SIAM J. CONTROL OPTIM.
Vol. 47, No. 4, pp. 1851–1878

INVARIANT CARNOT–CARATHEODORY METRICS ON S^3, $SO(3)$, $SL(2)$, AND LENS SPACES*

UGO BOSCAIN† AND FRANCESCO ROSSI‡

Abstract. In this paper we study the Carnot–Caratheodory metrics on $SU(2) \simeq S^3$, $SO(3)$, and $SL(2)$ induced by their Cartan decomposition and by the Killing form. Besides computing explicitly geodesics and conjugate loci, we compute the cut loci (globally), and we give the expression of the Carnot–Caratheodory distance as the inverse of an elementary function. We then prove that the metric given on $SU(2)$ projects on the so-called lens spaces $L(p, q)$. Also for lens spaces, we compute the cut loci (globally). For $SU(2)$ the cut locus is a maximal circle without one point. In all other cases the cut locus is a stratified set. To our knowledge, this is the first explicit computation of the whole cut locus in sub-Riemannian geometry, except for the trivial case of the Heisenberg group.

Key words. left-invariant sub-Riemannian geometry, Carnot–Caratheodory distance, global structure of the cut locus, lens spaces

AMS subject classifications. 22E30, 49J15, 53C17

DOI. 10.1137/070703727

1. Introduction.

In this paper we study the global structure of the cut locus (the set of points reached optimally by more than one geodesic) for the simplest sub-Riemannian structures on three-dimensional simple Lie groups (i.e., $SU(2)$, $SO(3)$, and $SL(2)$), namely, the left-invariant sub-Riemannian structures induced by their Cartan decomposition and by the Killing form.

Let G be a simple real Lie group of matrices with associated Lie algebra $\mathbf{L}$ and Killing form $\mathrm{Kil}(\cdot,\cdot)$. Let $\mathbf{L} = \mathbf{k} \oplus \mathbf{p}$ be its Cartan decomposition with the usual commutation relations $[\mathbf{k},\mathbf{k}] \subseteq \mathbf{k}$, $[\mathbf{p},\mathbf{p}] \subseteq \mathbf{k}$, $[\mathbf{k},\mathbf{p}] \subseteq \mathbf{p}$. If $\mathbf{L}$ is noncompact, we also require $\mathbf{k}$ to be the maximal compact subalgebra of $\mathbf{L}$. The most natural left-invariant sub-Riemannian structure that one can define on G is the one in which the distribution is generated by left translations of $\mathbf{p}$, and the sub-Riemannian metric $\langle\cdot,\cdot\rangle$ at the identity is generated by a scalar multiple of the Killing form restricted to $\mathbf{p}$. The scalar must be chosen positive or negative in such a way that the scalar product is positive definite. We call G, endowed with such a sub-Riemannian structure, a $\mathbf{k}\oplus\mathbf{p}$ **sub-Riemannian manifold**.

$\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifolds have very special features: there are no strict abnormal minimizers, and the Hamiltonian system given by the Pontryagin maximum principle (PMP) is integrable in terms of elementary functions (products of exponentials). More precisely, if we write the distribution at a point $g \in G$ as $\Delta(g) = g\mathbf{p}$, we have the following expression for geodesics parametrized by arclength, starting at time zero from g_0 (see [3, 10, 14, 22, 23]):

$$g(t) = g_0 e^{(A_k+A_p)t} e^{-A_k t}, \tag{1}$$

where $A_k \in \mathbf{k}$, $A_p \in \mathbf{p}$, and we have $\langle A_p, A_p\rangle = 1$. Thanks to left-invariance, without loss of generality we can always assume that g_0 is the identity, and we will do so throughout the paper.

*Received by the editors September 25, 2007; accepted for publication (in revised form) February 17, 2008; published electronically June 25, 2008. The first author was partially supported by a FABER grant from Région Bourgogne.
http://www.siam.org/journals/sicon/47-4/70372.html
†LE2i, CNRS UMR5158, Université de Bourgogne, 9, avenue Alain Savary - BP 47870, 21078 Dijon Cedex, France (boscain@sissa.it).
‡SISSA, via Beirut 2-4, 34014 Trieste, Italy (rossifr@sissa.it).

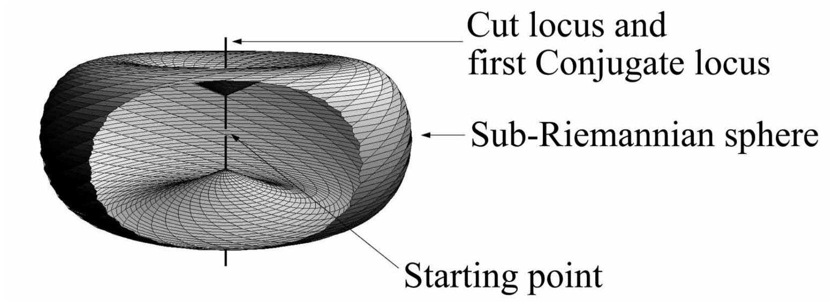

FIG. 1. *Local structure of sub-Riemannian spheres and of cut and conjugate loci for* 3-dim $\mathbf{k}\oplus\mathbf{p}$ *sub-Riemannian manifolds.*

In all three-dimensional cases (i.e., $SU(2)$, $SO(3)$, and $SL(2)$), $\mathbf{p}$ has dimension 2, while $\mathbf{k}$ has dimension 1. Writing $\mathbf{p} = \text{span}\,\{p_1, p_2\}$, where $\{p_1, p_2\}$ is an orthonormal frame for the sub-Riemannian structure (i.e., $\langle p_i, p_j\rangle = \delta_{ij}$) and $\mathbf{k} = \text{span}\,\{k\}$, we can write $A_p = \cos(\theta)p_1 + \sin(\theta)p_2$ and $A_k = ck$ with $\theta \in \mathbb{R}/2\pi,\ c \in \mathbb{R}$. The map associating to the triple (θ, c, t) the final point of the corresponding geodesic starting from the identity is called the *exponential map*,

$$\text{Exp} : \begin{array}{ccc} S^1 \times \mathbb{R} \times \mathbb{R}^+ & \to & G, \\ (\theta, c, t) & \mapsto & \text{Exp}(\theta, c, t) = e^{(A_k+A_p)t}e^{-A_k t}. \end{array}$$

For three-dimensional $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifolds, the local structure of the sub-Riemannian spheres, cut loci, and conjugate loci starting from the identity has been described by Agrachev (unpublished), and, due to cylindrical symmetry of the Killing form in the $\mathbf{p}$ subspace, it is very similar to that of the Heisenberg group. Indeed, locally, the cut locus coincides with the first conjugate locus (i.e., the set where local optimality is lost) and is made by two connected one-dimensional manifolds adjacent to the identity and transversal to the distribution; see Figure 1.

However, the global structure of the cut locus was still unknown. Indeed, to our knowledge, no global structure of the cut locus is known in sub-Riemannian geometry apart from that of the Heisenberg group.

The main result of our paper is the following.

THEOREM 1. *Let K_{Id} be the cut locus starting from the identity. We have the following:*

(i) *For $SU(2)$, K_{Id} is a maximal circle S^1 without one point (the identity).*

(ii) *For $SO(3)$, K_{Id} is a stratified set made by two manifolds glued in one point. The first manifold is $\mathbb{RP}^2$; the second manifold is a maximal circle S^1 without one point (the identity).*

(iii) *For $SL(2)$, K_{Id} is a stratified set made by two manifolds glued in one point. The first manifold is $\mathbb{R}^2$; the second manifold is a circle S^1 without one point (the identity).*

A picture of the three cut loci is given in Figure 2.

For all cases, the one-dimensional strata contain the cut locus appearing in the local analysis.

Notice that the $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifold $SU(2)$ has the structure of a CR (Cauchy–Riemann) manifold and is a tight structure [7, 16].

Once the cut locus is computed, one can obtain the expression of the sub-Riemannian distance from the identity. The following theorem gives the sub-Riemannian distance for $SU(2)$. The proof, given in section 5.1.1, can be adapted

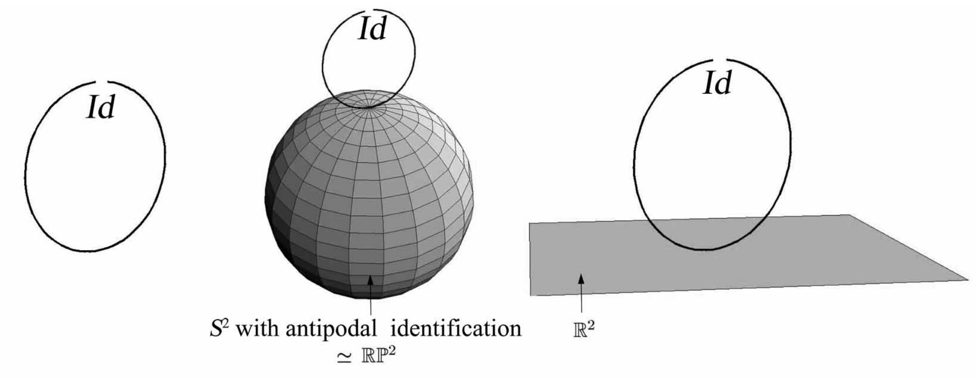

FIG. 2. *The cut loci for the* $\mathbf{k} \oplus \mathbf{p}$ *sub-Riemannian manifolds* $SU(2)$, $SO(3)$, *and* $SL(2)$.

to get similar results in the cases of $SO(3)$ and $SL(2)$.

THEOREM 2. *Let*

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \mid \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

Consider the sub-Riemannian distance from Id *defined by*

$$\dot{g} = g \left(\frac{u_1}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + \frac{u_2}{2} \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \right),$$

$$d(\mathrm{Id}, g_1) := \inf \left\{ \int_0^T \sqrt{u_1^2 + u_2^2} \mid g(0) = \mathrm{Id},\ g(T) = g_1 \right\}.$$

It holds that

$$d\left(\mathrm{Id}, \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}\right) = \begin{cases} 2\sqrt{\arg(\alpha)\,(2\pi - \arg(\alpha))} & \textit{if} \quad \beta = 0, \\ \psi(\alpha) & \textit{if} \quad \beta \neq 0, \end{cases} \tag{2}$$

where $\arg(\alpha) \in [0, 2\pi]$ *and* $\psi(\alpha) = t$ *is the unique solution of*

$$\begin{cases} -\frac{ct}{2} + \arctan\left(\frac{c}{\sqrt{1+c^2}} \tan\left(\frac{\sqrt{1+c^2}t}{2}\right)\right) = \arg(\alpha), \\ \frac{\sin\left(\frac{\sqrt{1+c^2}t}{2}\right)}{\sqrt{1+c^2}} = \sqrt{1 - |\alpha|^2}, \\ t \in \left(0, \frac{2\pi}{\sqrt{1+c^2}}\right). \end{cases} \tag{3}$$

This theorem and its analogues for $SO(3)$ and $SL(2)$ are useful to give estimates for the fundamental solutions of the hypoelliptic heat equation induced by the sub-Riemannian structure (see [5, 12, 17, 19]). Moreover, this theorem can be seen as the answer, in the case of $SU(2)$, to the question (formulated in [14]) about the possibility of inverting the matrix equation (1); i.e., for every matrix $g \in SU(2)$, find a matrix $A = A_k + A_p$, with $\langle A_p, A_p \rangle = 1$, being a solution to the equation $g = g_0 e^{(A_k + A_p)t} e^{-A_k t}$. If $\beta \neq 0$, then this equation has one and only one solution; otherwise it has more than one solution (indeed, infinitely many; see sections 3 and 5).

Then we study the most natural sub-Riemannian structures on the lens spaces $L(p,q)$ induced by the one on $SU(2)$. The lens space $L(p,q)$ (with p,q coprime integers, $p,q \neq 0$) is the quotient of $SU(2)$ by the equivalence relation

$$\left(\begin{array}{cc} \alpha_1 & \beta_1 \\ -\overline{\beta_1} & \overline{\alpha_1} \end{array} \right) \sim \left(\begin{array}{cc} \alpha_2 & \beta_2 \\ -\overline{\beta_2} & \overline{\alpha_2} \end{array} \right) \text{ if } \exists\ \omega \in \mathbb{C}\ p\text{th root of unity such that}$$

$$\left(\begin{array}{c} \alpha_2 \\ \beta_2 \end{array} \right) = \left(\begin{array}{cc} \omega & 0 \\ 0 & \omega^q \end{array} \right) \left(\begin{array}{c} \alpha_1 \\ \beta_1 \end{array} \right).$$

The lens spaces are three-dimensional manifolds, but they are neither Lie groups nor homogeneous spaces of $SU(2)$, except for the case $L(2,1) \simeq SO(3)$.

In the case of lens spaces, we get that the cut locus is much more complicated with respect to those on $SU(2)$ and $SL(2)$. It is still a stratified set, but in general with more strata. The precise description is given in section 5.2.

Sub-Riemannian structures on the lens space $L(4,1)$ are particularly interesting for mechanical applications and for problems of geometry of vision on the two-dimensional sphere. Indeed, $L(4,1) \simeq PTS^2$, the bundle of directions of S^2. These applications are the subject of a forthcoming paper.

The structure of this paper is the following. In section 2 we recall the definition of sub-Riemannian manifold, state the PMP (that is a first order necessary condition for optimality for problems of calculus of variations with nonholonomic constraints), and define the cut and conjugate loci. Then we define $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifolds. In section 3 we define $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian structures on $SU(2), SO(3)$, and $SL(2)$ and compute the corresponding geodesics and conjugate loci. In section 4 we give sub-Riemannian structures on lens spaces as quotients of the $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian structure on $SU(2)$. The core of the paper is section 5, where we compute the cut loci and the sub-Riemannian distance. The general idea is the following: we first identify the prolongation of the cut locus arising locally, then we compute the part of the cut locus due to the symmetries of the problem, and finally, we show that there is no other cut point.

2. Basic definitions.

2.1. Sub-Riemannian manifold.

An (n,m)-sub-Riemannian manifold is a triple $(M, \Delta, \mathbf{g})$, where

(i) M is a connected smooth manifold of dimension n;

(ii) Δ is a Lie bracket generating smooth distribution of constant rank $m < n$; i.e., Δ is a smooth map that associates to $q \in M$ an m-dim subspace $\Delta(q)$ of T_qM, and $\forall\ q \in M$, we have

(4)
$$\text{span}\,\{[f_1, [\dots[f_{k-1}, f_k]\dots]](q) \mid f_i \in \text{Vec}(M) \text{ and } f_i(p) \in \Delta(p)\ \forall\ p \in M\} = T_qM.$$

Here Vec(M) denotes the set of smooth vector fields on M.

(iii) $\mathbf{g}_q$ is a Riemannian metric on $\Delta(q)$, that is, smooth as a function of q.

The Lie bracket generating condition (4) is also known as the Hörmander condition.

