HAL
open science

Metrics with equatorial singularities on the sphere

Bernard Bonnard, Jean-Baptiste Caillau

To cite this version:

Bernard Bonnard, Jean-Baptiste Caillau. Metrics with equatorial singularities on the sphere. Annali di Matematica Pura ed Applicata, 2014, 193 (5), pp.1353-1382. 10.1007/s10231-013-0333-y . hal-00319299v7

HAL Id: hal-00319299

https://hal.science/hal-00319299v7

Submitted on 19 Mar 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Metrics with equatorial singularities on the sphere*

B. Bonnard† J.-B. Caillau‡

September 2012

Abstract

Motivated by optimal control of affine systems stemming from mechanics, metrics on the two-sphere of revolution are considered; these metrics are Riemannian on each open hemisphere whereas one term of the corresponding tensor becomes infinite on the equator. Length minimizing curves are computed and structure results on the cut and conjugate loci are given, extending those in [11]. These results rely on monotonicity and convexity properties of the quasi-period of the geodesics; such properties are studied on an example with elliptic transcendency. A suitable deformation of the round sphere allows to reinterpretate the equatorial singularity in terms of concentration of curvature and collapsing of the sphere onto a two-dimensional billiard.

Keywords. two-sphere of revolution, almost and sub-Riemannian metrics, cut and conjugate locus

MSC classification. 53C17, 49K15

Introduction

In the papers [11, 35] the authors investigate the cut and conjugate loci of a Riemannian metric on the two-sphere of revolution that can be written in the normal form $m(\varphi)\mathrm{d}\theta^2 + \mathrm{d}\varphi^2$—where φ is the angle along the meridian and θ the angle of revolution—, under the assumption that $m(\pi - \varphi) = m(\varphi)$ (symmetry with respect to the equator). Motivated by examples in optimal control, our aim is to extend these results to metrics of the form

$$XR(X)\mathrm{d}\theta^2 + \mathrm{d}\varphi^2, \quad X = \sin^2\varphi, \tag{1}$$

*Supported by ANR Geometric Control Methods (project no. NT09_504490).

†Math. Institute, Univ. Bourgogne & CNRS, 9 avenue Savary, F-21078 Dijon (bernard.bonnard@u-bourgogne.fr). On leave at INRIA, 2004 route des lucioles, F-06902 Sophia Antipolis.

‡Same address (jean-baptiste.caillau@u-bourgogne.fr). Also supported by Conseil Régional de Bourgogne (contract no. 2009-160E-160-CE-160T) and by SADCO Initial Training Network (FP7 grant no. 264735-SADCO).

where R is a rational fraction with a single pole at $X = 1$. The metric is Riemannian on both open hemispheres but has a singularity on the equator. Such metrics arise when one considers the time minimum control of a system

$$\dot{x}(t) = u_1(t)F_1(x(t)) + u_2(t)F_2(x(t)), \quad u_1^2(t) + u_2^2(t) \leq 1,$$

with fixed boundary conditions, $x(0) = x_0$, $x(t_f) = x_f$, and it is well known that minimizing time amounts to minimizing the length of the curve $t \mapsto x(t)$ provided the metric is defined by assuming F_1 and F_2 orthonormal. For such problems, singularities occur when the two vector fields become collinear. This leads to the concept of almost-Riemannian metrics [3, 4, 12, 15], and the analysis of such metrics is related to sub-Riemannian geometry [6, 26]. If the distribution $\{F_1, F_2\}$ is bracket generating, every pair of points can be joined by a length minimizing curve. Moreover, if there exists no abnormal trajectory, each minimizing curve is a geodesic, projection on the x-space of the Hamiltonian flow $\exp t\overrightarrow{h}$ where

$$h(z) = \frac{1}{2}\left(H_1^2(z) + H_2^2(z)\right), \quad z = (x, p),$$

and where $H_i(z) = \langle p, F_i(x)\rangle$ are the Hamiltonian lifts of the F_i's. Denoting $\exp_{x_0}$ the exponential mapping

$$\exp_{x_0}(t, p_0) = \Pi\left(\exp t\overrightarrow{h}(x_0, p_0)\right)$$

where Π is the projection $(x, p) \mapsto x$, the cut and conjugate loci are defined as in the Riemannian setting; the cut locus is the set of points where geodesic curves fail to be minimizing, while the conjugate locus is the image of the critical points of the exponential mapping.

In this paper, we generalize to the singular case (1) the results in [11] relating the structure of the cut and conjugate loci to the convexity of the quasi-period of the θ-coordinate: Under appropriate assumptions, the cut locus of a point is reduced to a single segment, and the conjugate locus has at most four cusps. Although similar to the Riemannian case, the singularity of the metric on the equator has two consequences; the injectivity radius is zero and the singularities of the conjugate locus of an equatorial point differ. In Section 2, two examples motivating this study are presented. One stems from quantum mechanics, the other from space mechanics. Both are limit cases of optimal control problems not linear but affine in the control. Section 3 is devoted to the integrability properties of the geodesics of (1), and preliminary computations of the period and quasi-period of coordinates φ and θ are made. The optimality status of the geodesics is studied in Section 4 where the main results on the structure of cut and conjugate loci are established. The second example of §2 is given a full treatment in Section 5. In order to be able to apply §4 results, a detailed analysis using a parameterization of geodesics by elliptic curves is presented. The last section accounts for concentration of curvature and collapsing phenomena that allow to reinterpretate the effect of the equatorial singularity on the metric.

1 Preliminaries

We consider the metric (1) on $\mathbf{S}^2$ where (θ, φ) are the coordinates induced by the covering of the sphere minus poles

$$\mathbf{R} \times (0, \pi) \ni (\theta, \varphi) \mapsto (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi).$$

The function R is a rational fraction with a single pole of order p at $X = 1$,

$$R(X) = \sum_{k=0}^{p} \frac{a_k}{(1-X)^k}, \quad a_0, \ldots, a_{p-1} \geq 0, \quad a_p > 0 \quad (p \geq 1). \tag{2}$$

We assume that the normalization condition $a_0 + \cdots + a_p = 1$ holds. The metric is Riemannian on hemispheres with a singularity on the equator when $\varphi = \pi/2$ since $R(X) \to \infty$ when $X \to 1$. It is a standard fact that Riemannian geometry can be recast in the Hamiltonian setting of optimal control: On each open hemisphere, finding a length minimizing curve connecting two points x_0, x_f, is equivalent to finding a measurable control function $u : [0, t_f] \to \mathbf{R}^2$ such that almost everywhere

$$\dot{x}(t) = u_1(t)F_1(x(t)) + u_2(t)F_2(x(t)), \quad u_1^2(t) + u_2^2(t) \leq 1, \tag{3}$$
$$x(0) = x_0, \quad x(t_f) = x_f,$$

and such that the final time t_f is minimized. In the previous coordinates, the vector fields are (because of the topology of the 2-sphere, one cannot provide an orthonormal frame of vector fields globally defined on $\mathbf{S}^2$)

$$F_1(\theta, \varphi) := \frac{1}{\sin\varphi}\sqrt{1/R(X)}\,\frac{\partial}{\partial\theta}, \quad F_2(\theta, \varphi) := \frac{\partial}{\partial\varphi} \tag{4}$$

as

Lemma 1. *$1/R(X)$ has a smooth square root.*

Proof. One has

$$\frac{1}{R(X)} = \frac{\cos^{2p}\varphi}{a_p + \cdots + a_0\cos^{2p}\varphi}$$

and $a_p + \cdots + a_0\cos^{2p}\varphi \geq a_p > 0$ by virtue of (2). □

This optimal control formulation makes sense on the whole sphere: This is how (1) will be understood as a metric defined on all $\mathbf{S}^2$ in this paper. Geometrically, the singularity $R(1) = \infty$ forces length minimizing curves to be vertical (directed along meridians) when crossing the equator.

Remark 1. The singularity here is not a degeneracy of the Riemannian tensor on the tangent space as in, *e.g.*, [30] but a degeneracy of a positive tensor on the cotangent space.

An appropriate algebraic setting for such a metric is the notion of generalized sub-Riemannian metrics (see [6, 24]) or *almost-Riemannian metrics* (see [1, 3, 4, 12]). Let g_0 be the round metric on the sphere; to (1) is associated the following morphism of fibre bundle:[1]

$$f : (T\mathbf{S}^2, g_0) \to T\mathbf{S}^2$$

[1] That is id-morphism, according to [16] terminology: Base points are unchanged by the morphism, and fibres are linearly sent to fibres.

such that f is the identity on the fibres above the poles and

$$f_{(\theta,\varphi)} : \frac{\partial}{\partial\theta} \mapsto \sqrt{1/R(X)}\,\frac{\partial}{\partial\theta}\,, \quad \frac{\partial}{\partial\varphi} \mapsto \frac{\partial}{\partial\varphi}$$

otherwise. This morphism induces an application between vector fields on the sphere, $f_* : \Gamma(T\mathbf{S}^2) \to \Gamma(T\mathbf{S}^2)$; if $F \in \Gamma(T\mathbf{S}^2)$,

$$f_*(F)(x) := f_x(F(x)), \quad x \in \mathbf{S}^2.$$

Although the direct image $f(T\mathbf{S}^2)$ is not a fibre bundle,

$$\Delta := f_*(\Gamma(T\mathbf{S}^2)) \subset \Gamma(T\mathbf{S}^2)$$

is a well defined submodule over smooth functions. For $x \in \mathbf{S}^2$ and $v \in \Delta_x := \{F(x),\ F \in \Delta\}$, define

$$g_x(v) := \inf\{g_{0_x}(u) \mid f_x(u) = v\}.$$

Outside poles,

$$g_x(v) = \inf\{u_1^2 + u_2^2 \mid u_1F_1(x) + u_2F_2(x) = v\}.$$

Then, given two points x_0, x_f on the sphere, set

$$d(x_0, x_f) := \inf \int_0^{t_f} \sqrt{g_{x(t)}(\dot{x}(t))}\,\mathrm{d}t$$

where the infimum is taken over all Lipschitz trajectories x such that

$$\dot{x}(t) \in \Delta_{x(t)}, \quad t \in [0, t_f] \quad \text{(a.e.)}$$

$$x(0) = x_0, \quad x(t_f) = x_f.$$

The set of such *horizontal* trajectories is not empty as is clear from the proof of

Proposition 1. *d defines a complete distance on $\mathbf{S}^2$ that induces the usual topology on the sphere.*

To prove this fact, one defines recursively the flag associated with Δ by means of the Lie bracket of vector fields,

$$\Delta^1 := \Delta, \quad \Delta^{k+1} := \Delta^k + [\Delta, \Delta^k]$$

with $[\Delta, \Delta^k] := \{[F, G] \mid F \in \Delta,\ G \in \Delta^k\}$, $k \geq 1$.

Lemma 2. *For all $x \in \mathbf{S}^2$, $\Delta_x^{p+1} = T_x\mathbf{S}^2$.*

Proof. Outside the equator, F_1 and F_2 have rank two, so the verification is restricted to $\varphi = \pi/2$ where F_1 vanishes. Since R has an order p pole,

$$(\mathrm{ad}^p F_2)F_1(\theta,\varphi) = \underbrace{\frac{\mathrm{d}^p}{\mathrm{d}\varphi^p}[(1/\sin\varphi)\sqrt{1/R(X)}]}_{\neq 0 \text{ at } \varphi=\pi/2}\,\frac{\partial}{\partial\theta}$$

so brackets of length at most $p+1$ span everywhere the whole tangent space. □

Proof of Proposition 1. According to lemma 2, Δ is bracket generating so there exist, for any pair of points on the sphere, Lipschitz curves almost everywhere tangent to Δ connecting them by Chow-Rashevsky theorem. On the compact manifold $\mathbf{S}^2$, Filippov theorem [5] then asserts existence of time minimizing curves among those. That the metric induces the canonical manifold topology is another consequence of Chow-Rashevsky [24]. □

Remark 2. As explained in [3], generic (in the sense of Whitney) almost-Riemannian metrics on the sphere are such that, for all $x \in \mathbf{S}^2$, either Δ_x (regular point), Δ_x^2 (*Grušin point*) or Δ_x^3 (*tangency point*) span the tangent space. According to Lemma 2, the situation we consider is not generic for $p \geq 3$.

2 Motivating examples

The following Hamiltonian on $\mathbf{S}^2$ is considered in [9]:

$$H(\theta, \varphi, p_\theta, p_\varphi) = -\delta \cos\varphi \sin\varphi \, p_\varphi + \frac{1}{2}\left(\frac{p_\theta^2}{\tan^2\varphi} + p_\varphi^2\right). \tag{5}$$

It originates in quantum mechanics and partly describes the energy minimum control of a spin 1/2 particle in a magnetic field. When the parameter δ is zero, this Hamiltonian corresponds exactly to the metric (1) obtained for $R(X) = 1/(1-X)$ (geodesics are integral curves of an appropriate quadratic Hamiltonian—see the beginning of §3). As explained in §4, a local model near the singularity $\varphi = \pi/2$ is the so-called *Grušin metric* on $\mathbf{R}^2$ [14, 22]

$$\mathrm{d}x^2 + \frac{\mathrm{d}y^2}{x^2}.$$

For this reason, the metric defined by (5) when $\delta = 0$ is called *Grušin metric on the sphere*. In contrast to the analogous metric on the plane, it has peculiarities (*e.g.*, meridional cusps of conjugate loci, see §4) due to the topology of the sphere. Like the second example, it is also connected with space mechanics (see [11]).

