

PROJECTIVE REEDS-SHEPP CAR ON S^2 WITH QUADRATIC COST

UGO BOSCAIN^{1, 2} AND FRANCESCO ROSSI²

Abstract. Fix two points $x, \bar{x} \in S^2$ and two directions (without orientation) $\eta, \bar{\eta}$ of the velocities in these points. In this paper we are interested to the problem of minimizing the cost

$$J[\gamma] = \int_0^T (\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) + K_{\gamma(t)}^2 \mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))) dt$$

along all smooth curves starting from x with direction η and ending in \bar{x} with direction $\bar{\eta}$. Here \mathbf{g} is the standard Riemannian metric on S^2 and K_γ is the corresponding geodesic curvature. The interest of this problem comes from mechanics and geometry of vision. It can be formulated as a sub-Riemannian problem on the lens space $L(4, 1)$. We compute the global solution for this problem: an interesting feature is that some optimal geodesics present cusps. The cut locus is a stratification with non trivial topology.

Mathematics Subject Classification. 49J15, 53C17.

Received May 30, 2008.

Published online December 19, 2008.

1. INTRODUCTION

Fix two points x and \bar{x} on a 2-D Riemannian manifold (M, \mathbf{g}) and two directions $\eta \in PT_x M$ and $\bar{\eta} \in PT_{\bar{x}} M$. Here by $PT_x M$ we mean the projective tangent space at the point x , *i.e.* the tangent space $T_x M \setminus \{0\}$ with the identification $v_1 \sim v_2$ if there exists $\alpha \in \mathbb{R} \setminus \{0\}$ such that $v_1 = \alpha v_2$. Given a vector $v \in T_x M \setminus \{0\}$ we call $\text{Dir}(v)$ its direction in $PT_x M$. We are interested in finding the path $\gamma : [0, T] \rightarrow M$ minimizing a compromise among length and geodesic curvature of a curve on M with fixed initial and final points x, \bar{x} , and fixed initial and final directions of the velocity $\eta, \bar{\eta}$. More precisely we consider the minimization problem:

$$J[\gamma] = \int_0^T (\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) + \beta^2 K_{\gamma(t)}^2 \mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))) dt \rightarrow \min, \quad (1.1)$$

along all smooth curves satisfying the boundary conditions $\gamma(0) = x$, $\text{Dir}(\dot{\gamma}(0)) = \eta$, $\gamma(T) = \bar{x}$, $\text{Dir}(\dot{\gamma}(T)) = \bar{\eta}$, as in Figure 1. Here $K_{\gamma(t)}$ is the geodesic curvature of the curve $\gamma(t)$, β is a fixed nonvanishing constant and the final time T is fixed.

Keywords and phrases. Carnot-Caratheodory distance, geometry of vision, lens spaces, global cut locus.

¹ LE2i, CNRS UMR5158, Université de Bourgogne, 9 avenue Alain Savary - BP 47870, 21078 Dijon Cedex, France.

² SISSA, via Beirut 2-4, 34014 Trieste, Italy. boscain@sissa.it; rossifr@sissa.it

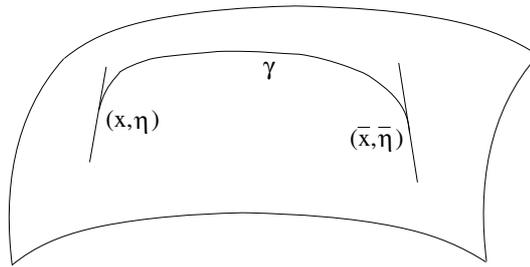


FIGURE 1. A curve γ satisfying the boundary conditions.

Notice that requiring γ smooth (indeed, $\gamma \in C^2$ is enough) is necessary for the geodesic curvature being defined in all points. However, we will formulate the problem in a wider class to be able to apply standard existence results and the Pontryagin Maximum Principle (PMP in the following, see for instance [4,21]), that is a first order necessary condition for optimality.

The problem stated above is extremely difficult in general. Indeed, one can see (see Sect. 2) that on the projective tangent bundle PTM the minimization problem (1.1) gives rise to a contact 3-D sub-Riemannian problem for which global solutions are known only for few examples. The typical difficulties one meets in such problems are the following:

1. One should apply first order necessary conditions for optimality (PMP) and solve an Hamiltonian system that generically is not integrable, to find candidate optimal trajectories (called *geodesics*).
2. Even if all solutions to the PMP are found, one has to evaluate their optimality.

In this paper we present the solution of this problem on the sphere S^2 , that we call the *projective Reeds-Shepp car with quadratic cost*. In this case the Hamiltonian system given by the PMP is Liouville integrable and one is faced to the problem of evaluating optimality of the geodesics. This second problem is solved with the computation of the cut locus K_x , that is the set of points reached by more than one optimal geodesic starting from x . For our specific problem, indeed the cut locus coincides with the set of points in which geodesics lose optimality.

The interest of this problems comes from mechanics, and more recently from problems of geometry of vision [14,20,22]. Indeed, in the spirit of the model of visual architecture due to Petitot, Citti-Sarti and Agrachev, the minimization problem stated above in the case $M = \mathbb{R}^2$ is the optimal control problem solved by the visual cortex $V1$ to reconstruct a contour that is partially hidden or corrupted in a planar image. The case $M = S^2$ can be seen as a modified version of this model, taking into account the curvature of the *retina* an/or the movements of the eyes.

1.1. Examples of problems of type (1.1)

1.1.1. The planar case

We present the case $M = \mathbb{R}^2$ endowed with the standard Riemannian metric: it is being studied in [19]. In this case the problem (1.1) is equivalent to the following optimal control problem

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{pmatrix} = u_1 \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \\ 0 \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \int_0^T (u_1^2 + \beta^2 u_2^2) dt \rightarrow \min,$$

$$(x(0), y(0), \theta(0)) = (x_0, y_0, \theta_0), \quad (x(T), y(T), \theta(T)) = (x_1, y_1, \theta_1).$$

The dynamics coincide with the Reeds-Shepp car (see [22]), except for the fact that here $\theta \in \mathbb{R}/\pi$. For this

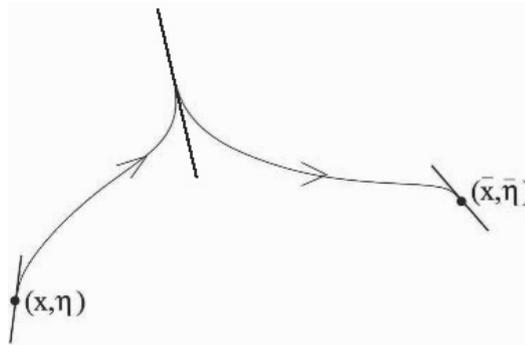


FIGURE 2. A geodesic with a cusp.

reason we call the problem (1.1) the **projective Reeds-Shepp car** on a Riemannian manifold, with quadratic cost.

1.1.2. *The case $M = S^2$*

The case $M = S^2$, endowed with the standard Riemannian metric, can be seen as the compactified version of the problem on the plane. It is the optimal control problem

$$\begin{pmatrix} \dot{a} \\ \dot{b} \\ \dot{\xi} \end{pmatrix} = u_1 \begin{pmatrix} \cos(\xi) \\ \frac{\sin(\xi)}{\sin(a)} \\ -\cot(a)\sin(\xi) \end{pmatrix} + u_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \int_0^T (u_1^2 + \beta^2 u_2^2) dt \rightarrow \min,$$

$$(a(0), b(0), \xi(0)) = (a_0, b_0, \xi_0), \quad (a(T), b(T), \xi(T)) = (a_1, b_1, \xi_1),$$

where (a, b) are standard spherical coordinates on the sphere (see (2.3)) and $\xi = \text{Dir}(\dot{\gamma})$. Also in this case the Hamiltonian system given by the Pontryagin Maximum Principle is Liouville integrable and one is faced to the problem of evaluating optimality of the geodesics.

This problem has special features. Indeed, PTS^2 is the 3-D manifold called *lens space* $L(4, 1)$, that can be seen as a suitable quotient of $SU(2)$ by a discrete group. Moreover, when $\beta = 1$ this problem is the projection of a left-invariant sub-Riemannian problem on $SU(2)$, called $\mathfrak{k} \oplus \mathfrak{p}$ problem, defined explicitly in Section 3.1. For $SU(2)$ an explicit expression for geodesics is given and we computed the cut locus and the Carnot-Caratheodory distance in [8] (results are recalled in Sect. 3.1). As a consequence, one can easily find geodesics for the problem on PTS^2 as projections of the geodesics for $SU(2)$. An interesting feature is that the projections of these geodesics on S^2 present cusps in some cases, see Figure 2. Observe that in cusp points the tangent direction is well defined.

$L(4, 1)$ can be described topologically as follows: consider a full 3-D ball $B := \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid \sum_{i=1}^3 x_i \leq 1\}$ and define the equivalence relation \approx on it as follows: the points $(x_1^+, x_2^+, x_3^+) \in \partial B^+ = \partial B \cap \{x_3 \geq 0\}$ and $(x_1^-, x_2^-, x_3^-) \in \partial B^- = \partial B \cap \{x_3 \leq 0\}$ are identified when

$$x_3^+ = -x_3^- \text{ and } \begin{pmatrix} x_1^+ \\ x_2^+ \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1^- \\ x_2^- \end{pmatrix},$$

as in Figure 3. The manifold B/\approx is topologically equivalent to $L(4, 1)$.

Observe that there exist different identification rules on $SU(2)$ such that the quotient manifold is topologically $L(4, 1)$ and the induced sub-Riemannian structure is well defined: indeed, in [8] Section 4, we already endowed $L(4, 1)$ with a sub-Riemannian structure that is different from the one we present in this paper. The structure

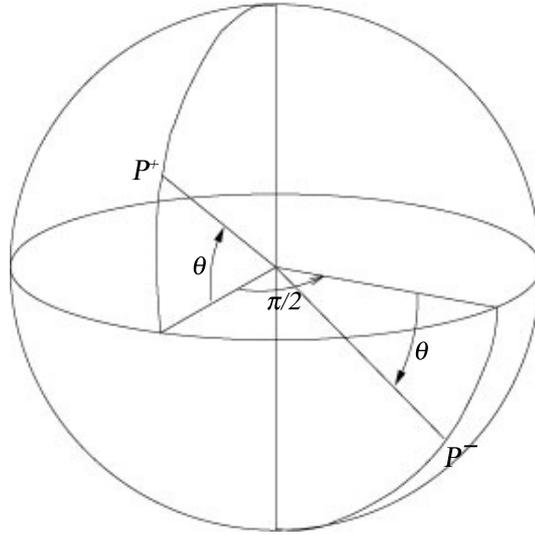


FIGURE 3. Identification rule on ∂B .

chosen in this paper (defined in Sect. 3) is the unique for which the minimal length problem on $L(4, 1)$ coincide with the problem (1.1). As a consequence, also the cut loci are different in the two cases. See Remark 3.2 for further details.