A Lipschitz continuous curve $\gamma : [0,T] \to M$ is said to be *horizontal* if $\dot\gamma(t) \in \Delta(\gamma(t))$ for almost every $t \in [0,T]$. Given a horizontal curve $\gamma : [0,T] \to M$, the *length of* γ is

$$l(\gamma) = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}\; dt. \tag{5}$$

The *distance* induced by the sub-Riemannian structure on M is the function

$$d(q_0, q_1) = \inf\{l(\gamma) \mid \gamma(0) = q_0, \gamma(T) = q_1, \gamma \text{ horizontal}\}. \tag{6}$$

The hypothesis of connectedness of M and the Lie bracket generating assumption for the distribution guarantee the finiteness and the continuity of $d(\cdot,\cdot)$ with respect to the topology of M (Chow's theorem; see, for instance, [3]).

The function $d(\cdot,\cdot)$ is called the Carnot–Caratheodory distance and gives to M the structure of metric space (see [6, 18]).

It is a standard fact that $l(\gamma)$ is invariant under reparametrization of the curve γ. Moreover, if an admissible curve γ minimizes the so-called *energy functional*

$$E(\gamma) = \int_0^T \mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))\; dt$$

with T fixed (and fixed initial and final point), then $v = \sqrt{\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))}$ is constant and γ is also a minimizer of $l(\cdot)$. On the other hand, a minimizer γ of $l(\cdot)$ such that v is constant is a minimizer of $E(\cdot)$ with $T = l(\gamma)/v$.

A *geodesic* for the sub-Riemannian manifold is a curve $\gamma : [0,T] \to M$ such that for every sufficiently small interval $[t_1, t_2] \subset [0,T]$, $\gamma_{|_{[t_1,t_2]}}$ is a minimizer of $E(\cdot)$. A geodesic for which $\mathbf{g}_{\gamma(t)}(\dot\gamma(t), \dot\gamma(t))$ is (constantly) equal to one is said to be parametrized by arclength.

Locally, the pair $(\Delta, \mathbf{g})$ can be given by assigning a set of m smooth vector fields that are orthonormal for $\mathbf{g}$, i.e.,

$$\Delta(q) = \operatorname{span}\{F_1(q), \dots, F_m(q)\}, \quad \mathbf{g}_q(F_i(q), F_j(q)) = \delta_{ij}. \tag{7}$$

When $(\Delta, \mathbf{g})$ can be defined as in (7) by m vector fields defined globally, we say that the sub-Riemannian manifold is *trivializable*.

Given a (n,m)-trivializable sub-Riemannian manifold, the problem of finding a curve minimizing the energy between two fixed points $q_0, q_1 \in M$ is naturally formulated as the optimal control problem

$$\dot q = \sum_{i=1}^m u_i F_i(q), \quad u_i \in \mathbb{R}, \quad \int_0^T \sum_{i=1}^m u_i^2(t)\; dt \to \min, \quad q(0) = q_0, \quad q(T) = q_1. \tag{8}$$

It is a standard fact that this optimal control problem is equivalent to the minimum time problem with controls $u_1, \dots, u_m$ satisfying $u_1^2 + \cdots + u_m^2 \le 1$.

When the manifold is analytic and the orthonormal frame can be assigned through m analytic vector fields, we say that the sub-Riemannian manifold is *analytic*.

In this paper we are concerned with sub-Riemannian manifolds that are trivializable and analytic since they are given in terms of left-invariant vector fields on Lie groups.

2.2. First order necessary conditions, cut locus, and conjugate locus. Consider a trivializable (n,m)-sub-Riemannian manifold. Solutions to the optimal control problem (8) are computed via the PMP (see, for instance, [3, 11, 21, 24]) that is a first order necessary condition for optimality and generalizes the Weierstraß conditions of the calculus of variations. For each optimal curve, the PMP provides a lift to the cotangent bundle that is a solution to a suitable pseudo-Hamiltonian system.

THEOREM 3 (PMP for the problem (8)). *Let M be an n-dimensional smooth manifold, and consider the minimization problem* (8)*, in the class of Lipschitz continuous curves, where F_i, $i=1,\ldots,m$ are smooth vector fields on M and the final time T is fixed. Consider the map $\mathbf{H}: T^*M\times\mathbb{R}\times\mathbb{R}^m\to\mathbb{R}$ defined by*

$$\mathbf{H}(q,\lambda,p_0,u) := \left\langle \lambda, \sum_{i=1}^m u_i F_i(q) \right\rangle + p_0 \sum_{i=1}^m u_i^2(t).$$

If the curve $q(.) : [0,T]\to M$ corresponding to the control $u(.) : [0,T]\to\mathbb{R}^m$ is optimal, then there exist a never vanishing Lipschitz continuous covector $\lambda(.) : t\in[0,T]\mapsto\lambda(t)\in T^*_{q(t)}M$ and a constant $p_0\le 0$ such that, for a.e. $t\in[0,T]$,

(i) $\dot q(t) = \frac{\partial \mathbf{H}}{\partial \lambda}(q(t),\lambda(t),p_0,u(t))$,

(ii) $\dot\lambda(t) = -\frac{\partial \mathbf{H}}{\partial q}(q(t),\lambda(t),p_0,u(t))$, *and*

(iii) $\frac{\partial \mathbf{H}}{\partial u}(q(t),\lambda(t),p_0,u(t)) = 0$.

Remark 1. A curve $q(.) : [0,T]\to M$ satisfying the PMP is said to be an *extremal*. In general, an extremal may correspond to more than one pair $(\lambda(.),p_0)$. If an extremal satisfies the PMP with $p_0\neq 0$, then it is called a *normal extremal*. If it satisfies the PMP with $p_0=0$ it is called an *abnormal extremal*. An extremal can be both normal and abnormal. For normal extremals one can normalize $p_0=-1/2$.

If an extremal satisfies the PMP only with $p_0=0$, then it is called a *strict abnormal extremal*. If a strict abnormal extremal is optimal, then it is called *a strict abnormal minimizer*. For a deep analysis of abnormal extremals in sub-Riemannian geometry, see [8, 15].

It is well known that all normal extremals are geodesics (see, for instance, [3]). Moreover, if there are no strict abnormal minimizers, then all geodesics are normal extremals for some fixed final time T. This always will be the case in this paper; indeed, we are concerned with sub-Riemannian manifolds of dimension 3, defined by a pair of vector fields F_1 and F_2 such that $\forall\, q\in M$, $\mathrm{span}\,\{F_1(q),F_2(q),[F_1(q),F_2(q)]\} = T_qM$, i.e., the so called three-dimensional *contact case*, for which there are no abnormal extremals (not even nonstrict).

In this case, from (iii) one gets $u_i(t) = \langle\lambda(t),F_i(t)\rangle$, $i=1,\ldots,m$, and the PMP becomes much simpler: a curve $q(.)$ is a geodesic if and only if it is the projection on M of a solution $(\lambda(t),q(t))$ for the Hamiltonian system on T^*M corresponding to

$$H(\lambda,q) = \frac{1}{2}\left(\sum_{i=1}^m \langle\lambda,F_i(q)\rangle^2\right), \qquad q\in M, \quad \lambda\in T_q^*M$$

satisfying $H(\lambda(0),q(0))\neq 0$.

Remark 2. Notice that H is constant along any given solution of the Hamiltonian system. Moreover, $H=\frac{1}{2}$ if and only if the geodesic is parametrized by arclength. In the following, for simplicity of notation we assume that all geodesics are defined for $t\in[0,+\infty)$.

Fix $q_0\in M$. For every $\lambda_0\in T^*_{q_0}M$ satisfying

$$H(\lambda_0,q_0) = 1/2 \tag{9}$$

and every $t > 0$, define the *exponential map* $\mathrm{Exp}(\lambda_0, t)$ as the projection on M of the solution, evaluated at time t, of the Hamiltonian system associated with H, with initial condition $\lambda(0) = \lambda_0$ and $q(0) = q_0$. Notice that condition (9) defines a hypercylinder $\Lambda_{q_0} \simeq S^{m-1} \times \mathbb{R}^{n-m}$ in $T^*_{q_0}M$.

DEFINITION 4. *The* **conjugate locus from** q_0 *is the set* C_{q_0} *of critical values of the map*

$$\mathrm{Exp}: \begin{array}{ccc} \Lambda_{q_0} \times \mathbb{R}^+ & \to & M, \\ (\lambda_0, t) & \mapsto & \mathrm{Exp}(\lambda_0, t). \end{array}$$

For every $\bar\lambda_0 \in \Lambda_{q_0}$*, let* $t(\bar\lambda_0)$ *be the nth positive time, if it exists, for which the map* $(\lambda_0, t) \mapsto \mathrm{Exp}(\lambda_0, t)$ *is singular at* $(\bar\lambda_0, t(\bar\lambda_0))$*. The* **nth conjugate locus** *from* q_0 $C^n_{q_0}$ *is the set* $\{\mathrm{Exp}(\bar\lambda_0, t(\bar\lambda_0)) \mid t(\bar\lambda_0) exists\}$.

The **cut locus** *from* q_0 *is the set* K_{q_0} *of points reached optimally by more than one geodesic, i.e., the set*

$$K_{q_0} = \left\{ q \in M \mid \exists \begin{array}{c} \lambda_1, \lambda_2 \in \Lambda_{q_0},\ \lambda_1 \neq \lambda_2, \\ t \in \mathbb{R}^+ \end{array} \ \textit{such that} \ \begin{array}{l} q = \mathrm{Exp}(\lambda_1, t), \\ q = \mathrm{Exp}(\lambda_2, t), \\ \mathrm{Exp}(\lambda_1, \cdot) \textit{ optimal in } [0,t], \\ \mathrm{Exp}(\lambda_2, \cdot) \textit{ optimal in } [0,t]. \end{array} \right\}$$

Remark 3. It is a standard fact that for every $\bar\lambda_0$ satisfying (9), the set $T(\bar\lambda_0) = \{\bar t > 0 \mid \mathrm{Exp}(\lambda, t)$ is singular at $(\bar\lambda_0, \bar t)\}$ is a discrete set (see, for instance, [3]).

Remark 4. Let $(M, \Delta, \mathbf{g})$ be a sub-Riemannian manifold. Fix $q_0 \in M$ and assume that

(i) Each point of M is reached by an optimal geodesic starting from q_0;

(ii) there are no abnormal minimizers.

The following facts are well known (a proof in the three-dimensional contact case can be found in [4]):

(i) The first conjugate locus $C^1_{q_0}$ is the set of points where the geodesics starting from q_0 lose local optimality;

(ii) if $q(.)$ is a geodesic starting from q_0, and $\bar t$ is the first positive time such that $q(\bar t) \in K_{q_0} \cup C^1_{q_0}$, then $q(.)$ loses optimality in $\bar t$; i.e., it is optimal in $[0, \bar t]$ and not optimal in $[0, t]$ for any $t > \bar t$;

(iii) if a geodesic $q(.)$ starting from q_0 loses optimality at $\bar t > 0$, then $q(\bar t) \in K_{q_0} \cup C^1_{q_0}$.

As a consequence, when the first conjugate locus is included in the cut locus (as in our cases; see section 5), the cut locus is the set of points where the geodesics lose optimality.

Remark 5. It is well known that, while in Riemannian geometry K_{q_0} is never adjacent to q_0, in sub-Riemannian geometry this is always the case. See [2].

2.3. $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifolds. For the sake of simplicity in the exposition, throughout the paper, when we deal with Lie groups and Lie algebras, we always consider that they are groups and algebras of matrices.

Let $\mathbf{L}$ be a simple Lie algebra and $\mathrm{Kil}(X, Y) = Tr(ad_X \circ ad_Y)$ its Killing form. Recall that the Killing form defines a nondegenerate pseudoscalar product on $\mathbf{L}$. In the following we recall what we mean by a Cartan decomposition of $\mathbf{L}$.

DEFINITION 5. *A Cartan decomposition of a simple Lie algebra* $\mathbf{L}$ *is any decomposition of the form*

$$\mathbf{L} = \mathbf{k} \oplus \mathbf{p}, \ \textit{where } [\mathbf{k}, \mathbf{k}] \subseteq \mathbf{k}, \quad [\mathbf{p}, \mathbf{p}] \subseteq \mathbf{k}, \quad [\mathbf{k}, \mathbf{p}] \subseteq \mathbf{p}. \tag{10}$$

DEFINITION 6. *Let G be a simple Lie group with Lie algebra $\mathbf{L}$. Let $\mathbf{L} = \mathbf{k} \oplus \mathbf{p}$ be a Cartan decomposition of $\mathbf{L}$. In the case in which G is noncompact, assume that $\mathbf{k}$ is the maximal compact subalgebra of $\mathbf{L}$.*

On G, consider the distribution $\Delta(g) = g\mathbf{p}$ endowed with the Riemannian metric $\mathbf{g}_g(v_1, v_2) = \langle g^{-1}v_1, g^{-1}v_2 \rangle$, where $\langle\ ,\ \rangle := \alpha \operatorname{Kil}\big|_{\mathbf{p}}(\ ,\ )$ and $\alpha < 0$ (resp., $\alpha > 0$) if G is compact (resp., noncompact).

In this case we say that $(G, \Delta, \mathbf{g})$ is a $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifold.

The constant α is clearly not relevant. It is chosen just to obtain good normalizations.

Remark 6. In the compact (resp., noncompact) case, the fact that $\mathbf{g}$ is positive definite on Δ is guaranteed by the requirement $\alpha < 0$ (resp., by the requirements $\alpha > 0$ and $\mathbf{k}$ maximal compact subalgebra).

Let $\{X_j\}$ be an orthonormal frame for the subspace $\mathbf{p} \subset \mathbf{L}$, with respect to the metric defined in Definition 6. Then the problem of finding the minimal energy between the identity and a point $g_1 \in G$ in fixed time T becomes the left-invariant optimal control problem

$$\dot{g} = g\left(\sum_j u_j X_j\right),\ \ u_j \in L^\infty(0,T), \int_0^T \sum_j u_j^2(t)\ dt \to \min,\quad g(0) = \mathrm{Id},\quad g(T) = g_1.$$

This problem admits a solution; see, for instance, Chapter 5 of [13].

For $\mathbf{k}\oplus\mathbf{p}$ sub-Riemannian manifolds, one can prove that strict abnormal extremals are never optimal, since the *Goh condition* (see [3]) is never satisfied. Moreover, the Hamiltonian system given by the PMP is integrable and the explicit expression of geodesics starting from the identity and parametrized by arclength is

$$g(t) = e^{(A_k + A_p)t} e^{-A_k t}, \tag{11}$$

where $A_k \in \mathbf{k}$, $A_p \in \mathbf{p}$, and $\langle A_p, A_p \rangle = 1$. This formula is well known in the community. It was used independently by Agrachev [1], Brockett [14], and Kupka (oral communication). The first complete proof was written by Jurdjevic in [22]. The proof that strict abnormal extremals are never optimal was first written in [10]. See also [3, 23].

