

COMMUNICATIONS IN
ANALYSIS AND GEOMETRY
Volume ?, Number ?, 1–33, 201?

Extremal polynomials in stratified groups

ENRICO LE DONNE, GIAN PAOLO LEONARDI, ROBERTO MONTI, AND
DAVIDE VITTONI

We introduce a family of extremal polynomials associated with the prolongation of a stratified nilpotent Lie algebra. These polynomials are related to a new algebraic characterization of abnormal sub-Riemannian extremals in stratified nilpotent Lie groups. They satisfy a set of remarkable structure relations that are used to integrate the adjoint equations, in both normal and abnormal case.

1	Introduction	1
2	Structure formulas for extremal polynomials	5
3	Characterization of abnormal extremals	19
4	The abnormal set in the free group of rank 2 and step 4	24
5	Spiral-like Goh extremals	29
	Acknowledgments	31
	References	31

1. Introduction

Let G be a stratified nilpotent Lie group of dimension n and rank r . The Lie algebra $\mathfrak{g} = \text{Lie}(G)$ has the stratification $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s$, where s is the step of the algebra, $\mathfrak{g}_i = [\mathfrak{g}_{i-1}, \mathfrak{g}_1]$ for $i = 2, \dots, s$, and $\mathfrak{g}_i = \{0\}$ for $i > s$. The rank of \mathfrak{g} is $r = \dim(\mathfrak{g}_1)$. Let X_1, \dots, X_n be a basis of \mathfrak{g} adapted to the stratification. When it is connected and simply connected, we can identify the group G with \mathbb{R}^n via exponential coordinates of the second type induced by the basis X_1, \dots, X_n , and we identify \mathfrak{g} with the corresponding Lie algebra of left invariant vector fields in \mathbb{R}^n .

Let $\text{Prol}(\mathfrak{g}) = \bigoplus_{k \leq s} \mathfrak{g}_k$ be the Tanaka prolongation of \mathfrak{g} , see [16]. Even though this is not strictly needed in our argument, an explicit construction is briefly recalled in Section 2. We extend X_1, \dots, X_n to a basis $\{X_j\}_{j \leq n}$ of the prolongation and to each $j \leq n$ we assign the degree $d(j) = k$ if and only if $X_j \in \mathfrak{g}_k$. Then we assume that the basis is adapted to the graduation: this means that $i < j$ implies $d(i) \leq d(j)$. When the prolongation is finite dimensional, the index j ranges in a finite set, $m \leq j \leq n$ for some $m \in \mathbb{Z}$. With abuse of notation, we denote the basis $\{X_j\}_{m \leq j \leq n}$ by $\{X_j\}_{j \leq n}$, as in the infinite dimensional case. Let $c_{ij}^k \in \mathbb{R}$ be the structure constants of $\text{Prol}(\mathfrak{g})$ associated with the basis $\{X_j\}_{j \leq n}$. Namely, for all $i, j \in \mathbb{Z}$ with $i, j \leq n$ we have

$$(1.1) \quad [X_i, X_j] = \sum_{k \leq n} c_{ij}^k X_k.$$

The sum is always finite, because each stratum \mathfrak{g}_k of $\text{Prol}(\mathfrak{g})$ is finite dimensional and $c_{ij}^k = 0$ if $d(k) \neq d(i) + d(j)$.

In this paper, we introduce a family of *extremal polynomials* $P_j^v(x)$, $j \leq n$ and $x \in \mathbb{R}^n$, associated with the basis $\{X_j\}_{j \leq n}$ of $\text{Prol}(\mathfrak{g})$. They depend linearly on a parameter $v \in \mathbb{R}^n$, see Definition 2.1. Extremal polynomials satisfy the following structure formulas.

Theorem 1.1. *For any $v \in \mathbb{R}^n$, $i = 1, \dots, n$, and $j \in \mathbb{Z}$ with $j \leq n$ there holds*

$$(1.2) \quad X_i P_j^v(x) = \sum_{k \leq n} c_{ij}^k P_k^v(x), \quad x \in \mathbb{R}^n.$$

Modulo the value at $x = 0$, extremal polynomials are uniquely determined by the family of identities (1.2). These structure formulas are the core of the paper and of our main technical result, Theorem 2.2. They first appeared in [8], but only in the case of *free groups*. In [8], the formulas were obtained a posteriori, as a consequence of certain algebraic identities, however only for $j = 1, \dots, n$ with no reference to the prolongation and only for the coordinates related to a Hall basis.

It is possible to give a coordinate-free definition of extremal polynomials, see [9]. However, for the applications considered in Sections 4 and 5, as well as for the applications given in [8], it was crucial to possess explicit formulae for the polynomials (Definition 2.1).

Our interest in extremal polynomials origins in the regularity problem of sub-Riemannian length minimizing curves, one of the main open problems

in the field (see [11], [2], [13]). Let M be a differentiable manifold and $\mathcal{D} \subset TM$ a bracket generating distribution. A Lipschitz curve $\gamma : [0, 1] \rightarrow M$ is horizontal if $\dot{\gamma}(t) \in \mathcal{D}(\gamma(t))$ for a.e. $t \in [0, 1]$. Fixing a quadratic form on \mathcal{D} , one can define the length of horizontal curves. Length minimizing curves may be either normal extremals or abnormal extremals: while normal extremals are always smooth, abnormal ones are a priori only Lipschitz continuous. Abnormal extremals depend only on the structure (M, \mathcal{D}) : they are precisely the singular points of the end-point mapping.

In Section 3, we use extremal polynomials and Theorem 1.1 to give an algebraic characterization of abnormal extremals in stratified nilpotent Lie groups (Carnot groups). This is of special interest because, by Mitchell’s theorem, Carnot groups are the infinitesimal model of equiregular sub-Riemannian structures.

Let $\vartheta_1, \dots, \vartheta_n$ be a basis of 1-forms of \mathfrak{g}^* , the dual of \mathfrak{g} . A curve $\lambda \in \text{Lip}([0, 1]; \mathfrak{g}^*)$ is given by coordinates $\lambda_1, \dots, \lambda_n \in \text{Lip}([0, 1])$ such that $\lambda = \lambda_1 \vartheta_1 + \dots + \lambda_n \vartheta_n$. If λ is the dual curve of a normal or abnormal extremal $\gamma : [0, 1] \rightarrow G = \mathbb{R}^n$, then

$$(1.3) \quad \dot{\lambda}_i = - \sum_{k=1}^n \sum_{j=1}^r c_{ij}^k \dot{\gamma}_j \lambda_k \quad \text{a.e. on } [0, 1], \quad i = 1, \dots, n.$$

See Section 3 and Theorem 3.1 for more details. In Theorem 3.4, we use the structure formulas (1.2) to integrate the system of adjoint equations (1.3). The solutions are

$$(1.4) \quad \lambda_i(t) = P_i^v(\gamma(t)), \quad i = 1, \dots, n,$$

where $v \in \mathbb{R}^n$ is such that $v_i = \lambda_i(0)$. Thus, we can prove the following theorem (see Theorem 3.8 for the complete statement).

Theorem 1.2. *Let $G = \mathbb{R}^n$ be a stratified nilpotent Lie group and let $\gamma : [0, 1] \rightarrow G$ be a horizontal curve with $\gamma(0) = 0$. Then, the following statements are equivalent:*

- (A) *The curve γ is an abnormal extremal.*
- (B) *There exist $v \in \mathbb{R}^n$, $v \neq 0$, such that $P_i^v(\gamma(t)) = 0$ for all $t \in [0, 1]$ and for all $i \leq r$.*

Abnormal extremals are precisely the horizontal curves lying inside algebraic varieties defined via extremal polynomials. This result extends [8, Theorem 1.1] because it applies to *nonfree* Carnot groups. It also improves that

result because abnormal curves are shown to be in algebraic varieties *smaller* than those considered in [8]. Notice that in (B) also indexes $i \leq 0$ related to the Tanaka prolongation are involved.

Extremal polynomials can be also used to find explicit prime integrals of the Hamiltonian system associated with the group. In fact, a normal extremal (γ, λ) , where γ is a curve in the group and λ is the dual curve in the cotangent bundle, solves the system of equations

$$(1.5) \quad \dot{\gamma} = \frac{\partial H}{\partial \lambda}(\gamma, \lambda), \quad \dot{\lambda} = -\frac{\partial H}{\partial \gamma}(\gamma, \lambda),$$

where H is the natural Hamiltonian function. Thanks to the Lie group structure, by Noether’s theorem the Hamiltonian system (1.5) has n prime integrals, where n is the dimension of the group. In the coordinates introduced above, the equation $\dot{\lambda} = -\partial H/\partial \gamma$ becomes precisely equation (1.3). Thus, the solutions (1.4) provide n prime integrals for the system (1.5).

The role of Tanaka prolongation in our theory is rather subtle. In a first attempt, we searched for an integration algorithm inverting the differentiation process that now is established by the structure formulas (1.2). Part of statement (B) in Theorem 1.2 is that an abnormal extremal γ satisfies $P_i^v(\gamma) = 0$ for some $v \neq 0$ and for all $i = 1, \dots, r$. Thus, for any $j = 1, \dots, n$ also the functions

$$(1.6) \quad B_{ij}^v(t) = \int_0^t P_i^v(\gamma(s)) \dot{\gamma}_j(s) ds, \quad t \in [0, 1],$$

vanish. Now, if the functions $B_{ij}^v(t)$ are *polynomials of the coordinates* $\gamma_1(t), \dots, \gamma_n(t)$, then we get new polynomials vanishing along the curve γ . However, this may or may not be the case and we were not able to understand this phenomenon in general. In Section 4, we show an example where the integration process (1.6) succeeds and, as a matter of fact, provides precisely the extremal polynomials associated with the layer \mathfrak{g}_0 of the Tanaka prolongation, the polynomials (4.49).

Tanaka prolongation has the following role in our argument. If P_j^v is an extremal polynomial associated with the layer $\ell \leq 0$ of the prolongation (i.e., $d(j) = \ell$) and γ is a horizontal curve, then by the structure formulas (1.2) we have

$$\frac{d}{dt} P_j^v(\gamma(t)) = \sum_{i=1}^r \sum_{k \leq n} c_{ij}^k \dot{\gamma}_i(t) P_k^v(\gamma(t)), \quad t \in [0, 1],$$

where the sum in k ranges over indices with $d(k) \geq \ell + 1$, because of the stratification. If the right hand side vanishes, then also the left hand side does and inductively, starting from the layer with $\ell = 1$ descending, we can prove that abnormal extremals are in the zero set of all extremal polynomials given by the layers with $\ell \leq 0$ of the prolongation (see Theorem 3.8).

In the final part of the paper, we show two applications of the theory.

In Section 4, we develop a technique to construct a nontrivial algebraic set containing *all* abnormal extremals passing through one point. This is related to the problem of estimating the size of the set of regular values of the end-point mapping (see [2] and [11, Section 10.2]). In sub-Riemannian manifolds with a distribution of corank 1, the image of the set of *length minimizing* abnormal extremals starting from one point has zero Lebesgue measure; on the other hand, independently from corank, the image of *strictly* abnormal length minimizing extremals has empty interior (see Corollary 3 in [15] and [1]).

Our technique seems to work when the prolongation is sufficiently large. In this case, for each abnormal curve γ there is at least one parameter $v \in \mathbb{R}^n$ and *many* indexes j such that $P_j^v(\gamma) = 0$. It is then possible to find a polynomial Q independent of v such that $Q(\gamma) = 0$ for any abnormal curve γ passing through one fixed point. We describe the technique in detail in the case of the free nilpotent Lie group of rank 2 and step 4. However, it can be implemented in many other examples and it is likely to work in any nonrigid group.

