\, xZ) = d_{\mathfrak{g}^1}(0,\, v_x) = d_N(e,\, \exp v_x).$$

Since $\varphi(xZ) = \varphi(x)Z$ for $\varphi \in \mathrm{I}_0(N)$, we obtain $\varphi(\exp v_x) = \exp v_{\varphi(x)}$. This completes the proof. $\square$

Because $[\mathfrak{h}^1,\, \mathfrak{h}^1] = \mathfrak{g}^2$, the set $M = \exp(\mathfrak{h}^1 + \mathfrak{g}^2)$ is a Carnot subgroup of N. Hence $M = \bigcup_{k \geq 1} (\exp \mathfrak{h}^1)^k$. By Sublemma 2, we obtain

Sublemma 4. M is invariant.

Again applying Sublemma 2 to the set $[N,\, N] = M \cap Z$, the commutator group $[N,\, N]$ is invariant. This completes the proof of Lemma 3.1.

3.2. Isometries

In this subsection, we will determine isometry group of 2-step Carnot group.

Lemma 3.4. *Every isometry φ satisfies that $d_{\mathfrak{g}^1}(p\varphi(x),\, p\varphi(y)) = d_{\mathfrak{g}^1}(p(x),\, p(y))$ for all $x,\, y \in N$.*

Proof. Let y' be another endpoint of the p-lift through x of the line segment joining $p(x)$ and $p(y)$. Then $p(y') = p(y)$. By Lemma 2.8,

$$d_{\mathfrak{g}^1}(p(x),\, p(y)) = d_N(x,\, y') = d_N(\varphi(x),\, \varphi(y')).$$

Applying Lemmas 3.1 and 2.8, we have

$$d_N(\varphi(x),\, \varphi(y')) \geq d_{\mathfrak{g}^1}(p\varphi(x),\, p\varphi(y')) = d_{\mathfrak{g}^1}(p\varphi(x),\, p\varphi(y)).$$

Therefore
$$d_{\mathfrak{g}^1}(p\varphi(x),\, p\varphi(y)) \le d_{\mathfrak{g}^1}(p(x),\, p(y)).$$
Similarly, for the inverse mapping φ^{-1},
$$d_{\mathfrak{g}^1}(p\varphi(x),\, p\varphi(y)) \ge d_{\mathfrak{g}^1}(p(x),\, p(y)).$$
This completes the proof. $\square$

Theorem 3.1. *Every isometry of 2-step Carnot group N fixing the identity e is an automorphism of Lie group, and $\mathrm{I}_0(N)$ is isomorphic to a subgroup of the orthogonal group $O(\mathfrak{g}^1)$ of $\mathfrak{g}^1$.*

In particular, $\mathrm{I}(N)$ and $\mathrm{I}_0(N)$ have the structure of Lie group, and $\mathrm{I}_0(N)$ is compact.

Proof. Let G be a subgroup of $O(\mathfrak{g}^1)$ such that
$$G = \{A^1 \in O(\mathfrak{g}^1) \mid \tilde{A}^1(\mathrm{Ker}\Pi_2) = \mathrm{Ker}\Pi_2\},$$
where $\tilde{A}^1$ is a linear transformation of $\tilde{\mathfrak{g}}^2 = \bigwedge^2 \mathfrak{g}^1$ associated with A^1.

Since the linear space $\mathfrak{g}^2$ is isomorphic to $\tilde{\mathfrak{g}}^2/\mathrm{Ker}\Pi_2$, every element in G induces an automorphism of the graded Lie algebra $\mathfrak{g}$. Conversely, every element of $O(\mathfrak{g}^1)$ that induces an automorphism belongs to G.

Let A be an automorphism of $\mathfrak{g}$ induced by $A^1 \in G$. We can define an isometry φ such that $\varphi(e^{\xi}) = e^{A\xi}$ $(\xi \in \mathfrak{g})$. Hence there is an injection from G to $\mathrm{I}_0(N)$. Now we must show that this injection is isomorphic.

By Lemmas 3.1 and 3.2, for any isometry φ, there exists an affine mapping A^1 of $\mathfrak{g}^1$ such that $p \circ \varphi = A^1 \circ p$. By Lemma 3.4, the mapping A^1 is an isometry. If we assume $\varphi(e) = e$, then $A^1 \in G$.