Remark 7. In the three-dimensional case, the Hamiltonian system given by the PMP is indeed integrable even if the cost is not built with the Killing form (biinvariant) but is only left-invariant. For the case of $SO(3)$ see [9].

3. $SU(2)$, $SO(3)$, $SL(2)$, their geodesics, and their conjugate loci.

In this section we fix coordinates on $SU(2)$, $SO(3)$, $SL(2)$, and we apply formula (11) in order to get the explicit expressions for geodesics and conjugate loci.

3.1. The $\mathbf{k} \oplus \mathbf{p}$ problem on $SU(2)$.

The Lie group $SU(2)$ is the group of unitary unimodular 2×2 complex matrices

$$SU(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{C}) \mid |\alpha|^2 + |\beta|^2 = 1 \right\}.$$

It is compact and simply connected. The Lie algebra of $SU(2)$ is the algebra of anti-Hermitian traceless 2×2 complex matrices

$$su(2) = \left\{ \begin{pmatrix} i\alpha & \beta \\ -\overline{\beta} & -i\alpha \end{pmatrix} \in \mathrm{Mat}(2, \mathbb{C}) \mid \alpha \in \mathbb{R}, \beta \in \mathbb{C} \right\}.$$

A basis of $su(2)$ is $\{p_1, p_2, k\}$, where

$$p_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad p_2 = \frac{1}{2}\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \quad k = \frac{1}{2}\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \tag{12}$$

whose commutation relations are $[p_1, p_2] = k$, $[p_2, k] = p_1$, and $[k, p_1] = p_2$. Recall that for $su(n)$ we have $\mathrm{Kil}(X, Y) = 2n\mathrm{Tr}(XY)$ (see [20, pp. 186, 516]); thus for $su(2)$, $\mathrm{Kil}(X, Y) = 4\mathrm{Tr}(XY)$ and, in particular, $\mathrm{Kil}(p_i, p_j) = -2\delta_{ij}$. The choice of the subspaces

$$\mathbf{k} = \mathrm{span}\,\{k\} \qquad \mathbf{p} = \mathrm{span}\,\{p_1, p_2\}$$

provides a *Cartan decomposition* for $su(2)$. Moreover, $\{p_1, p_2\}$ is an orthonormal frame for the inner product $\langle\cdot,\cdot\rangle = -\frac{1}{2}\mathrm{Kil}(\cdot,\cdot)$ restricted to $\mathbf{p}$.

Defining $\Delta(g) = g\mathbf{p}$ and $\mathbf{g}_g(v_1, v_2) = \langle g^{-1}v_1, g^{-1}v_2\rangle$, we have that $(SU(2), \Delta, \mathbf{g})$ is a $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifold.

Remark 8. Observe that all the $\mathbf{k} \oplus \mathbf{p}$ structures that one can define on $SU(2)$ are equivalent. For instance, one could set $\mathbf{k} = \mathrm{span}\,\{p_1\}$ and $\mathbf{p} = \mathrm{span}\,\{p_2, k\}$.

Recall that

$$SU(2) \simeq S^3 = \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in \mathbb{C}^2 \mid |\alpha|^2 + |\beta|^2 = 1 \right\}$$

via the isomorphism

$$\phi : \begin{array}{ccc} SU(2) & \to & S^3, \\ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} & \mapsto & \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \end{array}$$

In the following we always write elements of $SU(2)$ as pairs of complex numbers.

3.1.1. Expression of geodesics. We compute the explicit expression of geodesics using the formula (11). Consider an initial covector $\lambda = \lambda(\theta, c) = \cos(\theta)p_1 + \sin(\theta)p_2 + ck \in \Lambda_{\mathrm{Id}}$. The corresponding exponential map is

$$\mathrm{Exp}(\theta, c, t) := \mathrm{Exp}(\lambda(\theta, c), t) = e^{(\cos(\theta)p_1 + \sin(\theta)p_2 + ck)t} e^{-ckt} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

with

$$\begin{aligned} \alpha = {} & \frac{c\sin(\frac{ct}{2})\sin(\sqrt{1+c^2}\,\frac{t}{2})}{\sqrt{1+c^2}} + \cos\left(\frac{ct}{2}\right)\cos\left(\sqrt{1+c^2}\,\frac{t}{2}\right) \\ & + i\left(\frac{c\cos(\frac{ct}{2})\sin(\sqrt{1+c^2}\,\frac{t}{2})}{\sqrt{1+c^2}} - \sin\left(\frac{ct}{2}\right)\cos\left(\sqrt{1+c^2}\,\frac{t}{2}\right)\right), \\ \beta = {} & \frac{\sin(\sqrt{1+c^2}\,\frac{t}{2})}{\sqrt{1+c^2}}\left(\cos\left(\frac{ct}{2} + \theta\right) + i\sin\left(\frac{ct}{2} + \theta\right)\right). \end{aligned}$$

We have the following symmetry properties:

(i) *cylindrical symmetry*:

$$\mathrm{Exp}(\theta, c, t) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} \mathrm{Exp}(0, c, t);$$

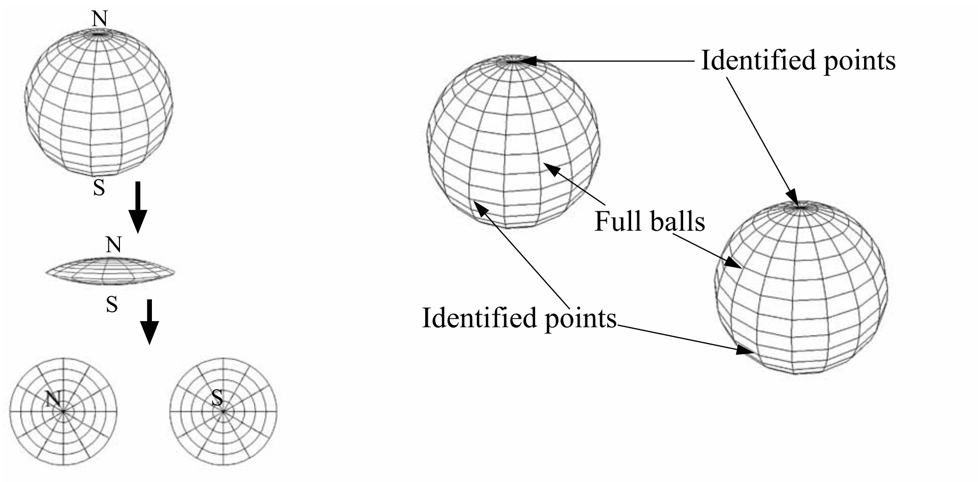

FIG. 3. *Left: construction of the* 2-dim *picture of* S^2. *Right: the* 3-dim *picture of* S^3.

(ii) *central symmetry*: Set $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \mathrm{Exp}(\theta, c, t)$. We have

$$\mathrm{Exp}(\theta, -c, t) = \begin{cases} \begin{pmatrix} \overline{\alpha} \\ e^{2i(\theta - \arg(\beta))}\beta \end{pmatrix} & \text{if} \quad \beta \neq 0, \\ \begin{pmatrix} \overline{\alpha} \\ 0 \end{pmatrix} & \text{if} \quad \beta = 0. \end{cases}$$

3.1.2. Pictures of S^2 and S^3. We recall a standard construction for representing S^2 in a two-dimensional space and S^3 in a three-dimensional space. For more details, see, e.g., [26]. Consider $S^2 \subset \mathbb{R}^3$ and flatten it on the equator plane, pushing the northern hemisphere down and the southern hemisphere up, getting two superimposed disks D^2 joined along their circular boundaries. The construction is drawn in Figure 3 (left). Similarly, consider $S^3 \subset \mathbb{C}^2 \simeq \mathbb{R}^4$: it can be viewed as two superimposed balls joined along their boundaries. In this case the boundaries are two spheres S^2. A picture of S^3 is given in Figure 3 (right).

3.1.3. The conjugate locus. Recall that all the partial derivatives of Exp evaluated in (θ, c, t) lie in $T_gSU(2) = g \cdot su(2)$ with $g = \mathrm{Exp}(\theta, c, t)$. One can easily check that the three vectors $g^{-1} \cdot \frac{\partial \mathrm{Exp}}{\partial \theta}_{|(\theta,c,t)}$, $g^{-1} \cdot \frac{\partial \mathrm{Exp}}{\partial c}_{|(\theta,c,t)}$, $g^{-1} \cdot \frac{\partial \mathrm{Exp}}{\partial t}_{|(\theta,c,t)} \in su(2)$ are linearly dependent (hence g is a conjugate point) if and only if

$$\sin\left(\sqrt{1+c^2}\frac{t}{2}\right)\left(2\sin\left(\sqrt{1+c^2}\frac{t}{2}\right) - \sqrt{1+c^2}t\cos\left(\sqrt{1+c^2}\frac{t}{2}\right)\right) = 0.$$

The first term is 0 if and only if $g \in e^{\mathbf{k}} = \left\{\begin{pmatrix} \alpha \\ 0 \end{pmatrix} \mid |\alpha| = 1\right\}$, while the second vanishes if and only if $\sqrt{1+c^2}\frac{t}{2} = \tan\left(\sqrt{1+c^2}\frac{t}{2}\right)$; hence we have two series of conjugate

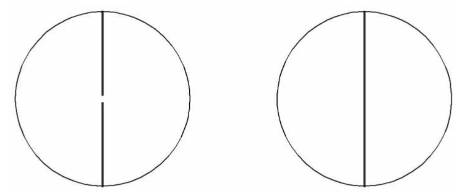

FIG. 4. $\mathbf{k} \oplus \mathbf{p}$ *problem on* $SU(2)$*: projection of the odd conjugate loci.*

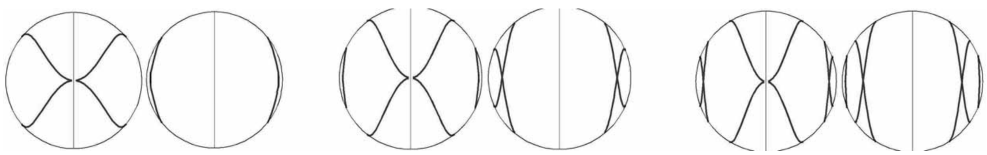

FIG. 5. $\mathbf{k} \oplus \mathbf{p}$ *problem on* $SU(2)$*: projection of the* 2*nd,* 4*th, and* 6*th conjugate loci.*

times as follows:

(i) first series: $t_{2n-1} = \frac{2n\pi}{\sqrt{1+c^2}}$, to which correspond the conjugate loci $C_{\text{Id}}^{2n-1} = e^{\mathbf{k}} \setminus \text{Id}$;

(ii) second series: $t_{2n} = \frac{2x_n}{\sqrt{1+c^2}}$, where $\{x_1, x_2, \ldots\}$ is the ordered set of the strictly positive solutions of $x = \tan(x)$, to which correspond the conjugate loci

$$C_{\text{Id}}^{2n} = \left\{ \left(\begin{array}{c} \frac{c\sin(x_n)}{\sqrt{1+c^2}} e^{i\left(\frac{\pi}{2} - y_n\right)} + \cos(x_n) e^{-iy_n} \\ \frac{\sin(x_n)}{\sqrt{1+c^2}} e^{i\theta} \end{array} \right) \middle| \begin{array}{c} c \in \mathbb{R}, \\ \theta \in \mathbb{R}/2\pi \end{array} \right\},$$

where $y_n = \frac{cx_n}{\sqrt{1+c^2}}$.

Remark 9. Notice that all the geodesics have a countable number of conjugate times.

We present some images of conjugate loci (Figures 4 and 5). For simplicity we present images of their sections with the plane $\text{Re}(\beta) = 0$. The complete images can be recovered using cylindrical symmetry.

Remark 10. Notice that the second conjugate locus is a two-dimensional submanifold of $SU(2)$, while the other even conjugate loci have self-intersections.

3.2. The $\mathbf{k} \oplus \mathbf{p}$ problem on $SO(3)$.

The Lie group $SO(3)$ is the group of special orthogonal 3×3 real matrices

$$SO(3) = \left\{ g \in \text{Mat}(3, \mathbb{R}) \mid gg^T = \text{Id}, \det(g) = 1 \right\}.$$

It is compact and its fundamental group is $\mathbb{Z}_2$. The Lie algebra of $SO(3)$ is the algebra of skew-symmetric 3×3 real matrices

$$so(3) = \left\{ \left(\begin{array}{ccc} 0 & -a & b \\ a & 0 & -c \\ -b & c & 0 \end{array} \right) \in \text{Mat}(3, \mathbb{R}) \right\}.$$

A basis of $so(3)$ is $\{p_1, p_2, k\}$, where

$$p_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad p_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

whose commutation relations are $[p_1, p_2] = k$, $[p_2, k] = p_1$, and $[k, p_1] = p_2$. Recall that $so(3)$ and $su(2)$ are isomorphic as Lie algebras, while $SU(2)$ is a double covering of $SO(3)$.

For $so(3)$ we have $\mathrm{Kil}(X, Y) = \mathrm{Tr}(XY)$ so, in particular, $\mathrm{Kil}(p_i, p_j) = -2\delta_{ij}$. The choice of the subspaces

$$\mathbf{k} = \mathrm{span}\,\{k\}\,, \qquad \mathbf{p} = \mathrm{span}\,\{p_1, p_2\}$$

gives a *Cartan decomposition* for $so(3)$. Moreover, $\{p_1, p_2\}$ is an orthonormal frame for the inner product $\langle\cdot,\cdot\rangle = -\frac{1}{2}\mathrm{Kil}(\cdot,\cdot)$ restricted to $\mathbf{p}$.

Defining $\Delta(g) = g\mathbf{p}$ and $\mathbf{g}_g(v_1, v_2) = \langle g^{-1}v_1, g^{-1}v_2\rangle$, we have that $(SO(3), \Delta, \mathbf{g})$ is a $\mathbf{k}\oplus\mathbf{p}$ sub-Riemannian manifold. As for $SU(2)$, all the $\mathbf{k}\oplus\mathbf{p}$ structures that one can define on $SO(3)$ are equivalent.