Finally, in Section 5 we construct a 64-dimensional Lie group possessing a spiral-like Goh extremal whose tangents at the singular point are all lines. This example points out a limitation of the shortening technique introduced in [10] and developed in [12] and [14]. It is also interesting in relation to the examples of *nonrectifiable* spiral-like rigid paths studied in [19] and [18], whereas our spiral-like extremal has finite length. As Example 6.4 in [8], it also shows that from the first and second order conditions (Pontryagin Maximum Principle and Goh condition) it is not possible to deduce any further regularity for abnormal minimizers beyond the Lipschitz regularity.

2. Structure formulas for extremal polynomials

Let X_1, \dots, X_n be a basis of $\mathfrak{g} = \text{Lie}(G)$ adapted to the stratification. We identify the group G with \mathbb{R}^n via exponential coordinates of the second type induced by the basis X_1, \dots, X_n . Namely, for any $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ we

have

$$(2.7) \quad x = e^{x_1 X_1} \circ \dots \circ e^{x_n X_n}(0) = \exp(x_n X_n) \cdot \dots \cdot \exp(x_1 X_1).$$

Above, $\exp : \mathfrak{g} \rightarrow G$ is the exponential mapping, \cdot is the group law in $G = \mathbb{R}^n$, and $\gamma(t) = e^{tX}(x)$, $t \in \mathbb{R}$, is the solution of the Cauchy Problem

$$\begin{cases} \dot{\gamma} = X(\gamma) \\ \gamma(0) = x. \end{cases}$$

In simply connected nilpotent groups, the exponential mapping is a global diffeomorphism. After the identification (2.7), the Lie algebra \mathfrak{g} is isomorphic to a Lie algebra of vector fields in \mathbb{R}^n that are left invariant with respect to the group law \cdot .

We recall the construction of the Tanaka prolongation. First, we define the vector space of all strata-preserving derivations of \mathfrak{g} :

$$\mathfrak{g}_0 = \left\{ \phi \in \bigoplus_{i=1}^s \mathfrak{g}_i \otimes \mathfrak{g}_i^* : \phi([X, Y]) = [\phi(X), Y] + [X, \phi(Y)], \quad X, Y \in \mathfrak{g} \right\}.$$

Recall that $\mathfrak{g}_i \otimes \mathfrak{g}_i^*$ is canonically isomorphic to $\text{End}(\mathfrak{g}_i)$. The direct sum vector space $\bigoplus_{i=0}^s \mathfrak{g}_i$ is a graded Lie algebra with the bracket

$$[\phi, X] = -[X, \phi] = \phi(X), \quad \text{for all } \phi \in \mathfrak{g}_0 \text{ and } X \in \mathfrak{g},$$

and with the natural bracket on \mathfrak{g} and \mathfrak{g}_0 .

By induction, assume that we have a vector space $\bigoplus_{i=1-k}^s \mathfrak{g}_i$ for some $k \geq 1$ and assume that the bracket $[\phi, X] = \phi(X)$ is already defined for all $\phi \in \bigoplus_{i=1-k}^s \mathfrak{g}_i$ and $X \in \mathfrak{g}$. Then we define the vector space of all derivations $\phi : \mathfrak{g} \rightarrow \mathfrak{g}_{1-k} \oplus \dots \oplus \mathfrak{g}_{s-k}$ such that $\phi(\mathfrak{g}_i) \subset \mathfrak{g}_{i-k}$:

$$\mathfrak{g}_{-k} = \left\{ \phi \in \bigoplus_{i=1}^s \mathfrak{g}_{i-k} \otimes \mathfrak{g}_i^* : \phi([X, Y]) = [\phi(X), Y] + [X, \phi(Y)], \quad X, Y \in \mathfrak{g} \right\}.$$

Recall that $\mathfrak{g}_{i-k} \otimes \mathfrak{g}_i^*$ is isomorphic to $\text{Hom}(\mathfrak{g}_i; \mathfrak{g}_{i-k})$. As above, the direct sum vector space $\bigoplus_{i=-k}^s \mathfrak{g}_i$ is a graded Lie algebra with the bracket

$$[\phi, X] = -[X, \phi] = \phi(X), \quad \text{for all } \phi \in \mathfrak{g}_{-k} \text{ and } X \in \mathfrak{g},$$

with the natural bracket on \mathfrak{g} , and with the bracket $[\phi, \psi]$ for $\phi, \psi \in \mathfrak{g}_{-k} \oplus \dots \oplus \mathfrak{g}_0$ defined by

$$[\phi, \psi](X) = [[\phi, X], \psi] + [\phi, [\psi, X]], \quad X \in \mathfrak{g}.$$

This inductive construction may or may not end after a finite number of steps, i.e., we have either $\mathfrak{g}_{-k} = \{0\}$ for some $k \geq 1$ or $\mathfrak{g}_{-k} \neq \{0\}$ for all $k \geq 1$. In both cases, we let

$$(2.8) \quad \text{Prol}(\mathfrak{g}) = \bigoplus_{k \leq s} \mathfrak{g}_k.$$

$\text{Prol}(\mathfrak{g})$ is a *graded* Lie algebra, called Tanaka prolongation of \mathfrak{g} . Namely, we have

$$(2.9) \quad [\mathfrak{g}_i, \mathfrak{g}_j] \subset \mathfrak{g}_{i+j}, \quad \text{for all } i, j \in \mathbb{Z} \text{ with } i, j, i+j \leq s.$$

This is the unique property that we need in the proof of the structure theorem of extremal polynomials, Theorem 2.2. In fact, we need (2.9) only for $j = 1, \dots, s$. We do not specifically need the Tanaka prolongation but only a graded Lie algebra extending \mathfrak{g} and satisfying (2.8) and (2.9). Among all such extensions $\text{Prol}(\mathfrak{g})$ is the largest one.

In general, the explicit computation of $\text{Prol}(\mathfrak{g})$ is difficult. When \mathfrak{g} is a *free* nilpotent Lie algebra of step $s \geq 3$, then we have $\text{Prol}(\mathfrak{g}) = \mathfrak{g}_0 \oplus \mathfrak{g}$, with the exceptional case of the step 3 and rank 2 free Lie algebra (see [17]).

Let $\{X_j\}_{j \leq n}$ be a basis of $\text{Prol}(\mathfrak{g})$. For integers $j_0, j_1, j_2, \dots, j_k \leq n$, we let

$$(2.10) \quad [X_{j_0}, X_{j_1}, X_{j_2}, X_{j_3}, \dots, X_{j_k}] = [\dots [[X_{j_0}, X_{j_1}], X_{j_2}], X_{j_3}], \dots, X_{j_k}].$$

Then, for $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathcal{I} = \mathbb{N}^n = \{0, 1, 2, \dots\}^n$, we define the iterated commutator of X_1, \dots, X_n

$$(2.11) \quad [\cdot, X_\alpha] = [\cdot, \underbrace{X_1, \dots, X_1}_{\alpha_1 \text{ times}}, \underbrace{X_2, \dots, X_2}_{\alpha_2 \text{ times}}, \dots, \underbrace{X_n, \dots, X_n}_{\alpha_n \text{ times}}].$$

Here, only elements X_1, \dots, X_n of the basis of \mathfrak{g} are involved. We agree that $[\cdot, X_{(0, \dots, 0)}] = \text{Id}$. The *generalized structure constants* $c_{i\alpha}^k \in \mathbb{R}$, with $\alpha \in \mathcal{I} =$

\mathbb{N}^n and $i, k \in \mathbb{Z}$ such that $i, k \leq n$, are defined via the relation

$$(2.12) \quad [X_i, X_\alpha] = \sum_{k \leq n} c_{i\alpha}^k X_k.$$

The above sum is always finite. In fact, letting $d(\alpha) = \sum_{j=1}^n \alpha_j d(j)$, we have $c_{i\alpha}^k = 0$ if $d(k) \neq d(i) + d(\alpha)$.

For $\alpha \in \mathcal{I}$ and $i \in \mathbb{Z}$ with $i \leq n$, we define the linear mapping $\phi_{i\alpha} \in \text{Hom}(\mathbb{R}^n; \mathbb{R})$

$$(2.13) \quad \phi_{i\alpha}(v) = \frac{(-1)^{|\alpha|}}{\alpha!} \sum_{k \leq n} c_{i\alpha}^k v_k, \quad v = (v_1, \dots, v_n) \in \mathbb{R}^n.$$

Above, we let $\alpha! = \alpha_1! \cdot \dots \cdot \alpha_n!$. Moreover, we agree that $v_k = 0$ for $k \leq 0$.

Definition 2.1. For each $i \in \mathbb{Z}$ with $i \leq n$ and $v \in \mathbb{R}^n$, we call the polynomial $P_i^v : \mathbb{R}^n \rightarrow \mathbb{R}$

$$(2.14) \quad P_i^v(x) = \sum_{\alpha \in \mathcal{I}} \phi_{i\alpha}(v) x^\alpha, \quad x \in \mathbb{R}^n,$$

extremal polynomial of the Lie algebra $\text{Prol}(\mathfrak{g})$ with respect to the basis $\{X_j\}_{j \leq n}$.

For any finite set of multi-indexes $\mathcal{A} \subset \mathcal{I}$, consider the polynomial

$$P(x) = \sum_{\alpha \in \mathcal{A}} c_\alpha x^\alpha,$$

where $c_\alpha \in \mathbb{R}$ for all $\alpha \in \mathcal{A}$. The homogeneous degree of the polynomial is $d(P) = \max \{d(\alpha) : \alpha \in \mathcal{A} \text{ such that } c_\alpha \neq 0\}$. We say that the polynomial is homogeneous of degree $k \geq 0$ if $d(\alpha) = k$ for all $\alpha \in \mathcal{A}$ such that $c_\alpha \neq 0$.

Extremal polynomials P_i^v are indeed polynomials. In fact, if $d(i) + d(\alpha) \neq d(k)$ then we have $c_{i\alpha}^k = 0$. Moreover, we have $d(k) \leq d(n) = s$ and therefore the sum in (2.14) ranges over the $\alpha \in \mathcal{I}$ such that $d(\alpha) \leq s - d(i)$. So we have $d(P_i^v) \leq s - d(i)$.

The following theorem is the main result of the paper.

Theorem 2.2. *For any $v \in \mathbb{R}^n$, $i = 1, \dots, n$, and $j \in \mathbb{Z}$ with $j \leq n$ there holds*

$$(2.15) \quad X_i P_j^v = \sum_{k \leq n} c_{ij}^k P_k^v.$$

Moreover, the polynomials $\{P_j^v\}_{j \leq n}$ are uniquely determined by (2.15) for $i = 1, \dots, r$ and $P_j^v(0) = v_j$ for $j \leq n$.

The identity $P_j^v(0) = v_j$ is proved in (2.36) below. The uniqueness follows from this observation: if f is a smooth function on $G = \mathbb{R}^n$ such that $X_i f = 0$ for all $i = 1, \dots, r$, then f is constant. Now, let $\{Q_j^v\}_{j \leq n}$ be a family of polynomials satisfying (2.15) for $i = 1, \dots, r$, and $Q_j^v(0) = v_j$ for $j \leq n$. Then, for $d(j) = s$ we have $X_i P_j^v = 0 = X_i Q_j^v$ and thus $P_j^v - Q_j^v = P_j^v(0) - Q_j^v(0) = 0$. Assume by induction that $P_j^v = Q_j^v$ for $d(j) \geq \ell + 1$ and take $j \leq n$ such that $d(j) = \ell$. From (2.15), we have

$$X_i P_j^v = \sum_{k \leq n} c_{ij}^k P_k^v = \sum_{k \leq n} c_{ij}^k Q_k^v = X_i Q_j^v, \quad i = 1, \dots, r,$$

and thus $P_j^v - Q_j^v = P_j^v(0) - Q_j^v(0) = 0$.