Consider a controlled dynamical system of the form

$$\frac{\mathrm{d}x}{\mathrm{d}l}(l) = \sum_{i=1}^{m} u_i(l) F_i(l, x(l)) \tag{6}$$

on a smooth n-dimensional manifold X where the $F_i : \mathbf{R} \times X \to X$ are vector fields parameterized by l and periodic, $F(l, x) = F(l + 2\pi, x)$ for all $l \in \mathbf{R}$ and $x \in X$. This dynamics is actually a particular case of an (autonomous) affine in the control system, as we may define the vector fields

$$\hat{F}_0(\hat{x}) := \omega(\hat{x})\frac{\partial}{\partial l}, \quad \hat{F}_i(\hat{x}) := \omega(\hat{x}) \sum_{j=1}^{n} \langle \mathrm{d}x_j, F_i \rangle \frac{\partial}{\partial x_j}, \quad i = 1, \ldots, m,$$

with $\hat{x} := (l, x) \in \hat{X} := \mathbf{R} \times X$ and write

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t}(t) = \hat{F}_0(\hat{x}(t)) + \sum_{i=1}^{m} u_i(t) \hat{F}_i(\hat{x}(t)).$$

The pulsation ω is any 2π-periodic in l positive function on $\hat{X}$ relating the two times of the system, t and l (the latter being understood as an angular length). The periodic vector fields F_i induce vector fields on $\mathbf{S}^1 \times X$ and it is appealing to use averaging [27] to define

$$\overline{F}_i(x) := \frac{1}{2\pi} \int_0^{2\pi} F_i(l, x)\, \mathrm{d}l.$$

We consider instead the Hamiltonian associated with the $\mathrm{L}^2\mathrm{d}t$-minimization of the control,

$$H(l, x, p) = \frac{\omega(l, x)}{2} \sum_{i=1}^{m} \langle p, F_i(l, x)\rangle^2.$$

Then,

$$\overline{H}(x, p) := \frac{1}{2\pi} \int_0^{2\pi} H(l, x, p)\, \mathrm{d}l$$

remains a positive quadratic form in the adjoint variable, p. Under the standard assumption that the $\hat{F}_i$, $i = 0, \ldots, m$, are bracket generating, one expects the rank of this quadratic form to be maximum and equal to n. The case being, one can find n independent vector fields $\overline{F}_i$ such that

$$\overline{H}(x, p) = \frac{1}{2} \sum_{i=1}^{n} \langle p, \overline{F}_i(x)\rangle^2.$$

Singularities of two types may exist though. First, such implicitly defined vector fields need not be smooth on X, even in the analytic situation.[2] In addition, the rank of the form may not be constant on X and drop at some points. This phenomenon accounts for the existence of an equatorial singularity in the following case.

On $X = \mathbf{R}_+^* \times \mathbf{D}$, $\mathbf{D}$ being the open unit Poincaré disk, set

$$F_1 := -\frac{3(1-e^2)w}{n^{1/3}(1+e\cos v)^2} \frac{\partial}{\partial n} + \frac{2(1-e^2)^2}{n^{4/3}(1+e\cos v)^2 w} \left[(e+\cos v)\frac{\partial}{\partial e} + \frac{\sin v}{e}\frac{\partial}{\partial \theta} \right]$$

with

$$v := l - \theta, \quad w := \sqrt{1 + 2e\cos(l-\theta) + e^2}\,,$$

and

$$\omega(l, x) := \frac{n(1+e\cos(l-\theta))^2}{(1-e^2)^{3/2}}.$$

Here, $x = (n, e, \theta) \in \mathbf{R}_+^* \times \mathbf{R}_+^* \times \mathbf{R}$ are coordinates on the positive line times the pointed disk. They are used in space mechanics to represent the geometry of plane elliptic trajectories in the controlled two-body problem [10]: n is the mean motion (that is $n = a^{-3/2}$ where a is the semi-major axis), e is the eccentricity, and θ is the argument of the pericenter. The averaging procedure just described leads to [10]

$$\overline{H}(n, e, \theta, p_n, p_e, p_\theta) = \frac{9n^{1/3}}{2} p_n^2 + \frac{1}{n^{5/3}} h(e, \theta, p_e, p_\theta)$$

[2] See, *e.g.*, §3.2 in Bonnard, B.; Caillau, J.-B.; Picot, G. Geometric and numerical techniques in optimal control of two and three-body problems. *Commun. Inf. Syst.* **10** (2010), no. 4, 239–278.

with

$$h(e,\theta,p_e,p_\theta)=\frac{1}{2}\left[\frac{4(1-e^2)^{3/2}}{1+\sqrt{1-e^2}}p_e^2+\frac{4(1-e^2)}{(1+\sqrt{1-e^2})}\frac{p_\theta^2}{e^2}\right].$$

Proposition 2. *$\overline{H}$, h and p_θ are independent first integrals in involution; in particular, $\overline{H}$ is Liouville integrable.*

Proof. The Poisson bracket of $\overline{H}$ and h is

$$\{\overline{H},h\}=\frac{\partial\overline{H}}{\partial p_n}\frac{\partial h}{\partial n}-\frac{\partial\overline{H}}{\partial n}\frac{\partial h}{\partial p_n}+\frac{1}{n^{5/3}}\{h,h\}=0,$$

the rest being obvious. □

The integration of $\overline{H}$ may so be performed by integrating h. This two-dimensional Hamiltonian (which retains the linear first integral p_θ) is lifted from the Poincaré disk to $\mathbf{S}^2$ using the following compactification: In the covering $(\theta,\varphi)\in\mathbf{R}\times(0,\pi)$ of the sphere minus poles where

$$e=\sin\varphi\sqrt{1+\cos^2\varphi}\,,$$

one has

$$h(\theta,\varphi,p_\theta,p_\varphi)=\frac{1}{2}\left[\frac{4\cos^4\varphi}{\sin^2\varphi(2-\sin^2\varphi)^2}p_\theta^2+p_\varphi^2\right].$$

The rank of this quadratic form in (p_θ,p_φ) drops from 2 to 1 when $\varphi=\pi/2$. Actually, h is exactly associated to the metric (1) with

$$R(X)=\left(\frac{1-X/2}{1-X}\right)^2=\frac{1}{4}\left[1+\frac{2}{1-X}+\frac{1}{(1-X)^2}\right].$$

This metric with an order two equatorial singularity will be referred to as the $(1,2,1)$ case (according to coefficients involved in the series) and studied in Section 5.

Remark 3. The Hamiltonian (5) in the first example can either be interpretated as stemming from an affine controlled system, or as defining a pseudo-Riemannian metric (see [9]). In constrast with the sub-Riemannian case that is characterized by linearity in the control, time minimization (with a bounded control) and minimization of the L^2-norm (with a prescribed final time) of affine systems cease to be equivalent.

3 Integrability properties

As time minimizing curves of the control system (3), geodesics satisfy Pontrjagin maximum principle [5, 13]: If x is a shortest time trajectory generated by the optimal control $u:[0,t_f]\to\mathbf{R}^2$, there exist a nonpositive constant $p^0\le 0$ and a Lipschitzian lift $(x,p):[0,t_f]\to T^*\mathbf{S}^2$ of the trajectory to the cotangent bundle such that $(p^0,p)\neq(0,0)$ and

$$\dot{x}(t)=\frac{\partial H}{\partial p}(x(t),p(t),u(t)),\quad \dot{p}(t)=-\frac{\partial H}{\partial x}(x(t),p(t),u(t)),\quad t\in[0,t_f]\quad\text{(a.e.)},$$

where

$$H : T^*\mathbf{S}^2 \times \mathbf{R}^2 \to \mathbf{R}, \quad H(x,p,u) := p^0 + u_1 H_1(x,p) + u_2 H_2(x,p).$$

The $H_i : T^*\mathbf{S}^2 \to \mathbf{R}$ are the Hamiltonian lifts of the two vector fields (4),

$$H_i(x,p) := \langle p, F_i(x)\rangle, \quad i = 1,2.$$

Besides, the following maximization condition holds,

$$H(x(t),p(t),u(t)) = \max_{|v|\leq 1} H(x(t),p(t),v), \quad t \in [0,t_f] \quad \text{(a.e.)}$$

As a result, the Hamiltonian evaluated along the *extremal* (x,p,u) is almost everywhere equal to a constant; since the final time is free, this constant is zero. The previous relations are homogeneous in (p^0,p) and there are two cases: Either $p^0 = 0$ (*abnormal case*), or $p^0 < 0$ (*normal case*).

Lemma 3. *Abnormal trajectories are stationary equatorial curves.*

Proof. Assume $p^0 = 0$. Then $H = u_1 H_1 + u_2 H_2$ has to be zero and maximized along the extremal; necessarily, the two Lipschitz functions H_1 and H_2 must vanish identically on $[0,t_f]$ (the maximum would otherwise be positive). As a result, for any time t, $p(t)$ is orthogonal to the span at $x(t)$ of F_1 and F_2. If there exists $t \in [0,t_f]$ such that $\varphi(t) \neq \pi/2$, the span at such an $x(t)$ is the whole tangent space, so $p(t) = 0$. As p is solution of the linear differential equation

$$\dot{p}(t) = -\frac{\partial H}{\partial x}(x(t),p(t),u(t)) = -p(t)\left[F_1'(x(t)) + F_2'(x(t))\right],$$

p is identically zero, which contradicts $(p^0,p) \neq (0,0)$. So $\varphi \equiv \pi/2$; then

$$\dot{\theta}(t) = u_1(t)\sqrt{1/R(X)}\,|_{X=1} = 0$$

so θ is also constant. □

We disregard theses trivial curves and normalize p^0 to -1.

Scholium. *Geodesics of (1) are integral curves of the quadratic Hamiltonian*

$$h(\theta,\varphi,p_\theta,p_\varphi) := \frac{1}{2}\left(\frac{p_\theta^2}{XR(X)} + p_\varphi^2\right) \tag{7}$$

restricted to the level set $\{h = 1/2\}$.

Proof. Let u be a minimum time control, and let (x,p,u) be the associated extremal. The normal Hamiltonian $-1+u_1H_1+u_2H_2$ is zero along the extremal, so (H_1,H_2) does not vanish. Because of the maximization condition,

$$u(t) = \frac{(H_1,H_2)}{\sqrt{H_1^2+H_2^2}}(x(t),p(t)) \quad \text{a.e.}$$

Then,

$$\begin{aligned} \dot{x}(t) &= u_1(t)F_1(x(t)) + u_2(t)F_2(x(t)) \\ &= H_1(x(t),p(t))F_1(x(t)) + H_2(x(t),p(t))F_2(x(t)) \end{aligned} \tag{8}$$

since $H_1^2 + H_2^2 = 1$ on $\{H = 0\}$ (that is $\{h = 1/2\}$), so

$$\dot{x}(t) = \frac{\partial h}{\partial p}(x(t), p(t)) \quad \text{a.e.}$$

One similarly verifies that

$$\dot{p}(t) = -\frac{\partial h}{\partial x}(x(t), p(t)) \quad \text{a.e.}$$

□

Remark 4. Such geodesics are arc length parameterized as (8) implies that

$$g_{x(t)}(\dot{x}(t)) = u_1^2(t) + u_2^2(t) = 1 \quad \text{a.e.}$$

As θ is a cyclic variable (symmetry of revolution—see [7] for a general reference) of h, p_θ is a linear first integral and

Proposition 3. *h is Liouville integrable. The coordinate φ is parameterized by a hyperelliptic curve of genus at most p.*

Proof. $X = \sin^2\varphi$, so $\dot{X}^2 = 4X(1-X)\dot{\varphi}^2$. On $\{h = 1/2\}$,

$$\dot{X}^2 = 4X(1-X)\left(1 - \frac{p_\theta^2}{XR(X)}\right),$$

so

$$Y^2 = 4(1-X)(a_p + \cdots + a_0(1-X)^p)[X(a_p + \cdots + a_0(1-X)^p) - p_\theta^2(1-X)^p] \quad (9)$$

with $Y = (a_p + \cdots + a_0(1-X)^p)\dot{X}$. The right hand side of (9) has degree at most $2(p+1)$, so the genus of the complex curve is at most p. □

Set

$$\Gamma(\varphi) := \frac{1}{XR(X)}, \quad \varphi \in (0, \pi).$$

Consider the extremal departing from $\varphi_0 \neq 0\,(\pi)$ (not a pole), θ_0 being normalized to 0 and defined by a positive p_θ (the degenerate case $p_\theta = 0$ corresponding to meridians—which are the only extremals passing through the poles) and non-negative $p_{\varphi_0} = \sqrt{1 - \Gamma(\varphi_0)p_\theta^2}$. Along the extremal, $\dot{\varphi}$ first vanishes when φ is equal to $\varphi_1 := \pi - \Gamma^{-1}(p_\theta^{-2})$ since