3.2.1. Expression of geodesics. Consider an initial covector $\lambda = \lambda(\theta, c) = \cos(\theta)p_1 + \sin(\theta)p_2 + ck \in \Lambda_{\mathrm{Id}}$. Using formula (11), we have that the exponential map is

$$\mathrm{Exp}(\theta, c, t) := \mathrm{Exp}(\lambda(\theta, c), t) = e^{(\cos(\theta)p_1 + \sin(\theta)p_2 + ck)t}e^{-ckt} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

with

$$\begin{aligned}
&a_{11} = K_1\cos(ct) + K_2\cos(2\theta + ct) + K_3 c\sin(ct),\\
&a_{12} = K_1\sin(ct) + K_2\sin(2\theta + ct) - K_3 c\cos(ct), \qquad a_{13} = K_4\cos(\theta) + K_3\sin(\theta),\\
&a_{21} = -K_1\sin(ct) + K_2\sin(2\theta + ct) + K_3 c\cos(ct),\\
&a_{22} = K_1\cos(ct) - K_2\cos(2\theta + ct) + K_3 c\sin(ct), \qquad a_{23} = -K_3\cos(\theta) + K_4\sin(\theta),\\
&a_{21} = K_4\cos(\theta + ct) - K_3\sin(\theta + ct), \qquad a_{22} = K_3\cos(\theta + ct) + K_4\sin(\theta + ct),\\
&a_{23} = \frac{\cos\left(\sqrt{1 + c^2}t\right) + c^2}{1 + c^2},\\
&K_1 = \frac{1 + \left(1 + 2c^2\right)\cos\left(\sqrt{1 + c^2}t\right)}{2\left(1 + c^2\right)}, \qquad K_2 = \frac{1 - \cos\left(\sqrt{1 + c^2}t\right)}{2\left(1 + c^2\right)},\\
&K_3 = \frac{\sin\left(\sqrt{1 + c^2}t\right)}{\sqrt{1 + c^2}}, \qquad K_4 = \frac{c\left(1 - \cos\left(\sqrt{1 + c^2}t\right)\right)}{1 + c^2}.
\end{aligned}$$

The set of geodesics has symmetry properties similar to the $SU(2)$ case. The conjugate locus can be obtained from that of the $SU(2)$ by the canonical projection $SU(2) \to SO(3)$. As for $SU(2)$, all the geodesics have a countable number of conjugate points.

3.3. The $\mathbf{k}\oplus\mathbf{p}$ problem on $\boldsymbol{SL(2)}$. The Lie group $SL(2)$ is the group of 2×2 real matrices with determinant 1,

$$SL(2) = \{g \in \mathrm{Mat}(2, \mathbb{R}) \mid \det(g) = 1\}\,.$$

It is a noncompact group and its fundamental group is $\mathbb{Z}$. The Lie algebra of $SL(2)$ is the algebra of traceless 2×2 real matrices

$$sl(2) = \left\{ \begin{pmatrix} a & b \\ c & -a \end{pmatrix} \in \mathrm{Mat}(2,\mathbb{R}) \right\}.$$

A basis of $sl(2)$ is $\{p_1, p_2, k\}$, where

$$p_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad p_2 = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad k = \frac{1}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

whose commutation relations are $[p_1,p_2] = -k$, $[p_2,k] = p_1$, and $[k,p_1] = p_2$. For $sl(n)$ we have $\mathrm{Kil}(X,Y) = 2n\mathrm{Tr}(XY)$ (see [20]); hence for $sl(2)$, $\mathrm{Kil}(X,Y) = 4\mathrm{Tr}(XY)$ and, in particular, $\mathrm{Kil}(p_i,p_j) = 2\delta_{ij}$. The choice of the subspaces

$$\mathbf{k} = \mathrm{span}\,\{k\}, \qquad \mathbf{p} = \mathrm{span}\,\{p_1,p_2\}$$

provides a *Cartan decomposition* for $sl(2)$. For $sl(2)$ the Cartan decomposition is unique, since $\mathbf{k}$ must be the maximal compact subalgebra. Moreover, $\{p_1,p_2\}$ is a orthonormal frame for the inner product $\langle\cdot,\cdot\rangle = \frac{1}{2}\mathrm{Kil}(\cdot,\cdot)$ restricted to $\mathbf{p}$.

Defining $\Delta(g) = g\mathbf{p}$ and $\mathbf{g}_g(v_1,v_2) = \langle g^{-1}v_1, g^{-1}v_2\rangle$, we have that $(SL(2),\Delta,\mathbf{g})$ is a $\mathbf{k}\oplus\mathbf{p}$ sub-Riemannian manifold.

3.3.1. Expression of geodesics. Consider an initial covector $\lambda = \lambda(\theta,c) = \cos(\theta)p_1 + \sin(\theta)p_2 + ck \in \Lambda_{\mathrm{Id}}$. Using formula (11), we have that the exponential map is

$$\mathrm{Exp}(\theta,c,t) := \mathrm{Exp}(\lambda(\theta,c),t) = e^{(\cos(\theta)p_1+\sin(\theta)p_2+ck)t}e^{-ckt} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

with

$$\begin{aligned}
a_{11} &= K_1\cos\left(c\frac{t}{2}\right) + K_2\left(\cos\left(\theta + c\frac{t}{2}\right) + c\sin\left(c\frac{t}{2}\right)\right),\\
a_{12} &= K_1\sin\left(c\frac{t}{2}\right) + K_2\left(\sin\left(\theta + c\frac{t}{2}\right) - c\cos\left(c\frac{t}{2}\right)\right),\\
a_{21} &= -K_1\sin\left(c\frac{t}{2}\right) + K_2\left(\sin\left(\theta + c\frac{t}{2}\right) + c\cos\left(c\frac{t}{2}\right)\right),\\
a_{22} &= K_1\cos\left(c\frac{t}{2}\right) + K_2\left(-\cos\left(\theta + c\frac{t}{2}\right) + c\sin\left(c\frac{t}{2}\right)\right),\\
K_1 &= \begin{cases} \mathrm{Cosh}\left(\sqrt{1-c^2}\,\frac{t}{2}\right), & c \in [-1,1],\\ \cos\left(\sqrt{c^2-1}\,\frac{t}{2}\right), & c \in (-\infty,-1)\cup(1,+\infty),\end{cases}\\
K_2 &= \begin{cases} \frac{\mathrm{Sinh}\left(\sqrt{1-c^2}\,\frac{t}{2}\right)}{\sqrt{1-c^2}}, & c \in (-1,1),\\ \frac{t}{2}, & c \in \{-1,1\},\\ \frac{\sin\left(\sqrt{c^2-1}\,\frac{t}{2}\right)}{\sqrt{c^2-1}}, & c \in (-\infty,-1)\cup(1,+\infty).\end{cases}
\end{aligned}$$

3.3.2. A useful decomposition of $SL(2)$.

PROPOSITION 7. *For every $g \in SL(2)$, there exists a unique pair $r \in e^{\mathbf{k}}$, $s \in e^{\mathbf{p}}$ such that $g = rs$.*

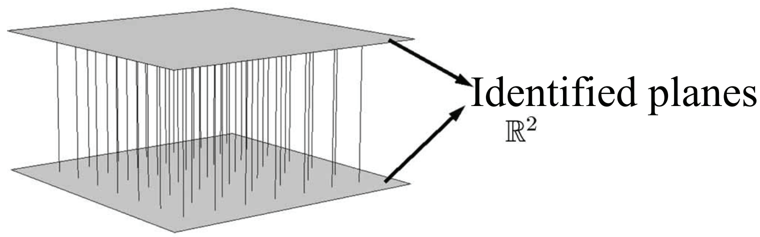

FIG. 6. *A picture of $SL(2)$.*

Proof. First, notice that $e^{\mathbf{k}} = SO(2)$ and that $e^{\mathbf{p}}$ is the set of 2×2 symmetric matrices with determinant 1 and positive trace.

Take

$$r = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \in e^{\mathbf{k}} \quad \text{and} \quad g = \begin{pmatrix} \alpha+\delta & \beta-\gamma \\ \beta+\gamma & \alpha-\delta \end{pmatrix} \in SL(2).$$

Notice that $(\alpha, \gamma) \neq (0,0)$. We have to prove that there exists a unique $\theta \in \mathbb{R}/2\pi$ such that $s = r^{-1}g$ is symmetric with positive trace. By direct computation, one gets that s is symmetric if and only if $\alpha \sin(\theta) = \gamma \cos(\theta)$. For any $(\alpha, \gamma) \in \mathbb{R}^2 \backslash (0,0)$, there exist two solutions of this equation $\theta_1, \theta_2 \in \mathbb{R}/2\pi$ with $\theta_2 = \theta_1 + \pi$. Thus

$$\mathrm{Tr}\left(\begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{pmatrix} g\right) = -\mathrm{Tr}\left(\begin{pmatrix} \cos(\theta_2) & \sin(\theta_2) \\ -\sin(\theta_2) & \cos(\theta_2) \end{pmatrix} g\right).$$

Observing that a symmetric matrix with determinant 1 has nonvanishing trace, either θ_1 or θ_2 provide $\mathrm{Tr}(s) > 0$. ◻

Topologically, $e^{\mathbf{k}} \simeq S^1$ and $e^{\mathbf{p}} \simeq \mathbb{R}^2$; hence $SL(2) \simeq S^1 \times \mathbb{R}^2$. In the following, we represent $SL(2)$ as the set $\mathbb{R}^2 \times [0,1]$ with the identification rule $(a,b,0) \sim (a,b,1)$. See Figure 6.

3.3.3. Symmetries in the $SL(2)$ problem. We have the following symmetry properties:

(i) *cylindrical symmetry*: $\mathrm{Exp}(\theta, c, t) = e^{z_0 k} e^{x p_1 + y p_2}$, where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

and (x_0, y_0, z_0) are defined by $\mathrm{Exp}(0, c, t) = e^{z_0 k} e^{x_0 p_1 + y_0 p_2}$;

(ii) *central symmetry*: $\mathrm{Exp}(\theta, -c, t) = e^{-z_0 k} e^{x p_1 + y p_2}$, where

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

and (x_0, y_0, z_0) are defined by $\mathrm{Exp}(\theta, c, t) = e^{z_0 k} e^{x_0 p_1 + y_0 p_2}$.

3.3.4. The conjugate locus. With arguments similar to those of section 3.1.3, one checks that $g = \mathrm{Exp}(\theta, c, t)$ is a conjugate point if and only if

$$\begin{cases} \mathrm{Sinh}\,(d)\left(2\mathrm{Sinh}\,(d) - t\sqrt{1-c^2}\,\mathrm{Cosh}\,(d)\right) = 0, & c \in (-1,1), \\ \pm \frac{t^4}{12} = 0, & c = \pm 1, \\ \sin\,(d)\left(2\sin\,(d) - t\sqrt{c^2-1}\cos\,(d)\right) = 0, & c \notin [-1,1] \end{cases}$$

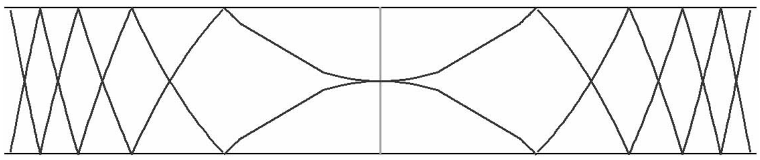

FIG. 7. $\mathbf{k}\oplus\mathbf{p}$ *problem on* $SU(2)$*: Section of the* 2*nd conjugate locus.*

with $d=\sqrt{1-c^2}\frac{t}{2}$ when $c\in(-1,1)$ and $d=\sqrt{c^2-1}\frac{t}{2}$ when $c\notin[-1,1]$.

The first two equations have only the trivial solution $t=0$. The third gives two series of conjugate times as follows:

(i) first series: $t_{2n-1}=\frac{2n\pi}{\sqrt{c^2-1}}$, to which correspond the conjugate loci $C_{\mathrm{Id}}^{2n-1}=e^{\mathbf{k}}\setminus\mathrm{Id}$;

(ii) second series: $t_{2n}=\frac{2x_n}{\sqrt{c^2-1}}$, where $\{x_1,x_2,\ldots\}$ is the ordered set of the strictly positive solutions of $x=\tan(x)$, to which correspond the conjugate loci

$$C_{\mathrm{Id}}^{2n}=\left\{\left(\begin{array}{cc} a_{11}^n(c,t) & a_{12}^n(c,t)\\ a_{21}^n(c,t) & a_{22}^n(c,t)\end{array}\right)\middle|\begin{array}{c} c\in\mathbb{R},\\ \theta\in\mathbb{R}/2\pi\end{array}\right\}$$

with

$$\begin{aligned}
a_{11}^n(c,t)&=\cos(x_n)\cos(y_n)+\frac{\sin(x_n)}{\sqrt{c^2-1}}\left(\cos(\theta)+c\sin(y_n)\right),\\
a_{12}^n(c,t)&=\cos(x_n)\sin(y_n)+\frac{\sin(x_n)}{\sqrt{c^2-1}}\left(\sin(\theta)-c\cos(y_n)\right),\\
a_{21}^n(c,t)&=-\cos(x_n)\sin(y_n)+\frac{\sin(x_n)}{\sqrt{c^2-1}}\left(\sin(\theta)+c\cos(y_n)\right),\\
a_{22}^n(c,t)&=\cos(x_n)\cos(y_n)+\frac{\sin(x_n)}{\sqrt{c^2-1}}\left(-\cos(\theta)+c\sin(y_n)\right),
\end{aligned}$$

where $y_n=\frac{cx_n}{\sqrt{c^2-1}}$.

Remark 11. Notice that not all geodesics have conjugate points. Indeed, $\mathrm{Exp}(\theta,c,\cdot)$ has a conjugate point if and only if $c\in(-\infty,-1)\cup(1,+\infty)$.

We present an image of the 2nd conjugate locus (Figure 7). For simplicity we present an image of its intersection with $\left\{e^{\mathbf{k}}e^{ap_1}|a\in\mathbb{R}\right\}$. The complete picture can be recovered using the cylindrical symmetry.

Remark 12. Notice that all even conjugate loci have self-intersection.

4. A sub-Riemannian structure on lens spaces.

4.1. Definition of $L(p,q)$.

Fix two coprime integers $p,q\in\mathbb{Z},\ p,q\neq0$. The **lens space** $L(p,q)$ is defined as the quotient of $SU(2)$ with respect to the identification rule

$$\left(\begin{array}{c}\alpha_1\\ \beta_1\end{array}\right)\sim\left(\begin{array}{c}\alpha_2\\ \beta_2\end{array}\right)\text{ if }\exists\ \omega\in\mathbb{C}\ p\text{th root of unity such that}$$

$$\left(\begin{array}{c}\alpha_2\\ \beta_2\end{array}\right)=\left(\begin{array}{cc}\omega & 0\\ 0 & \omega^q\end{array}\right)\left(\begin{array}{c}\alpha_1\\ \beta_1\end{array}\right).$$

Lens spaces are three-dimensional compact manifolds, but, except for $L(2,1) \simeq SO(3)$, they are neither Lie groups nor homogeneous spaces of $SU(2)$. The following topological equivalences hold: $\forall\ p,q,k \in \mathbb{Z}$, p,q coprime, $p,q \neq 0$, we have $L(p,q) \simeq L(p,-q) \simeq L(-p,q) \simeq L(p,q+kp)$. Lens spaces have highly nontrivial topology; for details we refer the reader to [25].

The following theorem permits us to choose a representative of $L(p,q)$ in $SU(2)$.