Before starting the proof of (2.15), we need three preliminary lemmas. For $i = 1, \dots, n$, let $G_i \subset G = \mathbb{R}^n$ be the subgroup of G

$$\begin{aligned} G_i &= \{x \in G = \mathbb{R}^n : x_1 = \dots = x_{i-1} = 0\} \\ &= \{e^{x_i X_i} \circ \dots \circ e^{x_n X_n}(0) \in G = \mathbb{R}^n : x_i, \dots, x_n \in \mathbb{R}\}. \end{aligned}$$

We clearly have $G_1 = G$. The following lemma is well-known and we omit the proof.

Lemma 2.3. *Let X_1, \dots, X_n be vector fields in \mathbb{R}^n satisfying (2.7). Then:*

- (i) *For all $i = 1, \dots, n$ and for all $x \in G_i$ we have $X_i(x) = \partial/\partial x_i$.*
- (ii) *There exist functions $f_{i\ell} : \mathbb{R}^n \rightarrow \mathbb{R}$, $i, \ell = 1, \dots, n$, such that*

$$X_i(x) = \frac{\partial}{\partial x_i} + \sum_{\ell: d(\ell) > d(i)} f_{i\ell}(x) \frac{\partial}{\partial x_\ell}, \quad x \in G = \mathbb{R}^n.$$

In the following two lemmas, we prove formula (2.15) in two simplified situations that will serve as base for a long induction argument.

Lemma 2.4. *For any $i = 1, \dots, n$, $j \in \mathbb{Z}$ with $j \leq n$, and $v \in \mathbb{R}^n$ there holds*

$$(2.16) \quad X_i P_j^v = \sum_{k \leq n} c_{ij}^k P_k^v \quad \text{on } G_i.$$

Proof. For $x \in G_i$, we have

$$(2.17) \quad P_j^v(x) = \sum_{\ell \leq n} \sum_{\alpha \in \mathcal{I}_i} c_\alpha c_{j\alpha}^\ell v_\ell x^\alpha,$$

where we let $\mathcal{I}_i = \{\alpha \in \mathcal{I} : \alpha_1 = \dots = \alpha_{i-1} = 0\}$ and

$$c_\alpha = \frac{(-1)^{|\alpha|}}{\alpha!}.$$

By (i) in Lemma 2.3, on G_i we have $X_i = \partial/\partial x_i$. Then, differentiating (2.17) we obtain

$$(2.18) \quad X_i P_j^v(x) = - \sum_{\ell \leq n} \sum_{\beta \in \mathcal{I}_i} c_\beta c_{j,\beta+e_i}^\ell v_\ell x^\beta, \quad x \in G_i.$$

On the other hand, for $x \in G_i$ and $v \in \mathbb{R}^n$ we have

$$(2.19) \quad \sum_{k \leq n} c_{ij}^k P_k^v(x) = \sum_{\ell, k \leq n} \sum_{\beta \in \mathcal{I}_i} c_\beta c_{ij}^k c_{k\beta}^\ell v_\ell x^\beta.$$

By (2.18) and (2.19), proving the claim (2.16) is equivalent to show that for all $\ell \leq n$ and for all $\beta \in \mathcal{I}_i$ we have

$$(2.20) \quad -c_{j,\beta+e_i}^\ell = \sum_{k \leq n} c_{ij}^k c_{k\beta}^\ell.$$

As $\beta \in \mathcal{I}_i$, there holds $-[X_j, X_{\beta+e_i}] = [[X_i, X_j], X_\beta]$, and thus, by (2.12) and (1.1), we obtain

$$(2.21) \quad - \sum_{\ell \leq n} c_{j,\beta+e_i}^\ell X_\ell = -[X_j, X_{\beta+e_i}] = [[X_i, X_j], X_\beta] = \sum_{k \leq n} c_{ij}^k [X_k, X_\beta] = \sum_{\ell, k \leq n} c_{ij}^k c_{k\beta}^\ell X_\ell.$$

Since $\{X_\ell\}_{\ell \leq n}$ is a basis, this proves (2.20). \square

Lemma 2.5. *For all $j \in \mathbb{Z}$ with $j \leq n$, $i = 1, \dots, n$ such that $d(i) = s$, and $v \in \mathbb{R}^n$ there holds*

$$(2.22) \quad X_i P_j^v = \sum_{k \leq n} c_{ij}^k P_k^v.$$

Proof. When $j = 1, \dots, n$, the left and right hand sides of (2.22) are both identically 0. Recall that we have $d(P_j^v) \leq s - d(j) < s$ and thus $X_i P_j^v = 0$. We prove the claim (2.22) for any $j \leq n$, in particular for $j \leq 0$. As $d(i) = s$, by (ii) in Lemma 2.3 we have $X_i = \partial/\partial x_i$ on $G = \mathbb{R}^n$. Then we have formula (2.18) with $\beta \in \mathcal{I}$ replacing $\beta \in \mathcal{I}_i$. We also have formula (2.19), again with $\beta \in \mathcal{I}$ replacing $\beta \in \mathcal{I}_i$. Then we are reduced to check identity (2.20) for all $\beta \in \mathcal{I}$.

We claim that, for any $\beta \in \mathcal{I}$, we have

$$(2.23) \quad [X_j, X_{\beta+e_i}] = -[[X_i, X_j], X_\beta].$$

We prove (2.23) by induction on the length of β and we assume without loss of generality that $i = n$. Notice that the vector field X_n commutes with any vector field X_k , $k = 1, \dots, n$.

Assume that $\beta_k \neq 0$ and $\beta_h = 0$ for $h > k$. Then, by the Jacobi identity, by $[X_k, X_n] = 0$, and by formula (2.23) for $\beta - e_k$ replacing β (and with $i = n$), we have

$$\begin{aligned} [X_j, X_{\beta+e_n}] &= [[X_j, X_{\beta-e_k}], X_k], X_n] \\ &= -[[X_n, [X_j, X_{\beta-e_k}]], X_k] \\ &= [[X_j, X_{\beta-e_k+e_n}], X_k] \\ &= -[[X_n, X_j], X_{\beta-e_k}], X_k] \\ &= -[[X_n, X_j], X_\beta], \end{aligned}$$

and we are finished.

Now, using (2.23) the proof can be concluded as in (2.20)-(2.21). \square

Proof of Theorem 2.2. The proof of (2.15) is a triple nested induction. We fix the vector $v \in \mathbb{R}^n$ and we let $P_j^v = P_j$. The claim in Theorem 2.2 reads

$$X_i P_j = \sum_{k \leq n} c_{ij}^k P_k \quad \text{on } G_\ell \text{ for all } i, \ell = 1, \dots, n \text{ and } j \leq n.$$

The first induction is descending on $i = 1, \dots, n$ and the base of induction is

$$X_n P_j = \sum_{k \leq n} c_{nj}^k P_k \quad \text{on } G \text{ for all } j \leq n.$$

This statement holds by Lemma 2.5, because $d(n) = s$.

The first inductive assumption is the following

$$(2.24) \quad X_h P_j = \sum_{k \leq n} c_{hj}^k P_k \quad \text{on } G \text{ for all } h > i \text{ and for all } j \leq n.$$

Our goal is to prove identity (2.24) also for $h = i$. The proof of this claim is a descending induction on $j \leq n$. The base of induction for $j = n$ is the following

$$X_i P_n = \sum_{k \leq n} c_{in}^k P_k \quad \text{on } G.$$

This identity holds because we have $X_i P_n = 0$ (P_n is constant), and $c_{in}^k = 0$ for any $i \geq 1$ and for all k .

The second inductive assumption is the following

$$(2.25) \quad X_i P_h = \sum_{k \leq n} c_{ih}^k P_k \quad \text{on } G_\ell \text{ for all } h > j \text{ and for all } \ell = 1, \dots, n.$$

The goal is to prove identity (2.25) also for $h = j$. The proof is a descending induction on $\ell = 1, \dots, i$. The base of induction for $\ell = i$ is the following

$$X_i P_j = \sum_{k \leq n} c_{ij}^k P_k \quad \text{on } G_i,$$

that is satisfied by Lemma 2.4.

The third inductive assumption is

$$(2.26) \quad X_i P_j = \sum_{k \leq n} c_{ij}^k P_k \quad \text{on } G_{\ell+1}$$

for $\ell + 1 \leq i$. Our goal is to prove identity (2.26) on G_ℓ .

In the following, we can use the three inductive assumptions (2.24)-(2.25)-(2.26).

We prove (2.26) at the point $x = (0, \dots, 0, x_\ell, \dots, x_n) \in G_\ell$. Notice that, by Lemma 2.3, we have $X_\ell = \partial/\partial x_\ell$ on G_ℓ . For any $s \in \mathbb{R}$, we let

$$\begin{aligned} y_s &= (0, \dots, 0, s, x_{\ell+1}, \dots, x_n) \in G_\ell \\ &= e^{sX_\ell}(y_0) \\ &= e^{(s-x_\ell)X_\ell}(x), \end{aligned}$$

where $y_0 = (0, \dots, 0, 0, x_{\ell+1}, \dots, x_n) \in G_{\ell+1}$. In particular, we have $y_{x_\ell} = x$. Integrating, we obtain the formula

$$\begin{aligned} (2.27) \quad P_j(x) &= P_j(y_0) + \int_0^{x_\ell} (X_\ell P_j)(y_s) ds \\ &= P_j(e^{-x_\ell X_\ell}(x)) + \int_0^{x_\ell} (X_\ell P_j)(e^{(s-x_\ell)X_\ell}(x)) ds. \end{aligned}$$

From the structure (ii) in Lemma 2.3 for the vector field X_i with $i > \ell$, we deduce that for any $z = (0, \dots, 0, z_\ell, \dots, z_n) \in G_\ell$ and for any $t \in \mathbb{R}$ we have

$$e^{tX_i}(z) = (0, \dots, 0, z_\ell, *, \dots, *) \in G_\ell,$$

and therefore $e^{tX_i}(x) = (0, \dots, 0, x_\ell, *, \dots, *)$ and $e^{tX_i}(y_s) = (0, \dots, 0, s, *, \dots, *)$. Replacing x with $e^{tX_i}(x) \in G_\ell$ into the identity (2.27), we get:

$$\begin{aligned} (2.28) \quad P_j(e^{tX_i}(x)) &= P_j(e^{-x_\ell X_\ell} \circ e^{tX_i}(x)) + \int_0^{x_\ell} (X_\ell P_j)(e^{(s-x_\ell)X_\ell} \circ e^{tX_i}(x)) ds \\ &= P_j(e^{-x_\ell X_\ell} \circ e^{tX_i} \circ e^{x_\ell X_\ell}(y_0)) + \int_0^{x_\ell} (X_\ell P_j)(e^{(s-x_\ell)X_\ell} \circ e^{tX_i} \circ e^{-(s-x_\ell)X_\ell}(y_s)) ds \\ &= P_j(y_0 \cdot \exp(x_\ell X_\ell) \cdot \exp(tX_i) \cdot \exp(-x_\ell X_\ell)) \\ &\quad + \int_0^{x_\ell} (X_\ell P_j)(y_s \cdot \exp(-(s-x_\ell)X_\ell) \cdot \exp(tX_i) \cdot \exp((s-x_\ell)X_\ell)) ds. \end{aligned}$$

For any $z \in G = \mathbb{R}^n$, we define the left translation $L_z : G \rightarrow G$ and the conjugate mapping $\Gamma_z : G \rightarrow G$

$$L_z y = z \cdot y \quad \text{and} \quad \Gamma_z(y) = z \cdot y \cdot z^{-1}, \quad y \in G.$$