The main result of this paper is the computation of the cut locus $K_{[Id]}$, presented in Theorem 4.1. It is a stratification and it has nontrivial topology. Indeed, it is the union of two sets $K_{[Id]} = K^{loc} \cup K^{sym}$, described as follows: K^{loc} is topologically equivalent to S^1 without the starting point $[Id]$. K^{sym} is given by the union of two 2-D manifolds glued along their boundaries

$$K^{sym} = K^a \cup K^b$$

where

- The manifold K^a is contained in the set $\left\{ [g] \mid g = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix} \text{ with } \text{Re}(\beta) = 0 \right\}$ and it is homeomorphic to a 2-D disc with only two points identified on the boundary. Observe that the boundary is thus homeomorphic to two circles S^1 glued in a point (we call them \mathcal{S}^1 and s^1). A part of K^{loc} (namely, the farthest part from the starting point) is contained in K^a .
- The manifold K^b is topologically equivalent to a 2-D disc whose boundary S^1 is glued to the boundary of K^a as presented in Figure 4.

A representation of the topology of the cut locus is given in Figure 4.

We give a picture of the cut locus in Figure 5. Here:

- The 1-D stratum K^{loc} is represented by two semi-diameters (one is the left-right line on the northern hemisphere, the other is the front-rear line on the southern hemisphere) without the poles (representing $[Id]$): due to identification rule \approx , the two lines are identified.
- The 2-D stratum K^a is represented by four “triangles” on the boundary of the sphere. The two dark gray triangles (on the front) are identified. Similarly, the two light gray triangles (on the rear) are identified.
- The 2-D stratum K^b , not illustrated in the figure, can be seen as a surface inside the sphere whose boundary is the closed line given by the concatenation of segments $\mathcal{S}\mathcal{R}\mathcal{S}\mathcal{R}srsr$.

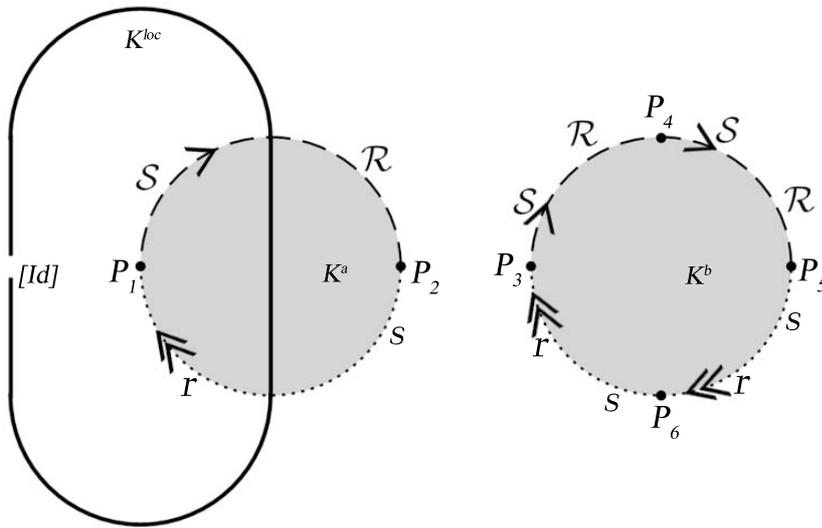


FIGURE 4. A scheme of the topology of the cut locus: points P_i with $i = 1, \dots, 6$ are identified, boundary of disks are identified according to arrows.

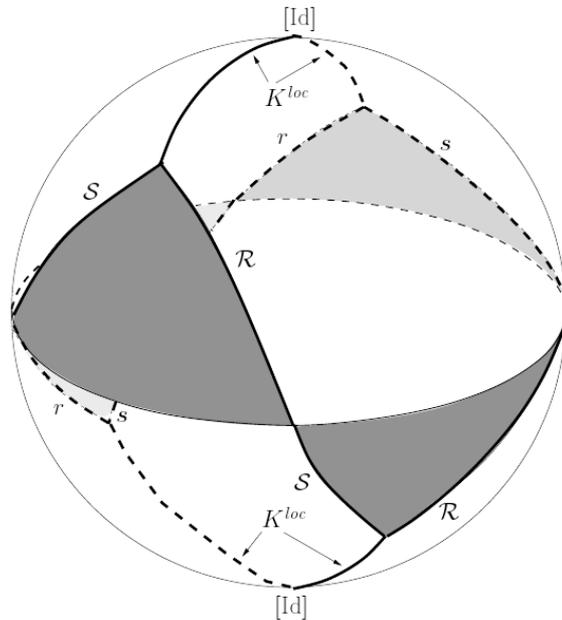


FIGURE 5. Picture of the cut locus: strata K^{loc} and K^a .

The paper is organized as follows: in Section 2 we define the mechanical problem on an orientable Riemannian manifold M and write the optimal control problem in the case $M = S^2$. In Section 3 we define $L(4, 1)$ and a sub-Riemannian structure on it such that the minimal length problem on $L(4, 1)$ coincide with the mechanical problem stated above. Finally, in Section 4 we give the solution of the problem.

2. THE PROBLEM OF MINIMIZING LENGTH AND CURVATURE ON A 2-D RIEMANNIAN MANIFOLD

In this section when we speak of *directions* we always mean directions without orientations. Let M be a 2-D manifold smooth manifold and PTM its *bundle of directions*, i.e.,

$$PTM := \bigcup_{x \in M} \{PT_x M\}, \text{ where}$$

$$PT_x M := \{(T_x M \setminus \{0\}) / \sim \text{ where } v_1 \sim v_2 \text{ if and only if } \exists \alpha \in \mathbb{R} \setminus \{0\} \mid v_1 = \alpha v_2\}.$$

Assume now that M is orientable and Riemannian, with Riemannian metric \mathbf{g} and let (x_1, x_2) be a *local orthogonal chart* on an open set $U \subset M$, i.e. a chart such that,

$$\mathbf{g}_{(x_1, x_2)}(\partial_{x_1}, \partial_{x_2}) \equiv 0 \text{ on } U.$$

In the following the symbol $\|\cdot\|$ indicates the norm with respect to $\mathbf{g}_{(x_1, x_2)}$.

In this case there is a canonical way to construct a chart (x_1, x_2, ξ) , with $\xi \in \mathbb{R}/\pi$ on PTU , as follows: write

$$v = v_1 \frac{\partial_{x_1}}{\|\partial_{x_1}\|} + v_2 \frac{\partial_{x_2}}{\|\partial_{x_2}\|} \in T_x U, \quad x \in U,$$

and define ξ as the angle between v and $\partial_{x_1}/\|\partial_{x_1}\|$, i.e. $\xi := \arctan\left(\frac{v_2}{v_1}\right)$.

In the following we call the chart (x_1, x_2, ξ) a *lifted-orthogonal chart*. Notice that $\|\partial_{x_1}\| \equiv 1$ and $\|\partial_{x_2}\| \equiv 1$ on U , implies \mathbf{g} flat on U .

Given a smooth curve $\gamma(\cdot) : [0, T] \rightarrow M$, its *lift* in PTM is the curve $(\gamma(\cdot), \xi(\cdot))$, where $\xi(t) \in PT_{\gamma(t)}M$ is the direction of $\dot{\gamma}(t)$. Writing

$$\dot{\gamma}(t) = \dot{\gamma}_1(t)\partial_{x_1} + \dot{\gamma}_2(t)\partial_{x_2} = (\dot{\gamma}_1(t)\|\partial_{x_1}\|)\partial_{x_1}/\|\partial_{x_1}\| + (\dot{\gamma}_2(t)\|\partial_{x_2}\|)\partial_{x_2}/\|\partial_{x_2}\|,$$

then

$$\xi(t) = \arctan\left(\frac{\dot{\gamma}_2(t)\|\partial_{x_2}\|}{\dot{\gamma}_1(t)\|\partial_{x_1}\|}\right). \quad (2.1)$$

Given a smooth curve $\Gamma(\cdot) : [0, T] \rightarrow PTM$, we say that it is *horizontal* if it is the lift of a curve on M . If $\Gamma(\cdot) = (\gamma_1(\cdot), \gamma_2(\cdot), \xi(\cdot))$ are its components in a lifted orthogonal chart, then $\Gamma(\cdot)$ is horizontal if and only if condition (2.1) holds.

The requirement that a smooth curve $\Gamma(\cdot)$ in PTM is horizontal, is equivalent to require that its velocity belongs to a 2-D distribution (see Appendix A). Let us describe this fact in detail: in a lifted orthogonal chart we have $\dot{\Gamma}(t) = \dot{\gamma}_1(t)\partial_{x_1} + \dot{\gamma}_2(t)\partial_{x_2} + \dot{\xi}(t)\partial_\xi$ and condition (2.1) is equivalent to

$$\dot{\Gamma}(t) \in \Delta(\Gamma(t)), \text{ where } \Delta(x_1, x_2, \xi) := \text{span}\{F_1, F_2\}, F_1 = \begin{pmatrix} \frac{\cos(\xi)}{\|\partial_{x_1}\|} \\ \frac{\sin(\xi)}{\|\partial_{x_2}\|} \\ \Xi(x_1, x_2, \xi) \end{pmatrix}, F_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The function $\Xi(x_1, x_2, \xi)$ is an arbitrary smooth function. We choose Ξ such that the vector field F_1 on PTM is the *geodesic spray*, i.e. the infinitesimal generator of the geodesic flow (see [12,25]).

Notice that $\text{span}\{F_1(x_1, x_2, \xi), F_2(x_1, x_2, \xi), [F_1, F_2](x_1, x_2, \xi)\} = T_{(x_1, x_2, \xi)}(PTM)$, hence (PTM, Δ) is a 3-D contact distribution.