PROPOSITION 8. *Consider the set*

$$E_p = \left\{ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in SU(2) \middle|\ \mathrm{Re}\,(\alpha) > 0,\ \frac{\mathrm{Im}\,(\alpha)^2}{\sin\left(\frac{\pi}{p}\right)^2} + |\beta|^2 < 1 \right\} \subset SU(2)$$

and define $\partial E_p^+ = \partial E_p \cap \{\mathrm{Im}\,(\alpha) \geq 0\}$, $\partial E_p^- = \partial E_p \cap \{\mathrm{Im}\,(\alpha) \leq 0\}$. *Endow* $\overline{E_p}$ *with the equivalence relation* $\div$ *defined as follows:*

1. *The relation is reflexive;*
2. *moreover, given* $\begin{pmatrix} \alpha^+ \\ \beta^+ \end{pmatrix} \in \partial E_p^+$ *and* $\begin{pmatrix} \alpha^- \\ \beta^- \end{pmatrix} \in \partial E_p^-$, *we have* $\begin{pmatrix} \alpha^+ \\ \beta^+ \end{pmatrix} \div \begin{pmatrix} \alpha^- \\ \beta^- \end{pmatrix}$ *if*
   (i) *either* $\mathrm{Im}\,(\alpha^+) = -\mathrm{Im}\,(\alpha^-) \neq 0$ *and* $\beta^+ = e^{2\pi i \frac{q}{p}} \beta^-$; *or*
   (ii) $\mathrm{Im}\,(\alpha^+) = \mathrm{Im}\,(\alpha^-) = 0$ *and* $\beta^+ = e^{2\pi i \frac{n}{p}} \beta^-$ *for some* $n \in \{1, \dots, p\}$.

The manifold $\overline{E_p}/\div$ *is diffeomorphic to* $L(p,q)$.

Proof. Take $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in SU(2)$ and let us look for ω pth root of unity such that $\begin{pmatrix} \omega\alpha \\ \omega^q\beta \end{pmatrix} \in \overline{E_p}$. This condition is equivalent to

$$\mathrm{Re}\,(\omega\alpha) \geq 0 \text{ and } \frac{\mathrm{Im}\,(\omega\alpha)^2}{\sin\left(\frac{\pi}{p}\right)^2} + |\omega^q\beta|^2 \leq 1. \tag{13}$$

Recalling that $|\omega^q\beta|^2 = |\beta|^2 = 1 - |\alpha|^2$ and that $\mathrm{Im}\,(\omega\alpha) = |\alpha|\sin(\arg(\omega\alpha))$ if $\alpha \neq 0$, we see that (13) is equivalent to

$$\arg\,(\omega\alpha) \in \left[-\frac{\pi}{p}, \frac{\pi}{p}\right] \quad \text{or} \quad \alpha = 0. \tag{14}$$

Thus,

(i) if $\alpha \neq 0$, there exists at least one solution ω_1 of $\arg\,(\omega\alpha) \in \left[-\frac{\pi}{p}, \frac{\pi}{p}\right]$. Moreover, we have two distinct solutions ω_1, ω_2 if and only if $\arg\,(\omega_1\alpha) = -\frac{\pi}{p}$ and $\arg\,(\omega_2\alpha) = \frac{\pi}{p}$. In this case,

$$\begin{pmatrix} \omega_1\alpha \\ \omega_1^q\beta \end{pmatrix} = \begin{pmatrix} |\alpha| e^{-i\frac{\pi}{p}} \\ \omega_1^q\beta \end{pmatrix} \text{ and } \begin{pmatrix} \omega_2\alpha \\ \omega_2^q\beta \end{pmatrix} = \begin{pmatrix} |\alpha| e^{i\frac{\pi}{p}} \\ \omega_2^q\beta \end{pmatrix};$$

observe that

$$\begin{pmatrix} \omega_1\alpha \\ \omega_1^q\beta \end{pmatrix} \div \begin{pmatrix} \omega_2\alpha \\ \omega_2^q\beta \end{pmatrix}.$$

(ii) if $\alpha = 0$, every ω pth root of unity satisfies $\begin{pmatrix} 0 \\ \omega^q\beta \end{pmatrix} \in \overline{E_p}$; observe that for all the pairs ω_1, ω_2 we have

$$\begin{pmatrix} 0 \\ \omega_1^q\beta \end{pmatrix} \div \begin{pmatrix} 0 \\ \omega_2^q\beta \end{pmatrix}.$$

Hence $\forall \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in SU(2)$ we have a unique

$$\left[\begin{pmatrix} \omega\alpha \\ \omega^q\beta \end{pmatrix}\right]_{\div} \in \overline{E_p}/\div;$$

i.e., the function

$$\psi : \begin{array}{ccc} L(p,q) = SU(2)/_{\sim} & \to & \overline{E_p}/\div, \\ \left[\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right] & \mapsto & \left[\begin{pmatrix} \omega\alpha \\ \omega^q\beta \end{pmatrix}\right]_{\div} \end{array}$$

is bijective. $\square$

Remark 13. A crucial observation for what follows is that the projection

$$\Pi : \begin{array}{ccc} SU(2) & \to & L(p,q), \\ g & \mapsto & [g] \end{array}$$

is a local diffeomorphism. Moreover, $\Pi_{|_{E_p}} : E_p \to L(p,q)\setminus[\partial E_p]$ is a diffeomorphism. In particular, E_p contains only one representative for each equivalence classes of $L(p,q)$; i.e., if $g, h \in E_p$ and $[g] = [h]$, then $g = h$.

Remark 14. Proposition 8 provides a picture of $L(p,q)$; recall that $SU(2)$ is drawn as two balls in $\mathbb{R}^3$ (see section 3.1.2). Hence $\overline{E_p} \subset SU(2)$ is drawn as a closed ellipsoid inside one of the two balls, via the map

$$\rho : \begin{array}{ccc} \overline{E_p} & \to & \overline{B_1(0)} \subset \mathbb{R}^3, \\ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} & \mapsto & (\mathrm{Re}\,(\beta), \mathrm{Im}\,(\beta), \mathrm{Im}\,(\alpha)). \end{array}$$

The picture of E_p is

$$F_p = \left\{ (x_1, x_2, x_3) \in \overline{B_1(0)} \mid x_1^2 + x_2^2 + \frac{x_3^2}{\sin\left(\frac{\pi}{p}\right)^2} < 1 \right\},$$

and the one of $\overline{E_p}$ is $\overline{F_p}$; see Figure 8 (left). The identification $\div$ induces the following identification on $\overline{F_p}$: given $(x_1^+, x_2^+, x_3^+) \in \partial F_p^+ = \partial F_p \cap \{x_3 \geq 0\}$ and $(x_1^-, x_2^-, x_3^-) \in \partial F_p^- = \partial F_p \cap \{x_3 \leq 0\}$, they are identified when

$$x_3^+ = -x_3^- \text{ and } \begin{pmatrix} x_1^+ \\ x_2^+ \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1^- \\ x_2^- \end{pmatrix}$$

with $\theta = \frac{2\pi q}{p}$; see Figure 8 (right).

Remark 15. Observe that the identification rule on $\overline{F}_p$ gives a 1-to-1 identification between $\partial F_p \cap \{x_3 > 0\}$ and $\partial F_p \cap \{x_3 < 0\}$, while there are, in general, more identified points on $\left\{x_1^2 + x_2^2 = 1\right\} \cap \{x_3 = 0\}$; see Figure 9.

4.2. Sub-Riemannian quotient structure on $L(p,q)$.

PROPOSITION 9. *The sub-Riemannian structure on $SU(2)$ given in section* 3.1 *induces a* 2-dim *sub-Riemannian structure on $L(p,q) = SU(2)/_{\sim}$ via the quotient map*

$$\Pi : \begin{array}{ccc} SU(2) & \to & L(p,q), \\ x & \mapsto & [x]; \end{array}$$

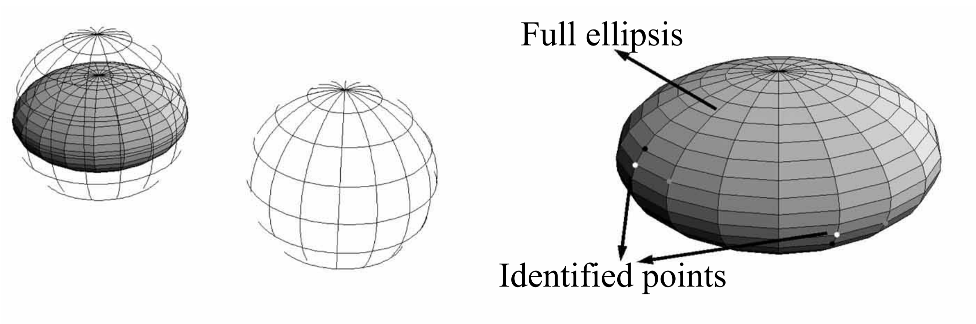

FIG. 8. *Left:* $\overline{F}_4$. *Right: the representation of* $L(4,1)$*, with some examples of the identification rule.*

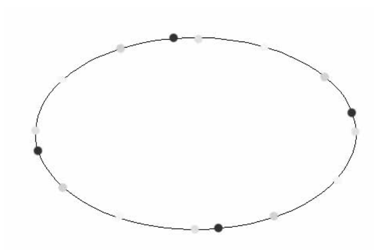

FIG. 9. $L(4,1)$*: some examples of the identification rule on* $\{x_1^2 + x_2^2 = 1\} \cap \{x_3 = 0\}$.

i.e.,

(i) *the map*

$$\widetilde{\Delta} : [g] \mapsto \Pi_* (\Delta(h)) \subset T_{[g]} L(p,q) \text{ with } h \in [g]$$

is a 2-dim *smooth distribution on* $L(p,q)$ *that is Lie bracket generating;*

(ii) $\tilde{\mathbf{g}}_{[g]}(v_*, w_*) = \langle v_*, w_* \rangle_{[g]} := \langle v, w \rangle_h$ *with* $h \in [g],\ v, w \in T_h SU(2),\ \Pi_*(v) = v_*,\ \Pi_*(w) = w_*$ *is a smooth positive definite scalar product on* $\widetilde{\Delta}$.

Proof. The role of the maps Π and $\Pi_{*_{|g}}$ is illustrated in Figure 10.

The map Π is a local diffeomorphism; thus $\Pi_{*_{|g}} : T_g SU(2) \to T_{[g]} L(p,q)$ is a linear isomorphism, and hence $\Pi_{*_{|g}} (\Delta(g))$ is a 2-dim subspace of $T_{[g]} L(p,q)$.

The following two statements are consequences of Lemma 10, presented below:

(i) the distribution $\widetilde{\Delta}([g])$ is well defined; i.e., $\forall\, h_1, h_2 \in [g]$ we have

$$\Pi_{*_{|h_1}} (\Delta(h_1)) = \Pi_{*_{|h_2}} (\Delta(h_2)) .$$

(ii) The positive definite scalar product $\langle v_*, w_* \rangle_{[g]}$ is well defined; i.e., $\forall\, h_1, h_2 \in [g],\ v_1, w_1 \in T_{h_1} SU(2),\ v_2, w_2 \in T_{h_2} SU(2)$ such that $\Pi_{*_{|h_1}}(v_1) = \Pi_{*_{|h_2}}(v_2)$ and $\Pi_{*_{|h_1}}(w_1) = \Pi_{*_{|h_2}}(w_2)$, we have $\langle v_1, w_1 \rangle_{h_1} = \langle v_2, w_2 \rangle_{h_2}$.

LEMMA 10. *Let* $h_1, h_2 \in [g]$ *with* $h_2 = \begin{pmatrix} \omega & 0 \\ 0 & \omega^q \end{pmatrix} h_1$. *The map*

$$\begin{array}{cccc} & \mathbf{p} & \to & \mathbf{p} \\ \phi : & \begin{pmatrix} n_1 \\ m_1 \\ 0 \end{pmatrix} & \mapsto & \begin{pmatrix} \mathrm{Re}\left(\omega^{q-1}\right) & -\mathrm{Im}\left(\omega^{q-1}\right) & 0 \\ \mathrm{Im}\left(\omega^{q-1}\right) & \mathrm{Re}\left(\omega^{q-1}\right) & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} n_1 \\ m_1 \\ 0 \end{pmatrix} \end{array}$$

$$\begin{tikzcd} T_g SU(2) \arrow[r, "\Pi_{*|g}"] \arrow[d] & T_{[g]} L(p,q) \arrow[d] \\ SU(2) \arrow[r, "\Pi"] & L(p,q). \end{tikzcd}$$

FIG. 10. *The role of the maps* Π *and* Π_*.

is bijective. Moreover, it is an isometry with respect to the positive definite scalar product $\langle\ ,\ \rangle$ *and satisfies,* $\forall\ \eta \in \mathbf{p}$,

$$\frac{d}{dt}_{|t=0}\left[h_1 e^{t\eta}\right] = \frac{d}{dt}_{|t=0}\left[h_2 e^{t\phi(\eta)}\right].$$

Proof. Let $h = \begin{pmatrix} a \\ b \end{pmatrix} \in SU(2)$ and $\eta = (n, m, 0) \in \mathbf{p}$. We have

$$he^{t\eta} = \begin{pmatrix} a\cos\left(\sqrt{n^2+m^2}\,\frac{t}{2}\right) - b\sin\left(\sqrt{n^2+m^2}\,\frac{t}{2}\right)\frac{n-im}{\sqrt{n^2+m^2}} \\ b\cos\left(\sqrt{n^2+m^2}\,\frac{t}{2}\right) + a\sin\left(\sqrt{n^2+m^2}\,\frac{t}{2}\right)\frac{n+im}{\sqrt{n^2+m^2}} \end{pmatrix}. \tag{15}$$

Take $h_1, h_2 \in [g]$ with $h_2 = \begin{pmatrix} \omega & 0 \\ 0 & \omega^q \end{pmatrix} h_1$ and $\eta_1, \eta_2 \in \mathbf{p}$ with coordinates $\eta_1 = (n_1, m_1, 0)$ and $\eta_2 = (n_2, m_2, 0)$. Consider the trajectories

$$\left[h_1 e^{t\eta_1}\right] = \left[\begin{pmatrix} a\cos\left(\sqrt{n_1^2+m_1^2}\,\frac{t}{2}\right) - b\sin\left(\sqrt{n_1^2+m_1^2}\,\frac{t}{2}\right)\frac{n_1-im_1}{\sqrt{n_1^2+m_1^2}} \\ b\cos\left(\sqrt{n_1^2+m_1^2}\,\frac{t}{2}\right) + a\sin\left(\sqrt{n_1^2+m_1^2}\,\frac{t}{2}\right)\frac{n_1+im_1}{\sqrt{n_1^2+m_1^2}} \end{pmatrix}\right]$$

and

$$\begin{aligned} [h_2 e^{t\eta_2}] &= \left[\begin{pmatrix} \omega a\cos\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right) - \omega^q b\sin\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right)\frac{n_2-im_2}{\sqrt{n_2^2+m_2^2}} \\ \omega^q b\cos\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right) + \omega a\sin\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right)\frac{n_2+im_2}{\sqrt{n_2^2+m_2^2}} \end{pmatrix}\right] \\ &= \left[\begin{pmatrix} a\cos\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right) - \omega^{q-1} b\sin\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right)\frac{n_2-im_2}{\sqrt{n_2^2+m_2^2}} \\ b\cos\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right) + \omega^{1-q} a\sin\left(\sqrt{n_2^2+m_2^2}\,\frac{t}{2}\right)\frac{n_2+im_2}{\sqrt{n_2^2+m_2^2}} \end{pmatrix}\right]. \end{aligned}$$

Thus $\frac{d}{dt}_{|t=0}[h_1 e^{t\eta_1}] = \frac{d}{dt}_{|t=0}[h_2 e^{t\eta_2}]$ in the case

$$\begin{cases} n_1^2 + m_1^2 = n_2^2 + m_2^2, \\ n_1 - im_1 = \omega^{q-1}(n_2 - im_2), \\ n_1 + im_1 = \omega^{1-q}(n_2 + im_2), \end{cases}$$

that is equivalent to $\omega^{q-1}(n_1 + im_1) = n_2 + im_2$. This equation is verified for $\eta_2 = \phi(\eta_1)$. $\square$

Since Π is a local diffeomorphism, $\forall\, g \in SU(2)\ \exists\, B(g)$ such that the map $\Pi_{*|_{B(g)}}$: $T_{B(g)}SU(2) \to T_{B([g])}L(p,q)$ is a diffeomorphism, and thus $\widetilde{\Delta}$ is smooth and Lie bracket generating, and $\langle v_*, w_* \rangle_{[g]}$ is smooth as a function of $[g]$. ∎

Proposition 9 implies that the sub-Riemannian structures on $SU(2)$ and $L(p,q)$ defined above are locally isometric via the map Π. As a consequence, the geodesics of $(L(p,q), \tilde{\Delta}, \tilde{\mathbf{g}})$ are the projection of geodesics of $(SU(2), \Delta, \mathbf{g})$. The conjugate locus for $L(p,q)$ can be obtained from that of $SU(2)$ by the projection Π.