With this notation, identity (2.28) reads

$$P_j(e^{tX_i}(x)) = P_j(L_{y_0} \Gamma_{\exp(x_\ell X_\ell)}(\exp(tX_i))) + \int_0^{x_\ell} (X_\ell P_j)(L_{y_s} \Gamma_{\exp(-(s-x_\ell)X_\ell)}(\exp(tX_i))) ds,$$

and thus

$$X_i P_j(x) = \lim_{t \rightarrow 0} \frac{P_j(e^{tX_i}(x)) - P_j(x)}{t} = A_{ij}(x) + B_{ij}(x),$$

where

$$\begin{aligned} A_{ij}(x) &= \lim_{t \rightarrow 0} \frac{1}{t} [P_j(L_{y_0} \Gamma_{\exp(x_\ell X_\ell)}(\exp(tX_i))) - P_j(y_0)], \quad \text{and} \\ B_{ij}(x) &= \lim_{t \rightarrow 0} \frac{1}{t} \int_0^{x_\ell} [X_\ell P_j(L_{y_s} \Gamma_{\exp(-(s-x_\ell)X_\ell)}(\exp(tX_i))) - X_\ell P_j(y_s)] ds. \end{aligned}$$

In the following we denote by Ad_z the differential of Γ_z . Recall that one has the formula $\text{Ad}_{\exp(Z)} = e^{\text{ad}(Z)}$, where $\text{ad}(Z)(Y) = [Z, Y]$, see [7, Proposition 1.91]. We have:

$$\begin{aligned} A_{ij}(x) &= [(dL_{y_0})(\text{Ad}_{\exp x_\ell X_\ell}(X_i))]P_j(y_0) \\ &= \{X_i + x_\ell[X_\ell, X_i] + \frac{1}{2}x_\ell^2[X_\ell, [X_\ell, X_i]] + \dots\}P_j(y_0) \\ &= \{X_i - x_\ell[X_i, X_\ell] + \frac{1}{2}x_\ell^2[[X_i, X_\ell], X_\ell] - \frac{1}{3!}x_\ell^3[[X_i, X_\ell], X_\ell], X_\ell] + \dots\}P_j(y_0) \\ &= \{X_i + \sum_{p=1}^s \frac{(-1)^p}{p!} x_\ell^p [X_i, X_{pe_\ell}]\}P_j(y_0) \\ &= \{X_i + \sum_{p=1}^s \sum_{h=1}^n \frac{(-1)^p}{p!} x_\ell^p c_{i,pe_\ell}^h X_h\}P_j(y_0). \end{aligned}$$

Notice that for any $p \geq 1$, we have $c_{i,pe_\ell}^h = 0$ for all $h \leq i$. For $h > i$ we can use the inductive assumption (2.24) and find

$$(2.29) \quad X_h P_j(y_0) = \sum_{k \leq n} c_{hj}^k P_k(y_0).$$

Moreover, since $y_0 \in G_{\ell+1}$, we can also use the inductive assumption (2.26) and find

$$(2.30) \quad X_i P_j(y_0) = \sum_{k \leq n} c_{ij}^k P_k(y_0).$$

By (2.29) and (2.30), we obtain the following expression for A_{ij} :

$$\begin{aligned} A_{ij}(x) &= \sum_{k \leq n} \left\{ c_{ij}^k + \sum_{p=1}^s \sum_{h=1}^n \frac{(-1)^p}{p!} x_\ell^p c_{hj}^k c_{i,pe_\ell}^h \right\} P_k^v(y_0) \\ &= \sum_{k \leq n} \left\{ c_{ij}^k + \sum_{p=1}^s \left(\frac{(-1)^p}{p!} x_\ell^p \sum_{h=1}^n c_{hj}^k c_{i,pe_\ell}^h \right) \right\} \left[\sum_{\alpha \in \mathcal{I}} \sum_{m \leq n} \frac{(-1)^{|\alpha|}}{\alpha!} c_{k\alpha}^m v_m y_0^\alpha \right]. \end{aligned}$$

A similar computation for $B_{ij}(x)$ shows that

$$\begin{aligned} B_{ij}(x) &= \int_0^{x_\ell} \lim_{t \rightarrow 0} \frac{1}{t} [X_\ell P_j(L_{y_s} \Gamma_{\exp(-(s-x_\ell)X_\ell)}(\exp(tX_i))) - X_\ell P_j(y_s)] ds \\ &= \int_0^{x_\ell} [(dL_{y_s})(\text{Ad}_{(x_\ell-s)X_\ell}(X_i))](X_\ell P_j)(y_s) ds \\ &= \int_0^{x_\ell} \left\{ X_i + \sum_{p=1}^s \sum_{u=1}^n \frac{(-1)^p}{p!} (x_\ell - s)^p c_{i,pe_\ell}^u X_u \right\} (X_\ell P_j)(y_s) ds. \end{aligned}$$

By (2.16), on G_ℓ we have the identity

$$X_\ell P_j = \sum_{h \leq n} c_{\ell j}^h P_h,$$

and since $y_s \in G_\ell$ for all s , we obtain

$$B_{ij}(x) = \int_0^{x_\ell} \sum_{h \leq n} c_{\ell j}^h \left\{ X_i + \sum_{p=1}^s \sum_{u=1}^n \frac{(-1)^p}{p!} (x_\ell - s)^p c_{i,pe_\ell}^u X_u \right\} P_h(y_s) ds.$$

Notice that $c_{\ell j}^h = 0$ for $h \leq j$ and $c_{i,pe_\ell}^u = 0$ for $p \geq 1$ and $u \leq i$. For $h > j$ and $u > i$ we can use the inductive assumptions (2.24) and (2.25). We obtain the following expression for B_{ij} :

$$\begin{aligned} B_{ij}(x) &= \int_0^{x_\ell} \sum_{h, k \leq n} c_{\ell j}^h \left\{ c_{ih}^k + \sum_{p=1}^s \sum_{u=1}^n \frac{(-1)^p}{p!} (x_\ell - s)^p c_{i,pe_\ell}^u c_{uh}^k \right\} P_k(y_s) ds \\ &= \int_0^{x_\ell} \left[\sum_{h, k \leq n} \sum_{u=1}^n \sum_{p=0}^s c_{\ell j}^h \frac{(-1)^p}{p!} (x_\ell - s)^p c_{i,pe_\ell}^u c_{uh}^k \right] \left[\sum_{\alpha \in \mathcal{I}} \sum_{m \leq n} v_m \frac{(-1)^{|\alpha|}}{\alpha!} c_{k,\alpha}^m y_s^\alpha \right] ds. \end{aligned}$$

Using the integrals

$$\begin{aligned} \int_0^{x_\ell} (x_\ell - s)^p y_s^\alpha ds &= \int_0^{x_\ell} (x_\ell - s)^p s^{\alpha_\ell} x_1^{\alpha_1} \cdots \widehat{x_\ell^{\alpha_\ell}} \cdots x_n^{\alpha_n} ds \\ &= \frac{p! \alpha_\ell!}{(\alpha_\ell + p + 1)!} x^{\alpha + (p+1)e_\ell}, \end{aligned}$$

we finally have

$$\begin{aligned} B_{ij}(x) &= \sum_{h,k,m \leq n} \sum_{u=1}^n \sum_{p=0}^s \sum_{\alpha \in \mathcal{I}_\ell} \frac{(-1)^{p+|\alpha|}}{p! \alpha!} c_{lj}^h c_{i,pe_\ell}^u c_{uh}^k c_{k\alpha}^m \frac{p! \alpha_\ell!}{(\alpha_\ell + p + 1)!} v_m x^{\alpha + (p+1)e_\ell} \\ &= - \sum_{h,k,m \leq n} \sum_{u=1}^n \sum_{p=0}^s \sum_{\alpha \in \mathcal{I}_\ell} \frac{(-1)^{|\alpha + (p+1)e_\ell|}}{(\alpha + (p+1)e_\ell)!} c_{lj}^h c_{i,pe_\ell}^u c_{uh}^k c_{k\alpha}^m v_m x^{\alpha + (p+1)e_\ell}. \end{aligned}$$

Analogously, we have:

$$A_{ij}(x) = \sum_{h,k,m \leq n} \sum_{p=0}^s \sum_{\alpha \in \mathcal{I}_{\ell+1}} \frac{(-1)^p}{p!} c_{i,pe_\ell}^h c_{hj}^k x^p \frac{(-1)^{|\alpha|}}{\alpha!} c_{k\alpha}^m v_m y_0^\alpha,$$

where the sum ranges over $\alpha \in \mathcal{I}_{\ell+1}$, because $y_0 = (0, \dots, 0, x_{\ell+1}, \dots, x_n)$. For $\alpha \in \mathcal{I}_{\ell+1}$, we have

$$\frac{(-1)^p}{p!} \frac{(-1)^{|\alpha|}}{\alpha!} = \frac{(-1)^{p+|\alpha|}}{(\alpha + pe_\ell)!} = c_{\alpha + pe_\ell},$$

and, moreover, $x_\ell^p y_0^\alpha = x^{\alpha + pe_\ell}$. Thus,

$$A_{ij}(x) = \sum_{k,h,m \leq n} \sum_{p=0}^s \sum_{\alpha \in \mathcal{I}_{\ell+1}} c_{\alpha + pe_\ell} c_{i,pe_\ell}^h c_{hj}^k c_{k\alpha}^m v_m x^{\alpha + pe_\ell}.$$

Our goal is to prove that the quantity $X_i P_j(x) = A_{ij}(x) + B_{ij}(x)$, $x \in G_\ell$, equals

$$C_{ij}(x) = \sum_{k \leq n} c_{ij}^k P_k(x) = \sum_{k,m \leq n} \sum_{\beta \in \mathcal{I}_\ell} c_\beta c_{ij}^k c_{k\beta}^m v_m x^\beta.$$

The coefficient of $c_\beta v_m x^\beta$ in $C_{ij}(x)$ is

$$\chi_C^{m,\beta} = \sum_{k \leq n} c_{ij}^k c_{k\beta}^m.$$

The equation $\alpha + p e_\ell = \beta$ with $\alpha_1 = \dots = \alpha_\ell = 0$ implies $p = \beta_\ell$ and $\alpha = \beta - p e_\ell = \beta - \beta_\ell e_\ell$. Thus, the coefficient of $c_\beta v_m x^\beta$ in $A_{ij}(x)$ is

$$\chi_A^{m,\beta} = \sum_{k,h \leq n} c_{i,\beta_\ell e_\ell}^h c_{hj}^k c_{k,\beta - \beta_\ell e_\ell}^m.$$

We compute the coefficient $\chi_B^{m,\beta}$ of $c_\beta v_m x^\beta$ in $B_{ij}(x)$. The equation $\alpha + (p+1)e_\ell = \beta$ implies the following: when $\beta_\ell = 0$ the equation has no solution and thus $\chi_B^{m,\beta} = 0$; when $\beta_\ell \geq 1$ we have $0 \leq p \leq \beta_\ell - 1$ and $\alpha = \beta - (p+1)e_\ell$, and therefore

$$\chi_B^{m,\beta} = - \sum_{h,k \leq n} \sum_{u=1}^n \sum_{p=0}^{\beta_\ell - 1} c_{\ell j}^h c_{i,p e_\ell}^u c_{uh}^k c_{k,\beta - (p+1)e_\ell}^m, \quad \beta_\ell \geq 1.$$

When $\beta_\ell = 0$, the identity $\chi_A^{m,\beta} + \chi_B^{m,\beta} = \chi_C^{m,\beta}$ is easily verified.