Lemma 4. *Γ is a strictly decreasing one-to-one mapping between $(0, \pi/2]$ and $\mathbf{R}_+$ such that $\Gamma(\pi - \varphi) = \Gamma(\varphi)$.*

Proof. One has

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}X} = -\frac{R(X) + XR'(X)}{(XR(X))^2}$$

with

$$R'(X) = \sum_{k=1}^{p} \frac{ka_k}{(1-X)^{k+1}}.$$

Since $\mathrm{d}X/\mathrm{d}\varphi = 2\sin\varphi\cos\varphi$ is positive on $(0, \pi/2)$, the conclusion follows. □

As Γ' does not vanish on $(0, \pi/2)$,

$$1 - \Gamma(\varphi)p_\theta^2 = O(\varphi_1 - \varphi)$$

in the neighbourhood of $\pi - \Gamma^{-1}(p_\theta^{-2})$, and the following integral (depending on p_θ and φ_0) is well-defined,

$$t_1 := \int_{\varphi_0}^{\pi - \Gamma^{-1}(p_\theta^{-2})} \frac{\mathrm{d}\varphi}{\sqrt{1 - \Gamma(\varphi)p_\theta^2}}.$$

Lemma 5. *The axial symmetry σ_1 with respect to $\theta(t_1)$ is an inner symmetry of the extremal.*

Proof. Set

$$\begin{array}{rclrcl} \hat\theta(t) & := & 2\theta(t_1) - \theta(2t_1 - t), & \hat p_\theta(t) & := & p_\theta(t), \\ \hat\varphi(t) & := & \varphi(2t_1 - t), & \hat p_\varphi(t) & := & -p_\varphi(2t_1 - t), \end{array}$$

and check that new curve is still an extremal, passing through the same point of the cotangent bundle at t_1 since $p_\varphi(t_1) = 0$. □

Necessarily, $\pi - \Gamma^{-1}(p_\theta^{-2}) \geq \pi - \varphi_0$, so there also exists $t_1' \leq t_1$ such that $\varphi(t_1') = \pi - \varphi_0$. Using the previous axial symmetry, we deduce the existence of $t_2 := 2t_1 - t_1' \geq t_1$ such that, again, $\varphi(t_2) = \pi - \varphi_0$. Using now the equatorial symmetry of Γ,

$$\Gamma(\pi - \varphi) = \Gamma(\varphi),$$

the following is clear.

Lemma 6. *The central symmetry s_2 with respect to $(\theta(t_2)/2, \pi/2)$ defines another extremal with the same initial condition.*

Proof. Set

$$\begin{array}{rclrcl} \hat\theta(t) & := & \theta(t_2) - \theta(t_2 - t), & \hat p_\theta(t) & := & p_\theta(t), \\ \hat\varphi(t) & := & \pi - \varphi(t_2 - t), & \hat p_\varphi(t) & := & p_\varphi(t_2 - t). \end{array}$$

The new curve is still an extremal since

$$\dot{\hat\theta}(t) = \Gamma(\pi - \hat\varphi(t))p_\theta = \Gamma(\hat\varphi(t))\hat p_\theta, \quad \dot{\hat p}_\varphi(t) = \frac{1}{2}\Gamma'(\pi - \hat\varphi(t))p_\theta^2 = -\frac{1}{2}\Gamma'(\hat\varphi(t))\hat p_\theta^2,$$

and $\hat\theta_0 = 0 = \theta_0$, $\hat\varphi_0 = \pi - (\pi - \varphi_0) = \varphi_0$. □

Finally denote t_3 the point such that $\varphi(t_3) = \pi/2 \leq \pi - \varphi_0$, and remark that the central symmetry σ_2 with respect to $(\theta(t_3), \pi/2)$ leaves the extremal invariant. Since the axial symmetry s_1 with respect to $\theta = 0$ obviously defines another extremal originating from the same point, we conclude that the group generated by s_1 and s_2 acts on the set of extremals with same initial condition, while the group generated by σ_1 and σ_2 defines inner symmetries of each extremal. The composition rules indicate in both cases that the underlying group is the four-order abelian Klein group,

$$\mathbf{V} = \mathbf{Z}/2\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} \simeq \{\mathrm{id}, s_1, s_2, s_1 s_2\} \simeq \{\mathrm{id}, \sigma_1, \sigma_2, \sigma_1\sigma_2\}.$$

Proposition 4. *Given any initial condition, the Klein group acts on the set of extremals issuing from the point. It also defines inner symmetries of any extremal.*

An extremal is said to be a *pseudo-equator* whenever $\dot{\varphi}(0) = p_\varphi(0)$ is equal to zero, whereas the equator itself cannot be an extremal because of the singularity.

Lemma 7. *Every extremal that is not a meridian is a pseudo-equator.*

Proof. For p_θ positive and p_{φ_0} nonnegative (the other cases are deduced by symmetry), there exists $\widetilde{\varphi}_0 = \Gamma^{-1}(p_\theta^{-2})$ such that, up to a time shift, the extremal is the pseudo-equator with initial condition $\widetilde{\varphi}_0$. □

Conversely, any pseudo-equator meets $\varphi = \pi/2$ as one understands from the analysis of symmetries. Taking $\widetilde{\varphi}_0 = \pi/2$ as new initial condition and retaining the same value for p_θ provides the same geodesic, up again to a time shift. As a result, rather than parameterizing extremals using both φ_0—we set $\theta_0 = 0$ thanks to the symmetry of revolution—and p_θ, one may either parameterize by $\varphi_0 \in (0, \pi/2)$ alone using the fact all geodesics (with the exception of meridians, $p_\theta = 0$) are pseudo-equators (then, implicitly, $p_\theta^2 = 1/\Gamma(\varphi_0)$, $\varphi_0 \neq \pi/2$ since the equator is not a geodesic), or parameterize by their *Clairaut constant* $p_\theta \in \mathbf{R}_+^*$, considering only the initial condition at singularity, $\varphi_0 = \pi/2$. The second point of view reduces the study of geodesics to those starting at singularity.

Proposition 5. *On every extremal, the coordinate φ is periodic with period*

$$T(p_\theta) = 4 \int_{\Gamma^{-1}(p_\theta^{-2})}^{\pi/2} \frac{\mathrm{d}\varphi}{\sqrt{1 - \Gamma(\varphi) p_\theta^2}},$$

and $\theta(t + T) = \theta(t) \pm \Delta\theta$ (sign depending on the sign of p_θ) with quasi-period

$$\Delta\theta(p_\theta) = 4 \int_{\Gamma^{-1}(p_\theta^{-2})}^{\pi/2} \frac{\Gamma(\varphi) p_\theta \mathrm{d}\varphi}{\sqrt{1 - \Gamma(\varphi) p_\theta^2}}.$$

Proof. According to the previous analysis, it is enough to check the result on pseudo-equators. But then, $t_1 = t_2 = t_3 = 2t_4$, so setting $T := 2t_1$ and using the axial symmetry with respect to $\theta(t_1)$ gives the result since $\varphi(T) = \varphi(0)$, $p_\varphi(T) = -p_\varphi(0) = 0 = p_\varphi(0)$. So $\dot{\theta} = \Gamma(\varphi) p_\theta$ is also periodic, which concludes the proof. □

As functions of φ_0,

$$T(\varphi_0) = 4 \int_{\varphi_0}^{\pi/2} \frac{\mathrm{d}\varphi}{\sqrt{1 - \Gamma(\varphi)/\Gamma(\varphi_0)}}, \tag{10}$$

and

$$\Delta\theta(\varphi_0) = 4 \int_{\varphi_0}^{\pi/2} \frac{\Gamma(\varphi) \mathrm{d}\varphi}{\sqrt{\Gamma(\varphi_0) - \Gamma(\varphi)}}. \tag{11}$$

These relations actually cover the case of meridians $p_\theta = 0$ (*i.e.* $\varphi_0 = 0$) for which $T = 2\pi$ and $\Delta\theta = 2\pi$ (two instantaneous rotations of angle π when crossing poles at $t = \pi$ and $t = 2\pi$). We end the section with the following result that will be used for §5 computations.

Proposition 6. $T' = p_\theta \Delta\theta'$.

Proof. Write as in [11]

$$T(p_\theta) = p_\theta \Delta\theta(p_\theta) + 4\int_{\Gamma^{-1}(p_\theta^{-2})}^{\pi/2} \sqrt{1-\Gamma(\varphi)p_\theta^2}\,\mathrm{d}\varphi,$$

so

$$\begin{aligned} T'(p_\theta) &= \Delta\theta(p_\theta) + p_\theta\Delta\theta'(p_\theta) + 4\int_{\Gamma^{-1}(p_\theta^{-2})}^{\pi/2} \frac{\partial}{\partial p_\theta}\left(\sqrt{1-\Gamma(\varphi)p_\theta^2}\right)\mathrm{d}\varphi \\ &\quad -8(\Gamma^{-1})'(p_\theta^{-2})p_\theta \underbrace{\sqrt{1-\Gamma(\varphi)p_\theta^2}\,|_{\varphi=\Gamma^{-1}(p_\theta^{-2})}}_{0} \\ &= p_\theta\Delta\theta'(p_\theta). \end{aligned}$$

□

4 Cut and conjugate loci

The *cut time* along a geodesic x, that is along an extremal of the minimum time problem (3), is the supremum of times t such that the curve restricted to $[0,t]$ is a shortest time trajectory between $x(0)$ and $x(t_f)$:

$$t_{\mathrm{cut}} := \sup\{t \geq 0 \mid x \text{ is minimizing on } [0,t]\}.$$

When $t_{\mathrm{cut}} < \infty$, $x(t_{\mathrm{cut}})$ is called a cut point. The set of all cut points on geodesics departing from a given intial point x_0 is the *cut locus* of x_0. A *separating point* along the geodesic is a point $x(t_M)$, $t_M > 0$, such that there exists a different geodesic, y, reaching the point at the same time: $x(t_M) = y(t_M)$. The exponential mapping of a fixed point $x_0 \in \mathbf{S}^2$ is

$$\exp_{x_0} : \mathbf{R}_+^* \times T_{x_0}^*\mathbf{S}^2 \cap \{h = 1/2\} \to \mathbf{S}^2 \tag{12}$$

$$(t, p_0) \mapsto x(t, x_0, p_0) := \Pi \circ \exp t\overrightarrow{h}\,(x_0, p_0)$$

where $\Pi : T^*\mathbf{S}^2 \to \mathbf{S}^2$ is the canonical projection and $\exp t\overrightarrow{h}$ the one-parameter global subgroup generated by the symplectic gradient of the quadratic Hamiltonian (7),

$$\overrightarrow{h}\,(x,p) = \frac{\partial h}{\partial p}(x,p)\frac{\partial}{\partial x} - \frac{\partial h}{\partial x}(x,p)\frac{\partial}{\partial p}.$$

The intersection of the fibre $T_{x_0}^*\mathbf{S}^2$ with $\{h = 1/2\}$ is a compact oval diffeomorphic to $\mathbf{S}^1$ outside the equator, or the union two lines $\{p_{\varphi 0} = \pm 1\}$ for an equatorial point. That the subgroup is globally defined in the second case comes from the analysis of Section 3: For any $p_\theta \in \mathbf{R}$ and $p_{\varphi 0} = \pm 1$, the coordinates φ and θ are periodic or quasi-periodic, respectively. A *conjugate point* along the geodesic x is a critical value of $\exp_{x(0)}$; if (t_c, p_0) is the corresponding critical point, t_c is the *conjugate time*. When $t_c > 0$ is the first conjugate time along the geodesic, $x(t_c)$ is called a *first conjugate point*. The set of first conjugate points on geodesics departing from x_0 is the *conjugate locus* of x_0. Results of

optimal control ensure that local optimality is lost after conjugate points [33]. Besides, extremals of such problems have to be smooth so that broken curves that are concatenations of minimizing geodesics cannot be minimizing, entailing that optimality cannot be preserved after separating points, such as those generated by the symmetries between extremals described in the previous section. Separating and conjugate times are so upper bounds of cut times. The following standard result remains valid in our setting with singularities.