Remark 16. One can check that the sub-Riemannian structure induced by $SU(2)$ on $L(2,1) \simeq SO(3)$ is equivalent to the $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian structure on $SO(3)$ defined in section 3.2.

5. Cut loci and distances.

In this section we prove the main theorems of the paper; i.e., we compute cut loci for $SU(2)$, $SO(3)$, lens spaces, and $SL(2)$, and we prove the formula (3) for the sub-Riemannian distance on $SU(2)$.

Recall that our problems satisfy the following assumptions:

(i) Each point of M is reached by an optimal geodesic starting from Id; see section 2.3;

(ii) we are in the three-dimensional contact case, and thus there are no abnormal minimizers. Hence Remark 4 applies.

PROPOSITION 11. *Let $T(\theta, c)$ be the cut time for* $\mathrm{Exp}(\theta, c, \cdot)$ *(possibly $+\infty$ if* $\mathrm{Exp}(\theta, c, \cdot)$ *is optimal on* $[0, +\infty)$*). Define*

$$\mathcal{D} = \left\{(\theta, c, t) \in \Lambda_{\mathrm{Id}} \times \mathbb{R}^+ \mid 0 < t < T(\theta, c)\right\}$$

and $M' = M \setminus (K_{\mathrm{Id}} \cup \mathrm{Id})$. The function $\mathrm{Exp}_{|_{\mathcal{D}}} : \mathcal{D} \to M'$ *is a diffeomorphism from $\mathcal{D}$ to M'.*

Proof. Let us first check that $\mathrm{Exp}(\mathcal{D}) \subset M'$. By contradiction, let $\mathrm{Exp}(\theta, c, t) \in M \setminus M'$; thus $t = 0$ or $t = T(\theta, c)$ or $\mathrm{Exp}(\theta, c, \cdot)$ is not optimal in $[0, t]$, i.e., $t > T(\theta, c)$. This is a contradiction. Let us verify that $\mathrm{Exp}_{|_{\mathcal{D}}}$ is injective; by contradiction, let $\mathrm{Exp}(\theta_1, c_1, t_1) = \mathrm{Exp}(\theta_2, c_2, t_2)$ with $(\theta_1, c_1, t_1) \neq (\theta_2, c_2, t_2)$. If $t_1 \neq t_2$, one of the two geodesics $\mathrm{Exp}(\theta_1, c_1, \cdot), \mathrm{Exp}(\theta_2, c_2, \cdot)$ has already lost optimality, and thus $t_i \geq T(\theta_i, c_i)$; hence $(\theta_i, c_i, t_i) \notin \mathcal{D}$, a contradiction. If $t_1 = t_2$, we have that $\mathrm{Exp}(\theta_1, c_1, t_1)$ is a cut point, and hence $t_1 \geq T(\theta_1, c_1)$, a contradiction. To verify that $\mathrm{Exp}_{|_{\mathcal{D}}}$ is surjective, take $g \in M'$ and observe that there is an optimal geodesic $\mathrm{Exp}(\theta, c, \cdot)$ reaching it at time $t \leq T(\theta, c)$. But $t = T(\theta, c)$ implies $g \in K_{\mathrm{Id}}$; thus $t < T(\theta, c)$.

The smoothness of $\mathrm{Exp}_{|_{\mathcal{D}}}$ and of its inverse follows from the facts that Exp is a local diffeomorphism outside the critical points (i.e., points where the differential of Exp is not of full rank) and that the critical points do not belong to $\mathcal{D}$. Indeed, by contradiction, let $(\theta, c, t) \in \mathcal{D}$ be a critical point, and hence t is a conjugate time as follows: it is either the first conjugate time that coincides with the cut time (i.e., $t = T(\theta, c)$) or a greater conjugate time (i.e., $t > T(\theta, c)$). In both cases $(\theta, c, t) \notin \mathcal{D}$, a contradiction. ∎

5.1. The cut locus for $SU(2)$.

THEOREM 12. *The cut locus for the* $\mathbf{k} \oplus \mathbf{p}$ *problem on* $SU(2)$ *is*

$$K_{\mathrm{Id}} = e^{\mathbf{k}} \setminus \mathrm{Id} = \left\{e^{ck} \mid c \in (0, 4\pi)\right\}.$$

Proof. Let

$$g \in e^{\mathbf{k}} \setminus \mathrm{Id} = \left\{\begin{pmatrix} \alpha \\ 0 \end{pmatrix} \mid \alpha \in \mathbb{C},\ |\alpha| = 1,\ \alpha \neq 1\right\},$$

and let $\mathrm{Exp}(\theta, c, \cdot)$ be the minimizing geodesic steering Id to g in time T. As a consequence of the cylindrical symmetry, we have that $\mathrm{Exp}(\psi, c, T) = g \ \ \forall\, \psi \in \mathbb{R}/2\pi$; thus $\left(e^{\mathbf{k}}\backslash \mathrm{Id}\right) \subset K_{\mathrm{Id}}$.

The core of the proof is to show that there are no cut points outside $e^{\mathbf{k}}$. Recall the expression of geodesics given in section 3.1.1. By contradiction, assume that $g \in SU(2)\backslash e^{\mathbf{k}}$ is reached by two different optimal trajectories $\mathrm{Exp}(\theta, c, \cdot)$ and $\mathrm{Exp}(\psi, d, \cdot)$ at time T. Observe that $\mathrm{Exp}(\theta, c, \frac{2\pi}{\sqrt{1+c^2}})$ and $\mathrm{Exp}(\psi, d, \frac{2\pi}{\sqrt{1+c^2}}) \in e^{\mathbf{k}} \subset K_{\mathrm{Id}}$; thus

$$0 < T < \min\left\{\frac{2\pi}{\sqrt{1+c^2}}, \frac{2\pi}{\sqrt{1+d^2}}\right\}. \tag{16}$$

Observe that $\mathrm{Exp}(\theta, c, T) = \mathrm{Exp}(\psi, d, T)$ implies that $|\beta|$ is equal in the two cases, i.e.,

$$\frac{\sin(\frac{\sqrt{1+c^2}T}{2})}{\sqrt{1+c^2}} = \frac{\sin(\frac{\sqrt{1+d^2}T}{2})}{\sqrt{1+c^2}}. \tag{17}$$

From this equation it follows that $|c| = |d|$. Indeed (17) is equivalent to

$$\frac{\sin(\frac{\sqrt{1+c^2}T}{2})}{\frac{\sqrt{1+c^2}T}{2}} = \frac{\sin(\frac{\sqrt{1+d^2}T}{2})}{\frac{\sqrt{1+d^2}T}{2}}.$$

From the facts that $\frac{\sqrt{1+c^2}T}{2}, \frac{\sqrt{1+d^2}T}{2} \in (0,\pi)$ and that the function $\frac{\sin p}{p}$ is injective for $p \in (0,\pi)$, it follows that $\frac{\sqrt{1+c^2}T}{2} = \frac{\sqrt{1+d^2}T}{2}$, and hence $|c| = |d|$.

Thus we consider the following two cases:

(i) $c = d \in \mathbb{R}$: the cylindrical symmetry implies that either $\theta = \psi$ (so the two geodesics coincide) or $g \in e^{\mathbf{k}}$. This is a contradiction.

(ii) $c = -d \in \mathbb{R}\backslash\{0\}$: with no loss of generality we assume $c > 0$. Since by the central and cylindrical symmetries, we have

$$\mathrm{Exp}(\psi, -c, t) = \begin{pmatrix} \overline{\alpha} \\ e^{i(\psi+\theta-\arg(\beta))}\beta \end{pmatrix}, \quad \text{where} \quad \mathrm{Exp}(\theta, c, t) = \begin{pmatrix} \alpha \\ \beta \end{pmatrix},$$

$\mathrm{Exp}(\theta, c, t) = \mathrm{Exp}(\psi, -c, t)$ implies $\mathrm{Im}\,(\alpha) = 0$. Hence

$$c\cos\left(\frac{ct}{2}\right)\sin\left(\frac{\sqrt{1+c^2}t}{2}\right) = \sqrt{1+c^2}\sin\left(\frac{ct}{2}\right)\cos\left(\frac{\sqrt{1+c^2}t}{2}\right). \tag{18}$$

The terms $c, \sin\left(\frac{\sqrt{1+c^2}t}{2}\right), \sqrt{1+c^2},\ \sin\left(\frac{ct}{2}\right)$ are nonzero because of (16) and $c < \sqrt{1+c^2}$. Thus $\cos\left(\frac{ct}{2}\right) = 0$ if and only if $\cos\left(\frac{\sqrt{1+c^2}t}{2}\right) = 0$, which is impossible because $0 < \frac{ct}{2} < \frac{\sqrt{1+c^2}t}{2} < \pi$. Hence we rewrite (18) as

$$\frac{\tan\left(\frac{\sqrt{1+c^2}t}{2}\right)}{\frac{\sqrt{1+c^2}t}{2}} = \frac{\tan\left(\frac{ct}{2}\right)}{\frac{ct}{2}}$$

and state that $(0, \tan(0)), \left(\frac{ct}{2}, \tan\left(\frac{ct}{2}\right)\right), \left(\frac{\sqrt{1+c^2}t}{2}, \tan\left(\frac{\sqrt{1+c^2}t}{2}\right)\right)$ are three distinct points aligned on the graph of the function tan in $[0,\pi)$, which is impossible. This is a contradiction. $\square$

The cut locus for the $\mathbf{k}\oplus\mathbf{p}$ sub-Riemannian manifold $SU(2)$ is given in Figure 11.

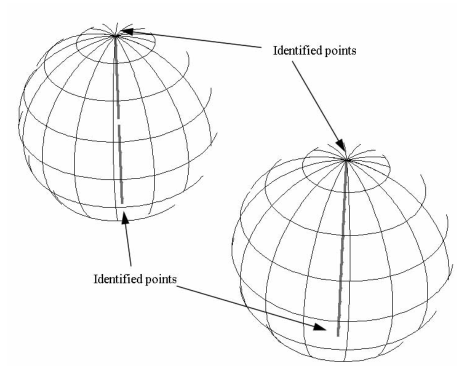

FIG. 11. *The cut locus for the* $\mathbf{k} \oplus \mathbf{p}$ *sub-Riemannian manifold* $SU(2)$.

5.1.1. The sub-Riemannian distance in $SU(2)$. In this section we compute the sub-Riemannian distance on $SU(2)$, i.e., we prove Theorem 2.

Let

$$g = \begin{pmatrix} \alpha \\ 0 \end{pmatrix} = \begin{pmatrix} e^{i\arg(\alpha)} \\ 0 \end{pmatrix} \in e^{\mathbf{k}}.$$

g is reached by a geodesic $\mathrm{Exp}(\theta, c, \cdot)$ at time $\frac{2\pi}{\sqrt{1+c^2}}$ for some $c \in \mathbb{R}$. Observe that

$$\mathrm{Exp}\left(\theta, c, \frac{2\pi}{\sqrt{1+c^2}}\right) = \mathrm{Exp}\left(\theta, \pm\sqrt{\frac{4\pi^2}{t^2} - 1}, t\right) \tag{19}$$

$$= \begin{pmatrix} -\cos\left(\sqrt{\pi^2 - \frac{t^2}{4}}\right) \mp i \sin\left(\sqrt{\pi^2 - \frac{t^2}{4}}\right) \\ 0 \end{pmatrix} = \begin{pmatrix} e^{i\left(\pi \pm \sqrt{\pi^2 - \frac{t^2}{4}}\right)} \\ 0 \end{pmatrix}.$$

Thus the distance $d(g, \mathrm{Id})$ is the smallest $t > 0$ such that $e^{i\left(\pi \pm \sqrt{\pi^2 - \frac{t^2}{4}}\right)} = e^{i\arg(\alpha)}$, whose solution is $t = 2\sqrt{\arg(\alpha)(2\pi - \arg(\alpha))}$, where $\arg(\alpha)$ is chosen in $[0, 2\pi]$.

Let $g = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \in SU(2)\backslash e^{\mathbf{k}}$. Applying Proposition 11, we have that $\mathrm{Exp}^{-1}_{|\mathcal{D}}(g)$ is well defined on $\mathcal{D} = \left\{(\theta, c, t) \in \Lambda_{\mathrm{Id}} \times \mathbb{R}^+ \mid 0 < t < \frac{2\pi}{\sqrt{1+c^2}}\right\}$. Thus the sub-Riemannian distance of g from the origin is $d(g, Id) = t$, where t is the third component of $\mathrm{Exp}^{-1}_{|\mathcal{D}}(g)$, i.e., the unique solution t of $\mathrm{Exp}(\theta, c, t) = g$ with $(\theta, c, t) \in \mathcal{D}$. Using the explicit form of Exp given in (3.1.1), one checks that the system

$$\begin{cases} \mathrm{Exp}(\theta, c, t) = g, \\ (\theta, c, t) \in \mathcal{D} \end{cases}$$

is equivalent to

$$\begin{cases} -\frac{ct}{2} + \arctan\left(\frac{c}{\sqrt{1+c^2}} \tan\left(\frac{\sqrt{1+c^2}\,t}{2}\right)\right) = \arg(\alpha), \\ \frac{\sin\left(\frac{\sqrt{1+c^2}\,t}{2}\right)}{\sqrt{1+c^2}} = \sqrt{1 - |\alpha|^2}, \\ \cos(\frac{ct}{2} + \theta) + i\sin(\frac{ct}{2} + \theta) = \arg(\beta). \end{cases}$$

The third equation has no role in the computation of distance as a consequence of the cylindrical symmetry.