When $\beta_\ell \geq 1$, the identity $\chi_A^{m,\beta} + \chi_B^{m,\beta} = \chi_C^{m,\beta}$ is proved as soon as we verify the following

(2.31)

$$\sum_{k \leq n} c_{ij}^k c_{k\beta}^m = \sum_{k,h \leq n} c_{i,\beta_\ell e_\ell}^h c_{hj}^k c_{k,\beta - \beta_\ell e_\ell}^m - \sum_{h,k \leq n} \sum_{u=1}^n \sum_{p=0}^{\beta_\ell - 1} c_{\ell j}^h c_{i,p e_\ell}^u c_{uh}^k c_{k,\beta - (p+1)e_\ell}^m.$$

Contracting with X_m the left hand side of (2.31) we find

$$(2.32) \quad \sum_{m,k \leq n} c_{ij}^k c_{k\beta}^m X_m = \sum_{k \leq n} c_{ij}^k [X_k, X_\beta] = [[X_i, X_j], X_\beta] \\ = [[[X_i, X_j], X_{\beta_\ell e_\ell}], X_{\beta - \beta_\ell e_\ell}].$$

Contracting with X_m the first term in the right hand side of (2.31) we find

$$(2.33) \quad \sum_{k,h,m \leq n} c_{i,\beta_\ell e_\ell}^h c_{hj}^k c_{k,\beta - \beta_\ell e_\ell}^m X_m = \sum_{k,h \leq n} c_{i,\beta_\ell e_\ell}^h c_{hj}^k [X_k, X_{\beta - \beta_\ell e_\ell}] \\ = \sum_{h=1}^n c_{i,\beta_\ell e_\ell}^h [[X_h, X_j], X_{\beta - \beta_\ell e_\ell}] \\ = [[[X_i, X_{\beta_\ell e_\ell}], X_j], X_{\beta - \beta_\ell e_\ell}].$$

Contracting the second term in the right hand side of (2.31) we find

$$\begin{aligned}
 & - \sum_{h,k,m \leq n} \sum_{u=1}^n \sum_{p=0}^{\beta_\ell-1} c_{\ell j}^h c_{i,pe_\ell}^u c_{uh}^k c_{k,\beta-(p+1)e_\ell}^m X_m \\
 (2.34) \quad & = \sum_{p=0}^{\beta_\ell-1} [[[[X_i, X_{pe_\ell}], [X_j, X_\ell]], X_{(\beta_\ell-p-1)e_\ell}], X_{\beta-\beta_\ell e_\ell}].
 \end{aligned}$$

We used $\beta_1 = \dots = \beta_\ell = 0$.

In (2.32), (2.33), and (2.34), commutators have the same tail $X_{\beta-\beta_\ell e_\ell}$. Thus the claim (2.31) follows provided that we prove the following identity: (2.35)

$$[[X_i, X_j], X_{\beta_\ell e_\ell}] = [[X_i, X_{\beta_\ell e_\ell}], X_j] + \sum_{p=0}^{\beta_\ell-1} [[[[X_i, X_{pe_\ell}], [X_j, X_\ell]], X_{(\beta_\ell-p-1)e_\ell}].$$

We prove (2.35) by induction on $\beta_\ell \geq 1$. The base of induction for $\beta_\ell = 1$ is

$$[[X_i, X_j], X_\ell] = [[X_i, X_\ell], X_j] + [X_i, [X_j, X_\ell]],$$

and this holds by the Jacobi identity.

By induction, we assume that (2.35) holds for β_ℓ and we prove it for $\beta_\ell + 1$. We have:

$$\begin{aligned}
 [[X_i, X_j], X_{(\beta_\ell+1)e_\ell}] &= [[[[X_i, X_j], X_{\beta_\ell e_\ell}], X_\ell] \\
 &= [[[[X_i, X_{\beta_\ell e_\ell}], X_j], X_\ell] + \sum_{p=0}^{\beta_\ell-1} [[[[X_i, X_{pe_\ell}], [X_j, X_\ell]], X_{(\beta_\ell-p-1)e_\ell}], X_\ell] \\
 &= [[X_i, X_{(\beta_\ell+1)e_\ell}], X_j] + [[X_i, X_{\beta_\ell e_\ell}], [X_j, X_\ell]] \\
 &\quad + \sum_{p=0}^{\beta_\ell-1} [[[[X_i, X_{pe_\ell}], [X_j, X_\ell]], X_{(\beta_\ell-p)e_\ell}] \\
 &= [[X_i, X_{(\beta_\ell+1)e_\ell}], X_j] + \sum_{p=0}^{\beta_\ell} [[[[X_i, X_{pe_\ell}], [X_j, X_\ell]], X_{(\beta_\ell-p)e_\ell}].
 \end{aligned}$$

This finishes the proof of Theorem 2.2. □

In the next proposition, we list some elementary properties of extremal polynomials that are used in Section 3.

Proposition 2.6. *Extremal polynomials have the following properties.*

- (i) For all $i = 1, \dots, n$ and $v \in \mathbb{R}^n$, we have $P_i^v(0) = v_i$.
- (ii) For all $i \in \mathbb{Z}$ with $i \leq 0$ and $v \in \mathbb{R}^n$, we have $P_i^v(0) = 0$.
- (iii) If $v \in \mathbb{R}^n$ is such that $P_i^v = 0$ for all $i = 1, \dots, n$ with $d(i) = 1$, then $v = 0$.

Proof. (i) For any $i = 1, \dots, n$, we have

$$(2.36) \quad P_i^v(0) = \phi_{i0}(v) = \sum_{k \leq n} c_{i0}^k v_k = v_i,$$

because $c_{i0}^k = \delta_{ik}$, the Kronecker symbol.

(ii) This follows as in (i) from the agreement that $v_k = 0$ for $k \leq 0$.

(iii) Assume that $P_k^v = 0$ when $d(k) = 1$. We claim that $P_k^v = 0$ for all $k = 1, \dots, n$. Because of (i), this will imply $v = 0$. The proof is by induction on $d(k)$. Assume that $P_k^v = 0$ for all $k = 1, \dots, n$ such that $d(k) \leq \ell$, where $\ell < s$. Take $k = 1, \dots, n$ such that $d(k) = \ell + 1$. By the stratification, there are constants $d_{ij}^k \in \mathbb{R}$, $d(i) = 1$ and $d(j) = \ell$, such that

$$X_k = \sum_{\substack{d(i)=1 \\ d(j)=\ell}} d_{ij}^k [X_i, X_j] = \sum_{\substack{d(i)=1 \\ d(j)=\ell}} \sum_{h \leq n} d_{ij}^k c_{ij}^h X_h,$$

and thus we have the identity

$$(2.37) \quad \sum_{\substack{d(i)=1 \\ d(j)=\ell}} d_{ij}^k c_{ij}^h = \delta_{hk}.$$

From $P_j^v = 0$ for $d(j) = \ell$, it follows that $X_i P_j^v = 0$. Now, from Theorem 2.2 and (2.37) we obtain

$$0 = \sum_{\substack{d(i)=1 \\ d(j)=\ell}} d_{ij}^k X_i P_j^v = \sum_{\substack{d(i)=1 \\ d(j)=\ell}} \sum_{h \leq n} d_{ij}^k c_{ij}^h P_h^v = P_k^v.$$

This finishes the proof. □

3. Characterization of abnormal extremals

In this section, we prove the theorems about the algebraic characterization of abnormal extremals. The Lie group G is identified with \mathbb{R}^n via exponential coordinates as in (2.7), where X_1, \dots, X_n is a basis of $\mathfrak{g} =$

$\text{Lie}(G)$ adapted to the stratification. The vector fields X_1, \dots, X_r are a basis for \mathfrak{g}_1 (and thus generators for \mathfrak{g}). The distribution of r -planes $\mathcal{D}(x) = \text{span}\{X_1(x), \dots, X_r(x)\}$, $x \in G$, is called horizontal distribution of G . A Lipschitz curve $\gamma : [0, 1] \rightarrow G = \mathbb{R}^n$ is \mathcal{D} -horizontal, or simply horizontal, if there exists a vector of functions $h = (h_1, \dots, h_r) \in L^\infty([0, 1]; \mathbb{R}^r)$ such that

$$\dot{\gamma} = \sum_{j=1}^r h_j X_j(\gamma), \quad \text{a.e. on } [0, 1].$$

The functions h are called controls of γ and, when $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ is given by the coordinates $\gamma = (\gamma_1, \dots, \gamma_n)$, we have $h_j = \dot{\gamma}_j$, $j = 1, \dots, r$. This follows from the structure of the vector fields X_1, \dots, X_r described in Lemma 2.3.

Let g_x be the quadratic form on $\mathcal{D}(x)$ making X_1, \dots, X_r orthonormal. The horizontal length of a horizontal curve γ is

$$L(\gamma) = \left(\int_0^1 g_{\gamma(t)}(\dot{\gamma}(t)) dt \right)^{1/2} = \left(\int_0^1 |h(t)|^2 dt \right)^{1/2}.$$

For any couple of points $x, y \in G$, we can define the function

$$(3.38) \quad d(x, y) = \inf \left\{ L(\gamma) : \gamma \text{ is horizontal, } \gamma(0) = x \text{ and } \gamma(1) = y \right\}.$$

The above set is always nonempty and d is a distance on G . A Lipschitz curve γ providing the minimum in (3.38) is called a *length minimizer*.

We denote by $\vartheta_1, \dots, \vartheta_n$ the basis of \mathfrak{g}^* dual to the basis X_1, \dots, X_n of \mathfrak{g} . A Lipschitz curve of 1-forms $\lambda \in \text{Lip}([0, 1]; \mathfrak{g}^*)$ is a curve

$$\lambda = \lambda_1 \vartheta_1 + \dots + \lambda_n \vartheta_n,$$

where $\lambda_i : [0, 1] \rightarrow \mathbb{R}$, $i = 1, \dots, n$, are Lipschitz functions. We call $\lambda_1, \dots, \lambda_n$ the coordinates of λ .

Theorem 3.1. *Let γ be a length minimizer in G . Then there exist a number $\lambda_0 \in \{0, 1\}$ and a Lipschitz curve $\lambda \in \text{Lip}([0, 1]; \mathfrak{g}^*)$ with coordinates $\lambda_1, \dots, \lambda_n$ such that:*

- (i) $\lambda_0 + |\lambda| \neq 0$ on $[0, 1]$;
- (ii) $\lambda_0 \dot{\gamma}_j + \lambda_j = 0$ on $[0, 1]$ for all $j = 1, \dots, r$;

(iii) for all $i = 1, \dots, n$, we have

$$(3.39) \quad \dot{\lambda}_i = - \sum_{k=1}^n \sum_{j=1}^r c_{ij}^k \dot{\gamma}_j \lambda_k \quad \text{a.e. on } [0, 1],$$

where c_{ij}^k are the structure constants (2.12).

Theorem 3.1 is a version of Pontryagin Maximum Principle adapted to the present setting. Equations (3.39) are called *adjoint equations*. We refer to [3, Chapter 12] for a proof of Theorem 3.1 in a more general framework. The version (3.39) of the adjoint equations is derived in [8], Theorem 2.6. The curve λ is called a *dual curve* of γ .

Definition 3.2. We say that a horizontal curve $\gamma : [0, 1] \rightarrow G$ is an extremal if there exist $\lambda_0 \in \{0, 1\}$ and a curve of 1-forms $\lambda \in \text{Lip}([0, 1]; \mathfrak{g}^*)$ such that i), ii), and iii) in Theorem 3.1 hold.

We say that γ is a normal extremal if there exists such a pair (λ_0, λ) with $\lambda_0 = 1$.

We say that γ is an abnormal extremal if there exists such a pair with $\lambda_0 = 0$.

We say that γ is a strictly abnormal extremal if γ is an abnormal extremal but not a normal one.

Remark 3.3. If γ is an abnormal extremal with dual curve λ , then we have

$$(3.40) \quad \lambda_1 = \dots = \lambda_r = 0 \quad \text{on } [0, 1].$$

This follows from condition ii) of Theorem 3.1 with $\lambda_0 = 0$.

Theorem 3.4. Let G be a stratified nilpotent Lie group, let $\gamma : [0, 1] \rightarrow G$ be a horizontal curve such that $\gamma(0) = 0$, and let $\lambda_1, \dots, \lambda_n : [0, 1] \rightarrow \mathbb{R}$ be Lipschitz functions. The following statements are equivalent:

(A) The functions $\lambda_1, \dots, \lambda_n$ solve a.e. the system of equations (3.39).

(B) There exists $v \in \mathbb{R}^n$ such that, for all $i = 1, \dots, n$, we have

$$(3.41) \quad \lambda_i(t) = P_i^v(\gamma(t)), \quad t \in [0, 1],$$

and in fact $v = (\lambda_1(0), \dots, \lambda_n(0))$.