Proposition 7. *Cut points of the metric (1) are either conjugate or separating points.*

Lemma 8. *Both T and $\Delta\theta$ vanish when $|p_\theta| \to \infty$.*

Proof. Directly follows from estimates of integrals (10-11) using the fact that Γ does not vanish identically at $\varphi = \pi/2$. □

Proof of Proposition 7. Let $\gamma(t_{\text{cut}})$ be the cut point along the geodesic γ starting from $x_0 = (\theta_0, \varphi_0)$ and generated by the adjoint vector p_0. If the cut point is not a conjugate point, the exponential mapping is a diffeomorphism in a small enough neighbourhood V_0 of (t_{cut}, p_0). Since the metric is complete, there are minimizing extremals γ_n joining x_0 to $\gamma_n(t_n) = \gamma(t_{\text{cut}} + 1/n)$, $t_n < t_{\text{cut}} + 1/n$, for $n \geq 1$. First assume that $\varphi_0 \neq \pi/2$. Then the oval $h^{-1}(\theta_0, \varphi_0, \cdot)(\{1/2\})$ is compact, and one can extract a converging subsequence of the $(p_{0_n})_n$ generating the extremals γ_n, and thus conclude classically (see, *e.g.*, [32]): $p_{0_n} \to \widetilde{p}_0$ and $\gamma_n(t_n) \to \gamma(t_{\text{cut}})$; assuming by contradiction that the point is not a separating one, that is assuming that $\widetilde{p}_0 = p_0$, implies that for n large enough $\gamma_n(t_n)$ belongs to $\exp_{x_0}(V_0)$, so that $t_n = t_{\text{cut}} + 1/n$, whence the contradiction. Let now $\varphi_0 = \pi/2$. Though $h^{-1}(\theta_0, \varphi_0, \cdot)(\{1/2\}) = \{p_{\varphi_0} = \pm 1\}$ is not compact anymore, the associated sequence $(p_{\theta_n})_n$ still has to be bounded otherwise there would exist a subsequence such that $|p_{\theta_n}| \to \infty$. Assume by contradiction that this is the case. Because of the central symmetry s_2, $(\Delta\theta(p_{\theta_n})/2, \pi/2)$ is a separating point on $\gamma_n = (\theta_n, \varphi_n)$, so $t_n < T(p_{\theta_n})/2$ and $\theta_n(t_n) < \Delta\theta(p_{\theta_n})/2$; as $|\dot{\varphi}_n| = |p_{\varphi_n}| \leq 1$, $|\varphi_n(t_n) - \pi/2| \leq T(p_{\theta_n})/2$ and $(\theta_n(t_n), \varphi_n(t_n)) \to (0, \pi/2)$ according to Lemma 8. Since $\gamma(t_{\text{cut}} + 1/n) = \gamma_n(t_n)$, this implies $\gamma(t_{\text{cut}}) = x_0$, which is contradictory. The sequence $(p_{\theta_n})_n$ is hence bounded, and we can conclude as previously. □

Cut loci for an analytical Riemannian metric on the sphere are known to be finite trees whose extremities are conjugate points after the work of Poincaré [28, 29, 31]. In our case, cut loci have peculiarities due to the symmetry of revolution and the equatorial singularity (see also Corollary 3 in §6; more generally on surface of revolutions, see [34]).

Theorem 1. *Under the assumption that $\Delta\theta$ is strictly decreasing for $p_\theta > 0$, cut loci of the metric (1) are antipodal subarcs. The cut locus of a pole is reduced to the opposite pole, is equal to the equator minus the point for an equatorial point, and to a proper closed subarc of the antipodal parallel otherwise.*

Proof. The case of poles is obvious and does not depend on any assumption on $\Delta\theta$ since the only extremals through them are meridians. Consider now the situation $\varphi_0 = \pi/2$, and show that the exponential mapping is injective on the quadrant

$$\bigcup_{p_\theta > 0} (0, T(p_\theta)/2] \times \{(p_\theta, 1)\},$$

that is show that subarcs of extremals defined by $t \in [0, T(p_\theta)/2]$, positive p_θ and $p_\varphi = +1$ do not intersect. If $p'_\theta > p_\theta$, the arc associated with p'_θ is strictly below the one associated with p_θ. Indeed, note that on the first half of such an arc $t \in [0, T(p_\theta)/4)$ and $\dot{\varphi}$ does not vanish so that the curve can be parameterized by φ instead of time. There,

$$f(\varphi, p_\theta) := \frac{\mathrm{d}\theta}{\mathrm{d}\varphi} = \frac{\Gamma(\varphi) p_\theta}{\sqrt{1 - \Gamma(\varphi) p_\theta^2}}$$

is an increasing function of p_θ since

$$\frac{\partial f}{\partial p_\theta}(\varphi, p_\theta) = \frac{\Gamma(\varphi)}{(1 - \Gamma(\varphi) p_\theta^2)^{3/2}} > 0.$$

As geodesics starting from $\varphi_0 = \pi/2$ cross again the equator at $\Delta\theta(p_\theta)/2$, the assumption ensures that the aforementioned subarcs do not intersect. We conclude by remarking that the full set of extremals is obtained by considering the action of the Klein group on geodesics with same initial condition (see §3). First, the central symmetry s_2 that generates intersections at $t = T(p_\theta)/2$, then the axial symmetry s_1 with respect to $\theta = 0$ that generates intersections at $\theta = \pi$, thus not prior to the previous ones since $\theta(T(p_\theta)/2) = \Delta\theta(p_\theta)/2$, and since $\Delta\theta(p_\theta) < 2\pi$ for $p_\theta > 0$ (by assumption, $\Delta\theta$ is decreasing, and equal to 2π on meridians, *i.e.* when $p_\theta = 0$). So extremals are optimal up to $t = T(p_\theta)/2$, and the corresponding point is a separating point. Since the metric is complete, each point of the equator is reached by such an extremal and the set of separating points, hence the cut locus, is the equator minus the initial point itself. Consider finally the case when the initial point is neither polar nor equatorial. Then p_θ^2 belongs to $(0, 1/\Gamma(\varphi_0))$, and extremals are again optimal up to $t = T(p_\theta)/2$. Indeed, there would otherwise exist shorter extremals, which would lead to the existence of shorter extremals for the initial condition $\varphi_0 = \pi/2$ too, contradicting the previous fact. The central symmetry s_2 still generates an intersection at $t = T(p_\theta)/2$, and $\varphi(T(p_\theta)/2) = \pi - \varphi_0$ so the corresponding separating point belongs to the antipodal parallel of the starting point. Since $\Delta\theta$ is decreasing, the extremities of the cut are obtained letting p_θ tend to $\pm(\Gamma(\varphi_0))^{-1/2}$ (now finite, since $\varphi_0 \neq \pi/2$), and the subarc is closed. □

To study the conjugate loci, we start with some properties of the local model at singularity. Setting $x := \pi/2 - \varphi$ and $y := \theta$, since $1 - \sin^2\varphi \sim (\pi/2 - \varphi)^2$ when φ tends to $\pi/2$, a local model for the metric (1) is

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \frac{\mathrm{d}y^2}{x^{2p}} \tag{13}$$

where p is the order of the pole. The equatorial symmetry of Γ translates into $(-x)^{2p} = x^{2p}$, so the discrete symmetry group is preserved. Such almost-Riemannian metrics are related to sub-Riemannian distributions. For $p = 1$, the local model is the Grušin metric $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2/x^2$, which is actually obtained by projecting the *Heisenberg* sub-Riemannian distribution [13]. This distribution is indeed defined, up to a renormalization, by the following two vector fields on $\mathbf{R}^3$,

$$F_1(x, y, z) := \frac{\partial}{\partial x} - y\frac{\partial}{\partial z}\,, \quad F_2(x, y, z) := \frac{\partial}{\partial x} + x\frac{\partial}{\partial z}\,,$$

and the corresponding sub-Riemannian Hamiltonian is

$$H(x,y,z,p_x,p_y,p_z) := \frac{1}{2}\left[(p_x^2+p_y^2)+2p_z(xp_y-yp_x)+(x^2+y^2)p_z^2\right],$$

which suggests to use cylindrical coordinates. In these variables

$$H(r,\theta,z,p_r,p_\theta,p_z) = \frac{1}{2}\left[p_r^2+(p_\theta/r+rp_z)^2\right].$$

As θ and z are cyclic, the system is integrable in dimension three, and projects onto a Hamiltonian in the (r,z)-space with the desired singularity,

$$h(r,z,p_r,p_z) := \frac{1}{2}(p_r^2+r^2p_z^2),$$

when restricting to $p_\theta = 0$. For $p=2$, the local model is $\mathrm{d}s^2 = \mathrm{d}x^2+\mathrm{d}y^2/x^4$, which is connected to the *flat Martinet* sub-Riemannian distribution. Consider indeed the two vector fields on $\mathbf{R}^3$ (see [2])

$$F_1(x,y,z) := \frac{\partial}{\partial x}+y^2\frac{\partial}{\partial z}, \quad F_2(x,y,z) := \frac{\partial}{\partial y},$$

so the sub-Riemannian Hamiltonian is

$$H(x,y,z,p_x,p_y,p_z) := \frac{1}{2}\left[p_y^2+(p_x+y^2p_z)^2\right].$$

The two coordinates x and z are cyclic, and the Hamiltonian projects onto $h(y,z,p_y,p_z) := (1/2)(p_y^2+y^4p_z^2)$ in the (y,z)-space when restricting to $p_x=0$, providing the higher order singularity. Going back to the general case, we compute geodesics issuing from the origin on the level set $\{h=1/2\}$ of the Hamiltonian

$$h(x,y,p_x,p_y) := \frac{1}{2}(p_x^2+x^{2p}p_y^2)$$

so that the initial adjoint state belongs to the union of the two lines, $\{p_x=\pm 1\}$. We set $\lambda := p_y$ and restrict to positive λ by symmetry (the trivial geodesics $(\pm t,0)$ being obtained for $\lambda=0$). The coordinate x is then

$$x(t) = \frac{1}{\sqrt[p]{\lambda}}q(t\sqrt[p]{\lambda}), \tag{14}$$

where q is the solution of

$$q'^2+q^{2p}=1, \quad q(0)=0, \quad q'(0)=1. \tag{15}$$

Equivalently,

$$q^{-1}(u) = \int_0^u \frac{\mathrm{d}v}{\sqrt[p]{1-v^{2p}}}, \quad u\in[-1,1].$$

For $p=1$, q is harmonic, elliptic for $p=2$, hyperelliptic and reciprocal to a hypergeometric function in general. More precisely,

$$q^{-1}(u) = {}_2\mathbf{F}_1(1/2,1/(2p);1+1/(2p);u^{2p})\cdot u \tag{16}$$

where ${}_2\mathbf{F}_1(a,b;c;z)$ is the hypergeometric series

$$ {}_2\mathbf{F}_1(a,b;c;z) = \sum_{n\geq 0} \alpha_n \frac{z^n}{n!}\,, $$

and

$$ \begin{array}{rcl} \alpha_0 & := & 1, \\ \alpha_{n+1}/\alpha_n & := & (n+a)(n+b)/(n+c). \end{array} $$

The reciprocal of q is hence equal to

$$ q^{-1}(u) = \sum_{n\geq 0} \alpha_n \frac{u^{2np+1}}{n!} $$

with $\alpha_n = (a)_n(b)_n/(c)_n$ for $a = 1/2$, $b = 1/(2p)$ and $c = 1 + 1/(2p)$, the notation $(a)_n$ standing for the Pochhammer symbol

$$ (a)_n := a(a+1)\cdots(a+n-1). $$

Here,

$$ \alpha_0 = 1, \quad \alpha_1 = \frac{1}{2}\cdot\frac{1}{2p+1}\,, \quad \alpha_2 = \frac{3}{4}\cdot\frac{2p+1}{8p^2+6p+1}\cdots $$

which gives the usual Taylor series of the reciprocal of the sine function for $p = 1$, $\arcsin u = u + u^3/6 + 3u^5/40 + 5u^7/112...$ Eventually, $\dot{y} = \lambda x^{2p}$, so

$$ y(t) = \frac{1}{(\sqrt[p]{\lambda})^{p+1}} r(t\sqrt[p]{\lambda}), \tag{17} $$

where r is defined by a second quadrature,

$$ r(s) := \int_0^s q^{2p}. $$

Lemma 9.

$$ r = \frac{1}{p+1}(s - qq'). $$

Proof. Differentiating and using (15),

$$ \begin{array}{rcl} \dfrac{1}{p+1}(s-qq')' & = & \dfrac{1}{p+1}(1 - q'^2 - qq'') \\ & = & q^{2p}, \end{array} $$

hence the result since $r(0) = 0$. □

For any $p \geq 1$, symmetry reasons imply that non-trivial minimizing geodesics emanating from the origin first intersect on the y-axis (see, *e.g.*, Fig. 1). As a consequence, the cut locus at the origin of the local model is the axis minus the origin itself (compare with the metric on the sphere, Theorem 1). The conjugate locus of the origin is obtained from the set of critical values of the exponential mapping

$$ \exp_{(0,0)}(t,\lambda) = (x(t,\lambda), y(t,\lambda)), \quad (t,\lambda) \in (\mathbf{R}_+^*)^2. $$

Let $t_{1c}(\lambda)$ denote the first conjugate time along the geodesic defined by $\lambda > 0$.