2.1. A sub-Riemannian problem on PTM

There is a natural sub-Riemannian problem associated with (PTM, Δ) . It comes from the problem of minimizing a compromise among length and geodesic curvature of a curve on M with fixed initial and final points, and fixed initial and final directions of the velocities. More precisely we consider the following problem:

(P): Fix $x, \bar{x} \in M$, $\eta \in PT_x M$ and $\bar{\eta} \in PT_{\bar{x}} M$. Minimize

$$\mathcal{J}[\Gamma] = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) + \beta^2 K_{\gamma(t)}^2 \mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))} dt \tag{2.2}$$

along all horizontal smooth curves $\Gamma(\cdot) = (\gamma(\cdot), \xi(\cdot)) : [0, T] \rightarrow PTM$ satisfying the boundary conditions $\Gamma(0) = (x, \eta)$, $\Gamma(T) = (\bar{x}, \bar{\eta})$, *i.e.* $\gamma(0) = x$, $\text{Dir}(\dot{\gamma}(0)) = \eta$, $\gamma(T) = \bar{x}$, $\text{Dir}(\dot{\gamma}(T)) = \bar{\eta}$. Here $K_{\gamma(t)}$ is the geodesic curvature of the curve $\gamma(t)$, and β is a fixed nonvanishing constant.

Problem **(P)** is illustrated in Figure 1. Its equivalence with the problem (1.1) (in which the square root is absent) is explained below. The cost (2.2) is the most natural cost associated with (PTM, Δ) . Indeed it is invariant by change of coordinates and by reparameterization of the curve γ . Moreover, as it will be clear from the example below, it can be interpreted as a length in PTM .

The term $\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))$ corresponds to the Riemannian length on M , while the term $\beta^2 K_{\gamma(t)}^2 \mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))$ corresponds to the geodesic curvature. The presence of $\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))$ in this second term guarantees the invariance by reparameterization. The constant β fixes the relative weight. It may happen that there is a natural choice for β or even that β is not relevant (see an example on \mathbb{R}^2 below), depending on the specific problem.

One can check that $\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t)) + \beta^2 K_{\gamma(t)}^2 \mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))$ defines a positive definite quadratic form on Δ , *i.e.* it endows (PTM, Δ) with a sub-Riemannian structure (see Appendix A). Indeed, if $x = (x_1, x_2, \xi)$ is a lifted orthogonal chart and Ξ is such that F_1 is the geodesic spray, we have that problem **(P)** becomes the optimal control problem:

$$\dot{x} = u_1 F_1 + u_2 F_2, \quad F_1 = \begin{pmatrix} \frac{\cos(\xi)}{\|\partial_{x_1}\|} \\ \frac{\sin(\xi)}{\|\partial_{x_2}\|} \\ \Xi(x_1, x_2, \xi) \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathcal{J}[\Gamma] = \int_0^T \sqrt{u_1^2 + \beta^2 u_2^2} dt \rightarrow \min$$

that is the minimal length problem in the sub-Riemannian manifold PTM .

Consider now the problem of minimization of the *energy functional* $E[\Gamma] := \int_0^T u_1^2 + \beta^2 u_2^2 dt$ with fixed final time T and fixed starting and ending points. If Γ is a minimizer of E , then it is a minimizer of \mathcal{J} and its *lifted velocity* $v(t) := \sqrt{u_1^2 + \beta^2 u_2^2}$ is constant. On the other side, a minimizer Γ of \mathcal{J} parameterized with constant lifted velocity v is also a minimizer for E with $T = \mathcal{J}[\Gamma]/v$. For details, see Appendix A.

Finding a complete solution for the problem **(P)**, *i.e.* finding the optimal trajectory connecting each initial and final condition, is a problem of optimal synthesis in dimension 3 that is extremely difficult in general. The case $M = \mathbb{R}^2$ with the standard euclidean structure is being studied in [19]. The complete solution in the case $M = S^2$ with the standard Riemannian metric is given as the solution of the sub-Riemannian problem on $L(4, 1)$ presented in Section 3.

2.1.1. The planar case

Consider $M = \mathbb{R}^2$ with the standard euclidean structure. In this case $PT\mathbb{R}^2 \simeq \mathbb{R}^2 \times S^1$ and a lifted orthogonal chart is (x_1, x_2, ξ) , where (x_1, x_2) are euclidean coordinates on \mathbb{R}^2 . In this case, since $\|\partial_{x_1}\| = \|\partial_{x_2}\| = 1$, writing

$\gamma(t) = (x_1(t), x_2(t))$, we have

$$\mathbf{g}(\dot{\gamma}, \dot{\gamma}) = \dot{x}_1^2 + \dot{x}_2^2, \quad \text{and} \quad K_\gamma = \frac{-\dot{x}_2 \ddot{x}_1 + \dot{x}_1 \ddot{x}_2}{(\dot{x}_1^2 + \dot{x}_2^2)^{\frac{3}{2}}} = \frac{\dot{\xi}}{\sqrt{\dot{x}_1^2 + \dot{x}_2^2}}, \quad \text{where } \xi(t) = \arctan\left(\frac{\dot{x}_2(t)}{\dot{x}_1(t)}\right).$$

Hence $\mathcal{J}[\Gamma] = \int_0^T \sqrt{\dot{x}_1^2 + \dot{x}_2^2 + \beta^2 \dot{\xi}^2} dt$. We choose $\Xi(x_1, x_2, \xi) \equiv 0$, then F_1 is the geodesic spray and we have

$$\Delta(x_1, x_2, \xi) := \text{span}\{F_1, F_2\}, \quad \text{where } F_1 = \begin{pmatrix} \cos(\xi) \\ \sin(\xi) \\ 0 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathcal{J}[\Gamma] = \int_0^T \sqrt{u_1^2 + \beta^2 u_2^2} dt.$$

The associated optimal control problem is

$$\begin{cases} \dot{x}_1 = u_1 \cos(\xi) \\ \dot{x}_2 = u_1 \sin(\xi) \\ \dot{\xi} = u_2, \end{cases} \quad u_1, u_2 \in \mathbb{R}, \quad \int_0^T \sqrt{u_1^2 + \beta^2 u_2^2} dt \rightarrow \min.$$

This problem can be seen as a left-invariant sub-Riemannian problem on the group $SE(2)/\sim$, the group of rototranslations of the plane

$$SE(2) := \left\{ \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & x_1 \\ \sin(\alpha) & \cos(\alpha) & x_2 \\ 0 & 0 & 1 \end{pmatrix} \mid \alpha \in \mathbb{R}/2\pi, x_i \in \mathbb{R} \right\},$$

endowed with the identification rule $(\alpha, x_1, x_2) \sim (\alpha + \pi, x_1, x_2)$. In this case the constant β can be set to 1, since it becomes irrelevant after the transformation $(x_1, x_2) \rightarrow (\beta x_1, \beta x_2)$. This problem is being studied in [19].

2.2. The case $M = S^2$

Let $S^2 = \{(y_1, y_2, y_3) \in \mathbb{R}^3 \mid y_1^2 + y_2^2 + y_3^2 = 1\}$ and let (a, b) be the spherical coordinates (see Fig. 6):

$$\begin{cases} y_1 = \sin(a) \cos(b) \\ y_2 = \sin(a) \sin(b) \\ y_3 = \cos(a), \end{cases} \quad a \in [0, \pi], \quad b \in [0, 2\pi). \tag{2.3}$$

In these coordinates the standard Riemannian metric on S^2 (induced by its embedding in \mathbb{R}^3) has the form

$$\mathbf{g}(a, b) = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2(a) \end{pmatrix}.$$

The geodesic curvature of a curve $\gamma(t) = (a(t), b(t))$ is:

$$K_\gamma = \frac{\sqrt{\sin(a)^2} \left(2 \cot(a) \dot{a}^2 \dot{b} + \cos(a) \sin(a) \dot{b}^3 - \dot{b} \ddot{a} + \dot{a} \ddot{b} \right)}{\left(\dot{a}^2 + \sin(a)^2 \dot{b}^2 \right)^{\frac{3}{2}}} = \frac{\cos(a) \dot{b}}{\sqrt{\dot{a}^2 + \sin(a)^2 \dot{b}^2}} + \frac{\dot{\xi}}{\sqrt{\dot{a}^2 + \sin(a)^2 \dot{b}^2}}$$

with $\xi = \arctan\left(\frac{\dot{b} \sin(a)}{\dot{a}}\right)$.

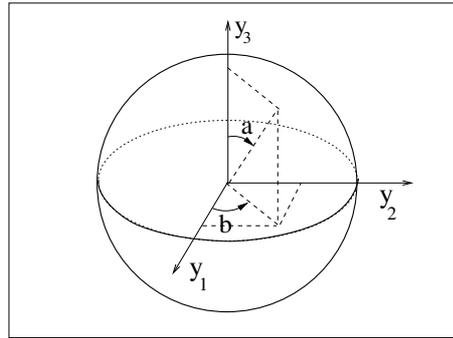


FIGURE 6. Spherical coordinates on S^2 .

Hence $\mathcal{J}[\Gamma] = \int_0^T \sqrt{\dot{a}^2 + \sin^2(a)\dot{b}^2 + \beta^2(\cos(a)\dot{b} + \dot{\xi})^2}$. In this case F_1 is the geodesic spray choosing $\Xi(a, b, \xi) = -\cot(a)\sin(\xi)$ for which we have

$$\Delta(a, b, \xi) := \text{span}\{F_1, F_2\}, \text{ where } F_1 = \begin{pmatrix} \cos(\xi) \\ \frac{\sin(\xi)}{\sin(a)} \\ \sin(a) \\ -\cot(a)\sin(\xi) \end{pmatrix}, \quad F_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathcal{J}[\Gamma] = \int_0^T \sqrt{u_1^2 + \beta^2 u_2^2} dt.$$

The associated optimal control problem is

$$\begin{cases} \dot{a} = u_1 \cos(\xi) \\ \dot{b} = u_1 \frac{\sin(\xi)}{\sin(a)} \\ \dot{\xi} = -u_1 \cot(a) \sin(\xi) + u_2, \end{cases} \quad u_1, u_2 \in \mathbb{R}, \quad \int_0^T \sqrt{u_1^2 + \beta^2 u_2^2} dt \rightarrow \min. \quad (2.4)$$

Remark 2.1. Observe that this optimal control problem admits trajectories for which $\dot{a} = \dot{b} = 0$. These trajectories in PTS^2 represent the possibility of changing the direction of the tangent vector on the sphere S^2 on a fixed point; mechanically, it represents the possibility of rotation on itself.