Proof. (B) \Rightarrow (A) Let $\lambda_1, \dots, \lambda_n$ be functions on $[0, 1]$ defined as in (3.41) for some $v \in \mathbb{R}^n$. Then, by Proposition 2.6 we have $v = (\lambda_1(0), \dots, \lambda_n(0))$. Moreover, by Theorem 2.2 we have, a.e. on $[0, 1]$ and for any $i = 1, \dots, n$,

$$(3.42) \quad \dot{\lambda}_i = \frac{d}{dt} P_i^v(\gamma) = \sum_{j=1}^r \dot{\gamma}_j X_j P_i^v(\gamma) = \sum_{j=1}^r \sum_{k=1}^n c_{ji}^k \dot{\gamma}_j P_k^v(\gamma) = - \sum_{j=1}^r \sum_{k=1}^n c_{ij}^k \dot{\gamma}_j \lambda_k.$$

(A) \Rightarrow (B) Let $v = (\lambda_1(0), \dots, \lambda_n(0)) \in \mathbb{R}^n$. Then the functions λ_i and the functions $\mu_i = P_i^v(\gamma)$, $i = 1, \dots, n$, solve the system of differential equation (3.39) with the same initial condition. By the uniqueness of the solution we have $\lambda_i = \mu_i$. \square

Next, we define the notion of corank for an abnormal extremal. For any $h \in L^2([0, 1]; \mathbb{R}^r)$, let γ^h be the solution of the problem

$$\dot{\gamma}^h = \sum_{j=1}^r h_j X_j(\gamma^h), \quad \gamma^h(0) = x_0.$$

The mapping $\mathcal{E} : L^2([0, 1]; \mathbb{R}^r) \rightarrow G = \mathbb{R}^n$, $\mathcal{E}(h) = \gamma^h(1)$, is called the *end-point mapping* with initial point $x_0 \in G$. A horizontal curve γ starting from x_0 with controls h is an abnormal extremal if and only if there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that

$$(3.43) \quad \langle d\mathcal{E}(h)w, v \rangle = 0$$

for all $w \in L^2([0, 1]; \mathbb{R}^r)$. Here, $d\mathcal{E}(h)$ is the differential of \mathcal{E} at the point h .

Definition 3.5. *The corank of an abnormal extremal $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ with controls h is the integer $\text{corank}(\gamma) = n - \dim(\text{Im } d\mathcal{E}(h)) \geq 1$.*

The corank of an abnormal extremal $\gamma : [0, 1] \rightarrow G$ is the integer $m \geq 1$ such that there exist precisely m linearly independent dual curves $\lambda^1, \dots, \lambda^m \in \text{Lip}([0, 1]; \mathfrak{g}^*)$ each solving the system of adjoint equations (3.39) with initial conditions $\lambda^1(0), \dots, \lambda^m(0) \in \mathbb{R}^n$ that are orthogonal to the image of the differential of the end-point mapping.

Definition 3.6. *Let $G = \mathbb{R}^n$ be a stratified nilpotent Lie group of rank r . For any $v \in \mathbb{R}^n$, $v \neq 0$, we call the set*

$$Z_v = \{x \in \mathbb{R}^n : P_j^v(x) = 0 \text{ for all } j \in \mathbb{Z} \text{ with } j \leq r\}$$

an abnormal variety of G of corank 1.

For linearly independent vectors $v_1, \dots, v_m \in \mathbb{R}^n$, $m \geq 2$, we call the set $Z_{v_1} \cap \dots \cap Z_{v_m}$ an abnormal variety of G of corank m .

Remark 3.7. The abnormal variety Z_v is the intersection of the zero sets of all the polynomials P_j^v with $j \leq r$. In this intersection, also indexes $j \leq 0$ are involved. With this respect, Definition 3.6 differs from the analogous definition in [8].

Theorem 3.8. Let $G = \mathbb{R}^n$ be a stratified nilpotent Lie group and let $\gamma : [0, 1] \rightarrow G$ be a horizontal curve with $\gamma(0) = 0$. The following statements are equivalent:

- (A) The curve γ is an abnormal extremal with $\text{corank}(\gamma) \geq m \geq 1$.
- (B) There exist $m \geq 1$ linearly independent vectors $v_1, \dots, v_m \in \mathbb{R}^n$ such that $\gamma(t) \in Z_{v_1} \cap \dots \cap Z_{v_m}$ for all $t \in [0, 1]$.

Proof. (A) \Rightarrow (B) Let γ be an abnormal extremal and let $\lambda = \lambda_1 \vartheta_1 + \dots + \lambda_n \vartheta_n$ be a Lipschitz curve of 1-forms with coordinates $\lambda_1, \dots, \lambda_n$ solving the system of differential equations (3.39) and such that $\lambda \neq 0$ pointwise. By Theorem 3.4, there is a $v \in \mathbb{R}^n$, $v \neq 0$, such that $\lambda_i = P_i^v(\gamma)$ for all $i = 1, \dots, n$. In particular, there holds $(\lambda_1(0), \dots, \lambda_n(0)) = v$. For abnormal extremals, we have $\lambda_1 = \dots = \lambda_r = 0$ on $[0, 1]$, i.e., $\gamma(t) \in \{x \in \mathbb{R}^n : P_j^v(x) = 0 \text{ for all } j = 1, \dots, r\}$ for all $t \in [0, 1]$.

We claim that $P_j^v(\gamma) = 0$ also for all $j \in \mathbb{Z}$ with $j \leq 0$. The proof is by induction on $d(j)$. The base of induction for $d(j) = 1$ is proved above. By induction, assume that $P_k^v(\gamma) = 0$ for all $k \in \mathbb{Z}$ such that $d(i) < d(k) \leq 1$. As in (3.42), by Theorem 2.2 we have

$$\frac{d}{dt} P_i^v(\gamma) = - \sum_{j=1}^r \sum_{k \leq n} c_{ij}^k \dot{\gamma}_j P_k^v(\gamma) = 0,$$

because $c_{ij}^k = 0$ for $d(k) \leq d(i)$, for any $j = 1, \dots, r$. It follows that $P_i^v(\gamma) = P_i^v(\gamma(0))$ is constant. By Proposition 2.6 and $\gamma(0) = 0$, it is the zero constant. This shows that $\gamma(t) \in Z_v$ for all $t \in [0, 1]$.

From the above argument, we conclude that if $\text{corank}(\gamma) \geq m$ then there exist at least m linearly independent vectors $v_1, \dots, v_m \in \mathbb{R}^n$ such that $\gamma(t) \in Z_{v_i}$ for all $t \in [0, 1]$ and $i = 1, \dots, m$.

(B) \Rightarrow (A) Let $v_1, \dots, v_m \in \mathbb{R}^n$ be linearly independent vectors such that (B) holds. For any $v = v_1, \dots, v_m$, the curve $\lambda = \lambda_1 \vartheta_1 + \dots + \lambda_n \vartheta_n$ with

coordinates $\lambda_i = P_i^v(\gamma)$ satisfies the system (3.39), by Theorem 3.4. Moreover, we have $\lambda_1 = \dots = \lambda_r = 0$ because γ is an abnormal extremal. From $v \neq 0$, it follows that $(\lambda_1(0), \dots, \lambda_n(0)) = v \neq 0$, and from the uniqueness of the solution to (3.39) with initial condition, it follows that $\lambda \neq 0$ pointwise on $[0, 1]$. As v_1, \dots, v_m are linearly independent, then the corresponding curves λ are also linearly independent. We conclude $\text{corank}(\gamma) \geq m$. \square

Remark 3.9. In the rank 2 case, an abnormal extremal γ satisfies $\lambda_1 = P_1^v(\gamma) = 0$ and $\lambda_2 = P_2^v(\gamma) = 0$. When γ is not a constant curve, these equations imply that also $\lambda_3 = 0$. Indeed, the equations (3.39) for λ_1 and λ_2 are

$$\dot{\lambda}_1 = \dot{\gamma}_2 \lambda_3 \quad \text{and} \quad \dot{\lambda}_2 = -\dot{\gamma}_1 \lambda_3.$$

Thus, if $\dot{\gamma}_1^2 + \dot{\gamma}_2^2 \neq 0$, it must be $\lambda_3 = 0$.

4. The abnormal set in the free group of rank 2 and step 4

In this section, we present a method to compute a set containing all abnormal extremals passing through one point. The method works when each abnormal extremal is in the zero sets of a sufficiently large number of extremal polynomials. This is the case when the prolongation is sufficiently large. In order to make the presentation clear, we focus on the free group of rank 2 and step 4. The key point is to show that a certain polynomial is nontrivial, see Theorem 4.1 below.

Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \dots \oplus \mathfrak{g}_4$ be the free nilpotent Lie algebra with rank $r = 2$ and step $s = 4$. This algebra is 8-dimensional. Let X_1, \dots, X_8 be a Hall basis of \mathfrak{g} (see [6] and [8]). This basis satisfies the following structure relations

$$(4.44) \quad \begin{aligned} X_3 &= [X_2, X_1], & X_4 &= [X_3, X_1], & X_5 &= [X_3, X_2], \\ X_6 &= [X_4, X_1], & X_7 &= [X_4, X_2], & X_8 &= [X_5, X_2]. \end{aligned}$$

In the Grayson-Grossman model [5], X_1, \dots, X_8 can be identified with vector fields in \mathbb{R}^8 where

$$(4.45) \quad \begin{aligned} X_1 &= \frac{\partial}{\partial x_1}, \\ X_2 &= \frac{\partial}{\partial x_2} - x_1 \frac{\partial}{\partial x_3} + \frac{x_1^2}{2} \frac{\partial}{\partial x_4} + x_1 x_2 \frac{\partial}{\partial x_5} - \frac{x_1^3}{6} \frac{\partial}{\partial x_6} - \frac{x_1^2 x_2}{2} \frac{\partial}{\partial x_7} - \frac{x_1 x_2^2}{2} \frac{\partial}{\partial x_8}. \end{aligned}$$

The Tanaka prolongation $\text{Prol}(\mathfrak{g})$ is finite dimensional and by Theorem 1 in [17], we have

$$\text{Prol}(\mathfrak{g}) = \mathfrak{g}_0 \oplus \mathfrak{g},$$

where \mathfrak{g}_0 is the vector space of strata preserving derivations of \mathfrak{g} . In fact, \mathfrak{g}_0 is isomorphic to $\text{End}(\mathfrak{g}_1)$ and thus \mathfrak{g}_0 is 4-dimensional, because $\dim(\mathfrak{g}_1) = 2$. We extend the basis for \mathfrak{g} to a basis $X_{-3}, X_{-2}, \dots, X_8$ of $\text{Prol}(\mathfrak{g})$ and we denote by $c_{j\alpha}^k$, with $j = -3, -2, \dots, 8$, $k = 1, \dots, 8$, and $\alpha \in \mathcal{I} = \mathbb{N}^8$, the generalized structure constants of $\text{Prol}(\mathfrak{g})$.

Let G be the free nilpotent Lie group of rank $r = 2$ and step $s = 4$. Via exponential coordinates of the second type, G is diffeomorphic to \mathbb{R}^8 . The extremal polynomials in \mathbb{R}^8 associated with the Tanaka prolongation of $\mathfrak{g} = \text{Lie}(G)$ are, for any fixed $v \in \mathbb{R}^8$,

$$(4.46) \quad P_j^v(x) = \sum_{\alpha \in \mathcal{I}} \frac{(-1)^{|\alpha|}}{\alpha!} \sum_{k=1}^8 c_{j\alpha}^k v_k x^\alpha = \sum_{k=1}^8 v_k Q_{jk}(x), \quad x \in \mathbb{R}^8,$$

where Q_{jk} are the polynomials in \mathbb{R}^8

$$(4.47) \quad Q_{jk}(x) = \sum_{\alpha \in \mathcal{I}} \frac{(-1)^{|\alpha|}}{\alpha!} c_{j\alpha}^k x^\alpha, \quad x \in \mathbb{R}^8.$$

Above, we have $j = -3, -2, \dots, 8$ and $k = 1, \dots, 8$. Notice that $c_{j\alpha}^k = 0$ if $d(j) + d(\alpha) \neq d(k)$. It follows that Q_{jk} is a polynomial with homogeneous degree $d(k) - d(j)$.