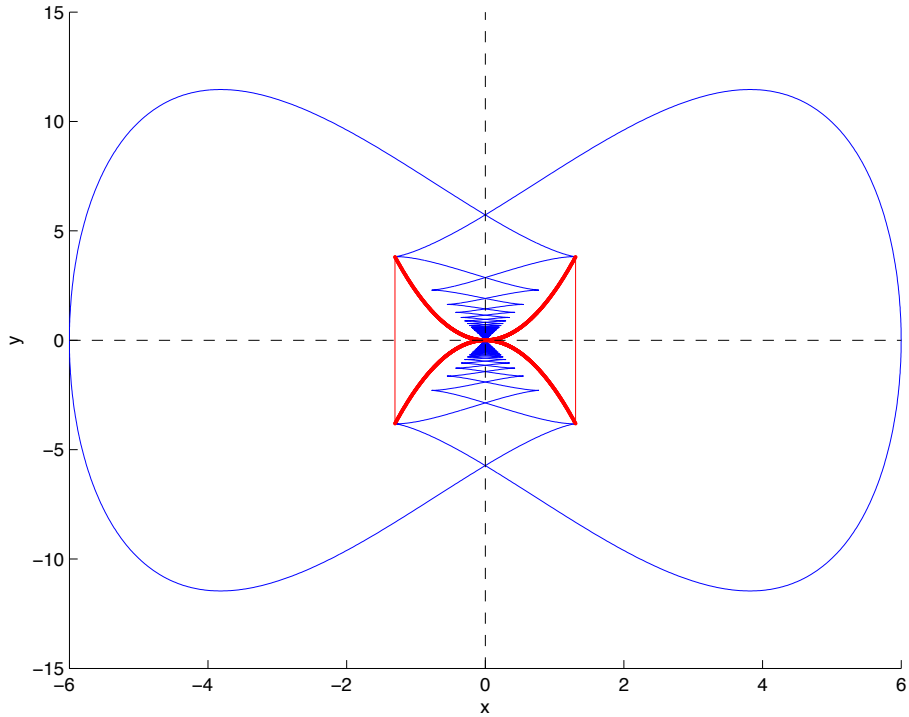

Figure 1: Grušin metric, $\mathrm{d}x^2 + \mathrm{d}y^2/x^2$ ($p = 1$); sphere, wavefront and conjugate locus of the origin. The wavefront (in blue) is the image of the exponential mapping for a fixed time, t; it is made of endpoints at time t of geodesics. The subset obtained by ruling out the part after the two first self-intersection points of the wavefront on the y-axis is the sphere; it is made of points at distance (minimum time) t of the origin. The conjugate locus (in red) is partly drawn (here, it is the set $y = \pm C_1 x^2$ minus the origin); it is made of critical values of the exponential (it contains so the first singularity of the wavefront portrayed). Compare with the Heisenberg metric in [21].

Lemma 10. *$t_{1c}(\lambda) = s_p/\sqrt[p]{\lambda}$ where s_p is the first positive root of the envelope equation $q = sq'$.*

Proof. Using (14) and (17), a pair (t, λ) is a critical point of the exponential if and only if $t\sqrt[p]{\lambda}$ is a solution to

$$qr' - (p+1)q'r = 0.$$

Expressing r according to Lemma 9, one gets the result. □

Proposition 8. *The conjugate locus at the origin of $\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2/x^{2p}$ is the set $y = \pm C_p x^{p+1}$ minus the origin where*

$$C_p \sim \frac{2\sqrt{2}}{p}, \quad p \to \infty.$$

Proof. According to Lemma 10,

$$x(t_{1c}(\lambda)) = \frac{1}{\sqrt[p]{\lambda}} q(s_p), \quad y(t_{1c}(\lambda)) = \frac{1}{(\sqrt[p]{\lambda})^{p+1}} r(s_p),$$

so points in the conjugate locus lie on the curve $y = \pm C_p x^{p+1}$ where $C_p = r(s_p)/q^{p+1}(s_p)$. The whole curve minus $(0,0)$ is obtained because of symmetries, and conjugate points accumulate towards the origin when $|\lambda| \to \infty$. It is clear from (15) that, as $p \to \infty$, q pointwisely converges towards the Lipschitz function equal to $s \mapsto s$ on $[0,1]$, $s \mapsto 2-s$ on $[1,3]$. In particular, the solution s_p to the envelope equation is such that $s_p \to 3-$; so $q'(s_p) \to -1/3$, $q^{2p}(s_p) = 1 - q'^2(s_p) \to 8/9$, and $q^{p-1}(s_p) \to 2\sqrt{2}/3$. Now, by virtue of Lemmas 9 and 10,

$$C_p = \frac{1}{p+1} \frac{s_p - q(s_p)q'(s_p)}{q^{p+1}(s_p)} = \frac{s_p}{p+1} q^{p-1}(s_p) \sim \frac{2\sqrt{2}}{p}, \quad p \to \infty.$$

□

Remark 5. For $p = 1$, $q(s) = \sin s$, $r(s) = (1/2)(s - \sin t \cos t)$; $t_{1c}(\lambda) = s_1/\lambda$ where s_1 is the first positive root of

$$\sin s = s \cos s.$$

For $p = 2$, one obtains

$$q(s) = \frac{\sqrt{2}}{2} \frac{\mathrm{sn}}{\mathrm{dn}}(s\sqrt{2}), \quad r(s) = \frac{1}{3}\left[s + \frac{\sqrt{2}}{2} \frac{\mathrm{cn}\,\mathrm{sn}}{\mathrm{dn}^3}(s\sqrt{2}) \right],$$

in terms of Jacobi functions of modulus $k = \sqrt{2}/2$; $t_{1c}(\lambda) = s_2/\sqrt{\lambda}$ where $\tau = s_2\sqrt{2}$ (with $s_2 < 3$) is the first root solution of

$$\mathrm{sn}\,\tau\,\mathrm{dn}\,\tau = \tau\,\mathrm{cn}\,\tau.$$

(Compare with the flat Martinet case in [2].)

Before stating the structure result on the conjugate locus, we finally recall the following.

Scholium. *Along a Jacobi field tangent to the level set of a Hamiltonian quadratic in the momentum, the Liouville form is constant.*

Proof. In coordinates $z = (x,p) \in \mathbf{R}^{2n}$, let $H(x,p) = (1/2)(A(x)p|p)$ with $A(x)$ symmetric; let $\gamma(t) = (z(t), \delta z(t))$ be a Jacobi field, that is $\dot{z}(t) = \overrightarrow{H}(z(t))$ and

$$\delta\dot{z}(t) = \overrightarrow{H}'(z(t))\delta z(t).$$

Along γ, the time derivative of the Liouville form $p\,\mathrm{d}x$ is

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}(p|\delta x) &= (\dot{p}|\delta x) + (p|\delta\dot{x}) \\ &= -(\nabla_x H|\delta x) + \underbrace{(p|\nabla^2_{xp} H\,\delta x)}_{2(\nabla_x H|\delta x)} + \underbrace{(p|\nabla^2_{pp} H\,\delta p)}_{(\nabla_p H|\delta p)} \\ &= (\nabla_x H|\delta x) + (\nabla_p H|\delta p). \end{aligned}$$

Let now $z_0(\sigma)$ be a local parameterization of the level set $\{H = c\}$; for a Hamiltonian curve $z(t, z_0(\sigma))$ with initial condition $z_0(\sigma)$, $H(z(t, z_0(\sigma))) = c$, so

$$H'(z(t, z_0(\sigma)))\underbrace{\frac{\partial z}{\partial z_0}(t, z_0(\sigma))z_0'(\sigma)}_{=:\delta z(t)} = 0.$$

Along the curve $(z(t, \sigma), \delta z(t))$, the Liouville form is hence constant. □

Here, the level set $\{h = 1/2\}$ of the quadratic Hamiltonian (7) on $\mathbf{S}^2$ is locally parameterized by p_θ and the exponential (12) writes

$$\exp_{\varphi_0}(t, p_\theta) = (\theta(t, p_\theta), \varphi(t, p_\theta)),$$

so

$$p_\theta \frac{\partial \theta}{\partial p_\theta}(t, p_\theta) + p_\varphi(t, p_\theta)\frac{\partial \varphi}{\partial p_\theta}(t, p_\theta) = 0 \tag{18}$$

since $\partial\theta(0, p_\theta)/\partial p_\theta = \partial\varphi(0, p_\theta)/\partial p_\theta = 0$.

Lemma 11. *Critical points (t, p_θ) of the exponential are characterized either by $\partial\theta(t, p_\theta)/\partial p_\theta = 0$ when $\dot\varphi \neq 0$, or by $\partial\varphi(t, p_\theta)/\partial p_\theta = 0$ when $p_\theta \neq 0$.*

Proof. The point (t, p_θ) is critical if and only if

$$\dot\theta(t, p_\theta)\frac{\partial \varphi}{\partial p_\theta}(t, p_\theta) - \dot\varphi(t, p_\theta)\frac{\partial \theta}{\partial p_\theta}(t, p_\theta) = 0.$$

When $p_\varphi(t, p_\theta) = \dot\varphi(t, p_\theta) \neq 0$, one can multiply both sides by p_φ and use (18) to get

$$\frac{\partial \theta}{\partial p_\theta}(t, p_\theta)\underbrace{\Big(p_\theta\dot\theta(t, p_\theta) + p_\varphi(t, p_\theta)\dot\varphi(t, p_\theta)\Big)}_{=h=1/2} = 0$$

whence the result. Similar computation when $p_\theta \neq 0$. □

Theorem 2. *Under the assumption that $\Delta\theta$ is strictly decreasing and convex for $p_\theta > 0$, conjugate loci of the metric (1) are reduced to the opposite pole for poles, have four cusps otherwise.*

Proof. Let $\varphi_0 = \pi/2$. Consider an extremal defined by a positive p_θ and $p_{\varphi_0} = +1$. For t in $(T(p_\theta)/4, 3T(p_\theta)/4)$, $\dot\varphi \neq 0$ and the extremal can be parameterized by φ according to

$$\theta(\varphi, p_\theta) = \frac{\Delta\theta(p_\theta)}{2} + \int_\varphi^{\pi/2} f(\phi, p_\theta)\,\mathrm{d}\phi,$$

where, as before,

$$f(\varphi, p_\theta) = \frac{\mathrm{d}\theta}{\mathrm{d}\varphi} = \frac{\Gamma(\varphi)p_\theta}{\sqrt{1 - \Gamma(\varphi)p_\theta^2}}.$$

The conjugacy condition is $\partial\theta/\partial p_\theta = 0$ (Lemma 11) so the coordinate $\varphi_{1c}(p_\theta)$ of the first conjugate point is solution of

$$\int_{\varphi_{1c}(p_\theta)}^{\pi/2} \frac{\partial f}{\partial p_\theta}(\varphi, p_\theta)\,\mathrm{d}\varphi = -\frac{\Delta\theta'(p_\theta)}{2} > 0,$$

in order that $\varphi_{1c}(p_\theta) < \pi/2$ (since $\partial f/\partial p_\theta > 0$). By differentiating the previous equality, one gets

$$\varphi_{1c}'(p_\theta) = \left[\frac{\partial f}{\partial p_\theta}(\varphi_{1c}(p_\theta), p_\theta)\right]^{-1} \left[\frac{\Delta\theta''}{2}(p_\theta) + \int_{\varphi_{1c}(p_\theta)}^{\pi/2} \frac{\partial^2 f}{\partial p_\theta^2}(\varphi, p_\theta)\,\mathrm{d}\varphi\right],$$

which is positive first because $\varphi_{1c}(p_\theta) < \pi/2$, then because

$$\frac{\partial^2 f}{\partial p_\theta^2} = \frac{3\Gamma^2(\varphi)p_\theta}{(1-\Gamma(\varphi)p_\theta^2)^{5/2}} > 0,$$

and by virtue of the nonnegativeness of $\Delta\theta''$. In particular, the parameterization $p_\theta \mapsto (\theta(\varphi_{1c}(p_\theta), p_\theta), \varphi_{1c}(p_\theta))$ of the conjugate locus is regular, and the set has no cusp for positive p_θ. The tangent vector along the conjugate locus is

$$\left(\frac{\partial\theta}{\partial\varphi}\varphi_{1c}' + \underbrace{\frac{\partial\theta}{\partial p_\theta}}_{0}\right)\frac{\partial}{\partial\theta} + \varphi_{1c}'\frac{\partial}{\partial\varphi},$$

which is proportional to $f(\varphi_{1c}(p_\theta), p_\theta)\partial/\partial\theta + \partial/\partial\varphi$. On the one hand, as in the regular case discussed in [11], $f(\varphi_{1c}(p_\theta), p_\theta)$ tends to 0 when $p_\theta \to 0+$, and the locus has a first meridional cusp because of the axial symmetry s_1. On the other hand, when p_θ goes to $+\infty$, the analysis on the local model (Proposition 8) shows that, contrary to the regular case, the conjugate locus has two tangential contacts with the meridian that, combined again with symmetry s_1, form a second meridional cusp at the initial point where conjugate points accumulate. The central symmetry s_2 gives the symmetric part, whence the result. When $\varphi_0 \neq \pi/2$, the same reasoning shows that there is no cusp for $p_\theta \in (0, 1/\sqrt{\Gamma(\varphi_0)})$; there are two meridional cusps due again to symmetry s_1, and two cusps tangent to the antipodal parallel containing the cut locus (see Theorem 3.6 of [11]). □

5 The $(1, 2, 1)$ case

The results of Section 4 are applied to the second example discussed in §3 (the application to the first example, the Grušin metric on the sphere, is straightforward; see [18]). Consider the metric (1) with $p = 2$ and $(a_0, a_1, a_2) = (1/4)(1, 2, 1)$, that is

$$R(X) = \left(\frac{1 - X/2}{1 - X}\right)^2 = \frac{1}{4}\left[1 + \frac{2}{1-X} + \frac{1}{(1-X)^2}\right].$$

Lemma 12. *The coordinate φ is parameterized by the family of elliptic curves*

$$\mathscr{C}_{p_\theta} : Y^2 = 4(1-X)[X(2-X)^2 - 4p_\theta^2(1-X)^2], \quad p_\theta \in \mathbf{R}. \tag{19}$$

Proof. Set $Y := (2 - X)\dot{X}$ on the level set 1/2 of the quadratic Hamiltonian (7). □

In the sequel, $\wp$ denotes the Weierstraß function with invariants g_2, g_3 (see [25]).