2.2.1. Existence of minimizers

To discuss the problem of existence of minimizers for the problem on the sphere S^2 , let us come back to the original problem of minimizing the cost (1.1), where $M = S^2$ with the standard Riemannian metric, along all smooth curves $\gamma : [0, T] \rightarrow S^2$. Here the class of smooth curves has been chosen to give a meaning to the geodesic curvature in the whole interval $[0, T]$ (indeed $C^2([0, T])$ was enough).

Unfortunately the class of smooth (or C^2) curves is too small to apply standard existence theorems for minimizers and first order necessary conditions for optimality (the PMP). We have to deal with a larger class of function.

This class arises naturally after the lift in PTS^2 , where the minimization problem takes the form (2.4). Here it is natural to look for minimizers $(a(\cdot), b(\cdot), \xi(\cdot))$ that are absolutely continuous and corresponding to controls $u_1(\cdot), u_2(\cdot)$ belonging to $L^2([0, T])$. In this class, standard existence theorems can be applied to guarantee existence of minimizers (see for instance [24], Thm. 5.1).

However the PMP works in the smaller class of L^∞ controls (corresponding to Lipschitz curves $(a(\cdot), b(\cdot), \xi(\cdot))$); hence, before applying PMP, one needs to prove that minimizers belong to this smaller class.

Thanks to a theorem of Sarychev and Torres (see [24], Thm. 5.2), one can prove that for the problem (2.4) all minimizers satisfy the conditions given by the PMP and that normal ones correspond to controls belonging

to $L^\infty([0, T])$. In our case there are no abnormal extremals (see definition in Appendix A.2) because the problem (2.4) is 3-D contact (for details see *e.g.* [2]). Then all optimal controls are indeed $L^\infty([0, T])$.

Finally, all trajectories satisfying the PMP are solutions of an Hamiltonian system, that in our case is analytic. Hence all minimizers are analytic and therefore smooth.

We conclude that it is equivalent to solve the original problem in the class of smooth curves or the lifted problem in the class of L^2 controls. A similar treatment has been presented with all details in [24] for a similar problem but in the presence of a drift.

3. A SUB-RIEMANNIAN PROBLEM ON $L(4, 1)$

In this section we define a $\mathfrak{k} \oplus \mathfrak{p}$ sub-Riemannian structure on $SU(2)$ and compute its cut locus and the sub-Riemannian distance. For details on Sub-Riemannian geometry and $\mathfrak{k} \oplus \mathfrak{p}$ manifolds, see Appendix A. We then define a sub-Riemannian structure on $L(4, 1)$ induced by the one on $SU(2)$ and prove that the sub-Riemannian problem on $L(4, 1)$ coincide with the optimal control problem (2.4) on PTS^2 .

3.1. The $\mathfrak{k} \oplus \mathfrak{p}$ sub-Riemannian structure on $SU(2)$

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

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

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

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

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

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

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

$$\mathfrak{k} = \text{span} \{k\} \quad \mathfrak{p} = \text{span} \{p_1, p_2\}$$

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

Defining $\Delta(g) = g\mathfrak{p}$ and $\mathfrak{g}_g(v_1, v_2) = \langle g^{-1}v_1, g^{-1}v_2 \rangle$, we have that $(SU(2), \Delta, \mathfrak{g})$ is a $\mathfrak{k} \oplus \mathfrak{p}$ sub-Riemannian manifold (for details see Appendix A.3). The sub-Riemannian manifold $SU(2)$ is thus endowed with the standard definition of sub-Riemannian length and distance.

3.1.1. Expression of geodesics

$\mathfrak{k} \oplus \mathfrak{p}$ manifolds have very special properties: there are no strict abnormal minimizers, hence all the geodesics starting from a point g_0 are parameterized by the initial covector (identified with a vector *via* the Killing form). The explicit expression of a geodesic starting from g_0 is

$$g(t) = g_0 e^{(A_k + A_p)t} e^{-A_k t}, \tag{3.1}$$

with $A_k \in \mathfrak{k}$, $A_p \in \mathfrak{p}$. The geodesic is parameterized by arclength when $\langle A_p, A_p \rangle = 1$. This condition defines a cylinder $\Lambda_{g_0} \subset T_{g_0}^*M$.

For the $SU(2)$ case, we compute the geodesics starting from the identity using formula (3.1). Consider an initial covector $\lambda = \lambda(\theta, c) = \cos(\theta)p_1 + \sin(\theta)p_2 + ck \in \Lambda_{\text{Id}}$. The corresponding exponential map is

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

with

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

Another property of $\mathfrak{k} \oplus \mathfrak{p}$ manifolds is the existence of a solution of the minimal length problem (see Rem. A.7), *i.e.* for each pair $g, h \in M$ there exists a trajectory γ steering g to h and minimizing the length: $l(\gamma) = d(g, h)$. Recall that this minimizing trajectory is a geodesic.

3.1.2. The cut locus and distance for $SU(2)$

In this section we recall the formula of the cut locus for the $\mathfrak{k} \oplus \mathfrak{p}$ manifold $SU(2)$ and of the sub-Riemannian distance. Both the results are proved in [8].

Theorem 3.1. *The cut locus for the minimal length problem on $SU(2)$ is*

$$K_{\text{Id}} = e^{\mathfrak{k}} \setminus \text{Id} = \{e^{ck} \mid c \in (0, 4\pi)\}.$$

The cut locus K_{Id} is topologically equivalent to a circle S^1 without a point, the starting point Id .

Theorem 3.2. *Let $g = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \in SU(2)$. Consider the sub-Riemannian distance from Id defined by the $\mathfrak{k} \oplus \mathfrak{p}$ structure on $SU(2)$. It holds*

$$d(\text{Id}, g) = \begin{cases} 2\sqrt{\arg(\alpha)(2\pi - \arg(\alpha))} & \text{if } \beta = 0 \\ \psi(\alpha) & \text{if } \beta \neq 0, \end{cases}$$

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

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

Remark 3.1. Notice that $\forall \alpha, \beta_1, \beta_2 \in \mathbb{C}, |\beta_1| = |\beta_2|$ we have $d\left(\text{Id}, \begin{pmatrix} \alpha \\ \beta_1 \end{pmatrix}\right) = d\left(\text{Id}, \begin{pmatrix} \alpha \\ \beta_2 \end{pmatrix}\right) = d\left(\text{Id}, \begin{pmatrix} \bar{\alpha} \\ \beta_1 \end{pmatrix}\right)$.

3.2. A sub-Riemannian problem on $L(4, 1)$

We define the 3-D sub-Riemannian manifold $L(4, 1)$ as a quotient of the $\mathfrak{k} \oplus \mathfrak{p}$ manifold $SU(2)$ defined above. $L(4, 1)$ is a lens space: for more details about lens spaces $L(p, q)$ see [23]. We prove that the quotient and the sub-Riemannian structure are compatible. We then compute the distance $\hat{d}(\cdot, \cdot)$ on $L(4, 1)$.

3.2.1. Definition of $L(4, 1)$

We define coordinates (a, b, c) on $SU(2)$ as follows:

Proposition 3.1. *Each $g \in SU(2)$ can be written as*

$$g = e^{-bp_2} e^{ap_1} e^{cp_2} \tag{3.2}$$

with $a \in [0, \pi]$, $b \in [0, 2\pi)$, $c \in \mathbb{R}/(4\pi)$. The value of a is uniquely determined by g .

If $a \neq 0$, $a \neq \pi$, then the values of b and c are uniquely determined by g .

If $a = 0$, then the value of $c - b \pmod{4\pi}$ is uniquely determined by g .

If $a = \pi$, then the value of $c + b \pmod{4\pi}$ is uniquely determined by g .

Proof. Observe that $e^{-bp_2} e^{ap_1} e^{cp_2} = \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix}$ with

$$\begin{cases} \alpha = \cos\left(\frac{a}{2}\right) \cos\left(\frac{c-b}{2}\right) + i \sin\left(\frac{a}{2}\right) \sin\left(\frac{c+b}{2}\right) \\ \beta = \sin\left(\frac{a}{2}\right) \cos\left(\frac{c+b}{2}\right) + i \cos\left(\frac{a}{2}\right) \sin\left(\frac{c-b}{2}\right). \end{cases}$$

The result follows from direct computation. □

In the following we will write $g = (a, b, c)$ for $g = e^{-bp_2} e^{ap_1} e^{cp_2}$, even though the vector (a, b, c) is not uniquely determined by g .

Proposition 3.2. *Define the following equivalence relation on $SU(2)$: $g_1 \approx g_2$ if there exist $(a_1, b_1, c_1), (a_2, b_2, c_2)$*

such that $g_1 = (a_1, b_1, c_1)$, $g_2 = (a_2, b_2, c_2)$ and
$$\begin{cases} a_1 = a_2 \\ b_1 = b_2 \\ c_1 = c_2 \pmod{\pi}. \end{cases}$$

The 3-D manifold $SU(2)/\approx$ is the lens space $L(4, 1)$.

Before proving the proposition, we recall the standard definition of $L(p, q)$ with $p, q \in \mathbb{Z} \setminus \{0\}$ coprime. Let

$$S^3 = \{(x_1, x_2) \in \mathbb{C}^2 \mid |x_1|^2 + |x_2|^2 = 1\}$$

and define on S^3 the following equivalence relation: $(x_1, x_2) \sim (y_1, y_2)$ if there exists $\omega \in \mathbb{C}$ p -th root of unity (i.e. $\omega^p = 1$) such that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & \omega^q \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

The quotient manifold S^3/\sim is the lens space $L(p, q)$.