Let $v \in \mathbb{R}^8$ be a vector such that $v_1 = v_2 = v_3 = 0$. The first three polynomials P_1^v , P_2^v , and P_3^v are the following

$$(4.48) \quad \begin{aligned} P_1^v(x) &= v_4 x_3 - v_5 \frac{x_2^2}{2} + v_6 x_4 + v_7 x_5 + v_8 \frac{x_2^3}{6}, \\ P_2^v(x) &= v_4 \frac{x_1^2}{2} + (x_3 + x_1 x_2) v_5 - \frac{x_1^3}{6} v_6 + \left(x_4 - \frac{x_1^2 x_2}{2}\right) v_7 + \left(x_5 - \frac{x_1 x_2^2}{2}\right) v_8, \\ P_3^v(x) &= -v_4 x_1 - v_5 x_2 + v_6 \frac{x_1^2}{2} + v_7 x_1 x_2 + v_8 \frac{x_2^2}{2}. \end{aligned}$$

These polynomials can be computed using the structure relations (4.44) and formulas (4.46).

In order to compute the four polynomials P_{-3}^v , P_{-2}^v , P_{-1}^v , and P_0^v associated with the stratum \mathfrak{g}_0 of $\text{Prol}(\mathfrak{g})$, we have to choose a basis $X_{-3}, X_{-2}, X_{-1}, X_0$

of \mathfrak{g}_0 . We can identify \mathfrak{g}_0 with $\text{End}(\mathfrak{g}_1)$, and \mathfrak{g}_1 with \mathbb{R}^2 via the basis X_1, X_2 . Hence, we can make the following choice:

$$X_{-3} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad X_{-2} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad X_{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad X_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The commutator $[X_j, X_i]$, with $j = -3, \dots, 0$ and $i = 1, 2$, is identified with $X_j(X_i)$ where X_j is a linear operator on \mathfrak{g}_1 . The polynomials associated with this basis are the following

$$(4.49) \quad \begin{aligned} P_{-3}^v(x) &= v_4(x_3x_2 + x_5) - v_5 \frac{x_2^3}{6} + v_6(x_4x_2 + x_7) + v_7(x_5x_2 + 2x_8) + v_8 \frac{x_2^4}{24}, \\ P_{-2}^v(x) &= v_4x_4 + v_5(x_2x_3 + 2x_5) + v_6x_6 + v_7(x_2x_4 + 2x_7) + v_8(x_2x_5 + 3x_8), \\ P_{-1}^v(x) &= v_4(x_1x_3 + 2x_4) + v_5\left(x_5 - \frac{x_1x_2^2}{2}\right) + v_6(x_1x_4 + 3x_6) \\ &\quad + v_7(x_1x_5 + 2x_7) + v_8\left(\frac{x_1x_2^3}{6} + x_8\right), \\ P_0^v(x) &= v_4 \frac{x_1^3}{6} + v_5\left(\frac{x_1^2x_2}{2} + x_1x_3 + x_4\right) - v_6 \frac{x_1^4}{24} \\ &\quad + v_7\left(-\frac{x_1^3x_2}{6} + x_1x_4 + 2x_6\right) + v_8\left(-\frac{x_1^2x_2^2}{4} + x_1x_5 + x_7\right). \end{aligned}$$

These formulas can be checked in the following way. According to (2.15), the polynomials (4.49) have to satisfy the structure identities

$$(4.50) \quad \begin{aligned} X_1P_{-3}^v &= 0, & X_1P_{-2}^v &= 0, & X_1P_{-1}^v &= P_1^v, & X_1P_0^v &= P_2^v, \\ X_2P_{-3}^v &= P_1^v, & X_2P_{-2}^v &= P_2^v, & X_2P_{-1}^v &= 0, & X_2P_0^v &= 0. \end{aligned}$$

Notice that we have the structure constants $c_{-3,1}^k = c_{-2,1}^k = c_{-1,2}^k = c_{0,2}^k = 0$, $c_{-3,2}^k = c_{-1,1}^k = \delta_{1k}$, $c_{-2,2}^k = c_{0,1}^k = \delta_{2k}$, $k = 1, 2$. Along with the condition $P_j^v(0) = 0$, $j = -3, \dots, 0$, the identities (4.50) uniquely determine the polynomials. Now, by a direct computation based on (4.45), it can be checked that the polynomials (4.49) do satisfy (4.50).

The polynomials (4.49) can also be computed using the following algorithm. Let $\gamma : [0, 1] \rightarrow \mathbb{R}^8$ be a horizontal curve such that $\gamma(0) = 0$ and, for $i, j = 1, 2$, let

$$B_{ij}^v(t) = \int_0^t P_i^v(\gamma) \dot{\gamma}_j ds, \quad t \in [0, 1].$$

Using the formulas (4.48) for P_1^v and P_2^v , the integrals in the right hand side can be computed. Using the explicit formulas for the coefficients of the vector

field X_2 in (4.45), the resulting functions can be shown to be polynomials of the coordinates of $\gamma(t)$. Namely, along γ we have $B_{12}^v = P_{-3}^v$, $B_{22}^v = P_{-2}^v$, $B_{11}^v = P_{-1}^v$, and $B_{21}^v = P_0^v$.

Let $\gamma : [0, 1] \rightarrow G = \mathbb{R}^8$ be an abnormal extremal such that $\gamma(0) = 0$. By Theorem 3.8 (see also Remark 3.9) there exists $v \in \mathbb{R}^8$, with $v \neq 0$, such that, for all $j = -3, -2, \dots, 3$, we have

$$(4.51) \quad P_j^v(\gamma(t)) = 0, \quad t \in [0, 1].$$

In particular, $v = (v_1, \dots, v_8)$ satisfies $v_1 = v_2 = v_3 = 0$ (because $\gamma(0) = 0$).

For each $j = -3, \dots, 3$, we define the following 5-dimensional vector of polynomials

$$Q_j(x) = (Q_{j4}(x), \dots, Q_{j8}(x)),$$

where Q_{jk} are defined in (4.47). By (4.46), identity (4.51) reads

$$\sum_{k=4}^8 v_k Q_{jk}(\gamma(t)) = 0, \quad t \in [0, 1].$$

In other words, along the curve γ the seven 5-dimensional vectors $Q_{-3}(x), \dots, Q_3(x)$ are orthogonal to the nonzero vector $(v_4, \dots, v_8) \in \mathbb{R}^5$. It follows that the 7×5 matrix

$$Q(x) = \begin{pmatrix} Q_{-3,4}(x) & \dots & Q_{-3,8}(x) \\ \vdots & & \vdots \\ Q_{3,4}(x) & \dots & Q_{3,8}(x) \end{pmatrix}$$

has rank at most 4 along the curve γ .

Let $M_1(x), \dots, M_{21}(x)$ be the 21 different 5×5 minors of $Q(x)$ and define the determinant functions

$$(4.52) \quad f_k(x) = \det(M_k(x)), \quad k = 1, \dots, 21, \quad x \in \mathbb{R}^8.$$

It can be proved that the functions f_1, \dots, f_{21} are homogeneous polynomials. By the discussion above, we have

$$f_k(\gamma(t)) = 0 \quad \text{for all } t \in [0, 1] \text{ and for all } k = 1, \dots, 21.$$

The polynomials f_1, \dots, f_{21} do not depend on v .

Theorem 4.1. *Let $G = \mathbb{R}^8$ be the free nilpotent Lie group of rank 2 and step 4. There exists a nonzero homogeneous polynomial $Q : \mathbb{R}^8 \rightarrow \mathbb{R}$ such that*

any abnormal curve $\gamma : [0, 1] \rightarrow \mathbb{R}^8 = G$ with $\gamma(0) = 0$ satisfies $Q(\gamma(t)) = 0$ for all $t \in [0, 1]$.

Proof. It is enough to show that at least one of the polynomials f_1, \dots, f_{21} in (4.52) is nonzero. Let us consider the polynomial with least degree

$$Q(x) = \det M(x) = \det \begin{pmatrix} Q_{-1,4}(x) & \dots & Q_{-1,8}(x) \\ Q_{0,4}(x) & \dots & Q_{0,8}(x) \\ \vdots & & \vdots \\ Q_{3,4}(x) & \dots & Q_{3,8}(x) \end{pmatrix}.$$

The homogeneous degree of Q is 14. We highlight the anti-diagonal of the matrix M :

$$M(x) = \begin{pmatrix} * & * & * & * & x_1x_2^3/6 + x_8 \\ * & * & * & -x_1^3x_2/6 + x_1x_4 + 2x_6 & * \\ * & * & x_4 & * & * \\ * & x_1x_2 + x_3 & * & * & * \\ -x_1 & * & * & * & * \end{pmatrix}.$$

Upon inspection of the polynomials P_{-1}^v, \dots, P_3^v in (4.48) and (4.49), we observe the following facts. The variable x_8 appears in the entry M_{ij} only when $i = j = 5$. When $i, j \leq 4$, the variable x_6 appears only in the entry M_{ij} with $i = j = 4$. When $i, j \leq 3$, the variable x_4 appears only in the entry M_{ij} with $i = j = 3$. When $i, j \leq 2$, the variable x_3 appears only in the entry M_{ij} with $i, j = 2$. Finally, we have $M_{11} = -x_1$. It follows that

$$Q(x) = -2x_1x_3x_4x_6x_8 + R(x),$$

where $R(x)$ is a polynomial that does not contain the monomial $x_1x_3x_4x_6x_8$. This proves that $Q \neq 0$. \square

Remark 4.2. Any abnormal curve passing through 0 is in the intersection of the zero sets of the 21 polynomials (4.52) in \mathbb{R}^8 . Even though all these polynomials are explicitly computable, the precise structure of this intersection is not clear.

For any $v \in \mathbb{R}^8$ with $v \neq 0$ and $v_1 = v_2 = v_3$, the equation $P_3^v(x) = 0$ determines two (in some exceptional cases four) different abnormal extremals γ^v parameterized by arc-length and such that $\gamma^v(0) = 0$. The mapping $v \mapsto \gamma^v(1)$ seems to parametrize a subset of \mathbb{R}^8 that is at most 5-dimensional. The abnormal set of $G = \mathbb{R}^8$ is then presumably 6-dimensional, while in

Theorem 4.1 abnormal extremals are shown to be in a set with dimension less than or equal to 7.

5. Spiral-like Goh extremals

The shortening technique introduced in [10] has two steps: a curve with a corner at a given point is blown up; the limit curve obtained in this way is shown not to be length minimizing. This provides an “almost- C^1 ” regularity result for sub-Riemannian length minimizing curves. The technique, however, fails when, in the blow up at the singular point, the curve has no proper angle. Here, we show that there do exist Goh extremals of this kind. The example is a generalization of [8, Example 6.4].

Let G be an n -dimensional stratified nilpotent Lie group with Lie algebra $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_s$, where $s \geq 3$ is step of the group. An abnormal extremal $\gamma : [0, 1] \rightarrow G$ with dual curve $\lambda = \lambda_1 \vartheta_1 + \cdots + \lambda_n \vartheta_n$ is said to be a Goh extremal if $\lambda_i = 0$ for all $i = 1, \dots, n$ such that $d(i) = 1$ or $d(i) = 2$. By Theorem 3.4, a horizontal curve $\gamma : [0, 1] \rightarrow G$ with $\gamma(0) = 0$ is a Goh extremal precisely when there exists $v \in \mathbb{R}^n$, $v \neq 0$, such that

$$(5.53) \quad P_i^v(\gamma(t)) = 0, \quad \text{for all } t \in [0, 1] \text{ and } d(i) \in \{1, 2\}.$$

By Theorem 2.2, the condition $P_i^v(\gamma) = 0$ for $d(i) = 2$ is equivalent to (5.53).