Proposition 9.

$$\sin^2\varphi = \frac{\wp(z) - 4/3}{\wp(z) - 1/3}\,, \quad z \in \mathbf{R},$$

$$t = z + \frac{1}{\wp'(a)}\left[2\zeta(a)z + \ln\frac{\sigma(z-a)}{\sigma(z+a)}\right]_0^z$$

with a such that $\wp(a) = 1/3$ and invariants

$$g_2 = \frac{16}{3} + 16p_\theta^2, \quad g_3 = \frac{64}{27} - \frac{16}{3}p_\theta^2. \tag{20}$$

Proof. The rational transform

$$u = \frac{1}{3} + \frac{1}{1-X}\,, \quad v = u^2 Y,$$

sends to infinity the fixed root $X = 1$ in the right hand side of (19), and allows to recast the equation of $\mathscr{C}_{p_\theta}$ under the canonical form $v^2 = 4u^3 - g_2 u - g_3$ with invariants (20). When parameterizing the elliptic curve $\mathscr{C}_{p_\theta}$ by the Weierstraß function, $(u, v) = (\wp(z), \wp'(z))$, only the unbounded component of the real cubic has to be used since $X = \sin^2\varphi \in (0, 1]$ (that is $\wp(z) > 4/3$), so $z \in \mathbf{R}$. In this parameterization, the change of time from z to t verifies

$$\frac{\mathrm{d}t}{\mathrm{d}z} = 1 + \frac{1}{\wp(z) - 1/3} > 0. \tag{21}$$

Introducing Weierstraß functions ζ and σ ($\wp = -\zeta'$ and $\zeta = \sigma'/\sigma$), one has (see [23])

$$\int \frac{\wp'(a)\,\mathrm{d}z}{\wp(z) - \wp(a)} = 2\zeta(a)z + \ln\frac{\sigma(z-a)}{\sigma(z+a)}.$$

For a geodesic originating from the singularity, $\varphi = \pi/2$ at $t = 0$ which corresponds to $z = 0$. □

As a function of z, the coordinate φ is a doubly periodic meromorphic function. Its lattice of periods $2\omega\mathbf{Z} + 2\omega'\mathbf{Z}$ depends on p_θ (the Weierstraß half-periods ω, ω' are functions of p_θ) and is real rectangular: $\omega \in \mathbf{R}$, $\omega' \in i\mathbf{R}$, and the periods can be choosen so that their ratio $\tau := \omega'/\omega$ belongs to the Poincaré upper half-plane $\mathbf{H} = \{x + iy \in \mathbf{C},\ y > 0\}$. Lattices are classified up to conformal transformations; these transformations are Möbius transforms in the Fuschian modular subgroup $\mathrm{PSL}(2, \mathbf{Z}) = \mathrm{SL}(2, \mathbf{Z})/\pm\mathrm{id}$ of automorphisms of $\mathbf{H}$, so the moduli space of congruences of lattices is $\mathbf{H}/\mathrm{PSL}(2, \mathbf{Z})$. The modular function j establishes a one-to-one correspondence between these moduli and the complex plane; in terms of the invariants of the elliptic curve,

$$j(\tau) = \frac{g_2^3}{\Delta}$$

where $\Delta = g_2^3 - 27g_3^2$ is the discriminant of the elliptic curve with ratio of periods τ. For the family of elliptic curves $\mathscr{C}_{p_\theta}$ (see (20)),

$$j(\tau(p_\theta)) = \frac{16(1 + 3p_\theta^2)^3}{27p_\theta^2(8 + 13p_\theta^2 + 16p_\theta^4)}. \tag{22}$$

Corollary 1. *The number of conformal classes associated with $\mathscr{C}_{p_\theta}$ is*

- *equal to* 1 *for* $p_\theta \in (0, 1/2) \cup \{2/3\}$,
- *equal to* 2 *for* $p_\theta \in \{1/2, \sqrt{2}\}$,
- *equal to* 3 *for* $p_\theta \in (1/2, 2/3) \cup (2/3, \sqrt{2}) \cup (\sqrt{2}, \infty)$.

The only square lattice is obtained for $p_\theta = 2/3$.

Proof. The rational fraction (22) has exactly one global minimum at $p_\theta = 2/3$ and one local maximum at $p_\theta = \sqrt{2}$. Moreover,

$$j(\tau(p_\theta)) = \frac{28}{27} - \frac{16(p_\theta^2 - 1/4)(p_\theta^2 - 2)^2}{27p_\theta^2(8 + 13p_\theta^2 + 16p_\theta^4)}$$

showing that the value of the local maximum is also attained when $p_\theta = 1/2$. □

Proposition 10.

$$\frac{T(p_\theta)}{4} = (1 + \beta)\omega - b\eta \mod \frac{\pi}{2}$$

where $\eta := \zeta(\omega)$, $b := \operatorname{Im} a$ *and* $\beta := \operatorname{Im} \zeta(a)$.

Lemma 13. *For all* $p_\theta > 0$, $(1/3, \pm 2i)$ *belongs to the bounded component of the imaginary cubic* $\mathscr{C}_{p_\theta}$.

Proof.

$$4(1/3)^3 - (1/3)g_2 - g_3 = \frac{4}{27} - \frac{1}{3}\left(\frac{16}{3} + 16p_\theta^2\right) - \left(\frac{64}{27} - \frac{16}{3}p_\theta^2\right) = (\pm 2i)^2.$$

□

Proof of Proposition 10. As $X = \sin^2 \varphi$, the period of X is half the period of φ so

$$\frac{T(p_\theta)}{2} = \int_0^{2\omega} \frac{\mathrm{d}t}{\mathrm{d}z} = 2\omega + \frac{1}{\wp'(a)}\left[2\zeta(a)z + \ln\frac{\sigma(z-a)}{\sigma(z+a)}\right]_0^{2\omega}. \tag{23}$$

According to [23, p. 170],

$$\ln\left[\frac{\sigma(2\omega - a)}{\sigma(-a)}\frac{\sigma(a)}{\sigma(2\omega + a)}\right] = \ln e^{-4\eta a} = -4\eta a \mod 2i\pi.$$

The term in brackets in (23) has to be positive because of (21), so $\wp'(a) = +2i$ by the previous lemma ($\wp(a) = 1/3$) and

$$\frac{T(p_\theta)}{4} = \omega + \frac{\omega\zeta(a) - \eta a}{i} \mod \frac{\pi}{2}.$$

The bounded component of the imaginary cubic is parameterized by $z \in \omega + i\mathbf{R}$ so $a = \omega + ib$ and $\zeta(a) = \eta + i\beta$ for some real b and β. □

In order to verify the assumptions of Theorems 1 and 2, one has to differentiate the quasi-period $\Delta\theta$ with respect to p_θ. By means of Proposition 6, it suffices to compute derivatives of the period T. Define

$$\partial := \frac{1}{\delta}\frac{\mathrm{d}}{\mathrm{d}p_\theta}, \quad \delta := \frac{\Delta}{256p_\theta} = p_\theta(8 + 13p_\theta^2 + 16p_\theta^4).$$

Lemma 14.

$$\begin{array}{rclrcl} \partial\omega & = & -A\omega - B\eta, & \partial b & = & -Ab - B\beta + D, \\ \partial\eta & = & C\omega + A\eta, & \partial\beta & = & Cb + A\beta + E, \end{array}$$

with A, B, C, D, E in $\mathbf{R}[p_\theta]$,

$$A = \frac{1}{3}(4 + 13p_\theta^2 + 24p_\theta^4), \quad B = -2 + p_\theta^2, \quad C = \frac{4}{9}(-2 - 5p_\theta^2 + 3p_\theta^4),$$

$$D = 2 + 6p_\theta^2, \quad E = \frac{1}{3}(-4 + 9p_\theta^2).$$

Proof. The derivatives of the (half) period and quasi-period ω and $\eta = \zeta(\omega)$ of $\wp$ and ζ, respectively, with respect to the invariants g_2, g_3 are known (see [23, p. 307]), and

$$\partial = \frac{32}{3}\left(3\frac{\partial}{\partial g_2} - \frac{\partial}{\partial g_3}\right)$$

according to (20). Moreover,

$$\wp(a) = \wp(a(p_\theta), p_\theta) = \frac{1}{3} \quad \text{implies} \quad \wp'(a)\partial a + \partial\wp(a) = 0,$$

$$\text{so} \qquad \partial a = -\frac{1}{2i}\partial\wp(a)$$

(because $\wp'(a) = 2i$), and one also knows the derivatives of $\wp$ (and ζ) with respect to g_2, g_3 (see [23, p. 298]). Then $\partial b = \operatorname{Im}\partial a$. Similarly,

$$\partial[\zeta(a(p_\theta), p_\theta)] = -\wp(a)\partial a + \partial\zeta(a) = -\frac{1}{3}\partial a + \partial\zeta(a)$$

and $\partial\beta$ is obtained taking the imaginary part. □

Lemma 15. *∂T is $\mathbf{R}[p_\theta]$-linear in (ω, η),*

$$\frac{\partial T(p_\theta)}{4} = -(A + E)\omega - (B + D)\eta.$$

Proof. Applying Lemma 14 rules to T, which is bilinear in (ω, η, b, β), one observes that the coefficients of b and β cancel. □

Proposition 11.

$$\Delta\theta'(p_\theta) = -\frac{4}{8 + 13p_\theta^2 + 16p_\theta^4}\left[\frac{2}{3}(11 + 12p_\theta^2)\omega + 7\eta\right],$$

$$\begin{aligned} \Delta\theta''(p_\theta) \quad = \quad & \frac{4}{p_\theta(8 + 13p_\theta^2 + 16p_\theta^4)^2} \times \\ & \left[\frac{2}{3}(24 + 181p_\theta^2 + 830p_\theta^4 + 480p_\theta^6)\omega + (-24 + 143p_\theta^2 + 400p_\theta^4)\eta\right]. \end{aligned}$$

Proof. Lemma 15 combined with Proposition 6 gives the first order derivative of the quasi-period; the second order one is obtained by applying the previous rules anew. □

The penultimate result of the section ensures that the cut and conjugate loci in the $(1,2,1)$ case have the structure described in Theorem 1 and Theorem 2.

Theorem 3. $\Delta\theta' < 0 \leq \Delta\theta''$ *in the* $(1,2,1)$ *case.*

Proof. The first derivative $\Delta\theta'$, expressed as linear combination of the nonnegative (quasi-)periods ω and η, is obviously negative. To obtain nonnegativity of $\Delta\theta''$, one has to check the sign of

$$\frac{2}{3}(24 + 181p_\theta^2 + 830p_\theta^4 + 480p_\theta^6)\omega + (-24 + 143p_\theta^2 + 400p_\theta^4)\eta$$

$$= 24(\frac{2}{3}\omega - \eta) + (\frac{2}{3}181\omega + 143\eta)p_\theta^2 + (\frac{2}{3}830\omega + 400\eta)p_\theta^4 + \frac{2}{3}480\omega p_\theta^6$$

$$= 24(\frac{2}{3}\omega - \eta + p_\theta^2) + (\frac{2}{3}181\omega + 143\eta - 24)p_\theta^2 + (\frac{2}{3}830\omega + 400\eta)p_\theta^4 + \frac{2}{3}480\omega p_\theta^6.$$

Let us denote

$$p_{\theta 1} := \frac{1}{20}\sqrt{\frac{-143 + 7\sqrt{1201}}{2}}$$

the positive root of $-24 + 143p_\theta^2 + 400p_\theta^4$, it is enough to verify that

$$\frac{2}{3}\omega - \eta + p_\theta^2 \geq 0 \quad \text{and} \quad \frac{2}{3}181\omega + 143\eta - 24 \geq 0$$

on $[0, p_{\theta 1}]$. Now, as p_θ tends to 0, η/ω degenerates to $(3/2)g_3/g_2|_{p_\theta=0} = 2/3$ (see [23]), so $2\omega/3 - \eta + p_\theta^2$ vanishes when $p_\theta \to 0$. Moreover, $(2\omega/3 - \eta + p_\theta^2)'$ is equal to

$$\frac{1}{p_\theta(8 + 13p_\theta^2 + 16p_\theta^4)}\left[(16 - \frac{2}{3}\omega - 5\eta)p_\theta^2 + (26 - \frac{20}{3}\omega - 8\eta)p_\theta^4 + 32p_\theta^6\right].$$

Using the coarse estimates $\omega \in [1,2]$ and $\eta \in [1/2, 1]$ on $[0, p_{\theta 1}]$, one has

$$16 - \frac{2}{3}\omega - 5\eta \geq \frac{25}{3} \quad \text{and} \quad 26 - \frac{20}{3}\omega - 8\eta \geq \frac{14}{3},$$

so nonnegativity of the derivative and the function $(2/3)\omega - \eta + p_\theta^2$ follows on this interval. The same holds for $(2/3)181\omega + 143\eta - 24$ that is bounded below by $1009/6$ on $[0, p_{\theta 1}]$, whence the result. □

The condition for a geodesic to be closed is $\Delta\theta \in \pi\mathbf{Q}$ (rationality of $\Delta\theta/\pi$), so that asymptotics when $p_\theta \to \infty$ measure the density of closed curves in the neighbourhood of the equator. This is also related to optimality conditions through injectivity radius (see [8] in the Riemannian case). The asymptotics below use quadratures by means of elliptic integrals detailed in the appendix.