Remark 17. The distance is a bounded function; this is due to its continuity and the compactness of $SU(2)$. The farthest point starting from Id is $-\mathrm{Id}$, whose distance is 2π.

Notice that $\forall\ \alpha, \beta_1, \beta_2 \in \mathbb{C}, |\beta_1| = |\beta_2|$, we have

$$d\left(\left(\begin{array}{c}\alpha\\ \beta_1\end{array}\right), \mathrm{Id}\right) = d\left(\left(\begin{array}{c}\alpha\\ \beta_2\end{array}\right), \mathrm{Id}\right) = d\left(\left(\begin{array}{c}\overline{\alpha}\\ \beta_1\end{array}\right), \mathrm{Id}\right).$$

This is due to the cylindrical and central symmetries.

5.2. The cut locus for $SO(3)$ and lens spaces. In this section we compute the cut locus for lens spaces $L(p,q)$. As a particular case, we get the cut locus for $SO(3) \simeq L(2,1)$.

THEOREM 13. *The cut locus for the sub-Riemannian problem on $L(p,q)$ defined in section* 9 *is a stratification*

$$K_{[\mathrm{Id}]} = K^{sym}_{[\mathrm{Id}]} \cup K^{loc}_{[\mathrm{Id}]}$$

with

$$K^{sym}_{[\mathrm{Id}]} = [\partial E_p] = \left\{ \left[\left(\begin{array}{c}\alpha\\ \beta\end{array}\right)\right] \mid a, b \in \mathbb{C}, \mathrm{Re}\,(\alpha) \geq 0,\ \frac{\mathrm{Im}\,(\alpha)^2}{\sin\left(\frac{\pi}{p}\right)^2} + |\beta|^2 = 1 \right\},$$

$$K^{loc}_{[\mathrm{Id}]} = \left[e^{\mathbf{k}}\right] \setminus [\mathrm{Id}] = \left\{ \left[\left(\begin{array}{c}\alpha\\ 0\end{array}\right)\right] \mid \alpha \in \mathbb{C}, |\alpha| = 1, \alpha^p \neq 1 \right\}.$$

Proof. Let us first prove the following lemma.

LEMMA 14. *A geodesic $\gamma(\cdot)$ in $L(p,q)$ steering* [Id] *to $[g]$ in minimum time T admits a unique lift $\gamma_0(\cdot)$ in $SU(2)$ starting from* Id.

Moreover, $\gamma_0(t) = \mathrm{Exp}(\theta_0, c_0, t)\ \ \forall\ t \in [0,T]$ *for some* $\theta_0 \in \mathbb{R}/2\pi,\ c \in \mathbb{R}$.

Proof. Take $\gamma(\cdot)$ as in the hypotheses. Since $L(p,q)$ and $SU(2)$ are locally diffeomorphic via Π, there is a unique lift $\gamma_0(\cdot)$ in $SU(2)$ starting from Id; i.e., $\gamma_0(0) = \mathrm{Id}$ and $[\gamma_0(t)] = \gamma(t)\ \ \forall\ t \in [0,T]$.

Let us prove that $\gamma_0(\cdot)$ is an optimal trajectory reaching $\gamma_0(T)$. By contradiction, there exists a trajectory $\gamma_1(\cdot)$ such that $\gamma_1(t_1) = \gamma_0(T)$ with $t_1 < T$. Hence, its projection $[\gamma_1(\cdot)]$ satisfies $[\gamma_1(t_1)] = [g]$ with $t_1 < T$, a contradiction.

Since $\gamma_0(\cdot)$ is an optimal trajectory, it is a geodesic of $SU(2)$, and there exist $\theta \in \mathbb{R}/2\pi,\ c \in \mathbb{R}$ such that $\gamma_0(t) = \mathrm{Exp}(\theta_0, c_0, t)\ \ \forall\ t \in [0,T]$. $\square$

Let us prove that $K^{loc}_{[\mathrm{Id}]} \subset K_{[\mathrm{Id}]}$. Consider $[g] \in K^{loc}_{[\mathrm{Id}]}$, a geodesic steering [Id] to $[g]$ in minimum time T with unique lift $\mathrm{Exp}(\theta_0, c_0, \cdot)$. By the definition of $K^{loc}_{[\mathrm{Id}]}$, we have $\mathrm{Exp}(\theta_0, c_0, T) \in e^{\mathbf{k}} \backslash \mathrm{Id} \subset SU(2)$; i.e., $\mathrm{Exp}(\theta_0, c_0, T)$ lies in the cut locus for the sub-Riemannian problem on $SU(2)$. Thus there exists another optimal geodesic $\mathrm{Exp}(\theta_1, c_1, \cdot)$ defined in $[0,T]$ such that $\mathrm{Exp}(\theta_1, c_1, T) = \mathrm{Exp}(\theta_0, c_0, T) \in [g]$. Thus the geodesic $[\mathrm{Exp}(\theta_1, c_1, \cdot)]$ reaches $[g]$ in minimum time. The geodesics in $SU(2)$ are distinct in a neighborhood of Id, so their projections in a neighborhood of [Id] are distinct as well.

Let us now prove that $K^{sym}_{[\mathrm{Id}]} \subset K_{[\mathrm{Id}]}$. Consider $[g] \in K^{sym}_{[\mathrm{Id}]}$, a geodesic steering [Id] to $[g]$ in minimum time T with unique lift $\mathrm{Exp}(\theta_0, c_0, \cdot)$; call $\mathrm{Exp}(\theta_0, c_0, T) = \left(\begin{smallmatrix} \alpha \\ \beta \end{smallmatrix}\right) \in [g]$. If $\beta = 0$, we have $[g] \in K^{loc}_{[\mathrm{Id}]}$ or $[g] = [\mathrm{Id}]$, so assume $\beta \neq 0$. Due to the cylindrical and central symmetries, we have

$$\mathrm{Exp}(\theta_0 + \psi, -c_0, T) = \begin{pmatrix} \overline{\alpha} \\ e^{2i(\theta_0 - \arg(\beta)) + i\psi}\beta \end{pmatrix}.$$

Consider $\psi^+ \in \mathbb{R}/2\pi$ being a solution of $e^{2i(\theta_0 - \arg(\beta)) + i\psi^+} = e^{2\pi i \frac{q}{p}}$ and $\psi^- \in \mathbb{R}/2\pi$ being a solution of $e^{2i(\theta_0 - \arg(\beta)) + i\psi^-} = e^{-2\pi i \frac{q}{p}}$. If $\mathrm{Exp}(\theta_0, c_0, T) \in \partial E_p^+$, we have $[\mathrm{Exp}(\theta_0 + \psi^+, -c_0, T)] = [\mathrm{Exp}(\theta_0, c_0, T)] = [g]$; if $\mathrm{Exp}(\theta_0, c_0, T) \in \partial E_p^-$, we have similarly $[\mathrm{Exp}(\theta_0 + \psi^-, -c_0, T)] = [\mathrm{Exp}(\theta_0, c_0, T)] = [g]$. If $c_0 \neq 0$, we have found two distinct trajectories reaching $[g]$ in optimal time; if $c_0 = 0$, we have $\mathrm{Exp}(\theta_0, c_0, T) \in \partial E_p^+ \cap \partial E_p^-$, and thus at least one of ψ^+ and ψ^- is not null, so at least one of $\mathrm{Exp}(\theta_0 + \psi^+, 0, \cdot)$ and $\mathrm{Exp}(\theta_0 + \psi^-, 0, \cdot)$ is distinct from $\mathrm{Exp}(\theta_0, 0, \cdot)$ in a neighborhood of Id, as are their projections in a neighborhood of [Id].

Finally, consider $[g] \in L(p,q) \setminus \left(K^{loc}_{[\mathrm{Id}]} \cup K^{sym}_{[\mathrm{Id}]} \cup [\mathrm{Id}]\right)$ and assume by contradiction that there exist two distinct geodesics steering [Id] to $[g]$ in minimum time T with distinct lifts $\mathrm{Exp}(\theta_0, c_0, \cdot)$, $\mathrm{Exp}(\theta_1, c_1, \cdot)$. There are two possibilities as follows:

(i) $\mathrm{Exp}(\theta_0, c_0, T) = \mathrm{Exp}(\theta_1, c_1, T)$. In this case, $\mathrm{Exp}(\theta_0, c_0, T)$ lies in the cut locus for the sub-Riemannian problem on $SU(2)$, and hence $[g] \in K^{loc}_{[\mathrm{Id}]}$, a contradiction.

(ii) $\mathrm{Exp}(\theta_0, c_0, T) \neq \mathrm{Exp}(\theta_1, c_1, T)$. Since by hypothesis $[g] \notin K^{sym}_{[\mathrm{Id}]}$, we have $\mathrm{Exp}(\theta_0, c_0, T), \mathrm{Exp}(\theta_1, c_1, T) \notin \partial E_p$. Recall that, if $[\mathrm{Exp}(\theta_0, c_0, T)] = [\mathrm{Exp}(\theta_0, c_0, T)]$ and $\mathrm{Exp}(\theta_0, c_0, T), \mathrm{Exp}(\theta_1, c_1, T) \in E_p$, then $\mathrm{Exp}(\theta_0, c_0, T) = \mathrm{Exp}(\theta_1, c_1, T)$, due to Remark 13. Thus we have that $\mathrm{Exp}(\theta_i, c_i, T) \in SU(2) \setminus \overline{E_p}$ for $i = 0$ or $i = 1$. We assume without loss of generality that $\mathrm{Exp}(\theta_0, c_0, T) \in SU(2) \setminus \overline{E_p}$; thus the geodesic $\mathrm{Exp}(\theta_0, c_0, t)$ with $t \in [0, T]$ steers $\mathrm{Id} \in E_p$ to $\mathrm{Exp}(\theta_0, c_0, T) \in SU(2) \setminus \overline{E_p}$, and hence $\exists\, \tilde{t} \in (0, T)$ such that $\mathrm{Exp}(\theta_0, c_0, \tilde{t}) \in \partial E_p$. Then we have that $\gamma_0(\tilde{t}) = [\mathrm{Exp}(\theta_0, c_0, \tilde{t})] \in K^{sym}_{[\mathrm{Id}]}$, and thus $\gamma_0(t)$ is no more optimal for $t \in [0, T]$, a contradiction. ◻

Remark 18. Notice that $K^{loc}_{[\mathrm{Id}]}$ is a manifold (a circle without a point), while $K^{sym}_{[\mathrm{Id}]}$ is not in general. Indeed, it is an orbifold. It can be seen as $S^2 \subset \mathbb{R}^3$ with the following identification: $(x_1^+, x_2^+, x_3^+) \in S^2 \cap \{x_3 \geq 0\}$ and $(x_1^-, x_2^-, x_3^-) \in S^2 \cap \{x_3 \leq 0\}$ are identified when $x_3^+ = -x_3^-$ and

$$\begin{pmatrix} x_1^+ \\ x_2^+ \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1^- \\ x_2^- \end{pmatrix}$$

with $\theta = \frac{2\pi q}{p}$. In the case $SO(3) \simeq L(2,1)$, we have that $K^{sym}_{[\mathrm{Id}]} = \mathbb{RP}^2$ (see Figure 12 (left)), while in the other cases it is not locally Euclidean; in fact, take a neighborhood of a point P on the equator and observe that it is topologically equivalent to a set of p half-planes with a common line as the boundary.

Next we give an idea of the topology of the cut locus for $L(4,1)$. Consider the space T_1 made by the two intersecting strips $\left\{(a,b,0) \in \mathbb{R}^3 \mid a,b \in [-1,1]\right\}$ and $\left\{(a,0,b) \in \mathbb{R}^3 \mid a,b \in [-1,1]\right\}$ with the following identification: $(-1,b,0) \sim (1,0,b)$ and $(-1,0,b) \sim (1,-b,0)$. The boundary of this set is topologically a circle S^1. Consider now a two-dimensional semisphere T_2. The cut locus $K^{sym}_{[\mathrm{Id}]}$ is topologically equivalent to the space given by gluing T_1 and T_2 along their boundaries S^1. The cut locus $K_{[\mathrm{Id}]}$ is given by gluing $K^{sym}_{[\mathrm{Id}]}$ with a circle S^1 along a point on T_2 and then removing

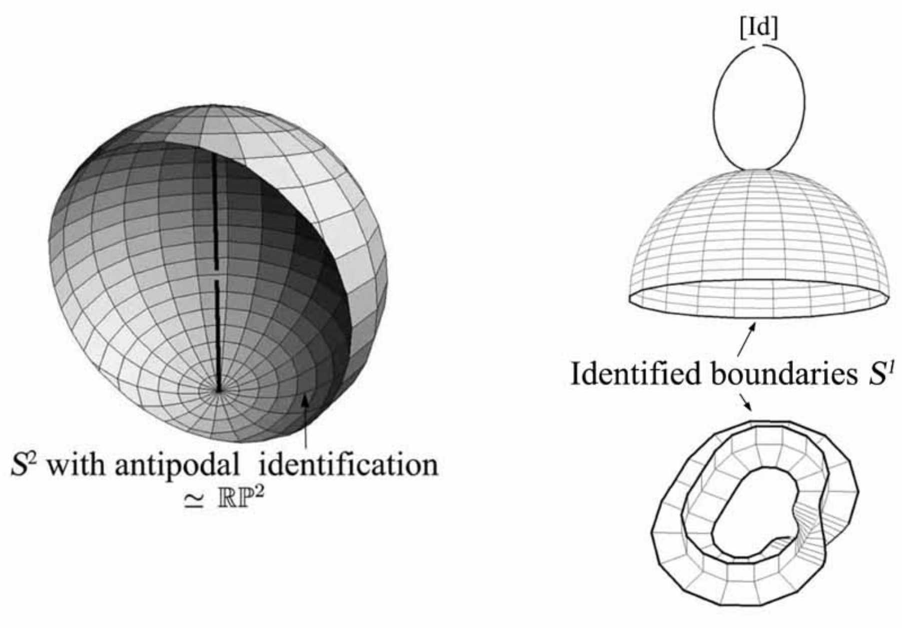

FIG. 12. *Left: The cut locus for the sub-Riemannian problem on* $SO(3)$. *Right: The cut locus for the sub-Riemannian problem on* $L(4, 1)$.

a point on S^1 (the starting point). See a picture of it in Figure 12 (right).