Proof of Proposition 3.2. Consider the isomorphism

$$\phi : \begin{pmatrix} SU(2) & \rightarrow & S^3 \\ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} & \mapsto & \left(\frac{\alpha + \beta}{\sqrt{2}}, \frac{\operatorname{Re}(\alpha) + i\operatorname{Im}(\beta) - \operatorname{Re}(\beta) - i\operatorname{Im}(\alpha)}{\sqrt{2}} \right). \end{pmatrix}$$

A straightforward computation gives that $g_1, g_2 \in SU(2)$ are equivalent ($g_1 \approx g_2$) if and only if $\phi(g_1), \phi(g_2) \in S^3$ are equivalent ($\phi(g_1) \sim \phi(g_2)$) with respect to the equivalence relation defined by $p = 4$, $q = 1$. Hence the manifolds $SU(2)/\approx$ and S^3/\sim are isomorphic. Thus $SU(2)/\approx$ is a 3-D manifold, the lens space $L(4, 1)$. □

3.2.2. Sub-Riemannian structure on $L(4, 1)$

We now define a sub-Riemannian structure on $L(4, 1)$ induced by the quotient map.

Proposition 3.3. *The sub-Riemannian structure on $SU(2)$ given in Section 3.1 induces a 2-dim sub-Riemannian structure on $L(4, 1)$ via the quotient map*

$$\Pi : \begin{array}{ccc} SU(2) & \rightarrow & L(4, 1), \\ g & \mapsto & [g], \end{array}$$

i.e.

- the map

$$\tilde{\Delta} : [g] \mapsto \Pi_* (\Delta(h)) \subset T_{[g]}L(4, 1) \text{ with } h \in [g]$$

is a 2-dim smooth distribution on $L(4, 1)$ that is Lie bracket generating;

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

Proof. The role of the map Π and $\Pi_{*|_g}$ is illustrated in the following diagram

$$\begin{array}{ccc} T_gSU(2) & \xrightarrow{\Pi_{*|_g}} & T_{[g]}L(4, 1) \\ \downarrow & & \downarrow \\ SU(2) & \xrightarrow{\Pi} & L(4, 1). \end{array}$$

The map Π is a local diffeomorphism, thus $\Pi_{*|_g} : T_gSU(2) \rightarrow T_{[g]}L(4, 1)$ is a linear isomorphism, hence $\Pi_{*|_g} (\Delta(g))$ is a 2-dim subspace of $T_{[g]}L(4, 1)$.

The following statements:

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

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

- 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 $\langle v_1, w_1 \rangle_{h_1} = \langle v_2, w_2 \rangle_{h_2}$;

are consequences of the following lemma:

Lemma 3.1. *Let $h_1, h_2 \in [g]$ with $h_1 = (a, b, c)$ and $h_2 = (a, b, c + k\pi)$, $k \in \mathbb{Z}$. The map*

$$\phi : \begin{array}{ccc} \mathbf{p} & \rightarrow & \mathbf{p}, \\ m = m_1p_1 + m_2p_2 & \mapsto & (-1)^k m_1p_1 + m_2p_2 \end{array}$$

is bijective and it is an isometry w.r.t. the positive definite scalar product $\langle \cdot, \cdot \rangle$. The following equality holds $\forall m \in \mathbf{p}$

$$\frac{d}{dt} \Big|_{t=0} [h_1 e^{tm}] = \frac{d}{dt} \Big|_{t=0} [h_2 e^{t\phi(m)}].$$

Proof. A direct computation gives that ϕ is bijective and that it is an isometry.

Consider now $h_1 e^{tm}$ and $h_2 e^{tn}$ with $m = m_1p_1 + m_2p_2 \in \mathbf{p}$, $n = n_1p_1 + n_2p_2 \in \mathbf{p}$. We have

$$h_1 e^{tm} = e^{-bp_2} e^{ap_2} e^{cp_2} e^{tm} = \begin{pmatrix} \alpha(a, b, c, t, m) & \beta(a, b, c, t, m) \\ -\bar{\beta}(a, b, c, t, m) & \bar{\alpha}(a, b, c, t, m) \end{pmatrix},$$

with

$$\begin{aligned} \alpha(a, b, c, t, m) &= K_0(t, m)\alpha(a, b, c, 0, 0) + (-m_1 + im_2)K_1(t, m)\beta(a, b, c, 0, 0), \\ \beta(a, b, c, t, m) &= K_0(t, m)\beta(a, b, c, 0, 0) + (m_1 + im_2)K_1(t, m)\alpha(a, b, c, 0, 0), \\ K_0(t, m) &:= \cos\left(\frac{t\sqrt{m_1^2 + m_2^2}}{2}\right), \quad K_1(t, m) := \frac{\sin\left(\frac{t\sqrt{m_1^2 + m_2^2}}{2}\right)}{\sqrt{m_1^2 + m_2^2}}. \end{aligned}$$

Observe that $h_2e^{tn} = e^{-bp_2}e^{ap_2}e^{cp_2+k\pi p_2}e^{tn} \simeq \begin{pmatrix} \alpha_2 & \beta_2 \\ \beta_2 & \bar{\alpha}_2 \end{pmatrix}$ with

$$\begin{aligned} \alpha_2 &= K_0(t, n)\alpha(a, b, c, 0, 0) + ((-1)^{k+1}n_1 + in_2)K_1(t, n)\beta(a, b, c, 0, 0), \\ \beta_2 &= K_0(t, n)\beta(a, b, c, 0, 0) + ((-1)^kn_1 + in_2)K_1(t, n)\alpha(a, b, c, 0, 0), \end{aligned}$$

hence $[h_1e^{tm}] = [h_2e^{tn}]$ if both the following conditions are satisfied:
$$\begin{cases} K_0(t, m) = K_0(t, n) \\ K_1(t, m) = K_1(t, n) \\ -m_1 + im_2 = (-1)^{k+1}n_1 + in_2 \\ m_1 + im_2 = (-1)^kn_1 + in_2. \end{cases}$$

These are verified when $n = \phi(m)$. Thus $\frac{d}{dt}|_{t=0} [h_1e^{tm}] = \frac{d}{dt}|_{t=0} [h_2e^{t\phi(m)}]$. □

Since Π is a local diffeomorphism, $\forall g \in SU(2) \exists B(g)$ such that the map $\Pi_{*|_{B(g)}} : T_{B(g)}SU(2) \rightarrow T_{B([g])}L(4, 1)$ is a diffeomorphism, thus $\tilde{\Delta}$ is smooth, Lie bracket generating, and $\langle v_*, w_* \rangle_{[g]}$ is smooth as a function of $[g]$. □

Remark 3.2. Observe that other definitions of a sub-Riemannian structure on $L(4, 1)$ induced by the one on $SU(2)$ are possible: for example, in [8] we defined the following identification rule on $SU(2)$:

$$\begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \sim \begin{pmatrix} \alpha' & \beta' \\ -\bar{\beta}' & \bar{\alpha}' \end{pmatrix} \text{ if exists } \omega \in \mathbb{C} \text{ such that } \omega^p = 1 \text{ and } \begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \begin{pmatrix} \omega & 0 \\ 0 & \omega^q \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (3.3)$$

and observed that the sub-Riemannian structure defined on $SU(2)$ in Section 3.1 induces a sub-Riemannian structure on the manifold $L(p, q) := SU(2)/\sim$.

The sub-Riemannian structures are locally isomorphic, but globally the structure may vary: indeed, in the following we will prove that the cut locus for the structure defined in Section 3.3 is globally different from the one defined by (3.3).

Remark 3.3. Notice that in this case the push-forward of a left invariant vector field is not a vector field, because the projections $\Pi_{*|_g}(gp)$, $\Pi_{*|_h}(hp)$ from different points g, h such that $[g] = [h]$ do not coincide. Nevertheless, the projections of the whole distribution $\Pi_{*|_g}(\Delta(g))$, $\Pi_{*|_h}(\Delta(h))$ coincide; then the sub-Riemannian structure on $L(4, 1)$ is well defined even though the projections of the left invariant field is not well defined.

We have standard definitions of sub-Riemannian length and distance on $L(4, 1)$, as presented in Appendix A, see (A.2) and (A.3). We indicate them respectively with \hat{l} and \hat{d} .

Observe that the geodesics for the sub-Riemannian manifold $L(4, 1)$ are the projections of the geodesics for $SU(2)$, due to the fact that the projection Π is a local isometry. In general we have the following.

Proposition 3.4. *Let $\Gamma : [0, T] \rightarrow SU(2)$ be a smooth curve. Define its projection $\underline{\Gamma} : [0, T] \rightarrow L(4, 1)$ by $\underline{\Gamma}(t) := \Pi(\Gamma(t))$. Then $\underline{\Gamma}$ is a smooth curve and its length coincide with the length of Γ :*

$$\hat{l}(\underline{\Gamma}) = l(\Gamma).$$

Recall now the definition of the **lift** of a curve in our case and a standard result about its optimality:

Definition 3.1. Let $\gamma : [0, T] \rightarrow L(4, 1)$ be a smooth curve in $L(4, 1)$ with $\gamma(0) = [g]$. Fix $g_1 \in SU(2)$ such that $g_1 \in [g]$. The lift of γ starting from g_1 is the unique smooth curve $\bar{\gamma} : [0, T] \rightarrow SU(2)$ satisfying $\bar{\gamma}(0) = g_1$ and $\Pi(\bar{\gamma}(t)) = \gamma(t)$ for any $t \in [0, T]$.

The length of the lift $\bar{\gamma}$ in $SU(2)$ coincide with the length of the curve γ in $L(4, 1)$:

$$l(\bar{\gamma}) = \hat{l}(\gamma).$$

Proposition 3.5. *Let $\gamma : [0, T] \rightarrow L(4, 1)$ be an optimal trajectory steering $[g]$ to $[h]$. Fix $g_1 \in SU(2)$ with $g_1 \in [g]$. Then the lift $\bar{\gamma} : [0, T] \rightarrow SU(2)$ starting from g_1 is an optimal trajectory steering g_1 to $\bar{\gamma}(T) \in [h]$.*

We end this section giving an expression to compute the sub-Riemannian distance on $L(4, 1)$.

Proposition 3.6. *The sub-Riemannian distance \hat{d} on $L(4, 1)$ can be computed via the sub-Riemannian distance d on $SU(2)$ as follows:*

$$\hat{d}([g], [h]) = \min_{h_1 \in [h]} \{d(g_1, h_1)\}$$

independently on the choice of $g_1 \in [g]$.

Proof. Consider a trajectory γ in $L(4, 1)$ that is solution of the minimal length problem between $[g]$ and $[h]$. Define its lift $\bar{\gamma}(\cdot)$ starting from a fixed $g_1 \in [g]$. It steers g_1 to a given $h_1 \in [h]$ and it is optimal, thus $\hat{d}([g], [h]) = d(g_1, h_1)$.