Proposition 12. *In the neighbourhood of meridians,*

$$T \sim 2\pi(1 - \frac{3\sqrt{2}}{8}p_\theta^2 + \frac{105\sqrt{2}}{512}p_\theta^4), \quad \Delta\theta \sim 2\pi(1 - \frac{3\sqrt{2}}{4}p_\theta + \frac{35\sqrt{2}}{128}p_\theta^3), \quad p_\theta \to 0.$$

and in the neighbourhood of the equator,

$$T \sim 4(2-\sqrt{2})K(3-2\sqrt{2})p_\theta^{-1/2}, \quad \Delta\theta \sim \frac{4}{3}(2-\sqrt{2})K(3-2\sqrt{2})p_\theta^{-3/2}, \quad p_\theta \to \infty.$$

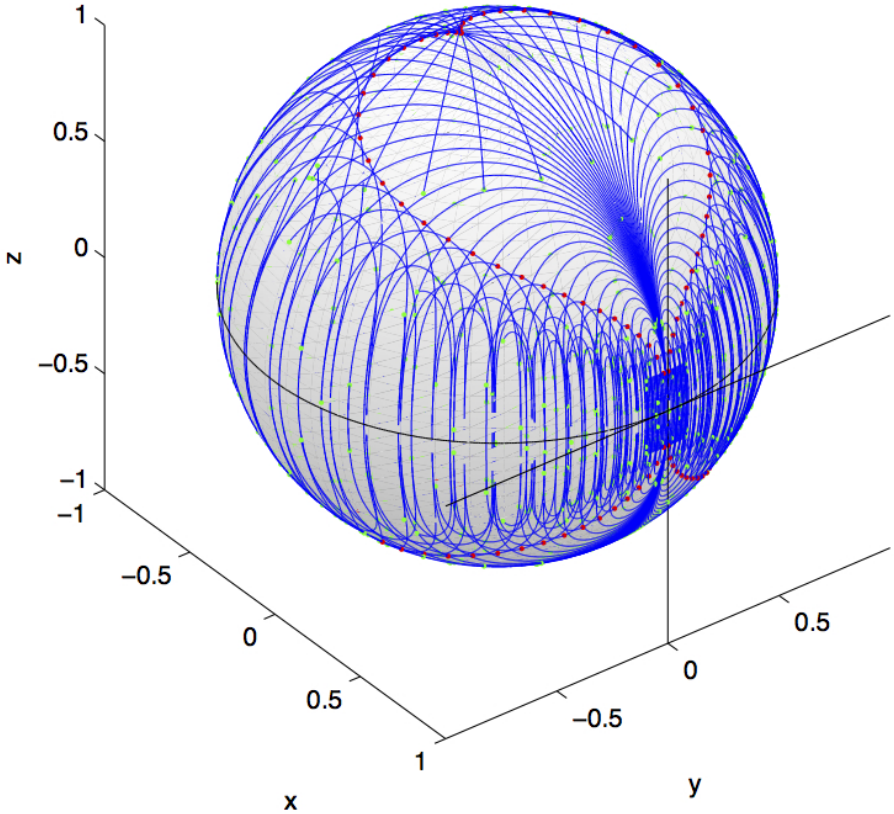

Figure 2: $(1, 2, 1)$ case; geodesics, cut and conjugate loci on the sphere. Symmetric geodesics (in blue) from opposite hemispheres intersect on the equator that, deprived of the origin, forms the cut locus (in black; Theorem 1). The envelope of geodesics generate the conjugate locus (red dots) which has for meridional cusps (Theorem 2).

6 Deformation of the round sphere

We consider the following deformation of the metric (1):

$$XR(\lambda X)\mathrm{d}\theta^2 + \mathrm{d}\varphi^2 \quad (X = \sin^2\varphi), \quad \lambda \in [0, 1]. \tag{24}$$

For $\lambda = 0$, one gets the canonical metric on the sphere as $R(0)$ is normalized to one, so we have an homotopy connecting the round metric to the singular one. For $\lambda < 1$, the metric is Riemannian (no equatorial singularity). In the case $R(X) = 1/(1 - X)$, such a deformation of the round metric appears in [8, 19]. See also [20] for perturbation results of the round sphere.

Proposition 13. *The curvature of the metric (24) is a rational fraction homogeneous in R; for $\lambda \in [0, 1)$,*

$$K_\lambda(X) = 1 + \lambda(4X - 3)\frac{R'(\lambda X)}{R(\lambda X)} + \lambda^2 X(1 - X)\left[\frac{R'^2(\lambda X)}{R^2(\lambda X)} - \frac{2R''(\lambda X)}{R(\lambda X)}\right].$$

Proof. The result follows from the fact that, given a metric $\Gamma(\varphi)\mathrm{d}\theta^2 + \mathrm{d}\varphi^2$ on $\mathbf{S}^2$, the Gaussian curvature is $K = -(\sqrt{\Gamma})''/\sqrt{\Gamma}$. □

Corollary 2 (Concentration of curvature)**.** *The following curvature estimates hold for the metric (24):*

$$K_\lambda(X) = 1 + \lambda R'(0)(4X - 3) + O(\lambda^2) \quad \textit{when} \quad \lambda \to 0,$$

$$K_1(X) \sim -\frac{p(p+1)}{1-X} \quad \textit{when} \quad X \to 1, \quad K_\lambda(1) \sim \frac{p}{1-\lambda} \quad \textit{when} \quad \lambda \to 1,$$

where p is the order of the pole of R.

Corollary 3. *Cut loci of the metric (24) are closed antipodal subarcs for λ close enough to zero.*

Proof. Since $K'_\lambda(X) \sim 4\lambda(R'/R)(0) + O(\lambda^2)$ with $(R'/R)(0) > 0$ by virtue of (2), the curvature is monotone non-decreasing along half-meridians and the result follows from [35] main theorem. □

The existence of conjugate points for metrics with singularities is typical of the almost-Riemannian setting where conjugate points may exist although curvature remains negative whenever defined (see, *e.g.*, [4]). In the $(1, 2, 1)$ case of §5 for instance,

$$K_1(X) = -\frac{(1+X)(4-X)}{(2-X)(1-X)} < 0 \quad (X = \sin^2\varphi).$$

Corollary 2 establishes that although curvature is negative in the neighbourhood of the singularity and even tends to $-\infty$ when $\varphi \to \pi/2$, there is a concentration of positive curvature on the singularity itself as $K_\lambda(1) \to +\infty$ when $\lambda \to 1-$, responsible for the existence of conjugate points after crossing the singularity (cut, and thus conjugate points being located after the equator by antipodality). An alternative interpretation of the singularity comes from the following fact.

Proposition 14. *For $R(X) = 1/(1-X)$ and $\lambda \in [0, 1)$, the metric (24) is conformal to the canonical metric on a oblate ellipsoid of revolution of unit semi-major axis and $\sqrt{1-\lambda}$ semi-minor axis.*

Proof. A parameterization of such an ellipsoid of semi-minor axis μ is

$$x = \sin\varphi\cos\theta, \quad y = \sin\varphi\sin\theta, \quad z = \mu\cos\varphi.$$

In these coordinates, the restriction of the flat $\mathbf{R}^3$ metric reads

$$\sin^2\varphi\,\mathrm{d}\theta^2 + (1 - (1-\mu^2)\sin^2\varphi)\mathrm{d}\varphi^2 = (1-\lambda X)[XR(\lambda X)\mathrm{d}\theta^2 + \mathrm{d}\varphi^2]$$

with $\lambda = 1 - \mu^2$. □

When λ tends to 1, $\mu = \sqrt{1-\lambda}$ tends to zero so the oblate ellipsoid collapses onto a two-sided Poincaré disk, each face being endowed with the flat metric.

Corollary 4. *For $R(X) = 1/(1-X)$, the metric (1) is conformal to the following constant curvature metrics on the Poincaré disk: (i) the flat metric $\mathrm{d}\rho^2 + \rho^2\mathrm{d}\theta^2$ $(K = 0)$, (ii) the canonical Poincaré metric $(\mathrm{d}\rho^2 + \rho^2\mathrm{d}\theta^2)/(1-\rho^2)^2$ $(K = -1)$, where (ρ, θ) are polar coordinates on $\mathbf{D}$.*

Proof. Setting $\rho = \sin\varphi$, one retrieves the standard polar coordinates on the disk. In these coordinates, $\mathrm{d}\rho^2 = (1-X)\mathrm{d}\varphi^2$, so

$$\frac{X}{1-X}\mathrm{d}\theta^2 + \mathrm{d}\varphi^2 = \frac{\mathrm{d}\rho^2 + \rho^2\mathrm{d}\theta^2}{1-\rho^2}.$$

□

In this case, crossing the equatorial singularity can thus be interpretated as crossing the boundary of the disk to go from one side of $\mathbf{D}$ to the other. This can also be seen, like in the flat case, as generating reflections of geodesics with the boundary. As for the canonical Poincaré metric, those reflections turn out to be orthogonal in general.

Proposition 15. *For the metric (1), crossing the equatorial singularity is interpretated on the Poincaré disk as reflecting on the boundary. Reflections on the boundary of the metric (24) are tangential (except for meridians) when $\lambda < 1$, and orthogonal when $\lambda = 1$.*

Proof. In polar coordinates on $\mathbf{D}$, the Hamiltonian (7) reads

$$h(\rho,\theta,p_\rho,p_\theta) = \frac{1}{2}\left[(1-X)p_\rho^2 + \frac{p_\theta^2}{XR(X)}\right], \quad X = \sin^2\varphi = \rho^2.$$

In particular,

$$\dot\theta \sim \frac{p_\theta}{a_p}(1-X)^p \quad \text{and} \quad |\dot\rho| = (1-X)|p_\rho| \sim \sqrt{1-X}$$

when $X \to 1$ since

$$p_\rho^2 = (1-X)^{-1}\left[1 - \frac{p_\theta^2}{XR(X)}\right] \sim (1-X)^{-1}$$

on $\{h = 1/2\}$. Reparameterizing time according to $\mathrm{d}\tau = \mathrm{d}t\sqrt{1-X}$ we get

$$\frac{\mathrm{d}\theta}{\mathrm{d}\tau} \sim \frac{p_\theta}{a_p}(1-X)^{p-1/2} \to 0 \quad \text{and} \quad \left|\frac{\mathrm{d}\rho}{\mathrm{d}\tau}\right| \sim 1, \quad X \to 1,$$

so contacts with $\partial\mathbf{D}$ (*i.e.* for $\rho^2 = X = 1$) are orthogonal reflections. Besides, using homotopy (24) to replace the singular metric by a Riemannian one when $\lambda < 1$ changes the contact for $p_\theta \neq 0$ (meridians, obtained for $p_\theta = 0$, obviously remain perpendicular to the boundary). Indeed, the deformed Hamiltonian is

$$h_\lambda(\rho,\theta,p_\rho,p_\theta) := \frac{1}{2}\left[(1-X)p_\rho^2 + \frac{p_\theta^2}{XR(\lambda X)}\right],$$

and, when $X \to 1$,

$$\dot\theta \sim \frac{p_\theta}{R(\lambda)} \quad \text{and} \quad \dot\rho = p_\varphi\sqrt{1-X}$$

since $p_\varphi = p_\rho\sqrt{1-X}$, which remains finite. So $\dot\theta \neq 0$ if $p_\theta \neq 0$ while $\dot\rho = 0$ at $X = 1$, and contacts with the boundary are tangential outside meridians. □

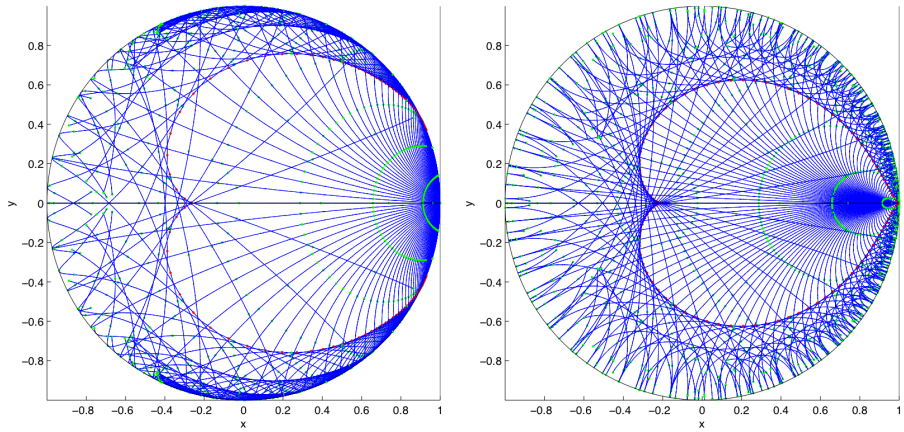

Figure 3: $(1,2,1)$ case; geodesics, cut and conjugate loci on the disk. On the left, metric obtained by homotopy from the round one for $\lambda < 1$; the initial condition is on the boundary, tangential contacts of the geodesics (in blue) with $\partial\mathbf{D}$ are observed. On the right, singular metric ($\lambda = 1$) for the same initial condition (compare with Fig. 2); contacts with $\partial\mathbf{D}$ are orthogonal. In both cases, the conjugate locus (red dots) is the catacaustic generated by reflections of geodesics on the boundary [17]. (In the case of the flat metric on the disk, the catacaustic of geodesics—that is of straight lines—originating from a point on the boundary, *i.e.* at the singularity, would be a cardioid.) Since reflections are orthogonal (or *specular*), the figure can also be interpretated as a billiard on the disk endowed with a particular Riemannian metric.