Let F be the free nilpotent Lie group of rank 3 and step 4 and consider the direct product $G = F \times F$. As F is diffeomorphic to \mathbb{R}^{32} , then G is a stratified Lie group of rank 6 and step 4 diffeomorphic to \mathbb{R}^{64} . We fix a basis X_1, \dots, X_{64} of $\text{Lie}(G)$ adapted to the stratification and we identify G with \mathbb{R}^{64} via exponential coordinates of the second type. We reorder and relabel the basis as $Y_1, \dots, Y_{32}, Z_1, \dots, Z_{32}$, where Y_1, \dots, Y_{32} is an adapted basis of a first copy of $\text{Lie}(F)$ and Z_1, \dots, Z_{32} is an adapted basis of a second copy of $\text{Lie}(F)$. We denote the corresponding coordinates on G by $(y, z) \in \mathbb{R}^{64}$ with $y, z \in \mathbb{R}^{32}$.

Notice that we have $[Y_i, Z_j] = 0$ for all $i, j = 1, \dots, 32$. Then any abnormal polynomial $P_i^v(x)$ of G , $i = 1, \dots, 64$ and $v \in \mathbb{R}^{64}$, splits as

$$P_i^v(x) = Q_j^{v_y}(y) + Q_k^{v_z}(z), \quad x = (y, z),$$

for some $j, k = 1, \dots, 32$ and $v_y, v_z \in \mathbb{R}^{32}$, where Q_1^v, \dots, Q_{32}^v , $v \in \mathbb{R}^{32}$, are the extremal polynomials of F . In particular, for $i = 4, 5, 6$ we have the

polynomials:

$$\begin{aligned}
 Q_4^v(y) &= v_4 - y_1 v_7 - y_2 v_8 - y_3 v_9 + \frac{y_1^2}{2} v_{15} + y_1 y_2 v_{16} + y_1 y_3 v_{17} + \frac{y_2^2}{2} v_{18} \\
 &\quad + y_2 y_3 v_{19} + \frac{y_3^2}{2} v_{20} + y_5 v_{30} + y_6 v_{31} \\
 Q_5^v(y) &= v_5 - y_1 v_{10} - y_2 v_{11} - y_3 v_{12} + \frac{y_1^2}{2} v_{21} + y_1 y_2 v_{22} + y_1 y_3 v_{23} + \frac{y_2^2}{2} v_{24} \\
 &\quad + y_2 y_3 v_{25} + \frac{y_3^2}{2} v_{26} - y_4 v_{30} + y_6 v_{32} \\
 Q_6^v(y) &= v_6 + y_1(v_9 - v_{11}) - y_2 v_{13} - y_3 v_{14} - \frac{y_1^2}{2} v_{17} - y_1 y_2 v_{19} - y_1 y_3 v_{20} + \frac{y_1^2}{2} v_{22} \\
 &\quad + y_1 y_2 v_{24} + y_1 y_3 v_{25} + \frac{y_3^2}{2} v_{29} + y_1^2 v_{30} + (y_1 y_2 - y_4) v_{31} - y_5 v_{32}.
 \end{aligned}$$

We omit the deduction of these formulas.

Let $v, w \in \mathbb{R}^{32}$ be such that $v_i = w_i = 0$ for all $i = 1, \dots, 6$ and $|v| + |w| \neq 0$. By Theorem 3.4, any horizontal curve $\gamma = (\gamma_y, \gamma_z) : [0, 1] \rightarrow G$ such that $\gamma(0) = 0$ and with support contained in the algebraic set

$$\Sigma_{v,w} = \{(y, z) \in \mathbb{R}^{64} : Q_i^v(y) = Q_i^w(z) = 0 \text{ for } i = 4, 5, 6\}$$

is a Goh extremal. Notice that, by (2.15), $Q_i^v(\gamma_y) = Q_i^w(\gamma_z) = 0$ for $i = 4, 5, 6$ implies $Q_i^v(\gamma_y) = Q_i^w(\gamma_z) = 0$ also for $i = 1, 2, 3$. We choose $v \in \mathbb{R}^{32}$ such that $v_7 = 1$, $v_{18} = 2$, and $v_j = 0$ otherwise. Then we have

$$Q_4^v(y) = y_2^2 - y_1, \quad Q_5^v(y) = Q_6^v(y) = 0, \quad y = (y_1, \dots, y_{32}).$$

For any pair of Lipschitz functions $\phi, \psi : [-1, 1] \rightarrow \mathbb{R}$ with $\phi(0) = \psi(0) = 0$, the horizontal curve $\gamma = (\gamma_y, \gamma_z) : [-1, 1] \rightarrow \mathbb{R}^{64}$ defined by $\gamma_y(t) = (t^2, t, \phi(t), *, \dots, *)$ and $\gamma_z(t) = (t^2, t, \psi(t), *, \dots, *)$ has support contained in $\Sigma_{v,v}$ and therefore it is a Goh extremal.

In particular, as in [4], we can choose

$$\phi(t) = t \cos(\log(1 - \log |t|)), \quad \psi(t) = t \sin(\log(1 - \log |t|)), \quad t \in [-1, 1],$$

where $\phi(0) = \psi(0) = 0$ in the limit sense. The curve γ is a kind of rectifiable spiral. Notice that $\max(|\phi'(t)|, |\psi'(t)|) \leq 2$ for all $t \in (0, 1)$. Moreover, for any $\alpha \in [0, 2\pi)$ there exists a positive and infinitesimal sequence $(\lambda_k)_{k \in \mathbb{N}}$

such that

$$\lim_{k \rightarrow \infty} \frac{1}{\lambda_k} (\phi(\lambda_k t), \psi(\lambda_k t)) = t(\cos \alpha, \sin \alpha), \quad t \in \mathbb{R}.$$

Viceversa, if the limit in the left hand side does exist for some infinitesimal sequence $(\lambda_k)_{k \in \mathbb{N}}$, then it is of the given form for some $\alpha \in [0, 2\pi)$.

The curve γ is not C^1 at $t = 0$ and all its tangents at $t = 0$ are lines. It is a natural question whether the Goh extremal γ is length minimizing or not. So far we were not able to answer this question. The unique result on the nonminimality of extremals with singular points is the removability of corners, see [10]. This was obtained through a shortening technique that consists of two steps. In a first step, the curve and the structure are blown up at the singular point getting in the limit a “canonical flat corner” (the union of two perpendicular half lines). In a second step, the flat corner is proved not to be length minimizing by constructing a shorter curve cutting the corner and adjusting the end-point. Because length minimality is preserved in the limit, this proves the non-minimality of the original corner. For the spiral γ , however, this shortening technique fails at the very first step because any blow-up of γ at the singular point is a line, that is indeed length minimizing.

Acknowledgments

It is a pleasure to thank Ben Warhurst and Alessandro Ottazzi for many illuminating discussions on Tanaka prolongation. We also thank Igor Zelenko and Francesco Fassò for some discussions on a preliminary version of the paper.

This work was partially supported by INDAM, University of Padova research project *Some analytic and differential geometric aspects in Nonlinear Control Theory, with applications to Mechanics*, and Fondazione CaRi-PaRo Project *Nonlinear Partial Differential Equations: models, analysis, and control-theoretic problems*.

References

- [1] A. AGRACHEV, *Any sub-Riemannian metric has points of smoothness*. (Russian) Dokl. Akad. Nauk 424 (2009), no. 3, 295–298; translation in Dokl. Math. 79 (2009), no. 1, 45–47.
- [2] A. AGRACHEV, *Some open problems*. Geometric Control Theory and sub-Riemannian Geometry, Springer, INdAM Series 5 (2014), 1–14. G.

- Stefani, U. Boscain, J.-P. Gauthier, A. Sarychev, M. Sigalotti (eds.).
- [3] A. AGRACHEV & Y. L. SACHKOV, *Control Theory from the Geometric Viewpoint*. Encyclopaedia of Mathematical Sciences, 87. Control Theory and Optimization, II. Springer-Verlag, Berlin, 2004.
 - [4] E. GONZALES, U. MASSARI, I. TAMANINI, *Boundaries of prescribed mean curvature*. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. 4 (1993), no. 3, 197–206.
 - [5] M. GRAYSON & R. GROSSMAN, *Models for free nilpotent Lie algebras*. J. Algebra 135 (1990), no. 1, 177–191.
 - [6] M. HALL JR., *A basis for free Lie rings and higher commutators in free groups*. Proc. Amer. Math. Soc. 1, (1950), 575–581.
 - [7] A. W. KNAPP, *Lie groups beyond an introduction*. Progress in Mathematics, 140. Birkhäuser 2002.
 - [8] E. LE DONNE, G. P. LEONARDI, R. MONTI, D. VITTONI, *Extremal curves in nilpotent Lie groups*. Geom. Funct. Anal. 23 (2013), no. 4, 1371–1401.
 - [9] E. LE DONNE, R. MONTGOMERY, A. OTTAZZI, P. PANSU, D. VITTONI, *Sard Property for the endpoint map on some Carnot groups*, in preparation.
 - [10] G. P. LEONARDI & R. MONTI, *End-point equations and regularity of sub-Riemannian geodesics*. Geom. Funct. Anal. 18 (2008), no. 2, 552–582.
 - [11] R. MONTGOMERY, *A Tour of Sub-Riemannian Geometries, Their Geodesics and Applications*. AMS, 2002.
 - [12] R. MONTI, *Regularity results for sub-Riemannian geodesics*. Calc. Var. Partial Differential Equations 49 (2014), no. 1–2, 549–582.
 - [13] R. MONTI, *The regularity problem for sub-Riemannian geodesics*. Geometric Control Theory and sub-Riemannian Geometry, Springer, INdAM Series 5 (2014), 313–332. G. Stefani, U. Boscain, J.-P. Gauthier, A. Sarychev, M. Sigalotti (eds.).
 - [14] R. MONTI, *A family of nonminimizing abnormal curves*. Ann. Mat. Pura Appl. 193 (2013), no. 6, 1577–1593.
 - [15] L. RIFFORD & E. TRELAT, *Morse-Sard type results in sub-Riemannian geometry*. Math. Ann. 332 (2005), 145–159.

- [16] N. TANAKA, *On differential systems, graded Lie algebras and pseudogroups*. J. Math. Kyoto Univ. 10 (1970), 1–82.
- [17] B. WARHURST, *Tanaka prolongation of free Lie algebras*. Geom. Dedicata, 130 (2007), 59–69.
- [18] I. ZELENKO, *Nonregular abnormal extremals of 2-distribution: existence, second variation and rigidity*. J. Dynamical and Control systems, 5 (1999), no. 3, 347–383.
- [19] I. ZELENKO & M. ZHITOMIRSKIĬ, *Rigid paths of generic 2-distributions on 3-manifolds*. Duke Math. J. 79 (1995), no. 2, 281–307.

[Le Donne]Department of Mathematics and Statistics, P.O. Box 35, FIN-40014, University of Jyväskylä, Finland *E-mail address*: ledonne@msri.org

[Leonardi]Università di Modena e Reggio Emilia, Dipartimento di Scienze Fisiche, Informatiche e Matematiche, via Campi 213/b, 41100 Modena, Italy *E-mail address*: gianpaolo.leonardi@unimore.it

[Monti and Vittonè]Università di Padova, Dipartimento di Matematica, via Trieste 63, 35121 Padova, Italy *E-mail address*: monti@math.unipd.it

[Vittonè]Universität Zürich, Institut für Mathematik, Winterthurerstrasse 190, 8057 Zürich, Switzerland *E-mail address*: vittonè@math.unipd.it

Received January 31, 2015