We prove now that for any $h_2 \in [h]$ we have $d(g_1, h_1) \leq d(g_1, h_2)$. By contradiction, assume that $\exists h_2 \in [h]$ such that $d(g_1, h_2) < d(g_1, h_1)$, thus there exists a smooth curve $\Gamma : [0, T] \rightarrow SU(2)$ steering g_1 to h_2 with $l(\Gamma) < d(g_1, h_1)$. Consider its projection $\underline{\Gamma}$: it steers $[g]$ to $[h]$ and it satisfies $\hat{l}(\underline{\Gamma}) < \hat{d}([g], [h])$. Contradiction. \square

3.3. Isometry between $L(4, 1)$ and PTS^2

This section is fundamental in the treatment of the mechanical problem presented in Section 2: we state that the problem (2.4) on PTS^2 is the problem of minimal length on the sub-Riemannian manifold $L(4, 1)$.

We start writing the expression in coordinates of the projection $S^3 \rightarrow S^2$, the well-known Hopf map.

Proposition 3.7. *Consider the Hopf map*

$$SU(2) \simeq S^3 \rightarrow S^2,$$

$$\psi : \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \mapsto \begin{pmatrix} 2(\operatorname{Re}(\alpha)\operatorname{Re}(\beta) + \operatorname{Im}(\alpha)\operatorname{Im}(\beta)) \\ 2(\operatorname{Re}(\alpha)\operatorname{Im}(\alpha) - \operatorname{Re}(\beta)\operatorname{Im}(\beta)) \\ \operatorname{Re}(\alpha)^2 + \operatorname{Im}(\beta)^2 - \operatorname{Re}(\beta)^2 - \operatorname{Im}(\alpha)^2 \end{pmatrix}.$$

Consider the coordinates (a, b, c) on $SU(2)$ defined in (3.2) and the spherical coordinates (a, b) on S^2 defined in (2.3): then the Hopf map expression in coordinates is

$$\psi : \begin{matrix} SU(2) & \rightarrow & S^2, \\ (a, b, c) & \mapsto & (a, b). \end{matrix}$$

Remark 3.4. Observe that both the (a, b, c) coordinates on $SU(2)$ and the spherical coordinates (a, b) are not well defined in $a = 0$ and $a = \pi$. Nevertheless, the map ϕ is well defined and has this expression in coordinates also in the degenerate cases.

The following theorem is the central result of this section: here we prove that our mechanical problem (2.4) coincide with the sub-Riemannian problem on $L(4, 1)$.

Theorem 3.3. *The lens space $L(4, 1)$ is diffeomorphic to PTS^2 , the bundle of directions of the sphere. The coordinates (a, b, c) on $L(4, 1)$ coincide with the coordinates $(a, b, -c)$ on PTS^2 . The diffeomorphism maps horizontal curves on $L(4, 1)$ to horizontal curves on PTS^2 and vice versa.*

The sub-Riemannian length of an horizontal curve $(a(\cdot), b(\cdot), c(\cdot))$ in $L(4, 1)$ is equal to the cost $\mathcal{J}[\Gamma]$ of the curve $\Gamma(\cdot) = (a(\cdot), b(\cdot), -c(\cdot))$ in PTS^2 .

Proof. Consider any horizontal smooth curve $\Gamma : (-\varepsilon, \varepsilon) \rightarrow SU(2)$ with $\Gamma(0) = (a, b, c)$. We have $\Gamma(t) = e^{-bp_2} e^{ap_1} e^{cp_2} e^{t\eta + o(t)}$ with $\eta \in \mathfrak{p}$ due to the fact that Γ is horizontal. Assume $\eta = \eta_1 p_1 + \eta_2 p_2$ satisfying $\eta_1 \neq 0$. The reason of this condition will be clear in the following.

Consider its projection $\gamma = \psi \circ \Gamma : (-\varepsilon, \varepsilon) \rightarrow S^2$: we have by computation

$$\gamma(t) = \begin{pmatrix} \sin(a) \cos(b) \\ \sin(a) \sin(b) \\ \cos(a) \end{pmatrix} + t\eta_1 \begin{pmatrix} \cos(a) \cos(b) \cos(c) + \sin(b) \sin(c) \\ \cos(a) \sin(b) \cos(c) - \cos(b) \sin(c) \\ -\sin(a) \cos(c) \end{pmatrix} + o(t).$$

Due to condition $\eta_1 \neq 0$, γ is smooth and differentiable in 0; the direction of its tangent vector in 0 depends only on $\Gamma(0) = (a, b, c)$. Its lift is $(a(\cdot), b(\cdot), \xi(\cdot))$ with $\xi(t) = \arccos\left(\frac{b \sin(a)}{a}\right)$: an explicit computation gives

$\xi(t) = -\arctan(\tan(c))$. Thus the function $f : \begin{matrix} SU(2) & \rightarrow & PTS^2, \\ (a, b, c) & \mapsto & (a, b, -\arctan(\tan(c))) \end{matrix}$ maps horizontal smooth curves of $SU(2)$ with $\eta_1 \neq 0$ to horizontal smooth curves of PTS^2 .

The case $\eta_1 = 0$ can be studied as a limit case: in this case $\dot{a} = \dot{b} = 0$, thus f maps horizontal curves of $SU(2)$ to the curves of PTS^2 representing rotations on a point (see Rem. 2.1). In this sense f maps all horizontal smooth curves of $SU(2)$ to horizontal smooth curves of PTS^2 .

Consider now $L(4, 1) = SU(2)/\approx$: the map f passes to the quotient and we have the explicit expression

$$f_{\approx} : \begin{matrix} L(4, 1) & \rightarrow & PTS^2, \\ [a, b, c] & \mapsto & (a, b, -c). \end{matrix}$$

Observe that f_{\approx} is a diffeomorphism. Moreover, it maps horizontal smooth curves of $L(4, 1)$ to horizontal smooth curves of PTS^2 . Then the differential df_{\approx} is a linear isomorphism $df_{\approx} : \tilde{\Delta}([g]) \subset L(4, 1) \rightarrow \Delta(a, b, -c) \subset PTS^2$ that is explicitly $df_{\approx}(\eta_1 \Pi_{*1_g}(gp_1) + \eta_2 \Pi_{*1_g}(gp_2)) = \eta_1 F_1(a, b, -c) - \eta_2 F_2(a, b, -c)$ with $g = (a, b, c) \in SU(2)$. Then we have that the length of a curve $\gamma : [0, T] \rightarrow L(4, 1)$ satisfies

$$\hat{l}(a(\cdot), b(\cdot), c(\cdot)) = \int_0^T \sqrt{\eta_1^2 + \eta_2^2} = J[a(\cdot), b(\cdot)]. \quad \square$$

4. PROOF OF THE MAIN RESULT

We give now the explicit computation of the cut locus for the problem **(P)** on S^2 . For this specific problem the first conjugate locus (see Appendix A) is completely contained in the cut locus since this is the case for the lifted problem on $SU(2)$ (see [8], Sects. 3.1.3 and 5.1). Hence the cut locus is the set of points where geodesics lose optimality (see Rem. A.4).

Theorem 4.1. *The cut locus $K_{[\text{Id}]}$ for the sub-Riemannian problem on $L(4,1)$ starting from the identity is a stratification $K_{[\text{Id}]} = K^{\text{loc}} \cup K^{\text{sym}}$ with*

$$\begin{aligned} K^{\text{loc}} &= [K_{\text{Id}}] \setminus [\text{Id}] = [e^{\mathbf{k}}] \setminus [\text{Id}], \\ K^{\text{sym}} &= \{[g] \in L(4,1) \mid \text{for a fixed } g^* \in [g] \exists x_1, x_2 \in [\text{Id}], x_1 \neq x_2 \text{ such that} \\ &\quad \forall x^* \in [\text{Id}] \ d(g^*, x_1) = d(g^*, x_2) \leq d(g^*, x^*)\}, \end{aligned}$$

where $d(\cdot, \cdot)$ is the sub-Riemannian distance on $SU(2)$ given in Theorem 3.2.

Proof. We first prove that K^{loc} lies in the cut locus. Consider $[g] \in K^{\text{loc}}$: due to the definition of $[g]$, there exists $g_1 \in [g]$ such that $g_1 = e^{n\mathbf{k}} = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}$ with $|\alpha| = 1$. Hence $[g] = \{g_1, -g_1, g_2, -g_2\}$ with $g_2 = \begin{pmatrix} 0 & i\alpha \\ i\alpha & 0 \end{pmatrix}$. Recall that $\hat{d}([\text{Id}], [g]) = \min_{g^* \in [g]} d(\text{Id}, g^*)$. Applying Theorem 3.2, we have $d(\text{Id}, g_1) = 2\sqrt{\arg(\alpha)(2\pi - \arg(\alpha))}$, $d(\text{Id}, -g_1) = \begin{cases} 2\sqrt{(\arg(\alpha) + \pi)(\pi - \arg(\alpha))} & \text{if } \alpha \in [0, \pi] \\ 2\sqrt{(\arg(\alpha) - \pi)(3\pi - \arg(\alpha))} & \text{if } \alpha \in [\pi, 2\pi] \end{cases}$, $d(\text{Id}, g_2) = d(\text{Id}, -g_2) = \pi$. A straightforward computation gives that the minimum is attained for $d(\text{Id}, g_1)$ or $d(\text{Id}, -g_1)$. We assume without loss of generality that we have $\hat{d}([\text{Id}], [g]) = d(\text{Id}, g_1)$. Recall that $g_1 \in K_{\text{Id}}$, thus there exist γ_1, γ_2 two different optimal trajectories on $SU(2)$ steering Id to g_1 . Thus their projections $\underline{\gamma}_1, \underline{\gamma}_2$ are trajectories on $L(4,1)$ steering $[\text{Id}]$ to $[g]$ and satisfying $\hat{l}(\underline{\gamma}_i) = l(\gamma_i) = d(\text{Id}, g_1) = \hat{d}([\text{Id}], [g])$ for $i = 1, 2$; hence they are both optimal. Observe that γ_1, γ_2 are distinct in a neighborhood of Id , thus $\underline{\gamma}_1, \underline{\gamma}_2$ are distinct in a neighborhood of $[\text{Id}]$. Thus $[g]$ is a cut point, due to the fact that it is reached by two different optimal trajectories $\underline{\gamma}_1, \underline{\gamma}_2$.