A Asymptotics in the $(1,2,1)$ case

Let α, β, γ and $\delta = 1$ be the roots of the degree four polynomial $P(X, p_\theta)$ involved in the computation (compare with (19)),

$$P(X, p_\theta) := (1-X)(X(2-X)^2 - 4p_\theta^2(1-X)^2).$$

An alternative quadrature for the period is

$$T(p_\theta) = \frac{4}{\sqrt{A_2 B_1}}\left[\Pi(\nu, k) + \frac{2-p}{p-q}K(k)\right] \tag{25}$$

where K and Π are respectively complete elliptic integrals of first and third kind,

$$\Pi(\nu, k) := \int_0^1 \frac{\mathrm{d}v}{(1-\nu v^2)\sqrt{1-v^2}\sqrt{1-k^2v^2}}\,, \quad K(k) := \Pi(0, k),$$

and

$$\Delta := 4(\beta-\alpha)(\beta-\delta)(\gamma-\alpha)(\gamma-\delta), \quad \sigma := (\alpha+\delta)(\beta+\gamma) - 2(\alpha\delta+\beta\gamma),$$

$$l_1 := \frac{\sigma - \sqrt{\Delta}}{(\beta-\gamma)^2}\,, \quad l_2 := \frac{\sigma + \sqrt{\Delta}}{(\beta-\gamma)^2}\,,$$

$$p := \frac{(\alpha+\delta) - l_1(\beta+\gamma)}{2(1-l_1)}\,, \quad q := \frac{(\alpha+\delta) - l_2(\beta+\gamma)}{2(1-l_2)}\,,$$

$$A_1 := -l_2\frac{1-l_1}{l_2-l_1}\,, \quad B_1 := -l_1\frac{1-l_2}{l_2-l_1}\,, \quad A_2 := \frac{1-l_1}{l_2-l_1}\,, \quad B_2 := \frac{1-l_2}{l_2-l_1}\,,$$

$$a := \sqrt{\frac{A_2}{B_2}}\,, \quad b := \sqrt{\frac{A_1}{B_1}}\,, \quad k := \frac{b}{a}\,, \quad \nu := b^2.$$

With the same notation as before,

$$\begin{aligned} \pm t \;&=\; \frac{1}{2\sqrt{A_2B_1}}\Bigg[\Pi(v,\nu,k) + \frac{2-p}{p-q}\,\mathrm{sn}^{-1}(v,k) \\ &+\; \sqrt{A_2B_1}\,\arctan\Big(\sqrt{A_1A_2}\sqrt{(1-v^2)(1-k^2v^2)} - \sqrt{B_1B_2}(1-\nu v^2)\Big)\Bigg]_v^1, \end{aligned}$$

where the elliptic integral of third kind is now incomplete, and where

$$v := b^{-1}\frac{X-q}{p-X} \in [-1,1] \quad (X = \sin^2\varphi).$$

Similarly,

$$\Delta\theta(p_\theta) = \frac{4p_\theta}{\sqrt{A_2B_1}}\left[\frac{2\Pi(\kappa,k)}{pq} - \frac{2\Pi(\mu,k)}{(2-p)(2-q)} + \frac{4(1-p)^2}{p(2-p)(p-q)}K(k)\right] \tag{26}$$

with, moreover,

$$c := \frac{q}{p}\,, \quad d := \frac{2-q}{2-p}\,, \quad \kappa := \frac{\nu}{c^2}\,, \quad \mu := \frac{\nu}{d^2}\cdot$$

In order to compute the asymptotics of Proposition 12, we need expansions of these roots in the neighbourhood of $p_\theta = 0$ and $p_\theta = \infty$. Such expansions, which are propagated to T using (25), are available in $\sqrt{\varepsilon}$-scale as both

$$Q(X,\varepsilon) := X(2-X)^2 - 4\varepsilon(1-X)^2 \quad \text{and} \quad \widetilde{Q}(X,\varepsilon) := 4(1-X)^2 - \varepsilon X(2-X)^2$$

possess either simple or order two roots for $\varepsilon = 0$, and allow to obtain the asymptotics in Proposition 12.

Lemma 16. *When $p_\theta \to 0$,*

$$\alpha = p_\theta^2 + o(p_\theta^2),$$

$$\beta = 2 - p_\theta\sqrt{2} + \frac{3}{2}p_\theta^2 - \frac{13\sqrt{2}}{16}p_\theta^3 + \frac{1}{2}p_\theta^4 + o(p_\theta^4),$$

$$\gamma = 2 + p_\theta\sqrt{2} + \frac{3}{2}p_\theta^2 + \frac{13\sqrt{2}}{16}p_\theta^3 + \frac{1}{2}p_\theta^4 + o(p_\theta^4).$$

When $p_\theta \to \infty$,

$$\alpha = 1 - \frac{1}{2}p_\theta^{-1} - \frac{1}{8}p_\theta^{-2} + o(p_\theta^{-2}), \quad \beta = 1 + \frac{1}{2}p_\theta^{-1} - \frac{1}{8}p_\theta^{-2} + o(p_\theta^{-2}),$$

$$\gamma = 4p_\theta^2 + 2 + \frac{1}{4}p_\theta^{-2} + o(p_\theta^{-2}).$$

References

[1] Agrachev, A. A. A Gauß-Bonnet formula for contact sub-Riemannian manifolds. *Dokl. Akad. Nauk.*, **381** (2001), 583–585.

[2] Agrachev, A.; Bonnard, B.; Chyba, M.; Kupka, I. Sub-Riemannian sphere in the Martinet flat case. *ESAIM Control Optim. and Calc. Var.* **2** (1997), 377–448.

[3] Agrachev, A.; Boscain, U.; Charlot, G.; Ghezzi, R.; Sigalotti, M. Two-dimensional almost-Riemannian structures with tangency points. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **27** (2010), 793–807.

[4] Agrachev, A.; Boscain, U.; Sigalotti, M. A Gauß-Bonnet like formula on two-dimensional almost-Riemannian manifolds. *Discrete Contin. Dyn. Syst.* **20** (2008), no. 4, 801–822.

[5] Agrachev, A. A.; Sachkov, Y. L. *Control Theory from the Geometric Viewpoint*, Springer, 2004.

[6] Bellaïche, A.; Risler, J.-J. *Sub-Riemannian geometry*, Birkhäuser, 1996.

[7] Besson, G. Géodésiques des surfaces de révolution. *Séminaire de théorie spectrale et géométrie*, **S9** (1991), 33–38.

[8] Bonnard, B.; Caillau, J.-B. Geodesic flow of the averaged controlled Kepler equation. *Forum Math.* **21** (2009), no. 5, 797–814.

[9] Bonnard, B.; Caillau, J.-B.; Cots, O. Energy minimization in two-level dissipative quantum control: The integrable case. *Discrete Contin. Dyn. Syst.* **suppl.** (2011), 198–208. Proceedings of 8th AIMS Conference on Dynamical Systems, Differential Equations and Applications, Dresden, May 2010.

[10] Bonnard, B.; Caillau, J.-B.; Dujol, R. Energy minimization of single-input orbit transfer by averaging and continuation. *Bull. Sci. Math.* **130** (2006), no. 8, 707–719.

[11] Bonnard, B.; Caillau, J.-B.; Sinclair, R.; Tanaka, M. Conjugate and cut loci of a two-sphere of revolution with application to optimal control. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **26** (2009), no. 4, 1081–1098.

[12] Bonnard, B.; Charlot, G.; Ghezzi, R.; Janin, G. The sphere and the cut locus at a tangency point in two-dimensional almost-Riemannian geometry. *J. Dynamical and Control Systems* **17** (2011), no. 1, 141–161.

[13] Bonnard, B.; Chyba, M. *Singular trajectories and their role in control theory.* Math. and Applications **40**, Springer, 2003.

[14] Boscain, U.; Chambrion, T.; Charlot, G. Nonisotropic 3-level quantum systems: Complete solutions for minimum time and minimal energy. *Discrete Contin. Dyn. Syst. Ser. B* **5** (2005), no. 4, 957–990.

[15] Boscain, U.; Charlot, G.; Ghezzi, R.; Sigalotti, M. Lipschitz classification of almost-Riemannian distances on compact oriented surfaces. *J. Geom. Anal.*, to appear.

[16] Bourbaki, N. *Variétés différentielles et analytiques*, Hermann, 1971.

[17] Bruce, J. W.; Giblin, P. J.; Gibson, C. G. *Topology*, **21** (1982), no. 2. 179–199.

[18] Caillau, J.-B.; Daoud, B.; Gergaud, J. On some Riemannian aspects of two and three-body controlled problems. *Recent Advances in Optimization and its Applications in Engineering*, 205–224, Springer, 2010. Proceedings of the 14th Belgium-Franco-German conference on Optimization, Leuven, September 2009.

[19] Faridi, A.; Schucking, E. Geodesics and deformed spheres. *Proc. Amer. Math. Soc.* **100** (1987), no. 3, 522–525.

[20] Figalli, A.; Rifford, L.; Villani, C. Nearly round spheres look convex. *Amer. J. Math.* **134** (2012), no. 1, 109–139.

[21] El Alaoui, C.; Gauthier, J.-P.; Kupka, I. Small sub-Riemannian balls on $\mathbf{R}^3$. *J. Dynam. Control Systems* **2** (1996), 359–421.

[22] Grušin, V. V. A certain class of elliptic pseudodifferential operators that are degenerate on a submanifold (Russian). *Mat. Sb. (N.S.)* **84** (1971), no. 126, 163–195. English translation: *Math. USSR-Sb.* **13** (1971), 155–185.

[23] Halphen, G.-H. *Traité des fonctions elliptiques et de leurs applications*, Première Partie. Gauthier-Villars, 1886.

[24] Jean, F. *Sub-Riemannian Geometry.* Lectures given at SISSA, 2003.

[25] Jones, G. A.; Singerman, D. *Complex Functions. An Algebraic and Geometric Viewpoint.* Cambridge University Press, 1987.

[26] Montgomery, R. *A tour of subriemannian geometries, their geodesics and applications.* American Mathematical Society, 2002.

[27] Moser, J. K. Regularization of Kepler's problem and the averaging method on a manifold. *Comm. Pure Appl. Math.* **23** (1970), 609–635.

[28] Myers, S. B. Connections between geometry and topology I. *Duke Math. J.* **1** (1935), 376–391.

[29] Myers, S. B. Connections between geometry and topology II. *Duke Math. J.* **2** (1936), 95–102.

[30] Pelletier, F. Sur le théorème de Gauß-Bonnet pour les pseudo-métriques singulières. *Séminaire de théorie spectrale et géométrie* (1987), no. 5, 99–105.

[31] Poincaré, H. Sur les lignes géodésiques des surfaces convexes. *Trans. Amer. Math. Soc.* **5** (1905), 237–274.

[32] Sakai, T. *Riemannian Geometry.* Amer. Math. Soc., 1995.

[33] Sarychev, A. V. The index of second variation of a control system, *Mat. Sb.* **41** (1982), 338–401.

[34] Shiohama, K.; Shioya, T.; Tanaka, M. *The geometry of total curvature on complete open surfaces.* Cambridge University Press, 2003.

[35] Sinclair, R.; Tanaka, M. The cut locus of a two-sphere of revolution and Toponogov's comparison theorem. *Tohoku Math. J.* **2** (2007), no. 59, 379–399.