5.3. The cut locus for $SL(2)$.

THEOREM 15. *The cut locus for the* $\mathbf{k} \oplus \mathbf{p}$ *problem on* $SL(2)$ *is a stratification*

$$K_{\mathrm{Id}} = K_{\mathrm{Id}}^{sym} \cup K_{\mathrm{Id}}^{loc}$$

with

$$K_{\mathrm{Id}}^{sym} = e^{2\pi k} e^{\mathbf{p}} = \left\{ g \in SL(2) \mid g = g^T, \mathrm{Tr} g < 0 \right\},$$
$$K_{\mathrm{Id}}^{loc} = e^{\mathbf{k}} \setminus \mathrm{Id} = \left\{ \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \mid \alpha \in R/2\pi, \alpha \neq 0 \right\}.$$

Proof. Let us first prove that $K_{\mathrm{Id}}^{loc} \subset K_{\mathrm{Id}}$. Let $g \in e^{\mathbf{k}} \setminus \mathrm{Id}$; it is reached optimally by a geodesic $\mathrm{Exp}(\theta, c, \cdot)$ at time T. Due to the cylindrical symmetry, we have $g = \mathrm{Exp}(\psi, c, T)\ \forall\ \psi \in \mathbb{R}/2\pi$; thus $g \in K_{\mathrm{Id}}$.

Let us now prove that $K_{\mathrm{Id}}^{sym} \subset K_{\mathrm{Id}}$. Let $g = e^{2\pi k} e^{x_0 p_1 + y_0 p_2} \in e^{2\pi k} e^{\mathbf{p}}$; it is reached optimally by a geodesic $\mathrm{Exp}(\theta, c, \cdot)$ at time T. If $x_0^2 + y_0^2 = 0$, we have $g = e^{2\pi k} \in K_{\mathrm{Id}}^{loc}$; thus it is a cut point. If $x_0^2 + y_0^2 \neq 0$, due to the cylindrical and central symmetry, we have $\mathrm{Exp}(\theta + \psi, -c, T) = e^{-2\pi k} e^{x p_1 + y p_2}$ with

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(2\theta + \psi) & \sin(2\theta + \psi) \\ \sin(2\theta + \psi) & -\cos(2\theta + \psi) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}.$$

Choose ψ in such a way that $\theta+\frac{\psi}{2}$ is the angle on the plane of the line passing through $(0,0)$ and (x_0,y_0). In this way we have $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. Observing that $e^{-2\pi k}=e^{2\pi k}$, we finally have that $g=\mathrm{Exp}(\theta,c,T)=\mathrm{Exp}(\theta+\psi,-c,T)$. Observe that $c\neq 0$ because $\mathrm{Exp}(\theta,0,\cdot)\in e^{\mathbf{p}}$; thus the two geodesics $\mathrm{Exp}(\theta,c,\cdot)$, $\mathrm{Exp}(\theta+\psi,-c,\cdot)$ are distinct.

We now prove that there is no cut point outside $K^{sym}_{\mathrm{Id}}\cup K^{loc}_{\mathrm{Id}}$. By contradiction, let $g\in SL(2)\setminus\left(K^{sym}_{\mathrm{Id}}\cup K^{loc}_{\mathrm{Id}}\cup \mathrm{Id}\right)$ be reached by two optimal trajectories $\mathrm{Exp}(\theta,c,\cdot)$ and $\mathrm{Exp}(\psi,d,\cdot)$ at time T. Writing

$$\mathrm{Exp}(\theta,c,t)=\begin{pmatrix} g_{11}(\theta,c,t) & g_{12}(\theta,c,t) \\ g_{21}(\theta,c,t) & g_{22}(\theta,c,t)\end{pmatrix},$$

we have

$$\begin{aligned} r(c,t) &:= \sqrt{(g_{11}-g_{22})^2+(g_{12}+g_{21})^2} \\ &= \begin{cases} \frac{\mathrm{Sinh}\left(\sqrt{1-c^2}\,\frac{t}{2}\right)}{\frac{\sqrt{1-c^2}}{2}}, & c\in(-1,1)\,, \\ t, & c\in\{-1,1\}\,, \\ \frac{\sin\left(\sqrt{c^2-1}\,\frac{t}{2}\right)}{\frac{\sqrt{c^2-1}}{2}}, & c\in(-\infty,-1)\cup(1,+\infty). \end{cases} \end{aligned} \tag{20}$$

The identity $\mathrm{Exp}(\theta,c,T)=\mathrm{Exp}(\psi,d,T)$ implies $r(c,T)=r(d,T)$, which implies $c^2=d^2$. Indeed, observe that in the three cases described in (20) we have, respectively, $r(c,t)>t$, $r(c,t)=t$, $r(c,t)<t$; thus $c,d\in(-1,1)$ or $c,d\in\{-1,1\}$ or $c,d\in(-\infty,-1)\cup(1,+\infty)$. In each of the three cases, the identity $r(c,T)=r(d,T)$ implies $c^2=d^2$. Indeed we have the following three cases.

Case $c,d\in(-1,1)$. In this case the conclusion follows from the fact that $\frac{\mathrm{Sinh}(p)}{p}$ is injective for $p\in(0,+\infty)$.

Case $c,d\in\{-1,1\}$. This case is straightforward.

Case $c,d\in(-\infty,-1)\cup(1,+\infty)$. Let us prove first that $\frac{\sqrt{c^2-1}T}{2}\in(0,\pi)$. By contradiction, assume $\frac{\sqrt{c^2-1}T}{2}\geq\pi$. There exists $t\in(0,T]$ such that $\frac{\sqrt{c^2-1}t}{2}=0$; hence $r(c,t)=0$, from which it follows that $\mathrm{Exp}(\theta,c,t)\in e^{\mathbf{k}}$.

Hence either $t<T$ (and $\mathrm{Exp}(\theta,c,\cdot)$ is not optimal on $[0,T]$, a contradiction) or $t=T$ (and $g\in K^{loc}_{\mathrm{Id}}\cup\mathrm{Id}$, a contradiction). Similarly we prove that $\frac{\sqrt{d^2-1}T}{2}\in(0,\pi)$.

Now observe that $r(c,T)=r(d,T)$ implies

$$\frac{\sin\left(\sqrt{c^2-1}\,\frac{T}{2}\right)}{\frac{\sqrt{c^2-1}T}{2}}=\frac{\sin\left(\sqrt{d^2-1}\,\frac{T}{2}\right)}{\frac{\sqrt{d^2-1}T}{2}}.$$

Recalling that $\frac{\sin p}{p}$ is injective for $p\in(0,\pi)$, we have $\frac{\sqrt{c^2-1}T}{2}=\frac{\sqrt{d^2-1}T}{2}$; hence $c^2=d^2$.

We have the following two cases:

(i) $c=d\in\mathbb{R}$. The identity $g_{11}(\theta,c,T)=g_{11}(\psi,c,T)$ implies either $\theta=\psi$ (i.e., the geodesics coincide) or $c\in(-\infty,-1)\cup(1,+\infty)$, and thus $\sin\left(\sqrt{c^2-1}\,\frac{T}{2}\right)=0$, i.e., $\exp(\theta,c,T)\in e^{\mathbf{k}}$; hence either $g=\mathrm{Id}$ or $g\in K^{loc}_{\mathrm{Id}}$, a contradiction.

(ii) $c=-d\in\mathbb{R}\setminus\{0\}$. Writing $g=e^{zk}e^{x_0p_1+y_0p_2}$, we have $\mathrm{Exp}(\psi,-c,T)=e^{-zk}e^{xp_1+yp_2}$. The identity $\mathrm{Exp}(\theta,c,T)=\mathrm{Exp}(\psi,-c,T)$ and the uniqueness of the

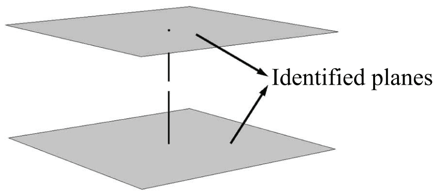

FIG. 13. *The cut locus for the* $\mathbf{k} \oplus \mathbf{p}$ *sub-Riemannian manifold* $SL(2)$.

decomposition from section 3.3.2 imply $e^{zk} = \pm \mathrm{Id}$. Thus g is symmetric, i.e., $g_{12}(\theta, c, T) = g_{21}(\theta, c, T)$.

If $c \in (-1, 1)$ this equation implies

$$\frac{\tan\left(c\frac{T}{2}\right)}{c\frac{T}{2}} = \frac{\mathrm{Tanh}\left(\sqrt{1-c^2}\frac{T}{2}\right)}{\sqrt{1-c^2}\frac{T}{2}}.$$

Choosing $c > 0$, observe that the first positive solution T_1 of the equation

$$\frac{\tan\left(c\frac{T_1}{2}\right)}{c\frac{T_1}{2}} = \frac{\mathrm{Tanh}\left(\sqrt{1-c^2}\frac{T_1}{2}\right)}{\sqrt{1-c^2}\frac{T_1}{2}}$$

satisfies $T_1 \in \left(\pi, 3\frac{\pi}{2}\right)$. The other cases, $c \in \{-1, 1\}$ and $c \in (-\infty, -1) \cup (1, +\infty)$, are treated similarly and lead to $T_1 \in \left(\pi, 3\frac{\pi}{2}\right)$.

Thus $\cos\left(c\frac{T_1}{2}\right) < 0$, and hence $\mathrm{Tr}\,(g) < 0$. But $\mathrm{Exp}(\theta, c, T_1)$ symmetric and $\mathrm{Tr}\,(g) < 0$ implies $\mathrm{Exp}(\theta, c, T_1) \in K_{\mathrm{Id}}^{sym}$; i.e., T_1 is a cut time. Thus either $T = T_1$ (meaning that $\mathrm{Exp}(\theta, c, t) \in K_{\mathrm{Id}}^{sym}$) or $T > T_1$ and $\mathrm{Exp}(\theta, c, \cdot)$ is not optimal in $[0, T]$, a contradiction. $\square$

We give a picture of the cut locus for the $\mathbf{k} \oplus \mathbf{p}$ sub-Riemannian manifold $SL(2)$ in Figure 13.

Acknowledgments. We are grateful to A. Agrachev for many illuminating discussions. We deeply thank G. Charlot for bringing to our attention some crucial properties of the cut locus, and we thank L. Paoluzzi for many explanations on lens spaces.

REFERENCES

[1] A. Agrachev, *Methods of control theory in nonholonomic geometry*, in Proceedings of the ICM-94, Birkhäuser, Zürich, 1996, pp. 1473–1483.

[2] A. Agrachev, *Compactness for sub-Riemannian length-minimizers and subanalyticity*, Rend. Sem. Mat. Univ. Politec. Torino, 56 (2001), pp. 1–12.

[3] A. A. Agrachev and Y. L. Sachkov, *Control Theory from the Geometric Viewpoint*, Encyclopaedia Math. Sci. 87, Springer-Verlag, Berlin, 2004.

[4] A. Agrachev, *Exponential mappings for contact sub-Riemannian structures*, J. Dynam. Control Systems, 2 (1996), pp. 321–358.

[5] R. Beals, B. Gaveau, and P. C. Greiner, *Hamilton–Jacobi theory and the heat kernel on Heisenberg groups*, J. Math. Pures Appl., 79 (2000), pp. 633–689.

[6] A. Bellaïche, *The tangent space in sub-Riemannian geometry*, in Sub-Riemannian Geometry, A. Bellaïche and J.-J. Risler, eds., Progr. Math. 144, Birkhäuser, Basel, 1996, pp. 1–78.

[7] A. Boggess, *CR Manifolds and the Tangential Cauchy–Riemann Complex*, Stud. Adv. Math., CRC Press, Boca Raton, FL, 1991.

[8] B. Bonnard and M. Chyba, *Singular Trajectories and Their Role in Control Theory*, Math. Appl. 40, Springer-Verlag, Berlin, 2003.

[9] U. Boscain, T. Chambrion, and G. Charlot, *Nonisotropic 3-level quantum systems: Complete solutions for minimum time and minimum energy*, Discrete Contin. Dyn. Syst. Ser. B, 5 (2005), pp. 957–990.
[10] U. Boscain, T. Chambrion, and J.-P. Gauthier, *On the $K + P$ problem for a three-level quantum system: Optimality implies resonance*, J. Dynam. Control Systems, 8 (2002), pp. 547–572.
[11] U. Boscain and B. Piccoli, *Optimal Synthesis for Control Systems on 2-D Manifolds*, Math. Appl. 43, Springer-Verlag, Berlin, 2004.
[12] U. Boscain and S. Polidoro, *Gaussian estimates for hypoelliptic operators via optimal control*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Mem. Mat. Appl., 18 (2007), pp. 333–342.
[13] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS Ser. Appl. Math., AIMS, Springfield, MO, 2007.
[14] R. W. Brockett, *Explicitly solvable control problems with nonholonomic constraints*, in Proceedings of the 38th Annual IEEE Conference on Decision and Control, Vol. 1, IEEE, Piscataway, NJ, 1999, pp. 13–16.
[15] Y. Chitour, F. Jean, and E. Trélat, *Genericity results for singular curves*, J. Differential Geom., 73 (2006), pp. 45–73.
[16] Y. Eliashberg, *Contact 3-manifolds twenty years since J. Martinet's work*, Ann. Inst. Fourier (Grenoble), 42 (1992), pp. 165–192.
[17] B. Gaveau, *Principe de moindre action, propagation de la chaleur et estimées sous elliptiques sur certains groupes nilpotents*, Acta Math., 139 (1977), pp. 95–153.
[18] M. Gromov, *Carnot–Caratheodory spaces seen from within*, in Sub-Riemannian Geometry, A. Bellaïche and J.-J. Risler, eds., Progr. Math. 144, Birkhäuser, Basel, 1996, pp. 79–323.
[19] L. Hörmander, *The analysis of linear partial differential operators*, Grundlehren Math. Wiss., Springer-Verlag, Berlin, 1983.
[20] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Pure Appl. Math., Academic Press, New York-London, 1978.
[21] V. Jurdjevic, *Geometric Control Theory*, Cambridge Stud. Adv. Math. 52, Cambridge University Press, Cambridge, UK, 1997.
[22] V. Jurdjevic, *Optimal Control, Geometry, and Mechanics*, in Mathematical Control Theory, J. Baillieul, and J. C. Willems, eds., Springer, New York, 1999, pp. 227–267.
[23] V. Jurdjevic, *Hamiltonian point of view on non-Euclidean geometry and elliptic functions*, Systems Control Lett., 43 (2001), pp. 25–41.
[24] L. S. Pontryagin, V. Boltianski, R. Gamkrelidze, and E. Mitchtchenko, *The Mathematical Theory of Optimal Processes*, Wiley Interscience, New York, London, 1962.
[25] D. Rolfsen, *Knots and Links*, Publish or Perish, Houston, 1990.
[26] J. R. Weeks, *The Shape of Space*, Monogr. Textbooks Pure Appl. Math., Marcel Dekker, New York, 1985.