We prove now that K^{sym} lies in the cut locus. Consider a point $[g] \in K^{\text{sym}}$ and a fixed $g^* \in [g]$. Recall that $d(g^*, x_1) = d(g^*, x_2)$. Due to the existence of a solution to the minimal length problem, there exist Γ_1 and Γ_2 geodesics in $SU(2)$ with the following property: Γ_i steers $x_i \in [\text{Id}]$ to $g^* \in [g]$ and it is optimal. Observe that $l(\Gamma_1) = l(\Gamma_2)$ and that Γ_1 and Γ_2 don't coincide in a neighborhood of g^* due to the fact that they are two different optimal geodesics. Now consider the two projections $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$: they are trajectories in $L(4,1)$ steering $[\text{Id}]$ to $[g]$ and their length is $\hat{l}(\underline{\Gamma}_i) = d(g^*, x_i) = \min_{x_j \in [\text{Id}]} d(g^*, x_j)$ due to the fact that $d(g^*, x_1) = d(g^*, x_2) \leq d(g^*, x^*)$. Hence $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$ are optimal geodesics. Notice that they are distinct in a neighborhood of $[g]$ because they are projections of two trajectories that are distinct in a neighborhood of g^* . Observe that the choice of $g^* \in [g]$ is not relevant, as a consequence of Theorem A.2.

We end proving that there is no cut point outside $K^{\text{loc}} \cup K^{\text{sym}}$. By contradiction, let $[g] \in L(4,1) \setminus (K^{\text{loc}} \cup K^{\text{sym}} \cup \{[\text{Id}]\})$ be a cut point: thus there are γ_1 and γ_2 two different optimal geodesics of $L(4,1)$ steering $[\text{Id}]$ to $[g]$. Consider their two lifts $\overline{\gamma}_1$ and $\overline{\gamma}_2$ starting from $\text{Id} \in SU(2)$. They reach respectively $g_1, g_2 \in [g]$ and they are optimal in $SU(2)$. We consider two distinct cases:

- $g_1 = g_2$: we have two distinct optimal trajectories $\overline{\gamma}_1$ and $\overline{\gamma}_2$ steering Id to the point g_1 . Hence $g_1 \in K_{\text{Id}}$. Observe that $g_1 \notin [\text{Id}]$ otherwise $[g] = [\text{Id}]$. Thus $g_1 \in K_{\text{Id}} \setminus [\text{Id}]$, hence $[g] \in K^{\text{loc}}$. Contradiction.
- $g_1 \neq g_2$: consider the two lifts of γ_1 and γ_2 reaching both the point g_1 and call them $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ respectively. Observe that $\tilde{\gamma}_1 = \overline{\gamma}_1$, hence $\tilde{\gamma}_1(0) = \text{Id}$. Instead, $\tilde{\gamma}_2(0) = x_2 \neq \text{Id}$: we prove it by contradiction. If $\tilde{\gamma}_2(0) = \text{Id}$, then $\tilde{\gamma}_2 = \overline{\gamma}_2$ by the uniqueness of the lift. Hence $\tilde{\gamma}_2(T) = g_2 \neq g_1$, that contradicts the definition of $\tilde{\gamma}_2$. Contradiction. Observe that $x_2 \in [\text{Id}]$ due to the fact that $\gamma_2(0) = [\text{Id}]$.

We have that both $\tilde{\gamma}_1$ and $\tilde{\gamma}_2$ are optimal, hence $d(\text{Id}, g_1) = l(\tilde{\gamma}_1) = \hat{l}(\gamma_1) = \hat{l}(\gamma_2) = l(\tilde{\gamma}_2) = d(x_2, g_1)$ with $x_2 \in [\text{Id}]$ and $x_2 \neq \text{Id}$. Observe that, choosing any $x^* \in [\text{Id}]$, we have $d(\text{Id}, g_1) \leq d(x^*, g_1)$ due to the optimality of γ_1 . Thus $[g] \in K^{\text{sym}}$. Contradiction. \square

We now show some pictures of the projections on S^2 of optimal geodesics on $L(4,1)$, i.e. the optimal trajectories for our problem.

In Figure 7 some optimal trajectories with cusps are shown: they meet in the same point with the same direction, where they lose optimality. This point with direction lies in the local cut locus K^{loc} .

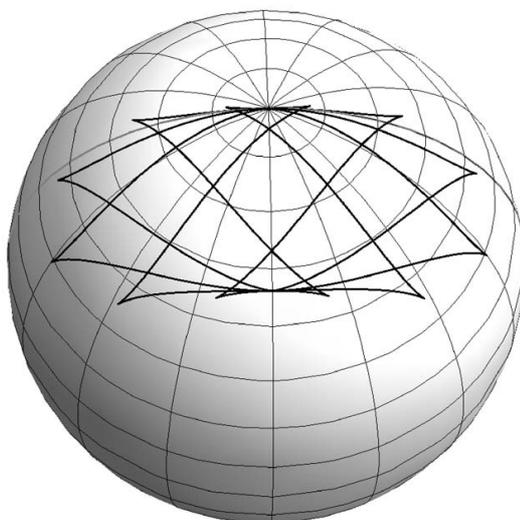


FIGURE 7. Projections of optimal trajectories with cusps.

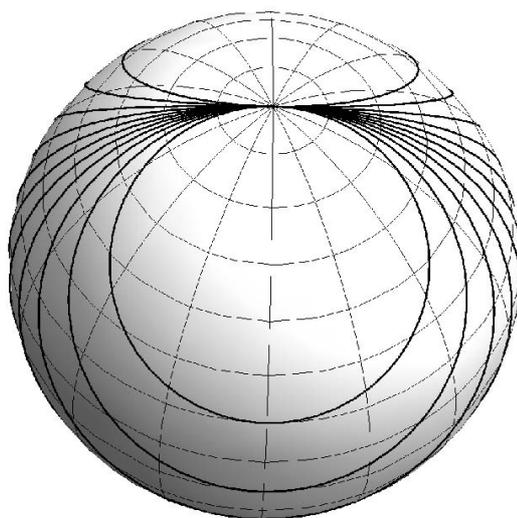


FIGURE 8. Projections of smooth optimal trajectories.

In Figure 8 we show smooth optimal trajectories, that are indeed half-circles: the points where they lose optimality lie in the intersection $K^{\text{loc}} \cap K^a$.

In Figure 9 we show optimal trajectories (both smooth and with cusps) meeting in one of the two points of intersections of the three strata $K^{\text{loc}} \cap K^a \cap K^b$.

APPENDIX A: BASIC RESULTS IN SUB-RIEMANNIAN GEOMETRY

In this appendix we recall some definitions and results about sub-Riemannian geometry, in particular about $\mathfrak{k} \oplus \mathfrak{p}$ manifolds. For a deeper presentation of sub-Riemannian geometry see *e.g.* [5,15].

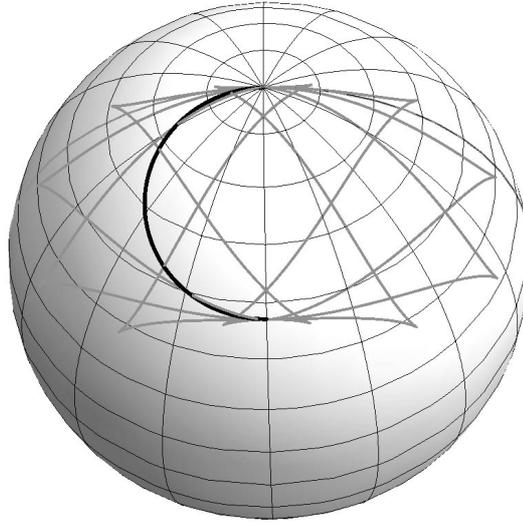


FIGURE 9. Projections of optimal trajectories (both smooth and with cusps) meeting in $K^{\text{loc}} \cap K^a \cap K^b$.

A.1. Sub-Riemannian manifold

A (n, m) -sub-Riemannian manifold is a triple (M, Δ, \mathbf{g}) , where

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

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

Here $\text{Vec}(M)$ denotes the set of smooth vector fields on M ;

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

The Lie bracket generating condition (A.1) is also known as Hörmander condition.

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

$$l(\gamma) = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))} dt. \tag{A.2}$$

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

$$d(q_0, q_1) = \inf\{l(\gamma) \mid \gamma(0) = q_0, \gamma(T) = q_1, \gamma \text{ horizontal}\}. \tag{A.3}$$

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 [4]). The function $d(\cdot, \cdot)$ is called the Carnot-Carathéodory distance and gives to M the structure of metric space (see [5,15]).

It is a standard fact that $l(\gamma)$ is invariant under reparameterization 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 side 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 $\gamma : [0, T] \rightarrow M$ 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 parameterized by arclength.

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

$$\Delta(q) = \text{span}\{F_1(q), \dots, F_m(q)\}, \quad \mathbf{g}_q(F_i(q), F_j(q)) = \delta_{ij}. \tag{A.4}$$

When (Δ, \mathbf{g}) can be defined as in (A.4) 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

$$\dot{q} = \sum_{i=1}^m u_i F_i(q), \quad u_i \in \mathbb{R}, \quad \int_0^T \sum_{i=1}^m u_i^2(t) dt \rightarrow \min, \quad q(0) = q_0, \quad q(T) = q_1. \tag{A.5}$$

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 $u_1^2 + \dots + u_m^2 \leq 1$.

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*.

A.2. First order necessary conditions, cut locus, conjugate locus

Consider a trivializable (n, m) -sub-Riemannian manifold. Solutions to the optimal control problem (A.5) are computed *via* the Pontryagin Maximum Principle (PMP for short, see for instance [4,7,16,21]) that is a first order necessary condition for optimality and generalizes the Weierstraß conditions of 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 A.1 (Pontryagin Maximum Principle for the problem (A.5)). *Let M be a n -dimensional smooth manifold and consider the minimization problem (A.5), in the class of Lipschitz continuous curves, where F_i , $i = 1, \dots, m$ are smooth vector fields on M and the final time T is fixed. Consider the map $\mathcal{H} : T^*M \times \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}$ defined by*

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

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

- (i) $\dot{q}(t) = \frac{\partial \mathcal{H}}{\partial \lambda}(q(t), \lambda(t), p_0, u(t));$
- (ii) $\dot{\lambda}(t) = -\frac{\partial \mathcal{H}}{\partial q}(q(t), \lambda(t), p_0, u(t));$
- (iii) $\frac{\partial \mathcal{H}}{\partial u}(q(t), \lambda(t), p_0, u(t)) = 0.$

Remark A.1. A curve $q(\cdot) : [0, T] \rightarrow M$ satisfying the PMP is said to be an *extremal*. In general, an extremal may correspond to more than one pair $(\lambda(\cdot), p_0)$. 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 [6,13].