It remains to show that φ coincides with the automorphism determined by A^1. Let x be a point in N such that
$$x = e^{X_1} e^{X_2} \cdots e^{X_k}, \qquad X_1, X_2, \ldots, X_k \in \mathfrak{g}^1.$$
If we consider the p-lift of a broken line in $\mathfrak{g}^1$ joining the origin, X_1, $X_1 + X_2$, $\ldots$, and $X_1 + X_2 + \cdots + X_k$, then we obtain
$$\varphi(x) = e^{A^1 X_1} e^{A^1 X_2} \cdots e^{A^1 X_k} = \varphi(e^{X_1})\varphi(e^{X_2}) \cdots \varphi(e^{X_k}).$$
Hence φ is an automorphism of Lie group N. $\square$

4. The r-step Carnot groups $(r > 2)$

4.1. The differentiability of isometries

In [5], the argument on isometry group needs the differntiability of isometries which she proved under the assumption that all the geodesics are normal.

In this section, to establish the differentiability of isometries of general Carnot groups without the normality assumption, we use the following theorem. (See [2])

Theorem 4.1 (Montgomery-Zippin). *Let X be a finite dimensional, locally compact, connected and locally connected metric space. If the group G of the isometries of X is transitive, then G is a Lie group with finitely many connected components.*

By the above theorem, the isometry group $\mathrm{I}(N)$ of a Carnot group N is a Lie group and $\mathrm{I}_0(N)$ is its closed subgroup. The group N itself can be identified with the subgroup of $\mathrm{I}(N)$, since the metric is left-invariant.

Proposition 4.1. *The isometric action of* $\mathrm{I}(N)$ *on N is smooth. In particular, every isometry is smooth.*

Proof. Let σ be the smooth injective mapping $N \to \mathrm{I}(N); g \mapsto L_g$ where L_g is the left translation by g. And let ρ be the canonical projection $\mathrm{I}(N) \to \mathrm{I}(N)/\mathrm{I}_0(N)$. Then it is easy to see that the composition $\rho \circ \sigma$ is a diffeomorphism $N \to \mathrm{I}(N)/\mathrm{I}_0(N)$. $\square$

Let $\varphi^* : T^*N \to T^*N$ be the bundle map induced by a smooth mapping $\varphi : N \to N$. If φ is an isometry, then φ^* preserves the Hamiltonian function H, which is defined in 2.3.

Therefore we have

Lemma 4.1. *Every isometry permutes the normal geodesics.*

Hence, by the argument in [5, §8], we obtain

Theorem 4.2. *Every isometry of Carnot groups fixing the identity e is a Lie group automorphism.*

Acknowledgement. The author wishes to express his gratitude to Kenji Fukaya for his encouragement and valuable suggestion, and also thanks the referee for the advice.

Japan Branch
Zurich Life Insurance Company LTD
Shinanomachi Rengakan
35 Shinanomachi, Shinjuku-ku
Tokyo 160-0016, Japan
e-mail: jpziki@zurich.co.jp

References

[1] C. Golé and R. Karidi, *A note on Carnot geodesics in nilpotent Lie groups*, preprint.

[2] M. Gromov, *Groups of polynomial growth and expanding maps*, Publ. Math. de l'IHES, 1980.

[3] ———, *Structures métriques pour les variétés riemanniennes*, rédigé par J. Lafontaine et P. Pansu, Cedic-Nathan, 1981. (English translation: Metric Structures for Riemannian and Non-Riemannian Spaces, Birkhäuser, 1999.)

[4] M. Gromov, *Carnot-Carathéodory spaces seen from within*, In: Sub-Riemannian Geometry, A. Bellaïche and J.-J. Risler (eds.), Birkhäuser, 1996.

[5] U. Hamenstädt, *Some regularity theorems for Carnot-Carathéodory metrics*, J. Differential Geom. **32** (1990), 819–850.

[6] L. Hsu, *Calculus of variations via the Griffiths formalism*, J. Differential Geom. **36** (1992), 551–589.

[7] I. Kishimoto, *The Morse index theorem for Carnot-Carathéodory spaces*, J. Math. Kyoto Univ. **38** (1998), 287–293.

[8] W. Liu and H. J. Sussmann, *Shortest paths for sub-Riemannian metrics on rank-two distributions*, Memoirs of AMS **564**, 1995.

[9] J. Mitchell, *On Carnot-Carathéodory metrics*, J. Differential Geom. **21** (1985), 35–45.

[10] D. Montgomery and L. Zippin, *Topological transformation groups*, Interscience, 1955.

[11] R. Montgomery, *Abnormal minimizers*, SIAM J. Control and Optim. **32** (1994), 1605–1620.

[12] P. Pansu, *Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. Math. **129** (1989), 1–60.

[13] R. S. Strichartz, *Sub-Riemannian geometry*, J. Differential Geom. **24** (1986), 221–263.

[14] ———, *Corrections to "Sub-Riemannian geometry"*, J. Differential Geom. **30** (1989), 595–596.