It is well known that all normal extremals are geodesics (see for instance [4]). Moreover if there are no strict abnormal minimizers then all geodesics are normal extremals for some fixed final time T . This is the case for the so called 3-D *contact case*, i.e. a sub-Riemannian manifold of dimension 3 for which $\Delta = \text{span}\{F_1, F_2\}$ where F_1, F_2 is a pair of vector fields such that for all $q \in M$, $\text{span}\{F_1(q), F_2(q), [F_1(q), F_2(q)]\} = T_qM$. Indeed for contact structures there are no abnormal extremals (even non strict).

In this case, from (iii) one gets $u_i(t) = \langle \lambda(t), F_i(t) \rangle$, $i = 1, \dots, m$ and the PMP becomes much simpler: a curve $q(\cdot)$ 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

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

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

Remark A.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 parameterized 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

$$H(\lambda_0, q_0) = 1/2 \tag{A.6}$$

and every $t > 0$ define the *exponential map* $\text{Exp}(\lambda_0, t)$ 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 (A.6) defines a hypercylinder $\Lambda_{q_0} \simeq S^{m-1} \times \mathbb{R}^{n-m}$ in $T_{q_0}^*M$.

Definition A.1. The **conjugate locus from q_0** is the set C_{q_0} of critical values of the map

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

For every $\bar{\lambda}_0 \in \Lambda_{q_0}$, let $t(\bar{\lambda}_0)$ be the n -th positive time, if it exists, for which the map $(\lambda_0, t) \mapsto \text{Exp}(\lambda_0, t)$ is singular at $(\bar{\lambda}_0, t(\bar{\lambda}_0))$. The **n -th conjugate locus from q_0** $C_{q_0}^n$ is the set $\{\text{Exp}(\bar{\lambda}_0, t(\bar{\lambda}_0)) \mid t(\bar{\lambda}_0) \text{ exists}\}$.

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

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

Remark A.3. It is a standard fact that for every $\bar{\lambda}_0$ satisfying (A.6), the set $T(\bar{\lambda}_0) = \{\bar{t} > 0 \mid \text{Exp}(\lambda, t) \text{ is singular at } (\bar{\lambda}_0, \bar{t})\}$ is a discrete set (see for instance [4]).

Remark A.4. Let (M, Δ, \mathbf{g}) be a sub-Riemannian manifold. Fix $q_0 \in M$ and assume: (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 3-D contact case can be found in [2]).

- The first conjugate locus $C_{q_0}^1$ is the set of points where the geodesics starting from q_0 lose local optimality;
- If $q(\cdot)$ is a geodesic starting from q_0 and \bar{t} is the first positive time such that $q(\bar{t}) \in K_{q_0} \cup C_{q_0}^1$, then $q(\cdot)$ loses optimality in \bar{t} , i.e. it is optimal in $[0, \bar{t}]$ and not optimal in $[0, t]$ for any $t > \bar{t}$;
- If a geodesic $q(\cdot)$ starting from q_0 loses optimality at $\bar{t} > 0$, then $q(\bar{t}) \in K_{q_0} \cup C_{q_0}^1$.

As a consequence, when the first conjugate locus is included in the cut locus, the cut locus is the set of points where the geodesics lose optimality.

Remark A.5. It is well known that, while in Riemannian geometry K_{g_0} is never adjacent to g_0 , in sub-Riemannian geometry this is always the case. See [3].

A.3 $\mathfrak{k} \oplus \mathfrak{p}$ sub-Riemannian manifolds

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

Definition A.2. A Cartan decomposition of a simple Lie algebra \mathbf{L} is any decomposition of the form:

$$\mathbf{L} = \mathfrak{k} \oplus \mathfrak{p}, \text{ where } [\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}, \quad [\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subseteq \mathfrak{p}.$$

Definition A.3. Let G be a simple Lie group with Lie algebra \mathbf{L} . Let $\mathbf{L} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan decomposition of \mathbf{L} . In the case in which G is non-compact assume that \mathfrak{k} is the maximal compact subalgebra of \mathbf{L} .

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

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

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

Remark A.6. In the compact (resp. non-compact) 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 \mathfrak{k} maximal compact subalgebra).

Let $\{X_j\}$ be an orthonormal frame for the subspace $\mathfrak{p} \subset \mathbf{L}$, with respect to the metric defined in Definition A.3. 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

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

Remark A.7. This problem admits a solution, see for instance Chapter 5 of [10].

For $\mathfrak{k} \oplus \mathfrak{p}$ sub-Riemannian manifolds, one can prove that strict abnormal extremals are never optimal, since the *Goh condition* (see [4]) is never satisfied. Moreover, the Hamiltonian system given by the Pontryagin Maximum Principle is integrable and the explicit expression of geodesics starting from g_0 and parameterized by arclength is

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

where $A_k \in \mathfrak{k}$, $A_p \in \mathfrak{p}$ and $\langle A_p, A_p \rangle = 1$. This formula has been known from long time. It was used independently by Agrachev [1], Brockett [11] and Kupka (oral communication). The first complete proof was written by Jurdjevic in [17]. The proof that strict abnormal extremals are never optimal was first written in [9]. See also [4,18].

Acknowledgements. We thank A. Agrachev for his help in recognizing the structure of the projective tangent bundle. We thank L. Paoluzzi for many explanations on lens spaces. We thank M. Saponi and G. Senaldi for their help with pictures.

REFERENCES

- [1] A. Agrachev, Methods of control theory in nonholonomic geometry, in *Proc. ICM-94*, Birkhauser, Zürich (1995) 1473–1483.
- [2] A. Agrachev, Exponential mappings for contact sub-Riemannian structures. *J. Dyn. Contr. Syst.* **2** (1996) 321–358.
- [3] A. Agrachev, Compactness for sub-Riemannian length-minimizers and subanalyticity. *Rend. Sem. Mat. Univ. Politec. Torino* **56** (2001) 1–12.
- [4] A.A. Agrachev and Yu.L. Sachkov, *Control Theory from the Geometric Viewpoint*, *Encyclopedia of Mathematical Sciences* **87**. Springer (2004).
- [5] A. Bellaïche, The tangent space in sub-Riemannian geometry, in *Sub-Riemannian Geometry*, *Progress in Mathematics* **144**, Birkhäuser, Basel (1996) 1–78.
- [6] B. Bonnard and M. Chyba, *Singular trajectories and their role in control theory*. Springer-Verlag, Berlin (2003).
- [7] U. Boscain and B. Piccoli, *Optimal Synthesis for Control Systems on 2-D Manifolds*, *SMAI* **43**. Springer (2004).
- [8] U. Boscain and F. Rossi, Invariant Carnot-Carathéodory metrics on S^3 , $SO(3)$, $SL(2)$ and lens spaces. *SIAM J. Contr. Opt.* **47** (2008) 1851–1878.
- [9] U. Boscain, T. Chambrion and J.P. Gauthier, On the K+P problem for a three-level quantum system: Optimality implies resonance. *J. Dyn. Contr. Syst.* **8** (2002) 547–572.
- [10] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, *Appl. Math. Series* **2**. American Institute of Mathematical Sciences (2007).
- [11] R.W. Brockett, Explicitly solvable control problems with nonholonomic constraints, in *Proceedings of the 38th IEEE Conference on Decision and Control* **1** (1999) 13–16.
- [12] Y. Chitour and M. Sigalotti, Dubins' problem on surfaces. I. Nonnegative curvature *J. Geom. Anal.* **15** (2005) 565–587.
- [13] Y. Chitour, F. Jean and E. Trélat, Genericity results for singular curves. *J. Differential Geometry* **73** (2006) 45–73.
- [14] G. Citti and A. Sarti, A cortical based model of perceptual completion in the roto-translation space. *J. Math. Imaging Vis.* **24** (2006) 307–326.
- [15] M. Gromov, Carnot-Carathéodory spaces seen from within, in *Sub-Riemannian Geometry*, *Progress in Mathematics* **144**, Birkhäuser, Basel (1996) 79–323.
- [16] V. Jurdjevic, *Geometric Control Theory*. Cambridge University Press (1997).
- [17] V. Jurdjevic, Optimal Control, Geometry and Mechanics, in *Mathematical Control Theory*, J. Bailleu and J.C. Willems Eds., Springer, New York (1999) 227–267.
- [18] V. Jurdjevic, Hamiltonian Point of View on non-Euclidean Geometry and Elliptic Functions. *System Control Lett.* **43** (2001) 25–41.
- [19] I. Moiseev and Yu.L. Sachkov, *Cut locus for the sub-Riemannian problem on $E(2)$* . [arXiv:0807.4731](https://arxiv.org/abs/0807.4731) (to appear).
- [20] J. Petitot, Vers une Neuro-géométrie. Fibrations corticales, structures de contact et contours subjectifs modaux, in *Mathématiques, Informatique et Sciences Humaines* **145**, Special issue, EHESS, Paris (1999) 5–101.
- [21] L.S. Pontryagin, V. Boltianski, R. Gamkrelidze and E. Mitchtchenko, *The Mathematical Theory of Optimal Processes*. John Wiley and Sons, Inc. (1961).
- [22] J.A. Reeds and L.A. Shepp, Optimal paths for a car that goes both forwards and backwards. *Pacific J. Math.* **145** (1990) 367–393.
- [23] D. Rolfsen, *Knots and links*. Publish or Perish, Houston (1990).
- [24] Yu.L. Sachkov, Maxwell strata in Euler's elastic problem. *J. Dyn. Contr. Syst.* **14** (2008) 169–234.
- [25] M. Spivak, *A comprehensive introduction to differential geometry*. Second edition, Publish or Perish, Inc., Wilmington, Del. (1979).